Matched predictions for Higgs production via heavy-quark loops in the SM and beyond

Johan Alwall,1 Qiang Li,2 and Fabio Maltoni3

1Theoretical Physics Department, Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, IL 60510, USA
2Paul Scherrer Institut, CH–5232 Villigen PSI, Switzerland
3Centre for Cosmology, Particle Physics and Phenomenology (CP3), Université Catholique de Louvain, Chemin du Cyclotron 2, B-1348 Louvain-la-Neuve, Belgium

The main Higgs production channel at hadron colliders is gluon fusion via heavy-quark loops. We present the results of a fully exclusive simulation of gluon fusion Higgs production based on the matrix elements for $h + 0, 1, 2$ partons including full heavy-quark loop dependence, matched to a parton shower. We consider a Higgs with standard model couplings as well as models where the Higgs has enhanced couplings to bottom quarks ($b$-philic). We study the most relevant kinematic distributions, such as jet and Higgs $p_T$ spectra and find that matched samples provide an accurate description of the final state. For the SM Higgs, we confirm the excellent accuracy of the large heavy-quark-mass approximation also in differential distributions over all phase space, with significant effects arising only at large $p_T$. For a $b$-philic Higgs however, the loops have a dramatic impact on the kinematics of the Higgs as well as of the jets and need to be accounted for exactly to achieve reliable event simulations.

I. INTRODUCTION

The CERN Large Hadron Collider (LHC) is running at a center of mass energy of 7 TeV and it has already accumulated several inverse femtobarns of integrated luminosity per experiment. One of its main goals is to explore the details of electroweak symmetry breaking, and in particular to establish the existence of a Higgs sector of or beyond the Standard Model (SM).

At the LHC, Higgs boson production mainly proceeds via a quantum effect, gluon fusion (GF) \cite{1}. This is induced by heavy-quark loops, in particular the bottom and the top quarks, the latter being by far the dominant one in the SM. For a not too heavy Higgs boson ($m_h \lesssim 2m_t$), and in appropriate kinematic regions ($p_T^h \lesssim m_t$), the top quark can be integrated out, resulting, to a very good approximation, in a simple, non-renormalizable effective field theory, $L_{\text{HEFT}} = \frac{1}{4} \frac{h}{v^2} F_{\mu\nu}^a F_{\mu\nu}^a$ (HEFT) \cite{2,3}, $v$ being the Higgs field vacuum expectation value and $F_{\mu\nu}^a$ the QCD field tensor. The next-to-leading order (NLO) QCD corrections \cite{8} were calculated decades ago in both HEFT as well as in the full SM and found to be very large ($\sigma^{\text{NLO}}/\sigma^{\text{LO}} \sim 2$). This motivated the formidable endeavour of the next-to-next-to-leading order (NNLO) QCD calculations, which have been fully evaluated in HEFT \cite{9,11}. The exact NNLO calculation involves three loop massive diagrams, and is currently out of reach. However, recently, the finite top-quark mass effects to the total NNLO prediction have been estimated through a power expansion \cite{12,15} and found to have a negligible impact on total rates. Soft gluon resummation effects have also been studied in HEFT at NNLL \cite{16,17}. On the other hand, it is known that in the hard tails of differential distributions or even in special kinematics regimes, such as at small-$x$, loop effects need to be accounted for exactly. So far, the recommended best predictions for Higgs GF inclusive production rates in the standard model \cite{19} are based on the NNLO+NNLL results in HEFT, while keeping the heavy quark mass dependence at NLO+NNLL \cite{20}.

In Beyond the SM (BSM) theories, GF becomes sensitive to all colored states in the spectrum with significant couplings to the Higgs(es). Even though the bookkeeping becomes more involved, as long as such states are heavy, an effective field approximation can still be used and QCD corrections can be computed as in the SM. The only genuine complication arises, not from additional heavy BSM particles, but from the possibility of bottom quarks to have enhanced couplings to the scalar (or pseudo-scalar) states of the theory. In SUSY, and more generally in type II two-Higgs-doublet-model scenarios, this corresponds to a large $\tan \beta$ scenario (where $\beta = v_1/v_2$, $v_{1,2}$ being the vacuum expectation values of the Higgs doublets coupling to down- and up-type fermions, respectively). In this case, the HEFT approximation cannot be employed and the accuracy of the best available predictions goes down to NLO \cite{19}.

Being of primary importance, total rates and Higgs kinematic distributions are now quite well predicted and also available via public codes such as ResBos \cite{21} and HqT \cite{22,23}. Differential $p_T$ distributions accurate to NLO can be also obtained via HIGLU \cite{24} as well as via HPro \cite{25}, which both keep the exact bottom- and top-quarks mass loop dependence and therefore can be used also for predictions of scalar Higgs in BSM. However, in experimental analyses, it is also crucial to get as precise predictions as possible for exclusive observables that involve extra jets, such as the jet $p_T$ spectra and the jet rates, at both parton and hadron level.
search strategies and in particular to curb the very large backgrounds, current analyses both at Tevatron and at the LHC select 0-, 1- and 2-jet events and perform independent analyses on each sample. The final systematic uncertainties are effected by both the theoretical and experimental ones of such a jet-bin based separation, see, e.g., Ref. [27]. In the HEFT, fully exclusive parton- and hadron-level calculations can be performed by Parton Shower (PS) programs such as Pythia [28], Herwig [29] and Sherpa [30] in the soft and collinear approximation, or with NLO QCD codes matched with parton showers: via the MC@NLO [31] and POWHEG [32, 33] methods. However, beyond the HEFT, no fully exclusive prediction has been available so far. The reason is that one needs to compromise between the validity of HEFT and the complexity of higher loop calculations. It is however possible to get full exclusive control at hadron level on the complex event topology at the LHC, while still reaching approximately NLL accuracy, with the help of recent sophisticated matching methods between matrix elements and parton showers [34, 35].

In PS programs, QCD radiation is generated in the collinear and soft approximation, using Markov chain techniques based on Sudakov form factors. Hard and widely separated jets are thus poorly described in this approach. On the other hand, tree-level fixed order amplitudes can provide reliable predictions in the hard region, while failing in the collinear and soft limits. To combine both descriptions and avoid double counting or gaps between samples with different multiplicity, an appropriate matching method is required. Several algorithms have been proposed over the years: the CKKW method, based on a shower veto and therefore on event re-weighting [36] and MLM schemes, based on event rejection [37, 38].

In this work, we report on the first matched simulation of Higgs production in gluon fusion that retains the full kinematic dependence on the heavy-quark loops, in the SM as well as in generic scenarios with enhanced Higgs couplings with bottom quarks, which we dub “b-philic Higgs”. The paper is organized as follows. We first describe our methodology. Then we present our results for a SM Higgs. We show that the matching procedure provides reliable results both at the Tevatron and especially at the LHC and that the effects from massive quark loops are indeed mild over all phenomenologically relevant phase space. The b-philic Higgs is considered in the following section, where it is shown that loop effects must be included exactly. Gluon fusion production is also compared to a matrix-element matched sample for $b\bar{b} \rightarrow h$, which is the dominant production mode in this scenario. We draw our conclusions in the last section.

II. METHOD

Our study is based on the $k_T$-MLM and shower-$k_T$ matching schemes implemented in MADGRAPH/MadEvent [39], interfaced with Pythia 6.4 for parton shower and hadronization. As explained below, we find it convenient to include the effects of the heavy-quark loop by simply reweighting the events generated via tree-level HEFT amplitudes.

In the $k_T$-jet MLM matching schemes [37, 40], matrix element multi-parton events are produced with a minimum separation $k_T$ cutoff of $Q_{\min}^{ME}$. For every event, the final-state partons are clustered according to the $k_T$ algorithm, and the $k_T$ value for each clustering vertex corresponding to a QCD emission is used as renormalization scale for $\alpha_S$ in that vertex. For the central hard $2 \rightarrow 1$ or $2 \rightarrow 2$ process, the transverse mass $m_T^2 = p_T^2 + m^2$ of the particle(s) produced in the central process is used as factorization and renormalization scale. Subsequently, this event is passed to the Pythia parton-shower generator. There, one of two schemes is employed. Either, the final partons (after parton showering) are clustered into jets, using the $k_T$ algorithm with a jet cutoff of $Q_{\min}^{jet} > Q_{\min}^{ME}$. The jets are considered to be matched to the original partons if $k_T$(parton, jet) is smaller than the cutoff $Q_{\min}^{jet}$. If any parton is not matched to a jet, the event is discarded. For events with parton multiplicity smaller than the highest multiplicity, the number of jets must be equal to the number of partons. We call this scheme the the $k_T$-MLM scheme. Alternatively, no matching between shower jets and partons is done. Instead, an event is retained provided that the hardest emission in the Pythia parton shower is below the scale $Q_{\min}^{jet}$ (or, for events from the highest multiplicity, below the scale $Q_{\min}^{parton}$ of the softest parton in the event). This is called the “shower-$k_T$” scheme, and allows for the matching scale $Q_{\min}^{ME}$ to be set equal to the matrix element cutoff scale $Q_{\min}^{ME}$. The two matching schemes have been shown to give equivalent results [40], but for the case of b-philic Higgs, the shower-$k_T$ scheme allows for lower matching scales and it is therefore more efficient.

In order to take into account the full kinematic dependence of the heavy quark loop in Higgs production, the full one-loop amplitudes for all possible subprocesses contributing to $h+0, 1, 2$ partons have been calculated. Analytic expressions have been generated with FeynArts 3.5 [41], and manipulated with FormCalc 5.3 [42]. The tensor integrals have been evaluated with the help of the LoopTools-2.5 package [43], which employs the reduction method introduced in Ref. [43] for pentagons, and Passarino-Veltman reduction for the lower point tensors. The resulting regular scalar integrals are evaluated with the FF package [44]. We have also implemented the reduction method for pentagon tensor integrals as proposed in Ref. [45] for better numerical stability. The codes have been used and validated against known results in a previous study [46]. The final implementation of the calculation includes the contributions from bottom and top quarks and their (destructive) interference. The evaluation of multi-parton loop amplitudes is, in general, computationally quite expensive. Moreover, in the case of inclusive matched samples, an ef-
Table I: Reference values for total cross sections for Higgs production in the SM and considering only $b$-loops, used for the normalization of the inclusive samples. Results have been obtained via the HNNLO [17] and BBH0 NNLO [18] codes, with $m_t = 173$ GeV, $m_b = 4.6$ GeV, $\mu_R = \mu_F = m_h$ and employing the MSTW2008NNLO pdf set [21].

<table>
<thead>
<tr>
<th>Cross section</th>
<th>Higgs mass [GeV]</th>
<th>Tevatron LHC @ 7 TeV</th>
</tr>
</thead>
<tbody>
<tr>
<td>$gg \rightarrow h$ (SM)</td>
<td>140</td>
<td>0.672 pb</td>
</tr>
<tr>
<td>$gg \rightarrow h$ (SM)</td>
<td>500</td>
<td>0.003 pb</td>
</tr>
<tr>
<td>$gg \rightarrow h$ (b-only)</td>
<td>140</td>
<td>3.0 fb</td>
</tr>
<tr>
<td>$b\bar{b} \rightarrow h$</td>
<td>140</td>
<td>4.55 fb</td>
</tr>
</tbody>
</table>

To illustrate the results of our simulations for the Tevatron and the LHC at 7 TeV for a standard model Higgs we show a few relevant observables in Figs. 1-3. We define jets via the $k_T$ algorithm, with the distance measure between parton $i$ and beam $B$, or partons $i$ and $j$ as $k_T^{ij} = p_T^i, k_T^{ij} = \min \left( p_T^i, p_T^j \right) \frac{\sqrt{2(\cosh \Delta y - \cos \Delta \phi)}}{D}$. Here $y$ is the rapidity and $\phi$ is the azimuthal angle around the beam direction. The resolution parameter is set to $D = 1$. Jets are required to satisfy $|y| < 4.5$ and $p_T^j > 30$ GeV. For sake of simplicity, we adopt Yukawa couplings corresponding to the pole masses, i.e., for the top quark $m_t = 173$ GeV and for the bottom-quark mass $m_b = 4.6$ GeV. Other quark masses are neglected. Throughout our calculation, we adopt the CTEQ6L1 parton distribution functions (PDFs) [50] with the core process renormalization and factorization scales $\mu_r = \mu_f = m_T^{\min}$, where $m_T^{\min} \equiv \sqrt{(p_T^k)^2 + m_T^2}$. For the matching performed in MADGRAPH/MADEVENT, the $k_T$-MLM scheme is chosen, with $Q_{\min}^{\text{ME}} = 30$ GeV and $Q_{\min}^{\text{jet}} = 50$ GeV.

In Figs. 1-3 we show the Higgs $p_T$ distributions for $m_h = 140$ GeV in gluon fusion production at the Tevatron. Results in the HEFT and with full loop dependence (LOOP) are compared to the default PYTHIA implementation (which accounts for $2 \rightarrow 2$ matrix element corrections) and to the NNLO+NNLL results as obtained by HqT [22, 23]. Curves normalized to the corresponding total cross sections of Table I.

### III. SM HIGGS PRODUCTION

To illustrate the results of our simulations for the Tevatron and the LHC at 7 TeV for a standard model Higgs we show a few relevant observables in Figs. 1-3. We define jets via the $k_T$ algorithm, with the distance measure between parton $i$ and beam $B$, or partons $i$ and $j$ as $k_T^{ij} = p_T^i, k_T^{ij} = \min \left( p_T^i, p_T^j \right) \frac{\sqrt{2(\cosh \Delta y - \cos \Delta \phi)}}{D}$. Here $y$ is the rapidity and $\phi$ is the azimuthal angle around the beam direction. The resolution parameter is set to $D = 1$. Jets are required to satisfy $|y| < 4.5$ and $p_T^j > 30$ GeV. For sake of simplicity, we adopt Yukawa couplings corresponding to the pole masses, i.e., for the top quark $m_t = 173$ GeV and for the bottom-quark mass $m_b = 4.6$ GeV. Other quark masses are neglected. Throughout our calculation, we adopt the CTEQ6L1 parton distribution functions (PDFs) [50] with the core process renormalization and factorization scales $\mu_r = \mu_f = m_T^{\min}$, where $m_T^{\min} \equiv \sqrt{(p_T^k)^2 + m_T^2}$. For the matching performed in MADGRAPH/MADEVENT, the $k_T$-MLM scheme is chosen, with $Q_{\min}^{\text{ME}} = 30$ GeV and $Q_{\min}^{\text{jet}} = 50$ GeV.

In Figs. 1-3 we show the Higgs $p_T$ distribution for Standard Model Higgs GF production at the Tevatron with $m_h = 140$ in a range of $p_T$ relevant for experimental analysis. We compare matched results in the HEFT theory and in the full theory (LOOP) with PYTHIA with $2 \rightarrow 2$ matrix element corrections. We also include the predictions from the analytic computation at NNLO+NNLL as obtained by HqT [22, 23]. The curves are all normalized to the NNLO+NNLL predictions. The three Monte-Carlo based predictions agree very well in all the shown range of $p_T$, suggesting that for this observable, higher multiplicity matrix element corrections (starting from $2 \rightarrow 3$) and loop effects are not important. This is the case also for jet $p_T$ distributions (not shown) in the same kinematical range. The NNLO+NNLL prediction, on the other hand, suggests a slightly softer Higgs spectrum.

In Figs. 1-3 we show the Higgs and jet $p_T$ distribu-
FIG. 2: SM Higgs $p_T$ distributions for $m_h = 140$ GeV and $m_h = 500$ GeV in gluon fusion production at 7 TeV LHC. In the upper plot results in the HEFT and with full loop dependence (LOOP) are compared over a large range of $p_T$ values to the default PYTHIA implementation, which accounts for $2 \to 2$ matrix element corrections. In the lower plot the low-$p_T$ range is compared to the NNLO+NNLL results as obtained by $HqT$[22,23]. Curves normalized to the corresponding total cross sections of Table I.

FIG. 3: Jet $p_T$ distributions for associated jets in gluon fusion production of $m_h = 140$ GeV and $m_h = 500$ GeV Higgs bosons at 7 TeV LHC.

butions, such as the $p_T$ of the Higgs, do agree remarkably well with the best available predictions, for example NNLO+NNLL at small Higgs $p_T$, and offer improved and easy-to-use predictions for other key observables such as the jet rates and distributions. In addition, for heavy Higgs masses and/or large $p_T$, loop effects, even though marginal for phenomenology, can also be taken into account in the same approach, if needed.

IV. B-PHILIC HIGGS PRODUCTION

In this section we present the results of a simulation of a $b$-philic Higgs. Parameters are the same as in the previous section, except that, as explained below, the top Yukawa coupling is set to zero and the matrix-element matching in MADGRAPH/MADEVENT is performed through the shower-$k_T$ matching scheme with $Q_{\text{match}} = 10$ GeV.

In Fig. 4 we show the $p_T^j$ distributions for GF production at the 7 TeV LHC of a $b$-philic Higgs with $m_h = 140$ GeV. We remind the reader that in our cal-
calculation the bottom-quark and top-quark masses can be chosen independently as well as the value of the corresponding Yukawa couplings. We can therefore study the production of scalars with arbitrary couplings to the heavy quarks such as those appearing in a generic two Higgs doublet model or in the minimal supersymmetric standard model. For the sake of illustration, we define a simplified scenario where the Higgs coupling to the top quark is set to zero. In so doing, we study the Higgs and jet distributions relative to a “large tan β” scenario with bottom-quark loops dominating. Note that for simplicity we keep the same normalization as in the standard model, i.e., \( m_b \sqrt{2} = m_t \) with \( m_b = 4.6 \text{ GeV} \), as the corresponding cross sections in enhanced scenarios can be easily obtained by rescaling.

In the \( b \)-philic Higgs production, the particle running in the loop is nearly massless, and there is no region in \( m_b \) or \( p_T \) where an effective description is valid. This also means that a parton-shower generator alone has no possibility of correctly describing the effects of jet radiation, and genuine loop matrix-elements plus a matched description are needed for achieving reliable simulations.

In fact, the largest production cross section for a \( b \)-philic Higgs does not come from loop induced gluon fusion, but from tree-level \( bb \) fusion. Phenomenologically, it is therefore very important to be able to also generate events for this kind of process, which typically leads to final states with more \( b \)-jets than the GF production. We do so by matching tree-level matrix elements for \( h+0, 1, 2 \) partons (with a \( hbb \) vertex) in the five flavor scheme to the parton shower. In so doing we provide a complete and consistent event simulation of inclusive Higgs production in a \( b \)-philic (or large \( \tan \beta \)) scenario. We note in passing that a four-flavor scheme, i.e., starting from the leading order process \( gg \to bbb \), could also be employed. While this latter approach has some important advantages, it also offers complications with respect to the simpler five flavor scheme. A detailed comparison between the two approaches, which are known to be compatible at the level of total cross sections (see, e.g., 19 and references therein), would certainly be welcome. Being beyond the

\[\begin{align*}
\frac{d\sigma}{dp_T} \mid \text{Tevatron} \\
\frac{d\sigma}{dp_T} \mid \text{LHC @ 7 TeV}
\end{align*}\]
V. CONCLUSIONS

In summary, we have presented the first fully exclusive simulation of gluon fusion inclusive Higgs production based on the exact one-loop matrix elements for $h + 0, 1, 2$ partons, matched to Pythia parton showers using multiple matching schemes implemented in MadGraph/MadEvent. We have compared the loop reweighted matched results with the corresponding HEFT results, Pythia results, and, when possible, with NNLO+NNLL predictions. We have considered both the SM Higgs and the case of scalar particles with enhanced couplings to bottom quarks and studied the most relevant kinematic distributions, such as jet and Higgs $p_T$ spectra. Our results highlight the relevance of a complete loop calculation at large $p_T$ for a standard model Higgs and in all phase space for $b$-philic Higgs. Such improved simulations might be particularly relevant in searches performed via multivariate analysis techniques where details about the kinematic distributions of the Higgs decay products and accompanying jets can have significant impact on the results.

We conclude by stressing that the method employed in this work, i.e., using tree-level amplitudes based on an effective theory to generate parton-level events and then reweighting them by the exact loop amplitudes before matching to the shower, is completely general and can therefore be applied to any loop-induced process. Work towards the automatization of this approach in MadGraph 5 via MadLoop is in progress.

Acknowledgments

We thank Paolo Torrielli, Simon de Visscher, Massimiliano Grazzini and Giuseppe Degrassi for helpful discussions. This work is supported by the European Community’s Marie-Curie Research Training Network HEPTOOLS under contract MRTN-CT-2006-035505, US DOE Contract No. DE-AC02-07CH11359, by the National Natural Science Foundation of China, by the IAP Program, BELSPO P6/11-P and the IISN convention 4.4511.10.


