

# Les Houches 2017: Physics at TeV Colliders

## Standard Model Working Group Report

### Conveners

*Higgs physics: SM issues*

J. Bendavid (CMS), F. Caola (Theory), R. Harlander (Theory), K. Tackmann (ATLAS)

*SM: Loops and Multilegs*

G. Heinrich (Theory), J. Huston (ATLAS), S. Kallweit (Theory), J. Thaler (Jets contact), K. Theofilatos (CMS)

*Tools and Monte Carlos*

V. Ciulli (CMS), E. Re (Theory), S. Prestel (Theory)

### Abstract

This Report summarizes the proceedings of the 2017 Les Houches workshop on Physics at TeV Colliders. Session 1 dealt with (I) new developments relevant for high precision Standard Model calculations, (II) theoretical uncertainties and dataset dependence of parton distribution functions, (III) new developments in jet substructure techniques, (IV) issues in the theoretical description of the production of Standard Model Higgs bosons and how to relate experimental measurements, (V) phenomenological studies essential for comparing LHC data from Run II with theoretical predictions and projections for future measurements, and (VI) new developments in Monte Carlo event generators.

### Acknowledgements

We would like to thank the organizers (N. Berger, F. Boudjema, C. Delaunay, M. Delmastro, S. Gascon, P. Gras, J. P. Guillet, B. Herrmann, S. Kraml, G. Moreau, E. Re, P. Slavich and D. Zerwas) and the Les Houches staff for the stimulating environment always present at Les Houches. We thank the Labex ENIGMASS, CNRS, LAPTh, LAPP, the Université Savoie Mont Blanc, and the Ministère des Affaires étrangères for support.

arXiv:1803.07977v1 [hep-ph] 21 Mar 2018

## Authors

J. R. Andersen<sup>1</sup>, J. Bellm<sup>2</sup>, J. Bendavid<sup>3</sup>, N. Berger<sup>4</sup>, D. Bhatia<sup>5</sup>, B. Biedermann<sup>6</sup>, S. Bräuer<sup>7</sup>, D. Britzger<sup>8</sup>, A. G. Buckley<sup>9</sup>, R. Camacho<sup>3</sup>, F. Caola<sup>1</sup>, G. Chachamis<sup>10</sup>, S. Chatterjee<sup>5</sup>, X. Chen<sup>11</sup>, M. Chiesa<sup>6</sup>, V. Ciulli<sup>12</sup>, J. R. Currie<sup>1</sup>, A. Denner<sup>6</sup>, F. Dreyer<sup>13</sup>, F. Driencourt-Mangin<sup>14</sup>, S. Forte<sup>15</sup>, M. V. Garzelli<sup>16</sup>, T. Gehrmann<sup>11</sup>, S. Gieseke<sup>17</sup>, E. W. N. Glover<sup>1</sup>, P. Gras<sup>18</sup>, N. Greiner<sup>11</sup>, C. Gütschow<sup>19</sup>, C. Gwenlan<sup>20</sup>, R. Harlander<sup>21</sup>, M. Heil<sup>1</sup>, G. Heinrich<sup>22</sup>, M. Herndon<sup>23</sup>, V. Hirschi<sup>24</sup>, A. H. Hoang<sup>25</sup>, S. Höche<sup>26</sup>, A. Huss<sup>27</sup>, J. Huston<sup>28</sup>, S. P. Jones<sup>22</sup>, S. Kallweit<sup>27</sup>, D. Kar<sup>29</sup>, A. Karlberg<sup>11</sup>, Z. Kassabov<sup>15,30</sup>, M. Kerner<sup>22</sup>, J. Klappert<sup>21</sup>, S. Kuttimalai<sup>26</sup>, J.-N. Lang<sup>11</sup>, A. Larkoski<sup>31</sup>, J. M. Lindert<sup>1</sup>, P. Loch<sup>32</sup>, K. Long<sup>23</sup>, L. Lönnblad<sup>2</sup>, G. Luisoni<sup>22</sup>, A. Maier<sup>1</sup>, P. Maierhöfer<sup>33</sup>, D. Maître<sup>1</sup>, S. Marzani<sup>34</sup>, J. A. McFayden<sup>3</sup>, I. Moulst<sup>35,36</sup>, M. Mozer<sup>37</sup>, S. Mrenna<sup>38</sup>, B. Nachman<sup>39</sup>, D. Napoletano<sup>1,40</sup>, C. Pandini<sup>3</sup>, A. Papaefstathiou<sup>41,42</sup>, M. Pellen<sup>6</sup>, L. Perrozzi<sup>43</sup>, J. Pires<sup>22,44</sup>, S. Plätzer<sup>25</sup>, S. Pozzorini<sup>11</sup>, S. Prestel<sup>38</sup>, S. Quackenbush<sup>45</sup>, K. Rabbertz<sup>37</sup>, M. Rauch<sup>17</sup>, E. Re<sup>27,46</sup>, C. Reuschle<sup>45</sup>, P. Richardson<sup>1,27</sup>, A. Gehrmann-De Ridder<sup>24</sup>, G. Rodrigo<sup>14</sup>, J. Rojo<sup>41,47</sup>, R. Röntsch<sup>48</sup>, L. Rottoli<sup>49</sup>, D. Samitz<sup>25</sup>, T. Samui<sup>5,50</sup>, G. Sborlini<sup>15</sup>, M. Schönherr<sup>27</sup>, S. Schumann<sup>7</sup>, L. Scyboz<sup>22</sup>, S. Seth<sup>1</sup>, H.-S. Shao<sup>51</sup>, A. Siódmok<sup>52</sup>, P. Z. Skands<sup>53</sup>, J. M. Smillie<sup>54</sup>, G. Soyez<sup>40</sup>, P. Sun<sup>28</sup>, M. R. Sutton<sup>55</sup>, F. J. Tackmann<sup>56</sup>, K. Tackmann<sup>56</sup>, J. Thaler<sup>13</sup>, K. Theofilatos<sup>43</sup>, S. Uccirati<sup>30</sup>, S. Weinzierl<sup>57</sup>, E. Yazgan<sup>58</sup>, C.-P. Yuan<sup>28</sup>, F. Yuan<sup>59</sup>

<sup>1</sup> Institute for Particle Physics Phenomenology, Durham University, Durham, DH1 3LE, UK

<sup>2</sup> Department of Astronomy and Theoretical Physics, Lund University, Sweden

<sup>3</sup> Experimental Physics Department, CERN, CH-1211 Geneva 23, Switzerland

<sup>4</sup> LAPP, CNRS/IN2P3 & Université Savoie Mont Blanc, 74940 Annecy, France

<sup>5</sup> Tata Institute of Fundamental Research, Mumbai, India

<sup>6</sup> Institut für Theoretische Physik und Astrophysik, Universität Würzburg, 97074 Würzburg, Germany

<sup>7</sup> Institute for Theoretical Physics, Georg-August-University Göttingen, 37077 Göttingen, Germany

<sup>8</sup> Physikalisches Institut, Universität Heidelberg, Germany

<sup>9</sup> School of Physics and Astronomy (SUPA), University of Glasgow, Glasgow, G12 8QQ, Scotland, UK

<sup>10</sup> Instituto de Física Teórica UAM/CSIC & Universidad Autónoma de Madrid, 28049 Madrid, Spain

<sup>11</sup> Physik Institut, Universität Zürich, CH-8057 Zürich, Switzerland

<sup>12</sup> Dipartimento di Fisica e Astronomia, Università di Firenze and INFN Sezione di Firenze, 50019 Sesto Fiorentino, Firenze, Italy

<sup>13</sup> Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

<sup>14</sup> Instituto de Física Corpuscular, Universitat de València – Consejo Superior de Investigaciones Científicas, Parc Científic, 46980 Paterna, Valencia, Spain

- 15 Tif Lab, Dipartimento di Fisica, Università di Milano and INFN, Sezione di Milano, 20133 Milan, Italy
- 16 II. Institute for Theoretical Physics, Hamburg University, 22761 Hamburg, Germany
- 17 Institute for Theoretical Physics, Karlsruhe Institute of Technology (KIT), 76131 Karlsruhe, Germany
- 18 IRFU, CEA, Université Paris-Saclay, Gif-sur-Yvette, France
- 19 Department of Physics and Astronomy, University College London, London, WC1E 6BT, UK
- 20 The University of Oxford, Oxford, UK
- 21 Institute for Theoretical Particle Physics and Cosmology, RWTH Aachen University, 52056 Aachen, Germany
- 22 Max Planck Institute for Physics, 80805 München, Germany
- 23 University of Wisconsin, Madison, WI 53706, USA
- 24 Institute for Theoretical Physics, ETH Zürich, CH-8093 Zürich, Switzerland
- 25 Particle Physics, Faculty of Physics, University of Vienna, 1090 Vienna, Austria
- 26 SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA
- 27 Theoretical Physics Department, CERN, CH-1211 Geneva 23, Switzerland
- 28 Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, USA
- 29 University of the Witwatersrand, Johannesburg, South Africa
- 30 Dipartimento di Fisica, Università di Torino e INFN, Sezione di Torino, 10125 Turin, Italy
- 31 Physics Department, Reed College, Portland, OR 97202, USA
- 32 Department of Physics, University of Arizona, Tucson, AZ 85721, USA
- 33 Physikalisches Institut, Albert-Ludwigs-Universität Freiburg, 79104 Freiburg, Germany
- 34 Dipartimento di Fisica, Università di Genova and INFN, Sezione di Genova, 16146 Genoa, Italy
- 35 Berkeley Center for Theoretical Physics, University of California, Berkeley, CA 94720, USA
- 36 Theoretical Physics Group, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
- 37 Institute of Experimental Particle Physics, Karlsruhe Institute of Technology (KIT), 76131 Karlsruhe, Germany
- 38 Fermi National Accelerator Laboratory, Batavia, IL, 60510-0500, USA
- 39 Physics Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

- <sup>40</sup> IPhT, CEA Saclay, CNRS, Université Paris-Saclay, 91191 Gif-sur-Yvette, France
- <sup>41</sup> Nikhef, Theory Group, Science Park 105, 1098 XG, Amsterdam, The Netherlands
- <sup>42</sup> Institute for Theoretical Physics Amsterdam and Delta Institute for Theoretical Physics, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands
- <sup>43</sup> Institute for Particle Physics and Astrophysics, ETH Zürich, CH-8093 Zürich, Switzerland
- <sup>44</sup> Centro de Física Teórica de Partículas (CFTP), Instituto Superior Técnico IST, Universidade de Lisboa, 1049-001 Lisboa, Portugal
- <sup>45</sup> Physics Department, Florida State University, Tallahassee, FL 32306, USA
- <sup>46</sup> LAPTh, CNRS, Université Savoie Mont Blanc, 74940 Annecy, France
- <sup>47</sup> Department of Physics and Astronomy, VU University, 1081 HV Amsterdam
- <sup>48</sup> Institute for Theoretical Particle Physics, Karlsruhe Institute of Technology (KIT), 76131 Karlsruhe, Germany
- <sup>49</sup> Rudolf Peierls Centre for Theoretical Physics, University of Oxford, OX1 3NP Oxford, UK
- <sup>50</sup> Indian Institute of Technology, Kanpure, India
- <sup>51</sup> Laboratoire de Physique Théorique et Hautes Énergies (LPTHE), UMR 7589, Sorbonne Université et CNRS, 75252 Paris Cedex 05, France
- <sup>52</sup> The Henryk Niewodniczański Institute of Nuclear Physics in Cracow, Polish Academy of Sciences
- <sup>53</sup> School of Physics and Astronomy, Monash University, Clayton, VIC 3800, Australia
- <sup>54</sup> Higgs Centre for Theoretical Physics, University of Edinburgh, Edinburgh EH9 3FD, UK
- <sup>55</sup> The University of Sussex, Falmer, Brighton, UK
- <sup>56</sup> Deutsches Elektronen-Synchrotron (DESY), 22607 Hamburg, Germany
- <sup>57</sup> Institut für Physik, Universität Mainz, 55099 Mainz, Germany
- <sup>58</sup> Institute of High Energy Physics, Chinese Academy of Sciences, Beijing, China
- <sup>59</sup> Nuclear Science Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

# Contents

<b>I</b>	<b>NLO automation and (N)NLO techniques</b>	<b>4</b>
1	Update on the precision Standard Model wish list <sup>1</sup> . . . . .	4
2	NTuples for NNLO processes <sup>2</sup> . . . . .	19
3	Progress in grid techniques for the fast reproduction of QCD calculations at NNLO <sup>3</sup>	21
4	The nested soft-collinear subtraction scheme <sup>4</sup> . . . . .	30
5	Loop-tree duality and the four-dimensional unsubtraction <sup>5</sup> . . . . .	35
6	Loop amplitudes: The numerical approach <sup>6</sup> . . . . .	44
7	NLO EW automation and technical comparisons <sup>7</sup> . . . . .	49
<b>II</b>	<b>Parton distribution functions</b>	<b>80</b>
1	Theoretical uncertainties and dataset dependence of parton distributions <sup>8</sup> . . . . .	80
<b>III</b>	<b>Jet substructure studies</b>	<b>87</b>
1	Towards extracting the strong coupling constant from jet substructure at the LHC <sup>9</sup>	87
2	Performance versus robustness: Two-prong substructure taggers for the LHC <sup>10</sup> . .	110
<b>IV</b>	<b>Standard Model Higgs</b>	<b>138</b>
1	Transverse momentum resummation in Higgs boson plus two jet production via weak boson fusion process <sup>11</sup> . . . . .	138
2	Higher order corrections in VBF Higgs production <sup>12</sup> . . . . .	140
3	The gluon-fusion component of Higgs boson-jet-jet production and its suppression in weak-boson fusion studies <sup>13</sup> . . . . .	143
4	The sensitivity of a $ZH/WH$ ratio measurement in the $H \rightarrow b\bar{b}$ channel <sup>14</sup> . . . . .	146
5	Parton shower and top quark mass effects in Higgs boson pair production <sup>15</sup> . . . . .	150
6	Treatment of theory uncertainties for Simplified Template Cross Sections <sup>16</sup> . . . . .	159

<sup>1</sup> F. Caola, G. Heinrich, J. Huston, S. Kallweit, K. Theofilatos

<sup>2</sup> D. Maître

<sup>3</sup> D. Britzger, C. Gwenlan, A. Huss, K. Rabbertz, M. R. Sutton

<sup>4</sup> R. Röntsch

<sup>5</sup> G. Chachamis, F. Driencourt-Mangin, G. Rodrigo, G. Sborlini

<sup>6</sup> S. Seth, S. Weinzierl

<sup>7</sup> S. Kallweit (section coordinator); B. Biedermann, S. Bräuer, M. Chiesa, A. Denner, N. Greiner, V. Hirschi, J.-N. Lang, J. M. Lindert, P. Maierhöfer, M. Pellen, S. Pozzorini, S. Quackenbush, C. Reuschle, M. Schönherr, S. Schumann, H.-S. Shao, S. Uccirati

<sup>8</sup> S. Forte, Z. Kassabov, J. Rojo, L. Rottoli

<sup>9</sup> I. Moutl, B. Nachman, G. Soyez, J. Thaler (section coordinators); S. Chatterjee, F. Dreyer, M. V. Garzelli, P. Gras, A. Larkoski, S. Marzani, A. Siódmok, A. Papaefstathiou, P. Richardson, T. Samui

<sup>10</sup> P. Loch, I. Moutl, B. Nachman, G. Soyez, J. Thaler (section coordinators); D. Bhatia, R. Camacho, G. Chachamis, S. Chatterjee, F. Dreyer, D. Kar, A. Papaefstathiou, T. Samui, A. Siódmok

<sup>11</sup> P. Sun, C.-P. Yuan, F. Yuan

<sup>12</sup> F. Dreyer, A. Karlberg, M. Rauch

<sup>13</sup> J. R. Andersen, M. Heil, A. Maier, J. M. Smillie

<sup>14</sup> R. Harlander, J. Klappert, C. Pandini, A. Papaefstathiou, L. Perrozzi

<sup>15</sup> G. Heinrich, S. P. Jones, M. Kerner, S. Kuttimalai, G. Luisoni

<sup>16</sup> N. Berger, F. J. Tackmann, K. Tackmann

<b>V</b>	<b>Phenomenological studies</b>	<b>173</b>
1	Top-pair production including multi-jet merging and EW corrections <sup>17</sup> . . . . .	173
2	Parton shower and off-shell effects in top quark mass determinations <sup>18</sup> . . . . .	178
3	Study of electroweak production of WZ in association with two jets at the LHC <sup>19</sup>	184
4	Considerations for the underlying event in jet calibration at the LHC <sup>20</sup> . . . . .	200
5	Precision comparisons of predictions for Higgs boson + jet production at the LHC as a function of jet size <sup>21</sup> . . . . .	206
<b>VI</b>	<b>Monte Carlo studies</b>	<b>215</b>
1	On renormalization scale variations in parton showers <sup>22</sup> . . . . .	215
2	On the cross-talk of parameter optimization and perturbative variations in event generators <sup>23</sup> . . . . .	224
3	Determination of nonperturbative correction factors and their dependence on Monte Carlo modeling <sup>24</sup> . . . . .	239
4	Is there room for improvement in the description of the gluon splitting to heavy quarks? <sup>25</sup> . . . . .	245

---

<sup>17</sup> C. Gütschow, J. M. Lindert, M. Schönherr

<sup>18</sup> G. Heinrich, L. Scyboz

<sup>19</sup> K. Long, M. Pellen (section coordinators); S. Bräuer, V. Ciulli, S. Gieseke, M. Herndon, M. Mozer, S. Plätzer, M. Rauch, E. Yazgan

<sup>20</sup> J. Huston, P. Loch

<sup>21</sup> J. Bellm, A. G. Buckley, X. Chen, J. R. Currie, A. Gehrmann-De Ridder, T. Gehrmann, E. W. N. Glover, S. Höche, A. Huss, J. Huston, S. Kuttimalai, J. Pires, S. Plätzer, M. Schönherr

<sup>22</sup> S. Höche, S. Mrenna, S. Prestel, M. Schönherr, P. Z. Skands

<sup>23</sup> J. Bellm, L. Lönnblad, S. Plätzer, S. Prestel, D. Samitz, A. Siodmok, A. H. Hoang

<sup>24</sup> J. Bellm, S. Kuttimalai

<sup>25</sup> J. A. McFayden, D. Napoletano, E. Re

As the experimental precision continues to improve during the 13 TeV running at the LHC, the requirements on the corresponding theoretical predictions have increased as well. The predictions include those defined at fixed-order, those resumming large logarithms due to kinematic thresholds and boundaries, and those involving parton showering, and subsequent hadronization. The latter allows for direct comparison at the hadron level to data. All levels of theoretical predictions are needed for a full exploration of LHC physics. We continue in these proceedings to discuss advances in theoretical predictions, while also examining their connections, their limitations, and their prospects for improvement.

Calculations for  $2 \rightarrow 2$  processes at NNLO start to become the new standard, with  $2 \rightarrow 3$  calculations on the horizon. One of the complications of such calculations is the dissemination of the results. In contrast to NLO, most of these programs are too difficult/lengthy for non-author users to be able to run independently. Two contributions discuss flexible storage possibilities for the results of calculations, either as ROOT ntuples or in grid form, both of which techniques have been very successful at NLO. Part of the impressive progress in NNLO calculations is due to the development of more efficient methods to treat infrared divergent real radiation at NNLO. Among these methods is a novel scheme, called “nested soft-collinear subtraction scheme”, which is presented in these proceedings. A different approach aims to avoid the occurrence of (dimensionally regulated) poles which need to be isolated and cancelled between real and virtual parts by combining all the contributions at integrand level and then performing the integrations completely numerically in four dimensions. Such four-dimensional frameworks are promising as they avoid some tedious technicalities related to calculations in  $D$  dimensions. However these methods face hurdles of different type, related to the purely numerical approach. Progress how to overcome these is reported here. Automation of electroweak NLO corrections has seen lots of progress in recent years, and a tuned comparison to validate the different available tools for obtaining EW one-loop amplitudes and complete fixed-order results at NLO EW accuracy are presented for off-shell  $ZZ$  and  $WW$  production, at the level of amplitudes as well as integrated and differential cross sections.

A key ingredient for any theoretical prediction at hadron colliders are parton distribution functions (PDFs). In the recent past, improved fit methodologies in combination with the use of new LHC data in the fits have led to a reduction of the nominal PDFs uncertainties. In many phenomenologically relevant cases, they are now at the level of few percent. Nevertheless, so far PDFs errors only reflect uncertainties coming from the fitted data, and not from the theoretical predictions used in the fit. Given the reduction of the former, the latter is no longer negligible. To address this issue, we perform preliminary studies of theory uncertainty in PDFs fits. More precisely, we study the convergence of PDFs fits at different orders in perturbation theory, and use perturbative stability as a handle on theoretical uncertainties.

Jet substructure techniques are widely applied at the LHC for testing QCD in extreme regions of phase space and for searching beyond the Standard Model for new physical phenomena. In these proceedings, we perform two studies that could extend the LHC physics program in both of these areas. To improve our understanding of QCD using jet substructure, we study the feasibility of combining precision jet substructure measurements and calculations to extract the strong coupling constant,  $\alpha_s$ . We find that 10% precision may already be feasible using existing technology in a region of phase space complementary to other methods for measuring  $\alpha_s$ . To improve searches for beyond the Standard Model physics using jet substructure, we study the trade-off between performance and robustness for tagging boosted  $W$  bosons. We identify two-prong tagging strategies that could enhance the signal significance and robustness to non-perturbative effects compared with the default methods used at ATLAS and CMS.

One of the pillars of the LHC program is the detailed study of the Higgs boson. Understanding if the discovered particle has the properties as predicted by the Standard Model, or if

deviations from the Standard Model predictions point to beyond-the-Standard-Model effects in the Higgs sector, requires increased precision both in theoretical predictions as well as experimental measurements. One contribution studies the soft gluon resummation effect in the Higgs boson plus two-jet production in the weak boson fusion process at the LHC. The results are also compared with the prediction of the Monte Carlo event generator `Pythia8`. We also include updated results for fully differential Higgs production in weak boson fusion at NNLO, with a particular emphasis on the structure of jet dynamics in the experimental fiducial region. Gluon fusion production with two jets is an important background to measurements of weak boson fusion. One of the contributions compares perturbative predictions for gluon fusion production with two jets and studies the suppression of this background. Another contribution studies the potential to extract signals from beyond the Standard Model by comparing the strongly related processes of  $WH$  and  $ZH$  production. In several loop induced processes, in particular  $H$ ,  $HH$  and  $ZZ$  production in gluon fusion, a substantial hardening of the  $p_T$  spectrum of the Higgs boson or the boson pair has been observed when matching the fixed order NLO result to a parton shower. An investigation of the reasons for this observation in the case of Higgs boson pair production, matched to `Powheg` combined with both `Pythia6` and `Pythia8`, `MG5_aMC@NLO+Pythia8` and `Sherpa` with both `CS`-shower and `Dire`-shower, is presented here. Furthermore, we develop a scheme for the parametrization of theory uncertainties for weak boson fusion and associated production with weak bosons in the context of the Simplified Template Cross Section (STXS) framework and discuss in which form to report theory uncertainties on STXS measurements to allow for their coherent treatment in the interpretation of the measurements.

Relevant progress on theoretical calculations has been achieved recently on many other important LHC physics processes. The results of several phenomenological studies are presented. A study for top quark pair production incorporating electroweak and QCD higher-order corrections into the multi-jet merging framework of `Sherpa` is presented. This is in particular targeted to allow for a reliable modelling of the top transverse momentum distribution, which is crucial for a multitude of new physics searches. Monte Carlo predictions at particle level are compared against a recent ATLAS measurement in the lepton+jet channel for the reconstructed top transverse momentum spectrum. Excellent agreement between Monte Carlo predictions and data is found when the electroweak corrections are included. Top quark pair production in the di-lepton channel is investigated comparing four different theoretical descriptions. The full NLO corrections to  $pp \rightarrow W^+W^-b\bar{b} \rightarrow (e^+\nu_e)(\mu^-\bar{\nu}_\mu)b\bar{b}$  production are compared to calculations in the narrow width approximation, where the production of a top quark pair is calculated at NLO and combined with different descriptions of the top quark decay: LO, NLO and via a parton shower. The study works out differences in the shape of the  $m_{lb}$  and  $m_{Wb}$  distributions in view of top quark mass determinations, in particular showing that corrections beyond the leading order in the decay play a significant role. A further study addresses the final state  $\ell^+\nu_\ell\ell^+\ell^-jj$ , which is of particular interest as it proceeds via diagrams featuring WZ scattering. The different contributions to this final state are first evaluated showing that the QCD contributions are overwhelming over the EW ones. Then, a comparison of theoretical predictions at LO accuracy and at LO matched with various parton showers is performed. While at LO all predictions are in good agreement, the inclusion of parton shower effects can introduce large differences especially for observables defined beyond LO (*e.g.* third-jet observables). In another contribution, we include a short pedagogical discussion of the treatment of the underlying event in measurements at high luminosity at the LHC, a treatment that is unknown even to most members of the experimental collaborations. We also examine how well fixed-order predictions, which often have the highest available accuracy, can describe the jet R-dependence for the inclusive jet, Z+jet and Higgs+jet final states. This tests how well fixed-order predictions, containing an increasing number of parton emissions as the perturbative order increases, can describe jet shapes, as compared to parton shower predictions. On the other hand, the non-perturbative corrections needed

by the fixed-order predictions for best comparison to data, can only be obtained from the parton shower Monte Carlo predictions for the relevant processes. This continues the study from Les Houches 2015, where good agreement was shown between fixed order and parton shower Monte Carlo predictions (for non-Sudakov observables) for Higgs+jets production with one value of  $R$  (0.4).

As parton shower Monte Carlo programs are crucial for physics at the LHC, we continue the direct comparisons of the different parton shower algorithms started in Les Houches 2015, in order to gain a better understanding of any differences that may result in intrinsic uncertainties for the resultant predictions. This concerns *e.g.* parton shower variations due to renormalization scale variations and their interplay with hadronization, or the correlation between PDF choices and the modeling of underlying event via multiparton interactions. Both of these points are addressed in a study that shows that correlations exist and that a sensible description of data sensitive to non-perturbative effects requires a retuning of the event generator when varying renormalization scales. This retuning does not necessarily mean a loss of predictivity in perturbatively dominated phase-space regions. Another study investigates how scale variations in the parton shower can be defined when considering known higher-order corrections. Keeping these higher-order corrections intact when performing scale variations requires the introduction of compensating terms. The use of similar compensation schemes leads to closer agreement of the variation bands of different parton-shower predictions, and interpolates between previous aggressive or conservative uncertainty estimates. The influence of Monte Carlo modelling on the extraction and calculation of nonperturbative correction factors is also considered in another study, where it is addressed if the choice of PDF sets and Monte Carlo models used for the extraction of nonperturbative correction factors may bias other measurements in which these correction factors are subsequently used. A short study that compares different Monte Carlo generators for a process involving  $B$ -hadrons in the final state is also included, with the aim of assessing to which extent NLO-based tools are successful in simulating kinematic configurations sensitive to the gluon splitting to pair of (massive)  $b$ -quarks.

# Chapter I

## NLO automation and (N)NLO techniques

### 1 Update on the precision Standard Model wish list <sup>1</sup>

Identifying key observables and processes that require improved theoretical input has been a key part of the Les Houches programme. In this contribution we briefly summarise progress since the previous report in 2015 and explore the possibilities for further advancements. For the first time, we also provide an estimate of the experimental uncertainties for key processes.

#### 1.1 Introduction

The period since the Les Houches 2015 report [1] is marked by significant progress in the automation of electroweak corrections and the production of NNLO results in an almost industrial way. The latter is mainly due to the development of methods which allow to treat the doubly unresolved real radiation parts occurring at NNLO in a largely automated manner, as well as due to the availability of two-loop integrals with an increasing number of kinematic scales.

On the parton shower side, NLO QCD matched results and matrix element improved multi-jet merging techniques have become a standard level of theoretical precision. The automation of full SM corrections including NLO electroweak predictions has also seen major improvements.

Another challenge is to make the NNLO  $2 \rightarrow 2$  predictions or complex NLO predictions publicly available to experimental analyses, and there has been major progress to achieve this goal. ROOT NTUPLES have been a useful tool for complicated final states at NLO and allow for very flexible re-weighting and analysis. The cost for this is the large disk space required to store the event information. A feasibility study using ROOT NTUPLES to store the much larger NNLO events in  $pp \rightarrow 2$  jets [2] is described in Sec. I.2. An extension of ApplGrid [3] and FastNLO [4] offers a simpler, but less flexible method to distribute higher order predictions. The latter option is likely to be used heavily in precision PDF fits, and new developments in the APPLfast project are described in Sec. I.3.

#### 1.2 Developments in theoretical methods

Precision predictions require a long chain of various tools and methods, all of which demand highly technical computations.

Computational methods for the amplitude level ingredients have seen substantial progress in the last few years. Scattering amplitudes at  $L$  loops are generally decomposed into a basis of integrals together with rational coefficients,

$$A_{2 \rightarrow n}^{(L)} = \sum (\text{coefficients})_i (\text{integrals})_i, \quad (\text{I.1})$$

one must then remove infrared singularities to obtain a finite cross-section,

$$d\sigma_{2 \rightarrow n} \text{N}^k \text{LO} = \text{IR}_k(A_{2 \rightarrow n}^k, A_{2 \rightarrow n+1}^{k-1}, \dots, A_{2 \rightarrow n+k}^0). \quad (\text{I.2})$$

where the function  $\text{IR}_k$  represents an infrared subtraction technique; a recently developed one is presented in Sec. I.4. Ultraviolet renormalisation must also be performed but in a (semi-)analytic approach presents no technical difficulties. There are also fully numerical approaches,

---

<sup>1</sup> F. Caola, G. Heinrich, J. Huston, S. Kallweit, K. Theofilatos

aiming to calculate higher order corrections without the separation into individually divergent components, such that 4-dimensional methods can be applied. Sections I.5 and I.6 are dedicated to a description of such methods.

### 1.2.1 Loop integrals

Most of the new analytic results for two-loop integrals and beyond have been calculated employing the differential equations technique [5, 6], which got a significant boost through Henn’s canonical form [7]. Since the last Les Houches report, several tools to find a canonical basis automatically have been developed: `epsilon` [8], `FUCHSIA` [9] and `CANONICA` [10]. Important new developments concerning the differential equations technique to calculate multi-loop integrals can be found in Refs. [11–14]. For a review on the method of differential equations we refer to Ref. [15].

Major progress since the last report has been made in the calculation of two-loop master integrals with massive propagators, for example the planar ones entering Higgs+jet [16, 17],  $gg \rightarrow \gamma\gamma$  via massive top quark loops [18], or HH [19].

Classes of integrals with one additional mass scale appearing in the propagators also have been calculated in the context of massive Bhabha scattering [20], electron-muon scattering (with  $m_e = 0, m_\mu \neq 0$ ) [21], the mixed QCD-EW corrections to the Drell-Yan process [22, 23] and three-loop corrections to the heavy flavour Wilson coefficients in DIS with two different masses [24, 25].

The integrals entering top quark pair production at NNLO [26] have been calculated numerically [27], while analytic results are only partially available [28, 29].

A major complication related to the (2-loop) integrals with massive propagators is related to the fact that the basis for an analytic representation of such integrals may go beyond the function class of generalized polylogarithms (GPLs), i.e. integrals of elliptic type occur. The latter have been subject of intense studies recently, see e.g. Refs. [13, 16, 17, 30–40].

For integrals which do not leave the class of GPLs, improvements in the understanding of the basis of multiple polylogarithms through symbol calculus and Hopf algebras (see e.g. [41, 42]) has led to a high degree of automation for these integral computations. This is a necessary step in order to apply such techniques to phenomenologically relevant cases, most notably e.g. of  $pp \rightarrow H$  at N<sup>3</sup>LO [43, 44], four-loop contributions to the cusp anomalous dimension or N<sup>3</sup>LO splitting functions [45–51]. At the multi-loop front, remarkable recent achievements are also the calculation of the five-loop QCD beta-function [52–55] and Higgs decays to hadrons and the R-ratio at N<sup>4</sup>LO [56].

There have also been developments in the direct evaluation of Feynman integrals with fewer scales, but higher loops. The `HYPERINT` [57] and `MPL` [58] packages have focused mainly on zero and one scale integrals with a high number of loops, but the algorithms employed have potential applications to a wider class of integrals. Another newly developed tool is `DREAM` [59], a program for the computation of multiloop integrals within the DRA (Dimensional Recurrence & Analyticity) method.

In order to facilitate the search for analytic results for multi-loop integrals in the literature, a new database `Loopedia` [60] has been created. At <https://loopedia.mpp.mpg.de> results for integrals can be searched for by topology. The webpage also allows to upload results for newly calculated integrals and literature information.

Direct numerical evaluation remains a powerful technique. Two contributions in these proceedings explain methods where loop and phase space integrations can be combined to cancel all poles at integrand level, such that the amplitudes can be evaluated in 4 dimensions, see Secs. I.6 and I.5.

Calculating multi-loop integrals numerically is also a promising strategy for integrals with a rather large number of kinematic scales. The sector decomposition algorithm [61] has seen a number of optimisations, implemented into the publicly available updates of the codes (PY)SECDEC [62, 63] and FIESTA [64]. For example, computations of analytically unknown 2-loop integrals entering  $pp \rightarrow HH$  [65, 66] and  $pp \rightarrow H + \text{jet}$  [67] at NLO including the full top quark mass dependence have been completed using a fully numerical approach.

### 1.2.2 Loop amplitudes and integrands

Following the analytic calculation of two-loop integrals with two massive legs, a complete set of helicity amplitudes has been obtained by two independent groups for  $pp \rightarrow VV'$  [68–71]. The results are publicly available from <http://vvamp.hepforge.org/>. Both approaches relied heavily on efficient implementations of integration-by-parts (IBP) reduction identities [72–77]. There are also new ideas how to improve multivariate functional reconstruction in the context of amplitude reduction [78–80]. Recent developments of tools in the IBP context are e.g. KIRA [81], a new program for IBP reduction identities, and FORCER, a FORM program for the reduction of four-loop massless propagator diagrams [82].

### 1.2.3 Generalised unitarity and integrand reduction

Extending the current multi-loop methods to higher multiplicity still represents a serious challenge. The increased complexity in the kinematics, and large amount of gauge redundancy in the traditional Feynman diagram approach, at one-loop has been solved numerically through on-shell and recursive off-shell methods. This breakthrough has led to the development of the now commonly used automated one-loop codes [83–89].

The  $D$ -dimensional generalised unitarity cuts algorithm [90–94] has been extended to multi-loop integrands using integrand reduction [95, 96]<sup>2</sup> and elements of computational algebraic geometry [98–106]. In contrast to the one-loop case, the basis of integrals obtained through this method is not currently known analytically and is much larger than the set of basis functions defined by standard integration-by-parts identities. The maximal unitarity method [107], which incorporates IBP identities, has been applied to a variety of two-loop examples in four dimensions [108–112]. Efficient algorithms to generate unitarity compatible IBP identities are a key ingredient in both approaches and have been the focus of on-going investigations [106, 113–117]. Very recently, automated tools for IBP reductions based on algebraic geometry have been developed, see e.g. CRISTAL and AZURITE [118–120].

Using cutting edge numerical unitarity methods, a result for the full 2-loop 4-gluon amplitude has been achieved [121]. Results for 2-loop 5-gluon amplitudes based on numerical unitarity are also advancing rapidly [122–124].

### 1.2.4 NNLO infrared subtraction methods for differential cross-sections

The construction of fully differential NNLO cross-sections for  $2 \rightarrow 2$  processes has been a major theoretical challenge over the last years. This programme has been a remarkable success with many different approaches now applied to LHC processes. We give a brief characterisation of the main methods below, as well as some of their LHC applications.

- Antenna subtraction [125, 126]:

Analytically integrated counter-terms, applicable to hadronic initial and final states. Almost completely local, requires averaging over azimuthal angles. Applied to  $e^+e^- \rightarrow$

---

<sup>2</sup>We do not attempt a complete review of integrand reduction here. Further information can be found in the review article [97] and references therein.

$3j$  [127, 128],  $pp \rightarrow j + X$  [129],  $pp \rightarrow 2j$  [2],  $pp \rightarrow Z + j$  [130, 131],  $pp \rightarrow W + j$  [132],  $pp \rightarrow H + j$  [133], dijets in DIS [134] and Higgs production in VBF [135].

- Sector Improved Residue Subtraction [136–138]:

Fully local counter-terms, based on a sector decomposition [61] approach for IR divergent real radiation [139–141] and an extension of the FKS approach at NLO [142, 143]. Numerically integrated counter-terms, for hadronic initial and final states. Recently formulated in a four-dimensional setting [144]. Applied to top-quark processes [26, 145–149] and to  $pp \rightarrow H + j$  [150, 151].

- $q_T$  [152]:

Phase-space slicing approach for colourless final states, applied to many  $pp \rightarrow VV'$  processes. An extension for  $t\bar{t}$  final states has been proposed [153]. A full list of processes is available in MATRIX [154], see also phenomenological studies [155–168].

Also applied to obtain NNLO differential results for  $VH$  [159, 169] and  $HH$  [170], as well as for  $HHW$  [171] and  $HHZ$  [172].

- $N$ -jettiness [173–175]:

Extension of the  $q_T$  method to final states including a jet, matching to soft-collinear effective theory (SCET) below the  $N$ -jettiness cut-off parameter. Applied to  $2 \rightarrow 2$  processes containing vector bosons or a boson plus one jet in the final state [174, 176–182]; for colourless final states see also MCFM version 8 [183]. Similar techniques also applied to top decay [184] and  $t$ -channel single top production [185].

- ColorFull [186]:

Fully local counter-terms extending the Catani-Seymour dipole method [187]. Analytically integrated for infrared poles, numerical integration for finite parts. Currently developed for hadronic final states such as  $H \rightarrow b\bar{b}$  [186] and  $e^+e^- \rightarrow 3$  jets [188–190].

- Nested Soft-Collinear Subtraction [191]:

Fully local subtraction terms, (partially) numerical cancellation of IR poles, allows matrix elements to be evaluated in four dimensions. Described in detail in Sec. I.4. Recently applied to  $pp \rightarrow WH$  with  $H \rightarrow b\bar{b}$  [192].

- Analytic local sector subtraction [193]:

Local subtraction, aiming at the minimal counterterm structure arising from a sector partition of the radiation phase space. Analytic integration of the counterterms. Proof of principle example from  $e^+e^- \rightarrow 2$  jets.

- Structure function approach/projection to Born [194]:

Range of applicability limited (it requires the knowledge of inclusive NNLO corrections). Applied to VBF Higgs production [194], and  $t$ -channel single top production [185].

### 1.3 The precision wish list

We break the list of precision observables into four sections: Higgs, jets, vector bosons and top quarks.

Corrections are defined with respect to the leading order, and we organise the perturbative expansion into QCD corrections, electroweak (EW) corrections and mixed  $\text{QCD} \otimes \text{EW}$ ,

$$d\sigma_X = d\sigma_X^{\text{LO}} \left( 1 + \sum_{k=1} \alpha_s^k d\sigma_X^{\delta\text{N}^k\text{LO}_{\text{QCD}}} + \sum_{k=1} \alpha^k d\sigma_X^{\delta\text{N}^k\text{LO}_{\text{EW}}} + \sum_{k,l=1} \alpha_s^k \alpha^l d\sigma_X^{\delta\text{N}^{(k,l)}\text{LO}_{\text{QCD} \otimes \text{EW}}} \right). \quad (\text{I.3})$$

We explicitly separate the mixed QCD and EW corrections to distinguish between additive predictions QCD+EW and mixed predictions QCD $\otimes$ EW. The definition above only applies in the case where the leading order process contains a unique power in each coupling constant. For example, in the case of  $q\bar{q} \rightarrow q\bar{q}Z$  two leading order processes exist: via gluon exchange of  $\mathcal{O}(\alpha_s^2\alpha)$ , via electroweak boson exchange of  $\mathcal{O}(\alpha^3)$  and the interference  $\mathcal{O}(\alpha_s\alpha^2)$ . In these cases it is customary to classify the Born process with highest power in  $\alpha_s$  (and typically the largest cross section) as the leading order, and label the others as subleading Born processes. The above classification is then understood with respect to the leading Born process, unless otherwise stated.

In the following we attempt to give a current snapshot of the available calculations of higher (fixed) order corrections in both QCD and EW theory. The main aim is to summarise computations that appeared in the 2015 wish list and that have now been completed, as well as to identify processes with a large mismatch between the (expected) experimental precision and the theoretical uncertainties<sup>3</sup>. We are aware that there are obvious difficulties in compiling such lists, which make it difficult to address every possible relevant computation. Specific approximations and/or extensions beyond fixed order are often necessary when comparing theory to data.

Following the 2015 wishlist we clarify that it is desirable to have a prediction that combines all the known corrections. For example N<sup>2</sup>LO<sub>QCD</sub>+NLO<sub>EW</sub> refers to a single code that produced differential predictions including  $\mathcal{O}(\alpha_s^2)$  and  $\mathcal{O}(\alpha)$  corrections. In most cases this is a non-trivial task and when considered in combination with decays can lead to a large number of different sub-processes.

### *Electroweak corrections*

Complete higher order corrections in the SM can quickly become technically complicated in comparison to the better known corrections in QCD. As a basic rule of thumb  $\alpha_s^2 \sim \alpha$ , so corrections at N<sup>2</sup>LO<sub>QCD</sub> and NLO<sub>EW</sub> are desirable together. Moreover, for energy scales that are large compared to the  $W$ -boson mass, EW corrections are enhanced by large logarithms (often called Sudakov logarithms). There has been progress towards a complete automation of NLO<sub>EW</sub> corrections within one-loop programs such as OPENLOOPS, GOSAM, RECOLA, MADLOOP and NLOX, which has led to the completion of many items from the 2013 list. A detailed tuned comparison of these amplitude generators for the production of the  $4\ell$  and  $2\ell 2\nu$  final states (off-shell  $ZZ$  and  $WW$  production) is presented at the level of amplitudes, and in combination with the Monte Carlo integration frameworks BBMC, MoCANLO, MUNICH, SHERPA and MADGRAPH5\_aMC@NLO at the level of integrated and differential cross sections — is presented in Sec. I.7.

### *Heavy top effective Higgs interactions and finite mass effects*

Many calculations of SM processes involving Higgs bosons use the effective gluon-Higgs couplings that arise in the  $m_t \rightarrow \infty$  limit, also called “Higgs Effective Field Theory” (HEFT). At high energy hadron colliders, gluon fusion is the most dominant production process for Higgs bosons. However, at high momentum transfers, where the top quark loops are resolved, the approximation will break down.

---

<sup>3</sup>Unfortunately, time has allowed a discussion of experimental uncertainties only for the Higgs sector and the inclusive  $W, Z$  and  $t\bar{t}$  processes. This will be rectified in future updates. Extrapolating to a data sample of 3000 fb<sup>-1</sup> can be problematic. Assuming a center-of-mass energy of 14 TeV for most of the running leads to a decrease in statistical errors by a factor of 10. We make the assumption that the systematic errors stay the same; this may be optimistic given the environment in the high luminosity LHC, so take this with a grain of salt. In almost all cases, the systematic errors will dominate over the statistical ones for this large data sample. We assume a luminosity uncertainty of 2%, as current.

For the data collected during Run II, and even more so at the HL/HE LHC, it is certainly true that they probe regions where the HEFT approximation becomes invalid and finite mass effects are important. Calculating the complete top mass dependence of such loop-induced processes at NLO is difficult since it involves two-loop integrals with several mass scales. While the analytic calculations of such integrals have seen much progress in the last two years, as reported here, the phenomenological results available so far for this class of processes mostly rely on either numerical methods or approximations. We list processes in the wishlist as  $N^k\text{LO}_{\text{HEFT}}\otimes N^l\text{LO}_{\text{QCD}}$  when re-weighting including the full top mass dependence up to order  $l$  has been performed.

### *Resummation*

We do not attempt a complete classification of all possible resummation procedures that have been considered or applied to the processes in the list. In many cases precision measurements will require additional treatment beyond fixed order, and since resummed predictions always match onto fixed order outside the divergent region it would be desirable for most predictions to be available this way. Since this is not feasible, some specific cases are highlighted in addition to the fixed order.

There are several important kinematic regions where perturbative predictions are expected to break down. Totally inclusive cross-sections often have large contributions from soft-gluon emission in which higher order logarithms can be computed analytically. The  $q_T$  and  $N$ -jettiness subtraction methods naturally match on to resummations of soft/collinear gluons, in the latter case through soft-collinear effective theory. A study using the  $q_T$  method has been applied in the case of  $pp \rightarrow ZZ$  and  $pp \rightarrow W^+W^-$  [195] where further details can be found. 0-jettiness resummations within SCET have also been considered for Higgs boson production [196], recently also extending to next-to-leading-logarithmic power corrections [197–200].

Observables with additional restrictions on jet transverse momenta can also introduce large logarithms and jet veto resummations have been studied extensively in the case of  $pp \rightarrow H$  and  $pp \rightarrow H + j$  [201–203]. More in general, the logarithmic structure of Higgs production in gluon fusion has been recently investigated in details, see e.g. [204–210].

With increasing precision of both experimental data and fixed order calculations other regions may also begin to play a role. A method for the resummation of logarithms from small jet radii has been developed e.g. in Refs. [203, 211–215]. A clear understanding of these effects is important as the most popular jet radius for physics analyses at the LHC is 0.4, a size for which resummation may start to become noticeable. These logarithms are implicitly resummed in parton shower Monte Carlos. In Sec. V.5, we continue our comparisons of fixed order and matrix element plus parton shower predictions for Higgs + jet production, started in Les Houches 2015, paying close attention to the dependence upon the jet radius.

These represent only a tiny fraction of the currently available tools and predictions with resummed logarithms. For a review the interested reader may refer to [216] and references therein.

### *Parton showering*

As in the case of resummation - we refrain from listing all processes in the wishlist to be desired with matching to a parton shower (PS).  $\text{NLO}_{\text{QCD}}+\text{PS}$  predictions are available in a largely automated way within MADGRAPH5\_aMC@NLO, SHERPA, POWHEG and HERWIG7. There have been many recent efforts in matching  $\text{N}^2\text{LO}_{\text{QCD}}$  corrections to parton showers for single boson production processes [217–220] and there are good prospects for extending these techniques to  $2 \rightarrow 2$  processes.

process	known	desired
$pp \rightarrow H$	$N^3\text{LO}_{\text{HEFT}}$ (incl.)	$N^3\text{LO}_{\text{HEFT}}$ (partial results available)
	$N^{(1,1)}\text{LO}_{\text{QCD}\otimes\text{EW}}$	$N^2\text{LO}_{\text{QCD}}$
	$N^2\text{LO}_{\text{HEFT}\otimes\text{NLO}_{\text{QCD}}}$	
$pp \rightarrow H + j$	$N^2\text{LO}_{\text{HEFT}}$	$N^2\text{LO}_{\text{HEFT}\otimes\text{NLO}_{\text{QCD}}+\text{NLO}_{\text{EW}}}$
	$\text{NLO}_{\text{QCD}}$	
$pp \rightarrow H + 2j$	$\text{NLO}_{\text{HEFT}\otimes\text{LO}_{\text{QCD}}}$	$N^2\text{LO}_{\text{HEFT}\otimes\text{NLO}_{\text{QCD}}+\text{NLO}_{\text{EW}}}$
	$N^3\text{LO}_{\text{QCD}}^{(\text{VBF}^*)}$ (incl.)	
	$N^2\text{LO}_{\text{QCD}}^{(\text{VBF}^*)}$	$N^2\text{LO}_{\text{QCD}}+\text{NLO}_{\text{EW}}^{(\text{VBF})}$
	$\text{NLO}_{\text{EW}}^{(\text{VBF})}$	
$pp \rightarrow H + 3j$	$\text{NLO}_{\text{HEFT}}$	$\text{NLO}_{\text{QCD}}+\text{NLO}_{\text{EW}}$
	$N^1\text{LO}_{\text{QCD}}^{(\text{VBF})}$	
$pp \rightarrow H + V$	$N^2\text{LO}_{\text{QCD}}+\text{NLO}_{\text{EW}}$	$\text{NLO}_{gg\rightarrow HZ}^{(t,b)}$
$pp \rightarrow HH$	$N^2\text{LO}_{\text{HEFT}\otimes\text{NLO}_{\text{QCD}}}$	$\text{NLO}_{\text{EW}}$
$pp \rightarrow H + t\bar{t}$	$\text{NLO}_{\text{QCD}}+\text{NLO}_{\text{EW}}$	
$pp \rightarrow H + t/\bar{t}$	$\text{NLO}_{\text{QCD}}$	$\text{NLO}_{\text{QCD}}+\text{NLO}_{\text{EW}}$

Table I.1: Precision wish list: Higgs boson final states.  $N^x\text{LO}_{\text{QCD}}^{(\text{VBF}^*)}$  means a calculation using the structure function approximation.

#### Decay sub-processes

The description of decay sub-processes is incomplete though we do list a few notable cases. Ideally all on-shell (factorised) decays would be available up to the order of the core process. In some cases this is potentially an insufficient approximation and full off-shell decays including background interference would be desirable, but are often prohibitive. The  $t\bar{t}$  final state is an obvious example where the off-shell decay to  $WWb\bar{b}$  at NNLO is beyond the scope of current theoretical methods.

Decays in the context of electroweak corrections are usually much more complicated. Full off-shell effects at NLO are expected to be small, but higher order corrections within factorisable contributions to the decay can be important, see e.g. Sec. V.2.

The case of vector boson pair production is particularly important given the completion of the  $N^2\text{LO}_{\text{QCD}}$  computation, and corrections are known at NLO within the double pole approximation [221] and beyond [222–226].

#### 1.4 Higgs boson associated processes

An overview of the status of Higgs boson associated processes is given in Table I.1.

*H*: Among the remarkable recent developments are first steps towards differential results for Higgs production in gluon fusion at  $N^3\text{LO}$  [227] as well as results going beyond

the threshold approximation [43, 44, 228]. Another important achievement is the recent calculation of the mixed QCD-EW corrections to this process at order  $\alpha\alpha_s^2$  in the soft gluon approximation [229, 230]. The latter corrections were found to increase the NLO QCD corrections by 5.4%.

The NNLO HEFT result has been supplemented by an expansion in  $1/m_t^n$ , and matched to a calculation of the high energy limit [231, 232].

A comprehensive phenomenological study has been presented in [233], and an updated version of the program `iHixs` [228] is available which combines the effective theory with:

- complete mass dependence at NLO including top, bottom and charm loops,
- $m_H/m_t$  corrections at NNLO,
- electro-weak corrections at NLO,
- re-scaling of the  $N^3\text{LO}_{\text{HEFT}}$  with the  $\text{LO}_{\text{QCD}}$  top loop.

Including the mixed order  $\alpha\alpha_s^2$  corrections calculated in [229, 230] should lead to a further improvement, such that the dominant uncertainties currently are expected to be PDF uncertainties and finite mass effects, estimated to be below 2%. Both of these effects can (probably) be further reduced.

The NNLO+PS computations [217, 234] have been extended to include finite top and bottom mass corrections at NLO [235].

The experimental uncertainty on the total Higgs boson cross section is currently of the order of 10-15% [236], based on a data sample of  $36 \text{ fb}^{-1}$ , and is expected to reduce to the order of 3% or less with a data sample of  $3000 \text{ fb}^{-1}$  [237]. To achieve the desired theoretical uncertainty, it may be necessary to calculate the finite mass effects to NNLO, as well as fully differential  $N^3\text{LO}_{\text{HEFT}}$  corrections.

$H + j$ : Known through to  $N^2\text{LO}_{\text{QCD}}$  in the infinite top mass limit [133, 150, 151, 176, 238]. The residual scale uncertainty in the NNLO HEFT is of the order of 5%. Very recently, this process has been calculated at NLO with full top quark mass dependence [67], based on numerical methods [62, 63], settling a longstanding question about the impact of the top quark mass dependence.

The top-bottom interference effects also have been calculated [239, 240], as well as the mass effects in the large transverse momentum expansion [241, 242]. Using high-energy resummation techniques at leading logarithmic accuracy, the Higgs boson transverse momentum spectrum with finite quark mass effects beyond the leading (fixed) order has been calculated in Ref. [206]. Parton shower predictions including finite mass effects in various approximations are also available [235, 243–245]

The full NLO calculation [67] revealed that the K-factor (NLO/LO) is fairly constant over the Higgs boson transverse momentum range above the top quark threshold when using the scale  $H_T/2$ . The full result is roughly 9% larger than in the HEFT approximation and 6% larger than in  $\text{FT}_{\text{approx}}$ , an approximation where the real radiation contains the full mass dependence, while the virtual part is calculated in the rescaled HEFT approximation. The Higgs  $p_T$  distribution in the full theory is significantly different from the one in the HEFT beyond  $p_T^H \sim 500 \text{ GeV}$ .

The current experimental uncertainty on the Higgs  $+ \geq 1$  jet differential cross section is of the order of 25-30%, dominated by the statistical error, for example the fit statistical errors for the case of the  $H \rightarrow \gamma\gamma$  analysis [246]. With a sample of

3000 fb<sup>-1</sup>, the statistical error will nominally decrease by about a factor of 10, resulting in a statistical error of the order of 2.5%. If the remaining systematic errors (dominated for the diphoton analysis by the spurious signal systematic error) remain the same, the resultant systematic error would be of the order of 9%, leading to a total error of approximately 9.5%. This is similar enough to the current theoretical uncertainty that it may motivate improvements on the  $H+j$  cross section calculation. The improvements could entail a combination of the NNLO HEFT results with the full NLO results, similar to the reweighting procedure that has been done one perturbative order lower.

$H+ \geq 2j$ : QCD corrections are an essential background to Higgs production in vector boson fusion (VBF). VBF production of a Higgs boson has recently been computed differentially to N<sup>2</sup>LO<sub>QCD</sub> accuracy [135, 194] in the “DIS” approximation. For the total cross section, results at N<sup>3</sup>LO<sub>QCD</sub> accuracy are also available [247].

In the gluon fusion channel, a detailed phenomenological study of Higgs boson production in association with up to 3 jets can be found in [248]. An assessment of the mass dependence of the various jet multiplicities has been performed in [249].

In the VBF channel, full NLO QCD corrections are available [250, 251].

EW corrections to VBF stable Higgs boson production have been calculated in Ref. [252] and are available in HAWK [253]. Complete NLO QCD+EW corrections to W<sup>+</sup>W<sup>+</sup> scattering have been calculated recently [226].

The current experimental error on the  $H+ \geq 2j$  cross section is on the order of 35% [246], again dominated by statistical errors, and again for the diphoton final state, by the fit statistical error. With the same assumptions as above, for 3000 fb<sup>-1</sup>, the statistical error will reduce to the order of 3.5%. If the systematic errors remain the same, at approximately 12%, (in this case the largest systematic error is from the jet energy scale uncertainty and the jet energy resolution uncertainty), a total uncertainty of approximately 12.5% would result, less than the current theoretical uncertainty. To achieve a theoretical uncertainty less than this value would require the calculation of  $H+ \geq 2j$  to NNLO in the HEFT.

$VH$ : Associated production of a Higgs boson with a vector boson is important to pin down the EW couplings of the Higgs, and also to access the  $H \rightarrow b\bar{b}$  coupling. First predictions at N<sup>2</sup>LO<sub>QCD</sub> have been available for some time [254, 255] and are implemented in the program `vh@nnlo` [256], where version 2 also can be used for calculations in the 2HDM and MSSM [257, 258]. NLO EW corrections have been calculated in [259–262], where Ref. [262] contains combined QCD+EW predictions including parton shower effects. Soft gluon resummation effects have been calculated and found to be small compared to the NNLO fixed order result [263]. Differential predictions at N<sup>2</sup>LO<sub>QCD</sub> for  $WH$  [169] and  $ZH$  [159] have recently been extended to include Higgs boson decays to bottom quarks [264]. The gluon initiated processes is particularly sensitive to new physics effects such as new particles in the loop or resonant additional Higgs bosons. It has been calculated at NLO in Ref. [265], where the K-factor is obtained in the limit  $m_t \rightarrow \infty$  and  $m_b = 0$ , and then used to rescale the full (one-loop) LO cross section. Threshold resummation for  $gg \rightarrow ZH$  has been calculated in Ref. [266]. Top quark mass effects at NLO, in the framework of an  $1/m_t$  expansion, have been considered in Ref. [267].

Fully-differential NNLO QCD results for associated  $VH$  production based on N-jettiness subtraction have been calculated in Ref. [180] and implemented in MCFM

including decays. Fully differential NNLO corrections to  $pp \rightarrow WH$  with  $H \rightarrow b\bar{b}$  based on nested soft-collinear subtraction [191] have been presented in [192]. A parton shower matched prediction using the MiNLO procedure in POWHEG has also been completed [220]. An implementation of dimension-six SMEFT operators related to  $VH$  production, which can be used for NLO QCD+PS accurate Monte Carlo event generation within the MG5\_aMC@NLO framework, is described in [268]. An implementation of the Higgs Pseudo-Observables framework for electroweak Higgs production in the HIGGSPO UFO model for Monte Carlo event generation at NLO in QCD is available in [269].

The total inclusive cross-section has been considered in the threshold limit at  $N^3\text{LO}_{\text{QCD}}$ , extracted from the inclusive Higgs cross-section [270].

*HH*: The  $N^2\text{LO}_{\text{QCD}}$  corrections were first computed in the infinite top mass limit [271] and have since been improved with threshold resummation to NLO+NNLL [272] and NNLO+NNLL [273]. Differential NNLO results in the  $m_t \rightarrow \infty$  limit are also available [170]. Power corrections to the NLO and NNLO cross sections in the  $m_t \rightarrow \infty$  limit have been computed [274, 275]. A complete computation at NLO including all finite top quark mass effects has been achieved using numerical methods [65, 66]. This calculation also has been matched to parton shower Monte Carlo programs [276, 277] and is publicly available within the POWHEX-BOX-V2 framework. On the analytical side, the planar two-loop integrals entering  $gg \rightarrow HH$  have been recently computed in the high-energy limit [19]. An interesting approach to reconstruct the top-quark mass dependence of the two-loop virtual amplitudes for HH production in gluon fusion (and possibly also other loop-induced processes) is presented in Ref. [278], where with Padé approximants based on the large- $m_t$  expansion of the amplitude in combination with analytic results near the top threshold leads to a result which comes very close to the full result. Very recently, top quark mass effects have been incorporated in the  $N^2\text{LO}_{\text{QCD}}$  calculation, combining one-loop double-real corrections with full top mass dependence with suitably reweighted real-virtual and double-virtual contributions evaluated in the large- $m_t$  approximation [168]. The residual uncertainty from missing  $m_t$  effects is estimated to be below 3% at 14 TeV and below 5% at 100 TeV.

*t $\bar{t}$ H*:  $N\text{LO}_{\text{EW}}$  corrections have been considered within the MadGraph5\_aMC@NLO framework [279, 280]. Moreover, NLO QCD corrections have been calculated for the process including the top quark decays [281]. Very recently, the calculation of combined NLO QCD and EW corrections to  $t\bar{t}H$  production, including top quark decays and full off-shell effects, has been achieved [282]. NLO+NNLL resummation for this process has been calculated in Ref. [283]. For results in the in the Standard Model Effective Field Theory at NLO in QCD see Ref. [284].

NLO QCD corrections to  $tH$  associated production are known [285, 286].

## 1.5 Jet final states

An overview of the status of jet final state is given in Table I.2.

*j+X*: Differential  $N^2\text{LO}_{\text{QCD}}$  corrections have been calculated in Ref. [129], scale choices have been studied in [287].

*2 jets*: The  $N^2\text{LO}_{\text{QCD}}$  corrections have been calculated in Ref. [2]. Complete NLO QCD+EW corrections are also available [288]. A remarkable result is also the calculation of 2-loop 4-gluon scattering based on numerical unitarity [121].

process	known	desired
$pp \rightarrow 2 \text{ jets}$	$\text{N}^2\text{LO}_{\text{QCD}}$	
	$\text{NLO}_{\text{QCD}}+\text{NLO}_{\text{EW}}$	
$pp \rightarrow 3 \text{ jets}$	$\text{NLO}_{\text{QCD}}$	$\text{N}^2\text{LO}_{\text{QCD}}$

Table I.2: Precision wish list: jet final states.

3 jets: A rapidly increasing number of results on 5-point two-loop amplitudes can be found in [105, 123, 124, 289–293].

## 1.6 Vector boson associated processes

The numerous decay channels for vector bosons and the possible inclusion of full off-shell corrections versus factorised decays in the narrow width approximation make vector boson processes complicated to classify. A full range of decays in the narrow width approximation would be a desirable minimum precision. In the meanwhile, for leptonic decays this goal is met for essentially all processes in the list. In terms of QCD corrections, full off-shell decays don't mean a significant complication of the respective QCD calculations and are available almost everywhere. This is no longer true for EW corrections, where leptonic decays increase the complexity of the calculation, and are thus not available for many high-multiplicity processes (involving more than four final-state particles) yet. Hadronic decays are even harder to classify because they are formally part of subleading Born contributions to processes involving jets and possibly further leptonically decaying vector bosons. Including higher-order corrections in a consistent way here will usually require full SM corrections to the complete tower of Born processes, as briefly discussed in Sec. 1.3. An overview of the status of vector boson associated processes is given in Table I.3, where leptonic decays are understood if not stated otherwise. Also  $\gamma$  induced processes become increasingly important in cases where EW corrections are highly relevant. While often included only at their leading order, first computations involving also full EW corrections to  $\gamma$ -induced channels were recently achieved.

*V*: Inclusive cross-sections and rapidity distributions in the threshold limit have been extracted from the  $pp \rightarrow V$  results [294, 295]. Parton shower matched  $\text{N}^2\text{LO}_{\text{QCD}}$  computations using both the MinLO method [296], SCET resummation [196] and via the UN<sup>2</sup>LOPS technique [219]. Completing the inclusive  $\text{N}^3\text{LO}_{\text{QCD}}$  computation beyond the threshold limit is an important step for phenomenological studies. The dominant factorisable corrections at  $\mathcal{O}(\alpha_s\alpha)$  ( $\text{N}^{(1,1)}\text{LO}_{\text{QCD}\otimes\text{EW}}$ ) are also now available [297].

The inclusive production cross section for  $W$  and  $Z$  bosons has been measured at the LHC using the leptonic decays of the vector bosons. The precision in those measurements already reached the barrier of the luminosity uncertainty  $\sim 2\%$ , which is not easy to further improve. For example, the most precise measurement of the  $W$  and  $Z$  bosons integrated fiducial cross sections is for the  $\sqrt{s} = 7$  TeV sample having  $\Delta\sigma_W/\sigma_W = 1.87\%$  and  $\Delta\sigma_Z/\sigma_Z = 1.82\%$  uncertainty, with the luminosity uncertainty ( $\sim 1.8\%$ ) accounting for most of it [298].

While the inclusive integrated cross sections have been already measured and compared fairly well with the present theoretical predictions, this is not the case for differential distributions. A key observable, both for precision studies as well as for

process	known	desired
$pp \rightarrow V$	$N^3\text{LO}_{\text{QCD}}^{(z \rightarrow 0)}$ (incl.)	
	$N^2\text{LO}_{\text{QCD}}$	$N^3\text{LO}_{\text{QCD}} + N^2\text{LO}_{\text{EW}} + N^{(1,1)}\text{LO}_{\text{QCD} \otimes \text{EW}}$
	$\text{NLO}_{\text{EW}}$	
$pp \rightarrow VV'$	$N^2\text{LO}_{\text{QCD}}$	$N^2\text{LO}_{\text{QCD}} + \text{NLO}_{\text{EW}}$
	$\text{NLO}_{\text{EW}}$	$\text{NLO}_{\text{QCD}}$ ( $gg$ channel, w/ massive loops)
	$\text{NLO}_{\text{QCD}}$ ( $gg$ channel)	
$pp \rightarrow V + j$	$N^2\text{LO}_{\text{QCD}} + \text{NLO}_{\text{EW}}$	hadronic decays
$pp \rightarrow V + 2j$	$\text{NLO}_{\text{QCD}} + \text{NLO}_{\text{EW}}$	$N^2\text{LO}_{\text{QCD}}$
	$\text{NLO}_{\text{EW}}$	
$pp \rightarrow V + b\bar{b}$	$\text{NLO}_{\text{QCD}}$	$N^2\text{LO}_{\text{QCD}} + \text{NLO}_{\text{EW}}$
$pp \rightarrow VV' + 1j$	$\text{NLO}_{\text{QCD}}$	$\text{NLO}_{\text{QCD}} + \text{NLO}_{\text{EW}}$
	$\text{NLO}_{\text{EW}}$ (w/o decays)	
$pp \rightarrow VV' + 2j$	$\text{NLO}_{\text{QCD}}$	$\text{NLO}_{\text{QCD}} + \text{NLO}_{\text{EW}}$
$pp \rightarrow W^+W^+ + 2j$	$\text{NLO}_{\text{QCD}} + \text{NLO}_{\text{EW}}$	
$pp \rightarrow VV'V''$	$\text{NLO}_{\text{QCD}}$	$\text{NLO}_{\text{QCD}} + \text{NLO}_{\text{EW}}$
	$\text{NLO}_{\text{EW}}$ (w/o decays)	
$pp \rightarrow \gamma\gamma$	$N^2\text{LO}_{\text{QCD}} + \text{NLO}_{\text{EW}}$	
$pp \rightarrow \gamma + j$	$N^2\text{LO}_{\text{QCD}} + \text{NLO}_{\text{EW}}$	
$pp \rightarrow \gamma\gamma + j$	$\text{NLO}_{\text{QCD}}$	$N^2\text{LO}_{\text{QCD}} + \text{NLO}_{\text{EW}}$

Table I.3: Precision wish list: vector boson final states.  $V = W, Z$  and  $V', V'' = W, Z, \gamma$ . Full leptonic decays are understood if not stated otherwise.

new physics searches, is the transverse momentum of the vector bosons, as well as the  $\phi^*$  variable which is also very much related with the momentum of the vector boson, without being affected by the leptons' energy scale uncertainties. For neutral Drell–Yan, those have been measured at 8 TeV with precision that is  $< 1\%$  for  $0 < p_{\text{T}} < 20$  GeV [299, 300]. At the same time, the high energy tail of the measured transverse momentum distribution is dominated by statistical uncertainties due to the sample size of the data. Refined measurements are expected both for the low and the high  $p_{\text{T}}$  part of the transverse momentum distribution in Drell–Yan. Special runs with very low pileup have been taken from the LHC, with the experiments targeting to measure with  $< 1\%$  accuracy in very fine grained bins the low  $p_{\text{T}} < 20$  GeV part of the distribution, seeking to understand the origin of the data over theory discrepancies in this part of the spectrum. For high  $p_{\text{T}}$  part of the  $p_{\text{T}}$  distribution, more experimental accuracy is also to be expected with higher luminosity data at  $\sqrt{s} = 13$  TeV. Data with high  $p_{\text{T}}$  vector bosons could be used

to study the strong coupling constant at  $N^2\text{LO}_{\text{QCD}}$  accuracy. For that, a complete  $N^3\text{LO}_{\text{QCD}}$  calculation is needed, given that at tree level the Drell-Yan production is not sensitive to the QCD coupling.

$V/\gamma + j$ : Both  $Z + j$  [130, 132, 177, 178, 301] and  $W + j$  [132, 174, 178, 179] have been completed through  $N^2\text{LO}_{\text{QCD}}$  including leptonic decays, via antenna subtraction and  $N$ -jettiness slicing. Also  $\gamma + j$  was calculated more recently through  $N^2\text{LO}_{\text{QCD}}$  in the  $N$ -jettiness slicing approach [181]. All processes of this class, and in particular their ratios were investigated in great detail in Ref. [302], combining  $N^2\text{LO}_{\text{QCD}}$  predictions with full NLO EW and leading NNLO EW effects in the Sudakov approximation, including also approximations for leading NLO  $\text{QCD} \otimes \text{EW}$  effects. Particular attention was devoted here to error estimates and correlations between the processes.

$V + \geq 2j$ : While fixed order  $\text{NLO}_{\text{QCD}}$  computations of  $V + \geq 2$  jet final states have been known for many years recent progress has been made for  $\text{NLO}_{\text{EW}}$  corrections [303] including merging and showering [304, 305].

$VV'$ : Complete  $N^2\text{LO}_{\text{QCD}}$  are now available for all vector-boson pair production processes, namely  $WW$  [160, 165],  $ZZ$  [161, 164],  $WZ$  [163, 166],  $Z\gamma$  [158, 162],  $W\gamma$  [162], using the  $q_T$  subtraction method. All these results have recently become publicly available in the MATRIX Monte Carlo framework [154]: Leptonic decays of the bosons are included throughout, consistently accounting for all resonant and non-resonant diagrams, off-shell effects and spin correlations. More recently,  $N^2\text{LO}_{\text{QCD}}$  results were achieved for  $Z\gamma$  [182] and  $ZZ$  [306] within the  $N$ -jettiness method. Leading  $N^3\text{LO}_{\text{QCD}}$  corrections, namely the  $\text{NLO}_{\text{QCD}}$  corrections to the loop-induced  $gg$  channels, became available for the neutral final states  $ZZ$  [307] and  $WW$  [308] involving full off-shell leptonic decays, based on the two-loop amplitudes of Refs. [69, 70]. More recently, also their interference effects with off-shell Higgs contributions were investigated [309, 310]. There has been good progress in calculating NLO EW corrections, and in the meanwhile all vector-boson pair production processes have been completed including full leptonic decays [222, 223, 225, 311–313]. In Ref. [314], the combined NLO  $\text{QCD} + \text{EW}$  corrections were presented for all  $2\ell 2\nu$  final states, taking into account the complete NLO corrections to the photon-induced channels for the first time. In these calculations, the recently developed automated approaches have been employed, which are validated against each other in a tuned comparison for the off-shell production of  $WW$  and  $ZZ$  pairs in Sec. I.7.

$VV' + j$ :  $\text{NLO}_{\text{QCD}}$  corrections have been known for many years. More recently,  $\text{NLO}_{\text{EW}}$  corrections became available for the some on-shell processes, with subsequent leptonic decays treated in narrow-width approximation [315, 316]. Full  $\text{NLO}_{\text{EW}}$  corrections including decays are clearly in reach of the automated tools described in Sec. I.7.

$VV' + 2j$ :  $\text{NLO}_{\text{QCD}}$  corrections have been known for several years, both in the QCD production and the EW VBS production modes. More recently, a first calculation of the full NLO SM corrections, i.e.  $\text{NLO}_{\text{QCD}}$ ,  $\text{NLO}_{\text{EW}}$  and mixed NLO corrections to all production channels, became available for  $W^+W^+ + 2j$  production including full leptonic decays [224, 226]: In particular the NLO EW corrections to the EW VBS production mode turn out to be remarkably large and negative if typical VBF kinematics are considered.

$VV'V''$ :  $\text{NLO}_{\text{QCD}}$  corrections have been known for many years. More recently,  $\text{NLO}_{\text{EW}}$  corrections became available for the on-shell processes involving three [317–321] and

process	known	desired
$pp \rightarrow t\bar{t}$	$N^2LO_{QCD}+NLO_{EW}$	
	$NLO_{QCD}$ (w/ decays, off-shell effects)	$N^2LO_{QCD}$ (w/ decays)
	$NLO_{EW}$ (w/ decays, off-shell effects)	
$pp \rightarrow t\bar{t} + j$	$NLO_{QCD}$ (w/ decays)	$N^2LO_{QCD}+NLO_{EW}$ (w/ decays)
	$NLO_{EW}$	
$pp \rightarrow t\bar{t} + 2j$	$NLO_{QCD}$ (w/ decays)	$NLO_{QCD}+NLO_{EW}$ (w/ decays)
$pp \rightarrow t\bar{t} + Z$	$NLO_{QCD}+NLO_{EW}$ (w/ decays)	
$pp \rightarrow t\bar{t} + W$	$NLO_{QCD}$	$NLO_{QCD}+NLO_{EW}$ (w/ decays)
	$NLO_{EW}$	
$pp \rightarrow t/\bar{t}$	$N^2LO_{QCD}^*$ (w/ decays)	$N^2LO_{QCD}+NLO_{EW}$ (w/ decays)

Table I.4: Precision wish list: top quark final states.  $N^2LO_{QCD}^*$  means a calculation using the structure function approximation.

two [322, 323] massive vector bosons, in some cases with their subsequent leptonic decays treated in narrow-width approximation. Given the comparable complexity to  $W^+W^+ + 2j$  production, where full NLO SM corrections could already be achieved,  $NLO_{EW}$  corrections to all these processes including full leptonic decays may be considered in reach. The processes involving two and three photons in the final state were completed at  $NLO_{QCD}$  and  $NLO_{EW}$  accuracy [324], taking into account full leptonic decays in case of  $V\gamma\gamma$  production.

$\gamma\gamma, \gamma\gamma + j$ : This process remains an important ingredient in Higgs measurements at Run II. Originally computed at  $N^2LO_{QCD}$  with  $q_T$  subtraction [325], it has recently been re-computed [326] using the  $N$ -jettiness subtraction implemented within MCFM. The  $q_T$  resummation at NNLL requested on the 2013 wish list are also now available [325]. Given the recent excitement in di-photon production a detailed understanding of these processes at high  $q_T$  will be important in the coming years. Prospects for  $N^3LO_{QCD}$  corrections remain closely connected with differential Higgs and Drell-Yan production at  $N^3LO_{QCD}$ . At high transverse momentum it may also be interesting to have  $N^2LO_{QCD}$  predictions for  $\gamma\gamma + j$ . Given that this is of equivalent complexity to  $3j$  production we add this process to the wish list.

## 1.7 Top quark associated processes

An overview of the status of top quark associated processes is given in Table I.4

$t\bar{t}$ : Fully differential predictions for  $t\bar{t}$  production at  $N^2LO_{QCD}$  are available [146, 147] are available, as well as **fastNLO** tables [327] and a study on scale choices [328]. The  $N^2LO_{QCD}$  corrections have recently been combined with NLO electroweak corrections [329]. Polarized double-virtual amplitudes for heavy-quark pair production are also available [330]. Complete  $NLO_{EW}$  corrections have been calculated in Ref. [331] for both the on-shell case and with complete off-shell effects. Electroweak correc-

$\sqrt{s}$	ATLAS	CMS	NNLO+NNLL
7 TeV	3.9%	3.6%	4.4%
8 TeV	3.6%	3.7%	4.1%
13 TeV	4.4%	5.3%	5.5%

Table I.5: Experimental uncertainty  $\Delta\sigma_{t\bar{t}}/\sigma_{t\bar{t}}$  on the inclusive  $t\bar{t}$  production cross section measurements, in the electron-muon channel at the LHC [368–371] compared to the precision of the NNLO+NNLL calculation [26, 372].

tions to multi-jet merged on-shell top quark pair production have been presented in Ref. [332].

Radiative corrections to top quark decays have been calculated in Refs. [333–335], and have been extended up to NNLO QCD [148, 184]. Resummation has been accomplished up to NNLL, together with other improvements going beyond fixed order [336–341].

NLO QCD corrections to  $W^+W^-b\bar{b}$  production with full off-shell effects have been performed in Refs. [342–345] including leptonic decays of the  $W$  bosons, and in Ref. [346] in the lepton plus jets channel. In Refs. [347, 348], NLO calculations in the 4-flavour scheme, i.e. with massive  $b$ -quarks, have been performed.

In the narrow-width-approximation (NWA) the NLO QCD calculation has been matched to a parton shower in Ref. [349] within an extension of the PowHeg [350, 351] framework, called `ttb_NLO_dec` in the POWHEG-BOX-V2. Within the Sherpa framework, NLO QCD predictions for top quark pair production with up to three jets matched to a parton shower are also available [352, 353]. A new NLO multi-jet merging algorithm relevant to top quark pair production is also available in Herwig 7.1 [354].

Based on an NLO calculation of  $W^+W^-b\bar{b}$  production combined with the Powheg framework, first results of the  $W^+W^-b\bar{b}$  calculation in the 5-flavour scheme matched to a parton shower have been presented in Ref. [355]. However, it has been noticed later that the matching of NLO matrix elements involving resonances of coloured particles to parton showers poses problems which can lead to artefacts in the top quark lineshape [356]. As a consequence, an improvement of the resonance treatment has been implemented in POWHEG-BOX-RES, called “resonance aware matching”, and combined with NLO matrix elements from OpenLoops [85], to arrive at the most complete description so far [357], based on the framework developed in Ref. [356] and the 4FNS calculation of Ref. [348]. An alternative algorithm to treat radiation from heavy quarks in the Powheg NLO+PS framework has been presented in Ref. [358].

NLO and off-shell effects in top quark mass determinations, comparing fixed order results with full off-shell effects, results based on the NWA including NLO corrections to both production and decay, and results based on NLO  $t\bar{t}$  matched to a parton shower, have been studied in Ref. [359]. Various aspects of the definition and extraction of the top quark mass have also been studied recently in Refs. [360–367].

In terms of experimental precision, the inclusive  $t\bar{t}$  production cross section has been measured by ATLAS and CMS Collaborations at  $\sqrt{s} = 7, 8$  and 13 TeV. The measurements’ uncertainty is a bit smaller than the corresponding theoretical calculations (Table I.5). Significant part of the theory uncertainty stems from PDFs and  $\alpha_s$ . For example, in the 13 TeV calculation  $\sim 4.2\%$  comes from PDFs and  $\alpha_s$ ,

while the scale uncertainty is about 3.5%. In terms of the total production cross section, the measurements agree with the theoretical predictions within the quoted uncertainties. However, a long standing problem related to the discrepancy observed in the transverse momentum distribution of the top-quarks (see e.g., [373, 374]), still misses from a complete resolution though higher order effects [147] seem to alleviate at least partially the effect. Understanding the origin of this discrepancy is important for the LHC physics programme since it affects directly (or indirectly) many physics analyses for which the  $t\bar{t}$  is a dominant source of background.

$t\bar{t}j$ : NLO QCD corrections have been calculated in [375–377] for on-shell top quarks, results matched to a parton shower are also available [378, 379]. In Ref. [380, 381] full off-shell effects have been included. Recent studies about the extraction of the top quark mass based on  $t\bar{t}$  jet production have been performed in [365, 382].

$t\bar{t}V$ : Will help to improve the constraints on anomalous EW couplings in the top quark sector during Run II. NLO<sub>QCD</sub> corrections to  $t\bar{t}Z$  including decays have been considered in Ref. [383, 384]. NLO QCD corrections to  $t\bar{t}\gamma\gamma$  production matched to a parton shower, with focus on observables which are sensitive to the polarisation of the top quarks, have been calculated in [385]. Both NLO<sub>EW</sub> and QCD corrections have been computed within the MADGRAPH5\_aMC@NLO framework [280]. Recent developments include the calculation of the NNLL corrections to the associated production of a top pair and a W boson [386] resp. a Z boson [387], and an investigation of the relative sizes of various types of QCD and EW corrections for  $t\bar{t}W$  and  $t\bar{t}t\bar{t}$  production [388].

$t/\bar{t}$ : Fully differential N<sup>2</sup>LO<sub>QCD</sub> corrections have been completed for the dominant  $t$ -channel production process, first for stable tops [149], and more recently including top-decays to NNLO accuracy, in the NWA [185]. Both the computations [149, 185] were performed in the structure function approximation. NLO QCD corrections to  $t$ -channel electroweak  $W + bj$  production, with finite top-width effects taken into account are available within MG5\_aMC@NLO [389], see also [390] for earlier work within MC@NLO. NLO QCD corrections to single top production in the in the  $t, s$  and  $tW$  channels are also available in SHERPA [391] and in POWHEG [392, 393]. Differential distributions for  $t$ -channel single top production and decay at NNLO in QCD have been presented recently in [394]. A determination of the top-quark mass from hadro-production of single top-quarks has been performed in [395], and in [396] using the Matrix Element Method at NLO QCD.

## 2 NTuples for NNLO processes <sup>4</sup>

NTuples files have proven very useful for the dissemination of NLO results for high multiplicity processes [397]. This contribution investigates the possibility of extending the use of event files for NNLO processes. As an example we investigate the dijet production at NNLO [2].

### 2.1 Introduction

A first investigation of event files for NNLO was performed in the previous Les Houches workshop [1] where the process  $e^+e^- \rightarrow 3j$  was considered at NNLO using EERAD3 [398]. The size of the event files were found to be in an acceptable region.

In this contribution we turn to the inclusive hadronic production of two jets [2]. Hadronic processes introduce a range of new complication compared to the original study:

---

<sup>4</sup> D. Maître

- the scale variation is more complex,
- we need to include information about the initial state flavours and momentum fractions.
- the infrared structure of the process is much more complex and therefore the number of subtractions is much enhanced.

The goal of this study is to ascertain whether using event files for hadronic processes is tractable from the point of view of the storage space needed to gather enough statistics.

## 2.2 Storage format

For this experiment we used a ROOT [399] file backend for storage. In the previous study the layout of the NLO nTuples described in [397]. For this study we have decided to modify the file layout from a quite specialised layout to a more general one. For a given phase-space point the form of a weight is given by:

$$\omega = \alpha_S^n \text{pdf}(x_1, id_1, \mu) \text{pdf}(x_2, id_2, \mu) \times \left( c_0 + c_1 \log(\mu^2) + c_2 \log^2(\mu^2) + \dots \right), \quad (\text{I.4})$$

where we have set the factorisation and renormalisation scale to be the same. Keeping track of both dependence separately is possible but not considered in this study. This omission is not relevant for the purpose of the study: the contribution that requires the largest storage is the double real radiation contribution, which has no logarithmic dependence on the scale.

So for each phase-space point we save the final state kinematic and flavour information, and an array of entries

$$n, x_1, x_2, id_1, id_2, j, c_j$$

corresponding to every coefficient of a logarithm in the weight in Eq. (I.4).

## 2.3 Event file sizes

The program NNLOjet [2] we used to estimate the file sizes separates the calculation in several parts:

- the born cross section (LO),
- the NLO virtual (V) and real (R) parts
- the two-loop contribution (VV)
- the squared one-loop three jet contribution (RV)
- the double real radiation contribution (RR), separated in two parts (RRa, RRb).

Each part is integrated separately and has different characteristics that impact on the storage size. Table I.6 shows the size of the storage required for each part of the calculation per event. These numbers were obtained by creating a small event file for each part. We use this estimate to extrapolate the size of the storage capacity needed for a realistic scenario.

Table I.6 shows the size writing out the weights and the momentum configurations in the order NNLOjet produces them. We can reduce the size of the event file by collecting all weights that share the same final state phase-space configuration and collect weights corresponding to the same value of  $n, x_1, x_2, id_1, id_2, j, c_j$ , the factor in size gained and the resulting estimated storage capacities needed are listed in Table I.7. From this table we can see that the order of magnitude of the storage space needed for the full process is of the order of 100 TB. This is somewhat higher than a comfortable size to work with, but not completely unmanageable.

There are optimisations that could be used to further reduce the storage requirements that we have not considered yet and are left for further studies.

part	$\alpha_s$	order	size [kB/event]	est. need [ $10^9$ events]	est. size [TB]
LO		2	0.21	10.0	1.9
V		3	3.86	5.0	18.0
R		3	8.95	5.0	41.7
VV		4	5.51	10.0	51.3
RV		4	154.09	2.0	287.0
RRa		4	463.91	5.0	2160.2
RRb		4	124.09	2.5	288.9

Table I.6: Size of the event files produced using NNLOjet for the dijet production process at NNLO.

part	$\alpha_s$	order	size [kB/event]	est. need [ $10^9$ events]	est. size [TB]	gain factor
LO		2	0.13	10.0	1.2	1.6
V		3	0.86	5.0	4.0	4.5
R		3	1.79	5.0	8.3	5.0
VV		4	1.32	10.0	12.3	4.2
RV		4	12.52	2.0	23.3	12.3
RRa		4	12.84	5.0	59.8	36.1
RRb		4	8.13	2.5	18.9	15.3

Table I.7: Size of the event files produced using NNLOjet for the dijet production process at NNLO, reordering the phase-space points and collecting weights with the same final state. The factor in storage space gained with respect to the naïve strategy is given in the last column.

## 2.4 Conclusions

In this contribution we investigated the size of the storage needed to produce event files for the dijet production at NNLO accuracy. Our current implementation forecasts storage requirements in the region of 100 TB. This requirement is quite high but not completely intractable. The code we used for the analysis is optimised for CPU consumption and not minimising the number of weights produced, there are therefore further improvement that can be considered to reduce the size requirement without affecting the accuracy, these will be considered in further studies. It should be noted that the information collected in the event files is exactly what is needed to produce fastNLO/APPLgrid/MCgrid tables [3, 400, 401] for specific histograms. NNLO event files can therefore be seen as a useful intermediate format.

## Acknowledgements

We would like to thank Nigel Glover and Juan Cruz-Martinez for providing us with the NNLOjet program and guidance for its usage.

## 3 Progress in grid techniques for the fast reproduction of QCD calculations at NNLO <sup>5</sup>

Techniques for generating interpolation grids for perturbative QCD calculations offer a fast and flexible way to reproduce those calculations with any choice of parton distribution function set or any value of the strong coupling constant. Such grids are suitable for the iterative fitting of

<sup>5</sup> D. Britzger, C. Gwenlan, A. Huss, K. Rabbertz, M. R. Sutton

parton distributions and standard model parameters, or detailed studies of scale dependence. Recent developments are briefly discussed for the APPLfast project which implements a common interface for the APPLgrid and fastNLO fast interpolation grid libraries with the NNLOJET calculation for general cross section calculations at next-to-next-to-leading order. Progress towards a more generic common lower level interface for all interaction with the fastNLO and APPLgrid backends suitable for use with other QCD calculations is also briefly discussed.

### 3.1 Introduction

The LHC performed extremely well both during Run 1 – with 7 and 8 TeV collisions – and, since 2015, during Run 2 with collisions at 13 TeV centre of mass energy. This has allowed measurements of the  $pp$  interaction cross sections over an unprecedentedly large kinematic region of many orders of magnitude in both the hard process scale, and proton momentum fraction,  $x$ . These increasingly precise measurements highlight the need for a more precise understanding of parton distributions functions (PDF) within the proton. Several important channels – such as searches for new, massive particles and Higgs production – suffer from large theoretical uncertainties dominated by PDFs, which also play a rôle in limiting the precision of various important parameters such as the  $W$ -boson mass, or strong coupling,  $\alpha_s(M_Z^2)$ . As LHC Run 2 continues, the expected increase in experimental precision in the new phase space region afforded by the 13 TeV collisions, renders it evermore important to have access to precise QCD calculations at high orders, and a correspondingly precise knowledge of the proton PDF. For a recent review, see elsewhere [402].

Higher order calculations at hadronic colliders generally require the numerical integration over the kinematics of the final state particles necessary for the cancellation of the infrared and collinear singularities. To achieve the required statistical precision, these calculations typically take many thousands of CPU hours. For an iterative fit for the proton PDF these calculations would need to be performed many hundreds, or thousands of times, once for each of the points in the PDF minimisation space, thus precluding the use of these higher order QCD calculations directly in PDF fits.

Since the turn of the current century, fast grid techniques have been developed which allow the storage of the weights of these higher order calculations on an interpolation grid. In this way, the convolution of the weights with the PDF can be performed *a posteriori*, such that the time consuming QCD calculation need be performed only once. These techniques were first used for jet production at next-to-leading order (NLO) in  $ep$  collisions at HERA [403, 404], and were then extended to jet production at NLO at the LHC with the APPLgrid [3, 405] and the fastNLO [4, 400] projects. The complete range of two-to-two processes for NLO QCD processes at the LHC became available in 2013, and in principle any NLO process can now be used with such grids.

The next-to-next-to-leading order (NNLO) corrections for inclusive electroweak boson production have been known for some time [131, 155, 406–411]. The recent, much anticipated completion of the inclusive jet [129] and dijet [2] cross sections, together with the NNLO cross section for  $Z$ +jets production [130] implemented within the unified framework of the NNLOJET calculation provides a valuable opportunity to be able to make use of these new QCD calculations in a PDF fit.

To this end, a project to implement a combined interface between the NNLOJET calculation and the fast grid technology in both the fastNLO and APPLgrid projects is underway. Known as APPLfast – a portmanteau of the APPLgrid and fastNLO project names – the intention is to provide an implementation of an interface that will allow for the production of full NNLO grids for any of the processes implemented within the NNLOJET program.

## 3.2 The APPLfast project

The APPLfast project was initiated at the QCD@LHC workshop at the end of 2015. The first promising results regarding the closure of the grids for the different contributions to the cross section have been shown previously. Here a very brief update on the project and the current status is provided together with some potential directions for future development.

The APPLfast code is written in C++ with Fortran callable sections. It is structured as a lightweight library which can be used to bridge between the NNLOJET code and the specific code for booking, and filling the grids. In line with many other unix based applications, including the APPLgrid and fastNLO packages themselves, the code is automatically configured and compiled with autotools.

Two of the major design principles for the interface to the NNLOJET code are that it should unify as much as possible the code that is required to interface with either of the underlying APPLgrid or fastNLO technologies, and that the interface itself should be as unobtrusive as possible in the NNLOJET code. In particular this last objective means that the interface should provide no additional performance overhead in terms of execution time when not filling a grid, and when filling a grid, should keep any additional overhead as low as possible. This is achieved by the use of a minimal set of hook functions that can be called from within the NNLOJET code itself and which can be left within the code with no impact on performance if the grid filling functionality is not required. This is in line with the implementation of the interfaces with other QCD calculations [412] and it is hoped that this may eventually be generalisable to simplify the production of an interface with any future calculations.

Fast interpolating grids themselves work by taking the PDF evaluated at each  $x_1, x_2, \mu_F^2$  phase space point, and generating interpolating coefficients which can be used to generate this PDF value using a smaller number of PDF evaluations at specific grid nodes. By then storing each weight from the hard process on each of the relevant grid nodes, but additionally weighted by these interpolating coefficients, summing the product of the PDF evaluated at the grid nodes with the stored weights, will result in the generation of the full cross section. This is a quadrature method the precision of which is determined predominantly by the quality of the interpolation. Typically a grid *warm up* stage is required in order to more precisely determine more optimal limits for the phase space in  $x$  and  $\mu_F^2$  of the grids. During this stage, the NNLOJET code runs in a custom mode which generates a uniform weight for each phase space point rather than the full weights from the matrix elements.

In the endeavour towards the implementation of a common interface, in line with the first principle, opportunities have been taken for the reuse of existing tools or implementation of common tools. One of the most significant aspects has been in the configuration of the grids themselves, and the mapping of the internal processes of the NNLOJET code to the different parton luminosity contributions within the grids in an efficient way, discussed briefly in the following section.

### 3.2.1 Implementation of parton luminosity contributions

In order to map between the internal processes of the NNLOJET code and the smaller space of parton-parton luminosity processes stored in the grids, both APPLgrid and fastNLO make use of the parton luminosity class from APPLgrid. In this class the different parton-parton contributions are each assigned to a specific *subprocess*. Ignoring top, in principle there can be 121 different subprocesses. However, the presence of different terms in a calculation which might have the same parton-parton input, means that the class must be able to map between the different internal processes in the NNLO code to these individual *subprocesses*. In addition, the presence of terms in the matrix elements which contain for example different CKM matrix

process	Internal NNLOJET subprocess	parton-parton composition
0	163 177 191 205 206 245 246 285 301 317 347 377 391 405 435 449 450 489 490 529 530 569 593 617 647 677 707 737	$(d, \bar{d}) + (s, \bar{s}) + (b, \bar{b})$
1	767 899 905 906 935 941 942 164 178 192 207 208 247 248 286 302 318 348 378 392 406 436 451 452 491 492 531 532 570 594 618 648 678 708 738	$(u, \bar{u}) + (c, \bar{c})$
2	165 179 193 209 210 249 250 287 303 319 349 379 393 407 437 453 454 493 494 533 534 571 595 619 649 679 709 739	$(\bar{d}, d) + (\bar{s}, s) + (\bar{b}, b)$
3	769 909 917 918 945 953 954 166 180 194 211 212 251 252 288 304 320 350 380 394 408 438 455 456 495 496 535 536 572 596 620 650 680 710 740	$(\bar{u}, u) + (\bar{c}, c)$
4	770 910 919 920 946 955 956 167 181 195 229 230 269 270 289 305 321 351 381 395 409 439 481 482 521 522 561 562 585 586 609 610 621 651 681	$(d, g) + (s, g) + (b, g)$
5	711 741 771 817 818 861 862 901 902 937 938 168 182 196 231 232 271 272 290 306 322 352 382 396 410 440 483 484 523 524 563 564 587 588 611 612 622 652 682	$(u, g) + (c, g)$
6	712 742 772 819 820 863 864 903 904 939 940 169 183 197 233 234 273 274 293 309 323 353 383 397 411 441 473 474 513 514 553 554 577 578 601 602 623 653 683	$(g, d) + (g, s) + (g, b)$
7	713 743 773 801 802 845 846 889 890 925 926 170 184 198 235 236 275 276 294 310 324 354 384 398 412 442 475 476 515 516 555 556 579 580 603 604 624 654 684	$(g, u) + (g, c)$
8	714 744 774 803 804 847 848 891 892 927 928 171 172 185 186 199 200 325 326 355 356 385 386 399 400 413 414 443 444 629 630 659 660 689 690 719 720 749 750	$(g, g)$
9	779 780 797 798 799 800 841 842 843 844 885 886 887 888 921 922 923 924 173 187 201 237 238 277 278 295 311 327 357 387 401 415 445 477 478 517 518 557 558 581 582 605 606 625 655 685	$(g, \bar{d}) + (g, \bar{s}) + (g, \bar{b})$
10	715 745 775 805 806 849 850 893 894 929 930 174 188 202 239 240 279 280 296 312 328 358 388 402 416 446 479 480 519 520 559 560 583 584 607 608 626 656 686	$(g, \bar{u}) + (g, \bar{c})$
11	716 746 776 807 808 851 852 895 896 931 932 175 189 203 241 242 281 282 297 313 329 359 389 403 417 447 485 486 525 526 565 566 589 590 613 614 627 657 687	$(\bar{d}, g) + (\bar{s}, g) + (\bar{b}, g)$
12	717 747 777 833 834 877 878 913 914 949 950 176 190 204 243 244 283 284 298 314 330 360 390 404 418 448 487 488 527 528 567 568 591 592 615 616 628 658 688	$(\bar{u}, g) + (\bar{c}, g)$
13	718 748 778 835 836 879 880 915 916 951 952 213 253 335 365 423 457 497 537 635 665 695 725 755 785 813 821 822 857 865 866	$(d, \bar{d}) + (d, \bar{s}) + (d, \bar{b}) + (s, \bar{d}) + (s, \bar{s}) + (s, \bar{b}) + (b, \bar{d}) + (b, \bar{s}) + (b, \bar{b})$
14	214 254 336 366 424 458 498 538 636 666 696 726 756 786 814 858 $(d, \bar{u}) + (d, \bar{c}) + (s, \bar{u}) + (s, \bar{c}) + (b, \bar{u}) + (b, \bar{c})$	
15	215 255 337 367 425 459 499 539 637 667 697 727 757 787 815 859 $(u, \bar{d}) + (u, \bar{s}) + (u, \bar{b}) + (c, \bar{d}) + (c, \bar{s}) + (c, \bar{b})$	
16	216 256 338 368 426 460 500 540 638 668 698 728 758 788 816 823 824 860 867 868 $(u, \bar{u}) + (u, \bar{c}) + (c, \bar{u}) + (c, \bar{c})$	
17	217 257 331 361 419 461 501 541 631 661 691 721 751 781 809 853 $(d, \bar{d}) + (d, \bar{s}) + (d, \bar{b}) + (s, \bar{d}) + (s, \bar{s}) + (s, \bar{b}) + (b, \bar{d}) + (b, \bar{s}) + (b, \bar{b})$	
18	218 258 332 362 420 462 502 542 632 662 692 722 752 782 810 854 $(d, \bar{u}) + (d, \bar{c}) + (s, \bar{u}) + (s, \bar{c}) + (b, \bar{u}) + (b, \bar{c})$	
19	219 259 333 363 421 463 503 543 633 663 693 723 753 783 811 855 $(u, \bar{d}) + (u, \bar{s}) + (u, \bar{b}) + (c, \bar{d}) + (c, \bar{s}) + (c, \bar{b})$	
20	220 260 334 364 422 464 504 544 634 664 694 724 754 784 812 856 $(u, \bar{u}) + (u, \bar{c}) + (c, \bar{u}) + (c, \bar{c})$	
21	221 261 343 373 431 465 505 545 643 673 703 733 763 793 829 873 $(\bar{d}, \bar{d}) + (\bar{d}, \bar{s}) + (\bar{d}, \bar{b}) + (\bar{s}, \bar{d}) + (\bar{s}, \bar{s}) + (\bar{s}, \bar{b}) + (\bar{b}, \bar{d}) + (\bar{b}, \bar{s}) + (\bar{b}, \bar{b})$	
22	222 262 344 374 432 466 506 546 644 674 704 734 764 794 830 874 $(\bar{d}, \bar{u}) + (\bar{d}, \bar{c}) + (\bar{s}, \bar{u}) + (\bar{s}, \bar{c}) + (\bar{b}, \bar{u}) + (\bar{b}, \bar{c})$	
23	223 263 345 375 433 467 507 547 645 675 705 735 765 795 831 875 $(\bar{u}, \bar{d}) + (\bar{u}, \bar{s}) + (\bar{u}, \bar{b}) + (\bar{c}, \bar{d}) + (\bar{c}, \bar{s}) + (\bar{c}, \bar{b})$	
24	224 264 346 376 434 468 508 548 646 676 706 736 766 796 832 876 $(\bar{u}, \bar{u}) + (\bar{u}, \bar{c}) + (\bar{c}, \bar{u}) + (\bar{c}, \bar{c})$	
25	225 265 339 369 427 469 509 549 639 669 699 729 759 789 825 837 838 869 881 882 $(\bar{d}, \bar{d}) + (\bar{d}, \bar{s}) + (\bar{d}, \bar{b}) + (\bar{s}, \bar{d}) + (\bar{s}, \bar{s}) + (\bar{s}, \bar{b}) + (\bar{b}, \bar{d}) + (\bar{b}, \bar{s}) + (\bar{b}, \bar{b})$	
26	226 266 340 370 428 470 510 550 640 670 700 730 760 790 826 870 $(\bar{d}, \bar{u}) + (\bar{d}, \bar{c}) + (\bar{s}, \bar{u}) + (\bar{s}, \bar{c}) + (\bar{b}, \bar{u}) + (\bar{b}, \bar{c})$	
27	227 267 341 371 429 471 511 551 641 671 701 731 761 791 827 871 $(\bar{u}, \bar{d}) + (\bar{u}, \bar{s}) + (\bar{u}, \bar{b}) + (\bar{c}, \bar{d}) + (\bar{c}, \bar{s}) + (\bar{c}, \bar{b})$	
28	228 268 342 372 430 472 512 552 642 672 702 732 762 792 828 839 840 872 883 884 $(\bar{u}, \bar{u}) + (\bar{u}, \bar{c}) + (\bar{c}, \bar{u}) + (\bar{c}, \bar{c})$	
29	291 307 573 597 897 933 $(d, \bar{d}) + (s, \bar{s}) + (b, \bar{b})$	
30	292 308 574 598 898 934 $(u, \bar{u}) + (c, \bar{c})$	
31	299 315 575 599 911 947 $(\bar{d}, \bar{d}) + (\bar{s}, \bar{s}) + (\bar{b}, \bar{b})$	
32	300 316 576 600 912 948 $(\bar{u}, \bar{u}) + (\bar{c}, \bar{c})$	

Fig. I.1: The NNLO contributions to the  $Z$ +jets calculation illustrating the 33 separate combinations of parton-parton input terms and their mapping to 768 separate internal NNLOJET contributions.

terms, means that for *a posteriori* setting of the CKM matrix, these different contributions should be mapped to different *subprocesses* even if their initial parton-parton contributions are the same.

The ability to map different internal NNLO calculation sub-contributions to the same, or different *subprocesses* is implemented in the class. Although separation into the different CMK contributions is not yet fully implemented, it need only be performed at the point of determining the mapping of processes in the NNLOJET initialisation, and should be considered to be straightforward.

By way of example, Fig. I.1 shows the set of parton-parton luminosities from the NNLO contribution to the  $Z$ +jets calculation, showing how the 768 internal NNLOJET processes are mapped to the separate 33 parton-parton luminosities stored in the grid.

### 3.2.2 Contributions at NNLO

The NNLOJET code itself generates the various contributions to the cross section individually, generating each of the leading order (LO), the NLO real (R), NLO virtual (V), and the NNLO double-real (RR), double-virtual (VV), and real-virtual (RV) contributions using a distinct run of the executable. The grids generated after each of these stages must then be combined to render the full cross section.

To produce a stable cross section at higher orders it is necessary to run a calculation generating a very large number of weights. This is particularly true for the double-real contribution since the large number of partons in the final state and the correspondingly more complicated infrared structure, typically requires hundreds of thousands of CPU hours. Because of this, it is necessary that many hundreds or thousands of separate jobs are required for each sub-

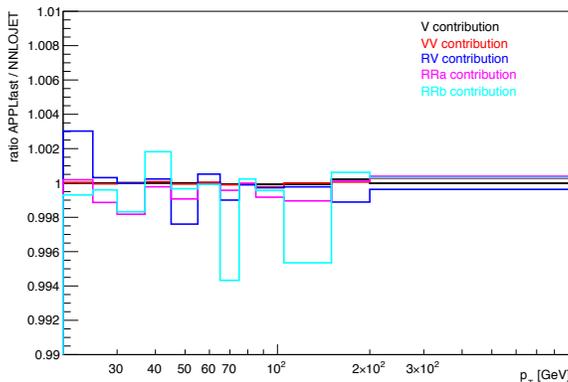


Fig. I.2: The ratio of the fast convolution from APPLfast with respect to the original NNLOJet calculation for various contributions to the  $Z$ +jets cross section at NLO and NNLO.

contribution. The resulting grids for each cross section from each job is typically  $\mathcal{O}(10 - 100)$  MBytes in size, depending on the number of bins in the observable and details of the interpolation used. Fortunately, the grids obtained by summing the output from the many thousands of separate jobs is typically of a similar size to the largest single grid.

In order to reproduce the cross section from the calculation with adequate precision, it is necessary that the grids themselves be combined using an analogous procedure to that used in the NNLOJET calculation itself. In NNLOJET a sophisticated combination procedure is used [413, 414] which allows individual cross section bins to be weighted independently from their neighbours. It was necessary to implement this functionality in the combination of the output grids before any detailed comparisons of the fast grid convolution and the NNLO cross sections could be made.

Figure I.2 shows the ratio of the fast convolution from APPLfast with respect to the original NNLOJET calculation for the NNLO contributions from the  $Z$ +jets cross section, together with the NLO virtual contribution shown for reference. In all cases the agreement is much better than 6%. Here the ratio is taken after combination of the grids following the NNLOJET combination procedure. It is worth mentioning the apparent worse performance of the ratio for the real contributions – particularly the double-real part. Although this is a quadrature method – each individual weight should be equally well stored on the grid – the precision of the convolution is affected by how well optimised the grid is after the initial warm up stage. For the real-virtual and double-real contributions shown here, a longer warmup phase would be beneficial.

Figure I.3 shows the combination in more detail specifically for the  $Z$ +jets double-virtual contribution. Note the reduced 1% maximum range of the vertical-axis. In the first panel, the ratio of the fast convolution with respect to the NNLOJET calculation is shown for two different types of combination; the first, a basic simple addition, where each grid and each bin is added with a weight of one; and the second, with the procedure from NNLOJET. In the case of the simply added grids the NNLOJET cross section has also been added with weight one. In both cases, the cross section is reproduced to within 0.4%. In the second panel, the fast convolution with both of these addition schemes is compared with the NNLOJET cross section combined with the full NNLOJET prescription. Clearly in this case, the simply added grid is seen to agree with the NNLOJET cross section only to within approximately 4%. This illustrates the correct function of the implementation the NNLOJET combination prescription when applied to the grids.

Figure I.4 shows the ratio of the the full convolution procedure for the combined cross

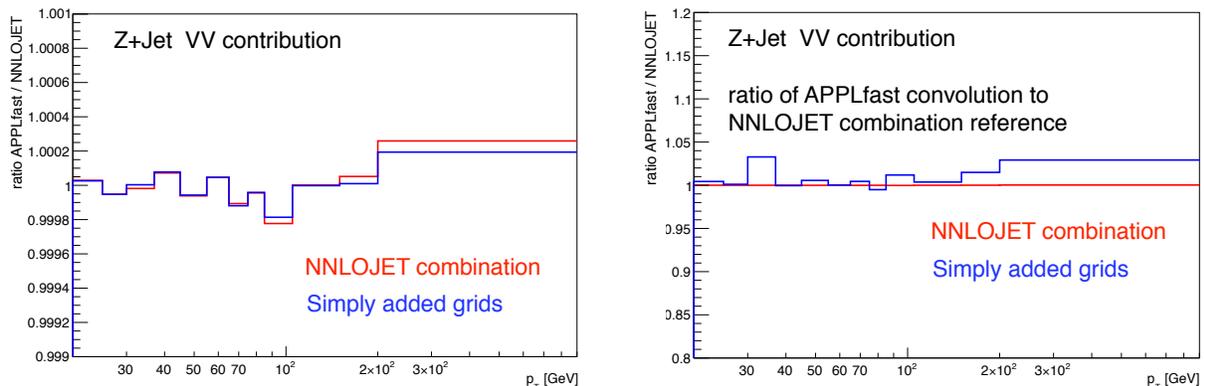


Fig. I.3: Ratios of the combined APPLfast convolution to the NNLOJET calculation for the NNLO double-virtual contribution. Different combination schemes have been used for illustration: on the left the grid convolution with either particular scheme is shown in ratio with the NNLOJET cross section, combined with the same scheme, while on the right, the different grid combinations are shown always with respect to the NNLOJET combination using the NNLOJET prescription.

section for each of the LO, NLO, and NNLO components for the inclusive jet cross section at the LHC, evaluated with the leading jet  $p_T$  scale. Note the 1% maximum vertical-axis range. Shown are the statistical uncertainties on the combined cross section and the inner shaded band represents the one per mille limit. The statistical uncertainties are larger on the NNLO contribution due to the NNLO real contributions. The fast convolution reproduces the NNLOJET cross section reliably at each of the orders, well within the statistical uncertainty on the calculation, but again, reasonably large excursions of up to 5% are observed in the NNLO contribution which have since been seen to improve with higher numbers of weights generated during the phase space warmup.

What can be concluded from the results so far is that all the major conceptual hurdles have been overcome and the primary development is reasonably complete. Currently a campaign of high statistical precision grid production is underway. This campaign is a significant undertaking given the number of different physics processes and the different phase space and binning schemes used by the different experimental collaborations.

### 3.3 Towards a standard grid interface

A combined interface capable of filling grids in either the fastNLO or APPLgrid formats, while useful, is only the beginning. A more beneficial approach would be to implement in each grid technology a common, lower level API for generic grid interactions between formats such that the bridging code or any client PDF fitting code could be fully agnostic about which format was being used.

In this way, instead of a single package with the bridging code including two distinct interfaces sharing common elements, the bridge code itself would consist of only a truly single interface interacting with the common API. Small differences would be handled within the implementation of the API for the particular grid technology being used. This should then render any change of the underlying grid technology essentially immaterial.

It is necessary to ensure some degree of backwards compatibility to ensure that existing grids which represent many thousands of hours of processing are still usable. Any such proposed new interface should then necessarily be seen as an evolving development implementation in the first instance, rather than the standard into which it will hopefully evolve.

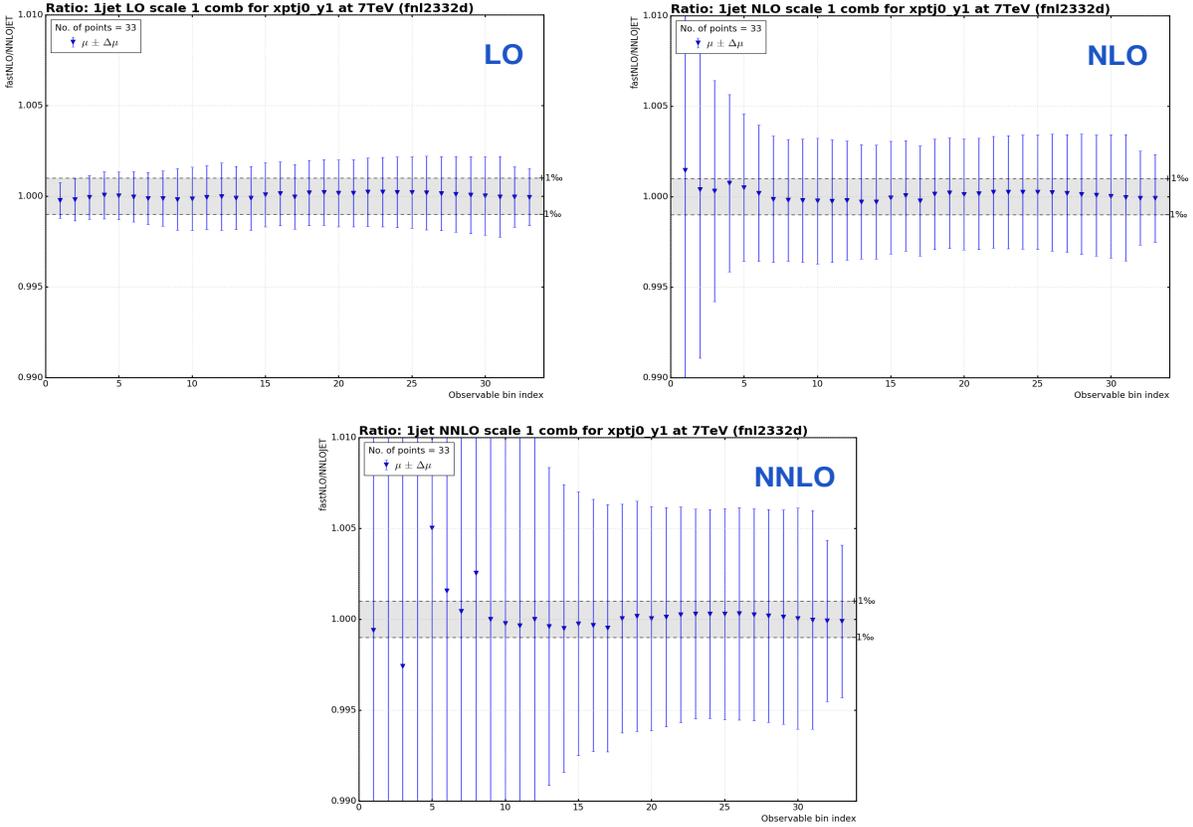


Fig. I.4: The ratio of the fast convolution from APPLfast with respect to the original NNLOJet calculation for LO, NLO and NNLO contributions from the inclusive jet cross section. The statistical error on the calculation is shown, with the 1% region shown as the shaded band.

In principle, stateless classes are to be preferred for the implementation of specific functionalities. Consequently, the interface or the filling of the grid should take all the parameters that are needed so that it can present the filling operation as an atomic action.

For a grid to be used for the fast convolution within a PDF fit, an interface for initiating the fast convolution to be called by the fitting code is also required, and it would be useful to define a separate, common interface for this purpose. This aspect of the interface will not be discussed here, except to say that it should fit into the general scheme. In this way, a number of application specific interfaces in a common namespace might be envisaged, for instance a common namespace, `lhgrid`, containing the interface to fill the grid in a specific API, `lhgrid::fillgrid_api`. Any eventual interface for the fast convolution could for instance be `lhgrid::convolute_api`. The full grid implementation would then inherit both these interfaces, or handle interaction between helper classes each of which could implement only one of the interfaces.

Figure I.5 illustrates an early implementation of the filling interface with filling methods to fill a single weight for a given phase space point for a given parton luminosity, or for filling the complete set of weights for all parton luminosities. A method to set a caching flag is also provided although not shown here.

This design has been chosen to facilitate grid filling with the different weight generation schemes implemented by the many different calculations currently interfaced but also be flexible enough to accommodate new calculations. The different schemes considered are:

```

/// production run filling interface

/// fill the grid with a single weight for a single parton luminosity
virtual int fill( const double& x1, const double& x2, const double& Q2,
                 const int&    order,
                 const int&    process_index,
                 const double& observable,
                 double weight ) = 0;

/// fill the grid with a single weight for a vector of weights - one
/// entry for each parton luminosity
virtual int fill( const double& x1, const double& x2, const double& Q2,
                 const int&    order,
                 const double& observable,
                 const std::vector<double>& weights ) = 0;

/// filling interface for the phase space warmup

/// fill the grid with a single weight for a single parton luminosity
virtual int fill_phasespace( const double& x1, const double& x2,
                             const double& Q2,
                             const int&    order,
                             const int&    process_index,
                             const double& observable,
                             double weight ) = 0;

/// fill the grid with a single weight for a vector of weights - one
/// entry for each parton luminosity
virtual int fill_phasespace( const double& x1, const double& x2,
                             const double& Q2,
                             const int&    order,
                             const double& observable,
                             const std::vector<double>& weights ) = 0;

```

Fig. I.5: Proposed interface functions for filling of a standardised grid, including a warm up stage

- i. some calculations generate the weights for all parton luminosities at the same time for each phase space point
- ii. some generate only one weight, for a single phase space point for a single parton luminosity

The interface must be able to handle each case. Both of these cases may be operated in two modes, where the code *a*) generates these weights in sequences for a number of different contributions but all with the same, or related phase points, or *b*) where each phase space point is completely distinct. For instance, many general purpose Monte Carlo generators [415], behave as in *ii*), whereas many NLO calculations [416–420] behave as in *ia*).

The opportunity for caching is then required since the grid filling operation itself is quite costly – for each phase space point, the interpolating coefficients in each of  $x_1$ ,  $x_2$  and  $\mu_F^2$  must be calculated. For case *ia*), no caching is required, but for *ib*) and particularly *ii*) it becomes

```

/// fill the reference histogram
virtual int fill_reference( const double& observable,
                           const double& weight ) = 0;

/// fill the reference histogram for a particular parton luminosity
virtual int fill_reference( const int& process_index,
                           const double& observable,
                           const double& weight ) = 0;

```

Fig. I.6: Proposed interface methods for filling of reference histograms.

very important to avoid the recalculation of the interpolating coefficients each time a weight is passed to the grid filling.

Unfortunately, the use of a cache to some extent breaks the desired stateless operation of the interface, but the filling process would still be presented as an atomic operation. There are methods for the filling of the grid during the phase space warmup stage, however these may in the end not be necessary, since the grid itself will know whether it is in a phase space warmup stage or not, and so can decide to fill with the weights or not. In principle the actual weights are not required during the phase space warmup – only the values of the  $x$ ,  $\mu_F^2$  etc. Similarly, the parton luminosity process index may not be required if the phase space is common to each parton luminosity.

To facilitate the closure testing of a grid, methods to fill a reference histogram are also provided, shown in Fig. I.6, such that the full weights, convoluted with the PDF when the grid was being filled can also be stored.

By storing in the general grid state information the identity of the PDF used in the calculation for the grid construction, it will be possible for the grid itself to perform the closure test on the quality of the implementation directly without input from the user.

### 3.4 Outlook

The prospect for creation of fast interpolation grids using the NNLOJET code is good – essentially all major hurdles have been overcome and high statistical precision grids are being produced.

An outline proposal for a common grid filling scheme that should simplify the process of creating an interface for any new calculations with the grid filling back ends has been proposed. Tests implementing this interface in development code have been made and the interface could be used in the next major release of both APPLgrid and fastNLO. Operational issues may lead to a necessary period of evolution of the interface until it is fully stable.

### Acknowledgements

The authors would like to thank the Institute of Particle Physics Phenomenology (IPPP) at Durham for the award of an IPPP Associateship to support this work, and gratefully acknowledges support from the state of Baden-Württemberg through bwHPC and the German Research Foundation (DFG) through grant no INST 39/963-1 FUGG.

## 4 The nested soft-collinear subtraction scheme <sup>6</sup>

### 4.1 Introduction

We consider NNLO subtraction schemes based on the framework of residue-improved sector decomposition, first developed in Refs. [136, 137]. This framework allows the highly successful NLO subtraction procedure of Frixione, Kunszt and Signer [142, 421] to be extended to NNLO by using sector decomposition to separate overlapping divergences. At present, it is the only approach to NNLO subtractions which is fully local in multi-particle phase space, and subtraction schemes based on this framework, such as STRIPPER [136, 137], have been used in a variety of non-trivial NNLO calculations [26, 145–147, 149–151].

There are, however, several aspects of the residue-improved sector decomposition framework which can be improved upon. First, the division of the phase space into sectors is useful to separate individual singular configurations, but it somewhat obscures the simplicity of the final result, and one expects simplifications upon recombining the sectors. Second, in the original method of Refs. [136, 137] an additional (artificial) sector was introduced to separate overlapping soft-collinear divergences which appear in Feynman diagrams but not in full amplitudes, and this further complicates the construction. As a result, the cancellation of the infrared poles is not transparent in this implementation. Initially this meant that the matrix elements had to be computed in  $d$ -dimensions, which adds to the computational difficulty, although a later development [144] allowed the computation of the matrix elements in four dimensions.

These issues can be addressed simultaneously by focusing on full matrix elements, and using the color coherence property of QCD amplitudes, which implies the decoupling of soft and collinear radiation. The soft and collinear limits may then be subtracted independently of one another, removing the need for the sector with correlated soft-collinear limits. Once this sector is discarded, there are clear simplifications on recombining sectors, which allow the integration over the unresolved phase space to be performed analytically for most of the singular limits, with one exception that requires numerical integration. This leads to an explicit (although partially numerical) cancellation of the IR poles for different kinematic structures, a compact expression for the finite integrated subtraction terms, and a straightforward method to remove singular regions, and allows all matrix elements to be evaluated in four dimensions. This approach to residue-improved sector decomposition, called the *nested soft-collinear subtraction scheme* [191], will be explained briefly in the following subsection. We refer the interested reader to Ref. [191] for a more detailed discussion.

### 4.2 Nested soft-collinear subtractions at NNLO

In the discussion of the subtraction scheme, we will focus on the hadroproduction of a color singlet final state  $V$ . This process allows us to confront most of the complexities of infrared singularities at a hadron collider, while also avoiding unnecessary complications. Our treatment is independent of the matrix element, so that the procedure may readily be adapted for any color singlet production process (e.g. Drell-Yan, diboson, associated  $VH$  production, etc.). We will focus on double-real emissions  $q\bar{q} \rightarrow V + gg$ , as this partonic configuration contains the most intricate singularity structures. We have performed similar calculations for all other partonic channels relevant for Drell-Yan production, and these will be presented in a future publication. We will, however, make some brief comments on these calculations at the end of this subsection, and show some preliminary results.

The differential cross section for the process  $q_1\bar{q}_2 \rightarrow V + g_4g_5$  is

$$2s \cdot d\sigma^{\text{RR}} = \frac{1}{2!} \int [dg_4][dg_5] F_{\text{LM}}(1, 2, 4, 5), \quad (\text{I.5})$$

---

<sup>6</sup> R. Röntsch

where  $s$  is the partonic center-of-mass energy, the phase space integration measure of the emitted gluon  $i$  is

$$[dg_i] = \frac{d^{d-1}p_i}{(2\pi)^{d-1}2E_i} \theta(E_{\max} - E_i), \quad (\text{I.6})$$

with  $E_{\max}$  an arbitrary energy parameter defined in the partonic center-of-mass frame, and

$$F_{\text{LM}}(1, 2, 4, 5) = \text{dLips}_V |\mathcal{M}(1, 2, 4, 5, V)|^2 \mathcal{F}_{\text{kin}}(1, 2, 4, 5, V). \quad (\text{I.7})$$

In the above,  $\text{dLips}_V$  is the Lorentz-invariant phase space for the colorless particles, including the momentum-conserving delta-function;  $\mathcal{M}(1, 2, 4, 5, V)$  is the matrix element for the process  $q_1 \bar{q}_2 \rightarrow V + g_4 g_5$ , and  $\mathcal{F}_{\text{kin}}$  defines an infrared-safe observable. This process has singularities which appear when either  $g_4$  or  $g_5$  becomes soft, or when either becomes collinear to either initial state parton, or when  $g_4$  and  $g_5$  become collinear to each other. A combination of these configurations may also occur, so one needs to consider different approaches to each singular limit. Thus, the key to extracting the poles is to separate these singular regions.

As mentioned above, the soft and collinear regions can be treated independently, as a consequence of color coherence (provided that one considers physical, i.e. gauge invariant and onshell, QCD amplitudes). We will therefore first regularize the soft singularities, followed by the collinear singularities. To do this, it is convenient to introduce the energy ordering  $E_4 > E_5$ . This removes the  $1/2!$  factor in Eq. (I.5), which can then be rewritten as

$$2s \cdot d\sigma^{\text{RR}} = \int [dg_4][dg_5] F_{\text{LM}}(1, 2, 4, 5) \theta(E_4 > E_5) \equiv \langle F_{\text{LM}}(1, 2, 4, 5) \rangle, \quad (\text{I.8})$$

where we have introduced the averaging sign  $\langle \dots \rangle$  to indicate integration over the final state phase space. We also introduce the soft and double-soft operators

$$S_i A = \lim_{E_i \rightarrow 0} A, \quad \mathcal{S} A = \lim_{E_4, E_5 \rightarrow 0} A, \quad \text{at fixed } E_5/E_4. \quad (\text{I.9})$$

We only need to consider the limits corresponding to  $\mathcal{S}$  and  $S_5$ ; the limit  $S_4$  does not occur due to the energy ordering. We can then write

$$\begin{aligned} \langle F_{\text{LM}}(1, 2, 4, 5) \rangle &= \langle \mathcal{S} F_{\text{LM}}(1, 2, 4, 5) \rangle + \langle S_5 (I - \mathcal{S}) F_{\text{LM}}(1, 2, 4, 5) \rangle \\ &+ \langle (I - S_5) (I - \mathcal{S}) F_{\text{LM}}(1, 2, 4, 5) \rangle. \end{aligned} \quad (\text{I.10})$$

The first term on the left-hand side in Eq. (I.10) corresponds to the double-soft limit, in which both gluons decouple completely. The second term captures the limit where  $g_5$  is soft but singularities from  $S_4$  are removed. In both terms, we can integrate over the phase space of the unresolved gluons to obtain explicit IR poles in  $1/\epsilon$ . In doing so, we note that the soft limits acting on  $F_{\text{LM}}$  remove the corresponding momentum from the momentum-conserving delta-function inside  $F_{\text{LM}}$ . However, the parameter  $E_{\max}$  in the gluonic phase space, cf. Eq. (I.6), prevents the energy integral from becoming unbounded from above. The final term of Eq. (I.10) has all soft singularities removed. All three terms, however, contain (potentially overlapping) collinear singularities, which must be disentangled and then subtracted.

We now introduce collinear and double-collinear operators

$$C_{ij} A = \lim_{\rho_{ij} \rightarrow 0} A, \quad \mathcal{C}_i A = \lim_{\rho_{4i}, \rho_{5i} \rightarrow 0} A, \quad \text{with non vanishing } \rho_{4i}/\rho_{5i}, \rho_{45}/\rho_{4i}, \rho_{45}/\rho_{5i}, \quad (\text{I.11})$$

where  $\rho_{ij} = 1 - n_i \cdot n_j$  and  $n_i$  is a unit vector that describes the direction of the momentum of the  $i$ -th particle in  $(d-1)$ -dimensional space. The task of separating the collinear singularities then proceeds in two stages. First, we partition the phase space by writing

$$w^{14,15} + w^{24,25} + w^{14,25} + w^{15,24} = 1, \quad (\text{I.12})$$

where the phase space partitions  $w$  are functions of the angles between the partons. They have the property that they vanish in various collinear limits, so that only certain limits need to be considered in each partition of the phase space. The triple-collinear partitions  $w^{i4,i5}$  (for  $i = 1, 2$ ) only have singularities corresponding to operators  $C_{4i}$ ,  $C_{5i}$  and  $C_{45}$ , while the double-collinear partitions  $w^{14,25}$  and  $w^{15,24}$  only have singularities corresponding to, respectively,  $C_{41}$  and  $C_{52}$ , and  $C_{51}$  and  $C_{42}$ .

The two double-collinear partitions are now free of overlapping singularities, but the triple-collinear partitions still have overlapping singularities that can occur, for example, in partition  $w^{14,15}$  when  $\vec{p}_1 \parallel \vec{p}_4 \parallel \vec{p}_5$ . To disentangle these, we make use of a sector decomposition based on angular ordering. For partition  $w^{i4,i5}$  (where  $i = 1, 2$ ), we write

$$\begin{aligned} 1 &= \theta\left(\eta_{5i} < \frac{\eta_{4i}}{2}\right) + \theta\left(\frac{\eta_{4i}}{2} < \eta_{5i} < \eta_{4i}\right) + \theta\left(\eta_{4i} < \frac{\eta_{5i}}{2}\right) + \theta\left(\frac{\eta_{5i}}{2} < \eta_{4i} < \eta_{5i}\right) \\ &\equiv \theta^{(a)} + \theta^{(b)} + \theta^{(c)} + \theta^{(d)}. \end{aligned} \quad (\text{I.13})$$

A parametrization of the angular phase space which allows both this decomposition and the factorization of singularities in hard amplitudes is given in Refs. [136, 137]. It is clear that the decomposition results in each sector containing only one collinear singularity: in partition  $w^{14,15}$ , for example, sectors  $a$  and  $c$  have limits  $C_{51}$  and  $C_{41}$ , respectively, while sectors  $b$  and  $d$  only have the  $C_{45}$  limit. Note that sectors  $b$  and  $d$  are not the same, as the energies of  $g_4$  and  $g_5$  are ordered.

Thus we have divided the phase space in such a way as to completely separate all the overlapping singularities. Note that since soft and collinear radiation is treated separately, the use of sectors is only required to separate overlapping *collinear* singularities, and there is no need for a sector in which the energies and emission angles of the radiated gluons vanish in a correlated manner. This should allow for additional flexibility in constructing the phase space for radiation, which we intend to explore in future work.

Using the sector decomposition described above, we can write the soft-regulated term as

$$\langle (I - S_5)(I - \mathcal{S})F_{\text{LM}}(1, 2, 4, 5) \rangle = \langle F_{\text{LM}}^{s_r c_s}(1, 2, 4, 5) \rangle + \langle F_{\text{LM}}^{s_r c_t}(1, 2, 4, 5) \rangle + \langle F_{\text{LM}}^{s_r c_r}(1, 2, 4, 5) \rangle, \quad (\text{I.14})$$

where the soft-regulated, single-collinear term  $\langle F_{\text{LM}}^{s_r c_s} \rangle$  reads

$$\begin{aligned} \langle F_{\text{LM}}^{s_r c_s} \rangle &= \sum_{(ij) \in dc} \left\langle [I - \mathcal{S}][I - S_5] \left[ C_{4i}[dg_4] + C_{5j}[dg_5] \right] w^{i4,j5} F_{\text{LM}}(1, 2, 4, 5) \right\rangle \\ &+ \sum_{i \in tc} \left\langle [I - \mathcal{S}][I - S_5] \left[ \theta^{(a)} C_{5i} + \theta^{(b)} C_{45} + \theta^{(c)} C_{4i} + \theta^{(d)} C_{45} \right] \right. \\ &\quad \left. \times [dg_4][dg_5] w^{i4,i5} F_{\text{LM}}(1, 2, 4, 5) \right\rangle, \end{aligned} \quad (\text{I.15})$$

the soft-regulated, triple-collinear terms  $\langle F_{\text{LM}}^{s_r c_t} \rangle$  reads

$$\begin{aligned} \langle F_{\text{LM}}^{s_r c_t} \rangle &= - \sum_{(ij) \in dc} \left\langle [I - \mathcal{S}][I - S_5] C_{4i} C_{5j} [dg_4][dg_5] w^{i4,j5} F_{\text{LM}}(1, 2, 4, 5) \right\rangle \\ &+ \sum_{i \in tc} \left\langle [I - \mathcal{S}][I - S_5] \left[ \theta^{(a)} \mathcal{C}_i [I - C_{5i}] + \theta^{(b)} \mathcal{C}_i [I - C_{45}] \right. \right. \\ &\quad \left. \left. + \theta^{(c)} \mathcal{C}_i [I - C_{4i}] + \theta^{(d)} \mathcal{C}_i [I - C_{45}] \right] [dg_4][dg_5] w^{i4,i5} F_{\text{LM}}(1, 2, 4, 5) \right\rangle, \end{aligned} \quad (\text{I.16})$$

and the fully regulated term  $\langle F_{\text{LM}}^{s_r c_r} \rangle$  reads

$$\begin{aligned}
\langle F_{\text{LM}}^{s_r c_r} \rangle &= \sum_{(ij) \in dc} \left\langle [I - \mathcal{S}] [I - S_5] \left[ (I - C_{5j})(I - C_{4i}) \right] \right. \\
&\quad \left. \times [dg_4][dg_5] w^{i4, j5} F_{\text{LM}}(1, 2, 4, 5) \right\rangle \\
&+ \sum_{i \in tc} \left\langle [I - \mathcal{S}] [I - S_5] \left[ \theta^{(a)} [I - \mathcal{C}_i] [I - C_{5i}] + \theta^{(b)} [I - \mathcal{C}_i] [I - C_{45}] \right. \right. \\
&\quad \left. \left. + \theta^{(c)} [I - \mathcal{C}_i] [I - C_{4i}] + \theta^{(d)} [I - \mathcal{C}_i] [I - C_{45}] \right] \right. \\
&\quad \left. \times [dg_4][dg_5] w^{i4, i5} F_{\text{LM}}(1, 2, 4, 5) \right\rangle. \tag{I.17}
\end{aligned}$$

This last term is manifestly finite, as all singularities are removed through the nested subtractions. As a result, it can be evaluated in four space-time dimensions and integrated numerically. It is the only term which contains the fully resolved double-real matrix element. The two terms describing the collinear limits, Eqs. (I.15) and (I.16), can be treated along the same lines as the soft and double-soft limits, i.e. the first two terms of Eq. (I.10). In each of these terms, the radiated gluons either decouple completely (in the case of a soft limit), or partially (in the case of a collinear limit). If a gluon decouples completely, we integrate over its angles and energies. If a gluon decouples partially, we integrate over its angles and rewrite its energy integral as a convolution with a splitting function (which is related but not identical to one of the standard Altarelli-Parisi splitting functions). As mentioned before, this integration over the unresolved gluons can be performed analytically in most cases. The exception is the triple-collinear limit in the triple-collinear partitions. For this configuration, the integration could be performed analytically with some effort, but is straightforward to do numerically, which we did.<sup>7</sup> In some cases (e.g. when we consider the  $S_5$  limit), the resulting expression has NLO-like kinematics, including leftover singular regions; in these cases, we repeat the subtraction procedure as at NLO. The result is expressions which have explicit IR poles in  $1/\epsilon$  multiplying manifestly finite  $F_{\text{LM}}$  structures of lower particle multiplicity convoluted with splitting functions. These expressions completely describe the singular structure of the double-real emission.

Having removed the soft-collinear correlated sector, it is now quite easy to see similarities between intermediate expressions for collinear singularities that appear in different partitions or sectors. By combining these expressions in a judicious manner, the results for the integrated subtraction terms can be simplified substantially. For example, restrictions on the phase space related to the parameter  $E_{\text{max}}$  in  $F_{\text{LM}}^{s_r c_s}$  can be entirely removed by combining intermediate results from the two double-collinear partitions with those from sectors  $a$  and  $c$  of the two triple-collinear partitions. As a result of this, the cancellation of the IR poles between double-real, real-virtual, and double-virtual corrections can be verified analytically to order  $1/\epsilon^2$ , and numerically at order  $1/\epsilon$ , and compact expressions are obtained for the finite part of the integrated subtraction terms.

The description presented in this subsection has focused on the partonic channel  $q\bar{q} \rightarrow V + gg$ , but we stress that similar calculations have been performed for all partonic channels contributing to Drell-Yan production. In all cases, the cancellation of the IR poles is demonstrated, and compact formulae are derived for the finite integrated subtraction terms. Indeed, the singularity structure of the remaining partonic channels is simpler than that considered

<sup>7</sup>The double-soft limit was initially treated in the same way [191], but we have since computed this limit analytically for color singlet production.

above. The full results will be presented in a forthcoming publication; here we make some brief comments about each channel:

- In the  $qg \rightarrow V + q'g$  channel, there are no divergences corresponding to the final state quark becoming soft. As a result, there is no need for energy ordering, and hence we parametrize the phase space in a slightly different way in order to take advantage of this. Moreover, there are no divergences when the final state quark is collinear to the initial state quark, and as a result, the partitioning of the phase space is done differently too.
- In the  $gg \rightarrow V + q\bar{q}$  channel, there are no soft singularities. The collinear singular structure is also simple, with only single collinear limits and no overlapping singularities. As a result, phase space partitioning is not needed, and the subtraction procedure is trivial.
- The four-quark channel  $q\bar{q}' \rightarrow V + q_1\bar{q}_2$  (including all crossings) may be divided into a singlet and non-singlet contribution. The latter contributes only to  $q\bar{q}$  and  $\bar{q}q$  channels, while the former contributes to these as well as the  $qq$  and  $\bar{q}\bar{q}$  channels. If the outgoing quarks are identical, there is an interference between the singlet and non-singlet contributions which has a triple-collinear singularity. These features make the bookkeeping for this channel quite complicated, but the actual singularities are relatively simple – for example, there are no single-soft divergences.

### 4.3 Numerical results for proof-of-concept calculation

In this section, we present some results for the process  $pp \rightarrow \gamma^* \rightarrow e^-e^+ + X$  at the 14 TeV LHC, comparing these with the analytic results of Ref. [406]. We consider lepton pairs with an invariant mass  $50 \text{ GeV} < Q < 350 \text{ GeV}$ , and use NNPDF3.0 parton distribution functions [422] with the renormalization and factorization scale  $\mu = 100 \text{ GeV}$ . We show representative results for the  $q\bar{q} \rightarrow V + gg$  partonic channel, as well as a subset of the four-quark contributions corresponding to the interference of  $s$ -channel and  $t$ -channel amplitudes labeled ‘A’ and ‘C’ in Ref. [406], which we call the  $4qAC$  contribution. The NNLO contributions  $d\sigma^{\text{NNLO}}$  for these channels are

$$d\sigma_{q\bar{q}}^{\text{NNLO}} = 14471(4) \text{ fb} \qquad d\sigma_{4qAC}^{\text{NNLO}} = -96.867(17) \text{ fb} \qquad (\text{I.18})$$

$$d\sigma_{q\bar{q}}^{\text{NNLO,analytic}} = 14470 \text{ fb} \qquad d\sigma_{4qAC}^{\text{NNLO,analytic}} = -96.866 \text{ fb.} \qquad (\text{I.19})$$

The NNLO contributions from the two channels differ by two orders of magnitude; nevertheless, the agreement between our results and the analytic results is better than one per mille. This indicates that our subtraction scheme is able to give extremely precise predictions even for numerically tiny contributions. This may appear unnecessary at first glance, but we also observe large cancellations between different partonic channels, in which case this degree of precision becomes important. The differential distributions in  $Q$  for these two NNLO contributions are shown in Fig. I.7. In both cases, the agreement is at the level of a few per mille to a few percent across several orders of magnitude. This level of agreement on the NNLO contribution indicates that this subtraction scheme is well suited to the task of providing high precision results for physical cross sections and kinematic distributions.

### Acknowledgments

I thank the organizers of the Les Houches Workshop series and the SM Working Group conveners in particular for a stimulating workshop. I am grateful for support from the German Federal Ministry for Education and Research (BMBF) under grant 05H15VKCCA.

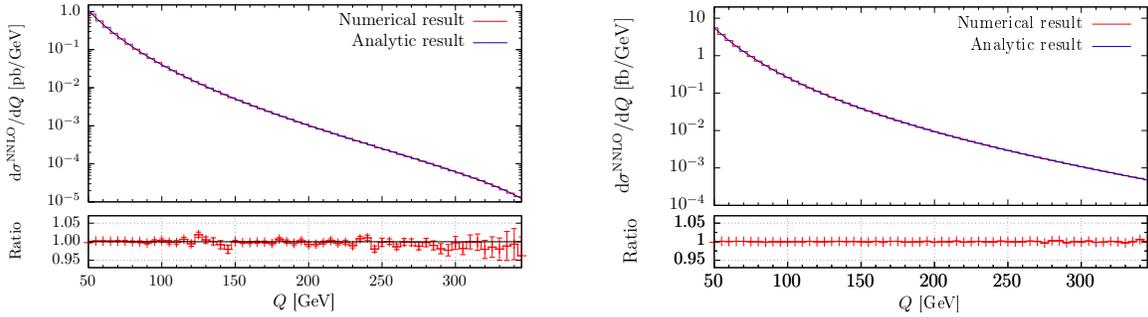


Fig. I.7: Comparison of the NNLO QCD contribution  $d\sigma^{\text{NNLO}}/dQ$  computed using the nested soft-collinear subtraction scheme with the analytic results in Ref. [406], for the  $q\bar{q} \rightarrow V + gg$  partonic channel (left), and for the  $4qAC$  contribution (right). The ratios between the results are shown in the lower panes.

## 5 Loop-tree duality and the four-dimensional unsubtraction <sup>8</sup>

In the context of four-dimensional methods for higher-order computations, we describe the main features of the four-dimensional unsubtraction (FDU) and the loop-tree duality (LTD) theorem. They grant a very powerful framework that allows to compute Feynman integrals and physical observables in four dimensions through a purely numerical implementation. Moreover, due to some remarkable mathematical properties of the intermediate expressions, the formalism is well suited for performing asymptotic expansions, circumventing many technical issues that may arise in the traditional approach. Here we will focus on the one-loop case, although we provide some insights of possible NNLO developments.

### 5.1 Introduction

The Loop-Tree Duality (LTD) formalism [423–440] turns  $N$ -leg loop integrals and amplitudes into a sum of connected tree-level-like diagrams with a remaining integration measure very similar to the  $(N + 1)$ -body phase-space [423]. Loop and tree-level radiative corrections of the same order in the perturbative expansion then, may in principle be computed numerically under a common integral sign with the use of a suitable integrator (usually a Monte Carlo routine) [433, 434]. The LTD approach fits into a wider effort in phenomenology aiming at fully automated next-to-leading order (NLO) computations. Many steps toward that direction have been taken in the recent years [84, 85, 87, 88, 415, 441–459]. Substantial progress has also been made at higher orders [460–462].

In dimensional regularization, a one-loop scalar diagram can be represented by

$$L^{(1)}(p_1, p_2, \dots, p_N) = -i \int \frac{d^d \ell}{(2\pi)^d} \prod_{i=1}^N G_F(q_i), \quad (\text{I.20})$$

where  $\ell = (\ell_0, \boldsymbol{\ell})$  is the loop momentum,  $G_F(q_i) = 1/(q_i^2 - m_i^2 + i0)$  are Feynman propagators and  $q_i$  are the momenta of the internal lines which depend on  $\ell$ . By applying the LTD, one essentially integrates over the energy component  $\ell_0$  using Cauchy’s residue theorem. The loop diagram then turns into a sum of integrals over the three-momentum  $\boldsymbol{\ell}$ , each of which is called a *dual contribution*. They emerge from the original integral after cutting one of the internal lines:

<sup>8</sup> G. Chachamis, F. Driencourt-Mangin, G. Rodrigo, G. Sborlini

$$L^{(1)}(p_1, p_2, \dots, p_N) = - \sum_{i=1}^N \text{Diagram} \quad (I.21)$$

where  $\tilde{\delta}(q_i) = 2\pi i \delta_+(q_i^2 - m_i^2)$ . The  $\tilde{\delta}(q_i)$  sets the internal lines on-shell by selecting the pole of the Feynman propagators with positive energy and negative imaginary part. Integrating the dual contributions over  $\ell$  requires most of the times a contour deformation due to the presence of the so-called *ellipsoid* and *hyperboloid* singularities [431]. These singularities are, in general, present at the integrand level.

On the other hand, the application of LTD implies modifying the propagators associated with the lines that are not being set on-shell. We define the dual propagators as

$$G_D(q_i; q_j) = \frac{1}{q_j^2 - m_j^2 - i0 \eta \cdot k_{ji}}, \quad (I.22)$$

in such a way that the original scalar integrals reads

$$L^{(1)}(p_1, p_2, \dots, p_N) = - \sum_{i=1}^N \int_{\ell} \tilde{\delta}(q_i) \prod_{j \neq i} G_D(q_i; q_j), \quad (I.23)$$

with  $i, j \in \{1, 2, \dots, N\}$ ,  $k_{ji} = q_j - q_i$  and  $\eta$  an arbitrary future-like or light-like vector,  $\eta^2 \geq 0$ , with positive definite energy  $\eta_0 > 0$ . This modified prescription is very important, since it is responsible for the cancellation of non-physical singularities manifesting in the different dual contributions. Moreover, as explained in Ref. [423], the dual prescription encodes the information contained in the different multiple cuts that appear when using the Feynman-Tree theorem (FTT) [463], thus allowing to show the equivalence between both theorems. It is worth to notice that we can chose  $\eta^\mu = (1, \mathbf{0})$  to simplify the computations; this is equivalent to applying Cauchy's residue theorem in the energy component of the loop momentum.

Assuming that there are only single powers of the Feynman propagators, the dual representation is straightforwardly valid for loop scattering amplitudes. Due to the fact that the single-cuts do not affect numerators, then the dual representation of scattering amplitudes is obtained by adding all possible *dual* single-cuts of the original loop amplitudes, and replacing the uncut Feynman propagators by dual propagators. In the case that higher powers of the propagators are present, we have to use the extended version of the LTD theorem [425], which consists in applying Cauchy's residue theorem with the well-known formula for poles of order  $n$ , i.e.

$$\text{Res}(\mathcal{A}, q_{i,0}^{(+)}) = \frac{1}{(n-1)!} \left. \frac{\partial^{n-1}}{\partial q_{i,0}^{n-1}} \left( \mathcal{A}(q_{i,0}) (q_{i,0} - q_{i,0}^{(+)})^n \right) \right|_{q_{i,0} = q_{i,0}^{(+)}} , \quad (I.24)$$

where  $q_{i,0}^{(+)} = \sqrt{\mathbf{q}_i^2 + m_i^2 - i0}$  is the positive energy solution of the corresponding on-shell dispersion relation. In that case, the explicit form of the scattering amplitude is relevant because the numerator is affected by the derivative. Specifically, this observation leads to non-trivial consequences when looking into the local form of the renormalization factors [435].

It is important to mention that the LTD theorem is valid beyond the one-loop level. For a generic  $L$ -loop integral, the theorem leads to a dual decomposition involving  $L$  cuts, in such a way that the remaining structure is topologically equivalent to real radiation processes with  $L$  additional final-state particles.

## 5.2 Computation of Feynman integrals

The numerical implementation of the LTD had been initially used to compute a huge amount of integrals with up to six external legs and the results were compared against reference values obtained with `LoopTools 2.10` [464] and `SecDec 3.0` [63]. Subsequently, results for one-loop diagrams with up to seven (heptagons) and eight (octagons) external legs were presented [431, 438]. Here we present the result for a non-trivial decagon (one-loop, ten external legs) and we comment on how the numerical implementation of the method can be optimised.

The LTD method is implemented in a C++ code [431] and relies on the `Cuba` library [465] for the numerical integration routines. At runtime, the code initially reads in and properly assigns internal masses and external momenta. Then it proceeds with an analysis of the ellipsoid and hyperboloid singularity structure to set up the details of the contour deformation, and finally performs the numerical integration using either the routine `Cuhre` or `Vegas` as are implemented in the `Cuba` library. We let our one-loop decagon take both a scalar and a tensor (rank-two) numerator. The external momenta configuration (the same for both the scalar and tensor case) is shown below:

$$\begin{aligned}
p_1 &= (-2.50000, & 0, & 0, & -2.50000) \\
p_2 &= (-2.50000, & 0, & 0, & 2.50000) \\
p_3 &= (-0.95848, & -0.38291, & 0.86652, & -0.14559) \\
p_4 &= (-0.26804, & 0.18418, & -0.17115, & -0.09288) \\
p_5 &= (-0.90712, & -0.17547, & -0.15156, & 0.87699) \\
p_6 &= (-0.79290, & 0.75301, & 0.21387, & 0.12617) \\
p_7 &= (-0.09296, & -0.02540, & -0.04121, & 0.07935) \\
p_8 &= (-0.72985, & -0.64952, & -0.24701, & 0.22314) \\
p_9 &= (-0.58078, & 0.08323, & -0.36994, & -0.43990) \\
p_{10} &= -p_1 - p_2 - p_3 - p_4 - p_5 - p_6 - p_7 - p_8 - p_9.
\end{aligned} \tag{I.25}$$

We work with internal propagators that have all different masses:

$$\begin{aligned}
m_1 &= 5.2020, m_2 = 4.2031, m_3 = 3.2042, m_4 = 7.2053, m_5 = 3.2064 \\
m_6 &= 1.2075, m_7 = 6.2086, m_8 = 8.2097, m_9 = 3.2008, m_{10} = 3.2019.
\end{aligned} \tag{I.26}$$

In Table I.8, we summarise the results for the scalar and tensor decagon. The running time in order to obtain the results on a typical Desktop machine (Intel i7 @ 3.4 GHz processor, 4-cores 8-threads), is around 30 seconds.

Propagator	Real Part	Imaginary Part
1	$2.530(4) \times 10^{-14}$	$+ i 8.514(1) \times 10^{-14}$
$\ell.p_3 \times \ell.p_5$	$8.08(4) \times 10^{-15}$	$+ i 6.144(5) \times 10^{-13}$

Table I.8: Scalar and tensor decagon with all internal masses different.

Recently, a new implementation code of the LTD was developed in `MATHEMATICA` which takes advantage of the system's internal numerical integration routines. The implementation in `MATHEMATICA` seems to present favorable features in various cases over the C++ code. However,

a thorough comparison is needed to select the strong points of both codes into a unified implementation. This will not only make the computation of one-loop diagrams much faster but it will also allow for rapid progress in attacking non-trivial two-loop diagrams.

### 5.3 Application to cross-section computations

Besides the possibility of computing Feynman integrals in an alternative way, LTD provides a robust framework to tackle cross-section computations in four space-time dimensions. As we mentioned in the Introduction, the dual contributions closely resembles PS integrals. This is a crucial point because the calculation of IR-safe observables involves adding both virtual and real corrections, with the last ones defined in a PS containing additional real particles. In consequence, if the on-shell internal lines could be interpreted as part of the real radiation, then the structure of the dual contributions might be directly related with the real-emission one.

The four-dimensional unsubtraction (FDU) framework is a fully local regularization method that can be used to compute physical observables directly in four dimensions. Since physical IR-safe quantities are associated with finite results, all the singularities appearing in intermediate steps of the calculation must be canceled in the final expression. So, the core idea of FDU is to achieve this cancellation directly at integrand level, and obtain a smooth integrable function by mapping the singular regions to the same points. This approach can be summarized in three steps:

1. Compute the dual contributions through the application of LTD to the virtual amplitudes. Locate the regions of the dual integration domain which are responsible of originating the physical IR in the loop amplitudes and demonstrate that these regions are compact.
2. Relate the real-radiation kinematics with the dual ones, forcing the internal lines that are on-shell to play the role of the real emission. Through a proper momentum mapping, describe the real and the dual contributions using the same integration variables: the combined expressions are smooth in any possible IR limit.
3. Rewrite the renormalization factors and the UV counter-terms in a local form, adjusting the sub-leading terms to reproduce the selected renormalization scheme. Adding these terms to the results obtained in the second step leads to an integrable behavior in the high-energy region.

In the following, we will center the discussion in the implementation of NLO corrections. In particular, we emphasize the importance of the momentum mapping and the local form of renormalization factors (i.e. their proper integrand-level expressions). The computation of the local UV subtraction counter-terms is obtained through an expansion around the UV propagator, as explained in Refs. [434, 435]. Finally, we conclude the section with a recompilation of benchmark results and a brief discussion about the NNLO extension.

#### 5.3.1 Real-virtual momentum mapping

Let's consider a pedagogical example: a generic decay process involving  $m$  particles in the final state. From the point of view of perturbation theory, the Born-level contribution is defined in a  $m$ -particle PS. On the other hand, the NLO corrections are obtained from the sum of the one-loop amplitudes integrated in a  $m$ -particle PS and the real-emission tree-level amplitudes in a  $(m+1)$ -particle PS. We denote the momenta associated to the Born kinematics as  $\{p_i^\mu\}_{i=1,\dots,m}$ , whilst we use  $\{p_i^{\prime\mu}\}_{i=1,\dots,m+1}$  for those momenta involved in the real radiation. Since the dual decomposition of the one-loop amplitudes includes an additional integration variable (namely the internal line that is set on-shell), both the real and the dual-virtual contributions can be

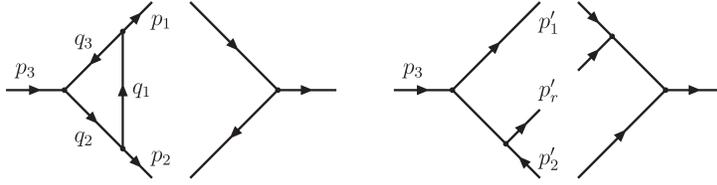


Fig. I.8: Topological correspondence among one-loop (left) and real-emission amplitudes (right). In this illustrative example, we consider a  $1 \rightarrow 2$  decay process at NLO. When  $q_1$  is set on-shell, the virtual diagram factorizes in the limit  $q_1 \parallel p_1$  exactly in the same way that the real contribution does in the limit  $p'_r \parallel p'_1$ .

written using  $(m+1)$  physical momenta. In order to relate both sets of momenta in such a way that the singularities are mapped to the same integration points, we follow the strategy applied in the dipole subtraction formalism [187, 466].

In the first place, we introduce a proper partition of the real-emission PS to isolate the IR singularities. So, we define

$$\mathcal{R}_i = \{y'_{ir} < \min y'_{jk}\}, \quad \sum_{i=1}^m \mathcal{R}_i = 1, \quad (\text{I.27})$$

where  $y'_{ij} = 2p'_i \cdot p'_j / Q^2$ ,  $r$  is the radiated parton from parton  $i$ , and  $Q$  is the typical hard scale of the scattering process. According to this definition, the only allowed collinear/soft configurations inside  $\mathcal{R}_i$  are  $i \parallel r$  or  $p'_r \rightarrow 0$ . So, different collinear singularities appear in non-overlapping regions of the real-emission PS.

When studying the virtual contribution through the LTD, we find  $m$  dual contributions, each one associated with a single cut of an internal line. It is worth appreciating that the different dual amplitudes include different IR singularities, which should be matched to those present in the real part. However, in the PS splitting defined in Eq. (I.27), the IR singularities are isolated in disjoint partitions; thus we need to identify a connection among *cuts* and *regions*. And, to infer the solution, we rely on the diagrammatic identification in the soft/collinear limit. As an explicit example, let's consider the  $1 \rightarrow 2$  decay process shown in Fig. I.8. First, we interpret the on-shell internal state in the dual amplitude as the extra-radiated particle in the real contribution, namely  $q_i \leftrightarrow p'_r$ . Then, we settle in one of the partitions, for instance  $\mathcal{R}_i$ . Because the only collinear singularity allowed is originated by  $i \parallel r$ , we distinguish particle  $i$  and call it the *emitter*. After that, we single out all the squared amplitude-level diagrams in the real contribution that become singular when  $i \parallel r$  and cut the line  $i$ . Following Fig. I.8, these have to be topologically compared with the dual-Born interference diagrams whose internal momenta  $q_i$  are on-shell. Due to the factorization properties [427, 467] and the topological analysis performed, we can guarantee that the dual contribution  $i$  and the real-emission in the region  $\mathcal{R}_i$  give rise to the same IR singular structure.

Finally, we propose an explicit connection among the dual and the real-emission momenta. Let us take the  $(m+1)$ -particle real-emission kinematics, with  $i$  as the emitter and  $r$  as the radiated particle, and we introduce a reference momentum, associated to the spectator  $j$ . Then, the momentum mapping with  $q_i = \ell + p_1 + \dots + p_i$  on-shell is given by

$$\begin{aligned} p_i^\mu &= p_i^\mu - q_i^\mu + \alpha_i p_j^\mu, & p_j^\mu &= (1 - \alpha_i) p_j^\mu, & p_k^\mu &= p_k^\mu \quad k \neq i, j, \\ p_r^\mu &= q_i^\mu, & \alpha_i &= \frac{(q_i - p_i)^2}{2p_j \cdot (q_i - p_i)}, \end{aligned} \quad (\text{I.28})$$

where all the partons are considered massless, i.e.  $p_i^2 = 0$  and  $p_i'^2 = p_j'^2 = p_r'^2 = 0$ . This construction fulfills momenta conservation, since the original Born-level kinematics also fulfills this physical constraint. Of course, this mapping can be generalized to multi-leg processes involving arbitrary masses [435].

### 5.3.2 Local expressions for renormalization counter-terms

Let's start with the well-known expression for the wave-function renormalization constant in the Feynman gauge with on-shell renormalization conditions, i.e.

$$\Delta Z_2 = \frac{g_S^2}{16\pi^2} C_F \left( -\frac{1}{\epsilon_{UV}} - \frac{2}{\epsilon_{IR}} + 3 \ln \frac{M^2}{\mu^2} - 4 \right), \quad (\text{I.29})$$

where we keep track of the IR and UV origin of the  $\epsilon$ -poles within DREG. This distinction is relevant because  $\Delta Z_2$  includes contributions originated in the calculation of self-energies, which partially cancels the IR singularities present in the squared terms of the real-emission contributions. Thus, we need to provide a proper *unintegrated* expression to locally cancel these singularities, i.e.

$$\begin{aligned} \Delta Z_2(p_1) &= -g_S^2 C_F \int_{\ell} G_F(q_1) G_F(q_3) \left( (d-2) \frac{q_1 \cdot p_2}{p_1 \cdot p_2} \right. \\ &\quad \left. + 4M^2 \left( 1 - \frac{q_1 \cdot p_2}{p_1 \cdot p_2} \right) G_F(q_3) \right), \end{aligned} \quad (\text{I.30})$$

which includes higher-order powers of the propagators [435] and reproduces the result shown in Eq. (I.29) after integration<sup>9</sup>. As we mentioned, the IR pole in  $\Delta Z_2$  will cancel when combined with the real contributions, but there are still UV divergences that need to be removed. To achieve the full regularization, we perform an expansion around the UV propagator, i.e.

$$G_F(q_{UV}) = \frac{1}{q_{UV}^2 - \mu_{UV}^2 + i0}, \quad (\text{I.31})$$

with  $\mu_{UV}$  the renormalization scale. Thus, the UV counter-term for the wave-function renormalization constant is given by

$$\begin{aligned} \Delta Z_2^{UV}(p_1) &= -(d-2) g_S^2 C_F \int_{\ell} (G_F(q_{UV}))^2 \left( 1 + \frac{q_{UV} \cdot p_2}{p_1 \cdot p_2} \right) \\ &\quad \times \left( 1 - G_F(q_{UV})(2 q_{UV} \cdot p_1 + \mu_{UV}^2) \right), \end{aligned} \quad (\text{I.32})$$

$$\equiv -\tilde{S}_{\epsilon} \frac{g_S^2}{16\pi^2} C_F \left( \frac{\mu_{UV}^2}{\mu^2} \right)^{-\epsilon} \frac{1 - \epsilon^2}{\epsilon}, \quad (\text{I.33})$$

whose integrated form exactly reproduces the UV pole present in Eq. (I.29). The finite remainders depend on the sub-leading terms proportional to  $\mu_{UV}^2$ , which can be adjusted to work in an specific renormalization scheme.

On the other hand, the vertex renormalization factors need also to be expressed in an *unintegrated* form. In the Feynman gauge, the generic vertex UV counter-term reads

$$\Gamma_{A,UV}^{(1)} = g_S^2 C_F \int_{\ell} (G_F(q_{UV}))^3 \left[ \gamma^{\nu} \not{q}_{UV} \Gamma_A^{(0)} \not{q}_{UV} \gamma_{\nu} - d_{A,UV} \mu_{UV}^2 \Gamma_A^{(0)} \right], \quad (\text{I.34})$$

<sup>9</sup>More details about this calculation can be found in Refs. [434, 435].

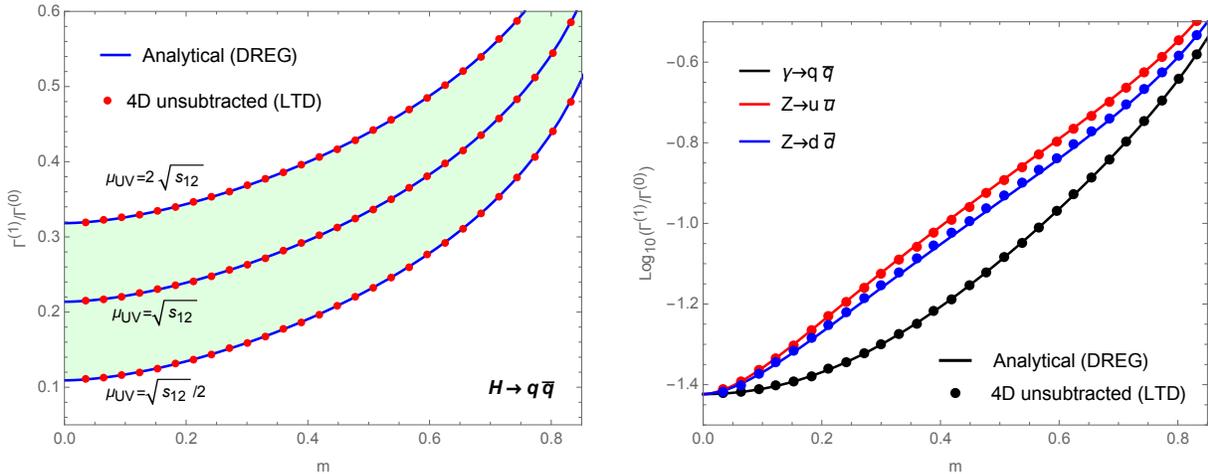


Fig. I.9: NLO QCD corrections to the decay rates  $Z, \gamma^* \rightarrow q\bar{q}$  (right) and  $H \rightarrow q\bar{q}$  (left), as a function of the normalized quark mass  $m = 2M/\sqrt{s_{12}}$ . The solid lines represent the results obtained within the DREG approach, and the colored dots were computed numerically through the application of the FDU technique. We find a complete agreement between these approaches and a smooth massless-limit transition. Moreover, the scale dependence for the Higgs decay is exactly reproduced, thanks to the introduction of local UV counter-terms.

where  $\Gamma_A^{(0)}$  represents the tree-level vertex. In the numerator, the term proportional to  $\mu_{UV}^2$  is sub-leading in the UV-limit. In this way, the coefficient  $d_{A,UV}$  can be adjusted in order to implement the desired renormalization scheme. For instance, we can tweak  $d_{A,UV}$  in order to reproduce the  $\overline{\text{MS}}$  scheme, in which the counter-term only cancels the  $\epsilon_{UV}$  pole in DREG.

It is important to emphasize that this construction of local UV-subtraction counter-terms is completely general and that the sub-leading terms can be always adjusted to reproduce the desired scheme-dependent contributions. Moreover, the tweaking of sub-leading terms in the UV-limit is *universal*, i.e. it only depends on the nature of the particles involved in the interaction (vertex or wave-function renormalization) but not on the specific process under consideration [434, 435, 457].

### 5.3.3 Selected examples

In order to have a proof of concept, we apply the FDU framework to compute the decay processes  $H \rightarrow q\bar{q}$  and  $Z, \gamma^* \rightarrow q\bar{q}$  at NLO in QCD with massive and massless quarks [434, 435]. The mentioned framework allows to obtain a combined real-virtual integrand with a non-singular behavior, i.e. numerically integrable in four space-time dimensions. The results are compared with the known expressions computed within the DREG framework. In Fig. I.9, the solid lines denote the analytical results computed in DREG as a function of  $m = 2M/\sqrt{s_{12}}$ , with  $M$  the quark mass and  $s_{12}$  the virtuality of the decaying particle. The colored dots are the values obtained through the FDU implementation.

In the first place, we emphasize that the FDU implementation is *purely numerical*. The agreement between both approaches is excellent; moreover, we can appreciate that the massless transition is smooth in both cases. Of course, the massless limit of the analytical DREG result is straightforward and leads to another analytical expression. The expressions for  $m > 0$  involve some logarithmic-enhanced terms in the real and virtual contributions, separately. Within DREG, these logarithms transform into  $\epsilon$ -poles when considering the limit  $m \rightarrow 0$ . However,

the FDU implementation manages to overcome the  $m \rightarrow 0$  limit in a surprisingly smooth way, as a consequence of the local regularization of the integrand and the smoothness of the real-virtual mapping in the massless limit. Finally, it is worth noticing that the renormalization scale dependence is successfully reproduced with the four-dimensional framework (left panel of Fig. I.9).

### 5.3.4 Towards a fully local NNLO framework

Previously, we have explicitly mentioned the details behind the implementation of FDU at NLO. However, it could be extended to deal with NNLO computations as well. Let's remember that the total NNLO cross-section is generically written as

$$\sigma^{\text{NNLO}} = \int_m d\sigma_{\text{VV}}^{(2)} + \int_{m+1} d\sigma_{\text{VR}}^{(2)} + \int_{m+2} d\sigma_{\text{RR}}^{(2)}, \quad (\text{I.35})$$

where the double virtual cross-section  $d\sigma_{\text{VV}}^{(2)}$  contains the interference of the two-loop with the Born scattering amplitudes and the square of the one-loop scattering amplitude with  $m$  final-state particles; the real-virtual cross-section  $d\sigma_{\text{VR}}^{(2)}$  includes the contributions from the interference of one-loop and tree-level scattering amplitudes with one extra external particle; and the double real cross-section  $d\sigma_{\text{RR}}^{(2)}$  are tree-level contributions integrated in a  $m + 2$  PS. The LTD representation of the two-loop scattering amplitude is obtained by setting two internal lines on-shell [424], whilst the squared one-loop introduces two independent loop three-momenta. In both contributions, the PS integration involves  $m$  external particles. On the other hand, the dual representation of the real-virtual contribution  $d\sigma_{\text{VR}}^{(2)}$  includes an additional particle from the real radiation and one independent loop-three momentum from the application of LTD. The double-real terms, contained within  $d\sigma_{\text{RR}}$ , directly involve two additional real-particles (that, however, are constrained by momentum conservation). In consequence, since the dual representation of the individual terms appearing in Eq. (I.35) always include two additional three-vectors to be integrated, we can find a proper momentum mapping to displace the different IR singularities to the same integration points. In this way, also a local cancellation will take place.

## 5.4 Asymptotic expansions of dual contributions

As stated before, the LTD theorem reduces the  $d$ -dimensional integration domain with a Minkowski metric into a  $(d - 1)$ -dimensional one with an Euclidian metric. This is another remarkable property of the FDU formalism because it allows to perform, *at integrand level*, naive and straightforward asymptotic expansions in any ratio of scales present in the process, without having to deal with most difficulties the traditional approach may encounter. This is a very interesting property because for processes involving many scales, and given a specific kinematical configuration, it is possible to perform the appropriate expansion, and consider and integrate only relevant terms, whose expression is much less complicated than the full integral. This is possible because the sum of all integrated terms of the expansion is the same as the expansion of the full integral, i.e. there is a perfect commutativity between integrating and expanding.

As an example, let's consider the two-point scalar process without applying LTD, and perform a naive expansion

$$\int_\ell \frac{1}{(\ell^2 - M^2 + i0)((\ell + p)^2 - M^2 + i0)} = \int_\ell \frac{1}{(\ell - M^2 + i0)^2} \left( 1 + \frac{2\ell \cdot p + p^2}{\ell - M^2 + i0} + \dots \right), \quad (\text{I.36})$$

for  $0 < p^2 \ll M^2$  and  $p_0 > 0$ . The issue here is that this expansion is not valid when  $\ell^2 \approx M^2$  because of the presence of  $p^2$  in the second Feynman propagator, therefore in the traditional approach, an additional step is required.

On the other hand, applying LTD and considering the first cut, the expansion

$$-\int_{\ell} \frac{\tilde{\delta}(\ell)}{2\ell \cdot p + p^2 - i0} = -\int_{\ell} \frac{\tilde{\delta}(\ell)}{2\ell \cdot p} \sum_{n=0}^{\infty} \left( \frac{-p^2}{2\ell \cdot p} \right)^n \quad (\text{I.37})$$

becomes valid for any value of  $\ell$ , because  $0 < 2\ell \cdot p = \mathcal{O}(M)$ , thanks to the presence of  $\tilde{\delta}(\ell)$ .

In Ref. [436], the Higgs boson production through gluon fusion and the Higgs decay to two photons processes have been computed using LTD. Asymptotic expressions have been obtained for the large and small mass limit of the particle inside the loop, reproducing well-known results in the literature, while showing the efficiency of this method.

## 5.5 Conclusions

The loop-tree duality is a very useful technique to decompose loop into tree-level amplitudes, with the additional property that the *dual* contributions are evaluated in an Euclidean space. In this article we discuss feasible applications that point towards the direction of skipping DREG and achieve a fully numerical implementation of higher-order computations in four space-time dimensions.

In the first place, we use this formalism to compute Feynman integrals in a more efficient way. By decomposing the loop integration into a sum of dual contributions we manage to simplify the treatment of intermediate expressions, since we avoid the presence of Gram determinants. We have tested this approach with one-loop multi-leg tensorial integrals, and compared the results with those provided by the traditional algorithms.

On the other hand, LTD allows to interpret the on-shell internal particles in each cut as a real external particle being radiated from the Born-level process. So, the formalism is well suited for performing an integrand-level combination of the real and virtual contributions. Moreover, through the introduction of a physically-motivated momentum mapping, the IR singular regions present in each term can be routed to the same integration points. This leads to a completely local cancellation of IR singularities, thus avoiding the introduction of IR counter-terms. Also, a local regularization of UV divergences is possible by performing an expansion around the UV propagator and calculating the dual expressions. The combination of all these ingredients leads to a four-dimensional framework, the FDU, which was successfully tested with some simple processes ( $H \rightarrow q\bar{q}$  and  $A^\mu \rightarrow q\bar{q}$  at NLO).

A spin-off of the local regularization properties of LTD is the possibility of simplifying asymptotic expansions. Since the dual contributions are defined in an Euclidean space, and all the IR/UV singularities have been locally regularized (thus transforming the integrands into actual *integrable* functions), the commutation among integration and series expansions is fulfilled. In this way, the formalism is expected to be more efficient than the traditional expansion-by-regions approach.

Hence, in the context of four-dimensional methods, the LTD/FDU approach provides many advantages for the implementation of physical computations at higher orders. Besides that, it could also shed light into deeper mathematical structures hidden behind the presence of IR/UV singularities.

## Acknowledgements

This work is supported by the Spanish Government and ERDF funds from European Commission (Grants No. FPA2014-53631-C2-1-P and SEV-2014-0398), by Generalitat Valenciana (Grant No. PROMETEO/2017/057), and by Consejo Superior de Investigaciones Científicas (Grant No. PIE-201750E021). GC acknowledges support from the Spanish Government grants

FPA2015-65480-P, FPA2016-78022-P and the Spanish MINECO Centro de Excelencia Severo Ochoa Programme (SEV-2016-0597). FDM acknowledges support from Generalitat Valenciana (GRISOLIA/2015/035) and GS from Fondazione Cariplo under the Grant No. 2015-0761.

## 6 Loop amplitudes: The numerical approach <sup>10</sup>

### 6.1 Introduction

Numerical methods enjoy a great flexibility. In most precision calculations they are used at least partially. A first example is the phase space integration through Monte Carlo integration. This provides the flexibility for arbitrary infrared-safe observables. A second example is the numerical computation of tree amplitudes through Berends-Giele recurrence relations. The approach based on recurrence relations is very efficient: The required CPU time scales polynomially like  $n^3$  with the number of external particles of the amplitude. The numerical approach is therefore the preferred method for processes with a large number of particles, starting already at moderate values of  $n$ .

Let us now look at higher-order corrections. Here we face the occurrence of divergences in intermediate stages of the calculation. It is common practice in NLO computations to use the subtraction method for the real part. This allows to treat the real part numerically. What about the virtual part? In the virtual part one has to supplement the subtraction method with an algorithm for contour deformation in loop momentum space [441–444, 446, 449, 451, 456, 457, 462, 468–472]. The contour deformation avoids regions, where individual loop propagators become singular, but Feynman’s  $i\delta$ -prescriptions allows a deformation into the complex plane. Related approaches are discussed in [423, 425, 427, 431, 433, 434, 458, 473–476]. In a condensed notation we have

$$\int_{n+1} d\sigma^{\text{R}} + \int_n d\sigma^{\text{V}} = \underbrace{\int_{n+1} (d\sigma^{\text{R}} - d\sigma_{\text{R}}^{\text{A}})}_{\text{convergent}} + \underbrace{\int_n (\mathbf{I} + \mathbf{L}) \otimes d\sigma^{\text{B}}}_{\text{finite}} + \underbrace{\int_{n+\text{loop}} (d\sigma^{\text{V}} - d\sigma_{\text{V}}^{\text{A}})}_{\text{convergent}}. \quad (\text{I.38})$$

The subtracted real part and the subtracted virtual part are integrable in four space-time dimensions. These integrations can be performed by Monte Carlo techniques. For a process with  $n$  final state particles at Born level, the integration for the subtracted real part is over the phase space of  $(n+1)$  final state particles, the integration for the subtracted virtual part is over the phase space of  $n$  final state particles and a four dimensional loop momentum space. The combination  $(\mathbf{I} + \mathbf{L})$  contains the subtraction terms added back. At NLO, all subtraction terms are rather simple and the integration over the unresolved phase space (for the real subtraction terms) and the integration over loop momentum space (for the virtual subtraction terms) can be performed analytically within dimensional regularisation. In the combination  $(\mathbf{I} + \mathbf{L})$  all explicit poles in the dimensional regularisation parameter  $\varepsilon$  cancel.

### 6.2 Cancellations at the integrand level

We are interested in extending the numerical approach towards NNLO. We will need the integrals over the subtraction terms. While at NLO the analytic integration of the subtraction terms is rather simple, this is no longer the case at NNLO. Since the sum of all subtraction terms is finite, we may ask if a numerical approach for the sum of these integrals is feasible. We therefore ask if a cancellation at the integrand level is possible. Since the individual subtraction terms live on different spaces, a few technical difficulties have to be addressed. It is therefore best, to study this issue at NLO first [457]. In this report we therefore focus on the sum of subtraction terms

<sup>10</sup> S. Seth, S. Weinzierl

from the real and the virtual part at NLO. We are interested in the integral

$$\int_n (\mathbf{I} + \mathbf{L}) = \int_n \left[ \int_1 d\sigma_R^A + \int_{\text{loop}} d\sigma_V^A + d\sigma_{\text{CT}}^V + d\sigma^C \right], \quad (\text{I.39})$$

where  $d\sigma_R^A$  is related to the real subtraction terms,  $d\sigma_V^A$  to the virtual subtraction terms,  $d\sigma_{\text{CT}}^V$  to the ultraviolet counterterm from renormalisation and  $d\sigma^C$  to the collinear counterterm from factorisation for initial-state partons. The first problem we face is that the unresolved phase space is  $(D - 1)$ -dimensional, while the loop momentum space is  $D$ -dimensional. This problem can be overcome with the help of the loop-tree duality [423, 425, 427, 431, 433, 434]: A cyclic-ordered one-loop amplitude

$$A_n = \int \frac{d^D k}{(2\pi)^D} \frac{P(k)}{\prod_{j=1}^n (k_j^2 - m_j^2 + i\delta)} \quad (\text{I.40})$$

can be written with Cauchy's theorem as a sum of  $n$  integrations over  $(D - 1)$ -dimensional forward hyperboloids

$$A_n = -i \sum_{i=1}^n \int \frac{d^{D-1} k}{(2\pi)^{D-1}} \frac{P(k)}{2k_i^0 \prod_{\substack{j=1 \\ j \neq i}}^n [k_j^2 - m_j^2 - i\delta (k_j^0 - k_i^0)]} \Bigg|_{k_i^0 = \sqrt{\vec{k}_i^2 + m_i^2}}. \quad (\text{I.41})$$

There exists an analogous formula for backward hyperboloids. The modified  $i\delta$ -prescription should be noted. This reduces the integration over  $D$ -dimensional loop momentum space to an integration in an  $(D - 1)$  dimensional space. In the next step we relate points in this space with points in real unresolved phase space: Given a set  $\{p_1, p_2, \dots, p_n\}$  of external momenta and an on-shell loop momentum  $k$  there is an invertible map

$$\{p_1, p_2, \dots, p_n\} \times \{k\} \rightarrow \{p'_1, p'_2, \dots, p'_n, p'_{n+1}\}. \quad (\text{I.42})$$

Note that taking the inverse map and projecting on the  $\{p_1, p_2, \dots, p_n\}$ -subspace

$$\{p'_1, p'_2, \dots, p'_n, p'_{n+1}\} \rightarrow \{p_1, p_2, \dots, p_n\} \quad (\text{I.43})$$

is the standard Catani-Seymour projection. The map in Eq. (I.42) singles out an emitter  $i$  and a spectator  $k$ . It therefore meshes well with dipole subtraction. For each dipole we have a map which maps all singular regions of this dipole onto each other.

Equipped with this map we may therefore map the singularities from the real subtraction terms onto the singularities of the virtual subtraction terms. Do they cancel locally? Not yet, there is a problem with the collinear singularities. Let us first focus on final state collinear singularities. The problem is easily understood in physical terms: Let us first consider the collinear singularities in the real subtraction terms. Since it is a real process, both partons have transverse polarisations. Aside from the collinear singularities in the splittings  $g \rightarrow gg$  and  $q \rightarrow qg$  there is also a collinear divergence in the splitting  $g \rightarrow q\bar{q}$ . Now let us look at the collinear singularities in the virtual subtraction terms. Here one finds that in the collinear limit one parton has a longitudinal polarisation. Furthermore there is no divergence in the splitting  $g \rightarrow q\bar{q}$ . Thus one cannot expect that the singularities cancel in the sum. The solutions of this problem comes from the field renormalisation constants. For massless particles, the

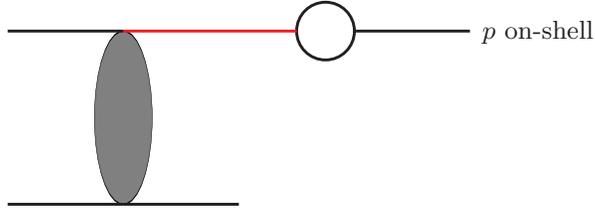


Fig. I.10: Self-energy corrections on external lines lead to an internal on-shell propagator (shown in red).

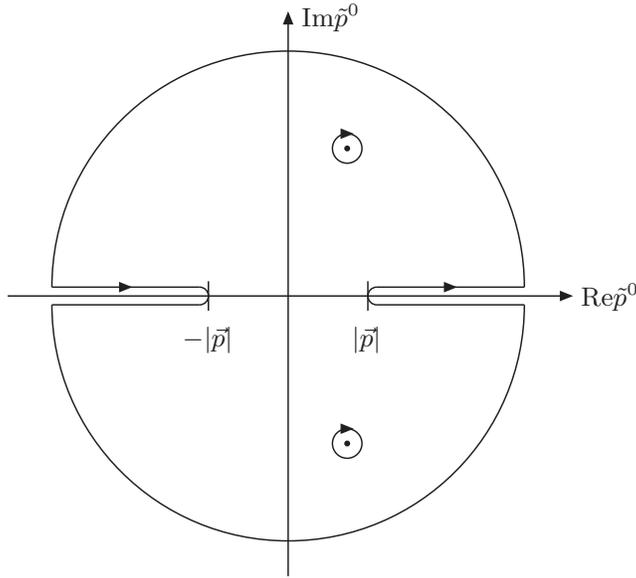


Fig. I.11: The contour for the dispersion integral for the self-energies.

field renormalisation constants are 1. However, they are only 1 due to a cancellation between (divergent) contributions from the ultraviolet and infrared region. We may write

$$\begin{aligned}
 Z_2 &= 1 = 1 + \frac{\alpha_s}{4\pi} C_F \left( \frac{1}{\varepsilon_{\text{IR}}} - \frac{1}{\varepsilon_{\text{UV}}} \right), \\
 Z_3 &= 1 = 1 + \frac{\alpha_s}{4\pi} (2C_A - \beta_0) \left( \frac{1}{\varepsilon_{\text{IR}}} - \frac{1}{\varepsilon_{\text{UV}}} \right).
 \end{aligned}
 \tag{I.44}$$

The infrared part of the field renormalisation constants will provide the missing terms which cancel against the collinear singularities from the real and virtual part.

Let us now try to implement the contributions from the field renormalisation constants into our numerical approach. The field renormalisation constants derived from self-energies and we are tempted to consider the integral representation for the self-energies. Here the next problem appears: For self-energy corrections on external lines, we have an internal (non-loop-like) propagator, which is on-shell. This is shown in Fig. I.10. This problem can be solved with the help of a dispersion relation. We re-write the contribution from the self-energy corrections as a dispersion integral with a contour as shown in Fig. I.11.

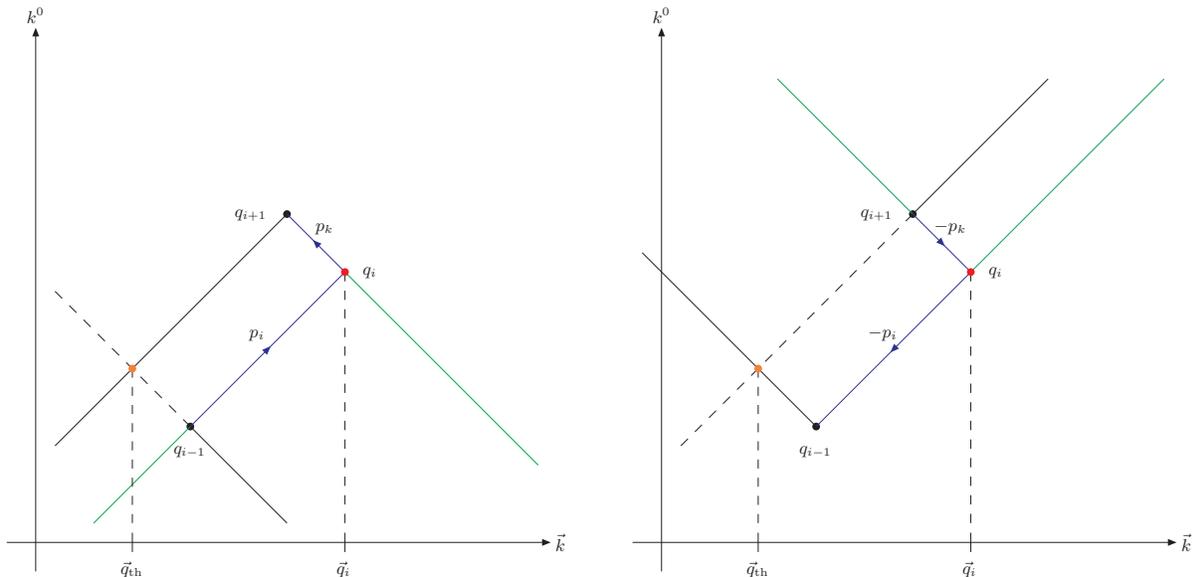


Fig. I.12: The integration regions for a final-final antenna. The left picture corresponds to the dipole with emitter  $i$  and spectator  $k$ , the right picture corresponds to the dipole with emitter  $k$  and spectator  $i$ . The soft region is indicated by a red dot, the collinear regions by blue line segments. There is a cancellation of singularities within the virtual dual contributions in the regions where two propagators are on-shell and have the same sign in the energy component. These regions are indicated in green. There is a threshold singularity (indicated by an orange dot) at  $\vec{q}_{\text{th}}$ . The threshold singularity is avoided by contour deformation.

For processes with only final-state infrared singularities (i.e. electron-positron annihilation) we now achieve a cancellation of all singularities at the integrand level. It is instructive to visualise the cancellations in loop momentum space. This is shown in Fig. I.12.

Now let us turn to processes with initial-state partons (i.e. hadron colliders like the LHC). For initial-state partons the local cancellation of collinear singularities is more involved. Here we face the problem that the regions of the collinear singularities from the real part and of the collinear singularities from the virtual part do not match in our unified space. The solution comes through the counterterm from factorisation. This counterterm has the form

$$d\sigma^{\text{C}} = \frac{\alpha_s}{4\pi} \int_0^1 dx_a \frac{2}{\varepsilon} \left( \frac{\mu_F^2}{\mu^2} \right)^{-\varepsilon} P^{a'a}(x_a) d\sigma^{\text{B}}(\dots, x_a p'_a, \dots), \quad (\text{I.45})$$

where  $P^{a'a}$  denotes a splitting function. For example, for the splitting  $g \rightarrow gg$  the splitting function  $P^{gg}$  is given by

$$P^{gg} = 2C_A \left[ \frac{1}{1-x} \Big|_+ + \frac{1-x}{x} - 1 + x(1-x) \right] + \frac{\beta_0}{2} \delta(1-x). \quad (\text{I.46})$$

This splitting function consists of an  $x$ -dependent part related to the real corrections and an end-point contribution proportional to  $\delta(1-x)$  related to the virtual corrections. We may now again use integral representations for these two contributions. This is sketched in Fig. I.13. Taking these into account, one achieves a local cancellation of all singularities.

Let us summarise how the singularities cancel locally. Soft singularities cancel between the virtual subtraction terms  $d\sigma_{\text{V,IR}}^{\text{A}}$  and the real subtraction terms  $d\sigma_{\text{R}}^{\text{A}}$ . The longitudinal parts

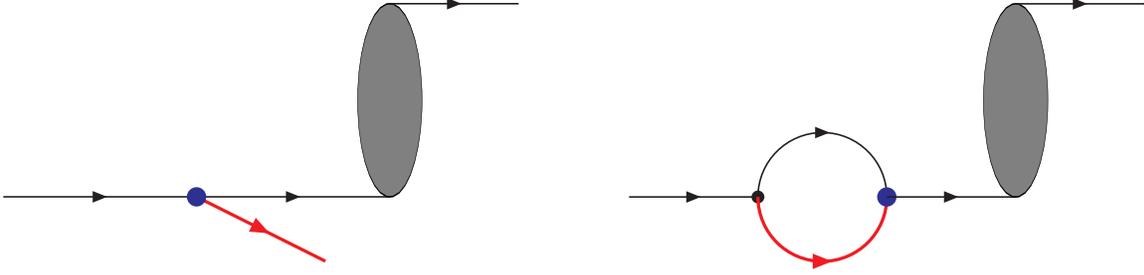


Fig. I.13: The integral representation of the collinear counterterm for initial-state partons can be split into two contributions: The  $x$ -dependent part (left) matches on the real contribution, the end-point contribution (right) matches on the virtual contribution.

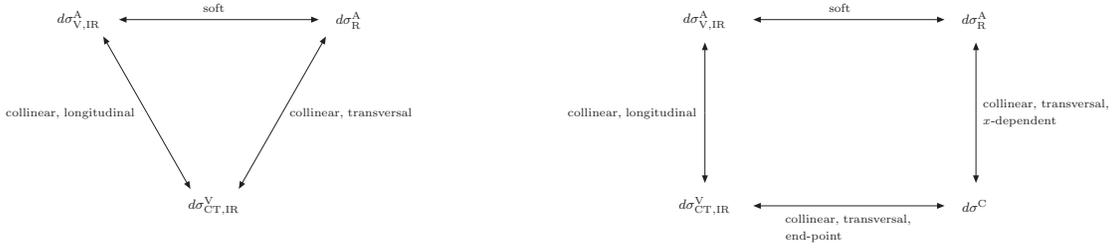


Fig. I.14: The local cancellation of infrared singularities for final-state singularities (left) and initial-state singularities (right).

of the collinear singularities present in virtual subtraction terms  $d\sigma_{V,IR}^A$  cancel against similar parts in the term  $d\sigma_{CT,IR}^V$  from field renormalisation. For the transverse parts of the collinear singularities we have to distinguish final-state collinear singularities and initial-state collinear singularities. In the former case these singularities cancel between  $d\sigma_R^A$  and  $d\sigma_{CT,IR}^V$ . In the latter case the  $x$ -dependent transverse part of the collinear singularities cancels between  $d\sigma_R^A$  and the counterterm  $d\sigma^C$  from factorisation, while the end-point contribution cancels between  $d\sigma_{CT,IR}^V$  and  $d\sigma^C$ . This is shown in Fig. I.14. To complete the discussion let us turn to the ultraviolet divergences. These are unproblematic, the singularities cancel locally between  $d\sigma_{V,UV}^A$  and  $d\sigma_{CT,UV}^V$ .

### 6.3 Conclusions

We have discussed the local cancellation of ultraviolet and infrared (i.e. soft and collinear) singularities at the integrand level at NLO for electron-positron colliders and hadron colliders. Although we have focused in the discussion on massless particles, the extension towards massive final-state particles is straightforward and does not pose any further conceptual problem. We may therefore perform the integrals over the subtraction terms numerically. At NLO there is no real advantage, as all integrated subtraction terms are known analytically. The application is the extension towards NNLO, where analytic integration of local subtraction terms is very challenging. We have seen that the cancellation at the integrand level involves several subtleties, which are best studied at NLO.

## 7 NLO EW automation and technical comparisons <sup>11</sup>

### 7.1 Introduction

Adequate predictions for scattering processes at particle colliders such as the Large Hadron Collider (LHC) are known to require the inclusion of at least next-to-leading order (NLO) perturbative corrections of the strong interactions. With the increasing experimental precision, and the extension of the kinematic reach thanks to larger collision energies, not only beyond-NLO predictions in QCD are required, but also the inclusion of perturbative corrections from electroweak (EW) interactions becomes inevitable. In particular, in the high-energy and more precisely the high transverse-momentum tails of kinematic distributions which are currently probed, EW corrections become sizeable.

The automated calculation of NLO QCD corrections started to be established almost a decade ago, and nowadays several independent and well-tested tools exist [83–85, 451, 464, 477–481], based on different approaches and implementations. In the recent years, also the automation of NLO EW corrections was pushed forward, and in the meanwhile a few tools are able to produce predictions at full NLO Standard Model (SM) accuracy [279, 304, 481, 482].

In this contribution we review the status of the existing tools for the automated calculation of EW one-loop amplitudes and their implementations into integration frameworks. We also give an overview of phenomenological studies at NLO EW precision carried out with these tools in the recent years. After the more qualitative comparison of some selected processes in the proceedings of the previous Les Houches 2015 workshop [1], we perform a detailed technical comparison for two  $2 \rightarrow 4$  processes with complicated resonance structures, namely the off-shell production of leptonically decaying  $ZZ$  and  $WW$  pairs: we start with a point-wise comparison of the EW 1-loop matrix elements between the different amplitude generators, and we study integrated cross sections and distributions at NLO EW accuracy.

### 7.2 Automated calculation of electroweak NLO corrections

Similar to NLO QCD calculations, the automation of EW corrections at NLO accuracy requires two main ingredients: a framework that takes care of the bookkeeping of partonic subprocesses contributing at the coupling order(s) under consideration, provides a subtraction procedure to treat IR QED singularities, and performs a stable phase-space integration to provide predictions for any observables of interest; and a one-loop provider (OLP) that generates one-loop transition amplitudes including appropriate EW renormalization procedures, and moreover guarantees their stable numerical evaluation in the complete phase space.

Since both tasks are to a wide extent complementary, and the combination of any integration framework with any OLP is in principle possible, we discuss them separately in the following. In the first part of this section, we give an overview of the existing Monte Carlo integration frameworks that have been applied for NLO EW calculations. In the second part, we briefly introduce the available OLPs, and discuss their recent applications in phenomenological studies.

---

<sup>11</sup> S. Kallweit (section coordinator); B. Biedermann, S. Bräuer, M. Chiesa, A. Denner, N. Greiner, V. Hirschi, J.-N. Lang, J. M. Lindert, P. Maierhöfer, M. Pellen, S. Pozzorini, S. Quackenbush, C. Reuschle, M. Schönherr, S. Schumann, H.-S. Shao, S. Uccirati

### 7.2.1 Integration frameworks for automated electroweak NLO corrections

#### BBMC

BBMC<sup>12</sup> is a multi-channel Monte Carlo integrator that has been designed for NLO corrections to processes with identified massless leptons and jets in the final state. The multi-channel kernel is a further development of the implementation in Ref. [483, 484]. The selection of the partonic channels for a given hadronic process and the NLO dipole subtraction [187, 485] is done in a fully automated way. The implementation supports typical NLO EW features like handling of collinear-unsafe photon radiation off charged leptons [486], fragmentation functions and photon-induced contributions. The program is interfaced with RECOLA [89] as matrix-element generator for both one-loop and tree-level amplitudes. BBMC has been used for all kinds of vector-boson pair-production processes with leptonic decays at NLO EW accuracy [222, 223, 225, 313] and for the calculation of the NLO EW and QCD corrections to vector-boson scattering in the same-sign WW channel [224, 226]. The code is intended to be publicly available in the future.

#### MoCANLO

MoCANLO<sup>13</sup> is a flexible Monte Carlo program that can compute arbitrary processes in the SM with NLO QCD and EW accuracy. The fast integration is ensured by using similar phase-space mappings to those of Refs. [483, 484, 487]. The infrared (IR) divergences appearing in virtual and real corrections are handled with the help of the Catani–Seymour dipole formalism [187, 485]. These tools have been successfully used for the computation of NLO corrections for high-multiplicity processes [224, 226, 281, 282, 331, 346].

#### MUNICH

MUNICH<sup>14</sup> is a fully general and very fast parton-level Monte Carlo integrator, written in C++, which was originally developed in the context of multi-leg NLO QCD calculations [342, 343, 348, 488]. It automatically performs the bookkeeping of partonic subprocesses, phase-space integration based on multi-channel techniques, and the treatment of IR singularities at NLO by means of dipole subtraction for massless [187] and massive [489] partons. Subsequently, this framework was extended to deal with full SM corrections at NLO accuracy, by imposing a generalized bookkeeping for subprocesses at arbitrary coupling orders. The subtraction of IR QED singularities is performed in the dipole-subtraction approach as detailed in Refs. [304, 314]. These developments allow not only NLO EW corrections to be performed, but full SM NLO corrections including all sub-leading coupling orders and photon-induced contributions for arbitrary SM processes. MUNICH relies on external amplitudes throughout: So far it has been applied in combination with OPENLOOPS, where the latter provides not only the one-loop amplitudes of QCD, EW, and mixed types, but also the tree amplitudes with all relevant colour and helicity correlations, as required in the IR subtraction. Corresponding applications of MUNICH+OPENLOOPS [302, 304, 305, 314] are detailed in the next section. MUNICH also constitutes the fundament of the public NNLO QCD framework MATRIX<sup>15</sup> [154], which was applied to perform the first calculations for almost all di-boson processes at NNLO QCD accuracy [158, 160–166, 170, 195]. These features will facilitate combined NNLO QCD+NLO EW predictions based on the MUNICH/MATRIX+OPENLOOPS framework in the near future.

---

<sup>12</sup>BBMC has been written by B. Biedermann, and is so far only a working title as the program is not yet ready for publication.

<sup>13</sup>MoCANLO stands for “MONte CARlo event generator for NLO computation“. It was originally written by Robert Feger and further developed by M. Pellen. In preparation.

<sup>14</sup>MUNICH is the abbreviation of “MULTI-chaNNel Integrator at swiss (CH) precision”— an automated parton level NLO generator by S. Kallweit. In preparation.

<sup>15</sup>MATRIX is available for download from: <http://matrix.hepforge.org>.

## SHERPA

SHERPA<sup>16</sup> [415, 490] is a multi-purpose Monte Carlo event generator aiming for the full simulation of high-energy collider scattering events. The SHERPA framework contains modules and algorithms for the various different phases of event evolution. This includes methods for the generation and integration of hard-scattering matrix elements, QCD parton showers [491, 492], and phenomenological models for the parton-to-hadron transition and the underlying event.

For the hard process generation it contains two tree-level matrix element generators for the full SM and New Physics scenarios [493], called AMEGIC [494] and COMIX [495]. Both matrix element generators are equipped with an automated subtraction of NLO QCD divergences in the Catani–Seymour dipole-factorization scheme [187, 489, 496]. This allows for the automated evaluation of NLO QCD corrections to arbitrary processes. Recently, AMEGIC was updated to be able to handle processes with NLO EW divergences as well [304, 497], which is used in these comparisons. To facilitate full NLO computations, interfaces to various one-loop matrix element codes are provided. In the following, SHERPA is used in combination with GOSAM [324, 498], OPENLOOPS [302, 304, 305, 314] and RECOLA [499]. In these combinations of tools, SHERPA provides the tree-level matrix elements for both the Born and real emission contributions, the IR subtraction, the process management and phase-space integration of all contributions to all processes considered. To analyse the generated parton-level events the RIVET [500] package is used. On-the-fly scale variations are available through an extension of the algorithm of Ref. [501].

## MADGRAPH5\_AMC@NLO

MADGRAPH5\_AMC@NLO<sup>17</sup> [88] is a single framework automating the computation of both LO (tree-level and loop-induced [502]) and NLO accuracy (differential) cross sections, including their matching to parton shower programs via the MC@NLO method [503]. It provides all the necessary elements for the SM and beyond SM (BSM) phenomenology studies. In particular, virtual contributions are numerically calculated by the MADLOOP module [479], which will be described later, while the real emission contributions are regulated using the Frixione-Kunszt-Signer (FKS) subtraction method [142, 143, 421]. Inclusive samples accurate at the NLO QCD level across many jet multiplicities can be obtained using the FxFx merging method [504]. External OLPs can also be interfaced to MADGRAPH5\_AMC@NLO, using the Binoth Les Houches Accord (BLHA) standard [505, 506] (as for example done with GOSAM [87] in Ref. [385]).

More recently, the MADGRAPH5\_AMC@NLO framework has been extended in order to support the computation of NLO EW corrections, including the case of mixed QCD and EW coupling order perturbative expansions. This allowed the computation of the EW corrections to the process  $t\bar{t} + H/Z/W^\pm$  [279, 280], and of the *complete* set of mixed-order NLO corrections to dijet [288],  $t\bar{t}W^\pm$  and  $t\bar{t}t\bar{t}$  [388]. The framework was also instrumental in combining NNLO QCD and NLO EW corrections to  $t\bar{t}$  [329, 507] production at both the LHC collision energies and 100 TeV. The corresponding new NLO EW capable version of MADGRAPH5\_AMC@NLO is planned for an imminent public release, together with a forthcoming reference paper [508]. Two NLO UFO models [509] with all the necessary rational  $R_2$  [510–514] and ultraviolet (UV) renormalization counterterms will be included, allowing for the computation of the complete NLO corrections to arbitrary scattering and decay processes in the SM. The two models correspond to two different EW renormalization schemes:  $\alpha(M_Z)$  [515] and  $G_\mu$  [515, 516]. For both setups, the use of renormalization conditions from either the on-shell or the complex-mass scheme [483, 517] has been validated. The introduction of the complex-mass scheme allows to include off-shell effects of unstable particles while retaining both gauge invariance and a well-defined perturbative

<sup>16</sup>The SHERPA program is available from <http://sherpa.hepforge.org>.

<sup>17</sup>The latest MADGRAPH5\_AMC@NLO version can be downloaded from <https://launchpad.net/mg5amcnlo>.

expansion. A new syntax is introduced in MADGRAPH5\_AMC@NLO for specifying the type of NLO corrections (QCD and/or EW) that must be accounted for:

```
MG5_aMC> set complex_mass_scheme True
MG5_aMC> import model <NLModel_with_qcd_qed>
MG5_aMC> generate <process> QCD=n QED=m [QCD QED]
MG5_aMC> output; launch
```

The above syntax implies that complex-mass scheme renormalization conditions are considered and that the following LO and NLO contributions must be included:

$$\begin{aligned}\sigma^{\text{LO}} &= \sum_{i \leq n, j \leq m, i+j=k_0} \alpha_s^i \alpha^j \sigma_{(i,j)}^{\text{LO}}, \\ \delta\sigma^{\text{NLO}} &= \sum_{i \leq n+1, j \leq m+1, i+j=k_0+1} \alpha_s^i \alpha^j \delta\sigma_{(i,j)}^{\text{NLO}}.\end{aligned}\tag{I.47}$$

One can choose the different integer values  $n$  and  $m$  and coupling order names in squared brackets to select specific terms in the perturbative series.

Also, numerous efforts in last few years have pushed the MADGRAPH5\_AMC@NLO framework so as to be able to perform various BSM studies at the NLO QCD accuracy [268, 284, 286, 518–537] with the help of FEYNRULES [538, 539] and NLOCT [540], which highlights the flexibility of the framework<sup>18</sup>. Indeed, these studies necessitate the support of many novel aspects absent in SM physics, like non-renormalizable operators [268], Majorana fermions [524] and spin-2 particles [531].

### 7.2.2 Generators for electroweak one-loop amplitudes and applications

RECOLA — BBMC/MoCANLO/SHERPA+RECOLA

RECOLA [89, 481] is a general one-loop amplitude provider which is publicly available.<sup>19</sup> The library combines the process generation and computation of matrix elements in a fully automated and recursive way, pushing limitations on the numbers of external particles to a yet unmatched level, while remaining a very flexible tool. In particular, RECOLA provides the functionality to compute tree and one-loop matrix elements, squared matrix elements, optionally summed over spin and colour, and spin- and/or colour-correlated matrix elements in the 't Hooft–Feynman gauge for arbitrary initial and final states. Standard renormalization schemes for the electroweak and strong gauge coupling are supported. Unstable particles are treated in the complex-mass scheme [483, 517, 541]. The recent upgrade, dubbed RECOLA2 [542, 543], extends the original version by new model files, as well as the possibility to perform computations in the Background-Field method, and a PYTHON interface. A key aspect of RECOLA/RECOLA2, compared to other automated NLO tools, is that no process source files are generated, and, thus, no intermediate compilation is required. Therefore, it naturally fits into the framework of modern event generators requiring flexible amplitude providers which can be used as black boxes.

Internally, RECOLA is based on the so-called Berends-Giele recursion [544] (BGR) allowing to construct tree-level amplitudes without referring to Feynman diagrams in the generation or in the computation phase. A. van Hameren showed [545] that, based on the decomposition of one-loop amplitudes  $\mathcal{M}_1$  in terms of tensor coefficients  $c$  and tensor integrals  $T$  as

$$\mathcal{M}_1 = \sum_k c_{k,\mu_1\dots} T_k^{\mu_1\dots},\tag{I.48}$$

<sup>18</sup>The available NLO QCD-ready UFO models for BSM studies are listed and can be downloaded at <http://feynrules.irmp.ucl.ac.be/wiki/NLOModels>.

<sup>19</sup>RECOLA can be obtained at <http://recola.hepforge.org>.

a recursion for tensor coefficients can be derived, similar to the BGR. RECOLA and RECOLA2 implement such an algorithm for computing tensor coefficients numerically for arbitrary processes in the SM and beyond, respectively, while the tensor integrals are obtained by means of the COLLIER library [546, 547]. Finally, the complete 1-loop renormalized amplitude  $\mathcal{M}_{\text{ren}}$  is constructed using the purely 4-dimensional part of Eq. (I.48) in addition to counterterm and rational parts,

$$\mathcal{M}_{\text{ren}} = \mathcal{M}_1 + \mathcal{M}_{\text{CT}} + \mathcal{M}_{\text{R}_2}, \quad (\text{I.49})$$

with  $\mathcal{M}_{\text{CT}}$  and  $\mathcal{M}_{\text{R}_2}$  being computed on equal footing with the tree-level amplitude, but using special Feynman rules.

RECOLA has been used with several Monte Carlo programs for the calculation of NLO QCD and EW corrections. In combination with the parton-level integration frameworks BBMC and MOCANLO, it has demonstrated to be particularly efficient for high-multiplicity processes up to  $2 \rightarrow 7$  scattering.

A first class of processes which have been calculated with the help of RECOLA is the production of massive di-bosons at NLO EW: WW [223], ZZ [222, 225], and WZ [313]. All these computations consider the off-shell production as well as all interferences and non-resonant contributions.

A second class of processes evaluated with RECOLA concerns the off-shell production of top quarks. Thus, the EW corrections [331] to the off-shell production of top-antitop pairs have been computed. Recently, QCD corrections to the same process but in the lepton+jets channel [346] have also been obtained. In addition, the associated production of a Higgs boson with off-shell top-antitop pairs has been calculated at both NLO QCD and EW accuracy [281, 282]. This constituted the first  $2 \rightarrow 7$  process obtained at NLO. For the EW corrections, for the first time, 9-point functions appear in the loop amplitude.

A further class covered is vector-boson scattering (VBS). First, NLO EW corrections to the same-sign WW VBS processes have been computed [224]. These turn out to be particularly large as an intrinsic feature of VBS processes at the LHC. Then, the full NLO QCD and EW corrections to the VBS process and its irreducible background have been obtained [226]. As the signature possesses three LO contributions, this amounts to compute four separate NLO contributions.

In order to fully exploit the high level of automation of RECOLA, it has been interfaced to the multi-purpose Monte Carlo generator SHERPA [415, 490]. Based on an automated subtraction of both QCD and EW divergences [496, 497], SHERPA+RECOLA [499] can compute any process at NLO QCD and EW accuracy in the SM. Some examples of the capabilities have been demonstrated by studying vector-boson production in association with jets<sup>20</sup>, on-shell  $t\bar{t}H$  production, and off-shell ZZ production for both QCD and EW corrections. This framework has been used *e.g.* for the computation of NLO QCD corrections to on-shell top-quark pair production [548]. Note that RECOLA has also been recently interfaced to WHIZARD [549, 550].

#### OPENLOOPS — MUNICH/SHERPA+OPENLOOPS

OPENLOOPS<sup>21</sup> provides scattering amplitudes at NLO QCD+EW based on the open-loops algorithm [85] – a fast hybrid tree-loop recursion for the numerical evaluation of tree and one-loop scattering amplitudes. At NLO QCD more than one hundred processes are publicly provided in the form of an automatically generated library that supports all interesting LHC processes, and

<sup>20</sup>The first application of RECOLA was the computation of NLO EW corrections to lepton pair production in association with two hard jets [303, 481].

<sup>21</sup>The OPENLOOPS one-loop generator is publicly available at <http://openloops.hepforge.org>.

which can be easily extended upon user request. A public NLO EW library is under development and will very soon be released as part of OPENLOOPS2 [551].

The recently achieved automation of EW corrections [304, 305, 314] is based on the well established QCD implementations and allows for NLO QCD+EW simulations for a vast range of SM processes, up to high particle multiplicity, at current and future colliders. To be precise, the new implementations allow for NLO calculations at any given order  $\alpha_s^n \alpha^m$ , with all relevant QCD–EW interference effects. Full NLO SM calculations that include all possible  $\mathcal{O}(\alpha_s^{n+k} \alpha^{m-k})$  contributions to a certain process are also supported.

The extension to NLO EW corrections required the implementation of all  $\mathcal{O}(\alpha)$  EW Feynman rules in the framework of the numerical open-loops recursion including counterterms associated with so-called  $R_2$  rational parts [512] and with the on-shell renormalization of UV singularities [516]. Additionally, for the treatment of heavy unstable particles the complex-mass scheme [517] has been implemented in a fully general way. Combined with the COLLIER tensor-reduction library [546], Denner:2016kdg, which implements the Denner–Dittmaier reduction techniques [552, 553] and the scalar integrals of Ref. [554], or with CUTTOOLS [555], which implements the OPP method [95], together with the ONELOOP library [556], the employed recursion permits to achieve very high CPU performance and a high degree of numerical stability. A new method for the automated construction of one-loop amplitudes and their on-the-fly reduction to scalar integrals building on the open-loops algorithm was introduced in Ref. [557], and will be part of OPENLOOPS2 [551]. This improved algorithms will further significantly enhance the numerical stability of the amplitude evaluation in particular relevant for real-virtual contributions in NNLO computations.

The frameworks MUNICH+OPENLOOPS and SHERPA+OPENLOOPS automate the full chain of operations – from process definition to collider observables – that enter NLO QCD+EW simulations at parton level. Employing the described frameworks, in Ref. [304] the simulation of  $W^+ + 1, 2, 3$  jets production at NLO EW+QCD was presented. To facilitate the calculation of the process with the highest jet multiplicity, these processes were factorized into a production part and a decay part. All sub-leading Born channels including interference-based and photon-induced ones were studied, and NLO corrections were calculated at relative  $\mathcal{O}(\alpha_s)$  and  $\mathcal{O}(\alpha)$  with respect to the leading Born contribution. The latter also involves interferences with sub-leading Born diagrams beside genuine EW contributions. A careful implementation of the narrow-width approximation is required in order to control numerical stability, given the appearance of pseudo-resonances for two or more associated jets: Here a pole regularization with a technical width parameter was chosen, which corresponds to a smooth and numerically negligible deformation with respect to the gauge-invariant on-shell limit. Phenomenologically it was found that  $V$ +multijet final states feature genuinely different EW effects as compared to the case of  $V$ +1jet.

Subsequently in Ref. [305] NLO QCD+EW simulations were presented for  $Z$ +jets and  $W^\pm$ +jets production including off-shell leptonic decays with 0, 1, and 2 charged leptons in the final states, applying the complex-mass scheme throughout. A naive-exclusive-sums approach was applied to  $V$ +1jet and  $V$ +2jets NLO QCD+EW predictions in order to derive an approximation that combines exact EW virtual corrections with an inclusive treatment of bremsstrahlung effects. This so-called *VI approximation* made it possible to include NLO EW corrections into the MEPS@NLO multijet merging framework of SHERPA, and so to recover perturbative convergence in a merged  $V$ +jet calculation. Recently this framework has also been applied for the calculation of NLO EW corrections to  $t\bar{t}$ (+jet) production and a corresponding MEPS@NLO QCD+EW<sub>VI</sub> multijet merging [332], as presented in Sec. V.1 of the proceedings at hand.

The work on  $V$ +jets was continued in the context of the dark-matter background study of Ref. [302], where the main focus was on providing not only the best available perturba-

tive prediction, but in particular a reasonable uncertainty estimate. The region of interest — as main background in dark-matter searches — is the high-energy tail of the transverse-momentum distribution of the vector boson in the invisible decay channel  $Z \rightarrow \nu\bar{\nu}$ . Most accurate predictions are achieved here by estimating  $Z(\nu\bar{\nu})+\text{jet}$  from related  $V+\text{jet}$  processes. Thus best possible theoretical predictions and uncertainty estimates were required for ratios between the different  $V+\text{jet}$  processes, with  $V = Z, W, \gamma$ . The frameworks MUNICH+OPENLOOPS and SHERPA+OPENLOOPS delivered all results involving EW corrections. The Sudakov approximation was used to formulate reasonable error estimates at NLO EW. Moreover, approximate NNLO EW predictions were achieved, based on an implementation of 2-loop Sudakov logarithms in OPENLOOPS, which was required to reduce the uncertainties originating from EW effects to a subdominant level. Moreover, mixed NNLO QCD $\times$ EW effects were estimated, based on the universality assumption of the leading Sudakov effects that promotes a multiplicative combination of QCD and EW corrections. Dedicated studies of the behaviour of NLO EW corrections on  $V + 2\text{jets}$  under variation of a jet resolution parameter were performed to achieve an improved uncertainty estimate beyond just taking the full relative  $\mathcal{O}(\alpha_s\alpha)$  contribution as unknown.

In Ref. [314], NLO QCD+EW corrections to  $2\ell 2\nu$  final states were investigated, both in different-flavour and same-flavour channels. The first is dominated by WW resonances, whereas the latter involves, depending on neutrino flavours, only ZZ resonances or both WW and ZZ resonances. In this last case, interference effects turn out to be completely negligible apart from the region around the  $Z \rightarrow 2\ell 2\nu$  peak. Throughout the calculation, photon-induced channels are included not only at LO, but also in the corrections at NLO EW accuracy. Both the overall impact of incoming-photon contributions and the dependence on the different photon PDFs available by then were investigated on a differential level. Moreover, the possibility of reproducing the exact NLO EW results by the above-mentioned EW VI approximation augmented with QED radiation effects was studied, where the latter were generated via YFS soft-photon resummation or, alternatively, by the Catani–Seymour dipole-based DGLAP-type resummation of the CSSHOWER. Both approaches describe the high-energy regions similarly well with deviations typically below 10%. The YFS resummation implementation in SHERPA also preserves the existing resonance structure.

In the framework of POWHEG+OPENLOOPS recently Monte Carlo generators for the production of  $HV(+\text{jet})$  at NLO QCD+EW have been presented [262]. They provide a consistent matching to the QCD and QED parton showers in PYTHIA8. Here  $V = \{W^\pm, Z\}$  denotes the corresponding leptonic off-shell processes, and the application of the improved MINLO method to  $HV+\text{jet}$  production allows for NLO QCD+EW accuracy for observables with both zero or one resolved jet. These generators have been used to study the behaviour of EW corrections for various kinematic distributions, relevant for experimental analyses of Higgs-strahlung processes. The OPENLOOPS tree and one-loop amplitudes are accessible in the POWHEG-BOX-RES framework [356] via a process-independent interface developed in Ref. [357]. This framework allows for the consistent resonance-aware parton-shower matching of off-shell processes at NLO including radiation off the decay products of (narrow) resonances, in particular relevant in the context of matched NLO EW corrections.

MADLOOP — MADGRAPH5\_AMC@NLO+MADLOOP

MADLOOP [479] is a module of the MADGRAPH5\_AMC@NLO [88] framework in charge of generating the code for one-loop matrix element computations<sup>22</sup>. It uses an original approach to the generation of one-loop Feynman diagrams that takes advantage of the existing tree-level diagram generation algorithm of MADGRAPH5\_AMC@NLO to directly generate  $L\text{-cut}$  dia-

<sup>22</sup>Procedure for obtaining a standalone code for one-loop computations in MADLOOP can be found at <http://cp3.irmp.ucl.ac.be/projects/madgraph/wiki/MadLoopStandaloneLibrary>.

grams, corresponding to loop diagrams with one loop propagator cut open, effectively turning it into a tree-level diagram with two additional final states compared to the starting loop topology.

Significant improvements have been made in MADLOOP5 [88] compared to its predecessor MADLOOP4 [479]. The generation of the loop numerator is rendered efficient by using an in-house implementation of the open-loops [85] technology, and is generically applicable to any model (spin-2 [531], supersymmetric models [522, 524], EFT [284, 520, 530, 536, 558, 559], vector-like quarks [532], dark matter [525, 528, 529] and others<sup>23</sup>) thanks to the generation of the vertex functions using ALOHA [560]. Moreover, MADLOOP offers the possibility of dynamically switching between interfaces to three different integrand-level reduction programmes [95, 117, 555, 561–564] and four different tensor integral reduction codes [547, 553, 565–568], each time providing a reliable estimate of the numerical uncertainty<sup>24</sup>. Whenever a numerically unstable evaluation is encountered, MADLOOP will re-evaluate the point on-the-fly in quadruple precision, not only within the loop reduction (in the case of NINJA [117, 563] and CUTTOOLS [555]) but also for the computation of the numerator tensor coefficients. Another key feature of MADLOOP is the flexibility it offers to the user of selecting only particular contributions of the complete one-loop matrix-element, for instance by limiting the allowed particle content, requiring certain propagators or selecting any arbitrary gauge-invariant subset of diagrams. As long as the input UFO model provides the relevant information, the correct corresponding  $R_2$  and UV counterterm contributions of Eq. (I.49) will be consistently accounted for.

MADLOOP can also provide independent results for multiple terms factorizing specific user-defined sets of couplings. This is useful in the context of BSM NLO QCD computations for studying particular interference contributions, while in the SM it is a necessary feature for computing the *complete* set of NLO corrections of order  $\mathcal{O}(\alpha_s^{b-n}\alpha^n)$  with  $n \leq b$  (see as for example dijet production [288], where  $b = 3$ ). It is also capable of providing spin- and colour-correlated matrix elements, allowing it to be used in the context of NNLO QCD computations. The internal projection of amplitudes onto colour-ordered ones also renders MADLOOP competitive for the computation of loop-induced matrix elements [502]. Finally, MADLOOP is not bound to be used within the MADGRAPH5\_AMC@NLO integration framework, and it can easily be used in standalone mode to generate computer libraries for arbitrary processes, ready to be used within any environment (including facilities for calling these libraries directly from within PYTHON).

## GOSAM — SHERPA+GOSAM

GOSAM [87, 480] is a publicly available tool for the automated generation of virtual amplitudes.<sup>25</sup> It is based on a Feynman diagrammatic approach, where the amplitude is generated algebraically in  $D$  dimensions. To generate the Feynman diagrams it uses QGRAF [569] to generate all the relevant topologies, and further employs FORM [570, 571] and SPINNEY [572] (a FORM library to handle the spinor-helicity formalism) to apply the appropriate Feynman rules and write an optimized FORTRAN output. Due to the algebraic approach no special treatment or additional Feynman rules for the rational  $R_2$  terms are needed. Internally GOSAM uses dimensional reduction (DRED) [573–576] to construct the amplitude, but it can be converted to the 't Hooft-Veltman scheme [577]. The output for the loop amplitude does not require a specific reduction technique but can be used for several reduction methods, either on the integral- or on the integrand level. At the moment three different reduction techniques are built in and supported. Per default it uses NINJA [562, 563, 578], a package that performs the reduction on

<sup>23</sup>See <http://feynrules.irmp.ucl.ac.be/wiki/NLOModels> for a list of available UFO models ready for NLO QCD computations.

<sup>24</sup>The one-loop scalar integrals in MADLOOP are numerically evaluated by the program ONELOOP [556] or directly by COLLIER [547] when it is used for the loop reduction.

<sup>25</sup>GOSAM can be obtained at <http://gosam.hepforge.org>.

the integrand level by applying a Laurent expansion. Alternatively one can choose between an OPP reduction [95, 510, 561] as implemented in SAMURAI [564], or the tensor reduction methods contained in the GOLEM95 library [567, 579–581]. The remaining scalar integrals can be evaluated by either using ONELOOP [556], or QCDLOOP [582, 583]. Both the reduction method and the library for the scalar integrals can be changed at any point on the fly.

To calculate physical observables like cross sections and differential distributions at the NLO level GOSAM can be combined with any Monte Carlo event generator that supports the BLHA interface [505, 506]. In practice, interfaces have been established with HERWIG7/HERWIG++ [584], MADGRAPH5\_AMC@NLO [385], POWHEG [585, 586], SHERPA [415] and WHIZARD [549] and has been used to calculate a large variety of different processes at NLO in QCD, EW and BSM. GOSAM itself contains model files for various versions of the SM (with or without CKM mixing, complex mass scheme, effective Higgs couplings). However, GOSAM is able to interpret model files in the UFO format [509] which allows for an immediate and straightforward implementation of BSM models.

In the context of EW corrections, the combination of GOSAM and SHERPA has recently been used to calculate diphoton processes, in association with 0, 1, and 2 jets [498] or in association with a third vector boson [324].

A recent implementation of incorporating the use of quadruple precision within GOSAM in connection with NINJA allows to automatically switch from double to quadruple precision when numerical instabilities are detected. This switch allows to maintain numerical stability in extreme regions of phase space when approaching singularities as is the case for the real–virtual contributions in a two-loop calculation [67].

GOSAM is the only automated tool that has been extended to calculate two loop corrections. It has successfully been applied to calculate the full NLO QCD corrections to double Higgs production, taking full top-quark mass dependence into account [65, 66], and very recently also to calculate the NLO QCD corrections to  $H+1\text{jet}$ , also including full top-quark mass dependence [67].

## NLOX

NLOX is the newest member in the set of matrix element providers for the automated generation of one-loop amplitudes. A non-public predecessor of the program had been available in the past, to calculate one-loop QCD corrections to selected processes [587]. NLOX has seen quite some progress in recent years. The current version of the program provides fully renormalized QCD and EW one-loop amplitudes in the SM, for all possible QCD+EW mixed coupling-power combinations to one-loop accuracy, including the full mass dependencies on initial- and final-state masses. NLOX is based on a Feynman diagrammatic approach, utilizing QGRAF [569], FORM [570, 571] and PYTHON, to algebraically generate C++ code for the virtual contribution to a certain process at a certain order of QCD+EW coupling powers, in terms of one-loop tensor coefficients. The tensor coefficients are calculated recursively at runtime through standard tensor reduction methods by the C++ library TRED, an integral part of NLOX. Several reduction techniques are available to TRED, many of which are found in Ref. [553, 565]. The scalar one-loop integrals are evaluated by either using ONELOOP [556], or QCDLOOP [582, 583]. UV and IR singularities are regularized in dimensional regularization. UV renormalization in NLOX is carried out by means of counterterm diagrams, which provide a flexible way to systematically include mass renormalization for massive propagators as well as Yukawa-type vertices. The renormalization constants in terms of which the EW counterterms are formulated are derived in the on-shell renormalization scheme, based on Ref. [516], or in the complex-mass scheme, based on Ref. [517], where the choice in NLOX is to expand self-energies with complex squared momenta around real squared momenta as described in Ref. [517]. The renormalization constants

in terms of which the QCD counterterms are formulated are currently derived in a mixed renormalization scheme: The field and mass renormalization constants for massive quarks are derived in the on-shell renormalization scheme, while the field renormalization constants for massless quarks and gluons are derived in the  $\overline{\text{MS}}$  renormalization scheme, however, with possible contributions from massive quarks decoupling. As EW input scheme NLOX provides the  $\alpha(0)$  EW input scheme or the  $G_\mu$  EW input scheme. NLOX is not yet publicly available and for more details we refer the reader to Ref. [588]. The current version of NLOX can be interfaced to a selection of in-house Monte Carlo integration routines. A general interface to the Monte Carlo integrators and event generators discussed in this study is currently under development, which is why no cross-section level comparison was undertaken with NLOX for this study. As NLOX is still under development, the photon-induced channels were not available as of the drafting of this document.

### 7.3 Technical comparison for $4\ell$ and $2\ell 2\nu$ production at NLO EW accuracy

In order to validate the automated tools for evaluating NLO EW corrections, described in the previous section, a technical comparison has been performed, at the level of amplitudes for a set of phase-space points, integrated cross sections, and a selection of differential distributions. In contrast to the study on EW automation in the proceedings of the previous Les Houches 2015 workshop [1], where a more qualitative comparison has been performed — *i.e.* without an exhaustive fine-tuning of all ingredients of the calculations, like details of *e.g.* lepton–photon recombination, jet–photon clustering descriptions, treatment of fragmentation contributions, etc. — the goal of this contribution is to pin down all involved codes as far as possible to a common set of input parameters and conventions. In that way, a point-wise agreement only limited by the numerical precision of the respective matrix-element generators, and a statistical agreement on the level of integrated cross sections and distributions should be achievable. To do so, we chose to investigate the off-shell production of ZZ and WW pairs with corresponding leptonic decays in the different-flavour channels, *i.e.*  $pp \rightarrow e^+e^-\mu^+\mu^-$  and  $pp \rightarrow e^+\nu_e\mu^-\bar{\nu}_\mu$ . More precisely, the production of the full 4-lepton final state is considered and, correspondingly, all double-, single-, and non-resonant contributions as well as spin correlations and interference effects are fully taken into account without any resonance approximation applied. These processes are very well suited for such a technical comparison as they are non-trivial  $2 \rightarrow 4$  processes that exhibit involved resonance structures, so the different implementations of intermediate resonances by means of the complex-mass scheme are probed. On the other hand, they do not involve jets at their leading order (LO), which facilitates the choice of a common definition of observables, as no treatment of jet–photon configurations is needed, which would in general allow for various self-consistent implementations. As a side effect, the number of partonic processes is easily manageable, and the relative  $\mathcal{O}(\alpha)$  contains only genuine EW corrections, but no interference contributions with sub-leading orders as it would be the case if LO jets were involved.

#### 7.3.1 Setup

In this section the setup of the calculations is detailed. As this study is not a phenomenological, but a technical one, we refrain from using the most up-to-date input parameters and parton distribution functions (PDF). We rather adopt the setup of the detailed comparison for off-shell ZZ production applied in Ref. [499]<sup>26</sup>. For off-shell WW production, we stick to the very same

<sup>26</sup>The benchmark for the results of Ref. [499] has actually been set in Ref. [225], where all ingredients of the calculations, in particular the virtual matrix elements, have been internally cross-checked by means of at least two fully independent implementations. The setup of this original calculation deviates in some subtleties from the one of Ref. [499]. For convenience, we adopt the setup of the latter, as a small part of the comparison presented here has already been carried out therein using this setup.

setup, and only minimally adapt the applied phase-space cuts to the new final state.

All predictions are for proton–proton collisions at a centre-of-mass energy of  $\sqrt{s} = 13$  TeV, like in the present LHC run. Following Ref. [499], we use the NNPDF-2.3QED NLO PDF set [589] throughout, *i.e.* both for the LO and the NLO EW calculations, and we do not consider incoming photons in the observables by simply setting the photon PDF to zero (non-vanishing lepton PDFs are not considered either). We note that all results presented here are independent of the strong coupling  $\alpha_s$  since no QCD corrections are considered and the leading coupling order is  $\mathcal{O}(\alpha_s^0\alpha^4)$  for both processes under consideration. Correspondingly, there is also no dependence on a QCD renormalization scale. The QCD factorization scale is chosen to be fixed, and we set  $\mu_F = M_{Z/W}$  for ZZ and WW production, respectively, where  $M_V$  denotes the pole mass of the respective weak boson (see below).

The electromagnetic coupling  $\alpha$  is calculated in the  $G_\mu$  scheme [590],

$$\alpha = \frac{\sqrt{2}}{\pi} G_\mu M_W^2 \left( 1 - \frac{M_W^2}{M_Z^2} \right), \quad \text{with} \quad G_\mu = 1.16637 \times 10^{-5} \text{ GeV}^{-2}, \quad (\text{I.50})$$

and  $M_V^2$  corresponds to the real parts of the squared pole masses. We use the complex-mass scheme [483, 517, 541] throughout the calculation, *i.e.* the weak mixing angle becomes a complex quantity, derived from the ratio  $\mu_W/\mu_Z$ , where  $\mu_V^2 = M_V^2 - iM_V\Gamma_V$ . For details on the complex renormalization of the EW parameters, we refer to Ref. [517]. Some subtleties of the EW renormalization procedure that cause differences of higher order in  $\alpha$  in the virtual matrix elements will be discussed in the next section.

As numerical values of the on-shell gauge-boson masses and widths we use

$$\begin{aligned} M_Z^{\text{OS}} &= 91.1876 \text{ GeV}, & \Gamma_Z^{\text{OS}} &= 2.4952 \text{ GeV}, \\ M_W^{\text{OS}} &= 80.385 \text{ GeV}, & \Gamma_Z^{\text{OS}} &= 2.085 \text{ GeV}, \end{aligned} \quad (\text{I.51})$$

and convert them to pole quantities according to the relations [591]

$$M_V = \frac{M_V^{\text{OS}}}{\sqrt{1 + \left( \Gamma_V^{\text{OS}}/M_V^{\text{OS}} \right)^2}}, \quad \Gamma_V = \frac{\Gamma_V^{\text{OS}}}{\sqrt{1 + \left( \Gamma_V^{\text{OS}}/M_V^{\text{OS}} \right)^2}}, \quad V = W, Z. \quad (\text{I.52})$$

For convenience, we also state the numerical values of the resulting gauge-boson pole masses and widths and of  $\alpha$  in double-precision accuracy,

$$\begin{aligned} M_Z &= 91.1534806191828 \text{ GeV}, & \Gamma_Z &= 2.49426637877282 \text{ GeV}, \\ M_W &= 80.3579736098775 \text{ GeV}, & \Gamma_W &= 2.08429899827822 \text{ GeV}, \end{aligned} \quad (\text{I.53})$$

and

$$\alpha = 0.00755525416742918. \quad (\text{I.54})$$

Higgs bosons and top quarks do not appear as internal resonances in our calculation<sup>27</sup>, so we can set their widths equal to zero and use the following values for their masses,

$$\begin{aligned} M_H &= 125 \text{ GeV}, & \Gamma_H &= 0 \text{ GeV}, \\ m_t &= 173 \text{ GeV}, & \Gamma_t &= 0 \text{ GeV}. \end{aligned} \quad (\text{I.55})$$

All the remaining quarks,  $q = u, d, c, s, b$ , and all charged leptons and neutrinos are considered as light particles with mass equal to zero throughout the calculation. The Cabibbo-Kobayashi-Maskawa (CKM) matrix is set to unity. We note, however, that a block-diagonal form of Cabibbo type with mixing only between the two light quark generations would not affect our results.

<sup>27</sup>We note that this is no longer true for the top quark if photon-induced processes or QCD corrections are considered.

In contrast to the calculation of Ref. [225], where the deep-inelastic scattering (DIS) factorization scheme [592, 593] is used, in our present calculation we apply the  $\overline{\text{MS}}$  factorization scheme. The numerical impact of this difference turns out to be far below the permille level and thus phenomenologically irrelevant. This difference could not be resolved in the comparison of Ref. [499]. However, in order to compare with our present calculations, the results for ZZ production therein have been converted to the  $\overline{\text{MS}}$  factorization scheme.

To be consistent with Ref. [499], we only apply phase-space cuts on the (dressed) charged leptons. Namely, we require 4 (2) charged leptons in our calculations of off-shell ZZ (WW) production that fulfil the requirements

$$p_{\text{T},\ell_i} > 15 \text{ GeV} , \quad |y_{\ell_i}| < 2.5 , \quad (\text{I.56})$$

and are separated in the rapidity–azimuthal-angle plane,

$$\Delta R(\ell_i, \ell_j) = \sqrt{(\Delta y(\ell_i, \ell_j))^2 + (\Delta \phi(\ell_i, \ell_j))^2} > 0.2 , \quad (\text{I.57})$$

independently of the flavours and charges of the leptons  $\ell_i$  and  $\ell_j$ . The above cuts are applied on the level of dressed leptons, and the dressing procedure is defined as follows: if a photon is emitted close to one of the bare charged leptons in the rapidity–azimuthal-angle plane, such that

$$\Delta R(\ell_i, \gamma) = \sqrt{(\Delta y(\ell_i, \gamma))^2 + (\Delta \phi(\ell_i, \gamma))^2} < 0.2 , \quad (\text{I.58})$$

the photon is combined with the respective bare lepton  $\ell_i$ , and the momentum of the resulting dressed lepton is defined by simple addition of the involved four-momenta. If the above condition holds for more than one of the leptons, the photon is combined with the lepton that is closer in the  $y - \phi$  plane.

### 7.3.2 Point-wise comparison of virtual EW matrix elements

We first validate the different OLPs at the level of one-loop amplitudes for certain kinematic configurations, namely the phase-space points specified in Table I.9. In order to be independent of the subtraction scheme used to cancel IR divergences, we found it useful to compare the coefficients of a Laurent expansion of these matrix elements, which are regularized in  $D = 4 - 2\epsilon$  space-time dimensions in the 't Hooft–Veltman scheme. The virtual renormalized matrix element, summed over colour and helicity configurations, including averaging factors for initial states, can be written as

$$2\Re \left\{ \overline{\mathcal{M}_V^* \cdot \mathcal{M}_{\text{Born}}} \right\} = \frac{(4\pi)^\epsilon}{\Gamma(1-\epsilon)} \left( \frac{V_2}{\epsilon^2} + \frac{V_1(\mu)}{\epsilon} + V_{\text{finite}}(\mu) + \mathcal{O}(\epsilon) \right) . \quad (\text{I.59})$$

The dependence on both  $\epsilon$  and the regularization scale  $\mu$  will cancel only upon combination with the real-emission corrections, *e.g.* by the corresponding dependencies of the  $I$ -operator in a Catani–Seymour dipole-subtraction approach [187]. In order to be independent of subtraction schemes, we shall compare the relevant coefficients of the Laurent expansion in  $\epsilon$ ,  $V_{\text{finite}}(\mu)$ ,  $V_1(\mu)$ , and  $V_2$  for a given regularization scale,

$$\mu = M_{\text{Z/W}} , \quad (\text{I.60})$$

for off-shell ZZ and WW production, respectively. We start with the squared Born matrix elements, summed over colour and helicity configurations, including averaging factors for initial states, denoted as

$$B = \overline{|\mathcal{M}_{\text{Born}}|^2} , \quad (\text{I.61})$$

	$p_1$	$p_2$		$p_3$	$p_4$	$p_5$	$p_6$
a)	u	$\bar{u}$	$\rightarrow$	$e^+$	$e^-$	$\mu^+$	$\mu^-$
b)	u	$\bar{u}$	$\rightarrow$	$e^+$	$\nu_e$	$\mu^-$	$\bar{\nu}_\mu$
c)	$\gamma$	$\gamma$	$\rightarrow$	$e^+$	$e^-$	$\mu^+$	$\mu^-$
d)	$\gamma$	$\gamma$	$\rightarrow$	$e^+$	$\nu_e$	$\mu^-$	$\bar{\nu}_\mu$

PSP 1	$E$	$P_x$	$P_y$	$P_z$
$p_1$	5.0000000000000E+02	0.0000000000000E+00	0.0000000000000E+00	5.0000000000000E+02
$p_2$	5.0000000000000E+02	0.0000000000000E+00	0.0000000000000E+00	-5.0000000000000E+02
$p_3$	8.85513330545030E+01	-2.21006902876900E+01	4.00803531916853E+01	-7.58054309569366E+01
$p_4$	3.28329419227098E+02	-1.03849611883456E+02	-3.01933755389540E+02	7.64949213871659E+01
$p_5$	1.52358109467431E+02	-1.05880959666592E+02	-9.77096383269757E+01	4.95483852267928E+01
$p_6$	4.30761138250968E+02	2.31831261837738E+02	3.59563040524830E+02	-5.02378756570221E+01
PSP 2	$E$	$P_x$	$P_y$	$P_z$
$p_1$	5.0000000000000E+02	0.0000000000000E+00	0.0000000000000E+00	5.0000000000000E+02
$p_2$	5.0000000000000E+02	0.0000000000000E+00	0.0000000000000E+00	-5.0000000000000E+02
$p_3$	1.17747708137171E+02	-6.07059218463615E+01	7.12310458623266E+01	7.14524452324688E+01
$p_4$	3.50954173077969E+02	-3.17881667026108E+01	8.39394286931734E+01	3.39282354933456E+02
$p_5$	3.49332228573790E+02	1.84009303995763E+02	-5.15277979392370E+01	-2.92435408257719E+02
$p_6$	1.81965890211070E+02	-9.15152154467912E+01	-1.03642676616263E+02	-1.18299391908206E+02
PSP 3	$E$	$P_x$	$P_y$	$P_z$
$p_1$	5.0000000000000E+02	0.0000000000000E+00	0.0000000000000E+00	5.0000000000000E+02
$p_2$	5.0000000000000E+02	0.0000000000000E+00	0.0000000000000E+00	-5.0000000000000E+02
$p_3$	2.71200874120175E+02	-3.73192604194911E+01	3.27628636758588E+01	-2.66615419075953E+02
$p_4$	2.28710119031214E+02	3.01476738097433E+01	-2.76873495467817E+01	-2.25017437071458E+02
$p_5$	3.68223563778399E+02	-3.45262681797353E+01	-6.56193561878981E+00	3.66542590606031E+02
$p_6$	1.31865443070211E+02	4.16978547894831E+01	1.48642148971272E+00	1.25090265541380E+02
PSP 4	$E$	$P_x$	$P_y$	$P_z$
$p_1$	5.0000000000000E+02	0.0000000000000E+00	0.0000000000000E+00	5.0000000000000E+02
$p_2$	5.0000000000000E+02	0.0000000000000E+00	0.0000000000000E+00	-5.0000000000000E+02
$p_3$	6.57509213776183E+01	-1.80605085177932E+01	-3.57198237923879E+00	-6.31208573766775E+01
$p_4$	4.34316531884259E+02	9.08209885688625E+01	8.42255785958083E+01	-4.16279293039596E+02
$p_5$	4.45341565495701E+02	-7.55663459270628E+01	-4.72204111505749E+01	4.36335960118077E+02
$p_6$	5.45909812424219E+01	2.80586587599355E+00	-3.34331850659945E+01	4.30641902981974E+01
PSP 5	$E$	$P_x$	$P_y$	$P_z$
$p_1$	5.0000000000000E+02	0.0000000000000E+00	0.0000000000000E+00	5.0000000000000E+02
$p_2$	5.0000000000000E+02	0.0000000000000E+00	0.0000000000000E+00	-5.0000000000000E+02
$p_3$	3.17520530505573E+02	6.02726046666900E+01	3.06151548398959E+02	-5.88024645074751E+01
$p_4$	1.82384412506608E+02	4.00234094798358E+01	1.75895774032594E+02	2.68863775258899E+01
$p_5$	2.46378670867285E+02	-4.83680566904740E+01	-2.40294803208887E+02	-2.49276573532170E+01
$p_6$	2.53716386120534E+02	-5.19279574560520E+01	-2.41752519222667E+02	5.68437443348023E+01
PSP 6	$E$	$P_x$	$P_y$	$P_z$
$p_1$	5.0000000000000E+02	0.0000000000000E+00	0.0000000000000E+00	5.0000000000000E+02
$p_2$	5.0000000000000E+02	0.0000000000000E+00	0.0000000000000E+00	-5.0000000000000E+02
$p_3$	3.67309882280838E+02	1.65971559565154E+01	3.39808610366238E+01	3.65357886350532E+02
$p_4$	1.32530757794944E+02	-1.52987445917027E+01	-3.08889911842803E+01	1.27969607326260E+02
$p_5$	2.97450839881748E+01	9.82801281377121E+00	1.54935079102153E+01	-2.34122061803670E+01
$p_6$	4.70414275936040E+02	-1.11264241785838E+01	-1.85853777625588E+01	-4.69915287496428E+02

Table I.9: Phase-space points used for comparison between different matrix-element generators.

a) PSP 1	$B/10^{-15}$	$V_{\text{finite}}/10^{-16}$	$V_1/10^{-17}$	$V_2/10^{-17}$
MADLOOP	5.26592465401088	6.60297993618509	2.63915540074976	-3.09566543908773
RECOLA	5.26592465401090	6.60088670209820	2.63915540075328	-3.09566543908732
OPENLOOPS	5.26592465401100	6.60088670210145	2.63915540078563	-3.09566543905505
GoSAM	5.26592465401086	6.60088670209788	2.63915540076095	-3.09566543909091
NLOX	5.26592465401084	6.60088670211436	2.63915540076702	-3.09566543908783
a) PSP 2	$B/10^{-12}$	$V_{\text{finite}}/10^{-13}$	$V_1/10^{-14}$	$V_2/10^{-14}$
MADLOOP	2.74057953273116	-3.10720743529659	2.47558966660999	-1.61109736655361
RECOLA	2.74057953273120	-3.10783717792090	2.47558966661119	-1.61109736655360
OPENLOOPS	2.74057953273113	-3.10783717792216	2.47558966660688	-1.61109736655762
GoSAM	2.74057953273109	-3.10783717792575	2.47558966661326	-1.61109736655355
NLOX	2.74057953273088	-3.10783717791578	2.47558966660321	-1.61109736655852
a) PSP 3	$B/10^{-4}$	$V_{\text{finite}}/10^{-6}$	$V_1/10^{-7}$	$V_2/10^{-7}$
MADLOOP	1.21906911746527	-4.79121605677418	-9.28399419983122	-7.16650993758228
RECOLA	1.21906911746653	-4.77231274104044	-9.28399419025240	-7.16650993468800
OPENLOOPS	1.21906911746730	-4.77231273844357	-9.28399415556438	-7.16650990014111
GoSAM	1.21906911746070	-4.77231359778343	-9.28399407990066	-7.16650993856488
NLOX	1.21906911748497	-4.77231281258676	-9.28399522232122	-7.16651015319136
a) PSP 4	$B/10^{-6}$	$V_{\text{finite}}/10^{-7}$	$V_1/10^{-8}$	$V_2/10^{-8}$
MADLOOP	4.77962555243898	1.65145000279798	-3.61194825362166	-2.80978604924244
RECOLA	4.77962555246723	1.63956377750150	-3.61194826344888	-2.80978605025647
OPENLOOPS	4.77962555244817	1.63956377748191	-3.61194826326975	-2.80978605006768
GoSAM	4.77962555243871	1.63956377924796	-3.61194825975445	-2.80978604980914
NLOX	4.77962555244696	1.63956377842641	-3.61194826211540	-2.80978605014797

Table I.10: Matrix-element comparison at the phase-space points given in Table I.9 for the partonic process  $u\bar{u} \rightarrow e^+e^-\mu^+\mu^-$ .

in order to validate that all input parameters have been set correctly. As we shall see, it is interesting to perform the comparison both on far off-shell and on-shell phase-space points for each of the two different partonic channels considered, namely the respective  $u\bar{u}$ -induced and  $\gamma\gamma$ -induced channels of each of the processes.

In Table I.9 we first list the two far off-shell kinematic configurations labelled PSP 1 and PSP 2 that apply to all processes. We then consider the points PSP 3 and PSP 4 for the  $e^+e^-\mu^+\mu^-$  production processes a) and c) where the invariant mass of the muon-anti-muon pair is exactly at the pole mass of the Z boson, and finally the points PSP 5 and PSP 6 for the  $e^+\nu_e\mu^-\bar{\nu}_\mu$  processes b) and d) where the invariant mass of the positron and its neutrino sits at exactly the pole mass of the W boson.

The numerical results for  $B$ ,  $V_{\text{finite}}$ ,  $V_1$ , and  $V_2$  obtained with the different involved OLPs (MADLOOP, RECOLA, OPENLOOPS, GoSAM, and NLOX) are displayed in Tables I.10–I.13. In general, an agreement among all OLPs of more than 10 digits for off-shell kinematics (typically a bit worse for on-shell kinematics) is found for the quantities  $B$ ,  $V_1$ , and  $V_2$ , whereas  $V_{\text{finite}}$  obtained with MADLOOP shows differences of up to the percent level compared to the corresponding results from the other programs, which show reasonable agreement among themselves. Such small deviations are to be expected due to differences in the details of the implementa-

b) PSP 1	$B/10^{-13}$	$V_{\text{finite}}/10^{-14}$	$V_1/10^{-15}$	$V_2/10^{-16}$
MADLOOP	1.26112388530353	-3.30864987248181	1.21073012235170	-4.38084412126839
RECOLA	1.26112388530350	-3.30882625694720	1.21073012235186	-4.38084412126840
OPENLOOPS	1.26112388530352	-3.30882625694700	1.21073012234675	-4.38084412131955
GoSAM	1.26112388530352	-3.30882625695049	1.21073012235248	-4.38084412126866
NLOX	1.26112388530352	-3.30882625694586	1.21073012235199	-4.38084412126837
b) PSP 2	$B/10^{-12}$	$V_{\text{finite}}/10^{-13}$	$V_1/10^{-14}$	$V_2/10^{-14}$
MADLOOP	9.19168012745872	-8.42372887318610	1.52668534752201	-3.19297083499976
RECOLA	9.19168012745880	-8.42533896054780	1.52668534751979	-3.19297083499981
OPENLOOPS	9.19168012745858	-8.42533896054679	1.52668534751231	-3.19297083500727
GoSAM	9.19168012745844	-8.42533896055772	1.52668534752199	-3.19297083499963
NLOX	9.19168012745785	-8.42533896056472	1.52668534750067	-3.19297083500005
b) PSP 5	$B/10^{-6}$	$V_{\text{finite}}/10^{-7}$	$V_1/10^{-8}$	$V_2/10^{-9}$
MADLOOP	2.66444601710804	-6.26911810540777	2.59441472204941	-9.25565109671323
RECOLA	2.66444601710734	-6.27314341024947	2.59441472291996	-9.25565109556038
OPENLOOPS	2.66444601710704	-6.27314341020842	2.59441471256317	-9.25565118709011
GoSAM	2.66444601710846	-6.27314341015885	2.59441472307111	-9.25565109542055
NLOX	2.66444601705951	-6.27314341916503	2.59441471119468	-9.25565111942051
b) PSP 6	$B/10^{-3}$	$V_{\text{finite}}/10^{-5}$	$V_1/10^{-5}$	$V_2/10^{-6}$
MADLOOP	2.08107330428286	-3.94303566134416	-1.76896206881657	-7.22915309068623
RECOLA	2.08107330429484	-3.96855804139705	-1.76896184875300	-7.22915318450819
OPENLOOPS	2.08107330428676	-3.96855798524525	-1.76896146354940	-7.22915318412123
GoSAM	2.08107330364288	-3.96852693676055	-1.76896416151949	-7.22915263418978
NLOX	2.08107341495760	-3.96853794693293	-1.76895854332064	-7.22915326559073

Table I.11: Matrix-element comparison at the phase-space points given in Table I.9 for the partonic process  $u\bar{u} \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu$ .

tion of the complex-mass scheme renormalization. In order to avoid complications in choosing the correct Riemann sheet when evaluating two-point functions of self-energies with complex arguments, it was proposed in Ref. [541] to expand these functions in the EW coupling (see Eq. (8) in Ref. [541]), and to truncate the series at the NLO EW level. This scheme has been adopted by RECOLA, OPENLOOPS, GoSAM and NLOX. Conversely, MADLOOP chooses not to perform such an expansion, and instead considers the *full* two-point function evaluated on the appropriate Riemann sheet. The details of this procedure will be documented in a forthcoming publication [508].

We stress that both approaches to handle the evaluation of self-energies in presence of the complex-mass scheme renormalization conditions are correct as they lead to identical results up to NLO EW accuracy. It is important, however, to verify that the differences between MADLOOP and the other involved OLPs are indeed formally beyond the NLO EW accuracy targeted. To this end we propose here to numerically investigate the asymptotic behaviour of  $V_{\text{finite}}$  in the limit of asymptotically vanishing EW couplings. This can be done numerically by rescaling the electromagnetic coupling constant and the related parameters (*i.e.*  $G_\mu, \alpha, \Gamma_Z, \Gamma_W$ ) by a parameter  $\lambda$ ,

$$G_\mu \rightarrow \lambda G_\mu, \quad \alpha \rightarrow \lambda \alpha, \quad \Gamma_Z \rightarrow \lambda \Gamma_Z, \quad \Gamma_W \rightarrow \lambda \Gamma_W, \quad (\text{I.62})$$

c) PSP 1	$B/10^{-13}$	$V_{\text{finite}}/10^{-14}$	$V_1/10^{-15}$	$V_2/10^{-15}$
MADLOOP	4.63762790127829	6.79330655006349	4.07216839247769	-2.23061748556626
RECOLA	4.63762790127830	6.79163662486900	4.07216839245629	-2.23061748556050
OPENLOOPS	4.63762790127838	6.79163662486753	4.07216839246097	-2.23061748560388
GoSAM	4.63762790127830	6.79163662486761	4.07216839247955	-2.23061748556541
c) PSP 2	$B/10^{-10}$	$V_{\text{finite}}/10^{-11}$	$V_1/10^{-12}$	$V_2/10^{-11}$
MADLOOP	2.26737153141645	1.88180804083847	2.38584215397888	-1.09056584355518
RECOLA	2.26737153141650	1.88096074990550	2.38584215406775	-1.09056584355509
OPENLOOPS	2.26737153141649	1.88096075053150	2.38584215383146	-1.09056584370294
GoSAM	2.26737153141644	1.88096075053592	2.38584215397731	-1.09056584355520
c) PSP 3	$B/10^{-6}$	$V_{\text{finite}}/10^{-6}$	$V_1/10^{-9}$	$V_2/10^{-9}$
MADLOOP	1.37978612284930	1.55018919031339	4.89785291501769	-6.63652265273678
RECOLA	1.37978612284863	1.55013518201790	4.89788114834105	-6.63652866815621
OPENLOOPS	1.37978612284923	1.55013518232261	4.89788114988480	-6.63652866830139
GoSAM	1.37978612284092	1.55011760547612	4.89816579319493	-6.63658173170046
c) PSP 4	$B/10^{-7}$	$V_{\text{finite}}/10^{-6}$	$V_1/10^{-10}$	$V_2/10^{-9}$
MADLOOP	2.19037672578717	1.68165624485106	7.61526199100670	-1.05353269570773
RECOLA	2.19037672578999	1.68164383557005	7.61526198688565	-1.05353269571203
OPENLOOPS	2.19037672578763	1.68164383554897	7.61526198357373	-1.05353269576734
GoSAM	2.19037672578690	1.68164383554095	7.61526199187791	-1.05353269572332

Table I.12: Matrix-element comparison at the phase-space points given in Table I.9 for the partonic process  $\gamma\gamma \rightarrow e^+e^-\mu^+\mu^-$ .

while other model parameters such as  $(M_Z, M_W, M_H, m_t)$  remain fixed. Within this setup, we then compute the relative difference of  $V_{\text{finite}}(\lambda)$  between RECOLA/OPENLOOPS<sup>28</sup> and MADLOOP,

$$\delta V_{\text{finite}}(\lambda) = 2 \left| \frac{V_{\text{finite}}^{\text{MADLOOP}}(\lambda) - V_{\text{finite}}^{\text{RECOLA/OPENLOOPS}}(\lambda)}{V_{\text{finite}}^{\text{MADLOOP}}(\lambda) + V_{\text{finite}}^{\text{RECOLA/OPENLOOPS}}(\lambda)} \right| \quad (\text{I.63})$$

at different values of  $\lambda$ , as displayed in Figs. I.15 and I.16. The expected scaling of the relative difference  $\delta V_{\text{finite}}(\lambda)$  with the parameter  $\lambda$  is dictated by the expected contribution of the UV mass counterterms  $\delta m$ , which we can schematically write as follows,

$$V_{\text{finite}}^{\delta m} \propto \frac{1}{p_V^2 - m_V^2 + im_V\Gamma_V} \delta m \frac{1}{p_V^2 - m_V^2 + im_V\Gamma_V}. \quad (\text{I.64})$$

The assumption is that the UV mass counterterms  $\delta m$  in MADLOOP only differs from the ones of the other OLPs by terms of order  $\mathcal{O}(\alpha\Gamma^2, \alpha^2\Gamma)$ . We can substitute this functional form of the difference in Eqs. (I.64) and (I.63) to deduce the expected dependence of  $\delta V_{\text{finite}}(\lambda)$  on the scaling parameter  $\lambda$  for  $\lambda \rightarrow 0$ ,

$$\delta V_{\text{finite}}(\lambda) = \lambda^\kappa \delta V_{\text{finite}}^{(\kappa)} + \mathcal{O}(\lambda^{\kappa+1}), \quad (\text{I.65})$$

<sup>28</sup>Given that the implementation in all involved OLPs except for MADLOOP is identical, we restrict ourselves to the results of RECOLA and OPENLOOPS in this comparison: We use their point-wise average for the central values, and the difference between them as an estimate for the numerical precision.

d) PSP 1	$B/10^{-13}$	$V_{\text{finite}}/10^{-15}$	$V_1/10^{-15}$	$V_2/10^{-16}$
MADLOOP	3.67352032122512	6.79828192861256	4.15882308803665	-8.83449344853984
RECOLA	3.67352032122520	6.78860727240780	4.15882308802252	-8.83449344853268
OPENLOOPS	3.67352032122515	6.78860727241689	4.15882308802947	-8.83449344826758
GoSAM	3.67352032122515	6.78860727251103	4.15882308806114	-8.83449344857124
d) PSP 2	$B/10^{-11}$	$V_{\text{finite}}/10^{-14}$	$V_1/10^{-12}$	$V_2/10^{-13}$
MADLOOP	9.78728443151128	2.83565594714426	1.43919663402090	-2.35375586979757
RECOLA	9.78728443151120	2.52764290726850	1.43919663402107	-2.35375586979726
OPENLOOPS	9.78728443151125	2.52764290753401	1.43919663402070	-2.35375586979729
GoSAM	9.78728443151106	2.52764290829874	1.43919663401982	-2.35375586979631
d) PSP 5	$B/10^{-5}$	$V_{\text{finite}}/10^{-5}$	$V_1/10^{-6}$	$V_2/10^{-7}$
MADLOOP	8.56103162287602	-2.08508067115892	1.40092813161870	-2.05885284890288
RECOLA	8.56103162287098	-2.08643204830236	1.40092813169621	-2.05885284880211
OPENLOOPS	8.56103162287173	-2.08643204831341	1.40092813323908	-2.05885283530621
GoSAM	8.56103162287777	-2.08643204828055	1.40092813175874	-2.05885284872070
d) PSP 6	$B/10^{-1}$	$V_{\text{finite}}/10^{-1}$	$V_1/10^{-3}$	$V_2/10^{-4}$
MADLOOP	2.30984277049480	1.15802573980481	2.61399257786962	-5.55496881434693
RECOLA	2.30984277049406	1.15474814667339	2.61399259319492	-5.55496881453983
OPENLOOPS	2.30984277049366	1.15474813504440	2.61399265092704	-5.55496797160493
GoSAM	2.30984277049006	1.15420829481100	2.61399257843170	-5.55496881427150

Table I.13: Matrix-element comparison at the phase-space points given in Table I.9 for the partonic process  $\gamma\gamma \rightarrow e^+\nu_e\mu^-\bar{\nu}_\mu$ .

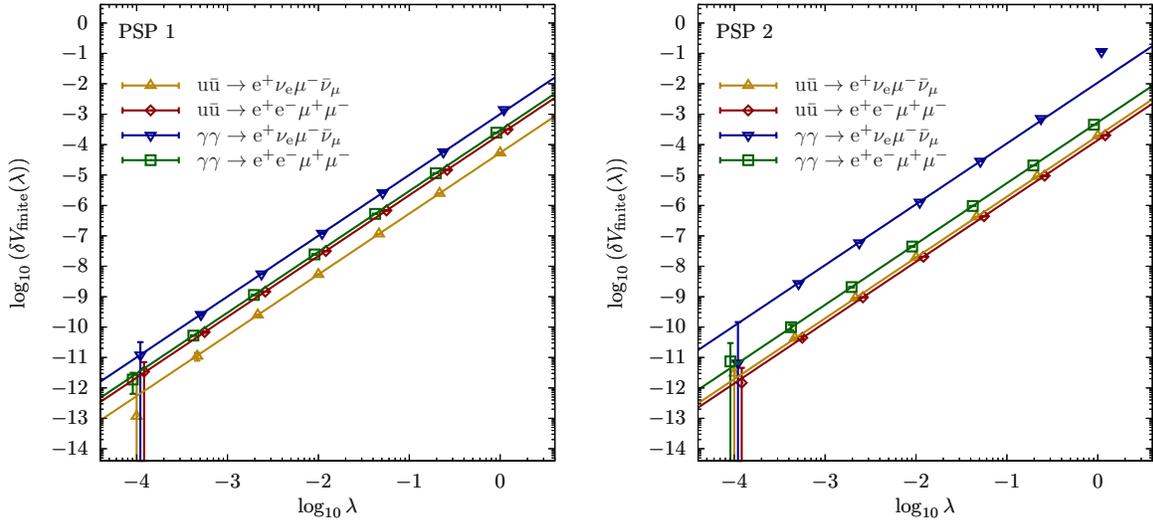


Fig. I.15: Relative differences of  $V_{\text{finite}}(\lambda)$  between MADLOOP and the average of OPENLOOPS and RECOLA,  $\delta V_{\text{finite}}(\lambda)$ , for the two far off-shell points PSP 1 and PSP 2, accompanied by lines corresponding to a quadratic  $\lambda$  dependence of  $\delta V_{\text{finite}}(\lambda)$ . See main text for more details.

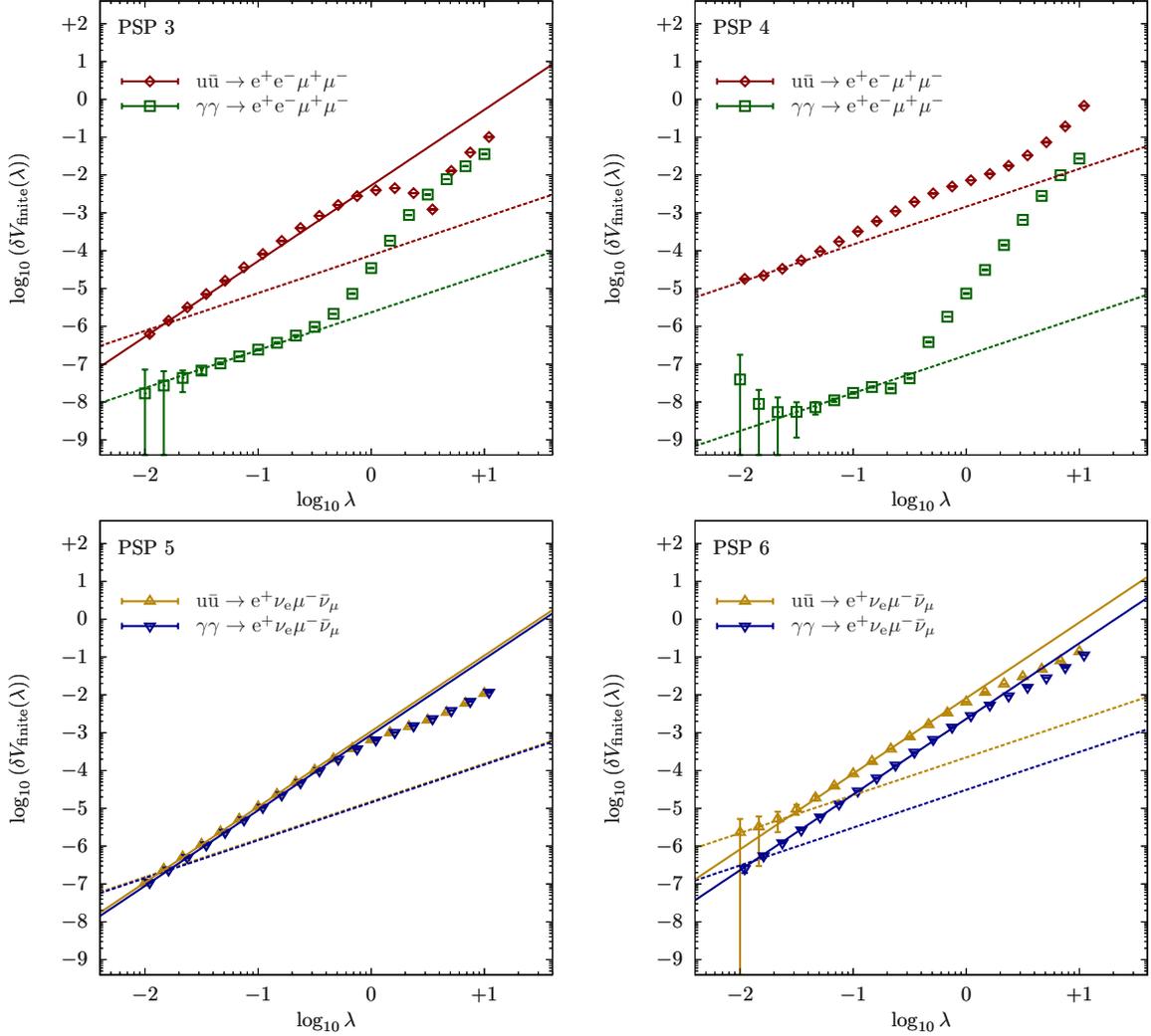


Fig. I.16: Relative differences of  $V_{\text{finite}}(\lambda)$  between MADLOOP and the average of OPENLOOPS and RECOLA,  $\delta V_{\text{finite}}(\lambda)$ , for the on-shell configurations PSP 3+4 (ZZ, upper plots) and PSP 5+6 (WW, lower plots), accompanied by solid or dashed lines corresponding to a quadratic or linear  $\lambda$  dependence of  $\delta V_{\text{finite}}(\lambda)$ , respectively. See main text for more details.

with  $\kappa = 1$  for exactly on-shell kinematic configurations ( $p_V^2 = m_V^2$ ) and  $\kappa = 2$  otherwise. The coefficients  $\delta V_{\text{finite}}^{(\kappa)}$  are by definition independent of  $\lambda$ .

Figure I.15 shows the results obtained for  $\delta V_{\text{finite}}(\lambda)$  at progressively smaller values of  $\lambda$  for the two off-shell points PSP 1+2. Figure I.16 deals with the on-shell points PSP 3+4 for ZZ production and PSP 5+6 for WW production, respectively. The chosen  $\lambda$  values are identical for all partonic processes in each frame, but slightly shifted horizontally in Figs. I.15 and I.16 to improve readability. The error bars on the relative difference  $\delta V_{\text{finite}}(\lambda)$  are extracted from the internal numerical estimate of MADLOOP on the one side, and as the difference between OPENLOOPS and RECOLA, which should be entirely due to numerical inaccuracies, on the other side, resulting in an overall error estimate on  $\delta V_{\text{finite}}(\lambda)$  itself. The numerical investigation of the on-shell phase-space points is particularly challenging because the propagators involved in both the numerator and denominator of Eq. (I.63) diverge when  $\lambda \rightarrow 0$  in a way that crucially depends on the numerical accuracy at which  $p_V^2 = m_V^2$  is realised. Moreover, the self-energy

evaluation can also be numerically less stable in that regime (in the case of the W boson *e.g.* it involves terms of order  $\mathcal{O}(\alpha \log(\Gamma))$ ). The above-mentioned estimates of the numerical stability point towards  $\log_{10} \lambda \gtrsim -4$  in the off-shell case and  $\log_{10} \lambda \gtrsim -2$  in the on-shell case being appropriate ranges for studying  $\delta V_{\text{finite}}(\lambda)$ .

Figures I.15 and I.16 demonstrate that  $\delta V_{\text{finite}}(\lambda)$  between MADLOOP and OPENLOOPS indeed follow the expected scaling of Eq. (I.65) (at some phase-space points with even greater values of  $\kappa$ ), confirming that they stand beyond the target NLO EW accuracy in *all* kinematic regimes and even beyond NNLO EW accuracy in the off-shell regions. We stress that within the complex-mass scheme, virtual matrix elements of NLO EW corrections always have contributions beyond  $\mathcal{O}(\alpha)$ , and there is *a priori* no argument for deciding which incomplete set of beyond NLO contributions better approximates the all-order result. The difference in the finite part of the virtual matrix element between MADLOOP and the other OLPs should therefore be considered, if anything, as a systematic theoretical uncertainty, similarly to renormalization scale variations in QCD. We note, however, that these small differences in the virtual matrix element do not lead to resolvable effects in any of the differential observables presented in the context of this comparison (see Sec. 7.3.4).

### 7.3.3 Integrated cross sections

The comparison of amplitudes at selected phase-space points in Sec. 7.3.2 essentially validates the matrix-element generators (apart from their correct implementation and initialisation within the corresponding integrators). In contrast, integrated cross sections focus on the validation of the phase-space integrators and the performance of their implementations of IR-subtraction techniques. Naturally, such comparisons on integrated level — even more so at the level of distributions (discussed in Sec. 7.3.4) — might also uncover issues in the stable evaluation of matrix elements in certain kinematical regions. However, to keep control of such issues, most amplitude generators introduced in Sec. 7.2.2 employ internal stability checks for each evaluated matrix element to guarantee that their numerical accuracy is sufficiently high such that phenomenological results are not affected by possible numerical instabilities.

In this spirit, the results obtained from RECOLA, OPENLOOPS, and GOSAM when interfaced to SHERPA have been combined, and are presented as a single result in the following presentation. This is justified by the fact that all contributions apart from the virtual correction within SHERPA are completely independent of the respective matrix-element generator, and their comparison would essentially not go beyond an internal check of SHERPA. Before combination, the statistical compatibility of the integrated virtual contributions obtained with different OLPs was verified. Consequently, only a single result labelled SHERPA+GOSAM/OPENLOOPS/RECOLA is shown. On the other side, it does make sense to involve integrated results obtained from different integration frameworks that employ the same one-loop matrix elements in the comparison.

We apply the input parameters and phase-space cuts specified in Sec. 7.3.1. The integrated results for off-shell ZZ production, which also involve the result of Refs. [225, 499]<sup>29</sup>, are collected in Table I.14, and those for off-shell WW production in Table I.15. For each of the applied frameworks we state integrated cross sections at LO ( $\sigma^{\text{LO}}$ ) and NLO EW ( $\sigma_{\text{EW}}^{\text{NLO}}$ ) accuracy with the corresponding statistical errors  $\delta\sigma^{\text{LO}}$  and  $\delta\sigma_{\text{EW}}^{\text{NLO}}$  indicated in square brackets, referring to the last stated digits of the respective results. Moreover, we provide a weighted average, defined

<sup>29</sup>We note again that those results, denoted here as MCB+RECOLA, have been converted from the DIS to the  $\overline{\text{MS}}$  factorization scheme. Besides, we recall that in Ref. [225] a different treatment of the electromagnetic coupling  $\alpha$  was chosen for real photon radiation. These effects have been compensated for, which explains the numerical difference to the cross sections presented here. The SHERPA+RECOLA results from Ref. [499] have not been used here, but were generated anew for the present comparison.

$pp \rightarrow e^+e^-\mu^+\mu^-$	$\sigma^{\text{LO}}$	$\sigma_{\text{EW}}^{\text{NLO}}$	$\Delta\sigma^{\text{LO}}$		$\Delta\sigma_{\text{EW}}^{\text{NLO}}$	
	[fb]	[fb]	[ $\sigma$ ]	[‰]	[ $\sigma$ ]	[‰]
average	11.49675[8]	10.88697[15]				
MCBB+RECOLA	11.49648[12]	10.88669[22]	-2.9	-0.02	-1.7	-0.03
MUNICH+OPENLOOPS	11.49702[11]	10.88720[25]	+3.2	+0.02	+1.2	+0.02
MoCANLO+RECOLA	11.49666[26]	10.88734[56]	-0.3	-0.01	+0.7	+0.03
SHERPA+GoSAM/OPENLOOPS/RECOLA	11.49670[34]	10.88737[77]	-0.1	-0.00	+0.5	+0.04
MADGRAPH5_AMC@NLO+MADLOOP	11.4956[22]	10.8860[63]	-0.5	-0.10	-0.1	-0.09

Table I.14: Comparison of integrated cross sections for hadronic  $e^+e^-\mu^+\mu^-$  (off-shell  $ZZ$ ) production, obtained with different integration frameworks.

$pp \rightarrow e^+\nu_e\mu^-\bar{\nu}_\mu$	$\sigma^{\text{LO}}$	$\sigma_{\text{EW}}^{\text{NLO}}$	$\Delta\sigma^{\text{LO}}$		$\Delta\sigma_{\text{EW}}^{\text{NLO}}$	
	[fb]	[fb]	[ $\sigma$ ]	[‰]	[ $\sigma$ ]	[‰]
average	448.5414[31]	438.1902[56]				
MUNICH+OPENLOOPS	448.5468[45]	438.1920[75]	+1.6	+0.01	+0.4	+0.00
MoCANLO+RECOLA	448.538[10]	438.193[13]	-0.4	-0.01	+0.2	+0.01
SHERPA+GoSAM/OPENLOOPS/RECOLA	448.5364[46]	438.186[11]	-1.4	-0.01	-0.4	-0.01
MADGRAPH5_AMC@NLO	448.541[40]	438.113[70]	-0.0	-0.00	-1.1	-0.18

Table I.15: Comparison of integrated cross sections for hadronic  $e^+\nu_e\mu^-\bar{\nu}_\mu$  (off-shell  $WW$ ) production, obtained with different integration frameworks.

from the individual cross sections  $\sigma_i^x$  and their standard deviations  $\delta\sigma_i^x$  via

$$\bar{\sigma}^x = (\delta\bar{\sigma}^x)^2 \sum_i \frac{\sigma_i^x}{(\delta\sigma_i^x)^2}, \quad \delta\bar{\sigma}^x = \left( \sum_i \frac{1}{(\delta\sigma_i^x)^2} \right)^{-\frac{1}{2}} \quad x = \text{LO, NLO EW}. \quad (\text{I.66})$$

The deviations for both LO ( $\Delta\sigma^{\text{LO}}$ ) and NLO EW ( $\Delta\sigma_{\text{EW}}^{\text{NLO}}$ ) cross sections with respect to the averaged cross sections are quantified both in terms of standard deviations [ $\sigma$ ]<sup>30</sup> and as relative deviations in permille [‰] in order to validate both the statistical agreement and the level of precision on which this agreement could be achieved.

For both processes we find good statistical agreement between all compared codes on the sub-permille level. For MCBB+RECOLA (only for off-shell  $ZZ$ ), MUNICH+OPENLOOPS, MoCANLO+RECOLA, and SHERPA+GoSAM/OPENLOOPS/RECOLA we are able to validate integrated results even on the remarkable level of better than a tenth of a permille throughout, which corresponds to a relative agreement of the integrated EW corrections at least at the level of a permille.

### 7.3.4 Differential cross sections

In order to validate the different calculations not only in the resonance regions that dominate the integrated cross sections discussed in the previous section, we perform also a comparison at the

<sup>30</sup>To compensate for the fact that this average depends on each single calculation  $i$ , this comparison is performed in units of

$$\sqrt{(\delta\sigma_i^x)^2 - (\delta\bar{\sigma}^x)^2}, \quad (\text{I.67})$$

which corresponds to comparing each single calculation with the average of the remaining ones, with result and standard deviation of this average defined according to Eq. (I.66).

level of differential distributions. We have selected (pseudo-)observables that investigate different regions of phase-space with certain peculiarities of the respective EW corrections: invariant-mass distributions involving possible resonances exhibit large corrections of QED type, which are driven by photon radiation off the involved leptons; high-energy tails of invariant-mass and transverse-momentum distributions undergo large negative corrections due to EW Sudakov logarithms, which are genuine weak effects and dominated by the virtual contributions; finally, we also investigate azimuthal-angle distributions as examples for observables whose EW corrections are only weakly phase-space dependent.

We start with a selection of distributions for off-shell ZZ production, which are shown in Figs. I.17–I.21. All plots are organised as follows: The top–left frame depicts the absolute predictions to give an idea of the overall size of the cross section in the respective region of phase space, both at LO and NLO EW accuracy; we note here that we consider observables whose cross sections cover a range of up to five orders of magnitude. The top–right frame shows the relative EW corrections, which vary between about  $-40\%$  in the high- $p_T$  tails of lepton transverse-momentum distributions and about  $+60\%$  in invariant-mass distributions below resonance regions. In both frames, all five individual predictions for the respective differential distribution are plotted, namely MCBB+RECOLA (blue), MUNICH+OPENLOOPS (red), MOCANLO+RECOLA (green), SHERPA+GOSAM/OPENLOOPS/RECOLA (yellow), and MADGRAPH5\_AMC@NLO+MADLOOP (cyan). The average of the five calculations, defined binwise in analogy to the integrated cross sections in the previous section, is plotted on top (black). Absolute predictions are shown both at LO (dashed) and NLO EW (solid) accuracy. To illustrate the relative as well as the statistical agreement, we perform pairwise comparisons between the individual predictions in the frames below: On the left-hand side, in each frame the ratio of two predictions  $i$  and  $j$  over the overall average is plotted, accompanied by their respective  $1\sigma$  error bands, *i.e.*

$$(d\sigma_{i/j} \pm \delta d\sigma_{i/j})/d\bar{\sigma}. \quad (\text{I.68})$$

On the right-hand side, we show the deviation between the same two individual predictions  $i$  and  $j$  in units of their combined standard deviation,

$$\delta d\sigma_{ij} = \sqrt{\delta d\sigma_i^2 + \delta d\sigma_j^2}, \quad (\text{I.69})$$

*i.e.* the plots are by construction symmetric. Since the available results from MCBB+RECOLA, MUNICH+OPENLOOPS, MOCANLO+RECOLA, and SHERPA+GOSAM/OPENLOOPS/RECOLA are typically about one order of magnitude more precise than the results provided by MADGRAPH5\_AMC@NLO+MADLOOP, we start with their pairwise comparisons in frames 2–7. In frame 8, we compare their average to the results from MADGRAPH5\_AMC@NLO+MADLOOP.

Figures I.17 and I.18 depict the invariant mass of the 4-lepton system, which exhibits the  $Z \rightarrow 4\ell$  peak around  $m_Z$  and the ZZ production threshold at  $m_{4\ell} \gtrsim 2m_Z$ , with the corresponding photon-radiation dominated EW corrections around the resonances, detailed in Fig. I.17. In the high-energy region, shown in Fig. I.18, the NLO EW corrections are driven by EW Sudakov logarithms, and grow negatively to  $-20\%$  at  $m_{4\ell} \approx 1$  TeV. In the full range, all five predictions agree very well on a statistical level. For the upper four calculations, this corresponds to a permille-level agreement throughout, apart from the region below the  $Z \rightarrow 4\ell$  resonance where the cross section is suppressed by about four orders of magnitude compared to the peak region.

Figure I.19 shows the invariant mass of the  $\mu^+\mu^-$  system, which is typically produced via an intermediate Z boson. The chosen range  $m_{\mu^+\mu^-} \in [60; 260]$  GeV details both the resonance region with its QED-dominated NLO EW corrections and an intermediate phase-space region where the pair is far off shell, but the dominance of EW Sudakov logarithms does not set in yet. We find agreement on a statistical level between all individual predictions, which again

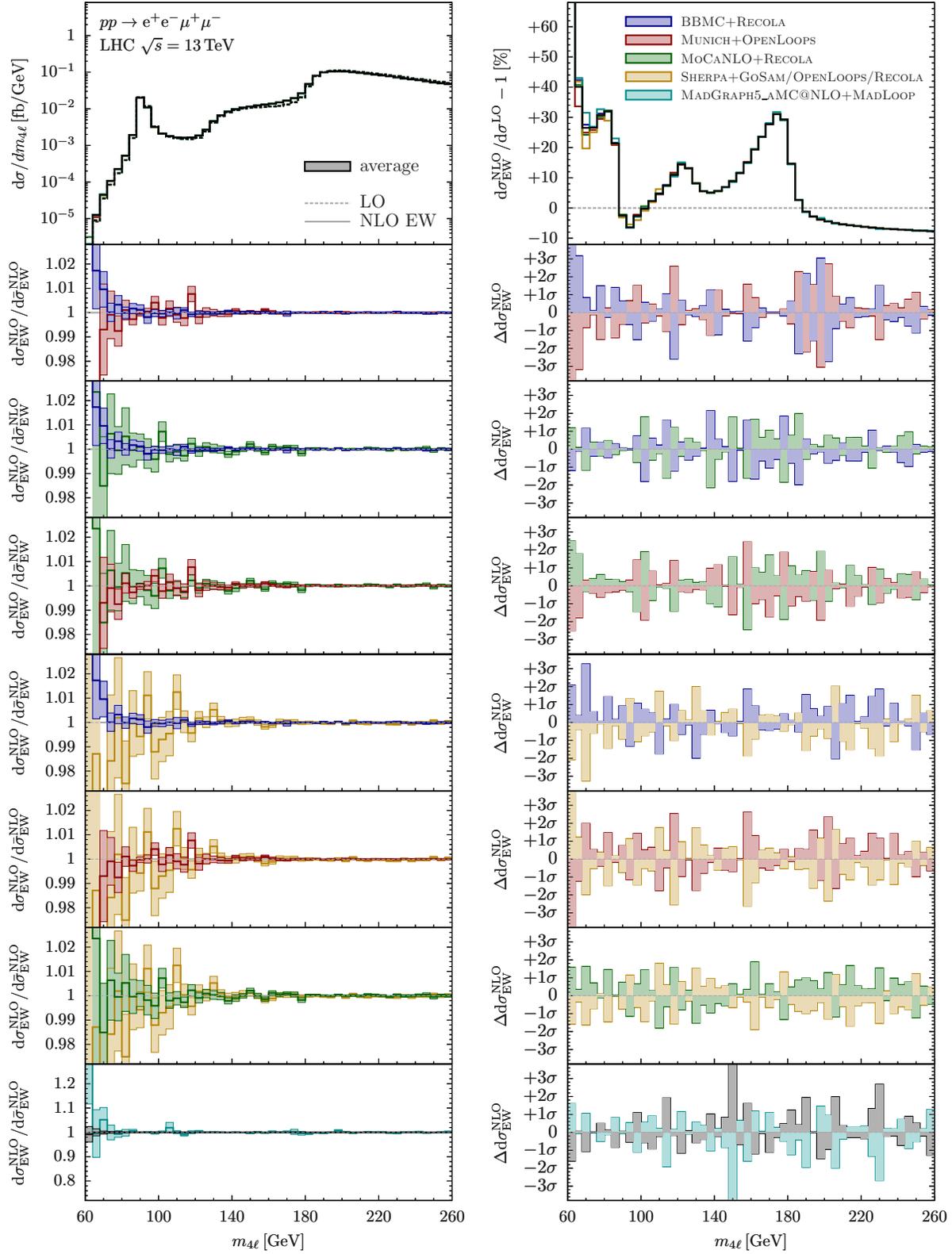


Fig. I.17: Technical comparison of NLO EW corrections to the distribution in the invariant mass of the  $4\ell$  system (resonance/threshold region) for hadronic  $e^+e^-\mu^+\mu^-$  (off-shell  $ZZ$ ) production. See main text of Sec. 7.3.4 for details.

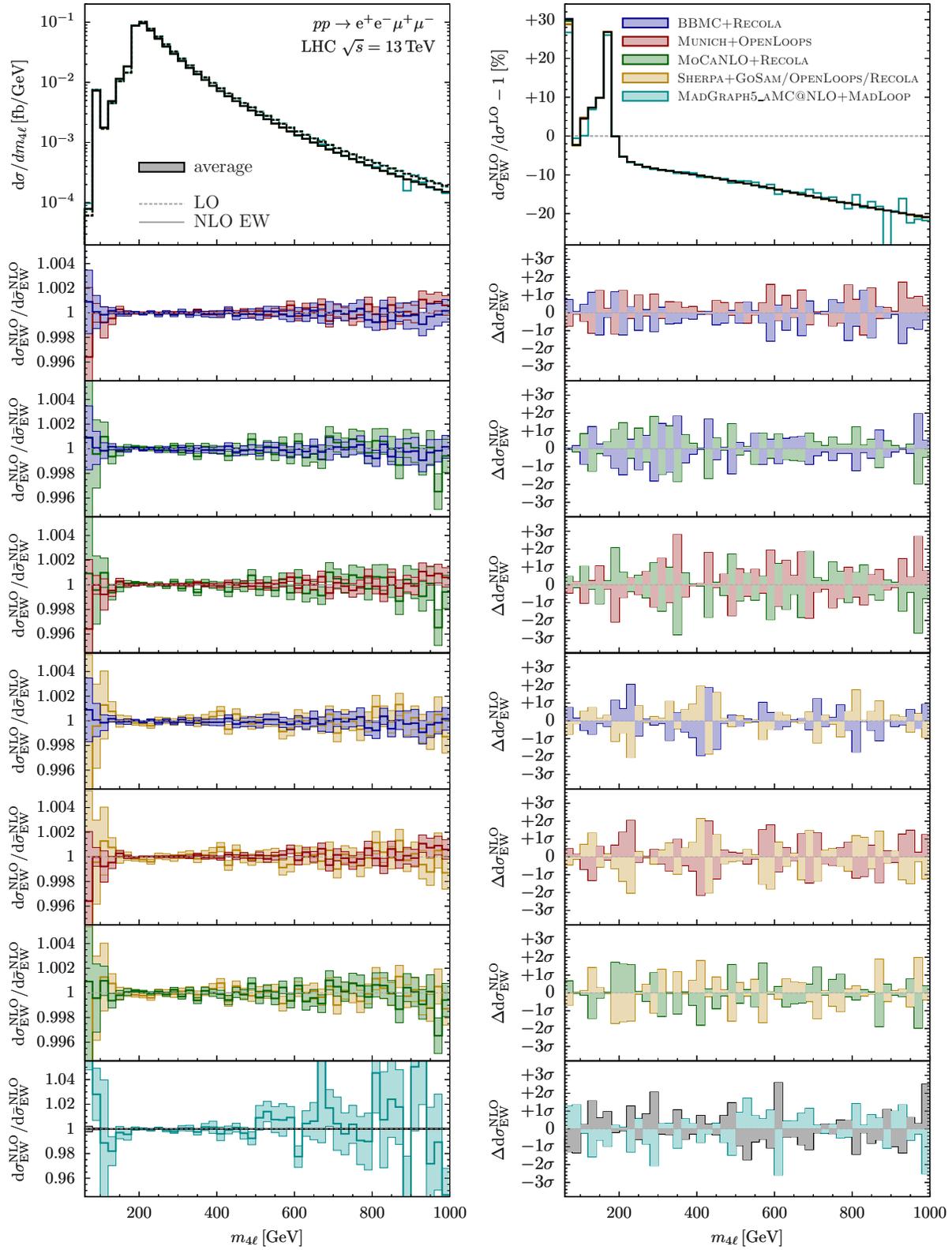


Fig. I.18: Technical comparison of NLO EW corrections to the distribution in the invariant mass of the  $4\ell$  system (high-energy region) for hadronic  $e^+e^-\mu^+\mu^-$  (off-shell  $ZZ$ ) production. See main text of Sec. 7.3.4 for details.

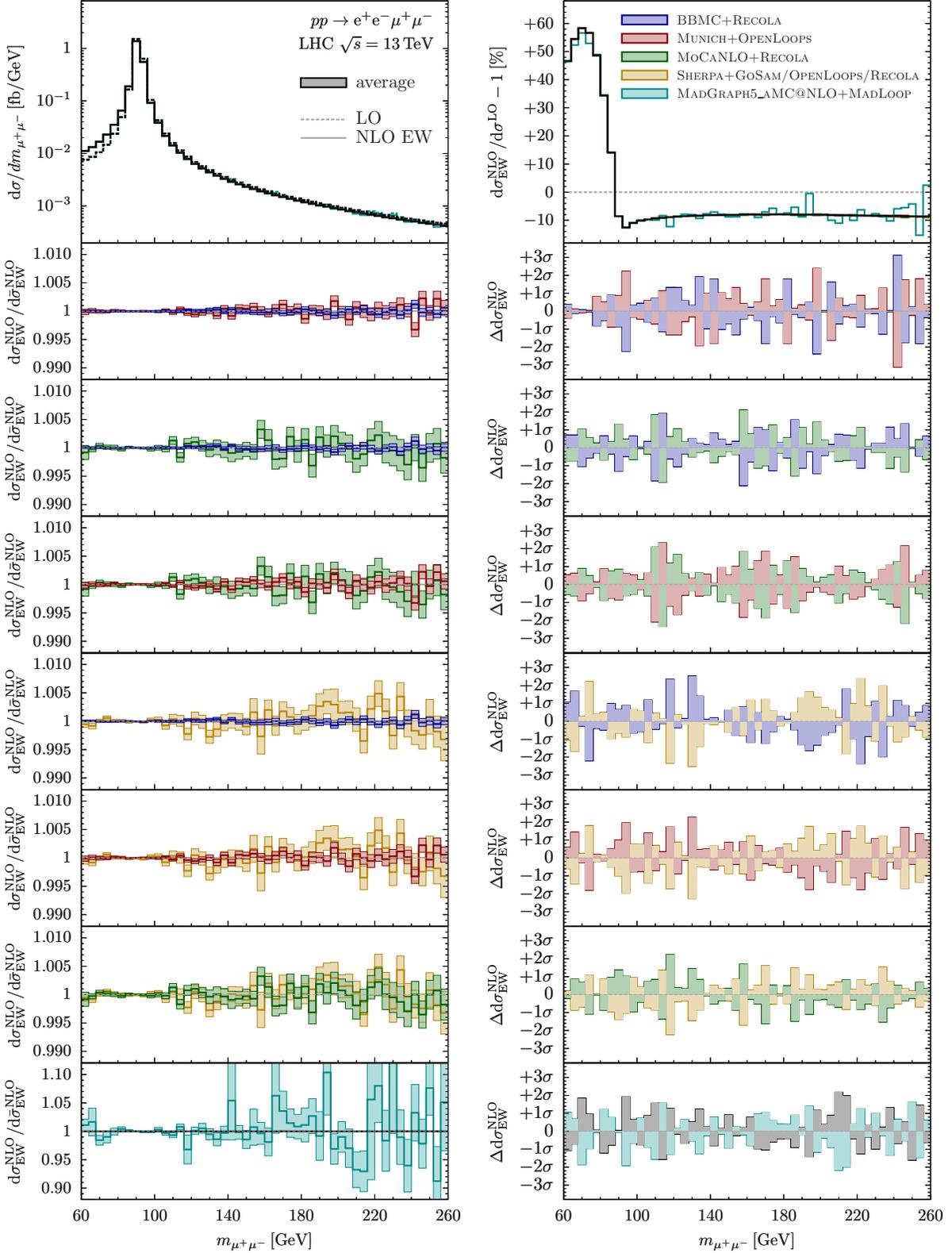


Fig. I.19: Technical comparison of NLO EW corrections to the distribution in the invariant mass of the  $\mu^+\mu^-$  pair for hadronic  $e^+e^-\mu^+\mu^-$  (off-shell  $ZZ$ ) production. See main text of Sec. 7.3.4 for details.

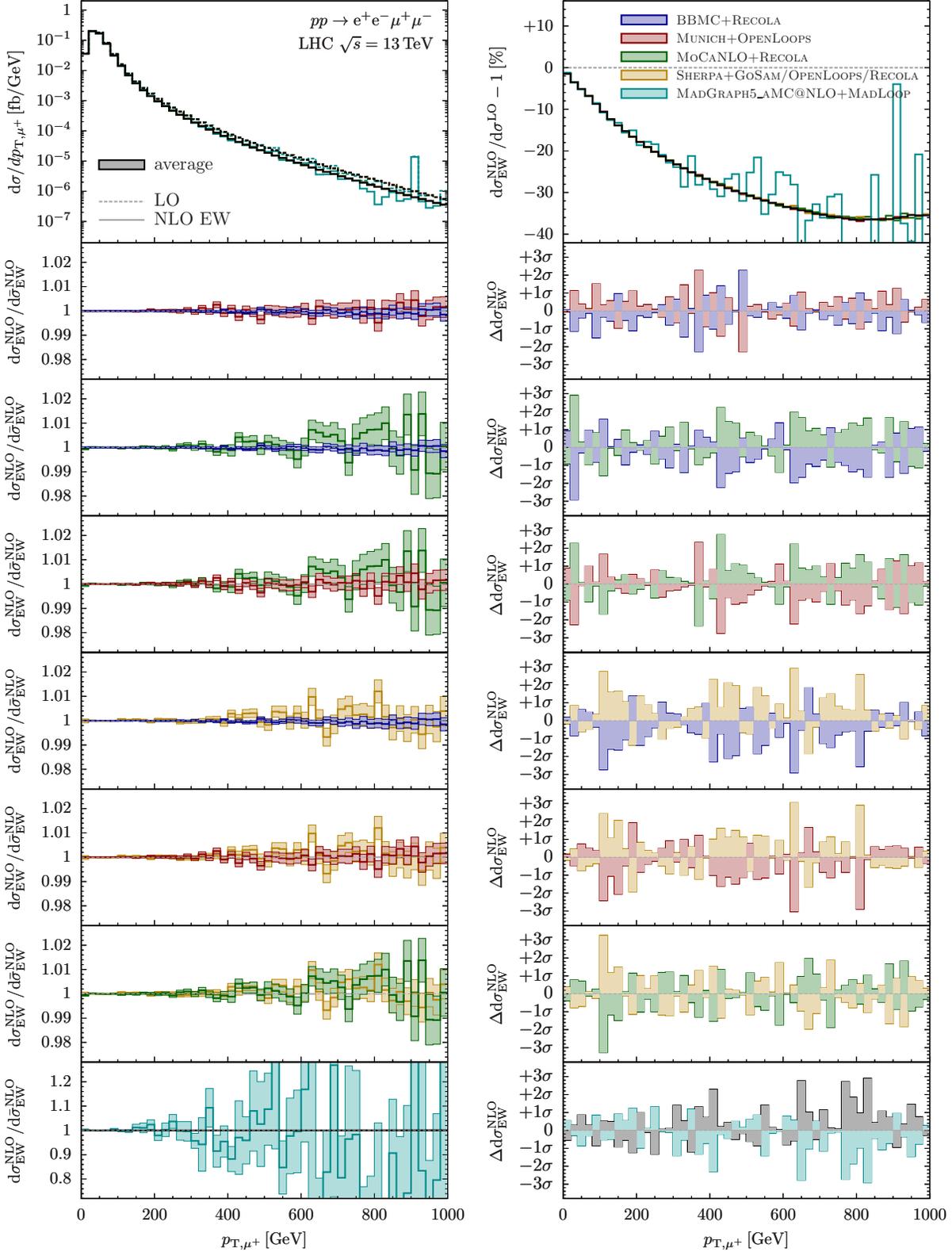


Fig. I.20: Technical comparison of NLO EW corrections to the distribution in the transverse momentum of the anti-muon,  $p_{T,\mu^+}$ , for hadronic  $e^+e^-\mu^+\mu^-$  (off-shell  $ZZ$ ) production. See main text of Sec. 7.3.4 for details.

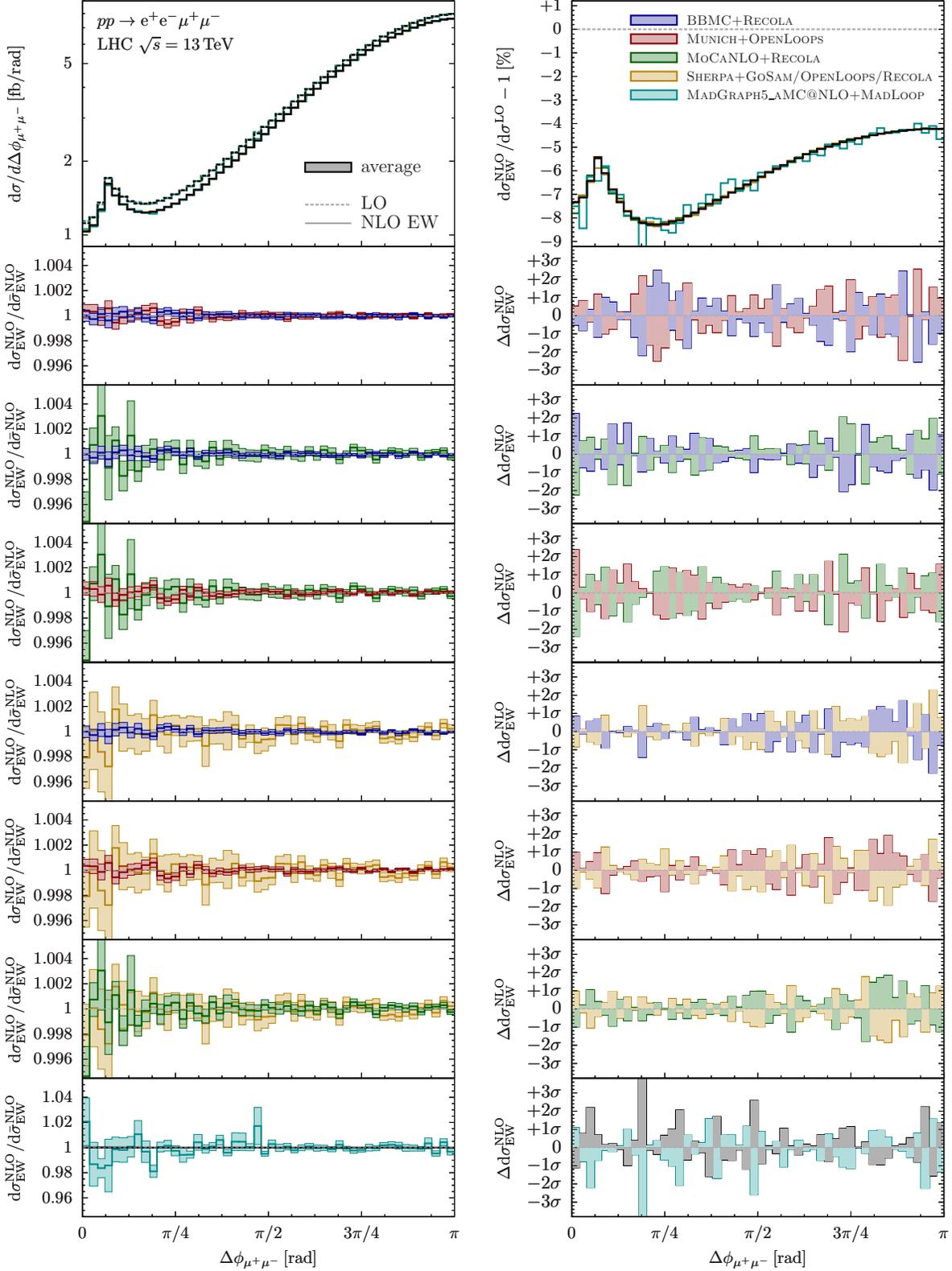


Fig. I.21: Technical comparison of NLO EW corrections to the distribution in the azimuthal angle between the anti-muon and the electron,  $\Delta\phi_{\mu^+\mu^-}$ , for hadronic  $e^+e^-\mu^+\mu^-$  (off-shell  $ZZ$ ) production. See main text of Sec. 7.3.4 for details.

corresponds to a permille level agreement in the full considered range for MCBB+RECOLA, MUNICH+OPENLOOPS, MoCANLO+RECOLA and SHERPA+GoSAM/OPENLOOPS/RECOLA.

As an example of a Sudakov-dominated observable, Fig. I.20 depicts the distribution in the transverse momentum of the produced anti-muon, with the typical large negative NLO EW corrections of  $-35\%$  at  $p_{T,\mu^+} \approx 1$  TeV. Statistically, the agreement is again good for all predictions in the full range. However, in the highly suppressed tail, a relative numerical precision better than a percent is only achieved by the most precise predictions.

The azimuthal-angle distribution of the  $\mu^+\mu^-$  pair in Fig. I.21 exhibits a preference for back-to-back configurations, while the collinear region dominated by  $\gamma^* \rightarrow \mu^+\mu^-$  splittings is suppressed by the applied lepton-lepton isolation cut. Overall, the distribution varies by less than one order of magnitude, and also the NLO EW corrections do not undergo particular enhancements and show variations of less than 5%. All individual predictions agree again very well on a statistical level, which corresponds to permille level agreement MUNICH+OPENLOOPS, MCBB+RECOLA, MoCANLO+RECOLA and SHERPA+GoSAM/OPENLOOPS/RECOLA, and at least percent level agreement for the MADGRAPH5\_AMC@NLO+MADLOOP prediction.

For off-shell WW production, we present a selection of distributions in Figs. I.22–I.25. The plots are organised analogously to those of ZZ production, but since with MUNICH+OPENLOOPS (red), MoCANLO+RECOLA (green), SHERPA+GoSAM/OPENLOOPS/RECOLA (yellow), and MADGRAPH5\_AMC@NLO+MADLOOP (cyan) there are only four predictions, pairwise comparisons are shown for all of them in frames 2–7. As for ZZ production, we find very good statistical agreement among the individual predictions for all considered observables.

We start with the pseudo-observable  $m_{2\ell 2\nu}$ , the invariant mass of the  $2\ell 2\nu$  system, in Fig. I.22. Above the WW threshold where the bulk of the cross section is, all individual predictions agree on the permille level, whereas the highly suppressed region below the threshold is by far less precise.

The distribution in the invariant mass of the  $\mu^+e^-$  system, shown in Fig. I.23, exhibits the typical Sudakov behaviour with negative NLO EW corrections of  $\approx -20\%$  at  $m_{\mu^+e^-} \approx 1$  TeV. The statistical agreement corresponds to a permille-level agreement for MUNICH+OPENLOOPS, MoCANLO+RECOLA and SHERPA+GoSAM/OPENLOOPS/RECOLA in the full range. under consideration.

The missing transverse-momentum distribution, depicted in Fig. I.24, falls particularly steeply since the double-resonant WW configuration is suppressed on the Born phase space for  $p_{T,\text{miss}} \gtrsim M_W$ . Correspondingly, the precision is restricted to the level of few percent in the tail of the distribution.

Finally, the distribution in the azimuthal-angle between the anti-muon  $\mu^+$  and the electron  $e^-$  exhibits no strong phase-space dependence, and the NLO EW corrections are quite flat. For MUNICH+OPENLOOPS, MoCANLO+RECOLA and SHERPA+GoSAM/OPENLOOPS/RECOLA, the good statistical agreement corresponds to a permille agreement, while deviations are well below one percent throughout for MADGRAPH5\_AMC@NLO+MADLOOP.

Beside the kinematic distributions discussed above, a number of observables was investigated in this comparison. The result is a reasonable statistical agreement for all distributions under consideration. Moreover, the observed deviations reflect the behaviour expected from a purely statistical distribution.

## 7.4 Conclusions

In this contribution we have performed a detailed comparison of calculations of NLO EW corrections between different automated fixed-order codes. As benchmark cases we considered off-shell WW and ZZ production, which offer a rich resonance structure. This allowed us to

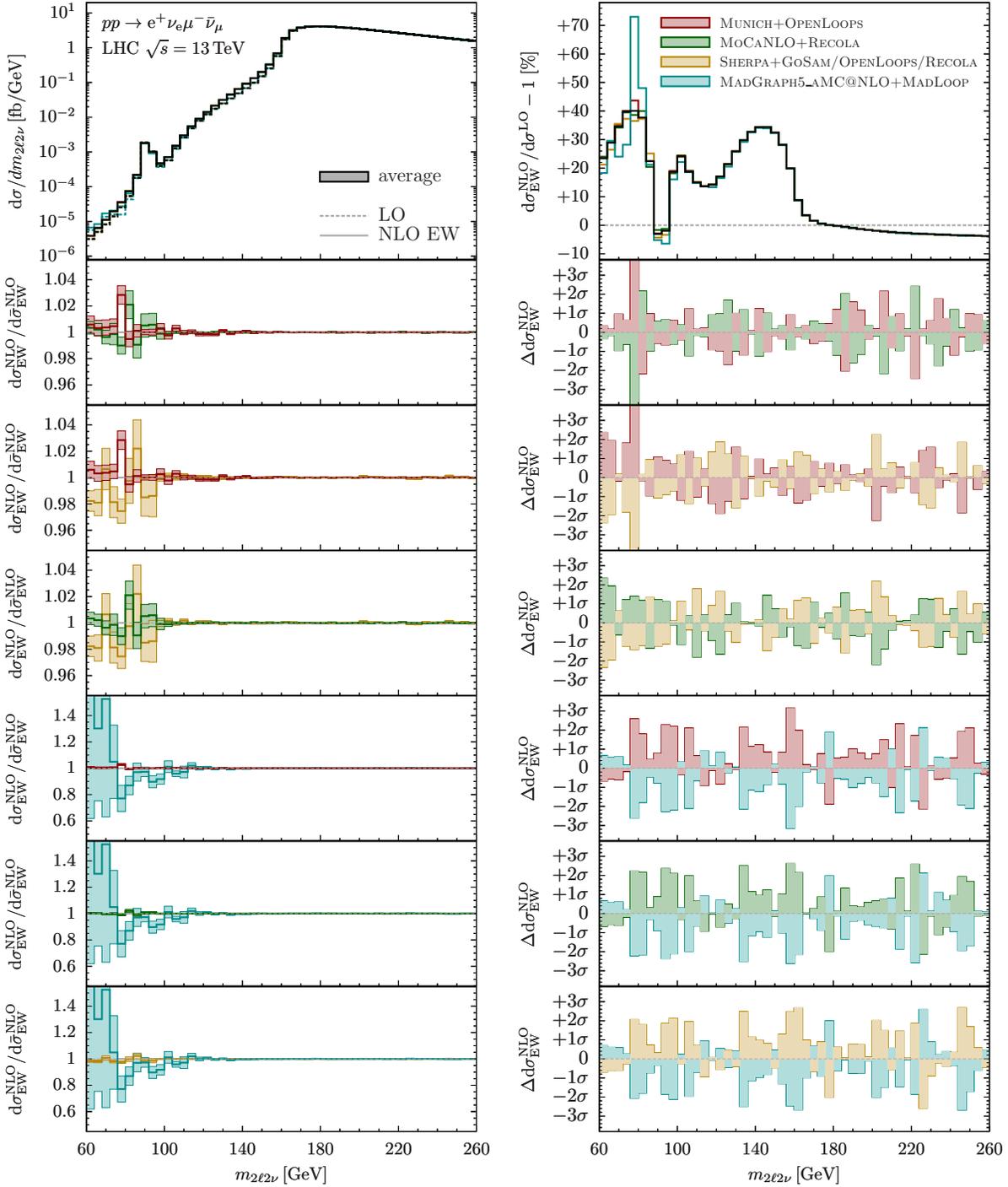


Fig. I.22: Technical comparison of NLO EW corrections to the distribution in the invariant mass of the  $2\ell 2\nu$  system (pseudo-observable) for hadronic  $e^+\nu_e\mu^-\bar{\nu}_\mu$  (off-shell  $WW$ ) production. See main text of Sec. 7.3.4 for details.

cross-validate many technical subtleties in the different implementations, in particular the required renormalization in the complex-mass scheme. First, we compared results obtained with the tools MADLOOP, RECOLA, OPENLOOPS, GoSAM and NLOX at the amplitude level and found very good agreement at NLO EW accuracy. Second, in order to compare results at the level of integrated cross sections, these amplitude providers have been interfaced in different

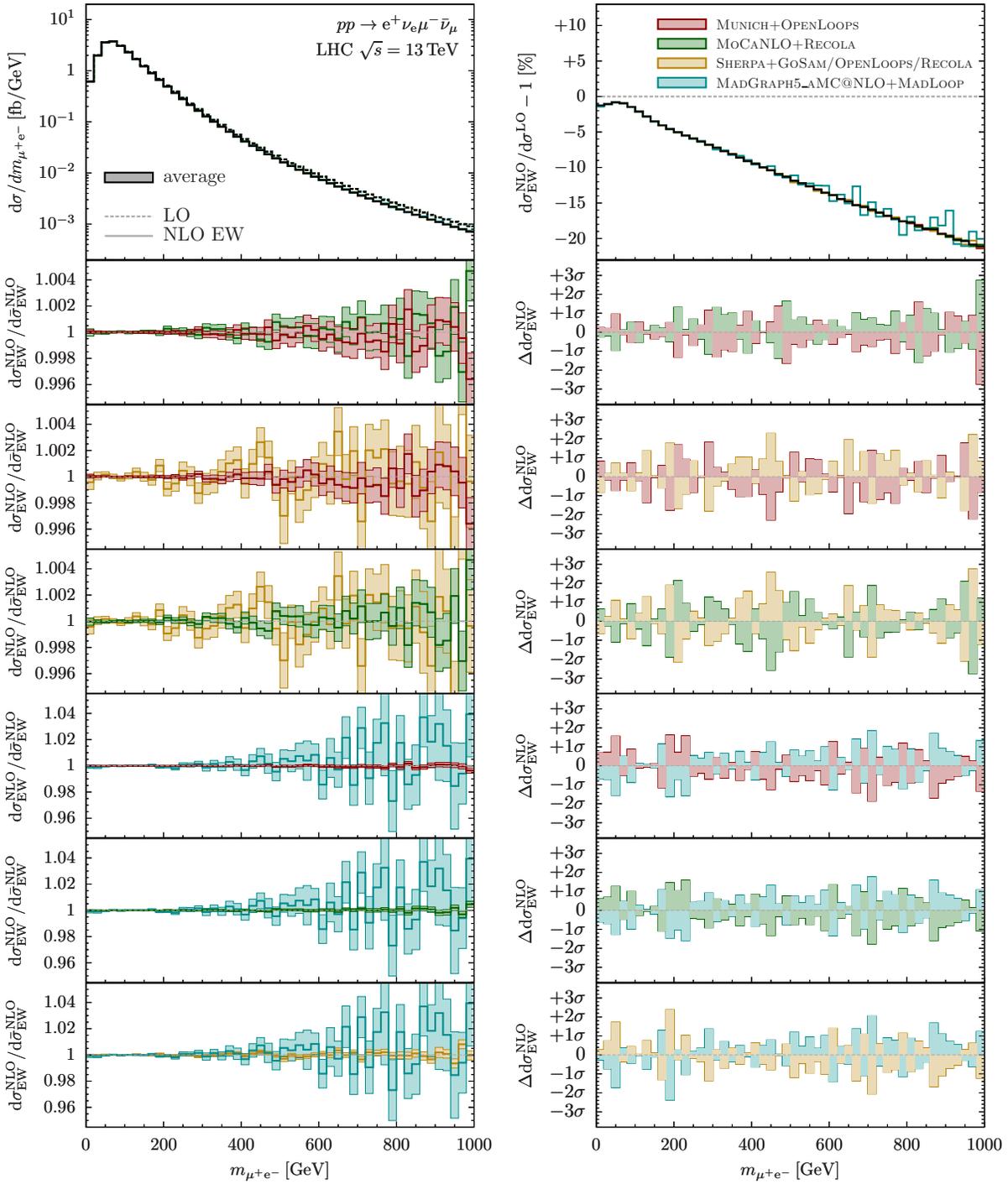


Fig. I.23: Technical comparison of NLO EW corrections to the distribution in the invariant mass of the  $\mu^+e^-$  system, for hadronic  $e^+\nu_e\mu^-\bar{\nu}_\mu$  (off-shell  $WW$ ) production. See main text of Sec. 7.3.4 for details.

combinations with the Monte Carlo frameworks BBMC, MoCANLO, MUNICH, SHERPA and MADGRAPH\_AMCNLO, which entail the subtraction of QED singularities. Again very good agreement has been found in various differential distributions including resonance peaks and high-energy tails. This study therefore strengthens our confidence in the correctness of existing and future predictions including NLO EW corrections from the tools considered here.

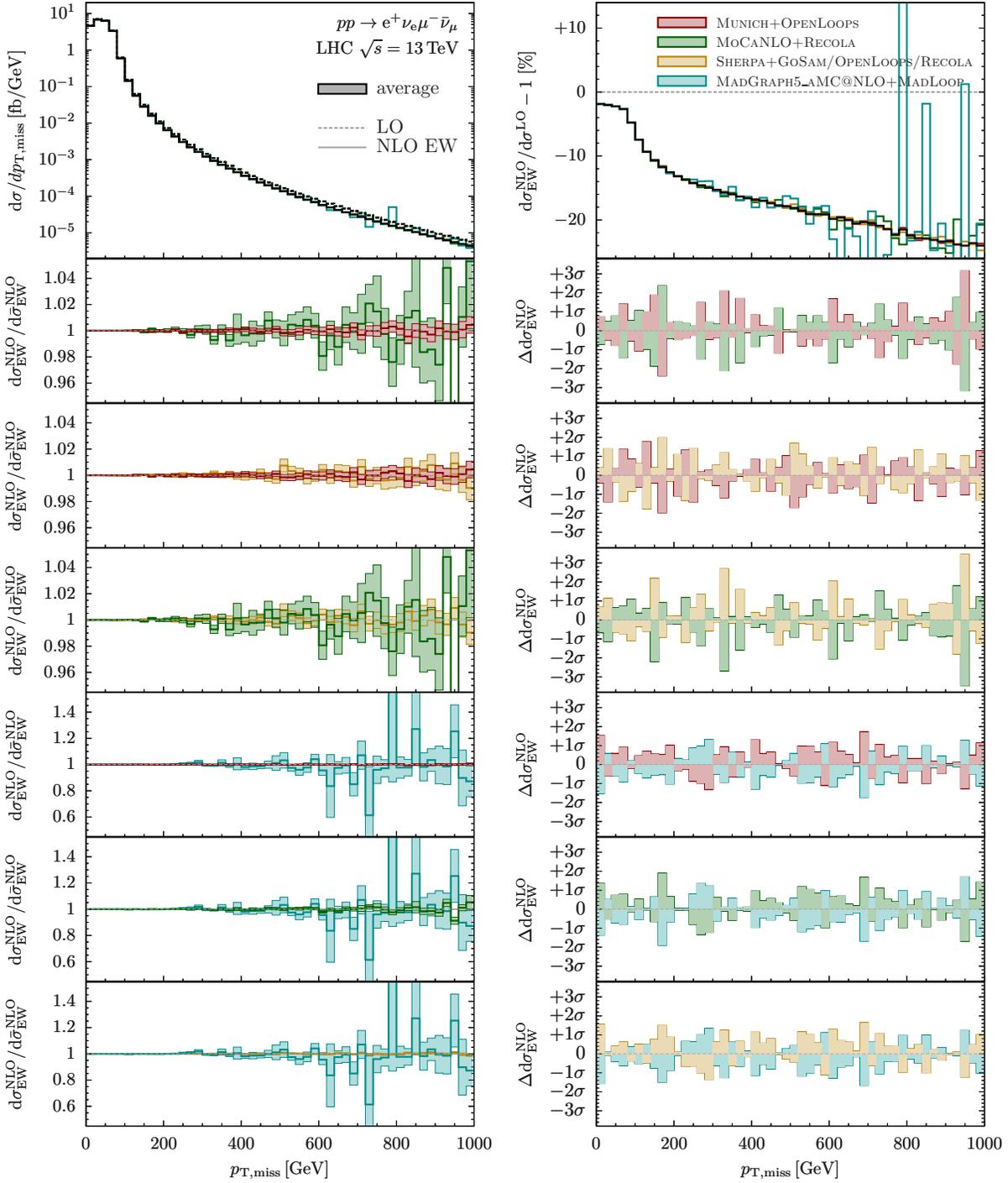


Fig. I.24: Technical comparison of NLO EW corrections to the distribution in the missing transverse momentum, for hadronic  $e^+\nu_e\mu^-\bar{\nu}_\mu$  (off-shell  $WW$ ) production. See main text of Sec. 7.3.4 for details.

## Acknowledgements

BB, AD, and MP acknowledge financial support by the German Federal Ministry for Education and Research (BMBF) under contract no. 05H15WWCA1 and the German Science Foundation (DFG) under reference number DE 623/6-1. The work of SB and SS has received funding from BMBF under contract 05H15MGCAA and from the European Union's Horizon 2020 re-

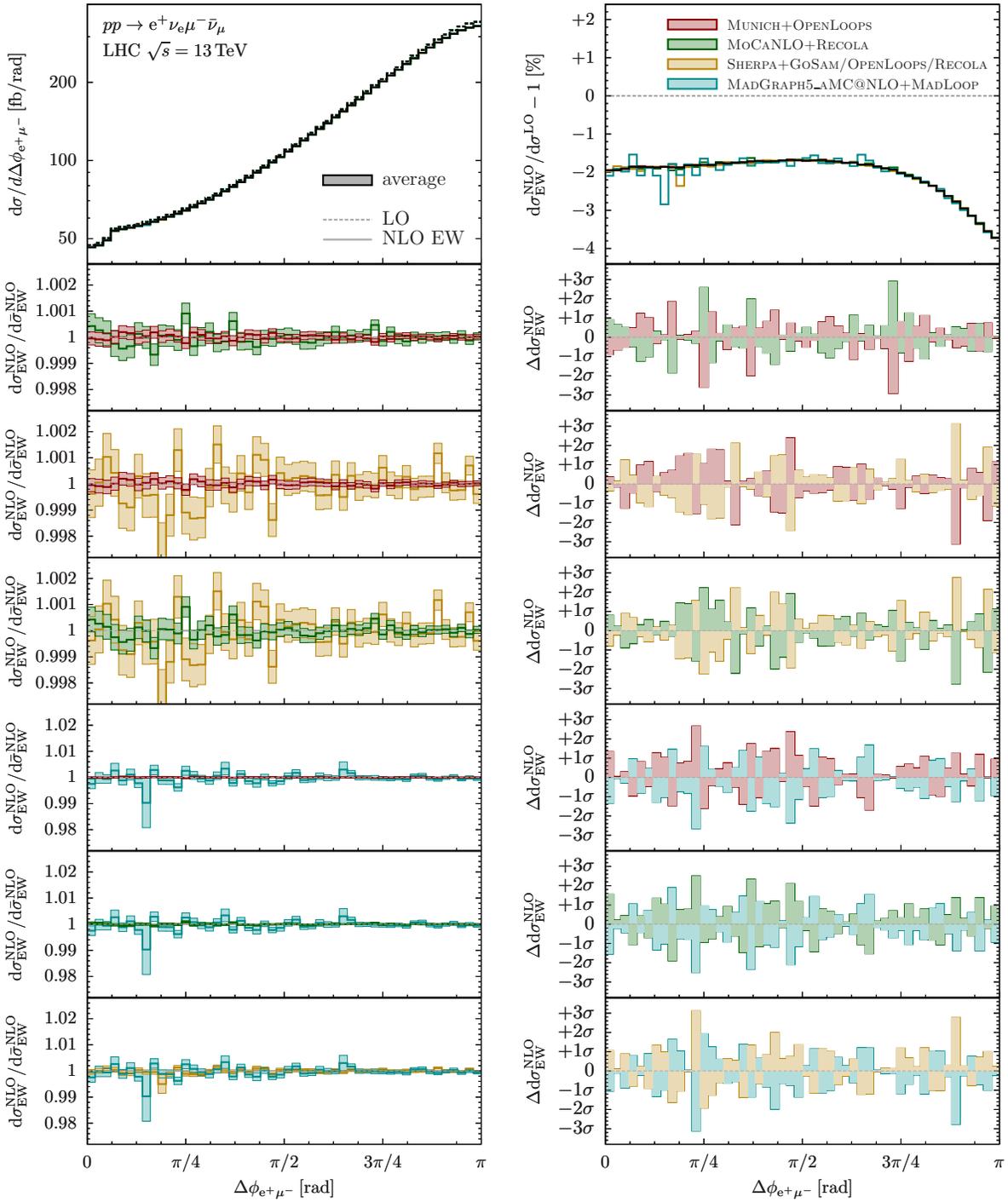


Fig. I.25: Technical comparison of NLO EW corrections to the distribution in the azimuthal angle between the anti-muon and the electron,  $\Delta\phi_{\mu+e-}$ , for hadronic  $e^+\nu_e\mu^-\bar{\nu}_\mu$  (off-shell  $WW$ ) production. See main text of Sec. 7.3.4 for details.

search and innovation programme as part of the Marie Skłodowska-Curie Innovative Training Network MCnetITN3 (grant agreement no. 722104). NG was supported by the Swiss National Science Foundation under contract PZ00P2\_154829. HSS is supported by the ILP Labex (ANR-11-IDEX-0004-02, ANR-10-LABX-63). The work of SQ and CR is supported by the U.S. Department of Energy under grant DE-SC0010102.

## Chapter II

### Parton distribution functions

#### 1 Theoretical uncertainties and dataset dependence of parton distributions <sup>1</sup>

We study theoretical uncertainties on parton distributions (PDFs) due to missing higher order (MHO) corrections by determining the change of PDFs when going from next-to-leading (NLO) to next-to-next-to-leading order (NNLO) theory. Based on the NNPDF3.1 framework, we compare PDF determinations obtained from different datasets, specifically a global, a proton-only, and a collider-only dataset. We show that PDF determinations obtained from a wider input dataset exhibit greater perturbative stability, and thus are likely to be affected by smaller theoretical uncertainties from MHOs. We also show that the effect of including deuterium nuclear corrections is smaller than that of excluding the deuterium data altogether.

##### 1.1 Parton distribution uncertainties

The accurate determination of the uncertainties on parton distribution functions (PDFs) of the proton [402] has been the main challenge in PDF studies at the LHC. For instance, PDFs represent one of the main sources of uncertainty in Higgs physics [236, 594]; they also have a significant impact on searches of physics beyond the standard model (see e.g. [595]), and on precision measurements such as the determination of the  $W$  boson mass [596]. Uncertainties on the current combined PDF4LHC15 PDF set [597] are typically of order 3-5% in the region covered by data, but the more recent NNPDF3.1 PDF set [598], which includes a wide array of LHC data, has PDF uncertainties typically between 1% and 3% in the data region (not including, in either case, the uncertainty on  $\alpha_s$ ).

This uncertainty — indeed, what is usually referred to as “PDF uncertainty” — includes the propagated uncertainty on the data used for PDF determinations, as well as further uncertainties due to the fitting methodology, but it does not include any theory error. Specifically, the current PDF uncertainty does not include a contribution accounting for the fact that PDFs are determined using fixed-order perturbative QCD, and thus surely one has to account for a missing higher order uncertainty (MHOU). As the uncertainty due to the data and methodology keeps decreasing, this theory error is bound to stop being negligible, and eventually become dominant.

So far, there has been a broad consensus that the use of the widest possible dataset for PDF determination — leading to so-called global PDF fits — is advantageous. Indeed, on the one hand, more data contain more information and thus allow for the accurate determination of the widest set of PDF in the most extended kinematic region. On the other hand, the use of multiple datasets provides a cross-check on both the theory and methodology. One may however ask whether the use of a wider dataset is also advantageous — or indeed not — in terms of theoretical uncertainties, specifically the MHOU. In particular, it is important to understand if MHOU are likely to be smaller with a global dataset or with a reduced dataset.

---

<sup>1</sup> S. Forte, Z. Kassabov, J. Rojo, L. Rottoli

## NNPDF3.1 global, NLO vs NNLO

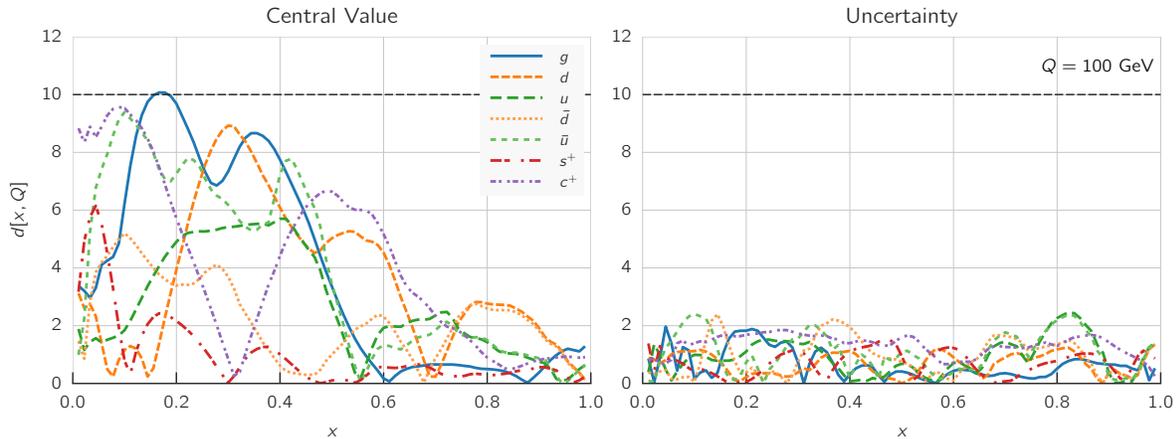


Fig. II.1: Distance between the central values (left) and the uncertainties (right) of NLO and NNLO global PDFs.

## 1.2 Missing higher order corrections and dataset dependence

### 1.2.1 Comparing NLO and NNLO PDFs

Quite in general, there is currently no way to reliably estimate the MHOUs. However, when several perturbative orders are known, we may at least study the behaviour of the perturbative expansion. Assuming reasonable convergence, the shift between, say, known NLO and NNLO results then provides a reasonable estimate of the MHOU on the NLO result.

We will thus address the problem of MHOU on PDFs by studying the way PDF change from NLO to NNLO. To this purpose, we have produced PDF determinations based essentially on the same dataset as in the NNPDF3.1 global analysis [598]. The only difference is that, while in NNPDF3.1 some jet data for which NNLO corrections were not yet available were treated approximately, here we only include both at NLO and NNLO jet data for which exact NNLO theory is available (see Sect. 4.4 of Ref. [598]), and using the exact NNLO corrections [129]. This ensures perturbative consistency.

In Fig. II.1 we show the distance between PDFs determined at NLO and at NNLO. Recall that the distance  $d$  is defined as the difference in units of standard deviation, normalized so that  $d = 10$  corresponds to an one- $\sigma$  shift (see Ref. [598] and references therein for a more detailed discussion). It is clear from Fig. II.1 that, whereas no distance between central values is greater than one- $\sigma$  in the data region, several distances (in particular for the gluon and light quarks) are of order one- $\sigma$ . But a one- $\sigma$  distance means that NLO-NNLO shift is the same size as the PDF uncertainty, so we conclude that, at NLO, MHOUs on central values are comparable to PDF uncertainties. Distances between uncertainties are instead of order one, *i.e.* comparable to a statistical fluctuation. This means that PDF uncertainties at NLO and NNLO do not differ by a statistically significant amount, consistent with the expectation that they reflect the uncertainty on the data and methodology, and thus, by and large, do not systematically depend on the perturbative order.

### 1.2.2 Dataset dependence

In order to study the issue that we set out in the introduction, namely, the dataset dependence of the MHOUs associated to a fit of parton distributions, we have repeated the previous PDF

	NLO	NNLO	NLO/NNLO	$\Delta$
global	1.279	1.253	1.02	1.16
proton	1.248	1.193	1.05	1.97
collider	1.181	1.114	1.06	2.07

Table II.1: Value of  $\chi^2$  per data point for the global, proton-only, and collider-only fits at NLO and NNLO. The NLO/NNLO ratio, and difference normalized to the standard deviation (see text) are also given. Note that the total number of data points  $N_{\text{dat}}$  in each of the three fits is different.

determination now based on two reduced datasets. First, we have produced a proton-only determination, in which we excluded all data with nuclear and deuterium targets (specifically fixed-target deep-inelastic and Drell-Yan production). Then, we have produced a collider-only determination, in which we have excluded all fixed-target DIS and DY data altogether.

These PDF determinations from smaller datasets were already discussed in Ref. [598] (as well as in previous NNPDF studies [422,599]) where it was argued that, even though these smaller datasets are in principle more consistent, the increased theoretical reliability does not make up for the great loss in accuracy: the PDF uncertainty increases monotonically when reducing the dataset, and there is no evidence of inconsistency between the data entering the global fit. With this motivation, the use of the more global dataset for the baseline PDF determination was advocated. These two PDF determinations have now been redone starting from the global dataset described in Sec. 1.2.1.

The values of the total  $\chi^2$  per data point for these NNPDF3.1-based PDF determinations are collected in Table II.1. In each case, we show the  $\chi^2$  per data point at NLO and NNLO, their ratio, and the difference  $\Delta = \frac{\chi_{\text{NLO}}^2 - \chi_{\text{NNLO}}^2}{\sqrt{2N_{\text{dat}}}}$  which is a measure of the improvement of the  $\chi^2$  in units of its standard deviation. Note that the results in Table II.1 only consider the dataset that was included in the fit in each case, and consequently the total number of data points  $N_{\text{dat}}$  in each of the three fits is different. Clearly, the PDF fits based on smaller datasets lead to a better  $\chi^2$ , due to the greater consistency of the dataset. Interestingly, however, the deterioration of the total  $\chi^2$  from NNLO to NLO, as measured both by the  $\chi^2$  ratios, and the difference in units of the standard deviation, is more severe for the fits based on a smaller datasets, and thus largest in the case of the collider-only fit. This provides a first indication that the use of a wider dataset may lead to greater perturbative stability.

To investigate this issue further, we have computed again the distance between the NLO and NNLO fits, as shown in Fig. II.1, but now for the proton-only and collider-only PDF sets. Results are shown in Fig. II.2. It is clear that while distances which were already sizable in the global fit (specifically for the gluon and down quark) are still big, now also the light quark PDFs (in particular also up and anti-up) display sizable NLO-to-NNLO shifts.

In order to achieve a fully quantitative comparison, in Fig. II.3 we display the shift of PDF central values between the NLO and NNLO fits, normalized to the NLO, for the gluon and light quarks, comparing the three PDF determinations. In order to facilitate visualization, the shifts are symmetrized about the  $x$  axis. It is clear that while for the gluon the shift is of similar size (and quite small) in the three PDF determinations, for the quarks there is a uniform hierarchy: the smallest dataset, *i.e.* the collider-only PDF set, nearly always displays the largest shifts, with very few localized exceptions.

A simple explanation of the greater perturbative stability of the more global fit seen in Table II.1 could be that hadron collider processes, which have larger perturbative corrections than deep-inelastic scattering, carry a greater weight in the fits to a reduced dataset: the global fit improves less because it is less consistent. But in this case, we would expect the global fit,

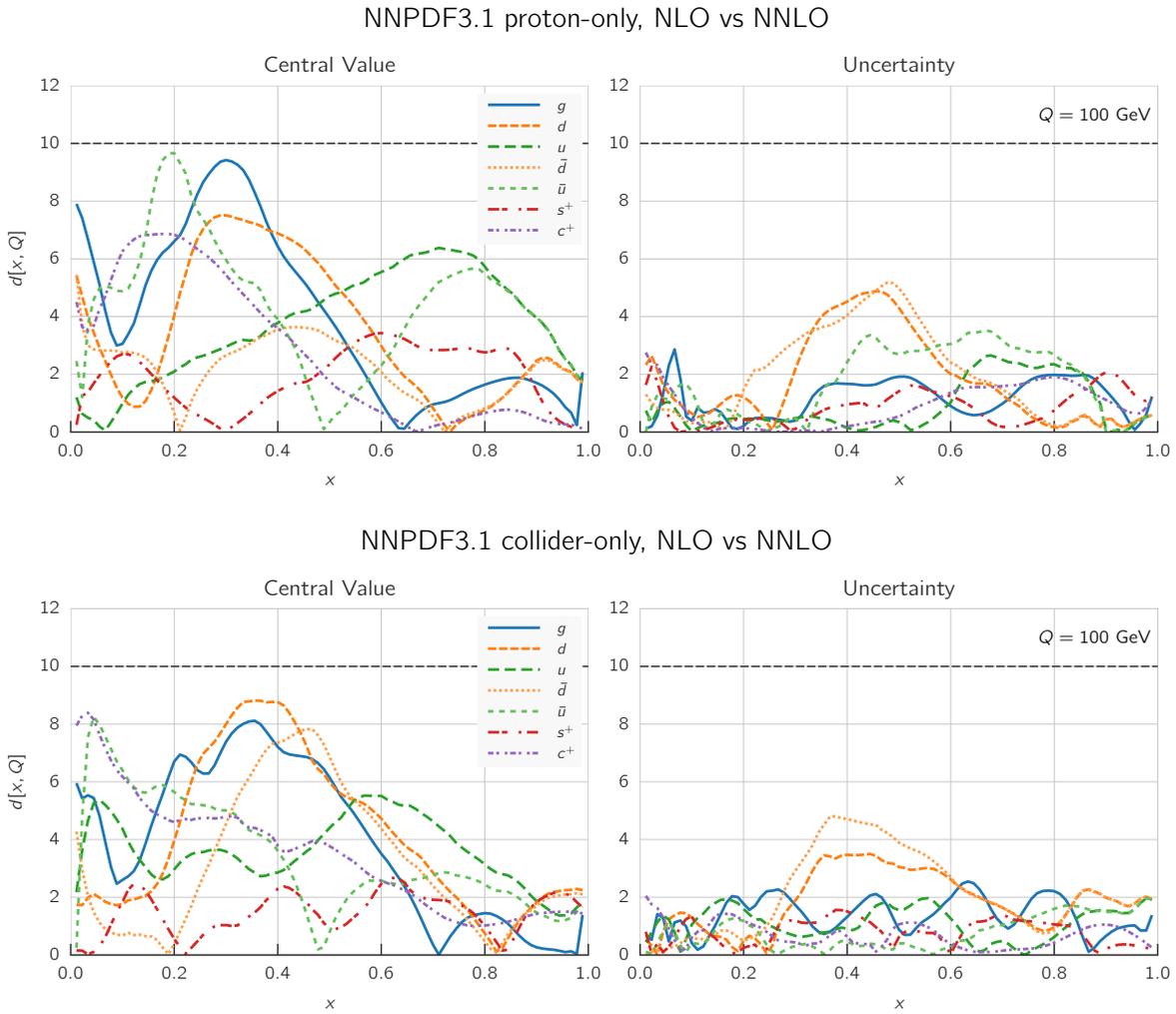


Fig. II.2: Same as Fig. II.1, but for the proton-only (top) and collider-only (bottom) PDF determinations.

due to its poorer consistency, not to show a significant improvement in PDF uncertainties, or to display a sizable change in results when going from NLO to NNLO. Instead, the opposite is the case. As extensively discussed in Ref. [598] the more global fit has significantly smaller PDF uncertainty, and as shown in Fig. II.3 it also changes less from NLO to NNLO. Hence the global fit is both less uncertain and more perturbatively stable.

An alternative explanation of the observed perturbative stability then seems more likely. Namely, that it is a consequence of the fact that missing higher order terms for different processes distort PDFs randomly by pulling them in different directions. Therefore, in a more global dataset in which the same PDF combination is determined from constraints by different processes, these uncertainties tend to average out.

### 1.3 Deuterium nuclear corrections

Perhaps the main advantage of the PDF fits based on smaller datasets is their greater consistency not only from the experimental, but also from the theoretical point of view. In particular, proton-only PDFs do not make use of any data that are affected by the poorly known nuclear corrections. It is then interesting to ask how the size of nuclear corrections compares to the uncertainties

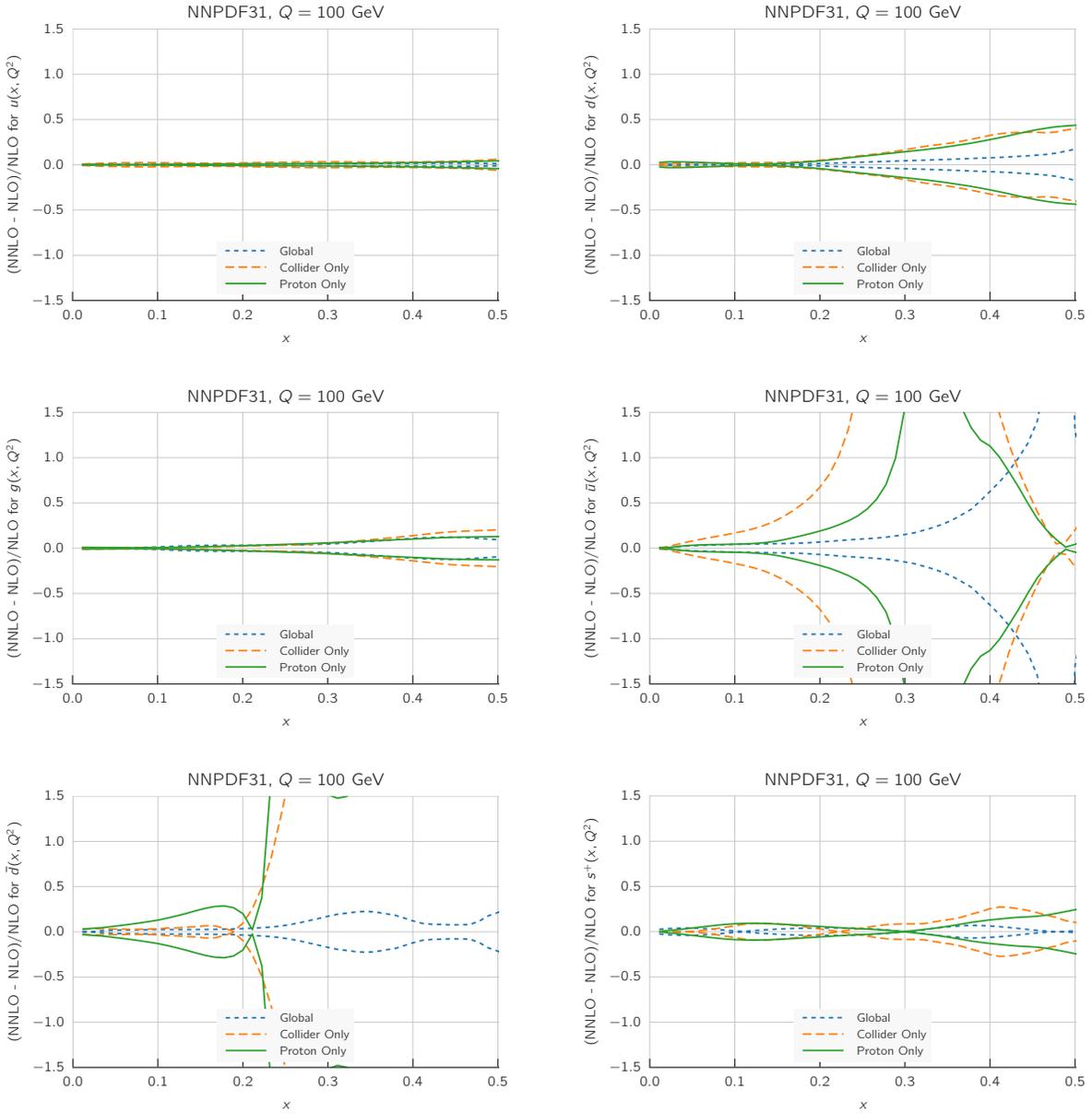
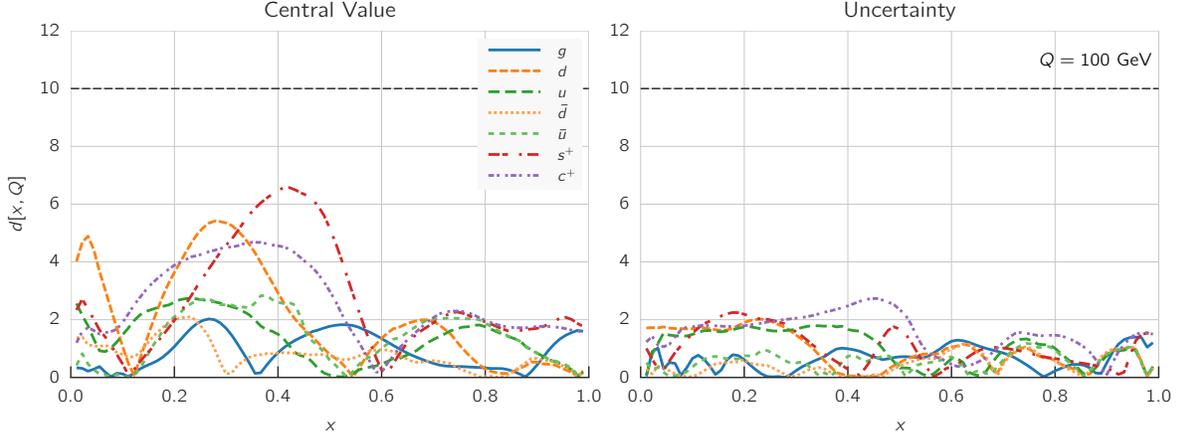


Fig. II.3: The relative shift between the central values of the NLO to NNLO PDFs normalized to the NLO, for the global, collider-only, and proton-only PDF determinations. Results are shown for, from left to right and from top to bottom, the gluon, up, down, anti-up, anti-down and total strange+antistrange PDFs. To facilitate visualization, the shifts are symmetrized about the  $x$  axis.

that we have discussed so far. Whereas existing determinations of heavy nuclear corrections are affected by large uncertainties [600], we may at least compare PDF determinations in which a model of nuclear effects for deuterium is included.

In order to isolate this effect, we have thus produced a PDF determination in which data using heavier nuclear targets have been excluded (hence specifically neutrino deep-inelastic scattering data) but data with deuterium targets (both deep-inelastic and fixed-target Drell-Yan) are kept. We then compare this fit either to the proton-only fit, or to itself with deuterium corrections included using the MMHT14 best-fit model of Ref. [601] (with default settings).

### NNPDF3.1 proton+deuteron, effect of deuteron corrections, NNLO



### NNPDF3.1 proton-only vs proton+deuteron, NNLO

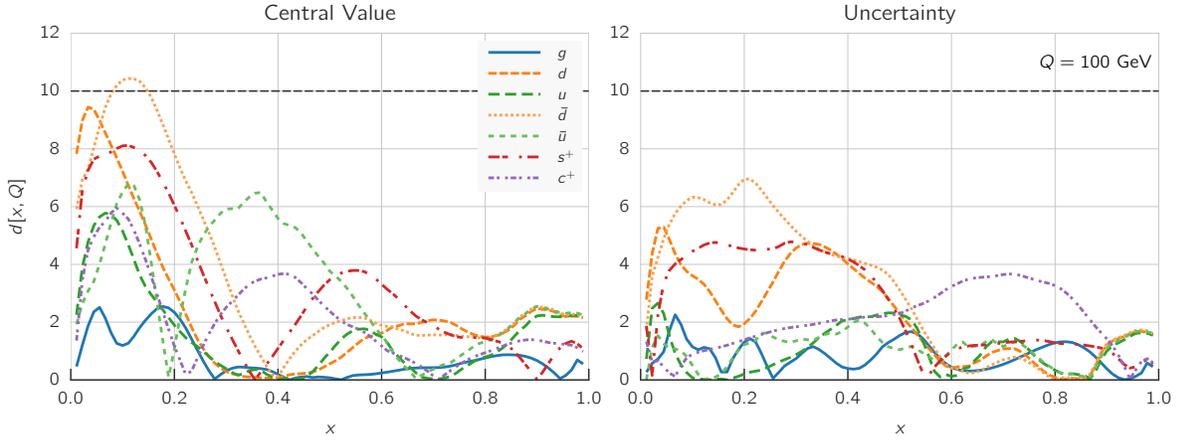


Fig. II.4: Distances between a fit with proton and deuterium target data and the proton-only fit (bottom) compared to the distances between a fit with proton and deuterium target data with and without linear corrections.

The corresponding distances between these fits are shown in Fig. II.4. It is clear that the effect of including the deuterium nuclear corrections is rather smaller than that of the deuterium data itself — whose impact is instead comparable to that of the NNLO corrections. We can thus conclude that the inclusion of deuterium data in the global PDF fit appears to be currently advantageous, in the sense that the impact of this data on the PDFs is greater than the likely size of their uncertainty due to missing nuclear corrections. On the other hand, if these data are included, some estimate of the associated theoretical uncertainty, arising both from deuterium nuclear corrections and from MHOs, should be performed, in view of the small size of the ensuing PDF uncertainty.

#### 1.4 Conclusions

We have demonstrated that PDF determinations based on a wider dataset are characterized by greater perturbative stability, and thus are most likely to exhibit smaller theoretical uncertainties related to missing higher orders. In addition, we have shown that the inclusion of data taken

with nuclear target, and specifically with deuterium ones, appears to be advantageous at present, in that the uncertainty related to the modeling of deuterium nuclear corrections appears to be rather smaller than both the MHOU and of the impact of this data on PDF uncertainties.

These results provide further evidence that the inclusion of theoretical uncertainties in PDF uncertainties, specifically those related to missing higher orders, but also to nuclear corrections, is now one of the highest priorities. An interesting observation in this respect is that a possible way to approach the determination of MHOU on PDFs might be to study the way they vary between different sets of experimental measurements. Indeed, our study suggests that these uncertainties tend to compensate when combining different classes of datasets. This, in turn, hints at the fact that the size of the MHOU might be estimated by performing dataset variations and studying the ensuing distribution of the best-fit PDFs. These topics are the subject of ongoing and forthcoming investigations.

### **Acknowledgements**

Stefano Forte thanks Daniel Maitre for interesting discussions on PDF uncertainties during the workshop. Stefano Forte is supported by the European Research Council under the European Union’s Horizon 2020 research and innovation Programme (grant agreement ERC-AdG-740006). Juan Rojo and Luca Rottoli are supported by an European Research Council Starting Grant “PDF4BSM”. The research of Juan Rojo is also partially supported by the Netherlands Organization for Scientific Research (NWO).

# Chapter III

## Jet substructure studies

### 1 Towards extracting the strong coupling constant from jet substructure at the LHC <sup>1</sup>

Recent advances in jet substructure have led to a new class of jet observables that are amenable to systematically-improvable calculations and are robust to the complex environment of the Large Hadron Collider (LHC). These observables exploit grooming to reduce non-perturbative contributions and simplify perturbative calculations. With these recent advances in both theory and experiment, we believe it is the appropriate time to begin investigating the possibility of extracting the strong coupling constant  $\alpha_s$  from jet substructure. In this section, we perform a proof-of-principle sensitivity study to demonstrate the suitability of such measurements to add useful information to the existing precision extractions of  $\alpha_s$ . We highlight a number of theoretical and experimental advantages of using groomed observables for  $\alpha_s$  extraction, and we discuss several difficulties of the LHC environment. Using a simplified approach, we show that a measurement of  $\alpha_s$  with an approximate 10% uncertainty should be feasible at the LHC. This result motivates a complete analysis with a full set of theoretical and experimental considerations, and we present a number of directions where improvements on both the theory and experiment sides could be made in the near future.

#### 1.1 Introduction

In quantum chromodynamics (QCD), the strong coupling constant ( $\alpha_s$ ) is responsible for the strength of interactions between quarks and gluons. Governed by this fundamental parameter, a plethora of strong-force phenomena emerge, including the binding of partons into massive hadrons that are responsible for most of the visible energy-density in the universe, and the creation of collimated sprays of hadrons known as jets that are ubiquitous at high-energy particle colliders. The internal structure of jets (jet substructure) has been extensively exploited to search for new particles at the Large Hadron Collider (LHC). Now that both the theoretical and experimental tools of jet substructure have reached a high degree of maturity [602–606], it is time to ask if the radiation patterns inside jets can be used to extract fundamental parameters of the Standard Model, such as  $\alpha_s$ .

There is a strong motivation (pun intended) for making a precise measurement of  $\alpha_s$ . The uncertainty in  $\alpha_s$  is a limiting factor in our predictions for the stability of the universe [607], and the uncertainty in  $\alpha_s$  at a variety of scales sets a model-independent sensitivity to strongly-interacting particles beyond the Standard Model [608, 609]. Now that many scattering processes have been calculated to a high perturbative order, the uncertainty on  $\alpha_s$  ( $\sigma_{\alpha_s}$ ) can be limiting in the overall accuracy; numerically  $\alpha_s^3 \sim \sigma_{\alpha_s}$ , so higher-order terms can be smaller than the leading-order correction uncertainty (see e.g.  $gg \rightarrow H$  at N<sup>3</sup>LO [610]). Determinations of  $\alpha_s$  probe a wide variety of physical phenomena, and the consistency between methods is both a crucial test of the theory and an important ingredient to make accurate predictions for the LHC and beyond.

The world-average value for  $\alpha_s$  at the  $Z$  boson mass ( $m_Z$ ) is  $0.118 \pm 0.0013$ , an impressive

---

<sup>1</sup> I. Moul, B. Nachman, G. Soyez, J. Thaler (section coordinators); S. Chatterjee, F. Dreyer, M. V. Garzelli, P. Gras, A. Larkoski, S. Marzani, A. Siódmok, A. Papaefstathiou, P. Richardson, T. Samui

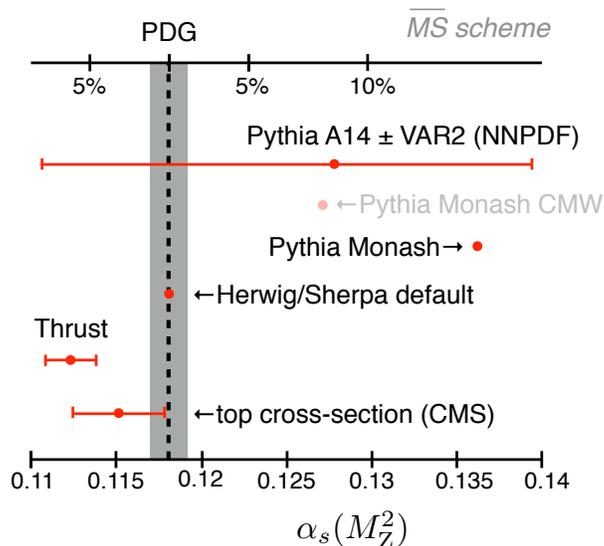


Fig. III.1: Various values of  $\alpha_s$  in the  $\overline{\text{MS}}$  scheme, including the world-average shown as a black dashed line with a grey uncertainty band [611]. The point labeled *Thrust* is from LEP data [618–627]. The measurement using the CMS top cross-section measurement is the first NNLO extraction at the LHC [628]. The other points are the  $\alpha_s$  parameter value used in the final state shower in various PS generators: HERWIG 7 [584] and SHERPA [415] use the world-average by default, while PYTHIA 8 [629, 630] employs a higher value in both the Monash default tune [631] and the A14 tunes (with VAR2 as an A14 uncertainty variation) [632]. The order at which  $\alpha_s$  is utilized is not the same in all cases.

1.1% total uncertainty [611].<sup>2</sup> There have been significant discussions in the community about the challenges and validity of various methods to extract  $\alpha_s$  (see for instance [611–617]). The most precise (and dominant) input to the world average is the  $\alpha_s$  value from lattice QCD calculations combined with measurements of  $B$ -hadron mass differences, with an uncertainty that is less than 1%. After the lattice, the most precise determination is from measurements and calculations of thrust and the  $C$ -parameter in  $e^+e^-$  collisions [618–627]. These methods are sensitive to very different regimes of QCD and, interestingly, differ from each other at about the 5% level, corresponding to a more than  $3\sigma$  tension.

Extractions based on  $e^+e^-$  event shapes are sensitive to soft and collinear regions of phase space, which are modeled with precision using higher-order resummation. Figure III.1 shows various values of  $\alpha_s$  extracted with next-to-next-to-leading order (NNLO) calculations including various levels of resummation, such as next-to-next-to-next-to-leading logarithmic ( $\text{N}^3\text{LL}$ ) for the thrust-based extraction. Also shown are the values of  $\alpha_s$  used as inputs to various Parton Shower (PS) generators.<sup>3</sup> The value of  $\alpha_s$  in the final-state PS is also sensitive to the soft and collinear regime of QCD (albeit in different ways); interestingly, the values used in the PYTHIA 8 program [629, 633], which are fit to  $e^+e^-$  event shape data, suggest a higher value of  $\alpha_s$  than the lattice result by about 15%. Another challenge with the event-shape extraction is that non-perturbative (NP) corrections are nearly degenerate with changes to  $\alpha_s$  [618]. This is in part because techniques to parametrically separate NP effects from perturbative effects (i.e. jet

<sup>2</sup>Unless otherwise specified, we always report  $\alpha_s$  at the  $Z$  boson mass in the  $\overline{\text{MS}}$  scheme.

<sup>3</sup>Note that these programs implement much more than just the QCD shower. For example, radiated photons are also generated, though the impact of those electroweak processes is quite negligible here (as opposed to photons from  $\pi^0 \rightarrow \gamma\gamma$ ).

grooming techniques to be discussed below) were not mature at the time of the Large Electron Positron (LEP) collider, and because the beam energies available at LEP did not allow for a large lever-arm between energy scales. The sensitivity to low-energy scales where QCD is in the NP regime is also a key challenge for the lattice determination. It would therefore be timely for a precision extraction of  $\alpha_s$  using jets at the LHC.

For most analyses at the LHC so far, jets have been used mainly as proxies for quark and gluon four-vectors. The multiplicity and kinematics of jets in purely hadronic final states can be predicted to high  $\alpha_s$  orders in perturbation theory with e.g. NNLOJET [129, 287] and NLOJet++ [417, 634]. As a result, measurements of jet multiplicities, energies, and angles can be used to extract  $\alpha_s$  [635–639]. Even though these measurements have achieved an uncertainty of around 5%, they have not yet been included in the Particle Data Group (PDG) combination [611] since they are using only next-to-leading-order (NLO) theory calculations. With recent progress in higher-order calculations, this will likely change soon, and it is noteworthy that an extraction using the recent NNLO  $t\bar{t}$  cross-section [26] is now included (see Fig. III.1).

In extractions based on fixed-order perturbation theory, the collinear region defining the internal structure of jets is avoided in part due to the lack of precision calculations.<sup>4</sup> It is important to remember that jets are not in one-to-one correspondence with individual quarks and gluons, but rather to clusters of hadrons resulting from many quarks and gluons; for this reason, the substructure of jets is also governed by  $\alpha_s$ . Recent theoretical and experimental advances in jet substructure have shown that the radiation pattern inside jets has great physics potential [602–606]. To date, the focus of jet substructures studies has been mostly on tagging the origin of jets and searching for physics beyond the Standard Model. However, present and future precision may be sufficient to make a useful measurement of  $\alpha_s$ .

A key tool to facilitate a substructure extraction of  $\alpha_s$  are jet grooming techniques, which systematically removes soft and wide-angle radiation from jets [640–645]. Jet grooming can parametrically separate the NP radiation in the jet from the hard perturbatively-described components. At a hadron collider, grooming also mitigates the contribution from the underlying event and additional nearly simultaneous interactions (pileup). Theoretical calculations for groomed observables have been performed at NNLL [646, 647] and NLL [648, 649] accuracy and will be extended to higher orders as calculations become available. Recently, ATLAS [650] and CMS [651] demonstrated 5-10% measurement uncertainties of these calculated quantities using existing technologies.

The purpose of this section is to study the feasibility of a measurement of  $\alpha_s$  using jet substructure at the LHC, using jet shapes that are infrared and collinear (IRC) safe. Similar techniques would of course also be interesting at both a low- and high-energy  $e^+e^-$  collider; we focus here on  $pp$  since high-quality data are now pouring out of the LHC. We view this work as part of a broader program to apply jet grooming techniques for precision QCD (see also Ref. [364] for applications to the top quark mass). There are significant experimental and theoretical challenges to achieve success in this program, but we believe it is possible with community synergy. An  $\alpha_s$  extraction represents a concrete goal to push the accuracy and understanding of jet substructure calculations in particular and QCD calculations more generally.

This remainder of this work is organized as follows. In Sec. 1.2, we define the observables and grooming strategies that will be considered in this study. In Sec. 1.3, we highlight a number of theoretical and experimental benefits of jet grooming, which make groomed observables particularly interesting for extractions of  $\alpha_s$ . In Sec. 1.4, we discuss the sensitivity of groomed jet mass to the value of  $\alpha_s$  using both analytic expressions and PS generators. In Sec. 1.5, an idealized setup is used to illustrate how an extraction of  $\alpha_s$  might work at the LHC, using real-

---

<sup>4</sup>Logarithms of the jet radius  $R$  may be important when the jet radius is small, and there are techniques to incorporate these corrections into fixed-order calculations [211, 212].

istic estimates of both the theoretical and experimental uncertainties. We conclude in Sec. 1.6 and discuss future directions for improving both the theoretical and experimental uncertainties for  $\alpha_s$  extractions from jet substructure.

## 1.2 Observable and Algorithm Definitions

In this subsection, we briefly review the definitions of the observables and grooming procedures that we will focus on in this study.

### 1.2.1 Two-point Correlators

A simple class of jet shape observables that have sensitivity to the value of  $\alpha_s$  are two-point correlation functions [652,653]. There has been significant theoretical study of these objects, and their perturbative behavior is understood to relatively high accuracy. For a set of constituents  $\{i\}$  in a jet  $J$ , the two-point correlation functions in  $pp$  collisions are defined as

$$e_2^{(\alpha)} \Big|_{pp} = \frac{1}{p_{TJ}^2} \sum_{i < j \in J} p_{Ti} p_{Tj} \left( \frac{R_{ij}}{R} \right)^\alpha, \quad (\text{III.1})$$

where  $p_{Ti}$  is the transverse momentum of particle  $i$  with respect to the beam axis,  $p_{TJ}$  is the transverse momentum of the jet,  $R_{ij}$  is the distance between particles  $i$  and  $j$  in the rapidity-azimuthal angle plane, and  $R$  is the jet radius. Note that some authors define the correlation functions without the  $R^\alpha$  factor in the denominator. The angular exponent  $\alpha$  is a parameter that controls the sensitivity to wide-angle emissions. If all emissions in the jet are nearly collinear, then the two-point energy correlation function reduces to a function of the jet mass ( $m$ ) when  $\alpha = 2$ :

$$e_2^{(2)} \sim \frac{m^2}{R^2 p_T^2}. \quad (\text{III.2})$$

In this study, we primarily restrict ourselves to the case of  $\alpha = 2$  as this has been the focus so far from both the theoretical [646–649] and experimental [650,651] communities. However, it is interesting to study if other values of  $\alpha$  could provide improved sensitivity to  $\alpha_s$ . Another benchmark value is  $\alpha = 1$ , which corresponds to  $k_T$  (or broadening) instead of mass. The two-point correlation functions are closely related to the jet angularities [654–657], with the latter being defined with respect to a jet axis.

### 1.2.2 Grooming Techniques

There are a variety of jet grooming techniques to mitigate soft and wide-angle jet contamination [640–645]. We focus here on the SoftDrop algorithm [645], defined using Cambridge/Aachen (C/A) reclustering [658–660]. Starting from a jet identified with an IRC-safe jet algorithm (such as anti- $k_t$  [661]), SoftDrop proceeds as follow:

1. Recluster the jet using the C/A clustering algorithm, producing an angular-ordered branching history for the jet.
2. Step through the branching history of the reclustered jet. At each step, check the SoftDrop condition

$$\frac{\min [p_{Ti}, p_{Tj}]}{p_{Ti} + p_{Tj}} > z_{\text{cut}} \left( \frac{R_{ij}}{R} \right)^\beta, \quad (\text{III.3})$$

where  $z_{\text{cut}}$  is a parameter defining the scale below which soft radiation is removed, and  $\beta$  is an exponent that controls the angular scale for removal. If the SoftDrop condition is not satisfied, then the softer of the two branches is removed from the jet. This process is then iterated on the harder branch.

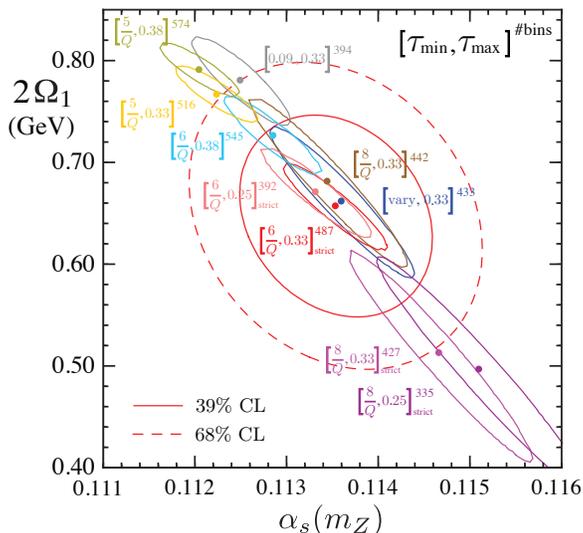


Fig. III.2: An illustration of the correlations between the first NP moment  $\Omega_1$  and  $\alpha_s$  for the thrust observable. NP effects for groomed observables have a significantly different structure, providing the possibility for complementary information to the extraction from (ungroomed) event shapes in  $e^+e^-$ . Figure from Ref. [618].

3. The SoftDrop procedure terminates once the SoftDrop condition is satisfied.

This procedure generalizes the modified Mass Drop Tagger (mMDT) [644], which corresponds to  $\beta = 0$  in the SoftDrop procedure described above. Note that non-global logarithms [662] are formally removed from the SoftDropped mass distribution, even when  $\beta > 0$  (though clearly they are restored as  $\beta \rightarrow \infty$ ).

### 1.3 Grooming Away Soft Complications

Having defined the observables and grooming procedures of interest, in this subsection we provide a general discussion of the potential advantages of using groomed jet observables for extractions of  $\alpha_s$ . The grooming procedure significantly simplifies a number of theoretical and experimental issues related to precision calculations and measurements in a  $pp$  environment. In particular, grooming provides:

- Mitigation of NP effects;
- Perturbative simplicity; and
- Improved detector resolution, due in large part to a reduced pileup sensitivity.

Each of these will be discussed in more detail below. Combined, we believe that they provide a strong motivation for the measurement of groomed jet observables, and the possibility of performing a precision extraction of  $\alpha_s$  using jet substructure at the LHC.

#### 1.3.1 Mitigating Non-perturbative Effects

One of the primary complications of  $\alpha_s$  extractions using event shapes is NP corrections, which cannot currently be calculated from first principles. Since event shape observables probe singular regions of the phase space, they are necessarily sensitive to NP corrections. For  $e^+e^-$  dijet

event shapes, such as thrust, these NP corrections can be given operator definitions in the dijet limit. Although these operators cannot be calculated, they can be modeled using *shape functions* [663–666].<sup>5</sup> These shape functions can be systematically expanded in moments, and for moderately large values of the observable, only the first moment ( $\Omega_1$ ) is required. This moment can be shown to be universal for a wide range of observables [670, 671], and can therefore be extracted from data along with  $\alpha_s$  [618, 619, 672].<sup>6</sup> This moment, however, is highly correlated with  $\alpha_s$  as shown in Fig. III.2. Therefore, observables with different, and preferably suppressed, NP effects could provide valuable complementary information for  $\alpha_s$  extractions.

Using a scaling argument, it is a straightforward exercise to compute the value of the groomed two-point correlators at which NP effects are expected to be important. This was considered in Ref. [644, 647] where it was shown that NP effects become important when (for  $\alpha \geq 1$ ),

$$e_2^{(\alpha)} \Big|_{\text{NP}} \simeq \left( \frac{\Lambda_{\text{QCD}}}{z_{\text{cut}} Q} \right)^{\frac{\alpha-1}{1+\beta}} \frac{\Lambda_{\text{QCD}}}{Q}, \quad (\text{III.4})$$

where  $Q = p_T R$  is the starting scale of jet fragmentation and  $\Lambda_{\text{QCD}} \simeq 1$  GeV is a typical NP scale. When  $\alpha < 1$ , the value is instead  $(\Lambda_{\text{QCD}}/(p_T R))^\alpha$ . If we consider for concreteness the jet mass ( $\alpha = 2$ ) we can learn a number of interesting lessons. First, by taking  $\beta \rightarrow \infty$ , we obtain the ungroomed result

$$e_2^{(2)} \Big|_{\text{NP}} \simeq \frac{\Lambda_{\text{QCD}}}{p_T R}, \quad (\text{III.5})$$

while the groomed result gives

$$e_2^{(2)} \Big|_{\text{NP}} \simeq \left( \frac{\Lambda_{\text{QCD}}}{z_{\text{cut}} p_T R} \right)^{\frac{1}{1+\beta}} \frac{\Lambda_{\text{QCD}}}{p_T R}. \quad (\text{III.6})$$

For  $z_{\text{cut}} p_T R \gg \Lambda_{\text{QCD}}$ , we see that the grooming significantly suppresses the scale of the NP physics, extending the range of perturbative validity by a factor of  $(\Lambda_{\text{QCD}}/z_{\text{cut}} p_T R)^{1/(1+\beta)}$ . For a 1 TeV jet with  $z_{\text{cut}} = 0.1$  and  $\beta = 0$ , this is an extension by a factor of  $\sim 100$ . Furthermore, we observe that, under the assumption that  $\beta \geq 0$ , this suppression is maximized for  $\beta = 0$ , motivating this choice in later studies.<sup>7</sup>

It is important to emphasize that just because the value of the mass at which NP physics enters is suppressed, this does not by itself imply that there is a larger region over which one has perturbative control. In principle the whole distribution could be shifted to lower values. We know, however, that for  $e_2^{(2)} \geq z_{\text{cut}}$ , the distribution is parametrically unaffected by the grooming procedure, and is therefore identical to the ungroomed jet mass.

In addition to being suppressed, the structure of the NP corrections is quite different for groomed jet mass, which may provide complementary information compared to standard event shapes. For event shape observables, NP corrections are typically large already at the Sudakov peak. In the case of a groomed observable, since the NP corrections are suppressed and the observable is predominantly single logarithmic (exactly true when  $\beta = 0$ ), we instead see that the NP corrections appear as a bump on the falling distribution. This is shown in Fig. III.3 along with a model using a shape function. Due to the fact that this is a completely different

<sup>5</sup>There are other approaches to analytically describe and fit NP effects that have similar features to shape functions. Another example is the dispersive model [667, 668] which has also been used to make extractions of  $\alpha_s$  using  $e^+e^-$  event shapes [669].

<sup>6</sup>See Ref. [673, 674] for subtleties related to hadron masses and the breakdown of universality.

<sup>7</sup>Negative values of  $\beta$  are also possible and could be interesting to study. However, one would need to dedicate specific attention in the vicinity of the endpoint of the distribution, which is  $e_2^{(\alpha)} = z_{\text{cut}}^{-\alpha\beta}$  at LL.

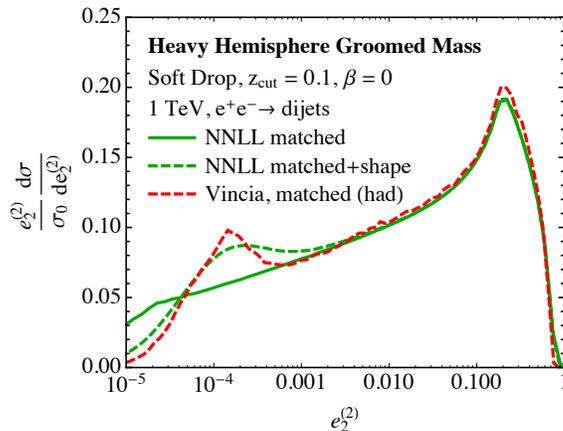


Fig. III.3: A plot showing the effects of NP corrections to the groomed jet mass. This illustrates first that the NP corrections are suppressed over a large range, and second that they take a very different form than for double-logarithmic observables. Figure from Ref. [647].

behavior, it may have different correlations as compared with standard event shapes, though a more complete understanding of this effect is required.

Finally, while we have so far focused on the effect of the grooming procedure on hadronization corrections and a comparison to  $e^+e^-$  event shapes, we must also emphasize that the grooming plays an important role in suppressing NP corrections from the underlying event. This is crucial for achieving a precision measurement in the  $pp$  environment, since there is a limited theoretical understanding of the underlying event. Here, we must rely purely on Monte Carlo model implementations of the underlying event. Monte Carlo studies have shown that in the regime where one has perturbative control, underlying event is highly suppressed after mMDT/SoftDrop [644, 645]. This is again a significant benefit of groomed observables.

### 1.3.2 Perturbative Simplicity

By studying the analytic behavior of grooming on the jet mass, the authors of Ref. [644] showed that grooming could be designed to enhance the simplicity and calculability of various jet shapes. In particular, grooming significantly simplifies the perturbative resummation in the  $pp$  environment in addition to suppressing NP corrections. Calculations of the standard jet mass in  $pp$  typically suffer from the following two perturbative difficulties:

- Global color correlations, which complicate the structure of higher-order soft functions; and
- Non-global logarithms [662], which complicate the perturbative structure for non-global measurements.

The jet mass was calculated in [675] ([676]) for  $H+$  jet with(out) a veto on additional radiation, supplemented with arguments that non-global effects are small. These above two features highlight that the ungroomed jet mass depends significantly on the particular process that created the jet, significantly reducing its universality, and complicating its structure in a busy  $pp$  environment. By contrast, the use of grooming can make the jet mass nearly universal.

In Ref. [647], it was shown that for  $e_2^{(\alpha)} \ll z_{\text{cut}} \ll 1$ , one can express the cross-section for

the groomed two-point correlator as

$$\frac{d\sigma^{pp}}{de_2^{(\alpha)}} = \sum_{k=q,\bar{q},g} D_k(p_T^{\min}, y_{\max}, z_{\text{cut}}, R) S_{C,k}(z_{\text{cut}}, e_2^{(\alpha)}) \otimes J_k(e_2^{(\alpha)}), \quad (\text{III.7})$$

where  $S_{C,k}$  is the soft function and  $J_k$  is the jet function. The function  $D_k$  depends on the parton flavor, and can be interpreted as the quark and gluon fractions. While the quark and gluon fractions are generally not IRC-safe quantities, in this limit the quark fraction can be defined jet-by-jet as

$$f_J = \sum_{i \in J_{\text{SD}}} f_i, \quad (\text{III.8})$$

with  $f_q = 1$ ,  $f_{\bar{q}} = -1$ ,  $f_g = 0$ , and  $J_{\text{SD}}$  are the parton-level constituents of the groomed jet. The grooming procedure makes this definition IRC safe at leading power (see related jet flavor approach in Ref. [677]). These fractions are independent of the value of the jet mass observable and can be extracted from fixed-order generators, for example MCFM [418, 420, 678].

The formula in Eq. (III.7) has a number of remarkable consequences. In the resummation region, we see that there are no non-global logarithms and there are no global color correlations. The collinear soft and jet functions depend only on the jet flavor, so they are therefore identical to in the case of groomed jets in  $e^+e^-$  collisions. In the resummation region, the grooming algorithm has therefore allowed the jet to be isolated from its  $pp$  environment, and the only required input from the hard scattering process are the quark and gluon fractions.

It is important to emphasize that this simplification does not occur throughout the entire distribution. In particular, it does not hold for  $m \gg Qz_{\text{cut}}$ , though this is the fixed order region, where we can match to standard fixed order perturbation theory. This matching has been performed to (N)NLO for dijets ( $V$ +jets) [647–649]. Ideally this would be done using  $2 \rightarrow 3$  matrix elements at NNLO, which are just now becoming available [123, 124, 289, 291, 679].

### 1.3.3 Pileup Resilience

Arguably the biggest experimental challenge for precision physics at a high-luminosity  $pp$  collider is noise from *pileup*: multiple nearly simultaneous proton-proton collisions. Jet shapes are particularly sensitive to pileup; for example, the jet mass scales as  $\mathcal{O}(A^2)$  [680] for the jet catchment area  $A$  [681] (whereas the jet  $p_T$  scales linearly with  $A$ ). The jet-area subtraction that works well for  $p_T$  has been extended to event shapes [682], but must be re-calibrated per observable. Constituent-based pileup subtraction schemes [683–687] show great promise and are actively being studied and adapted to the actual experimental settings [688–692]. Even without constituent-based subtraction techniques, though, there is a large reduction in pileup sensitivity to jet substructure from grooming [604, 688, 692, 693]. Grooming systematically removes soft and wide-angle radiation, which is exactly the profile characteristic of pileup. Even with extreme levels of pileup (up to 300 collisions), grooming can preserve the distribution of the jet mass distribution [694].

Despite the power of grooming for pileup suppression, there is still a residual degradation of resolution with increased levels of pileup which makes precision jet substructure measurements challenging at high instantaneous luminosity. Track-based observables are robust to pileup because their vertex of origin can be well-distinguished from pileup vertices. Precision track-based substructure observables have been calculated [695–698], but typically require universal NP input. It may be interesting to do a track- and jet-substructure-based extraction of  $\alpha_s$ , but this is left as a possibility for future work.

## 1.4 Observable Sensitivity to $\alpha_s$

In this subsection, we study the sensitivity of the groomed jet mass to variations in the value of  $\alpha_s$ . We begin with a discussion based on the analytic formulae at LL accuracy. We then perform a PS study, highlighting the interplay between the sensitivity of different parts of the distribution to variations in the value of  $\alpha_s$  and NP effects. Finally, we discuss the issue of Casimir scaling and the related issue of using normalized versus unnormalized distributions.

### 1.4.1 Analytic Understanding

To get an understanding of the sensitivity of the groomed mass distribution both to the value of  $\alpha_s$  as well as to the quark and gluon composition, it is enlightening to study the LL distribution. Here, for simplicity, we consider only the leading logs in the observable, in the resummation region; complete expressions can be found in Refs. [645, 647–649]. For  $\beta = 0$ , the LL result at fixed coupling for the cumulative distribution in the resummation region takes the schematic form

$$\Sigma(e_2^{(2)}) = \exp \left[ -\frac{\alpha_s C_i}{\pi} [\log(z_{\text{cut}}) - B_i] \log(e_2^{(2)}) \right], \quad (\text{III.9})$$

where  $B_i = -3/4$  for quarks and  $B_g = -\frac{11}{12} + \frac{n_f}{6C_A}$  for gluons ( $n_f$  is the number of active quark flavors). This highlights that for  $\beta = 0$ , the groomed jet mass is a single-logarithmic observable, contrasting with the standard double-logarithmic behavior of plain jet mass. Differentiating the cumulative distribution, we obtain the spectrum

$$\frac{e_2^{(2)}}{\sigma} \frac{d\sigma}{de_2^{(2)}} = -\frac{\alpha_s C_i}{\pi} [\log(z_{\text{cut}}) - B_i] \exp \left[ -\frac{\alpha_s C_i}{\pi} [\log(z_{\text{cut}}) - B_i] \log(e_2^{(2)}) \right]. \quad (\text{III.10})$$

Here, we immediately see several interesting consequences. In the resummation region, the slope of the distribution when plotted against  $\log e_2^{(2)}$  is set by the product  $\alpha_s C_i$ , where  $C_i$  is the Casimir factor, namely  $C_F = 4/3$  for quarks and  $C_A = 3$  for gluons. We therefore see that the groomed mass is indeed sensitive to the value of  $\alpha_s$ . Due to the larger color charge of gluons, we expect that samples of pure gluon jets would have a significantly higher sensitivity to the value of  $\alpha_s$ ; this expectation will be born out in our PS studies below. Because  $\alpha_s$  is always multiplied by a color factor, though, knowing the precise quark/gluon composition of a sample is essential, as discussed in Sec. 1.4.3. In practice, the PS studies and the analytic studies that follow (see Sec. 1.5) include higher-order effects, such as subleading terms in the splitting functions, that violate Casimir scaling.

### 1.4.2 Parton Shower Study

From the point of view of fitting for  $\alpha_s$ , a good observable is one whose probability distribution changes significantly with variations in  $\alpha_s$ . However, many observables that significantly change with  $\alpha_s$  are also very sensitive to NP effects, such as the constituent multiplicity inside jets. Here, we study many two-point correlators to quantify the tradeoff between the sensitivity to  $\alpha_s$  and the robustness to NP effects.

Given two probability distributions  $f$  and  $g$  in some observable  $\lambda$ , we define the separation power  $\Delta(f, g)$  [699] as

$$\Delta(f, g) = \frac{1}{2} \int d\lambda \frac{(f(\lambda) - g(\lambda))^2}{f(\lambda) + g(\lambda)}. \quad (\text{III.11})$$

The separation power is a number in  $[0, 1]$ , where  $\Delta = 0$  if and only if  $f = g$ . If  $f$  is the nominal probability distribution of some observable and  $g$  is the distribution of the same observable with a different value of  $\alpha_s$ , we would like  $\Delta(f, g)$  close to 1 (sensitivity). In contrast, if  $g$  is

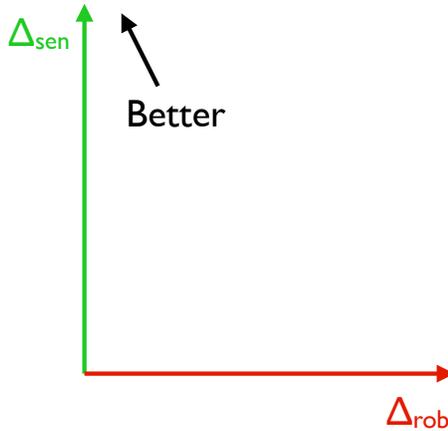


Fig. III.4: A schematic diagram to illustrate the sensitivity-robustness plane defined by the separation power in Eq. (III.11).

the same as  $f$  with some variation in the NP effects, then we would like  $\Delta(f, g)$  to be close to 0 (robustness). The plane used to study the tradeoff between sensitivity and robustness is shown in Fig. III.4. Not all information about  $\alpha_s$  sensitivity is captured by a single point in Fig. III.4, because sensitivity to NP effects could be in regions of low  $\alpha_s$  sensitivity and vice versa. Therefore, it is also useful to study the integrand of Eq. (III.11) as a function of different observables  $\lambda$ .

This sensitivity-robustness tradeoff is studied using two PS generators: HERWIG 7.1.0 [584, 700, 701] with its default settings<sup>8</sup> and PYTHIA 8.223 [629, 630] with the tune 4C [708]. These programs are formally LL accurate, though they include effects beyond Eq. (III.10). Results are presented separately for quark and gluon jets, using the  $Z + q$  and  $Z + g$  hard-scattering processes. We select jets with transverse momentum  $p_T > 500$  GeV and rapidity  $|y| < 2.5$ , clustered with the anti- $k_t$  algorithm [661] algorithm using a jet radius  $R = 0.8$ . A variety of two-point correlators (as defined in Sec. 1.2.1) are studied, with  $\alpha \in \{0.5, 1.0, 2.0\}$  corresponding to the Les Houches Angularity [709], width, and mass, respectively. Additionally, various SoftDrop grooming parameters are studied by varying  $\beta \in \{0, 1, 2\}$  and  $z_{\text{cut}} \in \{0.05, 0.1, 0.2\}$ . The plots in this section are shown for HERWIG 7.1 but the same qualitative conclusions hold for PYTHIA 8.2.

The sensitivity to  $\alpha_s$  is shown in Fig. III.5, in the case of normalized squared jet mass ( $e_2^{(2)}$ ) for quark and gluon jets.<sup>9</sup> To the left of the low-mass NP peak, the shape of the  $e_2^{(2)}$  distribution is nearly flat for quarks and nearly linear (in the log-space) for gluons, as expected from Eq. (III.10). Increasing  $\alpha_s$  shifts the quark distribution up but has nearly no impact on the shape of the distribution below the peak. The size of the NP peak is significantly impacted by the value of  $\alpha_s$ , with smaller  $\alpha_s$  values implying that NP effects are more important to the groomed mass distribution. In contrast, the slope of the distribution for gluons does change with the variations in  $\alpha_s$ , but there is no low-mass NP peak.

The robustness to NP effects is shown in Fig. III.6, using the same  $e_2^{(2)}$  distribution but now with variations in the modeling of hadronization and MPI effects. Overall, there is excel-

<sup>8</sup>This uses the MMHT 2014 PDFs sets [601], the angular-ordered parton shower [702], default settings for the cluster hadronization model [703], a Multiple Parton Interaction model [704–706], and plain color reconnections [707]. The factorization and renormalization scale is set to  $\mu_R^2 = \mu_F^2 = (\sum_i p_T^i)^2$  where  $p_T^i$  is the transverse momentum of the  $i$ -th jet.

<sup>9</sup>Here, we have varied the  $\alpha_s$  value used in the PS only. For future work, it would be interesting to test the sensitivity to  $\alpha_s$  in the hard matrix element and PDFs as well.

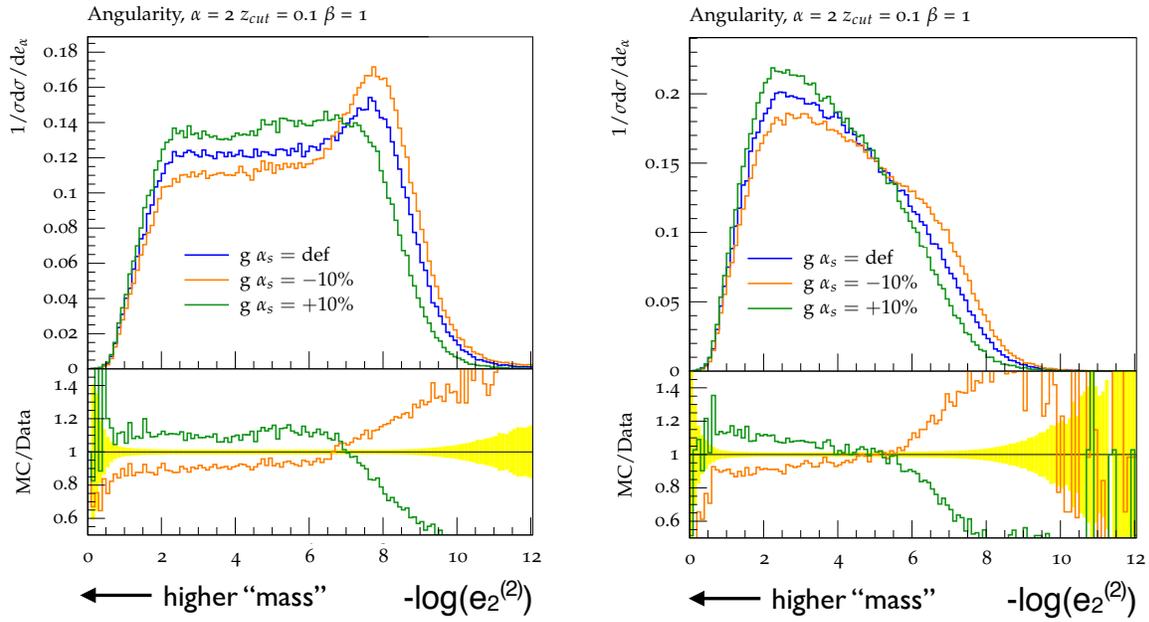


Fig. III.5: The sensitivity to  $\alpha_s$  in HERWIG 7. Shown is the distribution of the normalized squared jet mass ( $e_2^{(2)}$ ) for quark jets (left) and gluon jets (right), with higher values of the mass are on the left. The blue line uses  $\alpha_s = 0.118$  while the green and orange lines have the value of  $\alpha_s$  varied by 10%. The lower panels show the ratio with respect to the  $\alpha_s = 0.118$  curve.

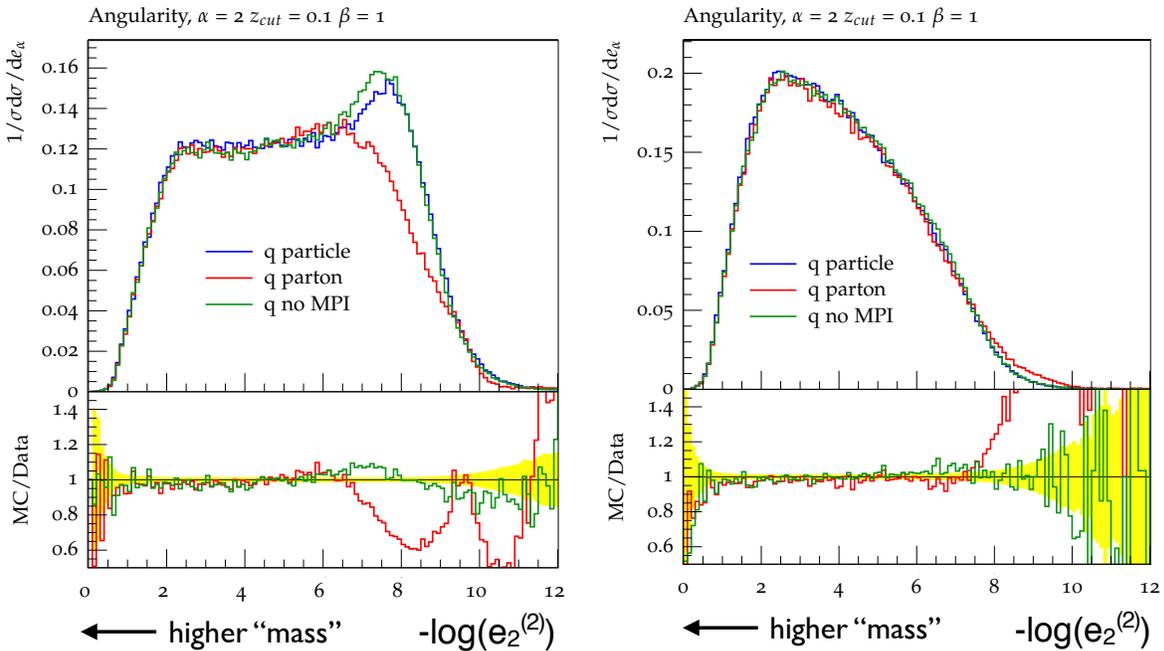


Fig. III.6: The robustness to NP effects in HERWIG 7, with the same distributions as Fig. III.5. The blue line shows the default particle-level simulation that includes the standard cluster hadronization model. The red curve has hadronization turned off and the green curve is the same as the blue, but with the HERWIG 7 model for multiple parton interactions (MPI) turned off. MPI is also off for the red curve.

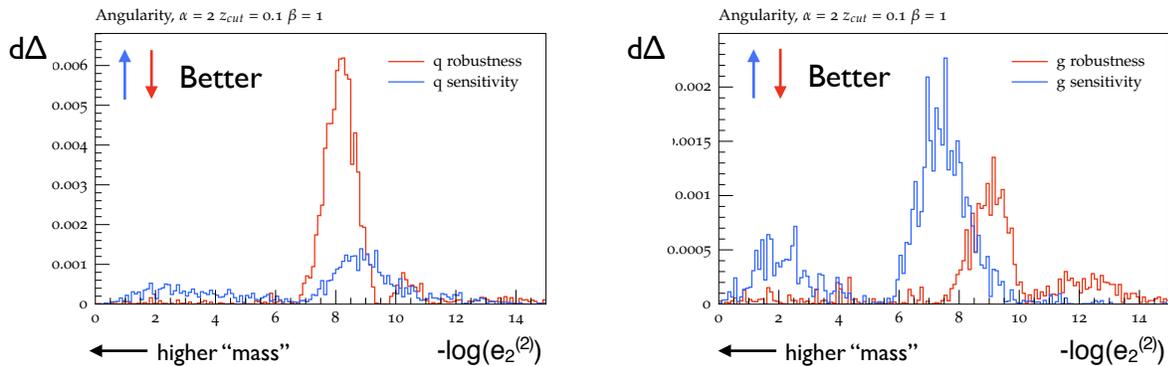


Fig. III.7: The integrand of Eq. (III.11) in HERWIG 7. Shown is the normalized squared jet mass ( $e_2^{(2)}$ ) for quark jets (left) and gluon jets (right), with higher values of the mass are on the left. The baseline  $f$  function from Eq. (III.11) is the same for the red and blue curves. For blue (sensitivity) the  $g$  function is from varying  $\alpha_s$  by 10%, whereas for red (robustness) the  $g$  function hadronization is turned off. Note the different vertical scale in the left and right plots.

lent stability in the  $e_2^{(2)}$  distribution. The low-mass peak for quarks is almost entirely due to hadronization effects. According to HERWIG 7, the impact of hadronization is much smaller for gluon jets, as expected since perturbative effects push the distribution away from the NP-sensitive region in Eq. (III.4).

Figure III.7 shows the distribution of the differential separation power (integrand of Eq. (III.11)), using the variation with  $\alpha_s$  as the sensitivity and the change from turning off hadronization as the robustness. For quark jets, Figs. III.5 and III.6 showed that the biggest variations with  $\alpha_s$  occurred at low mass which is also where NP effects are largest. This corresponds to the peak in the blue and red distributions in Fig. III.7 occurring in nearly the same location. In contrast, the peaks are more well-separated in Fig. III.7 for gluon jets and the blue is shifted to higher mass values where there is more perturbative control. This suggests that gluon-enriched samples are going to play an important role for  $\alpha_s$  determination from jet substructure.

A summary of the sensitivity-robustness tradeoff for many two-point correlators is presented in Fig. III.8. As already observed for the jet mass, gluons tend to have superior sensitivity and robustness compared with quarks. This is not surprising, as gluons have more perturbative radiation than quarks ( $C_A > C_F$ ). The jet mass has  $(\Delta_{\text{sensitivity}}, \Delta_{\text{robustness}}) = (0.097, 0.0043)$ ,  $(0.015, 0.0046)$  for quarks and gluons, respectively, using  $\beta = 0$  and  $z_{\text{cut}} = 0.1$ . The groomed two-point correlators with the best quark and gluon sensitivity and robustness have  $\alpha = 1$ ,  $z_{\text{cut}} = 0.05$  and  $\beta = 1$  or  $\beta = 2$ . The choice of  $\alpha = 1$  correspond to  $k_t$  (a.k.a. width) instead of mass, which may not be so surprising since the scale of  $\alpha_s$  is set by  $k_t$  and not mass. Interestingly,  $k_t$  with  $\beta = 0$  has significantly worse sensitivity than  $\beta = 1$  or  $\beta = 2$ , highlighting the importance of having some double-logarithmic information. This study suggests that  $k_t$  observables are important to include in future experimental and theoretical studies. The three-loop anomalous dimensions for the Energy-Energy Correlator (EEC) in  $e^+e^-$ , which is a  $k_t$  sensitive observable, were recently derived [710]. This will allow for a study of  $k_t$  sensitive observables at N<sup>3</sup>LL in  $e^+e^-$ , and a comparison with mass-type observables. It would also be interesting to study how the robustness versus sensitivity picture changes when considering only subsets of the available observable range.

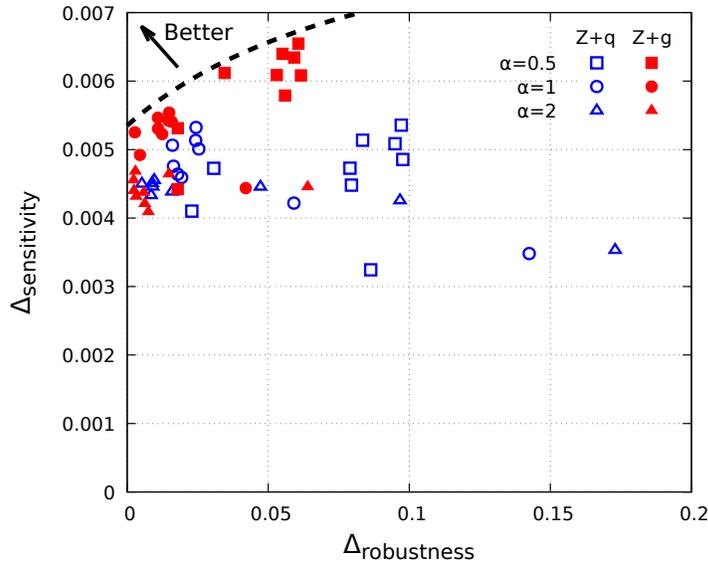


Fig. III.8: The tradeoff between sensitivity and robustness for 27 two-point correlator/jet grooming combinations. Open blue symbols represent quark jets and closed red symbols represent gluon jets.

### 1.4.3 The Issue of Casimir Scaling

While we have illustrated that groomed jet mass provides excellent sensitivity to  $\alpha_s$ , particularly for gluon jets, one problem that is immediately clear from Sec. 1.4.1 is that the leading behavior of the distributions is always dominated by the product  $\alpha_s C_i$ . While this is broken at higher perturbative orders, it implies that at lowest order there is a complete degeneracy of the value of  $\alpha_s$  and the quark versus gluon fraction of jets. This problem is not faced for dijet event shapes at  $e^+e^-$  colliders, which are almost entirely quark dominated.

There are a variety of different approaches to overcome this problem, each with their own advantages and disadvantages. First, the quark and gluon fractions are perturbatively calculable given the parton distribution functions (PDFs). Therefore, perturbatively calculating the quark and gluon fractions inputs the most possible information, and should correspondingly lead to the best sensitivity for  $\alpha_s$ . This has the downside, however, of also introducing sensitivity to the PDFs, which in principle should be fitted along with  $\alpha_s$  [711]. This difficulty also enters into other  $\alpha_s$  extractions at the LHC, for example the 3-2 jet rate [639] or the energy-energy correlator [635, 636]. One of the key hopes of using jet substructure was that the sensitivity to the PDFs could be minimized, but that is not the case if the quark and gluon fractions cannot be determined independently.

A second approach is to simultaneously fit for  $\alpha_s$  and the quark/gluon fraction. In the re-summation regime, the jet mass distribution only depends on  $\alpha_s$  and the quark/gluon fraction.<sup>10</sup> Due to the fact that the two fractions must add to unity, this introduces a single additional parameter in the fit (see Fig. III.9). The degeneracy between the quark/gluon fraction and  $\alpha_s$  is broken by higher-order effects. Furthermore, different  $e_2^{(\alpha)}$  have different dependence on  $\alpha_s$  and  $C_i$  at higher orders. Therefore, the measurement of multiple two-point correlators would

<sup>10</sup>At high mass, in the fixed-order regime, there is still residual explicit dependence on PDFs as is the case for the total cross-section. This regime also suffers from sample-dependent effects where the quark and gluon substructure distributions have residual dependence on the hard process.

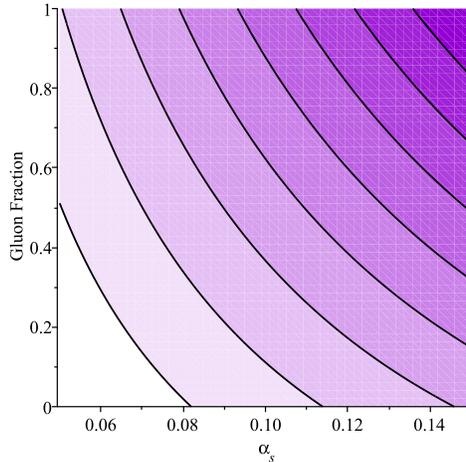


Fig. III.9: The slope of the probability distribution of  $\log(e_2^{(2)})$  is proportional to  $\alpha_s(C_A f_g + C_F(1 - f_g))$ , which is plotted above. The degeneracy between  $\alpha_s$  and  $f_g$  (gluon fraction) are represented by the banana-shaped isocontours.

allow the degeneracy to be broken. From a theoretical perspective, this significantly complicates the analysis, since it would require precise predictions to be made for the joint distribution of multiple two-point correlators.

In our fitting study in Sec. 1.5, we consider both of the above approaches. It would be interesting to develop other approaches to disentangling the quark and gluon fractions and  $\alpha_s$ . Without some kind of conceptual breakthrough, though, we expect that the quark/gluon fraction will be a limiting aspect of  $\alpha_s$  extractions from jet substructure at the LHC.

#### 1.4.4 Normalized vs. Unnormalized Distributions

In addition to the complication of quark and gluon fractions, another issue which appears for the extraction of  $\alpha_s$  from jet substructure is the issue of normalization.<sup>11</sup> Unlike for  $e^+e^-$  event shapes, the Born dijet cross-section in  $pp$  is sensitive to the value of  $\alpha_s$ . This implies that the rate itself, in particular the absolute quark and gluon jet rates, carries information regarding  $\alpha_s$ .

In our studies, we have decided to focus on using normalized distributions. At a conceptual level, this is because we want to perform a true measurement of  $\alpha_s$  from jet substructure, which is not dominated by the overall jet rate. But there are two other practical concerns that favor using normalized distributions.

First, it is currently only possible to perform precise measurements of the groomed mass using normalized distributions. Experimentally, the absolute rate is determined by the acceptance from kinematic requirements on the jet  $p_T$ . The sophisticated in-situ jet energy scale and resolution program carried out for small-radius jets [712–715] has not yet reached the same level of maturity for groomed large-radius jets. That said, this is simply a matter of time, and experimental efforts have already started in this direction [716]. A key challenge that still remains is to fully understand the correlations in the calibrations and uncertainties between the jet energy and the jet substructure observables.

Second, the use of normalized distributions minimizes the sensitivity to the PDFs. As

<sup>11</sup>We thank Gavin Salam for interesting discussions on this topic.

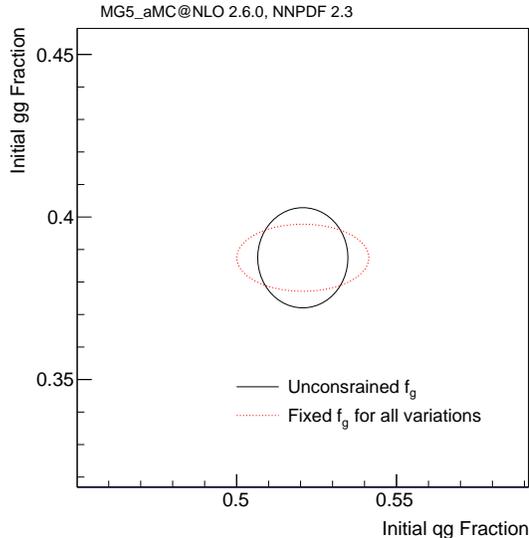


Fig. III.10: The fraction of  $qq$  and  $gg$  initial states for leading-order dijet production with  $p_T > 200$  GeV, simulated at leading order with MG5\_aMC 2.6.0 [88] using NNPDF 2.3 [599]. The ellipses correspond to the uncertainty from the 100 error PDF sets. For the red dashed line, the fraction of outgoing gluon jets  $f_g$  is constrained to be the same for all variations (and equal to the nominal PDF set).

discussed in Sec. 1.3.2, grooming renders the quark and gluon two-point-correlator distributions universal (in the resummation regime). Therefore, the measured distribution only depends on the fraction of gluon jets  $f_g$  that pass the event selection. In contrast, the total cross-section depends on the relative proportions of all possible partonic initial states, which introduces a source of uncertainty that is not present for the normalized cross-section. For example, the initial state could be one of  $qq$ ,  $qg$  or  $gg$ . The relative proportions can be parameterized by two numbers  $f_{qg}$  and  $f_{gg}$  where  $f_{qg} + f_{gg} = 1 - f_{qq}$ . Figure III.10 shows the uncertainty in  $f_{qg}$  and  $f_{gg}$  from leading-order PDFs. Fixing  $f_g$ , which carries all of the PDF sensitivity for the shape measurement, has little effect on the uncertainty in  $f_{qg}$  and  $f_{gg}$ . Thus, using unnormalized distributions results in additional PDF sensitivity from also measuring the total cross-section in addition to the shape of the substructure distribution.

From the perspective of perturbative accuracy, there is an important issue with using normalized distributions, which is that jet-shape observables start at  $\mathcal{O}(\alpha_s)$ . Specifically, the slope of the groomed mass distribution is  $\mathcal{O}(\alpha_s)$ . This implies that to have an  $\mathcal{O}(\alpha_s^2)$  uncertainty on the slope (as required to gain entry into the PDG world average), one needs to have  $\mathcal{O}(\alpha_s^3)$  control, namely the NNLO  $2 \rightarrow 3$  process for a jet with two constituents. For the case of  $e^+e^-$ , this level of accuracy has been achieved, where the NNLO corrections to  $e^+e^- \rightarrow 3$  jets are known and indeed used in extractions of  $\alpha_s$ . Due to recent progress in the calculation of the relevant amplitudes, we believe this is a realistic target for jet substructure, but clearly requires substantial further work.

## 1.5 Idealized Performance at the LHC

The purpose of this subsection is to make some numerical estimates regarding  $\alpha_s$  extraction from jet substructure at the LHC. Many simplifying assumptions are made, with the goal of motivating a more complete effort within the context of ATLAS and CMS in collaboration with

theorists. First, we illustrate how  $\alpha_s$  and the gluon fraction can be simultaneously extracted from the distribution of various two-point correlators. Next, we estimate the needed experimental precision required to make a useful measurement of  $\alpha_s$ . While both the theory and experimental precision will continue to improve over the next years, the community has already demonstrated that the work can begin with the first round of groomed jet mass results [646–651].

### 1.5.1 Extraction of Theory Templates

A complete extraction of  $\alpha_s$  will require matching resummed results to high fixed order and also estimating NP effects. The two sets of predictions for dijets thus far have been matched to LO [646, 647] and NLO [648, 649] and have used hadronization models to study NP corrections [648, 649]. Performing high-order fixed-order matching is conceptually straightforward but computationally expensive; while this will be required eventually, we focus here on a demonstration without matching. Therefore, we isolate the resummation regime,

$$e_2^{(\alpha)} \Big|_{\text{NP}} \leq e_2^{(\alpha)} \leq z_{\text{cut}} R^\alpha, \quad (\text{III.12})$$

where  $e_2^{(\alpha)} \Big|_{\text{NP}}$  is given in Eq. (III.4), such that regions of phase space that are highly sensitive to NP or fixed-order effects are removed. In this range, NLL calculations exist in analytic formulae that can be varied on-the-fly [648, 649]. Figure III.11 shows the quark and gluon templates for  $\alpha = \{1, 2\}$  and  $\beta = \{0, 1\}$ , truncated according to Eq. (III.12).

From these NLL distributions, pseudo-data are then generated from the binned analytic probability distribution  $t(\alpha_s, f_g)$ . These distributions are a superposition of the quark and gluon distributions and depend only on  $\alpha_s$  and  $f_g$ . Each pseudo-dataset has  $n$  events and its binned representation is denoted by  $h(\alpha_s, f_g, n)$ . For a given pseudo-dataset, the fitted values of  $\alpha_s$  and  $f_g$  are determined from a  $\chi^2$ -like fit:

$$\alpha_s, f_g = \text{argmin} \sum_i \frac{(h_i(\alpha_s, f_g, n) - t_i(\alpha_s, f_g))^2}{\sigma(h_i(\alpha_s, f_g, n))^2}, \quad (\text{III.13})$$

where  $t_i, h_i$  are the bin content of histograms  $t$  and  $h$ , and  $\sigma(h_i)$  is the statistical uncertainty in bin  $i$  of histogram  $h$ . In practice, there would also be systematic uncertainties (see Sec. 1.5.2), but the purpose of this study is to simply illustrate the sensitivity to  $\alpha_s$  and  $f_g$  for a given number of events.

An example fit is demonstrated in Fig. III.12 for the case of  $\alpha = 2$  and  $\beta = 0$ . The left plot of Fig. III.12 shows the  $\chi^2$  from Eq. (III.13) for two samples, one with 20% gluons and one with 80% gluons. The true value is taken to be  $\alpha_s = 0.1$ , and as expected, the  $\chi^2$  probability is high for  $f_g = 0.2$  and  $f_g = 0.8$ .<sup>12</sup> The banana shapes of the curves are a consequence of the degeneracy due to Casimir scaling discussed in Sec. 1.4.3. From one sample alone, there is essentially no ability to distinguish between a larger  $\alpha_s$  and a smaller  $f_g$ ; the only constraint comes from the fact  $0 \leq f_g \leq 1$  which results in a crude bound on  $\alpha_s$ . This is shown in the right plot of Fig. III.12 where the distribution is marginalized over  $f_g$  and normalized to unity. One can view this as the posterior probability of the fitted  $\alpha_s$ : the peak is the fitted value of  $\alpha_s$  and the width is the uncertainty. When  $f_g$  is known, the uncertainty in  $\alpha_s$  is significantly reduced; this is illustrated with pure quark and gluon samples. Due to the larger color factor, the measurement with pure gluon jets is more sensitive to  $\alpha_s$  than the fit using pure quark jets, as anticipated in Sec. 1.4.1. Using both the  $f_g = 0.2$  and  $f_g = 0.8$  samples to fit for  $\alpha_s$ , one can extract a  $\sim 30\%$  measurement of  $\alpha_s$ , but there is no clear peak at the correct value of  $\alpha_s$  due to

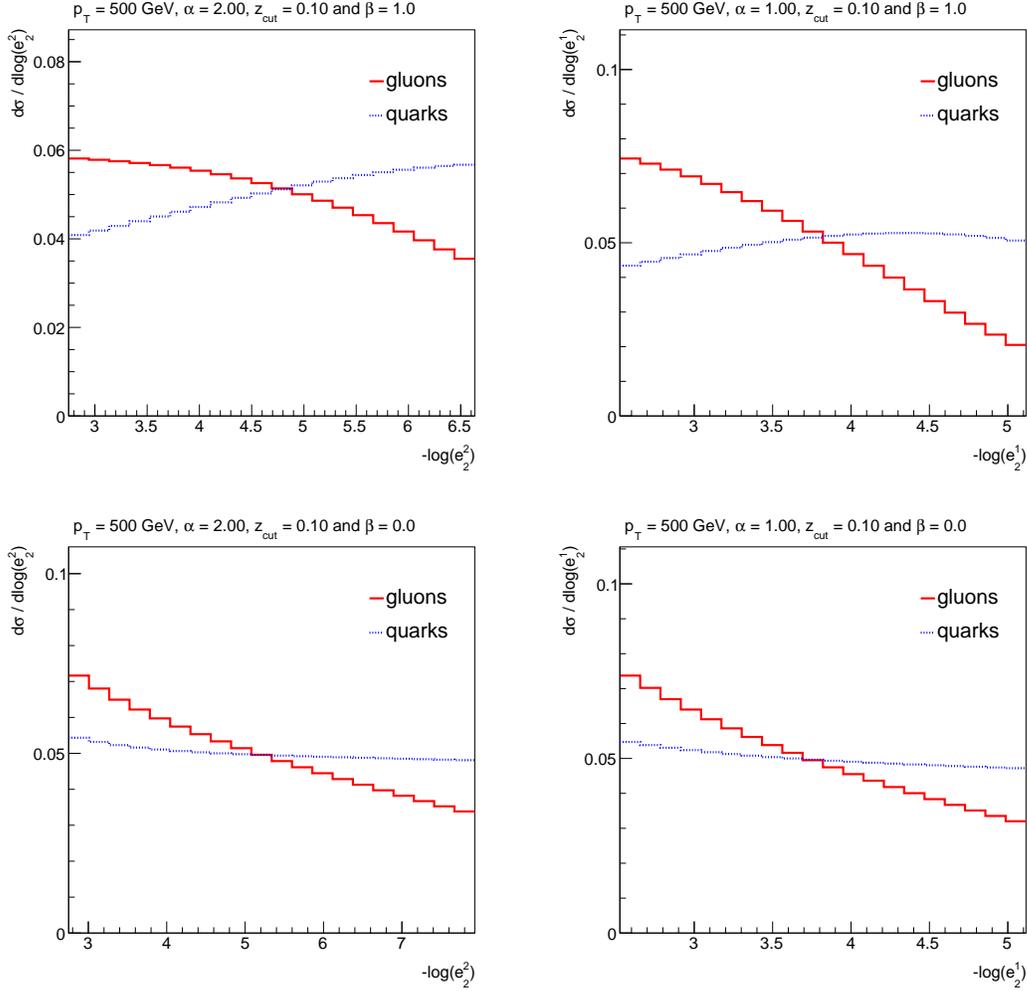


Fig. III.11: The quark and gluon templates, computed at NLL [648, 649] for  $\alpha = 2$  (left column) and  $\alpha = 1$  (right column), with grooming parameters  $z_{\text{cut}} = 0.1$  and  $\beta = 1$  (top row) and  $\beta = 0$  (bottom row). Note that larger masses are on the left so that the NP regime is on the right and the fixed-order regime is on the left.

the Casimir degeneracy.

One way to improve the situation is to combine multiple  $\alpha$ ,  $\beta$ , and  $z_{\text{cut}}$  values (only  $\alpha$  and  $\beta$  are varied here). Figure III.13 shows the  $f_g, \alpha_s$  fit for all of the  $\alpha, \beta$  values from Fig. III.11 that were not shown in Fig. III.12. As in Fig. III.12, there are two banana-shaped regions that correspond to the  $f_g = 20\%$  and  $f_g = 80\%$  samples. The tilt of the bananas is slightly different than the  $\alpha = 2, \beta = 0$  case and so there is a possibility to gain from combining the information in the observables. In practice, a challenge with a multi-observable extraction strategy is the need to understand correlations between observables. At the moment, joint distributions of two-point correlators are only known to NLL accuracy without grooming [657].

With that caveat in mind, Fig. III.14 shows the result of a combined fit assuming that all four distributions are statistically (and systematically) independent. This is a rather strong

<sup>12</sup>It is not necessarily peaked at this value because this is the result of one pseudo-experiment. Averaging over many pseudo-experiments results in peaks at  $f_g = 0.2$  and  $0.8$ .

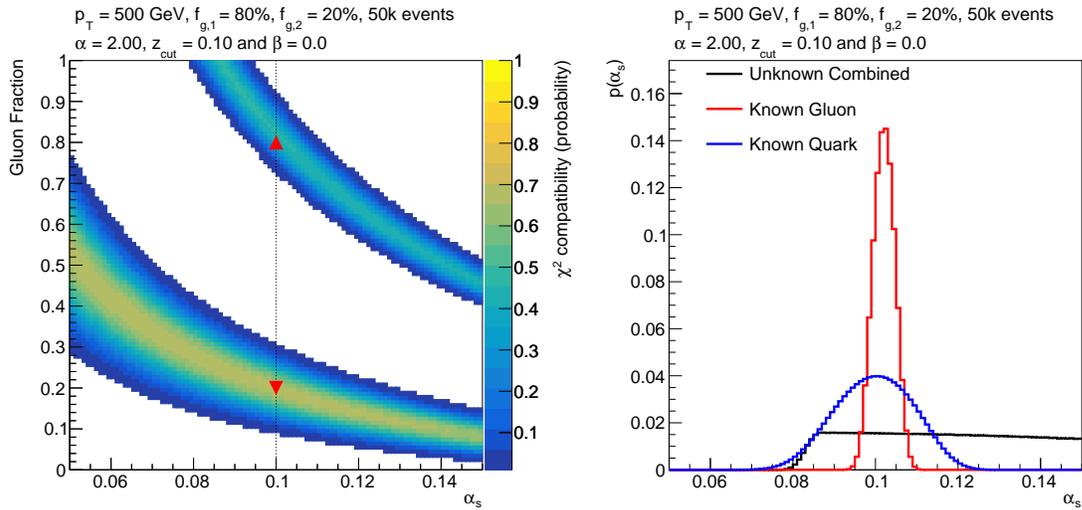


Fig. III.12: Left: the probability of the minimized  $\chi^2$  (assuming  $n_{\text{bins}} - 1$  degrees of freedom) from Eq. (III.13) as a function of  $f_g$  and  $\alpha_s$  for one sample with 80% gluons and another sample with 80% quarks. The true value of  $\alpha_s$  is 0.1, as indicated by triangle markers. Right: The right plot marginalized over  $f_g$  and normalized to unity. The three lines correspond to the fit performed on a pure sample of quarks, a pure sample of gluons, or a mixed sample of ( $f_g \in \{0.2, 0.8\}$ ) where the fractions are not known a priori. This is the result from one pseudo-experiment with 100k events.

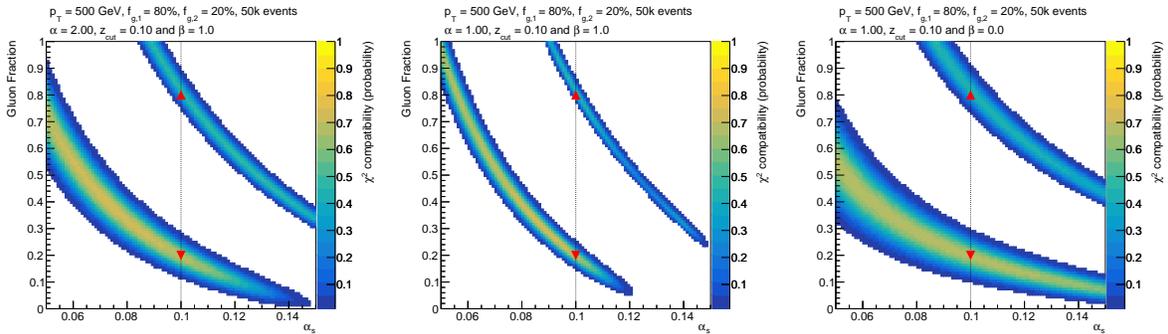


Fig. III.13: The same as the left plot of Fig. III.12, but for the remaining three observables from Fig. III.11.

assumption that is unlikely to be even approximately true in practice. However, the benefit of having different tilts in the  $f_g, \alpha_s$  plane is clearly shown and would be a generic feature of a multi-observable fit, even if the size of the gain is not as significant as shown here. It is important to emphasize that the black curve in Fig. III.14 assumes no prior knowledge of the gluon fraction of the event samples. The fit can of course be improved by using some knowledge of the gluon fractions from the hard scattering process convolved with PDFs, as discussed in Sec. 1.4.4.

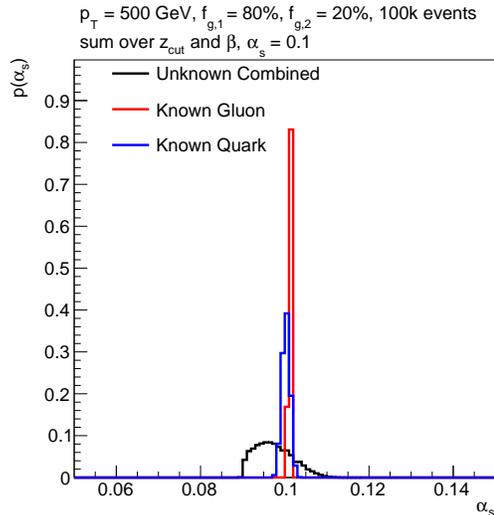


Fig. III.14: Marginalizing the  $f_g, \alpha_s$  fit over the gluon fraction for a fit that combines the four observables from Fig. III.11. The unknown combined mixture uses two samples with  $f_g = 20\%$  and  $f_g = 80\%$ .

### 1.5.2 Estimate of Experimental Resolution

To keep pace with precise theory predictions, the experimental resolution must be well-understood in order to ensure both precision and accuracy of jet substructure measurements. To estimate the impact of detector resolution on  $\alpha_s$  fit illustrated in Sec. 1.5.1, a fast simulation from DELPHES 3.4.1 [717] was studied using particle-level input from PYTHIA 8.210. This setup uses a CMS-like detector with jets built from particle-flow objects. There are generically two regimes for determining the experimental resolution. At high mass, there are well-resolved hard prongs in the jet, so the resolution is set by the jet energy resolution of those “sub-jets”. At low mass, the groomed jet is defined by nearly collinear splittings at which point the angular resolution can dominate.

Figure III.15 shows the fractional  $e_2^{(2)}$  resolution estimated from DELPHES. The resolution is smallest at high mass due to the excellent energy resolution of the CMS-like detector. The resolution at low mass is about 10% near 15 GeV and reaches 30% near the limit of  $\mathcal{O}(\text{few GeV})$ . Encouragingly, the fits in Sec. 1.5.1 relied only on the regime where  $\sim 10\%$  resolution seems to be achievable, which gives an indication that a 10% extraction of  $\alpha_s$  should be feasible. As a check that this estimated resolution is sensible, Fig. III.16 compares the migration matrix extracted from Delphes to the one published in the recent ATLAS measurement [650]. Here, the comparison is between the particle-level and detector-level groomed mass values. The migration is qualitatively the same, with an excellent diagonal behavior (low migrations) at high mass and a worse resolution at low mass.

To understand the sensitivity to jet mass scale and resolution uncertainties, we conduct pseudo-experiments by sampling from the templates described in Sec. 1.5.1 and smearing with the migration matrix shown in the left plot of Fig. III.16. We then unfold with another migration matrix [718, 719] that has the jet mass scale shifted or smeared by a fixed amount, and extract  $\alpha_s$  via Eq. (III.13) but for fixed and known  $f_g = 0$ . The results of this procedure are shown in Fig. III.17. There is little sensitivity to the jet mass resolution while there is a large uncertainty in the extracted  $\alpha_s$  if the uncertainty in the jet mass scale exceeds a few percent. Current jet

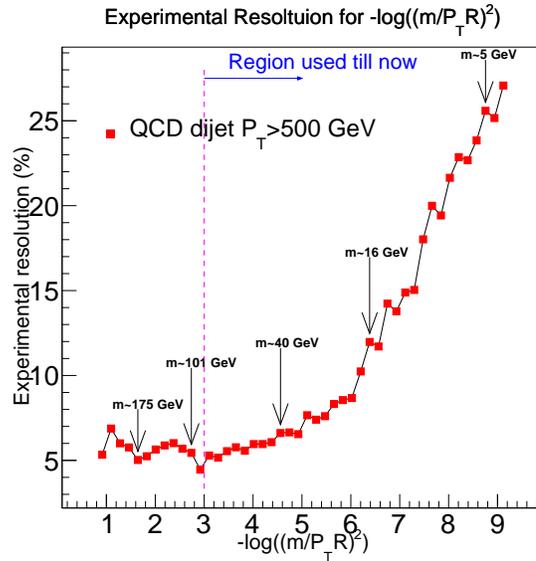


Fig. III.15: The fractional  $e_2^{(2)}$  distribution determined from dijet events simulated with PYTHIA 8 plus DELPHES. Arrows indicate the mass at given values of the two-point correlator. The upper bound of the resummation regime is indicated by a dashed line. Larger masses are to the left.

mass scale uncertainties are below 5% and resolution uncertainties are below 20% [716,720]. The mass scale uncertainty is as small as 2% in various regions of phase space. This again suggests that a  $\sim 10\%$  measurement of  $\alpha_s$  is feasible, as the uncertainty on the jet mass scale reaches the same level of maturity as the jet energy scale over the next few years.

## 1.6 Conclusions and Future Outlook

In this section, we performed a preliminary study assessing a possible extraction of the strong coupling constant  $\alpha_s$  from groomed jet substructure measurements at the LHC. This has been made possible by recent advances in the calculation of groomed event shapes, which allow them to be resummed to NNLL accuracy, as well as advances in the understanding of the experimental uncertainties for groomed jet observables.

We have highlighted a number of features that an extraction of  $\alpha_s$  from groomed substructure would offer. In particular, the grooming procedure suppresses NP hadronization corrections to the mass distribution, extending the range of perturbative control by up to a factor of  $\sim 100$  as compared to the ungroomed case, where hadronization corrections are a dominant uncertainty. Furthermore, the nature of the NP corrections is completely different, and therefore a groomed extraction could provide complementary information to event shape measurements at LEP and other  $e^+e^-$  colliders.

Using the groomed jet mass as a concrete example, we performed a feasibility study for the extraction of  $\alpha_s$  from groomed jet substructure. Considered separately, the groomed quark and gluon distributions both exhibit good sensitivity to  $\alpha_s$ . However, this study also highlighted a key issue relevant at the LHC. The leading behavior of the shape of the distribution is sensitive to the product  $\alpha_s C_i$ , where  $C_i$  is the color Casimir factor, namely  $C_A$  for gluon jets and  $C_F$  for quark jets. This immediately implies that there is a degeneracy between the value of  $\alpha_s$  and the quark/gluon fraction of the sample. We highlighted two possible ways of overcoming this degeneracy, either by a fixed-order calculation of the quark and gluon fractions, or by a

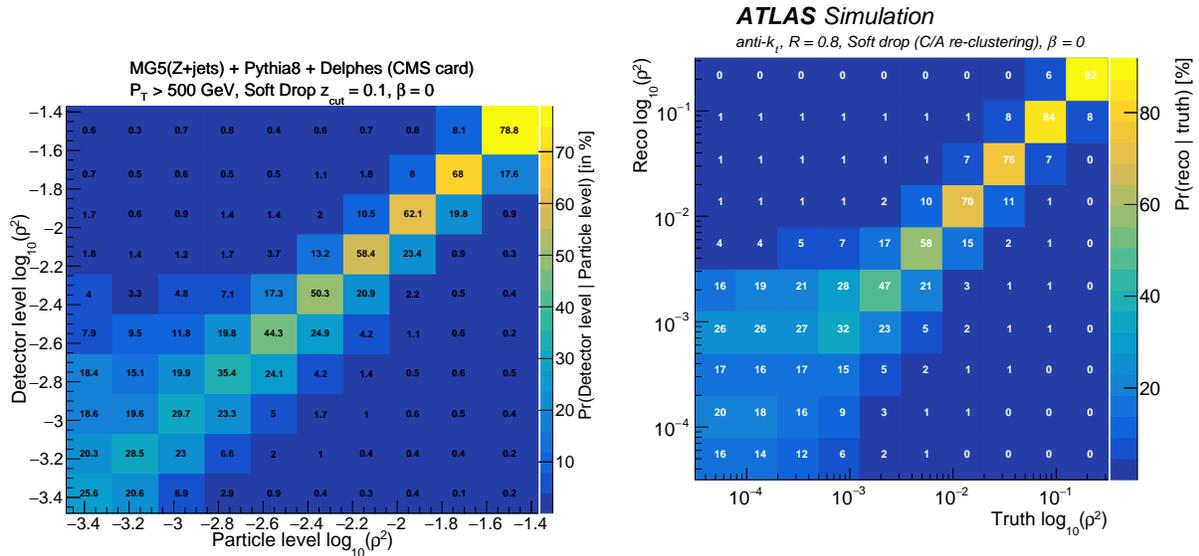


Fig. III.16: Left: The migration matrix between particle-level and detector-level using the Delphes simulation (left) and from the ATLAS measurement (right; reproduced from Ref. [650]). Unlike previous plots, larger masses are on the right (shown this way to match the ATLAS result).

simultaneous measurement of different substructure observables. This would ideally be combined with multiple samples with different (and relatively pure [721]) quark/gluon fractions. While we expect significantly better precision to be achievable by the explicit calculation, this would introduce a stronger dependence on the PDFs which then should ideally be fitted simultaneously.

With currently achievable experimental and theoretical uncertainties, we have shown that an extraction of  $\alpha_s$  at the 10% level is realistic using the currently available data at the LHC. We believe that this study motivates a serious effort to extract the strong coupling constant  $\alpha_s$  from jet substructure at the LHC using groomed jet shapes. A serious analysis will require advances on both the theory and experiment sides, so we conclude by discussing some of the major obstacles that must be overcome.

On the theory side, we can separately discuss three primary aspects of the calculation:

- **Resummation Accuracy:** For  $e^+e^-$  event shapes, the current state of the art is  $N^3\text{LL}$ . This has been achieved only for a few select observables using soft-collinear effective theory. The extension to  $N^3\text{LL}$  accuracy for the groomed jet mass would require only the calculation of the anomalous dimensions for the groomed soft function to three loops, which is possible using currently available techniques. At this level of accuracy, it will also be important to assess perturbative power corrections in  $z_{\text{cut}}$ . While Refs. [648, 649] have shown that these corrections are numerically small, they may become important if multiple  $z_{\text{cut}}$  values are used that are not much less than one.
- **Fixed-Order Matching:** A key aspect, and a potential complication to achieve theoretical accuracy competitive with  $e^+e^-$  determinations, is fixed-order matching. For  $e^+e^-$ , the perturbative corrections to  $e^+e^- \rightarrow 3$  jets are known to NNLO [722–725]. To achieve a similar perturbative accuracy for the matching for the groomed jet mass will require  $2 \rightarrow 3$  matrix elements at NNLO. While results for the amplitudes are just becoming available [123, 124, 289, 291, 679], it will be a while before numerically-efficient evaluations

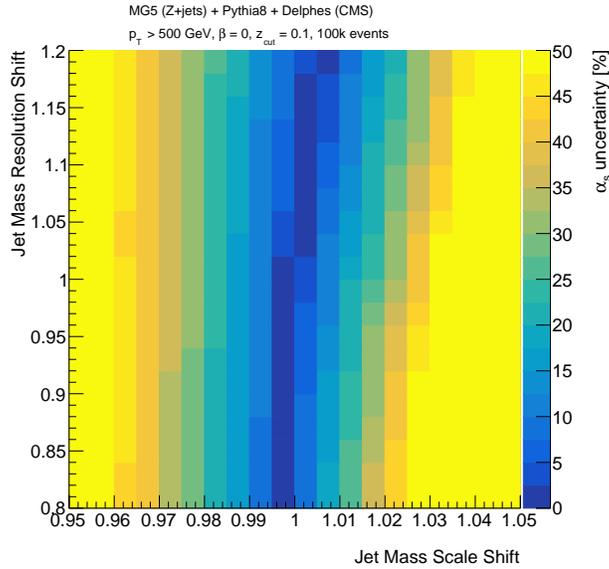


Fig. III.17: The impact of jet mass scale and jet mass resolution uncertainties on the uncertainty in the measured value of  $\alpha_s$ . See the text for details.

of the relevant cross-sections are available. We believe that this is currently the theoretically most difficult ingredient.

- **Non-Perturbative Corrections:** Finally, in addition to improving our understanding of the perturbative accuracy of the observable, it will also be crucial to improve our understanding of the NP aspects of groomed observables (mass and others). While it has been shown through PS studies that NP effects are suppressed throughout a large component of the distribution, it will be important to quantify this further. In the case of  $e^+e^-$  event shapes, operator definitions of soft matrix elements allow for field-theoretic definitions of the NP parameters. Ideally, this could also be performed for groomed observables, placing the groomed jet mass on a firmer theoretical footing.

Experimentally, there are three broad topics that need to be addressed to achieve a precision extraction of  $\alpha_s$ :

- **Mass Scale Uncertainties:** At the moment, ATLAS [650] and CMS [651] have very different approaches for determining the jet mass scale uncertainty, both with known limitations. CMS performs a fit to the hadronic  $W$  boson mass peak in one-lepton  $t\bar{t}$  events and takes the shift in the peak position as the uncertainty in the jet mass scale (which is negligible and thus ignored) [726]. Two challenges with this approach are that (a) the peak position is a convolution of particle-level and detector-level effects and (b) it is not clear that uncertainties derived for boosted  $W$  bosons should be the same as for generic quark and gluon jets at all masses. One can overcome (a) with a technique like forward-folding [727, 728]. Various PS generators can be studied, but it is likely not sufficient to have one global model comparison. The impact on the jet mass resolution in one topology may be completely different than the impact of the scale, resolution, or acceptance in another topology.

By contrast, the ATLAS measurement propagates constituent-based uncertainties through to the groomed mass. These uncertainties are derived from matching tracks to calorimeter-

cell clusters and studying the energy and angular matching. Studies have shown that this “bottom-up” approach works well for reproducing the jet energy scale [729], which has been validated also for groomed jets in Ref. [650]. However, this does not hold exactly for the mass, which is not linear in the constituent energies [730]. The uncertainties are validated using the standard ATLAS approach using track-jets [716, 731], but to achieve higher precision, a more detailed understanding of the impact of energy thresholds, fluctuation correlations, and calorimeter cluster merging will be required.

- **Mass Resolution Uncertainties:** ATLAS and CMS use their same respective approaches for the mass resolution as for the mass scale (i.e. bottom-up and  $W$  mass peak). ATLAS validates their approach in a similar manner as CMS, by using the  $W$  mass peak from  $t\bar{t}$  events.<sup>13</sup> From Fig. III.17, it seems that the mass resolution will not be the limiting factor, but it could be if the situation is not improved as the jet mass scale precision improves.
- **Pileup Modeling and Mitigation:** Grooming significantly reduces the impact of pileup, but if the Run 3+ data are to be used (pileup levels of 80+), then a significantly better and more detailed understanding of the degradation due to pileup will be required. Statistical uncertainties are currently not dominant, so it is conceivable that the higher instantaneous luminosity data will not be required for the precision  $\alpha_s$  extraction. This may change if one wants to exploit the largest lever-arm possible; to access the highest  $p_T$  jets, we will need more data.

We are optimistic that these difficulties on both the theory and experiment side can be overcome, enabling a precision probe of the strong coupling constant from jet substructure in the LHC environment. We also hope that this study motivates additional investigations to broaden the potential toolbox for  $\alpha_s$  extraction. At minimum, it would be interesting to study in more detail the use of other observables beyond the two-point correlators as well as alternative grooming techniques beyond mMDT/SoftDrop. More ambitiously, one potentially interesting direction is the use of track-based measurements, since it may be the case that it is ultimately experimental uncertainties that are the limiting factor. From the experimental perspective, this significantly reduces uncertainties, particularly in a high pile-up environment. From the theoretical side, however, it introduces the need for NP track functions. That said, only certain moments of the track functions are required for special observables, and it may be possible to perform a combined fits for the moments of the track functions and  $\alpha_s$ , much like how fits are performed for  $e^+e^-$  event shapes. Significantly more theoretical work is required to see if this is truly a viable possibility, although we believe that this is well motivated.

## Acknowledgments

We thank the participants of Les Houches 2017 for a lively environment and useful discussions. The work of GS is supported in part by the French Agence Nationale de la Recherche, under grant ANR-15-CE31-0016, and by the ERC Advanced Grant Higgs@LHC (No. 321133). The work of FD and JT is supported by the DOE under grant contract numbers DE-SC-00012567 and DE-SC-00015476. IM and BN are supported in part by the Office of High Energy Physics of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231, and the LDRD Program of LBNL. AS acknowledges support from the National Science Centre, Poland Grant No. 2016/23/D/ST2/02605, the MCnetITN3 H2020 Marie Curie Initial Training Network, contract 722104 and COST Action CA15213 THOR.

<sup>13</sup>This validation does not have the same complete forward-folding machinery as was used by ATLAS for trimmed jets.

## 2 Performance versus robustness: Two-prong substructure taggers for the LHC <sup>14</sup>

The ability to robustly identify, or “tag”, boosted hadronically-decaying resonances plays a central role at the LHC, both in searches for new physics, as well as for probing the Standard Model in extreme regions of phase space. While a variety of powerful and theoretically well-understood tagging approaches exist, their behavior in a realistic experimental environment is complicated by a number of issues including hadronization, underlying event, pileup, and detector effects. In this section, we perform a systematic study contrasting the robustness and performance of different theoretical approaches to designing jet substructure observables. These include standard jet shape observables as well as various grooming strategies currently used by the LHC experiments. We also introduce a number of new observables, based on the idea of “dichroic ratios”, which are designed to simultaneously maximize both robustness and performance. We discuss the different choices used by ATLAS and CMS, and we introduce reliable metrics for quantifying robustness and performance for substructure observables. Additionally, we study the dependence of taggers on the polarization of hadronically decaying  $W$  bosons, and identify strategies to perform polarimetry using the hadronic decay products. We conclude by making recommendations for future tagging strategies to ensure robust procedures based on sound theoretical organizing principles.

### 2.1 Introduction

With the ever-increasing dataset from the Large Hadron Collider (LHC), we are able to probe the Standard Model and search for physics beyond the Standard Model in increasingly extreme regions of phase space. Theoretically well-understood observables that are sensitive to phase space extremes are therefore playing an important role at the LHC. One of the most exciting new approaches which has emerged at the LHC are tools from jet substructure, which allow for the identification of boosted hadronically-decaying particles within jets through a detailed study of their radiation patterns. Techniques from jet substructure have now been widely used for Standard Model measurements [693, 732–742] as well as for searches for new physics [743–764].<sup>15</sup>

With the growing importance of jet substructure techniques, there has been a significant effort by both the theory and experimental communities to develop a more detailed understanding of the theoretical and experimental behavior of jet substructure observables. On the theory side, this has been pursued through explicit calculations [364, 644, 646–649, 657, 695, 696, 765–785], scaling arguments [786–788], and machine learning [789–795] approaches. On the experimental side, there have been detailed studies of the behavior of substructure observables in data, and their interplay with experimental realities, such as detector resolution and pileup. Summaries can be found in Refs. [602–605], and Ref. [606] provides a review of recent theoretical and machine learning developments in jet substructure.

As a result of these efforts, there now exist a number of theoretically well-motivated jet substructure tools. For tagging hadronically-decaying  $W/Z/H$  bosons, which decay primarily to jets with two well-resolved prongs (also referred to as subjets), a variety of powerful two-prong taggers have been developed. Modern two-prong taggers typically consist of two or three ingredients: a groomer which removes low-energy contamination, a two-prong finder which identifies two (or more) hard subjets, and a jet shape observable which constrains radiation patterns in the jet. Often, the groomer and the two-prong finder are taken identical. For jet shapes, it is well understood how to organize and study their behavior using power counting [787].

---

<sup>14</sup> P. Loch, I. Moulton, B. Nachman, G. Soyez, J. Thaler (section coordinators); D. Bhatia, R. Camacho, G. Chachamis, S. Chatterjee, F. Dreyer, D. Kar, A. Papaefstathiou, T. Samui, A. Siódmok

<sup>15</sup>More LHC studies using jet substructure can be found at <https://twiki.cern.ch/twiki/bin/view/AtlasPublic> and <http://cms-results.web.cern.ch/cms-results/public-results/publications/>.

A variety of different classes of observables exist, for example the energy correlation functions [653, 796, 797] and  $N$ -subjettiness observables [798, 799], and their relation is understood. The behavior of groomers is also now well understood, and a number of groomers with favorable experimental and theoretical properties have been introduced [644, 645].

The theoretical understanding of the behavior of tagging observables is primarily based on perturbation theory, and it is therefore not always clear how this translates to experimental reality, due to the presence of hadronization, underlying event, detector effects, and pileup. Indeed, the different LHC experiments have settled on different tagging combinations. For grooming and prong finding, ATLAS using trimming [643] uses while CMS uses the modified Mass Drop Tagger (mMDT) [640, 644] and its generalization, SoftDrop [645]. For jet shape observable, ATLAS uses  $D_2$  [780, 787], while CMS uses  $N$ -subjettiness ratio  $\tau_{2,1}$  [798, 799] or  $N_2$  [796] with DDT [800]. It is not clear whether these choices are driven by differences in the detectors, or an optimization with respect to different criteria, samples, or modeling. While detailed optimization must ultimately be performed by the experiments themselves, we believe that there is still much to be understood about the general organization and design of jet substructure observables.

In this section, we perform a comprehensive study of performance and robustness for two-prong tagging techniques. To frame the study, we use a theoretical organization into different tagging strategies based on the idea of dichroic observables [801], which generalize ratio observables to allow hybrid combinations of groomed and ungroomed shapes. These contain as special instances all the familiar observables used by the experiments, as well as new observables, such as dichroic versions of the  $N_2$  and  $D_2$  observables. We therefore place the ATLAS and CMS strategies as specific examples of broader classes of theoretical approaches for tagging two-prong substructure, about which we can draw general conclusions regarding robustness and performance.

The goal of this study is to highlight the interplay between performance and robustness, and assess the choices made by the different LHC experiments. Here, “performance” refers to the tagging efficiency (for a given background rejection) in the absence of systematic uncertainties. This has been the primary way to assess jet substructure observables in the past, but it does not capture the full set of considerations needed to apply jet substructure techniques in practice. By contrast, “robustness” refers to modifications of the substructure observables as different physics aspects are added. In particular, we consider robustness to non-perturbative effects, namely hadronization and underlying event, robustness to detector effects, and robustness to pileup radiation. In this study, we refine the metric introduced in Refs. [801, 802] for quantifying robustness. We believe that this dual assessment of performance and robustness will be useful in future studies of jet substructure observables. This allows us to study each tagging strategy in general, and the CMS and ATLAS approaches in particular. In all cases, we are able to identify observables with improved robustness and performance as compared with those currently used by the experiments.

As an additional aspect of this study, we also consider the robustness of the signal tagging efficiency to the polarization of the decaying boson. We show that while jet shape observables themselves are fairly robust to polarization, groomed mass cuts are not, so that the tagging efficiency depends strongly on the polarization. Furthermore, we propose that the momentum asymmetry of the subjects is a good discriminant between longitudinal and transverse polarizations, and can be used to perform polarimetry for boosted hadronic decays.

An outline of this study is as follows. In Sec. 2.2, we define our metrics for studying robustness. We discuss the key physics issues that we would like to assess robustness to, both theoretical and experimental, and describe a chain of different steps in our simulation process such that each physics issue can be isolated and studied. In Sec. 2.3, we define all jet substructure

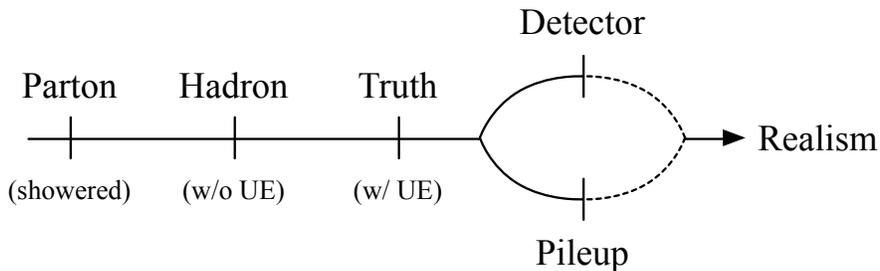


Fig. III.18: A summary of the different stages of physics considered in this study, from idealized parton-level events to fully realistic events including detector simulation and pileup. This allows us to address robustness to physics at each stage. Detailed definitions of each stage, and the physics probes used, are given in the text.

observables that will be used throughout this study, and provide details of our sample generation. In Sec. 2.4, we discuss theoretical approaches to designing robust two-prong taggers. We extend the approach of Ref. [801] and define several new dichroic observables formed from the energy correlation functions. In Sec. 2.5, we study the robustness of the observables to non-perturbative radiation both from hadronization and underlying event. In Sec. 2.6, we study robustness to detector and pileup. In Sec. 2.7, we study the robustness of substructure observables to the polarization of a decaying  $W/Z$  boson, and introduce observables to distinguish polarized samples. We summarize our results in Sec. 2.8 and make a number of recommendations for future jet substructure studies.

## 2.2 Quantifying Performance and Robustness

The goal of this study is to study the interplay between tagging performance and robustness for two-prong taggers. This requires a precise definition of the physics effects to which we are (or are not) robust, as well as a metric to quantify both performance and robustness.

Since we are able to generate pure signal and background samples, the tagging performance is straightforward to define using the signal and background efficiencies,  $\epsilon_S$  and  $\epsilon_B$ . In principle, we could evaluate the full receiver operating characteristic (ROC) curve relating  $\epsilon_B$  to  $\epsilon_S$ . For simplicity, we use the “tagging significance” as our metric for performance,

$$\epsilon = \frac{\epsilon_S}{\sqrt{\epsilon_B}}, \quad (\text{III.14})$$

evaluated at a fixed signal efficiency, typically  $\epsilon_S = 0.4$ .

We will approach robustness by moving from an idealized partonic description to a complete detector simulation including pileup; a “realistic” scenario representative of the LHC environment. This chain of realism is shown in Fig. III.18, which illustrates the following stages:

- **Parton Level:** We define the parton level result as the perturbative distribution for the active-active scattering (i.e. we do not include possible perturbative contributions to the underlying event). While this can be defined in an analytic calculation, it is more difficult in the context of parton shower generators, since there is necessarily a cut-off that must be imposed between perturbative and non-perturbative physics. Only the complete result is physical, and intermediate results should be interpreted with care. Nevertheless, to have some measure of the impact of non-perturbative effects, we will define parton level as generated by a parton shower generator with all hadronization effects turned off. We use

PYTHIA 8 [629,633] as our baseline generator, leaving a study of additional generators to future work.

- **Hadron Level:** We define the hadron level result as including hadronization in the shower, but not including any effects from the underlying event.
- **Truth Level:** We define truth level as the hadronized result including the underlying event as implemented in an event generator. Truth level therefore represents a complete hard scattering process in a hadron-hadron collider in isolation.
- **Detector Level:** Detector-level results are defined as truth level events passed through a detector simulation as implemented by TowerGrid. See Sec. 2.6.1 for details of the detector simulation.
- **Pileup Level:** Due to the high pileup environment of the LHC, we include in our study also the effects of uncorrelated proton interactions. We have done this separately from detector effects to be able to isolate and study the physics effects arising from pileup and pileup subtraction schemes. Our pileup subtraction scheme is described in Sec. 2.6.2.
- **Full Realism:** In the final stage of realism, we should consider events with pileup at detector level. These represent, to the level that we can consider in this study, realistic events as seen by the experiments at the LHC. Since our detector and pileup assessment is still fairly basic, we have left this final stage for future work.

Comparing the differences as we progress step by step through this sequence allows us to address at each stage the robustness to distinct physics issues, and we hope that our segmentation is sufficiently fine that we have a comprehensive view of robustness. In particular, the different steps in the chain allow us to study robustness to the following physics:

- **Parton → Hadron:** Changes in the distribution from parton level to hadron level probe non-perturbative physics associated with hadronization. For many event shapes, hadronization corrections can be given a field-theoretic definition in terms of a matrix element whose symmetry properties can be used to prove basic results. Ultimately, such corrections cannot be computed from first principles and must be included through models, such as those included in parton shower generators, or through dispersive approaches or shape functions in analytic calculations [663–668, 803]. To have the best theoretical control and understanding of jet substructure observables, it is therefore desirable that their performance is robust to the effects of hadronization.
- **Hadron → Truth:** Changes in the distribution from hadron level to truth level probe the impact of the underlying event, namely the physics associated with interactions of the colliding protons and their remnants. Such contributions are in principle both perturbative and non-perturbative. They are poorly understood theoretically, and it is currently not known how to treat them systematically, or define them field theoretically. It has been found empirically that the effects of underlying event are well modeled by a shape function [804], although the theoretical justification for this is not clear. Other simple theoretical models from which intuition can be gained have been proposed in Ref. [805]. The underlying event is implemented in parton shower event generators using models which are tuned to data, and we take these models as our definition of the underlying event. Due to this lack of theoretical understanding, as well as the fact that radiation from the underlying event is typically not associated with the physics that we are interested in probing, it is desirable that jet substructure observables be robust to the underlying event.

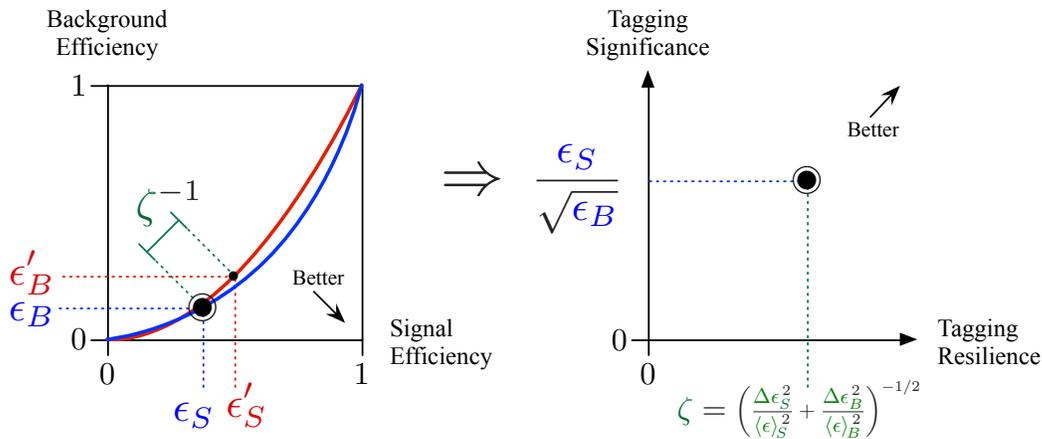


Fig. III.19: An illustration of the resilience metric  $\zeta$  used throughout the text to quantify robustness. In the left panel, we illustrate graphically  $\zeta$  as the change in ROC curve to a particular aspect of the underlying physics. In the right panel, we illustrate the tagging resilience vs. tagging performance plane which we will use to graphically illustrate our results. A simultaneously robust and performant tagger lives in the upper right hand corner of this space

- **Truth**  $\rightarrow$  **Detector**: Since we are ultimately interested in using jet substructure in experiments, the behavior of the detectors plays an essential role. The finite energy and spatial resolution of the detectors ultimately degrades the behavior of the observables. Furthermore, the detector response must be unfolded, and is therefore difficult to compute analytically, or to include to higher accuracy. Therefore, both for performance and calculability, it is desirable that jet substructure observables are robust to detector effects.
- **Truth**  $\rightarrow$  **Pileup**: Finally, due to the high pileup environment of the LHC, significant soft radiation can contaminate jet substructure observables. Since this radiation is not correlated with the underlying hard scattering process, it is not associated with the physics of interest, and therefore can only act to degrade the performance. Furthermore, it is difficult to model in an analytic calculation. While techniques exist to mitigate pileup, as reviewed in Sec. 2.6.2, it is desirable that the substructure observables used are as robust as possible to pileup contamination.

We will classify the first two of these as “Theory” issues, which will be discussed in Sec. 2.5, while the second two are classified as “Experimental” issues, and will be discussed in Sec. 2.6. This decomposition is of course somewhat arbitrary, since a coherent understanding involving the complete chain is required. However, this decomposition was chosen such that the “Theory” issues cover an idealized hadronic collision in isolation.

To compare the robustness to a particular step in this chain for different observables, we must introduce a metric. There is of course a high degree of arbitrariness in the definition of the metric. For example, one could base the metric on the shape of the signal or background distribution. Since we are ultimately interested in the performance of the observable, however, we introduce a metric which is based on the modification of the ROC curve. Consider a reference stage (unprimed) and a modified stage (primed); in the case of hadronization, the reference stage would be the hadron level and the modified stage would be the parton level. We first calculate the cut on the jet shape  $v_{\text{cut}}$  that yields a fixed signal efficiency  $\epsilon_S$  for the reference stage. We then compute the reference stage background efficiency  $\epsilon_B$  from that  $v_{\text{cut}}$ , as well as the modified

stage signal and background efficiencies  $\epsilon'_S$  and  $\epsilon'_B$ . We can then define a measure of robustness, which we refer to as resilience,  $\zeta$ , as

$$\zeta = \left( \frac{\Delta\epsilon_S^2}{\langle\epsilon\rangle_S^2} + \frac{\Delta\epsilon_B^2}{\langle\epsilon\rangle_B^2} \right)^{-1/2}, \quad (\text{III.15})$$

where

$$\Delta\epsilon_{S,B} = \epsilon_{S,B} - \epsilon'_{S,B}, \quad (\text{III.16})$$

$$\langle\epsilon\rangle_{S,B} = \frac{1}{2} (\epsilon_{S,B} + \epsilon'_{S,B}). \quad (\text{III.17})$$

This approach gives an estimate of how much our signal and background efficiencies have changed, for a given set of cuts, when going from one stage to another.

The meaning of  $\zeta$  is illustrated in Fig. III.19, where larger values of  $\zeta$  correspond to improved robustness. Because we anchor to a fixed  $v_{\text{cut}}$ , this method can even detect a uniform shift in both the signal and background distributions (even though such a shift would not change the ROC curve). We therefore believe that  $\zeta$  provides a reliable metric for assessing the robustness of the tagger. When presenting our results, we characterize observables in the plane of  $\epsilon$  (see Eq. (III.14)) versus  $\zeta$ , with better observables being in the upper-right corner. We find that this allows us to synthesize the information about a large number of observables in a compact manner. A similar metric and presentation style was used in Ref. [801, 802] to study robustness to non-perturbative effects. In addition to showing plots of the  $\epsilon$ - $\zeta$  plane, we will occasionally also show the modification of the distribution itself to provide additional insight into the robustness of the observables.

It is important to emphasize that it is impossible to completely characterize an optimal observable, particularly as jet substructure observables are being used for increasingly specific purposes. We hope that the issues that we have chosen to focus on are representative of the issues that will be important for a broad range of applications. Other aspects, such as the robustness of the substructure observable distributions to changes in the jet mass or  $p_T$  cuts, which are important for certain recent applications of jet substructure [806–811], and have received recent interest [812–814], are beyond the scope of the current project. However, it would be interesting to investigate them using similar techniques.

## 2.3 Observable and Sample Definitions

In this subsection, we define all the observables that will be studied throughout this study. This includes both the jet substructure shape observables and the grooming procedures. We also present the details of our sample generation.

### 2.3.1 Jet Shape Observables

The jet shape observables that we will consider are formed from ratios of the energy correlation functions [653, 796] or the  $N$ -subjettiness observables [798, 799]. The  $N$ -subjettiness observables are defined as [798, 799, 815]<sup>16</sup>

$$\tau_N^{(\beta)} = \sum_{1 \leq i \leq n_J} p_{T_i} \min \{ R_{i1}^\beta, \dots, R_{iN}^\beta \}. \quad (\text{III.18})$$

Here  $p_{T_i}$  is the  $p_T$  of particle  $i$ , and the sum is over all particles in the jet. The minimum is over the longitudinally-boost-invariant angle

$$R_{i,J}^2 = (\phi_i - \phi_J)^2 + (y_i - y_J)^2, \quad (\text{III.19})$$

<sup>16</sup>For this study, we have used the un-normalized definition of  $N$ -subjettiness.

between the particle  $i$  and the axis  $J$ .

Implicit in the definition of the  $N$ -subjettiness observable in Eq. (III.18) is the definition of the axes  $n_i$ . While their placement is unambiguous (up to power corrections) in the limit of resolved substructure, an algorithmic definition is required to determine their behavior in the unresolved limit. Two main approaches have been used for defining the axes. The first approach is to define the  $N$ -subjettiness axes as the axes found using an exclusive jet clustering algorithm. The second approach is to minimize the sum in Eq. (III.18) over possible light-like axes  $n_i$ . In this study, we defined the axes using subsets obtained by reclustering the jet, with choices motivated by the studies in Refs. [781, 816]. For  $\beta = 1$ , we recluster using the  $k_t$  algorithm [817] with the winner-take-all recombination scheme [818]. For  $\beta = 2$ , we use the generalized  $k_t$  algorithm [819] with  $p = 1/2$ .

For two-prong tagging, the relevant observable is the ratio [798]

$$\tau_{21}^{(\beta)} \equiv \frac{\tau_2^{(\beta)}}{\tau_1^{(\beta)}}. \quad (\text{III.20})$$

For a jet with a well resolved two-prong structure, we have  $\tau_{21}^{(\beta)} \ll 1$ , while for a jet without a well resolved substructure, we have  $\tau_{21}^{(\beta)} \sim 1$ . This observable has been extensively used at the LHC. It has been calculated to LL accuracy [781], and the effects of the axis definition on the perturbative behavior have been studied at NLO [820].

The second class of observables that we will consider are based on the energy correlation functions [653, 796]. Instead of correlating particles with axes, as is done for  $N$ -subjettiness, the idea of the energy correlation functions is to correlate  $n$ -tuples of particles. In discussing the energy correlation functions, it is convenient to work with dimensionless observables, written in terms of the angular variable,  $R_{ij}$  and the energy fraction variable  $z_i$ :

$$z_i \equiv \frac{p_{Ti}}{\sum_{j \in \text{jet}} p_{Tj}}, \quad (\text{III.21})$$

where  $p_{Ti}$  is the transverse momentum of particle  $i = 1, \dots, n_J$ . The generalized energy correlation function is defined as

$${}_v e_n^{(\beta)} = \sum_{1 \leq i_1 < i_2 < \dots < i_n \leq n_J} z_{i_1} z_{i_2} \dots z_{i_n} \prod_{m=1}^v \min_{s < t \in \{i_1, i_2, \dots, i_n\}}^{(m)} \{R_{st}^\beta\}, \quad (\text{III.22})$$

where  $\min^{(m)}$  denotes the  $m$ -th smallest element in the list. For a jet consisting of fewer than  $n$  particles,  ${}_v e_n$  is defined to be zero. More explicitly, the three arguments of the generalized energy correlation functions are as follows:

- The subscript  $n$ , which appears to the right of the observable, denotes the number of particles to be correlated.
- The subscript  $v$ , which appears to the left of the observable, denotes the number of pairwise angles entering the product. By definition, we take  $v \leq \binom{n}{2}$ , and the minimum then isolates the product of the  $v$  smallest pairwise angles.
- The angular exponent  $\beta > 0$  can be used to adjust the weighting of the pairwise angles as for  $N$ -subjettiness.

In this study, we use the 2-point energy correlation function,

$${}_1 e_2^{(\beta)} \equiv e_2^{(\beta)} = \sum_{1 \leq i < j \leq n_J} z_i z_j R_{ij}^\beta, \quad (\text{III.23})$$

as well as the 3-point correlators,

$$\begin{aligned}
{}_1e_3^{(\beta)} &= \sum_{1 \leq i < j < k \leq n_J} z_i z_j z_k \min \left\{ R_{ij}^\beta, R_{ik}^\beta, R_{jk}^\beta \right\}, \\
{}_2e_3^{(\beta)} &= \sum_{1 \leq i < j < k \leq n_J} z_i z_j z_k \min \left\{ R_{ij}^\beta R_{ik}^\beta, R_{ij}^\beta R_{jk}^\beta, R_{ik}^\beta R_{jk}^\beta \right\}, \\
e_3^{(\beta)} \equiv {}_3e_3^{(\beta)} &= \sum_{1 \leq i < j < k \leq n_J} z_i z_j z_k R_{ij}^\beta R_{ik}^\beta R_{jk}^\beta.
\end{aligned} \tag{III.24}$$

A number of 2-prong discriminants have been formed from the energy correlation functions [653, 787, 788, 796]. Here, we will focus on the observables

$$M_2^{(\beta)} = \frac{{}_1e_3^{(\beta)}}{e_2^{(\beta)}}, \quad N_2^{(\beta)} = \frac{{}_2e_3^{(\beta)}}{(e_2^{(\beta)})^2}, \quad D_2^{(\beta)} = \frac{e_3^{(\beta)}}{(e_2^{(\beta)})^3}, \tag{III.25}$$

each of which probes the correlations between particles within the jet in a slightly different manner. For a detailed discussion, see Ref. [796]. The  $N_2$  and  $D_2$  observables are powerful discriminants and have been used by CMS and ATLAS, respectively. The  $M_2$  observable is expected to have worse performance, except in particular scenarios, but we include it since it provides an example of a remarkably robust observable.

Beyond their discrimination power, these observables have nice theoretical properties. First, since they can be written as a sum over particles in the jet without reference to external axes, they are automatically “recoil-free” [652, 653, 818, 821, 822]. Second, since they have well-defined behavior in various soft and collinear limits, they are amenable to resummed calculations; in Ref. [780],  $D_2$  was calculated to next-to-leading-logarithmic (NLL) accuracy in  $e^+e^-$  for both signal (boosted  $Z$ ) and background (QCD) jets, and this was extended in Refs. [784, 785] to a hadron-collider environment by exploiting the simplifying properties of grooming.

### 2.3.2 Grooming Techniques

Groomers, which remove wide-angle soft radiation and contamination from a jet, also play an important role in two-prong tagging. While a variety of different grooming approaches have been defined [640–644, 767], we will focus on the mMDT/SoftDrop family, which is the most theoretically well understood, as well as trimming [643], which is used by the ATLAS experiment. In this subsection, we review the definition of the mMDT/SoftDrop and trimming algorithms and their parameters.

Starting from a jet identified with an IRC safe jet algorithm (such as anti- $k_t$  [661]), the SoftDrop algorithm proceeds as follows:

1. Recluster the jet using the Cambridge/Aachen (C/A) clustering algorithm [658–660], producing an angular-ordered branching history for the jet.
2. Step through the branching history of the reclustered jet. At each step, check the SoftDrop condition

$$\frac{\min [p_{Ti}, p_{Tj}]}{p_{Ti} + p_{Tj}} > z_{\text{cut}} \left( \frac{R_{ij}}{R} \right)^\beta. \tag{III.26}$$

Here,  $z_{\text{cut}}$  is a parameter defining the scale below which soft radiation is removed. If the SoftDrop condition is not satisfied, then the softer of the two branches is removed from the jet. This process is then iterated on the harder branch.

3. The procedure terminates once the SoftDrop condition is satisfied.

SoftDrop generalises the mMDT procedure [644] and the two are equivalent for  $\beta = 0$ . For this reason, we will often use the two names interchangeably. Any IRC-safe observable can be measured on a jet groomed with the SoftDrop procedure, without losing IRC safety (for  $\beta > 0$ ). The aggressivity of the SoftDrop grooming can be adjusted by tuning the parameters  $z_{\text{cut}}$  and  $\beta$ . Larger values of  $z_{\text{cut}}$  groom away more radiation within the jet for a fixed value of  $\beta$ . On the other hand, as  $\beta$  is increased, the grooming becomes less severe. Typical values of  $z_{\text{cut}}$  are around 0.1, while typical values of  $\beta$  are between 0 and 2.

Associated with the SoftDrop algorithm, in Sec. 2.7 we will study the observable  $z_g$  as a means for performing polarimetry. The  $z_g$  observable is also referred to as the groomed momentum fraction, and is defined as

$$z_g = \frac{\min [p_{Ti}, p_{Tj}]}{p_{Ti} + p_{Tj}} \quad (\text{III.27})$$

for the first declustering that satisfies the SoftDrop criteria. This observable probes the momentum sharing between the two prongs in the jet. For  $\beta \geq 0$ , this observable is Sudakov safe [823, 824] on QCD jets.

In addition to mMDT/SoftDrop, we also consider the trimming algorithm, since it is used by the ATLAS collaboration. Starting from a jet of radius  $R$  identified with an IRC-safe jet algorithm, trimming is defined by the following procedure:

1. Recluster the jet into subjets of radius  $R_{\text{sub}}$ .
2. Eliminate from the jet all particles in subjets that satisfy  $p_{T,\text{subjet}} > z_{\text{cut}} p_{T,\text{jet}}$ .
3. The trimmed jet is then defined to consist of the remaining particles.

Trimming has been experimentally shown to be a powerful grooming algorithm, and it exhibits excellent mass resolution. That said, trimming is known to suffer from non-global logarithms [662], and does not have a smooth spectrum as a function of the trimmed mass. The trimmed mass was analytically calculated in Ref. [644]. The trimming parameters used by ATLAS are  $R_{\text{sub}} = 0.2$ ,  $z_{\text{cut}} = 0.05$ , and the  $k_T$  algorithm is used to perform the reclustering.

### 2.3.3 Parton Shower Samples

For our QCD background jet samples we generated  $pp \rightarrow$  dijets in PYTHIA 8.230 [629, 633] with the Monash13 tune [631]. An explicit comparison to the HERWIG generator [584, 825] is left for future work. Samples were generated with hadronization off (parton level), with hadronization on but underlying event off (hadron level), and with hadronization and underlying event on (truth level). Pileup was included by superimposing uncorrelated minimum bias events, which are also generated in PYTHIA 8, this time using the 4C tune. Details of our pileup removal strategies will be described in Sec. 2.6.2.

For our polarized  $W$  samples, we considered a  $gg$  produced resonance,  $X$ , that decays to a pair of polarized  $W$  bosons. This kind of resonance decaying to longitudinally-polarized  $W$ s appear in warped extra-dimensional models, where the Standard Model fields propagate in the bulk. On the other hand, models with graviton-like tensors with minimal couplings yield only transversely-polarized  $W$  bosons. They were produced with the JHUGEN 3.1.8 [826, 827] generator, interfaced with PYTHIA 8 [633] for parton showering including the effect of hard gluon radiation. A resonance width of 1% was chosen. Table III.1 shows the coupling values used to generate the polarised  $W$  samples (see also Ref. [826] for more information).

Model	Production couplings	Decay couplings	Decay helicity amplitudes
$2_b^+$	$g_1 = 1$	$g_5 = 1$	$f_{00} = 0.98$
$2_m^+$	$g_1 = 1$	$g_1 = g_5 = 1$	$f_{00} = 0.08, f_{+-} = f_{-+} = 0.46$

Table III.1: A description of the production and decay constants used to produced the polarized  $W$  samples used in this studies.

## 2.4 Theory Approaches to Robust Ratio Observables

The observables defined in Sec. 2.3.1 are designed to distinguish boosted bosons from QCD jets using the detailed structure of the radiation within the jets. From this perspective, it is then immediately clear that there will be an interplay between performance and robustness. By using a maximal amount of radiation within the jet, the maximal information is available to identify the origin of the jet. However, an increased sensitivity to radiation also means an increased sensitivity to contamination, both theoretically from non-perturbative hadronization effects and underlying event, as well as experimentally from pileup. It also introduces sensitivity to the experimental reconstruction of soft momenta in the event. On the other hand, a tight grooming procedure, which removes radiation from within the jet, is expected to reduce the ability to distinguish boosted bosons from QCD jets. In that context, it is interesting to see that both ATLAS and CMS have currently adopted a tight grooming strategy.

In this subsection, we introduce a classification of different tagging strategies which will be an organizing framework for our study of robustness and performance. This organization will be based on the idea of dichroic observables [801], which are ratio observables constructed from combinations of groomed and ungroomed observables. In Sec. 2.4.1, we review the physics of the dichroic approach for the example of the  $\tau_{21}$  observable. In Sec. 2.4.2, we generalize the dichroic construction to energy correlation function based observables and give definitions of dichroic  $N_2$ ,  $D_2$ , and  $M_2$  observables. This allows us to have multiple concrete examples of observables for each of our general tagging strategies. Then, in Sec. 2.4.3 we present the complete set of tagging strategies that we will consider in this work, including the different parameters we will scan within each general class.

### 2.4.1 Review of the Dichroic Approach

In Ref. [801], it was proposed that in addition to considering the standard  $\tau_{21}$  observable and its groomed counterpart, one should also consider a mixed version, with a groomed denominator, and an ungroomed numerator. This observable was termed “dichroic”, since different levels of grooming are sensitive to different color structures. More explicitly, Ref. [801] considered the three observables:

$$\tau_{21}^{\text{large}} = \frac{\tau_2^{\text{large}}}{\tau_1^{\text{large}}}, \quad \tau_{21}^{\text{small}} = \frac{\tau_2^{\text{small}}}{\tau_1^{\text{small}}}, \quad \tau_{21}^{\text{dichroic}} = \frac{\tau_2^{\text{large}}}{\tau_1^{\text{small}}}, \quad (\text{III.28})$$

where “large” refers either to the plain jet of a lightly/loosely groomed jet and “small” denotes a more aggressively/tightly groomed jet. It was then argued that the dichroic combination was optimal for performance, while also increasing the robustness to non-perturbative contamination. Here we briefly summarize the motivation for the dichroic observable. Readers interested in a more detailed discussion of the dichroic approach, including perturbative calculations, are referred to Ref. [801].

The benefits of the dichroic approach can simply be illustrated by considering the distributions of the observables after aggressive grooming (tight), moderate grooming (loose), and in the dichroic approach. This is shown in Fig. III.20 for the particular case of  $\tau_{21}$ . Considerable

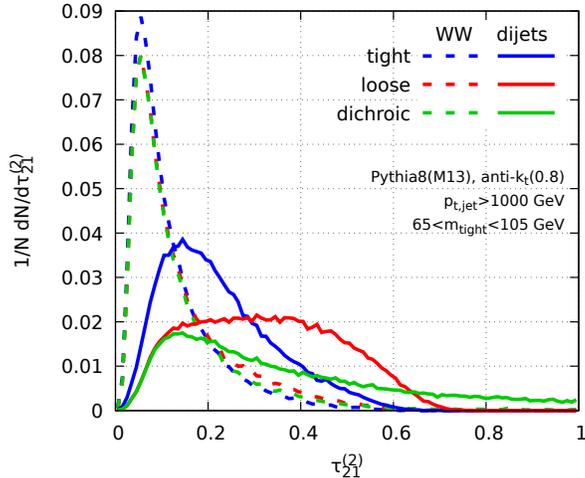


Fig. III.20: Distributions of the tight, loose, and dichroic  $\tau_{21}$  observables with a cut on the mMDT/SoftDropped mass. The dichroic approach offers considerably improved performance as compared with the tight grooming, almost preserving the performance of the loose grooming at high signal purity. Figure taken from Ref. [801].

performance is lost in going from loose to tight grooming, since the background distribution is pushed to lower values of  $\tau_{21}$ . For the dichroic observable, by contrast, the behavior of the loose and dichroic observables is identical at small values of  $\tau_{21}$ , but for larger values of  $\tau_{21}$ , the dichroic distribution is pushed to yet larger values than the loose distribution, leading to improved performance. Since the dichroic ratio observable uses a partially-groomed observable, it is expected to be less sensitive to non-perturbative effects due to hadronization. It therefore represents an interesting new class of observables to consider.

#### 2.4.2 New Dichroic Observables

It is almost straightforward to extend the dichroic definition from  $N$ -subjettiness ratios to observables formed from the energy correlation functions. For  $\tau_{21}$  it is immediately clear what the numerator and denominator of the observable are, however, this is initially less obvious for the energy-correlation-function-based observables that have a more complicated structure. Here, we define the dichroic variants of the  $M_2$ ,  $N_2$ , and  $D_2$  by isolating a single factor of a mass like observable ( $e_2$ ) as the denominator, and defining the remainder of the observable as the numerator. The definitions of the dichroic variants of the  $M_2$ ,  $N_2$ , and  $D_2$  observables are then

$$M_2^{\text{dichroic}} = \frac{(1e_3)^{\text{large}}}{e_2^{\text{small}}}, \quad (\text{III.29})$$

$$N_2^{\text{dichroic}} = \frac{(2e_3/e_2)^{\text{large}}}{e_2^{\text{small}}}, \quad (\text{III.30})$$

$$D_2^{\text{dichroic}} = \frac{(e_3/e_2^2)^{\text{large}}}{e_2^{\text{small}}}. \quad (\text{III.31})$$

The above prescription is most easy to see for the  $N_2$  observable. In the two-prong limit, the combination  $2e_3/e_2$  reduces to  $\tau_2$ , and therefore the dichroic  $N_2$  ratio behaves similarly to the dichroic  $\tau_{21}$  ratio.

Observable	Numerator	Denominator
$M_2$	$1e_3$	$e_2$
$N_2$	$2e_3/e_2$	$e_2$
$D_2$	$e_3/e_2^2$	$e_2$
$\tau_{21}$	$\tau_2$	$\tau_1$

Table III.2: Definitions of the numerators and denominators for the different jet substructure observables.

Notation: $m \otimes \frac{n}{d}$	$m$ (mass)	$n$ (numerator)	$d$ (denominator)
$p \otimes \frac{p}{p}$	plain	plain	plain
$\ell \otimes \frac{p}{p}$	loose	plain	plain
$\ell \otimes \frac{p}{\ell}$	loose	plain	loose
$\ell \otimes \frac{\ell}{\ell}$	loose	loose	loose
$t \otimes \frac{p}{p}$	tight	plain	plain
$t \otimes \frac{p}{\ell}$	tight	plain	loose
$t \otimes \frac{\ell}{\ell}$	tight	loose	loose
$t \otimes \frac{p}{t}$	tight	plain	tight
$t \otimes \frac{\ell}{t}$	tight	loose	tight
$t \otimes \frac{t}{t}$	tight	tight	tight
trim	trim	trim	trim

Table III.3: A summary of the different tagging strategies considered, including the notation, and the degree of grooming for the mass, and numerator and denominator of the shape observable. For simplicity, we have suppressed the jet radius,  $R$ . The definitions of plain, loose, tight and trim are given in the text.

### 2.4.3 Summary of Tagging Strategies

The tagging strategies we consider can be put into the general form of a (groomed) mass cut ( $m$ ) followed by a cut on a two-prong tagging observable, which takes the form

$$\mathcal{O} = \frac{\text{3-particle observable}}{\text{2-particle (mass) observable}} \equiv \frac{n}{d}. \quad (\text{III.32})$$

The explicit numerators ( $n$ ) and denominators ( $d$ ) for the different observables are summarized in Table III.2.

As our organizing principle for classes of jet substructure taggers, we use the type of grooming applied to the initial mass cut, and the type of grooming applied to the numerator and denominator of the two-prong observable. We use the notation

$$m \otimes \frac{n}{d} \quad (\text{III.33})$$

to denote the grooming applied to a particular observable, where  $m$ ,  $n$ , and  $d$  can take the values

- plain ( $p$ ): no grooming applied;
- loose ( $\ell$ ): SoftDrop with  $z_{\text{cut}} = 0.05$ ,  $\beta = 2$ ;
- tight ( $t$ ): mMDT with  $z_{\text{cut}} = 0.1$ .

Parameter	Values Scanned
Jet Radius $R$	0.8
Jet Shape	$D_2, N_2, M_2, \tau_{21}$
Jet Shape Angular Exponent $\beta$	1, 2
Jet $p_T$	500 GeV, 1000 GeV, 2000 GeV

Table III.4: Jet shapes and parameters scanned for each of the different strategies proposed in Table III.3.

- trim: trimming with  $R_{\text{sub}} = 0.2$ ,  $z_{\text{cut}} = 0.05$  and the  $k_T$  algorithm to perform the reclustering, as used by ATLAS.

In general, to have a good 2-prong tagging strategy, we want to have

$$m \geq d \geq n, \quad (\text{III.34})$$

where  $\geq$  refers to the aggressiveness of the groomer, with  $t > \ell > p$ .

The complete set of configurations that we consider is given in Table III.3. These constitute generic grooming strategies, and will be studied for each of  $N_2$ ,  $D_2$ , and  $\tau_{21}$ . They include the dichroic ratios, as well as the current ATLAS and CMS approaches as specific examples. This will allow us to draw general conclusions about the performance and robustness of jet substructure taggers.

While our general approach is based on studying the different classes of grooming strategies in Table III.3, for each of these different classes of strategies we will also scan different parameters. In particular, we will scan the jet shape, jet shape angular exponent, and jet  $p_T$ . The values scanned are summarized in Table III.4. This allows us to understand if the conclusions drawn are associated with specific observables within a given strategy, as well as to optimize over these parameters. Due to the large number of physics and detector parameters that can be varied in our study, only a subset of plots will be presented in the study. More specifically, we will primarily show results for the following 3 benchmark points

- an **ATLAS-like tagger** putting cuts on the trimmed mass and on  $D_2^{(1)}$  computed on the trimmed jet:  $D_2^{(1)}[\text{trim}]$ ;
- a **CMS-like tagger** putting cuts on the mMDT mass and on  $N_2^{(1)}$  computed on the tightly-groomed jet:  $N_2^{(1)}[t \otimes t/t]$ ;
- the **Les-Houches Dichroic Tagger (LHDT)** putting cuts on the mMDT mass and on the dichroic  $D_2^{(2)}$ :  $D_2^{(2)}[t \otimes l/t]$ .

To study variations, we will either fix the grooming strategy to one of the three benchmark points and vary the jet shape, or conversely, fix the type of jet shape to one of our benchmark points and vary the grooming strategy. A more global summary of the observables scanned, highlighting those which we find to be most promising, will be given in Fig. III.29.

## 2.5 Theory Robustness

In this subsection, we study robustness to hadronization and underlying event, which we categorize as “Theory” robustness. Robustness to the detector and to pileup are treated in Sec. 2.6.

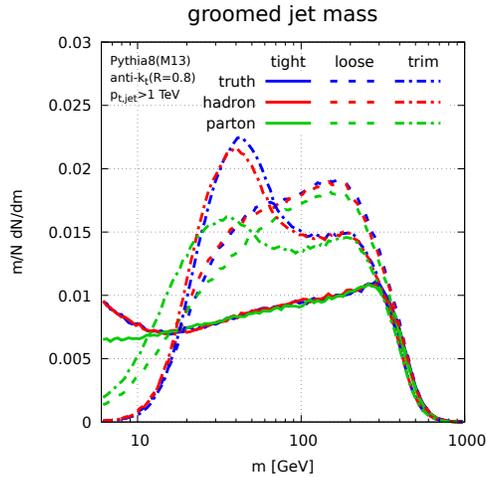


Fig. III.21: Mass distribution after grooming for the three groomers considered in this paper. The distributions are shown at parton level, at hadron level, and at truth level (i.e. including both hadronization and the underlying event).

### 2.5.1 Hadronization

We begin by studying the robustness of different tagging techniques to hadronization. As discussed in Sec. 2.2, this must be interpreted with some care, since the unhadronized events are not themselves physical. Nevertheless, the comparison of hadronized and unhadronized distributions is the best proxy for understand the impact of hadronization short of performing an analytic calculation. To ensure that our conclusions are robust, ideally we would consider parton shower generators which implement different hadronization models. For example, the PYTHIA shower uses the string model [828, 829], while HERWIG++ uses the cluster model [703, 830]. See for example Refs. [831–833] for a more detailed discussion. Due to the restricted scope of this report, here we only consider PYTHIA. The effect of hadronization on two-prong substructure observables has been studied in Refs. [780, 784, 785, 801].

Before studying robustness under hadronization quantitatively, we begin by showing several distributions with and without hadronization. This will help to introduce the different observables, as well as to give the reader a feeling for the robustness at the level of the shape of the distribution, and how this compares to our resilience measure.

In Fig. III.21, we show the jet-mass distribution for the three grooming strategies considered in this paper, for parton, hadron, and truth levels. Although we will not focus directly on the mass distribution in this paper, it plays an important role since all of our studies will be performed with jet mass cuts. Here we see two primary features. First, all three groomers give rise to significantly different mass distributions. This has been discussed in detail in Refs. [644, 645]. Second, with both tight and loose grooming, the distributions are robust to hadronization. This is particularly true for tight grooming where hadronization has almost no effect, except at extremely small values of the observable. On the other hand, the trimmed mass distribution is less robust to hadronization effects.

In Fig. III.22, we show distributions for our benchmark observables, namely  $D_2^{(1)}$ ,  $N_2^{(1)}$ , and a dichroic version of  $D_2^{(2)}$ , measured on both background and signal jets. In all cases, we see that hadronization has a sizable effect on the shape of the distribution, pushing it to larger values. For the  $D_2$  observable, hadronization is mostly isolated to small values of the observable, and at larger values reduces simply to a shift of the distribution. This has been discussed in detail for the case of  $D_2$  in Refs. [780, 784, 785]. For the  $N_2$  observable, hadronization effects are larger

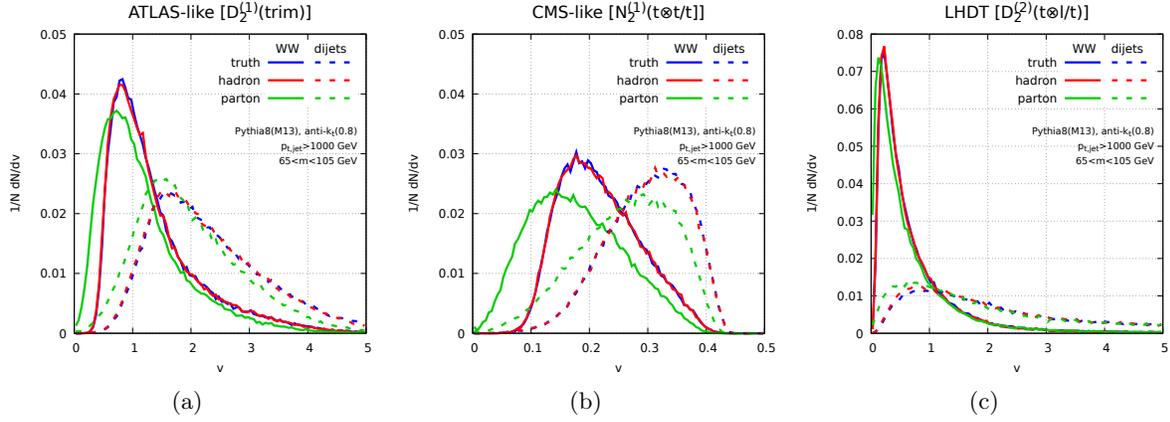


Fig. III.22: Distribution of three benchmark shapes:  $D_2^{(1)}$  computed on the trimmed jet (ATLAS-like),  $N_2^{(1)}$  computed on the tight (mMDT) jet (CMS-like), and the dichroic  $D_2^{(2)}$  with numerator computed on the loose jet and denominator computed on the tight jet (LHDT). The distributions are shown at parton level, at hadron level, and at “truth” level (i.e. including both hadronization and the underlying event), for both  $WW$  (solid) and dijet (dashed) events.

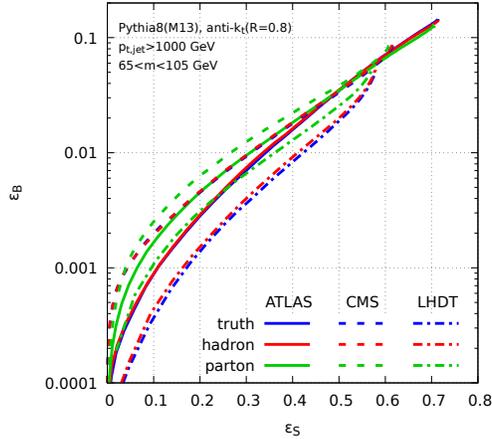


Fig. III.23: ROC curves corresponding to the shapes plotted on Fig. III.22. The three line types correspond to the three benchmark points.

and are significant throughout the entire distribution. Indeed, when we study performance and robustness quantitatively, we will find that while  $N_2$ -type observables tend to be more performant, they are also less resilient to hadronization effects.

Since the primary role of hadronization is to push the distributions to larger values at small values of the observable, the performance of the observables is typically highly sensitive to hadronization, particularly at high signal purity. In Fig. III.23 we illustrate this (lack of) robustness to hadronization at the level of the ROC curves for the different shape choices. In all cases, we see that hadronization considerably improves the performance of the observables. This is particularly true at high signal purity, and decreases as the signal purity is decreased. Since the region of high signal purity is typically that of interest for jet substructure studies, this emphasizes the importance of understanding the robustness of observables to hadronization effects.

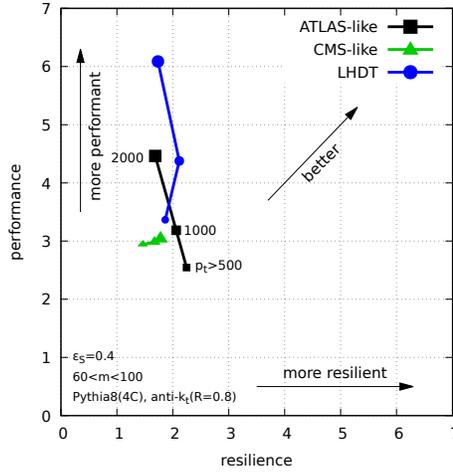


Fig. III.24: An illustration of the performance-resilience plane that will be used to illustrate our results. More performant observables lie to the top, more resilient observables to the right, and better observables in the upper-right corner.

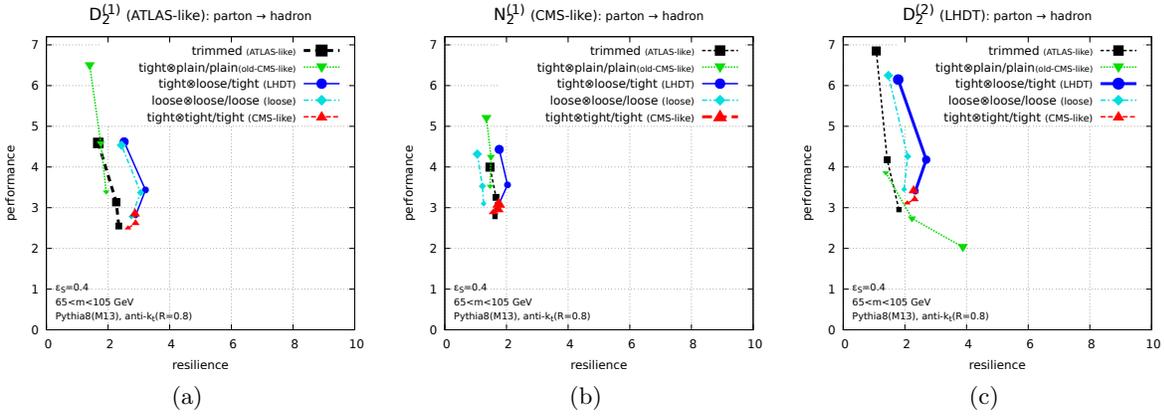


Fig. III.25: Plots of the performance-resilience plane under the addition of hadronization effects for the standard jet shape observables with different grooming strategies.

Having given a feel for the modifications due to hadronization at the level of both distributions and ROC curves, we now use our performance and resilience measures to perform a quantitative study. Since our visualization method allows a considerable amount of information to be condensed into a single plot, we first briefly review our presentation method with a sample plot. In Fig. III.24, we show a plot in the performance-resilience plane, in which we will display our results. More performant observables appear higher on the  $y$ -axis (towards the top), while more robust (resilient) observables appear higher on the  $x$ -axis (to the right), as indicated by the arrows. A performant and resilient observable will appear in the upper right corner. For each observable, we also perform a scan of  $p_T$ , from 500 – 2000 GeV, which are illustrated by the three connected points. This will be the default format in which we display our results throughout the rest of this study.

In Figs. III.25 and III.26, we show the performance-resilience plots for the effects of hadronization for our benchmark observables. Figure III.25 shows  $D_2$ ,  $N_2$ , and dichroic  $D_2$  for the different grooming strategies, while in Fig. III.26 we consider fixed grooming strategies

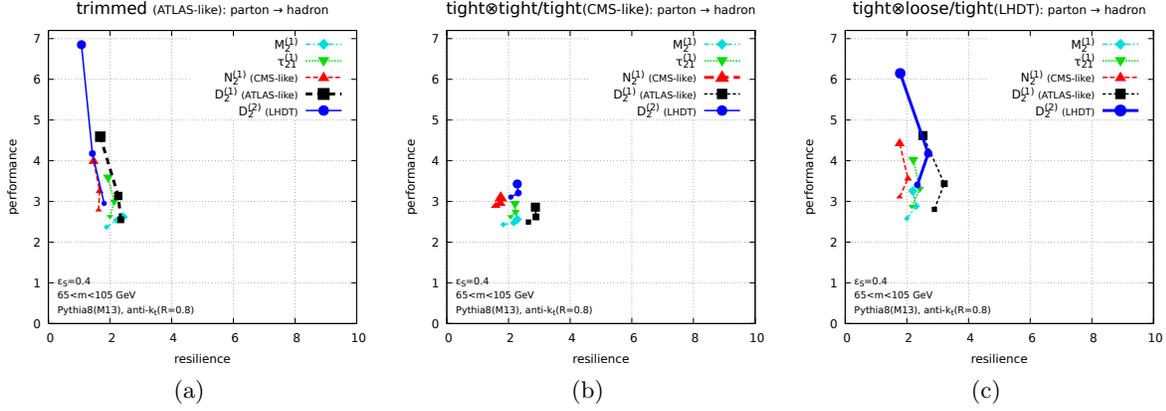


Fig. III.26: Plots of the performance-resilience plane under the addition of hadronization effects for different jet shape observables, with fixed grooming strategies.

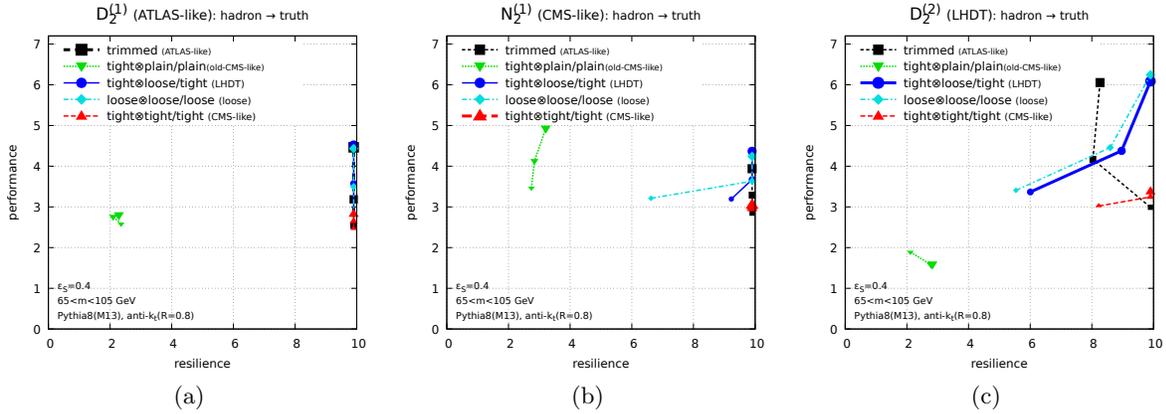


Fig. III.27: Plots of the performance-resilience plane going from hadron level to truth level (inclusion of underlying event) for the standard jet shape observables with different grooming strategies. In this and the next figure, resilience values have been cut at  $\zeta = 10$  for easier comparison with the corresponding hadronisation plots, Figs. III.25 and III.26.

in each plot, but different jet shape observables. We first notice that, in almost all cases, a dichroic form of the observable can be used to improve resilience while maintaining a similar level of performance. Among the shapes, we find that  $D_2$  tends to be the most robust and that  $N_2$  tends to be slightly more performant, although  $D_2$  becomes more performant than  $N_2$  at larger  $p_T$ . This agrees with what was seen by studying the distributions in Fig. III.22 by eye, however, we are now able to quantify this. We will see that this conclusion remains true under a larger scan of observables in Sec. 2.5.3. We also notice that for almost all the observables, the trends with  $p_T$  are similar.

### 2.5.2 Underlying Event

We can now repeat the same exercise performed for hadronization to study the robustness to underlying event. The results in the performance-resilience plane are shown in Figs. III.27 and III.28. These are identical to Figs. III.25 and III.26 but measure the robustness to underlying event instead of hadronization. The first thing that is clear from comparing Figs. III.25 and III.26 with Figs. III.27 and III.28 is that with modern grooming techniques, we are comparatively

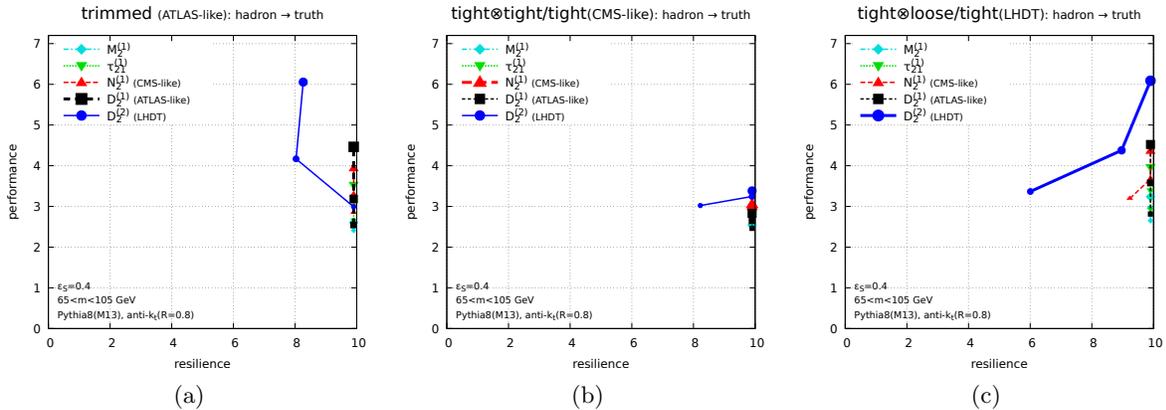


Fig. III.28: Plots of the performance-resilience plane going from hadron level to truth level (inclusion of underlying event) for different jet shape observables, with fixed grooming strategies.

much less sensitive to underlying event than to hadronization effects. With the exception of the  $\text{tight} \otimes \text{plain}$ /plain grooming strategy (which indeed does not groom the observable), all the standard observables are robust to underlying event for all the different grooming strategies. Similarly, all the different jet shape observables in this study are also robust to Underlying Event effects. The dichroid  $D_2^{(2)}$  observable (Figs. III.25c and III.27c) shows a smaller resilience against the Underlying Event although it remains much larger than the corresponding resilience to hadronisation effects. We believe that this should be viewed as a success of modern grooming tools. We also believe that it is desirable, since underlying event effects are much less under theoretical control than hadronization effects.

### 2.5.3 Towards Improved Performance and Robustness for ATLAS and CMS

The strategies used by ATLAS and CMS, namely trimmed  $D_2$  [780, 787] and SoftDropped  $N_2$  [796] with DDT [800] are specific examples of the broader approaches to two-prong tagging discussed above, and therefore our study allows us to gain insight into the different tagging strategies used by the collaborations,<sup>17</sup> as well as to suggest improved observables.

An overall scan of all the observables used in this study in the performance-resilience plane is shown in Fig. III.29. The current CMS and ATLAS observables are highlighted with the black and green markers, respectively. We can draw two interesting conclusions from this plot. First, for  $p_T > 500$  GeV, it appears as though ATLAS is using a more robust, but less performant observable, while CMS is using a more performant, but less robust observable. This must be caveated by the fact that we have not performed a full detector study, though it is suggestive. Second, we can choose from the observables considered in our study observables that are simultaneously more performant, and more resilient than those that are currently being used. At 500 GeV the gains in performance are fairly minimal, but at 1000 GeV, it seems that there are considerable gains in performance and robustness to be made. The names of a large variety of those observables which lie on the upper boundary of the performance-resilience space are marked in the figure.

Looking more closely at the observables along this upper boundary, we see that there is in fact considerable structure, and a number of general lessons can be learned. First, as we move along this boundary from least resilient to most resilient, we transition from  $N_2^{(1,2)}$  observables

<sup>17</sup>And perhaps even into the sociology of the different experiments! However, such conclusions should be taken with a grain of salt.

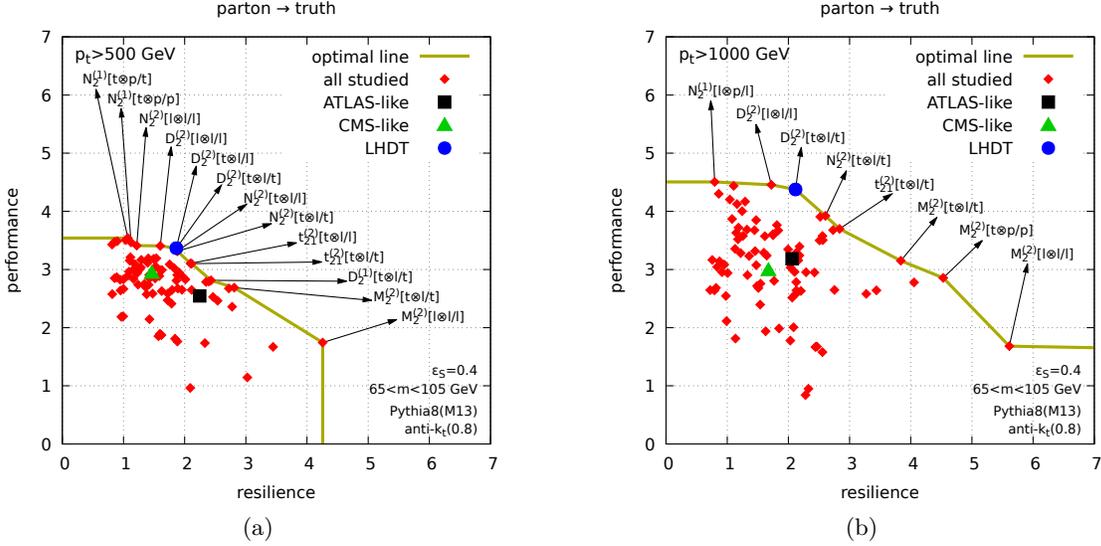


Fig. III.29: The performance-resilience plane for the different observables scanned in our study at (a) 500 GeV and (b) 1000 GeV. The ATLAS- and CMS-like observables are marked in black and green, respectively. In both cases, more robust, and more performant observables can be selected, and a number of such observables are marked.

which are the most performant, but less robust, through  $D_2^{(2)}$ , to  $M_2^{(2)}$ , which is more robust, but less performant. This was also clearly observed in the distributions of Fig. III.22. We believe that this is due to the fact that  $N_2$  has a hard phase space boundary, and therefore non-perturbative effects are not isolated at small values of the observable, although it would be interesting to understand this behavior in more detail.

A second pattern that is observed is that, in almost all cases, dichroic variants of the observables of the form  $t \otimes \frac{p}{t}$  or  $t \otimes \frac{\ell}{t}$  exhibit improved performance without significant loss in resilience. We believe that it is worthwhile for the experiments to consider some of the dichroic observables that were newly introduced in Sec. 2.4.2 with simultaneous performance and resilience in mind. We have highlighted in Fig. III.29 that a whole interesting phase space of such observables exist, which map out the performance-resilience plane. Different observables could be chosen depending on the particular needs of a given study. Note finally, that the above study has been carried using a specific choices of grooming strategies (loose and tight). There is therefore a potential additional gain that can be achieved by studying alternative (more or less aggressive) options.

## 2.6 Experimental Robustness

Having discussed robustness to theoretical issues, we now continue through our chain of realism of Fig. III.18 and consider robustness to detector effects. In Sec. 2.6.1 we describe our detector model. In Sec. 2.6.2 we describe pileup removal. In Sec. 2.6.3 we study the robustness of jet mass to detector effects. A more comprehensive study is left to a dedicated publication.

Signal	Model parameters			
	Coverage	$p_T^{\min}$	$p_T^{\max}$	Resolution function parameters
<b>Configuration A</b> ( $R_{\text{calo}} = 1150$ mm, $B_{\text{solenoid}} = 2$ T)				
Towers (EM)	$ \eta  < 2.5$	500 MeV	10 TeV	$a_{\text{calo}} = 10\% \times \sqrt{\text{GeV}}$ $b_{\text{calo}} = 0$ $c_{\text{calo}} = 0.7\%$
Towers (HAD)	$ \eta  < 4.9$	500 MeV	10 TeV	$a_{\text{calo}} = 50\% \times \sqrt{\text{GeV}}$ $b_{\text{calo}} = 0$ $c_{\text{calo}} = 3\%$
<b>Configuration C</b> ( $R_{\text{calo}} = 1290$ mm, $B_{\text{solenoid}} = 4$ T)				
Towers (EM)	$ \eta  < 2.5$	500 MeV	10 TeV	$a_{\text{calo}} = 3\% \times \sqrt{\text{GeV}}$ $b_{\text{calo}} = 0$ $c_{\text{calo}} = 0.5\%$
Towers (HAD)	$ \eta  < 4.9$	500 MeV	10 TeV	$a_{\text{calo}} = 100\% \times \sqrt{\text{GeV}}$ $b_{\text{calo}} = 0$ $c_{\text{calo}} = 5\%$
Tracks	$ \eta  < 2.5$	1 GeV	300 GeV	$a_{\text{track}} = 0.015\% \times \text{GeV}^{-1}$ $c_{\text{track}} = 0.5\%$

Table III.5: Principal properties and smearing parameters of the detector model configurations A and C. Electromagnetic (EM) towers have a tower size of  $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ , while hadronic (HAD) towers have  $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ . Tracks are not modeled in configuration A, while configuration C models both tracks and calorimeter signals to simulate the effect of a particle flow algorithm. The resolution function parameters  $a_{\text{calo}}$ ,  $b_{\text{calo}}$ , and  $c_{\text{calo}}$  are used together with the resolution function in Eq. (III.35) to determine the width of the Gaussian energy smearing. Similarly,  $a_{\text{track}}$  and  $c_{\text{track}}$  are used in Eq. (III.36) to determine the width of the track- $p_T$  resolution smearing.

### 2.6.1 Detector Models

The detector response is modeled by subjecting the particle-level objects<sup>18</sup> to a simple acceptance and smearing model representing tracking detectors as well as calorimeters. Basic detector feature descriptors are considered together with response features and signal reconstruction strategies similar to the ATLAS and CMS experiments at the LHC. The respective model configurations are A (ATLAS-like [834]) and C (CMS-like [835]). Effects from particles showering in calorimeters are not modeled.

Both configurations feature cylindrical detectors around the collision vertex with full azimuthal coverage  $-\pi < \phi < \pi$ . The detector modeling produces a final state represented by a list of *pseudoparticles* which are proxies for detector signals generated by stable particles. Particles emitted at pseudorapidities  $|\eta| > 4.9$  and non-detectable particles like neutrinos are excluded from this modeling and thus not part of this final state.

In configuration A, the pseudoparticles are produced by a calorimeter-only detector model with regular projective readout in  $(\eta, \phi)$  space. The energy of all stable particles generating

<sup>18</sup>Those are stable particles produced by the generator that potentially reach a detector. This principally requires a particle lifetime  $\tau$  in the laboratory frame given by  $c\tau > 10$  mm.

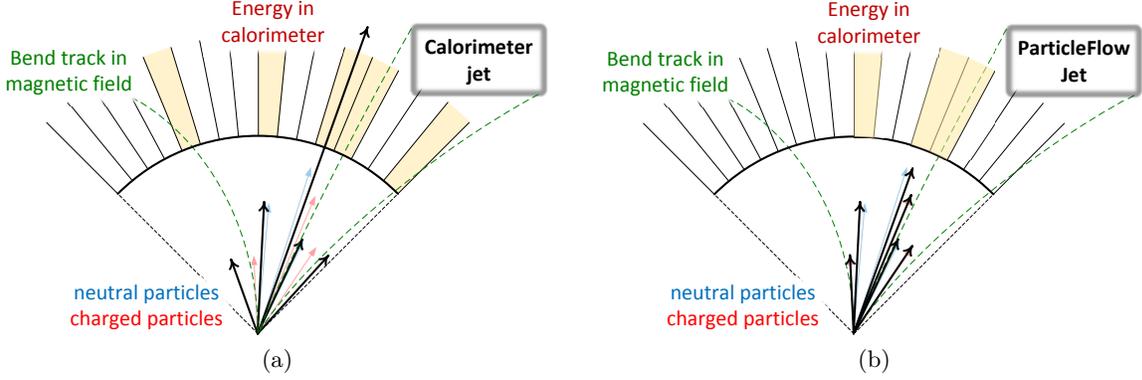


Fig. III.30: An illustration of the basic jet features for the configuration A (ATLAS-like, calorimeter only) and configuration C (CMS-like, particle flow using tracks and calorimeter) detector signal models. The solid black arrows indicate the jet composition from representations of neutral (pale blue) and charged (pale red) particles by calorimeter towers or reconstructed tracks. The green dashed curves show charged particle tracks bend in the magnetic field. Both illustrations show the same truth-level jet.

hadronic shower when interacting with the detector material is collected in hadronic (HAD) *calorimeter towers* of size  $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ . Electrons, positrons, and photons emitted within  $|\eta| < 2.5$  fill electromagnetic (EM) towers of size  $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$  with their energy, mimicking the typical coverage of the high granularity electromagnetic calorimeter. The direction of neutral particles is the generated direction at the vertex, while for all charged particles  $\phi$  is changed to the azimuth at the entry point of the particle into the calorimeter. This entry point is calculated using the particle trajectory in an axial uniform magnetic field of  $B_{\text{solenoid}} = 2$  T and the radius  $R_{\text{calo}} = 1150$  mm of the calorimeter front face. If the transverse momentum of the charged particle is too low to reach the calorimeter, i.e. its trajectory in the magnetic field does not exceed  $R_{\text{calo}}$ , the particle is considered invisible for the detector and thus ignored for further analysis. For particles reaching the calorimeter, the (bent) trajectory is not radial anymore at the front face. This suggests a distribution of the particle energy into more than one tower. In this model there is no energy sharing between towers, as this would require to at least model the longitudinal energy distribution in an electromagnetic or hadronic shower, with considerable computational effort.

After all energy is collected, the finite calorimeter resolution is modeled by smearing the tower energy  $E_{\text{tower}}$  following a Gaussian distribution with width  $\sigma_{E_{\text{tower}}}$  given by the canonical calorimeter resolution function:

$$\frac{\sigma_{E_{\text{tower}}}(E_{\text{tower}})}{E_{\text{tower}}} = \sqrt{\frac{a_{\text{calo}}^2}{E_{\text{tower}}} + \frac{b_{\text{calo}}^2}{E_{\text{tower}}^2} + c_{\text{calo}}^2}. \quad (\text{III.35})$$

The three components of this function are the stochastic term  $a_{\text{calo}}$  reflecting sampling and intrinsic shower fluctuations, the noise term  $b_{\text{calo}}$  quantifying the detector noise, and the constant term  $c_{\text{calo}}$  capturing fluctuations introduced in the process of the detector signal extraction. The values for  $a_{\text{calo}}$  and  $c_{\text{calo}}$  are shown in Table III.5. The detector noise is not modeled, thus the noise term is  $b_{\text{calo}} = 0$  for both configurations.

The tower energy after smearing is required to pass  $p_T^{\text{tower}} > p_T^{\text{min}}$ . Towers passing this requirement are converted into massless pseudoparticles using the nominal tower center  $(\eta_{\text{tower}}, \phi_{\text{tower}})$  and the tower energy  $E_{\text{tower}}$   $((E_{\text{tower}}, \eta_{\text{tower}}, \phi_{\text{tower}}) \mapsto (E_{\text{tower}}, \vec{p}_{\text{tower}})$  with

$|\vec{p}_{\text{tower}}| = E_{\text{tower}})$ .

Configuration C models a particle flow signal from a tracking detector combined with a calorimeter. Charged particles are bent in an axial uniform magnetic field with  $B_{\text{solenoid}} = 4 \text{ T}$ . The transverse momentum of the charged particles emitted within a tracking detector acceptance of  $|\eta| < 2.5$  is smeared along a Gaussian distribution function with width  $\sigma_{p_T^{\text{track}}}$  given by:

$$\frac{\sigma_{p_T^{\text{track}}}}{p_T^{\text{track}}} = \sqrt{(a_{\text{track}} \cdot p_T^{\text{track}})^2 + c_{\text{track}}^2}. \quad (\text{III.36})$$

If  $p_T^{\text{track}}$  after smearing is within  $p_T^{\text{min}} < p_T^{\text{track}} < p_T^{\text{max}}$ , the charged particle is added to the list of pseudoparticles with its direction at the interaction vertex. The trajectories of all charged particles within  $|\eta| < 2.5$  and outside of this transverse momentum range, and all charged particles with  $|\eta| > 2.5$ , are extrapolated to the front face of the calorimeter at  $R_{\text{calo}} = 1290 \text{ mm}$ . If the extrapolated particle trajectory reaches the calorimeter, the particle energy is added to a calorimeter tower at the extrapolated  $\phi$ , similar to the treatment of all charged particles in configuration A.

The energy of neutral particles is added to the calorimeter tower in the same way as in configuration A. The selection employing  $p_T^{\text{tower}} > p_T^{\text{min}}$  is applied as well after the tower energy smearing with the parameters for configuration C given in Table III.5. Figure III.30 shows the calorimeter-only composition of a given truth-level jet in configuration A together with the track-and-calorimeter jet composition of configuration C. The two model configurations produce significantly different jet compositions, in particular with respect to the energy flow from low- $p_T$  constituents.

The implementation of this detector model is available from the `DetectorModel` subdirectory of the `git` repository at <https://github.com/gsoyez/lh2017-2prongs/>.

### 2.6.2 Pileup Mitigation

Besides the detector response discussed above, LHC collisions are also contaminated by pileup. While pileup multiplicities remained reasonably low, around 20, during Run I of the LHC, they already increased to 40-60 in Run II and are expected to increase even further, in the 140-200 range, for the high-luminosity upgrade of the LHC.

To correct for the energy bias and smearing associated with the pileup contamination, one uses a variety of pileup mitigation techniques (see Ref. [836] for a recent review). The standard approach for most applications is the area–median subtraction method [455, 681, 682, 837] (potentially using the `ConstituentSubtractor` [686] for subtracting pileup from jet shapes). Recently, more complex pileup subtraction methods have been proposed [683–685, 687] (see also [838, 839]). In particular, PUPPI [685] and the `SoftKiller` [683] both provide event-wide pileup mitigation techniques that show good performance in terms of resolution, at the expense of requiring some degree of fine-tuning of their free parameters. PUPPI has already been used by the CMS collaboration in a series of jet-substructure studies.

For these proceedings, we will concentrate on the `SoftKiller` approach. This is motivated by its speed—for a typical LHC event with pileup would be subtracted and clustered in 200-700  $\mu\text{s}$ —by the fact that a public implementation is available, and by the fact that the `SoftKiller` and PUPPI have shown similar performance (see e.g. [840]).

The `SoftKiller` algorithm works by iteratively removing the softest particle in the event until the area-median pileup density estimate  $\rho$  is zero. This is equivalent to breaking the event in patches and iteratively removing the softest particle in the event until half of the patches are

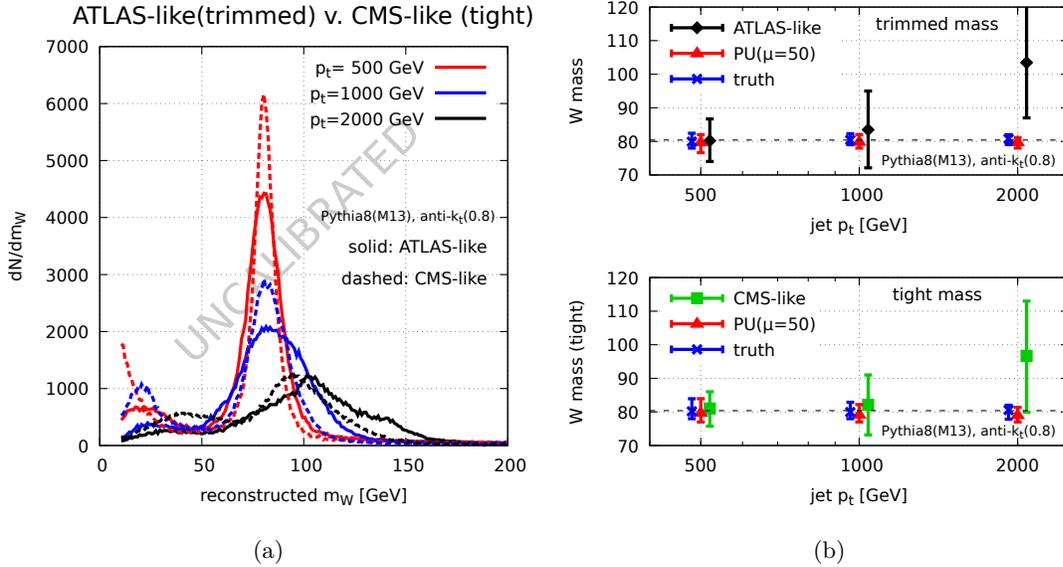


Fig. III.31: (a) Mass reconstruction with detector effects at different jet  $p_T$ . (b) The peak position and width of the mass distribution, after pileup and, separately, after detector effects.

empty, i.e. imposing a cut

$$p_T^{\text{cut}} = \text{median}_{p \in \text{patches}} \left\{ \max_{\text{particle } i \in p} \{p_{Ti}\} \right\}. \quad (\text{III.37})$$

SoftKiller then removes particles below a cutoff  $p_T^{\text{cut}}$ , chosen such that  $\rho = 0$ . We have

$$p_T^{\text{cut}} = \text{median}\{p_{Ti}^{\text{max}}\}. \quad (\text{III.38})$$

SoftKiller has been shown to provide good performance for removing pileup contamination.

### 2.6.3 Impact on Mass Resolution

The impact of detector resolution and pileup is far more complex than including the detector simulation discussed in Sec. 2.6.1. In particular, we have not performed any calibrations or in-situ corrections. We have therefore limited the scope of these proceedings to a discussion of their effects on mass resolution. This will highlight the need to perform a more sophisticated study when considering detector effects, which will be left to a future publication.

In Fig. III.31, we show distributions of the groomed mass for a hadronically-decaying boosted  $W$  boson in our CMS-like and ATLAS-like detectors at different values of the jet  $p_T$  (Fig. III.31a), together with the measure peak position and width, for the trimmed mass with the ATLAS-like detector simulation and the tight (mMDT) mass with the CMS-like simulation (Fig. III.31b).<sup>19</sup> We see that the detector has a significant impact on the distributions: while the  $W$  peak remains decent for 500-GeV jets, typical of most applications today, the raw resolution considerably degrades as the  $p_T$  is increased. We also see that both detectors show a similar trend, despite their different approaches. For  $p_T$ 's up to 1 TeV, it seems that the configuration C gives a slightly better resolution than the configuration A, an effect that is tentatively attributed to particle flow. However, at high  $p_T$  there is a trade off between tracking and calorimetry, which

<sup>19</sup>The position and width are defined as the median and width of the smallest mass window containing 40% of the events.

is better in configuration A, leading to comparable detector resolution. Furthermore, in the full experiments, the reconstruction algorithms are complementary and adapted to the experiments' detector technologies, so that the two experiments achieve similar performance. Due to the fact that the end results of the two detectors are quite similar, we find it unwise to draw any deeper conclusions, since there are potentially other effects, not included in our simulation, that could determine the final outcome. We will however point out that, in our (over-simplified) preliminary studies, the main patterns observed in Sec. 2.5.3, i.e. the overall good performance of  $N_2^{(1,2)}$  and  $D_2^{(2)}$  and the interest in investigating dichroic ratios, seem to remain valid.

Figure III.31b also includes the mass resolution in the presence of pileup, here Poisson-distributed with an average multiplicity of 50, mitigated using the SoftKiller method described in Sec. 2.6.2. Based on this simple analysis, it seems that the reconstruction of  $W$  mass peak still behaves properly in the presence of pileup. It is important to note, however, that this pileup result is in the absence of detector effects, and there is a non-trivial interplay between detector corrections and pileup corrections.

This brief study highlights the essential importance of understanding detector effects when designing jet substructure observables. Although we have found that the  $N_2$  and  $D_2$  observables emphasize, respectively, performance and robustness, it would also be interesting to understand if the different detectors played a role in the choice of one observable as compared with the other. As highlighted by the specific example of the groomed jet mass here, such a study requires considerable care to perform in a meaningful manner. The detectors have a large impact, and the two detectors appear qualitatively similar in our analysis setup. Differences may therefore be due to more subtle features, beyond those included in our study, and this is probably something that is best left to the experimental collaborations to study with realistic detector simulations. We hope that this study motivates the experimental collaborations to further investigate the performance and robustness of different two-prong tagging observables at each of the different steps along the chain of realism in order to understand the different choices in observables.

## 2.7 Polarization Dependence

For signal jets, there is another consideration related to the robustness of two-prong tagging, namely the specific nature of the decaying electroweak-scale resonance, which can also affect the substructure observables. Assuming that this electroweak-scale resonance is a color-singlet decaying to quarks, it is completely characterized by its spin structure.

All of our taggers consist of two conceptually distinct components, which will be affected in different ways by the polarization. We begin by making a cut on the (groomed) mass, followed by a cut on a particular jet shape which is sensitive to the two-prong structure. These two steps are associated with very different physics. A jet from a  $W \rightarrow q\bar{q}$  decay consists of two collimated sprays of radiation, which are proxies for the  $q$  and  $\bar{q}$ , along with radiation emitted from the dipole. Ideally, a groomer should terminate when it de-clusters the jet into two subjects corresponding to the jets initiated by the two quarks. If there is a large fraction of decays where the momentum sharing between the two subjects is hierarchical, though, then the groomer can remove one of the subjects. In this case, the jet will fail the groomed mass criteria. Since the polarization controls the momentum sharing of the subjects, this introduces a sensitivity on the polarization into the tagging procedure. On the other hand, the 2-prong tagging observables that we consider are all formed as ratios of an observable which is sensitive to radiation from the prongs, divided by a mass-type observable, which (largely) removes the dependence on the momentum sharing of the subjects.

In this subsection, we consider two main issues, namely the dependence of the standard jet substructure observables to the polarization of the signals, and the ability to perform polarimetry

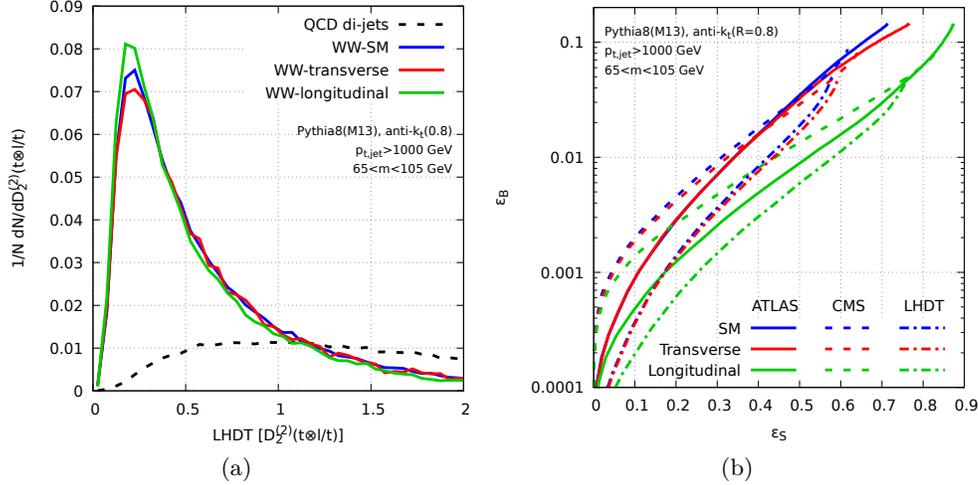


Fig. III.32: a) Distributions of the  $D_2$  observable as measured on the samples with different polarization compositions. These observables are found to be largely insensitive to the polarization. b) ROC curves comparing the different grooming strategies for different polarization samples. While the polarization has a limited effect on the jet shape observables, it modifies the groomed mass acceptance, and therefore the tagging efficiency, significantly.

using jet substructure observables measured on the hadronic decay products.

### 2.7.1 Performance Impact of Polarization

We begin by studying the sensitivity of two-prong taggers to polarization. To do this, we consider samples of hadronically decaying  $W$ s, which are either purely transversely polarized, purely longitudinally polarized, or have the Standard Model fraction (mostly transverse). The details of the sample generation were discussed in Sec. 2.3.3.

In Fig. III.32a, we show the  $D_2$  observable as measured on longitudinal and transverse hadronically-decaying  $W$  bosons, as well as the Standard Model mixture, as a representative example of the dependence of a two-prong tagger on polarization. From this figure, we see that the jet shape observable itself is remarkably insensitive to the polarization, due to its ratio nature. However, it is important to emphasize that this does not imply that the tagging performance is also independent of the polarization. As can be seen in Fig. III.32b, the tagging performance is significantly worse for transversely-polarized  $W$  bosons as compared with longitudinally-polarized  $W$  bosons. This is due to the fact that transversely-polarized  $W$  bosons have a more asymmetric energy sharing, and it is therefore more likely that one of the subjet prongs is groomed away, leading to the jet failing the mass cut. This difference between longitudinally and transversely-polarized bosons should be taken into account in studies at the LHC. It would also be interesting to develop tagging or reconstruction schemes that are less sensitive to polarization.

### 2.7.2 Tagging Longitudinal vs. Transverse Bosons

Having understood the dependence of standard jet substructure observables on polarization, it is interesting to know whether jet substructure observables are able to tag distinct polarizations on hadronically-decaying particles, and with what efficiency this can be done. Here, we do

not perform a comprehensive study, but restrict ourselves to studying a particular example of an observable which is sensitive to the  $W$  polarization, and we evaluate its performance for distinguishing transverse and longitudinal  $W$  bosons.

The sole impact of the polarization of the decaying object is to determine the kinematics of the decaying subjects. Indeed, this can be made rigorous in the sense of a factorization theorem for boosted jets in the two-prong limit. We are therefore interested in an observable that is sensitive to the kinematics of the two subjects. While a variety of different observables could be considered, here we consider the  $z_g$  observable [645, 824, 841], which measures the momentum sharing of the subjects. The precise definition of  $z_g$  was given in Sec. 2.3.2, which we recall for convenience

$$z_g = \frac{\min [p_{Ti}, p_{Tj}]}{p_{Ti} + p_{Tj}}. \quad (\text{III.39})$$

Here  $p_{Ti}$  and  $p_{Tj}$  are the momenta of the first set of subjects that pass the SoftDrop criteria.

Since we are focused on robustness in this paper, it is also worth commenting on the robustness of the  $z_g$  observable. For signal jets, since this observable measures global energy properties of the subjects, it is stable. Interestingly, it is also remarkably stable on the background, where it flows in the high- $p_T$  limit to the QCD splitting function [824].

In Fig. III.33a, we show the  $z_g$  distribution for different polarized  $W$  bosons. For reference, the  $z_g$  distribution for QCD dijets is also shown. The  $z_g$  distribution for transversely-polarized  $W$  bosons follows most closely the QCD distribution, as expected, being peaked at small values of  $z_g$ . On the other hand, for longitudinally-polarized  $W$  bosons, the  $z_g$  distribution is peaked at high values of  $z_g$ . This illustrates that the  $z_g$  observable is indeed behaving as expected, and is achieving sensitivity to the polarization of the decaying  $W$  boson only from its hadronic decay products. In Fig. III.33b, we show the ROC curve for separating longitudinal from transverse  $W$  bosons. Here we see that  $z_g$  provides moderate separation power for tagging the polarization of the decaying  $W$  bosons. It would be interesting to investigate this further to see if more ideal observables could be found, however, we are not optimistic that significant improvement can be achieved, since the primary imprint of the polarization should be in the kinematics of the subjects.

## 2.8 Summary and Recommendations

In this paper, we performed a comprehensive study of performance and robustness for two-prong tagging, and provided a unifying approach to understanding different classes of two-prong taggers based on dichroic observables, which allow for different amounts of grooming in the numerator and denominator of observables. We introduced measures of robustness, in addition to the standard measures of performance, and we used these measures to study the robustness of two-prong taggers to theory issues, namely hadronization and underlying event. We believe that these will be of more general utility in jet substructure, and could be applied also to study three-prong tagging, for example.

As a part of our study, we have also introduced a number of new dichroic observables, which generalize the dichroic  $N$ -subjettiness observables to observables formed from the energy correlation functions. This offers a general and unifying approach to designing new jet substructure observables which are simultaneously performant and resilient. We have shown that the  $N_2$ -style observables used by CMS tend to be more performant, but less resilient than the  $D_2$ -style observables used by ATLAS. For a given observable, we have found that moving to a dichroic variant can typically improve performance without significantly decreasing its robustness to hadronization.

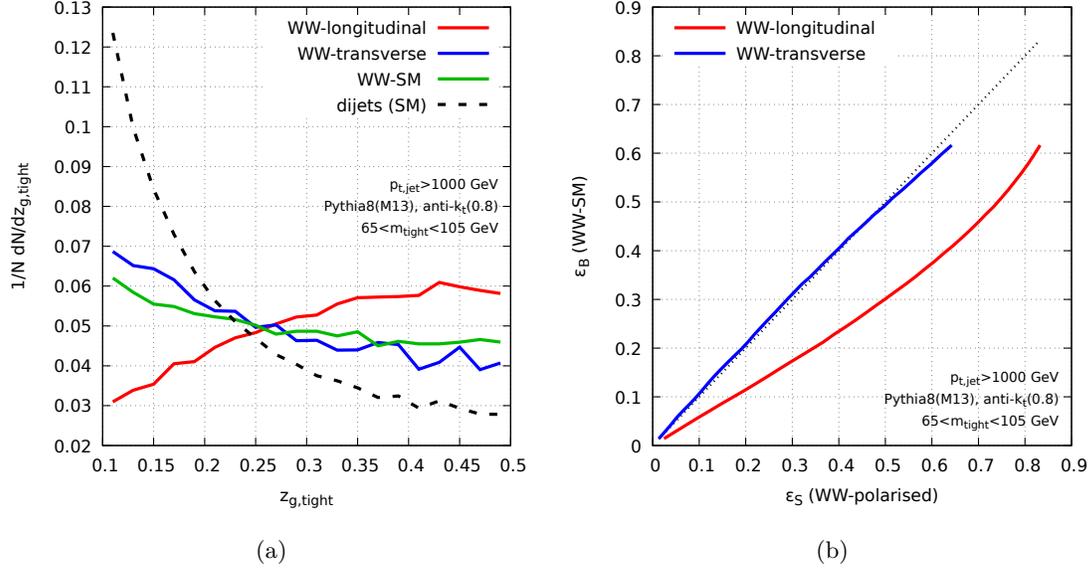


Fig. III.33: a) The  $z_g$  distribution as measured on boosted  $W$  samples with different polarization compositions. Transversely-polarized  $W$ s behave similarly to the QCD background, while decent separation is observed between longitudinally- and transversely-polarized  $W$ s. b) A ROC curve showing the separation between transversely- and longitudinally-polarized  $W$  bosons using the  $z_g$  observable. Moderate separation is observed.

We have also studied the effect of polarization on two-prong taggers. We found that while polarization has minimal effect on standard two-prong tagging observables, since they are typically defined as ratios, it has a large effect on their tagging efficiency due to applied mass cuts. Significantly better tagging performance is observed for longitudinally-polarized bosons. An interesting avenue beyond standard two-prong tagging is using jet substructure observables to tag the polarization of decaying Standard Model bosons using their hadronic decay products. We proposed the observable  $z_g$ , which measures the momentum sharing asymmetry between subjets, as an effective polarization tagger. We illustrated that separation between longitudinal and transverse bosons can be achieved. It would be interesting to study this problem in more detail.

We have identified a number of new two-prong taggers that outperform, in both robustness and tagging performance, those currently used by the ATLAS and CMS collaborations. These findings are summarized in Fig. III.29, which also lists a number of promising new observables. We therefore believe that further studies using more detailed simulations of the ATLAS and CMS detectors, and ultimately on real data, would be of significant interest. More generally, we expect that the emphasis on a simultaneous evaluation of the performance and robustness, as well as the particular metrics introduced in this paper, will play a significant role in future studies of jet substructure techniques at the LHC.

## Acknowledgments

We thank the participants of Les Houches 2017 for a lively environment and useful discussions. The work of GS is supported in part by the Paris-Saclay IDEX under the IDEOPTIMALJE grant, by the French Agence Nationale de la Recherche, under grant ANR-15-CE31-0016, and by the ERC Advanced Grant Higgs@LHC (No. 321133). The work of DK is supported by NRF

CSUR and Incentive funding The work of FD and JT is supported by the DOE under grant contract numbers DE-SC-00012567 and DE-SC-00015476. The work of IM is supported by Office of High Energy Physics of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231, and the LDRD Program of LBNL. AS acknowledges support from the National Science Centre, Poland Grant No. 2016/23/D/ST2/02605, the MCnetITN3 H2020 Marie Curie Initial Training Network, contract 722104 and COST Action CA15213 THOR.

# Chapter IV

## Standard Model Higgs

### 1 Transverse momentum resummation in Higgs boson plus two jet production via weak boson fusion process <sup>1</sup>

We study the soft gluon resummation effect in the Higgs boson plus two-jet production at the LHC. By applying the transverse momentum dependent factorization formalism, the large logarithms introduced by the small total transverse momentum of the Higgs boson plus two-jet final state system, are resummed to all orders in the expansion of the strong interaction coupling at the accuracy of Next-to-Leading Logarithm. We also compare our result with the prediction of the Monte Carlo event generator Pythia8, and found noticeable difference in the distributions of the total transverse momentum and the azimuthal angle correlations of the final state Higgs boson and two-jet system.

#### 1.1 Introduction

In this letter, we apply the transverse momentum dependent (TMD) resummation method to study the soft gluon resummation effect on the production of Higgs boson plus two jets via weak boson fusion (VBF) at the LHC. By applying the TMD factorization formalism, the large logarithms introduced by the small total transverse momentum ( $q_\perp$ ) of the Higgs boson plus two-jet final state system, are resummed to all orders in the expansion of the strong interaction coupling at the accuracy of Next-to-Leading Logarithm. In terms of our TMD factorization formalism, the  $q_\perp$  differential cross section is factorized into several individual factors, and each factor is calculated up to the one-loop order. As of today, the Monte Carlo (MC) event generators are the only available tools to predict the soft gluon shower effect for this channel. Our calculation, for the first time, provides an important test on the validity of the commonly used MC event generators.

#### 1.2 The Model

Our TMD resummation formula can be written as:

$$\frac{d^6\sigma}{dy_H dy_{J1} dy_{J2} dP_{J1\perp}^2 dP_{J2\perp}^2 d^2\vec{q}_\perp} = \sum_{ab} \left[ \int \frac{d^2\vec{b}}{(2\pi)^2} e^{-i\vec{q}_\perp \cdot \vec{b}} W_{ab \rightarrow Hcd}(x_1, x_2, b) + Y_{ab \rightarrow Hcd} \right], \quad (\text{IV.1})$$

where  $y_H$ ,  $y_{J1}$  and  $y_{J2}$  denote the rapidities of the Higgs boson and the jets, respectively,  $P_{J1\perp}$  and  $P_{J2\perp}$  are the jets transverse momentum, and  $\vec{q}_\perp = \vec{P}_{H\perp} + \vec{P}_{J1\perp} + \vec{P}_{J2\perp}$  is the imbalance transverse momentum of the Higgs boson and the two final state jets. The first term ( $W$ ) contains all order resummation effect and the second term ( $Y$ ) accounts for the difference between the fixed order result and the so-called asymptotic result which is given by expanding the resummation result to the same order in  $\alpha_s$  as the fixed order term.  $x_1$  and  $x_2$  are the momentum fractions of the incoming hadrons carried by the incoming partons.

---

<sup>1</sup> P. Sun, C.-P. Yuan, F. Yuan

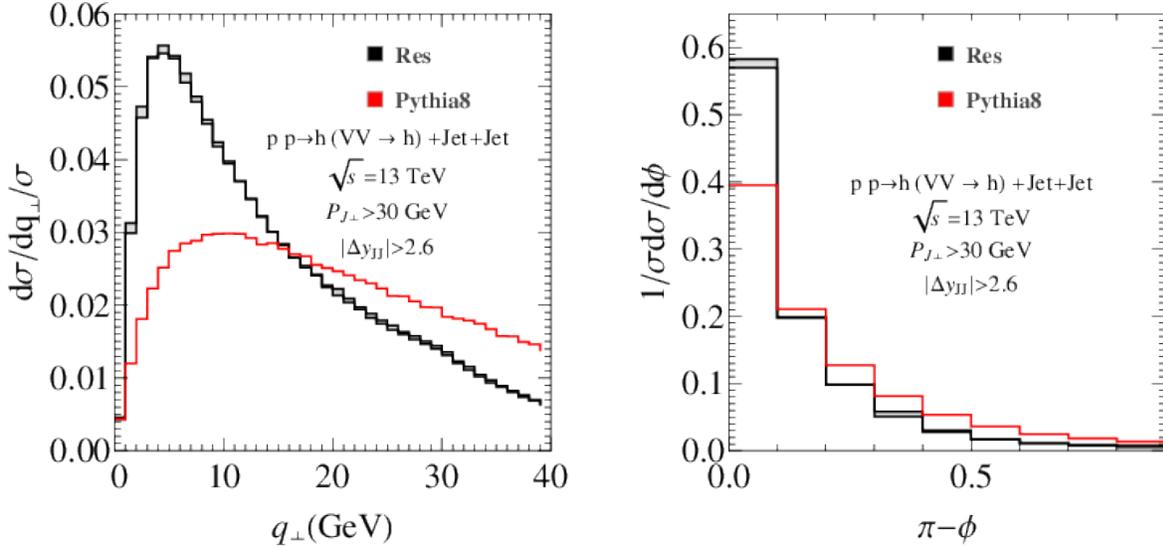


Fig. IV.1: The differential cross sections of Higgs boson plus two jet production at the LHC as functions of  $q_{\perp}$  and azimuthal angle  $\phi$  between Higgs boson and the final state two-jet system. In these plots, the  $\alpha_s$  order  $Y$  pieces are included in the resummation curves. The predictions from Pythia8 are based on the tree level scattering amplitudes with parton showers. The uncertainty of our resummation calculation is estimated by varying the resummation scale  $\hat{\mu}$  from  $P_{J\perp}^{lead}$  to  $P_{J\perp}^{sub}$ .

### 1.3 Numerical Results

Following the study in Ref [842], The uncertainty of our resummation calculation is estimated by varying the resummation scale  $\hat{\mu}$  from  $P_{J\perp}^{lead}$  to  $P_{J\perp}^{sub}$ , where the  $P_{J\perp}^{lead}$  and  $P_{J\perp}^{sub}$  are the transverse momenta of the final state leading jet and sub-leading jet, respectively. In addition, we set the renormalization scale in the hard factor to be  $\tilde{\mu} = m_H$ , and use the CT14 NNLO PDFs [843], with the mass of the Higgs boson ( $m_H$ ) set to be 125 GeV. Following the experimental analysis in Ref. [844], we require the rapidity of the observed jets to satisfy  $|y_J| < 4.4$ . We use the anti- $k_t$  algorithm to define the observed jets, and the jet size and the minimal transverse momentum are set at  $R = 0.4$  and  $P_{J\perp} > 30$  GeV. In our calculation we have applied the narrow jet approximation. We have also constrained the two final state jets to have a large rapidity separation with  $|\Delta y_{JJ}| > 2.6$ , which is the signal region for detecting the Higgs boson production via weak boson fusion process. In the same figure we also compare to the predictions from Pythia8 which are based on the tree level scattering amplitudes with parton showers.

As shown in Fig. IV.1, we find large difference between the Pythia8 and our predictions in the distributions of the total transverse momentum ( $q_{\perp}$ ) and the azimuthal angle ( $\phi$ ) correlations of the final state Higgs boson and two-jet system, after imposing the kinematic cuts used in the LHC data analysis. Specifically, Pythia8 predicts a flatter shape, in  $q_{\perp}$  distribution, than our resummation calculation. Another significant disagreement lies in the peak position of the  $q_{\perp}$  distribution. Pythia8 predicts a peak in  $q_{\perp}$  around 10 GeV, while ours is at about 5 GeV. Similarly, they also differ in predicting the  $\phi$  distribution. Pythia8 predicts a less back-to-back configuration, between Higgs boson and the final state two-jet system, than ours. In Ref. [844], the ATLAS Collaboration required the azimuthal angle separation ( $\phi$ ) between the Higgs boson and the di-jet system to be  $\phi > 2.6$ , and compared the measured fiducial cross section with the

Pythia8 prediction. For  $\phi > 2.6$ , Pythia8 and ours differ by about 8%, and our resummation calculation results in a larger total fiducial cross section. This implies a larger value in the coupling of Higgs boson to weak gauge bosons by about 4%. At the High-Luminosity LHC, with an integrated luminosity of up to  $3000 \text{ fb}^{-1}$ , the expected precision on the measurement of the production cross section of the SM-like Higgs boson via VBF mechanism is around 10% [845]. Hence, the difference found in our resummation and Pythia8 calculations of the fiducial cross sections could become important. Further comparisons on various event shapes between the experimental data and our resummation predictions could also be carried out in order to test the Standard Model and to search for New Physics.

## 2 Higher order corrections in VBF Higgs production <sup>2</sup>

Recently, a significant error was discovered [135] in the virtual matrix elements of the VBF Higgs plus three jet NLO cross section calculated in Ref. [846] and implemented in the program packages VBFNLO [477, 847, 848] and Powheg-Box [350, 351, 849, 850]. Since this result was used as part of the calculation of the fully differential NNLO QCD corrections [194], this bug impacts the size of the second order corrections. We report here on updated cross sections after implementation of the correct virtual corrections.

Because the issue is in the  $H + 3j$  calculation, there is no impact on the inclusive cross section. However, after VBF cuts, the cross section is slightly increased. We impose the following conditions on the tagging jets

$$m_{j_1, j_2} > 600 \text{ GeV}, \quad \Delta y_{j_1, j_2} > 4.5, \quad (\text{IV.2})$$

where the jets  $j_1, j_2$  are required to have transverse momentum  $p_{t,j} > 25 \text{ GeV}$  and rapidity  $|y_j| < 4.5$ , as well as being in opposite hemispheres ( $y_{j_1} y_{j_2} < 0$ ). The NNLO cross section after these cuts is  $\sigma = 0.844^{+0.008}_{-0.008} \text{ pb}$ , which is a  $\sim 2\%$  change to the result originally reported in Ref. [194].

In Fig. IV.2, we present the differential cross sections of the transverse momentum of the two leading jets and the Higgs boson, and the rapidity separation between the tagging jets. The correction of the bug in Ref. [846] leads to slightly harder jets, with an increase of  $\sim 2\%$  of the NNLO cross section at high transverse momentum. The difference is most visible on the rapidity separation, where the shape of the  $K$ -factor is changed for very widely separated jets.

Both the VBF Higgs NNLO QCD and the VBF Higgs plus three jet NLO cross sections have then been used to study the dependence of the NNLO result on the jet definition in Ref. [851]. For the dependence of the integrated cross section on the jet clustering radius  $R$ , shown in Fig. IV.3, the matching value between LO, NLO and NNLO results stays at approximately  $R=1.0$  after the correction. The slope of the NNLO cross section flattens slightly. In Fig. IV.4 we show the updated results on the differential distributions. The smallness of the NNLO corrections in the transverse momentum distributions of the subleading jet and the Higgs boson when using  $R=1.0$  is unchanged from our original result. In contrast, we had seen relevant remaining NNLO effects when using the larger clustering radius at small transverse momenta of the leading jet and at large rapidity differences of the two tagging jets. In both cases, the remaining effects are significantly reduced after the update.

---

<sup>2</sup> F. Dreyer, A. Karlberg, M. Rauch

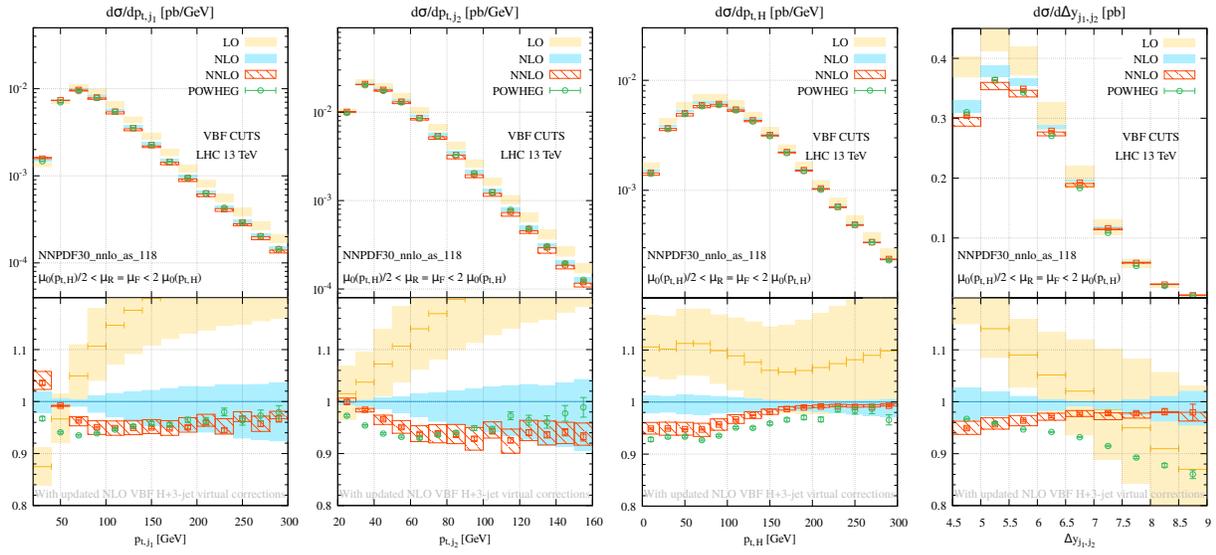


Fig. IV.2: Differential cross sections for the transverse momentum of the tagging jets and the higgs boson, as well as for the rapidity separation between the jets.

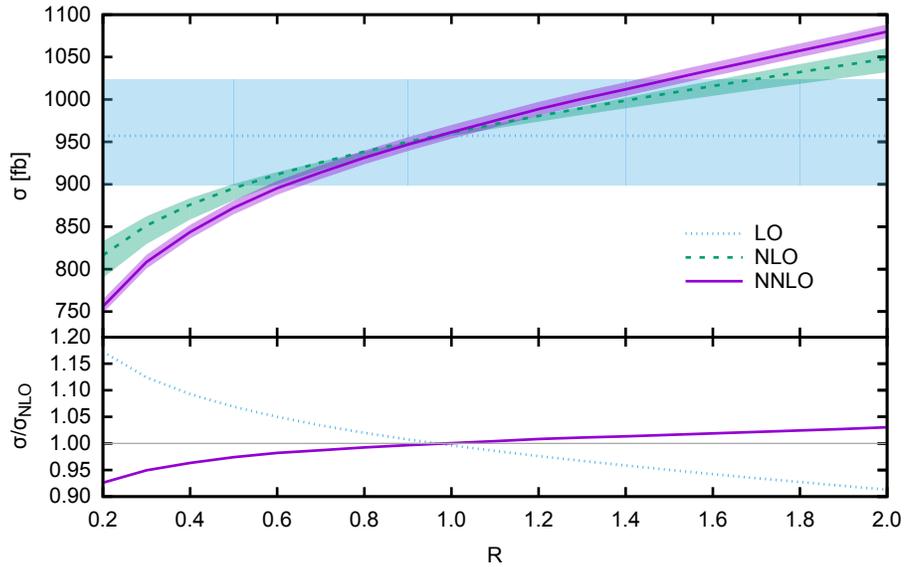


Fig. IV.3: Integrated cross section for VBF- $H$  production as function of the jet clustering radius  $R$ .

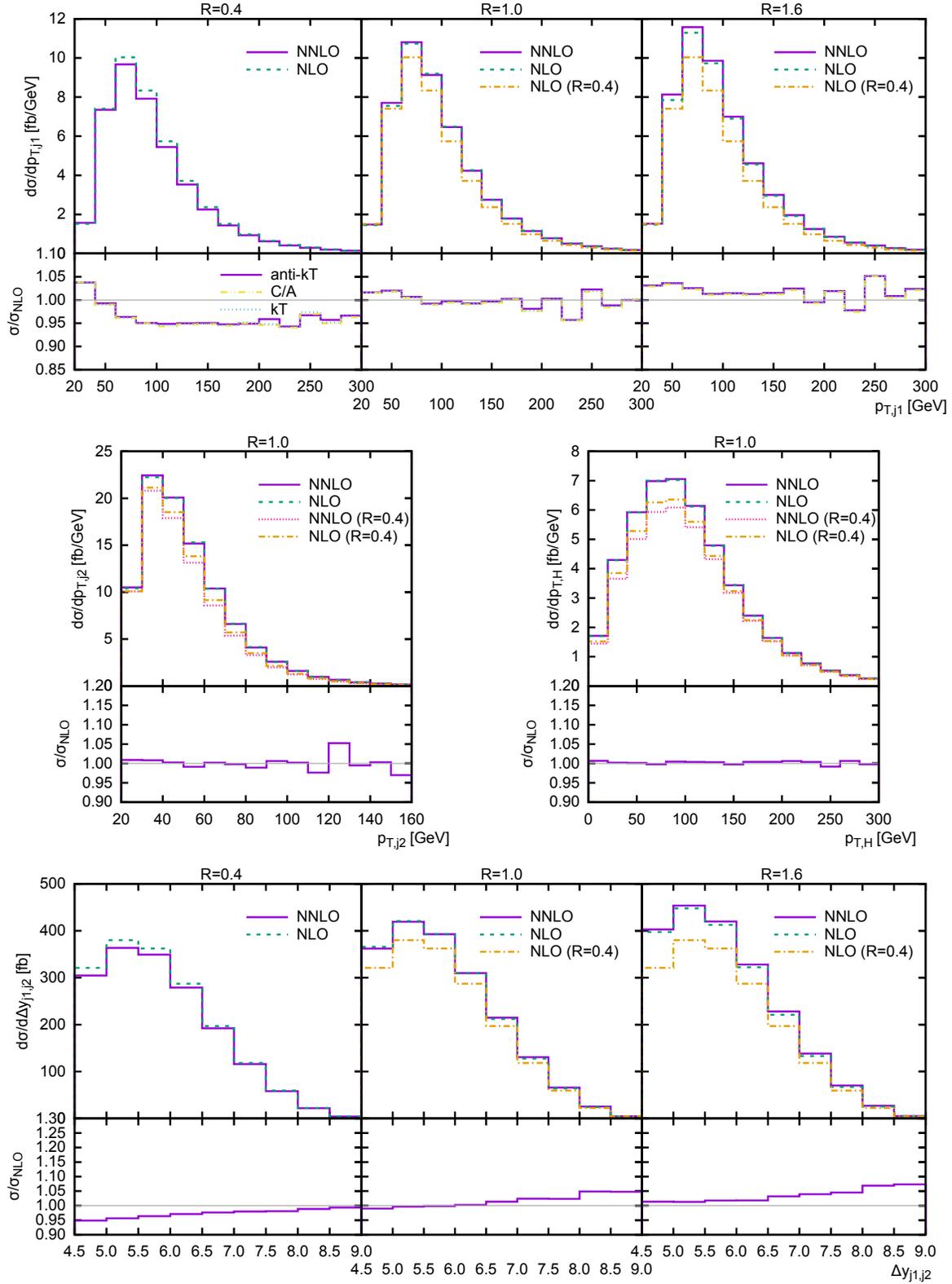


Fig. IV.4: Transverse momentum distribution for the leading jet (*top*), the subleading jet (*centre left*) and the Higgs boson (*centre right*) as well as the rapidity difference of the two tagging jets (*bottom*).

### 3 The gluon-fusion component of Higgs boson-jet-jet production and its suppression in weak-boson fusion studies <sup>3</sup>

We compare various perturbative predictions for the gluon-fusion component of Higgs-boson production in association with dijets (Hjj), in particular within the cuts designed to enhance the acceptance of Hjj from the production through weak-boson fusion.

The first sub-leading corrections from the high-energy limit have recently been included in the description of HEJ. The predictions are changed only minimally by the inclusion of these sub-leading corrections, and are here contrasted with those obtained at next-to-leading order accuracy. We show that a moderate cut on further central hard jet activity can reduce the gluon-fusion component in weak-boson fusion measurements by about 10–15%.

#### 3.1 Introduction

The study of the process of Higgs boson production in association with dijets is significant for several reasons. It can proceed through the Born level fusion of two weak bosons exchanged between quarks from each proton to a Higgs boson [852] (WBF) – or as a radiative correction to the top-loop mediated fusion of two gluons [853, 854] (GF). The quantum interference between the two processes is negligible [855], so the processes can in theory be studied independently. As such, the WBF process provides a direct and independent measurement of the Higgs boson coupling to weak bosons. Meanwhile, the GF process allows for tests of the  $CP$ -properties of the Higgs-boson couplings to gluons [856–858]; in particular the small admixtures from extended sectors with direct  $CP$ -violations can be measured by studying the azimuthal angle between the jets [858, 859].

The degree of success of all such studies will, however, depend not just on the standard suppression of background events unrelated to Higgs-boson production, but also on separating the WBF and GF component of the Hjj production. We will here report on an updated study on possibilities of suppressing the GF component during WBF studies. The GF component has the largest cross section at the inclusive level of Hjj, but it is dominated by the sub-processes involving two incoming gluons. This skews the partonic processes towards smaller total invariant masses than that for WBF, since the gluons predominantly carry less of the proton energy than the valence quarks do. A minimal cut on the invariant mass between the two jets (anti-kt,  $p_t > 30$  GeV) will therefore favour WBF over GF. We will consider cuts of  $m_{(j_1 j_2)} > 400$  GeV, in line with current experimental studies. Similarly, a minimal rapidity difference between the jets is often required, e.g.  $|y_{j_1} - y_{j_2}| > 2.8$ . This constitutes the WBF cuts.

A reliable calculation of the GF component within these cuts obviously rely on a successful description not just of the total rate, but also the spectrum in  $m_{j_1 j_2}$  and  $|y_{j_1} - y_{j_2}|$ . The description of QCD processes at large dijet invariant masses is traditionally difficult, and has been studied for both Wjj [860, 861] and Zjj [862], where also both a WBF and a QCD-dominated process contribute.

We will here be concerned with predictions from pure NLO [863, 864] and *High Energy Jets* (HEJ) [865–869], which both give a good description of e.g. the dijet invariant mass spectrum for the QCD-dominated processes in Wjj [860, 861, 870, 871]. The properties of the radiative corrections associated with the colour-octet exchange between jets separated by a large invariant mass is similar between various underlying processes, such that a test of the predictions for  $Zjj$  or  $Wjj$  can aid the investigation of  $Hjj$ . This is particularly useful, when further discrimination between the WBF and GF processes is sought by utilising the observation that the colour exchange associated with the GF process leads to increased perturbative jet-activity in the rapidity range of the colour-octet exchange compared to the WBF process [872]. The recent

---

<sup>3</sup> J. R. Andersen, M. Heil, A. Maier, J. M. Smillie

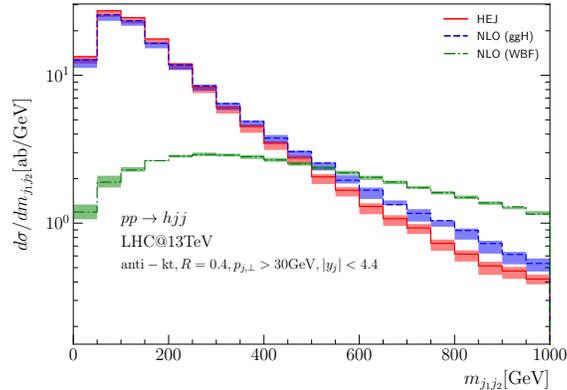


Fig. IV.5:  $d\sigma/dm_{j_1 j_2}$ : The spectrum of the invariant mass between the two hardest jets for WBF and GF HJJ calculated at NLO and with HEJ. The GF component dominates at small  $m_{j_1 j_2}$ . See text for further details.

measurement by ATLAS utilised this to isolate the WBF component of  $Zjj$ , but found very large variations in the predictions of the spectrum of  $m_{j_1 j_2}$  for the GF enhanced sample where a third jet was required, and between these predictions and data.

The reliability of the perturbative prediction for the jet veto can be jeopardised by succumbing to the temptation of using a small transverse momentum for the cut. In this study we will investigate how the suppression of the GF component can be obtained with perturbatively more stable cut on larger transverse momenta, but with the cut applied in a larger region of rapidity. The difference in jet radiation pattern between GF and WBF is valid not just for region in rapidity between the two hardest jets, but in the full region of colour octet exchange, which predominantly is between the forward ( $j_f$ ) and backward ( $j_b$ ) jet. Ref. [452] found stable predictions for  $d\sigma/dm_{j_f j_b}$  between shower, NLO and HEJ for a jet cut of  $p_\perp > 40\text{GeV}$ . Current analyses of LHC data frequently apply a jet veto cut  $25\text{GeV}$  to veto QCD contributions. In this contribution we will investigate the effect of applying a harder jet cut (which will reduce the effect of the veto), but in a larger region of rapidity (which will increase the effect of the veto).

### 3.2 Results

The predictions for the invariant mass spectrum between the two hardest jets for the GF (obtained at NLO and with HEJ) and WBF component (at NLO) of  $Hjj$  is displayed on Fig. IV.5. The cross sections for HEJ reported here have been normalised to NLO accuracy at the level of the inclusive cross section for  $Hjj$ . This decreases the scale variation of the results compared to those reported in Ref. [869]. Frequently applied WBF cuts designed to enhance the WBF component over that of GF apply a cut on the minimum invariant mass of the hardest two jets of  $400\text{GeV}$ . The spectrum for the GF component falls off slightly faster for HEJ than for NLO, which leads to differences for the prediction of the GF component within the WBF cuts. Predictions for the distribution in  $m_{j_1 j_2}$  with similar jet cuts as those applied in the studies of  $Hjj$  have been checked against data for  $Wjj$  [861, 871], albeit for centre-of-mass energies of  $7\text{TeV}$  and  $8\text{TeV}$ . A measurement of the distribution for pure dijet,  $Wjj$  or  $Zjj$  for similar jet cuts and parameters as those applied in  $Hjj$  would help in determining the shape at large invariant dijet mass, and thus reduce the uncertainty in the determination of the GF component. We note that it was found [862] for  $Zjj$  at  $13\text{TeV}$  that the predictions for the QCD-dominated contributions vary by more than a factor of 5 at large invariant masses.

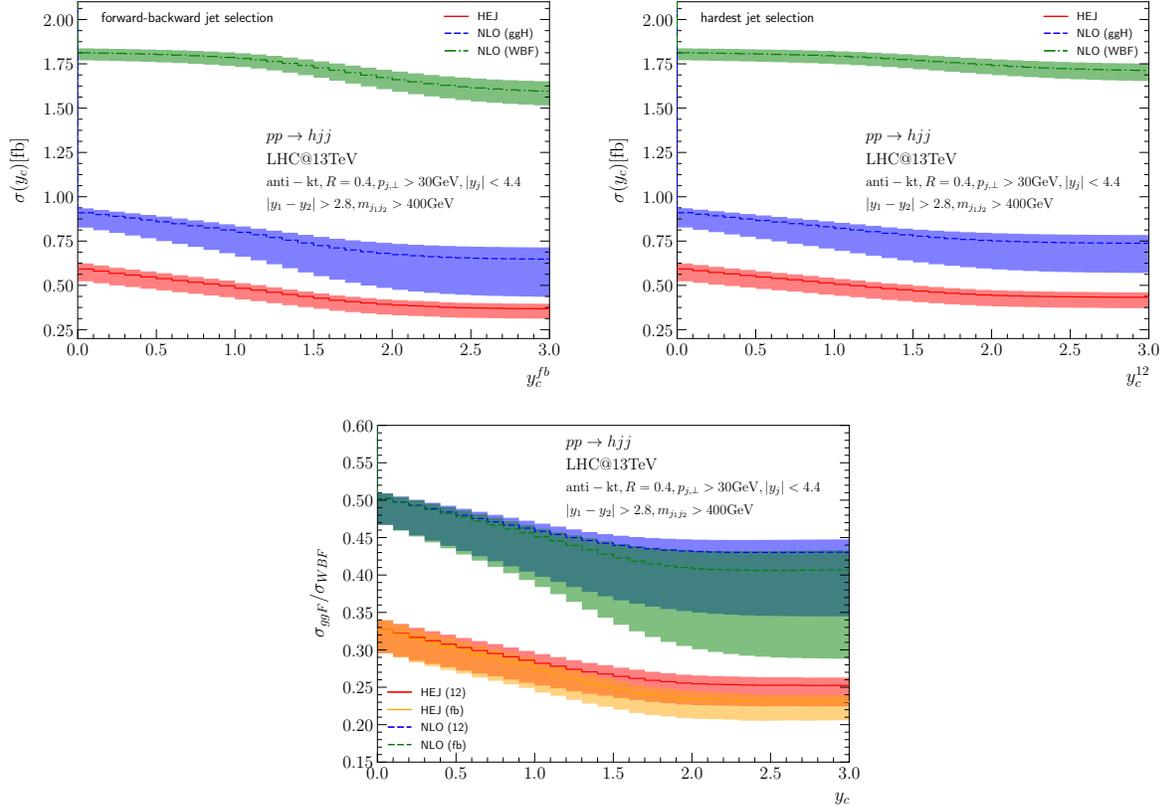


Fig. IV.6: The cross section passing the WBF cuts on the two hardest jets, and a further veto on central jet activity, either between the forward and backward jets (top left) or between the two hardest jets (top right). The former is rejecting more of the gluon-fusion component. The lower panel shows the relative size of the GF component compared to WBF.

The study of [862] discriminates further the GF and WBF components by referring to the dominance of further jet radiation from the GF component from the colour connection between the two jets [872], and thus applies a slightly softer jet veto in-between the two hardest jets. However, the discrimination need not apply just to the rapidity region in-between the two hardest jets. We will here investigate how a harder jet cut (thus perturbatively safer) can be applied to a larger region in rapidity to get the same effect. Instead of applying a jet veto to suppress the GF contribution only if a hard jet exists in-between the two hardest jets of the event, we will investigate the following procedure: Define the two tagging jets to be either the two hardest ( $j_1, j_2$ ) or the two furthest apart in rapidity ( $j_f, j_b$ ), and let  $y_0 = (y_{j_1} + y_{j_2})/2$  or  $y_0 = (y_{j_f} + y_{j_b})/2$ . The region of jet veto is then characterised by a rapidity  $y_c > 0$ , and an event is removed by the veto if it contains a further jet (with transverse momentum at least 30GeV) in-between the two tagging jets, and with a rapidity  $y_j$  with  $|y_j - y_0| < y_c$  (cf. [873]). This definition corresponds to applying no central jet veto at all for  $y_c = 0$ , whilst in the case of the tagging jets being the forward/backward pair the limit of large  $y_c$  corresponds to vetoing all additional jets, and  $\sigma(y_c = \infty)$  is the cross section for the *exclusive* production of a Higgs boson with two jets.

At the top of Fig. IV.6 we plot the result for the predictions of the GF and WBF components passing the cuts on a minimum invariant  $m_{j_1 j_2}$  and rapidity separation between the two hardest jets, as a function of a further veto of jet activity in an increasing region of rapidity

described as mentioned above by  $y_c$ , either between the forward/backward jet pair or just the two hardest jets (right). The difference between vetoing jets in-between the forward/backward pair or the hardest pair is modest, as illustrated on the bottom plot, which shows the relative GF component of WBF+GF passing the cuts, calculated at both NLO and with HEJ. We observe that for a veto characterised by  $y_c = 1$  there is almost no drop in the WBF component, whereas the relative size of the GF component compared to WBF can be lowered from 50% to 45% (NLO predictions, a 10% reduction in the GF contribution) or 33% to 28% (HEJ predictions, a 15% reduction in the GF contributions). We observe that the NLO scale dependence is comparable to the promised reduction in the GF contribution by such a central jet veto.

### 3.3 Conclusions

We find that while a jet veto certainly can serve as a useful tool for further suppressing the GF contribution to HJJ in WBF studies, the GF contribution at large dijet invariant mass in itself has an uncertainty comparable to the promised effects of a moderate central jet veto on the surviving cross section. The difference between the HEJ and NLO predictions is larger than the suppression, which in turn is roughly of the same size as the NLO scale dependence. It therefore seems prudent to perform detailed measurements of the dijet invariant mass spectrum for related processes, and utilise these to improve predictions compared to those contrasted to electroweak production of  $Zjj$  [862].

## 4 The sensitivity of a $ZH/WH$ ratio measurement in the $H \rightarrow b\bar{b}$ channel <sup>4</sup>

### 4.1 Introduction

Since no new fundamental particles beyond the SM have been found at the LHC up to now, the focus is turning towards indirect signals for New Physics. This requires the comparison of precise measurements with higher-order calculations. At a hadron collider, however, effects such as parton luminosities or jet tagging efficiencies influence both the accuracy of the theory prediction as well as the measurement. The identification of suitable observables where the sources of uncertainty are minimised is thus essential for a successful high-precision program at the LHC.

With this objective, we investigate the ratio of  $ZH$  to  $WH$  production. As suggested in [257], the close similarity in the leading-order production mechanisms for both processes should lead to a large cancellation of both the theoretical and the experimental uncertainties in this ratio. However, due to qualitatively different contributions to  $ZH$  production which, in the SM, occur at higher orders, the ratio reveals significant sensitivity to potential new physics. In the SM, the dominant difference between  $WH$  and  $ZH$  production is due to the partonic process  $gg \rightarrow ZH$ , which is mostly mediated by top- and bottom-quark loops. In extensions of the SM, this process may receive contributions from virtual loops of new particles which may alter its effect on the total cross section, as well as on its kinematics [258] (see also [874]). In addition, new  $s$ -channel effects due to the exchange of new Higgs bosons may occur [257].

A detailed analysis of such effects on various kinematical distributions will be presented elsewhere [875]. In this contribution, we study the potential of a dedicated measurement of the  $ZH/WH$  ratio at the LHC, based on recent results for the individual  $WH$  and  $ZH$  cross sections and assuming various degrees of correlation among their uncertainties.

---

<sup>4</sup> R. Harlander, J. Klappert, C. Pandini, A. Papaefstathiou, L. Perrozzi

## 4.2 The double ratio $R_R^{ZW}$

We consider first the experimental quantity:

$$R_{\text{exp}}^{ZW} = \frac{d\sigma^{ZH}}{d\sigma^{WH}}, \quad (\text{IV.3})$$

where  $d\sigma^{VH}$  represents the (differential) rate for vector boson ( $V = (Z, W)$ ) plus Higgs boson ( $H$ ) production. Despite the different final states in the numerator and the denominator, we expect that a number of systematic experimental uncertainties cancel. Here we derive rough estimates of the statistical and systematic uncertainties.

We can also define the contributions of the ‘‘gluon-fusion’’ (ggF) and ‘‘Drell-Yan’’-like (DY) component of the  $ZH$  process by:

$$R_{gg} = \frac{d\sigma^{gg}}{d\sigma_{\text{DY}}^{ZH}} = R_R^{ZW} - 1, \quad (\text{IV.4})$$

The ‘‘double ratio’’  $R_R^{ZW}$  can be constructed by dividing the experimental measurement of ratio of  $ZH$  and  $WH$  events with the theoretical prediction for the ratio between the DY  $ZH$  and  $WH$ :

$$R_R^{ZW} = \frac{R_{\text{exp}}^{ZW}}{R_{\text{DY,th}}^{ZW}}. \quad (\text{IV.5})$$

### 4.2.1 Statistical uncertainty

Error propagation on  $R_R^{ZW}$  yields:

$$\left(\frac{\delta R_R^{ZW}}{R_R^{ZW}}\right)^2 = \left(\frac{\delta R_{\text{DY}}^{ZW}}{R_{\text{DY}}^{ZW}}\right)_{\text{th}}^2 + \left(\frac{\delta R_{\text{exp}}^{ZW}}{R_{\text{exp}}^{ZW}}\right)_{\text{exp}}^2, \quad (\text{IV.6})$$

where the second term on the r.h.s. contains the experimental statistical and systematic uncertainties. In practice, the experimental ratio  $R_{\text{exp}}^{ZW}$  can be evaluated either by performing a fit, or explicitly, e.g.:

$$R_{\text{exp}}^{ZW} = \frac{dN^{ZH}}{dN^{WH}} = \frac{dN^{\ell^+\ell^-b\bar{b}} - dN_{\text{bkg}}^{\ell^+\ell^-b\bar{b}}}{dN^{\ell^\pm b\bar{b}} - dN_{\text{bkg}}^{\ell^\pm b\bar{b}}}, \quad (\text{IV.7})$$

where  $dN^{Xb\bar{b}}$  and  $dN_{\text{bkg}}^{Xb\bar{b}}$  represent the total and background events, respectively, in a certain bin for a given observable, in the final state  $Xb\bar{b}$ , where  $X$  is either one or two leptons. The background subtraction induces a systematic uncertainty which we absorb in the estimations of the next section. Assuming sufficiently large event samples, the (data) statistical uncertainty on the experimental ratio is given by:

$$\begin{aligned} \left(\frac{\delta R^{ZW}}{R^{ZW}}\right)_{\text{exp, stat}}^2 &= \left(\frac{\delta(dN^{\ell^+\ell^-b\bar{b}})}{dN^{\ell^+\ell^-b\bar{b}} - dN_{\text{bkg}}^{\ell^+\ell^-b\bar{b}}}\right)^2 + \left(\frac{\delta(dN^{\ell^\pm b\bar{b}})}{dN^{\ell^\pm b\bar{b}} - dN_{\text{bkg}}^{\ell^\pm b\bar{b}}}\right)^2 \\ &= \frac{dN^{\ell^+\ell^-b\bar{b}}}{(dN^{\ell^+\ell^-b\bar{b}} - dN_{\text{bkg}}^{\ell^+\ell^-b\bar{b}})^2} + \frac{dN^{\ell^\pm b\bar{b}}}{(dN^{\ell^\pm b\bar{b}} - dN_{\text{bkg}}^{\ell^\pm b\bar{b}})^2}, \end{aligned} \quad (\text{IV.8})$$

where we have assumed that the event sample sizes are sufficiently large.

The theoretical ratio  $R_{\text{DY,th}}^{ZW}$  is affected by theoretical systematic uncertainties due to parton density functions and higher-order corrections. Due to the similarity between the DY  $WH$  and  $ZH$  processes, these are expected to be highly correlated. Under the assumption that

the scale uncertainties are fully correlated, we have verified that the combined QCD scale and PDF uncertainties on this ratio are of the order of  $\sim 1\%$  and we neglect them here for simplicity. The effect of electroweak corrections is left to future work [875]. In the next section we describe an estimation of the systematic experimental uncertainty on  $R_{\text{exp}}^{ZW}$ .

#### 4.2.2 Systematic uncertainty

The systematic uncertainty on the  $R_{\text{exp}}^{ZW}$  includes all uncertainties contributing to an experimental measurement of this quantity. While a precise determination of the systematics would require a fully-fledged experimental analysis, we can nevertheless extract an estimate of the uncertainty from the results presented in [876], for the separate  $ZH$  and  $WH$  signal strengths:

$$\begin{aligned}\mu_{ZH} &= 1.12_{-0.33}^{+0.34}(\text{stat.})_{-0.30}^{+0.37}(\text{syst.}), \\ \mu_{WH} &= 1.35_{-0.38}^{+0.40}(\text{stat.})_{-0.45}^{+0.55}(\text{syst.})\end{aligned}\tag{IV.9}$$

The systematic term of the signal strength uncertainty includes all sources of experimental nature, related to the background and signal Monte Carlo simulation and data driven estimates, and to the finite size of the simulated samples.

We assume that these systematic uncertainties can be propagated directly on the experimental ratio defined by Eq. (IV.3). For simplicity, the uncertainties are symmetrized, and three different correlation scenarios are explored: uncertainties on  $\mu_{ZH}$  and  $\mu_{WH}$  completely uncorrelated, 50% correlated or 100% correlated, which we denote as:  $p_{WZ} = (0, 0.5, 1.0)$  respectively.

### 4.3 Analysis and results

We present here a proof-of-principle Monte-Carlo level analysis demonstrating the usefulness of the double ratio  $R_R^{ZW}$ . The  $VH(bb)$  selection follows the experimental cuts applied in [876] as closely as possible.

In the analysis, jets are reconstructed using the anti- $kt$  clustering algorithm with a jet-radius parameter of  $R = 0.4$ . The jet transverse momentum is required to be greater than 20 GeV for ‘central jets’ ( $|\eta| < 2.5$ ) and greater than 30 GeV for ‘forward jets’ ( $2.5 < |\eta| < 5$ ). Selected central jets are labeled as ‘ $b$ -tagged’ if a  $b$ -hadron is found within the jet. A  $b$ -tagging efficiency of 70% is considered, flat over transverse momentum of the jets, to reproduce the efficiency of the experimental  $b$ -tagging algorithm. The leading  $b$ -jet is required to have a transverse momentum larger than 45 GeV. The missing transverse energy is reconstructed by taking the transverse momentum of the negative sum of the four-momenta of all the visible particles. Electrons and muons were subject to isolation criteria, by requiring the scalar sum of the transverse momenta of tracks in  $R = 0.2$  around them to be less than 0.1 times their transverse momentum:  $\sum_{R < 0.2} p_{T,\text{tracks}} < 0.1 \times p_{T,\ell}$ .

Three event selections were considered in [876], corresponding to the  $Z \rightarrow \nu\nu$ , the  $W \rightarrow l\nu$ , and the  $Z \rightarrow ll$  channels. Here we only consider the 1- and 2-lepton channels. All selections require at least 2  $b$ -tagged central jets, used to define the invariant mass  $m_{bb}$ . For the  $W \rightarrow l\nu$  selections events with more than 3 central and forward jets are discarded.

To calculate the reconstructed top quark mass,  $m_{\text{top}}$ , an estimate of the four-momentum of the neutrino from the  $W$  boson decay is obtained, considering the neutrino’s transverse momentum components identical as the  $E_T^{\text{miss}}$  vector component, and constraining  $p_z^{\nu}$  from the  $W$  mass. Further details on 1-lepton and 2-lepton channels are as follows:

The  $Z \rightarrow ll$ -channel selection is further defined by:

- exactly 2 same-flavour (and opposite charge for muons) leptons  $p_T > 7$  GeV and  $|\eta| < 2.5$ , of which at least one with  $p_T > 25$  GeV

	$R_R^{ZW}$	stat. ( $\mathcal{L}/\text{fb}^{-1}$ )			syst. ( $p_{WZ}$ )		
		36.1	300	3000	0	0.5	1
all $m_{VH}$	1.32	$\pm 0.88$	$\pm 0.31$	$\pm 0.10$	$\pm 0.79$	$\pm 0.58$	$\pm 0.22$
high $m_{VH}$	1.51	$\pm 1.43$	$\pm 0.50$	$\pm 0.10$	$\pm 0.79$	$\pm 0.58$	$\pm 0.22$

Table IV.1: Results for the double ratio  $R_R^{ZW}$  following the analysis of [876]. The statistical uncertainty is shown for integrated luminosities of 36.6, 300, and 3000  $\text{fb}^{-1}$ , and the systematic uncertainty for fully uncorrelated, moderately-correlated, and fully correlated uncertainties in  $WH$  and  $ZH$  production, respectively ( $p_{WZ} = (0, 0.5, 1.0)$ ). In the second line, the  $VH$  invariant mass was restricted to  $m_{VH} \in (300, 600)$  GeV. The systematic uncertainties are assumed to be unchanged by this restriction.

- lepton invariant mass  $81 \text{ GeV} < m_{ll} < 101 \text{ GeV}$

The  $W \rightarrow l\nu$ -channel selection is further defined by:

- exactly 1 lepton with  $p_T > 25 \text{ GeV}$  and  $|\eta| < 2.5$
- $p_T^W > 150 \text{ GeV}$
- $E_T^{\text{miss}} > 30 \text{ GeV}$  in the electron sub-channel
- $m_{bb} > 75 \text{ GeV}$  or  $m_{\text{top}} \leq 225 \text{ GeV}$

Event samples at 13 TeV have been generated using MG5\_aMC@NLO [88, 877], interfaced to the general-purpose event generator HERWIG 7 [584, 700, 878–880] for the parton shower and non-perturbative effects, such as hadronization and multiple-parton interactions. No smearing was applied to the momenta of jets or leptons entering the analysis and no backgrounds originating from mis-tagging of charm jets or light jets were considered.

For simplicity, we concentrate on the “dijet-mass analysis” of [876], where the  $BDT_{VH}$  discriminant is replaced by the  $m_{bb}$  variable. This results in 10 signal regions, shown in the second and third rows of Table 12 in [876]. In the present analysis, we have only included signal regions with  $p_{T,V} > 150 \text{ GeV}$ . We have verified that the signal and background event yields are in reasonable agreement at the “selection” level with those of [876]. We have calculated the double ratio by constraining the invariant mass of the two  $b$ -jets to lie within  $m_{bb} \in [110, 140] \text{ GeV}$ .

### 4.3.1 Results

Table IV.1 shows results for the double ratio  $R_R^{ZW}$  after applying the analysis, as well as rough estimates of the statistical and systematic uncertainties. The ggF component of  $ZH$  was rescaled by a flat  $K$ -factor of  $K = 2$  to account for higher-order QCD corrections that are presently unavailable. The systematic uncertainty is given assuming a correlation between the individual  $ZH$  and  $WH$  sample systematic uncertainties of  $p_{WZ} = (0, 0.5, 1.0)$ , i.e. fully uncorrelated, moderately-correlated, and fully correlated, respectively. We show results for integrated luminosities of  $\mathcal{L} = 36.1, 300, 3000 \text{ fb}^{-1}$ , corresponding to Ref. [876], the LHC before LS3, and the HL-LHC, respectively.

It is evident from Table IV.1 that a measurement of  $R_R^{ZW} \neq 1$ , demonstrating the presence of a ggF component of  $ZH$ , will eventually be dominated by systematic uncertainties. However, the analysis of [876] has not been optimised to maximize the performance of  $R_R^{ZW}$ . Indeed,

if one restricts the range of the invariant mass of the reconstructed vector boson and Higgs boson system to  $m_{VH} \in (300, 600)$  GeV, the ratio increases to  $R_R^{ZW} \simeq 1.51$ , with an associated increased statistical error, as shown in the second line of Table IV.1. Such a cut is not applied in [876], but we expect its effect on the systematic uncertainty to be rather minor and leave it unchanged for the current discussion. We also stress that the numbers of Table IV.1 do not take into account any reduction of the systematic uncertainties to be expected from future improved analysis methods.

#### 4.4 Conclusions

We have presented rough results on the feasibility of measurement of the double ratio  $R_R^{ZW}$  in the  $VH$  process, defined by Eq. (IV.5). By exploiting this potential measurement, we have examined the prospects of constraining the gluon-fusion component of  $ZH$  production. Our preliminary estimates of statistical and systematic uncertainties, obtained through a hadron-level Monte Carlo analysis, indicate that both an optimised experimental analysis and a detailed investigation of systematic uncertainties would facilitate extraction of ggF-induced  $ZH$  production. This would allow constraints to be placed on several new physics models affecting ggF  $ZH$ .

#### Acknowledgements

AP acknowledges support by the ERC grant ERC-STG-2015-677323. CP acknowledges support by the University of Geneva. RH and JK are supported by BMBF under contract 05H15PACC1.

### 5 Parton shower and top quark mass effects in Higgs boson pair production <sup>5</sup>

We study the effects of various parton shower approaches on distributions relevant for Higgs boson pair production measurements. In particular, we investigate the behaviour of the Higgs boson pair transverse momentum spectrum which is particularly sensitive to both parton shower and top quark mass effects.

#### 5.1 Introduction

In order to obtain fully exclusive descriptions of final states, and to improve the accuracy of theoretical predictions in regions of the phase space where the convergence of the perturbative series is spoiled by large soft and collinear logarithms, matching Next-to-Leading Order (NLO) predictions to parton shower (PS) Monte Carlo programs is desirable and mostly represents the current state of the art. Parton shower Monte Carlo programs describe multiple soft and collinear emissions by resumming at least the leading tower of logarithms to all orders in perturbation theory. There are several different schemes which allow to match a fixed order calculation to a PS without spoiling the NLO accuracy of the former. One of the aspects which distinguishes them from each other is a different treatment of corrections which are formally of higher order, suppressed by higher powers of the strong coupling. It is however well known that these formally subleading terms can lead to large differences in the final prediction, especially for processes with large higher order corrections. A typical example which was already studied extensively in the past is the production of a Higgs boson in gluon fusion [881, 882].

More recently Higgs boson pair production in gluon fusion was computed at NLO accuracy with full top quark mass dependence [65, 66]. In a second step this was matched to different PS Monte Carlo programs [276, 277] following the three different matching schemes as implemented in the POWHEG-BOX [350, 351], MG5\_aMC@NLO [88, 502] and Sherpa [415]. The aim of the present

---

<sup>5</sup> G. Heinrich, S. P. Jones, M. Kerner, S. Kuttimalai, G. Luisoni

study is to identify where the differences between the various PS Monte Carlo descriptions come from and to assess how they depend on the matching and shower parameter settings.

## 5.2 MC@NLO Parton Shower Matching

In this report we will consider the two most widely used NLO parton shower matching schemes, POWHEG [849] and MC@NLO [503]. In order to point out the most relevant differences between them it is both sufficient and notationally convenient to consider the special case where the splitting kernels appearing in the fixed order subtraction terms  $D$  are identical to the splitting kernels used for parton shower evolution. In the MC@NLO formalism, the fixed-order NLO cross section is split up in two parts, one to be integrated over the born phase space  $\phi_B$  and one to be integrated over the real emission phase space  $\phi_R = \phi_B \times \phi_1$ . In terms of the leading order contributions  $B$ , the virtual corrections  $V$ , and the real emission corrections  $R$ , the seed cross sections for the two contributions are respectively given by

$$\bar{B}(\phi_B) = B(\phi_B) + V(\phi_B) + \int D(\phi_R) \Theta(\mu_{\text{PS}}^2 - t(\phi_R)) d\phi_1 \quad (\text{IV.10})$$

$$H(\phi_R) = R(\phi_R) - D(\phi_R) \Theta(\mu_{\text{PS}}^2 - t(\phi_R)) , \quad (\text{IV.11})$$

where  $t$  is the parton shower evolution variable and  $\mu_{\text{PS}}$  is the parton shower starting scale. Events of both types are then dressed with additional QCD radiation through a parton shower algorithm. Making this explicit in the hardest emission off the  $\bar{B}$  events, we thus have

$$\sigma_{\text{NLO+PS}} = \int \bar{B}(\phi_B) \left[ \Delta(t_0, \mu_{\text{PS}}^2) + \int \Delta(t, \mu_{\text{PS}}^2) \frac{D(\phi_B, \phi_1)}{B(\phi_B)} \Theta(\mu_{\text{PS}}^2 - t) \Theta(t - t_0) d\phi_1 \right] d\phi_B \quad (\text{IV.12})$$

$$+ \int H(\phi_R) d\phi_R, \quad (\text{IV.13})$$

with the Sudakov form factor  $\Delta(t_0, t_1) = \exp \left[ - \int_{t_0}^{t_1} \frac{D(\phi_R)}{B(\phi_B)} d\phi_1 \right]$  and the infrared cutoff scale of the parton shower  $t_0$ .

It is useful to consider the above expression in the kinematic regimes of soft and hard emissions separately. In the soft real emission phase space region where  $t \ll \mu_{\text{PS}}^2$  we have  $D \approx R$  and thus  $H \approx 0$  by construction of the subtraction terms. The remaining contribution is given by the term (IV.12) and gives a LO+PS like result that is rescaled by a local K-factor of  $\bar{B}/B$ . Considering only hard emissions with  $t \approx \mu_{\text{PS}}$ , on the other hand, we can set  $\Delta \approx 1$  and also drop the first term in the square bracket of (IV.12), where no emission occurs at all. After some re-arrangements, this gives

$$\sigma_{\text{NLO+PS}} = \int \left[ \bar{B}(\phi_B) - B(\phi_B) \right] \frac{D(\phi_B, \phi_1)}{B(\phi_B)} \Theta(\mu_{\text{PS}}^2 - t) d\phi_B d\phi_1 \quad (\text{IV.14})$$

$$+ \int R(\phi_R) d\phi_R, \quad (\text{IV.15})$$

where the term on the second line gives the fixed-order result and the term on the first line cancels up to differences between  $\bar{B}$  and  $B$ , which are of higher order in  $\alpha_s$ .

In MC@NLO matching, the scale  $\mu_{\text{PS}}$  therefore separates the real emission phase space in a resummation region that is populated by the parton shower through the  $\bar{B}$  events and a region that is populated mostly by the fixed-order real-emission contributions in  $H$ . Variations of this scale can thus be used in order to assess uncertainties associated with this separation.

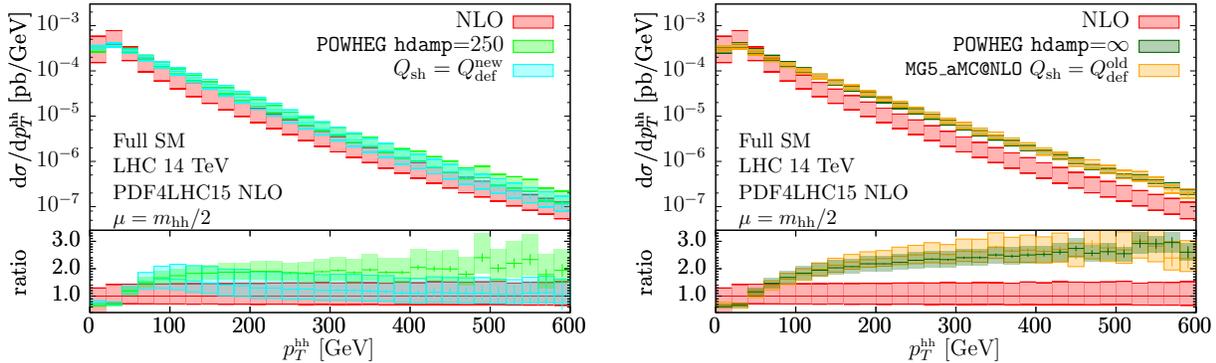


Fig. IV.7: Comparison between POWHEG and MG5\_aMC@NLO results Higgs boson pair transverse momentum distribution. Left: results based on the new definition of the shower starting scale  $Q_{\text{sh}}$  (MG5\_aMC@NLO version  $\geq 2.5.3$ ), compared to POWHEG with  $\text{hdamp}=250$ . Right: results based on the shower starting scale  $\mu_{\text{PS}}$  before version 2.5.3, compared to POWHEG with  $\text{hdamp}=\infty$ .

### 5.3 POWHEG Parton Shower Matching

In the formulation presented above, the POWHEG method can be understood as the limit in which  $\mu_{\text{PS}} \rightarrow \infty$  and  $D = R$ . This leads to  $H = 0$ . All real emission contributions are thus generated by parton shower emission off  $\bar{B}$  events. Taking  $\mu_{\text{PS}} \rightarrow \infty$  ensures that the full real-emission phase space is covered while setting  $D = R$  in the first emission ensures that the fixed-order radiation pattern is recovered. This method comes with the benefit that no negative event weights occur (in MC@NLO, the event weight (IV.11) can become negative). However, setting  $D = R$  also leads to the exponentiation of the full real-emission corrections in the Sudakov form factor. This is in general not justified since  $R$  contains hard, non-factorizing contributions. Instead of setting  $D = R$ , it has therefore been suggested in Ref. [881] to use

$$D = \frac{h_{\text{damp}}^2}{p_T^2 + h_{\text{damp}}^2} R, \quad (\text{IV.16})$$

where  $p_T$  is the transverse momentum of the Born final state ( $p_T = p_T^{\text{hh}}$  in our case). This implements a phase space separation as it is achieved in terms of the parton shower evolution variable in MC@NLO. Effectively, the damping term  $\frac{h_{\text{damp}}^2}{p_T^2 + h_{\text{damp}}^2}$  assumes the role of the Heaviside theta functions in Eqs. (IV.10) and (IV.11).

### 5.4 POWHEG and MG5\_aMC@NLO results

In Fig. IV.7 we clearly see the effect of damping the radiation as explained above by using a value for  $\text{hdamp}=\infty$  in POWHEG and by using a lower shower starting scale  $Q_{\text{sh}} = \mu_{\text{PS}}$  in the MC@NLO approach<sup>6</sup>. However, even with  $\text{hdamp}=250$ , the effects in the tail of the  $p_T^{\text{hh}}$  distribution are surprisingly large. This is why we made further investigations in order to understand its origin. First of all, we should mention that the  $p_T^{\text{hh}}$  distribution is particularly sensitive to extra radiation because NLO is the first order where a non-zero  $p_T^{\text{hh}}$  is generated at the level of the hard matrix element. Therefore, it is not too surprising that the extra emission generated by the parton shower strongly affects this distribution. However, the large effects are not expected to extend into a region where resummation should not play a role. Figure IV.8 shows that indeed the impact of the parton shower is small for other distributions

<sup>6</sup>In version 2.5.3 of MG5\_aMC@NLO, the shower starting scale is picked with some probability distribution to be in the interval `shower_scale_factor`  $\times [0.1 H_T/2, H_T/2]$  with  $H_T$  computed with Born kinematics, while previously it was picked in the interval `shower_scale_factor`  $\times [0.1 \sqrt{\hat{s}}, \sqrt{\hat{s}}]$ .

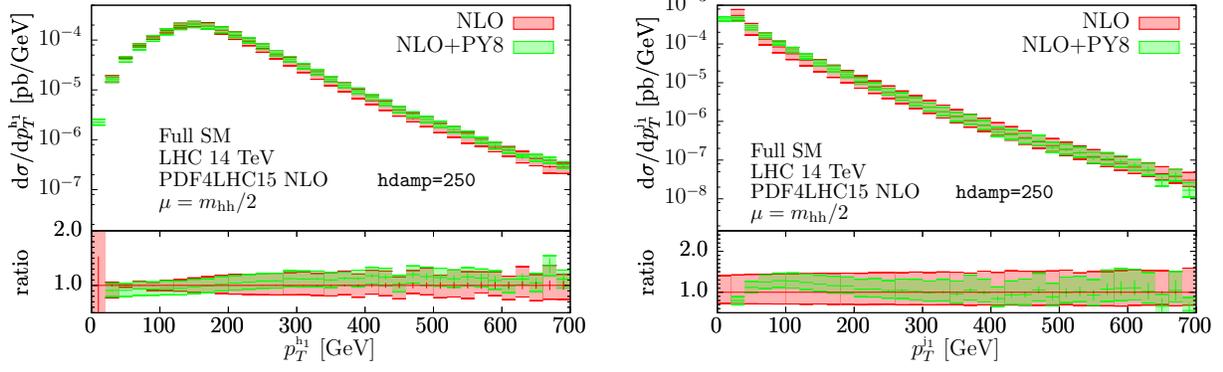


Fig. IV.8: Comparison between fixed order and POWHEG +Pythia 8 results. Left: transverse momentum distribution of the leading- $p_T$  Higgs boson. Right: transverse momentum distribution of the leading jet.

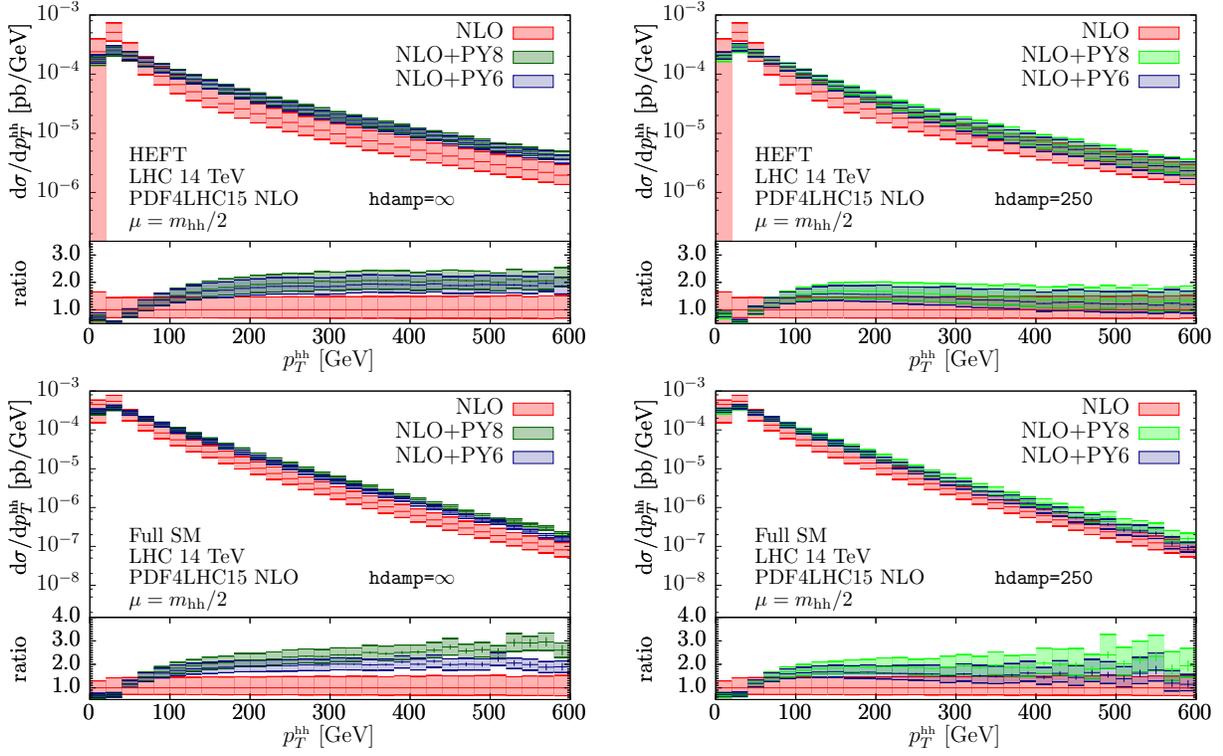


Fig. IV.9: Higgs boson pair transverse momentum distribution  $p_T^{hh}$ : left column with  $\text{hdamp}=\infty$ , right column with  $\text{hdamp}=250$ . We compare the fixed order result with showered results from both Pythia 6 and Pythia 8 in the basic HEFT approximation (upper row) and in the full SM (lower row).

like the transverse momentum of the leading Higgs boson, and even for the transverse momentum distribution of the leading jet. As the Higgs boson pair recoils against this jet in the fixed order calculation, while in the showered results more jets can be present, we again conclude that the enhancement in the tail of the  $p_T^{hh}$  distribution must be related to the fact that the shower generates a lot of radiation. From Eq. (IV.14) we can at least partly trace this back to the fact that  $\bar{B} - B$  is large for Higgs boson pair production, with an NLO K-factor of about 1.6. However, this cannot be the only reason, as we will see in the following.

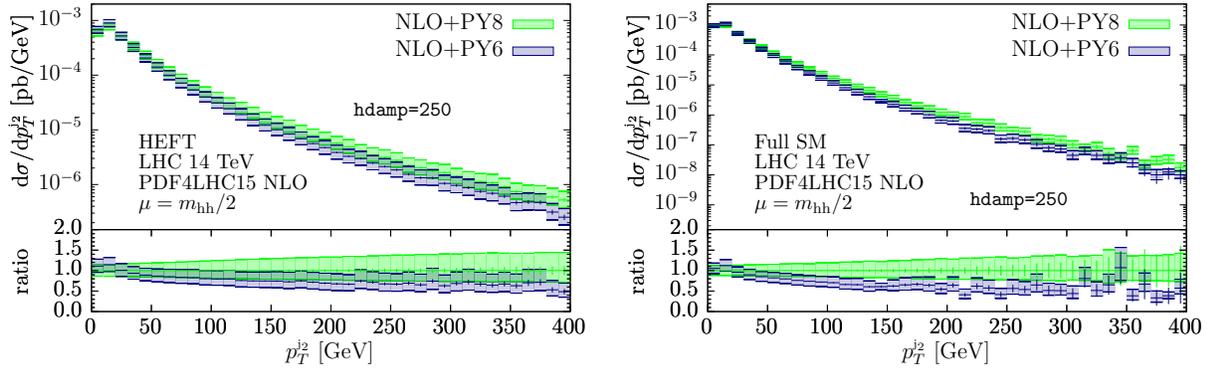


Fig. IV.10: Subleading jet transverse momentum distribution  $p_T^{j2}$  as obtained with Pythia8 and Pythia6 with  $hdamp=250$ . Left: basic HEFT approximation. Right: Full top quark mass dependence.

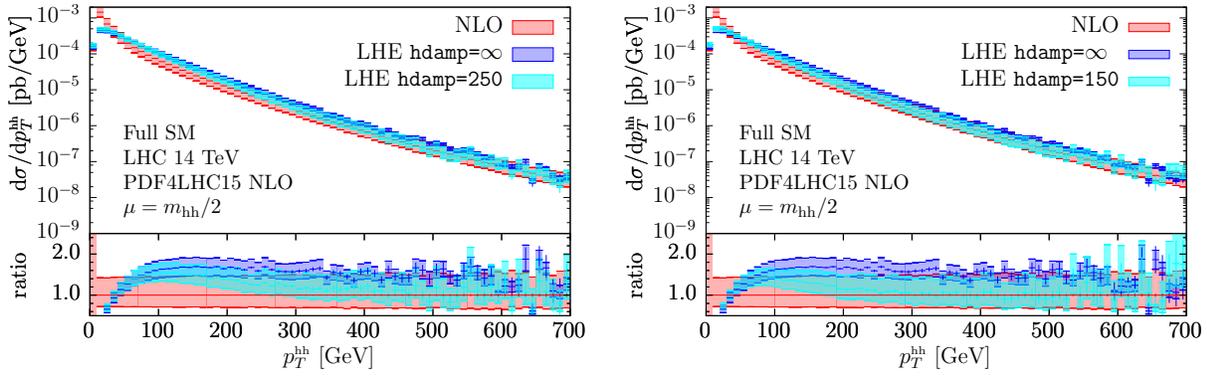


Fig. IV.11: Higgs boson pair transverse momentum distribution  $p_T^{hh}$  at LHE level within the basic HEFT approximation: left column with  $hdamp=250$ , right column with  $hdamp=150$ .

In Fig. IV.9 we compare the fixed order results to predictions obtained with the Pythia6 shower and the Pythia8 shower both in the basic HEFT approximation and in the full SM, for  $hdamp=250$  and  $hdamp=\infty$ . It is interesting to see that with Pythia6, the enhancement in the tail of the  $p_T^{hh}$  distribution is much less pronounced, which means that the hardness of the radiation recoiling against the Higgs-pair system is higher in Pythia8 than in the older Pythia6. Although this behaviour is known [708, 883], it has a particularly large impact in the case considered here<sup>7</sup>. To exhibit very clearly the difference in the hardness of the radiation generated by the two showers, in Fig. IV.10 we compare the transverse momentum spectrum of the subleading jet  $p_T^{j2}$  in basic HEFT and in the full theory. This jet is of pure shower origin, since the matrix elements in our calculations can describe only one jet. Above 200 GeV the difference between the two reaches a factor of 2 for predictions in the full SM. Even more evidence is given in Fig. IV.11, which compares POWHEG-BOX predictions at the Les Houches event level (LHE) for two different values of  $hdamp$ . These plots corroborate the hypothesis that the enhancement of the tail is not solely caused by the POWHEG matching procedure, since at the LHE level the tail of the distribution touches down on the fixed order result for  $hdamp=150$ ,  $hdamp=250$  and even for  $hdamp=\infty$  in the shown  $p_T^{hh}$  range. This is not the case after showering with Pythia8, as can be seen from Fig. IV.7.

<sup>7</sup>Very recent developments in Pythia8 [884] are likely to soften the observed behaviour in the tail of the  $p_T^{hh}$  distribution. We thank Stefan Prestel for pointing this out to us.

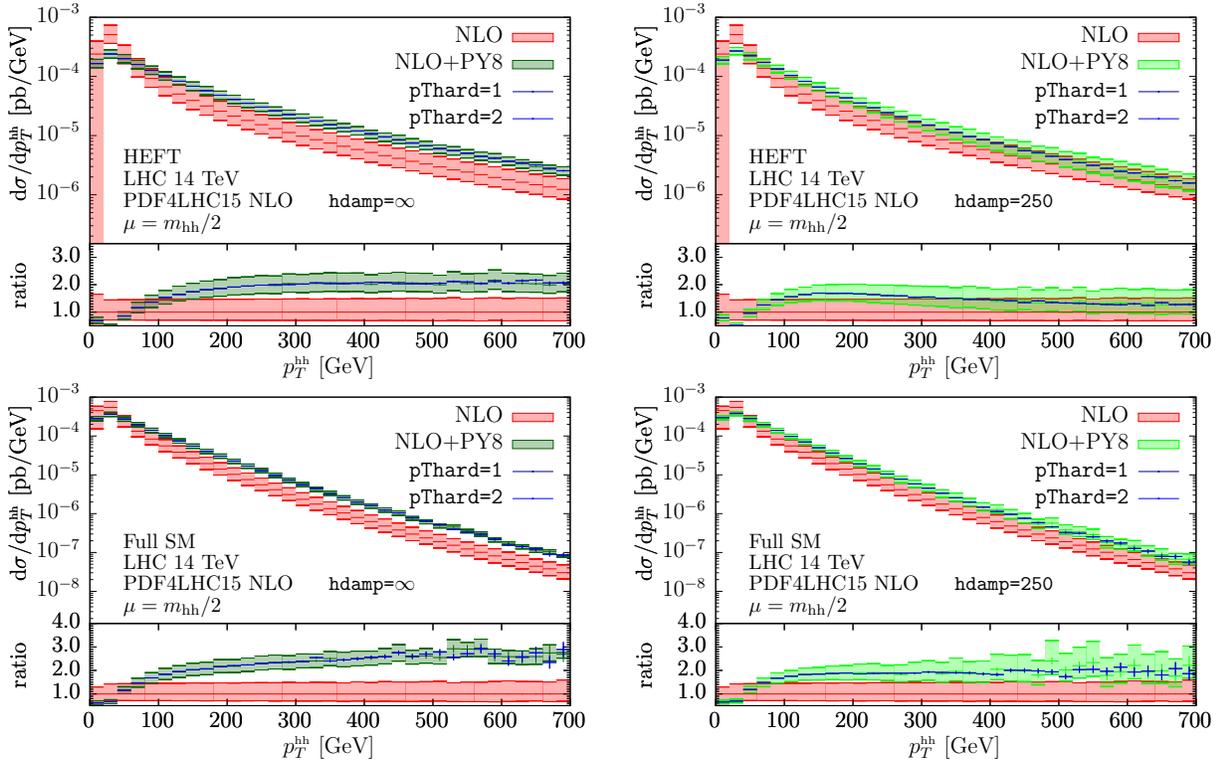


Fig. IV.12: Higgs boson pair transverse momentum distribution  $p_T^{\text{hh}}$  with variations of the parameter  $p_T^{\text{hard}}$ . The default corresponds to  $p_T^{\text{hard}}=0$ .

Therefore we investigated further the influence of the Pythia8 settings by varying the parameters `SCALUP` and `pThard`. Both these parameters control the hardness criterion. By default the POWHEG-BOX transfers this value in the `SCALUP` member of Les Houches Events. When `pThard=0` the default `SCALUP` value is used, leading to the results already discussed. In order to estimate the uncertainty due this choice it is possible to change the value of the `pThard` flag to be `pThard=1` or `pThard=2`. In the former case the transverse momentum of the POWHEG emission is tested against all other incoming and outgoing partons, and the minimal value is chosen, in the latter case the  $p_T$  of all final-state partons is tested against all other incoming and outgoing partons, and the minimal value is selected. Since in Higgs boson pair production there is only one coloured final state particle, and this coincides with the POWHEG emission, the two options lead to identical results, as shown in Fig. IV.12. The two curves where  $p_T^{\text{hard}} \neq 0$  are exactly superimposed. Changing `pThard` does not lead to significant changes because, as already mentioned above, in this process only one coloured final state parton is present, and hence there is only a single choice for a  $p_T$  reference value.

Another possibility is instead to keep the default value `pThard=0` and modify the `SCALUP` value within the POWHEG-BOX using the `changescalup` and `maxscalup` flags. In Fig. IV.13 we show results where we set `maxscalup=50`. This means that the shower starting scale in Pythia8 is lowered to be less than 50 GeV, irrespective of the hardness of the first POWHEG emission. To verify this we can look at the transverse momentum of the subleading jet, shown in Fig. IV.14, which is purely generated by shower emissions. The plots clearly show a sharp decrease at 50 GeV. Comparing the corresponding transverse momentum spectra of the Higgs-pair system we note a qualitative difference between basic HEFT predictions, shown on the left in Fig. IV.13, and full theory predictions, shown in the right plot in Fig. IV.13. While in the effective theory predictions the change in `maxscalup` is barely visible for both values of `hdamp`, in the full

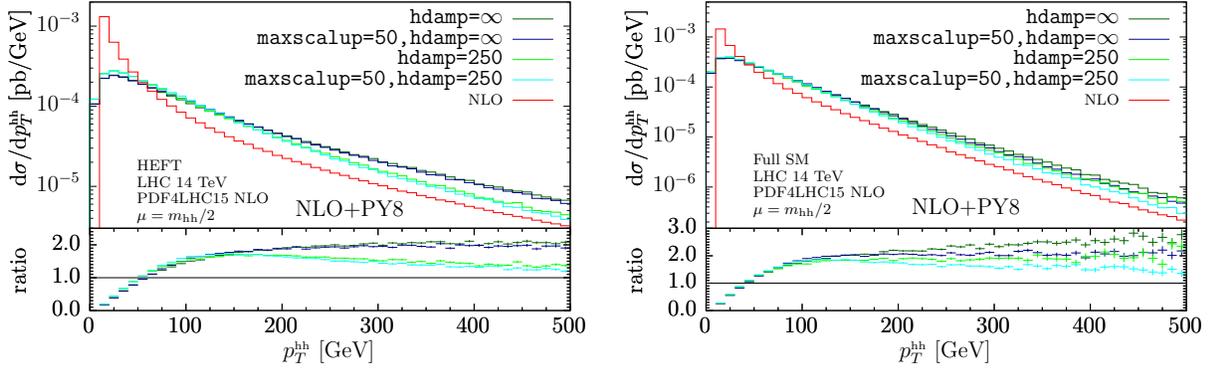


Fig. IV.13: Higgs boson pair transverse momentum distribution  $p_T^{\text{hh}}$  with variations of the `maxscalup` parameter in `Pythia8`. Left: basic HEFT approximation, Right: Full top quark mass dependence.

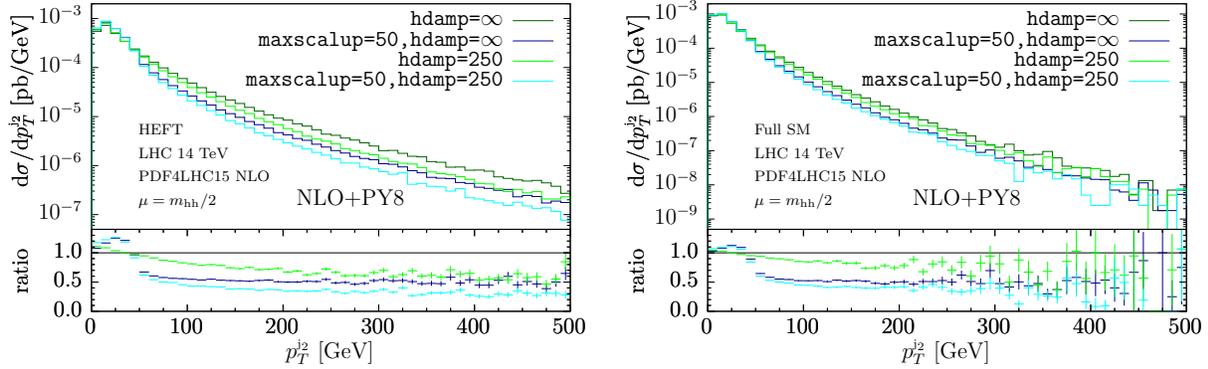


Fig. IV.14: Higgs boson pair transverse momentum distribution  $p_T^{j2}$  with variations of the `maxscalup` parameter in `Pythia8`. Left: HEFT approximation, Right: Full top quark mass dependence.

theory the curves start to deviate around 200 GeV. In both plots we also show fixed order NLO predictions, which we use as a reference in the ratio plots displayed in the lower inset. Imposing a smaller upper limit on the shower starting scale seems to cure the problem partially, at least in the full theory predictions, bringing the NLO+PS curves more in agreement with the fixed order NLO ones at large transverse momenta.

## 5.5 SHERPA results

Using an implementation in the `Sherpa` event generator it was shown in [277] that the large parton shower effects in the tail of the Higgs boson pair transverse momentum distribution can be traced back to the formally subleading term (IV.14). Here we will repeat the main arguments made in [277] and show the corresponding results obtained from the `Sherpa` implementation.

As stated above, in the full theory the term (IV.14) can become numerically large compared to the fixed-order result that is given by (IV.15). This is due to the fixed-order spectrum rapidly dropping above scales of the order of the top quark mass. The impact of the uncanceled parton shower contributions in (IV.14) therefore becomes large relative to the fixed-order prediction. In the HEFT approximation the fixed-order spectrum is considerably harder and the parton shower effects in the tail are thus of smaller significance.

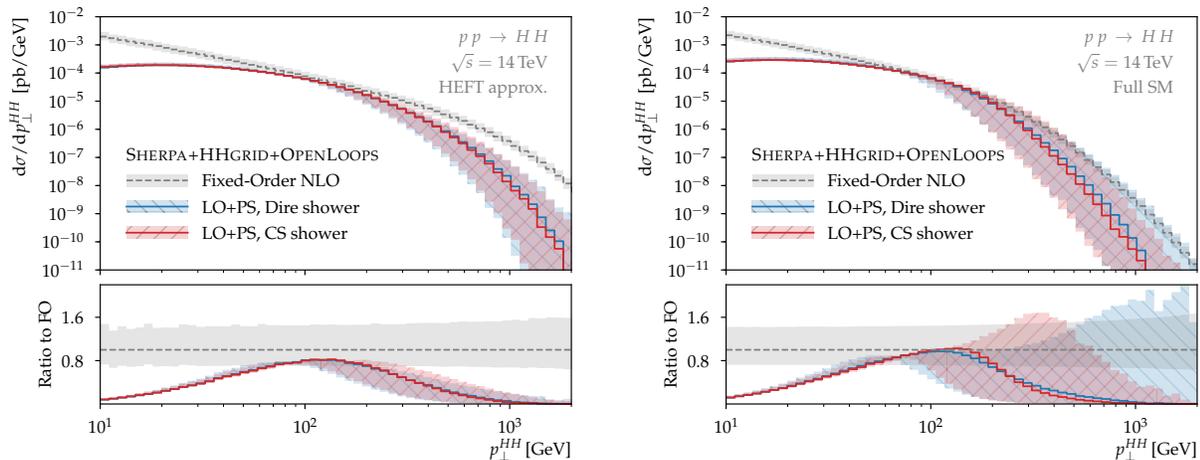


Fig. IV.15: LO+PS results for the Higgs pair transverse momentum spectrum in the HEFT approximation (left panel) and in the full theory (right panel). The results are compared to the fixed-order prediction shown with an uncertainty band that corresponds to variations of  $\mu_f$  and  $\mu_r$  by factors of two. The bands around the LO+PS results are obtained through variations of the parton shower starting scale  $\mu_{\text{PS}}$  by factors of two.

This effect is illustrated in Fig. IV.15, where the fixed-order transverse momentum spectrum is compared to predictions derived from LO+PS type simulations. The uncertainty bands on the LO+PS simulation correspond to variations of the parton shower starting scale  $\mu_{\text{PS}}$  by factors of 2. The central scale choice in case of the Dire shower [492] is  $\mu_{\text{PS}} = m_{\text{hh}}/4$ , whereas  $\mu_{\text{PS}} = m_{\text{hh}}/2$  is set for the CS shower [491]. As expressed by the Heaviside theta function in (IV.12), the parton shower starting scale implements a phase space restriction that restricts parton shower emissions to the phase space region where  $t < \mu_{\text{PS}}^2$ . If the hard region of phase space is made accessible to the parton shower by choosing a high parton shower starting scale (represented by the upper edge of the uncertainty band), both parton showers overestimate the fixed-order real-emission result in the tail of the distribution in the full theory. This is not the case in the HEFT approximation. In this approximation, the fixed-order spectrum is considerably harder and the parton shower therefore underestimates the fixed-order spectrum in the tail.

In MC@NLO, the parton shower effects in the tail are removed at order  $\alpha_s$  relative to the born through the modified subtraction prescription, i.e. by the term  $-D(\phi_R)\Theta(\mu_{\text{PS}}^2 - t(\phi_R))$  in Eq. (IV.11). The higher-order remainder (IV.14) that is left over is, however, relatively large due to the large numerical difference between  $\bar{B}$  and  $B$ . This is shown in Fig. IV.16. As indicated by the upper uncertainty bands, the left-over parton shower effects in the tail can be of order one if a high shower starting scale is chosen in the full theory. For more moderate choices, however, the fixed-order result is reproduced at large transverse momenta. In the HEFT approximation, parton shower effects are much smaller. This is to be expected, since, as shown in Fig. IV.15, the fixed-order spectrum is extremely hard. Even though  $\bar{B} - B$  is numerically large in the HEFT as well, any parton shower effects thus appear small when compared to the fixed-order spectrum.

## 5.6 Conclusions

We have studied parton shower effects on the Higgs boson pair transverse momentum spectrum based on NLO calculations matched to several parton shower Monte Carlo programs. Since these effects were found to be surprisingly large, we investigated in more detail the cause of the

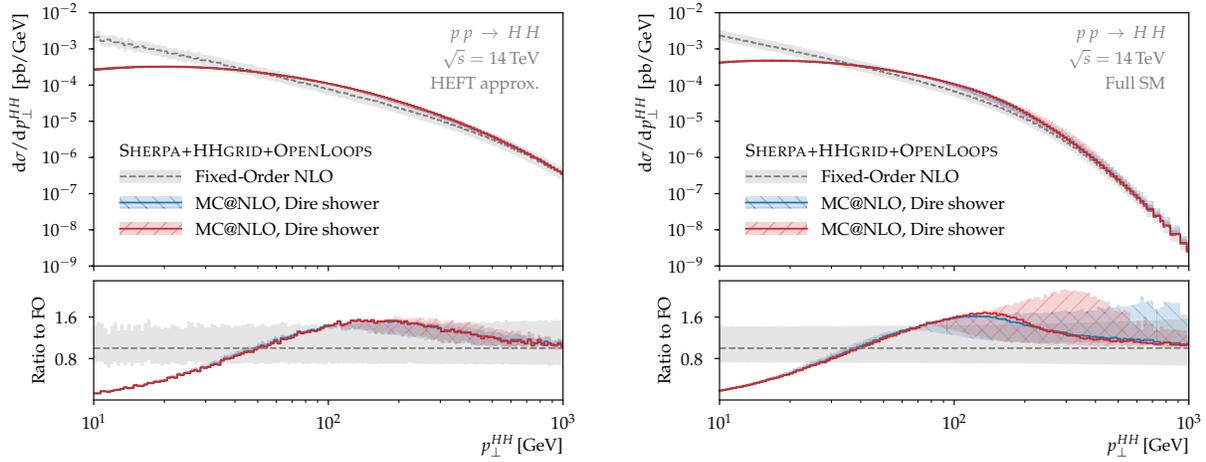


Fig. IV.16: MC@NLO results for the Higgs pair transverse momentum spectrum in the HEFT approximation (left panel) and in the full theory (right panel). The results are compared to the fixed-order prediction shown with an uncertainty band that corresponds to variations of  $\mu_f$  and  $\mu_r$  by factors of two. The bands around the MC@NLO results are obtained through variations of the parton shower starting scale  $\mu_{\text{PS}}$  by factors of two.

enhancement in the tail of the  $p_T^{\text{hh}}$  spectrum, comparing results in basic HEFT and in the full SM, obtained with three different matching schemes and various different shower settings.

We identified the following reasons for the growth of the  $p_T^{\text{hh}}$  tail in NLO+PS matched predictions. As can be seen from Eq. (IV.14), there are three factors which play a role: large K-factor ( $B - \bar{B}$ ), large splitting kernel, and the shower starting scale in the case of MC@NLO matching. The large NLO K-factor is certainly given for the process under consideration, further the contribution from the splitting kernel also seems to be large relative to the hard real radiation.

Depending on the matching scheme, the largeness of the (formally subleading) terms in Eq. (IV.14) can be controlled by tuning the shower starting scale or by separating the hard real radiation phase space from the Sudakov region. On the other hand, we observed important differences between various parton shower programs while using the same matching procedure. For example, the radiation pattern produced by the `Pythia8` and `Pythia6` showers seems to be quite different. These differences are harder to control without an in-depth knowledge of the parton shower programs. They are however a manifestation of the theoretical uncertainty related to the showering, which should be taken into account when estimating the total theoretical uncertainty related to an NLO+PS prediction.

## Acknowledgements

We are grateful to Stefan Prestel for useful comments about `Pythia`. We also thank the organizers of the “LH2017 Physics at TeV Colliders” workshop for creating a productive and pleasant environment. This research was supported in part by the Research Executive Agency (REA) of the European Union under the Grant Agreement PITN-GA2012316704 (HiggsTools) and by the U.S. Department of Energy under contract number DE-AC02-76SF00515.

## 6 Treatment of theory uncertainties for Simplified Template Cross Sections <sup>8</sup>

### 6.1 Introduction

Simplified Template Cross Sections (STXS) have been adopted as an evolution of the signal strength measurements performed during Run 1 of the LHC. Their goal is reduce the theoretical uncertainties that are directly folded into the measurements and provide more finely-grained measurements, while at the same time allowing and benefiting from the combination of measurements in many decay channels. For a detailed discussion, see Sec. III.3 of Ref. [1] and Sec. III.2 of Ref. [236].

The primary features to achieve these goals are as follows. First, they are cross sections (instead of signal strengths) defined in mutually exclusive regions of phase space. The bin definitions are abstracted and simplified compared to the exact fiducial volumes of specific analyses in different Higgs decay channels. The measurements are unfolded to these signal regions, the STXS “bins”, which are common for all analyses, allowing for a subsequent global combination of different decay channels as well as measurements from ATLAS and CMS. In particular, the STXS are defined inclusively in the Higgs boson decay, to naturally combine the various decay modes. While the bin definitions are simplified to allow for the combination of different decay channels (and also for ease of use), they nevertheless try to be as close as possible to the typical experimental selections to avoid any unnecessary extrapolations. The goal is to allow the use of advanced analysis techniques such as event categorization or multivariate techniques while still keeping the unfolding uncertainties small.

Second, the STXS bins are defined for specific production modes, with the SM production processes serving as kinematic templates. This separation into production modes is an essential aspect to reduce their model dependence, i.e., to eliminate the dependence of the measurements on the relative fractions of the production modes in the SM.

The number of separately measured bins can evolve with time, such that the measurements can become more fine-grained as the size of the available dataset increases. Several “stages” have been defined, and can be further extended. The stage 0 bin definitions essentially correspond to the production mode measurements of Run 1, while stage 1 has been defined as the target for the full Run 2 measurements. At intermediate stages individual bins can be merged and only their sum measured according to the sensitivity of each analysis and decay channel. While the full stage 1 granularity should become possible in the combination of all decay channels by the end of Run 2, individual decay channels will require some bins to be merged for the foreseeable future.

The measured STXS together with the partial decay widths are meant to serve as input for subsequent interpretations. Such interpretations could for example be the determination of signal strengths or coupling scale factors  $\kappa$  (providing compatibility with earlier results), EFT coefficients, tests of specific BSM models, and so forth. The treatment of theory uncertainties when reporting experimental measurements will be discussed in Sec. 6.5.

### 6.2 Overview of Theory Uncertainties

There are two places where theory uncertainties enter, which are important to distinguish. First, there are residual theory uncertainties in the measurements of the STXS bins, resulting from the assumed SM predictions of the kinematic distribution within each bin that enter in the unfolding of the experimental event categories to the STXS bins. Second, there are the theory uncertainties on the SM and beyond predictions of the STXS bins which are required in any subsequent interpretation. One of the primary purposes of the STXS bin definitions is to move the dominant theory uncertainties to the second interpretation step, since they are much easier

---

<sup>8</sup> N. Berger, F. J. Tackmann, K. Tackmann

to deal with there. Here, we are primarily concerned with the theory uncertainties at the interpretation level.

To combine the information from all measured bins in the interpretation, the theoretical predictions and their uncertainties must be evaluated separately for each bin, since different bins will in general contribute with different relative weights to the final result. In this context, the correlations of the theory uncertainties for different bins must be taken into account. This is particularly important whenever a certain binning cut induces an important additional source of perturbative uncertainties that affects each bin but should cancel in their sum. (A well-known example is the case of jet binning, where it is essential to treat the uncertainties induced by the jet-binning cut as anticorrelated between the jet bins [885, 886].)

Hence, to enable such interpretations in a flexible way, it is important to have a common treatment of theory uncertainties across all bins and production modes. The goal is to provide a general parameterization of theory uncertainties, which is physically motivated, in particular allowing to take into account the possible theory correlations between different bins, and is flexible enough to accommodate different (types of) theory predictions. In this way, the underlying uncertainty parameterization can be implemented once and used to easily test different predictions or include future theory improvements. For the gluon-fusion production process, such a parameterization has already been worked out [887]. Here we will discuss similar parameterizations for the vector-boson fusion (VBF) and associated (VH) production processes.

In general, to properly treat the theoretical uncertainties one should try to identify and distinguish different sources of uncertainties and take into account the correlation implied by each common uncertainty source. For the practical implementation, an essential requirement is to parametrize the uncertainties such that they can be implemented in terms of fully correlated and uncorrelated nuisance parameters. This is facilitated by parameterizing the uncertainties in terms of uncertainty sources, each of which is considered common and hence fully correlated among all bins, while the different sources are considered mutually independent and hence uncorrelated. The uncertainties can then be implemented via a single nuisance parameter for each source. For this purpose, we follow the generic parameterization strategy for uncertainties in kinematic bins discussed in Sec. I.4.2.a of Ref. [236], and which is briefly reviewed in Sec. 6.3 below.

It should be stressed that the theory uncertainty parameterization discussed here is not meant and not necessarily suitable to assess the residual theory uncertainties in the measurement step. However, whenever two bins are merged in the measurement, this in general reintroduces the dependence on the SM prediction for the (ratio of the) cross sections of the two merged bins with the resulting theory uncertainty. In this context, a common parameterization of the theory uncertainties between measurement and interpretation is essential to allow for a consistent treatment in the combination of measurements with differently merged bins as well as between measurements and interpretation.

Finally, we note that here we are only concerned with the perturbative theory uncertainties due to missing higher-order corrections. Uncertainties due to the imprecise knowledge of input parameters, like parton distribution functions or quark masses, are not considered here, since their correct treatment including correlations is straightforward.

### 6.3 Theory Uncertainties in Kinematic Bins

In this subsection, we review a general parameterization strategy for uncertainties in kinematic bins. For a more detailed discussion we refer to Sec. I.4.2.a of Ref. [236]. First, consider a single bin boundary  $a/b$  that splits the cross section  $\sigma_{ab} = \sigma_a + \sigma_b$  into two bins with cross sections  $\sigma_a$  and  $\sigma_b$ . In general, the binning cut has a nontrivial influence on the perturbative structure

of  $\sigma_a$  and  $\sigma_b$ , for example, it can introduce sensitivity to an additional energy scale or separate different jet multiplicities. This implies that the binning cut corresponds to an additional and a priori nonnegligible source of uncertainty that is not present in  $\sigma_{ab}$ .

The uncertainty matrix for  $\{\sigma_a, \sigma_b\}$  can be parametrized in terms of fully correlated and fully anticorrelated components as

$$C(\{\sigma_a, \sigma_b\}) = \begin{pmatrix} (\Delta_a^y)^2 & \Delta_a^y \Delta_b^y \\ \Delta_a^y \Delta_b^y & (\Delta_b^y)^2 \end{pmatrix} + \begin{pmatrix} \Delta_{a/b}^2 & -\Delta_{a/b}^2 \\ -\Delta_{a/b}^2 & \Delta_{a/b}^2 \end{pmatrix}. \quad (\text{IV.17})$$

This is a general parameterization of a  $2 \times 2$  symmetric matrix, and hence the uncertainties obtained with any prescription can always be written in this form, provided sufficient information or assumptions on the correlations are available. This parameterization is convenient for two reasons: First, the separation into independent components that are  $\pm 100\%$  correlated between the different bins allows for a straightforward implementation in terms of independent nuisance parameters for each component. That is, we can use two nuisance parameters  $\theta^y$  and  $\theta_{\text{cut}}$ , whose absolute uncertainty amplitudes for  $\{\sigma_{ab}, \sigma_a, \sigma_b\}$  are

$$\theta^y : \quad \{\Delta_{ab}^y, \Delta_a^y, \Delta_b^y\} \quad \theta_{a/b} : \quad \{0, \Delta_{a/b}, -\Delta_{a/b}\}, \quad (\text{IV.18})$$

where  $\Delta_{ab}^y = \Delta_a^y + \Delta_b^y$ . Second, this parameterization admits a simple physical interpretation: The first correlated component with superscript ‘‘y’’ can be interpreted as an overall yield uncertainty of a common source for all bins. The second anticorrelated component can be interpreted as a migration uncertainty between the two bins  $\sigma_a$  and  $\sigma_b$ , which is introduced by the  $a/b$  binning cut and must drop out in their sum. Having such a physical interpretation is very useful to identify and estimate each component in the actual theory calculation. For example, for jet-binning this interpretation has been explicitly exploited to construct uncertainty estimates for both fixed-order and resummed predictions [885, 886, 888].

Consider now the general case of multiple bins, where each bin can have more than one boundary and each boundary can be shared by different bins. To make this tractable in a systematic fashion, the idea is to use Eq. (IV.18) for any given single bin boundary  $a/b$  with all additional subdivisions removed. Then given the bin boundary  $a/b$  separating the cross section as  $\sigma_{ab} = \sigma_a + \sigma_b$ , we can include additional binning cuts that further subdivide  $\sigma_a$  and  $\sigma_b$  as  $\sigma_a = \sum_i \sigma_{ai}$  and  $\sigma_b = \sum_j \sigma_{bj}$ . Since we interpret the  $a/b$  boundary as a common uncertainty source, we can consider it as fully correlated among each set of sub-bins and implement it via a single nuisance parameter  $\theta_{a/b}$ . The corresponding uncertainty amplitudes for all the bins are now given as

$$\theta_{a/b} : \quad \Delta_{a/b} \times \left\{ \{x_{ai}\}, -\{x_{bj}\} \right\} \quad \text{with} \quad \sum_i x_{ai} = \sum_j x_{bj} = 1, \quad (\text{IV.19})$$

where the parameters  $x_{ai}$  and  $x_{bj}$  specify how the absolute uncertainty  $\Delta_{a/b}$  gets distributed among the sub-bins. That is, the impact on  $\sigma_{ai}$  is  $x_{ai}\Delta_{a/b}$  and on  $\sigma_{bj}$  it is  $-x_{bj}\Delta_{a/b}$ . The  $x_{ai}$  and  $x_{bj}$  parameters are specific to each nuisance parameter  $\theta_{a/b}$ .

In this way, we can consider each bin boundary as a potential source of an uncertainty with an associated nuisance parameter. Of course, in practice with sufficiently complicated bin boundaries, we have to apply some theoretical judgment in how to choose the relevant independent binning cuts. In addition, we can have one or more overall yield uncertainties correlated among all bins.

Mathematically speaking, this way of parameterizing the uncertainties has some redundancy. This is desired and makes it flexible enough to accommodate different scenarios while maintaining the simple physical interpretation in terms of the underlying uncertainty sources.

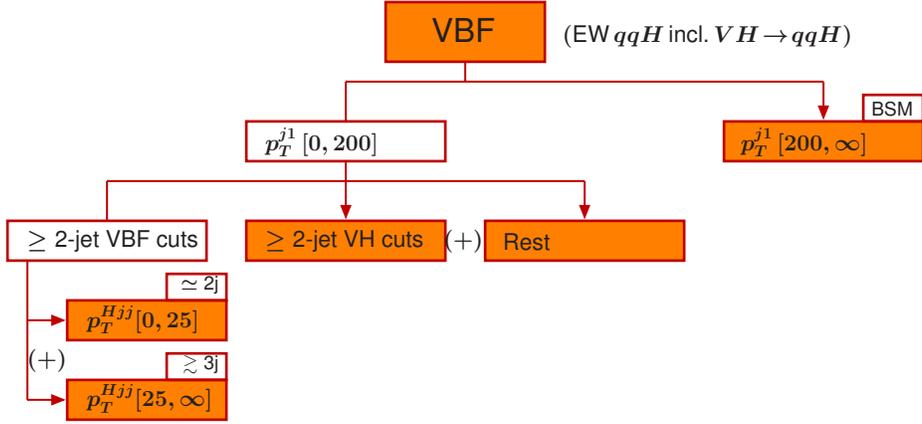


Fig. IV.17: Stage 1 bins for VBF production.

On the other hand, for the case of  $N$  bins, it is also desirable to have at least  $N$  independent nuisance parameters, such that in principle any generic  $N \times N$  uncertainty matrix can be reexpressed in this form.

## 6.4 Parameterization of Theory Uncertainties

To parametrize the perturbative uncertainties, we consider both QCD and EW sources. We give the physical interpretation of each source that should be followed when estimating the uncertainty in practice. While providing explicit uncertainty estimates is beyond the scope here, we will try to indicate in which cases estimates from fixed-order calculations are likely to be appropriate and where resummation (or parton-shower) uncertainties are expected to be relevant. We also point out where additional studies are needed to decide on whether certain sources should be treated as correlated or uncorrelated, or where additional sources might be needed in the future.

We hope that the parameterizations below will provide a sufficiently flexible and somewhat future-proof baseline for the experimental measurements and interpretations of STXS bins at stage 1.

### 6.4.1 VBF Production

The VBF template process is defined as electroweak  $qqH$  production, which includes the usual VBF topology but also  $q\bar{q} \rightarrow VH$  production with hadronic  $V \rightarrow q\bar{q}$  decays. Since they lead to the same final state, they essentially represent the  $t$ -channel and  $s$ -channel contributions to the same physical process, and can only be distinguished by enriching one or the other type of contribution via kinematic cuts.

The full stage 1 bins for vector-boson fusion are depicted in Fig. IV.17. They include a high- $p_T$  and low- $p_T$  bin, defined by a cut on the leading jet to be above or below  $p_T^{j1} = 200$  GeV. The low- $p_T$  bin is further split into three bins targeting the typical VBF and VH topologies and a bin for the rest. The VBF topology cuts are  $m_{jj} > 400$  GeV and  $\Delta\eta_{jj} > 2.8$  with both signal jets required to have  $p_T^j > 30$  GeV. The VH topology cuts are  $60 \text{ GeV} < m_{jj} < 120$  GeV. For our purposes we can essentially consider these as splitting the  $m_{jj}$  spectrum into three regions. (Technically, the Rest bin also includes events where one or zero jets pass the jet selection cuts.) Finally, the VBF topology bin is split into an exclusive 2-jet-like and inclusive 3-jet-like bin using a cut on  $p_T^{Hjj}$  to be above or below 25 GeV.

#### 6.4.1.1 QCD Uncertainties

For the QCD uncertainties, we identify the following sources/nuisance parameters:

- $\theta_{\text{VBF}}^y$ : The primary yield uncertainty for the high  $m_{jj} > 400$  GeV region.
- $\theta_{\text{Rest}}^y$ : The primary yield uncertainty for the low  $m_{jj} < 400$  GeV region.
- $\theta_{\text{VH}}^y$ : The overall yield uncertainty for the underlying VH production process.
- $\theta_{200}$ : The migration uncertainty related to the  $p_T^{j1} = 200$  GeV bin boundary.
- $\theta_{25}$ : The migration uncertainty related to the  $p_T^{Hjj} = 25$  GeV bin boundary.

For the VBF process, an important question that should be studied is how the uncertainties are correlated between the low and high  $m_{jj}$  regions, i.e., between the “VBF cuts” and the “Rest” bin. Here, we took the approach of considering them as primarily uncorrelated with each having their own nuisance parameter  $\theta_{\text{Rest}}$  and  $\theta_{\text{VBF}}$ . In principle, correlations between the bins can be taken into account by allowing these to also impact the corresponding other regions. An alternative would be to use an overall yield uncertainty and a migration uncertainty across the  $m_{jj} = 400$  GeV boundary. However, since the bins differ by more cuts, this seems less appropriate in this case.

For completeness, we have included the  $\theta_{\text{VH}}^y$  nuisance parameter here. It is the same parameter entering for the VH process below, such that the uncertainties for the hadronic and leptonic VH process are correlated. It will primarily impact the VH cuts bin, providing the dominant uncertainty there. In principle, it could also have a small impact on the other bins as well. In general, all the VH nuisance parameters discussed below can have some impact here. For example, the high  $p_T^j$  bin can have contributions from the boosted regime of VH production, where the hadronically decaying boosted vector boson produces a high- $p_T$  jet.

The overall impacts of the yield uncertainties ( $\Delta_{\text{VBF}}^y$ ,  $\Delta_{\text{Rest}}^y$ ) and of the  $\theta_{200}$  migration uncertainty ( $\Delta_{200}$ ) can be evaluated based on fixed-order uncertainties. Their separation into the individual bins for most bins could be evaluated from fixed-order predictions or based on Monte-Carlo predictions. (If different predictions are utilized to estimate the overall size and the separation into the bins, one must be careful to check that both predictions are consistent with each other.) The  $p_T^{Hjj}$  cut is in principle sensitive to resummation effects. Therefore, the size of the  $\theta_{25}$  migration uncertainty ( $\Delta_{25}$ ) as well as the separation of the other uncertainties within the  $p_T^{Hjj}$  bins should preferably be evaluated (or at least cross checked) based on parton-shower Monte Carlos or resummed calculations.

#### 6.4.1.2 EW Uncertainties

For the EW uncertainties, we identify the following sources/nuisance parameters:

- $\theta_{\text{Sud}}$ : The uncertainty related to EW Sudakov effects.
- $\theta_{\text{hard}}$ : The uncertainty related to hard EW (non-Sudakov) effects.

The reason to separate these two sources is that they correspond to structurally different types of EW corrections, and hence can be considered as largely uncorrelated. The Sudakov effects refer to the appearance of EW logarithms from the virtual exchange of EW gauge bosons. They are expected to be most relevant in the high- $p_T^j$  bin, while the non-Sudakov effects are likely to affect all bins. Since the EW uncertainties can be expected to be less relevant than the QCD ones, it should not be necessary to separate them further based on the individual bins,

	QCD uncertainties					EW uncertainties	
	$\Delta_{\text{VBF}}^y$	$\Delta_{\text{Rest}}^y$	$\Delta_{\text{VH}}^y$	$\Delta_{200}$	$\Delta_{25}$	$\Delta_{\text{Sud}}$	$\Delta_{\text{hard}}$
$p_T^{j1} [0, 200]$	$\approx 1$	$\approx 1$	$\approx 1$	$-1$		$y$	$y$
$\geq 2$ -jet VBF cuts	$\approx 1$	$\approx 0$	$\approx 0$	$-x_1$	$0$	$x_1 y$	$x_1 y$
$p_T^{Hjj} [0, 25]$	$(\approx 1)z$	$\dots$	$\dots$	$-x_1 z$	$+1$	$\dots$	$\dots$
$p_T^{Hjj} [25, \infty]$	$(\approx 1)(1-z)$	$\dots$	$\dots$	$-x_1(1-z)$	$-1$	$\dots$	$\dots$
$\geq 2$ -jet VH cuts	$\approx 0$	$\approx 0$	$\approx 1$	$-x_2$		$x_2 y$	$x_2 y$
Rest	$\approx 0$	$\approx 1$	$\approx 0$	$-x_3$		$x_3 y$	$x_3 y$
$p_T^{j1} [200, \infty]$	$\approx 0$	$\approx 0$	$\approx 0$	$+1$		$1 - y$	$1 - y$

Table IV.2: Structure and expected impact of different uncertainty sources for VBF production. Each column correspond to one independent nuisance parameter. The  $\Delta_i$  denote the absolute impact, which are multiplied by the corresponding parameters for each bin. Empty entries mean that a source has by definition no impact on a bin. The  $x_i, y, z$  parameters can be different in each column, and  $\sum_i x_i = 1$ .

and hence they effectively act as overall yield uncertainties correlated among all bins. Their evaluation should be based on the corresponding EW calculations.

The structure and expected impact of the different sources are illustrated in Table IV.2. Here, the  $\Delta_i$  denote the absolute uncertainties, and the  $x_i, y, z$  parameters in the different columns are different parameters, i.e., they are specific to each source (to each column). For notational simplicity, we did not adorn them with additional subscripts. The  $x_i$  satisfy  $\sum_i x_i = 1$ .

### 6.4.2 VH Production

The VH template process is formally defined as Higgs production in association with a leptonically decaying vector boson. It is separated into the three underlying processes  $q\bar{q}' \rightarrow WH$ ,  $q\bar{q} \rightarrow ZH$ , and  $gg \rightarrow ZH$ . The full stage 1 bins are depicted in Fig. IV.18. They include three  $p_T^V$  bins (two for  $gg \rightarrow ZH$ ) with binning cuts at  $p_T^V = 150$  GeV and  $p_T^V = 250$  GeV. The second  $p_T^V$  bin is further separated into an exclusive 0-jet bin and an inclusive  $\geq 1$ -jet bin, with a cut at  $p_T^j = 30$  GeV.

The hadronic VH processes  $q\bar{q} \rightarrow V(\rightarrow q\bar{q})H$  are included as part of the VBF template process. Similarly, the gluon-induced  $gg \rightarrow Z(\rightarrow q\bar{q})H$  is included as part of the  $gg \rightarrow H$  template process (where it corresponds to an electroweak real-emission correction).

#### 6.4.2.1 QCD Uncertainties

For the QCD uncertainties, we identify the following sources/nuisance parameters:

- $\theta_{\text{VH}}^y, \theta_{gg \rightarrow ZH}^y$ : The overall yield uncertainty for the underlying VH production process.
- $\theta_{150}^{\text{VH}}, \theta_{150}^{gg \rightarrow ZH}$ : The migration uncertainty related to the  $p_T^V = 150$  GeV bin boundary.
- $\theta_{250}^{\text{VH}}$ : The migration uncertainty related to the  $p_T^V = 250$  GeV bin boundary.
- $\theta_{0/1}^{\text{VH}}, \theta_{0/1}^{gg \rightarrow ZH}$ : The migration uncertainty related to the 0/1-jet bin boundary.

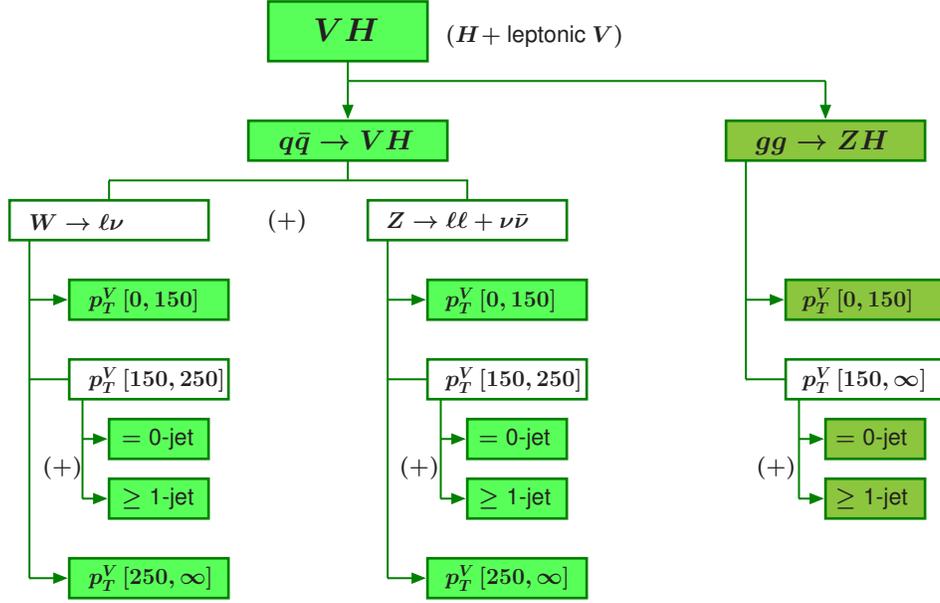


Fig. IV.18: Stage 1 bins for VH production.

These parameters are the same for the  $q\bar{q}' \rightarrow WH$  and  $q\bar{q} \rightarrow ZH$  subprocesses, whose QCD uncertainties are thus treated as correlated. The overall impacts ( $\Delta$ s) for the two subprocesses however do not have to be the same. The corresponding nuisance parameters for  $gg \rightarrow ZH$  are separate, such that the quark-induced and gluon-induced processes are treated as uncorrelated.

The overall impacts of the yield and  $p_T^V$  migration uncertainties ( $\Delta_{\text{WH}}^y$ ,  $\Delta_{\text{ZH}}^y$ ,  $\Delta_{gg \rightarrow ZH}^y$ ,  $\Delta_{150}^{\text{WH}}$ ,  $\Delta_{150}^{\text{ZH}}$ ,  $\Delta_{250}^{\text{WH}}$ ,  $\Delta_{250}^{\text{ZH}}$ ) can be evaluated based on fixed-order uncertainties. Their separation into the individual bins for most bins could be evaluated from fixed-order predictions or based on Monte-Carlo predictions. The 0/1-jet binning cut is likely to be sensitive to resummation effects. Therefore, the size of the  $\theta_{0/1}$  migration uncertainty ( $\Delta_{0/1}^{\text{ZH}}$ ,  $\Delta_{0/1}^{\text{WH}}$ ,  $\Delta_{0/1}^{gg \rightarrow ZH}$ ) as well as the separation of the other uncertainties within the jet bins should preferably be evaluated based on parton-shower Monte Carlos or resummed calculations.

For the  $gg \rightarrow ZH$  process it should be studied to what extent the uncertainties should be correlated with corresponding uncertainties in the  $gg \rightarrow H$  process. For example, most likely it would be most appropriate to identify the  $\theta_{0/1}$  for  $gg \rightarrow H$  and  $gg \rightarrow ZH$ , since the relevant resummation effects for both processes are the same.

#### 6.4.2.2 EW Uncertainties

For the EW uncertainties, we identify the following sources/nuisance parameters:

- $\theta_{\text{Sud}}^{\text{VH}}$ : The uncertainty related to EW Sudakov effects.
- $\theta_{\text{hard}}^{\text{WH}}$ ,  $\theta_{\text{hard}}^{\text{ZH}}$ : The uncertainty related to hard EW (non-Sudakov) effects.

As for the VBF process, we can distinguish EW Sudakov and hard (non-Sudakov) uncertainties. The Sudakov corrections are universal and are thus treated as correlated for WH and ZH production. On the other hand, the hard EW corrections are different with separate nuisance parameters for WH and ZH. One aspect to be studied is to what extent the EW uncertainties should be treated as correlated or uncorrelated between the VBF and VH processes. The naive expectation is that it would be most appropriate to treat the EW Sudakov uncertainties as correlated and the others as uncorrelated.

$qq' \rightarrow W$	QCD uncertainties				EW uncertainties		
	$\Delta_{\text{WH}}^y$	$\Delta_{150}^{\text{WH}}$	$\Delta_{250}^{\text{WH}}$	$\Delta_{0/1}^{\text{WH}}$	$\Delta_{\text{Sud}}^{\text{WH}}$	$\Delta_{\text{hard}}^{\text{WH}}$	
$p_T^V [0,150]$	$x_1$	-1	$-y$		$x_1$	$\dots$	
$p_T^V [150,250]$	$x_2$	$+1 - y$	$-(1 - y)$	0	$x_2$	$\dots$	
= 0-jet	$x_2 z$	$+(1 - y)z$	$-(1 - y)z$	+1	$\dots$	$\dots$	
$\geq 1$ -jet	$x_2(1 - z)$	$+(1 - y)(1 - z)$	$-(1 - y)(1 - z)$	-1	$\dots$	$\dots$	
$p_T^V [250, \infty]$	$x_3$	$y$	+1		$x_3$	$\dots$	
$q\bar{q} \rightarrow Z$	$\Delta_{\text{ZH}}^y$	$\Delta_{150}^{\text{ZH}}$	$\Delta_{250}^{\text{ZH}}$	$\Delta_{0/1}^{\text{ZH}}$	$\Delta_{\text{Sud}}^{\text{ZH}}$		$\Delta_{\text{hard}}^{\text{ZH}}$
$p_T^V [0,150]$	$x_1$	-1	$-y$		$x_1$		$\dots$
$p_T^V [150,250]$	$x_2$	$+1 - y$	$-(1 - y)$	0	$x_2$		$\dots$
= 0-jet	$x_2 z$	$+(1 - y)z$	$-(1 - y)z$	+1	$\dots$		$\dots$
$\geq 1$ -jet	$x_2(1 - z)$	$+(1 - y)(1 - z)$	$-(1 - y)(1 - z)$	-1	$\dots$		$\dots$
$p_T^V [250, \infty]$	$x_3$	$y$	+1		$x_3$		$\dots$

Table IV.3: Structure and expected impact of different uncertainty sources for VH production. Each column correspond to one independent nuisance parameter. The  $\Delta_i$  denote the absolute impact, which are multiplied by the corresponding parameters for each bin. Empty entries mean that a source has by definition no impact on a bin. The  $x_i, y, z$  parameters can be different in each column, and  $\sum_i x_i = 1$ . In addition,  $gg \rightarrow ZH$  production has an independent set of nuisance parameters with the analogous structure of uncertainties.

We do not yet consider EW uncertainties for  $gg \rightarrow ZH$ , since they are not yet relevant (due to the fact that the contribution of the subprocess itself is small with relatively large QCD uncertainties). Once they become relevant, they can easily be added as for the dominant quark-induced subprocesses. We also do not separately distinguish uncertainties due to mixed QCD-EW corrections, which for now should be included as part of the (hard) EW uncertainties. If it becomes necessary, one could add a separate source for these in the future. Uncertainties from photon-induced corrections are also not yet separately considered.

The structure and expected impact of the different sources are illustrated in Table IV.3. As before, the  $\Delta_i$  denote the absolute uncertainties, and the  $x_i, y, z$  parameters in the different columns are different parameters, i.e., they are specific to each source (to each column). The  $x_i$  satisfy  $\sum_i x_i = 1$ . The uncertainties for  $gg \rightarrow ZH$  production are not shown. They have the same structure but with an independent set of nuisance parameters.

## 6.5 Treatment of Theory Uncertainties in Reporting Experimental Results

The STXS measurements are performed with the SM production processes serving as kinematic templates. They can be interpreted in the context of the SM, but can also be recast in the context of beyond-the-SM models to help set constraints on new physics. In order to make optimal use of the measurement information, the interpretation should be performed using the full experimental likelihood.

This is however difficult in practice: full likelihoods are typically not made public by the

experimental collaborations, and also involve a large number of parameters including both the parameters of interest (POIs) and a large number of nuisance parameters (NPs) representing systematic uncertainties. Experimental results are usually reported in terms of the profile likelihood ratio (PLR), in which all nuisance parameters have been profiled. A Gaussian form is further assumed for the PLR, and results are reported in terms of the best-fit values of the POIs and their covariance matrix.

This presentation of results suffers from two limitations: firstly, non-Gaussian effects in the measurements are neglected; and secondly it is not possible to correlate systematic uncertainties in the interpretation stage with the same uncertainties in the experimental measurements, since the NPs representing the uncertainties in the latter case have been profiled in the reported results.

In this section, we present possible directions of improvement for the reporting of experimental results to address the second point.

### 6.5.1 Gaussian measurements with linear parameter response

We start by defining a simplified model in which the measurement is described as  $N$  independent Gaussian observations  $n_i$ , which can represent for instance event counts in  $N$  mutually exclusive signal regions. The corresponding likelihood is

$$L_n(\mathbf{n}) = \exp \left[ -\frac{1}{2}(\mathbf{n} - \hat{\mathbf{n}})^T H_n (\mathbf{n} - \hat{\mathbf{n}}) \right] \quad (\text{IV.20})$$

where  $\hat{\mathbf{n}}$  is the best-fit value of  $\mathbf{n}$  and  $H_n$  is the Hessian matrix, defined as the inverse of the covariance matrix

$$C = \begin{pmatrix} \sigma_1^2 & 0 & \cdots \\ 0 & \ddots & \ddots \\ \vdots & \ddots & \sigma_N^2 \end{pmatrix} \quad (\text{IV.21})$$

We further assume that the expected values of the  $n_i$  are linear functions  $n_i = \sum_{\alpha} k_{i\alpha} x_{\alpha}$  of the model parameters  $x_{\alpha}$ . The measurement of the  $x_{\alpha}$  is then described by the Gaussian likelihood

$$L_x(\mathbf{n}) = \exp \left[ -\frac{1}{2}(\mathbf{x} - \hat{\mathbf{x}})^T H_x (\mathbf{x} - \hat{\mathbf{x}}) \right] \quad (\text{IV.22})$$

where  $\hat{x}$  is the maximum-likelihood estimator of  $x$ , and

$$H_x = K^T H_n K, \quad (\text{IV.23})$$

$K$  denoting the matrix of the  $k_{i\alpha}$ . We furthermore explicitly split  $\mathbf{x}$  into POI and NP components,  $\mu$  and  $\theta$  respectively,  $\mathbf{x}^T = \left( \mu \mid \theta \right)$ , and assume that the  $\mathbf{n}$  similarly decompose into true experimental measurements and auxiliary measurements [889] providing constraints on the  $\theta$ . We also follow the usual convention [889] of orthogonalizing and normalizing the  $\theta$  so that the auxiliary measurements are all represented by standard Gaussians of unit width. We then have

$$H_n = \left( \begin{array}{c|c} \Sigma^{-1} & 0 \\ \hline 0 & I \end{array} \right) \quad K = \left( \begin{array}{c|c} R & \Delta \\ \hline 0 & I \end{array} \right) \quad (\text{IV.24})$$

where  $\Sigma^{-1}$  is the diagonal matrix of the uncertainties of the experimental measurements;  $R$  is the response matrix giving the impact of the POIs on the measurements; and  $\Delta$  provides the

impacts of the NPs on the measurements, representing systematic uncertainties. It follows that

$$H_{\mathbf{x}} = \left( \begin{array}{c|c} R^T \Sigma^{-1} R & R^T \Sigma^{-1} \Delta \\ \hline \Delta^T \Sigma^{-1} R & I + \Delta^T \Sigma^{-1} \Delta \end{array} \right) = \left( \begin{array}{c|c} C_{\text{stat}}^{-1} & \Lambda \\ \hline \Lambda^T & I + K_{\theta} \end{array} \right). \quad (\text{IV.25})$$

The  $C_{\text{stat}}$  matrix represents the statistical component of the measurement of  $\mu$ ; the bottom-right block describes the uncertainties of the  $\theta$ , which can be reduced from the initial unit values due to constraints from the measurements; and the off-diagonal blocks  $\Lambda$  represent the correlation between the measurements of  $\mu$  and  $\theta$  through systematic effects.

The measurement of the POIs  $\mu$  is performed using the profile likelihood ratio

$$\lambda(\mu) = \frac{L(\mu, \hat{\theta}(\mu))}{L(\hat{\mu}, \hat{\theta})} = \exp \left[ -\frac{1}{2} (\mu - \hat{\mu})^T H'_{\mu} (\mu - \hat{\mu}) \right]. \quad (\text{IV.26})$$

In terms of a general Hessian matrix

$$H_{\mathbf{x}} = \left( \begin{array}{c|c} H_{\mu\mu} & H_{\mu\theta} \\ \hline H_{\mu\theta}^T & H_{\theta\theta} \end{array} \right), \quad (\text{IV.27})$$

the profiled values of the nuisance parameters are

$$\hat{\theta}(\mu) = \hat{\theta} + H_{\theta\theta}^{-1} H_{\mu\theta}^T (\mu - \hat{\mu}) \quad (\text{IV.28})$$

and the profile likelihood ratio has a Gaussian form with the reduced Hessian matrix

$$H'_{\mu} = H_{\mu\mu} - H_{\mu\theta} H_{\theta\theta}^{-1} H_{\mu\theta}^T. \quad (\text{IV.29})$$

which for the Hessian of Eq. (IV.25) gives

$$H'_{\mu} = C_{\text{stat}}^{-1} - \Lambda (I + K_{\theta})^{-1} \Lambda^T. \quad (\text{IV.30})$$

In terms of the covariance matrix

$$C_{\mathbf{x}} = H_{\mathbf{x}}^{-1} = \left( \begin{array}{c|c} C_{\mu} & C_{\mu\theta} \\ \hline C_{\mu\theta}^T & C_{\theta} \end{array} \right) \quad (\text{IV.31})$$

of the full measurement, one has  $H'_{\mu} = C_{\mu}^{-1}$ : the sector of  $C_{\mathbf{x}}$  corresponding to the POI measurement is the inverse of  $H'_{\mu}$  computed above after profiling the NPs and thus already includes their impacts. Assuming these are small compared to statistical uncertainties ( $\Delta \Delta^T \ll \Sigma$ ), this can be written to linear order as

$$C_{\mu} = C_{\text{stat}} + C_{\text{stat}} \Lambda \Lambda^T C_{\text{stat}} \quad (\text{IV.32})$$

$$\equiv C_{\text{stat}} + C_{\text{syst}}, \quad (\text{IV.33})$$

where the first term represents statistical uncertainties and the second one systematics.

### 6.5.2 Link with experimental results

Higgs couplings measurements typically report uncertainties using the profile likelihood ratio, computed from the full experimental likelihood. They also often provide the covariance matrix  $C_\mu$ , obtained from the Hessian matrix  $H_{\mathbf{x}}$  of second derivatives of the likelihood evaluated at the best-fit position and truncated to POIs only. This allows the experimental results to be recast under specific theoretical models, accounting for correlation between the measured parameters. This suffers from the two limitations mentioned in the introduction: first the Gaussian assumption on the likelihood form, which is assumed to hold here; and secondly the non-inclusion of the nuisance parameters describing systematic uncertainties. If the same parameters also play a role at the level of the reinterpretation,  $C_\mu$  alone does not provide sufficient information to properly correlate the two effects.

Such a treatment can be performed using the full covariance matrix  $C_{\mathbf{x}} = H_{\mathbf{x}}^{-1}$ . To limit the size of the matrix, one can still truncate from  $C$  the uncertainties that are not relevant to the interpretation stage (for instance experimental uncertainties), leading to a reduced parameter set  $\tilde{\mathbf{x}} = (\mu, \theta_1 \cdots \theta_p)$ . This approach allows to rebuild an experimental likelihood as in Eq. (IV.22), assuming Gaussian assumptions.

A third option, inspired by Eq. (IV.33), is to report  $C_\mu$  together with  $C_{\text{stat}}$ , which allows to extract the effect of the relevant systematic uncertainties.  $C_{\text{stat}}$  can be obtained from the full likelihood in the same way as  $C_\mu$ , but keeping the systematics NPs fixed to their best-fit value. The method is most straightforwardly applied by reporting the matrices

$$C_p = C_\mu, C_{p-1} = C_{\mu|\theta_p}, C_{p-2} = C_{\mu|\theta_p, \theta_{p-1}} \cdots, C_0 = C_{\mu|\theta_p, \dots, \theta_1} \quad (\text{IV.34})$$

where the NPs after the vertical bar are fixed to their best-fit value when computing the correlation matrix.  $C_0$  therefore corresponds to the covariance matrix of the  $\mu$  without the effect of any of the  $\theta_1 \cdots \theta_p$  included,  $C_1$  corresponds to the covariance matrix including the effect of  $\theta_1$  only, and so on until all the  $\theta_1 \cdots \theta_p$  are included in  $C_p$ . Under the same assumptions as those leading to Eq. (IV.33), one then has

$$\delta C_i = C_i - C_{i-1} = C_{i-1} \Lambda_i \Lambda_i^T C_{i-1} \quad (\text{IV.35})$$

which allows to extract the impact  $\Lambda_i$  of the NP  $\theta_i$ . The sign of  $\Lambda_i$  is however ambiguous and needs to be provided in addition to the covariance matrices. The  $\Lambda_i$  can also be obtained without simplifying assumptions by re-expressing Eq. (IV.29) as

$$C_i^{-1} = C_{i-1}^{-1} - \Lambda_i (1 + K_i)^{-1} \Lambda_i^T. \quad (\text{IV.36})$$

The  $\Lambda_i$  (the impact of  $\theta_i$  on the POIs) and  $1 + K_i$  (which determines the uncertainty on  $\theta_i$ ) cannot be unambiguously determined since the covariance matrix is only sensitive to their combination. A possible convention is to require that the post-fit uncertainty on  $\theta_i$  is 1, as for the pre-fit case, by setting  $K_i = \Lambda_i^T C_{i-1} \Lambda_i$ . The components of  $\Lambda_i$  are then given by

$$[\Lambda_i]_\alpha = \pm \sqrt{(1 + K_i)[C_{i-1}^{-1} - C_i^{-1}]_{\alpha\alpha}}, \quad 1 + K_i = [1 - \text{Tr}(I - C_{i-1} C_i^{-1})]^{-1} \quad (\text{IV.37})$$

with the sign of the  $[\Lambda_i]_\alpha$  again provided as external input. The full matrix  $H_{\tilde{\mathbf{x}}}$  can then be reconstructed as

$$H_{\mathbf{x}} = \left( \begin{array}{c|c|c|c|c} C_0^{-1} & \Lambda_p & \Lambda_{p-1} & \cdots & \Lambda_1 \\ \hline \Lambda_p^T & I + K_p & 0 & \cdots & 0 \\ \hline \Lambda_{p-1}^T & 0 & I + K_{p-1} & \ddots & \vdots \\ \hline \vdots & \vdots & \ddots & \ddots & 0 \\ \hline \Lambda_1^T & 0 & \cdots & 0 & I + K_1 \end{array} \right). \quad (\text{IV.38})$$

Parameters	<i>0-jet</i>	<i>2-jet</i>
$\sigma_{\text{ggF}}$	102 fb $\times$ 96%	102 fb $\times$ 4%
$\sigma_{\text{VBF}}$	-	8 fb
Signal $A \times \epsilon$	40%	
$\sigma_{\text{Bkg}}^{\text{vis}}$	1500 fb	100 fb
$\mathcal{L}$	150 fb $^{-1}$	
Systematics	<i>0-jet</i>	<i>2-jet</i>
Experimental, <i>0-jet</i>	5%	-
Experimental, <i>2-jet</i>	-	5%
ggF theory on inclusive $A \times \epsilon$	2%	2%
VBF theory on inclusive $A \times \epsilon$	-	2%
ggF theory on <i>2-jet</i> $A \times \epsilon$	-	15%

Table IV.4: Summary of numerical parameters defining the example model.  $\sigma_{\text{ggF}}$  and  $\sigma_{\text{VBF}}$  are the total cross-sections for signal production in the ggF and VBF modes;  $A \times \epsilon$  is the overall efficiency of the signal selection;  $\mathcal{L}$  is the integrated luminosity of the data sample; and  $\sigma_{\text{Bkg}}^{\text{vis}}$  is the visible cross-section the continuum background (the ratio of the number of observed background events after all selections to the integrated luminosity). In the lower half of the table, each line corresponds to an independent source of systematic uncertainty, associated with a separate nuisance parameter.

This expression of  $H_{\mathbf{x}}$  can then be used to represent the experimental likelihood as above. It has the property that successively profiling away  $\theta_1 \cdots \theta_p$  leads to reduced Hessian matrices with top-left sectors equal to  $C_1^{-1} \cdots C_p^{-1}$ , so that profiling recovers at each stage the covariance matrices obtained from the full profile likelihood computation<sup>9</sup>.

### 6.5.3 Example

As an illustration, we consider a simple example inspired by the measurement of the gluon-fusion and VBF cross-sections in the  $H \rightarrow \gamma\gamma$  decay channel [890]. The measurement is modeled by counting experiments in two bins, denoted as *0-jet* and *2-jet*, and considering contributions from gluon-fusion Higgs production (ggF), VBF production and background. Theoretical and experimental uncertainties are considered on the experimental acceptance values, and implemented with log-normal profiles. The parameters of the model are summarized in Table IV.4. A profile likelihood ratio (PLR) scan in the  $(\sigma_{\text{ggF}}, \sigma_{\text{VBF}})$  plane for the SM hypothesis is performed, and the corresponding confidence level contours computed in the asymptotic approximation [891] are shown in Fig. IV.19a. The numerical results are

$$\sigma_{\text{ggF}} = 102_{-8.5}^{+8.2}(\text{stat})_{-4.1}^{+5.7}(\text{exp})_{-1.7}^{+2.3}(\text{theo}) \text{ fb} \quad (\text{IV.39})$$

$$\sigma_{\text{VBF}} = 8.0 \pm 2.1(\text{stat})_{-0.5}^{+0.7}(\text{exp})_{-0.4}^{+0.7}(\text{theo}) \text{ fb} \quad (\text{IV.40})$$

with a correlation coefficient of 17%. The covariance matrix of the fit is used to build a Gaussian likelihood following Eq. (IV.22). The PLR contours for this likelihood are also shown in Fig. IV.19a. Small differences are visible between the two due to the non-Gaussian form of the full likelihood. A similar comparison without systematic uncertainties included shows excellent agreement, which suggests the log-normal profiles of systematics is the main source of non-Gaussianity.

<sup>9</sup>Note that this works only if the parameters are profiled in the same order as for the computation of the  $C_i$ , since correlation information between the NPs is not included.

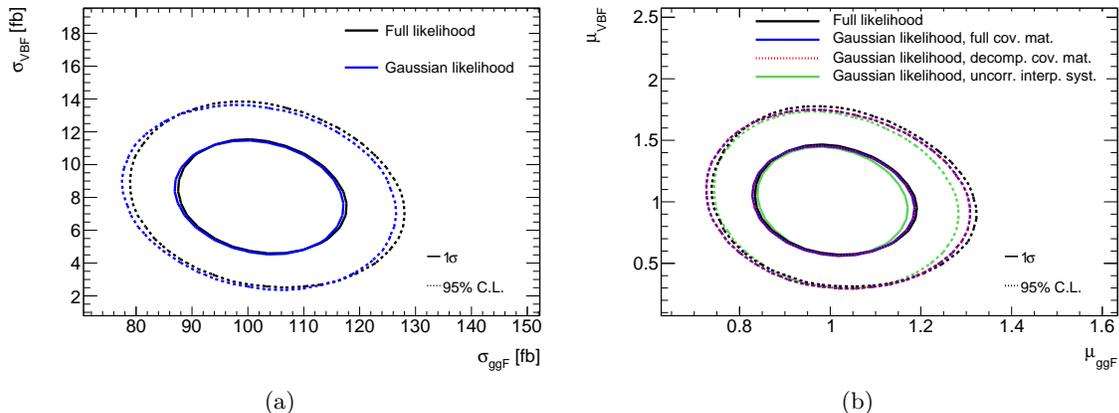


Fig. IV.19: Confidence level contours at the  $1\sigma$  (solid lines) and 95% CL (dashed lines) levels, computed from the profile likelihood ratio in the asymptotic approximation. (a) Contours in the  $(\sigma_{ggF}, \sigma_{VBF})$  plane of the original measurement, using the full likelihood (black) and the Gaussian likelihood built from the covariance matrix reported by HESSE at the best-fit point. (b) Contours in the  $(\mu_{ggF}, \mu_{VBF})$  plane computed from those in (a), using a reparameterization the full likelihood (black); the Gaussian likelihood built from the covariance matrix of the parameter set  $(\sigma_{ggF}, \sigma_{VBF}, \theta_{ggF}, \theta_{VBF})$  reported by HESSE (blue); the Gaussian likelihood built from the covariance matrix recomposed as in Eq. (IV.38) (red); and the Gaussian likelihood built from the covariance matrix of  $(\sigma_{ggF}, \sigma_{VBF})$ , with additional interpretation-stage uncertainties uncorrelated with those of the measurement.

The measurement is reinterpreted in the  $\mu$  parameterization of Higgs couplings, with parameters of interest defined as the ratios  $\mu_{ggF} = \sigma_{ggF}/\sigma_{ggF,SM}$  and  $\mu_{VBF} = \sigma_{VBF}/\sigma_{VBF,SM}$  of the measured cross-sections to their SM values. The parameterization includes uncertainties of 5% on the SM ggF and VBF cross-sections with log-normal profiles. They are considered to be fully correlated with the corresponding inclusive acceptance systematics ( $\theta_{ggF}$  and  $\theta_{VBF}$  respectively), making it possible to compare the possibilities described in the preceding section.

We consider the 3 cases of the model of Eq. (IV.27) in which the full covariance matrix is provided; the model of Eq. (IV.38) in which the covariance matrix is provided in decomposed form; and the case where the covariance matrix is only reported for the POIs, and the systematics on the interpretation are added separately without accounting for their correlation with those of the measurement.

The results are shown in Fig. IV.19b. The contours from the two methods of reporting the full covariance matrix are rather close to those of the exact likelihood, differing by the non-Gaussian effects already visible on Fig. IV.19a. No difference between the two methods is visible on the plot, as expected due to the smallness of the impact of NP/NP correlations. Both results provide a better description than the case where the systematics for the interpretation stage are not correlated with those of the measurement. The uncertainties are somewhat smaller in this case, since to first approximation the uncertainties at each stage add in quadrature, whereas they add linearly when fully correlated. The magnitude of the difference would however be reduced for smaller values of the theoretical uncertainties on the measured values.

## Acknowledgments

This writeup is based on discussions in Les Houches 2017 as well as subsequent discussions, with contributions and feedback from Josh Bendavid, Marco Delmastro, Frédéric Dreyer, Michael

Duehrssen-Debling, Markus Ebert, Maria Vittoria Garzelli, Nicolas Greiner, Stefan Kallweit, Predrag Milenovic, Luca Perrozzi, Carlo Pandini, Marco Zaro, and many others.

# Chapter V

## Phenomenological studies

### 1 Top-pair production including multi-jet merging and EW corrections <sup>1</sup>

#### 1.1 Introduction

The detailed study of top-quark properties, in particular in the dominant top-pair production mode, is amongst the main goals of the ongoing physics program at the LHC. After initial measurement at the inclusive cross section level [368, 892–896], where very good agreement with perturbative calculations at the NNLO+NNLL level in QCD [26, 336, 337] has been observed, the attention in the study of top-pair production has shifted towards differential measurements. In fact, the ATLAS and CMS collaborations already performed various measurements in the different top-decay channels, at 7 TeV [897–899], 8 TeV [373, 737] and 13 TeV [737, 900–903], presenting in particular results for the reconstructed top transverse momentum distribution, whose reliable modelling is of great importance for background estimates in a multitude of searches for Physics Beyond the Standard Model. These measurements have mostly been compared against predictions obtained with the Monte Carlo frameworks MC@NLO [904] (or more recently MadGraph\_aMC@NLO [88]) and POWHEG [905]. In these frameworks top-pair production at NLO QCD is matched to parton showers from PYTHIA [629] or HERWIG [906]. These measurements consistently indicate that the top quark transverse momentum distribution at low  $p_T$  is well predicted by the Monte Carlo programs, both in normalisation and shape, but these predictions exceed the data at high  $p_T$ . Comparing these measurements at the unfolded parton level to differential NNLO QCD predictions [146], this excess has been alleviated. This indicates the relevance of including higher jet multiplicities [352, 353, 375, 378–380, 907–917] in the modelling of the top quark transverse momentum distribution at high  $p_T$ . At the same time, it is well known that higher-order EW corrections alter the shape of the top transverse momentum distribution at high  $p_T$  due to the appearance of EW Sudakov logarithms [329, 331, 918–923] yielding corrections of about  $-10\%$  at  $p_{T,\text{top}} = 1$  TeV.

In this contribution [332], we first present a calculation of  $t\bar{t}$ (+jet) production at fixed-order NLO EW focussing on the top transverse momentum distribution, however, also showing results for the invariant mass distribution of the top-quark pair. Second, we present predictions based on the MEPS@NLO multi-jet merging in SHERPA [924–926], incorporating the EW corrections in an approximation that we show holds at the level of a few percent up the TeV range. In this approximation, the dominant virtual NLO EW corrections are incorporated exactly, while the NLO QED bremsstrahlung is first integrated out and subsequently incorporated via YFS multi-photon emission [927]. We compare the resulting MEPS@NLO QCD+EW<sub>virt</sub> predictions for the top-quark transverse momentum distribution at particle-level against a recent measurement performed by ATLAS in the lepton+jet channel, based on a selection of top-quark candidates in the boosted regime [737].

#### 1.2 From NLO EW to approximate EW corrections in the MEPS@NLO framework

Based on the automated SHERPA+OPENLOOPS framework fixed-order NLO EW corrections to  $t\bar{t}$  and  $t\bar{t}$ +jet production can readily be computed [85, 304, 497]. They comprise virtual corrections that are enhanced at large energies due to the appearance of EW Sudakov logarithms

---

<sup>1</sup> C. Gütschow, J. M. Lindert, M. Schönherr

and QED bremsstrahlung corrections. For  $t\bar{t}$ +jet production also mixed QCD-EW interference contributions arise in 4-quark channels and are included in our NLO EW predictions.

To be precise, at NLO EW, the cross section is defined as

$$d\sigma^{\text{NLO}} = d\Phi_B \left[ B(\Phi_B) + \tilde{V}(\Phi_B) \right] + d\Phi_R R(\Phi_R). \quad (\text{V.1})$$

Therein,  $B$  is the Born matrix element including all PDF and symmetry/averaging factors and  $\Phi_B$  is its accompanying phase space configuration.  $R$  and  $\Phi_R$  are defined analogously for the  $\mathcal{O}(\alpha)$  real correction and its phase space, while  $\tilde{V}$  contains the virtual  $\mathcal{O}(\alpha)$  correction  $V$  as well as the corresponding collinear counterterm of the PDF mass factorisation. At subleading orders of the strong coupling  $\alpha_S$ , further tree and one-loop contributions and photon-induced contributions arise. However, as shown in [329, 923] for  $t\bar{t}$  production these are numerically unimportant, given for the latter a determination of the photon content of the proton is based on the recent advances of [928]. In the following we consider the subleading tree contributions, but neglect any subleading one-loop or photon-induced contributions.

In order to approximate the full expression for the NLO EW contribution in the high-energy regime, by a form that is local in the Born phase space, and thus suitable for straightforward incorporation in the current MEPS@NLO framework in Sherpa, we define the  $W_{\text{virt}}$  approximation [305] as

$$d\sigma^{\text{NLO } W_{\text{virt}}} = d\Phi_B \left[ B(\Phi_B) + V(\Phi_B) + \int_1 d\Phi_1 R_{\text{approx}}(\Phi_B \cdot \Phi_1) \right]. \quad (\text{V.2})$$

Here, we introduce the approximated real-emission contribution  $R_{\text{approx}}$  such that its integral over the real-emission phase space equals the standard Catani-Seymour  $\mathbf{I}$ -operator. This construction is both finite and correctly reproduces the exact NLO EW corrections in the Sudakov limit, but also contains important non-logarithmic terms extending its validity in practice.

This result can now be straightforwardly used to incorporate approximate electroweak corrections in the MEPS@NLO method [924–926] of merging multiple samples of successive multiplicities at next-to-leading order accuracy in QCD. For any such multiplicity  $n$ ,  $n < n_{\text{max}}^{\text{NLO}}$ , its exclusive  $n$ -jet cross sections are defined as

$$d\sigma_n^{(\text{MEPS@NLO})} = \left[ d\Phi_n \bar{B}_n(\Phi_n) \bar{\mathcal{F}}_n(\mu_Q^2; < Q_{\text{cut}}) + d\Phi_{n+1} H_n(\Phi_{n+1}) \Theta(Q_{\text{cut}} - Q_{n+1}) \mathcal{F}_{n+1}(\mu_Q^2; < Q_{\text{cut}}) \right] \Theta(Q_n - Q_{\text{cut}}), \quad (\text{V.3})$$

with the familiar  $\bar{B}_n$  and  $H_n$  functions of the MC@NLO matching method [929–931] and their parton shower functionals  $\bar{\mathcal{F}}_n$  and  $\mathcal{F}_{n+1}$ . To this end, we modify the standard NLO QCD  $\bar{B}_n$  function to also incorporate the approximate NLO EW corrections now local in  $\Phi_n$ ,

$$\bar{B}_n(\Phi_n) \longrightarrow \bar{B}_{n,\text{QCD+W}}(\Phi_n) = \bar{B}_n(\Phi_n) + V_{n,\text{W}}(\Phi_n) + I_{n,\text{W}}(\Phi_n) + B_{n,\text{sub}}(\Phi_n). \quad (\text{V.4})$$

Additional multiplicities can be merged at leading order accuracy. They receive a differential  $K$ -factor [926, 932, 933], ensuring the continuity of the corrections applied throughout the spectrum. The term  $B_{n,\text{sub}}$  allows for the inclusion of additional subleading Born contributions in the merging.

$G_\mu = 1.1663787 \cdot 10^{-5} \text{ GeV}^2$	
$m_W = 80.385 \text{ GeV}$	$\Gamma_W = 2.0897 \text{ GeV}$
$m_Z = 91.1876 \text{ GeV}$	$\Gamma_Z = 2.4955 \text{ GeV}$
$m_h = 125 \text{ GeV}$	$\Gamma_h = 4.07 \text{ MeV}$
$m_t = 173.2 \text{ GeV}$	$\Gamma_t = 0$

Table V.1: Numerical values of all input parameters. While the masses are taken from [934], the widths are obtained from state-of-the-art calculations.

### 1.3 $t\bar{t}$ and $t\bar{t}$ +jet production at NLO EW

In Fig. V.1 we present predictions for the top-quark transverse momentum distribution and top-pair invariant mass distribution in top-pair and top-pair plus jet production at the LHC with 13 TeV including NLO EW corrections. For our predictions input parameters are chosen as listed in Table V.1 and the electroweak coupling  $\alpha$  is fixed and renormalized according to the  $G_\mu$ -scheme,  $\alpha = \frac{\sqrt{2}}{\pi} G_\mu \left| \mu_W^2 \sin^2 \theta_w \right|$ , where  $\mu_W$  denotes the complex-valued W mass, with  $\mu_V^2 = M_V^2 - i\Gamma_V M_V$  and  $\theta_w$  the also complex valued weak mixing angle, derived from the ratio  $\mu_W/\mu_Z$ . The massive vector bosons and the Higgs are renormalised in the complex-mass scheme, while the top-quark is kept stable and correspondingly renormalised in the on-shell scheme. The introduction of finite widths for the massive vector bosons is mandatory due to the appearance of otherwise singular resonant internal propagators in the  $\mathcal{O}(\alpha)$  bremsstrahlung. As renormalisation and factorisation scales for the strong coupling  $\alpha_S$  we use  $\mu = \mu_R = \mu_F = \frac{1}{2}(E_{T,t} + E_{T,\bar{t}})$ , where  $E_{T,t/\bar{t}}$  denotes the transverse energy of the top/anti-top. In the predictions for top-pair plus jet production we cluster jets according to the anti- $k_T$  algorithm implemented in FASTJET [819] and require  $p_{T,j} > 30 \text{ GeV}$  and  $|\eta_j| < 4.5$ .

The EW corrections to the top transverse momentum distribution shown in Fig. V.1 display a typical Sudakov behaviour and reach -10(-20)% at 1(2) TeV consistently in  $t\bar{t}$  and  $t\bar{t}$ +jet production. In fact, the difference in the relative NLO EW corrections between  $t\bar{t}$  and  $t\bar{t}$ +jet production is below 1(2)% up to about  $p_{T,t} = 1(2) \text{ TeV}$ . This indicates a factorisation of the EW correction from QCD radiation and supports a multiplicative combination of QCD and EW corrections for  $t\bar{t}$  production. The NLO EW<sub>virt</sub> approximation introduced in Eq. (V.2) reproduces the full NLO EW corrections in the transverse momentum of the top at the percent level, both for  $t\bar{t}$  and  $t\bar{t}$ +jet production. Subleading Born contributions of  $\mathcal{O}(\alpha_S\alpha)$  and  $\mathcal{O}(\alpha^2)$  in  $t\bar{t}$  production and  $\mathcal{O}(\alpha_S^2\alpha)$  and  $\mathcal{O}(\alpha_S\alpha^2)$  in  $t\bar{t}$ +jet production only contribute at the percent level.

In the invariant mass of the top-quark pair the corrections reach -5(-10)% at 1(5) TeV, where again we observe a universality of the EW corrections between  $t\bar{t}$  and  $t\bar{t}$ +jet production. Here the accuracy of the NLO EW<sub>virt</sub> approximation is slightly worse compared to the transverse momentum distribution. Still, the full NLO EW correction and the approximation agree at the percent level. Subleading Born contributions are marginally relevant for  $t\bar{t}$ +jet production, while for  $t\bar{t}$  production the subleading  $\mathcal{O}(\alpha^2)$  contribution yields a relevant contribution at very large top-pair invariant masses.

### 1.4 Multi-jet merged predictions for $t\bar{t}$ production including EW corrections

In Fig. V.2 we present on the left parton-level multi-jet merged MEPS@NLO predictions for the top-quark transverse momentum at the LHC with 13 TeV. The input parameters and settings are chosen as detailed in Sec. 1.3. We merge  $t\bar{t}$  plus zero and one jet production based on NLO matrix elements including EW corrections in the EW<sub>virt</sub> approximation of Eq. (V.4) and subleading Born contributions are included. Additional jet multiplicities up to  $t\bar{t}$ +4 jets are

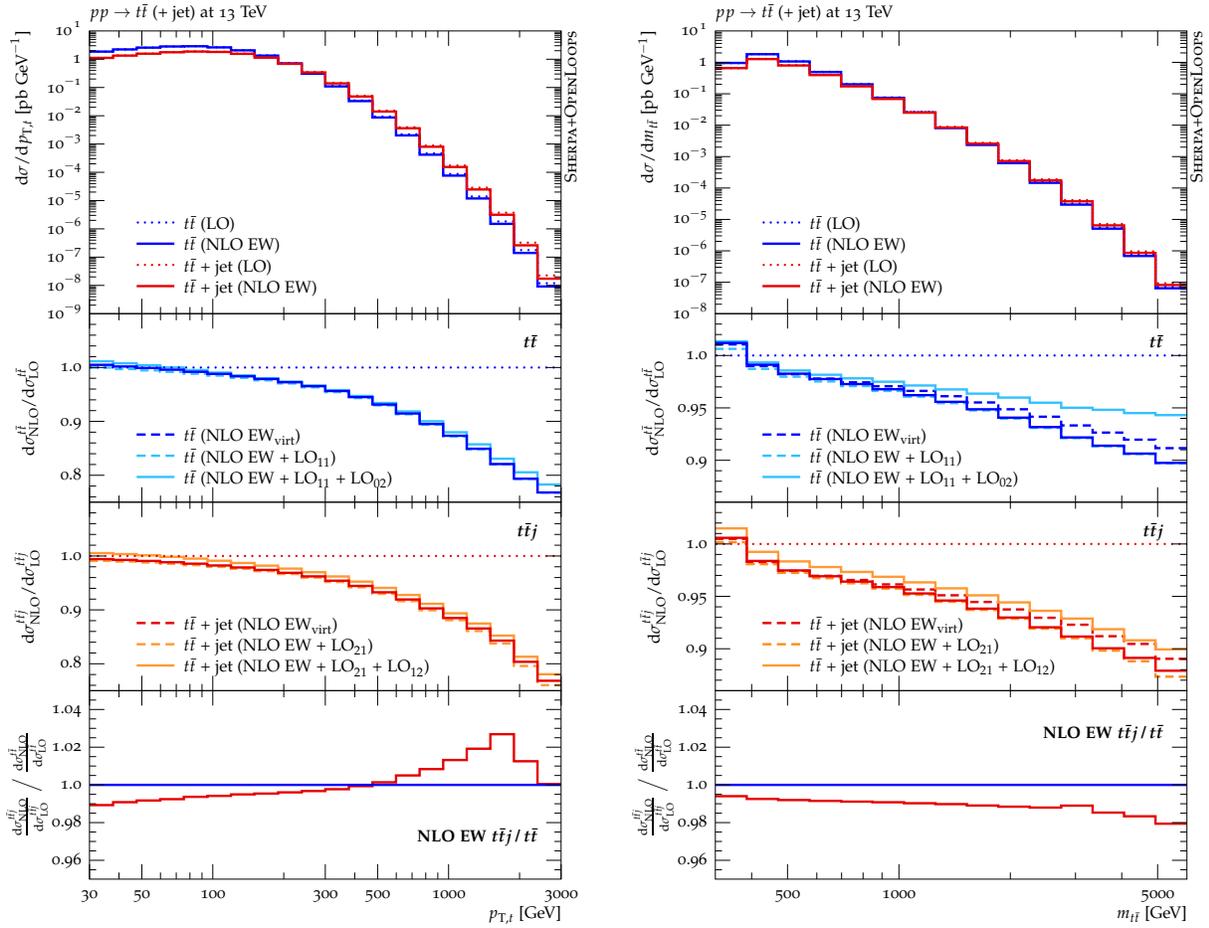


Fig. V.1: Top quark transverse momentum (left) and top-antitop invariant mass (right) distributions in inclusive  $t\bar{t}$  production (blue) and  $t\bar{t} + \text{jet}$  production (red) at NLO EW at 13 TeV at the LHC. The top panel shows the differential cross section, while the three lower panels show, from top to bottom, the NLO EW corrections to inclusive  $t\bar{t}$  production and  $t\bar{t} + \text{jet}$  production, respectively. The lowest panel shows the ratio of both corrections.

merged at LO accuracy. Comparing the MEPS@NLO QCD predictions with the MEPS@NLO QCD+EW<sub>virt</sub> predictions we recover an EW correction factor consistent with the fixed-order results presented in Sec. 1.3.

In the right plot of Fig. V.2 we finally compare particle-level MEPS@NLO predictions for the reconstructed top-quark transverse momentum distribution at the LHC with 8 TeV against a recent ATLAS measurement of differential top-pair production in the lepton+jets channel, based on a tagging of hadronically decaying top-quark candidates in the boosted regime [737]. In both the QCD and the QCD+EW<sub>virt</sub> predictions the top-quarks are decayed preserving spin correlations [493]. Non-perturbative effects, i.e. hadronization and MPI, are also included, as are higher order QED corrections [927]. Here, scales are set using the CKKW procedure [924, 935]. Thus,  $\mu_R = \mu_{\text{CKKW}}$  is defined through

$$\alpha_s^{2+n}(\mu_{\text{CKKW}}^2) = \alpha_s^2(\mu_{\text{core}}^2) \cdot \alpha_s(t_1) \cdots \alpha_s(t_n), \quad (\text{V.5})$$

and  $\mu_F = \mu_Q = \mu_{\text{core}}$ , where as core scale was have [352, 914]

$$\mu_{\text{core}} = \frac{1}{2} \left( \frac{1}{\hat{s}} + \frac{1}{m_t^2 - \hat{t}} + \frac{1}{m_t^2 - \hat{u}} \right)^{-\frac{1}{2}}, \quad (\text{V.6})$$

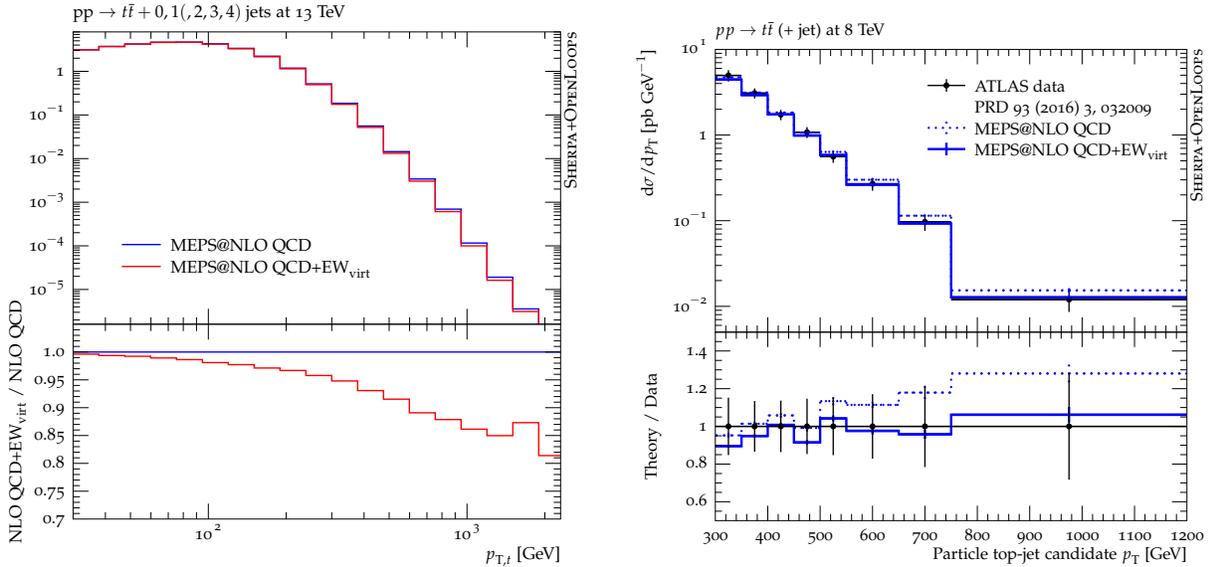


Fig. V.2: Top-quark transverse momentum distribution at parton-level (left) and particle-level (right), where for the latter hadronically decaying top candidates are reconstructed in a lepton+jets sample based on jet substructure techniques. Compared are MEPS@NLO QCD and MEPS@NLO QCD+ $W_{\text{virt}}$  predictions obtained with SHERPA+OPENLOOPS. The particle-level predictions are also compared against a recent measurement performed by ATLAS [737].

based on a reclustered core  $2 \rightarrow 2$  process. The remaining parameters are as discussed above. Again zero and one jet multiplicities are merged based on NLO matrix elements. In the high- $p_T$  tail the inclusion of the EW corrections results in a significantly improved agreement of the ATLAS data with the SHERPA+OPENLOOPS Monte Carlo prediction.

## 1.5 Conclusions

In this contribution we have presented predictions for the top-quark transverse momentum distribution and top-pair invariant mass distribution in top-pair production and top-pair plus jet production including NLO EW corrections. Subsequently, based on the MEPS@NLO multijet merging framework in Sherpa, we derived parton- and particle-level predictions for inclusive top-pair production including higher-order EW corrections. The EW corrections are incorporated in an approximation, based on exact virtual NLO EW contributions combined with integrated-out QED Bremsstrahlung. We showed that this approximation is able to reproduce the full NLO EW result for  $t\bar{t}$  and  $t\bar{t}$ +jet production at the percent level. Comparing our predictions against a recent measurement for the top-quark  $p_T$ -spectrum performed by ATLAS in the lepton+jets channel we find very good agreement between Monte Carlo predictions and data when the EW corrections are included.

## Acknowledgements

We thank Frank Krauss and Stefano Pozzorini for useful conversations and support. This work has received funding from the European Union’s Horizon 2020 research and innovation programme as part of the Marie Skłodowska-Curie Innovative Training Network MCnetITN3 (grant agreement no. 722104).

## 2 Parton shower and off-shell effects in top quark mass determinations <sup>2</sup>

We compare different theoretical descriptions of top quark pair production in the di-lepton channel. The full NLO corrections to  $pp \rightarrow W^+W^-b\bar{b} \rightarrow (e^+\nu_e)(\mu^-\bar{\nu}_\mu)b\bar{b}$  production are compared to calculations in the narrow width approximation, where the production of a top quark pair is calculated at NLO and combined with different descriptions of the top quark decay: LO, NLO and via a parton shower.

### 2.1 Introduction

A precise knowledge of the top quark mass is very important for many aspects of collider physics, ranging from electroweak precision tests and constraints on New Physics to investigations about the vacuum stability of the universe. Recent measurements at the LHC [936–939] reach an uncertainty below 1 GeV, which implies that theoretical work is demanded to further reduce it. As the top quark mass is not a physical observable, its definition is scheme dependent. The most commonly used mass definitions are the pole mass and the  $\overline{\text{MS}}$  mass, where the different masses are related by a perturbative series, see e.g. Refs. [940,941]. Recent theoretical studies with regards to the definition and extraction of the top quark mass can be found in Refs. [360–365,942].

For stable top quarks, NNLO corrections to differential distributions are known [146,327,328] and have been combined with NLO electroweak corrections recently [329]. Due to their very high complexity, the NNLO fixed-order calculations have so far only been combined with top quark decay in the narrow-width approximation (NWA), which factorises the production and decay processes. NLO corrections to top quark decays have been calculated in Refs. [333–335], and NNLO QCD corrections to the decay are also known meanwhile [148,943]. The combination of fixed-order NLO corrections to both production and decay with a parton shower in the narrow width approximation has been done in Ref. [349] within an extension of the PowHeg [350,351] framework, called `ttb_NLO_dec` in the POWHEG-BOX-V2. Within the SHERPA [415] framework, NLO QCD predictions based on the NWA for top quark pair production with up to three jets matched to a parton shower are also available, see Refs. [352,353].

However, a description of top quark production and decay which can describe the *shapes* of distributions to an accuracy required for improvements on current experimental precision needs to go beyond the narrow-width approximation. NLO QCD calculations of  $W^+W^-b\bar{b}$  production, including leptonic decays of the  $W$  bosons, have been performed in Refs. [342–345]. These calculations use the 5-flavour scheme, where the  $b$ -quarks are treated as massless partons. NLO electroweak corrections with complete off-shell effects have been calculated in the di-lepton channel in Ref. [331]. In the lepton+jets channel, NLO QCD corrections with full off-shell effects are available since recently [346].

The  $b$ -quark mass effects on observables like the invariant mass of a lepton- $b$ -quark pair ( $m_{l\bar{b}}$ ) are very small. However, the use of massive  $b$ -quarks (more precisely, the 4-flavour scheme, 4FNS) has the technically important feature that it avoids collinear singularities due to  $g \rightarrow b\bar{b}$  splittings. This implies that any phase space restrictions on the  $b$ -quarks can be made without destroying infrared safety, and thus allows to consider 0, 1- and 2-jet bins for  $pp \rightarrow e^+\nu_e\mu^-\bar{\nu}_\mu b\bar{b}$  in one and the same setup, which is important for cross sections defined by jet vetos. In Refs. [347,348], NLO calculations in the 4FNS have been performed.

Based on an NLO calculation of  $W^+W^-b\bar{b}$  production combined with the Powheg framework, first results of the  $W^+W^-b\bar{b}$  calculation in the 5-flavour scheme matched to a parton shower have been presented in Ref. [355]. However, it has been noticed subsequently that the matching of NLO matrix elements involving resonances of coloured particles to parton showers

---

<sup>2</sup> G. Heinrich, L. Scyboz

poses problems which can lead to artifacts in the top quark lineshape [356]. As a consequence, an improvement of the resonance treatment has been implemented in POWHEG-BOX-RES, called “resonance aware matching”, and combined with NLO matrix elements from OPENLOOPS [85], to arrive at the most complete description so far [357], based on the framework developed in Ref. [356] and the 4FNS calculation of Ref. [348]. An alternative algorithm to treat radiation from heavy quarks in the Powheg NLO+PS framework has been presented in Ref. [358]. An improved resonance treatment in the matching to parton showers for off-shell single top production at NLO has been worked out in Ref. [944], and for off-shell  $t\bar{t}$  and  $t\bar{t}H$  production in  $e^+e^-$  collisions in Ref. [945]. A recent study of top quark mass measurements at the LHC using various NLO+PS approaches can be found in Ref. [367], and a detailed assessment of fragmentation uncertainties in Ref. [366].

Here we investigate the impact of different theoretical approximations on distributions relevant to top quark mass measurements, as described in more detail in Ref. [359]. We compare the NLO calculation of  $W^+W^-b\bar{b}$  production of Ref. [345] with the calculation based on the narrow-width approximation where both  $t\bar{t}$  production *and* decay are calculated at NLO [334] and with a calculation in the narrow-width approximation where the Sherpa [415] parton shower is combined with our matrix elements of top quark pair production calculated at NLO.

## 2.2 Theoretical descriptions

We consider the following descriptions of the top quark pair production cross section in the di-lepton channel:

NLO<sub>full</sub>: full NLO corrections to  $pp \rightarrow W^+W^-b\bar{b}$  with leptonic  $W$ -decays,

NLO<sub>NWA</sub><sup>NLOdec</sup>: NLO  $t\bar{t}$  production  $\otimes$  NLO decay,

NLO<sub>NWA</sub><sup>LOdec</sup>: NLO  $t\bar{t}$  production  $\otimes$  LO decay,

NLO<sub>PS</sub>: NLO  $t\bar{t}$  production+shower  $\otimes$  decay via parton showering.

The calculations NLO<sub>full</sub> and NLO<sub>NWA</sub><sup>LOdec</sup> follow the ones described in Ref. [345], but now at  $\sqrt{s} = 13$  TeV and with cuts as described in Ref. [359]. The results for the NLO<sub>NWA</sub><sup>NLOdec</sup> calculation are obtained as described in Ref. [334], relying on the factorisation of the matrix elements according to

$$\mathcal{M}_{ij \rightarrow t\bar{t} \rightarrow b\bar{b}l_2l_2\nu}^{\text{NWA}} = \mathcal{P}_{ij \rightarrow t\bar{t}} \otimes \mathcal{D}_{t \rightarrow bl^+\nu} \otimes \mathcal{D}_{\bar{t} \rightarrow \bar{b}l^-\bar{\nu}}, \quad (\text{V.7})$$

where  $\mathcal{P}_{ij \rightarrow t\bar{t}}$  describes the  $t\bar{t}$  production process and  $\mathcal{D}_{t \rightarrow bl\nu}$  the top quark decay dynamics. Spin correlations are included as indicated by the symbol  $\otimes$ .

The NLO<sub>PS</sub> computations shown here are based on the NLO plus parton-shower matching scheme as implemented in SHERPA [914,929]. Using this scheme, we obtain an NLO+PS accurate description of  $t\bar{t}$  production. The top quark decays are attached afterwards in a way that spin correlations are preserved, and supplemented by their respective decay showers following the same procedure as proposed in [493].

In the following we will focus on the observable  $m_{lb}$  which is defined as the invariant mass

$$m_{lb}^2 = (p_l + p_b)^2, \quad (\text{V.8})$$

where  $p_l$  denotes the four-momentum of the lepton and  $p_b$  the four-momentum of the  $b$ -jet. As there are two top quarks, there are also two possible  $m_{lb}$  values per event. Since experimentally, it is not possible to reconstruct the  $b$ -quark charge on an event-by-event basis with sufficient

accuracy, one also needs a criterion to assign a pair of a charged lepton and a  $b$ -jet as the one stemming from the same top quark decay. Following [946], the algorithm applied here is to choose that ( $l^+b$ -jet,  $l^-b$ -jet') pairing which minimises the sum of the two  $m_{lb}$  values per event. Finally, the  $m_{lb}$  observable used in the analysis is the mean of the two  $m_{lb}$  values per event obtained when applying the above procedure.

The  $m_{lb}$  distribution has a kinematic edge at  $m_{lb}^{\text{edge}} = \sqrt{m_t^2 - M_W^2} = 152.6$  GeV, beyond which it is only populated by additional radiation, non-resonant contributions and incorrect  $b$ -lepton pairings.

### 2.3 Phenomenological results

We use the PDF4LHC15\_nlo\_30\_pdfas sets [422, 597, 601, 843] and a centre-of-mass energy of  $\sqrt{s} = 13$  TeV. Our default top quark mass is  $m_t = 172.5$  GeV. LO top quark and  $W$  boson widths are used in the LO calculations and the NLO  $t\bar{t} \otimes$  LO decay calculation, while NLO widths [947] are used in the remaining NLO calculations. NLO widths appearing in propagators are not expanded in  $\alpha_s$ . The QCD coupling in the NLO widths is varied according to the chosen scale. For more details on the input parameters and kinematic requirements we refer to Ref. [359]. The  $b$ -quarks are treated as massless in all fixed-order calculations.

We chose  $\mu_R = \mu_F = m_t$  as our central scale. The impact of choosing  $H_T/2$  (rather than  $m_t$ ) as the central scale on the top quark mass determined by our method has been shown to be very small [345]. Further it would be difficult to come up an  $H_T$  definition for the NLO<sub>PS</sub> approach that matches the one in the full  $W^+W^-b\bar{b}$  calculation.

For the NLO<sub>PS</sub> case, the standard  $\mu_R$  and  $\mu_F$  variations that we employ for our fixed-order calculations are not fully sufficient to assess the theory uncertainties, as the showering depends on further scale and parameter settings. Our variation in the NLO<sub>PS</sub> case, denoted by  $\mu_F\mu_R\alpha_s^{\text{PS}}$ , is a combination of simultaneously varying  $\mu_F$ ,  $\mu_R$  and  $\mu_R^{\text{PS}} \sim p_T^{\text{emit}}$  by a factor of two up and down. By evaluating  $\alpha_s(\mu_R^{\text{PS}})$ , the latter variation determines the shower uncertainty which we abbreviate by  $\alpha_s^{\text{PS}}$ . Other ways of uncertainty assessment include the variation of the shower starting scale  $\mu_Q$ , which regulates the overall size of the resummation regime primarily affecting the  $t\bar{t}$  production showers. The SHERPA default is to set this resummation scale equal to the factorisation scale, i.e. for our NLO<sub>PS</sub> results, we have set  $\mu_Q = \mu_F = m_t$ .

Experimentally, the top quark mass is measured most precisely by fitting normalised differential cross sections to Monte Carlo templates produced with different values for the top mass. This procedure is therefore sensitive to the modeling of the shape of the distributions entering the fit. Figure V.3 shows the differential cross section for the  $m_{lb}$  observable based on the four theoretical predictions NLO<sub>NWA</sub><sup>LOdec</sup>, NLO<sub>NWA</sub><sup>NLOdec</sup>, NLO<sub>PS</sub> and NLO<sub>full</sub>, where the latter is the baseline of the ratio plot, and its scale uncertainties are shown as grey bands. We observe that the simplest description, NLO<sub>NWA</sub><sup>LOdec</sup>, differs substantially from the NLO<sub>full</sub> prediction, in the range  $40 \text{ GeV} \leq m_{lb} \leq 150 \text{ GeV}$ , which comprises the bulk of the cross section, effectively moving the peak to greater values of  $m_{lb}$ . The NLO corrections to the top quark decay, NLO<sub>NWA</sub><sup>NLOdec</sup>, restore agreement with NLO<sub>full</sub> to within less than 10%, while the NLO<sub>PS</sub> prediction mostly lies between NLO<sub>NWA</sub><sup>LOdec</sup> and NLO<sub>NWA</sub><sup>NLOdec</sup> in this region. In contrast, the high-mass tail features substantial differences between the predictions, ranging from 70% for NLO<sub>NWA</sub><sup>LOdec</sup>, 50% for NLO<sub>NWA</sub><sup>NLOdec</sup>, to 20% for NLO<sub>PS</sub>. Taking only the range  $40 \text{ GeV} \leq m_{lb} \leq 150 \text{ GeV}$  into consideration, which is also within the range used for fits determining the top quark mass based on the  $m_{lb}$  distribution [359], the NWA predictions NLO<sub>NWA</sub><sup>NLOdec</sup> and NLO<sub>PS</sub>, which describe the top quark decay beyond the leading order, but do not contain non-resonant, singly-resonant and non-factorising contributions, come rather close to the full  $W^+W^-b\bar{b}$  calculation. Thus we conclude that the non-resonant and off-shell contributions are less important than corrections to the decay for top

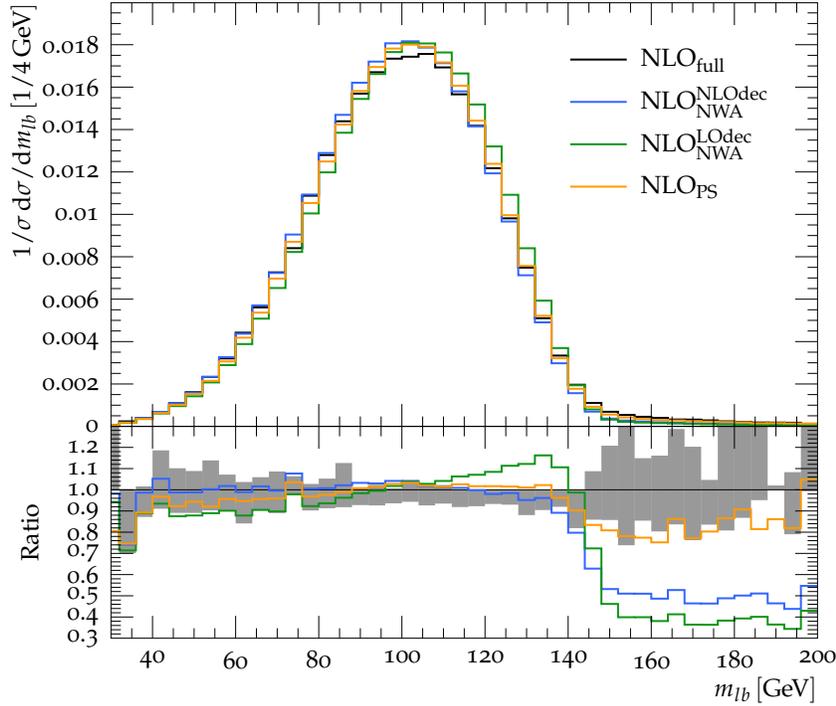


Fig. V.3: Normalised differential cross sections for the invariant mass  $m_{lb}$  at the 13 TeV LHC for four different theoretical descriptions: full NLO calculation including off-shell and non-resonant contributions ( $\text{NLO}_{\text{full}}$ ), NLO calculation in the NWA with and without NLO corrections to the top quark decay ( $\text{NLO}_{\text{NWA}}^{\text{NLOdec}}$  and  $\text{NLO}_{\text{NWA}}^{\text{LOdec}}$ , respectively) and NLO calculation for top quark pair production matched to a parton shower ( $\text{NLO}_{\text{PS}}$ ). The ratio of all descriptions to  $\text{NLO}_{\text{full}}$  including its scale uncertainty band is also shown.

quark mass determinations based on the  $m_{lb}$  observable.

In the case of the full  $W^+W^-b\bar{b}$  calculation, the NLO corrections introduce important changes in the shape of the  $m_{lb}$  distribution, as shown in Fig. V.4a. While this is expected for the tail of the distribution, it is remarkable that the result also gets corrections of up to 50% for lower  $m_{lb}$  values. The various descriptions of the top quark decay in the NWA exhibit a different behaviour, as shown in Fig. V.4b: when compared to the  $\text{NLO}_{\text{NWA}}^{\text{LOdec}}$  central prediction, both the  $\text{NLO}_{\text{NWA}}^{\text{NLOdec}}$  and the  $\text{NLO}_{\text{PS}}$  predictions soften the distribution up to the kinematic edge, while the region beyond the kinematic edge gets substantially enhanced by the parton shower. Notice that incorporating higher-order corrections to the decay pushes the  $m_{lb}$  distribution outside of the  $\text{NLO}_{\text{NWA}}^{\text{LOdec}}$  scale uncertainty bands, also in the peak region.

As a complementary study, and in order to disentangle the effect of additional emissions in the parton shower before and after the decay, we now consider the  $\text{NLO}_{\text{PS}}$  description in more detail. The parton shower can be artificially truncated after a given number of emissions  $n_{\text{prod}}$  and  $n_{\text{dec}}$  in the  $t\bar{t}$  production, respectively in the top quark decay. In Fig. V.5a, the unnormalised differential cross section for the  $m_{lb}$  observable is shown for four combinations of  $n_{\text{prod}}$  and  $n_{\text{dec}}$ , namely

- (A)  $(n_{\text{prod}} = \infty, n_{\text{dec}} = \infty)$
- (B)  $(n_{\text{prod}} = \infty, n_{\text{dec}} = 0)$
- (C)  $(n_{\text{prod}} = 0, n_{\text{dec}} = \infty)$
- (D)  $(n_{\text{prod}} = 1, n_{\text{dec}} = 1)$ ,

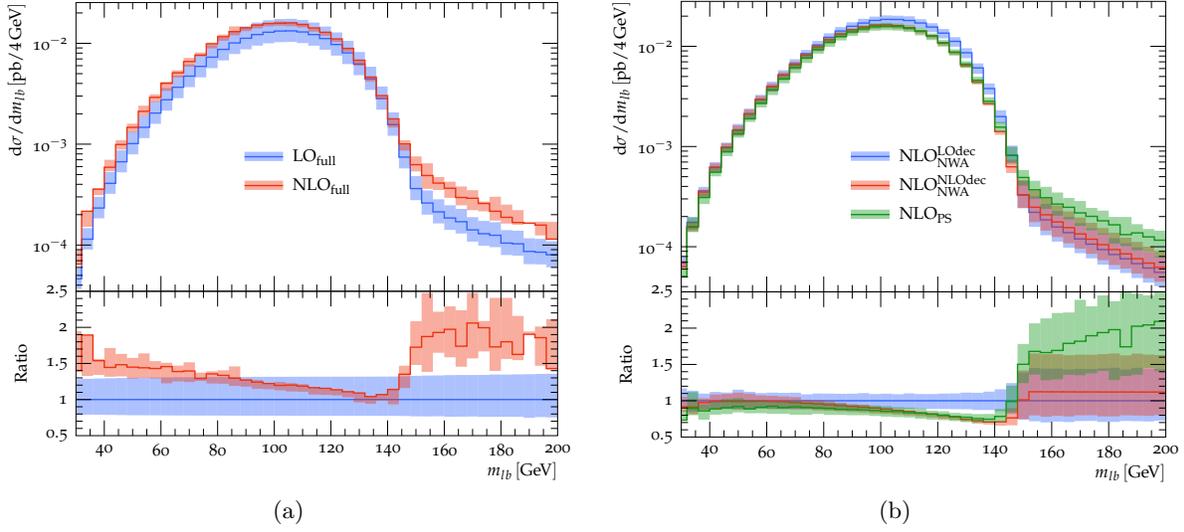


Fig. V.4: Results including scale variation bands for  $m_{lb}$ , for (a) the  $\text{LO}_{\text{full}}$  and  $\text{NLO}_{\text{full}}$  calculations, (b) the calculations based on the NWA. The ratio plot in (a) is with respect to  $\text{LO}_{\text{full}}$ , while the ratio plot in (b) is with respect to  $\text{NLO}_{\text{NWA}}^{\text{LOdec}}$ .

where  $n = \infty$  symbolically denotes the full shower emission (with default cutoff  $\mu_Q$ ). Comparing (A) to (B), that is the full shower and the deactivated shower in the top quark decay, we observe that the high- $m_{lb}$  tail is still described in a similar way in both cases, while the ratio (B)/(A) peaks sharply just before the turnover around the kinematic edge at  $\sim 145$  GeV. On the other hand, comparing (A) and (C), this time deactivating the shower in the top quark pair production, the agreement between the full and truncated shower is fair in the bulk of the distribution, while the ratio (C)/(A) drops in the high- $m_{lb}$  region. This is because there is no additional (mainly initial state) radiation in the production, which would further dilute the kinematic edge by radiating into the b-jet cone. If we consider case (D), where one emission is allowed in both the production and the decay shower, we can anticipate that allowing more emissions would progressively flatten the truncated curves shown in the ratio plot to the full shower baseline. We also observe that for  $m_{lb}$  values between 40 GeV and 140 GeV, the  $\text{NLO}_{\text{full}}$  result is within the uncertainty band of the  $\text{NLO}_{\text{PS}}$  result.

The underlying dynamics is exhibited more clearly if we look at a more theoretical observable, which does not suffer from the spread in phase space coming from the unknown neutrino four-momentum. We define the invariant mass of the  $W$  and b-jet system as

$$m_{Wb}^2 = (m_{lb} + p_\nu)^2 = (p_l + p_b + p_\nu)^2, \quad (\text{V.9})$$

where the same assignment for lepton and b-jets is used as in Eq. (V.8), and  $p_\nu$  designates the four-momentum of the neutrino corresponding to the lepton  $l$ . Again, the value of  $m_{Wb}$  used in the analysis is the mean of the two  $m_{Wb}$  values per event.

Figure V.5b describes the truncation of the parton shower for the  $m_{Wb}$  observable. In the case of the full shower (A), it peaks at the top mass  $m_t = 172.5$  GeV, and is spread by the real radiation on both sides. Going to the case (B), which corresponds to switching off the shower for the top quark decay, the spread of the peak drops to a constant baseline for  $m_{Wb}$  values below the peak, while even exceeding the  $n_{\text{prod}} = \infty$  result for  $m_{Wb}$  values beyond the peak. This behaviour can be understood from the fact that, if there is no radiation off the top quark at all, larger  $m_t$  values are preferred. In addition, initial state radiation being clustered into the

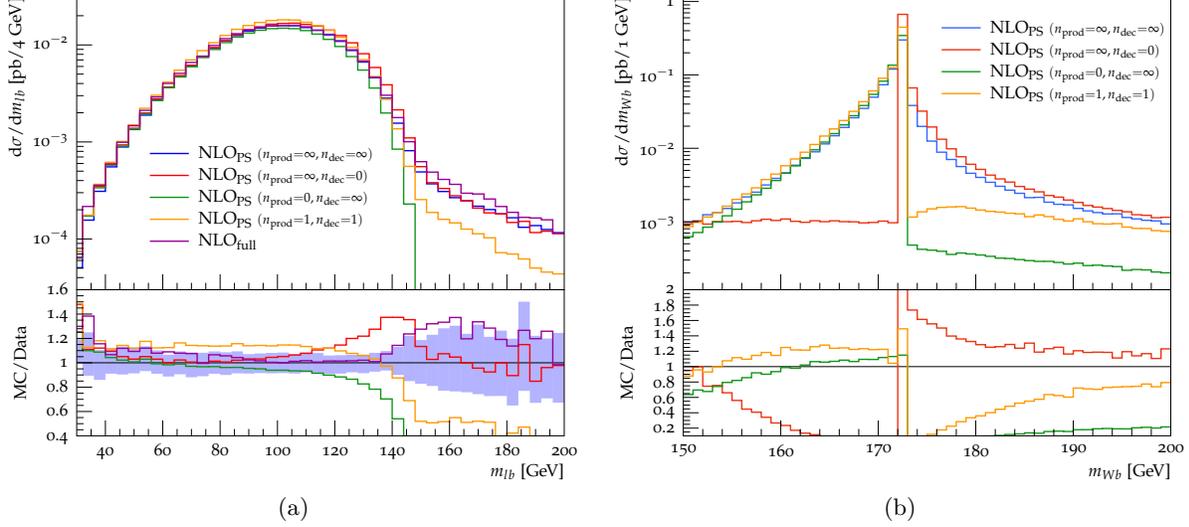


Fig. V.5: Truncated shower emissions for the observables (a)  $m_{lb}$ , (b)  $m_{Wb}$ . The shaded area in the ratio plot left denotes the scale uncertainty of the full NLO<sub>PS</sub> prediction.

b-jet increases the  $m_{Wb}$  values. Mirroring this situation, if the production shower is switched off as in case (C), the left side of the peak follows the fully-showered result (A) within 40%. The shower contribution above the top mass peak, in contrast, is reduced to almost zero, as additional radiation in the decay shower can only decrease the invariant mass of the  $Wb$  system. In case (D), we observe that one emission in the decay shower suffices to bring the curve below the peak to its final, fully-showered level. In contrast, the distribution only slowly builds up with each emission from the production shower in the high-mass tail.

These considerations show that the shape of observables sensitive to the top quark mass, like  $m_{lb}$  and  $m_{Wb}$ , is considerably affected by higher order corrections to both, production and decay.

## 2.4 Conclusions

We have investigated how various theoretical descriptions of top quark pair production translate into the behaviour of the distributions  $m_{lb}$  and  $m_{Wb}$  in view of measurements of the top quark mass in the di-lepton channel. In particular, we have compared the NLO QCD results for  $W^+W^-b\bar{b}$  production (NLO<sub>full</sub>) to results based on the narrow-width approximation, combining  $t\bar{t}$  production at NLO with (i) LO top quark decays (NLO<sub>NWA</sub><sup>LOdec</sup>), (ii) NLO top quark decays (NLO<sub>NWA</sub><sup>NLOdec</sup>) and (iii) a parton shower (NLO<sub>PS</sub>). We have found that corrections beyond the leading order in the *decay* (NLO<sub>NWA</sub><sup>NLOdec</sup> and NLO<sub>PS</sub> descriptions) are important, being always closer to the NLO<sub>full</sub> result than the leading order  $W^+W^-b\bar{b}$  result, even though the latter also contains non-resonant and non-factorising contributions.

We have further seen that the NLO<sub>NWA</sub><sup>NLOdec</sup> and NLO<sub>PS</sub> descriptions are pretty close to each other within an  $m_{lb}$  range that is usually taken into account for template fits to measure the top quark mass. We also considered parton showers with restricted emissions, restricting to zero or one emission in production and/or decay. The behaviour of the  $m_{lb}$  and  $m_{Wb}$  distributions in these cases shows that both the production and the decay shower play an important role to arrive at the fact that the NLO<sub>PS</sub> result, based on the narrow-width approximation, comes quite close to the NLO<sub>full</sub> result.

## Acknowledgements

We would like to thank our collaborators on this project, Andreas A. Maier, Richard Nisius, Johannes Schlenk and Jan Winter. We also thank the organizers of the “LH2017 Physics at TeV Colliders” workshop for creating a productive and pleasant environment. This research was supported in part by the Research Executive Agency (REA) of the European Union under the Grant Agreement PITN-GA2012316704 (HiggsTools).

## 3 Study of electroweak production of WZ in association with two jets at the LHC <sup>3</sup>

### 3.1 Introduction

The electroweak (EW) production of vector-boson pairs in association with two jets at the CERN Large Hadron Collider (LHC) forms an important class of processes both theoretically and experimentally. This signature includes vector-boson scattering (VBS) contributions, which constitute the principle example of processes where the scattering of two massive gauge bosons can be observed at the LHC. These processes provide a natural probe of the vector-boson quartic couplings, which arise due to the non-Abelian nature of the electroweak gauge group and are exactly predicted in the Standard Model (SM). The Higgs boson plays a unique role in these interactions, preventing the cross section from diverging in the high-energy limit and preserving the unitarity of the associated scattering amplitudes. Deviations in these channels could therefore indicate physics beyond the SM in the electroweak sector.

The measurements of these processes are particularly challenging due to their high multiplicities and small cross sections. They have been observed and even measured only recently by the experimental collaborations at the LHC. The most precise current measurement [948–951] concerns the scattering of two same sign W bosons (usually denoted  $W^\pm W^\pm jj$ ). This process has a unique signature due to the same-sign charged leptons in the final state. The cross section is well within reach with the large data sets being collected in the LHC Run II, and the signature is experimentally accessible due to the precise lepton identification, charge assignment, and momentum resolution of the LHC experiments. The nature of the final state is also attractive due to the low rate of background processes producing two prompt, same sign leptons associated with forward jets (referred to as the irreducible background).

Measurements of other VBS signatures such as ZZjj,  $W^+W^-jj$ , or  $W^\pm Zjj$  present additional challenges due to the lower cross section and larger irreducible backgrounds, which dominate over the VBS contribution in most regions of phase space. Nonetheless, studies have already been made at the LHC of both the ZZjj [952] and  $W^\pm Zjj$  [953] signature at  $\sqrt{s} = 8$  TeV. In these cases, separation of the electroweak component of the  $VVjj$  state depends on exploiting the different characteristics of the EW and non-EW components through kinematic selections. Accurate predictions for discriminating distributions and a detailed understanding of their associated uncertainties therefore directly impact such measurements. A theory-agnostic measurement that does not attempt separation of states by production mode avoids such dependencies, but cannot fully leverage the statistical tools and treatment of uncertainties available in an experimental analysis. The approaches are therefore complementary, and are often presented together in experimental results.

From a theoretical point of view, one of the challenges for the predictions of the EW di-boson production in association with two jets is its high multiplicity. This is why in the past, next-to-leading order (NLO) predictions have focused on VBS approximations. Only recently, full NLO computations became available at NLO QCD and EW for both the EW component

---

<sup>3</sup> K. Long, M. Pellen (section coordinators); S. Bräuer, V. Ciulli, S. Gieseke, M. Herndon, M. Mozer, S. Plätzer, M. Rauch, E. Yazgan

and the QCD-induced process for  $W^\pm W^\pm jj$  [226]. For this process, preliminary results [954] of a comparison of different theoretical predictions have shown that differences between the full computation and VBS-approximated ones are not significant given the present experimental accuracy. Qualitatively, one could expect a similar conclusion for other processes such as  $W^\pm Zjj$  but a quantitative check is still needed. Therefore, in these proceedings we aim at making a similar comparison of theoretical predictions implemented in Monte Carlo programs. This allows to infer the quality of the VBS approximation at leading-order (LO).

As such there is strong motivation for an investigation of the theoretical predictions, tools, and uncertainties used in a typical VBS analysis. We focus on the  $W^\pm Zjj$  [953] state, the measurement of which is strongly limited by the size of the data set, but which has not yet been studied at the 13 TeV LHC. Hence, a preliminary study on this process is of particular interest.

Due to its colour structure, the  $W^\pm Zjj$  signature possesses three different contributions at LO. The first, of order  $\mathcal{O}(\alpha^6)$ , is usually referred to as the EW component or even VBS component (even if it also possesses non-VBS contributions such as tri-boson production). The two quark lines can also be connected via a gluon while the gauge bosons are radiated off the quark lines. This contribution is of order  $\mathcal{O}(\alpha_s^2 \alpha^4)$  and is called QCD contribution/background. Finally, there exists a non-zero interference of order  $\mathcal{O}(\alpha_s \alpha^5)$ . For this signature (as opposed to *e.g.*  $W^\pm W^\pm jj$ ) the EW component is highly suppressed with respect to the QCD one. This implies having very exclusive experimental cuts in order to enhance the EW contribution and a good control over the description of both the EW signal and QCD background. To that end, the understanding of theoretical predictions and Monte Carlo programs is key. In particular, it is vital to ensure that all different programs can provide equivalent physics result. Such comparisons can also shed light on the true uncertainties of such predictions. Typically, uncertainty treatments assess the impact of parton density function (PDF) uncertainties or of missing higher orders in perturbative QCD. But in addition, Monte Carlo generators require many input parameters and the selection of which can have a significant impact on the predictions both in shape and normalisation. In the present study we use a common set-up for all predictions at LO accuracy. We additionally provide results with variations in some parameters and comment on their effects and motivations.

In these proceedings, we start with Sec. 3.2 where a short review of the theoretical state-of-the-art predictions is given. The programs used in the present work are also briefly described. In Sec. 3.3, the set-up of the calculation is presented. It amounts to give the input parameters as well as the event selection. Section 3.4 is devoted to the results of the study. It starts with LO predictions for the three different contributions to the  $e^+ \nu_e \mu^+ \mu^- jj$  final state at the LHC. Then the various comparisons are reported at both fixed-order and with parton shower. Finally, Sec. 3.5 contains a summary, concluding remarks as well as recommendations.

### 3.2 Theory and event generators

The study focuses on the  $W^\pm Zjj$  signature and more precisely on the partonic processes

$$pp \rightarrow e^+ \nu_e \mu^+ \mu^- jj, \quad (\text{V.10})$$

and

$$pp \rightarrow e^- \bar{\nu}_e \mu^+ \mu^- jj. \quad (\text{V.11})$$

The predictions presented are all at LO but NLO QCD corrections to the EW contribution and its irreducible background are known since 10 years in the VBS approximation [955] while the QCD corrections to the QCD-induced process have been computed more recently [956]. The NLO EW corrections are currently unknown. In Ref. [224], it has been argued that large NLO EW corrections to the EW contributions are an intrinsic feature of VBS at the LHC. Therefore,

they are expected to play a significant role for all VBS signatures. In Ref. [226], which focuses on the computation of the full NLO corrections to the  $W^\pm W^\pm jj$  process, it has been shown that the EW corrections to the EW process are the dominant NLO corrections. This means that the EW corrections to the EW contributions are expected to be at least of the same order as the QCD corrections for other VBS signatures. Parton-shower effects to VBS processes and the resulting uncertainties induced by a variation of factorisation, renormalisation, and shower-starting scale have also been studied recently [957, 958], taking the  $W^+W^-jj$  VBS process as an example. Thereby, the standard LHC jet definition with a cone radius of  $R=0.4$  leads to migration effects [851] and a reduction of the cross section by about 10% compared to the fixed-order results after VBS cuts.

In the following, the codes used for the predictions are briefly described.

### ***MADGRAPH5\_aMC@NLO***

MADGRAPH5\_aMC@NLO [88] is an automatic meta-code (a code that generates codes) which makes it possible to simulate any scattering process including NLO QCD corrections both at fixed-order and including matching to parton showers. The commands that have been used for the present computation (for the  $W^-$  case) are

```
> set complex_mass_scheme
> set gauge Feynman
> generate p p > e- ve~ mu+ mu- j j QED=6 QCD=0
> output WmZJJToENu2MuJJ
```

The version used is 2.6.0. We use the complex mass scheme and Feynman gauge for consistency with MoCaNLO +Recola, but verify that we obtain equivalent results if they are not explicitly specified. We note that the “deltaeta” variable in the MG5\_aMC run\_card also applies a requirement that the two jets have opposite rapidity sign. We do not make this requirement for our fiducial region definition, so we remove this condition for our comparisons. This is accomplished by modifying the cut behaviour in cuts.f in the “SubProcesses” folder, and validated by generating events without this condition and applying selections to the LHE-level partons using standalone code.

### ***HERWIG7***

Based on extensions of the previously developed MATCHBOX module [959], the HERWIG7 event generator [584, 825] facilitates the automated set-up of all ingredients necessary for a full NLO QCD calculation. It relies on an implementation of the Catani–Seymour dipole subtraction method [187, 489], as well as interfaces to a list of external matrix element providers – either at the level of squared matrix elements, based on extensions of the BLHA standard [505, 506, 594] or at the level of colour-ordered sub-amplitudes.

For this study the relevant tree-level matrix elements have been provided by an interface to VBFNLO [477, 847, 848] using an extension of the BLHA accord. We have matched these calculations to the angular ordered shower using the subtractive matching and default settings of the HERWIG7 7.1 release. The PDF sets that have been used to tune the parton shower are MMHT2014lo68cl and MMHT2014nlo68cl [601]. These sets are not the one used for this computation (see below). This can in particular impact initial-state radiations. We consider variations of the hard shower scale as detailed in Ref. [960], using the ‘resummation’ profile scale choice.

### *MoCaNLO + Recola*

The program MoCaNLO+RECOLA is made of a flexible Monte Carlo program dubbed MoCaNLO and of the general matrix element generator RECOLA [89, 481]. The program can compute arbitrary processes in the Standard Model with NLO QCD and EW accuracy. The fast integration is ensured by using similar phase-space mappings to those of Refs. [483, 484, 487]. The complex-mass scheme [483, 517] to treat unstable particles is always used. These tools have been successfully used for the computation of NLO corrections for high-multiplicity processes and in particular VBS processes [224, 226].

### *SHERPA*

SHERPA [415, 490] is a multipurpose event generator for high-energy particle collisions. It is built out of several algorithms and modules tailored to many different physics challenges of collider physics. In the SHERPA framework the infrared divergences appearing in the real-emission are treated by the Catani–Seymour dipole-subtraction method [187, 489, 496]. The default parton-shower algorithm of SHERPA is based on the Catani–Seymour factorisation [491, 961]. For NLO computations, interfaces to several one-loop generators exist, with the interface to RECOLA [89, 481] being the latest addition [499]. For correctly combining the matrix elements in multijet-production processes with the parton shower, SHERPA has adapted the MEPS@LO method [924] at LO and its generalisation MEPS@NLO [932] at NLO. In this study, the matrix elements are provided by COMIX [495], one of the two built-in generators next to AMEGIC [494]. Both of the studied partonic processes are then showered with the default shower.

### *VBFNLO*

VBFNLO [477, 847, 848] is a flexible Monte Carlo event generator for processes with electroweak bosons. Besides the Standard Model, selected processes can also be calculated in a variety of new-physics models, including effective field theories with dimension-6 and dimension-8 operators. The matrix elements for VBF and VBS processes are calculated in the VBS approximation. The corresponding  $s$ -channel contribution, which can be seen as triboson production where one vector boson decays hadronically, is available as well, but its contribution not included in the studies presented here. Its use of leptonic tensors in the calculation of the matrix elements can lead to a significant speed improvement compared to automatically generated code. For results with parton showers in this study, VBFNLO can serve as the matrix-element provider and phase-space generator for HERWIG7. The interface between the two programs is based on an extension of the BLHA standard [505, 506, 594]. When combining VBFNLO results with showering performed by Pythia, event files following the Les Houches LHE file standard [962, 963] are used instead.

## 3.3 Details of the set-up

In this section we describe default input parameters that have been used. Also, the event selection used in the comparison is described.

### *Input parameters*

All simulations are performed for the LHC running with a center-of-mass energy  $\sqrt{s} = 13$  TeV. The default PDF used is the NNPDF 3.0 set [422] with four active flavour at LO and a strong coupling constant  $\alpha_s(M_Z) = 0.130$ .<sup>4</sup> If this default PDF is not employed, it is explicitly stated. The masses and widths of the particle used in the simulations are

$$M_t = 173.21 \text{ GeV}, \quad \Gamma_t = 0 \text{ GeV},$$

---

<sup>4</sup>Its `lhaid` in LHAPDF6 [964] is 263400.

$$\begin{aligned}
M_Z^{\text{OS}} &= 91.1876 \text{ GeV}, & \Gamma_Z^{\text{OS}} &= 2.4952 \text{ GeV}, \\
M_W^{\text{OS}} &= 80.385 \text{ GeV}, & \Gamma_W^{\text{OS}} &= 2.085 \text{ GeV}, \\
M_H &= 125.0 \text{ GeV}, & \Gamma_H &= 4.07 \times 10^{-3} \text{ GeV}.
\end{aligned}
\tag{V.12}$$

The value of the mass and width of the Higgs boson are those recommended by the Higgs cross section working group [236]. The top quark does not appear at tree level in the simulations when the bottom-quarks in the initial state are neglected. Therefore its width is set to zero. The numerical values used in the simulation for the pole mass/width of the gauge bosons ( $V = W, Z$ ) are obtained from the measured on-shell (OS) values for the masses and widths according to Ref. [591] as:

$$M_V = M_V^{\text{OS}} / \sqrt{1 + (\Gamma_V^{\text{OS}} / M_V^{\text{OS}})^2}, \quad \Gamma_V = \Gamma_V^{\text{OS}} / \sqrt{1 + (\Gamma_V^{\text{OS}} / M_V^{\text{OS}})^2}.
\tag{V.13}$$

The EW coupling is renormalised in the  $G_\mu$  scheme [590] where

$$G_\mu = 1.16637 \times 10^{-5} \text{ GeV}^{-2}.
\tag{V.14}$$

The input parameters above yield a numerical value for  $\alpha$  of

$$\alpha = 7.555310522369 \times 10^{-3}.
\tag{V.15}$$

Note that for the EW contribution of order  $\mathcal{O}(\alpha^6)$ , no dependence on the strong coupling appears. For contributions (interference or QCD-induced contributions) where there is a dependence on  $\alpha_s$ , the numerical value used is the one extracted from the PDF set.

For the renormalisation and factorisation scales, two choices have been adopted. For the fixed scale, it is

$$\mu = \mu_{\text{fix}} = M_W,
\tag{V.16}$$

while the dynamical scale used is

$$\mu = \mu_{\text{dyn}} = \text{Max} [p_{T,j}].
\tag{V.17}$$

The latter should be understood as the maximum of the transverse momenta of the tagging jets (defined below). This observable is closely connected to the momentum transfer through the virtual vector boson, which has been shown to be a reasonable scale choice, and is also a reasonable choice when used as the starting scale of the parton shower [957]. These two scales are the default fixed and dynamical scale, respectively. If the scale used is different, this is explicitly stated in the text.

Photon-induced as well as bottom-induced contributions have been neglected. The photon contributions are expected to be small [226] while the bottom-induced contribution can lead to single-top resonant contributions. The later can in principle be isolated thanks to kinematic constraints.

### *Event selection*

Following experimental studies [953, 965], the event selection used for the present work is:

- All charged leptons are required to have

$$p_{T,\ell} > 20 \text{ GeV}, \quad |y_\ell| < 2.5.
\tag{V.18}$$

$\mu = \mu_{\text{dyn}} / \sigma_{\text{LO}}$ [fb]	$\text{pp} \rightarrow e^+ \nu_e \mu^+ \mu^- \text{jj}$	$\text{pp} \rightarrow e^- \bar{\nu}_e \mu^+ \mu^- \text{jj}$
$\mathcal{O}(\alpha^6)$	0.25416(6)	0.15003(3)
$\mathcal{O}(\alpha_s \alpha^5)$	0.006833(6)	0.003977(3)
$\mathcal{O}(\alpha_s^2 \alpha^4)$	0.9912(2)	0.6306(6)

Table V.2: Fiducial cross sections at LO for the process  $\text{pp} \rightarrow e^+ \nu_e \mu^+ \mu^- \text{jj}$  and  $\text{pp} \rightarrow e^- \bar{\nu}_e \mu^+ \mu^- \text{jj}$  at orders  $\mathcal{O}(\alpha^6)$ ,  $\mathcal{O}(\alpha_s \alpha^5)$ , and  $\mathcal{O}(\alpha_s^2 \alpha^4)$ . The predictions are expressed in fb and are for the LHC running at a centre-of-mass energy of  $\sqrt{s} = 13$  TeV. The scale used in the simulations is  $\mu = \mu_{\text{dyn}} = \text{Max} [p_{\text{T,j}}]$ . The integration errors of the last digits are given in parentheses.

- For the leptons of opposite charge and same flavour, an invariant mass cut to single out the Z-boson resonance is applied:

$$76 \text{ GeV} < m_{l_i^+ l_i^-} < 106 \text{ GeV}. \quad (\text{V.19})$$

- QCD jets are clustered thanks to the anti- $k_T$  algorithm [661] with radius parameter  $R = 0.4$ . At least two jets are required to have

$$p_{\text{T,j}} > 30 \text{ GeV}, \quad |y_j| < 4.7, \quad \Delta R_{j\ell} > 0.4, \quad (\text{V.20})$$

and are called tagging jets.

- On the two leading tagging jets, typical VBS cuts are applied:

$$m_{\text{jj}} > 500 \text{ GeV}, \quad |\Delta y_{\text{jj}}| > 2.5. \quad (\text{V.21})$$

These cuts have been issued either directly in the Monte Carlo programs or using a RIVET routine [500]. This file will be made public in order to make the present study easily reproducible.

### 3.4 Results

#### *Several contributions for one process*

As explained previously, the processes  $\text{pp} \rightarrow e^+ \nu_e \mu^+ \mu^- \text{jj}$  and  $\text{pp} \rightarrow e^- \bar{\nu}_e \mu^+ \mu^- \text{jj}$  possess at LO three contributions of orders  $\mathcal{O}(\alpha^6)$ ,  $\mathcal{O}(\alpha_s \alpha^5)$ , and  $\mathcal{O}(\alpha_s^2 \alpha^4)$ . As it can be seen in Table V.2, the EW component represents only about 20% of the total cross section<sup>5</sup>. This is in contrast with the  $W^\pm W^\pm \text{jj}$  signature where the EW component represents almost 90% of the cross section [226] in a comparable fiducial volume. Hence measuring the EW component is much more challenging. Therefore, inferring the shape of the signal and irreducible background is key. Note that the interference contribution is about 0.5% which is negligible with respect to current experimental accuracy.<sup>6</sup>

In Fig. V.6, several differential distributions are shown. In the upper plot the absolute predictions for each component as well as their sum are displayed. In the lower plot, each contribution is normalised to their sum and expressed in percent. These distributions reflect the same conclusion as for the cross section namely that the processes are largely dominated by QCD-induced contribution. The first two observables displayed (top) are the invariant mass and the rapidity difference of the two tagging jets. These two observables are used as cuts [see Eq. (V.21)]

<sup>5</sup>In this subsection and in particular for the results of Table V.2 and Fig. V.6, the cut  $|y_j| < 4.5$  as been used instead of 4.7 as everywhere else. This has a 1% effect on the fiducial cross section.

<sup>6</sup>For the  $W^\pm W^\pm \text{jj}$  signature, the interference contribution has been shown to be around 3% [226].

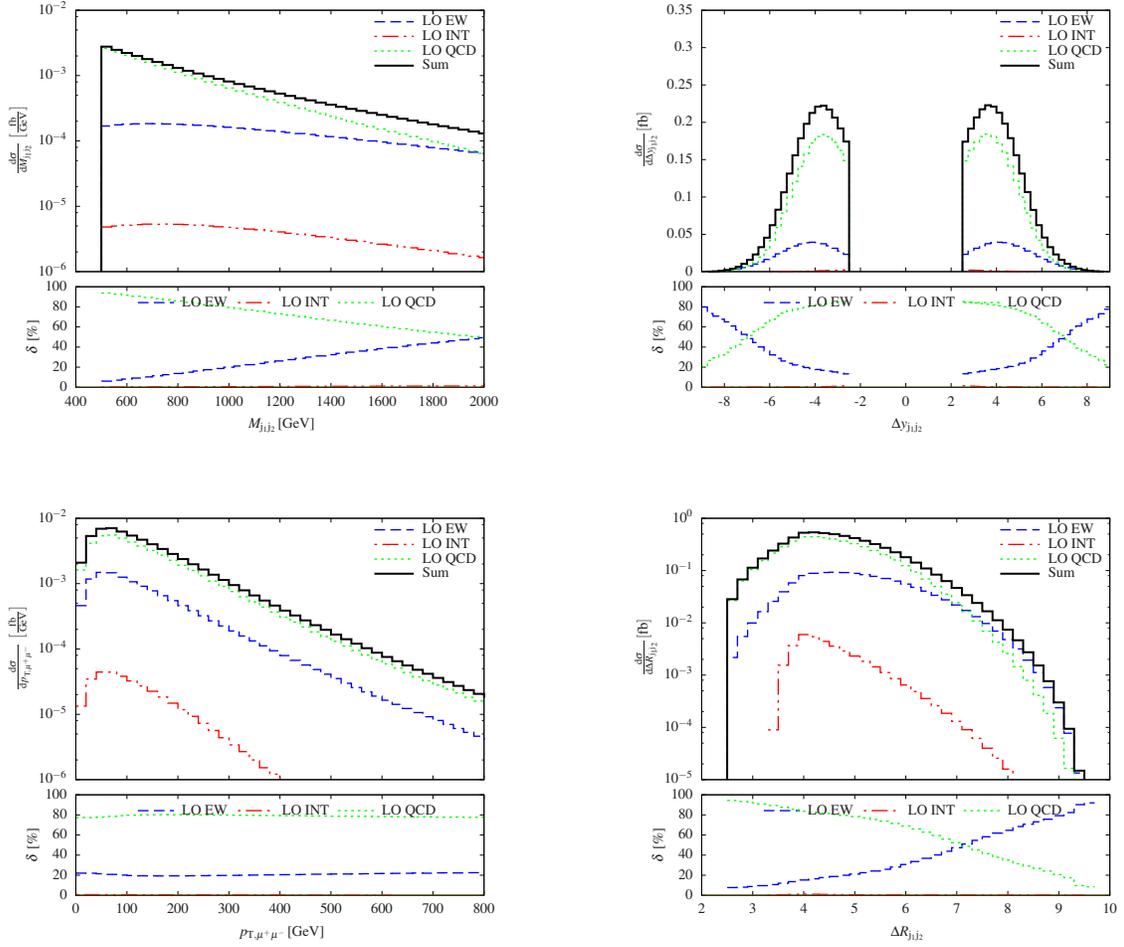


Fig. V.6: Differential distributions at a centre-of-mass energy  $\sqrt{s} = 13$  TeV at the LHC for  $pp \rightarrow e^+ \nu_e \mu^+ \mu^- jj$  with the contributions at orders  $\mathcal{O}(\alpha^6)$  (EW),  $\mathcal{O}(\alpha_s \alpha^5)$  (INT), and  $\mathcal{O}(\alpha_s^2 \alpha^4)$  (QCD): invariant mass of the two jets (top left), rapidity separation between the two jets (top right) transverse momentum of the anti-muon–muon system (bottom left), and distance between the two jets (bottom right). The upper panels show the LO predictions as well as their sum. The lower panels display the respective contributions normalised to their sum.

in order to enhance the EW component over the QCD background. The cuts are clearly visible on the plots and it is easily understandable why they are enhancing the EW contribution. Toward high invariant-mass, the EW component becomes more and more important. It even becomes of the same size as the QCD one for an invariant-mass of 2000 GeV. The same holds true for high rapidity separation for the two jets. On the other hand, the transverse invariant mass (bottom left) of muon–anti-muon pair does not display significant differences in the different contributions over the whole range. Other transverse-momentum distributions display the same pattern. Finally, we show the distance between the two jets, defined as  $\Delta R_{jj} = \sqrt{\Delta\eta(j_1, j_2)^2 + \Delta\phi(j_1, j_2)^2}$ . This observable also seems to possess a good discriminating power. In particular, for large distances, the EW component becomes dominant but with very low statistics. The interference effects are extremely suppressed due to the smallness of the cross section. Contribution of each of the studied observable does not exceed the per-cent level over the whole phase-space range.

$\mu = \mu_{\text{fix}} / \sigma_{\text{LO}}^{\text{EW}}$ [fb]	pp $\rightarrow e^+ \nu_e \mu^+ \mu^- \text{jj}$	pp $\rightarrow e^- \bar{\nu}_e \mu^+ \mu^- \text{jj}$
MADGRAPH5_aMC@NLO	0.2857(8)	0.1657(4)
MoCANLO+RECOLA	0.2885(1)	0.16718(3)
SHERPA	0.2885(2)	0.1670(2)
VBFNLO	0.2867(5)	0.1661(3)

Table V.3: Fiducial cross sections at LO for the process  $\text{pp} \rightarrow e^+ \nu_e \mu^+ \mu^- \text{jj}$  and  $\text{pp} \rightarrow e^- \bar{\nu}_e \mu^+ \mu^- \text{jj}$  at order  $\mathcal{O}(\alpha^6)$ . The predictions are expressed in fb and are for the LHC running at a centre-of-mass energy of  $\sqrt{s} = 13$  TeV. The scale used in the simulations is  $\mu = \mu_{\text{fix}} = M_W$ . The integration errors of the last digits are given in parentheses.

$\mu = \mu_{\text{dyn}} / \sigma_{\text{LO}}^{\text{EW}}$ [fb]	pp $\rightarrow e^+ \nu_e \mu^+ \mu^- \text{jj}$	pp $\rightarrow e^- \bar{\nu}_e \mu^+ \mu^- \text{jj}$
MADGRAPH5_aMC@NLO	0.2550(8)	0.1492(5)
MoCANLO+RECOLA	0.2574(2)	0.15003(3)
SHERPA	0.2574(2)	0.14998(6)

Table V.4: Fiducial cross sections at LO for the process  $\text{pp} \rightarrow e^+ \nu_e \mu^+ \mu^- \text{jj}$  and  $\text{pp} \rightarrow e^- \bar{\nu}_e \mu^+ \mu^- \text{jj}$  at order  $\mathcal{O}(\alpha^6)$ . The predictions are expressed in fb and are for the LHC running at a centre-of-mass energy of  $\sqrt{s} = 13$  TeV. The scale used in the simulations is  $\mu = \mu_{\text{dyn}} = \text{Max}[p_{T,j}]$ . The integration errors of the last digits are given in parentheses.

### Parton level comparisons

We start the comparison of the various predictions by a comparison at the level of the cross section. The fiducial volume is the one described in Eqs. (V.18)–(V.21). The results are documented in Tables V.3 and V.4, which give the cross sections for both processes  $\text{pp} \rightarrow e^+ \nu_e \mu^+ \mu^- \text{jj}$  and  $\text{pp} \rightarrow e^- \bar{\nu}_e \mu^+ \mu^- \text{jj}$  at fixed and dynamical scales. The predictions of MoCaNLO + RecoLa and SHERPA are generally in perfect agreement, but those of MADGRAPH5\_aMC@NLO show slight statistical disagreement for fixed scale. In this case the difference between MoCaNLO + RecoLa/SHERPA and MADGRAPH5\_aMC@NLO is about one percent as it can be seen in Table V.3. We note that the source of the difference is likely unchanged in the dynamic scale result, where the setup is otherwise unchanged, but is not clear due to the higher statistical uncertainty.

The predictions of VBFNLO are not expected to be in perfect agreement with the others as VBFNLO is not using full matrix elements but VBS-approximated ones. Nonetheless, the difference between the VBFNLO predictions and the ones of MoCaNLO + RecoLa/SHERPA amounts to 0.6% for the fixed scale.

As it can be seen in Fig. V.7, where the invariant mass and rapidity separation of the two tagging jets are displayed, the agreement between the predictions is good in both shape and normalisation. This is particularly true for the jet rapidity separation, where all predictions agree within statistical uncertainties over the whole range. On the other hand, for the di-jet invariant mass, subtle but statistically significant differences are more pronounced towards higher masses between MADGRAPH5\_aMC@NLO's predictions and the others. We do not see such discrepancies in other distributions, however, we see that this discrepancy is accentuated when relaxing the cuts applied at generation time through the run card and manually applying the selection to the LHE events, as shown in Fig. V.8.

Concerning the validation of the VBS approximation (MoCaNLO+RECOLA vs.

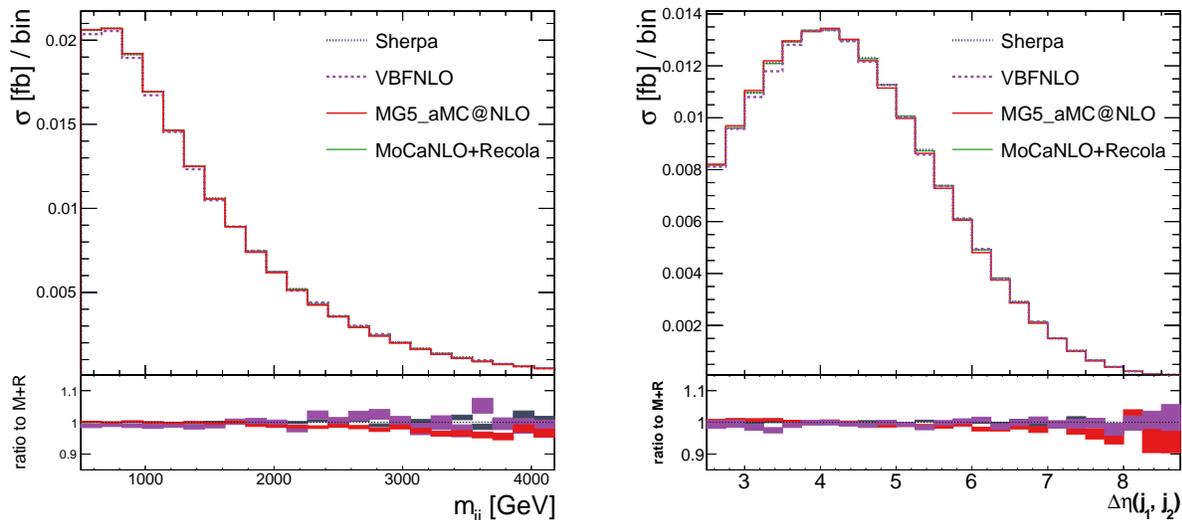


Fig. V.7: Differential distributions computed at fixed-order with centre-of-mass energy  $\sqrt{s} = 13$  TeV at the LHC for  $pp \rightarrow e^- \nu_e \mu^+ \mu^- jj$  at LO with fixed scaled  $\mu = M_W$ : invariant mass of the two jets (left), rapidity separation between the two jets (right). The predictions in the lower plot are normalised to the prediction of MoCaNLO +Recola. The shaded bands indicate the relative statistical uncertainty by bin for each sample. The statistical uncertainty on the MoCaNLO +Recola predictions is shown in grey, other samples are indicated in the ratio with the color indicated in the legend.

VBFNLO for example), both predictions are in very good agreement as at the level of the cross section. This supports the findings of Ref. [954] where preliminary results for similar comparisons for  $W^\pm W^\pm jj$  have been reported. This means that the VBS approximation (VBFNLO) approximates rather well the full computation (MOCANLO+RECOLA) in the fiducial region chosen.

We stress that differences of configuration should be considered independently of typical estimates of theoretical uncertainties such as QCD scale and PDF uncertainties. To illustrate this, we compute the PDF uncertainty for the NNPDF3.0 set and the two scale choices considered here. The PDF uncertainty is evaluated to be 3–5% using MADGRAPH5\_aMC@NLO. Scale uncertainties are evaluated using the typical prescription of varying  $\mu_R$  and  $\mu_F$  subject to the constraint  $1/2 \leq \mu_F/\mu_R \leq 2$ , using MADGRAPH5\_aMC@NLO and cross-checked with MoCaNLO +Recola, and are found to be between 7–10% for the scale choices considered. The full results obtained with MADGRAPH5\_aMC@NLO are shown in Table V.5.

Scale choice	$pp \rightarrow e^+ \nu_e \mu^+ \mu^- jj$	$pp \rightarrow e^- \bar{\nu}_e \mu^+ \mu^- jj$
$\mu_R = \mu_{\text{fix}} = m_W$	$0.286^{+9.2\%}_{-7.8\%} \pm 3.7\%$	$0.166^{+9.0\%}_{-7.7\%} \pm 4.3\%$
$\mu = \mu_{\text{dyn}} = \text{Max}[p_{T,j}]$	$0.255^{+8.0\%}_{-6.9\%} \pm 3.7\%$	$0.149^{+9.0\%}_{-7.7\%} \pm 4.3\%$

Table V.5: Fiducial cross sections at LO for the process  $pp \rightarrow e^+ \nu_e \mu^+ \mu^- jj$  and  $pp \rightarrow e^- \bar{\nu}_e \mu^+ \mu^- jj$  at order  $\mathcal{O}(\alpha^6)$  via MADGRAPH5\_aMC@NLO. The predictions are expressed in fb and are for the LHC running at a centre-of-mass energy of  $\sqrt{s} = 13$  TeV. Uncertainties are expressed as  $\sigma_{-\delta_{\text{scale}}}^{+\delta_{\text{scale}}} \pm \delta_{\text{PDF}}$ .

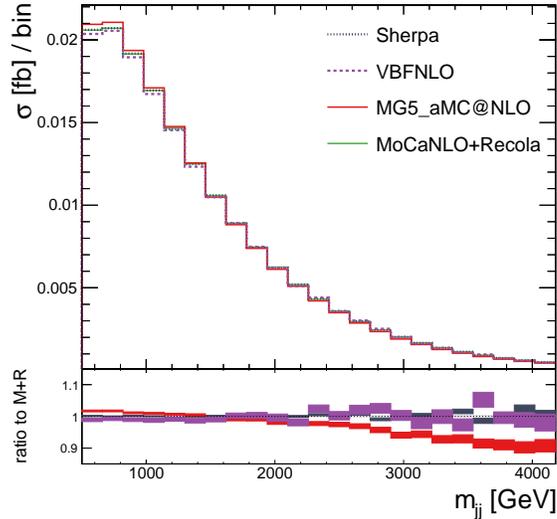


Fig. V.8: Invariant mass of the two jets computed at fixed-order with centre-of-mass energy  $\sqrt{s} = 13$  TeV at the LHC for  $pp \rightarrow e^- \nu_e \mu^+ \mu^- jj$  at LO with fixed scaled  $\mu = M_W$ . The predictions in the lower plot are normalised to the prediction of MoCaNLO +Recola. The shaded bands indicate the relative statistical uncertainty by bin for each sample. The statistical uncertainty on the MoCaNLO +Recola predictions is shown in grey, other samples are indicated in the ratio with the color indicated in the legend. MADGRAPH5\_aMC@NLO events are generated using loose selections in the run card before applying the selection on the LHE events. The trend of disagreement in high  $m_{jj}$  is accentuated for looser generation cuts, pointing to a possible phase-space integration effect.

We also compare the effect of configurations changes using the general-purpose generators SHERPA and MADGRAPH5\_aMC@NLO. These generators are able to compute arbitrary processes in the standard model, and therefore have default configurations designed to cover a wide range of processes. It is thus advisable to explicitly configure settings most appropriate to the process considered.

Using the same fiducial definition, but with the default parameter settings (defined in param\_card.dat) for MADGRAPH5\_aMC@NLO, we obtain a cross section of 0.1561(5) fb for the  $\text{Max}[p_{T,j}]$  scale, an increase of 4% from the nominal value. The primary source of this difference is the settings of the boson masses and widths, which are set to their LO values by default in MADGRAPH5\_aMC@NLO. Such a setting can be considered appropriate in cases where the partial widths of many decay channels must be considered together, but for an explicitly leptonic process where no gauge boson is set on-shell, we argue that the best measured values are more appropriate. We do not believe that this 4% difference constitutes an additional uncertainty, but warn that careful configuration of such parameters should be considered.

We similarly study the default dynamical scale choices in MADGRAPH5\_aMC@NLO and SHERPA. In both cases, this choice is motivated by a desire for broad application to a variety of processes. Sherpa uses the inverted parton shower to cluster the matrix element onto a core  $2 \rightarrow 2$  process. This procedure determines the scales and is especially suited for truncated shower merging. For MADGRAPH5\_aMC@NLO, the default scale depends on the order of the process and shower settings. At LO without merging of parton multiplicities, the scale is set to the central  $m_T$  scale after  $k_T$ -clustering of the event. If the last clustering is a  $t$ -channel

colourless exchange, the scale is set to the last  $m_T$  values on either side <sup>7</sup>. The resulting cross section is 5% greater than the  $\text{Max}[p_{T,j}]$  result. Using both the default scale and parameter settings (note however, that we still specify the PDF) we obtain a value of  $\sigma_{\text{default}} = 0.1638(5)$  fb, 9% greater than our configured  $\mu = \text{Max}[p_{T,j}]$  value. We reiterate that this difference, which is of comparable size to the phenomenological uncertainties, should not necessarily be considered on equal footing. The fact that the functional form of the dynamical scale choice can have a significant impact compared to the usual factor of two variation around a nominal value is well established. One should therefore take care to select an appropriately motivated choice. Additional uncertainties may be appropriate, but should be considered with care rather than taking a broad envelope of less motivated choices. This comparison additionally highlights the advantage of the reduced scale dependency at NLO-QCD accuracy.

### *Comparisons after parton shower*

In Figs. V.9 and V.10, a comparison of results obtained with different generators for the process  $pp \rightarrow e^- \nu_e \mu^+ \mu^- jj$  at LO supplemented with parton shower with fixed scaled  $\mu = M_W$  is shown. The predictions have been obtained from VBFNLO and MADGRAPH5\_aMC@NLO in association with PYTHIA8, VBFNLO with HERWIG7, and SHERPA with its own parton shower. In Fig. V.9 several differential distributions are shown: the invariant mass of the two jets, the rapidity separation between the two jets, the transverse momentum of the anti-muon–muon system, and the distance between the two jets. The error band in the plots represents the statistical error of the Monte Carlo integrations. The overall picture is that the predictions obtained from SHERPA, VBFNLO +PYTHIA8, and VBFNLO +HERWIG7 agree rather well over the whole kinematic range. This is not the case for the predictions of MADGRAPH5\_aMC@NLO +PYTHIA8 which differ by up to 20% in certain phase-space regions. Because of kinematic changes in jet and lepton kinematics due to the parton shower, it is necessary to generate events with a selection looser than the one used in the fixed-order fiducial definition and applied in the Rivet routine. The showered events for MADGRAPH5\_aMC@NLO are therefore generated with the conditions of Fig. V.8. The disagreement at high di-jet invariant mass, already seen at fixed order, is thus not an effect of the parton shower. The normalisation differences are also consistent with a reduction in acceptance into the fiducial region for the MADGRAPH5\_aMC@NLO sample due to this trend.

Differences are more pronounced in Fig. V.10 where the Zeppenfeld variable for the three charged leptons and the third jet as well as the number of jets are displayed. The Zeppenfeld variable for a given particle  $X$  is defined as

$$z_X = \frac{y_X - \frac{y_{j_1} + y_{j_2}}{2}}{|y_{j_1} - y_{j_2}|}, \quad (\text{V.22})$$

where  $y_{j_{1/2}}$  are the rapidity of the first and second hardest jet, respectively. The Zeppenfeld variable of the third jet, as well as the number of jets beyond two, are observables that are not defined at LO, and are only non-zero thanks to the emissions of the parton shower. It is thus expected that these feature a worse agreement than the previously discussed observables, as significantly different algorithms are employed by the parton shower generators considered. Here the differences can reach up to 100% in the central region. The fact that PYTHIA predicts more central jet activity has already been observed for the  $W^\pm W^\pm jj$  signature in preliminary results of a comparative study [954] but also for the Zjj signature [862, 966], where the predictions are compared with data. We note we have not tuned the parameters and algorithms of the shower but rather consider the spread of predictions as reflective of the uncertainty of parton

<sup>7</sup>The procedure is described at <https://cp3.irmp.ucl.ac.be/projects/madgraph/wiki/FAQ-General-13>.

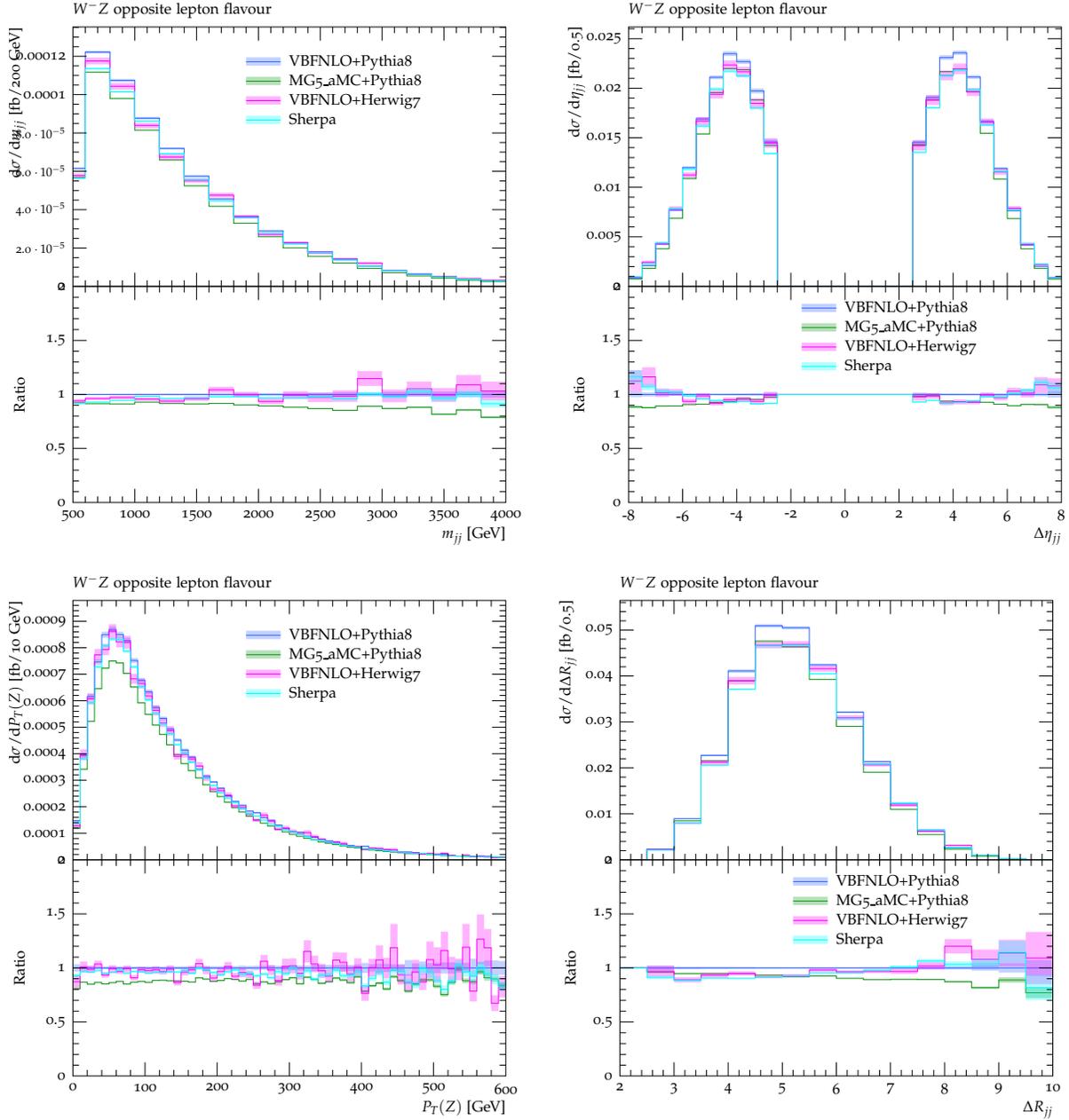


Fig. V.9: Differential distributions at a centre-of-mass energy  $\sqrt{s} = 13\text{TeV}$  at the LHC for  $pp \rightarrow e^- \nu_e \mu^+ \mu^- jj$  at LO matched with parton shower with fixed scaled  $\mu = M_W$ : invariant mass of the two jets (top left), rapidity separation between the two jets (top right), transverse momentum of the anti-muon-muon system (bottom left), and distance between the two jets (bottom right). In the lower plot, the normalisation is with respect to the VBFNLO +PYTHIA8 predictions. The error band represents the statistical error of the Monte Carlo integrations.

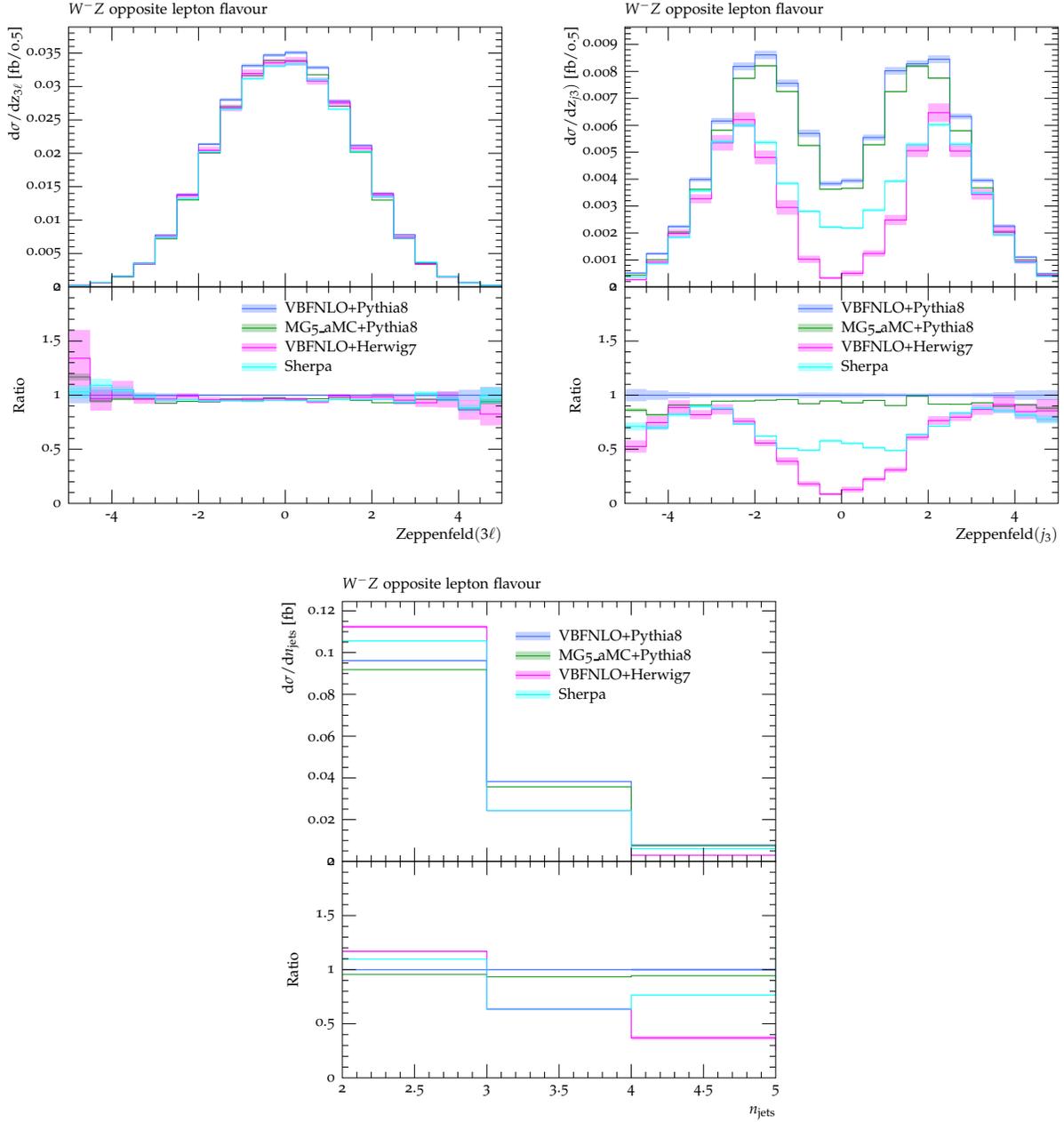


Fig. V.10: Differential distributions at a centre-of-mass energy  $\sqrt{s} = 13\text{TeV}$  at the LHC for  $pp \rightarrow e^- \nu_e \mu^+ \mu^- jj$  at LO matched with parton shower with fixed scaled  $\mu = M_W$ : Zeppenfeld variable for the three leptons (top left), Zeppenfeld variable for the third jet (top right), and number of jets (bottom). In the lower plot, the normalisation is with respect to the VBFNLO +PYTHIA8 predictions. The error band represents the statistical error of the Monte Carlo integrations.

shower dominated observables. While variables such as the Zeppenfeld of the third jet have known separation power between the EW and QCD induced production, tuning an experimental selection on this observable would introduce large theoretical uncertainties, especially if only LO predictions are considered. A similar argument holds for a veto on extra jet activity.

In Figs. V.11 and V.12 results obtained using the fixed scale  $\mu = M_W$  and dynamic scale  $\mu = \text{Max} [p_{T,j}]$  are compared. In particular, predictions for MADGRAPH5\_aMC@NLO and SHERPA for both scales are presented. For MADGRAPH5\_aMC@NLO this prediction is obtained by reweighting each event for the difference between the matrix-element calculation with the fixed scale and the dynamic scale. The statistical uncertainty is therefore largely correlated between the two predictions for this Monte Carlo. The observables displayed are the same as for the previous comparison. For the invariant mass of the two jets as well as the transverse momentum of the anti-muon–muon system, for both generators, the use of fixed scale enhances the predictions toward high transverse momentum. For the rapidity separation of the two jets as well as the distance between the two jets, the shape difference between fixed and dynamical scale is not present for MADGRAPH5\_aMC@NLO. On the other hand, in SHERPA, the use of fixed scale enhances the predictions for small separations.

Concerning the last set of observables (Zeppenfeld variable of the three leptons and third jet as well as the number of jets) in Fig. V.12, the use of fixed or dynamical scale does not have a large impact. The only clearly visible difference between fixed and dynamical scale is the normalisation. This effect is already observed at the level of the fiducial cross section in Tables V.3 and V.4. The effect of the different scale amounts to a change in normalisation of about 12% for both  $pp \rightarrow e^+ \nu_e \mu^+ \mu^- jj$  and  $pp \rightarrow e^- \bar{\nu}_e \mu^+ \mu^- jj$  processes.

### 3.5 Conclusions

The measurement of the EW component of the  $pp \rightarrow W^\pm Z jj$  process constitutes a real challenge for experimental collaborations. Such a measurement is complicated for several reasons including the low cross section and the large irreducible background. Therefore, these measurements have to rely on theoretical predictions in order to extract a signal, which makes theoretical predictions implemented in various Monte Carlo programs very important. It is thus key to have a good control over these predictions. To that end we have performed LO comparisons of different theoretical predictions with and without shower and hadronisation. The comparisons have all been performed with pre-defined input values for the matrix element calculation and a generic fiducial region for VBS measurements.

The first finding of the present study is that the EW component is overwhelmingly dominated by its irreducible QCD background. In the fiducial region that we have chosen, 80% of the fiducial cross section is accounted by the QCD background. This indicates the need for observables that enhance the EW contribution. Beyond the di-jet invariant mass and the rapidity separation between the two jets, we found that the distance between the two jets could have a good discriminating power. We also found that interferences between the QCD and EW amplitudes are also negligible in the fiducial region.

In the present proceedings, we have also performed a LO comparison. At parton level, the agreement between the various theoretical predictions is generally very good. This statement holds both for the inclusive cross section and differential distributions. Nonetheless differences can appear especially with respect to MADGRAPH5\_aMC@NLO predictions. These can probably be attributed to the subtle miss-configuration of the runs presented here or underestimated statistical errors. On the other hand, the differences between the full computations and the VBS-approximated one are relatively small. This implies that, as for the  $W^\pm W^\pm jj$  signature [954], given the present experimental precision, the VBS approximation is satisfactory in the typical

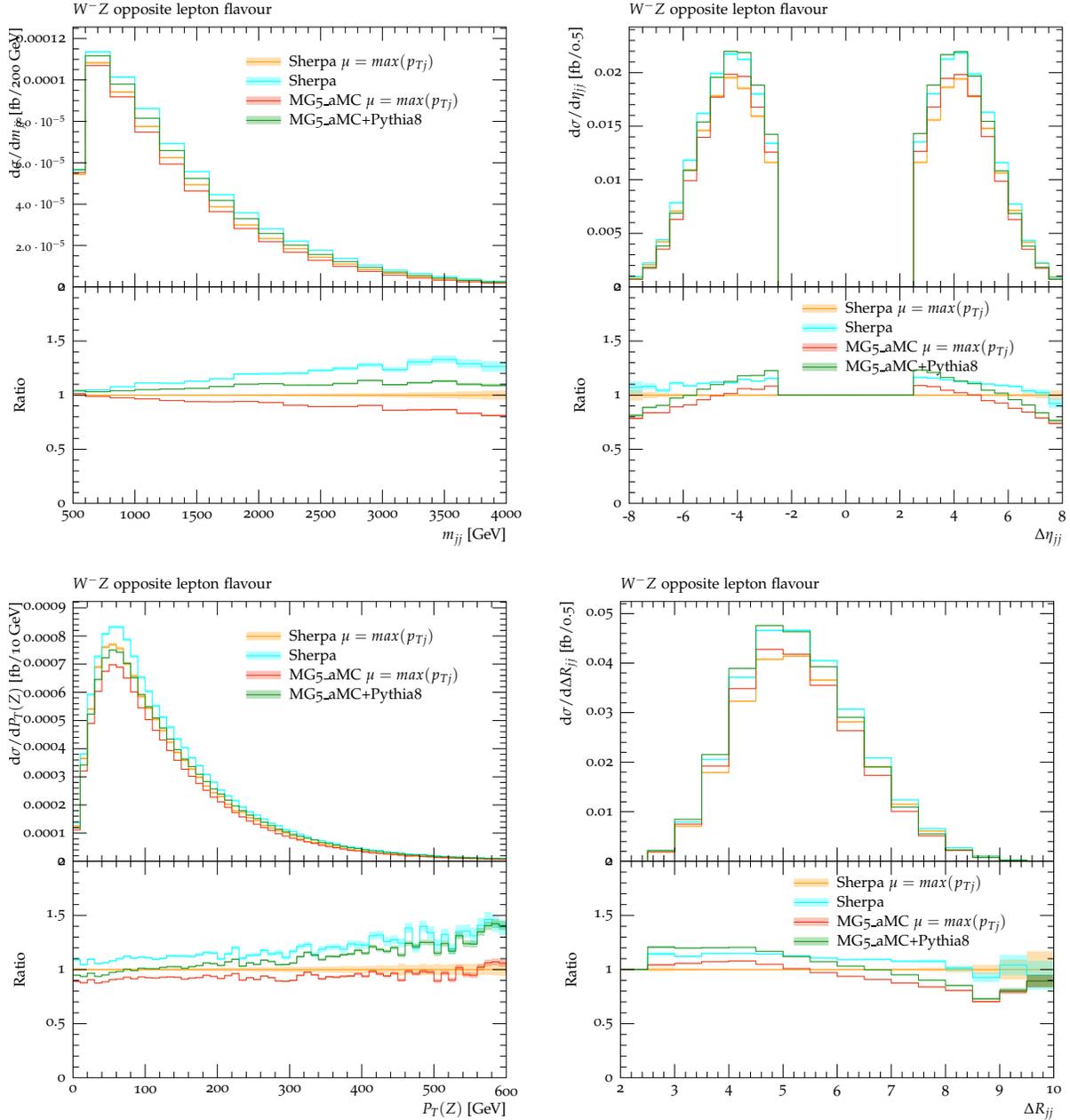


Fig. V.11: Differential distributions at a centre-of-mass energy  $\sqrt{s} = 13\text{TeV}$  at the LHC for  $pp \rightarrow e^- \nu_e \mu^+ \mu^- jj$  at LO matched with parton shower with fixed and dynamic scale values: invariant mass of the two jets (top left), rapidity separation between the two jets (top right), transverse momentum of the anti-muon-muon system (bottom left), and distance between the two jets (bottom right). In the lower plots, the predictions are normalised to the prediction of SHERPA with dynamical scale. The error band represents the statistical error of the Monte Carlo integrations. For the two MADGRAPH5\_aMC@NLO calculations this statistical error is largely correlated, as explained in the text.

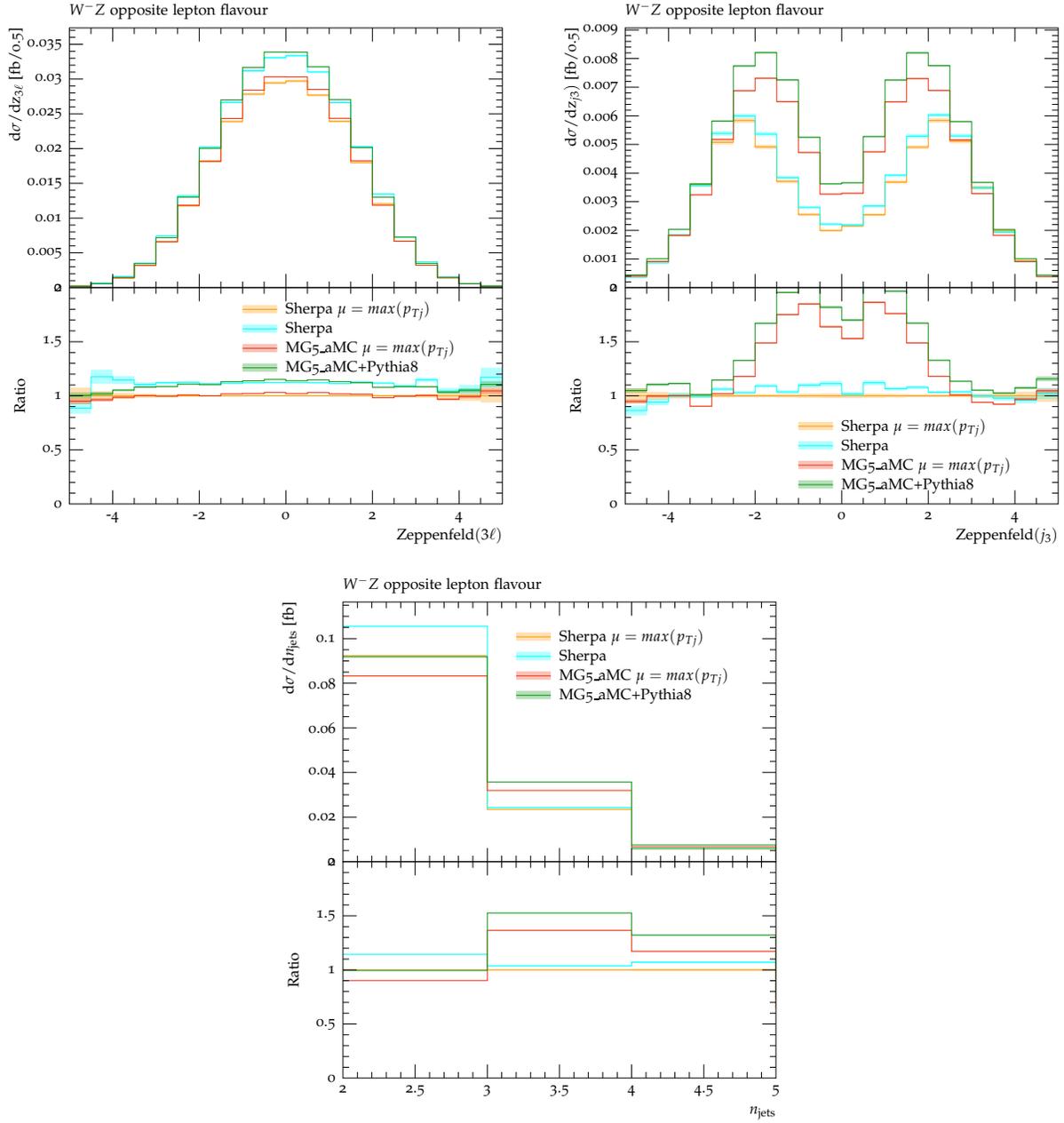


Fig. V.12: Differential distributions at a centre-of-mass energy  $\sqrt{s} = 13\text{TeV}$  at the LHC for  $pp \rightarrow e^- \nu_e \mu^+ \mu^- jj$  at LO matched with parton shower with fixed and dynamic scale values: Zeppenfeld variable for the three leptons (top left), Zeppenfeld variable for the third jet (top right), and number of jets (bottom). In the lower plots, the predictions are normalised to the predictions of SHERPA with dynamical scale. The error band represents the statistical error of the Monte Carlo integrations. For the two MADGRAPH5\_aMC@NLO calculations this statistical error is largely correlated, as explained in the text.

fiducial regions used by experimental collaborations for their measurements.

The predictions at LO supplemented with parton shower have relied on the fixed-order configuration. No tuning of the parton shower parameters have been performed and the default values have been taken. For PYTHIA8, the CUETP8M1 tune [967] which is based on the Monash tune [631] has been used. The agreement is in general worse, in particular for observables that are defined only beyond LO, where the theoretical predictions diverge significantly. Concerning the role of the fixed and dynamical scales, we have found that they can have a rather large influence on the predictions for the inclusive cross section and differential distributions. The inclusion of higher order predictions may cure this behaviour.

Finally, we would like to stress that these results constitute only a preliminary study of theoretical predictions for the process  $pp \rightarrow W^\pm Zjj$ . In particular, with a strictly fixed-order computation, some differences have been observed. These propagate to LO predictions matched to parton shower. In addition, for observables defined beyond LO, large differences due to the parton-shower algorithms appear. This study demonstrates the difficulty to obtain consistent Monte Carlo predictions. In particular, the choice of inputs and configuration can have a large impact on physics results. The effects of NLO QCD and EW corrections have not been studied in this work. They play a significant role in theoretical predictions for VBS and should be taken into account as much as possible in future studies. Overall, this study presents a benchmark for the performance of leading-order Monte Carlo generators for the simulation of vector boson scattering, and motivates further efforts in the above-mentioned directions that have not been fully addressed here.

## Acknowledgements

The work of SB is supported by German Federal Ministry for Education and Research (BMBF) under contract 05H15MGCAA. MP acknowledges financial support by the BMBF under contract no. 05H15WWCA1 and the German Science Foundation (DFG) under reference number DE 623/6-1. EY is supported by Chinese Academy of Sciences (CAS), MoST (Ministry of Science and Technology) through project no. 2013CB837801, and NSFC (National Natural Science Foundation of China) through project no. 11661141007.

## 4 Considerations for the underlying event in jet calibration at the LHC <sup>8</sup>

The jet reconstruction algorithms used by the two multi-purpose experiments ATLAS and CMS at the LHC provide jets calibrated at the level of stable particle jets. These particles are produced by the fragmentation of the hard scattered partons and by the underlying event generated in the same proton–proton collision. The corresponding calibration procedures apply corrections that not only remove the signal contribution from diffuse particle emissions from multiple proton collisions in the same bunch crossing, but also signal remnants from previous (or future) bunch crossings affecting the detector signals. These corrections typically reduce or remove the underlying event contribution to the experimental jets to an indeterminate degree. It is desirable to add the underlying event contribution back into the truth jet definition, as the detector jet is supposed to be calibrated to the full particle level for physics analysis. The recovery of this contribution in the applied jet calibration procedures is an inherent feature of the Monte Carlo based jet energy scale correction, which provides the absolute calibration for jets in simulation and experiment. The model-dependent nature of this approach is emphasized and the considerations of model variations in the determination of the systematic uncertainties are summarized.

---

<sup>8</sup> J. Huston, P. Loch

## 4.1 Introduction

The jet calibration applied by both ATLAS [834] and CMS [968] includes a necessary correction to remove signal contributions from multiple proton–proton interactions happening in the same LHC bunch crossing (*in-time pile-up*), together with contributions from preceding and possibly even following bunch crossings (*out-of-time pile-up*). In both Run 1 and Run 2, approaches based on the jet-area-based correction described in Ref. [837] have been used. Its stochastic nature does not distinguish between pile-up contributions and contributions from the underlying event (UE) when subtracting the estimated pile-up signal from a reconstructed jet. An additional consideration when discussing the UE signal inside and outside of jets is that the significant pile-up at high LHC intensities, with a typical average number of pile-up collisions per bunch crossing of about  $\langle\mu\rangle \approx 20$  towards the end of Run 1, and rising to around  $\langle\mu\rangle \approx 60$  during the recently concluded 2017 data taking period of Run 2, may render the UE contribution insignificant. Nevertheless, after corrections are applied to the detector jet to mitigate the effect of pile-up at any level, the UE contribution approaches scales of relevance, in particular for jets with low transverse momentum. This is even more important particular true when considering the typical precision reached for jet energy measurements at LHC, which in large regions of phase space approaches 1%.

This note briefly summarises a typical jet calibration scheme and the resulting representation of the UE in the calibrated jet signal. The strategies are similar for CMS and ATLAS, even though the detector signals entering the jet calibration are different. The discussion here focusses on the UE signal and the jet calibration procedures applied in ATLAS.

## 4.2 Pile-up mitigation in jet reconstruction

The jet reconstruction in the two multi-purpose experiments at the LHC uses particle-flow objects in CMS [969] and clusters of topologically connected calorimeter cell signals (topo-clusters) in ATLAS [970]. The principal pile-up mitigation technique applied in both experiments is the (scalar) jet-area-based pile-up subtraction method introduced in Ref. [837]. It is applied right after jet reconstruction, before any absolute calibration.

The set of topo-clusters in ATLAS reconstructed for a given event encompasses contributions from the hard scatter final state and the two softer contributions, UE and pile-up. Both pile-up and UE generally manifest themselves in generating additional topo-clusters with low energies (*direct* contribution). The inherent possibility of merging of signals of particles from the three sources in the topo-cluster formation algorithm, together with the splitting of topo-clusters implicit to this algorithm, leads to complex hard scatter signal modifications by the two soft contributions (*indirect* contribution) with significant variations from event to event, even with similar levels of soft activity. In addition, the overlap in phase space and the basically identical emission characteristics in terms of average transverse momentum flow between pile-up and UE makes it basically impossible to remove the pile-up signal contribution without affecting the UE contribution, in particular in high pile-up conditions like the ones observed at the end of Run 1 and during Run 2.

The formation of topo-clusters is guided by signal significance (signal-to-noise) patterns steered by thresholds adjusted to expected noise levels for a given run year at LHC. The increasing  $\langle\mu\rangle$  in LHC operations from early to late Run 1, from Run 1 to Run 2, and recently in Run 2 leads to a reduced sensitivity of the ATLAS calorimeter to UE in general.<sup>9</sup> The noise thresholds rise approximately proportional to  $\sqrt{\langle\mu\rangle}$ , which means that e.g. the energy needed to

---

<sup>9</sup>The transverse momentum scale of UE contributions is set by  $\hat{s}$  in each hard scatter proton–proton interaction, and thus changed only slightly in Run 1 when LHC moved from  $\sqrt{s} = 7$  TeV to  $\sqrt{s} = 8$  TeV. It changed a bit more significantly from Run 1 to Run 2, when  $\sqrt{s} = 8$  TeV increased to  $\sqrt{s} = 13$  TeV.

seed a topo-cluster at  $\langle\mu\rangle = 40$  is two times higher than at  $\langle\mu\rangle = 10$ . This renders a significant portion of the UE invisible at higher pile-up conditions, especially in signals outside of jets.

All topo-clusters with a positive energy signal are used for finding jets. This means that the topo-clusters collected into one jet include direct and indirect contributions from all three sources, and the total jet kinematics, which is reconstructed from the four-momentum recombination of all these clusters, is rather sensitive in particular to pile-up, which dominates at high proton intensities.

Early in Run 1, when pile-up was at a relatively low level ( $\langle\mu\rangle < 10$ ), ATLAS used a purely Monte Carlo (MC) simulation based mitigation technique, which actually formally preserved the UE contribution in a jet as much as it was visible in the detector at all. In this approach the transverse momentum ( $p_T$ ) contribution from pile-up was measured by the dependence of the ratio of the initial reconstructed jet  $-p_T$  to the expected  $p_T$  from a matched particle jet in MC simulations. This dependence was evaluated with respect to the in-time pile-up represented by the number of reconstructed primary vertices  $N_{PV}$  and to the out-of-time pile-up measured by the average number of pile-up collisions  $\mu$  around the signal event (see Ref. [970] for details). The correction was then applied to each jet in MC simulations and experimental data using the average slopes  $\partial p_T/\partial N_{PV}$  and  $\partial p_T/\partial\mu$ , together with the obvious assumption that the transverse momentum contribution from pile-up is expected to vanish for  $N_{PV} = 1$  and  $\mu = 0$ . As the UE contribution is expected to not depend on either  $N_{PV}$  or  $\mu$ , this approach in principle does not remove it from the jet signal. It was purely MC based and required complex validation in data [971]. In addition, this approach is only sensitive to the expected fluctuations in the in-time pile-up activity, where the number of additional proton–proton collisions fluctuates from bunch crossing to bunch crossing following a Poisson distribution around  $\langle\mu\rangle$ . Similarly, the out-of-time pile-up induced calorimeter signal modifications are only corrected at an average level. Further fluctuations in the pile-up contribution to the reconstructed jet- $p_T$  even in events recorded in bunch crossings with the same  $N_{PV}$  and  $\mu$ , which reflect varying  $p_T$ -flow patterns in the in-time pile-up collisions as well as in surrounding (previous) bunch crossings, are not addressed and limit the efficiency of this pile-up mitigation technique.

The recent approach implements the jet-area-based pile-up mitigation technique suggested in Ref. [837]. In this case the calorimeter signal itself is used to determine the level of pile-up in each event. This signal is sensitive to the in- and out-of-time  $p_T$ -flow patterns, and this technique thus reduces bunch crossing to bunch crossing fluctuations. The initially reconstructed small radius jet with transverse momentum  $p_T^{\text{jet,raw}}$  is corrected by using its area  $A^{\text{jet}}$  and the median transverse momentum density  $\rho$  to a basic transverse momentum measure  $p_T^{\text{jet,basic}}$  such that

$$p_T^{\text{jet,basic}} = p_T^{\text{jet,raw}} - \rho \times A^{\text{jet}}. \quad (\text{V.23})$$

This procedure removes a signal of the scale of the  $p_T$  contribution from pile-up, together with the part of the UE which is represented by topo-clusters outside of the jet.<sup>10</sup> The measurement of  $\rho$  uses topo-clusters that are most likely outside of jets, due to its construction as a median transverse momentum density. Residual *local* fluctuations of the signal contribution from pile-up to a given reconstructed jet are not corrected. Reconstructed calorimeter jets with  $p_T^{\text{jet,basic}} < p_T^{\text{min}}$  are considered as originating from pile-up and thus are dropped.<sup>11</sup>

This UE contribution to  $\rho$  in terms of topo-clusters is hard to determine independently in the experiment, even in low pile-up scenarios.<sup>12</sup> In the presence of significant pile-up, with

<sup>10</sup>CMS uses the same technique but based on particle flow signals [972].

<sup>11</sup>The actual value of  $p_T^{\text{min}}$  is adjusted to the pile-up conditions. It is typically of  $\mathcal{O}(20 \text{ GeV})$ .

<sup>12</sup>The  $p_T$ -flow from charged particles associated with the UE can be directly measured using reconstructed tracks from the hard scatter vertex, at least within the tracking detector acceptance of  $|\eta| < 2.5$ , following the strategy lined out in Ref. [973].

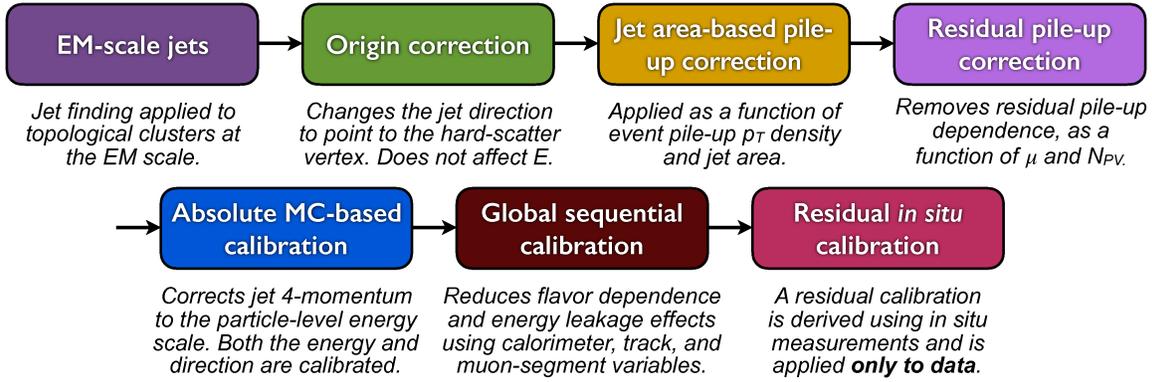


Fig. V.13: The ATLAS jet calibration scheme for early (2015) LHC Run 2 (taken from Ref. [713]).

the already discussed higher clustering thresholds, the  $\rho$  contribution from the UE is likely very small, in particular in terms of topo-clusters with a signal dominated by a particle from the UE. Here the indirect UE contribution merged into pile-up or hard scatter seeded clusters is more relevant, yet with a relatively small effect on the scale for  $\rho$ . It is also expected to be different inside and outside of jets, due to the clustering algorithm, which leads to a higher probability of small signal survival in the presence of the larger signals especially in the core of the jet. In this respect, UE and pile-up emissions close to the jet axis are likely to still contribute to the jet signal even after the jet-area-based correction, at a harder scale.

In addition to the jet-area-based correction given in Eq. (V.23), MC simulation based corrections similar to the ones discussed for low pile-up conditions are applied to  $p_T^{\text{jet,basic}}$  to mitigate residual dependencies on  $N_{PV}$  and  $\mu$ . These small corrections are again not expected to change the UE contribution.

### 4.3 Calibration of narrow jets

The full chain of jet calibration in ATLAS, as applied in 2015, is described in Ref. [713] and schematically shown in Fig. V.13. When jets are clustered from topo-clusters on the electromagnetic energy scale [970] (*EM-scale jets*), their direction is reconstructed with respect to the nominal detector centre. As a first step of the calibration chain, their direction is recalculated with respect to the reconstructed hard scatter vertex (*origin correction*). The *jet-area-based pile-up correction*, as discussed in the previous section, is applied next, followed by the previously mentioned *residual pile-up correction*. After this, the jets are subjected to the *absolute MC-based calibration*.

This absolute jet energy scale (JES) calibration is determined by matching calorimeter jets with truth particle jets in full MC simulations including pile-up, in QCD dijet calibration samples generated with a default MC generator configuration<sup>13</sup> and a detector simulation configuration. The match is based on the angular distance of the two jets and thus purely geometrical. The calorimeter jet energies  $E_{\text{calo}}^{\text{jet,raw}}$  include all contributions from in-time and out-of-time pile-up, and the UE. After all corrections up to the ones for pile-up are applied ( $E_{\text{calo}}^{\text{jet,raw}} \rightarrow E_{\text{calo}}^{\text{jet,basic}}$ ), the ratio of reconstructed calorimeter jet energy  $E_{\text{calo}}^{\text{jet,basic}}$  to the truth jet energy  $E_{\text{truth}}^{\text{jet}}$  is collected

<sup>13</sup>In 2015, the  $2 \rightarrow 3$  matrix element calculation in Powheg-Box 2.0 [350, 849] employing the CT10 PDF set [974] is used combined with parton showers and underlying event from Pythia8.1 [633], and the A14 tune of UE parameters [632].

from the calibration sample in bins of  $E_{\text{truth}}^{\text{jet}}$  and the truth jet pseudorapidity  $\eta_{\text{truth}}^{\text{jet}}$ ,

$$\mathcal{R}(E_{\text{truth}}^{\text{jet}}, \eta_{\text{truth}}^{\text{jet}}) = \frac{E_{\text{calo}}^{\text{jet,basic}}}{E_{\text{truth}}^{\text{jet}}}.$$

The average  $\langle \mathcal{R} \rangle(E_{\text{truth}}^{\text{jet}}, \eta_{\text{truth}}^{\text{jet}})$  is determined in each bin of  $E_{\text{truth}}^{\text{jet}}$  in a given slice of  $\eta_{\text{truth}}^{\text{jet}}$ . Using numerical inversion to project

$$\langle \mathcal{R} \rangle(E_{\text{truth}}^{\text{jet}}, \eta_{\text{truth}}^{\text{jet}}) \mapsto \langle \mathcal{R} \rangle(E_{\text{calo}}^{\text{jet,basic}}, \eta_{\text{calo}}^{\text{jet,basic}}),$$

and inverting the averages

$$\langle \mathcal{R} \rangle(E_{\text{calo}}^{\text{jet,basic}}, \eta_{\text{calo}}^{\text{jet,basic}}) \mapsto \langle \mathcal{R} \rangle^{-1}(E_{\text{calo}}^{\text{jet,basic}}, \eta_{\text{calo}}^{\text{jet,basic}}),$$

yields a calibration function when fitting a smooth functional form over all energies in an  $\eta_{\text{calo}}^{\text{jet,basic}}$  slice.

The truth jet includes all stable particles generated in MC simulations, except for neutrinos, muons and any particle<sup>14</sup> emerging from the overlaid in-time pile-up interactions. The remaining set comprises all particles with lifetimes  $\tau$  given by  $c\tau > 10$  mm in the laboratory frame. Among those are particles generated by the fragmentation of the parton showers associated with the hard scatter partons as well as the particles generated by radiation, and by multiple parton interactions, all characteristic for the UE. The total truth jet energy and, to a lesser extent, its direction are thus affected by the UE. With this respect, the UE energy contribution is reinstated in the calorimeter jet after this absolute calibration – the calibrated calorimeter jet energy corresponds on average to the truth jet energy in the same phase space, and thus contains an average UE at particle level.

This approach scales the calorimeter jet to a model dependent JES. Direct application of this calibration to data projects this model dependence into the jet measurement. To understand the bias introduced by this calibration method, alternative generators with different orders of, or approaches to, the hard process calculation and/or different choices of the parton distribution functions (PDFs), fragmentation models and sets of tuned UE parameters are used in the simulation of calorimeter jets, which are then subjected to the jet calibration derived with the default configuration. Recent (2015) generators considered are Sherpa 2.1 [415] employing multi-leg  $2 \rightarrow N$  matrix elements and their own default tuned set of soft event modelling parameter, and Herwig++ 2.7 [825,906] with a  $2 \rightarrow 2$  matrix element and their own default tuned parameter set. Any observed differences of the alternative calorimeter JES with respect to the default JES are absorbed into a systematic uncertainty. Due to the small contribution that can be expected from the UE in recent pile-up conditions, no particular uncertainty component is attributed to this modelling feature.

It is notable that the actual survival of the UE in the calorimeter signal space does not really enter into the jet calibration, as long as the jets in the calibration samples are subjected to the same reconstruction chain, up to and including the full pile-up correction, as any other jet in MC simulations and in experimental data. No matter how much UE signal might have been removed in this chain, the jets will consistently be calibrated including the UE. Significant changes in the default truth jet generator or in the detector simulation require a redetermination of the calibration functions.

In addition, significant changes to any of the steps leading up to calibration typically require a new round of determining the absolute MC calibration functions, such that calibrated calorimeter jets on average maintain the same truth JES. In cases where the changes in the

<sup>14</sup>Due to the large number of pile-up particles, those are not even stored in the calibration samples.

upstream corrections are insignificant, the calibration functions may not be changed and additional systematic uncertainty components may be injected. Changes in the pile-up activity e.g. due to increasing beam intensities at the LHC, may not require a full recalibration if the pile-up mitigation maintains its hard scatter jet finding efficiency. Smaller changes in the simulated detector response also add an uncertainty component. A rather complete evaluation is described in [712].

The next step after the absolute MC calibration is the *global sequential calibration*, which attempts to reduce fluctuations in the jet response by reducing its sensitivity to the jet flavour composition and possible longitudinal energy leakage in the calorimeters. This step is designed to not change the average response in any given phase space bin anymore, but to improve the resolution of the response function.

The origin correction is independently applied in data and MC simulations, as it only needs the reconstructed hard scatter vertex. The jet-area-based pile-up correction is also independently applied, as it only needs the reconstructed  $\rho$  in any given event. Assuming that this  $\rho$  reflects the pile-up (and possibly UE) contribution to jets in the same way in data and MC simulations, the remaining jet signal should have the same quality (e.g. with respect to the level of pile-up suppression) in both environments. In this context it is not required that  $\rho$  is actually the same in data and MC simulations (their distributions are typically found to be quite similar). This introduces a certain level of insensitivity to the actual pile-up simulation, which is typically far from perfect, see e.g. case studies in Ref. [970].

The residual pile-up correction, the absolute calibration and the global sequential calibration are all derived from MC simulations. A final set of in-situ measurements provide calibrations that are applied only to data and lead to the same average jet response in data and MC simulations.

#### 4.4 Large radius jets

At LHC, jets with large radius ( $R \geq 0.8$ ) are often considered in searches for new particles or boosted Standard Model particles. The large catchment area in  $(y, \phi)$ -space associated with these jets, especially their large extension in rapidity  $y$ , prohibits the straight forward application of a jet-area-based pile-up correction at the typically required levels of precision (e.g.,  $\mathcal{O}(1\%)$  for the  $p_T$  scale). This is largely due to the  $y$  range covered by these jets and the fact that both the number density of particles emerging from minimum bias (soft) proton–proton collisions and the average  $p_T$  are  $y$ -dependent. This suggests a significant variation of  $\rho$  inside the large radius jet. Additional complications may arise from a changing calorimeter readout granularity across such a large detector object.

Often large radius jets are groomed with e.g. trimming [] to extract hard radiation pattern inside the jets. Such techniques often remove pile-up and inadvertently UE contributions from the jet, with significant effects on the reconstructed jet kinematics, including the single jet mass. While the UE contribution can indeed affect the jet mass in ungroomed jets, in particular those without two- or three-prong decay substructure, the grooming reduces the impact of signals from any soft emission into the large radius jet.

The phase space of interest for large radius jet applications is usually determined by a search or decay tag goal. Assuming an attempt to tag hadronic  $W$  boson decays in a jet of  $R = 1.0$ , a typical (lower)  $p_T$  scale is given by  $p_T \approx 2 \times M_W$ , where  $M_W$  is the mass of the  $W$  boson. Higher masses introduce higher  $p_T$  scales accordingly. Even for the largest practical jet radii, the UE contribution to the overall  $p_T$  is insignificant, in particular in high pile-up conditions. Any loss of UE signal after grooming seems to be acceptable when jet-mass-based searches are conducted.

Any grooming attempt introduces a non-trivial modification of the jet composition, and as such needs careful reflections using MC simulations and careful validation with experimental data. In particular scenarios where UE may significantly affect a reconstructed jet feature, modelling dependencies have to be studied and incorporated into any relevant uncertainty or efficiency.

A cleaner inclusion of the UE into large radius jets by using jet re-clustering [975] is emerging for ATLAS (see first evaluations in Ref. [976]). The clustering of narrow jets into large radius jets provides fully corrected and calibrated inputs to any grooming technique and reconstruction of substructure-related observables for these large jets. Each narrow jet includes a contribution from the UE, as discussed in the previous section, and is corrected for pile-up as well. In analyses where softer radiation or energy flow between harder structural elements of the large radius jet are not explored, the re-summed large jet kinematics, including its mass, correctly include the UE within the physics model used for the narrow jet calibration.

## 4.5 Conclusions

Both ATLAS and CMS apply a MC simulation based absolute jet energy calibration that includes the particle-level underlying event contribution. Independent of how well this contribution is actually represented in the detector signals – clusters of topologically connected calorimeter cell signals in ATLAS and particle flow objects in CMS – the final jet energy scale reflects the underlying event as modelled by the physics models used in the determination of the absolute calibration. A limited but well motivated set of models and modelling parameters is used to evaluate systematic uncertainties introduced to the calibration of the jets in the experiment by the choice of a certain default MC simulation. The concern that any pile-up correction removes underlying event signals is valid but has no detectable influence on the jet energy scale, as long as all corrections and calibrations are derived and applied in a consistent manner. Final calibration functions derived from in-situ transverse momentum balance studies applied to data only shift the experimental jet energy scale to the one from the default MC simulation. In this the underlying event contribution is scaled in the same way as the overall jet energy scale.

The complexity of the jet calibration procedure in both CMS and ATLAS, and the misunderstandings arising from it, encouraged this pedagogical contribution to the proceedings.

## 5 Precision comparisons of predictions for Higgs boson + jet production at the LHC as a function of jet size <sup>15</sup>

We perform a phenomenological study of Higgs-boson production in association with a jet at the Large Hadron Collider. We investigate in particular the dependence of the leading jet cross section on the jet radius as a function of the jet transverse momentum. Theoretical predictions are obtained using perturbative QCD calculations up to next-to-next-to-leading order. They are compared to results obtained from matching next-to-leading order calculations to parton showers and possibly including higher-order real radiative effects through multi-jet merging.

### 5.1 Introduction

During Les Houches 2015 [1], a detailed comparison of fixed order (at NLO and NNLO) and matrix element plus parton shower (MEPS) predictions for differential Higgs boson (+jet) production was carried out. The goal was multi-fold: using identical starting configurations, to check the consistency of the MEPS predictions among themselves, and to demonstrate that the MEPS predictions revert to their underlying fixed order predictions in non-Sudakov kinematic

---

<sup>15</sup> J. Bellm, A. G. Buckley, X. Chen, J. R. Currie, A. Gehrmann-De Ridder, T. Gehrmann, E. W. N. Glover, S. Höche, A. Huss, J. Huston, S. Kuttimalai, J. Pires, S. Plätzer, M. Schönherr

regions. These comparisons largely showed good agreement among the ME+PS predictions and that the ME+PS predictions agreed reasonably well with their fixed-order counterparts.

We have continued these studies in Les Houches 2017, but now looking in finer detail at the level of agreement. In particular, the perturbative jet shape is not as well modeled at fixed order as with the NLO matrix element plus parton shower predictions that are available. This may have a quantitative impact on the cross section predictions, especially for small jet sizes.

The study is designed to use Higgs+jet, Z+jet and inclusive jet production, taking advantage of the NNLO calculations available for all three processes. The latter two processes are important for global PDF fits, where only fixed order predictions (along with the relevant non-perturbative corrections) have been used. Due to time constraints, only the Higgs+jet process will be considered in detail in these proceedings. The other two processes, along with Higgs+jet, will be included in a more comprehensive study in a future publication.

As a further test of the impact of parton showers versus fixed order, the jet size was varied over the values  $R \in [0.3, 0.4, 0.5, 0.6, 0.7, 1.0]$ , using the anti- $k_T$  jet algorithm [661]. Indirectly, this tests how well the one (two) extra gluon(s) at NLO (NNLO) reproduce the perturbative aspects of the jet shapes, as embodied by the parton shower. This is of particular interest as the Higgs boson + jets measurements that have been performed at the LHC in Run 2 have used a jet size of 0.4, which is slightly above the jet size region where small  $R$  effects may be important. Taking them into proper account would require resummation, as discussed in [212]. The MEPS predictions also basically provide this resummation, through the parton showers. The MEPS predictions at NLO can properly take these effects into account. However, the highest precision for  $H + \geq 1$  jet is from the NNLO predictions, and there is no MEPS formalism that works at this order. As an additional motivation, there has been recent speculation that the quark jet shape is not well-described at NLO. This study provides samples of both quark and gluon jets to test this hypothesis.

Predictions from MEPS programs were carried out at the parton shower level (for easiest comparison to fixed order).<sup>16</sup>

To the degree to which it was possible, the initial conditions have been constrained to be the same for all predictions. Each calculation used the PDF4LHCNNLO\_30 PDFs [597], with its central value of  $\alpha_s(m_Z)$  of 0.118. As its name implies, this PDF has 30 error PDF sets that completely determine the PDF uncertainty. The scale choices for all predictions have been designed to be as similar as possible, providing a greater level of control than was available in the 2015 Les Houches study. More detail will be provided in the sub-sections dealing with each prediction. A CMS Rivet [500] routine from the 13 TeV inclusive jet analysis [977], was modified to add the different  $R$  values, as well as additional variables. This Rivet routine was further modified for the Higgs boson + jet (and Z boson + jet) studies.

### 5.1.1 NNLOJET - NNLO Calculation of $pp \rightarrow H + j$

The NNLO corrections to  $pp \rightarrow H + j$  receive contributions from three types of parton-level sub-processes: the  $H+5$  parton tree-level process [978–980] (double-real correction), the  $H+4$  parton process at one-loop [981–983] (real-virtual correction), and the two-loop  $H+3$  parton process [984] (double-virtual correction). Each of these contributions are separately infrared (IR) divergent, while the divergences cancel upon integration over the unresolved phase space by virtue of the Kinoshita–Lee–Nauenberg theorem. In order to arrive at a fully differential pre-

<sup>16</sup>In the future publication, comparisons will also be made at hadron level, as a way of comparing the non-perturbative corrections as a function of jet radius. As a reminder, the non-perturbative corrections used for fixed order predictions are determined from a comparison of the parton shower predictions with and without the non-perturbative effects, as a function of jet radius. This implicitly requires the integrated jet shape determined by fixed order predictions to agree with those determined by parton showers.

diction, a procedure for the subtraction of IR singularities is required to make this cancellation manifest. To this end, we employ the antenna subtraction formalism [125, 126, 985–991] to extract the IR singularities from the various contributions and to achieve their cancellation prior to the phase-space integration. As a result, the phase-space integration over the contributions with different parton multiplicities are individually rendered finite and can therefore be performed numerically using Monte Carlo techniques. The antenna subtraction formalism is implemented in the NNLOJET framework. The application to NNLO corrections for  $H + j$  production and its validation against an independent calculation using the sector-improved residue subtraction method [150] was described in Refs. [236, 238, 992].

For the current study, we employ the dynamical scale

$$\mu_0 = \frac{1}{2}H_T = \frac{1}{2} \left( \sqrt{m_H^2 + p_{T,H}^2} + \sum_{jet} p_{T,jet} \right), \quad (\text{V.24})$$

as our central scale choice. The renormalisation ( $\mu_R$ ) and factorisation ( $\mu_F$ ) scales are varied independently around  $\mu_0$  by factors of  $\frac{1}{2}$  and 2 to estimate the size of missing higher-order contributions. Here, the two extreme variations are excluded such that we arrive at the custom 7-scale variation:

$$(\mu_R, \mu_F) = \left\{ (1, 1), (2, 2), \left(\frac{1}{2}, \frac{1}{2}\right), \left(\frac{1}{2}, 1\right), \left(1, \frac{1}{2}\right), (2, 1), (1, 2) \right\} \times \mu_0. \quad (\text{V.25})$$

The LO and NLO contributions using this dynamical scale choice were validated against Sherpa and Herwig7 as described in other sections in this report.

In order to obtain the results for the various cone sizes ( $R = 0.3, 0.4, 0.5, 0.6, 0.7$ , and 1.0) required in this study, we have exploited the fact that the Born-level kinematics is insensitive to  $R$ . As a result, the difference between two cone sizes can be obtained from a  $H + 2j$  calculation at one order lower:

$$\sigma_{H+j}^{NNLO}(R) - \sigma_{H+j}^{NNLO}(R') = \int \left[ d\sigma_{H+2j}^{NLO}(R) - d\sigma_{H+2j}^{NLO}(R') \right]_{njets \geq 1}. \quad (\text{V.26})$$

Note that the difference has to be taken at the level of the integrand, since one term acts as a local counter-term of the other in all IR-divergent limits where a jet becomes unresolved and  $H + 2j \rightarrow H + j$ . Using Eq. (V.26), all predictions for different  $R$  values can be obtained from a single NNLO computation by adding differences of  $H + 2j$  computations that are substantially cheaper to compute.

### 5.1.2 Herwig

We used HERWIG 7 [584, 825, 959] based on version 7.1.2 and ThePEG version 2.1.2 with minor changes to standard HERWIG 7 scale settings to the agreed ones. The NLO matching was performed with matrix elements from OpenLoops [85] and MadGraph [877] interfaced with the BLHA2 standard [506]. For parton distributions the PDF interface from LHAPDF6 [964] was used. In the results we show matched  $NLO \oplus PS$  predictions – to emphasise the similarities of matching and merging at high transverse momenta – with the  $\tilde{Q}$ -shower but tested with lower statistics that merging according to [993] and matching to the HERWIG 7 dipole shower [994] show similar behaviour. Hadronisation and MPI models are switched off and  $\alpha_S$  of the hard process is synchronized with the PDF set. We include effects of CMW scheme [995] by an enhanced shower  $\alpha_S = 0.124$ . The scale used for the core process in the Matching is defined as in Eq. (V.24). As shower starting scale we use the transverse momentum of the hardest jet.<sup>17</sup>

<sup>17</sup>The effect of changing this scale to the one in Eq. (V.24) is moderate and only noticeable for transverse momenta smaller or in the range of the Higgs mass.

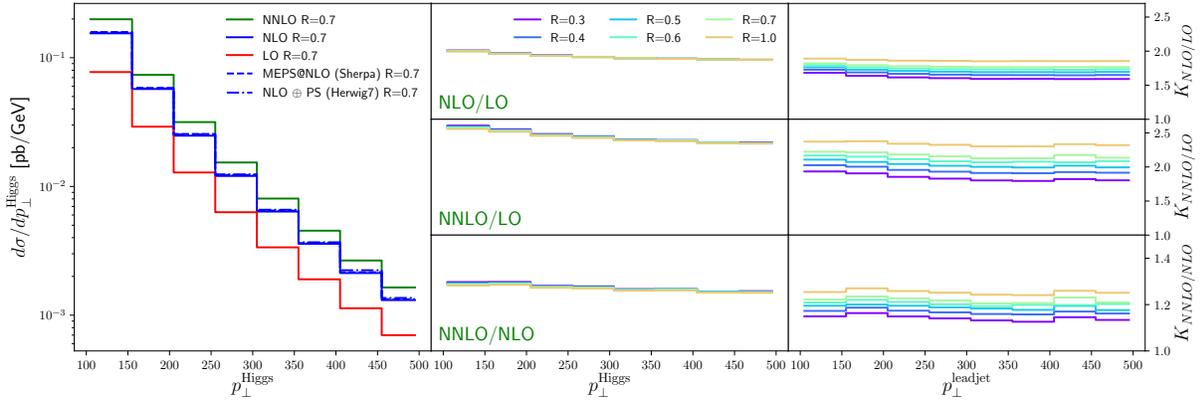


Fig. V.14: Shown on the left panel is the transverse momentum spectra of the Higgs boson ( $p_{\perp}^{\text{Higgs}}$ ) and the leading jet ( $p_{\perp}^{\text{leadjet}}$ ) in Higgs plus jet production at  $\sqrt{s} = 13$  TeV. We compare LO, NLO, NNLO and multi-jet merged predictions for a jet radius of 0.7. The  $K$ -factors (NLO/LO, NNLO/LO and NNLO/NLO) determined from the NNLOJET predictions in this study, as a function of  $p_{\perp}^{\text{Higgs}}$  ( $p_{\perp}^{\text{leadjet}}$ ) for various jet sizes is shown on the middle (right) panel.

### 5.1.3 Sherpa

We use a pre-release version of the SHERPA Monte Carlo event generator [415, 490], based on version Sherpa-2.2.4. The NLO matching is performed in the S-MC@NLO approach [929, 930], while the multi-jet merging employed is based on the MEPS@NLO formalism [925, 926]. We use a modified version of a parton shower algorithm [491], which is based on Catani-Seymour dipole subtraction [187, 489]. We use a running coupling consistent with the PDF, and we employ the CMW scheme to include the two-loop cusp anomalous dimension in the simulation [995]. In the merging procedure a core scale needs to be defined. To make the result comparable to the FO result we use the same definition as in Eq. (V.24). If no ordered clustering history can be defined, the couplings in the effective gluon-gluon-Higgs-vertex are evaluated at this core scale.

## 5.2 Results

The analyses use the anti- $k_T$  jet algorithm, with varying jet size as described above, with a jet transverse momentum threshold of 30 GeV, along with an (absolute) rapidity cut on the jets of 4.5. To avoid generation cut effects, the comparisons are performed above a jet transverse momentum value of 50 GeV.

Figure V.14 (left) shows the transverse momentum spectrum of the Higgs boson as predicted by the fixed-order LO, NLO and NNLO calculation, as well as the result from a multi-jet merged computation using the SHERPA event generator and the NLO-matched HERWIG result. We observe that above a  $p_T$  value of the order of the Higgs mass, the distributions agree in their shape, while the normalization differs slightly due to the higher-order effects included in the NNLO calculation and the differences between fixed-order and the MEPS results in their choice of renormalization and factorization scales. The Higgs boson  $p_T$  spectrum is the most inclusive observable in this study, and the prediction above a certain  $p_T$  may therefore serve as a useful normalization when comparing shapes of other distributions between fixed order and MEPS results.

Figure V.14 (middle) shows the  $K$ -factors as a function of the Higgs boson  $p_T$ ; as expected there is no jet size dependence for this variable. Figure V.14 (right) shows the local  $K$ -factors (NLO/LO, NNLO/LO, NNLO/NLO) for  $H + \geq 1$  jet production, as a function of the leading jet transverse momentum, for various jet sizes, obtained from NNLOJET. The  $K$ -factors are

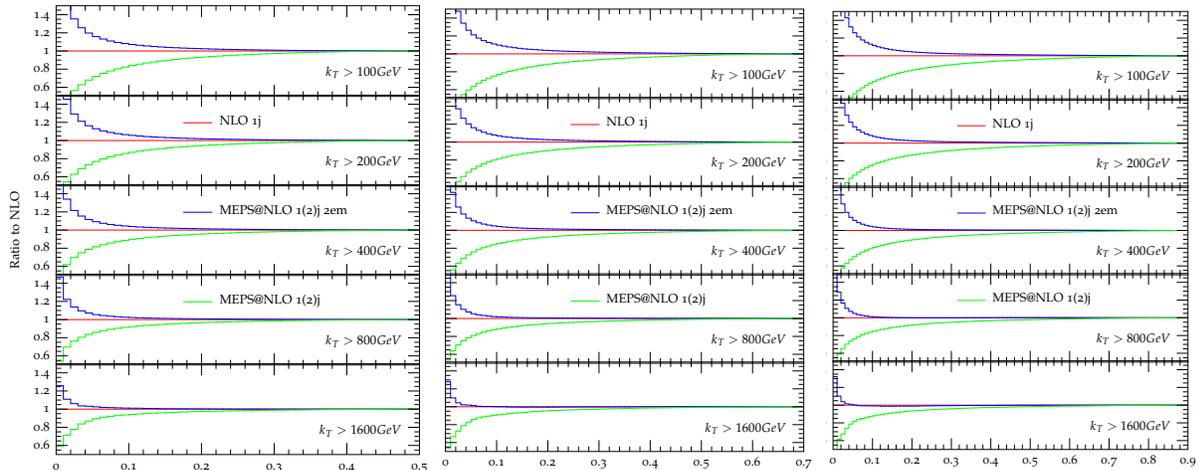


Fig. V.15: Ratios of integrated jet shapes (see Eq. V.27) as function of  $r$ , for  $R = 0.5$ ,  $R = 0.7$  and  $R = 0.9$  (left to right).

relatively flat as a function of jet  $p_T$ . They grow with increasing jet size, due to inclusion of additional real radiation.

In Fig. V.15 we investigate the difference between the fixed-order NLO and multi-jet merged predictions for integrated jet shapes (Eq. V.27). The integrated jet shape is defined as

$$\Psi(r) = \frac{1}{N_{\text{jet}}} \sum_{\text{jets}} \frac{p_T(0, r)}{p_T(0, R)}, \quad (\text{V.27})$$

with  $r$  being the radius of a cone which is concentric to the jet axis and  $p_T(r_1, r_2)$  being the magnitude of the vectorial momentum sum of all particles in the annulus between radius  $r_1$  and  $r_2$ . We also compare to a truncated merged prediction, where the number of final-state partons generated in the simulation is limited to at most two. This simulation presents the closest possible approximation to the fixed-order result that we are able to generate using the merging algorithms. It reflects the kinematical restrictions of the NLO calculation (i.e. that only up to one additional final-state parton can be present), but it also includes additional approximate higher-order virtual corrections by means of Sudakov factors. Nevertheless we observe that the full NLO result and the truncated merged result approach each other well within the jet cone, and the convergence is naturally faster for larger jet transverse momenta. This strongly suggests that the differences between the fixed-order predictions and merged results in Fig. V.19 below are due to higher-multiplicity final states. The discrepancies between fixed-order results and merged predictions for small and large  $R$  should therefore be reduced for higher-order perturbative calculations.

Figure V.16 shows the cross section scale variations at LO, NLO and NNLO for  $H + \geq 1$  jet production, as a function of the leading jet transverse momentum, for various jet sizes. The uncertainty band is given by the highest and lowest cross section predictions at each order. As expected, the uncertainty on the cross section decreases from LO to NLO to NNLO. It also decreases as the jet size decreases, perhaps not unsurprising given that larger jet radii lead to inclusion of more real radiative corrections. Also shown for comparison are the predictions from the two MEPS calculations (nominally of NLO accuracy). We scale the MEPS predictions with the  $K$ -factors derived from the Higgs  $p_T$  distribution above 150 GeV, see discussion of Fig. V.14.

Figure V.17 shows the leading jet  $p_T$  cross sections for the different scale choices, at NLO and NNLO, as a function of the jet size  $R$ . In this case, a minimum transverse momentum

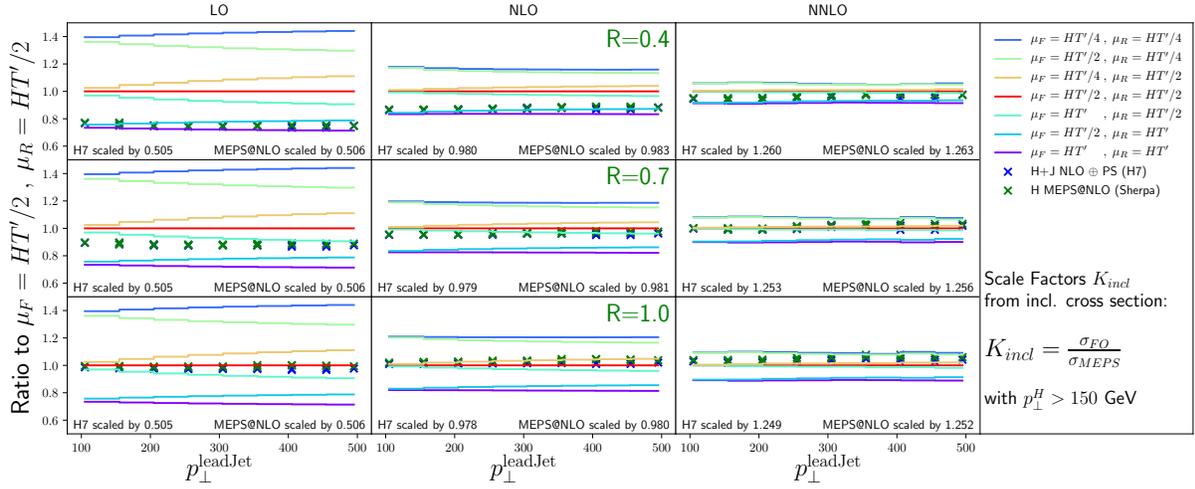


Fig. V.16: The scale variations at LO, NLO, and NNLO from NNLOJET for 3 jets size, as a function of leading jet transverse momentum, are shown. For comparison, the nominal NLO MEPS predictions are also shown. The generator predictions are scaled with the inclusive  $K_{incl}$  factor with Higgs  $p_{\perp}^H > 150$  GeV, see Fig. V.14.

requirement of 150 GeV has been placed on the leading jet. We assume this scale to be large enough such that  $M_H$  is not the large scale in the process. The dots for each scale choice have been fit to a functional form motivated by the expected behavior for jet-vetoed cross sections. We assume the leading functional form:

$$f(R) = a + b \log(R) + c \log^2(R) \quad (\text{V.28})$$

as we expect a logarithmic behaviour for the scales induced by the effective veto on the cross section by cutting with the jet cone  $R$ . The lines in Fig. V.17 are then interpolations with Eq. (V.28) and the fitted values.

Again, the scale variation band is given by the upper and lower curves at each order. It is notable that the scale uncertainty bands shrink as the jet size decreases, as mentioned earlier. For very low values of  $R$ , this improvement in the uncertainty can be regarded at least partially due to accidental cancellations that stem from the restrictions in phase space. Similar effects were pointed out in the context of exclusive jet rate measurements [886]. It can also be observed that for each particular scale, the slope is greater at NNLO than at NLO. The MEPS predictions are also plotted in the figure, and can be observed to have a greater slope than even the NNLO predictions. This can be seen as an effect of either including (at large  $R$ ) or not excluding (at small  $R$ ) additional semi-hard real emissions, which have a leading-order scale dependence and therefore induce a large change in the cross section.

Figure V.18 shows the cross sections for the Higgs transverse momentum (top) and leading jet transverse momentum (bottom) for several different jet sizes, at LO, NLO and NNLO (from NNLOJET) and from the two MEPS predictions. All cross sections have been scaled to their respective value for the reference jet size of  $R = 0.7$ . At this value we observe the best agreement between fixed-order and multi-jet merged results, save for an overall normalization which can be extracted from the Higgs transverse momentum spectrum, cf. Fig. V.14. The absolute value of the difference between the fixed-order and the multi-jet merged predictions away from  $R = 0.7$  increases roughly proportional to  $\log(R/0.7)$  (cf. Fig. V.19), which is expected due to the higher-multiplicity real-emission corrections included in the multi-jet merged calculation. Depending on kinematics they either enhance (at  $R > 0.7$ ) or reduce (at  $R < 0.7$ ) the cross section. The

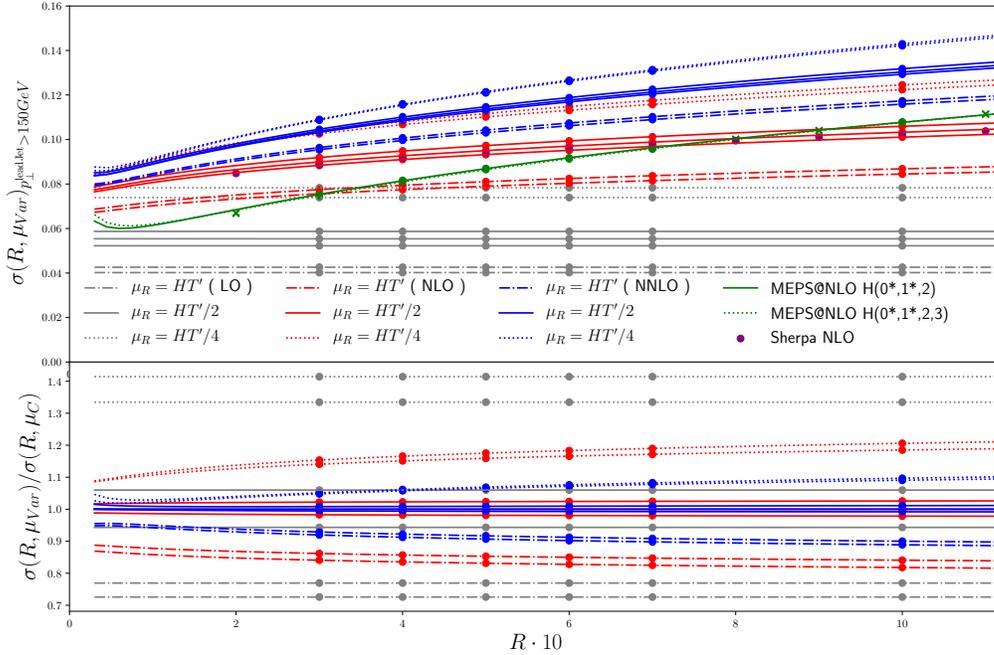


Fig. V.17: The  $R$ -dependence of the cross sections at NLO, NNLO and MEPS are shown, for particular scale values, as a function of the jet radius, for leading jet transverse momenta above 150 GeV.

differences between the MEPS predictions and those from NNLOJET decrease as the order is raised from NLO to NNLO. The difference is on the order of 5% for  $R = 0.5$  at NLO and of the order of a few percent at NNLO.

Given the better description of the jet shape provided by the MEPS predictions, this is an indication of the theoretical uncertainty associated with the truncation of the perturbative series. The uncertainty is reduced at NNLO as expected. It is noteworthy that the ratios in Fig. V.18 are relatively flat as a function of the transverse momenta.<sup>18</sup>

Figure V.19 shows the dependence of the relative difference between a NLO-matched, multi-jet merged prediction from Sherpa and the NLO fixed-order result as a function of the leading jet transverse momentum for varying jet radii. The ratio is flat as a function of the leading jet  $p_T$ . In Fig. V.17 we compared integrated cross sections, while here we observe interestingly a similar behaviour for the differential cross sections. In the right plot, the projection is with respect to the radius, and displays, in grey, the various transverse momentum intervals and, in coloured, the lowest and highest energies. Assuming the leading behaviour is given by Eq. (V.28), and with the flatness in the leading jet transverse momentum, the linear, (but slightly quadratic) behaviour in the logarithmic plot is expected. We note the zero crossing of the curve on the right-hand side, which corresponds to the best agreement between fixed-order and matched/merged result, is located at  $R \approx 0.7$  (see the discussion of Fig. V.18). In configurations where the jet rapidity is zero, this corresponds to a roughly equal partitioning of the rapidity phase-space into collinear sectors for color dipoles spanned between the initial-state partons and the final-state

<sup>18</sup> The flatness may in fact be somewhat accidental as we are using the EFT and needs to be confirmed upon including finite top mass effects.

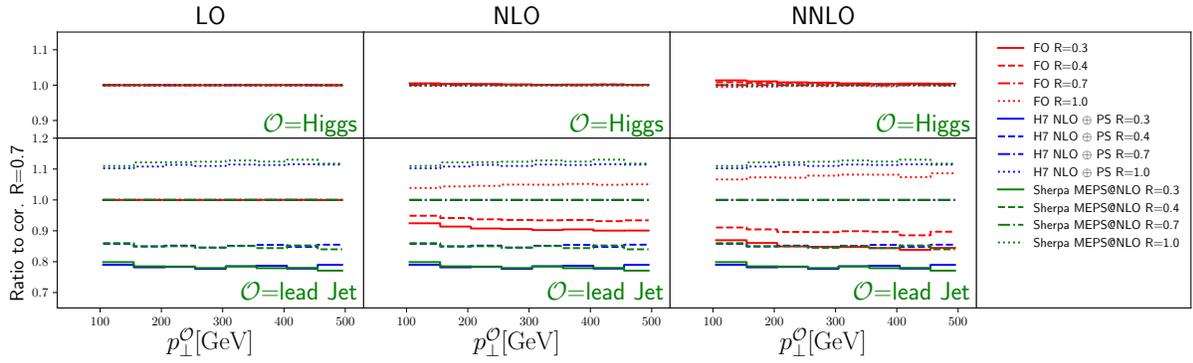


Fig. V.18: The ratio of each cross section (either Higgs  $p_T$  or lead jet  $p_T$ ) for specific jet sizes, scaled to the cross section for each prediction for a jet size of  $R = 0.7$ .

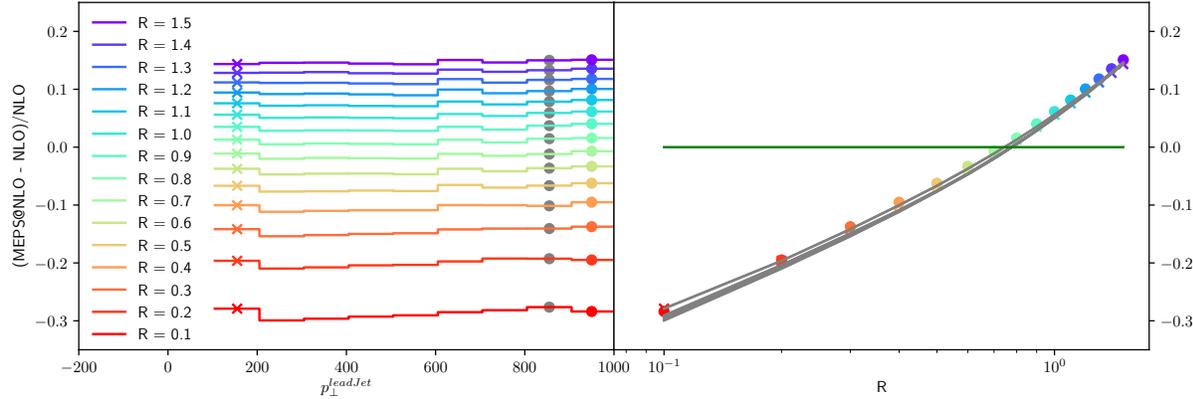


Fig. V.19: Relative difference between the NLO-matched and multi-jet merged prediction to the fixed-order result as a function of the leading jet transverse momentum and the jet radius.

jet.

Finally, for comparison, in Fig. V.20, we show a (simplified) plot of the comparison of predictions carried out in Les Houches 2015 [1] for the leading jet transverse momentum. For these particular scale choices, the NNLO [176] and NLO [249, 996, 997] give essentially the same result. The predictions from Sherpa and Herwig agree with NNLO at low  $p_T$ , but fall 10-15% below at higher  $p_T$ . This can be due partially to differences in the scale choices between each prediction in the figure, differences between the scale choices used in [1] compared to ours, and to the jet shape differences explored in detail in this note.

### 5.3 Outlook

Searches for new physics, as well as a better understanding of standard model physics, require an increasing level of precision, both for measurement and for theory. For differential distributions for  $H + \geq 1$  jet, the highest level of precision is obtained with NNLO predictions. Matched NLO plus parton shower predictions (MEPS) provide a more complete event description, but at a level of precision one order of  $\alpha_s$  lower (although with a more complete logarithmic treatment). Most physics measurements at the LHC make use of relatively small jet sizes (anti- $k_T$  with  $R = 0.4$ ), and  $H + \geq 1$  jet production is no exception. There can be differences between fixed order and MEPS predictions for the same observable just due to the different estimates of the amount of jet energy contained in a jet of radius  $R$ . These differences can be comparable to the size of the

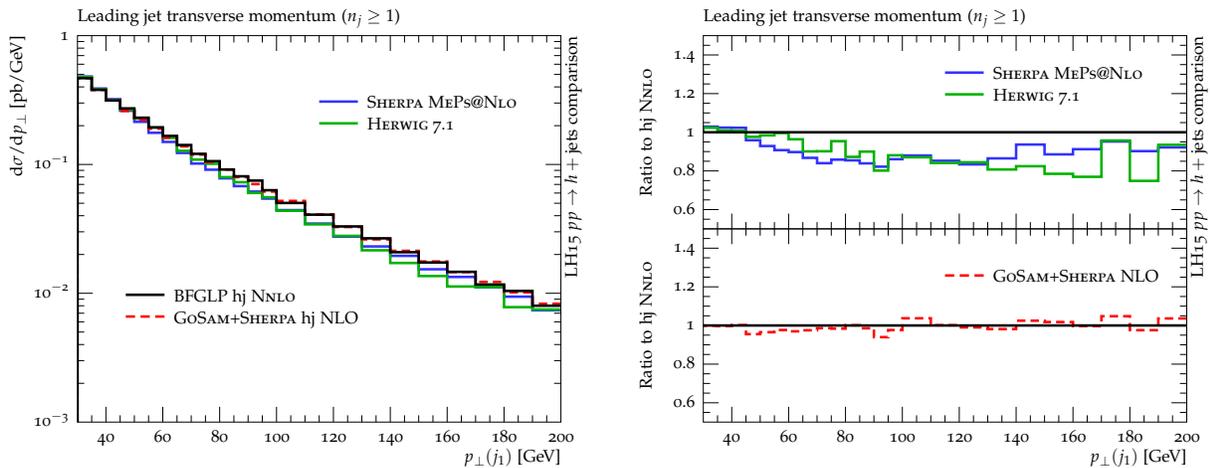


Fig. V.20: Predictions for the leading jet transverse momentum cross sections, for NLO, NNLO and MEPS calculations, from the 2015 Les Houches study [1].

scale uncertainty for the cross section at that order.

In this contribution, we have reported on an investigation of the impact of different jet sizes on Higgs boson plus jet physics at the LHC, paying close attention to the impact of the jet size on  $K$ -factors, on scale uncertainties, and on differences between fixed order and MEPS predictions. A more detailed study, intended for a future publication, will proceed from this contribution, expanded to also include inclusive jet and Z boson + jet final states. Better understanding of the issues described in this contribution may allow an improvement in the accuracy, and precision, of such predictions at the LHC.

## Acknowledgements

We thank the organizers for an inspiring workshop. This work was supported by the US Department of Energy under the contract DE-AC02-76SF00515. This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme, grant agreement No 668679. We thank the University of Zurich S3IT (<http://www.s3it.uzh.ch>) for providing support and computational resources. This research was supported in part by the UK Science and Technology Facilities Council, by the Swiss National Science Foundation (SNF) under contracts 200020-175595, 200021-172478, and CRSII2-160814, by a grant from the Swiss National Supercomputing Centre (CSCS) under project ID UZH10, by the Research Executive Agency (REA) of the European Union under the Grant Agreement PITN-GA-2012-316704 (‘‘IJHiggsTools’’) and the ERC Advanced Grant MC@NNLO (340983), and by the Fundacao para a Ciencia e a Tecnologia (FCT-Portugal), project UID/FIS/00777/2013. AB acknowledges support from Royal Society grants UF110191 and UF160548.

# Chapter VI

## Monte Carlo studies

### 1 On renormalization scale variations in parton showers <sup>1</sup>

We perform a phenomenological study of different schemes for varying the renormalization scale in parton shower calculations. To this end, we compare the conventional CMW scheme against a method which explicitly includes higher-order corrections proportional to the QCD beta function and the two-loop cusp anomalous dimension. The techniques are implemented in both SHERPA and PYTHIA, and compared to the default methods employed by these generators. In contrast to earlier studies of the same topic, we find that the uncertainty estimates agree well between the different parton-shower algorithms.

#### 1.1 Introduction

General-Purpose Event Generators are tools that combine perturbative calculations and phenomenological models to predict the final states of particle-particle collisions. The perturbative calculations include a fixed-order, matrix-element based component and an all-order (resummed) parton shower component. Precise, fixed-order predictions have become readily available due to improvements in calculation methods and numerical techniques. The limiting factor in testing predictions on data, particularly for observables that are not fully inclusive, are uncertainties in the parton shower evolution. Thus, realistic prescriptions to assess parton shower uncertainties are needed. They should ensure that only the effect of *uncalculated or ambiguous* higher-order terms are estimated, while preserving the leading higher-order corrections that have been included in the simulation due to their importance for practical applications. The latter include the use of transverse momenta as a scale in the evaluation of the strong coupling [998], and the use of the Catani-Marchesini-Webber (CMW) scheme [995] to recover known  $\mathcal{O}(\alpha_s^2)$  corrections to soft gluon emission.

Changes in the renormalization scale are diagnostic of parts of the parton shower uncertainty. Weighting techniques [961, 999, 1000] have made it very convenient to perform the related variations [501, 1001, 1002]. A previous analysis has thus focused on the comparison of uncertainties estimated by the various public parton-shower generators [1], with partially surprising results. It is timely to perform such a comparison again, and to question how parton shower variations should be implemented such as to obtain realistic, but not overly aggressive or overly conservative uncertainty estimates. The technique investigated here is motivated by comparison to analytic resummation, see for example [652]. We perform a dedicated comparison of different implementations of the method in the two event generators PYTHIA and SHERPA, using three different parton shower algorithms.

This contribution is organized as follows. First, we will motivate and define our renormalization scale variation scheme in Sec. 1.2. Next, we specify the implementation details for PYTHIA and SHERPA in Sec. 1.3. We present first results and compare the predicted uncertainties in jet production at LEP Run I and gluon-induced Higgs-boson production at LHC Run II in Sec. 1.4. The study concludes with a short summary in Sec. 1.5. Note that our aim is to better understand the impact of higher-order corrections related to soft-gluon emission in a phenomenological study. While the underlying computational technique is based on soft-gluon resummation at NLL, no claim of formal theoretical improvements of the parton-shower method

---

<sup>1</sup> S. Höche, S. Mrenna, S. Prestel, M. Schönherr, P. Z. Skands

is implied.

## 1.2 Motivation

The leading-order evolution kernels for a parton shower with evolution variable  $t$  and splitting variable  $z$  can be written in the universal form [831]

$$\mathcal{K}_{\tilde{ij} \rightarrow i,j}(t, z) = \frac{1}{t} \frac{\alpha_s(\mu)}{2\pi} P_{\tilde{ij} \rightarrow i,j}(t, z) + \mathcal{O}(\alpha_s^2) \quad (\text{VI.1})$$

In this context,  $P_{\tilde{ij} \rightarrow i,j}^{(i)}(t, z)$  are the model specific splitting probabilities for parton  $\tilde{ij}$  branching into partons  $i$  and  $j$ , in the presence of a parton (or set of partons)  $k$  that accounts for momentum conservation. This parton is conventionally called the recoil partner. Only the first-order pieces are usually considered in Eq. (VI.1), and a judicious scale setting is applied, based on the analysis in [998]. For transverse-momentum ordered showers, for which the evolution variable reduces to the relative transverse momentum  $p_{\perp}^2$  between the partons  $i$  and  $j$  in the soft and collinear limits, this leads to

$$\mathcal{K}_{\tilde{ij} \rightarrow i,j}(t, z) = \frac{1}{t} \frac{\alpha_s(bt)}{2\pi} P_{\tilde{ij} \rightarrow i,j}^{(0)}(t, z)$$

The prefactor  $b$  can be chosen according to the so-called CMW scheme [995] in order to reproduce the higher-logarithmic corrections arising from the two-loop cusp anomalous dimension [1003–1006]. This implies  $b(t) = \exp\{K(t)/\beta_0(t)\}$ , where

$$K(t) = \left( \frac{67}{18} - \frac{\pi^2}{6} \right) C_A - \frac{10}{9} T_R n_f(t), \quad \beta_0(t) = \frac{11}{3} C_A - \frac{4}{3} T_R n_f(t).$$

Note that  $n_f(t)$  is the number of active flavours at scale  $t$ . For five active flavors,  $b = 0.45$ , thus roughly halving the scale. For evolution kernels with soft gluon enhancement, this scale choice leads to sub-leading logarithmic corrections. At NLO, they take the expected form

$$\mathcal{K}_{\tilde{ij} \rightarrow i,j}(t, z) \Big|_{\mathcal{O}(\alpha_s^2)} = \frac{1}{t} \left( \frac{\alpha_s(t)}{2\pi} \right)^2 K(t) \left\{ \frac{2}{1-z} + \dots \right\}, \quad (\text{VI.2})$$

where the dots stand for terms that are non-singular in the soft-gluon limit. Alternatively, using  $b(t)$  can be interpreted as using a larger value of the strong coupling at the  $Z$  pole, i.e.  $\alpha_s(M_Z) \approx 0.130$  instead of the world average  $\alpha_s(M_Z) \approx 0.118$ . In the resummation of the transverse momentum of Drell-Yan lepton pairs and similar observables, Eq. (VI.2) would lead to the familiar resummation coefficient  $A_2$ . In addition, explicit logarithms of the renormalization scale and the transverse momentum in a dipole picture,  $k_T^2$ , appear in these calculations (see for example [652]), such that the double-logarithmic piece of the  $\mathcal{O}(\alpha_s^2)$  correction to the parton-shower splitting kernels should read

$$\mathcal{K}_{\tilde{ij} \rightarrow i,j}(t, z) \Big|_{\mathcal{O}(\alpha_s^2)} = \frac{1}{t} \left( \frac{\alpha_s(t)}{2\pi} \right)^2 \left[ \beta_0(t) \log \frac{k_T^2}{t} + K(t) \right] \frac{2}{1-z}. \quad (\text{VI.3})$$

The existing techniques for renormalization scale variations in parton showers implement this equation in different ways that are discussed in Sec. 1.3. Depending on the precise treatment of higher-logarithmic contributions, the related uncertainty estimates can exhibit a sizable variation. The phenomenological consequences of this effect will be discussed in Sec. 1.4. Our main demand is that variations around a central scale should not ruin the NLO accuracy in the soft

limit. It will be necessary to introduce a compensation factor to undo part of the naive scale variation.

In the current contribution, we will maintain a close correspondence with analytical results, but without any claim about the parton shower accuracy<sup>2</sup>. Thus, when varying the renormalization scale of parton shower splittings, we will implement variants of Eq. (VI.3), which differ only in how  $K(t)$  is included in the calculation. Note, however, that the definition of the soft enhanced part of the leading-order splitting function  $P^{(0)}(z, t)$  differs between parton showers with different definitions of  $t$  and  $z$ , and that the correspondingly different single-pole terms can also lead to visible differences. Nevertheless, we employ this minimalistic approach to define the least common denominator for three rather different parton showers, assuming our method will still provide insight that can be helpful for future studies.

### 1.3 Monte-Carlo Event Generators for this study

We have implemented the scheme outlined at the end of Sec. 1.2 in PYTHIA and SHERPA. Before presenting our results, we will describe the current default variation schemes and the precise technical realization of our new technique in the various generators.

#### 1.3.1 PYTHIA

PYTHIA 8 [630] is the most recent General-Purpose Event Generator in the PYTHIA family [629, 633, 1008]. It includes a native  $p_{\perp}$ -ordered parton shower supplemented with matrix-element corrections and multiple choices of recoil scheme, and further supports the VINCIA [1009, 1010] and DIRE [492] parton shower plugins. For this study, we will use the native shower in PYTHIA version 8.230 and the proper settings to enable the CMW scheme to set the argument of  $\alpha_s$ , by scaling  $\Lambda_{\text{QCD}}(n_f)$  appropriately.

The PYTHIA implementation of scale variations [1002] reflects the ambiguity in how compensation terms should be applied [1]. Similar to what is now assumed, it was decided that compensation terms should only apply to gluon emission, and not to  $g \rightarrow q\bar{q}$  splittings. Beyond that, the PYTHIA implementation of compensation terms was conservative, i.e. the effect of compensation was minimized when confronted with choices beyond  $\mathcal{O}(\alpha_s^2)$ . This led to modifying the parton-shower branching probability to

$$\mathcal{K}_{i\tilde{j} \rightarrow i,j}(t, z) = \frac{\alpha_s(kbt)}{2\pi} \left( 1 + (1 - \zeta) \frac{\alpha_s(b\mu_{\text{max}})}{2\pi} \beta_0 \ln k \right) \frac{P(z)}{t}, \quad (\text{VI.4})$$

for gluon emission, where  $P(z)$  is the full DGLAP splitting kernel, possibly including matrix-element corrections,  $\mu_{\text{max}} = \max(m_{\text{dip}}^2, kt)$  and

$$\zeta = \begin{cases} 1 - z & \text{for splittings with a } 1/(1 - z) \text{ singularity} \\ \min(z, 1 - z) & \text{for splittings with a } 1/(z(1 - z)) \text{ singularity} \end{cases}. \quad (\text{VI.5})$$

The inclusion of the factor  $(1 - \zeta)$  was motivated by the fact that the compensation is only theoretically motivated in the soft limit (i.e. for  $\zeta \rightarrow 0$ ). Thus, to remain conservative, the compensation was explicitly linearly damped outside of the soft limit.

To implement the new scheme proposed in Sec. 1.2, we change the branching probabilities to

$$\mathcal{K}_{i\tilde{j} \rightarrow i,j}(t, z) = \frac{\alpha_s(kbt)}{2\pi} \left( 1 + \frac{\alpha_s(bt)}{2\pi} \beta_0 \ln k \right) \frac{1}{t} \frac{2C_F}{1 - z} + \frac{\alpha_s(kbt)}{2\pi} \frac{1}{t} \left[ P_{qq}(z) - \frac{2C_F}{1 - z} \right] \quad (\text{VI.6})$$

<sup>2</sup>A formal comparison of other uncertainties and of the relation between parton showers and analytic NLL resummation for simple observables has been presented in [1007].

for gluon emission off (anti)quarks, and to

$$\begin{aligned} \mathcal{K}_{\tilde{ij} \rightarrow i,j}(t, z) &= \frac{\alpha_s(kbt)}{2\pi} \left( 1 + \frac{\alpha_s(bt)}{2\pi} \beta_0 \ln k \right) \frac{1}{t} \left[ \frac{2C_A}{1-z} + \frac{2C_A}{z} \right] \\ &+ \frac{\alpha_s(kbt)}{2\pi} \frac{1}{t} \left[ P_{gg}(z) - \frac{2C_A}{1-z} - \frac{2C_A}{z} \right] \end{aligned} \quad (\text{VI.7})$$

for gluon emission off gluons. For the first (and partially also subsequent) parton-shower emissions, Eqs. (VI.6) and (VI.7) are potentially rescaled with finite matrix-element correction factors. Note that the implementation does enforce that the compensation terms are only applied to pieces yielding double-logarithmic contributions. However, the two-loop cusp term is not only added to the soft contribution, but also enters in the hard- and collinear phase sectors, as is common for implementations of the CMW scheme. The difference w.r.t. Eq. (VI.3) is of  $\mathcal{O}(\alpha_s^3)$  in the soft region, and of  $\mathcal{O}(\alpha_s^2)$  in the hard and collinear regions.

### 1.3.2 Sherpa

The SHERPA Monte Carlo event generator [415] in its latest release, Sherpa-2.2.4, comprises two parton shower algorithms: CSS [491] and DIRE [492]. Both are based on Catani-Seymour [187, 489] dipole splitting functions. While CSS is constructed along the lines of a standard, transverse momentum ordered parton shower, DIRE combines the standard treatment of collinear configurations in parton showers with the resummation of soft logarithms in color dipole cascades. The evolution kernels of CSS can be written in the form of Eq. (VI.2), while in DIRE  $b(t) \rightarrow 1$  and the two-loop cusp anomalous dimension multiplies the soft enhanced terms of the splitting functions with  $1 + \alpha_s/(2\pi)K(t)$ .

When varying the argument of the strong coupling constant, i.e. replacing  $t \rightarrow kt$  with  $k$  a constant, the higher-logarithmic structure induced by the running of the coupling constant in the presence of the CMW scale factor needs to be preserved in order not to change the resummation implemented by the parton shower. Thus, in [1] the following naive replacement was used

$$\alpha_s(b(t)t) \rightarrow \alpha_s(b(t)kt) f(k,t), \quad f(k,t) = 1 + \sum_{i=0}^{n_{\text{th}}+1} \frac{\alpha_s(b(t)t)}{2\pi} \beta_0(n_f(t)) \log \frac{t_i}{t_{i-1}} \quad (\text{VI.8})$$

The sum runs over the number  $n_{\text{th}}$  of parton mass thresholds in the interval  $[t, kt]$  with  $t_0 = t$ ,  $t_{n_{\text{th}}+1} = kt$  and  $t_i$  are the encompassed parton mass thresholds. If  $k < 1$ , the ordering is reversed, recovering the correct sign. It is clear that this method will largely eliminate the dependence of the overall prediction on  $k$ . In order to obtain a realistic uncertainty estimate, we use Eq. (VI.3).

In the case of the CSS shower, we remove the explicit dependence on  $K(t)$  and reweight instead with a factor  $\alpha_s(b(t)t)/\alpha_s(t)$ , which leads to the definition of the final-state emitter, final-state spectator branching probabilities

$$\mathcal{K}_{\tilde{ij} \rightarrow i,j}(t, z) = \frac{\alpha_s(kbt)}{2\pi} \frac{1}{t} \frac{2C_F}{1-z(1-y)} f(k,t) + \frac{\alpha_s(kbt)}{2\pi} \frac{1}{t} \left[ V_{ij,k}(z,t) - \frac{2C_F}{1-z(1-y)} \right] \quad (\text{VI.9})$$

for gluon emission off (anti)quarks, and

$$\begin{aligned} \mathcal{K}_{\tilde{ij} \rightarrow i,j}(t, z) &= \frac{\alpha_s(kbt)}{2\pi} \frac{1}{t} \left[ \frac{2C_A}{1-z(1-y)} + \frac{2C_A}{y+z(1-y)} \right] f(k,t) \\ &+ \frac{\alpha_s(kbt)}{2\pi} \frac{1}{t} \left[ V_{ij,k}(z,t) - \frac{2C_A}{1-z(1-y)} - \frac{2C_A}{y+z(1-y)} \right] \end{aligned} \quad (\text{VI.10})$$

for gluon emission off gluons. The  $V_{ij,k}$  are the splitting kernels in [187]. This will generate higher-logarithmic corrections compared to Eq. (VI.3), which are, however, much smaller than the ones produced by Eq. (VI.8).

In the case of the DIRE shower, we use the branching probabilities

$$\mathcal{K}_{\tilde{ij} \rightarrow i,j}(t, z) = \frac{\alpha_s(kt) C_F}{2\pi} \frac{1}{t} \left[ \frac{2(1-z)}{(1-z)^2 + \kappa^2} \left( f(k, t) + \frac{\alpha_s(t)}{2\pi} K(t) \right) - (1+z) \right] \quad (\text{VI.11})$$

for gluon emission off (anti)quarks, and  $\mathcal{K}_{\tilde{ij} \rightarrow i,j}(t, z) = \mathcal{K}_{gg}(t, z) + (i \leftrightarrow j)$ , for gluon emission off gluons in the final state. The unsymmetrized gluon-to-gluon kernel is given by

$$\mathcal{K}_{gg}(t, z) = \frac{\alpha_s(kbt) 2C_A}{2\pi} \frac{1}{t} \left[ \frac{(1-z)}{(1-z)^2 + \kappa^2} \left( f(k, t) + \frac{\alpha_s(t)}{2\pi} K(t) \right) - 1 + \frac{z(1-z)}{2} \right]. \quad (\text{VI.12})$$

Note that the definition of  $\kappa^2$  depends on the type of dipole and is given in [492]. For gluon emission off gluons in the initial state we use

$$\mathcal{K}_{\tilde{ij} \rightarrow i,j}(t, z) = \frac{\alpha_s(kbt) 2C_A}{2\pi} \frac{1}{t} \left[ \frac{(1-z)}{(1-z)^2 + \kappa^2} \left( f(k, t) + \frac{\alpha_s(t)}{2\pi} K(t) \right) - 2 + \frac{1}{z} + z(1-z) \right] \quad (\text{VI.13})$$

## 1.4 Results

In the following, we will use the baseline parton-shower settings  $\alpha_s(M_Z) = 0.118$ , use two-loop running of  $\alpha_s$ , and employ the CMW prescription for all parton showers except DIRE. It is important to emphasize that these settings are only chosen to aid the comparison of parton shower variations, and are not based on the resulting quality of data description. With this in mind, we will also refrain from comparing our results to data, although we will employ the analyses [620] and [1011], as implemented in the RIVET framework [500]. We then perform variations of  $\alpha_s(kt)$  with  $k = \{1/4, 1, 4\}$  (and  $t$  of dimension  $\text{GeV}^2$ ) for both  $e^+e^-$  collisions at LEP I and gluon-induced Higgs-production at a 13 TeV LHC. For the latter, we do not include multiple parton interactions, primordial  $k_T$  modelling, QED effects or hadronization in the predictions. We use the parton distribution functions given in NNPDF30\_nlo\_as\_0118 [422] and implemented in LHAPDF [964]. Finally, for the LHC setup, we use  $\mu_Q = 125 \text{ GeV}$  as starting scale of the parton shower. Please note again that these settings define a theory study, and should *not* be considered a recommendation when performing a detailed experimental analysis.

The results will be presented as a main plot and three adjoint ratios. The latter show the predictions of one shower, compared to the baseline prediction of this same shower. The gray band represents the current default variation band of the tool, while the colored bands give the results of the schemes proposed in Sec. 1.2. The hatched gray bands show naive variations without any compensation terms, as a baseline. From these, it is immediately obvious that the inclusion of compensating terms has significant effects.

### 1.4.1 LEP results

In this section, we compare the MCEGs to each other, using the analyses of [1011] and [620]. We include both parton showering and hadronization into the simulations.

Figure VI.1 shows the separation between jets in the Durham algorithm. As expected, we observe that the default (gray) SHERPA variation is much more aggressive, while default (gray) PYTHIA is much more conservative. The variation bands of PYTHIA, CSS and DIRE are much more similar in size, highlighting that although different single-logarithmic terms in

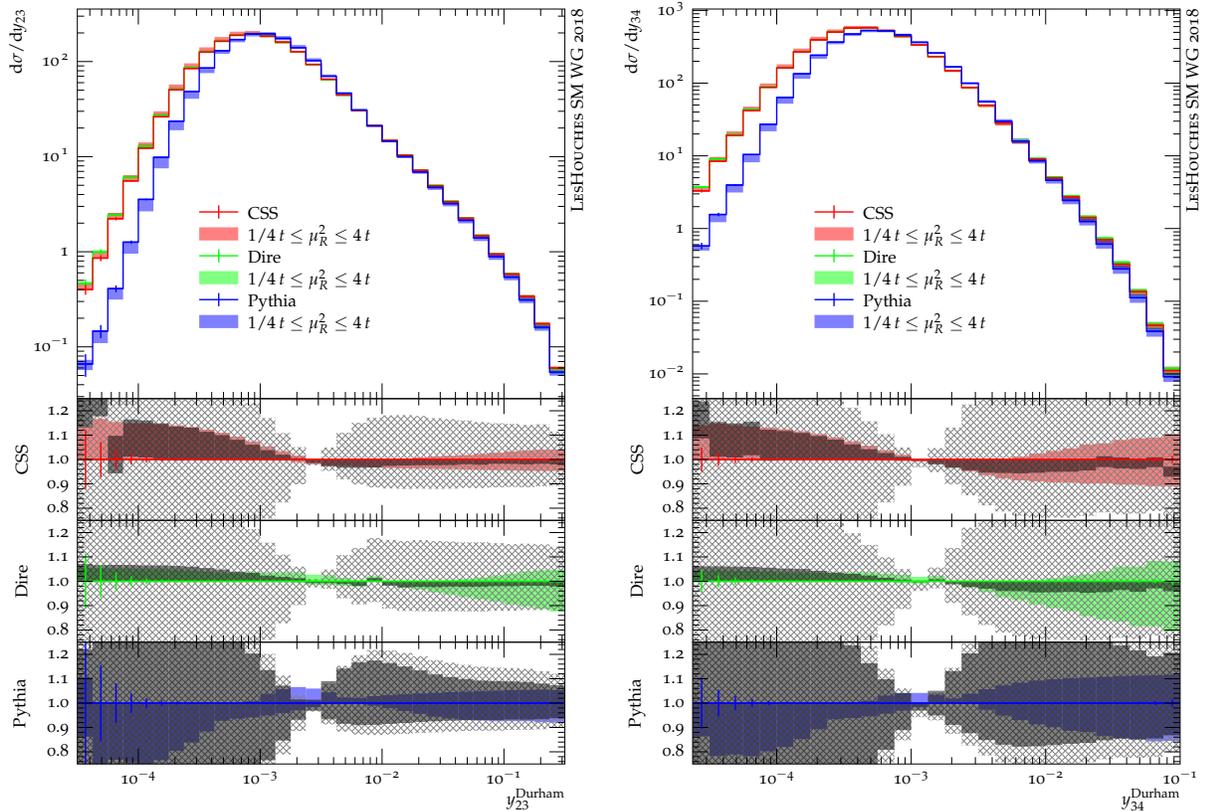


Fig. VI.1: Uncertainty estimates for the Durham  $2 \rightarrow 3$  jet rate (left) and  $3 \rightarrow 4$  jet rate (right) in the analysis setup of [1011]. Gray bands show the estimates using the method of [1] for SHERPA and DIRE, and of [1002] for PYTHIA. Blue bands give PYTHIA variations according to Sec. 1.3.1. Red and green bands give, respectively, SHERPA and DIRE variations according to Sec. 1.3.2. Hatched gray bands show naive variations without any compensation terms.

different showers, the minimal prescription leads to a convergence. The region below  $\sim 10^{-3}$  is dominated by hadronization effects, and thus, the differences between MCEGs here should not be considered problematic, since it heavily depends on how the event generator tune had been produced. It is however interesting to observe that the PYTHIA bands in the non-perturbative region are reduced in the minimal prescription, which will clearly aid in finding MCEG tunes including scale variations<sup>3</sup>

Similar conclusions can be drawn from Fig. VI.2. We again observe that the minimal prescription leads to a variation band that is more conservative than the default CSS and DIRE scheme, and more aggressive than the default PYTHIA variation. The width of the band becomes comparable in all three showers. Note again that the regions of  $\tau \sim 1$  and  $C < 0.25$  show large hadronization corrections, i.e. that the difference of the three tools in this region is dominated by tuning (and by using  $\alpha_s$  setups that are rather different from those used in the tunes). Again, it is worth mentioning that a smaller perturbative variation band in the non-perturbative region will ease tuning in the presence of perturbative variations.

These results – and other distributions not shown here – indicate that at LEP, varying the renormalization scale in the parton shower by factors of  $\frac{1}{4}, 1, 4$  (i.e.  $\frac{1}{2}, 1, 2$  when applied to GeV-valued scales) in the “minimal” prescription leads to uncertainty bands of  $\mathcal{O}(10-15\%)$  in all

<sup>3</sup>Such a study was performed elsewhere in the proceedings of this workshop.

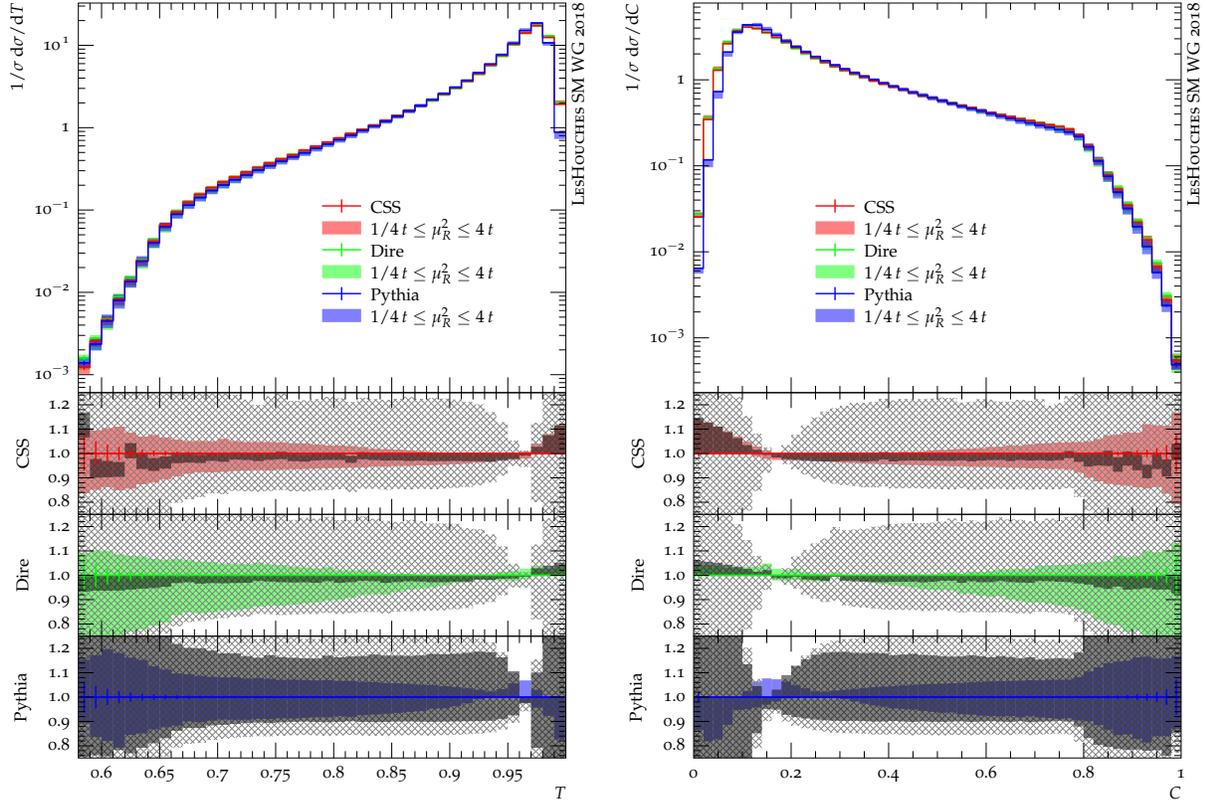


Fig. VI.2: Uncertainty estimates for the thrust (left) and C parameter (right) in the analysis setup of [620]. The gray bands show the estimates using the method of [1] for SHERPA and DIRE, and of [1002] for PYTHIA. Blue bands give PYTHIA variations according to Sec. 1.3.1. Red and green bands give, respectively, SHERPA and DIRE variations according to Sec. 1.3.2. Hatched gray bands show naive variations without any compensation terms.

three showers considered here. This improves over previous attempts that should be considered as too aggressive.

#### 1.4.2 LHC results

We will now discuss the impact of variations on distributions in gluon-induced Higgs-boson production at a 13 TeV LHC. The results of this section are at the parton level, i.e. we do not include the effect of multiple parton interactions, beam remnants, hadronization or other non-perturbative effects.

Figure VI.3 shows the Higgs-boson transverse momentum on logarithmic and linear scale, with the latter focussing on the low- $p_{\perp}$  region interesting for parton shower resummation. Before discussing the size of the bands, it is interesting to see that PYTHIA and the SHERPA CSS shower produce very similar results for  $2 \text{ GeV} < p_{\perp h} < 100 \text{ GeV}$ . The differences above 100 GeV would be ameliorated upon NLO matching or multi-leg merging, while the differences below 2 GeV are covered by non-perturbative effects in a hadron-level simulation. The DIRE shower predicts a rather different shape of the  $p_{\perp}$  spectrum above the Sudakov peak, which arises from the different choice of ordering variable compared to the CSS and PYTHIA. This difference is not covered by the respective uncertainty bands, because effects relating to momentum conservation and ordering variables are an inherently different degree of freedom of the parton shower implementation [1007].

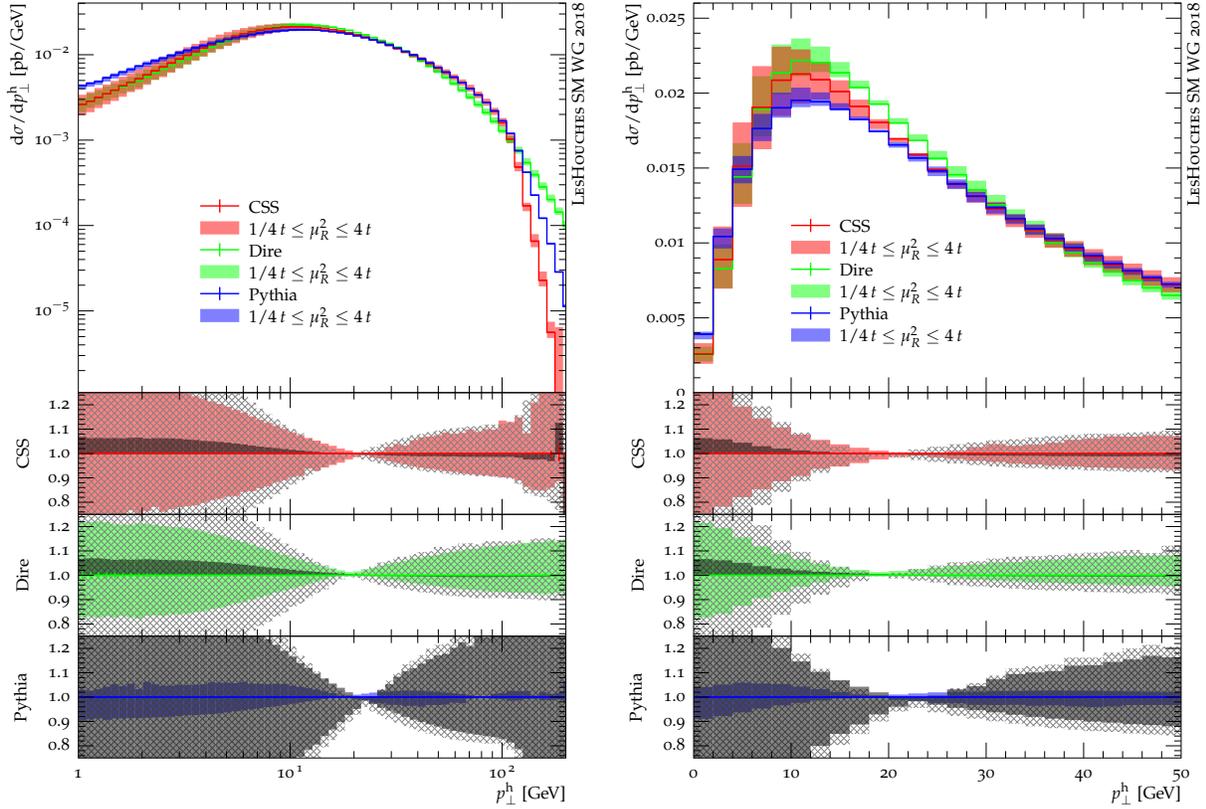


Fig. VI.3: Impact of shower variations on the Higgs  $p_{\perp}$  spectrum. The gray bands show the estimates using the method of [1] for SHERPA and DIRE, and of [1002] for PYTHIA. Blue bands give PYTHIA variations according to Sec. 1.3.1. Red and green bands give, respectively, SHERPA and DIRE variations according to Sec. 1.3.2. Hatched gray bands show naive variations without any compensation terms.

The effect of employing the minimal compensation scheme on the variation bands is very different in PYTHIA and SHERPA. As before, we see that the minimal scheme squeezes the band w.r.t. the default scheme in PYTHIA, while the the minimal scheme widens the band in CSS and DIRE. However, the PYTHIA band within the minimal scheme reduces to 10% or less, the CSS band increases beyond 50% at low  $p_{\perp h}$  and 20% at high values, and the DIRE band spans up to 35% variation at low  $p_{\perp h}$  and 20% at high values. These trends are reversed in the  $\Delta R$  and  $m_{jj}$  distributions shown in Fig. VI.4. Here, the PYTHIA band in the minimal scheme yields the largest variation, while the band of CSS is almost vanishing.

These effects should make clear that, as anticipated in the motivation, the effect of the compensation is observable-dependent. In particular, since the difference of parton shower evolution variables will lead to different single-logarithmic terms, the effect of compensation can, for one observable, be quantitatively different for different showers. Gluon-induced Higgs-boson production is, due to high parton-shower activity, an ideal laboratory to highlight these subtleties.

## 1.5 Discussion and Summary

In this study, we have investigated renormalization scale variations in parton showers. We have discussed compensation terms to ensure that such scale variations do not deteriorate expressly introduced higher-order corrections, and attempt to define a common baseline for applying such

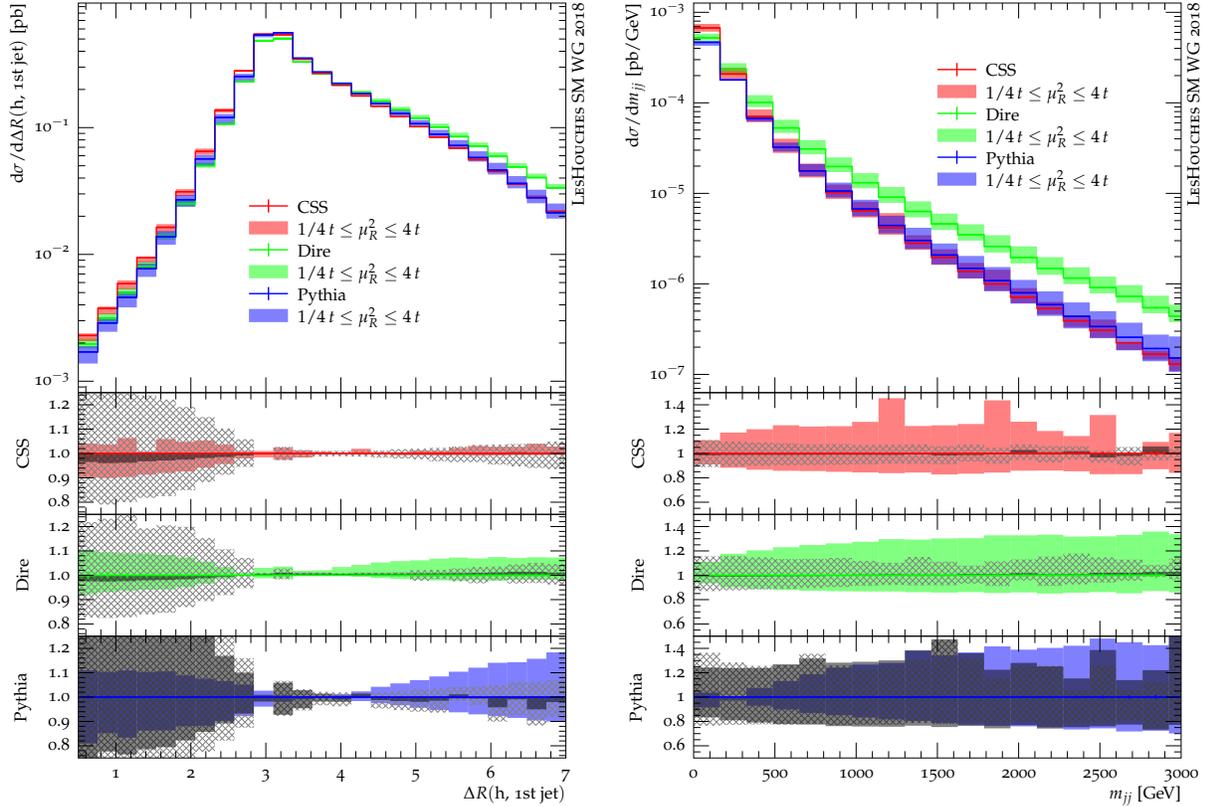


Fig. VI.4: Impact of shower variations on the Higgs-jet  $\Delta R$  separation and dijet invariant mass in Higgs + jets events. The gray bands show the estimates using the method of [1] for SHERPA and DIRE, and of [1002] for PYTHIA. Blue bands give PYTHIA variations according to Sec. 1.3.1. Red and green bands give, respectively, SHERPA and DIRE variations according to Sec. 1.3.2. Hatched gray bands show naive variations without any compensation terms.

compensation in PYTHIA and SHERPA. We have then compared this new “minimal” prescription to traditional LEP observables and to the result of the previous default variation schemes in the generators. This comparison suggests that in the minimal scale variations by factors  $k = \frac{1}{4}, 4$  lead to reasonable agreement between the three showers used in the study, while not impeding the ability to tune the event generator due to large perturbative variations in hadronization-dominated phase space regions. The study of Higgs-boson production at the LHC does not lead to conclusive results, since the shower- and observable-dependence of the variation bands is exposed by the high parton-shower activity. In this case, we recommend to use the most conservative uncertainty estimate available in the event generator.

## Acknowledgements

We thank the organizers for an inspiring workshop. This work was supported by the US Department of Energy under the contract DE-AC02-76SF00515. This manuscript has been authored by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics.

## 2 On the cross-talk of parameter optimization and perturbative variations in event generators <sup>4</sup>

We discuss the interplay of parton-shower variations and optimization of non-perturbative parameters in general-purpose event generators. To understand the correlation between infrared cut-offs and hadronization parameters with variations of  $\alpha_s$ , we investigate the interplay of tuning and  $\alpha_s$ -variations for  $e^+e^-$  colliders. For LHC collisions, we investigate the cross-talk of  $\alpha_s$ - and PDF variations.

### 2.1 Introduction

General-Purpose Monte Carlo Event Generators (MCEG) [831] are the backbone of collider-physics analysis prototyping, while at the same time providing the accurate background predictions to new-physics searches. With ever more detailed measurements at the LHC, the focus of MCEG developments has been providing and enabling precision calculations in perturbative QCD. At this point, it is important to stress that the most accurate simulation should predict unknown phenomena based on known effects. The latter will include the results of previous measurements as well as higher order calculations. Event generators combine both into a simulation that allows predictions for new analyses and collider setups by using higher order calculations and transferring, or extrapolating, a realistic model of previous experimental data. To this end, MCEG parameters are commonly “tuned” to many different data. This “tuning” helps constrain parameters that cannot be fixed by theory considerations alone, and encodes the global picture of previous measurements into the MCEG. On top of the (mostly non-perturbative QCD) parameters that are constrained by this exercise, MCEGs also heavily rely on pre-tabulated parton distribution functions (PDFs) [1012] to describe the longitudinal structure of colliding hadrons. The non-perturbative component of these PDFs is usually “extracted” by fitting the parameters of the PDF parametrizations to measurements of (mostly inclusive) scattering cross-sections. In either case, we expect strong correlations between non-perturbative parametrizations and perturbative input.

This has made a comprehensive error budget for MCEGs a long-standing problem. Tuning will thus have the effect of mixing perturbative and non-perturbative effects, and will potentially also account for deficiencies in the perturbative prediction. Thus, a well-defined event generator uncertainty should include both effects, and correctly incorporate the cross-talk between them. This is a daunting task in principle, but also in practice, since correlations and subtleties can easily be masked by the overall statistical uncertainty inherent to finite-time Monte Carlo simulation. With the advent of weighting techniques [501, 1001, 1002], some of the technical difficulties of correlated parton-shower variations were recently ameliorated. Thus, we will investigate the correlation of perturbative and non-perturbative MCEG pieces with these new tools. In this study, we do not attempt to provide a theoretical answer to the question of defining uncertainties. Instead, we provide a phenomenological study of some cross-talk, in hopes of gaining insight for the future.

This article is organized into two main sections followed by the summary. The main sections are in spirit intimately connected, in that we try to assess different perturbative-non-perturbative correlations from different angles. The first part is devoted to the cross-talk between the tuning of the parameters of hadronization models and final-state parton shower variations. The second part investigates the correlation between PDF choices and initial-state parton showering.

---

<sup>4</sup> J. Bellm, L. Lönnblad, S. Plätzer, S. Prestel, D. Samitz, A. Siodmok, A. H. Hoang

## 2.2 Generator tuning vs. parton-shower variations

In this section, we will investigate the cross-talk between non-perturbative hadronization parameters and changes of  $\alpha_s$  in the (perturbative) parton shower evolution. Parameters in hadronization models are commonly extracted at lepton colliders and then used unchanged at hadron colliders (with the primordial kT of partons inside of colliding hadrons being one notable exception). To not over-complicate the study, we thus discuss the cross-talk using LEP measurements.

### 2.2.1 Methodology

In order to investigate the interplay between generator tuning and parton-shower variations studied in [501, 960, 1002] we use two different parton shower Monte Carlo event generators HERWIG 7 [584] and PYTHIA 8 [629, 630] to produce a fully exclusive simulation of the events at the Leading Order. Both generators provide an option for using different shower modules, however at this stage of the project we only use default parton showers algorithms<sup>5</sup>: the angular ordered shower in HERWIG [702] and  $p_{\perp}$ -ordered shower in PYTHIA [1013]. For generator tuning we use the parametrization-based tune method provided by the PROFESSOR package [1014]. The starting point for the tuning procedure is the selection of a range  $[p_i^{\min}, p_i^{\max}]$  for each of the  $N$  tuning parameters  $p_i$ .

Since we want to explore the correlation of generator tuning and parton-shower variations our selection of parameters consist of the main parameters of non-perturbative hadronization models and the parton shower cut-off scale. To be more precise in HERWIG we adjust the shower cut-off (`pTMin`), and main hadronization parameters of the cluster model [703]: gluon constituent mass (`g:ConstituentMass`) and splitting parameter for clusters `PSplit` (not flavour specific<sup>6</sup>). Initially we also included into the procedure flavor depended parameters: `ClMaxs` and `ClPows`, however since the data selection was not really able to constrain the parameters we decided to keep them close to their default values and effectively remove from the tuning procedure. In PYTHIA except for the shower cut-off (`TimeShower:pTmin`) we also tune three main hadronization parameters of the Lund model [828]: a- and b-parameter (`StringZ:aLund` and `StringZ:bLund`) and non-perturbative pt-width (`StringPT:Sigma`).

For both generators, event samples are generated for random points in this multidimensional hypercube in their parameter space. The number of points sampled is chosen, depending on the number of input parameters, to ensure good control of the final tune. Each generated event is directly handed over to the Rivet package [500], which implements the experimental analyses. Thus the results for each observable are calculated at each set of parameter values. PROFESSOR parametrizes each bin of each histogram as a function of the input parameters. It is then able to find the set of parameters that fit the selected observables best. As a user, one simply has to choose the set of observables that one wishes to tune to and, optionally, their relative weights in the fit.

Usually, the hadronization and shower parameters are tuned to a wide range of experimental data, however, most of them are from LEP. Therefore, in our tuning effort, we also use LEP data. In particular we use ALEPH data from [620], however, we have not tuned to all available distributions<sup>7</sup>. We have, *e.g.* left out the higher jet multiplicities while including the lower ones. We also note that although the average multiplicity is included, it is only one bin and will therefore not constrain the tuning very much. It should be noted that the tunings done here are mainly for illustration and are not optimized in any way nor do they reflect the optimal

---

<sup>5</sup>In the future, we plan to use also Dipole Shower of HERWIG [994] and Dire [492]

<sup>6</sup>Cluster model implemented in HERWIG has possibility to set splitting parameters separately for light (`PSplitLight`), bottom (`PSplitBottom`) and charm (`PSplitCharm`) clusters, see [825] for details. However, in our exercise for simplicity we keep them all the same.

<sup>7</sup>The full list of selected observables and their weights are provided in Appendix 2.A.1.

outcome with respect to a global tuning effort.

To assess the effect of the correlation between infrared cut-offs and hadronization parameters with variations of  $\alpha_s$ , for each MCEG we perform two sets of tunes with the scale variation bands:

1. The central (reference) tune, is obtained by using a "central"  $\alpha_s(M_z)$  in the tuning procedure. This is simply chosen to be the default value in the respective generator. Its scale variation band (the envelope of the tune) is estimated by shifting  $\alpha_s(M_z) \rightarrow \alpha_s(\{\frac{1}{\xi}, 1, \xi\}M_z) \rightarrow \alpha'_s(M_z)$ , with<sup>8</sup>  $\xi = 2$  without additional retuning for each  $\alpha'_s(M_z)$  variation.
2. Two more tunes, called "retune" and the corresponding scale variation band, are obtained by the same  $\alpha_s(M_z)$  variation as for "central" tune but this time we also retune to ALEPH data for each  $\alpha'_s(M_z)$  variation.

In the next section we compare in details the results from the both MCEG using their "central" and "retune" versions of the tunes.

### 2.2.2 Results

We start by looking at the distributions that were actually tuned to. In Fig. VI.5 we show the distribution in jet resolutions where an event with three jets would be clustered into two,  $y_{23}$ , and the *thrust minor* ( $T_{\text{minor}}$ ) distribution (the thrust calculated wrt. the axis out of the event plane spanned by the thrust and thrust major axes), respectively. We see immediately that both generators are able to fit the distributions fairly well and, as expected, the scale variations of the central tune are much larger than the ones obtained when retuning for each scale choice.

We note that since the distributions are normalised to unity, the scale variations are artificially reduced around the center of the distributions. Looking more closely, we see some imperfections in the fits. For example, we see that both generators have problems describing the region of large  $T_{\text{minor}}$  region, corresponding to large activity out of the event plane driven by hard effects of  $\mathcal{O}(\alpha_s^2)$ , where neither generator includes matrix element corrections in our studies. Similarly, HERWIG has problems describing large  $y_{23}$  values (small  $-\ln(y_{23})$ ). However, this is as expected since the HERWIG simulations had switched-off matching to the  $\mathcal{O}(\alpha_s)$  tree-level matrix elements, while the PYTHIA simulation had it on. We would also like to note that in the region of large  $y_{23}$  (small  $-\ln(y_{23})$ ) the scale variations for PYTHIA are not reduced by the retuning. This is expected since there is no mechanism by which the non-perturbative parameters in PYTHIA can compensate for the effect of scale changes on hard jets. This is, however, not the case for HERWIG, where sometimes a cluster with a large mass is allowed to decay isotropically into two lighter clusters, then decaying into hadrons and ending up as jets in the Durham algorithm. This is regulated by the `PSplit` parameter which in this case can compensate somewhat for the scale variations at large  $y_{23}$ .

Turning now to the predictions for observables not tuned to, we first stay at the  $Z^0$  peak and look at the Durham  $y_{54}$  distribution in Fig. VI.6. Here we see again that the retuned scale variations are smaller than the central ones. We see the same pattern as for the  $y_{23}$  distribution, that for large  $y_{45}$  the scale variations are not much reduced for PYTHIA, while for HERWIG the reduction is significant. Comparing with the data, one could argue that the central scale variations overestimate the uncertainties in the predictions, while the retuned scale variations, where we use data to constrain the fit, give a more reasonable estimate of the uncertainties involved. Although in the region of low  $y_{45}$  where the retuned variations could be argued to be too small.

---

<sup>8</sup>We also performed studies for  $\xi = \sqrt{2}$ , however in order to show clearly the effect of the retuning we only present results for  $\xi = 2$ .

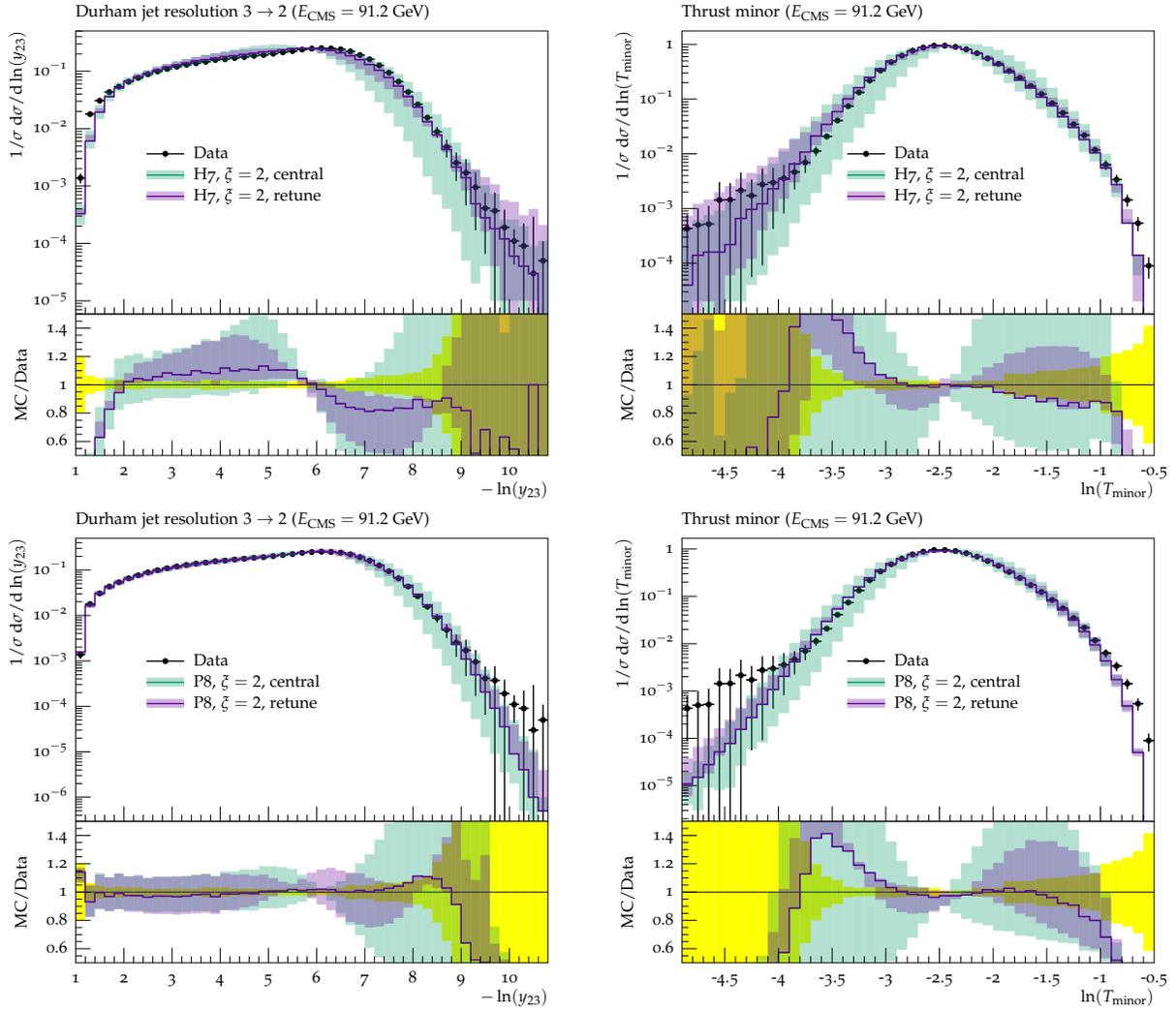


Fig. VI.5: Examples of the resulting tunes, "central" and "retune", with their uncertainty band compared to ALEPH data at 91.2 GeV. Top plots are HERWIG and bottom is PYTHIA. Left plots are the Durham jet resolution  $3 \rightarrow 2$  distribution, and the right plots are the Thrust minor distribution. The both observables were used in the tuning procedure.

Moving on to extrapolations in energy, we first look at the average charged multiplicity in Fig. VI.7. Here we see an unexpected behaviour for PYTHIA in the middle plot. The scale variations are actually increased with the tuning. This is because, even though the 91 GeV point is included in the tuning set, it is only one bin and the tuning procedure has clearly decided to sacrifice this bin in order to fit many others in other distributions. There are indication that one of the variations in HERWIG leave more space for other parameters to still adopt to the multiplicity bin, though this issue deserves a more detailed investigation. In the right plot, we have instead included the full multiplicity distribution from L3 [1015] in the tunings and there we get the expected behaviour, similar to that of HERWIG.

Finally we return to  $y_{23}$  and  $T_{\text{minor}}$  in Fig. VI.8, but at the higher collision energy, where we again see that the retuned scale variations in the predictions gives a more reasonable estimate of the uncertainties than the variations using the central tune. Again we see that for PYTHIA the variations are not reduced in the hard part of the spectra (small  $y_{23}$  and large  $T_{\text{minor}}$ ), while the retuning in HERWIG is able to compensate for the scale variations.

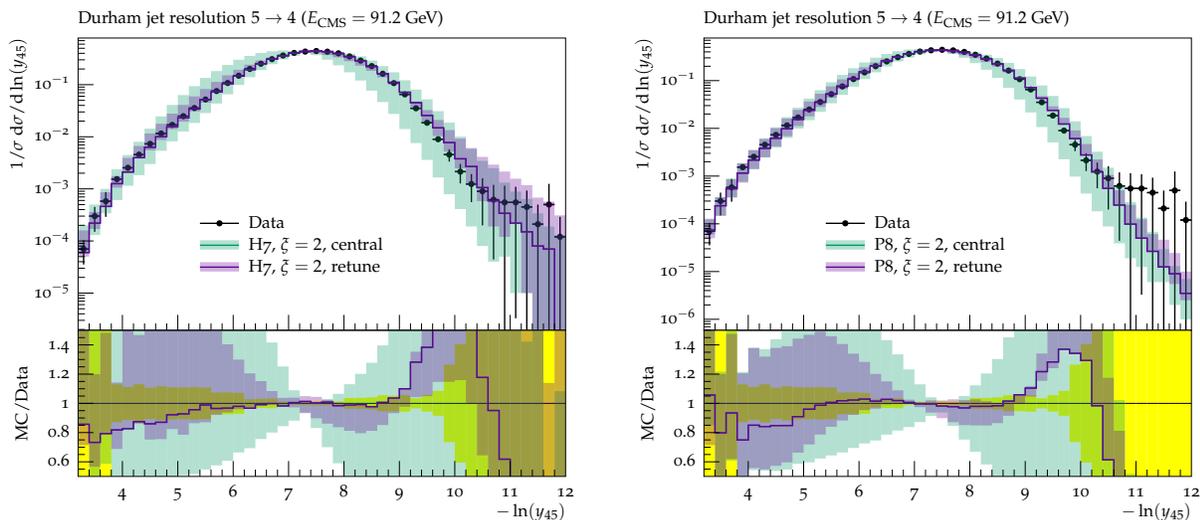


Fig. VI.6: The distribution of the Durham jet resolution scale  $5 \rightarrow 4$  at LEP for HERWIG (left) and PYTHIA (right).

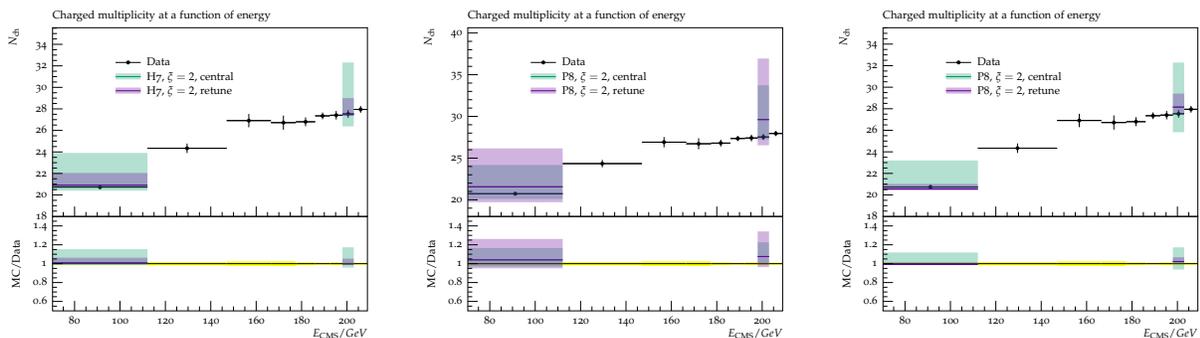


Fig. VI.7: Average charged multiplicity as a function of energy. Only two energies (91 and 200 GeV) are shown for HERWIG (left) and PYTHIA (middle). The right plot shows an alternative tuning of PYTHIA which includes the charged jet multiplicity at 91 GeV from L3 [1015].

### 2.2.3 Discussion

We think it is safe to say that using retuned scale variations gives a better estimate of the uncertainties in the predictions of shower/hadronisation models. This does however not mean that it gives a better estimate of the uncertainties in general. It is, in any case, an interesting exercise to make, as it can increase our understanding of how the models behave and how the tunings work in more detail. From our results it is *e.g.* clear that not using leading order matrix element matching in HERWIG can be compensated also in the hard regions by non-perturbative parameters in the hadronisation. Similarly, in PYTHIA it is possible to compensate some shortcomings in the description of several event shapes at the expense of an accurate description of the total multiplicity.

## 2.3 PDF fits and parton-shower variations

In this section, to complement the previous section, we will investigate the cross-talk between (collinear) non-perturbative initial-state parton distributions and changes of  $\alpha_s$  in the (perturbative, initial-state) parton shower evolution. We expect correlations due to the presence of PDF ratios in the evolution kernels used in backward parton-shower evolution [1016]. We can illus-

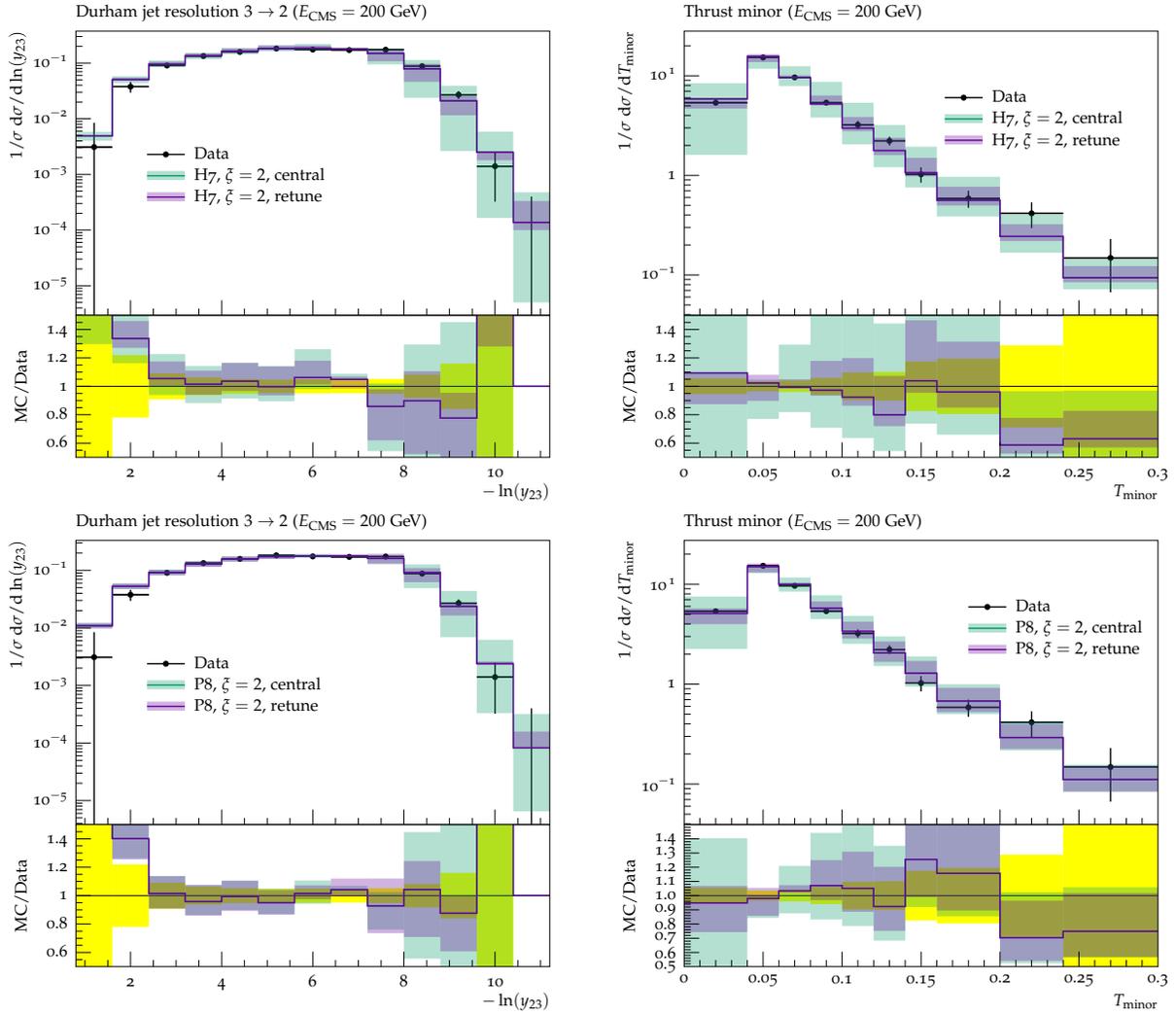


Fig. VI.8: Predictiona for 200 GeV from the tunes to ALEPH data at 91 GeV. Top plots are HERWIG and bottom is PYTHIA. Left plots are the tune Durham jet resolution  $3 \rightarrow 2$  distribution, and the right plots are the Thrust minor distribution.

trate this effect by dissecting the parton-shower result for the Drell-Yan lepton pair transverse momentum at a hadron collider. In an initial-state parton shower that is ordered in kinematical  $p_\perp$  with respect to the incoming proton beams, the first emission will, schematically, yield the distribution<sup>9</sup>

$$\begin{aligned}
 \frac{d^n \sigma}{d\Phi_0 dp_\perp} &= d\hat{\sigma}(pp \rightarrow Z; \Phi_0) f_{p_+}^{\sim}(x_+, \mu_F) f_{p_-}^{\sim}(x_-, \mu_F) \\
 &\otimes \left\{ \int dz_+ \Pi(\mu_Q, p_\perp) \frac{\alpha_s(p_\perp)}{2\pi} \frac{f_{p_+}^{\sim}(z_+, p_\perp)}{f_{p_+}^{\sim}(x_+, p_\perp)} \frac{f_{p_-}^{\sim}(x_-, p_\perp)}{f_{p_-}^{\sim}(x_-, p_\perp)} P_{p_+ \rightarrow p_+}(z_+, p_\perp) + (- \leftrightarrow +) \right\}
 \end{aligned} \tag{VI.14}$$

<sup>9</sup> Note that we only use this as a schematic illustration. For parton showers not ordered in (this particular definition of)  $p_\perp$ , the result will be more complicated, but the main message will still apply. The symbols in Eq. (VI.14) are defined by  $d\hat{\sigma}(pp \rightarrow Z; \Phi_0) \rightarrow$  differential partonic cross section for Z-boson production at a ‘‘Born’’ phase space point  $\Phi_0$ ;  $f_{p_i}(x_i, \mu) \rightarrow$  PDF for parton flavour  $\tilde{p}_i$  with momentum fraction  $x_i$ , evaluated at scale  $\mu$  with hadron moving in the  $i = \pm p_z$  direction;  $\mu_Q \rightarrow$  parton-shower starting or veto scale;  $z_i \rightarrow$  energy fraction carried away from the initial line by emission off parton moving in  $i$ -direction;  $P_{\tilde{p}_i \rightarrow p_i}^{\sim} \rightarrow$  probability of producing parton  $p_i$  from parton  $\tilde{p}_i$  by a parton-shower branching.

with the parton-shower no-emission probability

$$\begin{aligned} \Pi(\mu_Q, p_\perp) = & \tag{VI.15} \\ & \exp \left\{ - \int_{p_\perp}^{\mu_Q} d\bar{p}_\perp \int d\bar{z}_+ \frac{\alpha_s(p_\perp)}{2\pi} \frac{f_{p_+}^\sim(\frac{x_+}{\bar{z}_+}, \bar{p}_\perp)}{f_{p_+}^\sim(x_+, \bar{p}_\perp)} \frac{f_{p_-}^\sim(x_-, \bar{p}_\perp)}{f_{p_-}^\sim(x_-, \bar{p}_\perp)} P_{p_+ \rightarrow p_+}^\sim(\bar{z}_+, \bar{p}_\perp) + (- \leftrightarrow +) \right\} \end{aligned}$$

and where, to ease notation, the symbol  $P(z, p_\perp)$  collects splitting kernels, propagator- and Jacobian factors. Assuming that the parton shower is a faithful implementation of DGLAP evolution of  $f_{p_+}^\sim$  (and  $f_{p_-}^\sim$ ) from  $\mu_Q = \mu_F$  to  $p_\perp$ , we can write

$$\begin{aligned} \Pi(\mu_Q, p_\perp) &= \frac{f_{p_+}^\sim(x_+, p_\perp)}{f_{p_+}^\sim(x_+, \mu_Q)} \frac{f_{p_-}^\sim(x_-, p_\perp)}{f_{p_-}^\sim(x_-, \mu_Q)} \Delta(\mu_Q, p_\perp) \quad \text{where} \tag{VI.16} \\ \Delta(\mu_Q, p_\perp) &= \exp \left\{ - \int_{p_\perp}^{\mu_Q} d\bar{p}_\perp \int d\bar{z}_+ \frac{\alpha_s(p_\perp)}{2\pi} P_{p_+ \rightarrow p_+}^\sim(\bar{z}_+, \bar{p}_\perp) + (- \leftrightarrow +) \right\} \end{aligned}$$

which leads to a result reminiscent of analytic resummation<sup>10</sup>

$$\begin{aligned} \frac{d^n \sigma}{d\Phi_0 dp_\perp} &= d\hat{\sigma}(pp \rightarrow Z; \Phi_0) f_{p_+}^\sim(x_+, p_\perp) f_{p_-}^\sim(x_-, p_\perp) \tag{VI.17} \\ & \left\{ \int dz_+ \Delta(\mu_Q, p_\perp) \frac{\alpha_s(p_\perp)}{2\pi} P_{p_+ \rightarrow p_+}^\sim(z_+, p_\perp) + (- \leftrightarrow +) \right\} \end{aligned}$$

It is important to stress again that this result assumes that the parton shower is a faithful implementation of DGLAP evolution of  $f_p^\sim$ , which in particular suggests that the parton-shower evolution should be adjusted for different PDF sets  $f_p'$ . Conversely, it is sensible to assume that particular parton shower settings should correspond to a particular choice of PDF.

In the following, we will investigate the cross-talk between the choice of  $\alpha_s$  in the parton shower and (input) PDF sets. To prevent bias, we will investigate the effect of correlations on Higgs-boson and Drell-Yan lepton pair  $p_\perp$  distributions at two different energies at a  $pp$  collider. The uncertainties on boson transverse momentum distributions often contribute heavily to background modelling uncertainties in experimental new physics searches at the LHC. MCEGs are often used to study and model  $p_\perp$  distributions. We expect that significant changes in the uncertainty estimates due to the inclusion of correlations will provide useful information for the experimental communities. We will further check our findings on the complementary observable  $\langle N_{jets} \rangle$ , which should be very sensitive to the details of the parton shower, but only mildly PDF dependent.

### 2.3.1 Methodology

We will assess the correlations by investigating how  $\alpha_s(M_Z)$  choices in the parton shower and PDF choices interact. For this, we first define the ‘‘baseline’’  $\alpha_s^{(c)}(M_Z) = 0.118$  and produce parton-shower results with  $\alpha_s^{(c)}(k \cdot t)$  with  $k \in \{\frac{1}{\sqrt{2}}, 1, \sqrt{2}\}$ . We have verified that a) the result of  $\alpha_s^{(c)}(\frac{1}{\sqrt{2}}t)$  is numerically equivalent to using  $\alpha_s^\downarrow(t)$  with  $\alpha_s^\downarrow(M_Z) = \alpha_s^{(c)}(\frac{1}{\sqrt{2}}M_Z) \approx 0.124$ , and that b) the result of  $\alpha_s^{(c)}(\sqrt{2}t)$  is numerically equivalent to using  $\alpha_s^\uparrow(t)$  with  $\alpha_s^\uparrow(M_Z) = \alpha_s^{(c)}(\sqrt{2}M_Z) \approx 0.112$ , see gray lines in second ratio plots of Fig. VI.9.

This allows to reinterpret parton-shower renormalization scale variations as variations of  $\alpha_s(M_Z)$ . For judiciously assigned values of  $\alpha_s(M_Z)$ , we can find PDF sets that used this  $\alpha_s(M_Z)$

<sup>10</sup>See e.g. [1017, 1018] for overviews of analytic resummation of color-singlet boson  $p_\perp$  spectra.

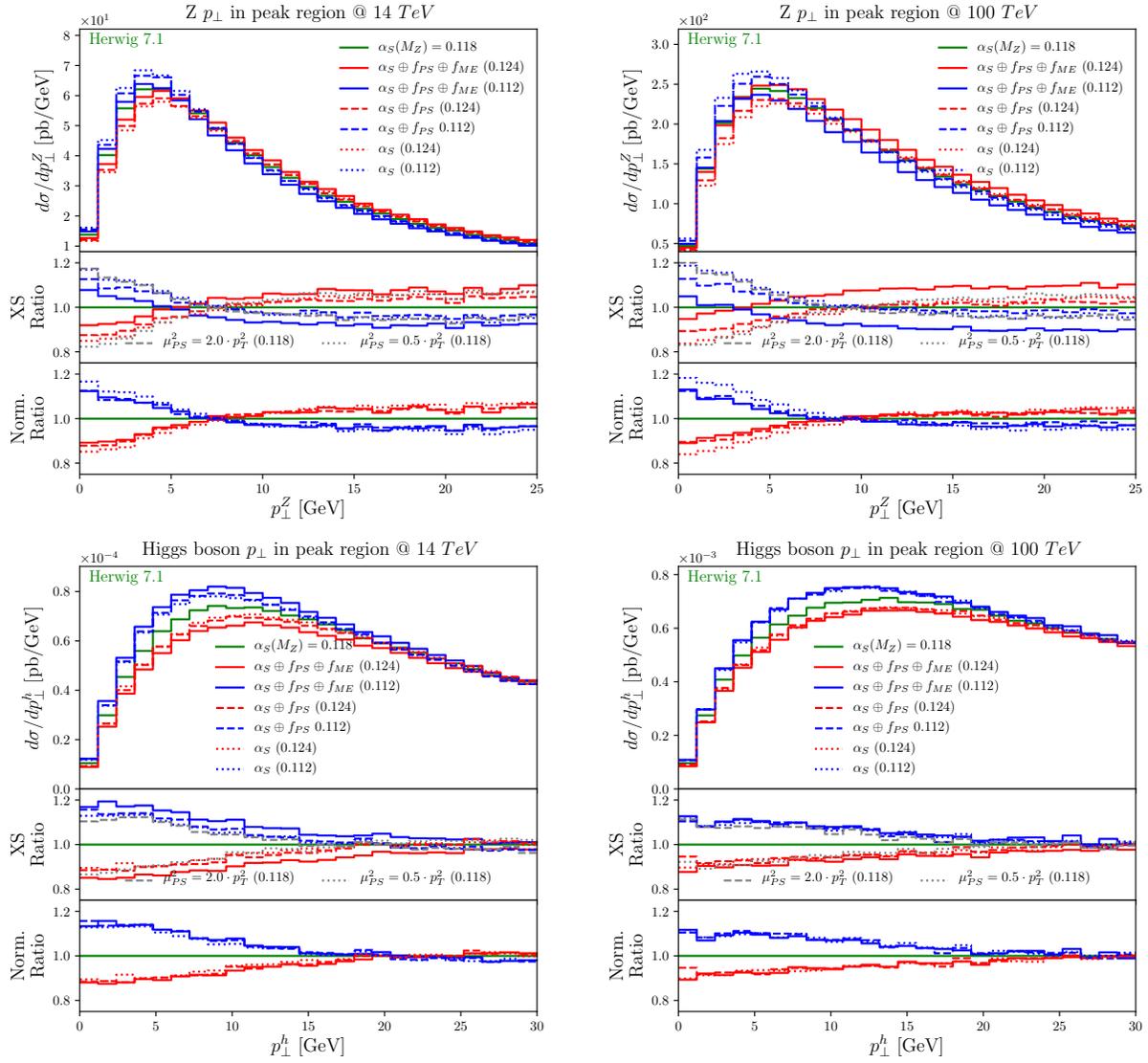


Fig. VI.9: Variation according to Table VI.1 for the bosons transverse momentum in Z boson (top) and Higgs (bottom) production at 14 (left) and 100 TeV (right). While the upper ratio plot shows the ratio to the baseline for differential cross section, the lower ratio plot illustrates the ratio w.r.t. normalized distributions. In the first ratio plot we add the lines for pure scale variations in the shower  $\alpha_S$  to display the similarity to value variations. See Sec. 2.3.4 for discussion.

when performing the fit. Such sets using a predetermined  $\alpha_s$  have recently become supplied by several PDF fitting groups. With these prerequisites, we compare the results with the settings given in Table VI.1.

As discussed above, we will focus the discussion on transverse momentum spectra. To investigate the effects of correlating  $\alpha_s$  and PDF further, we will check if changes in the  $p_\perp$  distribution also translate to the average number of jets  $\langle N_{jets} \rangle$ . The number of jets  $N_{jets}$  produced through parton showering is, by virtue of Eq. (VI.16), only sensitive to the value of the PDF at the parton-shower cut-off.  $\langle N_{jets} \rangle$  will thus only have a mild PDF dependence, but showcase the cross-talk between PDFs and  $\alpha_s$ .

Name	Description	$\alpha_s$ and PDF set	Legend in plots
Baseline:	Result of showering using only default settings and $\alpha_s^{(c)}(M_Z)$ .	$\alpha_s(M_Z) = 0.118$ , baseline PDF set, e.g. CT10nlo_as_0118	$\alpha_s(M_Z) = 0.118$ and green solid line
$\alpha_s^\downarrow$ variation	Result of changing $\alpha_s^{(c)}(M_Z) \rightarrow \alpha_s^{(c)}(\frac{1}{\sqrt{2}}M_Z)$ when showering.	$\alpha_s(M_Z) = 0.124$ , baseline PDF set, e.g. CT10nlo_as_0118	$\alpha_s(0.124)$ and red dotted line
$\alpha_s^\uparrow$ variation	Result of changing $\alpha_s^{(c)}(M_Z) \rightarrow \alpha_s^{(c)}(\sqrt{2}M_Z)$ when showering.	$\alpha_s(M_Z) = 0.112$ , baseline PDF set, e.g. CT10nlo_as_0118	$\alpha_s(0.112)$ and blue dotted line
Correlated PS variation	Result of simultaneously changing $\alpha_s^{(c)}(M_Z) \rightarrow \alpha_s^{(c)}(\frac{1}{\sqrt{2}}M_Z)$ and PDF set used for showering.	$\alpha_s(M_Z) = 0.124$ , PDF fitted with $\alpha_s(M_Z) = 0.124$ , e.g. CT10nlo_as_0124	$\alpha_s \oplus f_{PS}(0.124)$ and red dashed line
Correlated PS variation	Result of simultaneously changing $\alpha_s^{(c)}(M_Z) \rightarrow \alpha_s^{(c)}(\sqrt{2}M_Z)$ and PDF set used for showering.	$\alpha_s(M_Z) = 0.112$ , PDF fitted with $\alpha_s(M_Z) = 0.112$ , e.g. CT10nlo_as_0112	$\alpha_s \oplus f_{PS}(0.112)$ and blue dashed line
Correlated PS+ME variation	Result of simultaneously changing $\alpha_s^{(c)}(M_Z) \rightarrow \alpha_s^{(c)}(\frac{1}{\sqrt{2}}M_Z)$ , and the PDF set used for showering and in the calculation of the hard scattering cross section.	$\alpha_s(M_Z) = 0.124$ , PDF fitted with $\alpha_s(M_Z) = 0.124$ , e.g. CT10nlo_as_0124	$\alpha_s \oplus f_{PS} \oplus f_{ME}(0.124)$ and red solid line
Correlated PS+ME variation	Result of simultaneously changing $\alpha_s^{(c)}(M_Z) \rightarrow \alpha_s^{(c)}(\sqrt{2}M_Z)$ and the PDF set used for showering and in the calculation of the hard scattering cross section.	$\alpha_s(M_Z) = 0.112$ , PDF fitted with $\alpha_s(M_Z) = 0.112$ , e.g. CT10nlo_as_0112	$\alpha_s \oplus f_{PS} \oplus f_{ME}(0.112)$ and blue solid line

Table VI.1: List of settings to assess the correlations between PDFs and  $\alpha_s$  choices. In the text, we will refer to the envelope of  $\alpha_s^\downarrow$  and  $\alpha_s^\uparrow$  variation as “(uncorrelated)  $\alpha_s$  variation”, refer to the combination of the third and fourth items as “correlated PS variation”, and refer to the combination of the fifth and sixth curve as “correlated PS+ME” variation.

### 2.3.2 Simulation Setup

We use the default leading-order HERWIG setup for the simulations in this study. This includes matrix element corrections for the first jet in Higgs- and Z-boson production. As the baseline we use the central value of  $\alpha_s = 0.118$  by setting

```
set /Herwig/Shower/AlphaQCD:AlphaMZ 0.118;.
```

As PDF set, we use CT10nlo\_as\_0118, as implemented in LHAPDF [964]. Since our goal is to study the compensation effect of fitting in the evolution of parton showers, we limit most of the following to the ME+PS level. To turn off hadronisation and multiple parton interaction we set

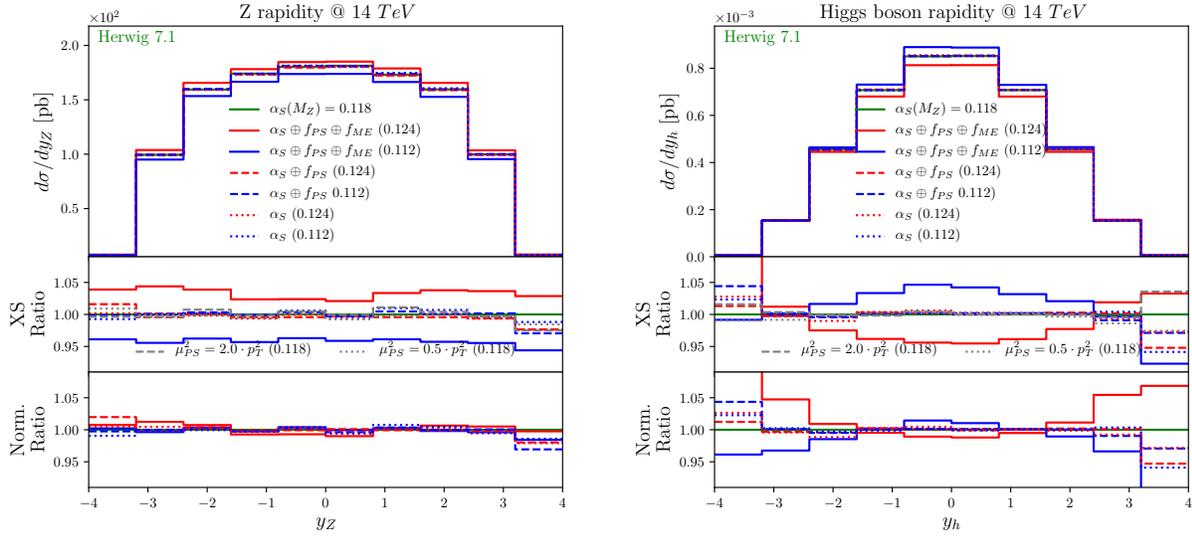


Fig. VI.10: Rapidity distributions for H-boson and Drell-Yan lepton pair production at leading-order, for 14 TeV LHC.

```
read Matchbox/PQCDLevel.in;
```

If desired we will further, as described in Table VI.1, correlate PDF set and  $\alpha_s$ -values by using  $\alpha_s^\uparrow(M_Z) = \alpha_s^{(c)}(\sqrt{2}M_Z) \approx 0.112$  together with CT10nlo\_as\_0112 PDF sets, and  $\alpha_s^\downarrow(M_Z) = \alpha_s^{(c)}(\frac{1}{\sqrt{2}}M_Z) \approx 0.124$  together with CT10nlo\_as\_0124 PDF sets.

### 2.3.3 Results

The effect of correlated  $\alpha_s + \text{PDF}$  variation and uncorrelated  $\alpha_s$  variation on color-singlet boson  $p_\perp$  spectra is illustrated in Fig. VI.9. The top ratio plots give the ratio to the baseline configuration on cross section level, while the lower ratios are taken by dividing the corresponding normalised distributions. For both for Higgs and Drell-Yan  $p_\perp$  spectra at the cross-section level (top ratio), the difference between correlated PS variation and uncorrelated  $\alpha_s$  variation is smaller than the difference correlated PS+ME variation and uncorrelated  $\alpha_s$  variation. The normalized differential distributions, however, show that the difference in correlated PS and PS+ME variations can be traced back to variations in the overall cross-section, which is induced by the difference in the PDFs  $f_{ME}$  used in the hard cross section calculation. It is important to note that the difference between uncorrelated  $\alpha_s$  variation and correlated PS variations remains visible at the level of normalized spectra, i.e. leads to  $p_\perp$ -dependent shape changes: The correlated PS variation leads to a squeezed band w.r.t. uncorrelated  $\alpha_s$  variation for the Drell-Yan  $p_\perp$ , and has little effect on the  $p_\perp$  of the Higgs boson.

The larger differences at the cross-section level can be understood by considering the rapidity distributions in Fig. VI.10, and the PDF changes in Fig. VI.11. As expected, the  $\alpha_s$  variation and the correlated PS variation lead to a vanishing uncertainty band<sup>11</sup>, whereas the correlated PS+ME variations yield a band. The normalization uncertainty of rapidity distributions in Fig. VI.10 is more uniform for the Drell-Yan process. At the same time, the uncertainty for Higgs production is very small at large rapidity, i.e. in the region that will most likely lead

<sup>11</sup>We do not vary the  $\alpha_s$  in the effective  $ggH$ -coupling.

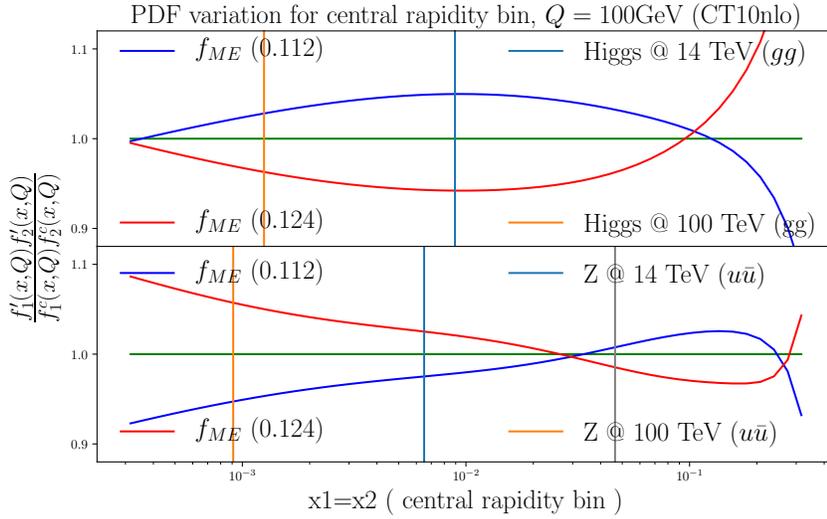


Fig. VI.11: Relative change CT10nlo\_as\_0112 and CT10nlo\_as\_0124 w.r.t. of CT10nlo\_as\_0118, as function of  $x$ .

to a significant contribution to the parton-shower radiation pattern and thus the Higgs  $p_{\perp}$  spectrum. This property then leads to smaller shape changes in the Higgs  $p_{\perp}$  spectrum, compared to the Drell-Yan  $p_{\perp}$  distribution - for Drell-Yan leptons, the correlated PS curves will uniformly be rescaled by a  $O(5\%)$  cross section differences, thus changing the apparent “shape” of the variation band, while for Higgs, the effect is smaller. Changes in the collider energy enhance this effect further, as can be seen in the second column of Fig. VI.9. In addition, for the typical  $x$ -values for different collider energies, gluon and quark PDF uncertainty behave differently. This is indicated in Fig. VI.11 for CT10nlo sets by plotting the PDF ratio

$$\frac{f'_1(x, Q) f'_2(x, Q)}{f^c_1(x, Q) f^c_2(x, Q)} \quad \text{and } Q = 100 \text{ GeV} , \quad (\text{VI.18})$$

for a configuration  $x = x_1 = x_2$  corresponding to a central rapidity bin in Fig. VI.10. The vertical lines represent the values at  $x = M_{Z/H}/\sqrt{S}$  for  $\sqrt{S} = 14$  and 100 TeV and in gray the central rapidity bin for Z-production at Tevatrons Run II energy  $\sqrt{S} = 1.96$  TeV. The spread in the PDF values for quark-initiated Z-boson production grows with collider energy, while it diminishes for gluon-initiated Higgs-boson production. This explains why the effect of PDF changes on the Z-boson  $p_{\perp}$  grow with collider energy, while they become smaller in the Higgs-boson  $p_{\perp}$  distribution. Note that, as indicated by the grey line, the uncertainty on the quark PDFs is small at Tevatron energies, since the PDF fit includes Tevatron data. Also, note that had we investigated compensation effects in Drell-Yan  $p_{\perp}$  spectra below Tevatron energies, the distributions would have painted a picture similar to what we now observe in the Higgs  $p_{\perp}$  case.

The normalized  $p_{\perp}$  distributions expose that the difference between correlated PS variation and correlated PS+ME variation is in fact minimal. The difference between correlated variation and uncorrelated  $\alpha_{s, PS}$  is most visible in the DY process, where the band shrinks. As desired, the inclusion of correlations reduces the overall shape uncertainty. No such effect is visible for Higgs. This effect can again be traced to the different PDF sets dominating the color-singlet boson production cross sections.

We now move on to the question if the effect in normalized distributions is robust to changing the observable, and when taking non-perturbative and soft-physics phenomena into consideration. The average number of jets defined in Sec. 2.3.1, binned in rapidity provides

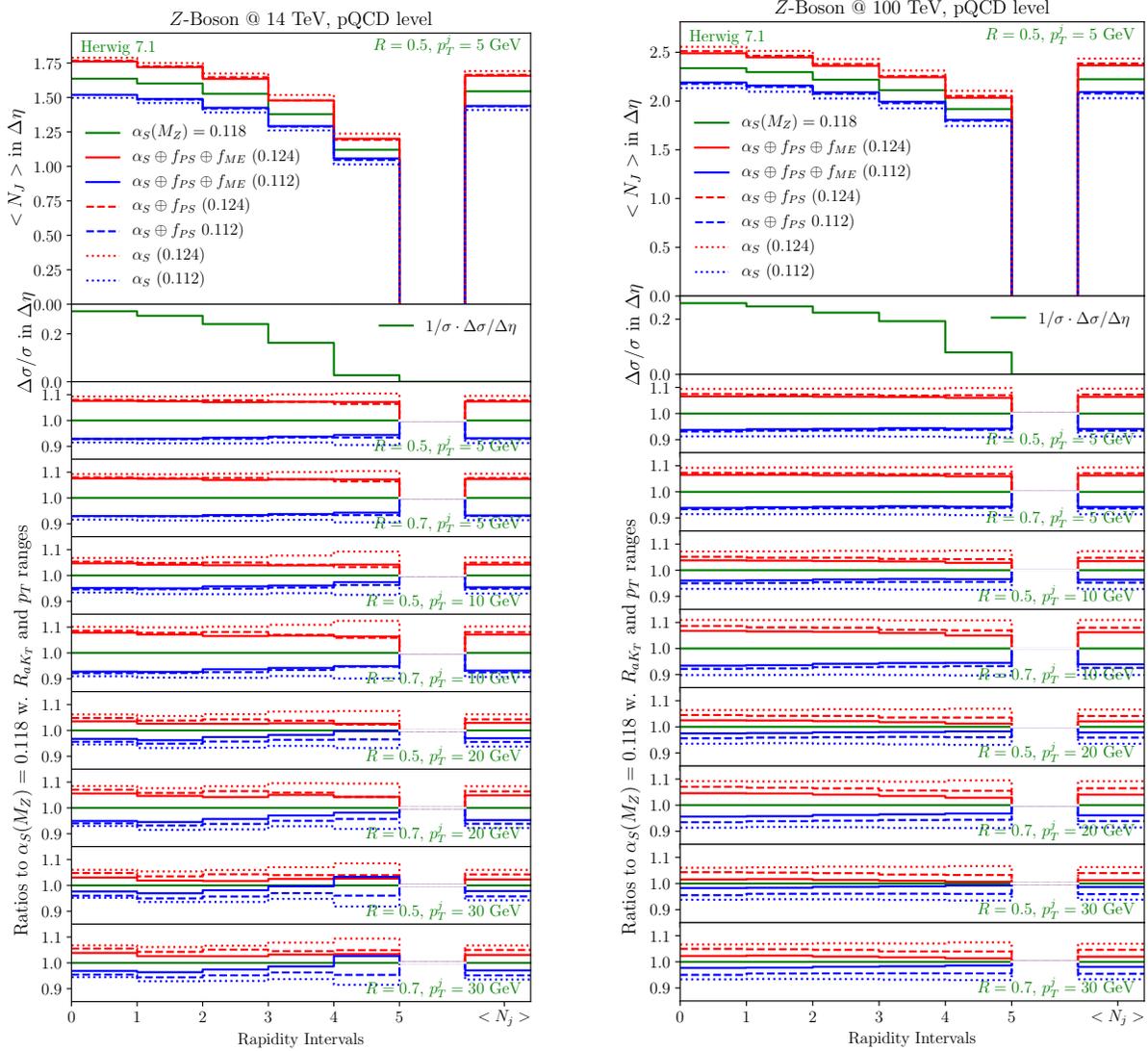


Fig. VI.12: Average number of jets in intervals of Z boson rapidity  $\Delta\eta$ , left for 14 TeV collider energy, right for 100 TeV. The second plot shows the relative size of the cross section in the intervals. The dependences of jet definition is illustrated as ratio plots w.r.t. the central prediction. The red and blue lines show the prediction for modified  $\alpha_S(M_Z)$  values,  $\alpha_S(M_Z) = 0.124$  and  $\alpha_S(M_Z) = 0.112$  respectively. For dotted lines we exclusively change the value of  $\alpha_S(M_Z)$  used in the showering process. To receive the dashed lines, also the pdf set used in the showering process was fitted with the corresponding  $\alpha_S(M_Z)$  value. The solid lines correspond to a full correlation of  $\alpha_S(M_Z)$  values in shower and the pdf sets used in hard cross section and shower calculation.

a complementary observable that is both very sensitive to the parton-shower evolution and to multiple parton interactions (MPI), while at the same time insensitive to the overall normalization.

In Figs. VI.12 and VI.13 we show the average number of jets for Z- and Higgs-boson production as a function of the boson rapidity. The first ratio plots (green line) is the relative cross section contribution in the rapidity intervals. The multiple ratios show the result when using different definitions of anti- $K_T$  jets:  $R_{aK_T} \in [0.5, 0.7]$  and  $p_T^j \in [5, 10, 20, 30]$  GeV. The last

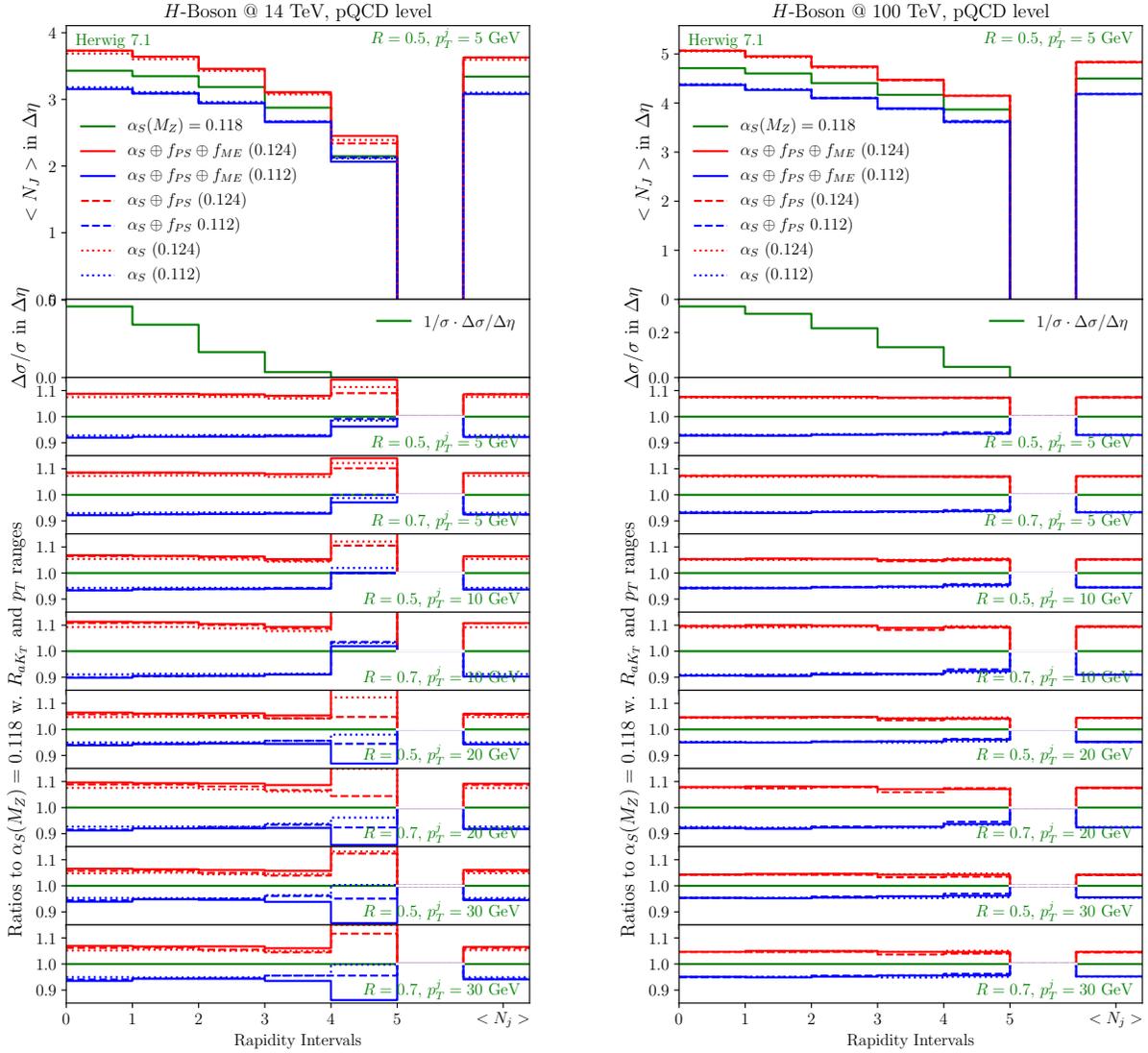


Fig. VI.13: Same as Fig. VI.12 here for Higgs boson production.

bin in the distribution represents the inclusive average number of jets. Figure VI.14 highlights the effect of including multiple parton interactions.

Compared to the transverse momenta this observable does not show the usual effect of crossing error bands due to parton shower unitarity. We further observe, depending on the jet definition, a drastic reduction of variation when using fully correlated variations in Figs. VI.12 and VI.14. In gluon dominated processes like Higgs and dijet production, this effect is not visible. The comparison between Figs. VI.12 and VI.14 also shows an interesting reduction of the variation bands when including the effects of hadronisation and multi parton interaction (MPI). Although the difference between correlated and uncorrelated variation is hardly altered by including MPI it is worth noting that if the shower produces fewer emissions because of a reduced strong coupling, the MPI is more active and adds more partons, since more of the hadron energy/momentum is left after evolution. Similar effects are expected with the interleaved MPI employed by Pythia.

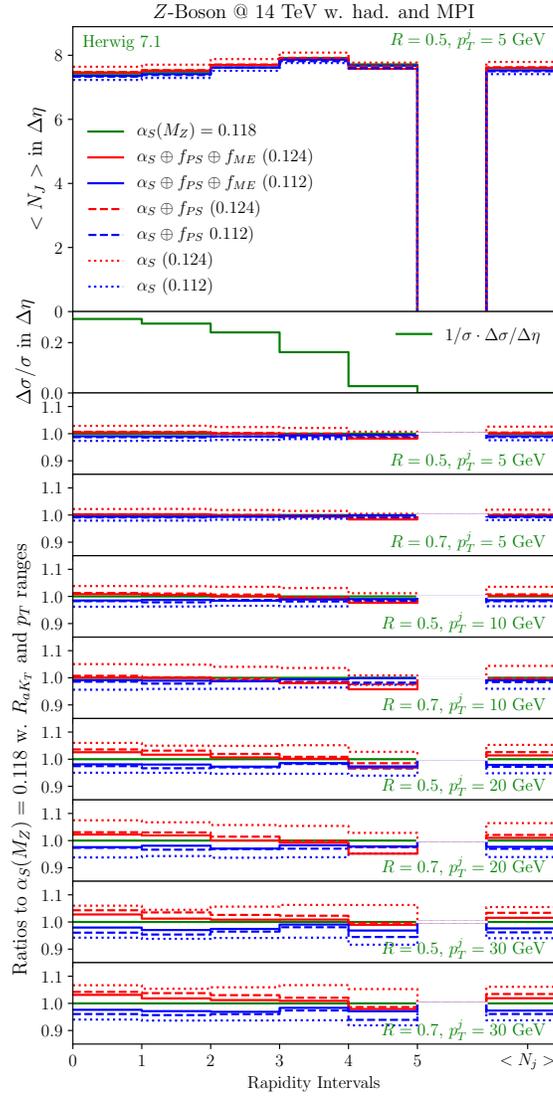


Fig. VI.14: As Fig. VI.12. Here including hadronisation effects as well as multiple parton interactions. Compared to Fig. VI.12 the average number of jets is less dependent on the value of the strong coupling constant. This is especially visible for low  $p_T$  jet definitions. This reduction of the error band is explained by the available energy left after hard process calculation and parton shower evolution. While the enhanced  $\alpha_S(M_Z)$  radiates more and therefore leaves less energy for additional MPI interactions, reducing the value of  $\alpha_S(M_Z)$  allows more additional MPI jets. This directly moderates the effect of value variation.

### 2.3.4 Discussion

Before concluding, let us summarise and discuss the findings of Sec. 2.3.3. We find that unnormalized distributions suggest large differences between  $\alpha_s$  variations and correlated PS+ME variations. These are directly linked to scaling with different PDFs used in the hard-scattering calculation. Since such uncertainties are beyond the scope of a (cross-section preserving) parton shower, it is prudent to base conclusions on normalized distributions. In this case, the effect of correlations is clearer. Including correlations by aligning the PDF with the  $\alpha_s$  used in the parton shower shrinks the envelope of variations. This effect is moderate for the Drell-Yan pair  $p_\perp$  distribution at a 14 TeV LHC, and small for the Higgs-boson  $p_\perp$  distribution at the same

setup. This might suggest that it is not necessary to respect the correlation for a 14 TeV LHC. However, aligning PDFs with  $\alpha_s$  variations can lead to more pronounced effects for observables that are sensitive to more details of the shower evolution. For the average number of jets, included correlations lead to larger differences between variations and to a more pronounced uncertainty reduction. Since multiple parton interactions are sensitive to the proton content *after* showering, the effect remains upon including MPI. In fact, the difference between correlated and uncorrelated variations is enhanced and develops intricate phase-space dependences. This is particularly visible for uncorrelated  $\alpha_S$  variations, which, after including MPI, exhibit a dependence on the  $p_{T,min}$  cut used to define jets. This is because larger shower variations lead to larger variations in MPI activity. The situation is improved when performing correlated PS variations. Thus, we conclude that for normalized distributions, it is reasonable to align the PDF and the  $\alpha_s$  when performing shower variations.

## 2.4 Summary

In this contribution, we have tried to illustrate and assess the cross-talk of perturbative parton-shower uncertainties and non-perturbative modelling in the context of MCEGs. We have argued that we should regard an MCEG as a tool that transfers our most detailed and well-constrained knowledge of the theory and previous data to new measurements. Adhering to this philosophy means that perturbative and non-perturbative MCEG aspects should be considered correlated. We have investigated this effect from two angles.

In the first part, we have argued that perturbative variations that clearly do not describe known data should be complemented with an adjustment of non-perturbative parameters to achieve a sensible description of known data. This can be achieved by retuning the MCEG’s description of hadronization-sensitive data for each parton-shower variation. We have compared the results of the envelope of such tunes with naive parton-shower scale variations around a “central” tune, for both HERWIG and PYTHIA. We find that, as expected, the retuned results exhibit a smaller variation in phase-space regions dominated by non-perturbative effects. The naive “central” variation bands of HERWIG and PYTHIA are rather different, with HERWIG being conservative-leaning and PYTHIA being more optimistic about the size of the variation band. The size of the retuned uncertainties is comparable between the two simulations. This convergence of two very different models suggests that comparing retuned variations might provide a better assessment of MCEG uncertainties in the future.

The second part of the study was motivated by cross-talk between PDF sets and parton-shower variations due to initial-state parton-shower backward evolution. In particular, we argued that the assumption of a faithful implementation of PDF evolution suggested that  $\alpha_s$  variations in the parton shower should be accompanied changing to a corresponding PDF set. We find, depending on the process, large changes in  $p_\perp$  distributions when including the changes in normalization due to changing PDF sets. These effects are much reduced when discussing normalized distributions. In these, as expected, using correlated PS variations leads, in some cases, to a reduced variation band. The impact is of order 5% for the Drell-Yan  $p_\perp$  spectrum at 14 TeV collider energy. Larger effects can be seen in the average number of jets in boson rapidity intervals of  $\Delta y = 1$ . In this observable, the effect of correlated PS variations is of 5% for the Drell-Yan process without the inclusion of multiple parton interactions, and more pronounced when including the latter. Multiple parton interactions enhance the impact of correlations, since the probability for MPI is very sensitive to parton-shower variations. Thus, we are led to conclude that it is prudent to change the parton-shower PDF consistently when varying  $\alpha_s$  in the parton shower.

## Acknowledgements

We thank the organizers for an inspiring workshop. This manuscript has been authored by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics. This project has also received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme, grant agreement No 668679 and the MCnetITN3 H2020 Marie Curie Initial Training Network, contract 722104. AHH and DS acknowledge partial support by the FWF Austrian Science Fund under the Doctoral Program No. W1252-N27 and the Project No. P28535-N27. AHH and SP also acknowledge support by the COST action CA16201 PARTICLEFACE. AS acknowledges support from the National Science Centre, Poland Grant No. 2016/23/D/ST2/02605 and COST Action CA15213 THOR. LL acknowledges support from the Swedish Research Council, contracts numbers 2016-03291 and 2016-05996.

## 2.A Appendix

### 2.A.1 Generator tuning vs. parton-shower variations: Weights

For the tuning to LEP data we used the following weights in the PROFESSOR tune, for both HERWIG and PYTHIA.

```
A_2004 := ALEPH_2004_S5765862
L_2004 := L3_2004_I652683
```

```
/A_2004/d01-x01-y01 1 # Charged multiplicity
/A_2004/d102-x01-y01 1 # Thrust minor
/A_2004/d110-x01-y01 1 # Jet mass difference
/A_2004/d118-x01-y01 1 # Aplanarity
/A_2004/d133-x01-y01 1 # Oblateness
/A_2004/d141-x01-y01 1 # Sphericity
/A_2004/d149-x01-y01 1 # Durham jet 2->1
/A_2004/d157-x01-y01 1 # Durham jet 3->2
/A_2004/d165-x01-y01 1 # Durham jet 4->3
/A_2004/d187-x01-y01 1 # 1-jet fraction
/A_2004/d195-x01-y01 1 # 2-jet fraction
/A_2004/d203-x01-y01 1 # 3-jet fraction
/A_2004/d211-x01-y01 1 # 4-jet fraction
/A_2004/d54-x01-y01 1 # Thrust
/A_2004/d62-x01-y01 1 # Heavy jet mass
/A_2004/d70-x01-y01 1 # Total jet broadening
/A_2004/d78-x01-y01 1 # Wide jet broadening
/A_2004/d86-x01-y01 1 # C-parameter
/A_2004/d94-x01-y01 1 # Thrust major
```

For PYTHIA we also did a separate tune where also the charged multiplicity distribution from L3 [1015] was included.

```
/L_2004/d59-x01-y01 1 # Charged multiplicity
```

## 3 Determination of nonperturbative correction factors and their dependence on Monte Carlo modeling<sup>12</sup>

In this contribution we discuss the influence of Monte Carlo modelling on the extraction and calculation of nonperturbative correction factors. In particular we address the question whether the choice of PDF sets and Monte Carlo models in the extraction of nonperturbative correction

---

<sup>12</sup> J. Bellm, S. Kuttimalai

factors may bias measurements in which the correction factors are used, e.g. PDF measurements at the LHC. We determine the non perturbative factors for various settings and explain the outcome in detail.

### 3.1 Introduction and Motivation

With an increasing amount of data being recorded and getting analyzed by the LHC experiments, it becomes possible to perform measurements more and more differentially in the event kinematics. This allows for the extraction of more information that can be used for a more precise determinations of key quantities, like the parton distribution functions (PDFs) of the proton.

At the same time, recent theoretical progress has lead to an increase in precision of fixed-order calculations [2, 129] which can, in many cases, be carried out at the next-to-next-to leading order in the strong coupling constant. Using such high-precision calculations and more differential observables for PDF fits requires the application of corrections that capture nonperturbative (NP) effects. They are not included in fixed-order calculations but their impact can be comparable to the remaining fixed-order uncertainty.

Modern Monte Carlo event generators (MCEGs) typically operate only at the leading order (LO) or the next-to-leading order (NLO) but supplement fixed-order calculations with an all-order treatment of soft and collinear QCD radiation through parton shower (PS) simulations and nonperturbative effects through models for hadronization and multiple parton interactions (MPI). While hadronization models account for the nonperturbative transition of QCD partons to observable colorless hadrons, MPI models account for the effect of interactions between the beam remnants which generally lead to an increase of softer jet activity in the events.

In order to correct high-precision fixed-order calculations for NP effects correction factors are often extracted MCEGs by performing separate runs with both hadronization and MPI disabled and enabled. The ratios of the corresponding predictions are then applied to fixed-order calculation as differential correction factors of the form

$$K_{\text{NP}}^{\mathcal{O}} = \frac{d\sigma^{\text{PS+HAD+MPI}}/d\mathcal{O}}{d\sigma^{\text{PS}}/d\mathcal{O}}, \quad (\text{VI.19})$$

where  $d\sigma^{\text{PS+HAD+MPI}}/d\mathcal{O}$  is the MCEG prediction differential in observable  $\mathcal{O}$  with parton showering, hadronization, and MPI [705, 1019] enabled and  $d\sigma^{\text{PS}}/d\mathcal{O}$  refers to the MCEG prediction with hadronization and MPI disabled. Here  $\mathcal{O}$  can be multi differential.

It can be shown that the NP corrections, defined in Eq. (VI.19), can be minimized by a smart choice of the jet radius in the observable definition. This is due to the hadronization correction and the MPI corrections exhibiting different scaling behaviors with  $R$  [1020]. While decreasing the jet radius  $R$  leads to larger hadronization corrections, it decreases corrections due to MPI. This can be intuitively understood since a large radius increases the chance of the jet receiving contributions from uncorrelated MPI interactions. A very small jet radius, on the other hand, leads to energy getting emitted outside of the jet cone in the transition from partons to hadrons, thereby increasing the impact of hadronization corrections. An optimal cone size can thus be chosen in order to minimize the sum MPI and hadronization effects.

When extracting PDFs from data it is crucial to ensure that the NP corrections discussed above do not introduce any bias. In particular, the NP correction factors should be independent of the PDFs and the MCEGs that were used to calculate them. In this note we study the effect of different PDFs, MCEGs, and the perturbative order at which the MCEGs operate (LO or NLO) on the the extracted NP factors and on the jet radius that minimizes the NP corrections. In order to do so we choose the NNLO PDF4LHC recommendation and the LO pdf

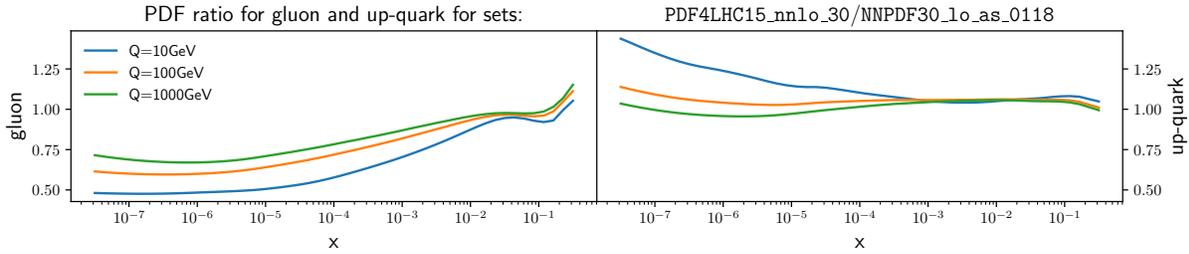


Fig. VI.15: Gluon and up-quark pdf ratio for the NNLO PDF set PDF4LHC15\_nnlo\_30 [597] w.r.t. the leading order set NNPDF30\_lo\_as\_0118 [422] this the same  $\alpha_S(M_Z) = 0.118$  value.

set from NNPDF30\_lo\_as\_0118 as the ratio of gluon and quark pdfs can be large, see Fig. VI.15. Despite the warning in [1021] and the discussion in [1022] we explicitly choose very diverse gluon distributions to estimate an upper bound on the effect.

In [1023] it was shown that parton shower evolution is usually less effected by PDF variations and more by modifications of the strong coupling, for example. More parton shower uncertainties have been studied in [501, 960]. Here we concentrate on the modification of the NP part of the simulation.

### 3.2 Analysis and Observable

As a benchmark observable we use the inclusive jet transverse momentum  $p_T$  at the LHC. It was measured by the CMS collaboration for two different jet radii  $R = 0.4$  and  $R = 0.7$  in several bins of jet rapidity [977]. For this purpose of the presented study we extended the corresponding validated Rivet [500] analysis to include more radii in the range  $R \in [0.1, 0.2 \dots 1.1]$ . In this analysis jets are defined using the anti- $k_T$  [661] jet algorithm. We perform our analysis for the LHC operating at a center-of-mass energy of  $\sqrt{s} = 13$  TeV.

### 3.3 Simulation Setup

For the extraction of NP correction factors we use the two general purpose MCEGs Herwig7 [584, 700] and Sherpa [415]. For parton showering we use the  $\bar{Q}$ -shower in Herwig7 [702] and the CS dipole shower [491]. We calculate the hard scattering process of inclusive jet production at LO or at NLO and use adaptations of the MC@NLO parton shower matching scheme [503] for consistent combination with the parton showers. In the calculation of the hard scattering we set the factorization and renormalization scales to  $\mu_f = \mu_r = H_T/2$ , where  $H_T$  is the scalar sum of all parton transverse momenta in the event. The hadronization and MPI models implemented in Sherpa are detailed in [415]. Herwig7's models are documented in [825]. For the study of the impact of PDFs we use the PDF4LHC15\_nnlo\_30 [597] and the NNPDF30\_lo\_as\_0118 [422] sets as implemented in LHAPDF [964]. We replace all PDFs sets used in the modelling (hard cross section, parton shower, MPI and proton remnant handling) and all instances of strong coupling constants to the corresponding  $\alpha_S(M_Z) = 0.118$  value.

We would like to stress that our generator settings are not tuned to data but rather aligned across both MCEGs for a comparison. The focus of this study is not the nominal NP correction factor but rather it's dependence on the parameters in the extraction.

### 3.4 Results

We start our discussion with results for the NP corrections for the differential inclusive jet transverse momentum spectrum in bins of pseudorapidity  $y$  as obtained from MCEG runs using

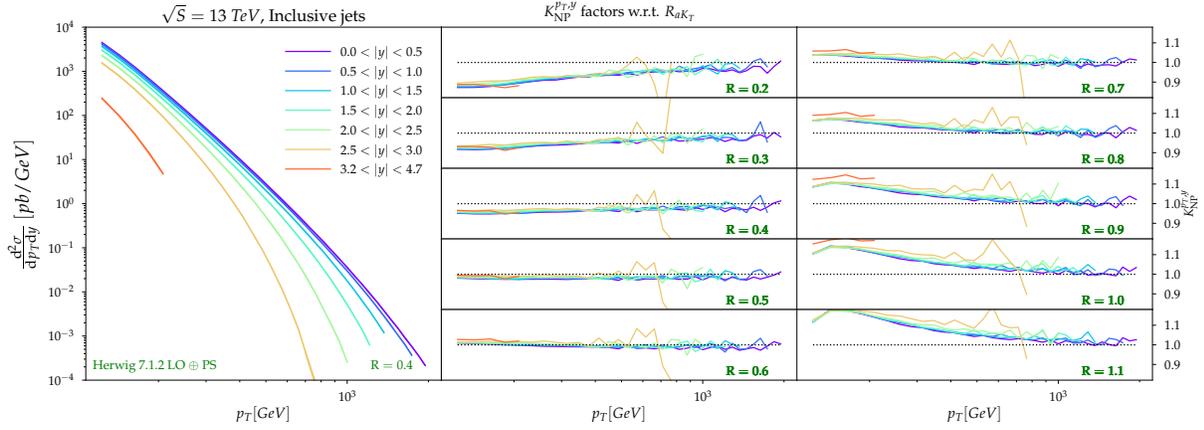


Fig. VI.16: Double differential cross section in transverse momentum and in intervals of rapidity. Left panel shows the differential cross section. The ratio plots in the center and to the right show the non perturbative factor for various jet opening angles  $R_{antiK_T}$ . While for small radii hadronisation effects tend to remove energy from thin jets, for large radii the jets tend to collect more MPI jets. Therefore the former get softer the latter harder. For broad jet definitions ( $R > 0.8$ ) even the generation cut is visible as the MPI addition allows to shift the cross section drop off into the analysis region.

LO matrix elements for the hard scattering process calculation. As shown in Fig. VI.16 we observe large negative NP corrections of the order of 10% in the region of small  $p_T$  when choosing a small jet radius of  $R \approx 0.1$ . The origin of these corrections is the energy that gets radiated out of the jet in the process of hadronization, as discussed above. At large values of  $R$ , we see large positive corrections of up to 20% due to additional MPI radiation entering the jets. At values of  $R \approx 0.5$  the NP corrections are minimized to the 1% level. We note that the dependence of NP effects on the pseudorapidity is very small and that NP effects generally vanish at very large transverse momenta of the order of 1 TeV.

Having shown the qualitative features of the NP corrections themselves, we now turn to the discussion of their dependence on the MCEG parameters used for the extraction. For this purpose, we show the ratio of NP factors extracted with Herwig7 and Sherpa at leading order in Fig. VI.17. The observed differences are below the 5% level and most pronounced at small jet radii. This suggests that the source of the small discrepancies are mainly differences in the hadronization and parton shower models. In addition we note that there are differences at large  $R > 0.8$  and very small  $p_T$  close to the analysis jet cut. We were able to trace these discrepancies back to different event generation cuts on the leading jet transverse momentum, which was set to 90 GeV and 100 GeV for Sherpa and Herwig7, respectively. This effect is illustrated in Fig. VI.18, where the event generation cuts are clearly visible as features in the NP factor. For larger jet radii this effect gets washed out, thus extending it into the region shown in Fig. VI.17.

In Fig. VI.19 we show the impact of using a different PDF set. Hadronization corrections can be expected to be independent of the PDF, as the hadronization models act only on final state of partons after the full evolution through parton showers has terminated. Correspondingly, we see barely any dependence on the PDF at small  $R$  where hadronization corrections dominate. At large  $R$ , we see larger effects. This is to be expected since the MPI modelling, which mostly affects large jet radii, crucially depends on the PDFs. We note, however, that MPI models are typically tuned to data based on a specific PDF set. It is thus very likely that the observed differences would have been much smaller, had we re-tuned the MPI models to the respective PDFs. In that sense, the MPI models in MCEGs are to be understood as a 'postdiction' or

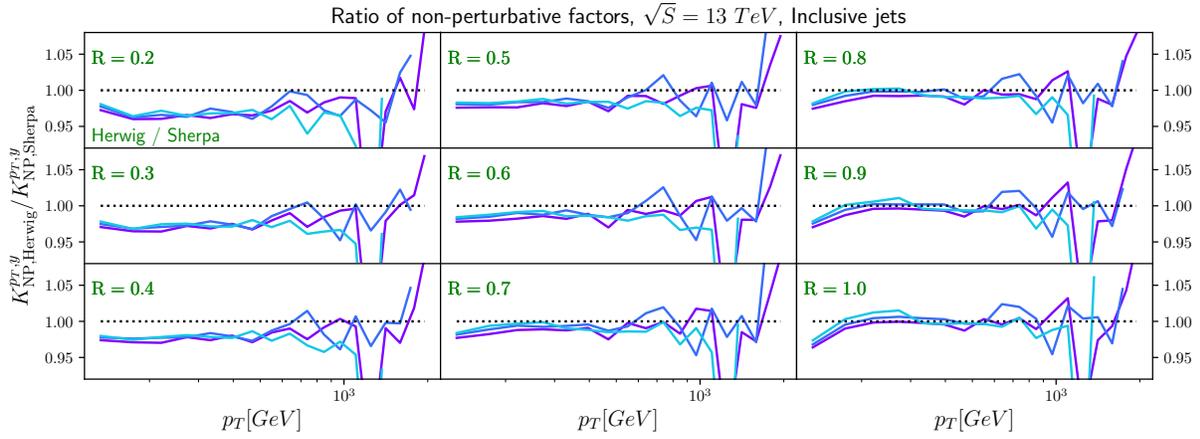


Fig. VI.17: Ratio of NP factors calculated using Herwig7 and Sherpa. With the setup in use the MPI modification area of broad jets is less modified while the hadronisation region is altered by the different shower and hadronisation models. We stress again that the findings are setup dependent and effects might shrink with tuned values.

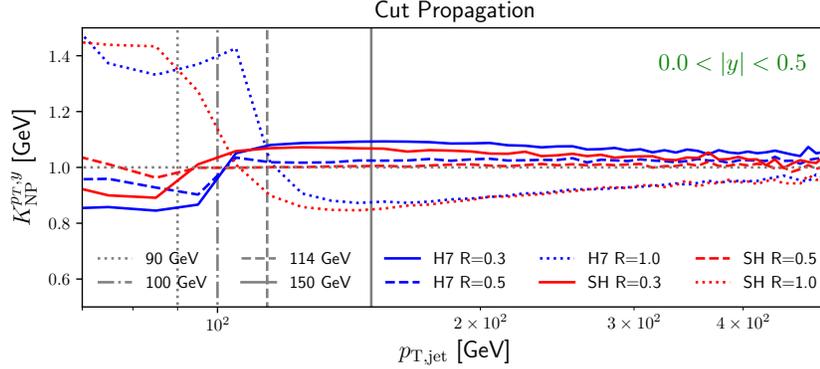


Fig. VI.18: Effect of event generation cuts on the NP factors. The cut on the leading jet transverse momentum was set to 100 GeV and 90 GeV for Herwig7 and Sherpa, respectively

parametrization of MPI effects as measured in data.

In order to show how the optimal choice of  $R$  is affected by the variables considered in this study, we quantify the overall size of NP corrections by the following measures

$$\mathcal{K}(R) = \int \left| \left( K_{\text{NP}}^{p_T, y}(R) - 1 \right) \right| \theta(p_T^{\text{max}} - p_T) dp_T dy. \quad (\text{VI.20})$$

Note that we do not apply a weight proportional to the double differential cross section  $d\sigma/dp_T dy$  in the definition above since we would like to minimize the NP corrections equally well in the entire phase space. As shown in Fig. VI.20, the value of  $R$  that minimizes the NP corrections we find<sup>13</sup> is  $R \approx 0.6$  for Herwig and  $R \approx 0.5$  for Sherpa, and it is independent of PDF or perturbative order within  $\Delta R = 0.05$ .

### 3.5 Conclusions

In this contribution, we investigated the effect of PDF set choice and simulation setup used to calculate nonperturbative correction factors. In order to do so, we chose strongly diverse PDFs

<sup>13</sup>Note that these can change with tuned values.

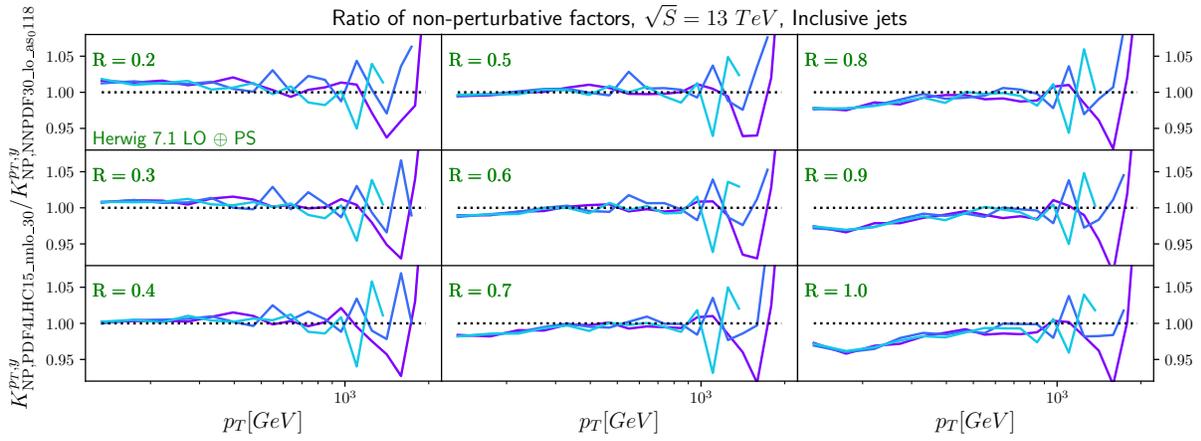


Fig. VI.19: Ratio of NP factors calculated using two different PDF sets. The effect of removing energy from the jet due to emissions outside of the small cone in the hadronisation process is hardly effected by changing the PDF set. The differences for higher R values can be explained by modification of the MPI addition with modified PDFs. The strongly reduced NNLO gluon pdf at small x values allows less MPI activity and reduced the additional energy in the broad jets. Therefore  $K_{\text{NP}}^{PT,y}(R)$  is reduced for large R values.

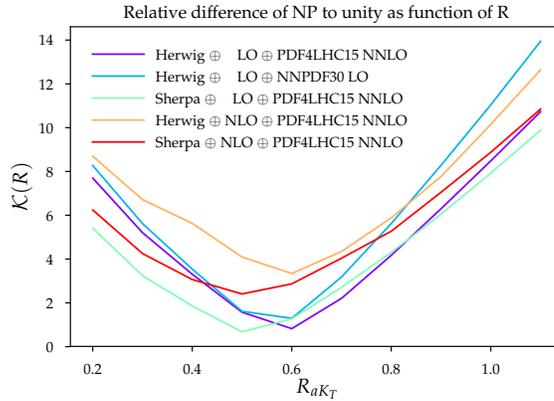


Fig. VI.20: For the simulation setup as described in Sec. 3.3 we calculate the measure according to Eq. (VI.20). For Herwig we get an optimal value for  $R \approx 0.6$ , for Sherpa slightly below. The pdf influence on the best R value is mild and matching to higher order matrix elements is even smaller. Note: The used setup is not representative for the tuned default setup. Therefore the absolute value of the optimal point might shift with standard parameters. Here we are interested in the shift due to PDF or higher perturbative accuracy. The height of the distribution is driven by MC statistics.

and calculated for two different shower and hadronisation models as well as LO and NLO plus parton shower simulations the NP factors. As expected we confirmed in this study that these factors are only mildly affected by the choice of model parameters or the perturbative order. The observed differences in the extracted NP factors are of the order of only 5% relative to the absolute value of the NP correction themselves. We found that jet radii of  $R \approx 0.5-0.6$  minimize the NP corrections to the level of only a few percent and the phenomenological impact of the differences in the extracted NP factors should therefore be very small. In addition we illustrated that event generation matrix element cuts can have an impact on NP correction factors if the opening angle of the jets allows the collection of much MPI activity.

## Acknowledgements

We would like to thank Joey Huston and Stefan Prestel for useful discussions and helpful comments. This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme, grant agreement No 668679. Furthermore, it was supported in part by the U.S. Department of Energy under contract number DE-AC02-76SF00515.

## 4 Is there room for improvement in the description of the gluon splitting to heavy quarks? <sup>14</sup>

We perform a comparison of three different Monte Carlo generators for a process involving  $B$ -hadrons in the final state, with the aim of assessing to which extent NLO tools, based on a 4-flavour scheme, are successful in simulating kinematic configurations sensitive to the gluon splitting to pair of (massive)  $b$ -quarks.

### 4.1 Introduction

The Monte Carlo modelling of processes featuring an heavy-quark pair in the final state, possibly in association with an heavy color-singlet, has been, for several years, one of the open issues for Monte Carlo developers.

Due to their large mass, top quarks are always treated differently with respect to the other QCD partons, i.e. they are always simulated using exact matrix elements, and, nowadays, whenever possible, this is done at NLO. Moreover, because of their large mass, it’s nearly impossible for a parton shower algorithm to produce a top-quark pair from a gluon splitting, hence, even in (multijet-merged) NLO+PS accurate predictions, the kinematics of top quarks is, to a very large extent, dominated by the matrix-element description.

The situation is very different for bottom (and charm) quarks, and, in fact, one of the open issues related to the treatment of heavy quarks in event generators has to do with processes where  $b$  quarks are present. As detailed below, in this context one of the main problem is the modelling of the  $g \rightarrow Q\bar{Q}$  splitting, especially when  $b$ -quarks are produced at small opening angles. It turns out that this region is relevant in several LHC analysis, not only when precise SM measurements are performed and compared with theory [1024–1028], but also when simulation tools become necessary to simulate Higgs production (*e.g.*  $b\bar{b} \rightarrow H$  [1029,1030]) or backgrounds with several  $b$ -jets. An important example belonging to the latter category is the modelling of QCD production of  $b\bar{b}$  pairs in association with vector bosons, which is one of the main backgrounds for measuring associated Higgs-boson production with  $H \rightarrow b\bar{b}$  decay. In this case in fact, complications arise from double  $b$ -tags in a single jets, which come from a collinear  $g \rightarrow b\bar{b}$  which is only seen as one jet. Similarly, it is known [913,917] that a proper description of the  $g \rightarrow b\bar{b}$  splitting (with mass effects) is very important for the simulation of  $t\bar{t} + b$  jets production, an irreducible background for measurements of  $t\bar{t}H$  production in the  $H \rightarrow b\bar{b}$  channel.

In this short contribution we want to have a critical look at how well the production of final states involving  $b$ -quarks is described, by means of a comparison among different generators, for the production process  $pp \rightarrow b\bar{b}$ .

### 4.2 Schematic outline of the problem

Broadly speaking, the problem of simulating processes with  $b$  quarks can be formulated as follows: as the mass of the  $b$  quark ( $m_b$ ) is relatively small, a  $b$ -quark is a quasi-massless particle. This means that, a priori, in the limit  $m_b \rightarrow 0$ , a parton shower algorithm is expected to simulate

---

<sup>14</sup> J. A. McFayden, D. Napoletano, E. Re

fairly well the kinematic regions dominated by the splitting  $g \rightarrow b\bar{b}$ . One should also notice that, even in the limit  $m_b \rightarrow 0$ , the splitting  $g \rightarrow b\bar{b}$  does not contain soft singularities, i.e. only a single collinear logarithm is present (the splitting probability  $dP_{g \rightarrow q\bar{q}}(q^2, z, \phi)$  is proportional to  $dq^2/q^2$ , with  $q^2$  being the invariant mass, or the transverse momentum of the splitting).

When one needs to keep  $m_b$  finite, it is natural to use a four-flavour (4F) scheme, that is, a scheme where there are no PDF for  $b$ -quarks (and hence no hard processes initiated by  $b$ -quarks). By using these matrix elements, power suppressed effects are, by definition, included, at the order of the matrix-element used. For phenomenological reasons, one also allows off-shell gluons, in a time-like shower, to split into massive  $b$ -pairs. This last step is intrinsically ambiguous, because the splitting is now, strictly speaking, non-singular, whereas the basic principle underlying parton shower algorithms is deeply connected to the logarithmic structure of QCD splittings. These ambiguities are related to the treatment of the kinematics for a quasi-collinear splitting, as well as to the choice of the renormalization scale  $\mu$  at which the strong coupling is evaluated. In modern parton showers, for massless splittings, the latter is chosen to be the transverse momentum of the splitting itself, as this choice guarantees that (known) subleading logarithmic terms are dynamically taken into account by the algorithm. As soon as the splitting contains massive quarks, several constraints based on theoretical arguments are lifted, leaving room for several a-priori legitimate choices [491, 584, 1031].

Despite the significant progress made in the last years to study and improve the simulation of processes with  $b$ -quarks (or  $b$ -jets) in the final state [586, 1029, 1030, 1032, 1033], it's not yet clear what is the best strategy to treat the  $g \rightarrow b\bar{b}$  splitting in final-state showers. For instance, the results presented in Ref. [1034] or refs. [1026, 1028] show that the data/theory agreement for observables sensitive to the kinematics of  $g \rightarrow b\bar{b}$  splitting is not always satisfactory, especially when the splitting is quasi-collinear. It's also worth noticing that this conclusion often applies not only to simulations based on LO matrix elements, but also on NLO ones. In short, it is fair to say that a better understanding of the aforementioned splittings is still needed.

In this contribution, we decided to make a small step forward, by performing a comparison, as fair as possible, among different generators and by using different observables. In particular, for the first time we will compare 4-flavour NLO predictions, matched to different parton showers, for observables sensitive to the underlying process  $pp \rightarrow b\bar{b}$  (+ jets) at the LHC. Data from a recent ATLAS measurements will also be included. In the following sections, we describe the tools used and the analysis considered, and comment on our findings. As we will observe, also for NLO-accurate tools, differences are observed mostly when the dominant kinematics is collinear  $g \rightarrow b\bar{b}$  splitting, calling therefore for further investigations.

## 4.3 Settings

### 4.3.1 Measurements and analysis

Monte Carlo predictions are compared to a recent ATLAS measurement of the production of  $B$ -hadron pairs in proton-proton collisions at  $\sqrt{s} = 8$  TeV [1034]. The measurement considers the  $B(\rightarrow J/\psi[\rightarrow \mu\mu] + X)B(\rightarrow \mu + X)$  decay mode where normalised differential cross sections are measured for observables designed to be sensitive to the different  $B$ -hadron production modes. For example,  $B$ -hadron production is measured as a function of  $\Delta\phi$  and  $\Delta R = \sqrt{\Delta\eta^2 - \Delta\phi^2}$ . This includes measurement down to  $\Delta R = 0$ , a region sensitive to gluon-splitting to  $b\bar{b}$ . As described above, these measurements can help developing more effective phenomenological models for the  $g \rightarrow Q\bar{Q}$  splittings.

These are the parameters that we will focus on for the remainder of this study. The predictions described below are analysed at  $B$ -hadron-level and the results are translated to the 3-muon-level via transfer functions provided by in Ref. [1034].

A similar study was carried out in the context of [1], where many set-ups and generator where compared against one another, and data, in the case of the associated production of a vector boson  $Z, W$  with  $b$ -quarks. However in that case the most relevant effects were coming from space-like (or initial state) splittings.

### 4.3.2 Configuration of event generators used

In the following we describe the generator setups used throughout. As we are mainly interested in the region where the two  $b$ -flavoured objects are not widely separated, we expect kinematical mass effects to play an important role. For this reason we chose to show only 4F scheme results. In the 4F scheme,  $b$ -quarks are decoupled from the evolution of QCD, and can only appear as, massive, final state particles as the result of a “hard”  $g \rightarrow b\bar{b}$  splitting.

SHERPA-MEPS@NLO [925, 926]:  $pp \rightarrow b\bar{b} + \text{up to } 3j$ , merging scale  $Q_{Cut} = 20$  GeV. Renormalization, factorization and shower starting scales are set according to SHERPA’s default choice. Each emission is clustered until a suitable core,  $2 \rightarrow 2$ , process is found. Scales appearing in the fixed order matrix elements are calculated as

$$\mu_R = \mu_F = \mu_Q = -\frac{1}{\sqrt{2}} \left( \frac{s+t+u}{stu} \right), \quad (\text{VI.21})$$

while for each emission from the shower the scale appearing in  $\alpha_s$ , or in the Jacobian PDF ratio, is set to be the  $k_T$  of the splitting. Instead for the  $g \rightarrow Q\bar{Q}$  splitting the scales are set equal to the invariant mass of the  $Q\bar{Q}$  pair. Jets are clustered using the anti- $k_T$  [661] algorithm with a  $p_T$  cut of 5 GeV. Finally, results are obtained using the NNPDF30 NNLO PDF [422] set with  $\alpha_s(m_Z) = 0.118$ , the default choice in SHERPA.

MADGRAPH5\_AMC@NLO [88]+Pythia8 [633]: MADGRAPH5\_AMC@NLOv2.5.5 is used to provide production of  $pp \rightarrow b\bar{b}$  events at NLO in the 4-flavour scheme with massive  $b$ -quarks that are subsequently showered by Pythia8.230. In the matrix element calculation a dynamical renormalization and factorization scale, defined as the sum of the transverse masses divided by two of all final state particles and partons, and the NNPDF30 NNLO PDF set with  $\alpha_s(m_Z) = 0.118$  are used. In Pythia8 the  $g \rightarrow b\bar{b}$  splitting scale is “Option 1” as described in Ref. [1031] which uses  $k_T$ -based scale. The A14 tune [632] is used.

HERWIG7 [584, 700, 825]: HERWIG7.1 is used to provide  $pp \rightarrow b\bar{b}$  events at NLO in the 4-flavour scheme with massive  $b$ -quarks. The MC@NLO matching scheme is used along with the angular-ordered parton shower. In the matrix element calculation the maximum jet  $p_T$  is used for the renormalization and factorization scales, and the MMHT14 PDF set [601] with  $\alpha_s(m_Z) = 0.118$  is used. The default HERWIG7 tune is used and the  $g \rightarrow b\bar{b}$  splitting scale is defined as the virtuality of the splitting.

## 4.4 Comparisons and discussion

Results described in the previous section, compared to experimental results are collected in Fig. VI.21, where we show normalized distributions for the azimuthal angular separation  $\Delta\phi = \phi_2 - \phi_1$  between the  $J/\psi$  and the third muon in the azimuthal plane (top left), and the  $\Delta R$  distance among the same two final states, both in the inclusive case (top right), and for  $p_T(J/\psi, \mu) \geq 20$  GeV (bottom). We start our discussions by noting that in all three plots two regions can be identified: a matrix element region and a parton shower one. Leaving, for a moment, the transition between the two regimes aside, it appears that while in the former a good level of agreement is shown among the three Monte Carlo predictions<sup>15</sup>, and between

<sup>15</sup>modulo large statistical uncertainties

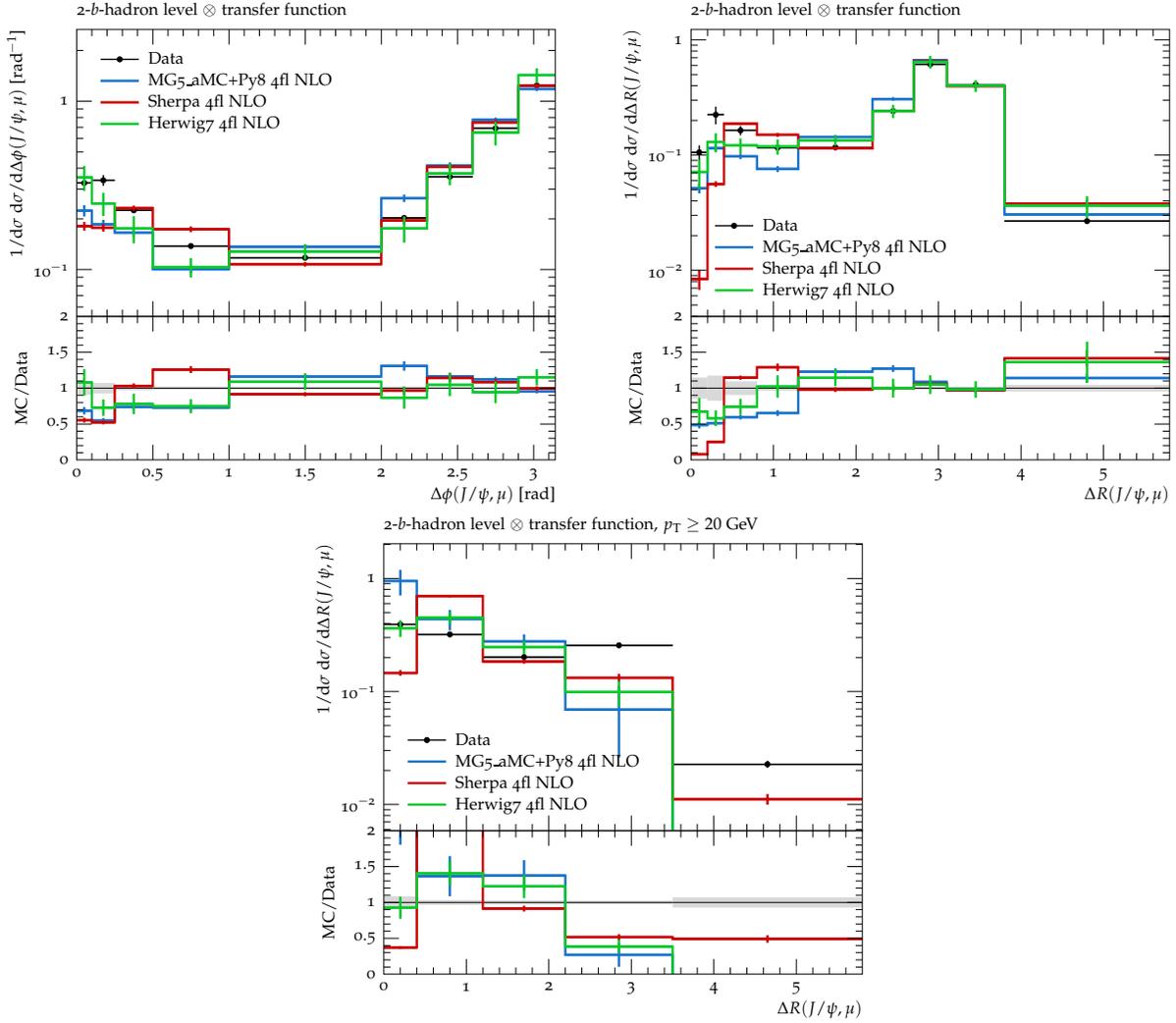


Fig. VI.21: Top-left plot:  $\Delta\phi$  distance between the  $J/\psi$  and the third muon; Top-right plot:  $\Delta R$  distance between the  $J/\psi$  and the third muon, in the fully inclusive case; Bottom plot:  $\Delta R$  distance for  $p_T(J/\psi, \mu) \geq 20$  GeV

Monte Carlo and data, the latter present visible discrepancies. The former statement is somehow obvious: apart from scale and PDF choices there are not many places in which the three generators can be much different. Note that scale/PDF variation bands are not shown, however it is to be expected that they would cover the difference in the matrix element region, *i.e.* the region  $\Delta\phi \gtrsim 0.3$ ,  $\Delta R \gtrsim 0.5$ .

On the other hand, in the shower dominated regime, one could argue that the spread in Monte Carlos is a testament of the size of theory uncertainties. To see how such large difference can come up, let us remind the reader that the splitting kernel for the  $g \rightarrow Q\bar{Q}$ , that enters the Sudakov form factor splitting can, symbolically, be written as [1031]

$$P_{g \rightarrow Q\bar{Q}} \propto \frac{\alpha_s(X)}{2\pi} \left[ P_{g \rightarrow q\bar{q}}^{(m_q=0)} + \frac{m_Q^2}{Y_{Q\bar{Q}}} \right]. \quad (\text{VI.22})$$

While the ambiguity in the exact choice of  $Y_{Q\bar{Q}}$  is a genuine subleading contribution in the shower, the actual value of  $X$ , while being formally higher order, can be quite large in specific

phase-space corners [1034], like the in the small  $\Delta\phi$  and  $\Delta R$  region.

Crucially, the three shower implementations we compare here, as discussed in the previous section, differ in the choice of both  $X$  and  $Y$ , and this is most likely the reason for the difference displayed in the plots we show here.

## 4.5 Conclusions

In this study we have observed that even 4-flavour NLO computations for  $pp \rightarrow b\bar{b}$ , matched to parton showers, don't describe very accurately the  $\Delta R$  distribution among  $B$ -hadrons, in the region that receives an important contribution from the kinematical configurations associated to a (quasi)-collinear  $g \rightarrow b\bar{b}$  splittings. Given that all the simulations considered contain the same (or very similar) perturbative accuracy at the level of hard matrix elements, it is very likely that a large part of the observed disagreement can be attributed to the parton showering part. This is also reflected by the spread of predictions obtained with different tools.

It would be highly desirable to perform further, dedicated studies to ameliorate the theoretical understanding of the  $g \rightarrow b\bar{b}$  splitting, and/or to find processes and observables that enhance the contribution of gluon splitting. In fact, having a better discrimination power among different choices for the modelling of gluon splitting into heavy quarks seems necessary in order to rule out some a-priori legitimate choices by means of a comparison with measured distributions. For instance the methods proposed in Ref. [1035] look very promising. We also note that to better understand the mechanics of  $b$ -quark pair production in regimes important for  $H \rightarrow b\bar{b}$  production, the  $p_T$  dependence of  $g \rightarrow b\bar{b}$  production would be important to measure. We hope that our study will help as a motivation and starting point for future studies on this topic.

## Acknowledgements

The work of DN has received funding from the European Union's Horizon 2020 research and innovation programme as part of the Marie Skłodowska-Curie Innovative Training Networks "MC-netITN3" (grant agreement no. 722104) and "HiggsTools" (grant agreements PITN-GA2012-316704) and by the ERC Advanced Grant MC@NNLO (340983). The work of ER is supported by a Maria Skłodowska-Curie Individual Fellowship of the European Commission's Horizon 2020 Programme under contract number 659147 PrecisionTools4LHC.

## References

- [1] J. R. Andersen et al., *Les Houches 2015: Physics at TeV Colliders Standard Model Working Group Report*, [arXiv:1605.04692 \[hep-ph\]](#). 4, 19, 49, 58, 159, 206, 213, 214, 215, 217, 218, 220, 221, 222, 223, 247
- [2] J. Currie, A. Gehrmann-De Ridder, T. Gehrmann, E. W. N. Glover, A. Huss, and J. Pires, *Precise predictions for dijet production at the LHC*, *Phys. Rev. Lett.* **119** (2017) no. 15, 152001, [arXiv:1705.10271 \[hep-ph\]](#). 4, 7, 13, 19, 20, 22, 240
- [3] T. Carli, D. Clements, A. Cooper-Sarkar, C. Gwenlan, G. P. Salam, F. Siegert, P. Starovoitov, and M. Sutton, *A posteriori inclusion of parton density functions in NLO QCD final-state calculations at hadron colliders: The APPLGRID Project*, *Eur. Phys. J.* **C66** (2010) 503–524, [arXiv:0911.2985 \[hep-ph\]](#). 4, 21, 22
- [4] T. Kluge, K. Rabbertz, and M. Wobisch, *FastNLO: Fast pQCD calculations for PDF fits*, in *Deep inelastic scattering. Proceedings, 14th International Workshop, DIS 2006, Tsukuba, Japan, April 20-24, 2006*, pp. 483–486. 2006. [arXiv:hep-ph/0609285 \[hep-ph\]](#). 4, 22
- [5] A. V. Kotikov, *Differential equations method: New technique for massive Feynman diagrams calculation*, *Phys. Lett.* **B254** (1991) 158–164. 5
- [6] T. Gehrmann and E. Remiddi, *Differential equations for two loop four point functions*, *Nucl. Phys.* **B580** (2000) 485–518, [arXiv:hep-ph/9912329 \[hep-ph\]](#). 5
- [7] J. M. Henn, *Multiloop integrals in dimensional regularization made simple*, *Phys. Rev. Lett.* **110** (2013) 251601, [arXiv:1304.1806 \[hep-th\]](#). 5
- [8] M. Prausa, *epsilon: A tool to find a canonical basis of master integrals*, *Comput. Phys. Commun.* **219** (2017) 361–376, [arXiv:1701.00725 \[hep-ph\]](#). 5
- [9] O. Gituliar and V. Magerya, *Fuchsia: a tool for reducing differential equations for Feynman master integrals to epsilon form*, *Comput. Phys. Commun.* **219** (2017) 329–338, [arXiv:1701.04269 \[hep-ph\]](#). 5
- [10] C. Meyer, *Algorithmic transformation of multi-loop master integrals to a canonical basis with CANONICA*, *Comput. Phys. Commun.* **222** (2018) 295–312, [arXiv:1705.06252 \[hep-ph\]](#). 5
- [11] M. Zeng, *Differential equations on unitarity cut surfaces*, *JHEP* **06** (2017) 121, [arXiv:1702.02355 \[hep-th\]](#). 5
- [12] M. Harley, F. Moriello, and R. M. Schabinger, *Baikov-Lee Representations Of Cut Feynman Integrals*, *JHEP* **06** (2017) 049, [arXiv:1705.03478 \[hep-ph\]](#). 5
- [13] R. N. Lee, A. V. Smirnov, and V. A. Smirnov, *Solving differential equations for Feynman integrals by expansions near singular points*, [arXiv:1709.07525 \[hep-ph\]](#). 5
- [14] J. Bosma, K. J. Larsen, and Y. Zhang, *Differential equations for loop integrals in Baikov representation*, [arXiv:1712.03760 \[hep-th\]](#). 5
- [15] J. M. Henn, *Lectures on differential equations for Feynman integrals*, *J. Phys.* **A48** (2015) 153001, [arXiv:1412.2296 \[hep-ph\]](#). 5
- [16] R. Bonciani, V. Del Duca, H. Frellesvig, J. M. Henn, F. Moriello, and V. A. Smirnov, *Two-loop planar master integrals for Higgs → 3 partons with full heavy-quark mass dependence*, *JHEP* **12** (2016) 096, [arXiv:1609.06685 \[hep-ph\]](#). 5
- [17] A. Primo and L. Tancredi, *On the maximal cut of Feynman integrals and the solution of their differential equations*, *Nucl. Phys.* **B916** (2017) 94–116, [arXiv:1610.08397 \[hep-ph\]](#). 5
- [18] M. Becchetti and R. Bonciani, *Two-Loop Master Integrals for the Planar QCD Massive Corrections to Di-photon and Di-jet Hadro-production*, *JHEP* **01** (2018) 048,

- [arXiv:1712.02537 \[hep-ph\]](#). 5
- [19] J. Davies, G. Mishima, M. Steinhauser, and D. Wellmann, *Double-Higgs boson production in the high-energy limit: planar master integrals*, [arXiv:1801.09696 \[hep-ph\]](#). 5, 13
- [20] J. M. Henn and V. A. Smirnov, *Analytic results for two-loop master integrals for Bhabha scattering I*, *JHEP* **11** (2013) 041, [arXiv:1307.4083 \[hep-th\]](#). 5
- [21] P. Mastrolia, M. Passera, A. Primo, and U. Schubert, *Master integrals for the NNLO virtual corrections to  $\mu e$  scattering in QED: the planar graphs*, *JHEP* **11** (2017) 198, [arXiv:1709.07435 \[hep-ph\]](#). 5
- [22] R. Bonciani, S. Di Vita, P. Mastrolia, and U. Schubert, *Two-Loop Master Integrals for the mixed EW-QCD virtual corrections to Drell-Yan scattering*, *JHEP* **09** (2016) 091, [arXiv:1604.08581 \[hep-ph\]](#). 5
- [23] A. von Manteuffel and R. M. Schabinger, *Numerical Multi-Loop Calculations via Finite Integrals and One-Mass EW-QCD Drell-Yan Master Integrals*, *JHEP* **04** (2017) 129, [arXiv:1701.06583 \[hep-ph\]](#). 5
- [24] J. Ablinger, J. Blümlein, A. De Freitas, A. Hasselhuhn, C. Schneider, and F. Wißbrock, *Three Loop Massive Operator Matrix Elements and Asymptotic Wilson Coefficients with Two Different Masses*, *Nucl. Phys.* **B921** (2017) 585–688, [arXiv:1705.07030 \[hep-ph\]](#). 5
- [25] J. Ablinger, J. Blümlein, A. De Freitas, C. Schneider, and K. Schönwald, *The two-mass contribution to the three-loop pure singlet operator matrix element*, *Nucl. Phys.* **B927** (2018) 339–367, [arXiv:1711.06717 \[hep-ph\]](#). 5
- [26] M. Czakon, P. Fiedler, and A. Mitov, *Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through  $O(\alpha_S^4)$* , *Phys. Rev. Lett.* **110** (2013) 252004, [arXiv:1303.6254 \[hep-ph\]](#). 5, 7, 18, 30, 89, 173
- [27] P. Bärnreuther, M. Czakon, and P. Fiedler, *Virtual amplitudes and threshold behaviour of hadronic top-quark pair-production cross sections*, *JHEP* **02** (2014) 078, [arXiv:1312.6279 \[hep-ph\]](#). 5
- [28] R. Bonciani, A. Ferroglia, T. Gehrmann, A. von Manteuffel, and C. Studerus, *Light-quark two-loop corrections to heavy-quark pair production in the gluon fusion channel*, *JHEP* **12** (2013) 038, [arXiv:1309.4450 \[hep-ph\]](#). 5
- [29] G. Abelof, A. Gehrmann-De Ridder, and I. Majer, *Top quark pair production at NNLO in the quark-antiquark channel*, *JHEP* **12** (2015) 074, [arXiv:1506.04037 \[hep-ph\]](#). 5
- [30] C. Bogner, *Generalizations of polylogarithms for Feynman integrals*, *J. Phys. Conf. Ser.* **762** (2016) no. 1, 012067, [arXiv:1603.00420 \[hep-ph\]](#). 5
- [31] L. Adams, E. Chaubey, and S. Weinzierl, *Simplifying Differential Equations for Multiscale Feynman Integrals beyond Multiple Polylogarithms*, *Phys. Rev. Lett.* **118** (2017) no. 14, 141602, [arXiv:1702.04279 \[hep-ph\]](#). 5
- [32] C. Bogner, A. Schweitzer, and S. Weinzierl, *Analytic continuation and numerical evaluation of the kite integral and the equal mass sunrise integral*, *Nucl. Phys.* **B922** (2017) 528–550, [arXiv:1705.08952 \[hep-ph\]](#). 5
- [33] E. Remiddi and L. Tancredi, *An Elliptic Generalization of Multiple Polylogarithms*, *Nucl. Phys.* **B925** (2017) 212–251, [arXiv:1709.03622 \[hep-ph\]](#). 5
- [34] J. Ablinger, J. Blümlein, A. De Freitas, M. van Hoeij, E. Imamoglu, C. G. Raab, C. S. Radu, and C. Schneider, *Iterated Elliptic and Hypergeometric Integrals for Feynman Diagrams*, [arXiv:1706.01299 \[hep-th\]](#). 5
- [35] M. Hidding and F. Moriello, *All orders structure and efficient computation of linearly*

- reducible elliptic Feynman integrals, [arXiv:1712.04441 \[hep-ph\]](#). 5
- [36] L.-B. Chen, J. Jiang, and C.-F. Qiao, *Two-Loop integrals for CP-even heavy quarkonium production and decays: Elliptic Sectors*, [arXiv:1712.03516 \[hep-ph\]](#). 5
- [37] L.-B. Chen, *Two-Loop master integrals for heavy-to-light form factors of two different massive fermions*, *JHEP* **02** (2018) 066, [arXiv:1801.01033 \[hep-ph\]](#). 5
- [38] J. Broedel, C. Duhr, F. Dulat, and L. Tancredi, *Elliptic polylogarithms and iterated integrals on elliptic curves I: general formalism*, [arXiv:1712.07089 \[hep-th\]](#). 5
- [39] J. Broedel, C. Duhr, F. Dulat, and L. Tancredi, *Elliptic polylogarithms and iterated integrals on elliptic curves II: an application to the sunrise integral*, [arXiv:1712.07095 \[hep-ph\]](#). 5
- [40] L. Adams and S. Weinzierl, *The  $\varepsilon$ -form of the differential equations for Feynman integrals in the elliptic case*, [arXiv:1802.05020 \[hep-ph\]](#). 5
- [41] C. Duhr, *Mathematical aspects of scattering amplitudes*, in *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders (TASI 2014): Boulder, Colorado, June 2-27, 2014*, pp. 419–476. 2015. [arXiv:1411.7538 \[hep-ph\]](#). 5
- [42] S. Abreu, R. Britto, C. Duhr, and E. Gardi, *Algebraic Structure of Cut Feynman Integrals and the Diagrammatic Coaction*, *Phys. Rev. Lett.* **119** (2017) no. 5, 051601, [arXiv:1703.05064 \[hep-th\]](#). 5
- [43] F. Dulat, B. Mistlberger, and A. Pelloni, *Differential Higgs production at  $N^3LO$  beyond threshold*, *JHEP* **01** (2018) 145, [arXiv:1710.03016 \[hep-ph\]](#). 5, 11
- [44] B. Mistlberger, *Higgs Boson Production at Hadron Colliders at  $N^3LO$  in QCD*, [arXiv:1802.00833 \[hep-ph\]](#). 5, 11
- [45] A. von Manteuffel, E. Panzer, and R. M. Schabinger, *On the Computation of Form Factors in Massless QCD with Finite Master Integrals*, *Phys. Rev.* **D93** (2016) no. 12, 125014, [arXiv:1510.06758 \[hep-ph\]](#). 5
- [46] J. Davies, A. Vogt, B. Ruijl, T. Ueda, and J. A. M. Vermaseren, *Large- $n_f$  contributions to the four-loop splitting functions in QCD*, *Nucl. Phys.* **B915** (2017) 335–362, [arXiv:1610.07477 \[hep-ph\]](#). 5
- [47] A. von Manteuffel and R. M. Schabinger, *Quark and gluon form factors to four-loop order in QCD: the  $N_f^3$  contributions*, *Phys. Rev.* **D95** (2017) no. 3, 034030, [arXiv:1611.00795 \[hep-ph\]](#). 5
- [48] J. Henn, A. V. Smirnov, V. A. Smirnov, M. Steinhauser, and R. N. Lee, *Four-loop photon quark form factor and cusp anomalous dimension in the large- $N_c$  limit of QCD*, *JHEP* **03** (2017) 139, [arXiv:1612.04389 \[hep-ph\]](#). 5
- [49] B. Ruijl, T. Ueda, J. A. M. Vermaseren, and A. Vogt, *Four-loop QCD propagators and vertices with one vanishing external momentum*, *JHEP* **06** (2017) 040, [arXiv:1703.08532 \[hep-ph\]](#). 5
- [50] R. N. Lee, A. V. Smirnov, V. A. Smirnov, and M. Steinhauser, *The  $n_f^2$  contributions to fermionic four-loop form factors*, *Phys. Rev.* **D96** (2017) no. 1, 014008, [arXiv:1705.06862 \[hep-ph\]](#). 5
- [51] S. Moch, B. Ruijl, T. Ueda, J. A. M. Vermaseren, and A. Vogt, *Four-Loop Non-Singlet Splitting Functions in the Planar Limit and Beyond*, *JHEP* **10** (2017) 041, [arXiv:1707.08315 \[hep-ph\]](#). 5
- [52] P. A. Baikov, K. G. Chetyrkin, and J. H. Kühn, *Five-Loop Running of the QCD coupling constant*, *Phys. Rev. Lett.* **118** (2017) no. 8, 082002, [arXiv:1606.08659 \[hep-ph\]](#). 5

- [53] F. Herzog, B. Ruijl, T. Ueda, J. A. M. Vermaseren, and A. Vogt, *The five-loop beta function of Yang-Mills theory with fermions*, *JHEP* **02** (2017) 090, [arXiv:1701.01404 \[hep-ph\]](#). 5
- [54] T. Luthe, A. Maier, P. Marquard, and Y. Schroder, *The five-loop Beta function for a general gauge group and anomalous dimensions beyond Feynman gauge*, *JHEP* **10** (2017) 166, [arXiv:1709.07718 \[hep-ph\]](#). 5
- [55] K. G. Chetyrkin, G. Falcioni, F. Herzog, and J. A. M. Vermaseren, *Five-loop renormalisation of QCD in covariant gauges*, *JHEP* **10** (2017) 179, [arXiv:1709.08541 \[hep-ph\]](#). [Addendum: JHEP12,006(2017)]. 5
- [56] F. Herzog, B. Ruijl, T. Ueda, J. A. M. Vermaseren, and A. Vogt, *On Higgs decays to hadrons and the R-ratio at  $N^4LO$* , *JHEP* **08** (2017) 113, [arXiv:1707.01044 \[hep-ph\]](#). 5
- [57] E. Panzer, *Algorithms for the symbolic integration of hyperlogarithms with applications to Feynman integrals*, *Comput. Phys. Commun.* **188** (2015) 148–166, [arXiv:1403.3385 \[hep-th\]](#). 5
- [58] C. Bogner, *MPL-A program for computations with iterated integrals on moduli spaces of curves of genus zero*, *Comput. Phys. Commun.* **203** (2016) 339–353, [arXiv:1510.04562 \[physics.comp-ph\]](#). 5
- [59] R. N. Lee and K. T. Mingulov, *DREAM, a program for arbitrary-precision computation of dimensional recurrence relations solutions, and its applications*, [arXiv:1712.05173 \[hep-ph\]](#). 5
- [60] C. Bogner, S. Borowka, T. Hahn, G. Heinrich, S. P. Jones, M. Kerner, A. von Manteuffel, M. Michel, E. Panzer, and V. Papara, *Loopedia, a Database for Loop Integrals*, *Comput. Phys. Commun.* **225** (2018) 1–9, [arXiv:1709.01266 \[hep-ph\]](#). 5
- [61] T. Binoth and G. Heinrich, *An automatized algorithm to compute infrared divergent multiloop integrals*, *Nucl. Phys.* **B585** (2000) 741–759, [arXiv:hep-ph/0004013 \[hep-ph\]](#). 6, 7
- [62] S. Borowka, G. Heinrich, S. Jahn, S. P. Jones, M. Kerner, J. Schlenk, and T. Zirke, *pySecDec: a toolbox for the numerical evaluation of multi-scale integrals*, *Comput. Phys. Commun.* **222** (2018) 313–326, [arXiv:1703.09692 \[hep-ph\]](#). 6, 11
- [63] S. Borowka, G. Heinrich, S. P. Jones, M. Kerner, J. Schlenk, and T. Zirke, *SecDec-3.0: numerical evaluation of multi-scale integrals beyond one loop*, *Comput. Phys. Commun.* **196** (2015) 470–491, [arXiv:1502.06595 \[hep-ph\]](#). 6, 11, 37
- [64] A. V. Smirnov, *FIESTA4: Optimized Feynman integral calculations with GPU support*, *Comput. Phys. Commun.* **204** (2016) 189–199, [arXiv:1511.03614 \[hep-ph\]](#). 6
- [65] S. Borowka, N. Greiner, G. Heinrich, S. P. Jones, M. Kerner, J. Schlenk, U. Schubert, and T. Zirke, *Higgs Boson Pair Production in Gluon Fusion at Next-to-Leading Order with Full Top-Quark Mass Dependence*, *Phys. Rev. Lett.* **117** (2016) no. 1, 012001, [arXiv:1604.06447 \[hep-ph\]](#). [Erratum: Phys. Rev. Lett.117,no.7,079901(2016)]. 6, 13, 57, 150
- [66] S. Borowka, N. Greiner, G. Heinrich, S. P. Jones, M. Kerner, J. Schlenk, and T. Zirke, *Full top quark mass dependence in Higgs boson pair production at NLO*, *JHEP* **10** (2016) 107, [arXiv:1608.04798 \[hep-ph\]](#). 6, 13, 57, 150
- [67] S. P. Jones, M. Kerner, and G. Luisoni, *NLO QCD corrections to Higgs boson plus jet production with full top-quark mass dependence*, [arXiv:1802.00349 \[hep-ph\]](#). 6, 11, 57
- [68] F. Caola, J. M. Henn, K. Melnikov, A. V. Smirnov, and V. A. Smirnov, *Two-loop helicity amplitudes for the production of two off-shell electroweak bosons in*

- quark-antiquark collisions, *JHEP* **11** (2014) 041, [arXiv:1408.6409 \[hep-ph\]](#). 6
- [69] F. Caola, J. M. Henn, K. Melnikov, A. V. Smirnov, and V. A. Smirnov, *Two-loop helicity amplitudes for the production of two off-shell electroweak bosons in gluon fusion*, *JHEP* **06** (2015) 129, [arXiv:1503.08759 \[hep-ph\]](#). 6, 16
- [70] A. von Manteuffel and L. Tancredi, *The two-loop helicity amplitudes for  $gg \rightarrow V_1 V_2 \rightarrow 4$  leptons*, *JHEP* **06** (2015) 197, [arXiv:1503.08835 \[hep-ph\]](#). 6, 16
- [71] T. Gehrmann, A. von Manteuffel, and L. Tancredi, *The two-loop helicity amplitudes for  $q\bar{q}' \rightarrow V_1 V_2 \rightarrow 4$  leptons*, *JHEP* **09** (2015) 128, [arXiv:1503.04812 \[hep-ph\]](#). 6
- [72] A. V. Smirnov, *Algorithm FIRE – Feynman Integral REduction*, *JHEP* **10** (2008) 107, [arXiv:0807.3243 \[hep-ph\]](#). 6
- [73] A. V. Smirnov and V. A. Smirnov, *FIRE4, LiteRed and accompanying tools to solve integration by parts relations*, *Comput. Phys. Commun.* **184** (2013) 2820–2827, [arXiv:1302.5885 \[hep-ph\]](#). 6
- [74] A. V. Smirnov, *FIRE5: a C++ implementation of Feynman Integral REduction*, *Comput. Phys. Commun.* **189** (2015) 182–191, [arXiv:1408.2372 \[hep-ph\]](#). 6
- [75] R. N. Lee, *Presenting LiteRed: a tool for the Loop InTEgrals REDuction*, [arXiv:1212.2685 \[hep-ph\]](#). 6
- [76] C. Studerus, *Reduze-Feynman Integral Reduction in C++*, *Comput. Phys. Commun.* **181** (2010) 1293–1300, [arXiv:0912.2546 \[physics.comp-ph\]](#). 6
- [77] A. von Manteuffel and C. Studerus, *Reduze 2 - Distributed Feynman Integral Reduction*, [arXiv:1201.4330 \[hep-ph\]](#). 6
- [78] A. von Manteuffel and R. M. Schabinger, *A novel approach to integration by parts reduction*, *Phys. Lett.* **B744** (2015) 101–104, [arXiv:1406.4513 \[hep-ph\]](#). 6
- [79] T. Peraro, *Scattering amplitudes over finite fields and multivariate functional reconstruction*, *JHEP* **12** (2016) 030, [arXiv:1608.01902 \[hep-ph\]](#). 6
- [80] J. Boehm, A. Georgoudis, K. J. Larsen, M. Schulze, and Y. Zhang, *Complete sets of logarithmic vector fields for integration-by-parts identities of Feynman integrals*, [arXiv:1712.09737 \[hep-th\]](#). 6
- [81] P. Maierhöfer, J. Usovitsch, and P. Uwer, *Kira - A Feynman Integral Reduction Program*, [arXiv:1705.05610 \[hep-ph\]](#). 6
- [82] B. Ruijl, T. Ueda, and J. A. M. Vermaseren, *Forcer, a FORM program for the parametric reduction of four-loop massless propagator diagrams*, [arXiv:1704.06650 \[hep-ph\]](#). 6
- [83] C. F. Berger, Z. Bern, L. J. Dixon, F. Febres Cordero, D. Forde, H. Ita, D. A. Kosower, and D. Maitre, *An Automated Implementation of On-Shell Methods for One-Loop Amplitudes*, *Phys. Rev.* **D78** (2008) 036003, [arXiv:0803.4180 \[hep-ph\]](#). 6, 49
- [84] G. Bevilacqua, M. Czakon, M. V. Garzelli, A. van Hameren, A. Kardos, C. G. Papadopoulos, R. Pittau, and M. Worek, *HELAC-NLO*, *Comput. Phys. Commun.* **184** (2013) 986–997, [arXiv:1110.1499 \[hep-ph\]](#). 6, 35, 49
- [85] F. Cascioli, P. Maierhöfer, and S. Pozzorini, *Scattering Amplitudes with Open Loops*, *Phys. Rev. Lett.* **108** (2012) 111601, [arXiv:1111.5206 \[hep-ph\]](#). 6, 18, 35, 49, 53, 56, 173, 179, 208
- [86] S. Badger, B. Biedermann, P. Uwer, and V. Yundin, *Numerical evaluation of virtual corrections to multi-jet production in massless QCD*, *Comput. Phys. Commun.* **184** (2013) 1981–1998, [arXiv:1209.0100 \[hep-ph\]](#). 6
- [87] G. Cullen et al., *GOSAM-2.0: a tool for automated one-loop calculations within the Standard Model and beyond*, *Eur. Phys. J.* **C74** (2014) no. 8, 3001, [arXiv:1404.7096](#)

- [hep-ph]. 6, 35, 51, 56
- [88] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **07** (2014) 079, [arXiv:1405.0301 \[hep-ph\]](#). 6, 35, 51, 55, 56, 101, 149, 150, 173, 186, 247
- [89] S. Actis, A. Denner, L. Hofer, J.-N. Lang, A. Scharf, and S. Uccirati, *RECOLA: REcursive Computation of One-Loop Amplitudes*, *Comput. Phys. Commun.* **214** (2017) 140–173, [arXiv:1605.01090 \[hep-ph\]](#). 6, 50, 52, 187
- [90] Z. Bern, L. J. Dixon, D. C. Dunbar, and D. A. Kosower, *One loop n point gauge theory amplitudes, unitarity and collinear limits*, *Nucl. Phys.* **B425** (1994) 217–260, [arXiv:hep-ph/9403226 \[hep-ph\]](#). 6
- [91] Z. Bern, L. J. Dixon, D. C. Dunbar, and D. A. Kosower, *Fusing gauge theory tree amplitudes into loop amplitudes*, *Nucl. Phys.* **B435** (1995) 59–101, [arXiv:hep-ph/9409265 \[hep-ph\]](#). 6
- [92] R. Britto, F. Cachazo, and B. Feng, *Generalized unitarity and one-loop amplitudes in  $N=4$  super-Yang-Mills*, *Nucl. Phys.* **B725** (2005) 275–305, [arXiv:hep-th/0412103 \[hep-th\]](#). 6
- [93] W. T. Giele, Z. Kunszt, and K. Melnikov, *Full one-loop amplitudes from tree amplitudes*, *JHEP* **04** (2008) 049, [arXiv:0801.2237 \[hep-ph\]](#). 6
- [94] D. Forde, *Direct extraction of one-loop integral coefficients*, *Phys. Rev.* **D75** (2007) 125019, [arXiv:0704.1835 \[hep-ph\]](#). 6
- [95] G. Ossola, C. G. Papadopoulos, and R. Pittau, *Reducing full one-loop amplitudes to scalar integrals at the integrand level*, *Nucl. Phys.* **B763** (2007) 147–169, [arXiv:hep-ph/0609007 \[hep-ph\]](#). 6, 54, 56, 57
- [96] R. K. Ellis, W. T. Giele, Z. Kunszt, and K. Melnikov, *Masses, fermions and generalized  $D$ -dimensional unitarity*, *Nucl. Phys.* **B822** (2009) 270–282, [arXiv:0806.3467 \[hep-ph\]](#). 6
- [97] R. K. Ellis, Z. Kunszt, K. Melnikov, and G. Zanderighi, *One-loop calculations in quantum field theory: from Feynman diagrams to unitarity cuts*, *Phys. Rept.* **518** (2012) 141–250, [arXiv:1105.4319 \[hep-ph\]](#). 6
- [98] P. Mastrolia and G. Ossola, *On the Integrand-Reduction Method for Two-Loop Scattering Amplitudes*, *JHEP* **11** (2011) 014, [arXiv:1107.6041 \[hep-ph\]](#). 6
- [99] S. Badger, H. Frellesvig, and Y. Zhang, *Hepta-Cuts of Two-Loop Scattering Amplitudes*, *JHEP* **04** (2012) 055, [arXiv:1202.2019 \[hep-ph\]](#). 6
- [100] Y. Zhang, *Integrand-Level Reduction of Loop Amplitudes by Computational Algebraic Geometry Methods*, *JHEP* **09** (2012) 042, [arXiv:1205.5707 \[hep-ph\]](#). 6
- [101] R. H. P. Kleiss, I. Malamos, C. G. Papadopoulos, and R. Verheyen, *Counting to One: Reducibility of One- and Two-Loop Amplitudes at the Integrand Level*, *JHEP* **12** (2012) 038, [arXiv:1206.4180 \[hep-ph\]](#). 6
- [102] B. Feng and R. Huang, *The classification of two-loop integrand basis in pure four-dimension*, *JHEP* **02** (2013) 117, [arXiv:1209.3747 \[hep-ph\]](#). 6
- [103] P. Mastrolia, E. Mirabella, G. Ossola, and T. Peraro, *Scattering Amplitudes from Multivariate Polynomial Division*, *Phys. Lett.* **B718** (2012) 173–177, [arXiv:1205.7087 \[hep-ph\]](#). 6
- [104] P. Mastrolia, E. Mirabella, G. Ossola, and T. Peraro, *Multiloop Integrand Reduction for Dimensionally Regulated Amplitudes*, *Phys. Lett.* **B727** (2013) 532–535,

- [arXiv:1307.5832 \[hep-ph\]](#). 6
- [105] S. Badger, H. Frellesvig, and Y. Zhang, *A Two-Loop Five-Gluon Helicity Amplitude in QCD*, *JHEP* **12** (2013) 045, [arXiv:1310.1051 \[hep-ph\]](#). 6, 14
- [106] P. Mastrolia, T. Peraro, and A. Primo, *Adaptive Integrand Decomposition in parallel and orthogonal space*, *JHEP* **08** (2016) 164, [arXiv:1605.03157 \[hep-ph\]](#). 6
- [107] D. A. Kosower and K. J. Larsen, *Maximal Unitarity at Two Loops*, *Phys. Rev.* **D85** (2012) 045017, [arXiv:1108.1180 \[hep-th\]](#). 6
- [108] K. J. Larsen, *Global Poles of the Two-Loop Six-Point  $N=4$  SYM integrand*, *Phys. Rev.* **D86** (2012) 085032, [arXiv:1205.0297 \[hep-th\]](#). 6
- [109] S. Caron-Huot and K. J. Larsen, *Uniqueness of two-loop master contours*, *JHEP* **10** (2012) 026, [arXiv:1205.0801 \[hep-ph\]](#). 6
- [110] H. Johansson, D. A. Kosower, and K. J. Larsen, *Two-Loop Maximal Unitarity with External Masses*, *Phys. Rev.* **D87** (2013) no. 2, 025030, [arXiv:1208.1754 \[hep-th\]](#). 6
- [111] H. Johansson, D. A. Kosower, and K. J. Larsen, *Maximal Unitarity for the Four-Mass Double Box*, *Phys. Rev.* **D89** (2014) no. 12, 125010, [arXiv:1308.4632 \[hep-th\]](#). 6
- [112] H. Johansson, D. A. Kosower, K. J. Larsen, and M. Sogaard, *Cross-Order Integral Relations from Maximal Cuts*, *Phys. Rev.* **D92** (2015) no. 2, 025015, [arXiv:1503.06711 \[hep-th\]](#). 6
- [113] J. Gluza, K. Kajda, and D. A. Kosower, *Towards a Basis for Planar Two-Loop Integrals*, *Phys. Rev.* **D83** (2011) 045012, [arXiv:1009.0472 \[hep-th\]](#). 6
- [114] R. M. Schabinger, *A New Algorithm For The Generation Of Unitarity-Compatible Integration By Parts Relations*, *JHEP* **01** (2012) 077, [arXiv:1111.4220 \[hep-ph\]](#). 6
- [115] H. Ita, *Two-loop Integrand Decomposition into Master Integrals and Surface Terms*, *Phys. Rev.* **D94** (2016) no. 11, 116015, [arXiv:1510.05626 \[hep-th\]](#). 6
- [116] K. J. Larsen and Y. Zhang, *Integration-by-parts reductions from unitarity cuts and algebraic geometry*, *Phys. Rev.* **D93** (2016) no. 4, 041701, [arXiv:1511.01071 \[hep-th\]](#). 6
- [117] V. Hirschi and T. Peraro, *Tensor integrand reduction via Laurent expansion*, *JHEP* **06** (2016) 060, [arXiv:1604.01363 \[hep-ph\]](#). 6, 56
- [118] A. Georgoudis, K. J. Larsen, and Y. Zhang, *Azurite: An algebraic geometry based package for finding bases of loop integrals*, *Comput. Phys. Commun.* **221** (2017) 203–215, [arXiv:1612.04252 \[hep-th\]](#). 6
- [119] K. J. Larsen and R. Rietkerk, *MultivariateResidues: a Mathematica package for computing multivariate residues*, *Comput. Phys. Commun.* **222** (2018) 250–262, [arXiv:1701.01040 \[hep-th\]](#). 6
- [120] A. Georgoudis, K. J. Larsen, and Y. Zhang, *Cristal and Azurite: new tools for integration-by-parts reductions*, in *13th International Symposium on Radiative Corrections: Application of Quantum Field Theory to Phenomenology (RADCOR 2017) St. Gilgen, Austria, September 24-29, 2017*. 2017. [arXiv:1712.07510 \[hep-ph\]](#). 6
- [121] S. Abreu, F. Febres Cordero, H. Ita, M. Jaquier, B. Page, and M. Zeng, *Two-Loop Four-Gluon Amplitudes from Numerical Unitarity*, *Phys. Rev. Lett.* **119** (2017) no. 14, 142001, [arXiv:1703.05273 \[hep-ph\]](#). 6, 13
- [122] S. Badger, G. Mogull, A. Ochirov, and D. O’Connell, *A Complete Two-Loop, Five-Gluon Helicity Amplitude in Yang-Mills Theory*, *JHEP* **10** (2015) 064, [arXiv:1507.08797 \[hep-ph\]](#). 6
- [123] S. Abreu, F. Febres Cordero, H. Ita, B. Page, and M. Zeng, *Planar Two-Loop Five-Gluon Amplitudes from Numerical Unitarity*, [arXiv:1712.03946 \[hep-ph\]](#). 6, 14,

- [124] S. Badger, C. Brønnum-Hansen, H. B. Hartanto, and T. Peraro, *First look at two-loop five-gluon scattering in QCD*, *Phys. Rev. Lett.* **120** (2018) no. 9, 092001, [arXiv:1712.02229 \[hep-ph\]](#). 6, 14, 94, 107
- [125] A. Gehrmann-De Ridder, T. Gehrmann, and E. W. N. Glover, *Antenna subtraction at NNLO*, *JHEP* **09** (2005) 056, [arXiv:hep-ph/0505111 \[hep-ph\]](#). 6, 208
- [126] J. Currie, E. W. N. Glover, and S. Wells, *Infrared Structure at NNLO Using Antenna Subtraction*, *JHEP* **04** (2013) 066, [arXiv:1301.4693 \[hep-ph\]](#). 6, 208
- [127] A. Gehrmann-De Ridder, T. Gehrmann, E. W. N. Glover, and G. Heinrich, *EERAD3: Event shapes and jet rates in electron-positron annihilation at order  $\alpha_s^3$* , *Comput. Phys. Commun.* **185** (2014) 3331, [arXiv:1402.4140 \[hep-ph\]](#). 7
- [128] T. Gehrmann, E. W. N. Glover, A. Huss, J. Niehues, and H. Zhang, *NNLO QCD corrections to event orientation in  $e^+e^-$  annihilation*, *Phys. Lett.* **B775** (2017) 185–189, [arXiv:1709.01097 \[hep-ph\]](#). 7
- [129] J. Currie, E. W. N. Glover, and J. Pires, *Next-to-Next-to Leading Order QCD Predictions for Single Jet Inclusive Production at the LHC*, *Phys. Rev. Lett.* **118** (2017) no. 7, 072002, [arXiv:1611.01460 \[hep-ph\]](#). 7, 13, 22, 81, 89, 240
- [130] A. Gehrmann-De Ridder, T. Gehrmann, E. W. N. Glover, A. Huss, and T. A. Morgan, *Precise QCD predictions for the production of a Z boson in association with a hadronic jet*, *Phys. Rev. Lett.* **117** (2016) no. 2, 022001, [arXiv:1507.02850 \[hep-ph\]](#). 7, 16, 22
- [131] A. Gehrmann-De Ridder, T. Gehrmann, E. W. N. Glover, A. Huss, and T. A. Morgan, *The NNLO QCD corrections to Z boson production at large transverse momentum*, *JHEP* **07** (2016) 133, [arXiv:1605.04295 \[hep-ph\]](#). 7, 22
- [132] A. Gehrmann-De Ridder, T. Gehrmann, E. W. N. Glover, A. Huss, and D. M. Walker, *NNLO QCD corrections to the transverse momentum distribution of weak gauge bosons*, [arXiv:1712.07543 \[hep-ph\]](#). 7, 16
- [133] X. Chen, J. Cruz-Martinez, T. Gehrmann, E. W. N. Glover, and M. Jaquier, *NNLO QCD corrections to Higgs boson production at large transverse momentum*, *JHEP* **10** (2016) 066, [arXiv:1607.08817 \[hep-ph\]](#). 7, 11
- [134] J. Currie, T. Gehrmann, A. Huss, and J. Niehues, *NNLO QCD corrections to jet production in deep inelastic scattering*, *JHEP* **07** (2017) 018, [arXiv:1703.05977 \[hep-ph\]](#). 7
- [135] J. Cruz-Martinez, T. Gehrmann, E. W. N. Glover, and A. Huss, *Second-order QCD effects in Higgs boson production through vector boson fusion*, Submitted to: *JHEP* (2018), [arXiv:1802.02445 \[hep-ph\]](#). 7, 12, 140
- [136] M. Czakon, *A novel subtraction scheme for double-real radiation at NNLO*, *Phys. Lett.* **B693** (2010) 259–268, [arXiv:1005.0274 \[hep-ph\]](#). 7, 30, 32
- [137] M. Czakon, *Double-real radiation in hadronic top quark pair production as a proof of a certain concept*, *Nucl. Phys.* **B849** (2011) 250–295, [arXiv:1101.0642 \[hep-ph\]](#). 7, 30, 32
- [138] R. Boughezal, K. Melnikov, and F. Petriello, *A subtraction scheme for NNLO computations*, *Phys. Rev.* **D85** (2012) 034025, [arXiv:1111.7041 \[hep-ph\]](#). 7
- [139] G. Heinrich, *A numerical method for NNLO calculations*, *Nucl. Phys. Proc. Suppl.* **116** (2003) 368–372, [arXiv:hep-ph/0211144 \[hep-ph\]](#). 7
- [140] C. Anastasiou, K. Melnikov, and F. Petriello, *A new method for real radiation at NNLO*, *Phys. Rev.* **D69** (2004) 076010, [arXiv:hep-ph/0311311 \[hep-ph\]](#). 7
- [141] T. Binoth and G. Heinrich, *Numerical evaluation of phase space integrals by sector*

- decomposition, *Nucl. Phys.* **B693** (2004) 134–148, [arXiv:hep-ph/0402265 \[hep-ph\]](#). 7
- [142] S. Frixione, Z. Kunszt, and A. Signer, *Three jet cross-sections to next-to-leading order*, *Nucl. Phys.* **B467** (1996) 399–442, [arXiv:hep-ph/9512328 \[hep-ph\]](#). 7, 30, 51
- [143] R. Frederix, S. Frixione, F. Maltoni, and T. Stelzer, *Automation of next-to-leading order computations in QCD: The FKS subtraction*, *JHEP* **10** (2009) 003, [arXiv:0908.4272 \[hep-ph\]](#). 7, 51
- [144] M. Czakon and D. Heymes, *Four-dimensional formulation of the sector-improved residue subtraction scheme*, *Nucl. Phys.* **B890** (2014) 152–227, [arXiv:1408.2500 \[hep-ph\]](#). 7, 30
- [145] M. Czakon, P. Fiedler, and A. Mitov, *Resolving the Tevatron Top Quark Forward-Backward Asymmetry Puzzle: Fully Differential Next-to-Next-to-Leading-Order Calculation*, *Phys. Rev. Lett.* **115** (2015) no. 5, 052001, [arXiv:1411.3007 \[hep-ph\]](#). 7, 30
- [146] M. Czakon, D. Heymes, and A. Mitov, *High-precision differential predictions for top-quark pairs at the LHC*, *Phys. Rev. Lett.* **116** (2016) no. 8, 082003, [arXiv:1511.00549 \[hep-ph\]](#). 7, 17, 30, 173, 178
- [147] M. Czakon, P. Fiedler, D. Heymes, and A. Mitov, *NNLO QCD predictions for fully-differential top-quark pair production at the Tevatron*, *JHEP* **05** (2016) 034, [arXiv:1601.05375 \[hep-ph\]](#). 7, 17, 19, 30
- [148] M. Brucherseifer, F. Caola, and K. Melnikov,  $\mathcal{O}(\alpha_s^2)$  corrections to fully-differential top quark decays, *JHEP* **04** (2013) 059, [arXiv:1301.7133 \[hep-ph\]](#). 7, 18, 178
- [149] M. Brucherseifer, F. Caola, and K. Melnikov, *On the NNLO QCD corrections to single-top production at the LHC*, *Phys. Lett.* **B736** (2014) 58–63, [arXiv:1404.7116 \[hep-ph\]](#). 7, 19, 30
- [150] R. Boughezal, F. Caola, K. Melnikov, F. Petriello, and M. Schulze, *Higgs boson production in association with a jet at next-to-next-to-leading order*, *Phys. Rev. Lett.* **115** (2015) no. 8, 082003, [arXiv:1504.07922 \[hep-ph\]](#). 7, 11, 30, 208
- [151] F. Caola, K. Melnikov, and M. Schulze, *Fiducial cross sections for Higgs boson production in association with a jet at next-to-next-to-leading order in QCD*, *Phys. Rev.* **D92** (2015) no. 7, 074032, [arXiv:1508.02684 \[hep-ph\]](#). 7, 11, 30
- [152] S. Catani and M. Grazzini, *An NNLO subtraction formalism in hadron collisions and its application to Higgs boson production at the LHC*, *Phys. Rev. Lett.* **98** (2007) 222002, [arXiv:hep-ph/0703012 \[hep-ph\]](#). 7
- [153] R. Bonciani, S. Catani, M. Grazzini, H. Sargsyan, and A. Torre, *The  $q_T$  subtraction method for top quark production at hadron colliders*, *Eur. Phys. J.* **C75** (2015) no. 12, 581, [arXiv:1508.03585 \[hep-ph\]](#). 7
- [154] M. Grazzini, S. Kallweit, and M. Wiesemann, *Fully differential NNLO computations with MATRIX*, [arXiv:1711.06631 \[hep-ph\]](#). 7, 16, 50
- [155] S. Catani, L. Cieri, G. Ferrera, D. de Florian, and M. Grazzini, *Vector boson production at hadron colliders: a fully exclusive QCD calculation at NNLO*, *Phys. Rev. Lett.* **103** (2009) 082001, [arXiv:0903.2120 \[hep-ph\]](#). 7, 22
- [156] S. Catani, L. Cieri, D. de Florian, G. Ferrera, and M. Grazzini, *Diphoton production at hadron colliders: a fully-differential QCD calculation at NNLO*, *Phys. Rev. Lett.* **108** (2011) 072001, [arXiv:1110.2375 \[hep-ph\]](#). 7
- [157] G. Ferrera, M. Grazzini, and F. Tramontano, *Higher-order QCD effects for associated WH production and decay at the LHC*, *JHEP* **04** (2014) 039, [arXiv:1312.1669 \[hep-ph\]](#). 7

- [158] M. Grazzini, S. Kallweit, D. Rathlev, and A. Torre, *Z $\gamma$  production at hadron colliders in NNLO QCD*, *Phys. Lett.* **B731** (2014) 204–207, [arXiv:1309.7000 \[hep-ph\]](#). 7, 16, 50
- [159] G. Ferrera, M. Grazzini, and F. Tramontano, *Associated ZH production at hadron colliders: the fully differential NNLO QCD calculation*, *Phys. Lett.* **B740** (2015) 51–55, [arXiv:1407.4747 \[hep-ph\]](#). 7, 12
- [160] T. Gehrmann, M. Grazzini, S. Kallweit, P. Maierhöfer, A. von Manteuffel, S. Pozzorini, D. Rathlev, and L. Tancredi, *W<sup>+</sup>W<sup>-</sup> Production at Hadron Colliders in Next to Next to Leading Order QCD*, *Phys. Rev. Lett.* **113** (2014) no. 21, 212001, [arXiv:1408.5243 \[hep-ph\]](#). 7, 16, 50
- [161] F. Cascioli, T. Gehrmann, M. Grazzini, S. Kallweit, P. Maierhöfer, A. von Manteuffel, S. Pozzorini, D. Rathlev, L. Tancredi, and E. Weihs, *ZZ production at hadron colliders in NNLO QCD*, *Phys. Lett.* **B735** (2014) 311–313, [arXiv:1405.2219 \[hep-ph\]](#). 7, 16, 50
- [162] M. Grazzini, S. Kallweit, and D. Rathlev, *W $\gamma$  and Z $\gamma$  production at the LHC in NNLO QCD*, *JHEP* **07** (2015) 085, [arXiv:1504.01330 \[hep-ph\]](#). 7, 16, 50
- [163] M. Grazzini, S. Kallweit, D. Rathlev, and M. Wiesemann, *W<sup>±</sup>Z production at hadron colliders in NNLO QCD*, *Phys. Lett.* **B761** (2016) 179–183, [arXiv:1604.08576 \[hep-ph\]](#). 7, 16, 50
- [164] M. Grazzini, S. Kallweit, and D. Rathlev, *ZZ production at the LHC: fiducial cross sections and distributions in NNLO QCD*, *Phys. Lett.* **B750** (2015) 407–410, [arXiv:1507.06257 \[hep-ph\]](#). 7, 16, 50
- [165] M. Grazzini, S. Kallweit, S. Pozzorini, D. Rathlev, and M. Wiesemann, *W<sup>+</sup>W<sup>-</sup> production at the LHC: fiducial cross sections and distributions in NNLO QCD*, *JHEP* **08** (2016) 140, [arXiv:1605.02716 \[hep-ph\]](#). 7, 16, 50
- [166] M. Grazzini, S. Kallweit, D. Rathlev, and M. Wiesemann, *W<sup>±</sup>Z production at the LHC: fiducial cross sections and distributions in NNLO QCD*, *JHEP* **05** (2017) 139, [arXiv:1703.09065 \[hep-ph\]](#). 7, 16, 50
- [167] S. Catani, L. Cieri, D. de Florian, G. Ferrera, and M. Grazzini, *Diphoton production at the LHC: a QCD study up to NNLO*, [arXiv:1802.02095 \[hep-ph\]](#). 7
- [168] M. Grazzini, G. Heinrich, S. Jones, S. Kallweit, M. Kerner, J. M. Lindert, and J. Mazzitelli, *Higgs boson pair production at NNLO with top quark mass effects*, [arXiv:1803.02463 \[hep-ph\]](#). 7, 13
- [169] G. Ferrera, M. Grazzini, and F. Tramontano, *Associated WH production at hadron colliders: a fully exclusive QCD calculation at NNLO*, *Phys. Rev. Lett.* **107** (2011) 152003, [arXiv:1107.1164 \[hep-ph\]](#). 7, 12
- [170] D. de Florian, M. Grazzini, C. Hanga, S. Kallweit, J. M. Lindert, P. Maierhöfer, J. Mazzitelli, and D. Rathlev, *Differential Higgs Boson Pair Production at Next-to-Next-to-Leading Order in QCD*, *JHEP* **09** (2016) 151, [arXiv:1606.09519 \[hep-ph\]](#). 7, 13, 50
- [171] H. T. Li and J. Wang, *Fully Differential Higgs Pair Production in Association With a W Boson at Next-to-Next-to-Leading Order in QCD*, *Phys. Lett.* **B765** (2017) 265–271, [arXiv:1607.06382 \[hep-ph\]](#). 7
- [172] H. T. Li, C. S. Li, and J. Wang, *Fully Differential Higgs Pair Production in Association With a Z Boson at Next-to-Next-to-Leading Order in QCD*, [arXiv:1710.02464 \[hep-ph\]](#). 7
- [173] R. Boughezal, X. Liu, and F. Petriello, *N-jettiness soft function at next-to-next-to-leading order*, *Phys. Rev.* **D91** (2015) no. 9, 094035, [arXiv:1504.02540 \[hep-ph\]](#). 7

- [174] R. Boughezal, C. Focke, X. Liu, and F. Petriello, *W-boson production in association with a jet at next-to-next-to-leading order in perturbative QCD*, *Phys. Rev. Lett.* **115** (2015) no. 6, 062002, [arXiv:1504.02131 \[hep-ph\]](#). 7, 16
- [175] J. Gaunt, M. Stahlhofen, F. J. Tackmann, and J. R. Walsh, *N-jettiness Subtractions for NNLO QCD Calculations*, *JHEP* **09** (2015) 058, [arXiv:1505.04794 \[hep-ph\]](#). 7
- [176] R. Boughezal, C. Focke, W. Giele, X. Liu, and F. Petriello, *Higgs boson production in association with a jet at NNLO using jettiness subtraction*, *Phys. Lett.* **B748** (2015) 5–8, [arXiv:1505.03893 \[hep-ph\]](#). 7, 11, 213
- [177] R. Boughezal, J. M. Campbell, R. K. Ellis, C. Focke, W. T. Giele, X. Liu, and F. Petriello, *Z-boson production in association with a jet at next-to-next-to-leading order in perturbative QCD*, *Phys. Rev. Lett.* **116** (2016) no. 15, 152001, [arXiv:1512.01291 \[hep-ph\]](#). 7, 16
- [178] R. Boughezal, X. Liu, and F. Petriello, *A comparison of NNLO QCD predictions with 7 TeV ATLAS and CMS data for V+jet processes*, *Phys. Lett.* **B760** (2016) 6–13, [arXiv:1602.05612 \[hep-ph\]](#). 7, 16
- [179] R. Boughezal, X. Liu, and F. Petriello, *W-boson plus jet differential distributions at NNLO in QCD*, *Phys. Rev.* **D94** (2016) no. 11, 113009, [arXiv:1602.06965 \[hep-ph\]](#). 7, 16
- [180] J. M. Campbell, R. K. Ellis, and C. Williams, *Associated production of a Higgs boson at NNLO*, *JHEP* **06** (2016) 179, [arXiv:1601.00658 \[hep-ph\]](#). 7, 12
- [181] J. M. Campbell, R. K. Ellis, and C. Williams, *Direct Photon Production at Next-to-Next-to-Leading Order*, *Phys. Rev. Lett.* **118** (2017) no. 22, 222001, [arXiv:1612.04333 \[hep-ph\]](#). 7, 16
- [182] J. M. Campbell, T. Neumann, and C. Williams, *Z $\gamma$  production at NNLO including anomalous couplings*, *JHEP* **11** (2017) 150, [arXiv:1708.02925 \[hep-ph\]](#). 7, 16
- [183] R. Boughezal, J. M. Campbell, R. K. Ellis, C. Focke, W. Giele, X. Liu, F. Petriello, and C. Williams, *Color singlet production at NNLO in MCFM*, *Eur. Phys. J.* **C77** (2017) no. 1, 7, [arXiv:1605.08011 \[hep-ph\]](#). 7
- [184] J. Gao, C. S. Li, and H. X. Zhu, *Top Quark Decay at Next-to-Next-to Leading Order in QCD*, *Phys. Rev. Lett.* **110** (2013) no. 4, 042001, [arXiv:1210.2808 \[hep-ph\]](#). 7, 18
- [185] E. L. Berger, J. Gao, C. P. Yuan, and H. X. Zhu, *NNLO QCD Corrections to t-channel Single Top-Quark Production and Decay*, *Phys. Rev.* **D94** (2016) no. 7, 071501, [arXiv:1606.08463 \[hep-ph\]](#). 7, 19
- [186] V. Del Duca, C. Duhr, G. Somogyi, F. Tramontano, and Z. Trócsányi, *Higgs boson decay into b-quarks at NNLO accuracy*, *JHEP* **04** (2015) 036, [arXiv:1501.07226 \[hep-ph\]](#). 7
- [187] S. Catani and M. H. Seymour, *A General algorithm for calculating jet cross-sections in NLO QCD*, *Nucl. Phys.* **B485** (1997) 291–419, [arXiv:hep-ph/9605323 \[hep-ph\]](#). [Erratum: *Nucl. Phys.*B510,503(1998)]. 7, 39, 50, 51, 60, 186, 187, 209, 218, 219
- [188] V. Del Duca, C. Duhr, A. Kardos, G. Somogyi, and Z. Trócsányi, *Three-Jet Production in Electron-Positron Collisions at Next-to-Next-to-Leading Order Accuracy*, *Phys. Rev. Lett.* **117** (2016) no. 15, 152004, [arXiv:1603.08927 \[hep-ph\]](#). 7
- [189] V. Del Duca, C. Duhr, A. Kardos, G. Somogyi, Z. Szőr, Z. Trócsányi, and Z. Tulipánt, *Jet production in the CoLoRFulNNLO method: event shapes in electron-positron collisions*, *Phys. Rev.* **D94** (2016) no. 7, 074019, [arXiv:1606.03453 \[hep-ph\]](#). 7
- [190] Z. Tulipánt, A. Kardos, and G. Somogyi, *Energy–energy correlation in electron–positron annihilation at NNLL + NNLO accuracy*, *Eur. Phys. J.* **C77** (2017) no. 11, 749, [arXiv:1708.04093 \[hep-ph\]](#). 7

- [191] F. Caola, K. Melnikov, and R. Röntsch, *Nested soft-collinear subtractions in NNLO QCD computations*, *Eur. Phys. J.* **C77** (2017) no. 4, 248, [arXiv:1702.01352 \[hep-ph\]](#). 7, 13, 30, 33
- [192] F. Caola, G. Luisoni, K. Melnikov, and R. Röntsch, *NNLO QCD corrections to associated WH production and  $H \rightarrow b\bar{b}$  decay*, [arXiv:1712.06954 \[hep-ph\]](#). 7, 13
- [193] L. Magnea, E. Maina, P. Torrielli, and S. Uccirati, *Towards analytic local sector subtraction at NNLO*, in *Proceedings, 13th International Symposium on Radiative Corrections: Application of Quantum Field Theory to Phenomenology (RADCOR2017): St. Gilgen, Austria, September 24-29, 2017*. 2018. [arXiv:1801.06458 \[hep-ph\]](#). 7
- [194] M. Cacciari, F. A. Dreyer, A. Karlberg, G. P. Salam, and G. Zanderighi, *Fully Differential Vector-Boson-Fusion Higgs Production at Next-to-Next-to-Leading Order*, *Phys. Rev. Lett.* **115** (2015) no. 8, 082002, [arXiv:1506.02660 \[hep-ph\]](#). 7, 12, 140
- [195] M. Grazzini, S. Kallweit, D. Rathlev, and M. Wiesemann, *Transverse-momentum resummation for vector-boson pair production at NNLL+NNLO*, *JHEP* **08** (2015) 154, [arXiv:1507.02565 \[hep-ph\]](#). 9, 50
- [196] S. Alioli, C. W. Bauer, C. Berggren, F. J. Tackmann, and J. R. Walsh, *Drell-Yan production at NNLL+NNLO matched to parton showers*, *Phys. Rev.* **D92** (2015) no. 9, 094020, [arXiv:1508.01475 \[hep-ph\]](#). 9, 14
- [197] I. Moulton, L. Rothen, I. W. Stewart, F. J. Tackmann, and H. X. Zhu, *Subleading Power Corrections for N-Jettiness Subtractions*, *Phys. Rev.* **D95** (2017) no. 7, 074023, [arXiv:1612.00450 \[hep-ph\]](#). 9
- [198] I. Moulton, L. Rothen, I. W. Stewart, F. J. Tackmann, and H. X. Zhu, *N-jettiness subtractions for  $gg \rightarrow H$  at subleading power*, *Phys. Rev.* **D97** (2018) no. 1, 014013, [arXiv:1710.03227 \[hep-ph\]](#). 9
- [199] M. Beneke, M. Garny, R. Szafron, and J. Wang, *Anomalous dimension of subleading-power N-jet operators*, *JHEP* **03** (2018) 001, [arXiv:1712.04416 \[hep-ph\]](#). 9
- [200] R. Boughezal, A. Isgro, and F. Petriello, *Next-to-leading-logarithmic power corrections for N-jettiness subtraction in color-singlet production*, [arXiv:1802.00456 \[hep-ph\]](#). 9
- [201] R. Boughezal, X. Liu, F. Petriello, F. J. Tackmann, and J. R. Walsh, *Combining Resummed Higgs Predictions Across Jet Bins*, *Phys. Rev.* **D89** (2014) no. 7, 074044, [arXiv:1312.4535 \[hep-ph\]](#). 9
- [202] R. Boughezal, C. Focke, Y. Li, and X. Liu, *Jet vetoes for Higgs production at future hadron colliders*, *Phys. Rev.* **D90** (2014) no. 5, 053001, [arXiv:1405.4562 \[hep-ph\]](#). 9
- [203] A. Banfi, F. Caola, F. A. Dreyer, P. F. Monni, G. P. Salam, G. Zanderighi, and F. Dulat, *Jet-vetoed Higgs cross section in gluon fusion at  $N^3LO+NNLL$  with small- $R$  resummation*, *JHEP* **04** (2016) 049, [arXiv:1511.02886 \[hep-ph\]](#). 9
- [204] P. F. Monni, E. Re, and P. Torrielli, *Higgs Transverse-Momentum Resummation in Direct Space*, *Phys. Rev. Lett.* **116** (2016) no. 24, 242001, [arXiv:1604.02191 \[hep-ph\]](#). 9
- [205] M. A. Ebert and F. J. Tackmann, *Resummation of Transverse Momentum Distributions in Distribution Space*, *JHEP* **02** (2017) 110, [arXiv:1611.08610 \[hep-ph\]](#). 9
- [206] F. Caola, S. Forte, S. Marzani, C. Muselli, and G. Vita, *The Higgs transverse momentum spectrum with finite quark masses beyond leading order*, *JHEP* **08** (2016) 150, [arXiv:1606.04100 \[hep-ph\]](#). 9, 11
- [207] M. A. Ebert, J. K. L. Michel, and F. J. Tackmann, *Resummation Improved Rapidity Spectrum for Gluon Fusion Higgs Production*, *JHEP* **05** (2017) 088, [arXiv:1702.00794 \[hep-ph\]](#). 9

- [208] W. Bizon, P. F. Monni, E. Re, L. Rottoli, and P. Torrielli, *Momentum-space resummation for transverse observables and the Higgs  $p_{\perp}$  at  $N^3LL+NNLO$* , **JHEP** **02** (2018) 108, [arXiv:1705.09127 \[hep-ph\]](#). 9
- [209] K. Melnikov and A. Penin, *On the light quark mass effects in Higgs boson production in gluon fusion*, **JHEP** **05** (2016) 172, [arXiv:1602.09020 \[hep-ph\]](#). 9
- [210] T. Liu and A. A. Penin, *High-Energy Limit of QCD beyond the Sudakov Approximation*, **Phys. Rev. Lett.** **119** (2017) no. 26, 262001, [arXiv:1709.01092 \[hep-ph\]](#). 9
- [211] M. Dasgupta, F. Dreyer, G. P. Salam, and G. Soyez, *Small-radius jets to all orders in QCD*, **JHEP** **04** (2015) 039, [arXiv:1411.5182 \[hep-ph\]](#). 9, 89
- [212] M. Dasgupta, F. A. Dreyer, G. P. Salam, and G. Soyez, *Inclusive jet spectrum for small-radius jets*, **JHEP** **06** (2016) 057, [arXiv:1602.01110 \[hep-ph\]](#). 9, 89, 207
- [213] D. W. Kolodrubetz, P. Pietrulewicz, I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, *Factorization for Jet Radius Logarithms in Jet Mass Spectra at the LHC*, **JHEP** **12** (2016) 054, [arXiv:1605.08038 \[hep-ph\]](#). 9
- [214] X. Liu, S.-O. Moch, and F. Ringer, *Threshold and jet radius joint resummation for single-inclusive jet production*, **Phys. Rev. Lett.** **119** (2017) no. 21, 212001, [arXiv:1708.04641 \[hep-ph\]](#). 9
- [215] X. Liu, S.-O. Moch, and F. Ringer, *Phenomenology of single-inclusive jet production with jet radius and threshold resummation*, [arXiv:1801.07284 \[hep-ph\]](#). 9
- [216] G. Luisoni and S. Marzani, *QCD resummation for hadronic final states*, **J. Phys.** **G42** (2015) no. 10, 103101, [arXiv:1505.04084 \[hep-ph\]](#). 9
- [217] K. Hamilton, P. Nason, E. Re, and G. Zanderighi, *NNLOPS simulation of Higgs boson production*, **JHEP** **10** (2013) 222, [arXiv:1309.0017 \[hep-ph\]](#). 9, 11
- [218] S. Alioli, C. W. Bauer, C. Berggren, F. J. Tackmann, J. R. Walsh, and S. Zuberi, *Matching Fully Differential NNLO Calculations and Parton Showers*, **JHEP** **06** (2014) 089, [arXiv:1311.0286 \[hep-ph\]](#). 9
- [219] S. Höche, Y. Li, and S. Prestel, *Drell-Yan lepton pair production at NNLO QCD with parton showers*, **Phys. Rev.** **D91** (2015) no. 7, 074015, [arXiv:1405.3607 \[hep-ph\]](#). 9, 14
- [220] W. Astill, W. Bizon, E. Re, and G. Zanderighi, *NNLOPS accurate associated HW production*, **JHEP** **06** (2016) 154, [arXiv:1603.01620 \[hep-ph\]](#). 9, 13
- [221] M. Billoni, S. Dittmaier, B. Jäger, and C. Speckner, *Next-to-leading order electroweak corrections to  $pp \rightarrow W+W^- \rightarrow 4$  leptons at the LHC in double-pole approximation*, **JHEP** **12** (2013) 043, [arXiv:1310.1564 \[hep-ph\]](#). 10
- [222] B. Biedermann, A. Denner, S. Dittmaier, L. Hofer, and B. Jäger, *Electroweak corrections to  $pp \rightarrow \mu^+\mu^-e^+e^- + X$  at the LHC: a Higgs background study*, **Phys. Rev. Lett.** **116** (2016) no. 16, 161803, [arXiv:1601.07787 \[hep-ph\]](#). 10, 16, 50, 53
- [223] B. Biedermann, M. Billoni, A. Denner, S. Dittmaier, L. Hofer, B. Jäger, and L. Salfelder, *Next-to-leading-order electroweak corrections to  $pp \rightarrow W^+W^- \rightarrow 4$  leptons at the LHC*, **JHEP** **06** (2016) 065, [arXiv:1605.03419 \[hep-ph\]](#). 10, 16, 50, 53
- [224] B. Biedermann, A. Denner, and M. Pellen, *Large electroweak corrections to vector-boson scattering at the Large Hadron Collider*, **Phys. Rev. Lett.** **118** (2017) no. 26, 261801, [arXiv:1611.02951 \[hep-ph\]](#). 10, 16, 50, 53, 185, 187
- [225] B. Biedermann, A. Denner, S. Dittmaier, L. Hofer, and B. Jäger, *Next-to-leading-order electroweak corrections to the production of four charged leptons at the LHC*, **JHEP** **01** (2017) 033, [arXiv:1611.05338 \[hep-ph\]](#). 10, 16, 50, 53, 58, 60, 67
- [226] B. Biedermann, A. Denner, and M. Pellen, *Complete NLO corrections to  $W^+W^+$*

- scattering and its irreducible background at the LHC, *JHEP* **10** (2017) 124, [arXiv:1708.00268 \[hep-ph\]](#). 10, 12, 16, 50, 53, 185, 186, 187, 188, 189
- [227] F. Dulat, S. Lionetti, B. Mistlberger, A. Pelloni, and C. Specchia, *Higgs-differential cross section at NNLO in dimensional regularisation*, *JHEP* **07** (2017) 017, [arXiv:1704.08220 \[hep-ph\]](#). 10
- [228] F. Dulat, A. Lazopoulos, and B. Mistlberger, *iHixs 2 - Inclusive Higgs Cross Sections*, [arXiv:1802.00827 \[hep-ph\]](#). 11
- [229] M. Bonetti, K. Melnikov, and L. Tancredi, *Three-loop mixed QCD-electroweak corrections to Higgs boson gluon fusion*, *Phys. Rev.* **D97** (2018) no. 3, 034004, [arXiv:1711.11113 \[hep-ph\]](#). 11
- [230] M. Bonetti, K. Melnikov, and L. Tancredi, *Higher order corrections to mixed QCD-EW contributions to Higgs production in gluon fusion*, [arXiv:1801.10403 \[hep-ph\]](#). 11
- [231] R. V. Harlander, H. Mantler, S. Marzani, and K. J. Ozeren, *Higgs production in gluon fusion at next-to-next-to-leading order QCD for finite top mass*, *Eur. Phys. J.* **C66** (2010) 359–372, [arXiv:0912.2104 \[hep-ph\]](#). 11
- [232] A. Pak, M. Rogal, and M. Steinhauser, *Production of scalar and pseudo-scalar Higgs bosons to next-to-next-to-leading order at hadron colliders*, *JHEP* **09** (2011) 088, [arXiv:1107.3391 \[hep-ph\]](#). 11
- [233] C. Anastasiou, C. Duhr, F. Dulat, E. Furlan, T. Gehrmann, F. Herzog, A. Lazopoulos, and B. Mistlberger, *High precision determination of the gluon fusion Higgs boson cross-section at the LHC*, *JHEP* **05** (2016) 058, [arXiv:1602.00695 \[hep-ph\]](#). 11
- [234] S. Höche, Y. Li, and S. Prestel, *Higgs-boson production through gluon fusion at NNLO QCD with parton showers*, *Phys. Rev.* **D90** (2014) no. 5, 054011, [arXiv:1407.3773 \[hep-ph\]](#). 11
- [235] K. Hamilton, P. Nason, and G. Zanderighi, *Finite quark-mass effects in the NNLOPS POWHEG+MiNLO Higgs generator*, *JHEP* **05** (2015) 140, [arXiv:1501.04637 \[hep-ph\]](#). 11
- [236] LHC Higgs Cross Section Working Group Collaboration, D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, [arXiv:1610.07922 \[hep-ph\]](#). 11, 80, 159, 160, 188, 208
- [237] J. Campbell, J. Huston, and F. Krauss, *The black book of quantum chromodynamics: a primer for the LHC era*. Oxford University Press, Oxford, 2018. <https://cds.cern.ch/record/2286381>. 11
- [238] X. Chen, T. Gehrmann, E. W. N. Glover, and M. Jaquier, *Precise QCD predictions for the production of Higgs + jet final states*, *Phys. Lett.* **B740** (2015) 147–150, [arXiv:1408.5325 \[hep-ph\]](#). 11, 208
- [239] K. Melnikov, L. Tancredi, and C. Wever, *Two-loop  $gg \rightarrow Hg$  amplitude mediated by a nearly massless quark*, *JHEP* **11** (2016) 104, [arXiv:1610.03747 \[hep-ph\]](#). 11
- [240] J. M. Lindert, K. Melnikov, L. Tancredi, and C. Wever, *Top-bottom interference effects in Higgs plus jet production at the LHC*, *Phys. Rev. Lett.* **118** (2017) no. 25, 252002, [arXiv:1703.03886 \[hep-ph\]](#). 11
- [241] J. M. Lindert, K. Kudashkin, K. Melnikov, and C. Wever, *Higgs bosons with large transverse momentum at the LHC*, [arXiv:1801.08226 \[hep-ph\]](#). 11
- [242] T. Neumann, *NLO Higgs+jet at large transverse momenta including top quark mass effects*, [arXiv:1802.02981 \[hep-ph\]](#). 11
- [243] R. Frederix, S. Frixione, E. Vryonidou, and M. Wiesemann, *Heavy-quark mass effects in Higgs plus jets production*, *JHEP* **08** (2016) 006, [arXiv:1604.03017 \[hep-ph\]](#). 11

- [244] T. Neumann and C. Williams, *The Higgs boson at high  $p_T$* , *Phys. Rev.* **D95** (2017) no. 1, 014004, [arXiv:1609.00367 \[hep-ph\]](#). 11
- [245] M. Buschmann, D. Goncalves, S. Kuttimalai, M. Schönherr, F. Krauss, and T. Plehn, *Mass Effects in the Higgs-Gluon Coupling: Boosted vs Off-Shell Production*, *JHEP* **02** (2015) 038, [arXiv:1410.5806 \[hep-ph\]](#). 11
- [246] ATLAS Collaboration, M. Aaboud et al., *Measurements of Higgs boson properties in the diphoton decay channel with  $36 \text{ fb}^{-1}$  of pp collision data at  $\sqrt{s} = 13 \text{ TeV}$  with the ATLAS detector*, [arXiv:1802.04146 \[hep-ex\]](#). 11, 12
- [247] F. A. Dreyer and A. Karlberg, *Vector-Boson Fusion Higgs Production at Three Loops in QCD*, *Phys. Rev. Lett.* **117** (2016) no. 7, 072001, [arXiv:1606.00840 \[hep-ph\]](#). 12
- [248] N. Greiner, S. Höche, G. Luisoni, M. Schönherr, and J.-C. Winter, *Full mass dependence in Higgs boson production in association with jets at the LHC and FCC*, *JHEP* **01** (2017) 091, [arXiv:1608.01195 \[hep-ph\]](#). 12
- [249] N. Greiner, S. Höche, G. Luisoni, M. Schönherr, J.-C. Winter, and V. Yundin, *Phenomenological analysis of Higgs boson production through gluon fusion in association with jets*, *JHEP* **01** (2016) 169, [arXiv:1506.01016 \[hep-ph\]](#). 12, 213
- [250] F. Campanario, T. M. Figy, S. Plätzer, and M. Sjö Dahl, *Electroweak Higgs Boson Plus Three Jet Production at Next-to-Leading-Order QCD*, *Phys. Rev. Lett.* **111** (2013) no. 21, 211802, [arXiv:1308.2932 \[hep-ph\]](#). 12
- [251] F. Campanario, T. M. Figy, S. Plätzer, M. Rauch, P. Schichtel, and M. Sjö Dahl, *Stress-testing the VBF approximation in multijet final states*, [arXiv:1802.09955 \[hep-ph\]](#). 12
- [252] M. Ciccolini, A. Denner, and S. Dittmaier, *Strong and electroweak corrections to the production of Higgs + 2jets via weak interactions at the LHC*, *Phys. Rev. Lett.* **99** (2007) 161803, [arXiv:0707.0381 \[hep-ph\]](#). 12
- [253] A. Denner, S. Dittmaier, S. Kallweit, and A. Mück, *HAWK 2.0: A Monte Carlo program for Higgs production in vector-boson fusion and Higgs strahlung at hadron colliders*, *Comput. Phys. Commun.* **195** (2015) 161–171, [arXiv:1412.5390 \[hep-ph\]](#). 12
- [254] O. Brein, A. Djouadi, and R. Harlander, *NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders*, *Phys. Lett.* **B579** (2004) 149–156, [arXiv:hep-ph/0307206 \[hep-ph\]](#). 12
- [255] O. Brein, R. Harlander, M. Wiesemann, and T. Zirke, *Top-Quark Mediated Effects in Hadronic Higgs-Strahlung*, *Eur. Phys. J.* **C72** (2012) 1868, [arXiv:1111.0761 \[hep-ph\]](#). 12
- [256] O. Brein, R. V. Harlander, and T. J. E. Zirke, *vh@nnlo - Higgs Strahlung at hadron colliders*, *Comput. Phys. Commun.* **184** (2013) 998–1003, [arXiv:1210.5347 \[hep-ph\]](#). 12
- [257] R. V. Harlander, S. Liebler, and T. Zirke, *Higgs Strahlung at the Large Hadron Collider in the 2-Higgs-Doublet Model*, *JHEP* **02** (2014) 023, [arXiv:1307.8122 \[hep-ph\]](#). 12, 146
- [258] R. V. Harlander, J. Klappert, S. Liebler, and L. Simon, *vh@nnlo-v2: New physics in Higgs Strahlung*, [arXiv:1802.04817 \[hep-ph\]](#). 12, 146
- [259] M. L. Ciccolini, S. Dittmaier, and M. Kramer, *Electroweak radiative corrections to associated WH and ZH production at hadron colliders*, *Phys. Rev.* **D68** (2003) 073003, [arXiv:hep-ph/0306234 \[hep-ph\]](#). 12
- [260] A. Denner, S. Dittmaier, S. Kallweit, and A. Mück, *Electroweak corrections to Higgs-strahlung off W/Z bosons at the Tevatron and the LHC with HAWK*, *JHEP* **03**

- (2012) 075, [arXiv:1112.5142 \[hep-ph\]](#). 12
- [261] P. Obul, S. Dulat, T.-J. Hou, A. Tursun, and N. Yalkun, *The next-to-leading order QCD and electroweak corrections to the Higgs-strahlung processes*, [arXiv:1801.06851 \[hep-ph\]](#). 12
- [262] F. Granata, J. M. Lindert, C. Oleari, and S. Pozzorini, *NLO QCD+EW predictions for HV and HV +jet production including parton-shower effects*, *JHEP* **09** (2017) 012, [arXiv:1706.03522 \[hep-ph\]](#). 12, 55
- [263] S. Dawson, T. Han, W. K. Lai, A. K. Leibovich, and I. Lewis, *Resummation Effects in Vector-Boson and Higgs Associated Production*, *Phys. Rev.* **D86** (2012) 074007, [arXiv:1207.4207 \[hep-ph\]](#). 12
- [264] G. Ferrera, G. Somogyi, and F. Tramontano, *Associated production of a Higgs boson decaying into bottom quarks at the LHC in full NNLO QCD*, [arXiv:1705.10304 \[hep-ph\]](#). 12
- [265] L. Altenkamp, S. Dittmaier, R. V. Harlander, H. Rzehak, and T. J. E. Zirke, *Gluon-induced Higgs-strahlung at next-to-leading order QCD*, *JHEP* **02** (2013) 078, [arXiv:1211.5015 \[hep-ph\]](#). 12
- [266] R. V. Harlander, A. Kulesza, V. Theeuwes, and T. Zirke, *Soft gluon resummation for gluon-induced Higgs Strahlung*, *JHEP* **11** (2014) 082, [arXiv:1410.0217 \[hep-ph\]](#). 12
- [267] A. Hasselhuhn, T. Luthe, and M. Steinhauser, *On top quark mass effects to  $gg \rightarrow ZH$  at NLO*, *JHEP* **01** (2017) 073, [arXiv:1611.05881 \[hep-ph\]](#). 12
- [268] C. Degrande, B. Fuks, K. Mawatari, K. Mimasu, and V. Sanz, *Electroweak Higgs boson production in the standard model effective field theory beyond leading order in QCD*, *Eur. Phys. J.* **C77** (2017) no. 4, 262, [arXiv:1609.04833 \[hep-ph\]](#). 13, 52
- [269] A. Greljo, G. Isidori, J. M. Lindert, D. Marzocca, and H. Zhang, *Electroweak Higgs production with HiggsPO at NLO QCD*, *Eur. Phys. J.* **C77** (2017) no. 12, 838, [arXiv:1710.04143 \[hep-ph\]](#). 13
- [270] M. C. Kumar, M. K. Mandal, and V. Ravindran, *Associated production of Higgs boson with vector boson at threshold  $N^3LO$  in QCD*, *JHEP* **03** (2015) 037, [arXiv:1412.3357 \[hep-ph\]](#). 13
- [271] D. de Florian and J. Mazzitelli, *Higgs Boson Pair Production at Next-to-Next-to-Leading Order in QCD*, *Phys. Rev. Lett.* **111** (2013) 201801, [arXiv:1309.6594 \[hep-ph\]](#). 13
- [272] D. Y. Shao, C. S. Li, H. T. Li, and J. Wang, *Threshold resummation effects in Higgs boson pair production at the LHC*, *JHEP* **07** (2013) 169, [arXiv:1301.1245 \[hep-ph\]](#). 13
- [273] D. de Florian and J. Mazzitelli, *Higgs pair production at next-to-next-to-leading logarithmic accuracy at the LHC*, *JHEP* **09** (2015) 053, [arXiv:1505.07122 \[hep-ph\]](#). 13
- [274] J. Grigo, J. Hoff, K. Melnikov, and M. Steinhauser, *On the Higgs boson pair production at the LHC*, *Nucl. Phys.* **B875** (2013) 1–17, [arXiv:1305.7340 \[hep-ph\]](#). 13
- [275] J. Grigo, J. Hoff, and M. Steinhauser, *Higgs boson pair production: top quark mass effects at NLO and NNLO*, *Nucl. Phys.* **B900** (2015) 412–430, [arXiv:1508.00909 \[hep-ph\]](#). 13
- [276] G. Heinrich, S. P. Jones, M. Kerner, G. Luisoni, and E. Vryonidou, *NLO predictions for Higgs boson pair production with full top quark mass dependence matched to parton showers*, *JHEP* **08** (2017) 088, [arXiv:1703.09252 \[hep-ph\]](#). 13, 150
- [277] S. Jones and S. Kuttimalai, *Parton Shower and NLO-Matching uncertainties in Higgs*

- Boson Pair Production*, [arXiv:1711.03319 \[hep-ph\]](#). 13, 150, 156
- [278] R. Gröber, A. Maier, and T. Rauh, *Reconstruction of top-quark mass effects in Higgs pair production and other gluon-fusion processes*, [arXiv:1709.07799 \[hep-ph\]](#). 13
- [279] S. Frixione, V. Hirschi, D. Pagani, H. S. Shao, and M. Zaro, *Weak corrections to Higgs hadroproduction in association with a top-quark pair*, *JHEP* **09** (2014) 065, [arXiv:1407.0823 \[hep-ph\]](#). 13, 49, 51
- [280] S. Frixione, V. Hirschi, D. Pagani, H. S. Shao, and M. Zaro, *Electroweak and QCD corrections to top-pair hadroproduction in association with heavy bosons*, *JHEP* **06** (2015) 184, [arXiv:1504.03446 \[hep-ph\]](#). 13, 19, 51
- [281] A. Denner and R. Feger, *NLO QCD corrections to off-shell top-antitop production with leptonic decays in association with a Higgs boson at the LHC*, *JHEP* **11** (2015) 209, [arXiv:1506.07448 \[hep-ph\]](#). 13, 50, 53
- [282] A. Denner, J.-N. Lang, M. Pellen, and S. Uccirati, *Higgs production in association with off-shell top-antitop pairs at NLO EW and QCD at the LHC*, *JHEP* **02** (2017) 053, [arXiv:1612.07138 \[hep-ph\]](#). 13, 50, 53
- [283] A. Kulesza, L. Motyka, T. Stebel, and V. Theeuwes, *Associated  $t\bar{t}H$  production at the LHC: theoretical predictions at NLO+NNLL accuracy*, [arXiv:1704.03363 \[hep-ph\]](#). 13
- [284] F. Maltoni, E. Vryonidou, and C. Zhang, *Higgs production in association with a top-antitop pair in the Standard Model Effective Field Theory at NLO in QCD*, *JHEP* **10** (2016) 123, [arXiv:1607.05330 \[hep-ph\]](#). 13, 52, 56
- [285] J. Campbell, R. K. Ellis, and R. Röntschi, *Single top production in association with a Z boson at the LHC*, *Phys. Rev.* **D87** (2013) 114006, [arXiv:1302.3856 \[hep-ph\]](#). 13
- [286] F. Demartin, F. Maltoni, K. Mawatari, and M. Zaro, *Higgs production in association with a single top quark at the LHC*, *Eur. Phys. J.* **C75** (2015) no. 6, 267, [arXiv:1504.00611 \[hep-ph\]](#). 13, 52
- [287] J. Currie, E. W. N. Glover, T. Gehrmann, A. Gehrmann-De Ridder, A. Huss, and J. Pires, *Single Jet Inclusive Production for the Individual Jet  $p_T$  Scale Choice at the LHC*, *Acta Phys. Polon.* **B48** (2017) 955–967, [arXiv:1704.00923 \[hep-ph\]](#). 13, 89
- [288] R. Frederix, S. Frixione, V. Hirschi, D. Pagani, H.-S. Shao, and M. Zaro, *The complete NLO corrections to dijet hadroproduction*, *JHEP* **04** (2017) 076, [arXiv:1612.06548 \[hep-ph\]](#). 13, 51, 56
- [289] T. Gehrmann, J. M. Henn, and N. A. Lo Presti, *Analytic form of the two-loop planar five-gluon all-plus-helicity amplitude in QCD*, *Phys. Rev. Lett.* **116** (2016) no. 6, 062001, [arXiv:1511.05409 \[hep-ph\]](#). [Erratum: *Phys. Rev. Lett.*116,no.18,189903(2016)]. 14, 94, 107
- [290] C. G. Papadopoulos, D. Tommasini, and C. Wever, *The Pentabox Master Integrals with the Simplified Differential Equations approach*, *JHEP* **04** (2016) 078, [arXiv:1511.09404 \[hep-ph\]](#). 14
- [291] D. C. Dunbar and W. B. Perkins, *Two-loop five-point all plus helicity Yang-Mills amplitude*, *Phys. Rev.* **D93** (2016) no. 8, 085029, [arXiv:1603.07514 \[hep-th\]](#). 14, 94, 107
- [292] D. Chicherin, J. Henn, and V. Mitev, *Bootstrapping pentagon functions*, [arXiv:1712.09610 \[hep-th\]](#). 14
- [293] R. H. Boels, Q. Jin, and H. Luo, *Efficient integrand reduction for particles with spin*, [arXiv:1802.06761 \[hep-ph\]](#). 14
- [294] T. Ahmed, M. Mahakhud, N. Rana, and V. Ravindran, *Drell-Yan Production at*

- Threshold to Third Order in QCD*, *Phys. Rev. Lett.* **113** (2014) no. 11, 112002, [arXiv:1404.0366 \[hep-ph\]](#). 14
- [295] T. Ahmed, M. K. Mandal, N. Rana, and V. Ravindran, *Rapidity Distributions in Drell-Yan and Higgs Productions at Threshold to Third Order in QCD*, *Phys. Rev. Lett.* **113** (2014) 212003, [arXiv:1404.6504 \[hep-ph\]](#). 14
- [296] A. Karlberg, E. Re, and G. Zanderighi, *NNLOPS accurate Drell-Yan production*, *JHEP* **09** (2014) 134, [arXiv:1407.2940 \[hep-ph\]](#). 14
- [297] S. Dittmaier, A. Huss, and C. Schwinn, *Dominant mixed QCD-electroweak  $O(\alpha_s\alpha)$  corrections to Drell-Yan processes in the resonance region*, *Nucl. Phys.* **B904** (2016) 216–252, [arXiv:1511.08016 \[hep-ph\]](#). 14
- [298] ATLAS Collaboration, M. Aaboud et al., *Precision measurement and interpretation of inclusive  $W^+$ ,  $W^-$  and  $Z/\gamma^*$  production cross sections with the ATLAS detector*, *Eur. Phys. J.* **C77** (2017) no. 6, 367, [arXiv:1612.03016 \[hep-ex\]](#). 14
- [299] ATLAS Collaboration, G. Aad et al., *Measurement of the transverse momentum and  $\phi_\eta^*$  distributions of Drell-Yan lepton pairs in proton-proton collisions at  $\sqrt{s} = 8$  TeV with the ATLAS detector*, *Eur. Phys. J.* **C76** (2016) no. 5, 291, [arXiv:1512.02192 \[hep-ex\]](#). 15
- [300] CMS Collaboration, A. M. Sirunyan et al., *Measurement of differential cross sections in the  $\phi^*$  variable for inclusive Z boson production in pp collisions at  $\sqrt{s} = 8$  TeV*, [arXiv:1710.07955 \[hep-ex\]](#). 15
- [301] R. Boughezal, X. Liu, and F. Petriello, *Phenomenology of the Z-boson plus jet process at NNLO*, *Phys. Rev.* **D94** (2016) no. 7, 074015, [arXiv:1602.08140 \[hep-ph\]](#). 16
- [302] J. M. Lindert et al., *Precise predictions for  $V + jets$  dark matter backgrounds*, *Eur. Phys. J.* **C77** (2017) no. 12, 829, [arXiv:1705.04664 \[hep-ph\]](#). 16, 50, 51, 54
- [303] A. Denner, L. Hofer, A. Scharf, and S. Uccirati, *Electroweak corrections to lepton pair production in association with two hard jets at the LHC*, *JHEP* **01** (2015) 094, [arXiv:1411.0916 \[hep-ph\]](#). 16, 53
- [304] S. Kallweit, J. M. Lindert, P. Maierhöfer, S. Pozzorini, and M. Schönherr, *NLO electroweak automation and precise predictions for  $W + multijet$  production at the LHC*, *JHEP* **04** (2015) 012, [arXiv:1412.5157 \[hep-ph\]](#). 16, 49, 50, 51, 54, 173
- [305] S. Kallweit, J. M. Lindert, P. Maierhöfer, S. Pozzorini, and M. Schönherr, *NLO QCD+EW predictions for  $V + jets$  including off-shell vector-boson decays and multijet merging*, *JHEP* **04** (2016) 021, [arXiv:1511.08692 \[hep-ph\]](#). 16, 50, 51, 54, 174
- [306] G. Heinrich, S. Jahn, S. P. Jones, M. Kerner, and J. Pires, *NNLO predictions for Z-boson pair production at the LHC*, [arXiv:1710.06294 \[hep-ph\]](#). 16
- [307] F. Caola, K. Melnikov, R. Röntsch, and L. Tancredi, *QCD corrections to ZZ production in gluon fusion at the LHC*, *Phys. Rev.* **D92** (2015) no. 9, 094028, [arXiv:1509.06734 \[hep-ph\]](#). 16
- [308] F. Caola, K. Melnikov, R. Röntsch, and L. Tancredi, *QCD corrections to  $W^+W^-$  production through gluon fusion*, *Phys. Lett.* **B754** (2016) 275–280, [arXiv:1511.08617 \[hep-ph\]](#). 16
- [309] F. Caola, M. Dowling, K. Melnikov, R. Röntsch, and L. Tancredi, *QCD corrections to vector boson pair production in gluon fusion including interference effects with off-shell Higgs at the LHC*, *JHEP* **07** (2016) 087, [arXiv:1605.04610 \[hep-ph\]](#). 16
- [310] J. M. Campbell, R. K. Ellis, M. Czakon, and S. Kirchner, *Two loop correction to interference in  $gg \rightarrow ZZ$* , *JHEP* **08** (2016) 011, [arXiv:1605.01380 \[hep-ph\]](#). 16
- [311] A. Denner, S. Dittmaier, M. Hecht, and C. Pasold, *NLO QCD and electroweak*

- corrections to  $W + \gamma$  production with leptonic  $W$ -boson decays, *JHEP* **04** (2015) 018, [arXiv:1412.7421 \[hep-ph\]](#). 16
- [312] A. Denner, S. Dittmaier, M. Hecht, and C. Pasold, *NLO QCD and electroweak corrections to  $Z + \gamma$  production with leptonic  $Z$ -boson decays*, *JHEP* **02** (2016) 057, [arXiv:1510.08742 \[hep-ph\]](#). 16
- [313] B. Biedermann, A. Denner, and L. Hofer, *Next-to-leading-order electroweak corrections to the production of three charged leptons plus missing energy at the LHC*, *JHEP* **10** (2017) 043, [arXiv:1708.06938 \[hep-ph\]](#). 16, 50, 53
- [314] S. Kallweit, J. M. Lindert, S. Pozzorini, and M. Schönherr, *NLO QCD+EW predictions for  $2\ell 2\nu$  diboson signatures at the LHC*, *JHEP* **11** (2017) 120, [arXiv:1705.00598 \[hep-ph\]](#). 16, 50, 51, 54, 55
- [315] W.-H. Li, R.-Y. Zhang, W.-G. Ma, L. Guo, X.-Z. Li, and Y. Zhang, *NLO QCD and electroweak corrections to  $WW$ +jet production with leptonic  $W$  -boson decays at LHC*, *Phys. Rev.* **D92** (2015) no. 3, 033005, [arXiv:1507.07332 \[hep-ph\]](#). 16
- [316] Y. Wang, R.-Y. Zhang, W.-G. Ma, X.-Z. Li, and L. Guo, *QCD and electroweak corrections to  $ZZ$ +jet production with  $Z$  -boson leptonic decays at the LHC*, *Phys. Rev.* **D94** (2016) no. 1, 013011, [arXiv:1604.04080 \[hep-ph\]](#). 16
- [317] D. T. Nhung, L. D. Ninh, and M. M. Weber, *NLO corrections to  $WWZ$  production at the LHC*, *JHEP* **12** (2013) 096, [arXiv:1307.7403 \[hep-ph\]](#). 16
- [318] Y.-B. Shen, R.-Y. Zhang, W.-G. Ma, X.-Z. Li, Y. Zhang, and L. Guo, *NLO QCD + NLO EW corrections to  $WZZ$  productions with leptonic decays at the LHC*, *JHEP* **10** (2015) 186, [arXiv:1507.03693 \[hep-ph\]](#). [Erratum: *JHEP*10,156(2016)]. 16
- [319] Y.-B. Shen, R.-Y. Zhang, W.-G. Ma, X.-Z. Li, and L. Guo, *NLO QCD and electroweak corrections to  $WWW$  production at the LHC*, *Phys. Rev.* **D95** (2017) no. 7, 073005, [arXiv:1605.00554 \[hep-ph\]](#). 16
- [320] H. Wang, R.-Y. Zhang, W.-G. Ma, L. Guo, X.-Z. Li, and S.-M. Wang, *NLO QCD + EW corrections to  $ZZZ$  production with subsequent leptonic decays at the LHC*, *J. Phys.* **G43** (2016) no. 11, 115001, [arXiv:1610.05876 \[hep-ph\]](#). 16
- [321] S. Dittmaier, A. Huss, and G. Knippen, *Next-to-leading-order QCD and electroweak corrections to  $WWW$  production at proton-proton colliders*, *JHEP* **09** (2017) 034, [arXiv:1705.03722 \[hep-ph\]](#). 16
- [322] C. Chen, W.-G. Ma, R.-Y. Zhang, Y. Zhang, L.-W. Chen, and L. Guo, *Electroweak radiative corrections to  $W^+W^-\gamma$  production at the ILC*, *Eur. Phys. J.* **C74** (2014) no. 11, 3166, [arXiv:1409.4900 \[hep-ph\]](#). 17
- [323] Y. Wang, R.-Y. Zhang, W.-G. Ma, X.-Z. Li, S.-M. Wang, and H.-Y. Bi,  *$ZZ\gamma$  production in the NLO QCD+EW accuracy at the LHC*, *J. Phys.* **G44** (2017) no. 8, 085002, [arXiv:1707.03534 \[hep-ph\]](#). 17
- [324] N. Greiner and M. Schönherr, *NLO QCD+EW corrections to diphoton production in association with a vector boson*, *JHEP* **01** (2018) 079, [arXiv:1710.11514 \[hep-ph\]](#). 17, 51, 57
- [325] L. Cieri, F. Coradeschi, and D. de Florian, *Diphoton production at hadron colliders: transverse-momentum resummation at next-to-next-to-leading logarithmic accuracy*, *JHEP* **06** (2015) 185, [arXiv:1505.03162 \[hep-ph\]](#). 17
- [326] J. M. Campbell, R. K. Ellis, Y. Li, and C. Williams, *Predictions for diphoton production at the LHC through NNLO in QCD*, *JHEP* **07** (2016) 148, [arXiv:1603.02663 \[hep-ph\]](#). 17
- [327] M. Czakon, D. Heymes, and A. Mitov, *fastNLO tables for NNLO top-quark pair differential distributions*, [arXiv:1704.08551 \[hep-ph\]](#). 17, 178

- [328] M. Czakon, D. Heymes, and A. Mitov, *Dynamical scales for multi-TeV top-pair production at the LHC*, *JHEP* **04** (2017) 071, [arXiv:1606.03350 \[hep-ph\]](#). 17, 178
- [329] M. Czakon, D. Heymes, A. Mitov, D. Pagani, I. Tsirikos, and M. Zaro, *Top-pair production at the LHC through NNLO QCD and NLO EW*, *JHEP* **10** (2017) 186, [arXiv:1705.04105 \[hep-ph\]](#). 17, 51, 173, 174, 178
- [330] L. Chen, M. Czakon, and R. Poncelet, *Polarized double-virtual amplitudes for heavy-quark pair production*, [arXiv:1712.08075 \[hep-ph\]](#). 17
- [331] A. Denner and M. Pellen, *NLO electroweak corrections to off-shell top-antitop production with leptonic decays at the LHC*, *JHEP* **08** (2016) 155, [arXiv:1607.05571 \[hep-ph\]](#). 17, 50, 53, 173, 178
- [332] C. Gütschow, J. M. Lindert, and M. Schönherr, *Multi-jet merged top-pair production including electroweak corrections*, [arXiv:1803.00950 \[hep-ph\]](#). 18, 54, 173
- [333] W. Bernreuther, A. Brandenburg, Z. G. Si, and P. Uwer, *Top quark pair production and decay at hadron colliders*, *Nucl. Phys.* **B690** (2004) 81–137, [arXiv:hep-ph/0403035 \[hep-ph\]](#). 18, 178
- [334] K. Melnikov and M. Schulze, *NLO QCD corrections to top quark pair production and decay at hadron colliders*, *JHEP* **08** (2009) 049, [arXiv:0907.3090 \[hep-ph\]](#). 18, 178, 179
- [335] J. M. Campbell and R. K. Ellis, *Top-quark processes at NLO in production and decay*, *J. Phys.* **G42** (2015) no. 1, 015005, [arXiv:1204.1513 \[hep-ph\]](#). 18, 178
- [336] M. Beneke, P. Falgari, S. Klein, and C. Schwinn, *Hadronic top-quark pair production with NNLL threshold resummation*, *Nucl. Phys.* **B855** (2012) 695–741, [arXiv:1109.1536 \[hep-ph\]](#). 18, 173
- [337] M. Cacciari, M. Czakon, M. Mangano, A. Mitov, and P. Nason, *Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation*, *Phys. Lett.* **B710** (2012) 612–622, [arXiv:1111.5869 \[hep-ph\]](#). 18, 173
- [338] A. Ferroglia, S. Marzani, B. D. Pecjak, and L. L. Yang, *Boosted top production: factorization and resummation for single-particle inclusive distributions*, *JHEP* **01** (2014) 028, [arXiv:1310.3836 \[hep-ph\]](#). 18
- [339] A. Broggio, A. S. Papanastasiou, and A. Signer, *Renormalization-group improved fully differential cross sections for top pair production*, *JHEP* **10** (2014) 98, [arXiv:1407.2532 \[hep-ph\]](#). 18
- [340] N. Kidonakis, *High-order threshold corrections for top-pair and single-top production*, in *Proceedings, Meeting of the APS Division of Particles and Fields (DPF 2015): Ann Arbor, Michigan, USA, 4-8 Aug 2015*. 2015. [arXiv:1509.07848 \[hep-ph\]](#). 18
- [341] B. D. Pecjak, D. J. Scott, X. Wang, and L. L. Yang, *Resummed differential cross sections for top-quark pairs at the LHC*, *Phys. Rev. Lett.* **116** (2016) no. 20, 202001, [arXiv:1601.07020 \[hep-ph\]](#). 18
- [342] A. Denner, S. Dittmaier, S. Kallweit, and S. Pozzorini, *NLO QCD corrections to WWbb production at hadron colliders*, *Phys. Rev. Lett.* **106** (2011) 052001, [arXiv:1012.3975 \[hep-ph\]](#). 18, 50, 178
- [343] A. Denner, S. Dittmaier, S. Kallweit, and S. Pozzorini, *NLO QCD corrections to off-shell top-antitop production with leptonic decays at hadron colliders*, *JHEP* **10** (2012) 110, [arXiv:1207.5018 \[hep-ph\]](#). 18, 50, 178
- [344] G. Bevilacqua, M. Czakon, A. van Hameren, C. G. Papadopoulos, and M. Worek, *Complete off-shell effects in top quark pair hadroproduction with leptonic decay at next-to-leading order*, *JHEP* **02** (2011) 083, [arXiv:1012.4230 \[hep-ph\]](#). 18, 178

- [345] G. Heinrich, A. Maier, R. Nisius, J. Schlenk, and J. Winter, *NLO QCD corrections to  $W^+W^-b\bar{b}$  production with leptonic decays in the light of top quark mass and asymmetry measurements*, *JHEP* **1406** (2014) 158, [arXiv:1312.6659 \[hep-ph\]](#). 18, 178, 179, 180
- [346] A. Denner and M. Pellen, *Off-shell production of top-antitop pairs in the lepton+jets channel at NLO QCD*, *JHEP* **02** (2018) 013, [arXiv:1711.10359 \[hep-ph\]](#). 18, 50, 53, 178
- [347] R. Frederix, *Top Quark Induced Backgrounds to Higgs Production in the  $WW^{(*)} \rightarrow ll\nu\nu$  Decay Channel at Next-to-Leading-Order in QCD*, *Phys. Rev. Lett.* **112** (2014) no. 8, 082002, [arXiv:1311.4893 \[hep-ph\]](#). 18, 178
- [348] F. Cascioli, S. Kallweit, P. Maierhöfer, and S. Pozzorini, *A unified NLO description of top-pair and associated  $Wt$  production*, *Eur. Phys. J.* **C74** (2014) no. 3, 2783, [arXiv:1312.0546 \[hep-ph\]](#). 18, 50, 178, 179
- [349] J. M. Campbell, R. K. Ellis, P. Nason, and E. Re, *Top-Pair Production and Decay at NLO Matched with Parton Showers*, *JHEP* **04** (2015) 114, [arXiv:1412.1828 \[hep-ph\]](#). 18, 178
- [350] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, *JHEP* **11** (2007) 070, [arXiv:0709.2092 \[hep-ph\]](#). 18, 140, 150, 178, 203
- [351] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043, [arXiv:1002.2581 \[hep-ph\]](#). 18, 140, 150, 178
- [352] S. Höche, F. Krauss, P. Maierhöfer, S. Pozzorini, M. Schönherr, and F. Siegert, *Next-to-leading order QCD predictions for top-quark pair production with up to two jets merged with a parton shower*, *Phys. Lett.* **B748** (2015) 74–78, [arXiv:1402.6293 \[hep-ph\]](#). 18, 173, 176, 178
- [353] S. Höche, P. Maierhöfer, N. Moretti, S. Pozzorini, and F. Siegert, *Next-to-leading order QCD predictions for top-quark pair production with up to three jets*, *Eur. Phys. J.* **C77** (2017) no. 3, 145, [arXiv:1607.06934 \[hep-ph\]](#). 18, 173, 178
- [354] J. Bellm, K. Cormier, S. Gieseke, S. Plätzer, C. Reuschle, P. Richardson, and S. Webster, *Top Quark Production and Decay in Herwig 7.1*, [arXiv:1711.11570 \[hep-ph\]](#). 18
- [355] M. V. Garzelli, A. Kardos, and Z. Trócsányi, *Hadroproduction of  $W^+W^-b\bar{b}$  at NLO accuracy matched with shower Monte Carlo programs*, *JHEP* **08** (2014) 069, [arXiv:1405.5859 \[hep-ph\]](#). 18, 178
- [356] T. Ježo and P. Nason, *On the Treatment of Resonances in Next-to-Leading Order Calculations Matched to a Parton Shower*, *JHEP* **12** (2015) 065, [arXiv:1509.09071 \[hep-ph\]](#). 18, 55, 179
- [357] T. Ježo, J. M. Lindert, P. Nason, C. Oleari, and S. Pozzorini, *An NLO+PS generator for  $t\bar{t}$  and  $Wt$  production and decay including non-resonant and interference effects*, *Eur. Phys. J.* **C76** (2016) no. 12, 691, [arXiv:1607.04538 \[hep-ph\]](#). 18, 55, 179
- [358] L. Buonocore, P. Nason, and F. Tramontano, *Heavy quark radiation in NLO+PS POWHEG generators*, *Eur. Phys. J.* **C78** (2018) no. 2, 151, [arXiv:1711.06281 \[hep-ph\]](#). 18, 179
- [359] G. Heinrich, A. Maier, R. Nisius, J. Schlenk, M. Schulze, L. Scyboz, and J. Winter, *NLO and off-shell effects in top quark mass determinations*, [arXiv:1709.08615 \[hep-ph\]](#). 18, 179, 180
- [360] M. Beneke, P. Marquard, P. Nason, and M. Steinhauser, *On the ultimate uncertainty of the top quark pole mass*, [arXiv:1605.03609 \[hep-ph\]](#). 18, 178

- [361] M. Butenschoen, B. Dehnadi, A. H. Hoang, V. Mateu, M. Preisser, and I. W. Stewart, *Top Quark Mass Calibration for Monte Carlo Event Generators*, *Phys. Rev. Lett.* **117** (2016) no. 23, 232001, [arXiv:1608.01318 \[hep-ph\]](#). 18, 178
- [362] S. Kawabata and H. Yokoya, *Top-quark mass from the diphoton mass spectrum*, *Eur. Phys. J.* **C77** (2017) no. 5, 323, [arXiv:1607.00990 \[hep-ph\]](#). 18, 178
- [363] A. H. Hoang, C. Lepenik, and M. Preisser, *On the Light Massive Flavor Dependence of the Large Order Asymptotic Behavior and the Ambiguity of the Pole Mass*, *JHEP* **09** (2017) 099, [arXiv:1706.08526 \[hep-ph\]](#). 18, 178
- [364] A. H. Hoang, S. Mantry, A. Pathak, and I. W. Stewart, *Extracting a Short Distance Top Mass with Light Grooming*, [arXiv:1708.02586 \[hep-ph\]](#). 18, 89, 110, 178
- [365] G. Bevilacqua, H. B. Hartanto, M. Kraus, M. Schulze, and M. Worek, *Top quark mass studies with  $t\bar{t}j$  at the LHC*, [arXiv:1710.07515 \[hep-ph\]](#). 18, 19, 178
- [366] G. Corcella, R. Franceschini, and D. Kim, *Fragmentation Uncertainties in Hadronic Observables for Top-quark Mass Measurements*, [arXiv:1712.05801 \[hep-ph\]](#). 18, 179
- [367] S. Ferrario Ravasio, T. Ježo, P. Nason, and C. Oleari, *A Theoretical Study of Top-Mass Measurements at the LHC Using NLO+PS Generators of Increasing Accuracy*, [arXiv:1801.03944 \[hep-ph\]](#). 18, 179
- [368] ATLAS Collaboration, G. Aad et al., *Measurement of the  $t\bar{t}$  production cross-section using  $e\mu$  events with  $b$ -tagged jets in  $pp$  collisions at  $\sqrt{s} = 7$  and 8 TeV with the ATLAS detector*, *Eur. Phys. J.* **C74** (2014) no. 10, 3109, [arXiv:1406.5375 \[hep-ex\]](#). [Addendum: *Eur. Phys. J.*C76,no.11,642(2016)]. 18, 173
- [369] *Measurement of the production cross-section using  $e\mu$  events with  $b$ -tagged jets in  $pp$  collisions at with the ATLAS detector*, *Physics Letters B* **761** (2016) 136 – 157. 18
- [370] CMS Collaboration, V. Khachatryan et al., *Measurement of the  $t$ - $t$ bar production cross section in the  $e$ - $\mu$  channel in proton-proton collisions at  $\sqrt{s} = 7$  and 8 TeV*, *JHEP* **08** (2016) 029, [arXiv:1603.02303 \[hep-ex\]](#). 18
- [371] CMS Collaboration, V. Khachatryan et al., *Measurement of the  $t\bar{t}$  production cross section using events in the  $e\mu$  final state in  $pp$  collisions at  $\sqrt{s} = 13$  TeV*, *Eur. Phys. J.* **C77** (2017) 172, [arXiv:1611.04040 \[hep-ex\]](#). 18
- [372] M. Czakon and A. Mitov, *Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders*, *Comput. Phys. Commun.* **185** (2014) 2930, [arXiv:1112.5675 \[hep-ph\]](#). 18
- [373] CMS Collaboration, V. Khachatryan et al., *Measurement of the differential cross section for top quark pair production in  $pp$  collisions at  $\sqrt{s} = 8$  TeV*, *Eur. Phys. J.* **C75** (2015) no. 11, 542, [arXiv:1505.04480 \[hep-ex\]](#). 19, 173
- [374] ATLAS Collaboration, G. Aad et al., *Measurements of top-quark pair differential cross-sections in the lepton+jets channel in  $pp$  collisions at  $\sqrt{s} = 8$  TeV using the ATLAS detector*, *Eur. Phys. J.* **C76** (2016) no. 10, 538, [arXiv:1511.04716 \[hep-ex\]](#). 19
- [375] S. Dittmaier, P. Uwer, and S. Weinzierl, *NLO QCD corrections to  $t$  anti- $t$  + jet production at hadron colliders*, *Phys. Rev. Lett.* **98** (2007) 262002, [arXiv:hep-ph/0703120 \[HEP-PH\]](#). 19, 173
- [376] K. Melnikov and M. Schulze, *NLO QCD corrections to top quark pair production in association with one hard jet at hadron colliders*, *Nucl. Phys.* **B840** (2010) 129–159, [arXiv:1004.3284 \[hep-ph\]](#). 19
- [377] K. Melnikov, A. Scharf, and M. Schulze, *Top quark pair production in association with a jet: QCD corrections and jet radiation in top quark decays*, *Phys. Rev.* **D85** (2012) 054002, [arXiv:1111.4991 \[hep-ph\]](#). 19

- [378] A. Kardos, C. Papadopoulos, and Z. Trócsányi, *Top quark pair production in association with a jet with NLO parton showering*, *Phys. Lett.* **B705** (2011) 76–81, [arXiv:1101.2672 \[hep-ph\]](#). 19, 173
- [379] S. Alioli, S.-O. Moch, and P. Uwer, *Hadronic top-quark pair-production with one jet and parton showering*, *JHEP* **01** (2012) 137, [arXiv:1110.5251 \[hep-ph\]](#). 19, 173
- [380] G. Bevilacqua, H. B. Hartanto, M. Kraus, and M. Worek, *Top Quark Pair Production in Association with a Jet with Next-to-Leading-Order QCD Off-Shell Effects at the Large Hadron Collider*, *Phys. Rev. Lett.* **116** (2016) no. 5, 052003, [arXiv:1509.09242 \[hep-ph\]](#). 19, 173
- [381] G. Bevilacqua, H. B. Hartanto, M. Kraus, and M. Worek, *Off-shell Top Quarks with One Jet at the LHC: A comprehensive analysis at NLO QCD*, *JHEP* **11** (2016) 098, [arXiv:1609.01659 \[hep-ph\]](#). 19
- [382] J. Fuster, A. Irlles, D. Melini, P. Uwer, and M. Vos, *Extracting the top-quark running mass using  $t\bar{t} + 1$ -jet events produced at the Large Hadron Collider*, *Eur. Phys. J.* **C77** (2017) no. 11, 794, [arXiv:1704.00540 \[hep-ph\]](#). 19
- [383] R. Röntsch and M. Schulze, *Constraining couplings of top quarks to the Z boson in  $t\bar{t} + Z$  production at the LHC*, *JHEP* **07** (2014) 091, [arXiv:1404.1005 \[hep-ph\]](#). [Erratum: *JHEP*09,132(2015)]. 19
- [384] R. Röntsch and M. Schulze, *Probing top-Z dipole moments at the LHC and ILC*, *JHEP* **08** (2015) 044, [arXiv:1501.05939 \[hep-ph\]](#). 19
- [385] H. van Deurzen, R. Frederix, V. Hirschi, G. Luisoni, P. Mastrolia, and G. Ossola, *Spin Polarisation of  $t\bar{t}\gamma\gamma$  production at NLO+PS with GoSam interfaced to MadGraph5\_aMC@NLO*, *Eur. Phys. J.* **C76** (2016) no. 4, 221, [arXiv:1509.02077 \[hep-ph\]](#). 19, 51, 57
- [386] A. Broggio, A. Ferroglia, G. Ossola, and B. D. Pecjak, *Associated production of a top pair and a W boson at next-to-next-to-leading logarithmic accuracy*, *JHEP* **09** (2016) 089, [arXiv:1607.05303 \[hep-ph\]](#). 19
- [387] A. Broggio, A. Ferroglia, G. Ossola, B. D. Pecjak, and R. D. Sameshima, *Associated production of a top pair and a Z boson at the LHC to NNLL accuracy*, *JHEP* **04** (2017) 105, [arXiv:1702.00800 \[hep-ph\]](#). 19
- [388] R. Frederix, D. Pagani, and M. Zaro, *Large NLO corrections in  $t\bar{t}W^\pm$  and  $t\bar{t}t\bar{t}$  hadroproduction from supposedly subleading EW contributions*, *JHEP* **02** (2018) 031, [arXiv:1711.02116 \[hep-ph\]](#). 19, 51
- [389] A. S. Papanastasiou, R. Frederix, S. Frixione, V. Hirschi, and F. Maltoni, *Single-top t-channel production with off-shell and non-resonant effects*, *Phys. Lett.* **B726** (2013) 223–227, [arXiv:1305.7088 \[hep-ph\]](#). 19
- [390] S. Frixione, E. Laenen, P. Motylinski, and B. R. Webber, *Single-top production in MC@NLO*, *JHEP* **03** (2006) 092, [arXiv:hep-ph/0512250 \[hep-ph\]](#). 19
- [391] E. Bothmann, F. Krauss, and M. Schönherr, *Single top-quark production with SHERPA*, [arXiv:1711.02568 \[hep-ph\]](#). 19
- [392] S. Alioli, P. Nason, C. Oleari, and E. Re, *NLO single-top production matched with shower in POWHEG: s- and t-channel contributions*, *JHEP* **09** (2009) 111, [arXiv:0907.4076 \[hep-ph\]](#). [Erratum: *JHEP*02,011(2010)]. 19
- [393] E. Re, *Single-top Wt-channel production matched with parton showers using the POWHEG method*, *Eur. Phys. J.* **C71** (2011) 1547, [arXiv:1009.2450 \[hep-ph\]](#). 19
- [394] E. L. Berger, J. Gao, and H. X. Zhu, *Differential Distributions for t-channel Single Top-Quark Production and Decay at Next-to-Next-to-Leading Order in QCD*, *JHEP* **11** (2017) 158, [arXiv:1708.09405 \[hep-ph\]](#). 19

- [395] S. Alekhin, S. Moch, and S. Thier, *Determination of the top-quark mass from hadro-production of single top-quarks*, *Phys. Lett.* **B763** (2016) 341–346, [arXiv:1608.05212 \[hep-ph\]](#). 19
- [396] T. Martini and P. Uwer, *The Matrix Element Method at next-to-leading order QCD for hadronic collisions: Single top-quark production at the LHC as an example application*, [arXiv:1712.04527 \[hep-ph\]](#). 19
- [397] Z. Bern, L. J. Dixon, F. Febres Cordero, S. Höche, H. Ita, D. A. Kosower, and D. Maitre, *Ntuples for NLO Events at Hadron Colliders*, *Comput. Phys. Commun.* **185** (2014) 1443–1460, [arXiv:1310.7439 \[hep-ph\]](#). 19, 20
- [398] A. Gehrmann-De Ridder, T. Gehrmann, E. W. N. Glover, and G. Heinrich, *Infrared structure of  $e^+e^- \rightarrow 3$  jets at NNLO*, *JHEP* **11** (2007) 058, [arXiv:0710.0346 \[hep-ph\]](#). 19
- [399] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, *Nucl. Instrum. Meth.* **A389** (1997) 81–86. 20
- [400] fastNLO Collaboration, D. Britzger, K. Rabbertz, F. Stober, and M. Wobisch, *New features in version 2 of the fastNLO project*, in *Proceedings, 20th International Workshop on Deep-Inelastic Scattering and Related Subjects (DIS 2012): Bonn, Germany, March 26-30, 2012*, pp. 217–221. 2012. [arXiv:1208.3641 \[hep-ph\]](#). 21, 22
- [401] E. Bothmann, N. Hartland, and S. Schumann, *Introducing MCgrid 2.0: Projecting cross section calculations on grids*, *Comput. Phys. Commun.* **196** (2015) 617–618. 21
- [402] J. Gao, L. Harland-Lang, and J. Rojo, *The Structure of the Proton in the LHC Precision Era*, Submitted to *Phys. Rep.* (2017), [arXiv:1709.04922 \[hep-ph\]](#). 22, 80
- [403] C. Adloff *et al*, H1 Collaboration, *Measurement and QCD analysis of jet cross-sections in deep inelastic positron - proton collisions at  $s^{*(1/2)}$  of 300-GeV*, *Eur. Phys. J.* **C19** (2001) 289–311, [arXiv:hep-ex/0010054 \[hep-ex\]](#). 22
- [404] S. Chekanov *et al*, ZEUS Collaboration, *An NLO QCD analysis of inclusive cross-section and jet-production data from the zeus experiment*, *Eur. Phys. J.* **C42** (2005) 1–16, [arXiv:hep-ph/0503274 \[hep-ph\]](#). 22
- [405] T. Carli, G. P. Salam, and F. Siegert, *A Posteriori inclusion of PDFs in NLO QCD final-state calculations*, in *HERA and the LHC: A Workshop on the Implications of HERA for LHC Physics (Startup Meeting, CERN, 26-27 March 2004; Working Group Meeting, CERN, 17-21 January 2005; Final Meeting 21-24 Mar 2005) CERN, Geneva, Switzerland, October 11-13, 2004*. 2005. [arXiv:hep-ph/0510324 \[hep-ph\]](#). 22
- [406] R. Hamberg, W. L. van Neerven, and T. Matsuura, *A complete calculation of the order  $\alpha_s^2$  correction to the Drell-Yan  $K$  factor*, *Nucl. Phys.* **B359** (1991) 343–405. [Erratum: *Nucl. Phys.* B644,403(2002)]. 22, 34, 35
- [407] C. Anastasiou, L. J. Dixon, K. Melnikov, and F. Petriello, *Dilepton rapidity distribution in the Drell-Yan process at NNLO in QCD*, *Phys. Rev. Lett.* **91** (2003) 182002, [arXiv:hep-ph/0306192 \[hep-ph\]](#). 22
- [408] S. Catani, G. Ferrera, and M. Grazzini, *W Boson Production at Hadron Colliders: The Lepton Charge Asymmetry in NNLO QCD*, *JHEP* **05** (2010) 006, [arXiv:1002.3115 \[hep-ph\]](#). 22
- [409] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, *W Physics at the LHC with FEWZ 2.1*, *Comput. Phys. Commun.* **184** (2013) 208–214, [arXiv:1201.5896 \[hep-ph\]](#). 22
- [410] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, *FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order*, *Comput. Phys. Commun.* **182** (2011) 2388–2403, [arXiv:1011.3540 \[hep-ph\]](#). 22
- [411] C. Anastasiou, L. J. Dixon, K. Melnikov, and F. Petriello, *High precision QCD at*

- hadron colliders: Electroweak gauge boson rapidity distributions at NNLO*, *Phys. Rev.* **D69** (2004) 094008, [arXiv:hep-ph/0312266](#) [[hep-ph](#)]. 22
- [412] V. Bertone, R. Frederix, S. Frixione, J. Rojo, and M. Sutton, *aMCfast: automation of fast NLO computations for PDF fits*, *JHEP* **08** (2014) 166, [arXiv:1406.7693](#) [[hep-ph](#)]. 23
- [413] T. A. Morgan, *Precision Z Boson Phenomenology at the LHC*. PhD thesis, 2016, University of Durham, ISNI: 0000 0004 5993 8474, [[etheses.dur.ac.uk/11773](#)]. [http://etheses.dur.ac.uk/11773](#). 25
- [414] A. Gehrmann-De Ridder, T. Gehrmann, N. Glover, A. Huss, and T. A. Morgan, *NNLO QCD corrections for Z boson plus jet production*, *PoS RADCOR2015* (2016) 075, [arXiv:1601.04569](#) [[hep-ph](#)]. 25
- [415] T. Gleisberg, S. Höche, F. Krauss, M. Schönherr, S. Schumann, F. Siegert, and J. Winter, *Event generation with SHERPA 1.1*, *JHEP* **02** (2009) 007, [arXiv:0811.4622](#) [[hep-ph](#)]. 28, 35, 51, 53, 57, 88, 150, 178, 179, 187, 204, 209, 218, 241
- [416] Z. Nagy and Z. Trócsányi, *Next-to-leading order calculation of four jet observables in electron positron annihilation*, *Phys. Rev.* **D59** (1999) 014020, [arXiv:hep-ph/9806317](#) [[hep-ph](#)]. [Erratum: *Phys. Rev.* **D62** (2001) 099902]. 28
- [417] Z. Nagy, *Three jet cross-sections in hadron hadron collisions at next-to-leading order*, *Phys. Rev. Lett.* **88** (2002) 122003, [arXiv:hep-ph/0110315](#) [[hep-ph](#)]. 28, 89
- [418] J. M. Campbell and R. K. Ellis, *An Update on vector boson pair production at hadron colliders*, *Phys. Rev.* **D60** (1999) 113006, [arXiv:hep-ph/9905386](#) [[hep-ph](#)]. 28, 94
- [419] J. M. Campbell and R. K. Ellis, *Radiative corrections to Z b anti-b production*, *Phys. Rev.* **D62** (2000) 114012, [arXiv:hep-ph/0006304](#) [[hep-ph](#)]. 28
- [420] J. M. Campbell and R. K. Ellis, *MCFM for the Tevatron and the LHC*, *Nucl. Phys. Proc. Suppl.* **205-206** (2010) 10–15, [arXiv:1007.3492](#) [[hep-ph](#)]. 28, 94
- [421] S. Frixione, *A General approach to jet cross-sections in QCD*, *Nucl. Phys.* **B507** (1997) 295–314, [arXiv:hep-ph/9706545](#) [[hep-ph](#)]. 30, 51
- [422] NNPDF Collaboration, R. D. Ball et al., *Parton distributions for the LHC Run II*, *JHEP* **04** (2015) 040, [arXiv:1410.8849](#) [[hep-ph](#)]. 34, 82, 180, 187, 219, 241, 247
- [423] S. Catani, T. Gleisberg, F. Krauss, G. Rodrigo, and J.-C. Winter, *From loops to trees by-passing Feynman’s theorem*, *JHEP* **09** (2008) 065, [arXiv:0804.3170](#) [[hep-ph](#)]. 35, 36, 44, 45
- [424] I. Bierenbaum, S. Catani, P. Draggiotis, and G. Rodrigo, *A Tree-Loop Duality Relation at Two Loops and Beyond*, *JHEP* **10** (2010) 073, [arXiv:1007.0194](#) [[hep-ph](#)]. 35, 42
- [425] I. Bierenbaum, S. Buchta, P. Draggiotis, I. Malamos, and G. Rodrigo, *Tree-Loop Duality Relation beyond simple poles*, *JHEP* **03** (2013) 025, [arXiv:1211.5048](#) [[hep-ph](#)]. 35, 36, 44, 45
- [426] I. Bierenbaum, P. Draggiotis, S. Buchta, G. Chachamis, I. Malamos, and G. Rodrigo, *News on the Loop-tree Duality*, *Acta Phys. Polon.* **B44** (2013) 2207–2213. 35
- [427] S. Buchta, G. Chachamis, P. Draggiotis, I. Malamos, and G. Rodrigo, *On the singular behaviour of scattering amplitudes in quantum field theory*, *JHEP* **11** (2014) 014, [arXiv:1405.7850](#) [[hep-ph](#)]. 35, 39, 44, 45
- [428] S. Buchta, G. Chachamis, I. Malamos, I. Bierenbaum, P. Draggiotis, and G. Rodrigo, *The loop-tree duality at work*, *PoS LL2014* (2014) 066, [arXiv:1407.5865](#) [[hep-ph](#)]. 35
- [429] S. Buchta, G. Chachamis, P. Draggiotis, I. Malamos, and G. Rodrigo, *Towards a Numerical Implementation of the Loop-Tree Duality Method*, *Nucl. Part. Phys. Proc.* **258-259** (2015) 33–36, [arXiv:1509.07386](#) [[hep-ph](#)]. 35

- [430] G. F. R. Sborlini, R. Hernández-Pinto, and G. Rodrigo, *From dimensional regularization to NLO computations in four dimensions*, PoS **EPS-HEP2015** (2015) 479, [arXiv:1510.01079 \[hep-ph\]](#). 35
- [431] S. Buchta, G. Chachamis, P. Draggiotis, and G. Rodrigo, *Numerical implementation of the loop–tree duality method*, *Eur. Phys. J.* **C77** (2017) no. 5, 274, [arXiv:1510.00187 \[hep-ph\]](#). 35, 36, 37, 44, 45
- [432] S. Buchta, G. Chachamis, I. Malamos, I. Bierenbaum, P. Draggiotis, and G. Rodrigo, *The loop-tree duality at NLO and beyond*, *Nucl. Part. Phys. Proc.* **273-275** (2016) 2009–2014. 35
- [433] R. J. Hernandez-Pinto, G. F. R. Sborlini, and G. Rodrigo, *Towards gauge theories in four dimensions*, *JHEP* **02** (2016) 044, [arXiv:1506.04617 \[hep-ph\]](#). 35, 44, 45
- [434] G. F. R. Sborlini, F. Driencourt-Mangin, R. Hernandez-Pinto, and G. Rodrigo, *Four-dimensional unsubtraction from the loop-tree duality*, *JHEP* **08** (2016) 160, [arXiv:1604.06699 \[hep-ph\]](#). 35, 38, 40, 41, 44, 45
- [435] G. F. R. Sborlini, F. Driencourt-Mangin, and G. Rodrigo, *Four-dimensional unsubtraction with massive particles*, *JHEP* **10** (2016) 162, [arXiv:1608.01584 \[hep-ph\]](#). 35, 36, 38, 40, 41
- [436] F. Driencourt-Mangin, G. Rodrigo, and G. F. R. Sborlini, *Universal dual amplitudes and asymptotic expansions for  $gg \rightarrow H$  and  $H \rightarrow \gamma\gamma$  in four dimensions*, [arXiv:1702.07581 \[hep-ph\]](#). 35, 43
- [437] G. Chachamis, S. Buchta, P. Draggiotis, and G. Rodrigo, *Attacking One-loop Multi-leg Feynman Integrals with the Loop-Tree Duality*, PoS **DIS2016** (2016) 067, [arXiv:1607.00875 \[hep-ph\]](#). 35
- [438] G. Chachamis and G. Rodrigo, *The Loop-Tree Duality: Progress Report*, PoS **DIS2017** (2018) 178, [arXiv:1709.02646 \[hep-ph\]](#). 35, 37
- [439] N. Selomit Ramírez-Uribe, R. J. Hernández-Pinto, and G. Rodrigo, *QED and QCD self-energy corrections through the loop-tree duality*, *J. Phys. Conf. Ser.* **912** (2017) no. 1, 012013, [arXiv:1709.07802 \[hep-ph\]](#). 35
- [440] J. L. Jurado, G. Rodrigo, and W. J. Torres Bobadilla, *From Jacobi off-shell currents to integral relations*, *JHEP* **12** (2017) 122, [arXiv:1710.11010 \[hep-ph\]](#). 35
- [441] D. E. Soper, *QCD calculations by numerical integration*, *Phys. Rev. Lett.* **81** (1998) 2638–2641, [arXiv:hep-ph/9804454 \[hep-ph\]](#). 35, 44
- [442] D. E. Soper, *Techniques for QCD calculations by numerical integration*, *Phys. Rev.* **D62** (2000) 014009, [arXiv:hep-ph/9910292 \[hep-ph\]](#). 35, 44
- [443] D. E. Soper, *Choosing integration points for QCD calculations by numerical integration*, *Phys. Rev.* **D64** (2001) 034018, [arXiv:hep-ph/0103262 \[hep-ph\]](#). 35, 44
- [444] M. Krämer and D. E. Soper, *Next-to-leading order numerical calculations in Coulomb gauge*, *Phys. Rev.* **D66** (2002) 054017, [arXiv:hep-ph/0204113 \[hep-ph\]](#). 35, 44
- [445] A. Ferroglia, M. Passera, G. Passarino, and S. Uccirati, *All purpose numerical evaluation of one loop multileg Feynman diagrams*, *Nucl. Phys.* **B650** (2003) 162–228, [arXiv:hep-ph/0209219 \[hep-ph\]](#). 35
- [446] Z. Nagy and D. E. Soper, *General subtraction method for numerical calculation of one loop QCD matrix elements*, *JHEP* **09** (2003) 055, [arXiv:hep-ph/0308127 \[hep-ph\]](#). 35, 44
- [447] Z. Nagy and D. E. Soper, *Numerical integration of one-loop Feynman diagrams for N-photon amplitudes*, *Phys. Rev.* **D74** (2006) 093006, [arXiv:hep-ph/0610028 \[hep-ph\]](#). 35

- [448] M. Moretti, F. Piccinini, and A. D. Polosa, *A Fully Numerical Approach to One-Loop Amplitudes*, [arXiv:0802.4171 \[hep-ph\]](#). 35
- [449] W. Gong, Z. Nagy, and D. E. Soper, *Direct numerical integration of one-loop Feynman diagrams for N-photon amplitudes*, *Phys. Rev.* **D79** (2009) 033005, [arXiv:0812.3686 \[hep-ph\]](#). 35, 44
- [450] W. Kilian and T. Kleinschmidt, *Numerical Evaluation of Feynman Loop Integrals by Reduction to Tree Graphs*, [arXiv:0912.3495 \[hep-ph\]](#). 35
- [451] S. Becker, C. Reuschle, and S. Weinzierl, *Numerical NLO QCD calculations*, *JHEP* **12** (2010) 013, [arXiv:1010.4187 \[hep-ph\]](#). 35, 44, 49
- [452] SM and NLO Multileg Working Group Collaboration, T. Binoth et al., *The SM and NLO Multileg Working Group: Summary report*, in *Physics at TeV colliders. Proceedings, 6th Workshop, dedicated to Thomas Binoth, Les Houches, France, June 8-26, 2009*, pp. 21–189. 2010. [arXiv:1003.1241 \[hep-ph\]](#). 35, 144
- [453] S. Becker, C. Reuschle, and S. Weinzierl, *Efficiency Improvements for the Numerical Computation of NLO Corrections*, *JHEP* **07** (2012) 090, [arXiv:1205.2096 \[hep-ph\]](#). 35
- [454] A. De Roeck et al., *From the LHC to Future Colliders*, *Eur. Phys. J.* **C66** (2010) 525–583, [arXiv:0909.3240 \[hep-ph\]](#). 35
- [455] SM AND NLO MULTILEG and SM MC Working Groups Collaboration, J. Alcaraz Maestre et al., *The SM and NLO Multileg and SM MC Working Groups: Summary Report*, [arXiv:1203.6803 \[hep-ph\]](#). 35, 131
- [456] S. Becker and S. Weinzierl, *Direct contour deformation with arbitrary masses in the loop*, *Phys. Rev.* **D86** (2012) 074009, [arXiv:1208.4088 \[hep-ph\]](#). 35, 44
- [457] S. Seth and S. Weinzierl, *Numerical integration of subtraction terms*, *Phys. Rev.* **D93** (2016) no. 11, 114031, [arXiv:1605.06646 \[hep-ph\]](#). 35, 41, 44
- [458] C. Gnendiger et al., *To d, or not to d: recent developments and comparisons of regularization schemes*, *Eur. Phys. J.* **C77** (2017) no. 7, 471, [arXiv:1705.01827 \[hep-ph\]](#). 35, 44
- [459] S. Frixione and B. R. Webber, *The MC and NLO 3.4 Event Generator*, [arXiv:0812.0770 \[hep-ph\]](#). 35
- [460] G. Passarino, *An Approach toward the numerical evaluation of multiloop Feynman diagrams*, *Nucl. Phys.* **B619** (2001) 257–312, [arXiv:hep-ph/0108252 \[hep-ph\]](#). 35
- [461] C. Anastasiou, S. Beerli, and A. Daleo, *Evaluating multi-loop Feynman diagrams with infrared and threshold singularities numerically*, *JHEP* **05** (2007) 071, [arXiv:hep-ph/0703282 \[hep-ph\]](#). 35
- [462] S. Becker and S. Weinzierl, *Direct numerical integration for multi-loop integrals*, *Eur. Phys. J.* **C73** (2013) no. 2, 2321, [arXiv:1211.0509 \[hep-ph\]](#). 35, 44
- [463] R. P. Feynman, *Quantum theory of gravitation*, *Acta Phys. Polon.* **24** (1963) 697–722. 36
- [464] T. Hahn and M. Perez-Victoria, *Automatized one loop calculations in four-dimensions and D-dimensions*, *Comput. Phys. Commun.* **118** (1999) 153–165, [arXiv:hep-ph/9807565 \[hep-ph\]](#). 37, 49
- [465] T. Hahn, *CUBA: A Library for multidimensional numerical integration*, *Comput. Phys. Commun.* **168** (2005) 78–95, [arXiv:hep-ph/0404043 \[hep-ph\]](#). 37
- [466] S. Catani and M. H. Seymour, *The Dipole formalism for the calculation of QCD jet cross-sections at next-to-leading order*, *Phys. Lett.* **B378** (1996) 287–301, [arXiv:hep-ph/9602277 \[hep-ph\]](#). 39

- [467] S. Catani, D. de Florian, and G. Rodrigo, *Space-like (versus time-like) collinear limits in QCD: Is factorization violated?*, *JHEP* **07** (2012) 026, [arXiv:1112.4405 \[hep-ph\]](#). 39
- [468] M. Assadsolimani, S. Becker, and S. Weinzierl, *A simple formula for the infrared singular part of the integrand of one-loop QCD amplitudes*, *Phys. Rev.* **D81** (2010) 094002, [arXiv:0912.1680 \[hep-ph\]](#). 44
- [469] M. Assadsolimani, S. Becker, C. Reuschle, and S. Weinzierl, *Infrared singularities in one-loop amplitudes*, *Nucl. Phys. Proc. Suppl.* **205-206** (2010) 224–229, [arXiv:1006.4609 \[hep-ph\]](#). 44
- [470] S. Becker, D. Götz, C. Reuschle, C. Schwan, and S. Weinzierl, *NLO results for five, six and seven jets in electron-positron annihilation*, *Phys. Rev. Lett.* **108** (2012) 032005, [arXiv:1111.1733 \[hep-ph\]](#). 44
- [471] S. Becker, C. Reuschle, and S. Weinzierl, *Efficiency improvements for the numerical computation of NLO corrections*, *JHEP* **1207** (2012) 090, [arXiv:1205.2096 \[hep-ph\]](#). 44
- [472] D. Götz, C. Reuschle, C. Schwan, and S. Weinzierl, *NLO corrections to Z production in association with several jets*, *PoS LL2014* (2014) 009, [arXiv:1407.0203 \[hep-ph\]](#). 44
- [473] R. Pittau, *A four-dimensional approach to quantum field theories*, *JHEP* **11** (2012) 151, [arXiv:1208.5457 \[hep-ph\]](#). 44
- [474] R. Pittau, *QCD corrections to  $H \rightarrow gg$  in FDR*, *Eur. Phys. J.* **C74** (2014) no. 1, 2686, [arXiv:1307.0705 \[hep-ph\]](#). 44
- [475] A. M. Donati and R. Pittau, *FDR, an easier way to NNLO calculations: a two-loop case study*, *Eur. Phys. J.* **C74** (2014) 2864, [arXiv:1311.3551 \[hep-ph\]](#). 44
- [476] B. Page and R. Pittau, *Two-loop off-shell QCD amplitudes in FDR*, *JHEP* **11** (2015) 183, [arXiv:1506.09093 \[hep-ph\]](#). 44
- [477] K. Arnold et al., *VBFNLO: A Parton level Monte Carlo for processes with electroweak bosons*, *Comput. Phys. Commun.* **180** (2009) 1661–1670, [arXiv:0811.4559 \[hep-ph\]](#). 49, 140, 186, 187
- [478] S. Badger, B. Biedermann, and P. Uwer, *NGluon: A Package to Calculate One-loop Multi-gluon Amplitudes*, *Comput. Phys. Commun.* **182** (2011) 1674–1692, [arXiv:1011.2900 \[hep-ph\]](#). 49
- [479] V. Hirschi, R. Frederix, S. Frixione, M. V. Garzelli, F. Maltoni, and R. Pittau, *Automation of one-loop QCD corrections*, *JHEP* **05** (2011) 044, [arXiv:1103.0621 \[hep-ph\]](#). 49, 51, 55, 56
- [480] G. Cullen, N. Greiner, G. Heinrich, G. Luisoni, P. Mastrolia, G. Ossola, T. Reiter, and F. Tramontano, *Automated One-Loop Calculations with GoSam*, *Eur. Phys. J.* **C72** (2012) 1889, [arXiv:1111.2034 \[hep-ph\]](#). 49, 56
- [481] S. Actis, A. Denner, L. Hofer, A. Scharf, and S. Uccirati, *Recursive generation of one-loop amplitudes in the Standard Model*, *JHEP* **04** (2013) 037, [arXiv:1211.6316 \[hep-ph\]](#). 49, 52, 53, 187
- [482] M. Chiesa, N. Greiner, and F. Tramontano, *Automation of electroweak corrections for LHC processes*, *J. Phys.* **G43** (2016) no. 1, 013002, [arXiv:1507.08579 \[hep-ph\]](#). 49
- [483] A. Denner, S. Dittmaier, M. Roth, and D. Wackerroth, *Predictions for all processes  $e^+e^- \rightarrow \text{fermions} + \gamma$* , *Nucl. Phys.* **B560** (1999) 33–65, [arXiv:hep-ph/9904472 \[hep-ph\]](#). 50, 51, 52, 59, 187
- [484] S. Dittmaier and M. Roth, *LUSIFER: A LUCid approach to six FERmion production*, *Nucl. Phys.* **B642** (2002) 307–343, [arXiv:hep-ph/0206070 \[hep-ph\]](#). 50, 187
- [485] S. Dittmaier, *A General approach to photon radiation off fermions*, *Nucl. Phys.* **B565**

- (2000) 69–122, [arXiv:hep-ph/9904440 \[hep-ph\]](#). 50
- [486] S. Dittmaier, A. Kabelschacht, and T. Kasprzik, *Polarized QED splittings of massive fermions and dipole subtraction for non-collinear-safe observables*, *Nucl. Phys.* **B800** (2008) 146–189, [arXiv:0802.1405 \[hep-ph\]](#). 50
- [487] F. A. Berends, R. Pittau, and R. Kleiss, *All electroweak four fermion processes in electron - positron collisions*, *Nucl. Phys.* **B424** (1994) 308–342, [arXiv:hep-ph/9404313 \[hep-ph\]](#). 50, 187
- [488] A. Denner, L. Hošeková, and S. Kallweit, *NLO QCD corrections to  $W^+W^+jj$  production in vector-boson fusion at the LHC*, *Phys. Rev.* **D86** (2012) 114014, [arXiv:1209.2389 \[hep-ph\]](#). 50
- [489] S. Catani, S. Dittmaier, M. H. Seymour, and Z. Trócsányi, *The Dipole formalism for next-to-leading order QCD calculations with massive partons*, *Nucl. Phys.* **B627** (2002) 189–265, [arXiv:hep-ph/0201036 \[hep-ph\]](#). 50, 51, 186, 187, 209, 218
- [490] T. Gleisberg, S. Höche, F. Krauss, A. Schälicke, S. Schumann, and J.-C. Winter, *SHERPA 1.α: A Proof of concept version*, *JHEP* **02** (2004) 056, [arXiv:hep-ph/0311263 \[hep-ph\]](#). 51, 53, 187, 209
- [491] S. Schumann and F. Krauss, *A Parton shower algorithm based on Catani-Seymour dipole factorisation*, *JHEP* **03** (2008) 038, [arXiv:0709.1027 \[hep-ph\]](#). 51, 157, 187, 209, 218, 241, 246
- [492] S. Höche and S. Prestel, *The midpoint between dipole and parton showers*, *Eur. Phys. J.* **C75** (2015) no. 9, 461, [arXiv:1506.05057 \[hep-ph\]](#). 51, 157, 217, 218, 219, 225
- [493] S. Höche, S. Kuttimalai, S. Schumann, and F. Siegert, *Beyond Standard Model calculations with Sherpa*, *Eur. Phys. J.* **C75** (2015) no. 3, 135, [arXiv:1412.6478 \[hep-ph\]](#). 51, 176, 179
- [494] F. Krauss, R. Kuhn, and G. Soff, *AMEGIC++ 1.0: A Matrix element generator in C++*, *JHEP* **02** (2002) 044, [arXiv:hep-ph/0109036 \[hep-ph\]](#). 51, 187
- [495] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, *JHEP* **12** (2008) 039, [arXiv:0808.3674 \[hep-ph\]](#). 51, 187
- [496] T. Gleisberg and F. Krauss, *Automating dipole subtraction for QCD NLO calculations*, *Eur. Phys. J.* **C53** (2008) 501–523, [arXiv:0709.2881 \[hep-ph\]](#). 51, 53, 187
- [497] M. Schönherr, *An automated subtraction of NLO EW infrared divergences*, *Eur. Phys. J.* **C78** (2018) no. 2, 119, [arXiv:1712.07975 \[hep-ph\]](#). 51, 53, 173
- [498] M. Chiesa, N. Greiner, M. Schönherr, and F. Tramontano, *Electroweak corrections to diphoton plus jets*, *JHEP* **10** (2017) 181, [arXiv:1706.09022 \[hep-ph\]](#). 51, 57
- [499] B. Biedermann, S. Bräuer, A. Denner, M. Pellen, S. Schumann, and J. M. Thompson, *Automation of NLO QCD and EW corrections with Sherpa and Recola*, *Eur. Phys. J.* **C77** (2017) 492, [arXiv:1704.05783 \[hep-ph\]](#). 51, 53, 58, 59, 60, 67, 187
- [500] A. Buckley, J. Butterworth, L. Lönnblad, D. Grellscheid, H. Hoeth, J. Monk, H. Schulz, and F. Siegert, *Rivet user manual*, *Comput. Phys. Commun.* **184** (2013) 2803–2819, [arXiv:1003.0694 \[hep-ph\]](#). 51, 189, 207, 219, 225, 241
- [501] E. Bothmann, M. Schönherr, and S. Schumann, *Reweighting QCD matrix-element and parton-shower calculations*, *Eur. Phys. J.* **C76** (2016) no. 11, 590, [arXiv:1606.08753 \[hep-ph\]](#). 51, 215, 224, 225, 241
- [502] V. Hirschi and O. Mattelaer, *Automated event generation for loop-induced processes*, *JHEP* **10** (2015) 146, [arXiv:1507.00020 \[hep-ph\]](#). 51, 56, 150
- [503] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, *JHEP* **06** (2002) 029, [arXiv:hep-ph/0204244 \[hep-ph\]](#). 51, 151, 241

- [504] R. Frederix and S. Frixione, *Merging meets matching in MC@NLO*, **JHEP** **12** (2012) 061, [arXiv:1209.6215 \[hep-ph\]](#). 51
- [505] T. Binoth et al., *A Proposal for a standard interface between Monte Carlo tools and one-loop programs*, **Comput. Phys. Commun.** **181** (2010) 1612–1622, [arXiv:1001.1307 \[hep-ph\]](#). [1(2010)]. 51, 57, 186, 187
- [506] S. Alioli et al., *Update of the Binoth Les Houches Accord for a standard interface between Monte Carlo tools and one-loop programs*, **Comput. Phys. Commun.** **185** (2014) 560–571, [arXiv:1308.3462 \[hep-ph\]](#). 51, 57, 186, 187, 208
- [507] M. Czakon, D. Heymes, A. Mitov, D. Pagani, I. Tsirikos, and M. Zaro, *The top-quark charge asymmetry at the LHC and Tevatron through NNLO QCD and NLO EW*, [arXiv:1711.03945 \[hep-ph\]](#). 51
- [508] R. Frederix, S. Frixione, V. Hirschi, D. Pagani, H.-S. Shao, and M. Zaro, *The automation of next-to-leading order electroweak calculations*, in preparation . 51, 63
- [509] C. Degrande, C. Duhr, B. Fuks, D. Grellscheid, O. Mattelaer, and T. Reiter, *UFO - The Universal FeynRules Output*, **Comput. Phys. Commun.** **183** (2012) 1201–1214, [arXiv:1108.2040 \[hep-ph\]](#). 51, 57
- [510] G. Ossola, C. G. Papadopoulos, and R. Pittau, *On the Rational Terms of the one-loop amplitudes*, **JHEP** **05** (2008) 004, [arXiv:0802.1876 \[hep-ph\]](#). 51, 57
- [511] P. Draggiotis, M. V. Garzelli, C. G. Papadopoulos, and R. Pittau, *Feynman Rules for the Rational Part of the QCD 1-loop amplitudes*, **JHEP** **04** (2009) 072, [arXiv:0903.0356 \[hep-ph\]](#). 51
- [512] M. V. Garzelli, I. Malamos, and R. Pittau, *Feynman rules for the rational part of the Electroweak 1-loop amplitudes*, **JHEP** **01** (2010) 040, [arXiv:0910.3130 \[hep-ph\]](#). [Erratum: JHEP10,097(2010)]. 51, 54
- [513] M. V. Garzelli, I. Malamos, and R. Pittau, *Feynman rules for the rational part of the Electroweak 1-loop amplitudes in the  $R_{xi}$  gauge and in the Unitary gauge*, **JHEP** **01** (2011) 029, [arXiv:1009.4302 \[hep-ph\]](#). 51
- [514] H.-S. Shao, Y.-J. Zhang, and K.-T. Chao, *Feynman Rules for the Rational Part of the Standard Model One-loop Amplitudes in the 't Hooft-Veltman  $\gamma_5$  Scheme*, **JHEP** **09** (2011) 048, [arXiv:1106.5030 \[hep-ph\]](#). 51
- [515] S. Dittmaier and M. Krämer, *Electroweak radiative corrections to W boson production at hadron colliders*, **Phys. Rev.** **D65** (2002) 073007, [arXiv:hep-ph/0109062 \[hep-ph\]](#). 51
- [516] A. Denner, *Techniques for calculation of electroweak radiative corrections at the one loop level and results for W physics at LEP-200*, **Fortsch. Phys.** **41** (1993) 307–420, [arXiv:0709.1075 \[hep-ph\]](#). 51, 54, 57
- [517] A. Denner, S. Dittmaier, M. Roth, and L. H. Wieders, *Electroweak corrections to charged-current  $e^+e^- \rightarrow 4$  fermion processes: Technical details and further results*, **Nucl. Phys.** **B724** (2005) 247–294, [arXiv:hep-ph/0505042 \[hep-ph\]](#). [Erratum: Nucl. Phys.B854,504(2012)]. 51, 52, 54, 57, 59, 187
- [518] P. Artoisenet et al., *A framework for Higgs characterisation*, **JHEP** **11** (2013) 043, [arXiv:1306.6464 \[hep-ph\]](#). 52
- [519] F. Maltoni, K. Mawatari, and M. Zaro, *Higgs characterisation via vector-boson fusion and associated production: NLO and parton-shower effects*, **Eur. Phys. J.** **C74** (2014) no. 1, 2710, [arXiv:1311.1829 \[hep-ph\]](#). 52
- [520] C. Degrande, F. Maltoni, J. Wang, and C. Zhang, *Automatic computations at next-to-leading order in QCD for top-quark flavor-changing neutral processes*, **Phys. Rev.** **D91** (2015) 034024, [arXiv:1412.5594 \[hep-ph\]](#). 52, 56

- [521] F. Demartin, F. Maltoni, K. Mawatari, B. Page, and M. Zaro, *Higgs characterisation at NLO in QCD: CP properties of the top-quark Yukawa interaction*, *Eur. Phys. J.* **C74** (2014) no. 9, 3065, [arXiv:1407.5089 \[hep-ph\]](#). 52
- [522] C. Degrande, B. Fuks, V. Hirschi, J. Proudom, and H.-S. Shao, *Automated next-to-leading order predictions for new physics at the LHC: the case of colored scalar pair production*, *Phys. Rev.* **D91** (2015) no. 9, 094005, [arXiv:1412.5589 \[hep-ph\]](#). 52, 56
- [523] G. Durieux, F. Maltoni, and C. Zhang, *Global approach to top-quark flavor-changing interactions*, *Phys. Rev.* **D91** (2015) no. 7, 074017, [arXiv:1412.7166 \[hep-ph\]](#). 52
- [524] C. Degrande, B. Fuks, V. Hirschi, J. Proudom, and H.-S. Shao, *Matching next-to-leading order predictions to parton showers in supersymmetric QCD*, *Phys. Lett.* **B755** (2016) 82–87, [arXiv:1510.00391 \[hep-ph\]](#). 52, 56
- [525] M. Backović, M. Krämer, F. Maltoni, A. Martini, K. Mawatari, and M. Pellen, *Higher-order QCD predictions for dark matter production at the LHC in simplified models with s-channel mediators*, *Eur. Phys. J.* **C75** (2015) no. 10, 482, [arXiv:1508.05327 \[hep-ph\]](#). 52, 56
- [526] C. Degrande, M. Ubiali, M. Wiesemann, and M. Zaro, *Heavy charged Higgs boson production at the LHC*, *JHEP* **10** (2015) 145, [arXiv:1507.02549 \[hep-ph\]](#). 52
- [527] C. Degrande, K. Hartling, H. E. Logan, A. D. Peterson, and M. Zaro, *Automatic predictions in the Georgi-Machacek model at next-to-leading order accuracy*, *Phys. Rev.* **D93** (2016) no. 3, 035004, [arXiv:1512.01243 \[hep-ph\]](#). 52
- [528] O. Mattelaer and E. Vryonidou, *Dark matter production through loop-induced processes at the LHC: the s-channel mediator case*, *Eur. Phys. J.* **C75** (2015) no. 9, 436, [arXiv:1508.00564 \[hep-ph\]](#). 52, 56
- [529] M. Neubert, J. Wang, and C. Zhang, *Higher-Order QCD Predictions for Dark Matter Production in Mono-Z Searches at the LHC*, *JHEP* **02** (2016) 082, [arXiv:1509.05785 \[hep-ph\]](#). 52, 56
- [530] D. Buarque Franzosi and C. Zhang, *Probing the top-quark chromomagnetic dipole moment at next-to-leading order in QCD*, *Phys. Rev.* **D91** (2015) no. 11, 114010, [arXiv:1503.08841 \[hep-ph\]](#). 52, 56
- [531] G. Das, C. Degrande, V. Hirschi, F. Maltoni, and H.-S. Shao, *NLO predictions for the production of a spin-two particle at the LHC*, *Phys. Lett.* **B770** (2017) 507–513, [arXiv:1605.09359 \[hep-ph\]](#). 52, 56
- [532] B. Fuks and H.-S. Shao, *QCD next-to-leading-order predictions matched to parton showers for vector-like quark models*, *Eur. Phys. J.* **C77** (2017) no. 2, 135, [arXiv:1610.04622 \[hep-ph\]](#). 52, 56
- [533] C. Degrande, R. Frederix, V. Hirschi, M. Ubiali, M. Wiesemann, and M. Zaro, *Accurate predictions for charged Higgs production: Closing the  $m_{H^\pm} \sim m_t$  window*, *Phys. Lett.* **B772** (2017) 87–92, [arXiv:1607.05291 \[hep-ph\]](#). 52
- [534] C. Arina et al., *A comprehensive approach to dark matter studies: exploration of simplified top-philic models*, *JHEP* **11** (2016) 111, [arXiv:1605.09242 \[hep-ph\]](#). 52
- [535] O. Bessidskaia Bylund, F. Maltoni, I. Tsirikos, E. Vryonidou, and C. Zhang, *Probing top quark neutral couplings in the Standard Model Effective Field Theory at NLO in QCD*, *JHEP* **05** (2016) 052, [arXiv:1601.08193 \[hep-ph\]](#). 52
- [536] C. Zhang, *Single Top Production at Next-to-Leading Order in the Standard Model Effective Field Theory*, *Phys. Rev. Lett.* **116** (2016) no. 16, 162002, [arXiv:1601.06163 \[hep-ph\]](#). 52, 56
- [537] C. Degrande, O. Mattelaer, R. Ruiz, and J. Turner, *Fully-Automated Precision*

- Predictions for Heavy Neutrino Production Mechanisms at Hadron Colliders*, *Phys. Rev. D* **94** (2016) no. 5, 053002, [arXiv:1602.06957 \[hep-ph\]](#). 52
- [538] N. D. Christensen and C. Duhr, *FeynRules - Feynman rules made easy*, *Comput. Phys. Commun.* **180** (2009) 1614–1641, [arXiv:0806.4194 \[hep-ph\]](#). 52
- [539] A. Alloul, N. D. Christensen, C. Degrande, C. Duhr, and B. Fuks, *FeynRules 2.0 - A complete toolbox for tree-level phenomenology*, *Comput. Phys. Commun.* **185** (2014) 2250–2300, [arXiv:1310.1921 \[hep-ph\]](#). 52
- [540] C. Degrande, *Automatic evaluation of UV and R2 terms for beyond the Standard Model Lagrangians: a proof-of-principle*, *Comput. Phys. Commun.* **197** (2015) 239–262, [arXiv:1406.3030 \[hep-ph\]](#). 52
- [541] A. Denner and S. Dittmaier, *The Complex-mass scheme for perturbative calculations with unstable particles*, *Nucl. Phys. Proc. Suppl.* **160** (2006) 22–26, [arXiv:hep-ph/0605312 \[hep-ph\]](#). [,22(2006)]. 52, 59, 63
- [542] A. Denner, J.-N. Lang, and S. Uccirati, *NLO electroweak corrections in extended Higgs Sectors with RECOLA2*, *JHEP* **07** (2017) 087, [arXiv:1705.06053 \[hep-ph\]](#). 52
- [543] A. Denner, J.-N. Lang, and S. Uccirati, *Recola2: REcursive Computation of One-Loop Amplitudes 2*, *Comput. Phys. Commun.* **224** (2018) 346–361, [arXiv:1711.07388 \[hep-ph\]](#). 52
- [544] F. A. Berends and W. T. Giele, *Recursive Calculations for Processes with n Gluons*, *Nucl. Phys.* **B306** (1988) 759–808. 52
- [545] A. van Hameren, *Multi-gluon one-loop amplitudes using tensor integrals*, *JHEP* **07** (2009) 088, [arXiv:0905.1005 \[hep-ph\]](#). 52
- [546] A. Denner, S. Dittmaier, and L. Hofer, *COLLIER - A fortran-library for one-loop integrals*, *PoS LL2014* (2014) 071, [arXiv:1407.0087 \[hep-ph\]](#). 53, 54
- [547] A. Denner, S. Dittmaier, and L. Hofer, *COLLIER: a fortran-based Complex One-Loop Library in Extended Regularizations*, *Comput. Phys. Commun.* **212** (2017) 220–238, [arXiv:1604.06792 \[hep-ph\]](#). 53, 56
- [548] D. Buarque Franzosi, F. Fabbri, and S. Schumann, *Constraining scalar resonances with top-quark pair production at the LHC*, [arXiv:1711.00102 \[hep-ph\]](#). 53
- [549] W. Kilian, T. Ohl, and J. Reuter, *WHIZARD: Simulating Multi-Particle Processes at LHC and ILC*, *Eur. Phys. J.* **C71** (2011) 1742, [arXiv:0708.4233 \[hep-ph\]](#). 53, 57
- [550] W. Kilian, S. Brass, T. Ohl, J. Reuter, V. Rothe, P. Stienemeier, and M. Utsch, *New Developments in WHIZARD Version 2.6*, in *International Workshop on Future Linear Collider (LCWS2017) Strasbourg, France, October 23-27, 2017*. 2018. [arXiv:1801.08034 \[hep-ph\]](#). 53
- [551] F. Buccioni, J. M. Lindert, P. Maierhöfer, S. Pozzorini, and M. Zoller, *The OpenLoops2 amplitude generator*, in preparation . 54
- [552] A. Denner and S. Dittmaier, *Reduction of one loop tensor five point integrals*, *Nucl. Phys.* **B658** (2003) 175–202, [arXiv:hep-ph/0212259 \[hep-ph\]](#). 54
- [553] A. Denner and S. Dittmaier, *Reduction schemes for one-loop tensor integrals*, *Nucl. Phys.* **B734** (2006) 62–115, [arXiv:hep-ph/0509141 \[hep-ph\]](#). 54, 56, 57
- [554] A. Denner and S. Dittmaier, *Scalar one-loop 4-point integrals*, *Nucl. Phys.* **B844** (2011) 199–242, [arXiv:1005.2076 \[hep-ph\]](#). 54
- [555] G. Ossola, C. G. Papadopoulos, and R. Pittau, *CutTools: A Program implementing the OPP reduction method to compute one-loop amplitudes*, *JHEP* **03** (2008) 042, [arXiv:0711.3596 \[hep-ph\]](#). 54, 56
- [556] A. van Hameren, *OneLoop: For the evaluation of one-loop scalar functions*, *Comput.*

- Phys. Commun. **182** (2011) 2427–2438, [arXiv:1007.4716 \[hep-ph\]](#). 54, 56, 57
- [557] F. Buccioni, S. Pozzorini, and M. Zoller, *On-the-fly reduction of open loops*, *Eur. Phys. J. C* **78** (2018) no. 1, 70, [arXiv:1710.11452 \[hep-ph\]](#). 54
- [558] C. Zhang, *Constraining qtt operators from four-top production: a case for enhanced EFT sensitivity*, *Chin. Phys. C* **42** (2018) no. 2, 023104, [arXiv:1708.05928 \[hep-ph\]](#). 56
- [559] C. Zhang, *Effective field theory approach to top-quark decay at next-to-leading order in QCD*, *Phys. Rev. D* **90** (2014) no. 1, 014008, [arXiv:1404.1264 \[hep-ph\]](#). 56
- [560] P. de Aquino, W. Link, F. Maltoni, O. Mattelaer, and T. Stelzer, *ALOHA: Automatic Libraries Of Helicity Amplitudes for Feynman Diagram Computations*, *Comput. Phys. Commun.* **183** (2012) 2254–2263, [arXiv:1108.2041 \[hep-ph\]](#). 56
- [561] P. Mastrolia, G. Ossola, C. Papadopoulos, and R. Pittau, *Optimizing the Reduction of One-Loop Amplitudes*, *JHEP* **0806** (2008) 030, [arXiv:0803.3964 \[hep-ph\]](#). 56, 57
- [562] P. Mastrolia, E. Mirabella, and T. Peraro, *Integrand reduction of one-loop scattering amplitudes through Laurent series expansion*, *JHEP* **1206** (2012) 095, [arXiv:1203.0291 \[hep-ph\]](#). 56
- [563] T. Peraro, *Ninja: Automated Integrand Reduction via Laurent Expansion for One-Loop Amplitudes*, *Comput. Phys. Commun.* **185** (2014) 2771–2797, [arXiv:1403.1229 \[hep-ph\]](#). 56
- [564] P. Mastrolia, G. Ossola, T. Reiter, and F. Tramontano, *Scattering AMplitudes from Unitarity-based Reduction Algorithm at the Integrand-level*, *JHEP* **08** (2010) 080, [arXiv:1006.0710 \[hep-ph\]](#). 56, 57
- [565] G. Passarino and M. J. G. Veltman, *One Loop Corrections for  $e^+ e^-$  Annihilation Into  $\mu^+ \mu^-$  in the Weinberg Model*, *Nucl. Phys. B* **160** (1979) 151–207. 56, 57
- [566] A. I. Davydychev, *A Simple formula for reducing Feynman diagrams to scalar integrals*, *Phys. Lett. B* **263** (1991) 107–111. 56
- [567] T. Binoth, J.-P. Guillet, G. Heinrich, E. Pilon, and T. Reiter, *Golem95: A Numerical program to calculate one-loop tensor integrals with up to six external legs*, *Comput. Phys. Commun.* **180** (2009) 2317–2330, [arXiv:0810.0992 \[hep-ph\]](#). 56, 57
- [568] J. Fleischer, T. Riemann, and V. Yundin, *One-Loop Tensor Feynman Integral Reduction with Signed Minors*, *J. Phys. Conf. Ser.* **368** (2012) 012057, [arXiv:1112.0500 \[hep-ph\]](#). 56
- [569] P. Nogueira, *Automatic Feynman graph generation*, *J. Comput. Phys.* **105** (1993) 279–289. 56, 57
- [570] J. Vermaseren, *New features of FORM*, [arXiv:math-ph/0010025 \[math-ph\]](#). 56, 57
- [571] J. Kuipers, T. Ueda, J. Vermaseren, and J. Vollinga, *FORM version 4.0*, *Comput. Phys. Commun.* **184** (2013) 1453–1467, [arXiv:1203.6543 \[cs.SC\]](#). 56, 57
- [572] G. Cullen, M. Koch-Janusz, and T. Reiter, *Spinney: A Form Library for Helicity Spinors*, *Comput. Phys. Commun.* **182** (2011) 2368–2387, [arXiv:1008.0803 \[hep-ph\]](#). 56
- [573] W. Siegel, *Supersymmetric Dimensional Regularization via Dimensional Reduction*, *Phys. Lett. B* **84** (1979) 193. 56
- [574] W. Siegel, *Inconsistency of Supersymmetric Dimensional Regularization*, *Phys. Lett. B* **94** (1980) 37. 56
- [575] I. Jack, D. Jones, and K. Roberts, *Equivalence of dimensional reduction and dimensional regularization*, *Z. Phys. C* **63** (1994) 151–160, [arXiv:hep-ph/9401349 \[hep-ph\]](#). 56

- [576] D. Stöckinger, *Regularization by dimensional reduction: consistency, quantum action principle, and supersymmetry*, *JHEP* **0503** (2005) 076, [arXiv:hep-ph/0503129 \[hep-ph\]](#). 56
- [577] G. 't Hooft and M. J. G. Veltman, *Regularization and Renormalization of Gauge Fields*, *Nucl. Phys.* **B44** (1972) 189–213. 56
- [578] H. van Deurzen, G. Luisoni, P. Mastrolia, E. Mirabella, G. Ossola, et al., *Multi-leg One-loop Massive Amplitudes from Integrand Reduction via Laurent Expansion*, *JHEP* **1403** (2014) 115, [arXiv:1312.6678 \[hep-ph\]](#). 56
- [579] G. Heinrich, G. Ossola, T. Reiter, and F. Tramontano, *Tensorial Reconstruction at the Integrand Level*, *JHEP* **1010** (2010) 105, [arXiv:1008.2441 \[hep-ph\]](#). 57
- [580] G. Cullen, J. P. Guillet, G. Heinrich, T. Kleinschmidt, E. Pilon, et al., *Golem95C: A library for one-loop integrals with complex masses*, *Comput. Phys. Commun.* **182** (2011) 2276–2284, [arXiv:1101.5595 \[hep-ph\]](#). 57
- [581] J. P. Guillet, G. Heinrich, and J. von Soden-Fraunhofen, *Tools for NLO automation: extension of the golem95C integral library*, *Comput. Phys. Commun.* **185** (2014) 1828–1834, [arXiv:1312.3887 \[hep-ph\]](#). 57
- [582] R. K. Ellis and G. Zanderighi, *Scalar one-loop integrals for QCD*, *JHEP* **02** (2008) 002, [arXiv:0712.1851 \[hep-ph\]](#). 57
- [583] S. Carrazza, R. K. Ellis, and G. Zanderighi, *QCDLoop: a comprehensive framework for one-loop scalar integrals*, *Comput. Phys. Commun.* **209** (2016) 134–143, [arXiv:1605.03181 \[hep-ph\]](#). 57
- [584] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, *Eur. Phys. J.* **C76** (2016) no. 4, 196, [arXiv:1512.01178 \[hep-ph\]](#). 57, 88, 96, 118, 149, 186, 208, 225, 241, 246, 247
- [585] G. Luisoni, P. Nason, C. Oleari, and F. Tramontano,  *$HW^\pm/HZ + 0$  and 1 jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO*, *JHEP* **10** (2013) 083, [arXiv:1306.2542 \[hep-ph\]](#). 57
- [586] G. Luisoni, C. Oleari, and F. Tramontano,  *$Wb\bar{b}j$  production at NLO with POWHEG+MiNLO*, *JHEP* **04** (2015) 161, [arXiv:1502.01213 \[hep-ph\]](#). 57, 246
- [587] L. Reina and T. Schutzmeier, *Towards  $W b \bar{b} + j$  at NLO with an Automated Approach to One-Loop Computations*, *JHEP* **09** (2012) 119, [arXiv:1110.4438 \[hep-ph\]](#). 57
- [588] S. Honeywell, S. Quackenbush, L. Reina, and C. Reuschle, *NLOX release note and manual*, in preparation . 58
- [589] NNPDF Collaboration, R. D. Ball, V. Bertone, S. Carrazza, L. Del Debbio, S. Forte, A. Guffanti, N. P. Hartland, and J. Rojo, *Parton distributions with QED corrections*, *Nucl. Phys.* **B877** (2013) 290–320, [arXiv:1308.0598 \[hep-ph\]](#). 59
- [590] A. Denner, S. Dittmaier, M. Roth, and D. Wackerroth, *Electroweak radiative corrections to  $e^+e^- \rightarrow WW \rightarrow 4$  fermions in double-pole approximation: The RACOONWW approach*, *Nucl. Phys.* **B587** (2000) 67–117, [arXiv:hep-ph/0006307 \[hep-ph\]](#). 59, 188
- [591] D. Yu. Bardin, A. Leike, T. Riemann, and M. Sachwitz, *Energy-dependent width effects in  $e^+e^-$ -annihilation near the Z-boson pole*, *Phys. Lett.* **B206** (1988) 539–542. 59, 188
- [592] K. P. O. Diener, S. Dittmaier, and W. Hollik, *Electroweak higher-order effects and theoretical uncertainties in deep-inelastic neutrino scattering*, *Phys. Rev.* **D72** (2005) 093002, [arXiv:hep-ph/0509084 \[hep-ph\]](#). 60
- [593] S. Dittmaier and M. Huber, *Radiative corrections to the neutral-current Drell-Yan process in the Standard Model and its minimal supersymmetric extension*, *JHEP* **01** (2010) 060, [arXiv:0911.2329 \[hep-ph\]](#). 60

- [594] J. R. Andersen et al., *Les Houches 2013: Physics at TeV Colliders: Standard Model Working Group Report*, [arXiv:1405.1067 \[hep-ph\]](#). 80, 186, 187
- [595] W. Beenakker, C. Borschensky, M. Krämer, A. Kulesza, E. Laenen, S. Marzani, and J. Rojo, *NLO+NLL squark and gluino production cross-sections with threshold-improved parton distributions*, *Eur. Phys. J.* **C76** (2016) no. 2, 53, [arXiv:1510.00375 \[hep-ph\]](#). 80
- [596] ATLAS Collaboration, M. Aaboud et al., *Measurement of the W-boson mass in pp collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector*, [arXiv:1701.07240 \[hep-ex\]](#). 80
- [597] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, *J. Phys.* **G43** (2016) 023001, [arXiv:1510.03865 \[hep-ph\]](#). 80, 180, 207, 241
- [598] NNPDF Collaboration, R. D. Ball et al., *Parton distributions from high-precision collider data*, *Eur. Phys. J.* **C77** (2017) no. 10, 663, [arXiv:1706.00428 \[hep-ph\]](#). 80, 81, 82, 83
- [599] R. D. Ball et al., *Parton distributions with LHC data*, *Nucl. Phys.* **B867** (2013) 244–289, [arXiv:1207.1303 \[hep-ph\]](#). 82, 101
- [600] K. J. Eskola, P. Paakkinen, H. Paukkunen, and C. A. Salgado, *EPPS16: Nuclear parton distributions with LHC data*, *Eur. Phys. J.* **C77** (2017) no. 3, 163, [arXiv:1612.05741 \[hep-ph\]](#). 84
- [601] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*, *Eur. Phys. J.* **C75** (2015) no. 5, 204, [arXiv:1412.3989 \[hep-ph\]](#). 84, 96, 180, 186, 247
- [602] A. Abdesselam et al., *Boosted objects: a probe of beyond the standard model physics*, *EPHJA,C71,1661.2011* **C71** (2011) 1661, [arXiv:1012.5412 \[hep-ph\]](#). 87, 89, 110
- [603] A. Altheimer et al., *Jet Substructure at the Tevatron and LHC: New results, new tools, new benchmarks*, *J. Phys.* **G39** (2012) 063001, [arXiv:1201.0008 \[hep-ph\]](#). 87, 89, 110
- [604] A. Altheimer et al., *Boosted objects and jet substructure at the LHC. Report of BOOST2012, held at IFIC Valencia, 23rd-27th of July 2012*, *Eur. Phys. J.* **C74** (2014) no. 3, 2792, [arXiv:1311.2708 \[hep-ex\]](#). 87, 89, 94, 110
- [605] D. Adams et al., *Towards an Understanding of the Correlations in Jet Substructure*, *Eur. Phys. J.* **C75** (2015) no. 9, 409, [arXiv:1504.00679 \[hep-ph\]](#). 87, 89, 110
- [606] A. J. Larkoski, I. Moult, and B. Nachman, *Jet Substructure at the Large Hadron Collider: A Review of Recent Advances in Theory and Machine Learning*, [arXiv:1709.04464 \[hep-ph\]](#). 87, 89, 110
- [607] A. Andreassen, W. Frost, and M. D. Schwartz, *Scale Invariant Instantons and the Complete Lifetime of the Standard Model*, [arXiv:1707.08124 \[hep-ph\]](#). 87
- [608] D. E. Kaplan and M. D. Schwartz, *Constraining Light Colored Particles with Event Shapes*, *Phys. Rev. Lett.* **101** (2008) 022002, [arXiv:0804.2477 \[hep-ph\]](#). 87
- [609] D. Becciolini, M. Gillioz, M. Nardecchia, F. Sannino, and M. Spannowsky, *Constraining new colored matter from the ratio of 3 to 2 jets cross sections at the LHC*, *Phys. Rev.* **D91** (2015) no. 1, 015010, [arXiv:1403.7411 \[hep-ph\]](#). [Addendum: *Phys. Rev.* **D92**,no.7,079905(2015)]. 87
- [610] C. Anastasiou, C. Duhr, F. Dulat, F. Herzog, and B. Mistlberger, *Higgs Boson Gluon-Fusion Production in QCD at Three Loops*, *Phys. Rev. Lett.* **114** (2015) 212001, [arXiv:1503.06056 \[hep-ph\]](#). 87
- [611] Particle Data Group Collaboration, C. Patrignani et al., *Review of Particle Physics*, *Chin. Phys.* **C40** (2016) no. 10, 100001. 88, 89
- [612] *Workshop on Precision Measurements of alphas*. 2011. [arXiv:1110.0016 \[hep-ph\]](#).

- [http://lss.fnal.gov/cgi-bin/find\\_paper.pl?conf-11-611](http://lss.fnal.gov/cgi-bin/find_paper.pl?conf-11-611). 88
- [613] G. Altarelli, *The QCD Running Coupling and its Measurement*, PoS **Corfu2012** (2013) 002, [arXiv:1303.6065 \[hep-ph\]](#). 88
- [614] A. Pich, *Review of  $\alpha_s$  determinations*, [arXiv:1303.2262 \[hep-ph\]](#). [PoSConfinementX,022(2012)]. 88
- [615] S. Moch et al., *High precision fundamental constants at the TeV scale*, [arXiv:1405.4781 \[hep-ph\]](#). 88
- [616] D. d'Enterria, ed., *Proceedings, High-Precision  $\alpha_s$  Measurements from LHC to FCC-ee*, CERN. CERN, Geneva, 2015. [arXiv:1512.05194 \[hep-ph\]](#). <https://inspirehep.net/record/1409920/files/arXiv:1512.05194.pdf>. 88
- [617] G. P. Salam, *The strong coupling: a theoretical perspective*, [arXiv:1712.05165 \[hep-ph\]](#). 88
- [618] R. Abbate, M. Fickinger, A. H. Hoang, V. Mateu, and I. W. Stewart, *Thrust at  $N^3LL$  with Power Corrections and a Precision Global Fit for  $\alpha_s(m_Z)$* , *Phys. Rev.* **D83** (2011) 074021, [arXiv:1006.3080 \[hep-ph\]](#). 88, 91, 92
- [619] A. H. Hoang, D. W. Kolodrubetz, V. Mateu, and I. W. Stewart, *Precise determination of  $\alpha_s$  from the  $C$ -parameter distribution*, *Phys. Rev.* **D91** (2015) no. 9, 094018, [arXiv:1501.04111 \[hep-ph\]](#). 88, 92
- [620] ALEPH Collaboration, A. Heister et al., *Studies of QCD at  $e^+e^-$  centre-of-mass energies between 91 and 209 GeV*, *Eur. Phys. J.* **C35** (2004) 457–486. 88, 219, 221, 225
- [621] DELPHI Collaboration, J. Abdallah et al., *The Measurement of  $\alpha(s)$  from event shapes with the DELPHI detector at the highest LEP energies*, *Eur. Phys. J.* **C37** (2004) 1–23, [arXiv:hep-ex/0406011 \[hep-ex\]](#). 88
- [622] DELPHI Collaboration, P. Abreu et al., *Measurement of event shape and inclusive distributions at  $S^{*(1/2)} = 130\text{-GeV}$  and  $136\text{-GeV}$* , *Z. Phys.* **C73** (1997) 229–242. 88
- [623] DELPHI Collaboration, P. Abreu et al., *Energy dependence of event shapes and of  $\alpha(s)$  at LEP-2*, *Phys. Lett.* **B456** (1999) 322–340. 88
- [624] JADE Collaboration, O. Biebel, P. A. Movilla Fernandez, and S. Bethke,  *$C$  parameter and jet broadening at PETRA energies*, *Phys. Lett.* **B459** (1999) 326–334, [arXiv:hep-ex/9903009 \[hep-ex\]](#). 88
- [625] L3 Collaboration, B. Adeva et al., *Studies of hadronic event structure and comparisons with QCD models at the  $Z^0$  resonance*, *Z. Phys.* **C55** (1992) 39–62. 88
- [626] OPAL Collaboration, G. Abbiendi et al., *Measurement of event shape distributions and moments in  $e^+e^- \rightarrow$  hadrons at 91-GeV - 209-GeV and a determination of  $\alpha(s)$* , *Eur. Phys. J.* **C40** (2005) 287–316, [arXiv:hep-ex/0503051 \[hep-ex\]](#). 88
- [627] SLD Collaboration, K. Abe et al., *Measurement of  $\alpha_s(M(Z)^2)$  from hadronic event observables at the  $Z^0$  resonance*, *Phys. Rev.* **D51** (1995) 962–984, [arXiv:hep-ex/9501003 \[hep-ex\]](#). 88
- [628] CMS Collaboration, S. Chatrchyan et al., *Determination of the top-quark pole mass and strong coupling constant from the  $t$   $t$ -bar production cross section in  $pp$  collisions at  $\sqrt{s} = 7$  TeV*, *Phys. Lett.* **B728** (2014) 496–517, [arXiv:1307.1907 \[hep-ex\]](#). [Erratum: *Phys. Lett.* **B738**,526(2014)]. 88
- [629] T. Sjöstrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, *JHEP* **05** (2006) 026, [arXiv:hep-ph/0603175 \[hep-ph\]](#). 88, 96, 113, 118, 173, 217, 225
- [630] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, *An Introduction to PYTHIA 8.2*,

- Comput. Phys. Commun. **191** (2015) 159–177, [arXiv:1410.3012 \[hep-ph\]](#). 88, 96, 217, 225
- [631] P. Skands, S. Carrazza, and J. Rojo, *Tuning PYTHIA 8.1: the Monash 2013 Tune*, *Eur. Phys. J.* **C74** (2014) no. 8, 3024, [arXiv:1404.5630 \[hep-ph\]](#). 88, 118, 200
- [632] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, public note **ATL-PHYS-PUB-2014-021** (2014) . 88, 203, 247
- [633] T. Sjöstrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, *Comput. Phys. Commun.* **178** (2008) 852–867, [arXiv:0710.3820 \[hep-ph\]](#). 88, 113, 118, 203, 217, 247
- [634] Z. Nagy, *Next-to-leading order calculation of three jet observables in hadron hadron collision*, *Phys. Rev.* **D68** (2003) 094002, [arXiv:hep-ph/0307268 \[hep-ph\]](#). 89
- [635] ATLAS Collaboration, G. Aad et al., *Measurement of transverse energy-energy correlations in multi-jet events in pp collisions at  $\sqrt{s} = 7$  TeV using the ATLAS detector and determination of the strong coupling constant  $\alpha_s(m_Z)$* , *Phys. Lett.* **B750** (2015) 427–447, [arXiv:1508.01579 \[hep-ex\]](#). 89, 99
- [636] ATLAS Collaboration, M. Aaboud et al., *Determination of the strong coupling constant  $\alpha_s$  from transverse energy-energy correlations in multijet events at  $\sqrt{s} = 8$  TeV using the ATLAS detector*, *Eur. Phys. J.* **C77** (2017) no. 12, 872, [arXiv:1707.02562 \[hep-ex\]](#). 89, 99
- [637] CMS Collaboration, V. Khachatryan et al., *Constraints on parton distribution functions and extraction of the strong coupling constant from the inclusive jet cross section in pp collisions at  $\sqrt{s} = 7$  TeV*, *Eur. Phys. J.* **C75** (2015) no. 6, 288, [arXiv:1410.6765 \[hep-ex\]](#). 89
- [638] CMS Collaboration, V. Khachatryan et al., *Measurement of the inclusive 3-jet production differential cross section in proton-proton collisions at 7 TeV and determination of the strong coupling constant in the TeV range*, *Eur. Phys. J.* **C75** (2015) no. 5, 186, [arXiv:1412.1633 \[hep-ex\]](#). 89
- [639] CMS Collaboration, S. Chatrchyan et al., *Measurement of the ratio of the inclusive 3-jet cross section to the inclusive 2-jet cross section in pp collisions at  $\sqrt{s} = 7$  TeV and first determination of the strong coupling constant in the TeV range*, *Eur. Phys. J.* **C73** (2013) no. 10, 2604, [arXiv:1304.7498 \[hep-ex\]](#). 89, 99
- [640] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, *Jet substructure as a new Higgs search channel at the LHC*, *Phys. Rev. Lett.* **100** (2008) 242001, [arXiv:0802.2470 \[hep-ph\]](#). 89, 90, 111, 117
- [641] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, *Techniques for improved heavy particle searches with jet substructure*, *Phys. Rev.* **D80** (2009) 051501, [arXiv:0903.5081 \[hep-ph\]](#). 89, 90, 117
- [642] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, *Recombination Algorithms and Jet Substructure: Pruning as a Tool for Heavy Particle Searches*, *Phys. Rev.* **D81** (2010) 094023, [arXiv:0912.0033 \[hep-ph\]](#). 89, 90, 117
- [643] D. Krohn, J. Thaler, and L.-T. Wang, *Jet Trimming*, *JHEP* **02** (2010) 084, [arXiv:0912.1342 \[hep-ph\]](#). 89, 90, 111, 117
- [644] M. Dasgupta, A. Fregoso, S. Marzani, and G. P. Salam, *Towards an understanding of jet substructure*, *JHEP* **09** (2013) 029, [arXiv:1307.0007 \[hep-ph\]](#). 89, 90, 91, 92, 93, 110, 111, 117, 118, 123
- [645] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, *Soft Drop*, *JHEP* **05** (2014) 146, [arXiv:1402.2657 \[hep-ph\]](#). 89, 90, 93, 95, 111, 123, 135
- [646] C. Frye, A. J. Larkoski, M. D. Schwartz, and K. Yan, *Precision physics with pile-up*

- insensitive observables*, [arXiv:1603.06375 \[hep-ph\]](#). 89, 90, 102, 110
- [647] C. Frye, A. J. Larkoski, M. D. Schwartz, and K. Yan, *Factorization for groomed jet substructure beyond the next-to-leading logarithm*, [JHEP 07 \(2016\) 064](#), [arXiv:1603.09338 \[hep-ph\]](#). 89, 90, 92, 93, 94, 95, 102, 110
- [648] S. Marzani, L. Schunk, and G. Soyez, *A study of jet mass distributions with grooming*, [JHEP 07 \(2017\) 132](#), [arXiv:1704.02210 \[hep-ph\]](#). 89, 90, 94, 95, 102, 103, 107, 110
- [649] S. Marzani, L. Schunk, and G. Soyez, *The jet mass distribution after Soft Drop*, [arXiv:1712.05105 \[hep-ph\]](#). 89, 90, 94, 95, 102, 103, 107, 110
- [650] ATLAS Collaboration, M. Aaboud et al., *A measurement of the soft-drop jet mass in pp collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, [arXiv:1711.08341 \[hep-ex\]](#). 89, 90, 102, 105, 107, 108, 109
- [651] CMS Collaboration Collaboration, *Measurement of the differential jet production cross section with respect to jet mass and transverse momentum in dijet events from pp collisions at  $\sqrt{s} = 13$  TeV*, Tech. Rep. CMS-PAS-SMP-16-010, CERN, Geneva, 2017. <https://cds.cern.ch/record/2273393>. 89, 90, 102, 108
- [652] A. Banfi, G. P. Salam, and G. Zanderighi, *Principles of general final-state resummation and automated implementation*, [JHEP 03 \(2005\) 073](#), [arXiv:hep-ph/0407286 \[hep-ph\]](#). 90, 117, 215, 216
- [653] A. J. Larkoski, G. P. Salam, and J. Thaler, *Energy Correlation Functions for Jet Substructure*, [JHEP 1306 \(2013\) 108](#), [arXiv:1305.0007 \[hep-ph\]](#). 90, 111, 115, 116, 117
- [654] C. F. Berger, T. Kucs, and G. F. Sterman, *Event shape / energy flow correlations*, [Phys. Rev. D68 \(2003\) 014012](#), [arXiv:hep-ph/0303051 \[hep-ph\]](#). 90
- [655] L. G. Almeida, S. J. Lee, G. Perez, G. F. Sterman, I. Sung, and J. Virzi, *Substructure of high- $p_T$  Jets at the LHC*, [Phys. Rev. D79 \(2009\) 074017](#), [arXiv:0807.0234 \[hep-ph\]](#). 90
- [656] S. D. Ellis, C. K. Vermilion, J. R. Walsh, A. Hornig, and C. Lee, *Jet Shapes and Jet Algorithms in SCET*, [JHEP 11 \(2010\) 101](#), [arXiv:1001.0014 \[hep-ph\]](#). 90
- [657] A. J. Larkoski, J. Thaler, and W. J. Waalewijn, *Gaining (Mutual) Information about Quark/Gluon Discrimination*, [JHEP 11 \(2014\) 129](#), [arXiv:1408.3122 \[hep-ph\]](#). 90, 103, 110
- [658] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, *Better jet clustering algorithms*, [JHEP 08 \(1997\) 001](#), [arXiv:hep-ph/9707323 \[hep-ph\]](#). 90, 117
- [659] M. Wobisch and T. Wengler, *Hadronization corrections to jet cross-sections in deep inelastic scattering*, [arXiv:hep-ph/9907280 \[hep-ph\]](#). 90, 117
- [660] M. Wobisch, *Measurement and QCD analysis of jet cross-sections in deep inelastic positron proton collisions at  $\sqrt{s} = 300$  GeV*, 2000. 90, 117
- [661] M. Cacciari, G. P. Salam, and G. Soyez, *The anti- $k_t$  jet clustering algorithm*, [JHEP 04 \(2008\) 063](#), [arXiv:0802.1189 \[hep-ph\]](#). 90, 96, 117, 189, 207, 241, 247
- [662] M. Dasgupta and G. P. Salam, *Resummation of nonglobal QCD observables*, [Phys. Lett. B512 \(2001\) 323–330](#), [arXiv:hep-ph/0104277 \[hep-ph\]](#). 91, 93, 118
- [663] G. P. Korchemsky and G. F. Sterman, *Power corrections to event shapes and factorization*, [Nucl. Phys. B555 \(1999\) 335–351](#), [arXiv:hep-ph/9902341 \[hep-ph\]](#). 92, 113
- [664] G. Korchemsky and S. Tafat, *On power corrections to the event shape distributions in QCD*, [JHEP 0010 \(2000\) 010](#), [arXiv:hep-ph/0007005 \[hep-ph\]](#). 92, 113
- [665] A. H. Hoang and I. W. Stewart, *Designing gapped soft functions for jet production*,

- Phys. Lett. **B660** (2008) 483–493, [arXiv:0709.3519 \[hep-ph\]](#). 92, 113
- [666] Z. Ligeti, I. W. Stewart, and F. J. Tackmann, *Treating the  $b$  quark distribution function with reliable uncertainties*, Phys. Rev. **D78** (2008) 114014, [arXiv:0807.1926 \[hep-ph\]](#). 92, 113
- [667] Y. L. Dokshitzer, G. Marchesini, and B. R. Webber, *Dispersive approach to power behaved contributions in QCD hard processes*, Nucl. Phys. **B469** (1996) 93–142, [arXiv:hep-ph/9512336 \[hep-ph\]](#). 92, 113
- [668] Y. L. Dokshitzer and B. R. Webber, *Calculation of power corrections to hadronic event shapes*, Phys. Lett. **B352** (1995) 451–455, [arXiv:hep-ph/9504219 \[hep-ph\]](#). 92, 113
- [669] T. Gehrmann, G. Luisoni, and P. F. Monni, *Power corrections in the dispersive model for a determination of the strong coupling constant from the thrust distribution*, Eur. Phys. J. **C73** (2013) no. 1, 2265, [arXiv:1210.6945 \[hep-ph\]](#). 92
- [670] C. Lee and G. F. Sterman, *Universality of nonperturbative effects in event shapes*, eConf **C0601121** (2006) A001, [arXiv:hep-ph/0603066 \[hep-ph\]](#). 92
- [671] C. Lee, *Universal nonperturbative effects in event shapes from soft-collinear effective theory*, Mod. Phys. Lett. **A22** (2007) 835–851, [arXiv:hep-ph/0703030 \[hep-ph\]](#). 92
- [672] R. Abbate, M. Fickinger, A. H. Hoang, V. Mateu, and I. W. Stewart, *Precision Thrust Cumulant Moments at  $N^3LL$* , Phys. Rev. **D86** (2012) 094002, [arXiv:1204.5746 \[hep-ph\]](#). 92
- [673] G. P. Salam and D. Wicke, *Hadron masses and power corrections to event shapes*, JHEP **05** (2001) 061, [arXiv:hep-ph/0102343 \[hep-ph\]](#). 92
- [674] V. Mateu, I. W. Stewart, and J. Thaler, *Power Corrections to Event Shapes with Mass-Dependent Operators*, Phys. Rev. **D87** (2013) no. 1, 014025, [arXiv:1209.3781 \[hep-ph\]](#). 92
- [675] T. T. Jouttenus, I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, *Jet mass spectra in Higgs boson plus one jet at next-to-next-to-leading logarithmic order*, Phys. Rev. **D88** (2013) no. 5, 054031, [arXiv:1302.0846 \[hep-ph\]](#). 93
- [676] M. Dasgupta, K. Khelifa-Kerfa, S. Marzani, and M. Spannowsky, *On jet mass distributions in  $Z$ +jet and dijet processes at the LHC*, JHEP **10** (2012) 126, [arXiv:1207.1640 \[hep-ph\]](#). 93
- [677] A. Banfi, G. P. Salam, and G. Zanderighi, *Infrared safe definition of jet flavor*, Eur. Phys. J. **C47** (2006) 113–124, [arXiv:hep-ph/0601139 \[hep-ph\]](#). 94
- [678] J. M. Campbell, R. K. Ellis, and C. Williams, *Vector boson pair production at the LHC*, JHEP **07** (2011) 018, [arXiv:1105.0020 \[hep-ph\]](#). 94
- [679] S. Badger, B. Biedermann, P. Uwer, and V. Yundin, *Next-to-leading order QCD corrections to five jet production at the LHC*, Phys. Rev. **D89** (2014) no. 3, 034019, [arXiv:1309.6585 \[hep-ph\]](#). 94, 107
- [680] G. P. Salam, *Towards Jetography*, Eur. Phys. J. **C67** (2010) 637–686, [arXiv:0906.1833 \[hep-ph\]](#). 94
- [681] M. Cacciari, G. P. Salam, and G. Soyez, *The Catchment Area of Jets*, JHEP **04** (2008) 005, [arXiv:0802.1188 \[hep-ph\]](#). 94, 131
- [682] G. Soyez, G. P. Salam, J. Kim, S. Dutta, and M. Cacciari, *Pileup subtraction for jet shapes*, Phys. Rev. Lett. **110** (2013) no. 16, 162001, [arXiv:1211.2811 \[hep-ph\]](#). 94, 131
- [683] M. Cacciari, G. P. Salam, and G. Soyez, *SoftKiller, a particle-level pileup removal method*, Eur. Phys. J. **C75** (2015) no. 2, 59, [arXiv:1407.0408 \[hep-ph\]](#). 94, 131
- [684] D. Krohn, M. D. Schwartz, M. Low, and L.-T. Wang, *Jet Cleansing: Pileup Removal at*

- High Luminosity*, *Phys. Rev.* **D90** (2014) no. 6, 065020, [arXiv:1309.4777 \[hep-ph\]](#). **94**, **131**
- [685] D. Bertolini, P. Harris, M. Low, and N. Tran, *Pileup Per Particle Identification*, *JHEP* **10** (2014) 059, [arXiv:1407.6013 \[hep-ph\]](#). **94**, **131**
- [686] P. Berta, M. Spousta, D. W. Miller, and R. Leitner, *Particle-level pileup subtraction for jets and jet shapes*, *JHEP* **06** (2014) 092, [arXiv:1403.3108 \[hep-ex\]](#). **94**, **131**
- [687] P. T. Komiske, E. M. Metodiev, B. Nachman, and M. D. Schwartz, *Pileup Mitigation with Machine Learning (PUMML)*, *JHEP* **12** (2017) 051, [arXiv:1707.08600 \[hep-ph\]](#). **94**, **131**
- [688] CMS Collaboration Collaboration, *Pileup Removal Algorithms*, Tech. Rep. CMS-PAS-JME-14-001, CERN, Geneva, 2014. <https://cds.cern.ch/record/1751454>. **94**
- [689] CMS Collaboration Collaboration, *Pileup per particle identification: preparation for Run 2*, . **94**
- [690] ATLAS Collaboration Collaboration, *Constituent-level pile-up mitigation techniques in ATLAS*, Tech. Rep. ATLAS-CONF-2017-065, CERN, Geneva, Aug, 2017. <https://cds.cern.ch/record/2281055>. **94**
- [691] ATLAS Collaboration Collaboration, *Impact of Alternative Inputs and Grooming Methods on Large-R Jet Reconstruction in ATLAS*, Tech. Rep. ATL-PHYS-PUB-2017-020, CERN, Geneva, Dec, 2017. <https://cds.cern.ch/record/2297485>. **94**
- [692] ATLAS Collaboration, G. Aad et al., *Performance of pile-up mitigation techniques for jets in pp collisions at  $\sqrt{s} = 8$  TeV using the ATLAS detector*, *Eur. Phys. J.* **C76** (2016) no. 11, 581, [arXiv:1510.03823 \[hep-ex\]](#). **94**
- [693] ATLAS Collaboration, G. Aad et al., *Identification of boosted, hadronically decaying W bosons and comparisons with ATLAS data taken at  $\sqrt{s} = 8$  TeV*, *Eur. Phys. J.* **C76** (2016) no. 3, 154, [arXiv:1510.05821 \[hep-ex\]](#). **94**, **110**
- [694] ATLAS Collaboration Collaboration, *Jet substructure at very high luminosity*, Tech. Rep. JetSubstructureECFA2014, CERN, Geneva, Oct, 2014. <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/JetSubstructureECFA2014>. **94**
- [695] D. Krohn, M. D. Schwartz, T. Lin, and W. J. Waalewijn, *Jet Charge at the LHC*, *Phys. Rev. Lett.* **110** (2013) no. 21, 212001, [arXiv:1209.2421 \[hep-ph\]](#). **94**, **110**
- [696] W. J. Waalewijn, *Calculating the Charge of a Jet*, *Phys. Rev.* **D86** (2012) 094030, [arXiv:1209.3019 \[hep-ph\]](#). **94**, **110**
- [697] H.-M. Chang, M. Procura, J. Thaler, and W. J. Waalewijn, *Calculating Track-Based Observables for the LHC*, *Phys. Rev. Lett.* **111** (2013) 102002, [arXiv:1303.6637 \[hep-ph\]](#). **94**
- [698] B. T. Elder, M. Procura, J. Thaler, W. J. Waalewijn, and K. Zhou, *Generalized Fragmentation Functions for Fractal Jet Observables*, *JHEP* **06** (2017) 085, [arXiv:1704.05456 \[hep-ph\]](#). **94**
- [699] BaBar Collaboration, D. Boutigny et al., *The BABAR physics book: Physics at an asymmetric B factory*, in *Workshop on Physics at an Asymmetric B Factory (BaBar Collaboration Meeting) Pasadena, California, September 22-24, 1997*. 1998. **95**
- [700] J. Bellm et al., *Herwig 7.1 Release Note*, [arXiv:1705.06919 \[hep-ph\]](#). **96**, **149**, **241**, **247**
- [701] D. Reichelt, P. Richardson, and A. Siodmok, *Improving the Simulation of Quark and*

- Gluon Jets with Herwig 7*, *Eur. Phys. J.* **C77** (2017) no. 12, 876, [arXiv:1708.01491 \[hep-ph\]](#). 96
- [702] S. Gieseke, P. Stephens, and B. Webber, *New formalism for QCD parton showers*, *JHEP* **12** (2003) 045, [arXiv:hep-ph/0310083 \[hep-ph\]](#). 96, 225, 241
- [703] B. R. Webber, *A QCD Model for Jet Fragmentation Including Soft Gluon Interference*, *Nucl. Phys.* **B238** (1984) 492–528. 96, 123, 225
- [704] M. Bähr, J. M. Butterworth, and M. H. Seymour, *The Underlying Event and the Total Cross Section from Tevatron to the LHC*, *JHEP* **01** (2009) 065, [arXiv:0806.2949 \[hep-ph\]](#). 96
- [705] M. Bähr, S. Gieseke, and M. H. Seymour, *Simulation of multiple partonic interactions in Herwig++*, *JHEP* **07** (2008) 076, [arXiv:0803.3633 \[hep-ph\]](#). 96, 240
- [706] S. Gieseke, F. Loshaj, and P. Kirchgaeßer, *Soft and diffractive scattering with the cluster model in Herwig*, *Eur. Phys. J.* **C77** (2017) no. 3, 156, [arXiv:1612.04701 \[hep-ph\]](#). 96
- [707] S. Gieseke, C. Rohr, and A. Siodmok, *Colour reconnections in Herwig++*, *Eur. Phys. J.* **C72** (2012) 2225, [arXiv:1206.0041 \[hep-ph\]](#). 96
- [708] R. Corke and T. Sjöstrand, *Interleaved Parton Showers and Tuning Prospects*, *JHEP* **03** (2011) 032, [arXiv:1011.1759 \[hep-ph\]](#). 96, 154
- [709] P. Gras, S. Höche, D. Kar, A. Larkoski, L. Lonnblad, S. Plätzer, A. Siodmok, P. Skands, G. Soyez, and J. Thaler, *Systematics of quark/gluon tagging*, *JHEP* **07** (2017) 091, [arXiv:1704.03878 \[hep-ph\]](#). 96
- [710] I. Mout and H. X. Zhu, *Simplicity from Recoil: The Three-Loop Soft Function and Factorization for the Energy-Energy Correlation*, [arXiv:1801.02627 \[hep-ph\]](#). 98
- [711] A. Accardi et al., *A Critical Appraisal and Evaluation of Modern PDFs*, *Eur. Phys. J.* **C76** (2016) no. 8, 471, [arXiv:1603.08906 \[hep-ph\]](#). 99
- [712] ATLAS Collaboration, *Jet energy measurement and its systematic uncertainty in proton-proton collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector*, *Eur. Phys. J.* **C75** (2015) 17, [arXiv:1406.0076 \[hep-ex\]](#). 100, 205
- [713] ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, *Phys. Rev.* **D96** (2017) no. 7, 072002, [arXiv:1703.09665 \[hep-ex\]](#). 100, 203
- [714] CMS Collaboration, V. Khachatryan et al., *Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV*, *JINST* **12** (2017) no. 02, P02014, [arXiv:1607.03663 \[hep-ex\]](#). 100
- [715] CMS Collaboration, *Jet energy scale and resolution performances with 13TeV data*, CMS-DP-2016-020 (2016) . 100
- [716] ATLAS Collaboration, *In-situ measurements of the ATLAS large-radius jet response in 13 TeV pp collisions*, Tech. Rep. ATLAS-CONF-2017-063, CERN, Geneva, Jul, 2017. <http://cds.cern.ch/record/2275655>. 100, 106, 109
- [717] DELPHES 3 Collaboration, J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaitre, A. Mertens, and M. Selvaggi, *DELPHES 3, A modular framework for fast simulation of a generic collider experiment*, *JHEP* **02** (2014) 057, [arXiv:1307.6346 \[hep-ex\]](#). 105
- [718] G. D’Agostini, *A Multidimensional unfolding method based on Bayes’ theorem*, *Nucl. Instrum. Meth.* **A362** (1995) 487–498. 105
- [719] T. Adye, *Unfolding algorithms and tests using RooUnfold*, in *Proceedings, PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments*

- and Unfolding, CERN, Geneva, Switzerland 17-20 January 2011, pp. 313–318, CERN. CERN, Geneva, 2011. [arXiv:1105.1160 \[physics.data-an\]](#). 105
- [720] CMS Collaboration Collaboration, *Jet algorithms performance in 13 TeV data*, Tech. Rep. CMS-PAS-JME-16-003, CERN, Geneva, 2017. <https://cds.cern.ch/record/2256875>. 106
- [721] J. Gallicchio and M. D. Schwartz, *Pure Samples of Quark and Gluon Jets at the LHC*, *JHEP* **10** (2011) 103, [arXiv:1104.1175 \[hep-ph\]](#). 107
- [722] A. Gehrmann-De Ridder, T. Gehrmann, E. W. N. Glover, and G. Heinrich, *NNLO corrections to event shapes in  $e^+e^-$  annihilation*, *JHEP* **12** (2007) 094, [arXiv:0711.4711 \[hep-ph\]](#). 107
- [723] A. Gehrmann-De Ridder, T. Gehrmann, E. W. N. Glover, and G. Heinrich, *Second-order QCD corrections to the thrust distribution*, *Phys. Rev. Lett.* **99** (2007) 132002, [arXiv:0707.1285 \[hep-ph\]](#). 107
- [724] S. Weinzierl, *NNLO corrections to 3-jet observables in electron-positron annihilation*, *Phys. Rev. Lett.* **101** (2008) 162001, [arXiv:0807.3241 \[hep-ph\]](#). 107
- [725] S. Weinzierl, *Event shapes and jet rates in electron-positron annihilation at NNLO*, *JHEP* **06** (2009) 041, [arXiv:0904.1077 \[hep-ph\]](#). 107
- [726] CMS Collaboration, A. M. Sirunyan et al., *Search for massive resonances decaying into WW, WZ or ZZ bosons in proton-proton collisions at  $\sqrt{s} = 13$  TeV*, *JHEP* **03** (2017) 162, [arXiv:1612.09159 \[hep-ex\]](#). 108
- [727] *Measurement of large radius jet mass reconstruction performance at  $\sqrt{s} = 8$  TeV using the ATLAS detector*, Tech. Rep. ATLAS-CONF-2016-008, CERN, Geneva, Mar, 2016. <http://cds.cern.ch/record/2139642>. 108
- [728] ATLAS Collaboration Collaboration, *Jet mass reconstruction with the ATLAS Detector in early Run 2 data*, Tech. Rep. ATLAS-CONF-2016-035, CERN, Geneva, Jul, 2016. <http://cds.cern.ch/record/2200211>. 108
- [729] ATLAS Collaboration, M. Aaboud et al., *A measurement of the calorimeter response to single hadrons and determination of the jet energy scale uncertainty using LHC Run-1 pp-collision data with the ATLAS detector*, *Eur. Phys. J.* **C77** (2017) no. 1, 26, [arXiv:1607.08842 \[hep-ex\]](#). 109
- [730] B. Nachman, *Investigating the Quantum Properties of Jets and the Search for a Supersymmetric Top Quark Partner with the ATLAS Detector*. PhD thesis, Stanford U., Phys. Dept., 2016. [arXiv:1609.03242 \[hep-ex\]](#). <https://inspirehep.net/record/1486232/files/arXiv:1609.03242.pdf>. 109
- [731] ATLAS Collaboration, G. Aad et al., *Performance of jet substructure techniques for large- $R$  jets in proton-proton collisions at  $\sqrt{s} = 7$  TeV using the ATLAS detector*, *JHEP* **09** (2013) 076, [arXiv:1306.4945 \[hep-ex\]](#). 109
- [732] CMS Collaboration, S. Chatrchyan et al., *Search for a Higgs boson in the decay channel  $H$  to  $ZZ^*$  to  $q \bar{q} \ell^- l^+$  in pp collisions at  $\sqrt{s} = 7$  TeV*, *JHEP* **1204** (2012) 036, [arXiv:1202.1416 \[hep-ex\]](#). 110
- [733] CMS Collaboration, *Search for a Standard Model-like Higgs boson decaying into  $WW$  to  $l \nu q \bar{q}$  in pp collisions at  $\sqrt{s} = 8$  TeV*, Tech. Rep. CMS-PAS-HIG-13-008, 2013. 110
- [734] ATLAS Collaboration, G. Aad et al., *Measurement of jet charge in dijet events from  $\sqrt{s} = 8$  TeV pp collisions with the ATLAS detector*, *Phys. Rev.* **D93** (2016) no. 5, 052003, [arXiv:1509.05190 \[hep-ex\]](#). 110
- [735] ATLAS Collaboration, G. Aad et al., *Measurement of colour flow with the jet pull angle in  $t\bar{t}$  events using the ATLAS detector at  $\sqrt{s} = 8$  TeV*, *Phys. Lett.* **B750** (2015)

- 475–493, [arXiv:1506.05629 \[hep-ex\]](#). 110
- [736] *Performance of jet substructure techniques in early  $\sqrt{s} = 13$  TeV pp collisions with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2015-035, CERN, Geneva, Aug, 2015. <http://cds.cern.ch/record/2041462>. 110
- [737] ATLAS Collaboration, G. Aad et al., *Measurement of the differential cross-section of highly boosted top quarks as a function of their transverse momentum in  $\sqrt{s} = 8$  TeV proton-proton collisions using the ATLAS detector*, *Phys. Rev. D* **93** (2016) no. 3, 032009, [arXiv:1510.03818 \[hep-ex\]](#). 110, 173, 176, 177
- [738] *Studies of b-tagging performance and jet substructure in a high  $p_T$   $g \rightarrow b\bar{b}$  rich sample of large- $R$  jets from pp collisions at  $\sqrt{s} = 8$  TeV with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2016-002, CERN, Geneva, Feb, 2016. <http://cds.cern.ch/record/2135187>. 110
- [739] ATLAS Collaboration Collaboration, *Boosted Higgs ( $\rightarrow b\bar{b}$ ) Boson Identification with the ATLAS Detector at  $\sqrt{s} = 13$  TeV*, Tech. Rep. ATLAS-CONF-2016-039, CERN, Geneva, Aug, 2016. <http://cds.cern.ch/record/2206038>. 110
- [740] ATLAS Collaboration Collaboration, *Discrimination of Light Quark and Gluon Jets in pp collisions at  $\sqrt{s} = 8$  TeV with the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2016-034, CERN, Geneva, Jul, 2016. <http://cds.cern.ch/record/2200202>. 110
- [741] CMS Collaboration Collaboration, *Measurement of the  $t\bar{t}$  production cross section at 13 TeV in the all-jets final state*, Tech. Rep. CMS-PAS-TOP-16-013, CERN, Geneva, 2016. <https://cds.cern.ch/record/2161138>. 110
- [742] CMS Collaboration Collaboration, *Search for  $t\bar{t}H$  production in the  $H \rightarrow b\bar{b}$  decay channel with  $\sqrt{s} = 13$  TeV pp collisions at the CMS experiment*, Tech. Rep. CMS-PAS-HIG-16-004, CERN, Geneva, 2016. <https://cds.cern.ch/record/2139578>. 110
- [743] CMS Collaboration, *Search for BSM  $t\bar{t}$  Production in the Boosted All-Hadronic Final State*, Tech. Rep. CMS-PAS-EXO-11-006, 2011. 110
- [744] ATLAS, CMS Collaboration, S. Fleischmann, *Boosted top quark techniques and searches for  $t\bar{t}$  resonances at the LHC*, *J. Phys. Conf. Ser.* **452** (2013) no. 1, 012034. 110
- [745] ATLAS, CMS Collaboration, J. Pilot, *Boosted Top Quarks, Top Pair Resonances, and Top Partner Searches at the LHC*, *EPJ Web Conf.* **60** (2013) 09003. 110
- [746] ATLAS Collaboration, *Performance of boosted top quark identification in 2012 ATLAS data*, Tech. Rep. ATLAS-CONF-2013-084, ATLAS-COM-CONF-2013-074, 2013. 110
- [747] CMS Collaboration, S. Chatrchyan et al., *Search for Anomalous  $t\bar{t}$  Production in the Highly-Boosted All-Hadronic Final State*, *JHEP* **1209** (2012) 029, [arXiv:1204.2488 \[hep-ex\]](#). 110
- [748] CMS Collaboration Collaboration, *Search for pair-produced vector-like quarks of charge  $-1/3$  decaying to  $bH$  using boosted Higgs jet-tagging in pp collisions at  $\sqrt{s} = 8$  TeV*, Tech. Rep. CMS-PAS-B2G-14-001, CERN, Geneva, 2014. <https://cds.cern.ch/record/1752557>. 110
- [749] CMS Collaboration Collaboration, *Search for top-Higgs resonances in all-hadronic final states using jet substructure methods*, Tech. Rep. CMS-PAS-B2G-14-002, CERN, Geneva, 2014. <https://cds.cern.ch/record/1706121>. 110
- [750] CMS Collaboration, V. Khachatryan et al., *Search for vector-like  $T$  quarks decaying to top quarks and Higgs bosons in the all-hadronic channel using jet substructure*, *JHEP* **06** (2015) 080, [arXiv:1503.01952 \[hep-ex\]](#). 110
- [751] CMS Collaboration, V. Khachatryan et al., *Search for a massive resonance decaying*

- into a Higgs boson and a  $W$  or  $Z$  boson in hadronic final states in proton-proton collisions at  $\sqrt{s} = 8$  TeV, *JHEP* **02** (2016) 145, [arXiv:1506.01443 \[hep-ex\]](#). 110
- [752] ATLAS Collaboration, G. Aad et al., *Search for high-mass diboson resonances with boson-tagged jets in proton-proton collisions at  $\sqrt{s} = 8$  TeV with the ATLAS detector*, *JHEP* **12** (2015) 055, [arXiv:1506.00962 \[hep-ex\]](#). 110
- [753] ATLAS Collaboration, M. Aaboud et al., *Searches for heavy diboson resonances in pp collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, [arXiv:1606.04833 \[hep-ex\]](#). 110
- [754] ATLAS Collaboration, M. Aaboud et al., *Search for heavy resonances decaying to a  $Z$  boson and a photon in pp collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, *Phys. Lett.* **B764** (2017) 11–30, [arXiv:1607.06363 \[hep-ex\]](#). 110
- [755] ATLAS Collaboration, M. Aaboud et al., *Search for dark matter produced in association with a hadronically decaying vector boson in pp collisions at  $\sqrt{s}=13$  TeV with the ATLAS detector*, [arXiv:1608.02372 \[hep-ex\]](#). 110
- [756] ATLAS Collaboration Collaboration, *Search for resonances with boson-tagged jets in  $15.5 \text{ fb}^{-1}$  of pp collisions at  $\sqrt{s} = 13$  TeV collected with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2016-055, CERN, Geneva, Aug, 2016. <http://cds.cern.ch/record/2206137>. 110
- [757] *Search for diboson resonances in the  $llq$  final state in pp collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2015-071, CERN, Geneva, Dec, 2015. <http://cds.cern.ch/record/2114843>. 110
- [758] *Search for diboson resonances in the  $\nu\nu qq$  final state in pp collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2015-068, CERN, Geneva, Dec, 2015. <http://cds.cern.ch/record/2114840>. 110
- [759] CMS Collaboration Collaboration, *Search for dark matter in final states with an energetic jet, or a hadronically decaying  $W$  or  $Z$  boson using  $12.9 \text{ fb}^{-1}$  of data at  $\sqrt{s} = 13$  TeV*, Tech. Rep. CMS-PAS-EXO-16-037, CERN, Geneva, 2016. <http://cds.cern.ch/record/2205746>. 110
- [760] CMS Collaboration Collaboration, *Search for new physics in a boosted hadronic monotop final state using  $12.9 \text{ fb}^{-1}$  of  $\sqrt{s} = 13$  TeV data*, Tech. Rep. CMS-PAS-EXO-16-040, CERN, Geneva, 2016. <https://cds.cern.ch/record/2205286>. 110
- [761] CMS Collaboration, V. Khachatryan et al., *Search for dark matter in proton-proton collisions at 8 TeV with missing transverse momentum and vector boson tagged jets*, *JHEP* **12** (2016) 083, [arXiv:1607.05764 \[hep-ex\]](#). [Erratum: *JHEP*08,035(2017)]. 110
- [762] CMS Collaboration Collaboration, *Searches for invisible Higgs boson decays with the CMS detector.*, Tech. Rep. CMS-PAS-HIG-16-016, CERN, Geneva, 2016. <https://cds.cern.ch/record/2201136>. 110
- [763] CMS Collaboration Collaboration, *Search for top quark-antiquark resonances in the all-hadronic final state at  $\sqrt{s}=13$  TeV*, Tech. Rep. CMS-PAS-B2G-15-003, CERN, Geneva, 2016. <https://cds.cern.ch/record/2160237>. 110
- [764] CMS Collaboration Collaboration, *Search for dark matter in association with a boosted top quark in the all hadronic final state*, Tech. Rep. CMS-PAS-EXO-16-017, CERN, Geneva, 2016. <https://cds.cern.ch/record/2160266>. 110
- [765] I. Feige, M. D. Schwartz, I. W. Stewart, and J. Thaler, *Precision Jet Substructure from Boosted Event Shapes*, *Phys. Rev. Lett.* **109** (2012) 092001, [arXiv:1204.3898 \[hep-ph\]](#). 110
- [766] M. Field, G. Gur-Ari, D. A. Kosower, L. Mannelli, and G. Perez, *Three-Prong*

- Distribution of Massive Narrow QCD Jets*, *Phys. Rev.* **D87** (2013) no. 9, 094013, [arXiv:1212.2106 \[hep-ph\]](#). 110
- [767] M. Dasgupta, A. Fregoso, S. Marzani, and A. Powling, *Jet substructure with analytical methods*, *Eur. Phys. J.* **C73** (2013) no. 11, 2623, [arXiv:1307.0013 \[hep-ph\]](#). 110, 117
- [768] M. Dasgupta, A. Powling, and A. Siodmok, *On jet substructure methods for signal jets*, *JHEP* **08** (2015) 079, [arXiv:1503.01088 \[hep-ph\]](#). 110
- [769] M. Seymour, *Jet shapes in hadron collisions: Higher orders, resummation and hadronization*, *Nucl. Phys.* **B513** (1998) 269–300, [arXiv:hep-ph/9707338 \[hep-ph\]](#). 110
- [770] H.-n. Li, Z. Li, and C.-P. Yuan, *QCD resummation for jet substructures*, *Phys. Rev. Lett.* **107** (2011) 152001, [arXiv:1107.4535 \[hep-ph\]](#). 110
- [771] A. J. Larkoski, *QCD Analysis of the Scale-Invariance of Jets*, *Phys. Rev.* **D86** (2012) 054004, [arXiv:1207.1437 \[hep-ph\]](#). 110
- [772] M. Jankowiak and A. J. Larkoski, *Angular Scaling in Jets*, *JHEP* **1204** (2012) 039, [arXiv:1201.2688 \[hep-ph\]](#). 110
- [773] Y.-T. Chien and I. Vitev, *Jet Shape Resummation Using Soft-Collinear Effective Theory*, *JHEP* **1412** (2014) 061, [arXiv:1405.4293 \[hep-ph\]](#). 110
- [774] Y.-T. Chien, *Resummation of Jet Shapes and Extracting Properties of the Quark-Gluon Plasma*, *Int. J. Mod. Phys. Conf. Ser.* **37** (2015) 1560047, [arXiv:1411.0741 \[hep-ph\]](#). 110
- [775] J. Isaacson, H.-n. Li, Z. Li, and C. P. Yuan, *Factorization for substructures of boosted Higgs jets*, [arXiv:1505.06368 \[hep-ph\]](#). 110
- [776] A. J. Larkoski, I. Moulton, and D. Neill, *Toward Multi-Differential Cross Sections: Measuring Two Angularities on a Single Jet*, *JHEP* **09** (2014) 046, [arXiv:1401.4458 \[hep-ph\]](#). 110
- [777] M. Procura, W. J. Waalewijn, and L. Zeune, *Resummation of Double-Differential Cross Sections and Fully-Unintegrated Parton Distribution Functions*, *JHEP* **02** (2015) 117, [arXiv:1410.6483 \[hep-ph\]](#). 110
- [778] D. Bertolini, J. Thaler, and J. R. Walsh, *The First Calculation of Fractional Jets*, *JHEP* **1505** (2015) 008, [arXiv:1501.01965 \[hep-ph\]](#). 110
- [779] B. Bhattacharjee, S. Mukhopadhyay, M. M. Nojiri, Y. Sakaki, and B. R. Webber, *Associated jet and subjet rates in light-quark and gluon jet discrimination*, *JHEP* **1504** (2015) 131, [arXiv:1501.04794 \[hep-ph\]](#). 110
- [780] A. J. Larkoski, I. Moulton, and D. Neill, *Analytic Boosted Boson Discrimination*, *JHEP* **05** (2016) 117, [arXiv:1507.03018 \[hep-ph\]](#). 110, 111, 117, 123, 127
- [781] M. Dasgupta, L. Schunk, and G. Soyez, *Jet shapes for boosted jet two-prong decays from first-principles*, *JHEP* **04** (2016) 166, [arXiv:1512.00516 \[hep-ph\]](#). 110, 116
- [782] Z.-B. Kang, F. Ringer, and I. Vitev, *Jet substructure using semi-inclusive jet functions within SCET*, [arXiv:1606.07063 \[hep-ph\]](#). 110
- [783] A. Hornig, Y. Makris, and T. Mehen, *Jet Shapes in Dijet Events at the LHC in SCET*, *JHEP* **04** (2016) 097, [arXiv:1601.01319 \[hep-ph\]](#). 110
- [784] A. J. Larkoski, I. Moulton, and D. Neill, *Analytic Boosted Boson Discrimination at the Large Hadron Collider*, [arXiv:1708.06760 \[hep-ph\]](#). 110, 117, 123
- [785] A. J. Larkoski, I. Moulton, and D. Neill, *Factorization and Resummation for Groomed Multi-Prong Jet Shapes*, [arXiv:1710.00014 \[hep-ph\]](#). 110, 117, 123
- [786] J. R. Walsh and S. Zuberi, *Factorization Constraints on Jet Substructure*, [arXiv:1110.5333 \[hep-ph\]](#). 110

- [787] A. J. Larkoski, I. Moult, and D. Neill, *Power Counting to Better Jet Observables*, *JHEP* **12** (2014) 009, [arXiv:1409.6298 \[hep-ph\]](#). 110, 111, 117, 127
- [788] A. J. Larkoski, I. Moult, and D. Neill, *Building a Better Boosted Top Tagger*, *Phys. Rev. D* **91** (2015) no. 3, 034035, [arXiv:1411.0665 \[hep-ph\]](#). 110, 117
- [789] J. Cogan, M. Kagan, E. Strauss, and A. Schwartzman, *Jet-Images: Computer Vision Inspired Techniques for Jet Tagging*, *JHEP* **02** (2015) 118, [arXiv:1407.5675 \[hep-ph\]](#). 110
- [790] L. de Oliveira, M. Kagan, L. Mackey, B. Nachman, and A. Schwartzman, *Jet-images — deep learning edition*, *JHEP* **07** (2016) 069, [arXiv:1511.05190 \[hep-ph\]](#). 110
- [791] L. G. Almeida, M. Backović, M. Cliche, S. J. Lee, and M. Perelstein, *Playing Tag with ANN: Boosted Top Identification with Pattern Recognition*, *JHEP* **07** (2015) 086, [arXiv:1501.05968 \[hep-ph\]](#). 110
- [792] P. Baldi, K. Bauer, C. Eng, P. Sadowski, and D. Whiteson, *Jet Substructure Classification in High-Energy Physics with Deep Neural Networks*, *Phys. Rev. D* **93** (2016) no. 9, 094034, [arXiv:1603.09349 \[hep-ex\]](#). 110
- [793] D. Guest, J. Collado, P. Baldi, S.-C. Hsu, G. Urban, and D. Whiteson, *Jet Flavor Classification in High-Energy Physics with Deep Neural Networks*, [arXiv:1607.08633 \[hep-ex\]](#). 110
- [794] J. S. Conway, R. Bhaskar, R. D. Erbacher, and J. Pilot, *Identification of High-Momentum Top Quarks, Higgs Bosons, and W and Z Bosons Using Boosted Event Shapes*, [arXiv:1606.06859 \[hep-ex\]](#). 110
- [795] J. Barnard, E. N. Dawe, M. J. Dolan, and N. Rajcic, *Parton Shower Uncertainties in Jet Substructure Analyses with Deep Neural Networks*, [arXiv:1609.00607 \[hep-ph\]](#). 110
- [796] I. Moult, L. Necib, and J. Thaler, *New Angles on Energy Correlation Functions*, *JHEP* **12** (2016) 153, [arXiv:1609.07483 \[hep-ph\]](#). 111, 115, 116, 117, 127
- [797] P. T. Komiske, E. M. Metodiev, and J. Thaler, *Energy flow polynomials: A complete linear basis for jet substructure*, [arXiv:1712.07124 \[hep-ph\]](#). 111
- [798] J. Thaler and K. Van Tilburg, *Identifying Boosted Objects with  $N$ -subjettiness*, *JHEP* **1103** (2011) 015, [arXiv:1011.2268 \[hep-ph\]](#). 111, 115, 116
- [799] J. Thaler and K. Van Tilburg, *Maximizing Boosted Top Identification by Minimizing  $N$ -subjettiness*, *JHEP* **1202** (2012) 093, [arXiv:1108.2701 \[hep-ph\]](#). 111, 115
- [800] J. Dolen, P. Harris, S. Marzani, S. Rappoccio, and N. Tran, *Thinking outside the ROCs: Designing Decorrelated Taggers (DDT) for jet substructure*, *JHEP* **05** (2016) 156, [arXiv:1603.00027 \[hep-ph\]](#). 111, 127
- [801] G. P. Salam, L. Schunk, and G. Soyez, *Dichroic subjettiness ratios to distinguish colour flows in boosted boson tagging*, *JHEP* **03** (2017) 022, [arXiv:1612.03917 \[hep-ph\]](#). 111, 112, 115, 119, 120, 123
- [802] M. Dasgupta, A. Powling, L. Schunk, and G. Soyez, *Improved jet substructure methods:  $Y$ -splitter and variants with grooming*, *JHEP* **12** (2016) 079, [arXiv:1609.07149 \[hep-ph\]](#). 111, 115
- [803] S. Bosch, B. Lange, M. Neubert, and G. Paz, *Factorization and shape function effects in inclusive  $B$  meson decays*, *Nucl. Phys. B* **699** (2004) 335–386, [arXiv:hep-ph/0402094 \[hep-ph\]](#). 113
- [804] I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, *Dissecting Soft Radiation with Factorization*, *Phys. Rev. Lett.* **114** (2015) no. 9, 092001, [arXiv:1405.6722 \[hep-ph\]](#). 113

- [805] M. Cacciari, G. P. Salam, and S. Sapeta, *On the characterisation of the underlying event*, **JHEP** **04** (2010) 065, [arXiv:0912.4926 \[hep-ph\]](#). 113
- [806] CMS Collaboration, A. M. Sirunyan et al., *Inclusive search for a highly boosted Higgs boson decaying to a bottom quark-antiquark pair*, [arXiv:1709.05543 \[hep-ex\]](#). 115
- [807] CMS Collaboration Collaboration, *Inclusive search for the standard model Higgs boson produced in pp collisions at  $\sqrt{s} = 13$  TeV using  $H \rightarrow b\bar{b}$  decays*, Tech. Rep. CMS-PAS-HIG-17-010, CERN, Geneva, 2017. <http://cds.cern.ch/record/2266164>. 115
- [808] CMS Collaboration Collaboration, *Search for light vector resonances decaying to a quark pair produced in association with a jet in proton-proton collisions at  $\sqrt{s} = 13$  TeV*, Tech. Rep. CMS-PAS-EXO-17-001, CERN, Geneva, 2017. <http://cds.cern.ch/record/2264843>. 115
- [809] CMS Collaboration, A. M. Sirunyan et al., *Search for Low Mass Vector Resonances Decaying to Quark-Antiquark Pairs in Proton-Proton Collisions at  $\sqrt{s} = 13$  TeV*, **Phys. Rev. Lett.** **119** (2017) no. 11, 111802, [arXiv:1705.10532 \[hep-ex\]](#). 115
- [810] CMS Collaboration, A. M. Sirunyan et al., *Search for low mass vector resonances decaying into quark-antiquark pairs in proton-proton collisions at  $\sqrt{s} = 13$  TeV*, [arXiv:1710.00159 \[hep-ex\]](#). 115
- [811] ATLAS Collaboration, M. Aaboud et al., *Search for light resonances decaying to boosted quark pairs and produced in association with a photon or a jet in proton-proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, [arXiv:1801.08769 \[hep-ex\]](#). 115
- [812] C. Shimmin, P. Sadowski, P. Baldi, E. Weik, D. Whiteson, E. Goul, and A. Sogaard, *Decorrelated Jet Substructure Tagging using Adversarial Neural Networks*, **Phys. Rev. D** **96** (2017) no. 7, 074034, [arXiv:1703.03507 \[hep-ex\]](#). 115
- [813] J. A. Aguilar-Saavedra, J. H. Collins, and R. K. Mishra, *A generic anti-QCD jet tagger*, **JHEP** **11** (2017) 163, [arXiv:1709.01087 \[hep-ph\]](#). 115
- [814] I. Moutl, B. Nachman, and D. Neill, *Convolved Substructure: Analytically Decorrelating Jet Substructure Observables*, [arXiv:1710.06859 \[hep-ph\]](#). 115
- [815] I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, *N-Jettiness: An Inclusive Event Shape to Veto Jets*, **Phys. Rev. Lett.** **105** (2010) 092002, [arXiv:1004.2489 \[hep-ph\]](#). 115
- [816] I. W. Stewart, F. J. Tackmann, J. Thaler, C. K. Vermilion, and T. F. Wilkason, *XCone: N-jettiness as an Exclusive Cone Jet Algorithm*, **JHEP** **11** (2015) 072, [arXiv:1508.01516 \[hep-ph\]](#). 116
- [817] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, *Longitudinally-invariant  $k_{\perp}$ -clustering algorithms for hadron-hadron collisions*, **Nucl. Phys. B** **406** (1993) 187–224. 116
- [818] A. J. Larkoski, D. Neill, and J. Thaler, *Jet Shapes with the Broadening Axis*, **JHEP** **04** (2014) 017, [arXiv:1401.2158 \[hep-ph\]](#). 116, 117
- [819] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet User Manual*, **Eur. Phys. J. C** **72** (2012) 1896, [arXiv:1111.6097 \[hep-ph\]](#). 116, 175
- [820] A. J. Larkoski and I. Moutl, *The Singular Behavior of Jet Substructure Observables*, **Phys. Rev. D** **93** (2016) 014017, [arXiv:1510.08459 \[hep-ph\]](#). 116
- [821] S. Catani, G. Turnock, and B. R. Webber, *Jet broadening measures in  $e^+e^-$  annihilation*, **Phys. Lett. B** **295** (1992) 269–276. 117
- [822] Y. L. Dokshitzer, A. Lucenti, G. Marchesini, and G. P. Salam, *On the QCD analysis of jet broadening*, **JHEP** **01** (1998) 011, [arXiv:hep-ph/9801324 \[hep-ph\]](#). 117

- [823] A. J. Larkoski and J. Thaler, *Unsafe but Calculable: Ratios of Angularities in Perturbative QCD*, *JHEP* **09** (2013) 137, [arXiv:1307.1699 \[hep-ph\]](#). 118
- [824] A. J. Larkoski, S. Marzani, and J. Thaler, *Sudakov Safety in Perturbative QCD*, *Phys. Rev. D* **91** (2015) no. 11, 111501, [arXiv:1502.01719 \[hep-ph\]](#). 118, 135
- [825] M. Bähr et al., *Herwig++ Physics and Manual*, *Eur. Phys. J. C* **58** (2008) 639–707, [arXiv:0803.0883 \[hep-ph\]](#). 118, 186, 204, 208, 225, 241, 247
- [826] Y. Gao, A. V. Gritsan, Z. Guo, K. Melnikov, M. Schulze, and N. V. Tran, *Spin determination of single-produced resonances at hadron colliders*, *Phys. Rev. D* **81** (2010) 075022, [arXiv:1001.3396 \[hep-ph\]](#). 118
- [827] S. Bolognesi, Y. Gao, A. V. Gritsan, K. Melnikov, M. Schulze, N. V. Tran, and A. Whitbeck, *On the spin and parity of a single-produced resonance at the LHC*, *Phys. Rev. D* **86** (2012) 095031, [arXiv:1208.4018 \[hep-ph\]](#). 118
- [828] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, *Parton Fragmentation and String Dynamics*, *Phys. Rept.* **97** (1983) 31–145. 123, 225
- [829] B. Andersson, *The Lund model*, *Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.* **7** (1997) 1–471. 123
- [830] G. Marchesini and B. R. Webber, *Monte Carlo Simulation of General Hard Processes with Coherent QCD Radiation*, *Nucl. Phys.* **B310** (1988) 461–526. 123
- [831] A. Buckley et al., *General-purpose event generators for LHC physics*, *Phys. Rept.* **504** (2011) 145–233, [arXiv:1101.2599 \[hep-ph\]](#). 123, 216, 224
- [832] P. Z. Skands, *QCD for Collider Physics*, in *Proceedings, High-energy Physics. Proceedings, 18th European School (ESHEP 2010): Raseborg, Finland, June 20 - July 3, 2010*. 2011. [arXiv:1104.2863 \[hep-ph\]](#). 123
- [833] P. Skands, *Introduction to QCD*, [arXiv:1207.2389 \[hep-ph\]](#). 123
- [834] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003. 129, 201
- [835] CMS Collaboration, *The CMS experiment at the CERN LHC*, *JINST* **3** (2008) S08004. 129
- [836] G. Soyez, *Pileup mitigation at the LHC: a theorist’s view*. PhD thesis, 2018. [arXiv:1801.09721 \[hep-ph\]](#). <http://inspirehep.net/record/1651233/files/arXiv:1801.09721.pdf>. 131
- [837] M. Cacciari and G. P. Salam, *Pileup subtraction using jet areas*, *Phys. Lett.* **B659** (2008) 119–126, [arXiv:0707.1378 \[hep-ph\]](#). 131, 201, 202
- [838] J. Tseng and H. Evans, *Sequential recombination algorithm for jet clustering and background subtraction*, *Phys. Rev. D* **88** (2013) 014044, [arXiv:1304.1025 \[hep-ph\]](#). 131
- [839] M. Cacciari, G. P. Salam, and G. Soyez, *Use of charged-track information to subtract neutral pileup*, *Phys. Rev. D* **92** (2015) no. 1, 014003, [arXiv:1404.7353 \[hep-ph\]](#). 131
- [840] *Mitigation of Pileup Effects at the LHC*, <https://indico.cern.ch/event/306155/>, CERN, April 2014. 131
- [841] A. J. Larkoski and J. Thaler, *Aspects of jets at 100 TeV*, *Phys. Rev. D* **90** (2014) no. 3, 034010, [arXiv:1406.7011 \[hep-ph\]](#). 135
- [842] P. Sun, J. Isaacson, C.-P. Yuan, and F. Yuan, *Resummation of High Order Corrections in Higgs Boson Plus Jet Production at the LHC*, *Phys. Lett.* **B769** (2017) 57–61, [arXiv:1602.08133 \[hep-ph\]](#). 139
- [843] S. Dulat, T.-J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt, D. Stump, and C.-P. Yuan, *New parton distribution functions from a global analysis of*

- quantum chromodynamics, *Phys. Rev.* **D93** (2016) no. 3, 033006, [arXiv:1506.07443 \[hep-ph\]](#). 139, 180
- [844] ATLAS collaboration, *Measurement of fiducial, differential and production cross sections in the  $H \rightarrow \gamma\gamma$  decay channel with  $13.3 \text{ fb}^{-1}$  of 13 TeV proton-proton collision data with the ATLAS detector*, ATLAS-CONF-2016-067, 2016. 139
- [845] CMS Collaboration, *Projected performance of Higgs analyses at the HL-LHC for ECFA 2016*, CMS-PAS-FTR-16-002, 2017. 140
- [846] T. Figy, V. Hankele, and D. Zeppenfeld, *Next-to-leading order QCD corrections to Higgs plus three jet production in vector-boson fusion*, *JHEP* **02** (2008) 076, [arXiv:0710.5621 \[hep-ph\]](#). 140
- [847] J. Baglio et al., *VBFNLO: A Parton Level Monte Carlo for Processes with Electroweak Bosons – Manual for Version 2.7.0*, [arXiv:1107.4038 \[hep-ph\]](#). 140, 186, 187
- [848] J. Baglio et al., *Release Note - VBFNLO 2.7.0*, [arXiv:1404.3940 \[hep-ph\]](#). 140, 186, 187
- [849] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040, [arXiv:hep-ph/0409146 \[hep-ph\]](#). 140, 151, 203
- [850] B. Jäger, F. Schissler, and D. Zeppenfeld, *Parton-shower effects on Higgs boson production via vector-boson fusion in association with three jets*, *JHEP* **07** (2014) 125, [arXiv:1405.6950 \[hep-ph\]](#). 140
- [851] M. Rauch and D. Zeppenfeld, *Jet clustering dependence of Higgs boson production in vector-boson fusion*, *Phys. Rev.* **D95** (2017) no. 11, 114015, [arXiv:1703.05676 \[hep-ph\]](#). 140, 186
- [852] R. N. Cahn and S. Dawson, *Production of Very Massive Higgs Bosons*, *Phys. Lett.* **136B** (1984) 196. [Erratum: *Phys. Lett.*138B,464(1984)]. 143
- [853] V. Del Duca, W. Kilgore, C. Oleari, C. Schmidt, and D. Zeppenfeld, *Higgs + 2 jets via gluon fusion*, *Phys. Rev. Lett.* **87** (2001) 122001, [arXiv:hep-ph/0105129 \[hep-ph\]](#). 143
- [854] V. Del Duca, W. Kilgore, C. Oleari, C. Schmidt, and D. Zeppenfeld, *Gluon fusion contributions to  $H + 2$  jet production*, *Nucl. Phys.* **B616** (2001) 367–399, [arXiv:hep-ph/0108030 \[hep-ph\]](#). 143
- [855] J. R. Andersen, T. Binoth, G. Heinrich, and J. M. Smillie, *Loop induced interference effects in Higgs Boson plus two jet production at the LHC*, *JHEP* **02** (2008) 057, [arXiv:0709.3513 \[hep-ph\]](#). 143
- [856] V. Hankele, G. Klämke, D. Zeppenfeld, and T. Figy, *Anomalous Higgs boson couplings in vector boson fusion at the CERN LHC*, *Phys. Rev.* **D74** (2006) 095001, [arXiv:hep-ph/0609075 \[hep-ph\]](#). 143
- [857] G. Klämke and D. Zeppenfeld, *Higgs plus two jet production via gluon fusion as a signal at the CERN LHC*, *JHEP* **04** (2007) 052, [arXiv:hep-ph/0703202 \[hep-ph\]](#). 143
- [858] J. R. Andersen, K. Arnold, and D. Zeppenfeld, *Azimuthal Angle Correlations for Higgs Boson plus Multi-Jet Events*, *JHEP* **06** (2010) 091, [arXiv:1001.3822 \[hep-ph\]](#). 143
- [859] T. Plehn, D. L. Rainwater, and D. Zeppenfeld, *Determining the structure of Higgs couplings at the LHC*, *Phys. Rev. Lett.* **88** (2002) 051801, [arXiv:hep-ph/0105325 \[hep-ph\]](#). 143
- [860] ATLAS Collaboration, G. Aad et al., *Measurements of the W production cross sections in association with jets with the ATLAS detector*, *Eur. Phys. J.* **C75** (2015) no. 2, 82, [arXiv:1409.8639 \[hep-ex\]](#). 143
- [861] ATLAS Collaboration, M. Aaboud et al., *Measurements of electroweak  $Wjj$  production*

- and constraints on anomalous gauge couplings with the ATLAS detector, *Eur. Phys. J.* **C77** (2017) no. 7, 474, [arXiv:1703.04362 \[hep-ex\]](#). 143, 144
- [862] ATLAS Collaboration, M. Aaboud et al., *Measurement of the cross-section for electroweak production of dijets in association with a Z boson in pp collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, *Phys. Lett.* **B775** (2017) 206–228, [arXiv:1709.10264 \[hep-ex\]](#). 143, 144, 145, 146, 194
- [863] J. M. Campbell, R. K. Ellis, and G. Zanderighi, *Next-to-Leading order Higgs + 2 jet production via gluon fusion*, *JHEP* **10** (2006) 028, [arXiv:hep-ph/0608194 \[hep-ph\]](#). 143
- [864] J. M. Campbell, R. K. Ellis, and C. Williams, *Hadronic production of a Higgs boson and two jets at next-to-leading order*, *Phys. Rev.* **D81** (2010) 074023, [arXiv:1001.4495 \[hep-ph\]](#). 143
- [865] J. R. Andersen and J. M. Smillie, *Constructing all-order corrections to multi-jet rates*, *JHEP* **01** (2010) 039, [arXiv:0908.2786 \[hep-ph\]](#). 143
- [866] J. R. Andersen and J. M. Smillie, *Multiple Jets at the LHC with High Energy Jets*, *JHEP* **06** (2011) 010, [arXiv:1101.5394 \[hep-ph\]](#). 143
- [867] J. R. Andersen, T. Hapola, and J. M. Smillie, *W Plus Multiple Jets at the LHC with High Energy Jets*, *JHEP* **1209** (2012) 047, [arXiv:1206.6763 \[hep-ph\]](#). 143
- [868] J. R. Andersen, J. J. Medley, and J. M. Smillie,  *$Z/\gamma^* \rightarrow \ell^+ \ell^-$  plus multiple hard jets in high energy collisions*, *JHEP* **05** (2016) 136, [arXiv:1603.05460 \[hep-ph\]](#). 143
- [869] J. R. Andersen, T. Hapola, A. Maier, and J. M. Smillie, *Higgs Boson Plus Dijets: Higher Order Corrections*, *JHEP* **09** (2017) 065, [arXiv:1706.01002 \[hep-ph\]](#). 143, 144
- [870] D0 Collaboration, V. M. Abazov et al., *Studies of W boson plus jets production in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96$  TeV*, *Phys. Rev.* **D88** (2013) 092001, [arXiv:1302.6508 \[hep-ex\]](#). 143
- [871] ATLAS Collaboration, M. Aaboud et al., *Measurement of differential cross sections and  $W^+/W^-$  cross-section ratios for W boson production in association with jets at  $\sqrt{s} = 8$  TeV with the ATLAS detector*, [arXiv:1711.03296 \[hep-ex\]](#). 143, 144
- [872] Y. L. Dokshitzer, V. A. Khoze, and T. Sjöstrand, *Rapidity gaps in Higgs production*, *Phys. Lett.* **B274** (1992) 116–121. 143, 145
- [873] D. L. Rainwater, R. Szalapski, and D. Zeppenfeld, *Probing color singlet exchange in Z + two jet events at the CERN LHC*, *Phys. Rev.* **D54** (1996) 6680–6689, [arXiv:hep-ph/9605444 \[hep-ph\]](#). 145
- [874] C. Englert, M. McCullough, and M. Spannowsky, *Gluon-initiated associated production boosts Higgs physics*, *Phys. Rev.* **D89** (2014) no. 1, 013013, [arXiv:1310.4828 \[hep-ph\]](#). 146
- [875] R. Harlander, J. Klappert, C. Pandini, and A. Papaefstathiou, *Exploiting the WH/ZH symmetry in the search for New Physics*, to appear . 146, 148
- [876] ATLAS Collaboration, M. Aaboud et al., *Evidence for the  $H \rightarrow b\bar{b}$  decay with the ATLAS detector*, *JHEP* **12** (2017) 024, [arXiv:1708.03299 \[hep-ex\]](#). 148, 149, 150
- [877] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, *MadGraph 5 : Going Beyond*, *JHEP* **06** (2011) 128, [arXiv:1106.0522 \[hep-ph\]](#). 149, 208
- [878] S. Gieseke et al., *Herwig++ 2.5 Release Note*, [arXiv:1102.1672 \[hep-ph\]](#). 149
- [879] K. Arnold et al., *Herwig++ 2.6 Release Note*, [arXiv:1205.4902 \[hep-ph\]](#). 149
- [880] J. Bellm et al., *Herwig++ 2.7 Release Note*, [arXiv:1310.6877 \[hep-ph\]](#). 149
- [881] S. Alioli, P. Nason, C. Oleari, and E. Re, *NLO Higgs boson production via gluon fusion*

- matched with shower in POWHEG, *JHEP* **04** (2009) 002, [arXiv:0812.0578 \[hep-ph\]](#). 150, 152
- [882] E. Bagnaschi, R. V. Harlander, H. Mantler, A. Vicini, and M. Wiesemann, *Resummation ambiguities in the Higgs transverse-momentum spectrum in the Standard Model and beyond*, *JHEP* **01** (2016) 090, [arXiv:1510.08850 \[hep-ph\]](#). 150
- [883] R. Corke and T. Sjöstrand, *Improved Parton Showers at Large Transverse Momenta*, *Eur. Phys. J.* **C69** (2010) 1–18, [arXiv:1003.2384 \[hep-ph\]](#). 154
- [884] B. Cabouat and T. Sjöstrand, *Some Dipole Shower Studies*, [arXiv:1710.00391 \[hep-ph\]](#). 154
- [885] C. F. Berger, C. Marcantonini, I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, *Higgs Production with a Central Jet Veto at NNLL+NNLO*, *JHEP* **04** (2011) 092, [arXiv:1012.4480](#). 160, 161
- [886] I. W. Stewart and F. J. Tackmann, *Theory Uncertainties for Higgs and Other Searches Using Jet Bins*, *Phys. Rev.* **D85** (2012) 034011, [arXiv:1107.2117 \[hep-ph\]](#). 160, 161, 211
- [887] See [https://indico.cern.ch/.../WG1\\_March\\_meeting\\_followup.pdf](https://indico.cern.ch/.../WG1_March_meeting_followup.pdf). 160
- [888] I. W. Stewart, F. J. Tackmann, J. R. Walsh, and S. Zuberi, *Jet  $p_T$  resummation in Higgs production at NNLL'+NNLO*, *Phys. Rev.* **D89** (2014) no. 5, 054001, [arXiv:1307.1808 \[hep-ph\]](#). 161
- [889] The ATLAS Collaboration, The CMS Collaboration, The LHC Higgs Combination Group Collaboration, *Procedure for the LHC Higgs boson search combination in Summer 2011*, Tech. Rep. CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11, CERN, Geneva, Aug, 2011. <https://cds.cern.ch/record/1379837>. 167
- [890] ATLAS Collaboration, *Measurements of Higgs boson properties in the diphoton decay channel with  $36 \text{ fb}^{-1}$  of  $pp$  collision data at  $\sqrt{s} = 13 \text{ TeV}$  with the ATLAS detector*, [arXiv:1802.04146 \[hep-ex\]](#). 170
- [891] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, *Eur. Phys. J.* **C71** (2011) 1554, [arXiv:1007.1727 \[physics.data-an\]](#). [Erratum: *Eur. Phys. J.*C73,2501(2013)]. 170
- [892] CMS Collaboration, S. Chatrchyan et al., *Measurement of the  $t\bar{t}$  production cross section in the dilepton channel in  $pp$  collisions at  $\sqrt{s} = 7 \text{ TeV}$* , *JHEP* **11** (2012) 067, [arXiv:1208.2671 \[hep-ex\]](#). 173
- [893] CMS Collaboration, S. Chatrchyan et al., *Measurement of the  $t\bar{t}$  production cross section in the dilepton channel in  $pp$  collisions at  $\sqrt{s} = 8 \text{ TeV}$* , *JHEP* **02** (2014) 024, [arXiv:1312.7582 \[hep-ex\]](#). [Erratum: *JHEP*02,102(2014)]. 173
- [894] CMS Collaboration, V. Khachatryan et al., *Measurement of the top quark pair production cross section in proton-proton collisions at  $\sqrt{s} = 13 \text{ TeV}$* , *Phys. Rev. Lett.* **116** (2016) no. 5, 052002, [arXiv:1510.05302 \[hep-ex\]](#). 173
- [895] ATLAS Collaboration, M. Aaboud et al., *Measurement of the  $t\bar{t}$  production cross-section using  $e\mu$  events with  $b$ -tagged jets in  $pp$  collisions at  $\sqrt{s}=13 \text{ TeV}$  with the ATLAS detector*, *Phys. Lett.* **B761** (2016) 136–157, [arXiv:1606.02699 \[hep-ex\]](#). [Erratum: *Phys. Lett.*B772,879(2017)]. 173
- [896] ATLAS, CMS Collaboration, A. K. Nayak, *Recent inclusive  $t\bar{t}$  cross section measurements*, *PoS CKM2016* (2017) 117. 173
- [897] CMS Collaboration, S. Chatrchyan et al., *Measurement of differential top-quark pair production cross sections in  $pp$  collisions at  $\sqrt{s} = 7 \text{ TeV}$* , *Eur. Phys. J.* **C73** (2013) no. 3, 2339, [arXiv:1211.2220 \[hep-ex\]](#). 173

- [898] ATLAS Collaboration, G. Aad et al., *Measurements of normalized differential cross sections for  $t\bar{t}$  production in  $pp$  collisions at  $\sqrt{s} = 7$  TeV using the ATLAS detector*, *Phys. Rev.* **D90** (2014) no. 7, 072004, [arXiv:1407.0371 \[hep-ex\]](#). 173
- [899] ATLAS Collaboration, G. Aad et al., *Differential top-antitop cross-section measurements as a function of observables constructed from final-state particles using  $pp$  collisions at  $\sqrt{s} = 7$  TeV in the ATLAS detector*, *JHEP* **06** (2015) 100, [arXiv:1502.05923 \[hep-ex\]](#). 173
- [900] CMS Collaboration, V. Khachatryan et al., *Measurement of differential cross sections for top quark pair production using the lepton+jets final state in proton-proton collisions at 13 TeV*, *Phys. Rev.* **D95** (2017) no. 9, 092001, [arXiv:1610.04191 \[hep-ex\]](#). 173
- [901] CMS Collaboration, A. M. Sirunyan et al., *Measurement of normalized differential  $t$ - $t$ -bar cross sections in the dilepton channel from  $pp$  collisions at  $\sqrt{s} = 13$  TeV*, [arXiv:1708.07638 \[hep-ex\]](#). 173
- [902] ATLAS Collaboration, M. Aaboud et al., *Measurements of top-quark pair differential cross-sections in the  $e\mu$  channel in  $pp$  collisions at  $\sqrt{s} = 13$  TeV using the ATLAS detector*, *Eur. Phys. J.* **C77** (2017) no. 5, 292, [arXiv:1612.05220 \[hep-ex\]](#). 173
- [903] ATLAS Collaboration, M. Aaboud et al., *Measurements of  $t\bar{t}$  differential cross-sections of highly boosted top quarks decaying to all-hadronic final states in  $pp$  collisions at  $\sqrt{s} = 13$  TeV using the ATLAS detector*, [arXiv:1801.02052 \[hep-ex\]](#). 173
- [904] S. Frixione, P. Nason, and B. R. Webber, *Matching NLO QCD and parton showers in heavy flavor production*, *JHEP* **08** (2003) 007, [arXiv:hep-ph/0305252 \[hep-ph\]](#). 173
- [905] S. Frixione, P. Nason, and G. Ridolfi, *A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction*, *JHEP* **09** (2007) 126, [arXiv:0707.3088 \[hep-ph\]](#). 173
- [906] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour, and B. R. Webber, *HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, *JHEP* **01** (2001) 010, [arXiv:hep-ph/0011363 \[hep-ph\]](#). 173, 204
- [907] A. Bredenstein, A. Denner, S. Dittmaier, and S. Pozzorini, *NLO QCD corrections to  $pp \rightarrow t$  anti- $t$   $b$  anti- $b$  +  $X$  at the LHC*, *Phys. Rev. Lett.* **103** (2009) 012002, [arXiv:0905.0110 \[hep-ph\]](#). 173
- [908] A. Bredenstein, A. Denner, S. Dittmaier, and S. Pozzorini, *NLO QCD Corrections to Top Anti-Top Bottom Anti-Bottom Production at the LHC: 2. full hadronic results*, *JHEP* **03** (2010) 021, [arXiv:1001.4006 \[hep-ph\]](#). 173
- [909] G. Bevilacqua, M. Czakon, C. G. Papadopoulos, R. Pittau, and M. Worek, *Assault on the NLO Wishlist:  $pp \rightarrow t$  anti- $t$   $b$  anti- $b$* , *JHEP* **09** (2009) 109, [arXiv:0907.4723 \[hep-ph\]](#). 173
- [910] G. Bevilacqua, M. Czakon, C. G. Papadopoulos, and M. Worek, *Dominant QCD Backgrounds in Higgs Boson Analyses at the LHC: A Study of  $pp \rightarrow t$  anti- $t$  + 2 jets at Next-To-Leading Order*, *Phys. Rev. Lett.* **104** (2010) 162002, [arXiv:1002.4009 \[hep-ph\]](#). 173
- [911] G. Bevilacqua, M. Czakon, C. G. Papadopoulos, and M. Worek, *Hadronic top-quark pair production in association with two jets at Next-to-Leading Order QCD*, *Phys. Rev.* **D84** (2011) 114017, [arXiv:1108.2851 \[hep-ph\]](#). 173
- [912] A. Kardos and Z. Trócsányi, *Hadroproduction of  $t$  anti- $t$  pair with a  $b$  anti- $b$  pair using PowHel*, *J. Phys.* **G41** (2014) 075005, [arXiv:1303.6291 \[hep-ph\]](#). 173
- [913] F. Cascioli, P. Maierhöfer, N. Moretti, S. Pozzorini, and F. Siegert, *NLO matching for  $t\bar{t}b\bar{b}$  production with massive  $b$ -quarks*, *Phys. Lett.* **B734** (2014) 210–214,

- [arXiv:1309.5912 \[hep-ph\]](#). 173, 245
- [914] S. Höche, J. Huang, G. Luisoni, M. Schönherr, and J. Winter, *Zero and one jet combined next-to-leading order analysis of the top quark forward-backward asymmetry*, *Phys. Rev.* **D88** (2013) no. 1, 014040, [arXiv:1306.2703 \[hep-ph\]](#). 173, 176, 179
- [915] M. Czakon, H. B. Hartanto, M. Kraus, and M. Worek, *Matching the Nagy-Soper parton shower at next-to-leading order*, *JHEP* **06** (2015) 033, [arXiv:1502.00925 \[hep-ph\]](#). 173
- [916] G. Bevilacqua, M. V. Garzelli, and A. Kardos,  *$t\bar{t}b\bar{b}$  hadroproduction with massive bottom quarks with PowHel*, [arXiv:1709.06915 \[hep-ph\]](#). 173
- [917] T. Ježo, J. M. Lindert, N. Moretti, and S. Pozzorini, *New NLOPS predictions for  $t\bar{t} + b$ -jet production at the LHC*, [arXiv:1802.00426 \[hep-ph\]](#). 173, 245
- [918] J. H. Kühn, A. Scharf, and P. Uwer, *Electroweak effects in top-quark pair production at hadron colliders*, *Eur. Phys. J.* **C51** (2007) 37–53, [arXiv:hep-ph/0610335 \[hep-ph\]](#). 173
- [919] W. Bernreuther, M. Fuecker, and Z.-G. Si, *Weak interaction corrections to hadronic top quark pair production*, *Phys. Rev.* **D74** (2006) 113005, [arXiv:hep-ph/0610334 \[hep-ph\]](#). 173
- [920] J. H. Kühn, A. Scharf, and P. Uwer, *Weak Interactions in Top-Quark Pair Production at Hadron Colliders: An Update*, *Phys. Rev.* **D91** (2015) no. 1, 014020, [arXiv:1305.5773 \[hep-ph\]](#). 173
- [921] W. Bernreuther and Z.-G. Si, *Distributions and correlations for top quark pair production and decay at the Tevatron and LHC.*, *Nucl. Phys.* **B837** (2010) 90–121, [arXiv:1003.3926 \[hep-ph\]](#). 173
- [922] W. Hollik and D. Pagani, *The electroweak contribution to the top quark forward-backward asymmetry at the Tevatron*, *Phys. Rev.* **D84** (2011) 093003, [arXiv:1107.2606 \[hep-ph\]](#). 173
- [923] D. Pagani, I. Tsiniikos, and M. Zaro, *The impact of the photon PDF and electroweak corrections on  $t\bar{t}$  distributions*, *Eur. Phys. J.* **C76** (2016) no. 9, 479, [arXiv:1606.01915 \[hep-ph\]](#). 173, 174
- [924] S. Höche, F. Krauss, S. Schumann, and F. Siegert, *QCD matrix elements and truncated showers*, *JHEP* **05** (2009) 053, [arXiv:0903.1219 \[hep-ph\]](#). 173, 174, 176, 187
- [925] S. Höche, F. Krauss, M. Schönherr, and F. Siegert, *QCD matrix elements + parton showers: The NLO case*, *JHEP* **04** (2013) 027, [arXiv:1207.5030 \[hep-ph\]](#). 173, 174, 209, 247
- [926] T. Gehrmann, S. Höche, F. Krauss, M. Schönherr, and F. Siegert, *NLO QCD matrix elements + parton showers in  $e^+e^- \rightarrow$  hadrons*, *JHEP* **01** (2013) 144, [arXiv:1207.5031 \[hep-ph\]](#). 173, 174, 209, 247
- [927] M. Schönherr and F. Krauss, *Soft Photon Radiation in Particle Decays in SHERPA*, *JHEP* **12** (2008) 018, [arXiv:0810.5071 \[hep-ph\]](#). 173, 176
- [928] A. Manohar, P. Nason, G. P. Salam, and G. Zanderighi, *How bright is the proton? A precise determination of the photon parton distribution function*, *Phys. Rev. Lett.* **117** (2016) no. 24, 242002, [arXiv:1607.04266 \[hep-ph\]](#). 174
- [929] S. Höche, F. Krauss, M. Schönherr, and F. Siegert, *A critical appraisal of NLO+PS matching methods*, *JHEP* **09** (2012) 049, [arXiv:1111.1220 \[hep-ph\]](#). 174, 179, 209
- [930] S. Höche, F. Krauss, M. Schönherr, and F. Siegert, *W + n-Jet predictions at the Large Hadron Collider at next-to-leading order matched with a parton shower*, *Phys. Rev. Lett.* **110** (2013) no. 5, 052001, [arXiv:1201.5882 \[hep-ph\]](#). 174, 209

- [931] S. Höche and M. Schönherr, *Uncertainties in next-to-leading order plus parton shower matched simulations of inclusive jet and dijet production*, *Phys. Rev.* **D86** (2012) 094042, [arXiv:1208.2815 \[hep-ph\]](#). 174
- [932] S. Höche, F. Krauss, M. Schönherr, and F. Siegert, *NLO matrix elements and truncated showers*, *JHEP* **08** (2011) 123, [arXiv:1009.1127 \[hep-ph\]](#). 174, 187
- [933] S. Höche, F. Krauss, S. Pozzorini, M. Schönherr, J. M. Thompson, and K. C. Zapp, *Triple vector boson production through Higgs-Strahlung with NLO multijet merging*, *Phys. Rev.* **D89** (2014) no. 9, 093015, [arXiv:1403.7516 \[hep-ph\]](#). 174
- [934] Particle Data Group Collaboration, K. A. Olive et al., *Review of Particle Physics*, *Chin. Phys.* **C38** (2014) 090001. 175
- [935] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber, *QCD matrix elements + parton showers*, *JHEP* **11** (2001) 063, [arXiv:hep-ph/0109231 \[hep-ph\]](#). 176
- [936] CMS Collaboration, V. Khachatryan et al., *Measurement of the top quark mass using proton-proton data at  $\sqrt{s} = 7$  and 8 TeV*, *Phys. Rev.* **D93** (2016) no. 7, 072004, [arXiv:1509.04044 \[hep-ex\]](#). 178
- [937] ATLAS Collaboration, *Measurement of the top quark mass in the  $t\bar{t} \rightarrow$  lepton+jets channel from  $\sqrt{s} = 8$  TeV ATLAS data*, ATLAS-CONF-2017-071, 2017. <https://cds.cern.ch/record/??> 178
- [938] B. Pearson, *Top quark mass in ATLAS*, in *10th International Workshop on Top Quark Physics (TOP2017) Braga, Portugal, September 17-22, 2017*. 2017. [arXiv:1711.09763 \[hep-ex\]](#). 178
- [939] CMS Collaboration, A. Castro, *Recent Top Quark Mass Measurements from CMS*, in *10th International Workshop on Top Quark Physics (TOP2017) Braga, Portugal, September 17-22, 2017*. 2017. [arXiv:1712.01027 \[hep-ex\]](#). 178
- [940] K. G. Chetyrkin and M. Steinhauser, *Short distance mass of a heavy quark at order  $\alpha_s^3$* , *Phys. Rev. Lett.* **83** (1999) 4001–4004, [arXiv:hep-ph/9907509 \[hep-ph\]](#). 178
- [941] K. Melnikov and T. v. Ritbergen, *The Three loop relation between the  $\overline{MS}$ -bar and the pole quark masses*, *Phys. Lett.* **B482** (2000) 99–108, [arXiv:hep-ph/9912391 \[hep-ph\]](#). 178
- [942] A. H. Hoang, A. Jain, C. Lepenik, V. Mateu, M. Preisser, I. Scimemi, and I. W. Stewart, *The  $\overline{MS}$  Mass and the  $\mathcal{O}(\Lambda_{\text{QCD}})$  Renormalon Sum Rule*, [arXiv:1704.01580 \[hep-ph\]](#). 178
- [943] J. Gao and A. S. Papanastasiou, *Top-quark pair-production and decay at high precision*, *Phys. Rev.* **D96** (2017) 051501, [arXiv:1705.08903 \[hep-ph\]](#). 178
- [944] R. Frederix, S. Frixione, A. S. Papanastasiou, S. Prestel, and P. Torrielli, *Off-shell single-top production at NLO matched to parton showers*, *JHEP* **06** (2016) 027, [arXiv:1603.01178 \[hep-ph\]](#). 179
- [945] B. Chokoufi Nejad, W. Kilian, J. M. Lindert, S. Pozzorini, J. Reuter, and C. Weiss, *NLO QCD predictions for off-shell  $t\bar{t}$  and  $t\bar{t}H$  production and decay at a linear collider*, *JHEP* **12** (2016) 075, [arXiv:1609.03390 \[hep-ph\]](#). 179
- [946] ATLAS Collaboration, M. Aaboud et al., *Measurement of the top quark mass in the  $t\bar{t} \rightarrow$  dilepton channel from  $\sqrt{s} = 8$  TeV ATLAS data*, *Phys. Lett.* **B761** (2016) 350–371, [arXiv:1606.02179 \[hep-ex\]](#). 180
- [947] M. Jezabek and J. H. Kühn, *Semileptonic Decays of Top Quarks*, *Phys. Lett.* **B207** (1988) 91–96. 180
- [948] ATLAS Collaboration, G. Aad et al., *Evidence for Electroweak Production of  $W^\pm W^\pm jj$  in  $pp$  Collisions at  $\sqrt{s} = 8$  TeV with the ATLAS Detector*, *Phys. Rev. Lett.* **113** (2014)

- no. 14, 141803, [arXiv:1405.6241 \[hep-ex\]](#). 184
- [949] CMS Collaboration, V. Khachatryan et al., *Study of vector boson scattering and search for new physics in events with two same-sign leptons and two jets*, *Phys. Rev. Lett.* **114** (2015) no. 5, 051801, [arXiv:1410.6315 \[hep-ex\]](#). 184
- [950] CMS Collaboration, A. M. Sirunyan et al., *Observation of electroweak production of same-sign  $W$  boson pairs in the two jet and two same-sign lepton final state in proton-proton collisions at  $\sqrt{s} = 13$  TeV*, [arXiv:1709.05822 \[hep-ex\]](#). 184
- [951] ATLAS Collaboration, M. Aaboud et al., *Measurement of  $W^\pm W^\pm$  vector-boson scattering and limits on anomalous quartic gauge couplings with the ATLAS detector*, *Phys. Rev.* **D96** (2017) 012007, [arXiv:1611.02428 \[hep-ex\]](#). 184
- [952] CMS Collaboration, A. M. Sirunyan et al., *Measurement of vector boson scattering and constraints on anomalous quartic couplings from events with four leptons and two jets in proton-proton collisions at  $\sqrt{s} = 13$  TeV*, *Phys. Lett.* **B774** (2017) 682–705, [arXiv:1708.02812 \[hep-ex\]](#). 184
- [953] ATLAS Collaboration, G. Aad et al., *Measurements of  $W^\pm Z$  production cross sections in  $pp$  collisions at  $\sqrt{s} = 8$  TeV with the ATLAS detector and limits on anomalous gauge boson self-couplings*, *Phys. Rev.* **D93** (2016) no. 9, 092004, [arXiv:1603.02151 \[hep-ex\]](#). 184, 185, 188
- [954] C. F. Anders et al., *VBSCan Split 2017 Workshop Summary*, 2018. [arXiv:1801.04203 \[hep-ph\]](#). 185, 192, 194, 197
- [955] Bozzi, Giuseppe and Jäger, Barbara and Oleari, Carlo and Zeppenfeld, Dieter, *Next-to-leading order QCD corrections to  $W+Z$  and  $W-Z$  production via vector-boson fusion*, *Phys. Rev.* **D75** (2007) 073004, [arXiv:hep-ph/0701105 \[hep-ph\]](#). 185
- [956] F. Campanario, M. Kerner, L. D. Ninh, and D. Zeppenfeld, *WZ Production in Association with Two Jets at Next-to-Leading Order in QCD*, *Phys. Rev. Lett.* **111** (2013) no. 5, 052003, [arXiv:1305.1623 \[hep-ph\]](#). 185
- [957] M. Rauch and S. Plätzer, *Parton Shower Matching Systematics in Vector-Boson-Fusion WW Production*, *Eur. Phys. J.* **C77** (2017) no. 5, 293, [arXiv:1605.07851 \[hep-ph\]](#). 186, 188
- [958] M. Rauch, *Vector-Boson Fusion and Vector-Boson Scattering*, [arXiv:1610.08420 \[hep-ph\]](#). 186
- [959] S. Plätzer and S. Gieseke, *Dipole Showers and Automated NLO Matching in Herwig++*, *Eur. Phys. J.* **C72** (2012) 2187, [arXiv:1109.6256 \[hep-ph\]](#). 186, 208
- [960] J. Bellm, G. Nail, S. Plätzer, P. Schichtel, and A. Siódmok, *Parton Shower Uncertainties with Herwig 7: Benchmarks at Leading Order*, *Eur. Phys. J.* **C76** (2016) no. 12, 665, [arXiv:1605.01338 \[hep-ph\]](#). 186, 225, 241
- [961] S. Höche, S. Schumann, and F. Siegert, *Hard photon production and matrix-element parton-shower merging*, *Phys. Rev.* **D81** (2010) 034026, [arXiv:0912.3501 \[hep-ph\]](#). 187, 215
- [962] E. Boos et al., *Generic user process interface for event generators*, in *Physics at TeV colliders. Proceedings, Euro Summer School, Les Houches, France, May 21-June 1, 2001*. 2001. [arXiv:hep-ph/0109068 \[hep-ph\]](#). 187
- [963] J. Alwall et al., *A standard format for Les Houches Event Files*, *Comput. Phys. Commun.* **176** (2007) 300–304, [hep-ph/0609017](#). 187
- [964] A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page, M. Rüfenacht, M. Schönherr, and G. Watt, *LHAPDF6: parton density access in the LHC precision era*, *Eur. Phys. J.* **C75** (2015) 132, [arXiv:1412.7420 \[hep-ph\]](#). 187, 208, 219, 232, 241

- [965] CMS Collaboration, *Prospects for the study of vector boson scattering in same sign WW and WZ interactions at the HL-LHC with the upgraded CMS detector*, Tech. Rep. CMS-PAS-SMP-14-008, CERN, Geneva, 2016. <https://cds.cern.ch/record/2220831>. 188
- [966] CMS Collaboration, A. M. Sirunyan et al., *Electroweak production of two jets in association with a Z boson in proton-proton collisions at  $\sqrt{s} = 13$  TeV*, Submitted to: Eur. Phys. J. C (2017) , [arXiv:1712.09814 \[hep-ex\]](https://arxiv.org/abs/1712.09814). 194
- [967] CMS Collaboration, V. Khachatryan et al., *Event generator tunes obtained from underlying event and multiparton scattering measurements*, Eur. Phys. J. **C76** (2016) no. 3, 155, [arXiv:1512.00815 \[hep-ex\]](https://arxiv.org/abs/1512.00815). 200
- [968] CMS Collaboration, *The CMS Experiment at the CERN LHC*, JINST **3** (2008) S08004. 201
- [969] CMS Collaboration, *Particle-flow reconstruction and global event description with the CMS detector*, JINST **12** (2017) no. 10, P10003, [arXiv:1706.04965 \[physics.ins-det\]](https://arxiv.org/abs/1706.04965). 201
- [970] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. **C77** (2017) 490, [arXiv:1603.02934 \[hep-ex\]](https://arxiv.org/abs/1603.02934). 201, 202, 203, 205
- [971] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, public note **ATLAS-CONF-2014-018** (2014) . 202
- [972] CMS Collaboration, *Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS*, JINST **6** (2011) P11002, [arXiv:1107.4277 \[physics.ins-det\]](https://arxiv.org/abs/1107.4277). 202
- [973] CDF Collaboration, R. Field, *The underlying event in large transverse momentum charged jet and  $Z^-$  boson production*, Int. J. Mod. Phys. **A16S1A** (2001) 250–254. 202
- [974] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin, and C. P. Yuan, *New parton distributions for collider physics*, Phys. Rev. **D82** (2010) 074024, [arXiv:1007.2241 \[hep-ph\]](https://arxiv.org/abs/1007.2241). 203
- [975] B. Nachman, P. Nef, A. Schwartzman, M. Swiatlowski, and C. Wanotayaroj, *Jets from Jets: Re-clustering as a tool for large radius jet reconstruction and grooming at the LHC*, JHEP **02** (2015) 075, [arXiv:1407.2922 \[hep-ph\]](https://arxiv.org/abs/1407.2922). 206
- [976] ATLAS Collaboration, *Jet reclustering and close-by effects in ATLAS run II*, public note **ATLAS-CONF-2017-062** (2017) . 206
- [977] CMS Collaboration, V. Khachatryan et al., *Measurement of the double-differential inclusive jet cross section in proton-proton collisions at  $\sqrt{s} = 13$  TeV*, Eur. Phys. J. **C76** (2016) no. 8, 451, [arXiv:1605.04436 \[hep-ex\]](https://arxiv.org/abs/1605.04436). 207, 241
- [978] V. Del Duca, A. Frizzo, and F. Maltoni, *Higgs boson production in association with three jets*, JHEP **05** (2004) 064, [arXiv:hep-ph/0404013 \[hep-ph\]](https://arxiv.org/abs/hep-ph/0404013). 207
- [979] L. J. Dixon, E. W. N. Glover, and V. V. Khoze, *MHV rules for Higgs plus multi-gluon amplitudes*, JHEP **12** (2004) 015, [arXiv:hep-th/0411092 \[hep-th\]](https://arxiv.org/abs/hep-th/0411092). 207
- [980] S. D. Badger, E. W. N. Glover, and V. V. Khoze, *MHV rules for Higgs plus multi-parton amplitudes*, JHEP **03** (2005) 023, [arXiv:hep-th/0412275 \[hep-th\]](https://arxiv.org/abs/hep-th/0412275). 207
- [981] S. Badger, E. W. Nigel Glover, P. Mastrolia, and C. Williams, *One-loop Higgs plus four gluon amplitudes: Full analytic results*, JHEP **01** (2010) 036, [arXiv:0909.4475 \[hep-ph\]](https://arxiv.org/abs/0909.4475). 207
- [982] S. Badger, J. M. Campbell, R. K. Ellis, and C. Williams, *Analytic results for the one-loop NMHV Hqgg amplitude*, JHEP **12** (2009) 035, [arXiv:0910.4481 \[hep-ph\]](https://arxiv.org/abs/0910.4481).

- [983] L. J. Dixon and Y. Sofianatos, *Analytic one-loop amplitudes for a Higgs boson plus four partons*, **JHEP** **08** (2009) 058, [arXiv:0906.0008 \[hep-ph\]](#). 207
- [984] T. Gehrmann, M. Jaquier, E. W. N. Glover, and A. Koukoutsakis, *Two-Loop QCD Corrections to the Helicity Amplitudes for  $H \rightarrow 3$  partons*, **JHEP** **02** (2012) 056, [arXiv:1112.3554 \[hep-ph\]](#). 207
- [985] A. Gehrmann-De Ridder, T. Gehrmann, and E. W. N. Glover, *Gluon-gluon antenna functions from Higgs boson decay*, **Phys. Lett.** **B612** (2005) 49–60, [arXiv:hep-ph/0502110 \[hep-ph\]](#). 208
- [986] A. Gehrmann-De Ridder, T. Gehrmann, and E. W. N. Glover, *Quark-gluon antenna functions from neutralino decay*, **Phys. Lett.** **B612** (2005) 36–48, [arXiv:hep-ph/0501291 \[hep-ph\]](#). 208
- [987] A. Daleo, T. Gehrmann, and D. Maître, *Antenna subtraction with hadronic initial states*, **JHEP** **04** (2007) 016, [arXiv:hep-ph/0612257 \[hep-ph\]](#). 208
- [988] A. Daleo, A. Gehrmann-De Ridder, T. Gehrmann, and G. Luisoni, *Antenna subtraction at NNLO with hadronic initial states: initial-final configurations*, **JHEP** **01** (2010) 118, [arXiv:0912.0374 \[hep-ph\]](#). 208
- [989] T. Gehrmann and P. F. Monni, *Antenna subtraction at NNLO with hadronic initial states: real-virtual initial-initial configurations*, **JHEP** **12** (2011) 049, [arXiv:1107.4037 \[hep-ph\]](#). 208
- [990] R. Boughezal, A. Gehrmann-De Ridder, and M. Ritzmann, *Antenna subtraction at NNLO with hadronic initial states: double real radiation for initial-initial configurations with two quark flavours*, **JHEP** **02** (2011) 098, [arXiv:1011.6631 \[hep-ph\]](#). 208
- [991] A. Gehrmann-De Ridder, T. Gehrmann, and M. Ritzmann, *Antenna subtraction at NNLO with hadronic initial states: double real initial-initial configurations*, **JHEP** **10** (2012) 047, [arXiv:1207.5779 \[hep-ph\]](#). 208
- [992] X. Chen, T. Gehrmann, N. Glover, and M. Jaquier, *Higgs plus one jet production at NNLO*, **PoS RADCOR2015** (2016) 056, [arXiv:1604.04085 \[hep-ph\]](#). 208
- [993] J. Bellm, S. Gieseke, and S. Plätzer, *Merging NLO Multi-jet Calculations with Improved Unitarization*, [arXiv:1705.06700 \[hep-ph\]](#). 208
- [994] S. Plätzer and S. Gieseke, *Coherent Parton Showers with Local Recoils*, **JHEP** **01** (2011) 024, [arXiv:0909.5593 \[hep-ph\]](#). 208, 225
- [995] S. Catani, B. R. Webber, and G. Marchesini, *QCD coherent branching and semiinclusive processes at large  $x$* , **Nucl. Phys.** **B349** (1991) 635–654. 208, 209, 215, 216
- [996] H. van Deurzen, N. Greiner, G. Luisoni, P. Mastrolia, E. Mirabella, et al., *NLO QCD corrections to the production of Higgs plus two jets at the LHC*, **Phys. Lett.** **B721** (2013) 74–81, [arXiv:1301.0493 \[hep-ph\]](#). 213
- [997] G. Cullen, H. van Deurzen, N. Greiner, G. Luisoni, P. Mastrolia, et al., *NLO QCD corrections to Higgs boson production plus three jets in gluon fusion*, **Phys. Rev. Lett.** **111** (2013) 131801, [arXiv:1307.4737 \[hep-ph\]](#). 213
- [998] D. Amati, A. Bassetto, M. Ciafaloni, G. Marchesini, and G. Veneziano, *A treatment of hard processes sensitive to the infrared structure of QCD*, **Nucl. Phys.** **B173** (1980) 429. 215, 216
- [999] S. Plätzer and M. Sjö Dahl, *The Sudakov Veto Algorithm Reloaded*, **Eur. Phys. J. Plus** **127** (2012) 26, [arXiv:1108.6180 \[hep-ph\]](#). 215
- [1000] L. Lönnblad, *Fooling Around with the Sudakov Veto Algorithm*, **Eur. Phys. J.** **C73** (2013) 2350, [arXiv:1211.7204 \[hep-ph\]](#). 215

- [1001] J. Bellm, S. Plätzer, P. Richardson, A. Siódmok, and S. Webster, *Reweighting Parton Showers*, *Phys. Rev.* **D94** (2016) no. 3, 034028, [arXiv:1605.08256 \[hep-ph\]](#). 215, 224
- [1002] S. Mrenna and P. Skands, *Automated Parton-Shower Variations in Pythia 8*, *Phys. Rev.* **D94** (2016) no. 7, 074005, [arXiv:1605.08352 \[hep-ph\]](#). 215, 217, 220, 221, 222, 223, 224, 225
- [1003] J. Kodaira and L. Trentadue, *Summing Soft Emission in QCD*, *Phys. Lett.* **112B** (1982) 66. 216
- [1004] C. T. H. Davies and W. J. Stirling, *Nonleading Corrections to the Drell-Yan Cross-Section at Small Transverse Momentum*, *Nucl. Phys.* **B244** (1984) 337–348. 216
- [1005] C. T. H. Davies, B. R. Webber, and W. J. Stirling, *Drell-Yan Cross-Sections at Small Transverse Momentum*, *Nucl. Phys.* **B256** (1985) 413. [1,I.95(1984)]. 216
- [1006] S. Catani, E. D’Emilio, and L. Trentadue, *The Gluon Form-factor to Higher Orders: Gluon Gluon Annihilation at Small  $Q^{\perp}$  transverse*, *Phys. Lett.* **B211** (1988) 335–342. 216
- [1007] S. Höche, D. Reichelt, and F. Siegert, *Momentum conservation and unitarity in parton showers and NLL resummation*, *JHEP* **01** (2018) 118, [arXiv:1711.03497 \[hep-ph\]](#). 217, 221
- [1008] T. Sjöstrand, *The Lund Monte Carlo for Jet Fragmentation*, *Comput. Phys. Commun.* **27** (1982) 243. 217
- [1009] W. T. Giele, D. A. Kosower, and P. Z. Skands, *Higher-Order Corrections to Timelike Jets*, *Phys. Rev.* **D84** (2011) 054003, [arXiv:1102.2126 \[hep-ph\]](#). 217
- [1010] N. Fischer, S. Prestel, M. Ritzmann, and P. Skands, *Vincia for Hadron Colliders*, *Eur. Phys. J.* **C76** (2016) no. 11, 589, [arXiv:1605.06142 \[hep-ph\]](#). 217
- [1011] JADE and OPAL Collaboration, P. Pfeifenschneider et al., *QCD analyses and determinations of  $\alpha_s$  in  $e^+ e^-$  annihilation at energies between 35-GeV and 189-GeV*, *Eur. Phys. J.* **C17** (2000) 19–51, [arXiv:hep-ex/0001055](#). 219, 220
- [1012] J. Rojo et al., *The PDF4LHC report on PDFs and LHC data: Results from Run I and preparation for Run II*, *J. Phys.* **G42** (2015) 103103, [arXiv:1507.00556 \[hep-ph\]](#). 224
- [1013] T. Sjöstrand and P. Z. Skands, *Transverse-momentum-ordered showers and interleaved multiple interactions*, *Eur. Phys. J.* **C39** (2005) 129–154, [arXiv:hep-ph/0408302 \[hep-ph\]](#). 225
- [1014] A. Buckley, H. Hoeth, H. Lacker, H. Schulz, and J. E. von Seggern, *Systematic event generator tuning for the LHC*, *Eur. Phys. J.* **C65** (2010) 331–357, [arXiv:0907.2973 \[hep-ph\]](#). 225
- [1015] L3 Collaboration, P. Achard et al., *Studies of hadronic event structure in  $e^+e^-$  annihilation from 30-GeV to 209-GeV with the L3 detector*, *Phys. Rept.* **399** (2004) 71–174, [arXiv:hep-ex/0406049 \[hep-ex\]](#). 227, 228, 239
- [1016] T. Sjöstrand, *A Model for Initial State Parton Showers*, *Phys. Lett.* **157B** (1985) 321–325. 228
- [1017] S. Catani, L. Cieri, D. de Florian, G. Ferrera, and M. Grazzini, *Universality of transverse-momentum resummation and hard factors at the NNLO*, *Nucl. Phys.* **B881** (2014) 414–443, [arXiv:1311.1654 \[hep-ph\]](#). 230
- [1018] T. Becher and M. Neubert, *Drell-Yan Production at Small  $q_T$ , Transverse Parton Distributions and the Collinear Anomaly*, *Eur. Phys. J.* **C71** (2011) 1665, [arXiv:1007.4005 \[hep-ph\]](#). 230
- [1019] S. Alekhin et al., *HERA and the LHC - A workshop on the implications of HERA for LHC physics: Proceedings Part A*, [hep-ph/0601012](#). 240

- [1020] M. Dasgupta, L. Magnea, and G. P. Salam, *Non-perturbative QCD effects in jets at hadron colliders*, *JHEP* **02** (2008) 055, [arXiv:0712.3014 \[hep-ph\]](#). 240
- [1021] T. Sjöstrand, ed., *Why MB/UE physics prefer LO PDFs*.  
<http://home.thep.lu.se/~torbjorn/pdfdoc/pdfwarning.pdf>. 241
- [1022] A. Buckley, *Sensitivities to PDFs in parton shower MC generator reweighting and tuning*, [arXiv:1601.04229 \[hep-ph\]](#). 241
- [1023] S. Gieseke, *Uncertainties of Sudakov form-factors*, *JHEP* **01** (2005) 058,  
[arXiv:hep-ph/0412342 \[hep-ph\]](#). 241
- [1024] CMS Collaboration, V. Khachatryan et al., *Measurements of the associated production of a Z boson and b jets in pp collisions at  $\sqrt{s} = 8$  TeV*, *Eur. Phys. J.* **C77** (2017) no. 11, 751, [arXiv:1611.06507 \[hep-ex\]](#). 245
- [1025] CMS Collaboration, V. Khachatryan et al., *Measurement of the production cross section of a W boson in association with two b jets in pp collisions at  $\sqrt{s} = 8$  TeV*, *Eur. Phys. J.* **C77** (2017) no. 2, 92, [arXiv:1608.07561 \[hep-ex\]](#). 245
- [1026] ATLAS Collaboration, G. Aad et al., *Measurement of differential production cross-sections for a Z boson in association with b-jets in 7 TeV proton-proton collisions with the ATLAS detector*, *JHEP* **10** (2014) 141, [arXiv:1407.3643 \[hep-ex\]](#). 245, 246
- [1027] ATLAS Collaboration, G. Aad et al., *Measurement of the cross-section for W boson production in association with b-jets in pp collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector*, *JHEP* **06** (2013) 084, [arXiv:1302.2929 \[hep-ex\]](#). 245
- [1028] CMS Collaboration, S. Chatrchyan et al., *Measurement of the cross section and angular correlations for associated production of a Z boson with b hadrons in pp collisions at  $\sqrt{s} = 7$  TeV*, *JHEP* **12** (2013) 039, [arXiv:1310.1349 \[hep-ex\]](#). 245, 246
- [1029] F. Krauss, D. Napoletano, and S. Schumann, *Simulating b-associated production of Z and Higgs bosons with the SHERPA event generator*, *Phys. Rev.* **D95** (2017) no. 3, 036012, [arXiv:1612.04640 \[hep-ph\]](#). 245, 246
- [1030] F. Krauss and D. Napoletano, *Towards a fully massive five-flavour scheme*, [arXiv:1712.06832 \[hep-ph\]](#). 245, 246
- [1031] T. Sjöstrand, *On the  $g \rightarrow Q\bar{Q}$  Rate*,  
<http://home.thep.lu.se/~torbjorn/pdfdoc/g2qqbarsplit.pdf>. 246, 247, 248
- [1032] A. Gehrmann-De Ridder, M. Ritzmann, and P. Z. Skands, *Timelike Dipole-Antenna Showers with Massive Fermions*, *Phys. Rev.* **D85** (2012) 014013, [arXiv:1108.6172 \[hep-ph\]](#). 246
- [1033] R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, R. Pittau, and P. Torrielli, *W and Z/ $\gamma^*$  boson production in association with a bottom-antibottom pair*, *JHEP* **09** (2011) 061, [arXiv:1106.6019 \[hep-ph\]](#). 246
- [1034] ATLAS Collaboration, M. Aaboud et al., *Measurement of b-hadron pair production with the ATLAS detector in proton-proton collisions at  $\sqrt{s} = 8$  TeV*, *JHEP* **11** (2017) 062, [arXiv:1705.03374 \[hep-ex\]](#). 246, 249
- [1035] P. Ilten, N. L. Rodd, J. Thaler, and M. Williams, *Disentangling Heavy Flavor at Colliders*, *Phys. Rev.* **D96** (2017) no. 5, 054019, [arXiv:1702.02947 \[hep-ph\]](#). 249