

# UNIVERSITÄT BONN

## Physikalisches Institut

**Measurements of the Top Quark Pair Production  
Cross Section in Lepton+Jets Final States using  
a Topological Multivariate Technique as well as  
Lifetime  $b$ -Tagging  
in Proton-Antiproton Collisions at  $\sqrt{s} = 1.96$  TeV  
with the DØ Detector at the Tevatron**

by  
Tobias Golling

**Abstract:** Two alternative measurements of the  $t\bar{t}$  production cross section at  $\sqrt{s} = 1.96$  TeV in proton-antiproton collisions in the lepton+jets channel are presented. The  $t\bar{t}$  production cross section is extracted by combining the kinematic event information in a multivariate discriminant. The measurement yields  $\sigma_{p\bar{p}\rightarrow t\bar{t}+X} = 5.13_{-1.57}^{+1.76}$  (stat)  $_{-1.10}^{+0.96}$  (syst)  $\pm 0.33$  (lumi) pb in the muon+jets channel, using  $229.1 \text{ pb}^{-1}$ , and in the combination with the electron+jets channel ( $226.3 \text{ pb}^{-1}$ )  $\sigma_{p\bar{p}\rightarrow t\bar{t}+X} = 6.60_{-1.28}^{+1.37}$  (stat)  $_{-1.11}^{+1.25}$  (syst)  $\pm 0.43$  (lumi) pb. The second measurement presented reconstructs explicitly secondary vertices to do lifetime  $b$ -tagging. The measurement combines the muon+jets and the electron+jets channel, using  $158.4 \text{ pb}^{-1}$  and  $168.8 \text{ pb}^{-1}$ , respectively:  $\sigma_{p\bar{p}\rightarrow t\bar{t}+X} = 8.24_{-1.25}^{+1.34}$  (stat)  $_{-1.63}^{+1.89}$  (syst)  $\pm 0.54$  (lumi) pb.

Post address:  
Nussallee 12  
D-53115 Bonn  
Germany



BONN-IR-2005-01  
Bonn University  
January 2005



UNIVERSITÄT BONN  
Physikalisches Institut

**Measurements of the Top Quark Pair Production  
Cross Section in Lepton+Jets Final States using  
a Topological Multivariate Technique as well as  
Lifetime  $b$ -Tagging  
in Proton-Antiproton Collisions at  $\sqrt{s} = 1.96$  TeV  
with the DØ Detector at the Tevatron**

by  
Tobias Golling

Dieser Forschungsbericht wurde als Dissertation von der Mathematisch -  
Naturwissenschaftlichen Fakultät der Universität Bonn angenommen.

Angenommen am: 22. Februar 2005  
Referent: Prof. Dr. N. Wermes  
Korreferent: Prof. Dr. E. Hilger



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Theoretical Aspects</b>	<b>5</b>
2.1	The Physics of the Top Quark . . . . .	6
2.1.1	Properties of the Top Quark . . . . .	6
2.1.2	The Virtual Top Quark . . . . .	9
2.1.3	Top Quark Pair Production via the Strong Interaction	12
2.1.4	Production of the Top Quark via the Weak Interaction	21
2.1.5	Decay of the Top Quark . . . . .	23
2.1.6	Signature of $t\bar{t}$ Events . . . . .	25
2.2	The Physics of the Bottom Quark . . . . .	28
2.3	Signal and Background Signature . . . . .	29
2.3.1	$t\bar{t}$ Events in the Lepton + Jets Final State . . . . .	29
2.3.2	Electroweak $W$ Boson Production in Association with Jets . . . . .	30
2.3.3	QCD Multijet Production . . . . .	34
2.3.4	Small Additional Background Processes . . . . .	35
2.4	Monte Carlo Simulation . . . . .	37
2.4.1	$t\bar{t}$ Signal Simulation . . . . .	40
2.4.2	$W$ -plus-jets Background Simulation . . . . .	40
2.4.3	$Z$ -plus-jets Background Simulation . . . . .	44
2.4.4	Vector Boson Pair Production Background Simulation	44
2.4.5	Single Top Background Simulation . . . . .	44
2.5	Measurement of the Luminosity . . . . .	44
2.5.1	The Total $p\bar{p}$ Cross Section and the Optical Theorem	45
<b>3</b>	<b>Experimental Environment</b>	<b>49</b>
3.1	The Fermilab Accelerator Complex . . . . .	49
3.2	Interactions of Energetic Particles with Matter . . . . .	50
3.2.1	Electrons and Photons . . . . .	51
3.2.2	Muons . . . . .	52
3.2.3	Hadronic Particles . . . . .	52
3.2.4	Neutrinos . . . . .	53

3.3	The DØ Detector . . . . .	53
3.3.1	Coordinate System . . . . .	54
3.3.2	Luminosity System . . . . .	55
3.3.3	The Central Tracking System . . . . .	56
3.3.4	The Calorimeter System . . . . .	60
3.3.5	The Muon System . . . . .	67
3.3.6	The Forward Proton Detector . . . . .	73
3.3.7	The Trigger System . . . . .	74
<b>4</b>	<b>Event Reconstruction and Object Identification</b>	<b>79</b>
4.1	Charged Tracks . . . . .	79
4.2	Primary Vertex . . . . .	80
4.3	Muons . . . . .	83
4.3.1	Muon Identification Criteria . . . . .	84
4.3.2	Muon Momentum Correction . . . . .	85
4.3.3	Muon Momentum Scale and Resolution . . . . .	87
4.3.4	Muon Isolation Criteria . . . . .	88
4.4	Electrons . . . . .	90
4.5	Jets . . . . .	93
4.5.1	The Jet Reconstruction Algorithm . . . . .	93
4.5.2	The T42 Algorithm . . . . .	95
4.5.3	Jet Identification . . . . .	96
4.5.4	Separation of Jets from Electromagnetic Objects . . . . .	101
4.5.5	Jet Energy Scale . . . . .	102
4.5.6	Jet Energy Resolution . . . . .	103
4.6	Missing $E_T$ . . . . .	109
4.6.1	$\cancel{E}_T$ Resolution . . . . .	109
4.7	$b$ Jets . . . . .	110
4.7.1	Taggability . . . . .	111
4.7.2	Secondary Vertex Tagger Algorithm . . . . .	112
<b>5</b>	<b>Event Trigger</b>	<b>115</b>
5.1	Trigger Efficiency Measurement . . . . .	116
5.1.1	Muon Trigger . . . . .	117
5.1.2	Jet Trigger . . . . .	119
5.1.3	Event Trigger Efficiency . . . . .	121
<b>6</b>	<b>Data Sample</b>	<b>123</b>
6.1	Data Quality Selection . . . . .	123
6.2	Integrated Luminosity of the Data Samples . . . . .	125

<b>7</b>	<b>The Topological <math>t\bar{t}</math> Analysis</b>	<b>127</b>
7.1	$t\bar{t}$ Event Preselection . . . . .	127
7.1.1	Jet Multiplicity Cut . . . . .	128
7.1.2	Muon Acceptance, Reconstruction Efficiency and Cosmic Veto . . . . .	130
7.1.3	Muon Tracking Efficiency . . . . .	135
7.1.4	Efficiency of the Muon $\eta$ Cut . . . . .	137
7.1.5	Efficiency of the Muon Transverse Momentum Cut . . . . .	138
7.1.6	Efficiency of the Muon Separation from Jets . . . . .	138
7.1.7	Efficiency of the Track $\chi_{trk}^2$ Cut . . . . .	138
7.1.8	Efficiency of the Muon $dca$ Significance Cut . . . . .	139
7.1.9	Tight Muon Isolation Efficiency . . . . .	140
7.1.10	Efficiencies for the Missing Transverse Energy Cut . . . . .	140
7.1.11	Efficiency of the Second Lepton Veto (SLV) . . . . .	142
7.1.12	Efficiency of the $Z \rightarrow \mu\mu$ Veto . . . . .	143
7.1.13	PV Reconstruction Efficiency . . . . .	145
7.1.14	Efficiency of the $ \Delta z(\mu, PV) $ Cut . . . . .	147
7.1.15	Jet Trigger Efficiencies . . . . .	147
7.1.16	Muon Trigger Efficiencies . . . . .	148
7.1.17	Total $t\bar{t}$ Preselection Efficiency . . . . .	148
7.2	Contributions to the Data Set . . . . .	150
7.2.1	$t\bar{t}$ Signal . . . . .	150
7.2.2	$t\bar{t}$ Dilepton Background . . . . .	151
7.2.3	Matrix Method - Evaluation of the QCD Background . . . . .	151
7.2.4	$W$ -plus-jets Background . . . . .	154
7.2.5	$Z$ -plus-jets Background . . . . .	155
7.3	Topological Multivariate Analysis . . . . .	156
7.3.1	Modeling of Signal and Backgrounds . . . . .	157
7.3.2	Topological Variables . . . . .	157
7.3.3	The Topological Likelihood Discriminant . . . . .	160
7.3.4	Cross-Section Extraction Procedure . . . . .	162
7.3.5	Topological Optimization . . . . .	165
7.3.6	Result of the Topological Optimization . . . . .	166
7.4	The $t\bar{t}$ Production Cross-Section . . . . .	173
7.4.1	$W$ -plus-jets and $t\bar{t}$ Contamination in QCD-multijets Sample . . . . .	173
7.4.2	Result of the $t\bar{t}$ Production Cross-Section Measurement . . . . .	175
7.5	Systematic Uncertainties . . . . .	181
7.5.1	Uncertainties on the Preselection Efficiency . . . . .	181
7.5.2	Uncertainties on the Fitted Number of $t\bar{t}$ Events . . . . .	181
7.5.3	Summary of the Systematic Uncertainties . . . . .	187
7.5.4	Result of the $t\bar{t}$ Production Cross-Section Measurement in the $\mu$ +Jets Channel . . . . .	188

7.5.5	$t\bar{t}$ Production Cross-Section Measurement in the $e$ +Jets Channel . . . . .	189
7.5.6	$t\bar{t}$ Production Cross-Section Measurement in the Lepton+Jets Channel . . . . .	191
<b>8</b>	<b>The <math>b</math>-Tagging Analysis</b>	<b>195</b>
8.1	$t\bar{t}$ Event Preselection . . . . .	196
8.2	The Jet Tagging Efficiencies . . . . .	197
8.2.1	Taggability . . . . .	198
8.2.2	$b$ -Tagging Efficiency . . . . .	200
8.2.3	$c$ -Tagging Efficiency . . . . .	206
8.2.4	Mis-Tagging Efficiency . . . . .	207
8.3	Contributions to the Data Set . . . . .	209
8.3.1	$t\bar{t}$ Signal . . . . .	211
8.3.2	Matrix Method - Evaluation of the QCD Background	212
8.3.3	Small Backgrounds . . . . .	215
8.3.4	$W$ -plus-jets and $Z$ -plus-jets Background . . . . .	217
8.4	Cross-Section Extraction Procedure . . . . .	225
8.5	Systematic Uncertainties . . . . .	226
8.5.1	Non- $b$ -Tag Related Sources of Systematic Uncertainties	228
8.5.2	$b$ -Tag Related Sources of Systematic Uncertainties . .	228
8.6	The $t\bar{t}$ Production Cross-Section . . . . .	229
<b>9</b>	<b>Summary and Conclusion</b>	<b>231</b>
<b>A</b>	<b>Sanity Checks for the Topological Analysis</b>	<b>237</b>
<b>B</b>	<b>Sanity Checks for the <math>b</math>-Tagging Analysis</b>	<b>269</b>

# Chapter 1

## Introduction

The modern times basically denote the beginnings of scientific experimenting along with its most crucial tools, the logic, the consequent application of rules and conclusions comprehensible by the mind, and the empiricism, the practical experience based on precise observations. Herewith, the ground was laid to systematically address the fundamental questions, such as “Was die Welt im Innersten zusammenhält” [1], to build dedicated experiments on the search for their answers, and to confront the mind with the findings, so that interpretations of the results can be formulated in terms of theories which in turn make predictions that can be tested experimentally.

The culmination of this healthy interaction between experimental observations and theoretical interpretations is the *Standard Model* of particle physics describing the fundamental building blocks of matter and their interaction. With the exception of the tau neutrino, the top quark, being the protagonist of the analyses presented, is the most recently discovered of these building blocks. The history of the top quark is a success for the *Standard Model*, beginning with the prediction of its existence as the electroweak isospin partner of the bottom quark, discovered in 1977, and the prediction of its mass in consideration of the electroweak precision data since 1992, and finally crowned by its recent discovery during FERMILAB Tevatron’s Run I in 1995 by the CDF and DØ collaborations [2, 3, 4].

Not much is known about the top quark so far, its properties are only roughly measured, and the Tevatron is presently the only place where it can be produced. Top physics gives insights into perturbative Quantum Chromodynamics (QCD) [5], there is a possible connection to the mechanism of electroweak symmetry breaking, and it provides the playground for the search for *New Physics* at the same time. These extensive implications make the top quark one of the hottest topics in elementary particle physics of our time.

The subject of this dissertation is the measurement of the rate, expressed in terms of the cross section, at which a top-antitop quark pair is produced

in  $p\bar{p}$  collisions at a center-of-mass energy of  $\sqrt{s} = 1.96$  TeV. The data are provided by the FERMILAB Tevatron collider during Run II and collected with the DØ detector, a multi-purpose particle detector that allows for the measurement of the momentum and the energy as well as the identification of the particles.

This measurement is compared with its theoretical prediction from the *Standard Model* and thus represents a basic test of the latter, in particular the perturbative Quantum Chromodynamics [5].

The top quark is not observable directly. Due to the very short lifetime, it decays to a  $b$  quark and a  $W$  boson. The analyses presented deal with the final state where one  $W$  boson decays hadronically to two quarks and the other to a charged lepton and a neutrino. In a preselection of the data, all particles in the final state are reconstructed in the DØ detector: the charged lepton, the neutrino as missing transverse energy and the quarks as jets. Subsequently, two ways are pursued to extract the  $t\bar{t}$  cross section. Due to the large top mass and the fact that the top-antitop quark pair is produced mainly at rest at the Tevatron center-of-mass energy of  $\sqrt{s} = 1.96$  TeV, the final state objects are high-energetic and isotropic, whereas the final state objects of the backgrounds, dominated by the electroweak  $W$  boson production in association with four or more jets (referred to as  $W$ -plus-jets) originating from QCD bremsstrahlung, are low energetic and in the forward direction. The kinematical and topological information is combined in a likelihood discriminant whose shape found in data is described by a linear combination of the shapes of the  $t\bar{t}$  signal and the backgrounds. The relative number of events are fitted in order to extract the  $t\bar{t}$  cross section; this method is referred to as the *topo* analysis. On the other hand, the so-called *btag* analysis does not make use of any topological information but it requires, in addition to the above mentioned event selection, to identify one or two of the two  $b$  quarks in a  $t\bar{t}$  event using lifetime  $b$ -tagging by reconstructing explicitly secondary vertices. This selection is approximately 60 % efficient for  $t\bar{t}$  events, while approximately 95 % of the background is removed, since it contains no heavy flavor quarks in the final state, which makes this analysis so powerful.

In an earlier version of the *topo* analysis [6, 7], the  $W$ -plus-jets background was estimated by using the Berends scaling [8]. However, this method is characterized by large uncertainties. The introduction of the multivariate technique mentioned above improved the sensitivity of the analysis considerably [9]. In the *topo* analysis presented here, the separation power of the likelihood discriminant and therefore the statistical uncertainty, and the insensitivity to the dominant systematic uncertainties are further improved by a detailed optimization study, and the luminosity is extended to  $229 \text{ pb}^{-1}$ . The results are in the process of being published in a refereed journal (Phys. Rev. Lett.).

Between Run I and Run II the DØ detector was significantly upgraded.

In particular, it was equipped with a central tracking system and a solenoid magnetic field, which allow to do tracking and thus lifetime  $b$ -tagging for the first time with the  $D\bar{O}$  detector. The first  $D\bar{O}$  measurement ever shown outside of  $D\bar{O}$  that made use of lifetime  $b$ -tagging was an earlier version of the  $btag$  analysis [10, 11]. An improved understanding of the systematic uncertainties and a larger integrated luminosity of approximately  $160 \text{ pb}^{-1}$  lead to the  $btag$  analysis presented here [12, 13]. The results are in the process of being published in a refereed journal (Phys. Rev. Lett.) with an even increased luminosity of approximately  $230 \text{ pb}^{-1}$ .

This dissertation is organized as follows:

- The *Theoretical Aspects*, relevant to the analyses presented, are detailed in Chapt. 2, including the physics of the top and the bottom quarks, the event signatures, the MC simulation, and the measurement of the luminosity.
- The *Experimental Environment*, comprising the Tevatron accelerator facilities and the  $D\bar{O}$  detector, is described in Chapt. 3.
- The *Event Reconstruction and Object Identification* of the final state objects is given in Chapt. 4.
- The *Event Trigger* is discussed in Chapt. 5.
- The *Data Sample* used for the analyses is presented in Chapt. 6.
- The *Topological  $t\bar{t}$  Analysis* is discussed in Chapt. 7 and consists of the  $t\bar{t}$  event preselection, a list of the contributions to the data set, the illustration of the multivariate technique, the extraction of the production cross section and the discussion of the systematic uncertainties.
- The  *$b$ -Tagging Analysis* follows in Chapt. 8 starting with the  $t\bar{t}$  event preselection, and the discussion of the tagging efficiencies, then the contributions to the data set are listed, the cross section extraction procedure is explained, the systematic uncertainties are discussed and the  $t\bar{t}$  production cross section is presented.
- *Summary and Conclusion* are given in Chapt. 9.



## Chapter 2

# Theoretical Aspects

The “*Standard Model*” of elementary particle physics is the model which describes our present understanding of quarks, leptons and their interactions. The model is enormously successful in predicting a wide range of fundamental phenomena. The *Standard Model* is a renormalizable gauge theory with local gauge invariance [14], it is characterized in part by the spectrum of elementary particles which consist of 6+6 fundamental particles (fermions) and 4 fundamental forces, mediated by gauge bosons, as shown in Table 2.1. An introduction and discussion of the basic concepts of the *Standard Model* can be found in [17] and [18], particularly in [19] concerning the electroweak theory and in [5] concerning Quantum Chromodynamics (QCD).

In Sect. 2.1 the physics of the top quark in the framework of the *Standard Model* is discussed. Section 2.2 deals with the bottom quark which is a decay product of the top quark. The signature of signal and background processes is discussed in Sect. 2.3. The Monte Carlo simulation of the events is described in Sect. 2.4, and a discussion about the measurement of the luminosity at hadron colliders is given in Sect. 2.5.

	Generation		
	I	II	III
leptons:	$\nu_e$ (1953) $e$ (1897)	$\nu_\mu$ (1962) $\mu$ (1936)	$\nu_\tau$ (2000) $\tau$ (1975)
quarks:	u (1968) d (1968)	c (1974) s (1964)	t (1995) b (1977)
gauge bosons:	$g_1, \dots, g_8$ (1979) $\gamma$ (1900) $W^\pm, Z^0$ (1983)		

Table 2.1: The three generations of the constituents of matter and the gauge bosons. The dates of discovery are given in parentheses [2, 3, 4, 15, 16, 17].

## 2.1 The Physics of the Top Quark

The top quark is the heaviest known elementary particle with a mass equal to the mass of a gold atom and approximately 40 times larger than the next heaviest quark, the bottom quark. It differs from the other quarks not only by the much larger mass, but also by its lifetime which is too short to build hadronic bound states.

The top quark was predicted since the discovery of the bottom quark 1977 as its electroweak isospin partner, the top quark mass was predicted since 1992 with continuously increasing precision in consideration of the electroweak precision data from LEP, SLD, NuTeV and the  $p\bar{p}$  colliders [20], and the top quark was discovered in Run I of the FERMILAB Tevatron in 1995 [2, 3, 4], completing the quark sector of the three-generation structure of the *Standard Model*.

The top quark is one of the least well-studied components of the *Standard Model* and the Tevatron, with a center-of-mass energy of  $\sqrt{s} = 1.96$  TeV, is at present the only collider where the top quark can be produced and studied. This is of strong interest due to the following considerations:

- The measurement of the production rate and the properties of the top quark represent one of the most important tests of the *Standard Model*, in particular of the QCD at the Tevatron.
- The top quark mass is an input for many electroweak analyses, in particular an accurate measurement can be used for a more precise prediction of the Higgs mass.
- The top quark is expected to play an important role in the discovery of new particles. This is indicated by the Higgs boson coupling to the top quark, which is stronger than for all other fermions.
- Background to *New Physics* consists mainly of events containing top quarks. The understanding of this main background, its signature and production rate, is crucial for the discovery potential of *New Physics*.
- The electroweak single top production allows for the first time to measure the CKM matrix element  $V_{tb}$  directly.

### 2.1.1 Properties of the Top Quark

In the *Standard Model*, the top quark is defined as the weak isospin partner of the bottom quark. As such, it is a spin- $\frac{1}{2}$  fermion of electric charge  $+\frac{2}{3}$  and transforms as a color anti-triplet under the SU(3) gauge group of strong interactions. Its quantum numbers have not yet been measured directly, although a large amount of direct and indirect evidence supports the *Standard Model* assignments [21, 22, 23, 24, 25].

### 2.1.1.1 Top Quark Mass

The mass of the top quark is heavier than the mass of any other elementary particle and is measured with better relative precision than any other quark, see Fig. 2.1. The Tevatron experiments have measured the top quark mass in

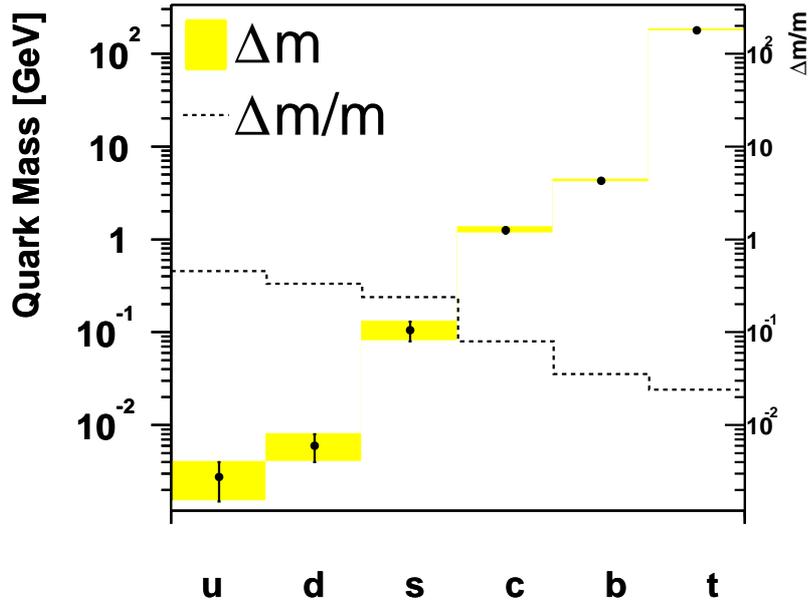


Figure 2.1: Quark masses and their uncertainties.

Run I to be  $m_t = 178.0 \pm 2.7(stat.) \pm 3.3(syst.)$  GeV [26]. Preliminary results for Run II [27] are in good agreement with this value. The perspectives for the full data set of Run II are to measure the top quark mass to a precision better than 3 GeV. Together with a precision of the  $W$  mass of 20 MeV the Higgs mass is expected to be constrained to better than 40 % [28]. At the Large Hadron Collider (LHC) the mass of the top quark will be measured with a precision of about 1 GeV [29].

### 2.1.1.2 Yukawa Coupling

Yukawa coupling is the Higgs coupling to fermions and thus relates the fermionic matter content of the *Standard Model* to the source of mass generation, the Higgs sector [30]. In the *Standard Model*, the Yukawa coupling to the top quark is very close to unity, a theoretically interesting value, leading to speculations that *New Physics* may be accessed via top quark physics [31]. The coupling will be measured in the associated  $t\bar{t}$  Higgs production at the LHC.

### 2.1.1.3 Spin Correlation

One of the remarkable features of the top quark is that it is the only quark whose spin is directly observable. This is a consequence of its very short lifetime of about  $5 \times 10^{-25}$  s, which is  $\mathcal{O}(10)$  times shorter than the characteristic time scale for hadronization. As a result the top quark is the only quark that decays before it has a chance to hadronize. The spin information is transferred to the top quark decay products,  $t \rightarrow Wb$ , and can be measured in the leptonic  $W$  boson decay. Bounds on  $t\bar{t}$  spin correlations were found in Run I [24].

### 2.1.1.4 $W$ Helicity

The helicity of  $W$  bosons from top quark decays is predicted by masses and couplings in the *Standard Model*. Studies of the decay angular distributions allow a direct analysis of the  $V - A$  nature of the  $Wtb$  coupling. In the *Standard Model*, the fraction of decays to longitudinally polarized  $W$  bosons is expected to be  $f_0 = \frac{m_t^2}{2m_W^2 + m_t^2 + m_b^2} \approx 70\%$ . Deviations from this value would bring into question the validity of the Higgs mechanism of spontaneous symmetry breaking. Measurements in Run I [32, 33] and Run II [34], using the angular distribution of the objects in the final state, are in good agreement with the expectation.

### 2.1.1.5 Top Quark Charge

The top quark charge is easily accessible in  $e^+e^-$  production by scanning the center-of-mass energy of two times the top quark mass and measuring the ratio  $R = \frac{e^+e^- \rightarrow \text{hadrons}}{e^+e^- \rightarrow \mu^+\mu^-}$ , however, this region of center-of-mass energy has not been accessible so far. At hadron colliders, as the Tevatron, the top quark charge is directly measured by measuring the charge of its decay products, where the charge of the  $W$  boson is easily accessible through its leptonic decay products, but the charge measurement of the associated  $b$  jet is challenging. Indirectly, the top quark charge can be inferred by assuming the *Standard Model* electromagnetic coupling and by measuring the rate at which the top quark radiates a photon.

### 2.1.1.6 Resonances

The invariant mass of the  $t\bar{t}$  system may reveal s-channel resonances, predicted by various models. Scans of the  $t\bar{t}$  mass spectrum were performed in Run I and Run II and no significant peaks were observed [21, 22, 23].

### 2.1.2 The Virtual Top Quark

The precise electroweak measurements performed at LEP, SLD, NuTeV and the  $p\bar{p}$  colliders can be used to check the validity of the *Standard Model* and within its framework, to infer valuable information about its fundamental parameters [35]. The accuracy of the measurements makes them sensitive to the mass of the top quark and the Higgs boson through radiative corrections [20].

All electroweak quantities depend only on five parameters. At leading order this dependence is reduced to only 3 parameters, two gauge couplings and the Higgs-field vacuum-expectation value. The three best-measured electroweak quantities can be used to determine these three parameters: The electromagnetic coupling constant  $\alpha$ , measured in low-energy experiments, the Fermi constant  $G_F$ , determined from the  $\mu$  lifetime [36], and the mass of the  $Z$  boson, measured from  $e^+e^-$  annihilation [37]. By defining the Weinberg angle  $\theta_W$  through:

$$\sin^2 \theta_W \equiv 1 - \frac{m_W^2}{m_Z^2} \quad (2.1)$$

the  $W$  mass at tree level can be expressed by:

$$m_W^2 = \frac{\frac{\pi\alpha}{\sqrt{2}G_F}}{\sin^2 \theta_W}. \quad (2.2)$$

Considering also one-loop corrections, the expression in Eq. 2.2 is modified to:

$$m_W^2 = \frac{\frac{\pi\alpha}{\sqrt{2}G_F}}{\sin^2 \theta_W (1 - \Delta r)} \quad (2.3)$$

where  $\Delta r$  contains all the one-loop corrections. Contributions to  $\Delta r$  originate from the top quark by the one-loop diagrams shown in Fig. 2.2, which contribute to the  $W$  and  $Z$  masses:

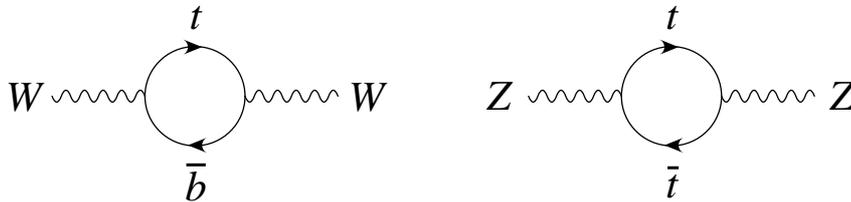


Figure 2.2: Virtual top quark loops contributing to the  $W$  and  $Z$  boson masses.

$$(\Delta r)_{top} \simeq -\frac{3G_F}{8\sqrt{2}\pi^2 \tan^2 \theta_W} \frac{1}{m_t^2}. \quad (2.4)$$

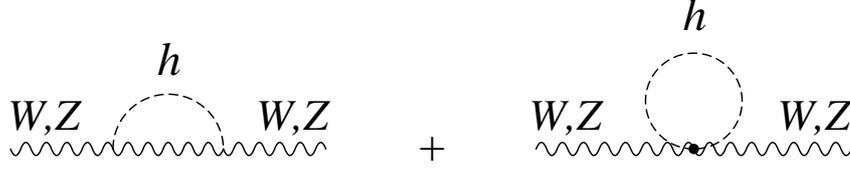


Figure 2.3: Virtual Higgs boson loops contributing to the W and Z boson masses.

Also the Higgs boson contributes to  $\Delta r$  via the one-loop diagrams shown in Fig. 2.3:

$$(\Delta r)_{Higgs} \simeq \frac{11G_F m_Z^2 \cos^2 \theta_W}{24\sqrt{2}\pi^2} \ln \frac{m_H^2}{m_Z^2}. \quad (2.5)$$

While the leading  $m_t$  dependence is quadratic, the leading  $m_H$  dependence is logarithmic. Therefore the inferred constraints on  $m_H$  are much weaker than those on  $m_t$ . This was used to successfully predict the top quark mass several years before it was discovered, which is shown in Fig. 2.4. Neutral

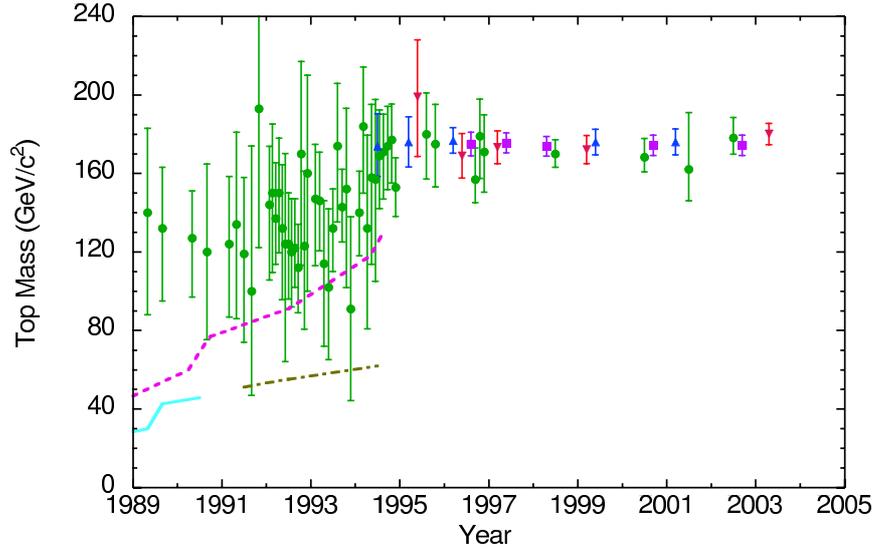


Figure 2.4: Evolution of the top quark mass prediction and measurement with time [38]. Until 1995 only predictions from electroweak fits ( $\bullet$ ) have been done. These are consistent with the direct observation by CDF ( $\blacktriangle$ ) and DØ ( $\blacktriangledown$ ). The world average from both experiments is also shown ( $\blacksquare$ ) as well as the lower bounds from hadron colliders (dashed lines) and  $e^+e^-$  colliders (solid line).

current weak interaction data, such as  $e^+e^-$  annihilation near the Z mass,  $\nu N$  and  $eN$  deep-inelastic scattering,  $\nu e$  elastic scattering and atomic parity

violation [39] can be used to predict the top quark mass [20], see Fig. 2.5. The most recent indirect measurement of the top quark mass using the  $Z$ -

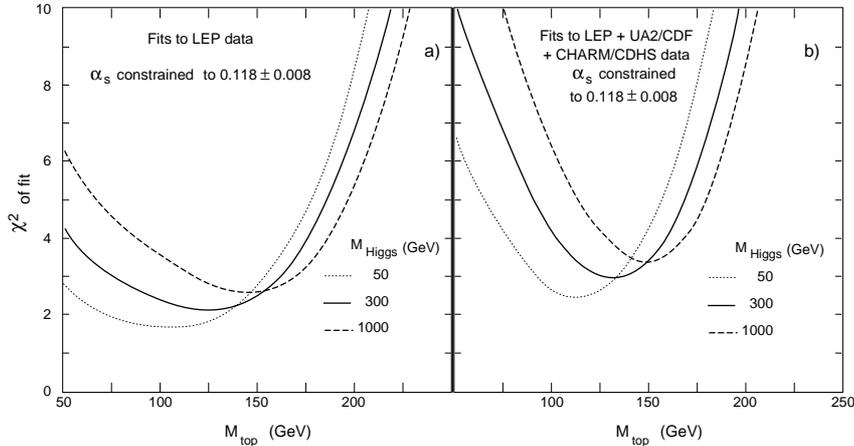


Figure 2.5:  $\chi^2$  of the fit to the electroweak data as a function of the top quark mass using LEP data (left) and LEP, hadron collider and neutrino experiment data (right) [20]. The dependence on the Higgs mass is weak, since  $m_H$  enters only logarithmically in the electroweak fit, whereas  $m_t$  enters quadratically.

pole data together with the direct measurements of the  $W$  mass and total width yields [37]:

$$m_t = 179_{-9}^{+11} \text{ GeV}, \quad (2.6)$$

which is in very good agreement with the direct measurement [26]:

$$m_t = 178.0 \pm 4.3 \text{ GeV}. \quad (2.7)$$

The global fit to both the direct and the indirect measurements yields [37]:

$$m_t = 178.2 \pm 3.9 \text{ GeV}. \quad (2.8)$$

The successful prediction of the mass of the top quark before its discovery gives a good reason to have confidence in the prediction for the Higgs boson mass from the electroweak global fit to all data, including the direct measurement of  $m_t$ . Figure 2.6 (left) shows the result of this fit for the Higgs mass which is the corresponding plot to Fig. 2.5 for the top quark mass. The most likely value of the Higgs mass, corresponding to the minimum of the curve in Fig. 2.6 (left), is at 114 GeV, with an experimental uncertainty of +69 and -45 GeV [37].

Another representation of the electroweak global fit is given by Fig. 2.6 (right) [37]. It shows the direct and indirect measurement of  $m_t$  and  $m_W$ .

Also shown are the *Standard Model* predictions of Higgs masses between 114 and 1000 GeV. As can be seen in the figure, the direct and indirect measurements are in good agreement and prefer a low value of the *Standard Model* Higgs mass.

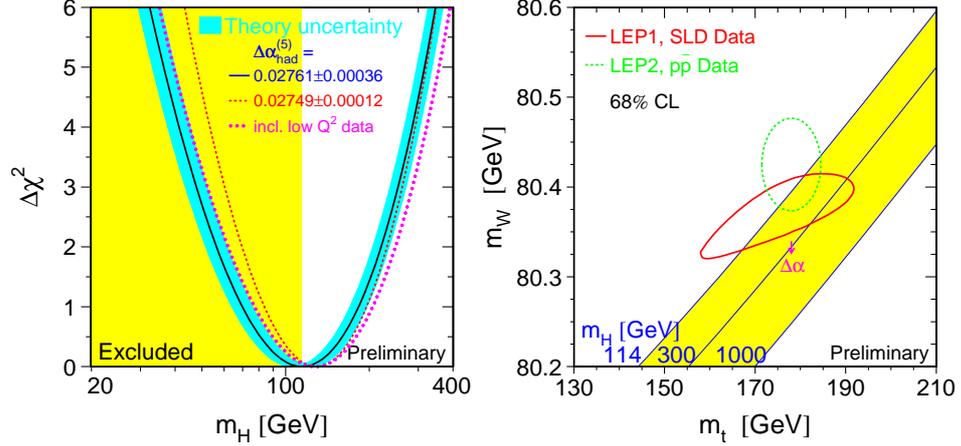


Figure 2.6: Left: Blueband plot, showing the indirect measurement of the Higgs boson mass from all electroweak data together with the 95% confidence lower limit on the Higgs mass of 114.4 GeV [35]. Right: Contour curves of 68% confidence level in the  $(m_t, m_W)$  plane, for the corresponding indirect (LEP1, SLD data) and direct (LEP2,  $p\bar{p}$  data) determination in a global fit to electroweak precision data [35]. Also shown is the correlation between  $m_t$  and  $m_W$  as expected in the *Standard Model* for different Higgs boson masses.

### 2.1.3 Top Quark Pair Production via the Strong Interaction

The  $t\bar{t}$  production in high energy interactions of a  $p\bar{p}$  collision at the Tevatron is described by QCD. In this theory a hard scattering process between a proton and an antiproton is the result of an interaction between the quarks and the gluons which are the constituents of the incoming hadrons. The incoming hadrons provide broad band beams of partons which possess varying fractions  $x$  of the momenta of their parent hadrons. An illustration of such a  $p\bar{p}$  collision is shown in Fig. 2.7.

#### 2.1.3.1 Factorization of the Cross Section

The  $p\bar{p}$  interaction can be separated into a short distance (hard scattering) cross section of partons of type  $i$  and  $j$ ,  $\hat{\sigma}^{ij}$ , and into long distance

pieces which are factored into the parton momentum distribution functions  $f_i(x_i, \mu_F^2)$  (see Fig. 2.7). This separation is called factorization and is set by the factorization scale  $\mu_F^2$ .  $\mu_F^2$  decides at what scale the separation is made. The remaining short distance cross section involves only high momentum transfer and is calculable in perturbation theory. It is insensitive to the physics of low momentum scale, in particular it is not dependent on the hadron wave functions or the type of the incoming hadrons. This factorization property of the cross section can be proven to all orders in perturbation theory [40], the more terms are included in the perturbative expansion, the weaker the dependence on this arbitrary scale  $\mu_F^2$ .

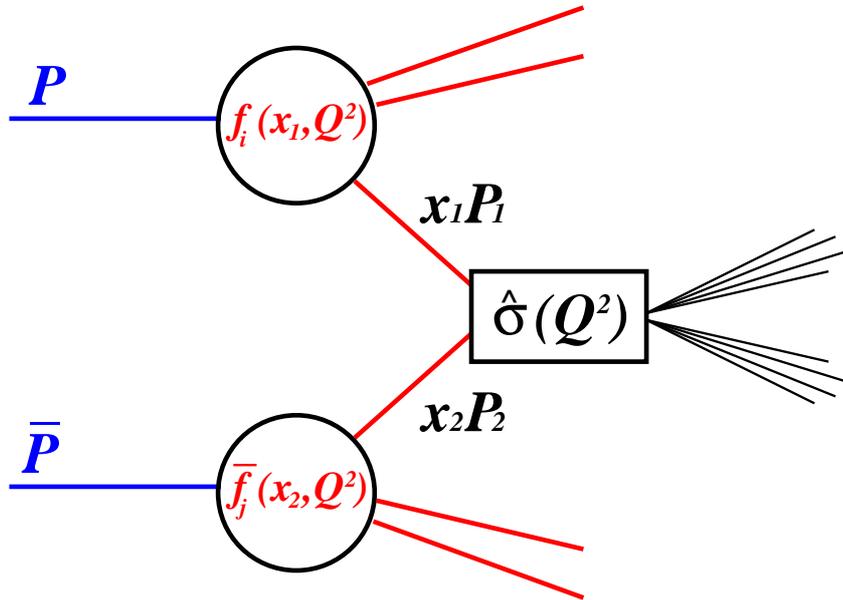


Figure 2.7: The parton model description of a hard scattering process.

The parton distribution function (PDF)  $f_i(x_i, \mu_F^2)$  is the probability density that the parton of flavor  $i$  is participating in the hard scattering interaction with longitudinal momentum fraction  $x_i$  of the incoming hadron, when probed at a scale  $\mu_F^2$ . Since the parton distribution functions can presently not be calculated in perturbative QCD, they are extracted in global QCD fits at next-to-leading order from data [41, 42]. An example parameterization, obtained by the CTEQ collaboration [43], is shown in Figure 2.8.

### 2.1.3.2 Renormalization and the Running Coupling Constant

In order to calculate a physical observable such as the  $t\bar{t}$  cross section in perturbation theory one must include all the virtual loop diagrams. If one tries to calculate this series of diagrams one obtains infinities (ultra-violet di-

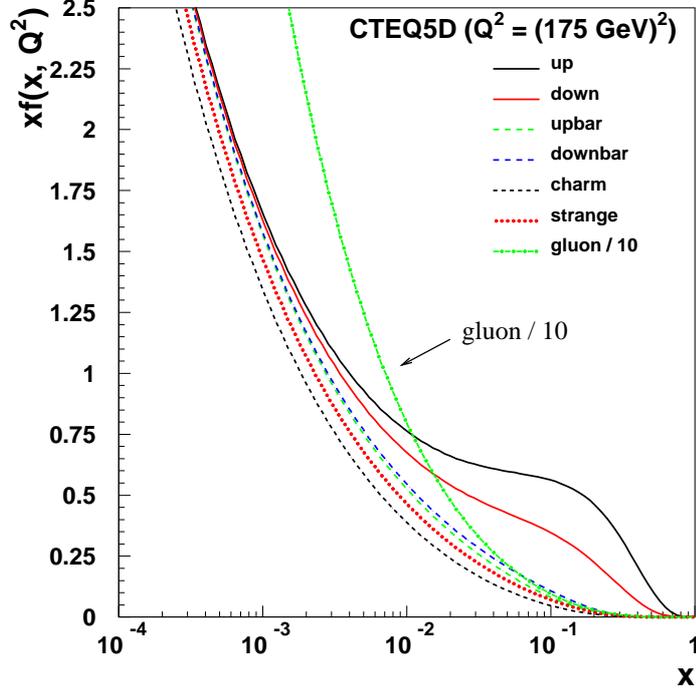


Figure 2.8: The quark, anti-quark and gluon momentum densities in the proton as a function of the longitudinal momentum fractions  $x$  at  $Q^2 = m_t^2$  from CTEQ5D [43].

vergences). These divergences are removed by a renormalization procedure, which introduces the artificial scale  $\mu_R^2$ . However, it is clear that the physical quantities cannot depend on the arbitrary scale  $\mu_R^2$  and this independence is expressed in terms of a Renormalization Group Equation [44, 5]. It can be shown [45, 5] that the solution of the Renormalization Group Equation implies the running coupling constant  $\alpha_s(Q^2)$ . Neglecting higher orders of  $\alpha_s(Q^2)$ , one finds:

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu_R^2)}{1 + \alpha_s(\mu_R^2)b \ln \frac{Q^2}{\mu_R^2}}, \quad (2.9)$$

with  $b = (33 - 2n_f)/12\pi$  and  $n_f$  is the number of active flavors. If  $Q^2/\mu_R^2$  becomes very large, the running coupling  $\alpha_s(Q^2)$  decreases to zero, which is known as asymptotic freedom. A summary of  $\alpha_s$  measurements at various scales is shown in Fig. 2.9.

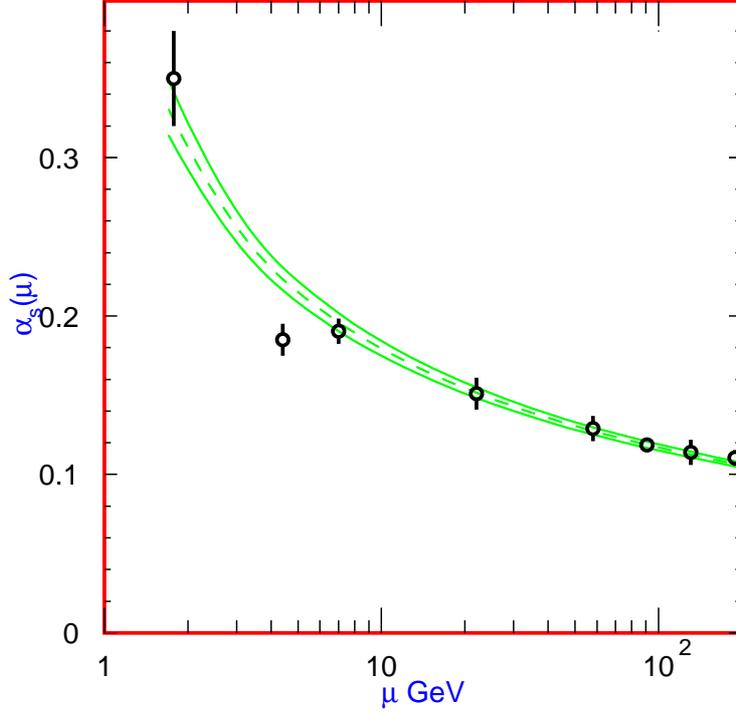


Figure 2.9: Summary of the values of  $\alpha_s$  at various scales  $\mu$  and a fit and the  $\pm 1\sigma$  limits [39]. The data are in increasing order of  $\mu$ ,  $\tau$  width,  $\Upsilon$  decays, deep inelastic scattering,  $e^+e^-$  event shapes at 22 GeV from the JADE data, shapes at TRISTAN at 58 GeV,  $Z$  width and  $e^+e^-$  event shapes at 135 and 189 GeV.

### 2.1.3.3 The $t\bar{t}$ Production Cross Section

It is common to choose the same scale  $\mu$  both for the factorization ( $\mu_F^2$ ) and the renormalization ( $\mu_R^2$ ) scale. This convention will be used in the following:  $\mu_F^2 = \mu_R^2 \equiv \mu^2$ .

The total top quark pair production cross section for a hard scattering process initiated by a  $p\bar{p}$  collision at the center-of-mass energy  $\sqrt{s}$  can be expressed as

$$\sigma^{p\bar{p} \rightarrow t\bar{t}+X}(s, m_t) = \sum_{i,j=q,\bar{q},g} \int dx_i dx_j f_i(x_i, \mu^2) \bar{f}_j(x_j, \mu^2) \hat{\sigma}^{ij \rightarrow t\bar{t}}(\rho, m_t^2, \alpha_s(\mu^2), \mu^2). \quad (2.10)$$

$f_i(x_i, \mu^2)$  and  $\bar{f}_j(x_j, \mu^2)$  are the PDF's for the proton and the antiproton.

The summation indices  $i$  and  $j$  run over all  $q\bar{q}$  and gluon pairs.  $\rho = \frac{4m_t^2}{\hat{s}}$  and  $\hat{s} = x_i \cdot x_j \cdot s$  is the effective center-of-mass energy squared for the partonic process, shown in Fig. 2.10.

It can be seen that at the Tevatron with a  $p\bar{p}$  center-of-mass energy of 1.96 TeV, the  $t\bar{t}$  quark pair is mainly produced just above the kinematically allowed threshold,  $\sqrt{\hat{s}} \gtrsim 2m_t$ . Figure 2.11 shows the transverse momentum and the  $\eta$  (given by Eq. 3.6) distribution of the  $t\bar{t}$  system, where values of  $p_{T_{t\bar{t}}} > 0$  occur due to initial state gluon radiation. The  $\eta_{t\bar{t}}$  distribution is strongly peaked in the forward direction, which demonstrates that the longitudinal momentum fractions  $x$  of the two partons which take part in the hard scattering process are very different most of the time. The transverse

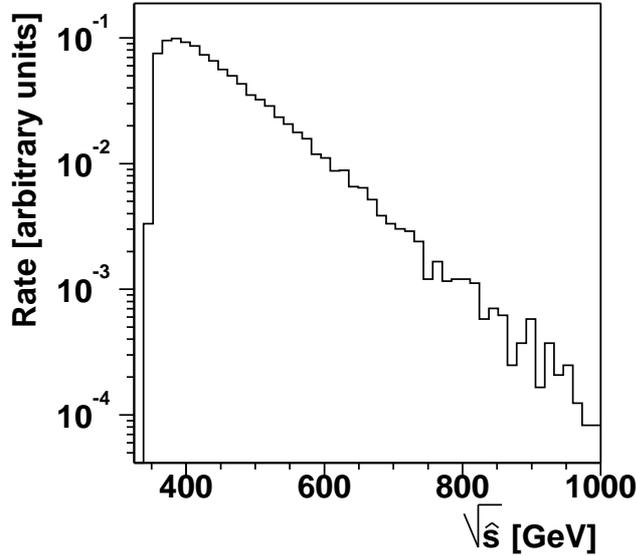


Figure 2.10:  $\sqrt{\hat{s}} = \sqrt{x_i \cdot x_j \cdot s}$ , the partonic center-of-mass energy for the production of a  $t\bar{t}$  quark pair at  $\sqrt{s} = 1.96$  TeV.

momenta of the top quarks in a  $t\bar{t}$  process are of the order of the top quark mass. This implies that the top quarks are mainly emitted in the central  $\eta$  region at the Tevatron, see Fig. 2.12. At the LHC, due to the higher center-of-mass energy, top quarks will be produced at much larger values of  $|\eta|$ .

The total short distance cross section can be written as:

$$\hat{\sigma}^{ij}(\rho, m_t^2, \alpha_s(\mu^2), \mu^2) \equiv \frac{\alpha_s^2(\mu^2)}{m_t^2} f_{ij}(\rho, \alpha_s(\mu^2), \mu^2/m_t^2). \quad (2.11)$$

The dimensionless functions  $f_{ij}$  are perturbatively computable and have the

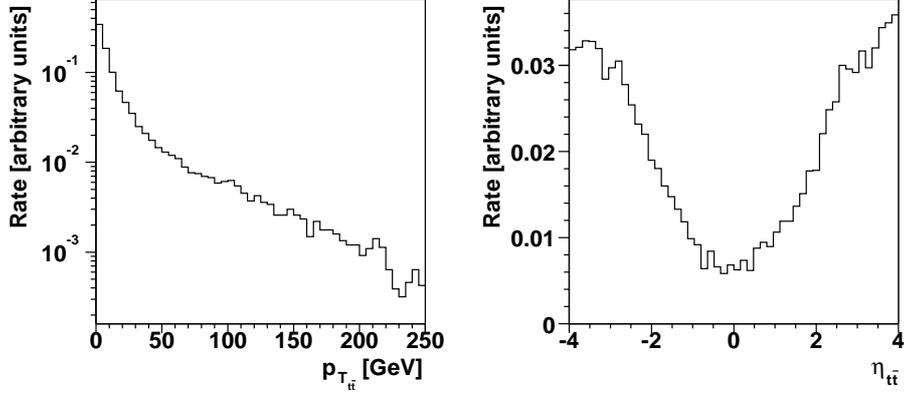


Figure 2.11: Transverse Momentum and  $\eta$  distributions of the  $t\bar{t}$  system at  $\sqrt{s} = 1.96$  TeV.

following expansion in  $\alpha_s$ :

$$\begin{aligned}
 f_{ij}(\rho, \alpha_s(\mu^2), \mu^2/m_t^2) &= f_{ij}^{(0)}(\rho) + 4\pi\alpha_s(\mu^2) \left[ f_{ij}^{(1)}(\rho) + \bar{f}_{ij}^{(1)}(\rho) \ln \frac{\mu^2}{m_t^2} \right] \\
 &+ \sum_{n=2}^{\infty} \alpha_s^n(\mu^2) f_{ij}^{(n)}(\rho, \mu^2/m_t^2). \quad (2.12)
 \end{aligned}$$

The Feynman diagrams of the leading order (LO) subprocesses are shown in Figure 2.13. The corresponding LO terms  $f_{ij}^{(0)}$  are explicitly given by:

$$\begin{aligned}
 f_{q\bar{q}}^{(0)}(\rho) &= \frac{1}{27}\pi\beta\rho [2 + \rho] \\
 &\stackrel{\rho \rightarrow 1}{\simeq} \frac{1}{9}\pi\beta \rightarrow 0, \quad (2.13)
 \end{aligned}$$

$$\begin{aligned}
 f_{gg}(\rho) &= \frac{1}{12}\pi\rho \left[ \left( 1 + \rho + \frac{\rho^2}{16} \right) \ln \frac{1+\beta}{1-\beta} - \beta \left( \frac{7}{4} + \frac{31}{16}\rho \right) \right] \\
 &\stackrel{\rho \rightarrow 1}{\simeq} \frac{7}{192}\pi\beta \rightarrow 0, \quad (2.14)
 \end{aligned}$$

$$f_{gq}^{(0)}(\rho) = f_{q\bar{q}}^{(0)}(\rho) = 0, \quad (2.15)$$

where  $\beta = \sqrt{1-\rho}$  is the velocity of the top quarks in the  $t\bar{t}$  center-of-mass frame. The limit is given for the  $t\bar{t}$  production near the threshold region,  $\hat{s} \rightarrow (2m_t)^2$ , where  $\rho = \frac{4m_t^2}{\hat{s}} \rightarrow 1$  and

$$\frac{f_{q\bar{q}}^{(0)}(\rho)}{f_{gg}(\rho)} \simeq 3. \quad (2.16)$$

Near threshold and for the case when the parton momenta are equal, one finds  $x_i = x_j = 2m_t/\sqrt{s} \simeq 0.18$  for a top quark mass  $m_t = 175$  GeV and

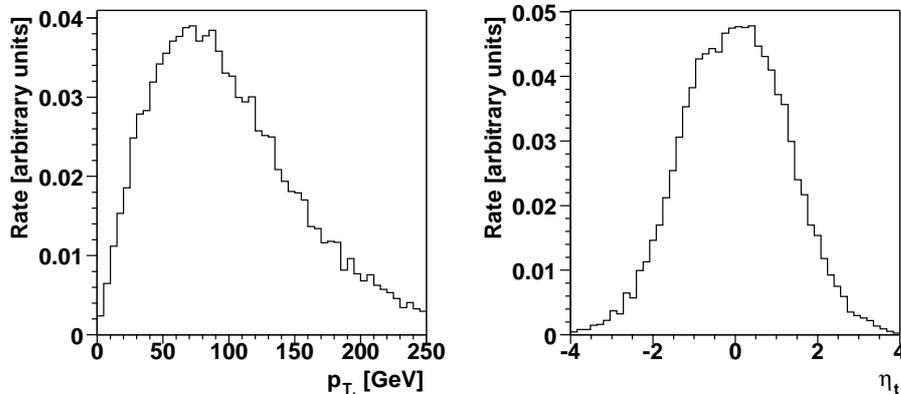


Figure 2.12: Transverse momentum and  $\eta$  distributions of the top quarks in a  $t\bar{t}$  process at  $\sqrt{s} = 1.96$  TeV.

$\sqrt{s} = 1.96$  TeV. For these values of  $x$  the  $q$  and  $\bar{q}$  momentum densities are much larger than the gluon momentum density (see Fig. 2.8) which leads to further enhancement of the  $q\bar{q}$  partonic cross section over the  $gg$  one. As a result, at Tevatron energies the  $q\bar{q} \rightarrow t\bar{t}$  process dominates, contributing 85% of the cross section. The  $gg \rightarrow t\bar{t}$  process contributes 15%.

The remaining next-to-leading order (NLO) contributions  $f_{ij}^{(1)}(\rho)$  and  $\bar{f}_{ij}^{(1)}(\rho)$  can only be evaluated numerically [46, 47]. While the LO functions  $f_{ij}^{(0)}(\rho)$  vanish for  $\rho \rightarrow 1$  because of phase-space suppression, the NLO functions do not and it can be shown [46, 47] that soft gluon emissions give the bulk of the NLO correction.

For the most recent calculations of the top quark production cross section the parton-level cross sections are carried out with the inclusion of the full NLO matrix elements [46]. The complete NLO calculation of the  $t\bar{t}$  cross section is improved with the resummation of leading (LL) [48] and next-to-leading soft logarithms (NLL) [49] appearing at all orders of perturbation theory, which is equivalent to the incorporation of the dominant contributions originating from the emission of soft gluons. The numeric next-to-leading order (NLO) results with resummation correction (NLL) are summarized in Table 2.2. All cross sections are quoted for a top quark mass of 175 GeV, however, the world average is now  $m_t = 178.0 \pm 4.3$  GeV [26]. This change in the top quark mass corresponds to a drop of the  $t\bar{t}$  cross section by approximately 10 %. The considered sources of theoretical uncertainties on the  $t\bar{t}$  cross section are:

**Scale Uncertainty (purely theoretical):** Dependence on the choice of renormalization ( $\mu_R$ ) and factorization ( $\mu_F$ ) scales, considering the conventional range  $m_t/2 < \mu < 2m_t$  setting  $\mu_R = \mu_F \equiv \mu$ . The

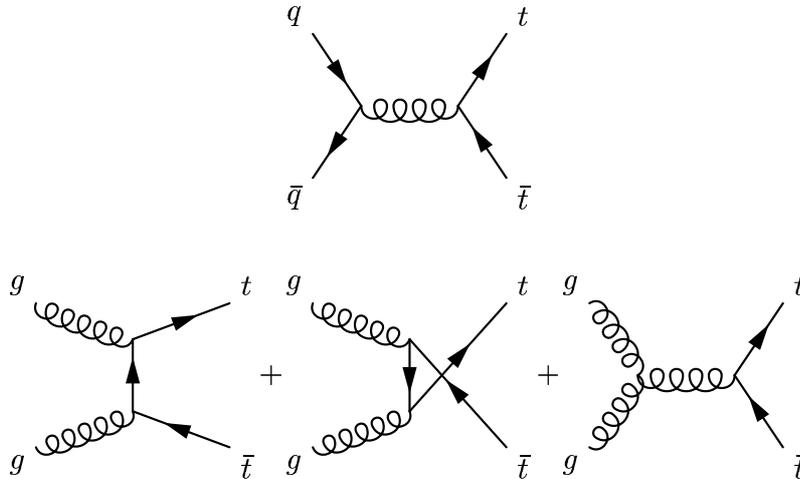


Figure 2.13: Lowest order Feynman diagrams for the production of  $t\bar{t}$  pairs at the Tevatron. At Tevatron energies, the diagram involving quark-antiquark annihilation dominates over those involving gluon-gluon fusion.

possibility of varying the values of renormalization and factorization scale ( $0.5 < \mu_R/\mu_F < 2$ , with  $0.5 < \mu_{R,F}/m_t < 2$ ) is also considered and found to have a small impact.

**PDF Uncertainty:** The sets of PDF parameterizations considered are CTEQ6 [41] and MRST 2002 [50]. The CTEQ and MRST collaborations provide sets of  $\sim 40$  PDF's which allow to assign systematic uncertainties.

$\alpha_s$  **Uncertainty:** The  $\pm 1\sigma$  uncertainty of  $\alpha_s(M_Z)$  is considered, where  $M_Z$  is the mass of the  $Z$  boson.

	[49], as updated in [51]	[52]
Tevatron Run I ( $\sqrt{s} = 1.8$ TeV, $p\bar{p}$ )	$5.19^{+0.52}_{-0.68}$ pb	$5.24 \pm 0.31$ pb
Tevatron Run II ( $\sqrt{s} = 1.96$ TeV, $p\bar{p}$ )	$6.70^{+0.71}_{-0.88}$ pb	$6.77 \pm 0.42$ pb
LHC ( $\sqrt{s} = 14.0$ TeV, $pp$ )	$833^{+52}_{-39}$ pb	$872.8^{+2.3}_{-27.6}$ pb

Table 2.2:  $t\bar{t}$  production cross section for  $m_t = 175$  GeV for Tevatron Run I, Run II and LHC, column 1 from [49], as updated in [51] (the value from the last row is from [49]), column 2 from [52]. The uncertainties in [49] and [51] are dominated by the PDF and  $\alpha_s$  uncertainties. Whereas in [52] the dominant uncertainties arise from the kinematics ambiguity (one-particle inclusive versus pair-invariant mass), see [53] for a more detailed discussion.

The dependence of the  $t\bar{t}$  production cross section on the scale  $\mu$  is shown in Fig. 2.14 [52], it can be seen that the dependence decreases with increasing number of orders considered. The top quark mass dependence of the  $t\bar{t}$  cross

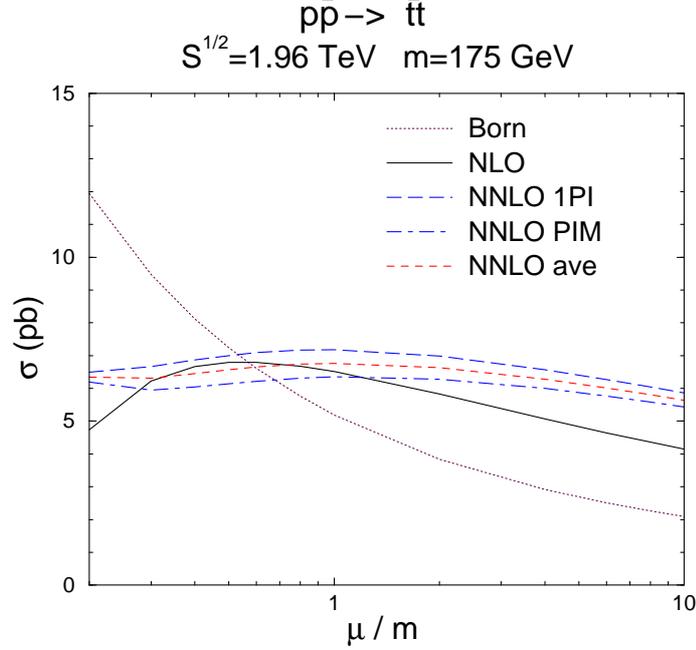


Figure 2.14: The scale dependence of the  $t\bar{t}$  cross section. The exact definition of the terms which are considered in the perturbative expansion referred to as “NNLO” can be found in [52].

section is shown in Fig. 2.15 [52]. For the constant center-of-mass energy of 1.96 TeV, the cross section drops with increasing top quark mass for the following reasons:

- The reduced phase space available for the partonic short distance cross section.
- The lower probability to find a parton with larger  $x \simeq 2m_t/\sqrt{s}$  from the incoming hadron.
- The running of  $\alpha_s(Q^2)$ , where  $Q^2 \simeq m_t^2$ .

The center-of-mass energy dependence of the  $t\bar{t}$  cross section is shown in Fig. 2.16. Here the cross section rises with increasing center-of-mass energy due to the increased phase space available and the higher probability to find a parton with larger  $x \simeq 2m_t/\sqrt{s}$  from the incoming hadron. In particular the contributing fractions from the gluon fusion and the  $q\bar{q}$  partonic subprocesses change as a result of the different PDF's shown in Fig. 2.8. Table 2.3 shows

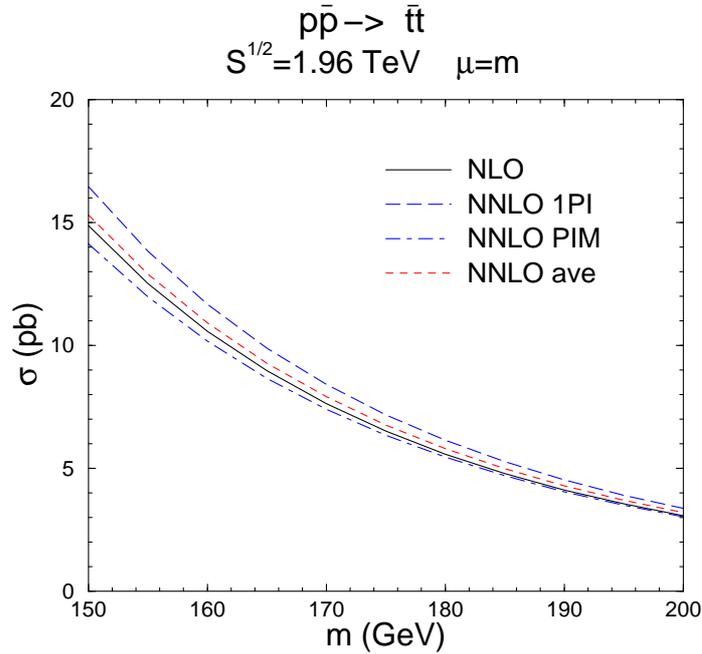


Figure 2.15: The top quark mass dependence of the  $t\bar{t}$  cross section. The exact definition of the terms which are considered in the perturbative expansion referred to as “NNLO” can be found in [52].

the relative contributions to the total  $t\bar{t}$  cross section both for the Tevatron in Run I and Run II and the LHC.

Deviations of the measured cross section from the theoretical prediction could indicate effects beyond QCD perturbation theory. Explanations might include substantial non-perturbative effects, new production mechanisms or additional top-decay modes beyond the *Standard Model*.

#### 2.1.4 Production of the Top Quark via the Weak Interaction

So far, only the top quark pair production has been observed in experiments. Nevertheless, also single top quarks can be produced by electroweak interactions in conjunction with a bottom quark. In this case, a virtual  $W$  boson interacts with a bottom quark producing a top quark in the final state, see Fig. 2.17.

Despite the name, the electroweak single top production is of similar strength as the strong  $t\bar{t}$  production, since there is no phase space suppression and the production of a real  $W$  boson is allowed, and there is no CKM mixing suppression, since  $V_{tb} \approx 1$ . However, the single top production is experimentally more challenging, since its topology is closer to the background topology and there is more background since there are less jets produced.

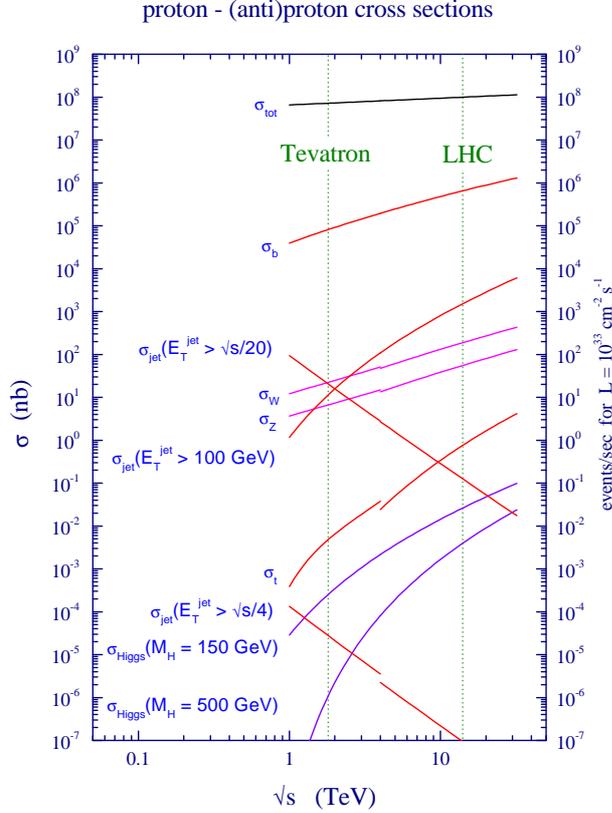


Figure 2.16: QCD predictions for hard scattering cross sections at the Tevatron and the LHC [54].  $\sigma_t$  stands for the  $t\bar{t}$  production cross section. The steps in the curves mark the transition from  $p\bar{p}$  (Tevatron) to  $pp$  (LHC) scattering.

The s-channel process (left diagram in Fig. 2.17) has the least theoretical uncertainties. This process is very similar to the direct production of an on-shell  $W$  boson. However, to decay into a top quark and a bottom quark, the  $W$  boson has to be off-shell. The large virtuality of the  $W$  boson heavily reduces the cross section. It is expected to be  $0.88 \pm 0.07$  pb [59] at the Tevatron. In the  $W$ -gluon fusion t-channel process, a gluon splits into a  $b\bar{b}$  pair. The bottom quark interacts with a virtual  $W$  boson emitted from a quark from the other hadron and turns into a top quark. This process has larger theoretical uncertainties in the gluon PDF, but a larger cross section as well. Its predicted value is  $1.98 \pm 0.21$  pb [59]. In Run I, an upper limit on the electroweak top quark production has been set by the CDF and DØ experiments [55, 56], first Run II limits exist with a sensitivity comparable

	$q\bar{q} \rightarrow t\bar{t}$	$gg \rightarrow t\bar{t}$
Tevatron Run I ( $\sqrt{s} = 1.8$ TeV, $p\bar{p}$ )	90 %	10 %
Tevatron Run II ( $\sqrt{s} = 1.96$ TeV, $p\bar{p}$ )	85 %	15 %
LHC ( $\sqrt{s} = 14.0$ TeV, $pp$ )	10 %	90 %

Table 2.3: Relative contributions to the leading order  $t\bar{t}$  cross section for the Tevatron Run I and Run II and the LHC. The higher the center-of-mass energy, the higher the contribution from the gluon fusion process, which is due to the increased gluon probability density at lower values of accessible  $x$ .

to Run I:

$$\begin{aligned} D\mathcal{O} & : \sigma_{\text{single top}}(\text{s- +t-channel}) < 23 \text{ pb [57]} \\ \text{CDF} & : \sigma_{\text{single top}}(\text{s- +t-channel}) < 17.8 \text{ pb [58]}. \end{aligned}$$

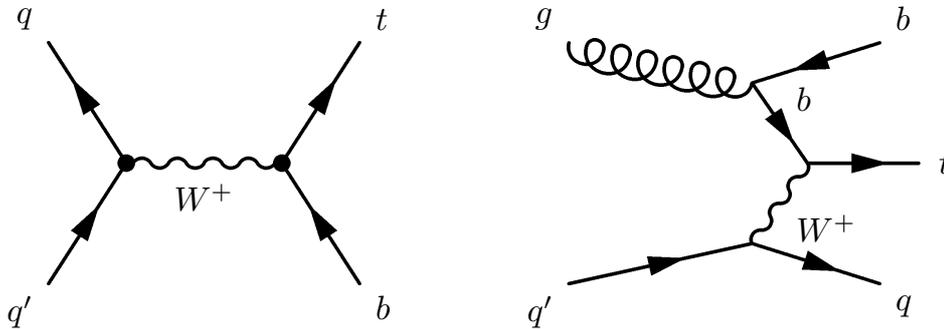


Figure 2.17: Single top quark production via the weak interaction. The main contributions at the Tevatron arise from the s-channel process (left) and the t-channel (right).

### 2.1.5 Decay of the Top Quark

The top quark decays via weak interaction, according to  $\frac{-ig}{2\sqrt{2}} \bar{t} \gamma^\mu (1 - \gamma^5) V_{tb} b W_\mu$ . The *Standard Model* predicts a branching fraction  $BR(t \rightarrow Wb) > 0.998$ , because  $|V_{tb}| \simeq 1$ . Neglecting the mass of the  $b$  quark and higher order terms the total width of the top quark,  $\Gamma_t$ ,

which increases with the top quark mass, can be expressed as:

$$\begin{aligned}\Gamma_t &= \frac{G_F m_t^3}{8\pi\sqrt{2}} \left(1 - \frac{m_W^2}{m_t^2}\right)^2 \left(1 + 2\frac{m_W^2}{m_t^2}\right) \left[1 - \frac{2\alpha_s}{3\pi} \left(\frac{2\pi^2}{3} - \frac{5}{2}\right)\right] \\ &\simeq 1.50 \text{ GeV}, \text{ (for } m_t = 178.0 \text{ GeV),}\end{aligned}\tag{2.17}$$

with a precision better than 2%. The two-loop QCD corrections have also been calculated [60], thereby improving the overall theoretical accuracy to better than 1%.  $\Gamma_t$  corresponds to the very short lifetime of the top quark of about  $5 \times 10^{-25}$  s.

### 2.1.5.1 Top Quark Branching Ratios and $V_{tb}$

The next most likely decay modes are the off-diagonal CKM decays  $t \rightarrow Ws$  and  $t \rightarrow Wd$ . A measurement of the ratio of branching fractions,  $R = BR(t \rightarrow Wb)/BR(t \rightarrow Wq)$  was performed in Run I [61] and preliminary results in Run II exist [62, 63], both consistent with the *Standard Model* expectation. Within the *Standard Model*,  $R$  can be expressed in terms of CKM matrix elements:

$$R = \frac{|V_{tb}|^2}{|V_{tb}|^2 + |V_{ts}|^2 + |V_{td}|^2}.\tag{2.18}$$

Assuming three generations of quarks and unitarity of the CKM matrix,  $R$  represents a measurement of  $|V_{tb}|$ .

The magnitude of  $V_{tb}$  can also be extracted directly by measuring the cross section for single top quark production via the weak interaction, which is proportional to  $|V_{tb}|^2$  (see also Sect. 2.1.4).

### 2.1.5.2 FCNC

The *Standard Model* predicts very small rates for flavor changing neutral current (FCNC) decays of the top quark. Their observation would indicate the presence of *New Physics*. Limits on the top quark decay modes  $t \rightarrow q\gamma$  and  $t \rightarrow qZ$  were set in Run I [64], the decay  $t \rightarrow qg$  will be accessible at the LHC [29].

### 2.1.5.3 Non-Standard Model Top Quark Decay

Searches for decays into charged Higgs bosons  $t \rightarrow H^+ b$  in  $t\bar{t}$  production were performed in Run I [65], where the  $H^+$  preferably decays to either  $cs$  or  $\tau\nu$ , resulting in a final state different from the *Standard Model* expectation. As a consequence, a significant contribution from  $t \rightarrow H^+ b$  would give rise to a visible *Standard Model* cross sections lower than the expectation (assuming that non-*Standard Model* contributions to  $t\bar{t}$  production are negligible).

### 2.1.6 Signature of $t\bar{t}$ Events

As discussed in Sect. 2.1.5 the top quark decays to a  $W$  boson and a  $b$  quark nearly 100% of the time. The  $W$  boson can decay to a charged lepton and a neutrino or to a  $q\bar{q}'$  pair. While all three lepton generations ( $e, \mu, \tau$ ) are allowed, the hadronic  $W$  decay modes are kinematically limited to the production of first and second generation  $q\bar{q}'$  pairs. At Born level all three leptonic  $W$  decay modes have the same probability, each of the two hadronic modes is three times more likely to occur than a leptonic decay due to the color factor of three. Altogether there are nine potential decay paths, which have all the same probability of  $1/9$  at Born level. Due to higher order corrections this symmetry between the decay modes is slightly broken. A summary of the  $W$  decay modes is shown in Table 2.4. The

decay mode		$BR$ at Born level	$BR$ [39]
$W^+$	$\rightarrow e^+\nu_e$	$1/9$	$(10.68 \pm 0.12)\%$
$W^+$	$\rightarrow \mu^+\nu_\mu$	$1/9$	$(10.68 \pm 0.12)\%$
$W^+$	$\rightarrow \tau^+\nu_\tau$	$1/9$	$(10.68 \pm 0.12)\%$
$W^+$	$\rightarrow u\bar{d}, c\bar{s}$	$2 \cdot 3 \cdot 1/9 = 6/9$	$(67.96 \pm 0.35)\%$

Table 2.4: Leading order and best known branching fractions [39] of the real  $W^+$  boson decay, assuming lepton universality. Identical for the charge conjugates of the modes above ( $W^-$ ).

resulting final states for  $t\bar{t}$  events are shown in Fig. 2.18. Three signatures can be distinguished:

**dilepton** Both  $W$  bosons decay leptonically. The signature is two charged leptons, two neutrinos and two  $b$  quarks, i.e.  $l\nu l'\nu' + \geq 2$  jets.

**lepton-plus-jets** One  $W$  boson decays leptonically and one hadronically. The signature is one charged lepton, one neutrino, a  $q\bar{q}'$  pair and two  $b$  quarks, i.e.  $l\nu + \geq 4$  jets.

**all-jets** both  $W$  bosons decay hadronically. The signature is two  $q\bar{q}'$  pairs and two  $b$  quarks, i.e.  $\geq 6$  jets.

Only the all-jets channel and the decay channels where the charged lepton is an electron or a muon have been analyzed so far (the identification of  $\tau$  leptons is difficult and work is in progress). However, a fraction of the  $\tau$  leptons decays leptonically to an electron or muon and two neutrinos. These events have the same signature as the events where the  $W$  boson decays directly to an electron or a muon and are treated as part of the signal in these channels (the same signature arises from the weak interaction of the neutrinos which does not allow to observe them directly in the detector).

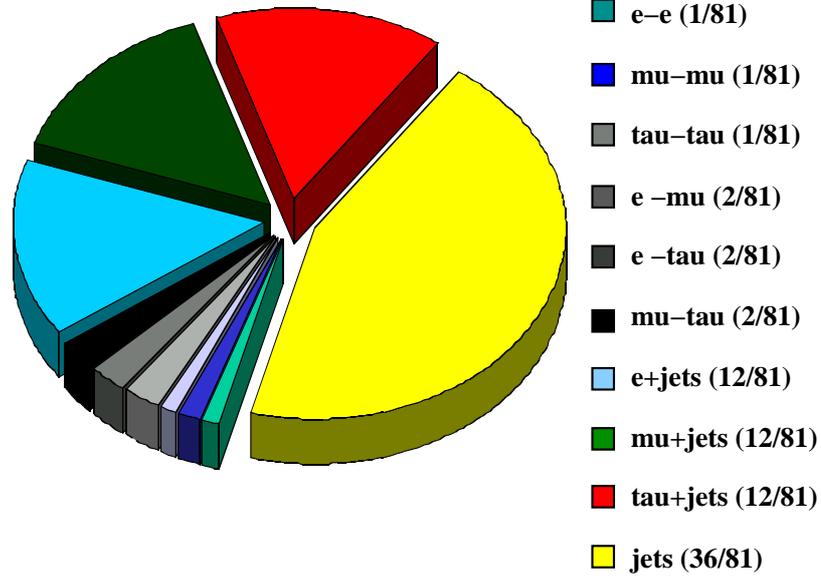


Figure 2.18: Pie chart of the  $t\bar{t}$  event decay channels at Born level.

The leptonic  $\tau$  decay modes are summarized in Table 2.5. As a consequence

decay mode		$BR$
$\tau^-$	$\rightarrow e^- \bar{\nu}_e \nu_\tau$	$(17.84 \pm 0.06)\%$
$\tau^-$	$\rightarrow \mu^- \bar{\nu}_\mu \nu_\tau$	$(17.36 \pm 0.06)\%$

Table 2.5: Branching fractions [39] of the real  $\tau^-$  lepton decay. Identical for the charge conjugates of the modes above ( $\tau^+$ ).

there are three  $t\bar{t}$  decay channels of the dilepton category,  $ee$ ,  $\mu\mu$  and  $e\mu$ , there are two of the lepton-plus-jets category,  $e$ +jets and  $\mu$ +jets, and there is the all-jets category which are all considered for analyses. Table 2.6 summarizes these.

channel	decay mode	$BR$ at Born level	$BR$ [39]
	$t\bar{t} \rightarrow e^+ \nu_e e^- \bar{\nu}_e b\bar{b}$	1/81	$(1.14 \pm 0.02)\%$
	$t\bar{t} \rightarrow e^+ \nu_e \mu^- \bar{\nu}_\mu b\bar{b}$	2/81	$(2.28 \pm 0.04)\%$
	$t\bar{t} \rightarrow \mu^+ \nu_\mu \mu^- \bar{\nu}_\mu b\bar{b}$	1/81	$(1.14 \pm 0.02)\%$
$ee$	$t\bar{t} \rightarrow e_{(\tau)}^+ e_{(\tau)}^- b\bar{b} + \nu' s$	-	$(1.58 \pm 0.03)\%$
$\mu\mu$	$t\bar{t} \rightarrow e_{(\tau)}^+ \mu_{(\tau)}^- b\bar{b} + \nu' s$	-	$(3.16 \pm 0.06)\%$
$e\mu$	$t\bar{t} \rightarrow \mu_{(\tau)}^+ \mu_{(\tau)}^- b\bar{b} + \nu' s$	-	$(1.57 \pm 0.03)\%$
	$t\bar{t} \rightarrow e^+ \nu_e q\bar{q}' b\bar{b}$	12/81	$(14.52 \pm 0.09)\%$
	$t\bar{t} \rightarrow \mu^+ \nu_\mu q\bar{q}' b\bar{b}$	12/81	$(14.52 \pm 0.09)\%$
$e$ +jets	$t\bar{t} \rightarrow e_{(\tau)}^+ q\bar{q}' b\bar{b} + \nu' s$	-	$(17.11 \pm 0.11)\%$
$\mu$ +jets	$t\bar{t} \rightarrow \mu_{(\tau)}^+ q\bar{q}' b\bar{b} + \nu' s$	-	$(17.04 \pm 0.11)\%$
all-jets	$t\bar{t} \rightarrow q\bar{q}' q\bar{q}' b\bar{b}$	36/81	$(46.19 \pm 0.46)\%$
	$t\bar{t} \rightarrow \tau$ final states	17/81	$(20.21 \pm 0.13)\%$

Table 2.6:  $t\bar{t}$  decay channels, Born level and best known branching fractions [39]. The charge conjugated final states are implied.  $\ell_{(\tau)}$ , with  $\ell = (e, \mu)$ , include both the decay modes  $W \rightarrow \ell\nu$  and  $W \rightarrow \tau\nu \rightarrow \ell + \nu\nu$ . The branching fractions considered for the  $D\mathcal{O}$  analyses are denoted by  $ee$ ,  $\mu\mu$  and  $e\mu$  [66],  $e$ +jets and  $\mu$ +jets in the analyses presented here, and all-jets [67].

## 2.2 The Physics of the Bottom Quark

There are two  $b$  quarks in the final state of a  $t\bar{t}$  event (the top quark decays to a  $W$  boson and a  $b$  quark nearly 100% of the time). The presence of  $b$  quarks distinguishes the  $t\bar{t}$  final state from most of the background processes. As a consequence, identifying the bottom flavor of the corresponding jet can be used as a selection criteria. The hadronization of the bare  $b$  quark leads to  $b$  hadrons. In many aspects, these  $b$  hadrons are significantly different from all other particles, they have a

- long lifetime,
- large mass,
- high decay multiplicity,
- substantial leptonic branching ratio, typically  $\sim 10\%$  per lepton.

The most important property for the selection of  $b$  hadrons is their lifetime, which is around 1.6 ps ( $c\tau \sim 450\mu\text{m}$ ). This means that flight distances are of order 5 mm for a  $b$  hadron of 50 GeV, typical for  $b$  hadrons from a top quark decay, see Fig. 2.19.

The  $b$  quark contains most of the momentum of the  $b$  hadron. In the spectator model the decay of the  $b$  hadron is independent of the light quark in the hadron. This implies that the lifetime of all  $b$  hadrons should be the same, which is approximately the case [39]. The  $b$  quark decays via the weak interaction into a  $c$  quark and a  $W$  boson of high virtuality, due to the large mass difference of the  $b$  quark and the  $W$  boson. These circumstances lead to the long lifetime.

Hadrons composed of  $c$  quarks have a 2-3 times shorter lifetime. Due to the smaller mass difference of the  $c$  quark and the light quark in the hadron, a  $W$  boson can be exchanged allowing for additional decay modes and therefore leading to a shorter lifetime. Also the CKM matrix elements are larger than the allowed CKM matrix elements for the  $b$  quark (the top quark decay in the latter is highly suppressed due to the limited phase space).

The top quark, however, is much heavier than the  $W$  boson and decays promptly due to the large phase space available and due to  $V_{tb} \approx 1$ . Light quarks dominantly hadronize by building neutral and charged pions. The neutral pions decay through the electromagnetic interaction ( $\pi^0 \rightarrow \gamma\gamma$ ) in about  $10^{-16}$  seconds, whereas the charged pions cannot. Being the lightest hadron the charged pion decays through the weak interactions resulting in a very long lifetime of  $2.6 \times 10^{-8}$  seconds, so that the charged pions interact with the detector before they decay.

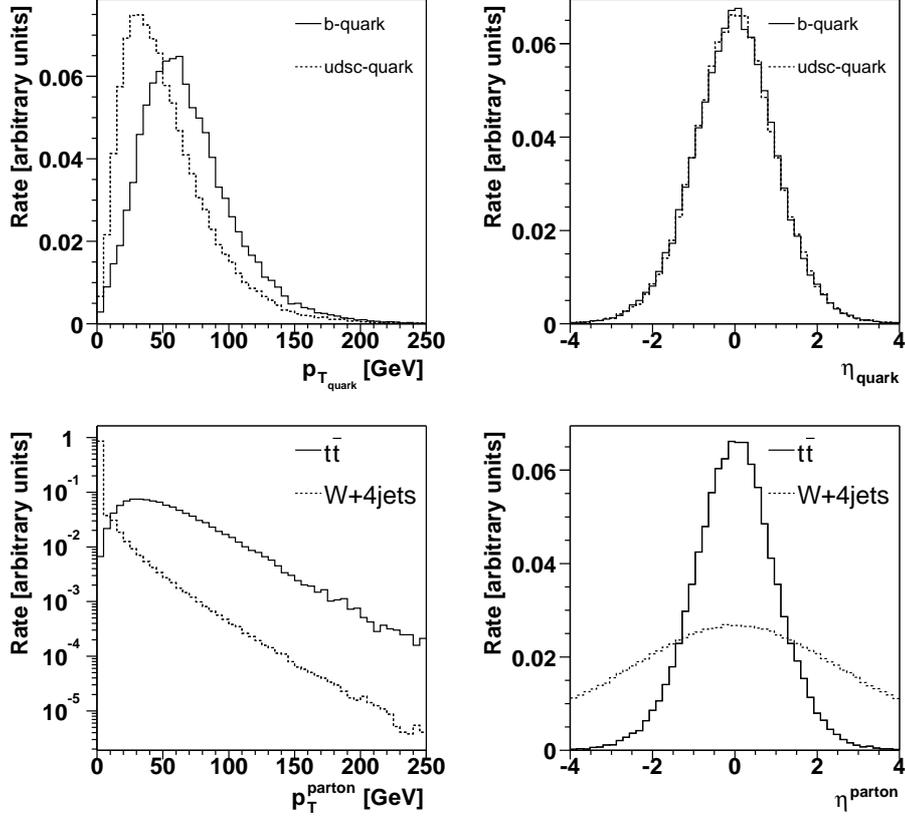


Figure 2.19: Transverse momentum and  $\eta$  distribution of quarks and gluons which will hadronize and will be detected as jets in the calorimeter. Top:  $b$  quarks from the top quark decay and  $udsc$  quarks from the hadronic  $W$  boson decay in  $t\bar{t}$  events. Bottom:  $udsc$  quarks from the hadronic  $W$  boson decay in  $t\bar{t}$  events and  $uds$  quarks and gluons produced in the electroweak  $W$  boson production.

## 2.3 Signal and Background Signature

In this thesis the lepton-plus-jets final states are discussed. Its experimental signature and the background processes which have the same signature are discussed in the following sections.

### 2.3.1 $t\bar{t}$ Events in the Lepton + Jets Final State

The signature of  $t\bar{t}$  events with lepton-plus-jets final states can be described as:

- One charged lepton ( $e$  or  $\mu$ ) from a leptonic  $W$  boson decay with high transverse momentum.

- Missing transverse energy ( $\cancel{E}_T$ ) from the neutrino emission of the leptonic  $W$  boson decay.
- Two  $b$  jets, from the hadronization of the  $b$  quarks.
- Two non- $b$  jets ( $u, d, s, c$ ) from the hadronic  $W$  decay.
- Additional jets due to initial (ISR) and final state radiation (FSR).

A sketch of a  $t\bar{t}$  event with a muon in the final state is shown in Fig. 2.20. Due to the  $V - A$  structure of the  $W$  decay [45], the neutrino has on average

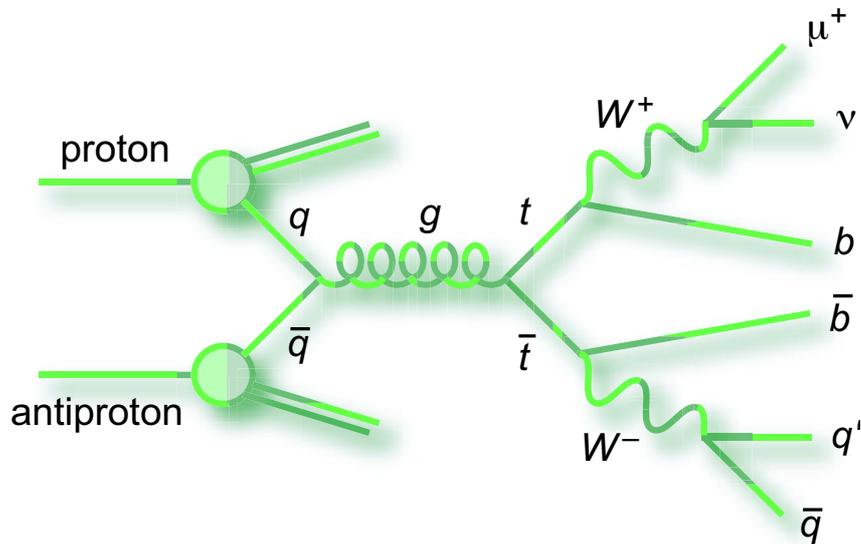


Figure 2.20: Sketch of the production and decay of a  $t\bar{t}$  pair in the  $\mu$ -plus-jets channel.

a higher  $p_T$  than the charged lepton, shown for the muonic  $W$  boson decay in Fig. 2.21. Both the charged lepton and the neutrino are produced in the central region. The bottom quark from the top quark decay has a considerably larger transverse momentum than the light quarks from the hadronic decay of the second  $W$  boson, shown in Fig. 2.19.

### 2.3.2 Electroweak $W$ Boson Production in Association with Jets

The main source of  $W$  bosons at a hadron collider like the Tevatron is the exclusive direct electroweak production through  $q\bar{q}'$  annihilation, as shown in the Feynman diagram in Fig. 2.22. Bosons produced in this process (Drell-Yan process in case of leptonic decay) have no transverse momentum and are almost fully polarized along the antiproton direction due to the  $V - A$

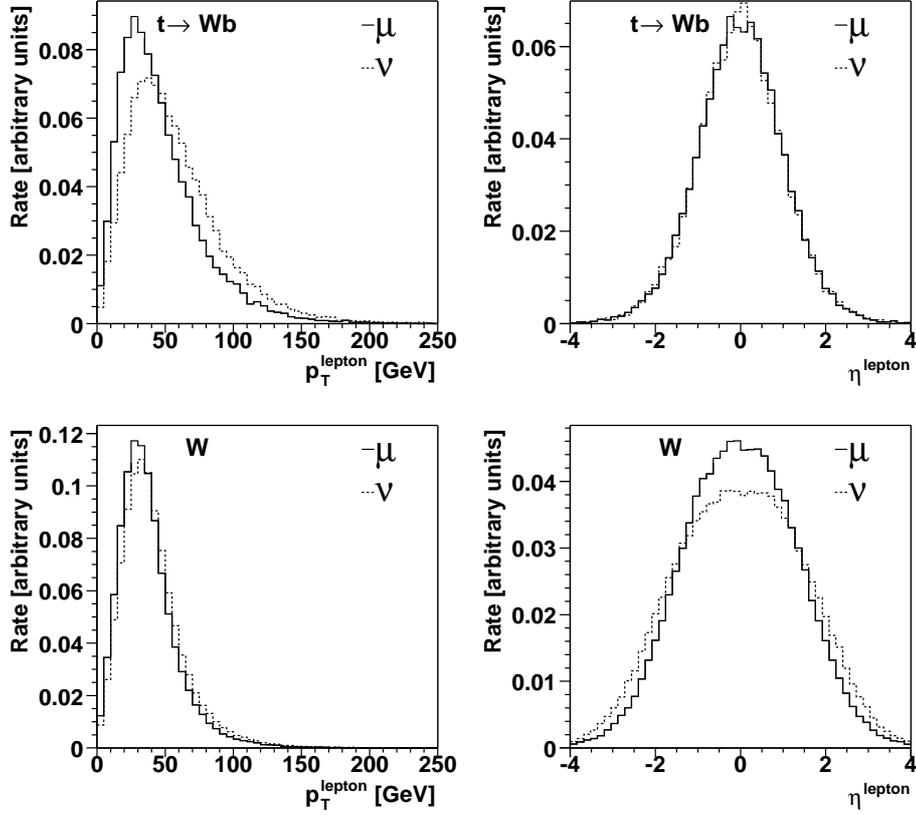


Figure 2.21: Transverse momentum and  $\eta$  distribution of the leptons from the leptonic  $W$  boson decay in a  $t\bar{t}$  process (top) and in the electroweak  $W$  production (bottom).

coupling [45]. However, a gluon emitted from one of the initial quarks can generate the transverse momentum of the  $W$  boson and adds at the same time an additional gluon in the final state. Alternatively, an initial gluon can split into a quark-antiquark pair and one of these quarks interacts with an initial quark from the other hadron to produce the  $W$  boson together with a quark in the final state. The corresponding Feynman diagrams are depicted in Fig. 2.23.  $W$  boson production with two partons in the final state is shown in Fig. 2.24.

With each additional parton of the strong coupling a vertex proportional to  $\sqrt{\alpha_s}$  is added. The lowest order matrix elements for the production of a  $W$  boson in association with up to four partons at hadron colliders have been computed [8] using various techniques to control the rapid growth of the number of contributing Feynman diagrams as the number of partons increases. The cross section is computed by Monte Carlo integration of the final state parton phase space, however, with a large uncertainty. An

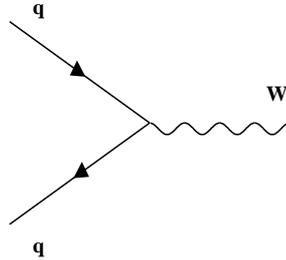


Figure 2.22: Feynman diagram for the  $W + 0$  parton process. In case of a leptonic  $W$  decay this process is also called ‘Drell-Yan production’.

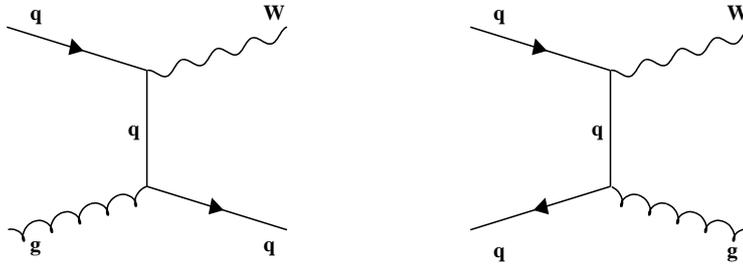


Figure 2.23: Some examples of Feynman diagrams for the  $W + 1$  parton process.

effective scaling law (Berends scaling [8]) has been found which relates the cross section of  $W + n$  jets production with  $W + (n + 1)$  jets:

$$\frac{\sigma(W + (n + 1)_{\text{jets}})}{\sigma(W + n_{\text{jets}})} = \alpha, \quad (2.19)$$

where  $\alpha$  is in lowest order related to  $\alpha_s$ . The numerical value of  $\alpha$ , however, depends on the  $p_T$  cuts applied to the jets in the sample as they set the effective scale for the scaling. At higher orders the picture gets more complicated as more diagrams get involved in the production of  $n$  or  $n + 1$  jets.

The signature of the electroweak  $W$  boson production with a subsequent leptonic decay in association with four jets in the final state is identical to the lepton-plus-jets signature of the  $t\bar{t}$  decay, discussed in Sect. 2.3.1. It is found to be the dominant background to the  $t\bar{t}$  process in the lepton-plus-jets channel.

Up quarks tend to carry, on average, a larger momentum fraction of the proton than down quarks. The same holds for anti-up and anti-down quarks

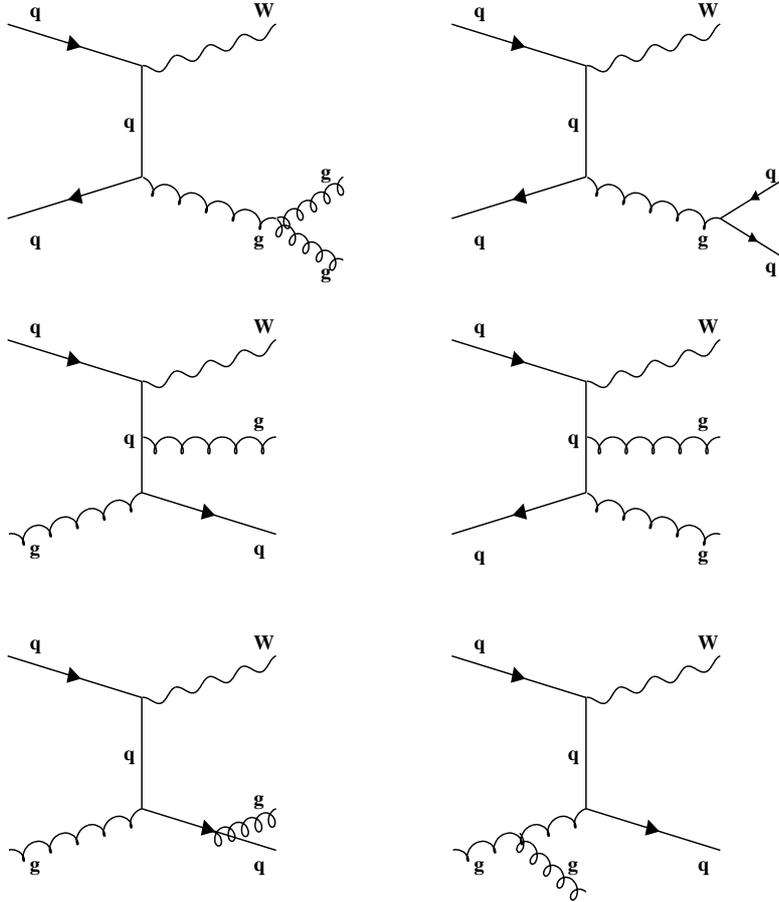


Figure 2.24: Some examples of Feynman diagrams for the  $W + 2$  partons process.

in the antiproton. Therefore  $W^+$  bosons, produced in a  $u(p) + \bar{d}(\bar{p})$  reaction, have more likely a boost in the proton direction, while  $W^-$  bosons, produced in a  $d(p) + \bar{u}(\bar{p})$  reaction, have more likely a boost in the antiproton direction (see Fig. 2.25). This charge asymmetry for directly produced  $W$  bosons also results in a corresponding charge asymmetry of the decay leptons from the  $W$ .  $W$  bosons produced in top decays do not show this asymmetry. As decay products of the top or anti-top quark they follow the direction of their mother particles, which are produced symmetrically in  $\eta$  without any preference for large  $\eta$  values, i.e. the forward direction. Therefore the  $W$  bosons from top decays are produced centrally (see Fig. 2.25). The  $W$  bosons from the top quark decay have, on average, a larger transverse momentum than the  $W$  bosons produced via the weak interaction due to the large top quark mass. These properties are transferred to the leptons from the  $W$

boson decay which have a larger transverse momentum and are produced at lower  $|\eta|$  in  $t\bar{t}$  events. The jets produced in association with the electroweak  $W$  boson originate mainly from QCD bremsstrahlung. The corresponding cross section is infrared and collinear divergent, resulting in jets with low transverse momenta and with large values of  $|\eta|$ , see Fig. 2.19.

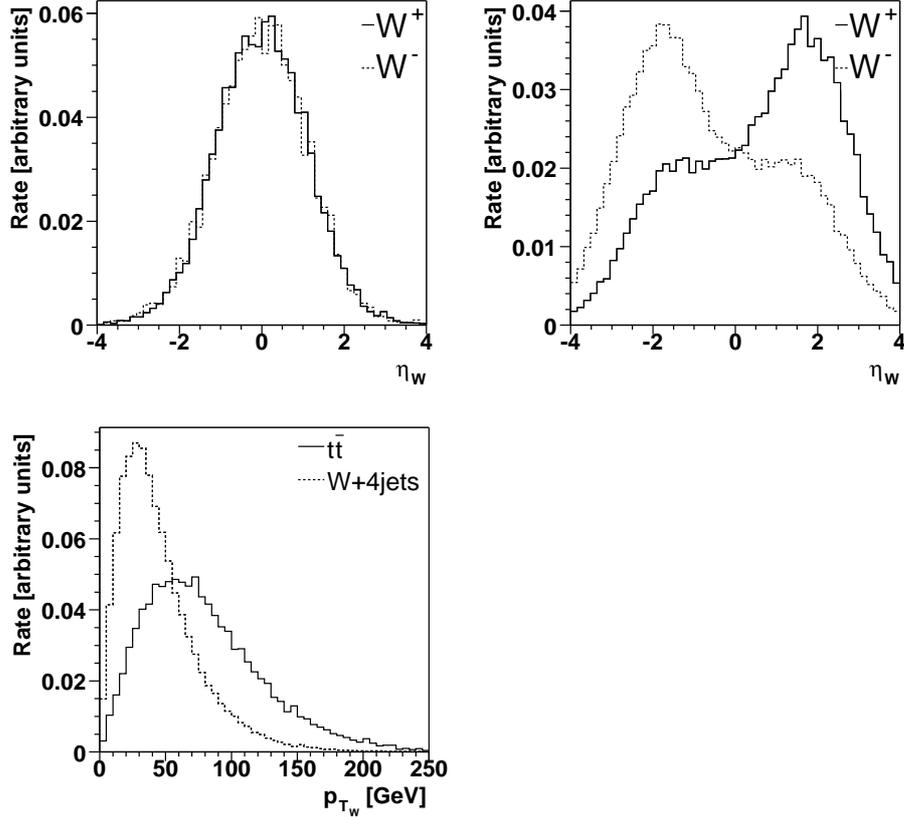


Figure 2.25:  $\eta$  distribution of the  $W$  boson in a  $t\bar{t}$  process (left) and in the electroweak  $W$  production (right) and the transverse momentum distributions of the  $W$  boson in the two different production modes (bottom).

### 2.3.3 QCD Multijet Production

One of the largest cross sections at the Tevatron is the QCD multijet production through the strong interaction. The production mechanisms of additional jets are the same as in the  $W$ -plus-jets production, through QCD bremsstrahlung, i.e. gluon emission and gluon to  $q\bar{q}$  splitting. The cross section for QCD-multijets production decreases with each additional jet due to  $\alpha_s$  being  $< 1$  ( $0.1 \lesssim \alpha_s \lesssim 0.2$  for jets that can be reconstructed in the detector).

Electromagnetically fluctuating jets are jets containing a large fraction of electromagnetically interacting particles<sup>1</sup>. These jets can fake electrons in the detector.  $c$  and  $b$  quarks have a substantial branching ratio for decay modes which involve both a charged lepton and a neutrino in the final state. The muon originating from this semi-leptonic heavy quark decay can appear to be isolated if the remaining part of the jet is too soft to be reconstructed. The neutrino from the heavy quark decay, the misreconstruction of the jet and the mismeasurement of the muon momentum lead to  $\cancel{E}_T$  in the detector. These effects can lead to the same signature as the signal  $t\bar{t}$  events.

### 2.3.4 Small Additional Background Processes

In the following sections further processes are discussed which have only a very small contribution to the background of the  $t\bar{t}$  process with lepton-plus-jets final states.

#### 2.3.4.1 $Z/\gamma^*$ Production in Association with Jets

The  $Z/\gamma^*$  production, has the same Feynman diagrams as the electroweak  $W$  boson production (Fig. 2.22-2.24, by replacing the  $W$  by a  $Z/\gamma^*$ ). The  $Z$  production rate, where the  $Z$  boson decays to two charged leptons of the same lepton generation ( $e, \mu, \tau$ ), is roughly a factor of ten smaller than the corresponding  $W$  rate as shown in Fig. 2.26:  $\frac{\sigma_{p\bar{p} \rightarrow Z+X} \cdot BR(Z \rightarrow \ell\ell)}{\sigma_{p\bar{p} \rightarrow W+X} \cdot BR(W \rightarrow \ell\nu)} \approx 0.1$ . In the case when one of the two charged leptons is not detected due to the limited acceptance, then the detector signature of the  $Z/\gamma^*$  production in association with  $\geq 4$  jets is the same as the lepton-plus-jets signature of the  $t\bar{t}$  decay.

#### 2.3.4.2 Single Top Production

The production of a single top quark via the weak interaction is discussed in Sect. 2.1.4. For a leptonic decay of the  $W$  boson from the top quark decay the experimental signature is one charged lepton,  $\cancel{E}_T$  from the neutrino and two  $b$  jets. For the t-channel there is an additional light jet. Due to ISR and FSR even more jets can be produced, which leads to the same detector signature as the lepton-plus-jets final state of the  $t\bar{t}$  decay.

#### 2.3.4.3 Vector Boson Pair Production

Pair production of  $W$  bosons occurs primarily through interactions represented by diagrams in Fig. 2.27. The cross section is measured in Run II [69]

---

<sup>1</sup>Most of the hadrons in a jet are pions of which on average one third are  $\pi^0$ 's which in turn decay to two photons; the actual fraction, however, can fluctuate strongly.

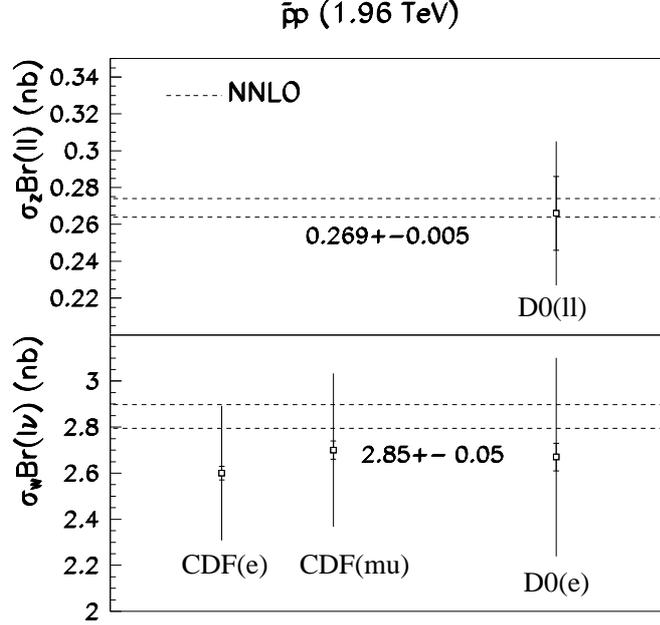


Figure 2.26: The NNLO  $W$  and  $Z$  boson production rates in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96$  TeV compared to the preliminary results from Run II of the Tevatron. The area between the dashed lines gives the  $1\sigma$  band uncertainty in the calculation [68].

to be

$$\sigma_{p\bar{p} \rightarrow W^+W^-+X} = 13.8_{-3.8}^{+4.3} \text{ (stat.) }_{-0.9}^{+1.2} \text{ (syst.) } \pm 0.9 \text{ (lumi.) pb,}$$

and is in good agreement with the next-to-leading order calculation [70], given in Table. 2.7. The cross sections for the production of  $Z$  boson pairs, and for the production of a  $W$  boson in association with a  $Z$  boson are also calculated at next-to-leading order and given in Table. 2.7. They have not yet been observed, but there are ongoing analyses.

$W^+W^-$	$ZW^+$ or $ZW^-$	$ZZ$
13.0-13.5 pb	1.95-2.01 pb	1.56-1.60 pb

Table 2.7: Total NLO cross section for the given di-boson processes for a  $p\bar{p}$  center-of-mass energy of 2 TeV.

Decay modes where one of the two bosons decays leptonically and one hadronically can lead to similar final state signatures as the one of the signal  $t\bar{t}$  process.

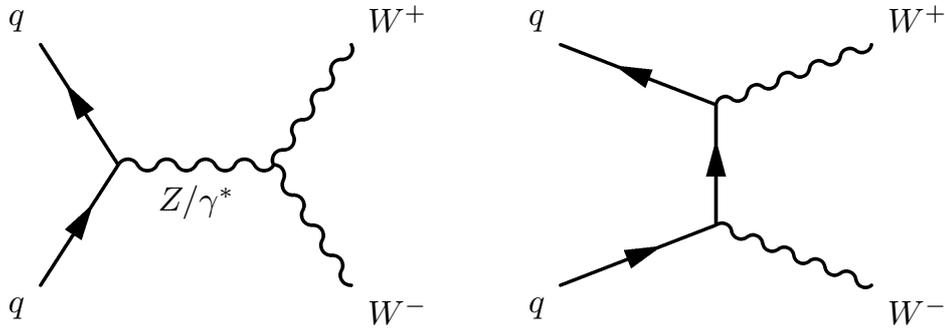


Figure 2.27: Feynman diagrams which represent leading order  $W$  boson pair production processes.

## 2.4 Monte Carlo Simulation

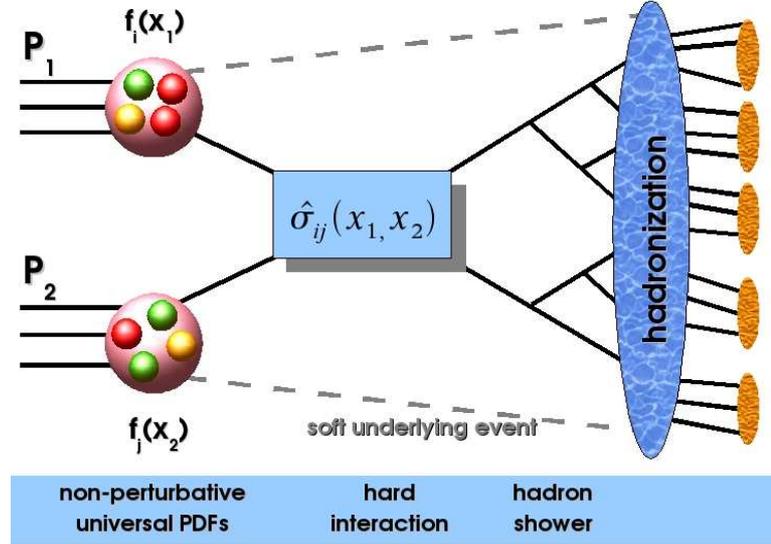
The study of the signature of the signal and background processes, the optimization of the event selection, as well as the accurate measurement of selection efficiencies require the Monte Carlo simulation of the data events, including the hard scattering interaction, hadronization, detector response and digitization.

The simulation of a  $p\bar{p}$  interacting makes use of the factorization (see Sect. 2.1.3.1) in a short distance hard scattering interaction, calculable in perturbative QCD, and the long range physics, including the parton momentum distributions and further soft physics interaction, referred to as the *underlying event*. As discussed in Sect. 2.1.3.1 this separation introduces the artificial factorization scale  $Q^2$ . Additional effects, such as *multiple proton interactions* and *pile-up* in the detector can occur.

Figure 2.28 shows a sketch of the  $p\bar{p}$  interactions. The full chain of the simulation is described in the following.

**The hard scatter interaction** is described by calculating the leading order matrix element using ALPGEN [71]. The set of parton distribution functions used is CTEQ 5L [43] and CTEQ 6.1M [41], the latter derived in NLO, which is in principle not adequate to be used in association with a leading order matrix element, however, proper PDF uncertainties are at present only available for NLO PDFs, and numerically the change in the  $t\bar{t}$  cross section is found to be small.

**The underlying event** is comprised of a hard component and of a soft component. The hard component describes the particles that arise from initial and final state radiation and from the outgoing hard scattered partons. The soft component consists of *beam-beam remnants* and *multiple parton interactions*.

Figure 2.28: Sketch of a  $p\bar{p}$  interaction

**The beam-beam remnant** describes the outgoing partons of the  $p\bar{p}$  interaction, which did not take part in the hard scattering process. The color connection between these spectator partons and the two partons from the hard scattering is the origin of this soft interaction and is hard to model.

**Multiple parton interactions** describe the possibility that a hard scattering event also contains “semi-hard” interactions between the remaining partons from the identical  $p\bar{p}$  pair. Again, there is a color connection between the “semi-hard” and the hard scattering partons, and in addition a dependence on the  $p_T$  of the hard scattering partons.

The *transverse* region, defined as the phase space around the plane orthogonal to the jet with highest  $p_T$  in the event, is sensitive to the *underlying event*. A data to Monte Carlo comparison of the average charged particle density and  $p_T$  distribution in the *transverse* region leads to a measurement of the *underlying event* and a tuning of its MC modeling, so-called “Tune A” [72].

PYTHIA 6.202 and JETSET [73], including *multiple parton interactions*, are used to model the *underlying event*. PYTHIA models the soft component of the underlying event with color string fragmentation.

**Multiple proton interactions** can occur when more than one  $p\bar{p}$  interaction in the two colliding bunches of hadrons takes place.

The multiple proton interactions are simulated by superimposing min-

imum bias data events to the event. Minimum bias events are defined as events which show a minimum activity in the detector, i.e. not being triggered by a high  $p_T$  lepton, jet or  $\cancel{E}_T$ . The number of added events is taken from a Poisson distribution with a mean between 0.5 and 0.8 events. This number is luminosity dependent.

**Pile-up** describes overlapping  $p\bar{p}$  interactions from consecutive bunch crossings in the detector, which are reconstructed in one event.

The pile-up is currently simulated, in the future randomly recorded data events (so-called zero bias) will be used for the modeling.

**Hadronization:** The collections of partons must then be hadronized into colorless mesons and baryons. Different approaches are used by the event generators. The model implemented in PYTHIA [73] splits gluons into  $q\bar{q}$  pairs and turns them into hadrons via the string fragmentation model. HERWIG [74] uses an approach where colorless clusters are formed from quarks and gluons with low invariant mass, which are turned into hadrons.

**Detector response:** The passage of the particles through and the response of the different detector components is simulated using the software package *dogstar*. The description of the detector material and geometry is described by the GEANT3 program [75], taking into account the present understanding of the detector and test beam results.

**Digitization:** The next step is the simulation of the digitization of the detector response, realized by the software package *d0sim*. After this step, data and MC simulation are on the same level, called raw data, and can be treated identically.

**Reconstruction:** The digitized information of nearly one million channels of detector response is processed by the reconstruction program *reco* to define higher level objects and properties from which the event kinematics can be inferred as described in Sect. 4.

**Analysis tool:** The analysis tool used for the analyses presented is Top-Analyze [76, 77], a software-package which processes the reconstructed events further and produces manageable ROOT files.

All software packages used are from versions of the so-called production release number 14 (p14).

The Monte Carlo samples used for the analyses presented are discussed in the following.

Generation Parameters	$t\bar{t}$	$W$ -plus-jets		$Z$ -plus-jets
		<i>topo</i>	<i>btag</i>	
PDF	CTEQ6.1M	CTEQ5L	CTEQ6.1M	CTEQ5L
$Q^2$	$m_t^2$	$M_W^2 + \sum p_{T_j}^2$	$(\frac{M_W}{2})^2$	$(0.5 \times m_Z)^2$
Underlying event	Tune A	Tune A	none	Tune A
$p_T(\text{parton})$	none	$> 8$ GeV	$> 12$ GeV	$> 12$ GeV
$ \eta(\text{parton}) $	none	$< 3.0$	$< 2.7$	$< 2.7$
$\Delta R(\text{parton}, \text{parton})$	none	$> 0.4$	$> 0.4$	$> 0.4$
$p_T(l)$	none	none	$> 12$ GeV	$> 12$ GeV
$ \eta(l) $	none	none	$< 2.7$	$< 10.0$
$p_T(\nu)$	none	none	8 GeV	none
No. of min bias events	0.8	0.5	0.5	0.5

Table 2.8: Main generation parameters for  $t\bar{t}$ ,  $W$ -plus-jets and  $Z$ -plus-jets. The  $W$ -plus-jets MC generation parameters for the later performed *topo* analysis are looser and in particular the choice of  $Q^2$  is found to describe the data even better than the respective parameters used in the *btag* analysis.

### 2.4.1 $t\bar{t}$ Signal Simulation

The production and decay of the  $t\bar{t}$  signal is simulated using ALPGEN 1.2 [71], which includes the complete  $2 \rightarrow 6$  Born level matrix elements, followed by PYTHIA to simulate the *underlying event*, including ‘‘Tune A’’, and the hadronization. This procedure takes advantage of the full spin correlation information for top quarks that is provided in ALPGEN 1.2. The top quark mass is set to 175 GeV. EVTGEN [78], known to successfully describe the spin correlations between the decay particles, is used to provide the various branching fractions and lifetimes for the following  $b$  quark states:  $B^0$ ,  $B^+$ ,  $B_s^0$ ,  $B_c^+$ , and  $\Lambda_b$ . The factorization scale for the calculation of the  $t\bar{t}$  process is  $Q = m_t$ .

MC samples are generated separately for the three signatures, dilepton, lepton-plus-jets and all-jets, according to the decay of the  $W$  bosons. Leptons include  $e$ ,  $\mu$ , and  $\tau$ , with  $\tau$ s decaying inclusively using TAUOLA [79].

For the mass dependence of the cross-section measurement, samples similar to the above one are generated for the following top quark mass values: 150, 160, 170, 175, 180, 190, 200 GeV.

The main generation parameters are summarized in Table 2.8.

### 2.4.2 $W$ -plus-jets Background Simulation

The  $W$ +jets background is simulated using ALPGEN 1.2 [71] followed by PYTHIA to simulate the *underlying event* and the hadronization. It includes the correct masses for  $c$  and  $b$  quarks. In the *btag* analysis each data sample ( $\mu$ -plus-jets and electron-plus-jets) is subdivided into four disjoint event

samples with 1, 2, 3 and  $\geq 4$  jets in the final state, and each sample represents an individual counting experiment. In the *topo* analysis only events with  $\geq 4$  jets are considered. This motivates the generation of the following exclusive samples:

- $Wjjjj$ ,  $Wcjjj$ ,  $Wc\bar{c}Jj$  and  $Wb\bar{b}Jj$ ,
- $Wjjj$ ,  $Wcjj$ ,  $Wc\bar{c}J$  and  $Wb\bar{b}J$ ,
- $Wjj$ ,  $Wcj$ ,  $Wc\bar{c}$  and  $Wb\bar{b}$ ,
- $Wj$  and  $Wc$ ,

where  $j$  is any of  $u, d, s, g$ , and  $J$  is any of  $u, d, s, g, c$  parton ( $Wc\bar{c}c\bar{c}$ ,  $Wb\bar{b}c\bar{c}$ , and  $Wb\bar{b}b\bar{b}$  processes are not included; their cross sections are negligible). The main generation parameters are summarized in Table 2.8. No parton-level cuts are applied on the heavy quarks ( $c$  or  $b$ ), except the  $c$  quark in the single  $c$  quark production. The corresponding production cross sections for the  $W$ -plus-jets MC samples are summarized in Table 2.9.  $W$  bosons are forced to decay to leptons, combining  $W \rightarrow e\nu$ ,  $W \rightarrow \mu\nu$ , and  $W \rightarrow \tau\nu$  decays;  $\tau$ s are forced to semileptonic decays using TAUOLA [79] (with the respective fraction of  $W \rightarrow \tau\nu$  adjusted in the overall sample to correctly reflect the contributions to the  $e$  and  $\mu$  channels).

For the determination of the systematic uncertainty associated to the factorization scale in the *topo* analysis, another sample is created with the same generation parameters but using  $Q^2 = \langle p_{T_j} \rangle^2$ .

process	$\sigma$ (pb)	process	$\sigma$ (pb)	process	$\sigma$ (pb)	process	$\sigma$ (pb)
$Wj$	424.90	$Wjj$	126.81	$Wjjj$	32.48	$Wjjjj$	8.89
$Wc$	16.01	$Wcj$	7.60	$Wcjj$	2.38	$Wcjjj$	0.64
		$Wb\bar{b}$	4.61	$Wb\bar{b}J$	2.00	$Wb\bar{b}Jj$	0.81
		$Wc\bar{c}$	11.43	$Wc\bar{c}J$	4.68	$Wc\bar{c}Jj$	1.93

Table 2.9:  $W$ +jets processes in ALPGEN and their cross-sections for the leptonic  $W$  boson decay,  $\sigma \equiv \sigma_{p\bar{p} \rightarrow W+jets} BR(W \rightarrow e\nu)$ , where  $j = u, d, s, g$  and  $J = u, d, s, g, c$ .

### 2.4.2.1 Jet-Parton Matching

The subject of this section is the discussion of the consistent combination of the leading order parton level calculations, performed by ALPGEN, with the partonic evolution given by the shower MC programs PYTHIA and JETSET. The problem that occurs in this merging process can be denoted as the problem of *double counting* of configurations leading to the same final state. This technique is only used for the MC of the *btag* analysis. Since this

analysis treats exclusive jet multiplicity bins as independent channels, the jet-parton matching procedure is a simplification of the matching schemes proposed and used for the combination of inclusive  $W/Z + n$  jet MC files [80, 81, 82].

**Double counting:** The parton shower MC programs serve to model the higher order corrections to the leading order matrix element calculation in all orders of  $\alpha_s$ . Two sources of double counting can be identified:

The parton shower MC programs transform the leading order matrix element of  $W + (n - m)$  jets with  $(n - m)$  jets exclusively in the final state into a final state involving  $n$  jets by adding  $m$  jets due to initial and final state radiation. This implies a *double counting* of events with  $n$  reconstructed jets in the final state from all leading order matrix element calculations of  $W + (n - m)$  jets, where the parton evolution leads to the reconstruction of  $m$  additional jets, where  $0 \leq m \leq n$ . This is illustrated in Fig. 2.29 for  $n = 3$  and  $m = 0, 1$ .

The detector acceptance, response and the jet selection criteria lead also to the migration of  $W + (n + k)$  jet events to events with  $n$  reconstructed jets in the final state, where  $k \geq 0$ . In particular the  $k$  jets can be too soft or too forward to be reconstructed or selected.

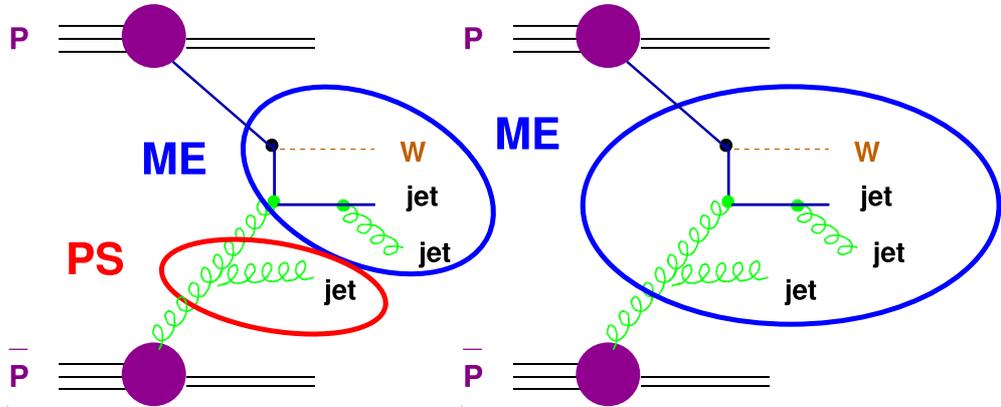


Figure 2.29: Left:  $W + 2$  jets process calculated by the matrix element (ME) and one additional jet generated by the parton shower (PS). Right:  $W + 3$  jets process calculated by the matrix element and no additional jet generated by the parton shower. Both processes lead to the same final state.

A matching of partons, produced by the matrix element calculation, and reconstructed jets is performed in order to eliminate the *double counting*. This matching procedure also reduces the sensitivity of the parton-level cross sections, predicted by ALPGEN, to the parton generation cuts. Two matching procedures can be applied:

**CKKW matching:** The multijet matrix elements are merged with the shower development by reweighting the matrix elements weights with Sudakov form factors and vetoing shower emissions in regions of phase space already covered by the parton level configurations [80, 81]. This matching procedure is named after the initials of the authors (S. Catani, F. Krauss, R. Kuhn and B. R. Webber).

**MLM matching:** Matrix element partons are matched to parton jets [82], proposed and named after M. L. Mangano.

An approximation of the MLM matching is used in the *btag* analysis. Since the parton jets are not available, instead the matching is performed between matrix element partons and reconstructed jets. To explain the matching procedure it is useful to classify the  $W$ -plus-jets MC samples based on the number of heavy flavor ( $c$  or  $b$ ) jets:

- $W$ + light jets, events without  $c$  or  $b$  jets
- $Wc$ : events with one  $c$  jet due to single  $c$  production,
- $W(c\bar{c})$ : events with one  $c$  jet due to double  $c$  production where two  $c$  quarks are merged in one jet or one of the  $c$  jets is outside of the acceptance region,
- $Wc\bar{c}$ : events with two  $c$  jets, and possibly a third  $c$  jet,
- $W(b\bar{b})$ : events with one  $b$  jet due to double  $b$  production where two  $b$  quarks are merged in one jet or one of the  $b$  jets is outside of the acceptance region (single  $b$  production is highly suppressed and neglected),
- $Wb\bar{b}$ : events with two  $b$  jets, and possibly an additional  $c$  jet.

The jet flavor in the MC is determined by a matching with the generated heavy flavor hadrons:

**$b$  jet:** A reconstructed jet is considered a  $b$  jet if it is matched to a  $b$  hadron,  $\Delta R(\text{jet}, \text{hadron}) < 0.5$ ,

**$c$  jet:** A reconstructed jet is considered a  $c$  jet if it is matched to a  $c$  hadron,  $\Delta R(\text{jet}, \text{hadron}) < 0.5$ , and not to a  $b$  hadron,

**light jet:** A reconstructed jet is considered a light jet if it is neither matched to a  $c$  nor to a  $b$  hadron.

### Approximated MLM Jet-Parton Matching

- The number of reconstructed jets is required to be equal to the number of matrix element partons, where  $(c\bar{c})$  and  $(b\bar{b})$  are treated as one parton, otherwise the event is rejected.

- Light jets and  $c$  jets from  $Wc$  events are required to be matched to matrix element partons, i.e.  $\Delta R(\text{jet}, \text{parton}) < 0.5$ , otherwise the event is rejected.
- The fourth jet multiplicity bin is inclusive, i.e. all events with  $\geq 4$  jets reconstructed are considered, therefore, events where additional non-matched light jets exist, are not rejected.

### 2.4.3 $Z$ -plus-jets Background Simulation

The  $Z$ -plus-jets background is simulated using ALPGEN 1.2 [71] followed by PYTHIA to simulate the *underlying event* and the hadronization. The main generation parameters are summarized in Table 2.8. All virtual photon processes (Drell-Yan production) and the interference between the diagrams involving virtual photons and  $Z$  bosons are taken into account. Samples are generated in 3 separate bins of the invariant mass of the dilepton system: 15-60 GeV, 60-130 GeV, and  $> 130$  GeV. The  $Z$  boson is forced to decay to  $ee$ ,  $\mu\mu$ , and  $\tau\tau$ ; the  $\tau$ 's are forced into the semileptonic decay mode using TAUOLA [79] (with the respective fraction of  $Z \rightarrow \tau\tau$  adjusted in the overall sample to correctly reflect the contributions to the  $e$  and  $\mu$  channels).

### 2.4.4 Vector Boson Pair Production Background Simulation

The  $WW$ ,  $WZ$  and  $ZZ$  backgrounds are simulated using ALPGEN 1.2 [71] followed by PYTHIA to simulate the *underlying event* and the hadronization with inclusive  $\tau$  decays.

### 2.4.5 Single Top Background Simulation

The background from single top production is simulated using CompHEP [83], followed by PYTHIA to simulate the *underlying event* and the hadronization with inclusive  $\tau$  decays.

## 2.5 Measurement of the Luminosity

The event rate  $R$  in a collider is proportional to the interaction cross section  $\sigma_{int}$ , and the factor of proportionality is called the instantaneous luminosity  $\mathcal{L}$ :

$$R = \mathcal{L} \cdot \sigma_{int} . \quad (2.20)$$

For the measurement of a cross section, as presented in this thesis for the  $t\bar{t}$  production cross section, both the interaction rate and the instantaneous luminosity have to be measured.

The idea of the measurement of the instantaneous luminosity is to use a reference interaction with known cross section and to measure its rate. The process of choice at the Tevatron is the inelastic  $p\bar{p}$  cross section, which is related to the total and the elastic cross section through:

$$\sigma_{inelastic} \equiv \sigma_{total} - \sigma_{elastic}. \quad (2.21)$$

In contrast to  $e^+e^-$  or  $e\gamma$  colliders, where the reference cross section can be calculated with high precision, at hadron colliders both the total and the elastic cross section are measured separately to determine the inelastic cross section. The total cross section can be determined from event rates only without knowing the luminosity. The key is the usage of the optical theorem.

The measurement of the inelastic event rate with the DØ luminosity system is discussed in Sect. 3.3.2.

### 2.5.1 The Total $p\bar{p}$ Cross Section and the Optical Theorem

According to Eq. 2.20 and Eq. 2.21 the total cross section can be written in terms of the instantaneous luminosity  $\mathcal{L}$  and the event rate  $R$ :

$$\sigma_{total} = \frac{1}{\mathcal{L}}(R_{el} + R_{inel}). \quad (2.22)$$

On the other hand the optical theorem relates the total cross section to the imaginary part of the forward elastic scattering amplitude:

$$\sigma_{total} = \frac{4\pi}{k} \text{Im}F(\Theta)^{\Theta=0}, \quad (2.23)$$

where  $k$  is the momentum of the incoming hadron. Squaring of Eq. 2.23 yields:

$$\sigma_{total}^2 = \frac{16\pi^2}{k^2} \frac{\text{Im}F(0)^2}{\text{Im}F(0)^2 + \text{Re}F(0)^2} \cdot |F(0)|^2. \quad (2.24)$$

The scattered outgoing flux in the solid angle  $d\Omega$  is the product of the scattering cross section and the incident flux. For elastic scattering incident and outgoing flux are the same and therefore:

$$|F(\Theta)|^2 = \left( \frac{d\sigma}{d\Omega} \right)_{el}^{\Theta}. \quad (2.25)$$

For forward scattering,  $\Theta = 0$ , and by introducing the Mandelstam variable  $t = -2k^2(1 - \cos \Theta)$  and using  $d\Omega = 2\pi d \cos \Theta = \frac{2\pi}{2k^2} dt$ , Eq. 2.25 can be written as:

$$|F(0)|^2 = \left( \frac{d\sigma}{d\Omega} \right)_{el}^{\Theta=0} = \frac{1}{\mathcal{L}} \frac{2k^2}{2\pi} \left( \frac{dR_{el}}{dt} \right)_{t=0}. \quad (2.26)$$

Introducing  $\rho = \frac{\text{Re}F(0)}{\text{Im}F(0)}$  as the ratio of the real to imaginary part of the forward scattering amplitude, Eq. 2.24 can be written as:

$$\sigma_{total}^2 = \frac{16\pi}{1 + \rho^2} \frac{1}{\mathcal{L}} \left( \frac{dR_{el}}{dt} \right)_{t=0}, \quad (2.27)$$

and by dividing Eq. 2.27 by Eq. 2.22 the total cross section can be expressed as a function of measurable quantities:

$$\sigma_{total} = \frac{16\pi}{1 + \rho^2} \frac{(\frac{dR_{el}}{dt})_{t=0}}{R_{el} + R_{inel}}, \quad (2.28)$$

without knowing the instantaneous luminosity. Measurements of  $\rho$  are shown in Fig. 2.30, measured total cross sections are shown in Fig. 2.31 as a function of the center-of-mass energy. Also shown is the elastic cross section in the latter figure, which can be determined experimentally. Both cross sections can be inserted in Eq. 2.21 to get the inelastic  $p\bar{p}$  cross section, which is measured in Run I at  $\sqrt{s} = 1.8$  TeV by the E710, the E811 and the CDF collaboration. The average of the three measurements is scaled up to  $\sqrt{s} = 1.96$  TeV [84] and is found to be [85]:

$$\sigma_{inel} = 60.7 \pm 2.4 \text{ mb}. \quad (2.29)$$

This value is then used via Eq. 2.20 to measure the instantaneous luminosity.

For many precision measurements the uncertainty on the luminosity measurement (currently an uncertainty of 6.5% [85] is assigned) will be the limiting factor at the end of Run II. The aim is to use simultaneous  $W$  boson and  $Z$  boson event counts [86] instead which can be done with higher accuracy. The theoretical cross sections are very well known [68] and have been measured [87, 88], however, the cross sections are orders of magnitude smaller than the inelastic  $p\bar{p}$  cross section and thus the rates are lower, leading to larger statistical errors.

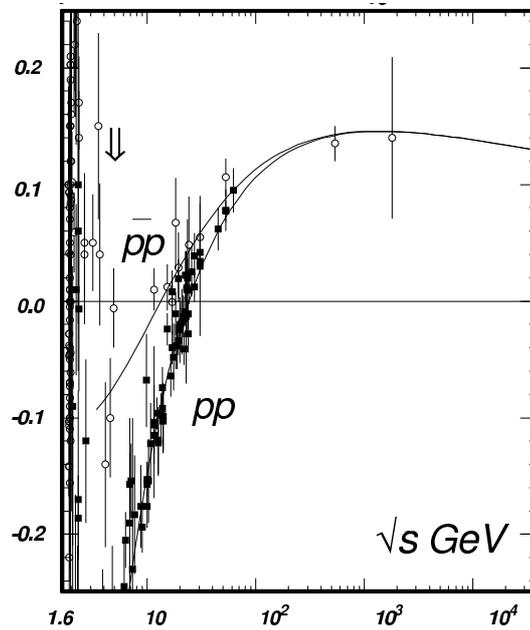


Figure 2.30: Ratio of the real to imaginary parts of the forward hadronic scattering amplitudes,  $\rho$ , as a function of  $\sqrt{s}$  [39]. The curves represent fits to the data points; the vertical arrow indicates the lower end of the fit range.

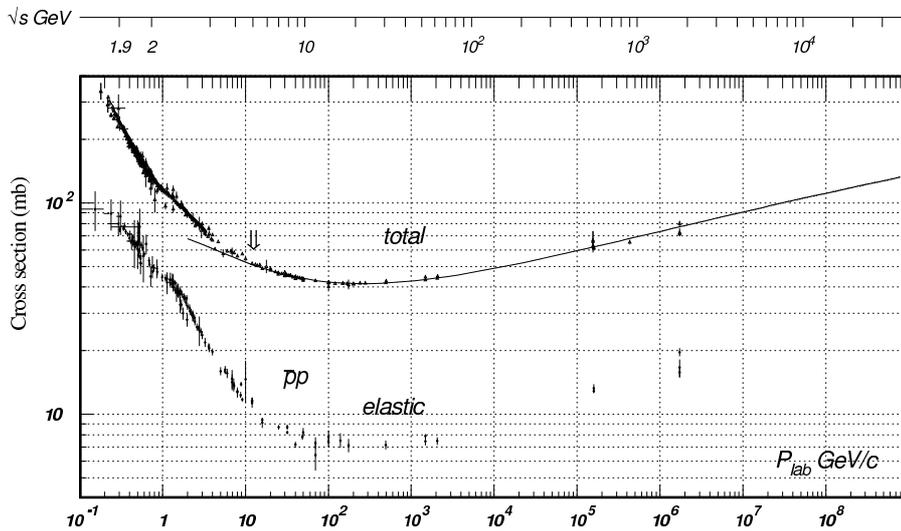


Figure 2.31: Total and elastic cross sections for  $p\bar{p}$  collisions as a function of laboratory beam momentum and total center-of-mass energy [39]. The curve represents a fit to the data points; the vertical arrow indicates the lower end of the fit range.



## Chapter 3

# Experimental Environment

The data analyzed in this document were collected between 2002 and 2004 from collisions at the Tevatron of protons and antiprotons at a center-of-mass energy of 1.96 TeV and recorded by the DØ detector. At present, the FERMILAB Tevatron is the world's highest-energy collider. During the data-taking period from 1992 to 1996 (Run I), the Tevatron experiments CDF and DØ each collected about  $125 \text{ pb}^{-1}$  of  $p\bar{p}$  collision data at a center-of-mass energy of 1.8 TeV, leading to the discovery of the top quark, and measurement of its mass, a precision measurement of the mass of the  $W$  boson, detailed analyses of gauge boson couplings, studies of jet production and vastly improved limits on the production of new phenomena, such as leptoquarks and supersymmetric particles, among many other accomplishments.

The new data-taking period (Run II) has started in March 2001 and is expected to deliver between  $4 \text{ fb}^{-1}$  and  $9 \text{ fb}^{-1}$  by the year 2009. The accelerator and the detector are described in turn.

### 3.1 The Fermilab Accelerator Complex

Fermilab maintains a series of eight accelerators of increasing energies, culminating in the Tevatron, which collides protons and antiprotons at a center-of-mass energy of 1.96 TeV, see Fig. 3.1. The design and operation of this accelerator chain is described in detail in [89, 90], a brief description is given here.

The protons used in the collisions are extracted from hydrogen ions. The ions are accelerated to 750 keV by a Cockroft-Walton accelerator and injected into a linear accelerator which boosts their energy to 400 MeV. These ions are stripped off their electrons as they pass through a sheet of graphite and are injected into the Booster, a synchrotron which brings their energy to 8 GeV.

Protons from the Booster are sent to the Main Injector, where they are

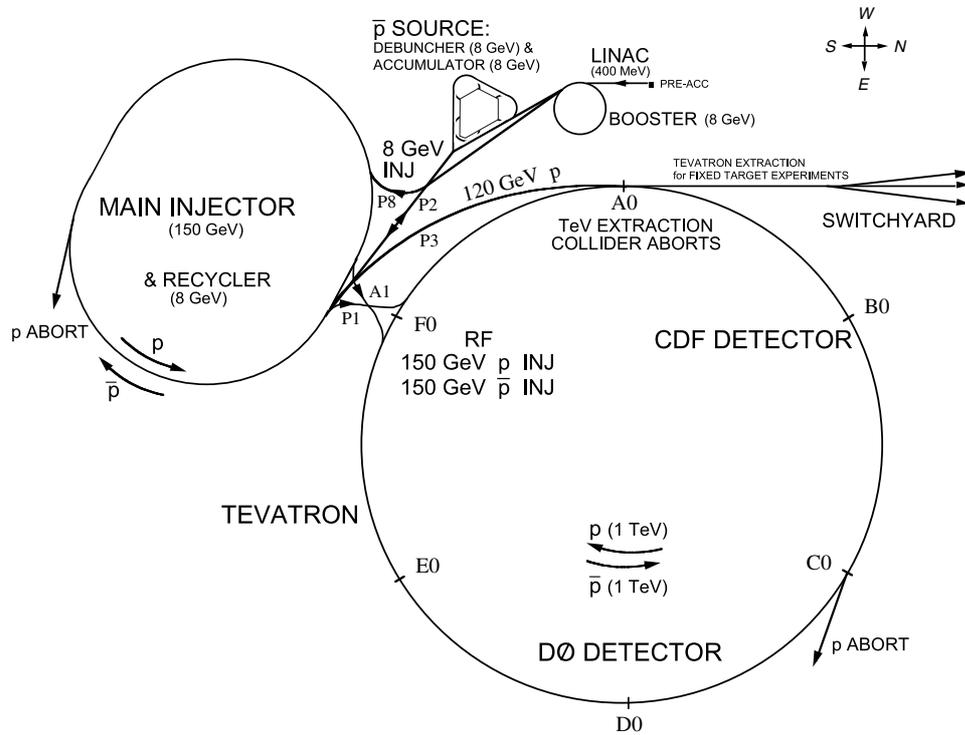


Figure 3.1: The Fermilab accelerator complex.

further accelerated to 150 GeV. Antiprotons used in the collisions are collected from the interaction products of a portion of the 150 GeV proton beam incident on a Nickel-Copper target. Antiprotons are cooled and debunched in the Debuncher and Accumulator, and accelerated to 150 GeV by the Main Injector.

Protons and antiprotons are injected into the Tevatron, where they are accelerated to their final energy of 980 GeV before colliding at the center of the DØ detector.

Collisions occur in bunches, with 36 bunches each separated by 396 nsec.

Figure 3.2 shows the integrated luminosity per week and total integrated luminosity for Run II from May 2001 until August 2004. The initial luminosity for each store is shown in Figure 3.3.

### 3.2 Interactions of Energetic Particles with Matter

The DØ detector surrounds the collision point and records the kinematics of the collision by examining its long lived products. The most prevalent, and those relevant to this thesis, are electrons, photons, muons, hadronic

### 3.2. INTERACTIONS OF ENERGETIC PARTICLES WITH MATTER<sup>51</sup>

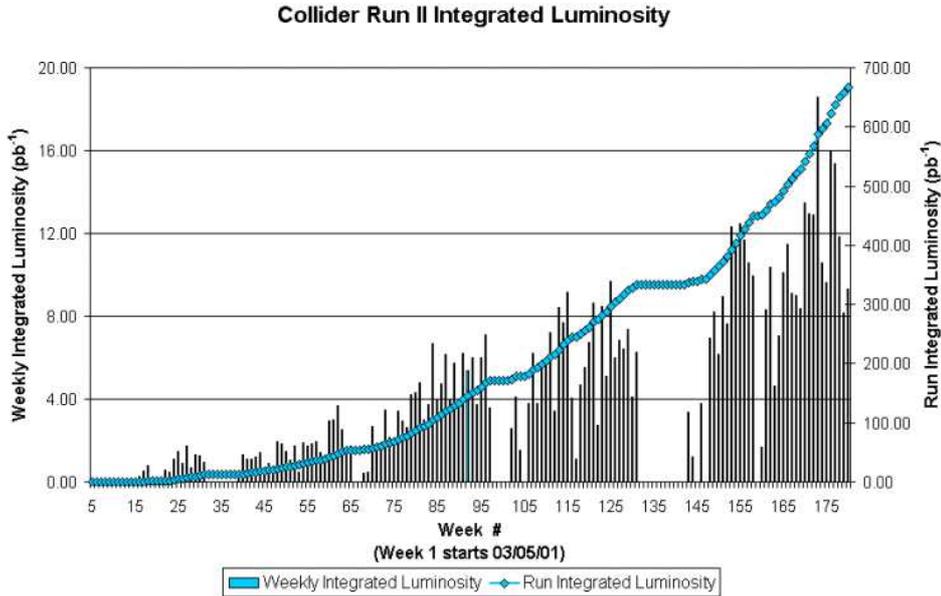


Figure 3.2: The integrated luminosity per week and total integrated luminosity for Run II from May 2001 until August 2004 in  $\text{pb}^{-1}$ .

particles and neutrinos.

The interaction of these particles with detector subsystems results in energy loss which can be detected and measured. Tracking detectors are designed to measure the particle positions with minimal energy loss. Calorimeters are constructed to fully absorb the particles and their showers and thus their energy in the process of measurement.

The modes of interactions of the relevant particles with the detector are discussed in turn.

#### 3.2.1 Electrons and Photons

Electrons passing through matter lose energy primarily through ionization and through bremsstrahlung. Above a critical energy [39],

$$E_c = \frac{800}{Z + 1.2} \text{ MeV}, \quad (3.1)$$

where  $Z$  is the atomic number of the medium, bremsstrahlung is the dominant process. The emitted photons produce electron-positron pairs, which in turn emit photons. The resulting shower of electrons and photons grows until the energy of the electrons falls below the critical energy, where they interact primarily through ionization. The mean distance over which an

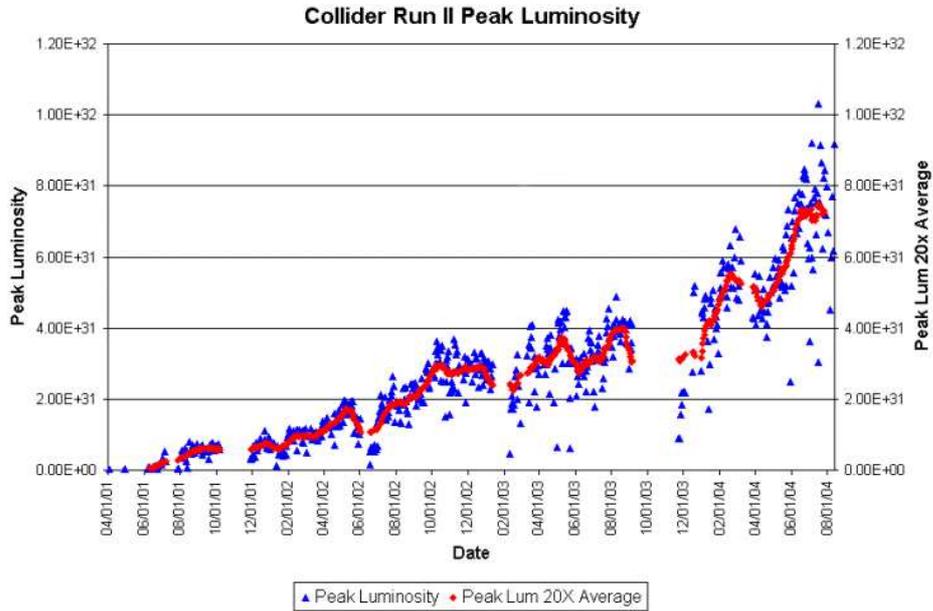


Figure 3.3: The initial luminosity for each store for the Tevatron Run II from May 2001 until August 2004.

electron loses all but  $1/e$  of its energy is called the radiation length  $X_0$  [39],

$$X_0 = \frac{716.4A}{Z(Z+1)\ln(287/\sqrt{Z})} \text{ g cm}^{-2}, \quad (3.2)$$

where  $A$  is the atomic mass of the medium in  $\text{g mol}^{-1}$ .

Photons interacting with matter produce electron-positron pairs, and hence an electromagnetic shower.

### 3.2.2 Muons

Muons interact through bremsstrahlung at a much lower rate than electrons due to their larger mass. Their energy loss is primarily through ionization. Figure 3.4 shows the energy loss per unit of material for muons in various energy regimes. Muons at the Tevatron have energies of the order of GeV and hence are minimum ionizing particles, also called MIP. They deposit only minimal energy in the central detector and leave it essentially unperturbed, in contrast to all other particles (with the exception of neutrinos, see Sect. 3.2.4).

### 3.2.3 Hadronic Particles

Hadronic particles interact inelastically with the nuclei of the detector element, producing primarily pions and nucleons. At high energies, the re-

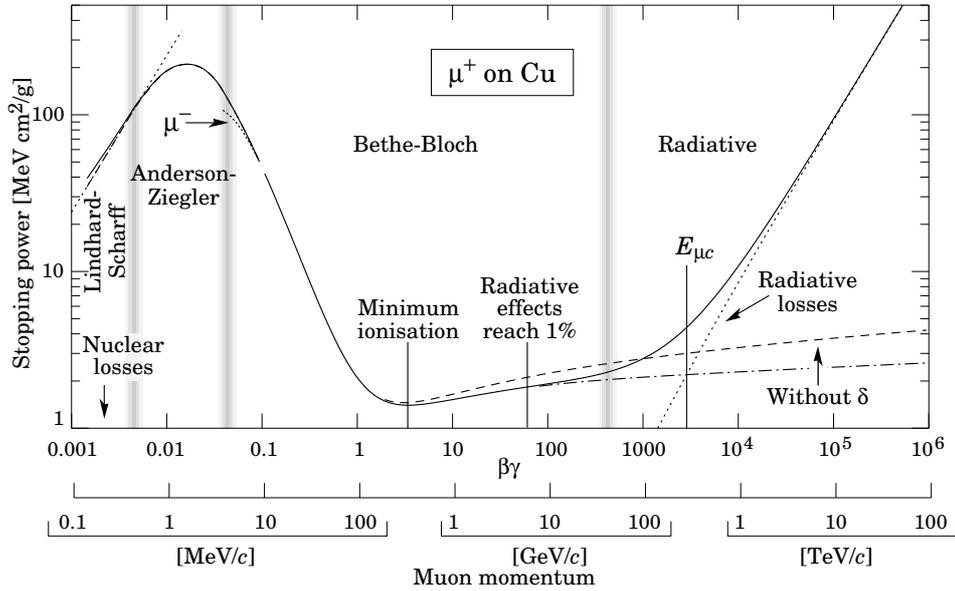


Figure 3.4: Energy loss through ionization of muons in various energy regimes [39].

sulting particles interact similarly with nearby nuclei, producing a shower of hadronic particles. The characteristic length scale is the nuclear interaction length, which is dependent on the material density and atomic mass and is roughly independent of energy:

$$\lambda_I \approx 35 A^{1/3} \text{ g cm}^{-2}. \quad (3.3)$$

A significant fraction of energy of the initial hadron escapes the hadronic cascade in form of neutral pions, which produce a secondary cascade. A smaller fraction results in invisible energy loss through unbinding of nuclei by spallation, non-ionizing collisions and uncaptured energy of neutrinos.

### 3.2.4 Neutrinos

As uncharged leptons, neutrinos interact only weakly via  $W$  and  $Z$  boson exchange, making their energy loss negligible and their direct detection practically impossible at DØ. Their presence can be inferred, however, from transverse momentum conservation requirements.

## 3.3 The DØ Detector

The DØ detector [91, 92, 93] is a large multi-purpose detector. It has been designed to identify and to precisely measure the four-momenta of the particles discussed in Sect. 3.2. To serve this purpose the DØ detector consists of

three major subsystems. At the core of the detector, a magnetized tracking system records precisely the angles of charged particles and measures their transverse momenta. A hermetic, finely grained Uranium and Liquid Argon calorimeter measures the energy of electromagnetic and hadronic showers, and a muon spectrometer detects and measures the momenta of escaping muons. Figure 3.5 gives an overview of the detector.

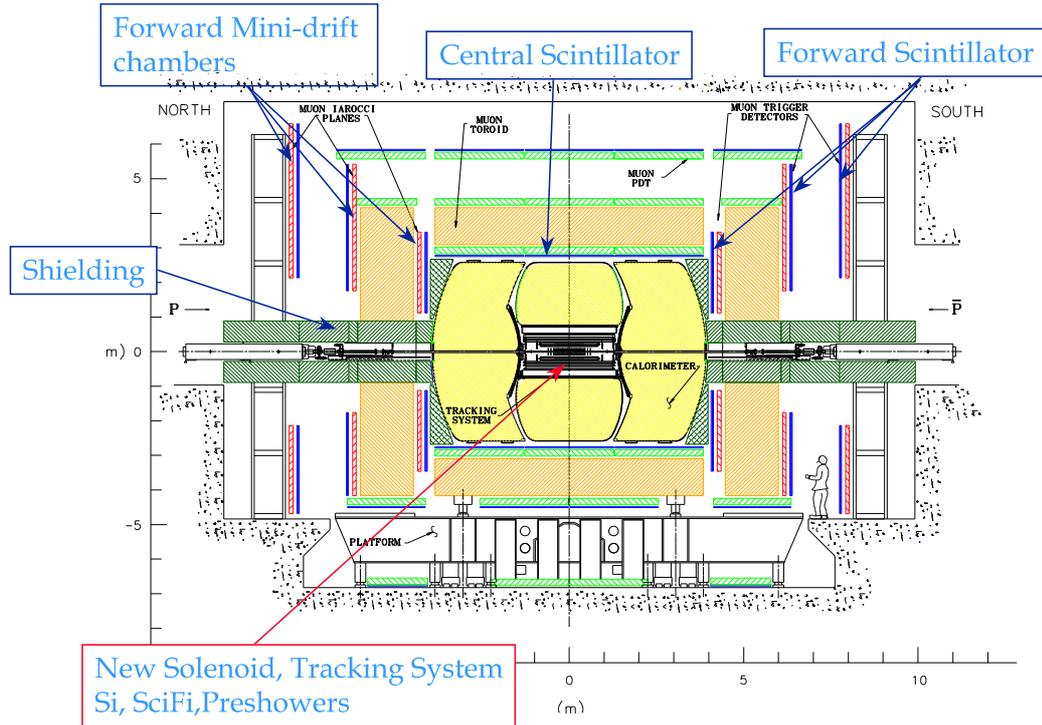


Figure 3.5: Side view of the DØ detector.

### 3.3.1 Coordinate System

The DØ coordinate system is right-handed with the  $z$  axis aligned with the direction of the beam such that the protons flow in the positive  $z$  direction. The  $y$  axis is then vertical, and the  $x$  axis points towards the center of the accelerator ring.

Another useful set of coordinates are the standard polar coordinates ( $r$ ,  $\phi$ ). The coordinate  $r$  denotes the perpendicular distance from the  $z$  axis,

$$r = \sqrt{x^2 + y^2}, \quad (3.4)$$

and  $\phi$  is the azimuthal angle

$$\phi = \arctan \frac{y}{x}. \quad (3.5)$$

In addition a particular reformulation of the polar angle  $\theta = \arccos \frac{z}{\sqrt{x^2+y^2+z^2}}$  is utilized, the pseudo-rapidity  $\eta$ :

$$\eta = -\ln \tan \frac{\theta}{2}. \quad (3.6)$$

The pseudo-rapidity, resulting from the rapidity  $y = \frac{1}{2} \ln \left( \frac{E+p_z}{E-p_z} \right)$  when particle masses are neglected, is a convenient choice at a hadron collider as the multiplicity of high energy particles is roughly constant in  $\eta$ . Additionally, rapidity intervals are Lorentz-invariant under boosts along the  $z$  axis.

Depending on the choice of the origin of the coordinate system, the coordinates are referred to as physics coordinates, when the origin is the reconstructed vertex of the interaction ( $\phi$  and  $\eta$ ), and they are referred to as detector coordinates ( $\phi_{det}$  and  $\eta_{det}$ ), when the origin is chosen to be the center of the DØ detector.

### 3.3.2 Luminosity System

The primary purpose of the Luminosity Monitor (LM) [94] is to make an accurate determination of the Tevatron collider luminosity at the DØ interaction region [85].

The LM detector consists of two arrays of twenty-four plastic scintillation counters with photomultiplier readout. A schematic drawing of an array is shown in Fig. 3.6. The arrays are located in front of the end calorimeters

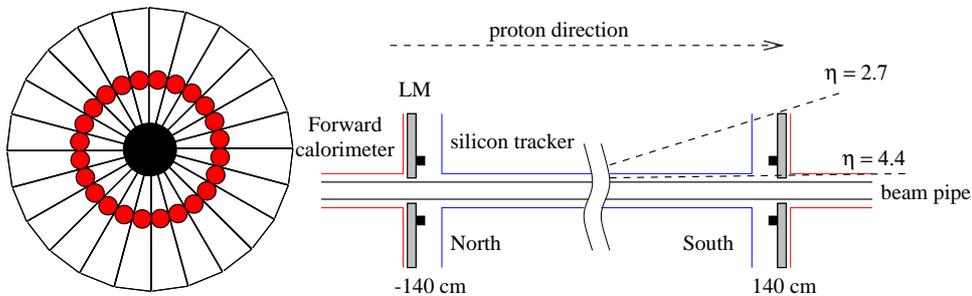


Figure 3.6: Luminosity Monitor layout. The  $r - \phi$  view is shown on the left, the  $r - z$  view of the two arrays is shown on the right.

at  $z = \pm 140$  cm, and occupy the region between the beam pipe and the Forward Preshower Detector (see Sect. 3.3.4.1). The counters are 15 cm long and cover the pseudo-rapidity range  $2.7 < |\eta_{det}| < 4.4$ .

The luminosity measurement is accomplished by detecting inelastic  $p\bar{p}$  collisions in the dedicated LM detector through

$$\mathcal{L} = \frac{R}{\varepsilon A \sigma_{inel}}. \quad (3.7)$$

The determination of  $\sigma_{inel}$  is discussed in Sect. 2.5 and given by Eq. 2.29. The cross section has to be corrected for the efficiency  $\varepsilon$  and the acceptance  $A$  of the LM detector for inelastic  $p\bar{p}$  collisions.

Multiple  $p\bar{p}$  collisions can occur in a single beam crossing. The number of interactions per bunch crossing is given by Poisson statistics. Collision products arrive at each set of scintillators roughly in coincidence, while beam halo products passing through the detector appear distinctly separated. Time-of-flight information from the two luminosity arrays and the  $z$  vertex distribution (see Fig. 3.7) is utilized to separate these processes. The rate  $R$  is corrected for these two effects.

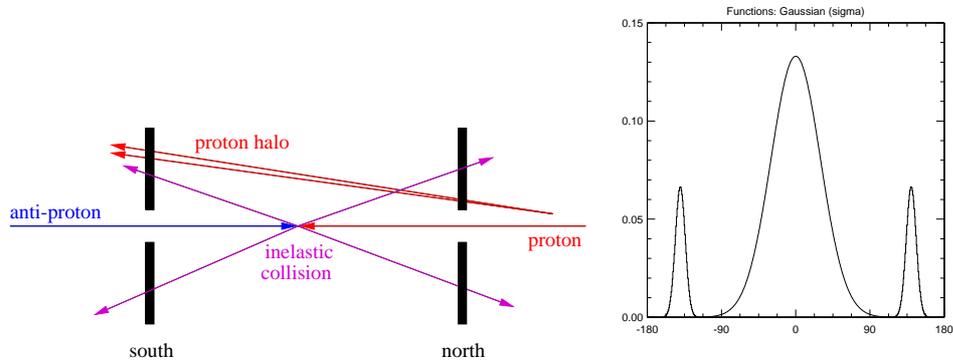


Figure 3.7: The sketch on the left shows the differentiation between inelastic collisions and beam halo. Expected  $z$  vertex distribution for inelastic collisions, centered at  $z = 0$  cm,  $p$  halo centered at  $z = -140$  cm and  $\bar{p}$  halo centered at  $z = 140$  cm (right).

### 3.3.3 The Central Tracking System

The central tracking system is surrounded by a solenoid which provides a nearly uniform magnetic field of  $B = 2$  T parallel to the beam axis. Charged particles produced in the collision are bent around the field lines. The radius  $r$  of the curvature allows for a measurement of the transverse momentum through:

$$p_T[\text{GeV}] = 0.3 \cdot r[\text{m}] \cdot B[\text{T}]. \quad (3.8)$$

Closest to the beam pipe itself is the Silicon Microstrip Tracker (SMT), which allows for the precision measurements, crucial for accurate measurement of impact parameter and identification of secondary vertices. Surrounding the SMT is the Central Fiber Tracker (CFT), comprised of 16 layers of scintillating fiber. The CFT extends to a radius of 50 cm, giving a lever arm long enough to provide effective transverse momentum resolution. The  $D\bar{O}$  central tracking system is depicted in Fig. 3.8.

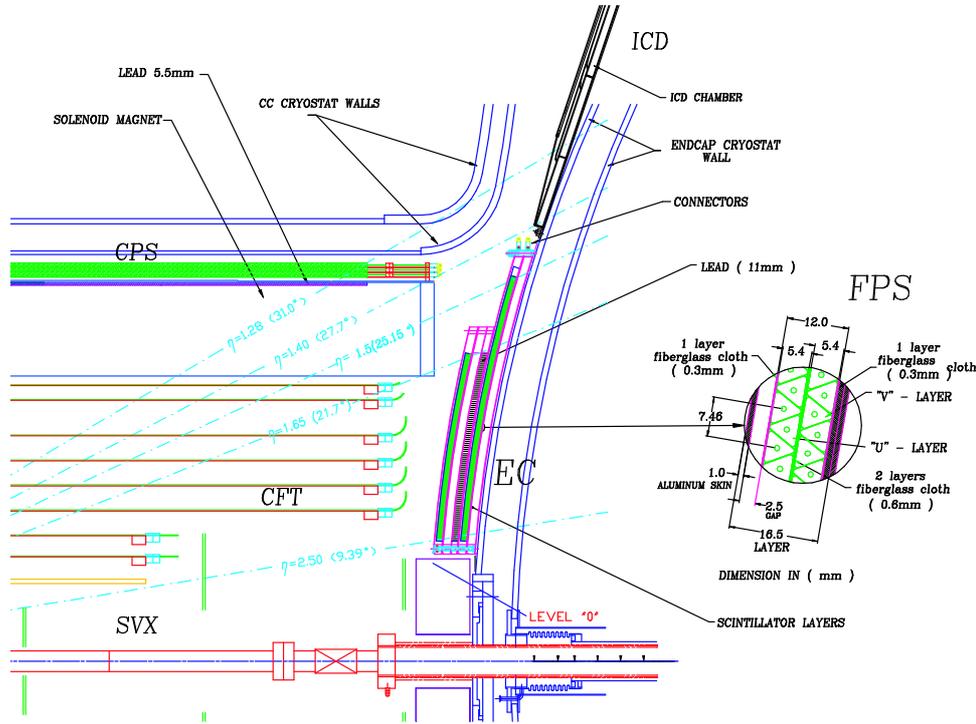


Figure 3.8: Cross-sectional view of the  $D\emptyset$  tracking and preshower system.

### 3.3.3.1 The Silicon Microstrip Tracker

Charged particles passing through the  $300\ \mu\text{m}$  thick wafers of  $n$ -type silicon which comprises the SMT [95] produce pairs of electrons and holes. The ionized charge is collected by strips of  $p$ -type or  $n^+$ -type silicon strips, whose minute constructions (between  $\sim 50\ \mu\text{m}$  and  $\sim 150\ \mu\text{m}$  pitch) provide for the measurement of the ionization.

The length of the interaction region ( $\sigma \approx 25\ \text{cm}$ ) sets the length scale of the device in  $z$ , up to  $\pm 3\ \text{cm}$  for the barrel. With a long interaction region, it is difficult to deploy detectors such that the tracks are generally perpendicular to detector surfaces for all  $\eta$ . This led to the design of barrel modules interspersed with disks in the center and assemblies of disks in the forward and backward regions. The barrel detectors measure primarily the  $r - \phi$  coordinate and the disk detectors measure  $r - z$  as well as  $r - \phi$ . Thus vertices for high  $\eta$  particles are reconstructed in three dimensions by the disks, and vertices of particles at small values of  $\eta$  are measured in the barrel.

An isometric view of the SMT is shown in Fig. 3.9. The detector has six barrels in the central region. Each has four silicon readout layers, each layer having two staggered and overlapping sub-layers, see Fig. 3.10. The outer barrels have single sided and double sided  $2^\circ$  stereo ladders. The four

inner barrels have double sided  $90^\circ$  stereo and double sided  $2^\circ$  stereo ladders. Each barrel is capped at high  $|z|$  with a disk of twelve double sided wedge detectors, called an “F-disk”. In the far forward and backward regions, a unit consisting of three F-disks and two large-diameter “H-disks” provides tracking at high  $|\eta_{det}| < 3.0$ . The F-disks are made of twelve wedges of double sided stereo detectors. The H-disks are made of 24 pairs of single sided detectors glued back to back. Table 3.1 summarizes the SMT design numbers.

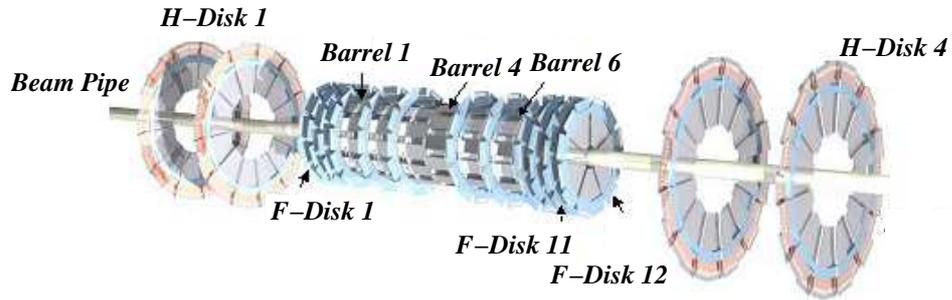


Figure 3.9: Isometric view of the  $D\bar{O}$  silicon vertex detector.

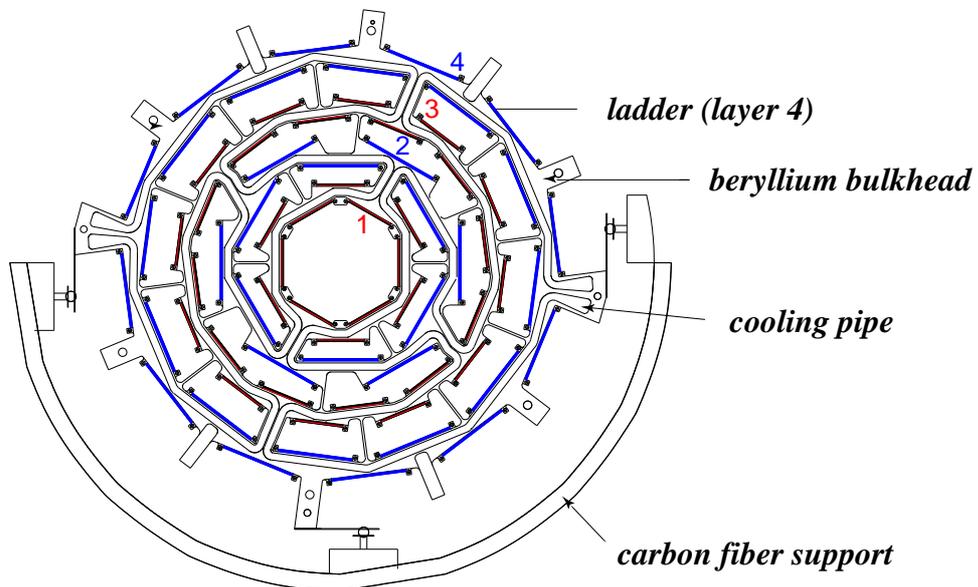


Figure 3.10: Cross section of the silicon vertex detector barrel structure.

The axial hit resolution is on the order of  $10 \mu\text{m}$ , the  $z$  hit resolution is  $35 \mu\text{m}$  for  $90^\circ$  stereo and  $450 \mu\text{m}$  for  $2^\circ$  stereo ladders.

	Barrels	F-Disks	H-Disks
#Channels	387 072	258 048	147 456
Sensors	s/d sided	double sided	single sided
Stereo	$0^\circ, 2^\circ, 90^\circ$	$\pm 15^\circ$	$\pm 7.5^\circ$
#Modules	432	144	96 pairs
Si area	$1.3 \text{ m}^2$	$0.4 \text{ m}^2$	$1.3 \text{ m}^2$
Inner radius	2.7 cm	2.6 cm	9.5 cm
Outer radius	9.4 cm	10.5 cm	26 cm
Maximal $ z $	38.4 cm	54.8 cm	120 cm

Table 3.1: Specifications of the Silicon Vertex Detector [96].

### 3.3.3.2 Central Fiber Tracker

The Central Fiber Tracker [92, 93] consists of 835  $\mu\text{m}$  scintillating fibers mounted on eight concentric support cylinders and occupies the radial space from 20 to 52 cm from the center of the beam pipe. The two innermost cylinders are 1.66 m long, the outer six cylinders are 2.52 m long. Each cylinder supports one doublet layer of fibers oriented along the beam direction and a second doublet layer at a stereo angle of alternating  $+3^\circ$  and  $-3^\circ$ . The two layers of fibers are offset by half a fiber width to provide improved coverage. The small fiber diameter gives the CFT a cluster resolution of about 100  $\mu\text{m}$  per doublet layer.

Light production in the fibers is a multistep process. When a charged particle traverses one of the fibers, the scintillator emits light at  $\lambda = 340 \text{ nm}$  through a rapid fluorescence decay. A wave-shifting dye absorbs the light well at  $\lambda = 340 \text{ nm}$  and emits at  $\lambda = 530 \text{ nm}$ . The light is then transmitted by total internal reflexion to the end of the scintillating fibers, where the light is transferred through an optical connection to clear fiber waveguides of identical diameter which are 7.8 to 11.9 m long. The light is only observed from one end of each scintillating fiber. The opposite end of each of the scintillating fibers is mirrored by sputtering with an aluminum coating that provides a reflectivity of 85 to 90 %. The clear fiber waveguides carry the scintillation light to visible light photon counters (VLPCs) which convert it into an electronic pulse which is readout.

The visible light photon counters are situated in a liquid Helium cryostat and operate at a temperature of 9 K. They detect photons with a quantum efficiency of 85 % and provide charge of about 30 to 60 k electrons per photon. A minimum ionizing particle creates an average of eight photoelectrons per layer, depending on the angle between the scintillating fiber and particle trajectory.

### 3.3.3.3 Solenoid Magnet

The superconducting solenoid magnet [98] is designed to optimize the momentum resolution,  $\Delta p_T/p_T$ , and tracking recognition. It is 2.73 m in length and 1.42 m in diameter, corresponding to 0.9 radiation lengths, and provides a 2 T magnetic field with an integrated field homogeneity of 0.5 %. It operates at 10 K, the current is 4820 A and the stored energy is 5.6 MJ.

### 3.3.3.4 Tracking Performance

Hits from both tracking detectors are combined to reconstruct tracks. The momentum resolution of the tracker for minimal ionizing particles can be parameterized as:

$$\sigma(p^{-1}) = \frac{\sqrt{(S \cdot \sqrt{\cosh \eta})^2 + (C \cdot p_T)^2}}{p}, \quad (3.9)$$

where  $p$  is the particle momentum and  $\eta$  is the pseudo-rapidity.  $S$  accounts for the multiple scattering term and  $C$  represents the resolution term. A study [99] of  $Z \rightarrow \mu^+ \mu^-$  events has found  $S = 0.015$  and  $C = 0.0018$ .

To calibrate the tracker, meson and baryon resonances are used. The process  $K_s^0 \rightarrow \pi^+ \pi^-$  is measured with a width of  $\sigma = 7.3$  MeV and  $\Lambda^0 \rightarrow p^+ \pi^-$  with  $\sigma = 2.6$  MeV (Fig. 3.11). The reconstruction of secondary vertices is crucial to identify  $b$  hadrons. The impact parameter resolution is shown as a function of transverse momentum in Fig. 3.12. As an example, the reconstruction of processes  $\Xi^\pm \rightarrow \Lambda^0 \pi^\pm$  and  $\Omega^\pm \rightarrow \Lambda^0 K^\pm$  is shown in Fig. 3.13, where multiple tracks with impact parameter of the order of centimeters are reconstructed.

## 3.3.4 The Calorimeter System

The function of the calorimeter system is to measure the energy of particles by inducing them to produce electromagnetic and hadronic showers. Inert passive layers of dense material in which the shower begins are followed by active layers, where the surviving fraction of the shower energy is sampled through ionization.

### 3.3.4.1 The Preshower Detectors

The preshower detectors (see Fig. 3.8) aid in electron identification and resolution, and background rejection. They function as calorimeters as well as tracking detectors, enhancing the spatial matching between tracks and calorimeter showers [101].

**Central Preshower detector** The Central Preshower detector (CPS) consists of three concentric cylindrical layers of triangular scintillator

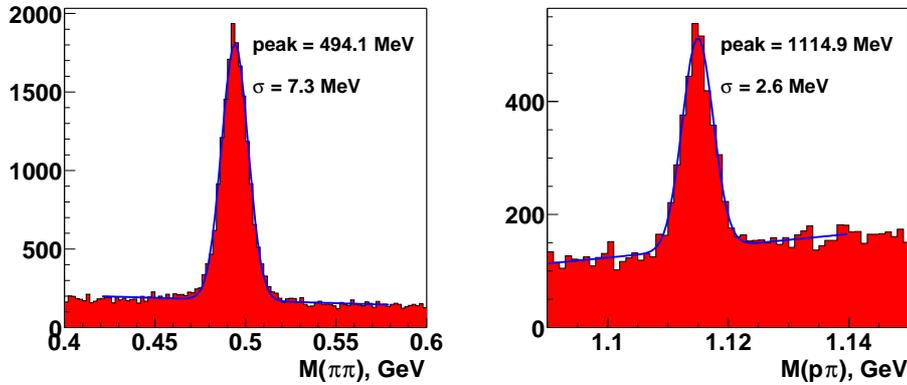


Figure 3.11: Invariant mass distributions for the processes  $K_s^0 \rightarrow \pi^+\pi^-$  (left) and  $\Lambda^0 \rightarrow p^+\pi^-$  (right) [100]. Also shown a Gauss distribution plus a straight line fit to the signal and background. The mean and the width of the Gauss fit are also shown.

strips (axial and stereo  $\pm 23^\circ$ ) and is located in the nominal 5 cm gap between the solenoid and the central calorimeter, covering the region  $|\eta_{det}| < 1.3$ . Between the solenoid and the CPS is a Pb radiator approximately one radiation length thick in order to discriminate photons and electrons from pions by converting them into showers.

**Forward Preshower detector** The Forward Preshower detectors (FPS) are each mounted on the spherical heads of the end calorimeter cryostats (see Sect. 3.3.4.2), occupying the region between the luminosity counters (see Sect. 3.3.2) at the inner edge and the intercryostat detectors (see Sect. 3.3.4.3) at the outer edge, covering the region  $1.5 < |\eta_{det}| < 2.5$ . Each detector is made from two layers, at different  $z$ , of a double layer of scintillator strips (stereo  $\pm 22.5^\circ$ ), separated by a  $2X_0$ -thick lead-stainless-steel absorber. All charged particles passing through the detector register a three-dimensional hit in the first, so-called MIP layer. Photons do not generally interact in the MIP layer. Electrons and photons shower in the absorber, leading to a cluster of energy in the second, so-called shower layer. These signatures can be used to distinguish muons, electrons and photons.

Scintillating light produced from the passage of charged particles is collected by wavelength-shifting fibers and piped through clear fibers to the visible photon counters in the same way as explained for the central fiber detector in Section 3.3.3.2.

The pre-shower detectors are discussed in more detail in [102] and [103].

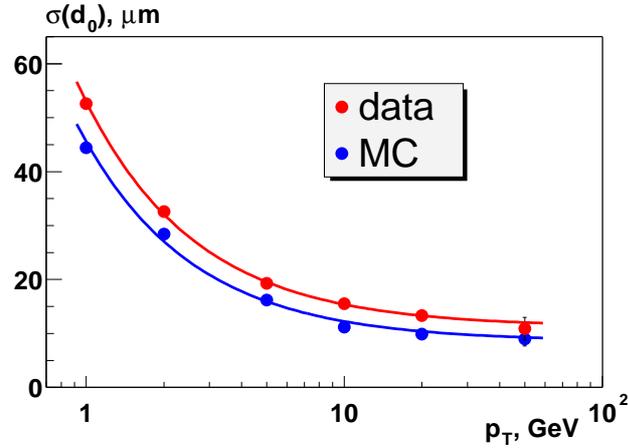


Figure 3.12: Impact parameter resolution measured in data and in simulated single muon events. The lines show fits to data and MC [100].

### 3.3.4.2 The Liquid Argon Calorimeter

The Liquid Argon calorimeter [92, 93, 104] is designed to provide the energy measurement for and assist in the identification of electrons, photons, taus and jets and establish the transverse energy balance in an event. The device is also sensitive to MIPs and therefore can serve to identify muons. The calorimeter itself (i.e. the modules) is unchanged from Run I and depicted in Fig. 3.14. However, there is significantly more material in front of the calorimeter ( $2 \lesssim X_0 \lesssim 4$ , depending on the  $\eta$ ) and the electronics is rebuilt.

As shown in Fig. 3.15, the Liquid Argon calorimeter is subdivided into the central calorimeter (CC) covering roughly  $|\eta_{det}| < 1$  and two end calorimeters (EC) extending the coverage to  $|\eta_{det}| \approx 4$ . Each calorimeter contains an electromagnetic section closest to the interaction region followed by fine and coarse hadronic sections whose module size increases with the distance from the interaction region. The active medium for all of the calorimeters is liquid Argon and each of the three calorimeters is located within a cryostat that maintains the temperature at approximately 80 K. In order to achieve the same energy response for electromagnetic and hadronic particles,  $e/h \approx 1$  (compensating calorimeter), different absorber plates are used in different locations. The  $e/\pi$  ratio is measured for charged pions  $\pi$  in test beam data of Run I and the calorimeter is found to be nearly compensating, with  $1 < e/\pi < 1.05$  above 30 GeV [105]. The electromagnetic sections (EM) use thin (3 or 4 mm) plates, made from nearly pure depleted Uranium. The fine hadronic sections are made from 6-mm-thick Uranium-Niobium alloy. The coarse hadronic modules contain relatively thick (46.5 mm) plates of either Copper in the CC or stainless steel in the EC.

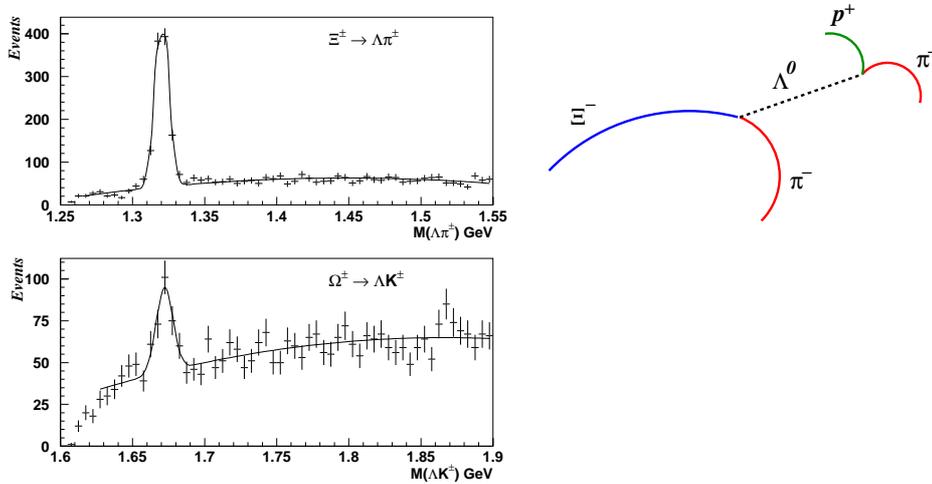


Figure 3.13: Invariant mass distributions for the processes  $\Xi^\pm \rightarrow \Lambda^0 \pi^\pm$  (top left) and  $\Omega^\pm \rightarrow \Lambda^0 K^\pm$  (bottom left). The reconstruction of the processes are examples of events with multiple tracks with impact parameter in the order of centimeters. A sketch of the decay  $\Xi^- \rightarrow \Lambda^0 \pi^- \rightarrow p^+ \pi^- \pi^-$  is also shown (right) [100].

The elements are combined in a basic unit, a calorimeter cell, depicted in Fig. 3.16, which contains the absorber plates, the active medium and a Copper readout pad laminated to G10 and covered in resistive epoxy for collecting the ionization. The electric field is established by grounding the absorber plates and holding the resistive surface of the pad at typically 1.6 kV. The electron drift time across the 2.3 mm gap is approximately 450 ns. Several such pads are ganged together in depth to form a readout cell (see Fig. 3.15).

The readout cells are arranged and sized such that each covers roughly an area of  $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ , comparable to the transverse sizes of showers: 1-2 cm for EM showers and about 10 cm for hadronic showers. Typical cone sizes of jets are  $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \approx 0.5$ . Segmentation finer than this is useful in probing the shapes of jets. Longitudinal subdivisions are useful since longitudinal shower profiles help distinguish electrons and hadrons. There are four separate depth layers for the EM modules in the CC and EC. The first two layers are approximately two radiation lengths ( $X_0$ ) thick and help measure the transverse shower development. The third layer is placed where the shower is expected to reach its maximum and the cells measure  $\Delta\eta \times \Delta\phi = 0.05 \times 0.05$  to provide improved spatial resolution. The electromagnetic section contains in total 65.6 mm of Uranium, which represents approximately 20 radiation lengths ( $X_0^{Ur} \approx 3.2$  mm) to capture the over-

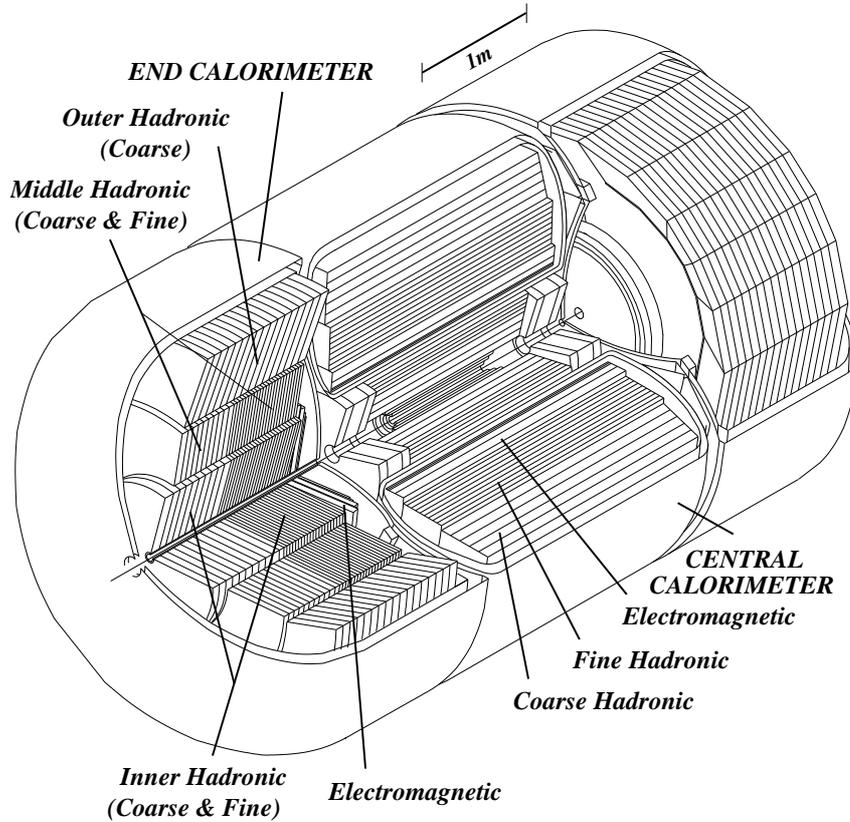


Figure 3.14: Isometric view of the central and two end calorimeters.

whelming fraction of the electromagnetic energy. As the nuclear interaction length is much larger than the radiation length, ( $\lambda_I^{Ur} \approx 10.5 \text{ cm} \approx 30X_0^{Ur}$ ), hadronic particles typically deposit most of their energy in the hadronic section of the calorimeter, which contains approximately  $6.4\lambda_I$  of Uranium and Copper.

### 3.3.4.3 The Inter-Cryostat Detectors

Since the calorimeter system is contained in three separate cryostats, it provides incomplete coverage in the pseudorapidity region  $0.8 < |\eta_{det}| < 1.4$ , as can be seen in Fig. 3.15. Additional layers of sampling are added in form of scintillating counters between the CC and EC cryostats (called Inter-Cryostat Detector or ICD [92, 93, 104]). The segmentation of  $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$  matches exactly the Liquid Argon calorimeter geometry. In addition, separate single-cell structures, called *massless gaps* [92, 93, 104] are installed, both in the central calorimeter and in the end cap calorimeters.



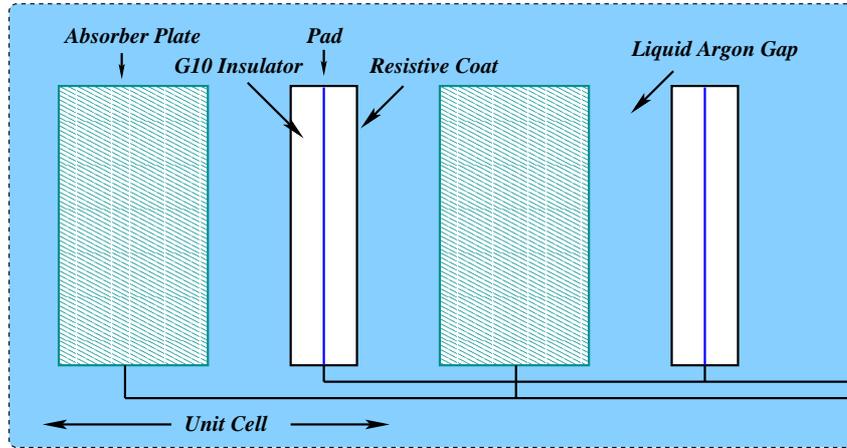


Figure 3.16: Schematic view of a calorimeter cell.

Particle	$C$	$S$	$N$
$e$	$0.0115^{+0.0027}_{-0.0036}$	$(0.135 \pm 0.005)\sqrt{\text{GeV}}$	0.43 GeV
$\pi$	$0.032 \pm 0.004$	$(0.45 \pm 0.04)\sqrt{\text{GeV}}$	0.975 GeV

Table 3.2: Calorimeter energy resolution parameters, measured with electrons for the electromagnetic section [107] ( $C$  from the  $Z \rightarrow ee$  mass resolution,  $S$  from the test beam and  $N$  from  $W \rightarrow e\nu$ ) and with pions from the test beam for the hadronic section [106]. Uncertainties on  $N$  were not determined.

calorimeter can be performed utilizing the three resonances  $J/\Psi$ ,  $\Upsilon$  and  $Z$  with a leptonic decay to two electrons. The central tracking system provides the momentum measurement from which the energy can be inferred. Figure 3.17 shows the di-electron invariant mass spectrum measured with the calorimeter (left) and the central tracker (right). Figure 3.18 shows the di-electron invariant mass spectrum, using both calorimeter and tracking information, in the region of the  $Z$  boson mass. The events can be selected with high statistics and with high purity and represent three independent measurements allowing to constrain the  $S$ ,  $C$  and  $N$  terms in Eq. 3.10.

The energy resolution of the hadronic calorimeter in Run II can be measured by selecting collisions randomly (so-called “zero-bias”). Most of the time no hard-scatter interaction will occur. The central tracking system can be used to identify single tracks representing isolated charged particles, which are dominantly charged pions. The tracking system provides a momentum measurement which can be used to determine the expected energy deposition of the particle in the calorimeter.

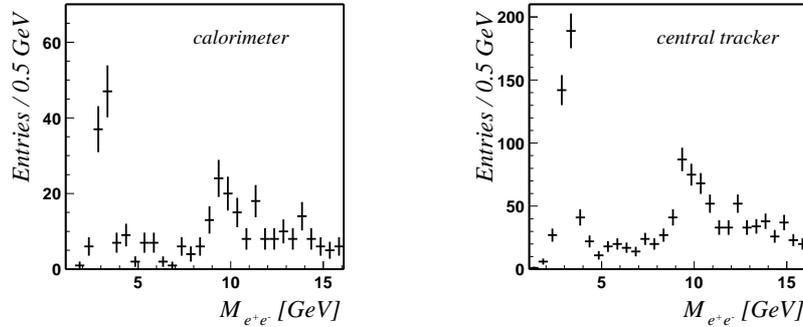


Figure 3.17: Di-electron invariant mass spectrum measured with the calorimeter (left) and the central tracker (right). The  $J/\Psi$  and  $\Upsilon$  can clearly be seen [108].

The preliminary studies [108] suggest a decreased calorimeter response in comparison to Run I, mainly due to the following reasons:

- The electron drift time of 450 ns provides a challenge for signal charge integration for a roughly 5 times larger beam crossing frequency occurring every 396 ns in Run II compared to  $2 \mu\text{s}$  in Run I. Only 70 % of the signal charge is integrated for the shorter time in Run II leading to a larger sensitivity to fluctuations. In particular the contribution from slow nuclear products from the hadronic shower is lost. Also the mechanical tolerances with which the calorimeter was built meet the requirements of the Run I integration time and are too large for Run II. All this involves a degradation of the  $S$  and  $N$  terms in Eq. 3.10.
- The new electronics used in Run II is found to have a worse noise performance, affecting mainly the  $N$  term in Eq. 3.10.
- The energy response of the cells differs by up to 10 % [108]. This cell to cell miscalibration affects the  $C$  term in Eq. 3.10.
- The new tracking system detector components, the solenoid magnet and the preshower detector installed for Run II in front of the calorimeter have a radiation lengths  $2 \lesssim X_0 \lesssim 4$ , depending on the  $\eta$  of the particle. As a consequence, the term  $S$  in Eq. 3.10 degrades.

### 3.3.5 The Muon System

As charged particles, which do not cause electromagnetic or hadronic showers, the muons originating from a  $p\bar{p}$  collision penetrate the tracking system

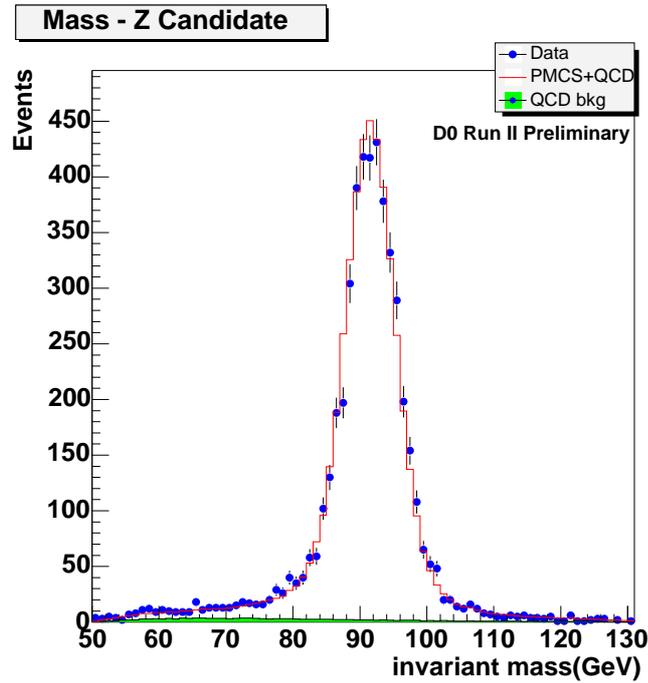


Figure 3.18: Invariant di-electron mass distribution of data (dots) and expected background (line) for the  $Z$  peak region. Shaded is QCD background [87].

and the calorimeter essentially unperturbed. The  $D\bar{O}$  muon detection system, placed around the calorimeter and depicted in Fig. 3.19, serves to identify and trigger on these muons and measure their momenta and their charge.

For this purpose it consists of a system of proportional drift tubes (PDTs), mini drift tubes (MDTs) and scintillation counters. The PDTs cover  $|\eta_{det}| < 1.0$ , the planes of MDTs extend the muon detection to  $|\eta_{det}| = 2.0$ . The scintillator counters are used for triggering and for cosmic and beam related muon rejection. Toroidal magnets and special shielding complete the muon system. Each sub-system has three layers called A, B, and C. The A layer is innermost and located between the calorimeter and the iron of the toroid magnet. B and C layers are located outside the iron. In the region directly below the calorimeter, only partial coverage by muon detectors is possible. The support structure for the  $D\bar{O}$  detector and readout electronics is located in this region.

The most probable value for the energy loss of a muon in the calorimeter is 1.6 GeV, and about 1.7 GeV in the iron. The momentum measurement is corrected for this energy loss.

In the following the subsystems of the muon spectrometer are discussed.

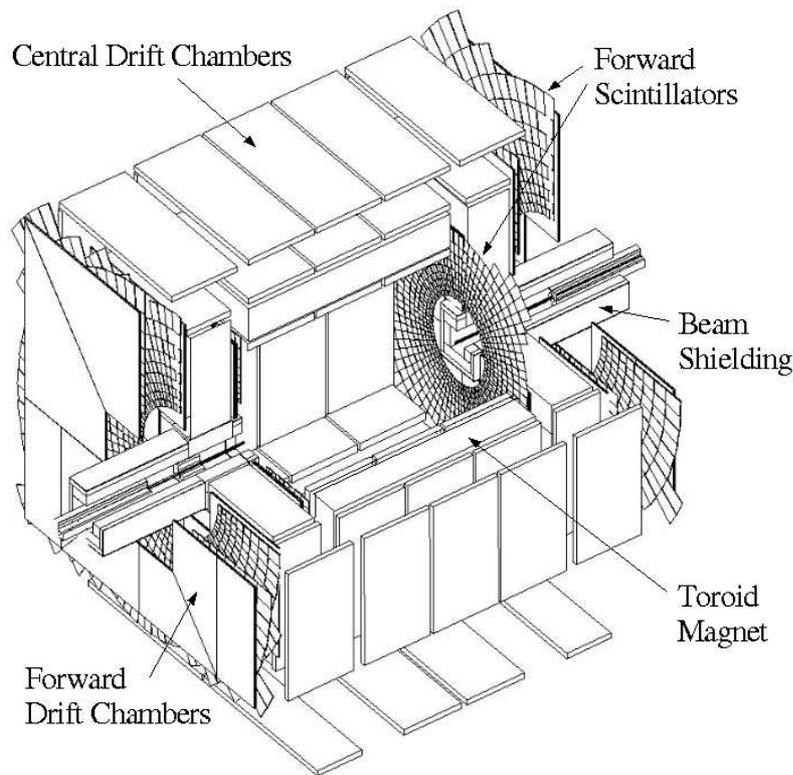


Figure 3.19: A cut-away view of the muon system.

### 3.3.5.1 Toroid Magnet

The 1973 ton toroid magnet [98], located between layer A and B, allows a measurement of the muon momentum. It is a square annulus 109 cm thick, in radial distance of 317.5 cm from the beam line. In order to give access to the inner portions of the  $D\bar{O}$  detector, the central toroid is split in 3 parts. A bottom section gives support for the calorimeter and tracking detectors. The central toroid is completed by two movable c-shaped shells. The two forward toroid magnets are located at  $447 \leq |z| \leq 600$  cm. The coils carry currents of 1500 A and result in an internal field of 1.8 T.

### 3.3.5.2 Drift Tubes

Drift tubes are rectangular volumes, filled with gas. The ionization created by a passing charged particle is collected and amplified by a 50  $\mu\text{m}$  gold-plated tungsten sense wire which runs through the center of the chamber. Vernier cathode pads are located above and below the wires to provide information on the hit position along the wire. Figure 3.20 shows the geometry of an example drift tube, and the arrangement of central drift tubes (PDTs)

in the A, B and C layers. Figure 3.21 shows an exploded view of the drift tubes.

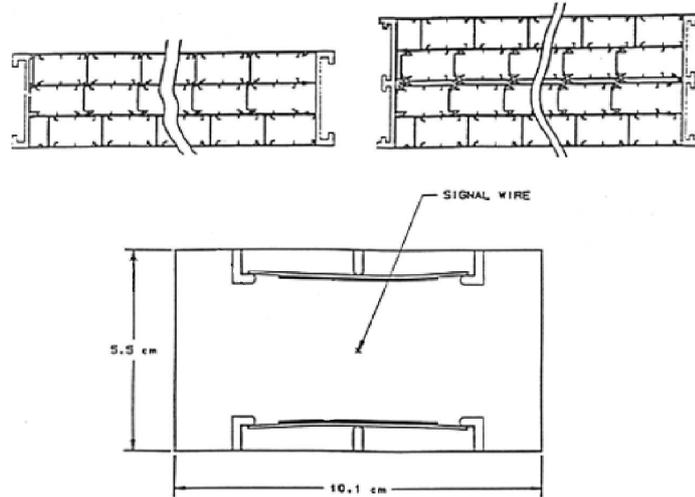


Figure 3.20: Drift tubes in the central muon system. An example tube (bottom) and stacking of tubes in the A (top left) and BC layers (top right).

A measurement of the arrival time of the pulse from the sense wire and a calibration of the drift time of the gas allows for calculation of the radial distance from the sense wire. A comparison of arrival times from adjacent wires provides a rough measurement of the position of the ionization along the wire.

**Proportional Drift Tubes** Central drift tubes are constructed of extruded aluminum coated with steel foil and filled with a mixture of 84 % Argon, 8 %  $\text{CH}_4$  and 8 %  $\text{CF}_4$ . The gas flow rate is 500 liters per hour. The PDTs are  $5.5 \times 10.0 \text{ cm}^2$  in cross section and 240 cm long. The operating high voltage is 2.3 kV for the pads and 4.7 kV for the wires. The drift velocity is approximately  $10 \text{ cm}/\mu\text{s}$ , for a maximum drift time of about 500 ns. The coordinate resolution of the radial distance to the sense wire is approximately 3 mm. The PDTs are discussed in more detail in [109].

**Mini Drift Tubes** Mini drift tubes have a shorter electron drift time (40-60 ns, depending on the inclination of the tracks), a better coordinate resolution ( $\approx 0.7 \text{ mm}$ ), they are radiation hard, have a high segmentation and thus a low occupancy. They are filled with a mixture of 90 %  $\text{CF}_4$  and 10 %  $\text{CH}_4$ . The MDTs are  $0.94 \times 0.94 \text{ cm}^2$  in cross section and have a maximum tube length of 583 cm. The anode wire is grounded and  $-3.2 \text{ kV}$  is applied to the cathode. The MDTs are discussed in more detail in [110].

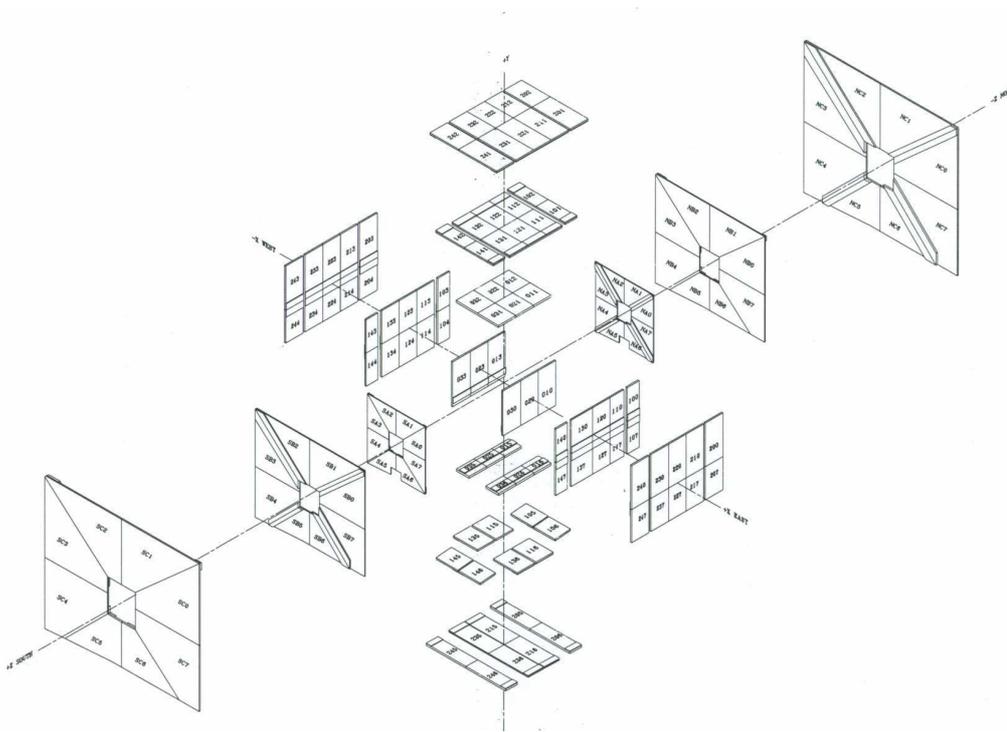


Figure 3.21: Exploded view of the three drift chamber layers of the local muon system.

### 3.3.5.3 Scintillation Counters

Sheets of scintillating pixels accompany each layer of drift tubes, with the exception of the B layer in the central system. Designed to cover  $4.5^\circ$  in  $\phi$ , they provide additional position measurement, and are used for triggering, cosmic ray veto, beam related muon rejection and track reconstruction. Figure 3.22 shows an exploded view of the scintillation counters.

The pixels consist of a slab of scintillator in which light-collecting fibers are set in grooves. A photomultiplier tube collects the light and provides an analog voltage pulse to the digitizing electronics.

The scintillator counters are discussed in more details in [111].

### 3.3.5.4 Shielding

Three sources contribute to background in the central and forward muon system:

- Scattered proton and antiproton fragments that interact with the end of the calorimeter or with the beam pipe produce background in the central and forward A layer.

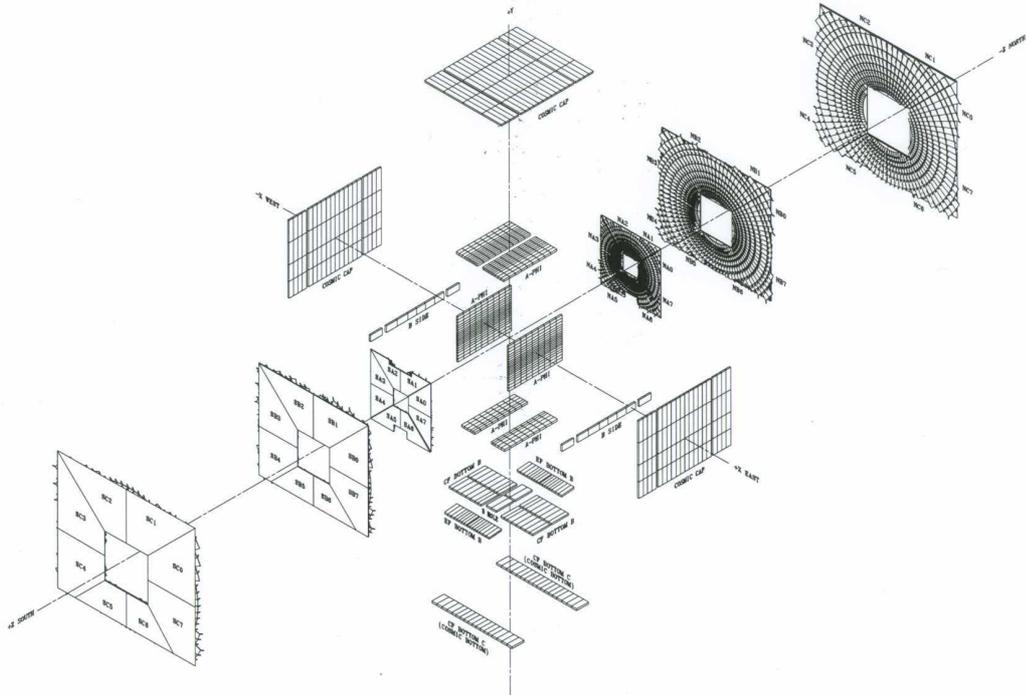


Figure 3.22: Exploded view of the three scintillator layers of the local muon system.

- Proton and antiproton fragments, mostly muons from pion decays created by proton and antiproton interactions upstream of the detector, interacting with the low beta quadrupole magnets produce hits in the B and C layers of the forward system.
- Beam halo interactions affect both the central and the forward muon system.

The shielding consists of layers of iron, polyethylene and lead in a steel structure surrounding the beam pipe and low beta quadrupole magnets. Iron is used as the hadronic and electromagnetic absorber, polyethylene is a good absorber for neutrons due to its high hydrogen content, and lead is used to absorb gamma rays.

### 3.3.5.5 Performance

The performance of the muon system combined with the central tracker is shown in Fig. 3.23 and Fig. 3.24 where the di-muon invariant mass is shown at the  $\omega$ ,  $\phi$ ,  $J/\Psi$ ,  $\Psi'$ ,  $\Upsilon$  and at the  $Z$  mass.

The momentum resolution of the muon system has been studied using

reconstructed muons for which a central track was associated. The momentum resolution for muons as measured by the muon system,  $\sigma(p_T)/p_T$ , varies between 0.1 for low-momentum muons and 0.5 for muons with  $p_T > 50$  GeV [112]. The overall muon momentum resolution, including information from the silicon vertex detector (see Sect. 3.3.3.1) and the central fiber detector (see Sect. 3.3.3.2), is defined by the central tracking system for muons with momentum up to approximately 100 GeV, the muon system improves the resolution only for higher momentum muons [93].

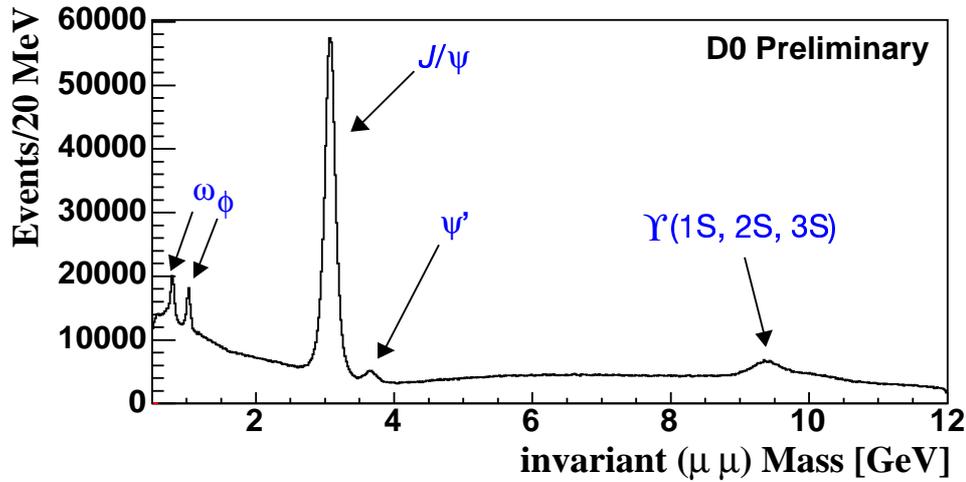


Figure 3.23:  $\mu^+\mu^-$  invariant mass distributions of various meson resonances:  $\omega$ ,  $\phi$ ,  $J/\Psi$ ,  $\Psi'$  and  $\Upsilon$  [113].

### 3.3.6 The Forward Proton Detector

The Forward Proton Detector (FPD) [114] is a series of momentum spectrometers that make use of accelerator magnets in conjunction with position detectors along the beam line in order to determine the kinematic variables  $t$  and  $\xi$  of the scattered  $p$  and  $\bar{p}$ , where  $|t|$  is the four-momentum transfer of the scattered proton or antiproton, and  $\xi = 1 - x_p$ , where  $x_p$  is the fractional longitudinal momentum of the scattered particle with respect to the incoming proton. The FPD covers the region  $0 \leq t \leq 4.5$  GeV<sup>2</sup>.

The position detectors must operate a few millimeters from the beam and have to be moved away during the injection of protons in the accelerator. Special devices, called Roman pots [115], are designed to house the position detectors allowing for remotely controlled movement with an accuracy in the order of tens of microns.

The Roman pots are housed in stainless steel chambers called castles. The FPD, shown in Fig. 3.25, consists of 18 Roman pots arranged in six castles, where the detectors placed in each castle can approach the beam

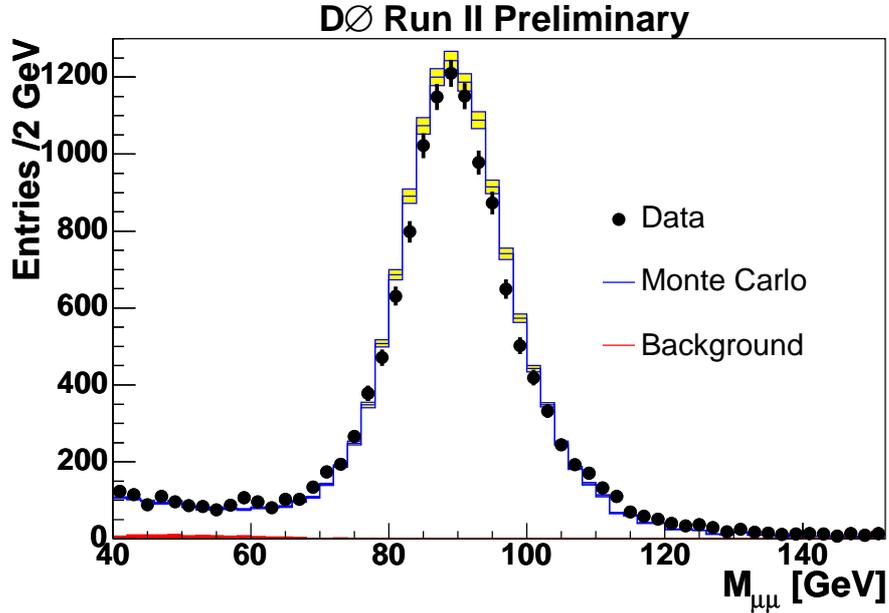


Figure 3.24:  $\mu^+\mu^-$  invariant mass spectrum at the  $Z$  mass [88].

from up, down, inside and outside directions with respect to the Tevatron ring. Four castles are located downstream of the low beta quadrupole magnets on each side of the colliding point: two on the proton side (P1 and P2) and two on the antiproton side (A1 and A2).

Figure 3.26 shows first results of the FPD system. The  $\xi$  distribution before and after various levels of corrections and the  $t$  distribution are shown.

### 3.3.7 The Trigger System

The overwhelming majority of proton antiproton encounters result in collisions of little interest. In particular, collisions which produce massive particles such as  $W$ ,  $Z$ , the top quark or those which might provide evidence of *New Physics* occur extremely rarely, as can be seen in Fig. 2.16. To accumulate a large sample of events of interest without having to store and reconstruct a staggering number of uninteresting collisions, DØ employs an event trigger which decides whether to store an event or to disregard it. At the Tevatron an input bunch crossing rate of 2.5 MHz must be reduced to the final rate of 50 Hz, a limit given by the offline reconstruction capabilities. The trigger system is a three tiered pipelined system; each tier examines the event in more detail than lower tiers and restricts the rate of events to higher tiers.

An event can fail the trigger because it was recognized as a less interest-

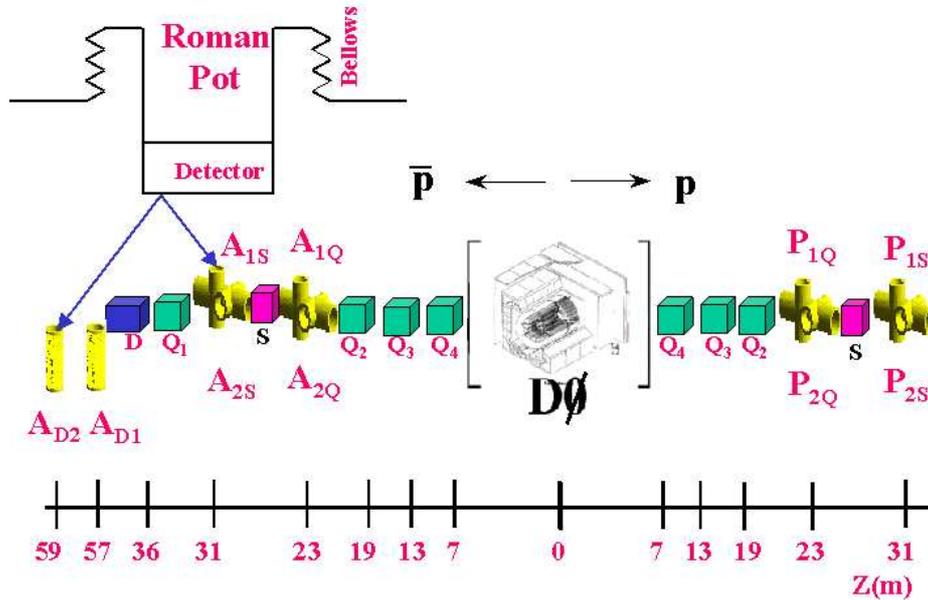


Figure 3.25: The FPD: Quadrupole Roman pot detectors are named P or A when placed on the  $p$  or  $\bar{p}$  side, respectively. Dipole pots, located on the  $\bar{p}$  side, are named D.

ing process, because it was mistaken for a less interesting process (trigger inefficiency), or because the trigger or data acquisition systems were busy processing previous collisions (dead time).

### 3.3.7.1 The Level 1 Trigger

Collisions occur at a rate of 2.5 MHz. The first trigger stage (Level 1 or L1) has a pipeline, which allows to make a decision within  $4.2 \mu\text{s}$ , resulting in a trigger accept rate of about 2 kHz. The trigger decision is made by a framework built of field programmable gate arrays (FPGAs), which take inputs consisting of simple objects created in the individual subdetectors from the luminosity monitor, the calorimeter and the muon system.

The luminosity system provides an indication that a collision occurred with a position on the  $z$  axis which would place it within  $D\bar{O}$ 's volume.

The calorimeter employs a special path which performs a very quick summation of electromagnetic and hadronic towers at a resolution of  $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ , excluding the coarse hadronic section due to a higher noise rate. The trigger requires that the energy in these towers be above a certain threshold. Based on signal to noise considerations, only the trigger towers for  $|\eta_{det}| < 3.2$  are used. Some of the data used in the analyses presented

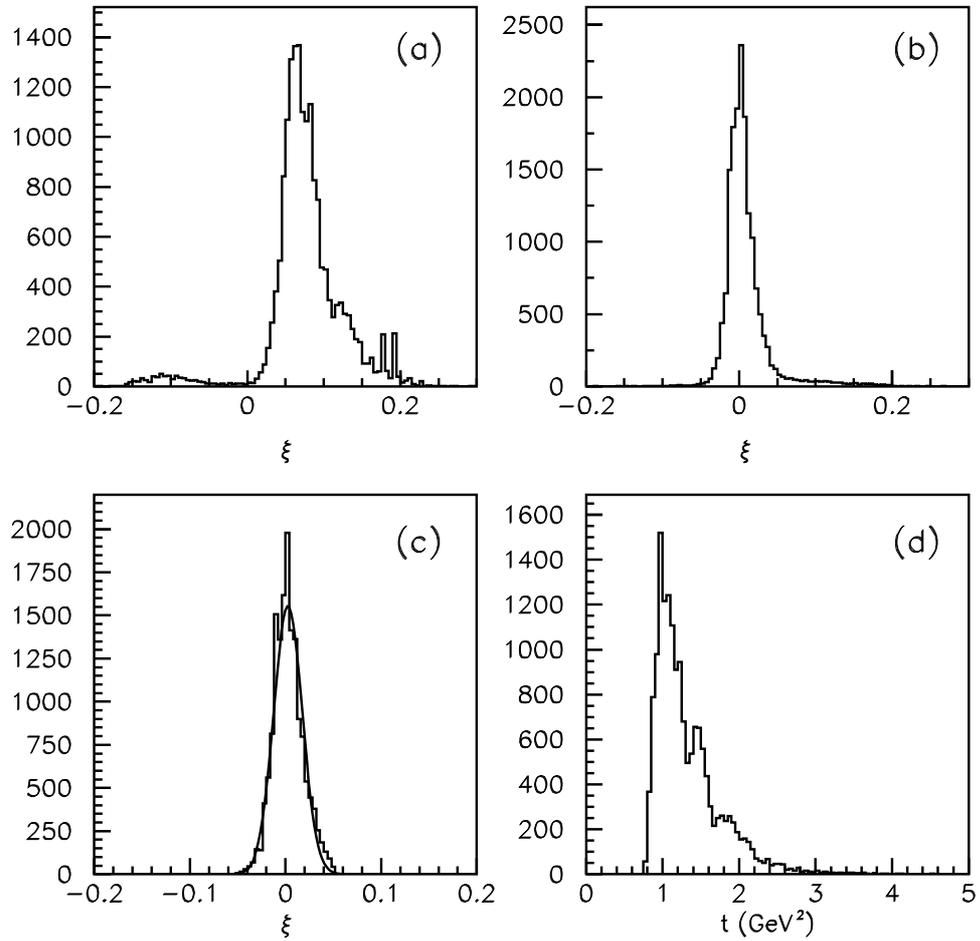


Figure 3.26:  $\xi = 1 - x_p$ , where  $x_p$  is the fractional longitudinal momentum of the scattered particle with respect to the incoming proton. a)  $\xi$  distribution before alignment; b)  $\xi$  distribution after alignment; c) fitted  $\xi$  distribution; d)  $t$  distribution [116].

are triggered with a limited trigger coverage,  $|\eta_{det}| < 2.4$ . Additional trigger terms are possible for global quantities such as the total sum of all tower energies,  $\sum E_T$ , and the missing transverse energy,  $\cancel{E}_T$ .

The muon trigger in the analyses presented requires a coincidence between the scintillators in the A and B or C layers. Additionally, the drift tubes and the central fiber tracker can be used to provide trigger information.

### 3.3.7.2 The Level 2 Trigger

In the second stage (Level 2 or L2), hardware engines associated with specific subdetectors process information used by a global processor in determining

correlations between different detectors, e.g. matching tracks and leptons. Level 2 has an accept rate of 1 kHz at a maximum dead-time of 5% and a maximal latency of 100  $\mu$ s. The Level 2 trigger passes events to the Level 3 system.

Figure 3.27 shows the design of the DØ Level 1 and Level 2 trigger system. The components currently being commissioned are shown in light gray.

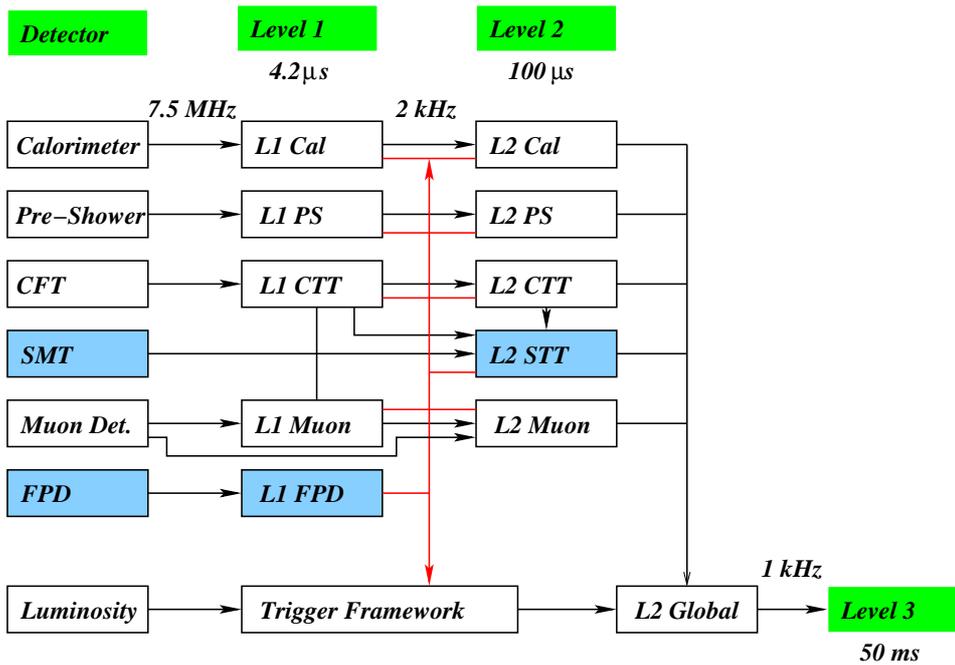


Figure 3.27: The design of the DØ Level 1 and Level 2 trigger system. The components currently being commissioned are shown in light gray.

### 3.3.7.3 The Level 3 Trigger

The third stage (Level 3 or L3) uses a collection of approximately 100 farm nodes to perform a limited reconstruction of the event and make a trigger decision using the full event information (complete physics objects as well as their correlations). The nominal 1 kHz input rate is reduced to 50 Hz for data recorded for offline analyses.



## Chapter 4

# Event Reconstruction and Object Identification

The data as collected consist of nearly a million channels of immediate detector response. These channels have to be processed carefully for evidence of the products of a collision, from which the kinematics of the collision can be inferred.

Algorithms are used to reduce the huge amount of information to define basic physics objects and their properties, representing the particles from the collision. The detector design allows to distinguish and to define the following fundamental objects: tracks, primary vertex, electrons, photons, muons, jets and their flavor, and missing transverse energy  $\cancel{E}_T$ .

Generally, the object reconstruction and identification is optimized for efficiency, purity and to provide the best possible measurement of the magnitude and direction of the object's momentum. Other objects in the event, particles not originating from the collision or noise in the detector components or the readout electronics can mimic the object signature, degrading the purity.

More specific to the analyses presented, the object identification is optimized in order to isolate  $t\bar{t}$  events from background events.

In this chapter, strategies for reconstructing these objects from the detector responses and the respective selection criteria are described.

### 4.1 Charged Tracks

Charged particles, curving through the magnetic field, leave traces in the central tracking system from which tracks are reconstructed. The reconstruction is abstracted in two pieces: hit clustering, which groups individual channels which are likely to represent the passage of an individual particle, and track finding, which finds groups of clusters located along a physical path.

**Hit Clustering** Particles passing through the silicon vertex detector will deposit charge in a number of strips. A particle traversing the central fiber detector will illuminate a number of fibers. In both cases a cluster is defined by a group of adjacent strips above a noise threshold.

**Track Finding** The track finding is subdivided into two algorithms: pattern recognition and track fitting. The pattern recognition creates sets of clusters which lie along physical paths. The track fitting uses sophisticated algorithms (Kalman fitter [117]) to fit a candidate charged particle track to a physical path, using a  $\chi^2$  test.

## 4.2 Primary Vertex

The precise determination of the primary vertex (PV) position is crucial for all  $b$ -tagging algorithms and in order to determine if a lepton originates from the PV. The PV reconstruction has been significantly modified in the underlying production release p14 compared to the previous release. Details of the new algorithm and its performance on Monte Carlo and data can be found in [118].

The new approach to the PV reconstruction consists of two major steps (two-pass algorithm):

- First, the algorithm locates the position of the beam spot center. At this step, all tracks with distance of closest approach ( $dca$ ) significance  $S_{(0,0)} < 100$ , calculated with respect to  $(x, y) = (0, 0)$  in the transverse plane, are fitted to the PV. The  $dca$  is the shortest distance between the track and the PV, the  $dca$  significance is defined by dividing the  $dca$  by its uncertainty. The result of the fit to the PV is the list of possible primary vertices.
- At the second pass, the track  $dca$  significance is calculated with respect to the position of these first-pass vertices. Only tracks with at least two SMT hits, surviving a tight  $dca$  significance cut, are fitted to the final primary vertices.

The next step is the selection of the hard scatter vertex from the list of reconstructed vertices. The method used in p14 is described in detail in Ref. [119]. It is based on the fact that tracks from minimum bias interactions have smaller transverse momenta than tracks from hard scatter interactions. The  $\log_{10} p_T$  distribution of tracks from minimum bias processes is used to define a probability for a track to come from a minimum bias vertex. For each vertex, the product of the probabilities of each track divided by the total number of tracks is calculated, thus forming the probability for

a vertex to originate from a minimum bias interaction. The PV with the lowest minimum bias probability is chosen as the hard scatter PV.

There are two implementations of the PV algorithms, so-called “DØreco” and “DØroot”. The momenta of calorimeter objects in the event, jets, electrons and the  $\cancel{E}_T$ , are reconstructed with respect to the DØreco PV. For tracking related quantities, the  $dca$ 's of leptons, and the reconstruction of secondary vertices for  $b$ -tagging, the slightly different DØroot PV algorithm is applied, since it has a better performance and since the DØreco PV used only a one-pass algorithm for part of the data set used.

The DØreco vertex finder and the DØroot vertex finder share the vertex selection procedure but differ in the track-selection and fitting techniques. However, they show comparable performance in p14 data in terms of reconstruction efficiency and in the reconstructed  $x$ ,  $y$  and  $z$  coordinate of the vertex, see Fig. 4.1. The main characteristics of both algorithms regarding track selection at the second pass are summarized in Table 4.1.

Track variable	DØreco	DØroot
$p_T$	$\geq 0.5$ GeV	$\geq 0.5$ GeV
SMT hits	$\geq 2$ (Data) $\geq 0$ (Monte Carlo)	$\geq 2$
$dca$ significance	$\leq 5.0$	$\leq 3.0$

Table 4.1: Track selection of the two vertex algorithms at the second pass.

The DØroot algorithm has an additional step which is performed before the two-pass approach. It starts from clustering tracks along the  $z$ -coordinate. The clustering algorithm starts from the track with highest  $p_T$  and adds the track with closest vertex, which is within 2 cm from the first one. The position of the cluster of tracks is recalculated with every additional track. Clusters of tracks are the input for the two-pass algorithm.

The average number of tracks in generic QCD multijet events is 20 and the average efficiency of the PV reconstruction is 98%. The efficiency of the PV reconstruction is about 100% in the central  $|z|$  region and drops quickly outside the SMT fiducial volume ( $|z| < 36$  cm for the barrel) due to the requirement of two SMT hits per track, forming the PV, see Fig. 4.2. The resolutions (convoluted with the beam spot size) of  $x$ ,  $y$  and  $z$  coordinates of the reconstructed PV depend on the number of tracks,  $Ntrks_{PV}$ , fitted into the PV. They are completely dominated by the beam spot starting from  $Ntrks_{PV} > 15$  and the resolution is about 35  $\mu\text{m}$  in the transverse plane.

To ensure a high reconstruction quality the following additional PV selection is required:

- $|z_{PV}| \leq 60$  cm.

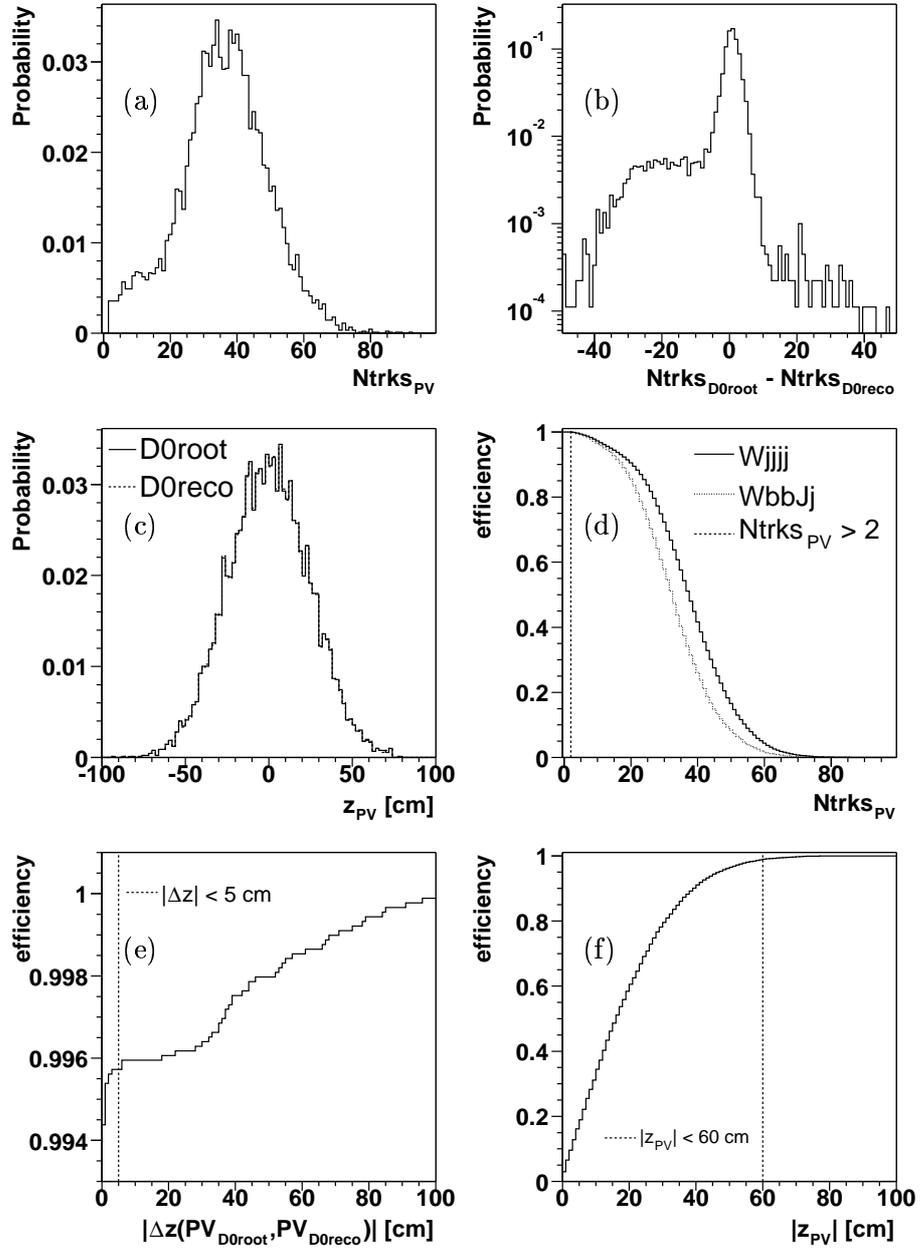


Figure 4.1: PV distributions and survival efficiencies for  $t\bar{t} \rightarrow \mu\nu_\mu b\bar{b}qq$  if not stated explicitly. Number of tracks associated to the PV (a), difference in number of tracks associated to the DØreco and the DØroot PV (b),  $z$  position of the DØreco and the DØroot PV, which is essentially identical (c), survival efficiency for a cut on the number of tracks associated to the DØroot PV (d), survival efficiency for a cut on the  $z$ -separation between the DØroot and the DØreco PV (e), survival efficiency for a cut on the  $z$  position of the DØroot PV (f).

- At least three tracks fitted into the PV:  $Ntrks_{PV} > 2$ .
- The DØreco PV and the DØroot PV are required to have a  $z$ -separation of less than 5 cm.

Figure 4.1 shows distributions and efficiencies for the number of tracks associated to the PV, the  $z$  position of the vertex and the  $z$ -separation of the DØreco and the DØroot PV. If not stated explicitly the DØroot PV is denoted by PV. The DØroot PV contains less tracks than the DØreco PV due to the tighter track selection given in Table 4.1. Events containing heavy flavor jets ( $WbbJj$ ) have less tracks associated to the PV than events containing only light flavor jets ( $Wjjjj$ ) since there are many tracks which are significantly displaced from the PV and therefore not associated to it (see also Table 4.1).

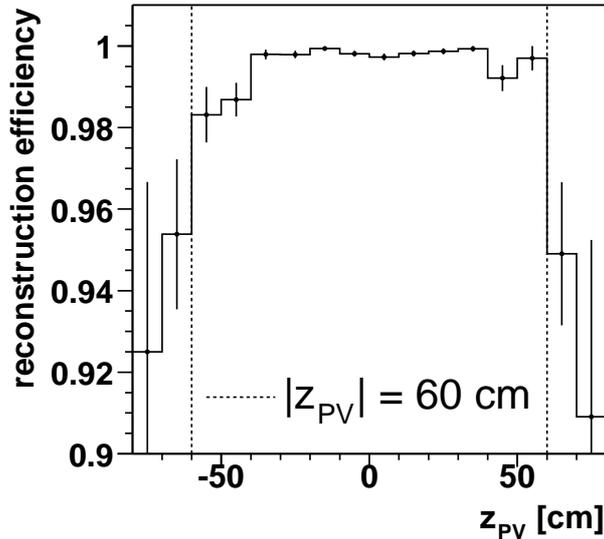


Figure 4.2: PV reconstruction efficiency as a function of  $z$  position of the PV.

### 4.3 Muons

Muons are reconstructed using information from two independent detector systems: the muon detector and the central tracker. A 'local' track in the muon system is the basis of the muon identification. Muons are required to have hits in all layers of the muon system (both inside and outside the toroid steel). The exact muon identification criteria, and a study of the Muon ID efficiency are presented in Sect. 4.3.1. The superior track resolution from the central tracker is used to improve the knowledge of the kinematic properties

of the muon, and to have a confirmation that the muon originated from the primary vertex. A refinement of the track  $p_T$  measurement is presented in Sect. 4.3.2. The muon momentum scale and resolution in MC is adjusted to the values found in the data, as discussed in Sect. 4.3.3. Section 4.3.4 describes the variables used to determine how well a muon is isolated from other physics objects in the event.

### 4.3.1 Muon Identification Criteria

The following standard Muon ID requirements are used:

- Muon candidates are required to be of ' $nseg = 3$  medium' quality, according to the certified p14 Muon ID criteria [120]:
  - at least two wire hits in the A segment,
  - at least one scintillator hit in the A segment,
  - at least two wire hits in the BC segment,
  - at least one scintillator hit in the BC segment, (except for central muons with less than four wire hits in the BC segment).
- The standard 'loose' cut against cosmics is applied, based on timing information from scintillator hits associated with the muon. The default timing cut values from the p14 Muon ID [120] are used, requiring A-layer and BC-layer scintillator times  $|t_A| < 10$  ns and  $|t_{BC}| < 10$  ns.

In addition a 'central track match' is required, which denotes that the muon tracks are extended to the point of closest approach (PCA) to the beam and their parameters are compared with those of central tracks at the PCA. A global fit is performed with all central tracks within 1 radian in azimuthal and polar angle of a muon track at PCA. The central track with the highest  $\chi^2$ -probability is considered as the muon candidate. The measurement of the muon track parameters are taken from the central tracking system. The following additional quality requirements are applied to the central track:

- $\chi^2_{track}/NDF < 4$  for the central track fit, to remove bad track fits.
- A distance  $|\Delta z(\mu, PV)| < 1$  cm between the track and the primary vertex, to further reduce background from cosmics and badly reconstructed tracks
- $dca$  significance less than 3 standard deviations away from zero,  $dca/\sigma(dca) < 3$ , in order to reject muons from semi-leptonic heavy flavor decays, see Fig. 4.3.

A reliable central track measurement is important since the central track  $p_T$  is used in the analysis, e.g. to build some of the topological variables.

Tracks that fail the  $\chi_{track}^2/NDF$  or  $\Delta z(\mu, PV)$  quality cuts are observed to have a larger probability to give a very high  $p_T$  measurement ( $> 200$  GeV), see Fig. 4.4. Applying these cuts significantly reduced the tail of very high  $p_T$  muons. The remaining high  $p_T$  muons can be explained by genuine high  $p_T$  muons, for which the track curvature  $q/p_T$  is well measured within the finite resolution of the central tracker.

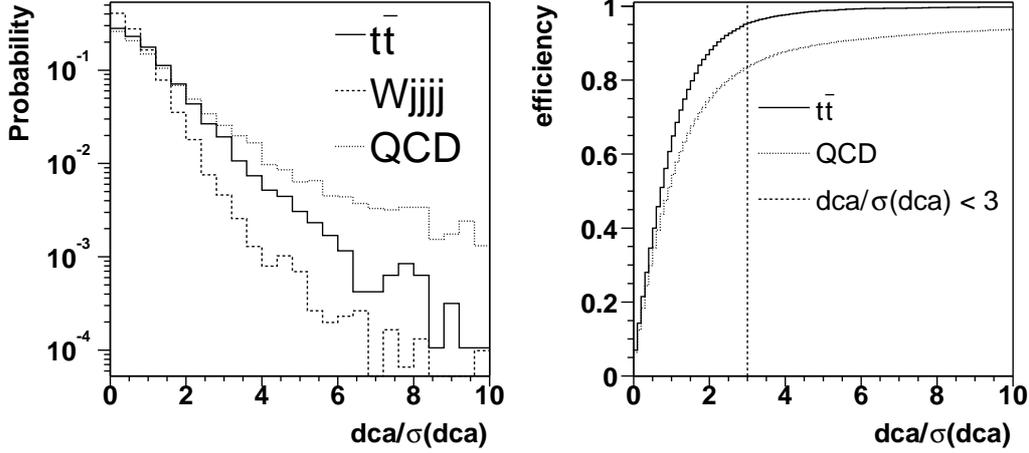


Figure 4.3: Left: Distributions for the distance of closest approach significance of the muon with highest  $p_T$  in  $t\bar{t} \rightarrow \mu\nu_\mu b\bar{b}qq$  MC,  $Wjjjj$  MC and QCD data (selected by requiring non-isolated muons) events. The presence of the two heavy flavor jets in the  $t\bar{t} \rightarrow \mu\nu_\mu b\bar{b}qq$  events leads to a larger  $dca$  significance than for  $Wjjjj$  events due to the finite lifetime of the  $b$  hadrons. Muons from semileptonic heavy flavor decays (QCD) are more likely to be significantly displaced from the PV. Right: Survival efficiencies as a function of the cut on the  $dca/\sigma(dca)$ .

### 4.3.2 Muon Momentum Correction

A correction to the momentum of muons matched to CFT-only tracks, tracks where no hits are found in the SMT subdetector, is employed. The procedure adopted from the Muon ID group considers the primary vertex as a constraint for the fit. The track is refitted such that the  $dca$  in  $x$  and  $y$  remains 0 [121]. The correction factor is given by:

$$SF_{CFT-only} = 1 - dca/qopt * (ERR(r, qopt)/ERR(r, r)),$$

where  $dca$  is the muon  $r$ - $\phi$  distance of closest approach to the primary vertex,  $qopt$  is the muon charge divided by the muon  $p_T$ ,  $ERR(r, qopt)$  and  $ERR(r, r)$  represent the respective error matrix entries where  $r$  is the  $r$ - $\phi$  impact parameter relative to  $(x, y, z) = (0, 0, 0)$ .

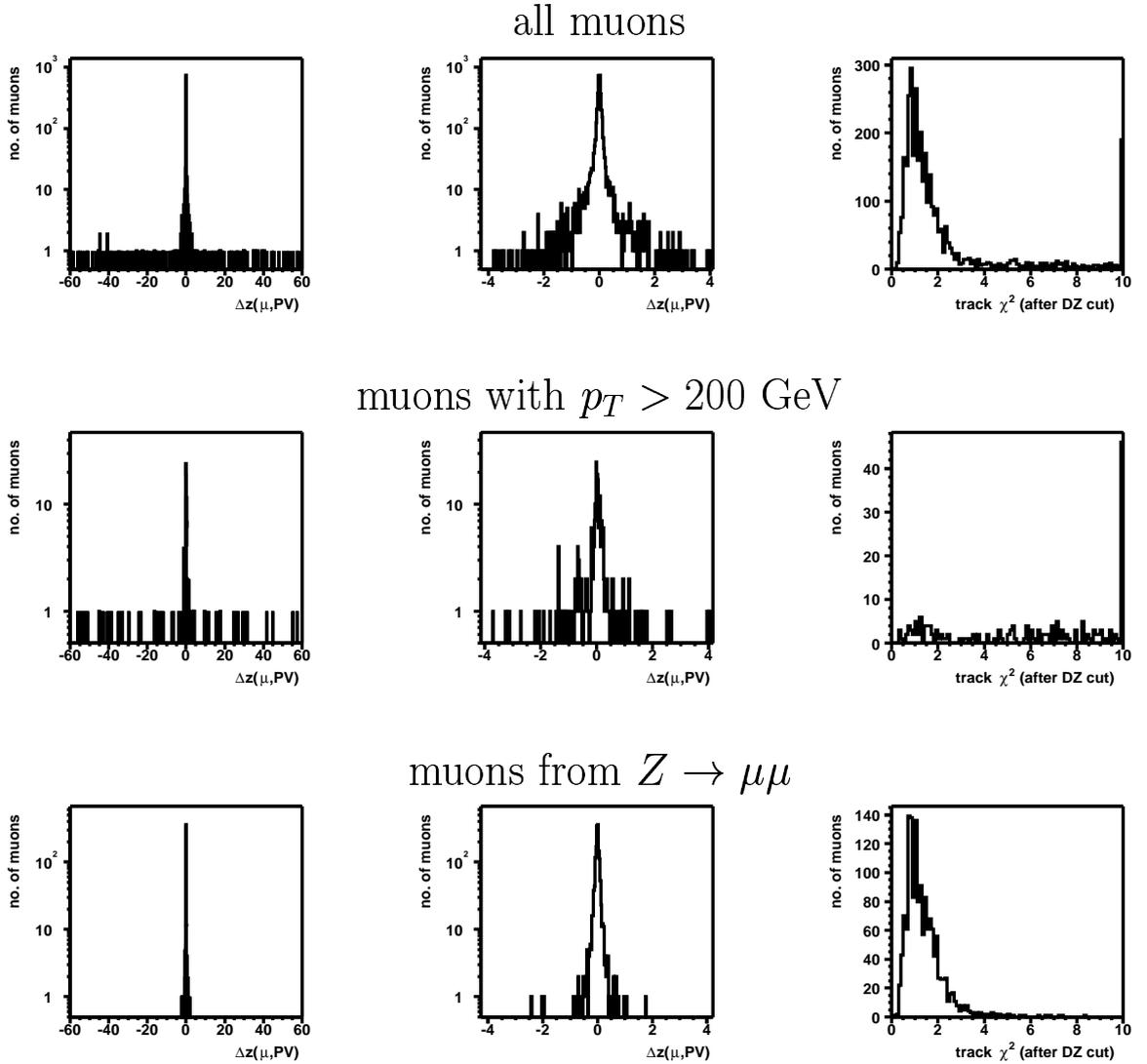


Figure 4.4: The distributions of the distance  $\Delta z$  between the vertex of the central track matched to the muon, and the primary vertex (left and middle on different  $x$ -scales) and  $\chi^2/NDF$  of the central track (right) for three different event selections: all muon quality cuts applied except for the cuts on  $\Delta z$  and  $\chi^2/NDF$  (top), idem for muons with a central track  $p_T > 200$  GeV (middle), and for muons identified to come from a  $Z$  boson decay,  $Z \rightarrow \mu\mu$  (bottom). Clearly the fraction of bad quality central tracks is enhanced for muons with very high  $p_T$ .

Roughly 18% of the muons are subject to this correction, with a strong  $\eta_{det}$  dependence, in the central region much less muons are affected. The size of the correction is shown in Fig. 4.5.

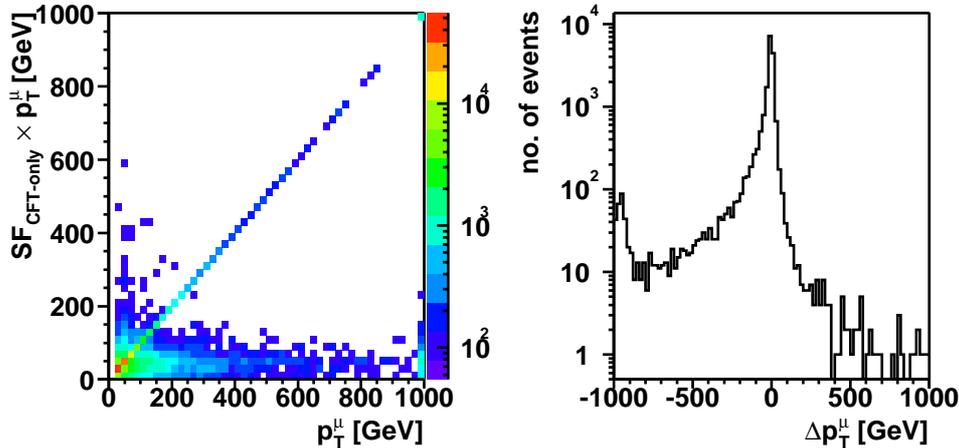


Figure 4.5: Left: Corrected muon  $p_T$  versus uncorrected muon  $p_T$ , most of the time the correction is small and the  $p_T$ 's are very similar. In particular muons with a very high  $p_T$  ( $p_T > 200$  GeV) are corrected down. Right:  $p_T$  difference:  $SF_{CFT-only} \times p_T - p_T$

### 4.3.3 Muon Momentum Scale and Resolution

The muon momentum scale and resolution in the Monte Carlo simulation is found not to reproduce the values measured in data. Corrections are applied to the Monte Carlo to bring data and MC to an agreement, which is crucial for efficiencies or shapes determined from the MC.

A data to MC comparison of the fitted width of the  $Z$  mass reconstructed from  $Z \rightarrow \mu\mu$  events allows to infer the difference in resolution. Since the resolution in the MC is found to be overestimated a simple oversmearing of the muon  $p_T$  in the MC can be used as a correction. The fitted mean of the  $Z$  mass allows to constrain the scale correction which has to be applied to the MC. The fitted mean and width are summarized in Table 4.2.

	fitted $Z$ mass	fitted $Z$ width
data	$89.9 \pm 0.5$	$11.1 \pm 0.5$
MC	$91.7 \pm 0.4$	$8.8 \pm 0.3$

Table 4.2: Fitted  $Z$  mass and width in data and MC [122].

The muon momentum correction is realized by the following transformation:

$$\frac{1}{p_T} \rightarrow \frac{1}{\alpha p_T} + G(0, \sigma_\xi), \quad (4.1)$$

where  $\alpha$  is the scale factor which accounts for the overall calibration, and  $G(0, \sigma_\xi)$  is a random variable drawn from a Gaussian distribution with a mean of 0 and a width  $\sigma_\xi$ .

The transformation from Eq. 4.1 is performed for different parameter values of  $(\alpha, \sigma_\xi)$ ; this is done by systematically scanning the relevant parameter space in small steps. The optimum scale  $\alpha$  and the width of the Gaussian,  $\sigma_\xi$ , are obtained simultaneously from a Kolmogorov-Smirnov goodness-of-fit test [123] for the  $Z$  mass distribution in data and Monte Carlo.

The level of agreement between data and Monte Carlo is found to be dependent on the muon detector pseudo-rapidity,  $\eta_{det}$ . Due to the limited statistics only two pseudo-rapidity regions are considered. First, the procedure is applied only to the events with both muons in the central region, defined as  $|\eta_{det}| < 1.62$ . After the scale factor and the additional smearing correction is determined for the central muons, the rest of the sample is treated in the following way. For the events with both muons in the forward region the procedure is the same as above. For those events with one muon in the central and the other in the forward region, the central muon  $p_T$  is corrected using the fixed parameters, obtained from the central-central case, while the scale and amount of additional smearing for the forward muon  $p_T$  are allowed to vary.

The Kolmogorov-Smirnov matching probability is determined as a function of  $\sigma_\xi$  and  $\alpha$ . The test is repeated ten times per set of parameters  $(\alpha, \sigma_\xi)$ , in order to be less dependent on statistical fluctuations which enter through the random variable drawn from the Gaussian distribution  $G(0, \sigma_\xi)$ . The selection cut on the dimuon invariant mass is varied from  $70 < m_{\mu\mu} < 110$  GeV to  $75 < m_{\mu\mu} < 105$  GeV; the optimal oversmearing and scaling correction are found to be independent of the choice of mass range.

Table 4.3 summarizes the result of the Kolmogorov-Smirnov goodness-of-fit test both for the muons in the central region and in the forward region.

The transverse momenta of muons in the simulation are corrected according to Eq. 4.1 using the numbers for  $\sigma_\xi$  and  $\alpha$  from Table 4.3.

#### 4.3.4 Muon Isolation Criteria

The muons coming from the leptonic decay of a  $W$  boson tend to be isolated from jets and to have a relatively high transverse momentum  $p_T$ , see Fig. 4.6. The main background for identification of such muons comes from semileptonic decays in heavy quark jets. Those muons tend to be non-isolated and

	$\alpha$	$\sigma_\xi$
central	0.991	0.0025 GeV <sup>-1</sup>
forward	0.996	0.0043 GeV <sup>-1</sup>

Table 4.3: Result of the Kolmogorov-Smirnov goodness-of-fit tests for muons in the central region ( $|\eta_{det}| < 1.62$ ) and muons in the forward region ( $1.62 \leq |\eta_{det}| \leq 2.00$ ).  $\alpha$  represents the transverse momentum scale correction and  $\sigma_\xi$  is the width of the Gaussian oversmearing. [122].

to have a lower transverse momentum, shown in Fig. 4.6, labeled as “QCD”. Muon isolation is used to distinguish muons from these two different sources.

A loose isolation criteria is defined by demanding that a muon is separated from a jet,  $\Delta R(\mu, jet) > 0.5$ , see Fig. 4.6.

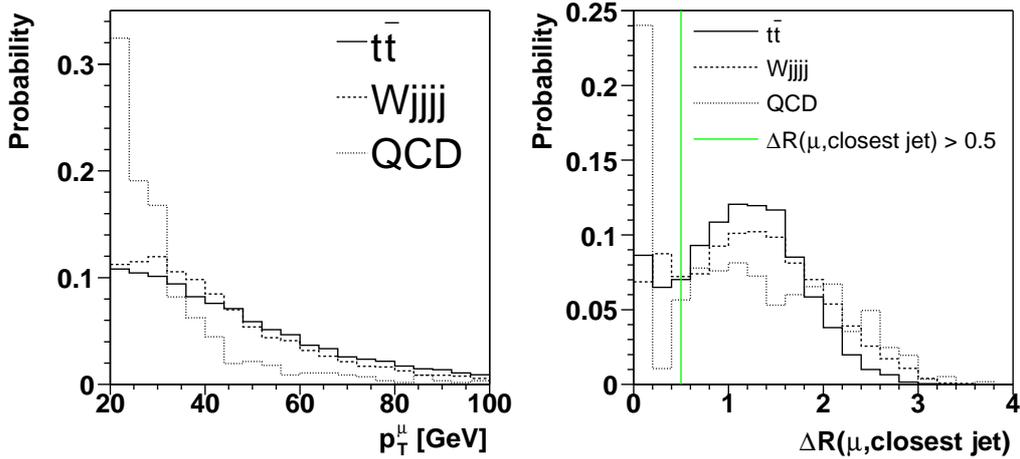


Figure 4.6: Transverse momentum of the high  $p_T$  muon in  $t\bar{t}$  and  $Wjjjj$  MC and in QCD data (left).  $\Delta R(\mu, jet)$  in  $t\bar{t}$  and  $Wjjjj$  MC and in QCD data (right)

In previous  $t\bar{t}$  cross-section measurements [124] the following variables and cuts were used as a measure of muon isolation [125]:

- $\text{Halo}(0.1,0.4) < 2.5$  GeV, where  $\text{Halo}(0.1,0.4)$  is the sum of the  $E_T$  of calorimeter clusters in a hollow cone between  $\Delta R = 0.1$  and  $\Delta R = 0.4$  away from the muon, as illustrated in Figure 4.7. In forming this sum, cells in the electromagnetic and fine hadronic calorimeters are considered, due to an enhanced noise level the coarse hadronic calorimeter is excluded from the sum.
- $\text{TrkCone}(0.5) < 2.5$  GeV, where  $\text{TrkCone}(0.5)$  is the sum of the  $p_T$  of

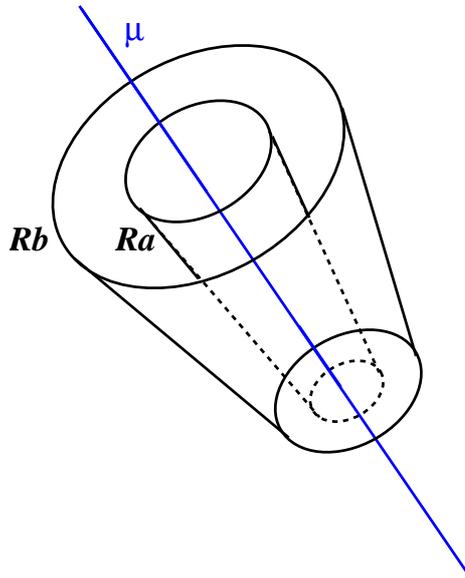


Figure 4.7: A hollow cone in  $R$ , with the inner edge  $R_a$  and the outer edge  $R_b$  surrounding the muon. Halo(0.1,0.4) is calculated by summing the transverse energies of all calorimeter cells  $i$  with  $R_i$ :  $R_a > R_i > R_b$ .

all tracks within a cone of radius  $\Delta R = 0.5$  surrounding the muon. The track matched to the muon is excluded from this sum.

The distributions for Halo(0.1,0.4) and TrkCone(0.5) and the survival efficiencies are shown in Fig. 4.8 for  $t\bar{t} \rightarrow \mu\nu_\mu b\bar{b}q\bar{q}'$  MC,  $Wjjjj$  MC and QCD data.

It was found that the separation between the two classes of muons mentioned above could be improved by using the difference between their  $p_T$  spectra. Thus, more powerful isolation variables were defined, taking the ratio of the above isolation variables and the transverse momentum of the muon,  $p_{T\mu}$ :

- Rat11  $\equiv$  Halo(0.1,0.4)/ $p_{T\mu} < 0.08$  and
- Rattrk  $\equiv$  TrkCone(0.5)/ $p_{T\mu} < 0.06$ .

The distributions for Rat11 and Rattrk and the survival efficiencies are shown in Fig. 4.9 for  $t\bar{t} \rightarrow \mu\nu_\mu b\bar{b}q\bar{q}'$  MC,  $Wjjjj$  MC and QCD data.

#### 4.4 Electrons

At the reconstruction stage, an EM cluster is defined as a set of towers in a cone of radius  $R = 0.2$  around an initial tower selected on the basis of its

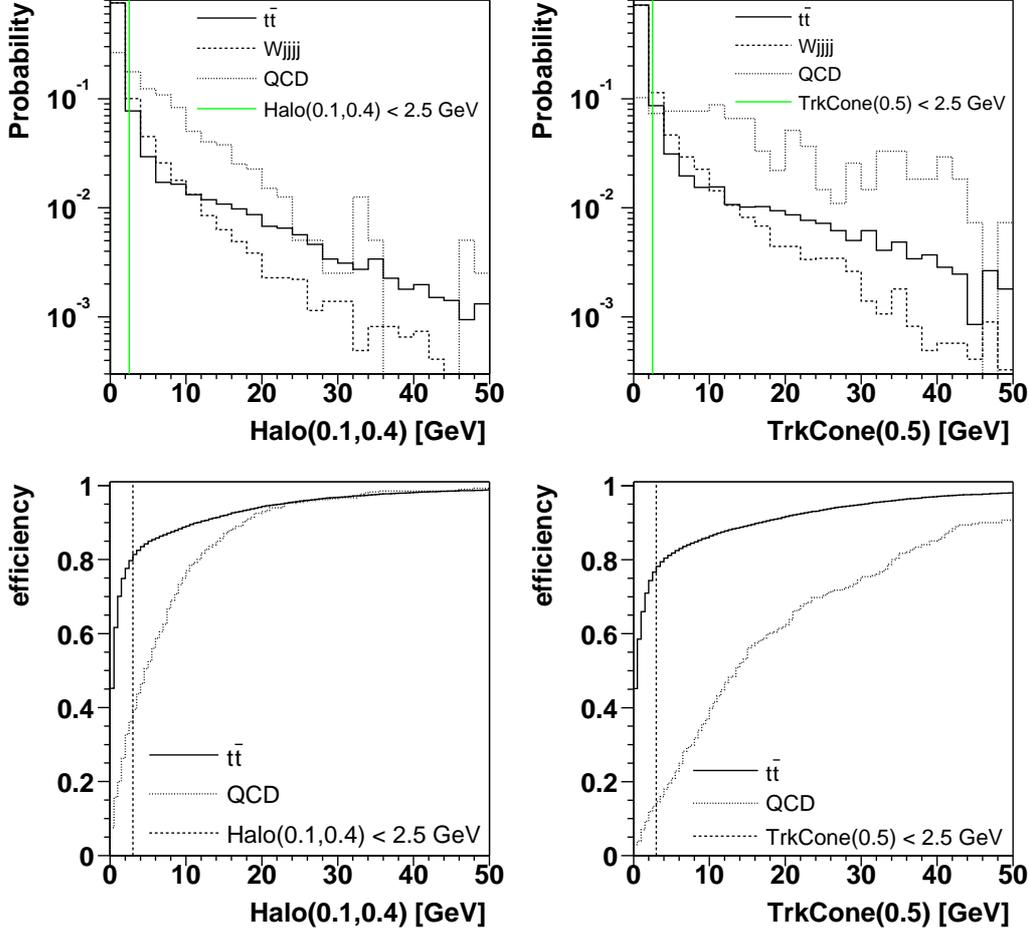


Figure 4.8: Top: Distributions for Halo(0.1,0.4) (left) and TrkCone(0.5) (right) for  $t\bar{t} \rightarrow \mu\nu_\mu b\bar{b}qql$  MC,  $Wjjjj$  MC and QCD data. Bottom: Survival efficiencies as a function of the cut on the isolation variables.

energy content. Among all reconstructed clusters, genuine EM showers are expected to have a large EM fraction  $EMF \equiv E_{EM}/E_{tot}$ , where  $E_{EM}$  is the cluster energy in the EM section of the calorimeter and  $E_{tot}$  is its total energy within the cone. Furthermore, they are expected to have a longitudinal and lateral shower development compatible with those of an electron. Each cluster is attributed a  $\chi_{cal}^2$  (*H-matrix7*) based on seven variables which compares the values of the energy deposited in each layer of the EM calorimeter and the total energy of the shower with average distributions obtained from the simulation. Additionally, a variable is defined that quantifies the cluster isolation:

$$f_{iso} = \frac{E_{tot}(R < 0.4) - E_{EM}(R < 0.2)}{E_{EM}(R < 0.2)} \quad (4.2)$$

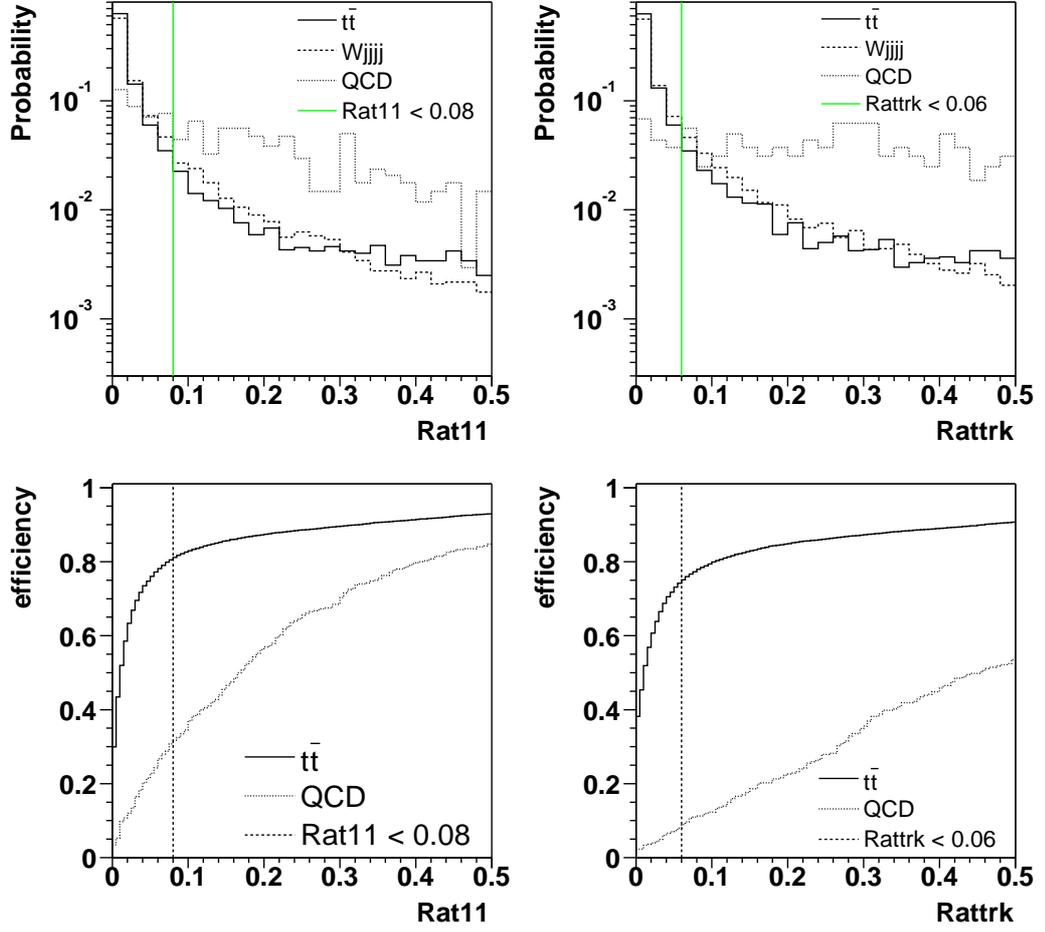


Figure 4.9: Top: Distributions for  $Rat11$  (left) and  $Rattrk$  (right) for  $t\bar{t} \rightarrow \mu\nu_\mu b\bar{b}q\bar{q}l$  MC,  $Wjjjj$  MC and QCD data. Bottom: Survival efficiencies as a function of the cut on the isolation variables.

Electron candidates are selected by requiring

- $EMF > 0.9$ ,
- $f_{iso} < 0.15$  and
- $H-matrix7 < 75$ .

**Track Match:** For electrons, an associated track candidate is required.

An associated track candidate is a track in a road satisfying:

- $|\Delta\phi(EM, track)| < 0.05$ ,
- $|\Delta\eta(EM, track)| < 0.05$ .

**Electron Likelihood:** It is required that the electron resembles a canonical sample of electrons by selecting those with a large electron likelihood [126], which includes both calorimeter and tracking information. Only electrons with a likelihood discriminant larger than 0.85 are used.

In the analyses presented electrons are required to be reconstructed in the CC calorimeter:  $|\eta_{det}| < 1.1$  and have a transverse momentum  $p_T > 20$  GeV.

## 4.5 Jets

Hadronic particles are reconstructed as jets in the calorimeter; the algorithm for the reconstruction is presented in Sect. 4.5.1. An additional algorithm is used to reduce contributions from noisy cells to the jet; the T42 algorithm is discussed in Sect. 4.5.2. Reconstructed jets have to meet identification criteria in order to be considered a good jet, these are given in Sect. 4.5.3. Electromagnetic objects might also be reconstructed as jets, the method to separate them is subject of Sect. 4.5.4. The reconstructed energy in a jet cone is not equal to the original particle level energy, the necessary corrections are discussed in Sect. 4.5.5. The measurement of the jet energy resolution is presented in Sect. 4.5.6.

### 4.5.1 The Jet Reconstruction Algorithm

The interaction of hadronic particles with the calorimeter is explained in Sect. 3.2, resulting in a shower of hadronic particles which has typically the shape of a cone. The ideal jet algorithm should reconstruct the kinematic properties of the initial hadronic particle. In particular, it should be infrared (see Fig. 4.10) and collinear safe (see Fig. 4.11), i.e. it should not depend on the fact that the cross section for infrared and collinear parton emission diverges, it should be independent of the detailed detector geometry and granularity, have a maximal reconstruction efficiency and require a minimal CPU time.

The jet algorithm associates adjacent particles, reconstructed as clusters of energy in the calorimeter, into jets. Following the recommendation of the Run II QCD workshop, the *improved legacy cone* algorithm [127] comes closest to the ideal jet reconstruction, as discussed above, and is used for the analyses presented.

The first step of the algorithm is to find seeds, represented by calorimeter towers above a minimum seed threshold,  $E_T^{tower} > E_T^{seed}$ , where a tower is defined as the sum of all cells sharing the same pseudo-rapidity and azimuthal angle. Cells in the coarse hadronic calorimeter, the end cap massless gap or the end cap hadronic layer 16 or 17 are not considered as seeds due to an enhanced noise level.

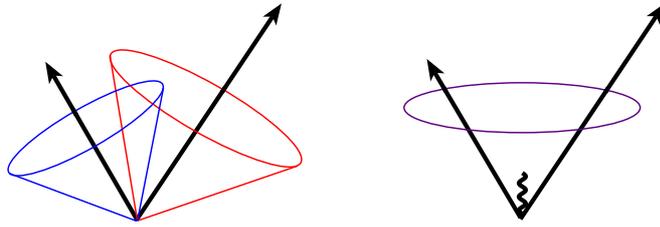


Figure 4.10: An illustration of infrared sensitivity in cone jet clustering. Clustering begins around seed particles, shown as arrows with length proportional to energy. Soft radiation from either of the two particles, represented by the curly line, (right sketch) may cause a merging of the jets.

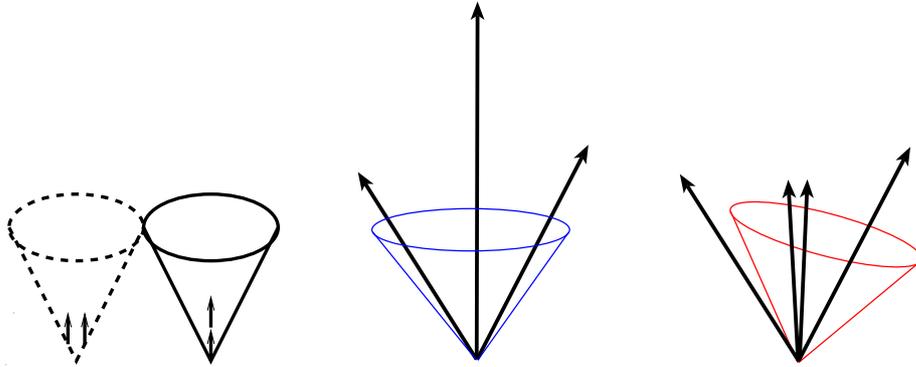


Figure 4.11: An illustration of collinear sensitivity in cone jet clustering. Left two sketches: No seed particle is produced due to the energy splitting among several detector towers (left), whereas a seed is produced without the splitting (right). The right two sketches show the sensitivity of the jet reconstruction to an  $E_T$  ordering of seed particles

Proto-jets are created with a simple cone algorithm from an  $E_T$  ordered list of seeds.

A seedless algorithm is found to be infrared safe and is therefore preferred over an algorithm with seeds, however, it is expensive computationally. An approximation of a seedless algorithm is achieved by the addition of midpoints,  $E_T$  weighted centers between pairs of proto-jets. These centroids are also considered as proto-jets.

Overlapping cones are separated with a split and merge procedure.

The jet algorithm specifications can be summarized as:

- $R_{cone} = 0.5$
- $E_T^{seed} = 0.5 \text{ GeV}$
- add midpoints after the cone clustering

- split and merge proto-jets
- keep all jets with more than  $E_T^{reco} = 8$  GeV.

### 4.5.2 The T42 Algorithm

The T42 algorithm has been proposed [128] to obtain a finer and uniform treatment of the calorimeter noise which leads to an improvement in the reconstruction of different objects (electrons, photons, jets,  $\cancel{E}_T$ ), whose identification and energy measurement relies mainly on the calorimeter. Calorimeter noise, generally defined as energy deposition not related to the hard interaction, can be schematically classified as “hot”, “warm” or “normal”:

**Hot noise:** Hot cells are related to detector problems (hardware failure, abnormal electronic noise), or to physics processes like backscattering of particles interacting in the beam pipe outside of the vertex interaction region into the calorimeter. The energy is typically large,  $> 1$  GeV.

**Warm noise:** Warm cells are due to pedestal subtraction problems or hardware deficiencies, the cell energy levels are typically lower, on the order of hundred MeV, however, they might appear in great numbers in a definite region of the detector, creating so-called warm zones.

**Normal noise:** Normal noise cells are due to Gaussian electronic noise surviving the zero suppression. They are at lower energies, typically below  $4\text{--}5\sigma$ , where  $\sigma$  is the RMS of the pedestal. Typically, between 1000 and 3000 such cells appear per event.

The T42 algorithm is implemented in the TopAnalyze code [76, 77] and is processed before reconstructing the calorimeter objects. It aims at the rejection of the “normal” noise cells. For the T42 algorithm, an isolated cell is considered a noise cell and thus discarded if it is not “signal-like”. A cell is considered “signal-like” if its energy is positive (negative energy cells can originate from electronics noise and from pile-up which is baseline subtracted) and above a high threshold of  $+4\sigma$ , or if its energy is above  $+2\sigma$  and in addition the energy of a neighboring cell is above  $+4\sigma$ . The acronym T42 stands for “threshold  $4\sigma - 2\sigma$ ”, however, the current implementation corresponds to “threshold  $4\sigma - 2.5\sigma$ ”, resulting in the name T42.5.

The first electromagnetic layer (layer 1), and the layers 8, 9, 10 of the intercryostat region are not considered by the algorithm, so all their cells with positive energy are kept in the event, and are not used as neighbors. A detailed description of the current implementation can be found in [129].

The ratio of rejected cells by T42 over the number of cells in the event ranges from 30% to 60%. In the main part of the calorimeter ( $|\eta| < 3.2$ ), the fraction of cells rejected by T42 corresponds to the number of cells expected

from noise between  $2.5$  and  $4\sigma$ , assuming a Gaussian distribution [130]. This is a good indication that T42 is indeed reducing mainly noise cells. In the forward region, more cells are rejected since cells from pile-up effects accumulate close to the beam pipe [130], but this has no influence on high  $p_T$  physics, which is subject of the analyses presented.

### 4.5.3 Jet Identification

Once jets are clustered following the cone algorithm, further quality selection cuts are applied to each jet. These criteria are aimed at removing jets which are not reconstructed from hadronic particles from the hard interaction:

- To remove isolated electromagnetic particles a cut on the energy fraction, deposited in the electromagnetic section of the calorimeter ( $EMF$ ), is applied at  $0.05 < EMF < 0.95$ .
- To remove jets which predominantly deposit their energy in the coarse hadronic section of the calorimeter, a cut on the fraction of the jet energy deposited therein ( $CHF$ ) is applied at  $CHF < 0.4$ . The noise level is higher in the coarse hadronic section; this cut is essentially aimed at removing those jets which are clustered around noise in the coarse hadronic section.
- To remove jets clustered from hot cells, a cut on the ratio of the highest to the next-to-highest transverse energy cell in the calorimeter ( $HotF$ ) is applied at  $HotF < 10$ .
- To remove those jets clustered from a single hot tower, the number of towers containing 90% of the jet energy ( $n90$ ) is required to be greater than 1.
- L1 jet confirmation, discussed in Sect. 4.5.3.1.
- The minimum  $p_T$  requirement for jets after the jet energy scale correction (see Sect. 4.5.5) is 15 GeV.

The cut values for these variables were determined in the data by defining samples which contain predominantly the jets which should be kept (denoted as “good” jets), and samples which are enriched in the jets not reconstructed from hadronic particles from the hard interaction but from noise as described in Sect. 4.5.2 or from noise in the electronics of the readout chain. The latter jets are aimed to be rejected (denoted as “fake” jets), while keeping a high identification efficiency for the good jets.

Physics processes with a clear detector signature are used to select the samples of good and fake jets. For the good jets the “tag-and-probe” method is used:

**Tag-and-probe method:** Processes with two objects in the final state are selected by requiring that one of the two objects passes all the quality criteria, which might be even tightened to guarantee a high purity. This is called the “tag” object. The probe object is identified with a lower quality, where the selection criteria under study is not required.

The tag and probe objects are back-to-back in the azimuthal angle. This leads to the fact that the tag-and-probe method tends to overestimate the efficiency in the case of correlated inefficiencies which are also back-to-back.

The samples of good jets used are “dijet” events and “QCD Compton” events, example Feynman diagrams are shown in Fig. 4.12:

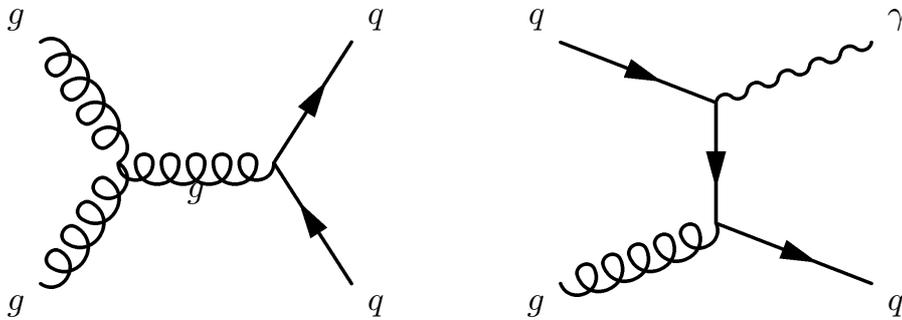


Figure 4.12: Example Feynman diagram for a dijet event (left) and a QCD Compton event (right).

**Dijet events:** to probe high  $p_T$  jets, events are selected which satisfy:

- single jet trigger
- two jets reconstructed
- $\cancel{E}_T < 5$  GeV
- the tag jet (the jet which is not considered for the efficiency measurement) passes all jet ID criteria and is required to fire the trigger
- $\Delta\phi$  (tag jet, probe jet)  $> 2.9$

**QCD Compton events:** to probe low  $p_T$  jets, events are selected which satisfy:

- single EM trigger
- one EM object and one jet reconstructed
- $\cancel{E}_T < 5$  GeV

- $\Delta\phi$  (EM object, probe jet)  $> 2.9$

The sources of noise resulting in fake jets are manifold and might depend on the luminosity, the event topology, they might be run dependent and have a strong dependence on the  $\eta_{det}$  and the  $p_T$  of the jet. As a consequence, it is difficult to define a clear signature for the presence of a certain class of fake jets in an event. Two sample selections suggest the presence of fake jets, so-called “third-jet” events and “multijet” events:

**Third-jet events:** dijet or QCD Compton events are selected which satisfy:

- $p_T$  balance of the two objects:  $\Delta p_T < 5$  GeV
- $\cancel{E}_T > 10$  GeV
- the probe fake jet is defined as a third-jet reconstructed in the event with no tracks with  $p_T > 0.5$  GeV matched to the jet in a cone of  $\Delta R < 0.5$

The first requirement guarantees that the selected events are very likely to be clean dijet or QCD Compton events with only two objects in the final state. This di-object system is used to “tag” the event. The second requirement, however, violates momentum conservation in the transverse plane. This can either be explained by an undetected particle from the hard scatter interaction (e.g. a neutrino) or it has not the hard scatter interaction as its origin. The latter is further supported by the third requirement, which selects an additional jet which did not leave any trace in the tracking system and thus, is likely to have noise as its origin.

**Multijet events:** events are selected which satisfy:

- $\geq 6$  jets reconstructed
- $\cancel{E}_T > 20$  GeV
- the probe fake jets are defined as jets with no tracks with  $p_T > 0.5$  GeV matched to the jet in a cone of  $\Delta R < 0.5$

The dominant production mechanism for a signature of six or more jets in an event is the QCD multijet production, where no or very little invisible particles are expected in the final state. As for the “third-jet” events the second requirement suggests the presence of noise jets, and the third requirement selects those which are not confirmed by the tracking system.

The survival efficiency for good and for fake jets is studied as a function of the cut value. The distributions for these quantities are shown in Fig. 4.13. Since the level of purity and representativeness of the fake jet samples is not

known precisely, the cut values are optimized for a high selection efficiency, while trying to reject some of the alleged fake jets. Another criteria for the

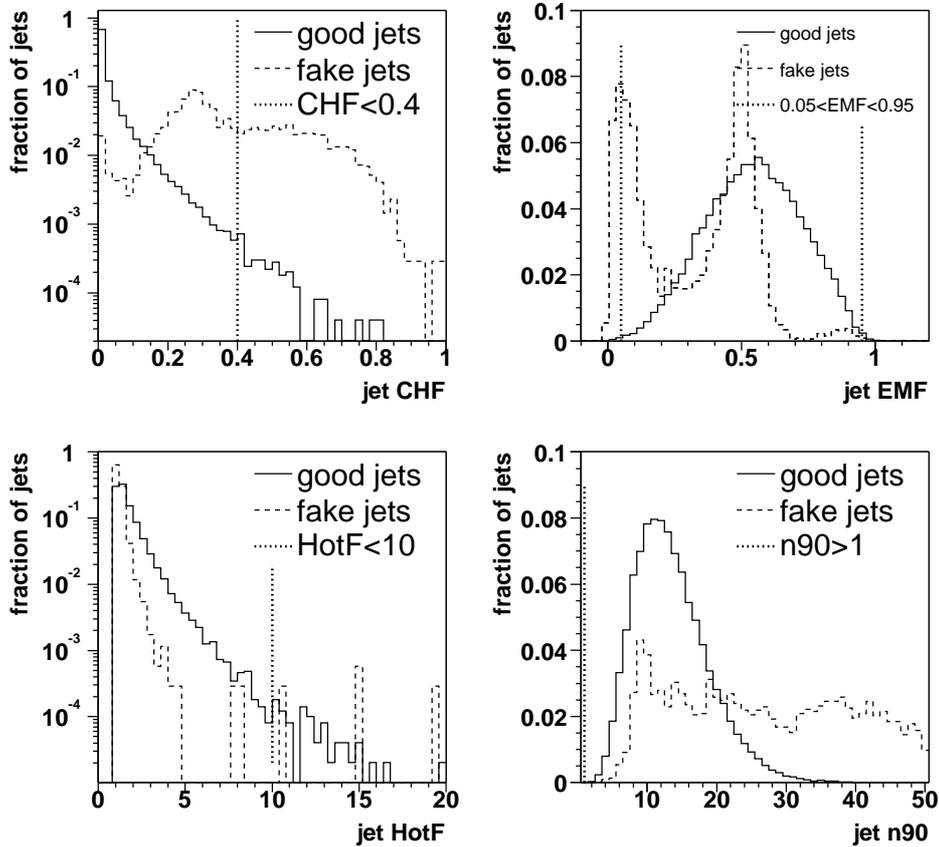


Figure 4.13: The distributions for central jets (CC) for the quantities which serve to identify good jets while rejecting fake jets:  $CHF$ ,  $EMF$ ,  $HotF$  and  $n90$  are plotted for good jets (dijet sample) and for fake jets (multijet sample). Also shown are the jet ID cut values.

identification cuts is the safety of the cuts, e.g. it was decided not to apply a upper cut on  $n90$ , since there might be interesting signals which produce jets in this  $n90$  region.

#### 4.5.3.1 L1 Jet Confirmation

Despite the numerous quality requirements applied to avoid clustering jets on noise in the calorimeter, a significant number of fake jets are selected. The application of the T42 algorithm reduces the number of fake jets originating from the discussed calorimeter noise roughly by a factor of 2 [131]; further

criteria must be applied to achieve the highest rejection of fake jets. In the 2003 analyses, this rejection was based on the  $f90$  variable [132], but this introduced a significant inefficiency on good jets (4%) [132].

Fake jets originating from noise in the calorimeter readout is found to have the largest contribution to the remaining fake jets. This contribution is not eliminated by the T42 algorithm. Instead an alternative calorimeter readout, the L1 trigger readout chain, is utilized to confirm the presence of good jets.

The comparison of the energy in the L1 compared to precision readout is found to be the most powerful discriminant against noise, which does not appear simultaneously in the two readout chains. For the trigger version v8 (Run  $\leq 172359$ ) the L1 calorimeter system was instrumented up to  $|\eta_{det}| \leq 2.4$ , for trigger versions 9 and higher up to  $|\eta_{det}| \leq 3.2$ , see also Sect. 3.3.7.1 and [133].

A new variable,  $L1SET$ , is defined for a given jet as the scalar sum of the trigger towers  $E_T$  inside the jet cone of  $R = 0.5$ . To quantify the agreement of the transverse energy measurement of the two readout chains and thus, to confirm the quality of the jet, the ratio of the two energy measurements is built:

$$\frac{L1SET}{E_T^{reco} \times (1 - CHF)}. \quad (4.3)$$

The  $L1SET$  measurement is not corrected for the JES and it does not include the coarse hadronic calorimeter section. In order to get the best correlation between the energy measured in the precision and L1 readout, the uncorrected jet  $E_T$  from the precision readout is used for the ratio,  $E_T^{reco}$ , subtracting the coarse hadronic energy fraction  $CHF$ .

The level of noise is found to be dependent of the  $|\eta_{det}|$  of the jet; three regions of pseudo-rapidity are defined which correspond to the geometry of the calorimeter, CC, ICD and EC, see Sect. 3.3.4.2 and Sect. 3.3.4.3: The CC region is defined as  $|\eta_{det}| < 0.8$ , the ICD region is defined as  $0.8 < |\eta_{det}| < 1.5$  and the EC region is defined as  $|\eta_{det}| > 1.5$ .

Figure 4.14 shows the distribution for  $\frac{L1SET}{E_T^{reco} \times (1 - CHF)}$  for good jets and for fake jets and the survival efficiencies for cutting on this variable. The cut values on  $\frac{L1SET}{E_T^{reco} \times (1 - CHF)}$  were chosen to have a high efficiency,  $> 99\%$  both in CC, ICD and EC and to reject the maximum amount of alleged fake jets (see Fig. 4.14); the cut values are summarized in Table 4.4.

The rejection is increased by a factor of 2 compared to the rejection obtained with the previous  $f90$  cut, when T42 is not applied, and by a larger factor (about 4) when T42 is applied, since then the  $f90$  cut is less efficient [131].

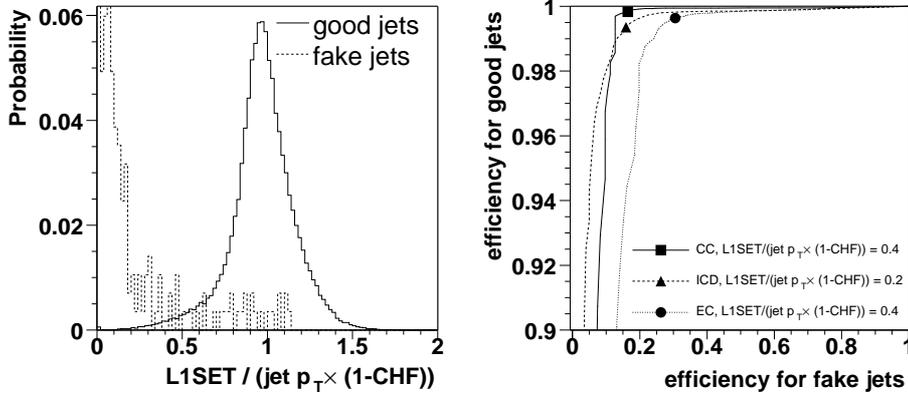


Figure 4.14: Left:  $\frac{L1SET}{E_T^{reco} \times (1-CHF)}$  distribution for good jets from the dijet sample and for fake jets from the multijet sample, integrated over the full pseudo-rapidity region. Right: The corresponding survival efficiency for the good vs. the fake jets varying the cut on  $\frac{L1SET}{E_T^{reco} \times (1-CHF)}$ . The CC, ICD and EC regions are defined in the text, the chosen cut values for the three  $\eta_{det}$  regions are highlighted by the solid points.

	CC	ICD	EC
$\frac{L1SET}{E_T^{reco} \times (1-CHF)}$	0.4	0.2	0.4

Table 4.4: Additional jet identification criterion: cut values on  $\frac{L1SET}{E_T^{reco} \times (1-CHF)}$  for the L1 jet confirmation.

#### 4.5.4 Separation of Jets from Electromagnetic Objects

Electrons and photons with transverse energy greater than 8 GeV are also reconstructed as jet objects in the calorimeter. This represents a problem of how to discriminate properly between real jets and electromagnetic objects, but on the other hand to apply the appropriate energy correction for the type of interaction the particles in the energy cluster undergo within the calorimeter (i.e. electromagnetic vs. nuclear), see Sect. 4.5.5 and Sect. 4.6. Reconstructed and identified (see Sect. 4.5.3) jet candidates are not considered as jets but as electron candidates if they overlap with an electromagnetic object ( $\Delta R(jet, EM) < 0.5$ ) with the following selection criteria:

- $p_T > 15.0$  GeV (after jet energy scale correction),
- $|\eta_{det}| < 2.5$ ,

- $EMF > 0.9$ ,
- $f_{\text{iso}} < 0.15$ ,
- $H\text{-matrix}7 < 50$ .

#### 4.5.5 Jet Energy Scale

The calorimeter is very effective at absorbing the hadronic energy of the jet. However there are several mechanisms which cause the energy of the cells clustered into a jet to deviate from the energy of the initial parton. The most important of these are:

**Calorimeter Response R:** hadronic showers may lose energy in ways which do not leave any ionization, electromagnetic and hadronic particles may therefore be imbalanced. Furthermore, the measured jet energy can be distorted due to a different response of the calorimeter to different particles, a non-linear response of the calorimeter to the particle energies, un-instrumented regions of the detector or dead material.

**Energy Offset O:** energy in the clustered cells which is due to the underlying event, multiple interactions, energy pile-up, electronics noise and noise from the Uranium absorber can provide an offset to the energy of the jet.

**Showering Corrections S:** a fraction of the jet energy is excluded due to the finite size of the cone used for clustering.

The Jet Energy Scale (JES) corrections attempt to correct the reconstructed jet energy,  $E^{\text{reco}}$ , back to the particle level energy,  $E^{\text{corr}}$ , as it would have been before interacting with the calorimeter. The correction may be written as

$$E^{\text{corr}} = \frac{E^{\text{reco}} - O}{R \times S},$$

where  $R$  is the calorimeter response to a jet,  $O$  is the energy offset and  $S$  is the fraction of shower leakage outside the jet cone in the calorimeter.

$R$  is determined by examining QCD Compton events. The energy of the photon is purely electromagnetic and its electromagnetic energy scale can be calibrated independently using  $Z \rightarrow ee$  events. The transverse energy of the jet should therefore balance the transverse energy of the photon.

$O$  is determined from energy densities in events which are triggered when a minimum activity in the luminosity monitor is reported (so-called minimum bias triggered events).

$S$  is determined from measured energy profiles of jets.

Since the MC may not model all these effects accurately, there can also be a difference in the scale of jets in data vs those in MC. In the current analysis the JetCorr v5.1 [134] package is used which incorporates the above mentioned components and provides corrections separately for data and MC, see Fig. 4.15 and Fig. 4.16.

The systematic uncertainty assigned to the JES correction comprises both the statistical and the systematic uncertainties on the data and the MC measurement, and can be expressed as:

$$\sigma = \sqrt{\sigma_{stat,data}^2 + \sigma_{syst,data}^2 + \sigma_{stat,MC}^2 + \sigma_{syst,MC}^2}. \quad (4.4)$$

The statistical and the total (quadratic sum of statistical and systematic) uncertainties are shown in Fig. 4.15 for the data and in Fig. 4.16 for the MC. The quadratic sum of the systematic uncertainties on the data  $\sigma_{syst,data}^2$  and the MC  $\sigma_{syst,MC}^2$  in Eq. 4.4 is known to lead to an overestimation of the systematic uncertainty, since the systematics are correlated for the most part among data and MC, and the systematics relevant to the cross section analyses presented arise only from the relative systematic uncertainty between data and MC. These correlations are currently unknown and therefore, conservatively, no correlations are assumed in Eq. 4.4. The correlations are in the process of being evaluated and will be available for the JetCorr v5.3 [134] package.

The JES corrections are expected to depend also on the flavor of the jet, in particular  $b$  jets may have a different response and also a different showering correction due to their harder fragmentation. For the analyses presented, only one additional class of jets is considered: jets containing a muon ( $\Delta R(\mu, jet) < 0.5$ ). The hypothesis is that the muon originates from a semileptonic  $b$  decay, which produces a neutrino along with the muon. As an approximation, it is assumed that the neutrino carries the same momentum as the muon and the jet is corrected in addition for these two particles, invisible to the calorimeter.

#### 4.5.6 Jet Energy Resolution

For high energy jets, above  $p_T \sim 50$  GeV, the jet energy resolutions are measured using the dijet event sample as defined in Sect. 4.5.3. However, no  $\cancel{E}_T$  cut is imposed and both jets pass all jet ID cuts. The details on the event selection can be found in Ref. [134]. The single jet trigger with lowest  $E_T$  threshold used, JT\_25TT\_NG, becomes fully efficient for a jet  $p_T$  above 40 GeV, therefore the dijet data allows to measure the jet energy resolutions only above this threshold.

The sample is split in several bins of average  $p_T$  of the dijet system,  $\langle p_T \rangle = (p_{T1} + p_{T2})/2$ , and for each  $\langle p_T \rangle$  bin the distribution of the transverse

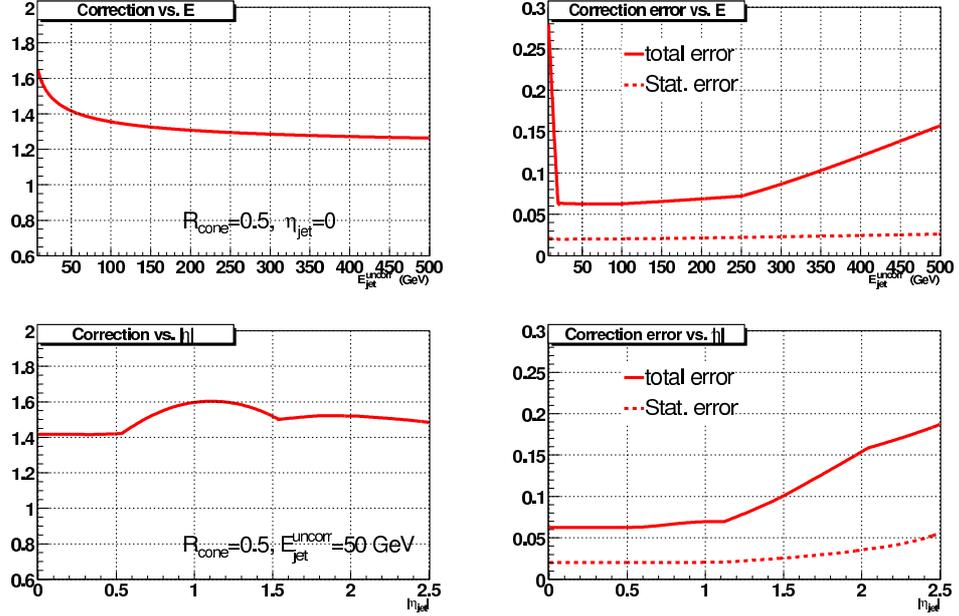


Figure 4.15: JES measurement in data. Left: JES correction as a function of uncorrected jet energy (top) and as a function of jet  $\eta_{det}$  (bottom). The respective statistical and total uncertainties are shown on the right [134].

momentum asymmetry variable

$$A = \frac{|p_{T1} - p_{T2}|}{p_{T1} + p_{T2}}, \quad (4.5)$$

is studied. The width of the  $A$  distribution,  $\sigma_A$ , obtained from a Gaussian fit with a mean value set to zero, gives the jet  $p_T$  resolution through the formula:

$$\frac{\sigma_{p_T}}{p_T} = \sqrt{2} \sigma_A. \quad (4.6)$$

In order to derive the resolution for the jet  $p_T$  range below 50 GeV, QCD Compton events are used as discussed in Sect. 4.5.3. However, no  $\cancel{E}_T$  cut is imposed and the jet passes all jet ID cuts. The events are triggered by a single EM trigger with no track requirement. In the case of QCD Compton events, the asymmetry variable is defined as:

$$A_{pj} = \frac{p_T^{jet} - p_T^\gamma}{p_T^\gamma}. \quad (4.7)$$

Given that the resolution of the photon is much better than the resolution of the hadronic jet,  $\sigma_{p_T^\gamma}$  can be ignored compared to  $\sigma_{p_T^{jet}}$ , and the jet resolution can be expressed as:

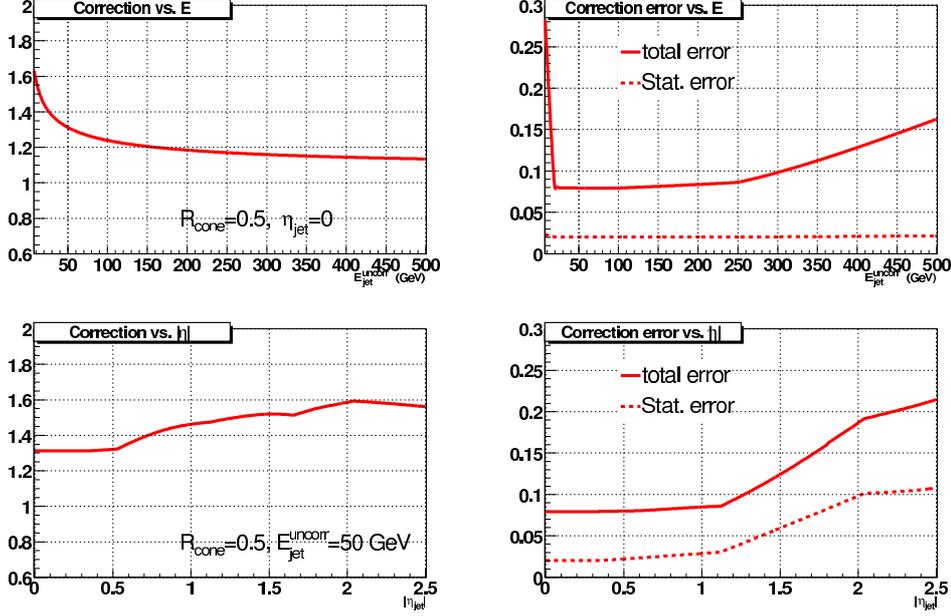


Figure 4.16: JES measurement in MC. Left: JES correction as a function of uncorrected jet energy (top) and as a function of jet  $\eta_{det}$  (bottom). The respective statistical and total uncertainties are shown on the right [134].

$$\frac{\sigma_{p_T^{jet}}}{p_T^{jet}} = \sigma_{A_{pj}} \times R_{pj} \quad (4.8)$$

where  $R_{pj} = p_T^\gamma / p_T^{jet}$  is a factor to correct the imbalance between the average jet  $p_T$  and the photon  $p_T$  in each  $p_T$  bin.  $R_{pj}$  is found to be compatible with unity given the systematic uncertainties on the JES.

The results obtained from the dijet and the QCD Compton samples are put together and fitted using the formula:

$$\frac{\sigma(p_T)}{p_T} = \sqrt{C^2 + \left(\frac{S}{\sqrt{p_T}}\right)^2 + \left(\frac{N}{p_T}\right)^2}; \quad (4.9)$$

see Eq. 3.10 and Sect. 3.3.4.4 for the meaning of the coefficients  $C$ ,  $S$  and  $N$ . The combined results are shown in Fig. 4.17, and Fig. 4.18, for the data and MC, respectively. The fit parameters are summarized in the Tables 4.5 and 4.6.

Jets in the simulation are corrected by applying an additional smearing according to a random Gaussian distribution with a width:

$$\sigma(p_T) = \sqrt{\sigma_{p_T}^{data}(p_T)^2 - \sigma_{p_T}^{MC}(p_T)^2} \quad (4.10)$$

$ \eta_{det} $ range	$N$	$S$	$C$
$0.0 <  \eta_{det}  < 0.5$	5.05	0.753	0.0893
$0.5 <  \eta_{det}  < 1.0$	0.	1.20	0.0870
$1.0 <  \eta_{det}  < 1.5$	2.24	0.924	0.135
$1.5 <  \eta_{det}  < 2.0$	6.42	0.	0.0974

Table 4.5: Jet energy resolution coefficients for data.

$ \eta_{det} $ range	$N$	$S$	$C$
$0.0 <  \eta_{det}  < 0.5$	4.26	0.658	0.0436
$0.5 <  \eta_{det}  < 1.0$	4.61	0.621	0.0578
$1.0 <  \eta_{det}  < 1.5$	3.08	0.816	0.0729
$1.5 <  \eta_{det}  < 2.0$	4.83	0.	0.0735

Table 4.6: Jet energy resolution coefficients for MC.

with the resolution found in the simulation,  $\sigma_{p_T}^{MC}(p_T)$ , and in the data,  $\sigma_{p_T}^{data}(p_T)$ .

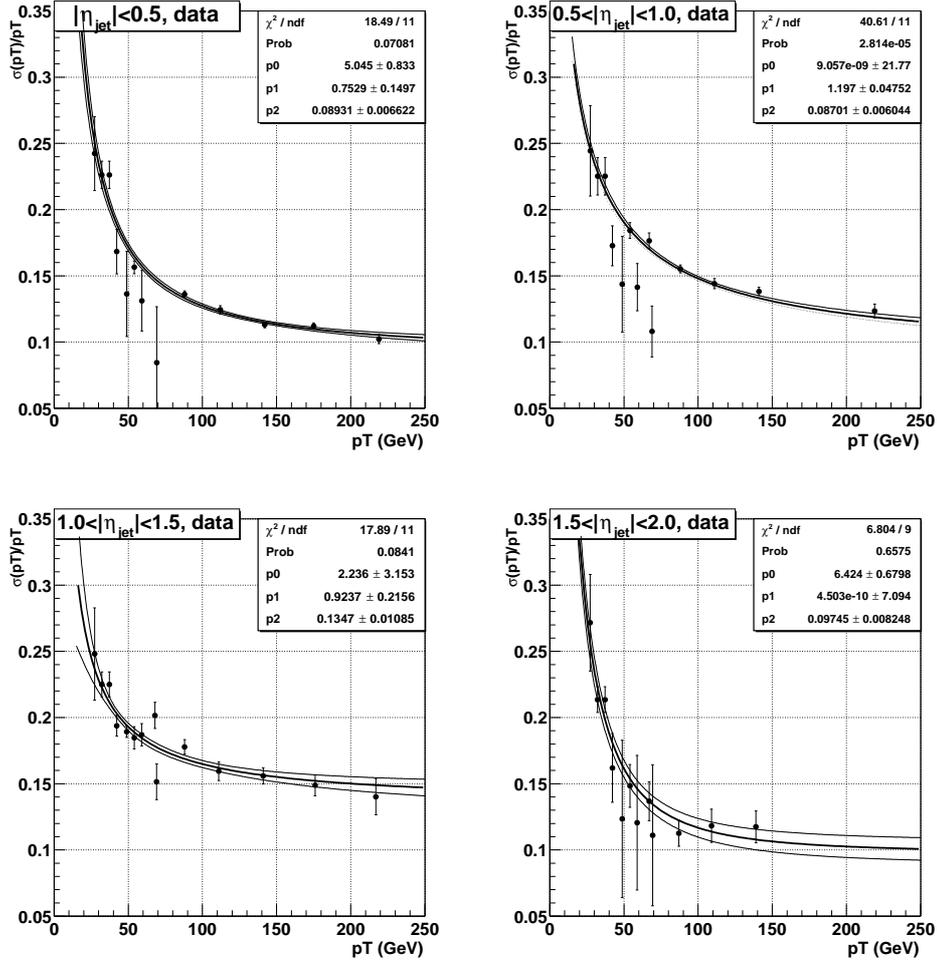


Figure 4.17: Jet  $p_T$  resolutions for different  $\eta_{det}$  regions in data. The points below  $\sim 50$  GeV are obtained using  $\gamma$ +jet events, whereas for  $p_T \gtrsim 50$  GeV resolutions are measured using dijet data. Bands of  $\pm 1\sigma$  statistical error are also shown.

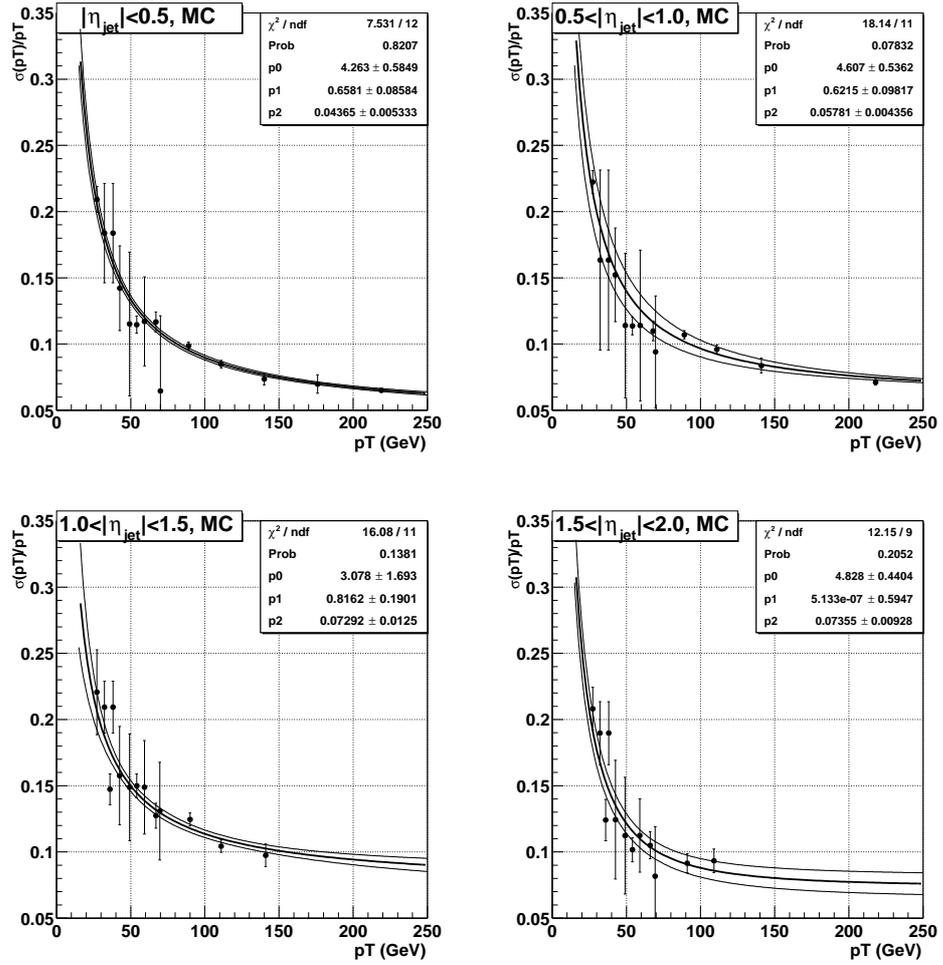


Figure 4.18: Jet  $p_T$  resolutions for different  $\eta_{det}$  regions in MC. The points below  $\sim 50$  GeV are obtained using  $\gamma$ +jet events, whereas for  $p_T \gtrsim 50$  GeV resolutions are measured using dijet data. Bands of  $\pm 1\sigma$  statistical error are also shown.

## 4.6 Missing $E_T$

The presence of a neutrino in the final state can be detected only from the imbalance of an event in the transverse plane. It is reconstructed from the vector sum of the transverse energies of all cells surviving the T42 algorithm, except for those in the coarse hadronic layer which are treated separately due to their high level of noise. The only cells of the coarse hadronic calorimeter which are used in the  $\cancel{E}_T$  sum are those clustered within good jets. The vector opposite to this total visible momentum vector is denoted the missing energy vector and its modulus is the raw missing transverse energy ( $\cancel{E}_{Traw}$ ).

The response of electromagnetic particles such as photons, electrons or  $\pi^0$ 's is different from that of hadrons and, in particular, from that of jets. In events with both electromagnetic objects and jets, this imbalance translates directly into missing transverse energy. As a JES correction is derived for all good jets, it can also be applied to the missing transverse energy. In order to do so, the JES correction (limited to the response part of such correction) applied to all good jets is subtracted from the  $\cancel{E}_T$  vector. In the same way the EM correction for electromagnetic objects ( $p_T > 15.0$  GeV,  $|\eta_{det}| < 2.5$ ,  $EMF > 0.9$ ,  $f_{iso} < 0.15$ ,  $H\text{-matrix7} < 50$ ) is applied to the  $\cancel{E}_T$  vector. The resulting modulus is denoted the calorimeter missing transverse energy ( $\cancel{E}_{TCAL}$ ).

As a muon is a minimum ionizing particle throughout the entire detector, it will deposit only a small amount of energy in the calorimeter. Its presence can thus also fake missing transverse energy in the calorimeter. The transverse momentum of all track matched muons (as measured in the central tracking detector) present in the event is subtracted from the missing transverse energy vector after deduction of the expected energy deposition of the muon in the calorimeter (taken from GEANT look-up tables).

### 4.6.1 $\cancel{E}_T$ Resolution

The  $\cancel{E}_T$  resolution is studied in events where no  $\cancel{E}_T$  is expected.  $Z$ -plus-jets events with a muonic  $Z$  decay can be selected with a high efficiency and purity and are expected to have no  $\cancel{E}_T$ . The  $\cancel{E}_T$  resolution in data is found to be worse than in MC as shown in Fig. 4.19 for events with  $\geq 0$  and  $\geq 2$  jets. The Monte Carlo is further smeared in order to bring it in agreement with data, as described in [135]. A  $\cancel{E}_T$  oversmearing parameter is defined by calculating the difference in quadrature of the  $\cancel{E}_T$  resolutions in data and Monte Carlo and then fitting this vs  $\Sigma E_{Tunclus}$ , the event scalar  $E_T$  (clustered and unclustered) minus the  $E_T$ 's of all the reconstructed (clustered) objects. No jet multiplicity dependence is observed. The  $\cancel{E}_T$  oversmearing correction is realized by the following transformation of  $\Sigma E_{Tunclus}$ :

$$\Sigma E_{Tunclus} \rightarrow \Sigma E_{Tunclus} + G(0, \sigma_{E_T}(\Sigma E_{Tunclus})), \quad (4.11)$$

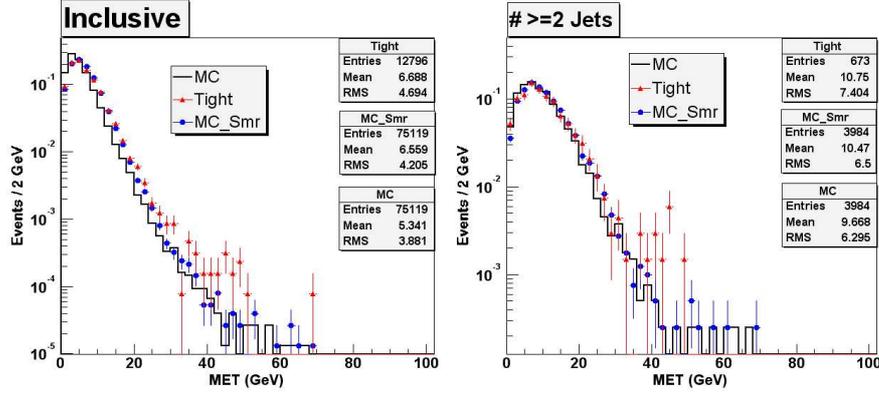


Figure 4.19:  $\cancel{E}_T$  resolution in  $Z \rightarrow \mu\mu$  events for data (Tight), MC and smeared MC with  $\geq 0$  jets (left) and  $\geq 2$  jets (right) [135].

where  $G(0, \sigma_{E_T}(\Sigma E_{Tunclus}))$  is a random variable drawn from a Gaussian distribution with a mean of 0 and a width  $\sigma_{E_T}(\Sigma E_{Tunclus})$ , parameterized as

$$\sigma_{E_T}(\Sigma E_{Tunclus}) = 2.553 + 0.008951 \times \Sigma E_{Tunclus}.$$

The resulting oversmeared  $\cancel{E}_T$  is also shown in Fig. 4.19. The smearing of the Monte Carlo improves the agreement with data particularly in the core of the  $\cancel{E}_T$  distribution. However, the effect of the oversmearing is most pronounced for events with few jets in the event (left plot in Fig. 4.19), the effect is much less pronounced for events with  $\geq 2$  jets in the event (right plot in Fig. 4.19) and is expected to be even less pronounced for events with  $\geq 4$  jets, the relevant region for the  $t\bar{t}$  signal events. For this reason no  $\cancel{E}_T$  oversmearing is applied.

## 4.7 $b$ Jets

Reconstructed and identified jets can be further classified by their flavor, depending if they originate from the decay and hadronization of a light flavor quark or gluon ( $u, d, s, g$ ), a  $c$  quark, or a  $b$  quark. The corresponding jets are called light flavor jets,  $c$  jets or  $b$  jets. At least two techniques can be used to distinguish a heavy flavor jet ( $c$  or  $b$  jets) from a light flavor jet:

- **Soft Lepton Tagging (SLT):** the presence of a soft electron or muon within the jet cone indicates a semileptonic  $b$  or  $c$  hadron decay with a branching ratio of typically  $\sim 10\%$  per lepton.
- **Lifetime tagging:** identifying charged tracks which are significantly displaced from the primary vertex due to the finite lifetime of the  $b$  or  $c$  hadron decay.

The probability to identify a jet as a  $b$  jet using lifetime tagging is broken down into two components:

- The probability for a jet to be “taggable” (also called “taggability”).
- The probability for a taggable jet to be tagged (also called “tagging efficiency”).

Three lifetime tagging algorithms exist at DØ for the identification of  $b$  jets:

- **Counting Signed Impact Parameter (CSIP):** require a minimum number of tracks with large impact parameter significance with respect to the primary vertex.
- **Jet Lifetime Probability (JLIP):** for each calorimeter jet, a probability is computed that it does not originate from the PV using the impact parameter information of tracks seen in the SMT layers.
- **Secondary Vertex Tagger (SVT):** explicit reconstruction of a secondary vertex with a large decay length significance with respect to the primary vertex [136].

In the analyses presented, the SVT algorithm is used to identify  $b$  jets.

#### 4.7.1 Taggability

The taggability serves to decouple the tagging efficiency from issues related to tracking inefficiencies and calorimeter noise problems, which are therefore absorbed into the taggability. According to the  $b$ -ID group definition [137], a calorimeter jet is taggable if it is matched within  $\Delta R \leq 0.5$  to a track-based jet (track-jet).

##### 4.7.1.1 Track-Jets

Jets consist on average to 2/3 of charged particles which can be reconstructed as tracks in the central tracking system. The algorithm to reconstruct these track-jets from charged tracks can be subdivided in three steps:

1.  **$z$  pre-clustering:** cluster tracks according to their  $z$  of closest approach with respect to  $z = 0$ . Looping in descendant order of track  $p_T$ , tracks are added to the pre-cluster if  $\Delta z < 2$  cm, where  $\Delta z$  is the difference between the  $z$  of the closest approach of the track and the pre-cluster.
2. **Track selection:** for every pre-cluster, identify the closest reconstructed primary vertex and select tracks satisfying the following criteria:  $p_T > 0.5$  GeV,  $\geq 1$  SMT hits,  $|dca| < 0.20$  cm and  $|zdca| < 0.4$  cm.

3. **Jet clustering:** for every pre-cluster, the selected tracks are clustered in the  $(\eta, \phi)$ -plane using the simple cone jet algorithm with seed  $p_T > 1$  GeV, requiring at least two tracks.

#### 4.7.2 Secondary Vertex Tagger Algorithm

The secondary vertex algorithm consists of three main steps: reconstruction and identification of the primary interaction vertex (see Sect. 4.2), reconstruction of track-jets (see Sect. 4.7.1.1), and secondary vertex finding, a sketch is shown in Fig. 4.20.

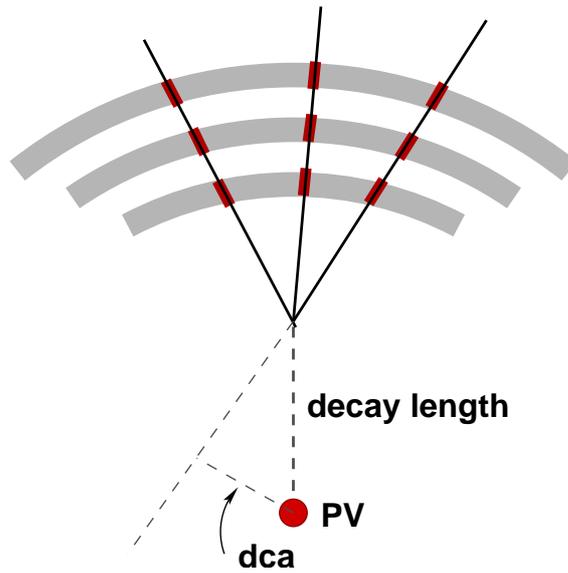


Figure 4.20: The secondary vertex algorithm explicitly reconstructs vertices inside of jets. The decay length significance cut is tuned to optimize the performance of the algorithm.

The secondary vertex finding is performed for every track-jet. The different steps involved are:

1. **Track selection:** only tracks with  $p_T > 1.0$  GeV and (signed) impact parameter significance  $(|dca|/\sigma(dca)) > 3.5$  are considered. The sign of the impact parameter is given by the sign of the projection of the impact parameter onto the track-jet axis, see Fig. 4.21 for a schematic drawing.  $dca$  and  $\sigma(dca)$  are computed with respect to the selected primary vertex. The signed impact parameter significance distribution is also shown in Fig. 4.21.
2. **Vertex finding and fitting:** for every track-jet with at least two selected tracks, one tries to find a secondary vertex from those tracks

by using the *Build-Up* algorithm.

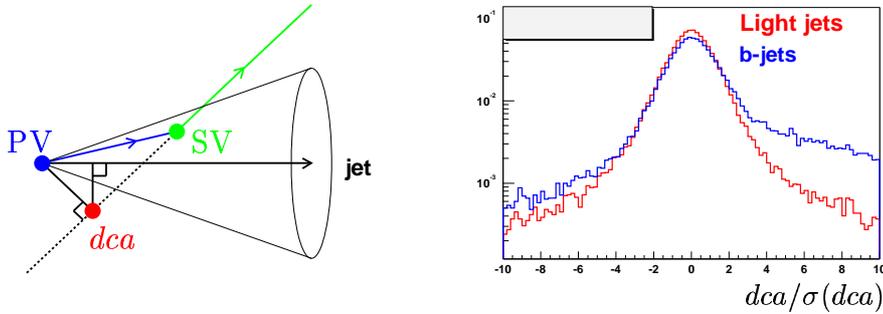


Figure 4.21: Left: A schematic drawing showing the definition of the impact parameter ( $dca$ ), its sign is given by the sign of the inner product of  $\vec{dca}$  and the track-jet momentum. Right: The impact parameter significance distributions for tracks from light and  $b$  jets from Monte Carlo. One can observe a clear excess at positive impact parameter significance for  $b$  jets.

- Build-Up algorithm:** The *Build-Up* algorithm starts finding *seed* vertices by fitting all combinations of pairs of selected tracks in track-jets. Then, it attempts to attach additional tracks to the *seed* vertices according to the resulting  $\chi^2$  contribution to the vertex. The process is repeated until no more tracks can be associated to seeds. This procedure is such that the resulting vertices might share tracks.
3. **Vertex selection:** the secondary vertices found are required to have track multiplicity  $\geq 2$ ,  $|L_{xy}| < 2.6$  cm,  $|L_{xy}/\sigma(L_{xy})| > 7.0$ ,  $|\text{collinearity}| > 0.9$  and  $\chi^2/dof < 10$ . The vertex transverse decay length is defined as  $|L_{xy}| = |\vec{r}_{SV} - \vec{r}_{PV}|$ . The collinearity is defined as the inner product of  $\vec{L}_{xy}$  and the vertex momentum, computed as the vector sum of the momenta of all attached tracks after the constrained fit. The sign of the transverse decay length is given by the sign of the collinearity.
  4.  **$V^0$  removal:** secondary vertices composed of two tracks with opposite sign are required to be inconsistent with a  $V^0$  hypothesis. The hypotheses tested by the algorithm include  $K_s^0 \rightarrow \pi^+\pi^-$ ,  $\Lambda^0 \rightarrow p^+\pi^-$  (the higher  $p_T$  track is assumed to be the proton, according to the observation in the MC) and the photon conversion ( $\gamma \rightarrow e^+e^-$ ). Secondary vertices are rejected if the invariant di-track mass, given the  $V^0$  hypothesis, is consistent with the corresponding  $V^0$  mass in a mass window defined by  $\pm 3\sigma$  of the measured  $V^0$  mass resolution, shown in Fig. 3.11 for the  $K_s^0$  and the  $\Lambda^0$ .

A calorimeter jet is identified as a  $b$  jet (also called “tagged” or “having a  $b$ -tag”) if it is taggable and it contains a selected secondary vertex with  $L_{xy} > 0$  within  $\Delta R < 0.5$ . Events containing one or more tagged jets are also called “tagged events”.

# Chapter 5

## Event Trigger

Collisions occur every 396 ns, i.e. at a rate far beyond the capabilities of the data recording or analysis structures. As discussed in Sect. 3.3.7, a trigger system selects the events of interest. The  $t\bar{t}$  signature of the  $\mu$ -plus-jets events is one muon, one neutrino and  $\geq 4$  jets as discussed in Sect. 2.3. To ensure that the relevant data are recorded, a trigger to select both, a muon and a jet, was designed by balancing efficiency and simplicity with the need for rejection.

The data sample analyzed was collected using five different trigger list versions: v8 to v12 (a trigger list version is defined for a certain run range and consists of a set of well defined triggers). A summary of the trigger names and conditions is given in Table 5.1. The trigger conditions are:

### Level 1:

- `mulptxatxx`: Level 1 tight muon scintillator coincidence,
- `CJT(1,X)`: Calorimeter trigger tower with uncorrected  $E_T > X$  GeV total energy,  $X = 5$  GeV for v8.2-v11, and  $X = 3$  GeV for v12,  $|\eta_{det}| < 2.4$  for v8.2-v10, and  $|\eta_{det}| < 3.2$  for v11-v12,

### Level 2:

- `MUON(1,med)`: Level 2 medium muon as described in [138],
- `JET(1,10)`: Level 2 jet with uncorrected  $E_T > 10$  GeV,

### Level 3:

- `JET(1,X)`: Level 3 jet with uncorrected  $E_T > X$  GeV,  $X = 20$  GeV for v8.2-v11, and  $X = 25$  GeV for v12.

Trigger List	v8.2-v11	v12
Trigger Name	MU_JT20_L2M0	MU_JT25_L2M0
L1	mu1ptxatxx_CJT(1,5)	mu1ptxatxx_CJT(1,3)
L2	MUON(1,med)	MUON(1,med)_JET(1,10)
L3	JET(1,20)	JET(1,25)

Table 5.1: Summary of triggers used in trigger lists v8 to v12.

## 5.1 Trigger Efficiency Measurement

The trigger efficiency can be measured in two different ways:

- simulate the trigger requirements on Monte Carlo simulated events, using the program `TrigSim`,
- measure single object efficiencies of satisfying individual trigger conditions at Level 1, Level 2 and Level 3 in the data, parameterize the efficiencies as a function of the relevant kinematic variables and fold the parameterizations into the Monte Carlo events.

Although correlations and overlap between triggers are automatically taken into account using the first method, currently, the Monte Carlo modeling of trigger objects and trigger quantities is not adequate to be used for precision measurements of the trigger efficiency. Therefore the second method based on trigger efficiencies derived from data is chosen.

The approach used to combine single object trigger efficiencies to calculate the probability of an event to satisfy a specific trigger is described in [139] and briefly summarized in the following.

The total event probability ( $P(L1, L2, L3)$ ) is calculated as the product of the probabilities for the event to satisfy the trigger conditions at each triggering level,

$$P(L1, L2, L3) = P(L1) \cdot P(L2|L1) \cdot P(L3|L1, L2) \quad (5.1)$$

where  $P(L2|L1)$  and  $P(L3|L1, L2)$  represent the conditional probability for an event to satisfy a set of criteria given it has already passed the requirements imposed at the previous triggering level(s).

The total probability of an event to satisfy a set of trigger requirements is obtained assuming that the probability for a single object to satisfy a specific trigger condition is independent of the presence of other objects in the event. Under this assumption, the contributions from different types of objects to the total event probability can be factored out such that

$$P_{(object_1 \& object_2)} = P_{object_1} \cdot P_{object_2}. \quad (5.2)$$

Furthermore, under this assumption, the probability ( $P$ ) for at least one object to satisfy a particular trigger condition, out of a total of  $N$  objects present in an event, is given by

$$P = 1 - \prod_{i=1}^N (1 - P_i), \quad (5.3)$$

where  $P_i$  represents the single object probability (the probability for an electron to fire a jet trigger and vice versa is also considered in the corresponding product).

The total trigger efficiency is calculated as the luminosity weighted average of the event probability associated to the trigger requirements contained in each individual trigger list.

The probability of a single object to satisfy a particular trigger requirement is measured using the following general procedure [140]. The first step consists of identifying a sample of events, unbiased with respect to the trigger requirement under study. Offline reconstructed objects are then identified in the events. The efficiency is obtained by calculating the fraction of these offline reconstructed objects that satisfy the trigger condition under study. Single object efficiencies are in general parameterized as a function of the kinematic variables  $p_T$ ,  $\eta$  and  $\phi$  of the offline reconstructed objects.

In order to take into account major changes in the trigger system and thus in the trigger response, many single object trigger efficiencies are measured separately for the different trigger lists v8 to v12. Data recorded using different trigger lists, for which no changes to a particular subdetector and associated trigger system occurred, are combined.

### 5.1.1 Muon Trigger

The Level 1 and Level 2 trigger conditions and implementation have not changed in the period of data taking using the trigger list v8 to v12. The same efficiency parameterizations are therefore used for the trigger efficiency calculations in all trigger list versions.

The highest statistics event sample used to obtain the muon trigger efficiencies consists of  $Z \rightarrow \mu\mu$  events. The “tag-and-probe” method, see Sect. 4.5.3, is then used to calculate the fraction of offline muons that pass the trigger requirement under study. Events triggered by one of the single muon triggers in each of the trigger list versions considered are further selected by requiring the presence of two offline muons. Muons are identified in this sample of events using the offline selection criteria (see Sect. 4.3). The invariant mass of the two offline muons is required to be within a small window around the  $Z$  mass:  $80 \text{ GeV} < m_{\mu\mu} < 100 \text{ GeV}$ . One muon is randomly chosen (“tag”) and required to satisfy the L1 single muon trigger requirement. The second offline muon (“probe”) is then used to calculate the efficiency of a particular trigger criteria.

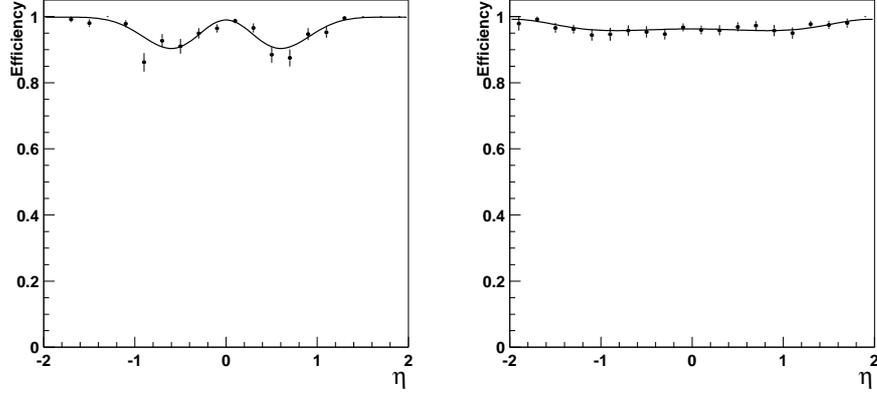


Figure 5.1: Muon trigger efficiency as a function of the offline muon  $\eta$  for the Level 1 condition `mulptxatxx` (left) and for the Level 2 condition `MUON(1,med)` given that the muon already satisfies the Level 1 condition (right).

		A0	A1	A2	A3
L1	<code>mulptxatxx</code>	$-0.8 \pm 0.2$	$2.8 \pm 0.5$	$0.1 \pm 0.1$	$0.99 \pm 0.01$
L2	<code>MUON(1,med)</code>	$0 \pm 0$	$8.7 \pm 7.0$	$1.779 \pm 0.006$	$0.981 \pm 0.007$

Table 5.2: Fit results for the Level 1 and Level 2 muon trigger efficiencies for data recorded with trigger list v8 to v12. The function used to fit the  $\eta$  dependence is given in Eq. 5.4.

Figure 5.1 shows the probability of an offline muon to satisfy the Level 1 requirement `mulptxatxx` (left) and the Level 2 requirement `MUON(1,med)` given it has fired the Level 1 condition (right) as a function of the offline muon  $\eta$ . For a muon with  $p_T > 15$  GeV, the Level 1 and Level 2 trigger efficiencies are found to be constant as a function of  $p_T$ . The function used to parameterize the  $\eta$  dependence is given by

$$f(\eta) = A3 + A0 \cdot \exp(-A1 \cdot (\eta^2 - A2^2)) \sin(\eta^2 - A2^2). \quad (5.4)$$

Results of the fit are summarized in Table 5.2.

In order to check for any possible geometrical biases associated to the use of the “tag-and-probe” method (see Sect. 4.5.3), both Level 1 and Level 2 muon trigger efficiencies are also calculated using a sample of  $W \rightarrow \mu\nu$  events. This sample of events is selected using the following criteria

- require at least one isolated medium muon with a track matched and  $p_T > 20$  GeV.
- $\Delta R(\mu, jet) > 0.5$
- Primary vertex reconstructed with at least 3 tracks,  $|z_{PV}| < 60$  cm
- $|dca|/\sigma_{dca} < 3$
- $\cancel{E}_T > 20$  GeV

Both the Level 1 and the Level 2 muon trigger efficiencies are found to be compatible with those obtained using  $Z \rightarrow \mu\mu$  events, given the associated uncertainties, which vary between 1 and 5 %.

### 5.1.2 Jet Trigger

The jet trigger efficiencies are parameterized as a function of the JES corrected jet  $p_T$  in three regions of the calorimeter: CC ( $|\eta_{det}| < 0.8$ ), ICR ( $0.8 \leq |\eta_{det}| < 1.5$ ) and EC ( $|\eta_{det}| \geq 1.5$ ).

In order to check for the possible presence of biases associated with a particular event selection, two different methods are used to calculate the jet trigger efficiencies. The first approach consists of measuring the jet trigger efficiency on a sample of events that pass at least one of the many muon triggers present in the trigger list version of interest. This method has the advantage of not depending on the calorimeter trigger response. The second approach consists of measuring the jet trigger efficiencies using a sample of events that satisfies any one of the single electron triggers present in the different trigger lists considered. In order to reduce the contamination of electrons and therefore obtain a pure sample of jets, only events that contain exactly one offline reconstructed electron, matching both a Level 1 and Level 2 electron in the event, are considered. Both methods are found to give compatible results, given the associated uncertainties.

The assumption that the probability for a single object to satisfy a specific trigger condition is independent of the presence of other objects in the event is verified for the jet triggers. The trigger efficiency is derived for events with 2, 3 and  $\geq 4$  jets in the event; no dependence on the jet topology is found. The L1, L2 and L3 trigger efficiencies for the dijet sample are shown in Fig. 5.2 for various  $|\eta_{det}|$  regions and trigger list versions. The function used to parameterize the jet trigger efficiency for L1, L2 and L3 is

$$f(p_T) = 0.5 \cdot A2 \cdot \left( 1 + \frac{2}{\sqrt{\pi}} \cdot \int_0^{\frac{p_T - A0}{\sqrt{p_T} \cdot A1}} \exp(-t^2) dt \right). \quad (5.5)$$

Jets containing muons are also corrected for the energy invisible to the calorimeter, as discussed in Sect. 4.5.5. However, using the non-muonic JES

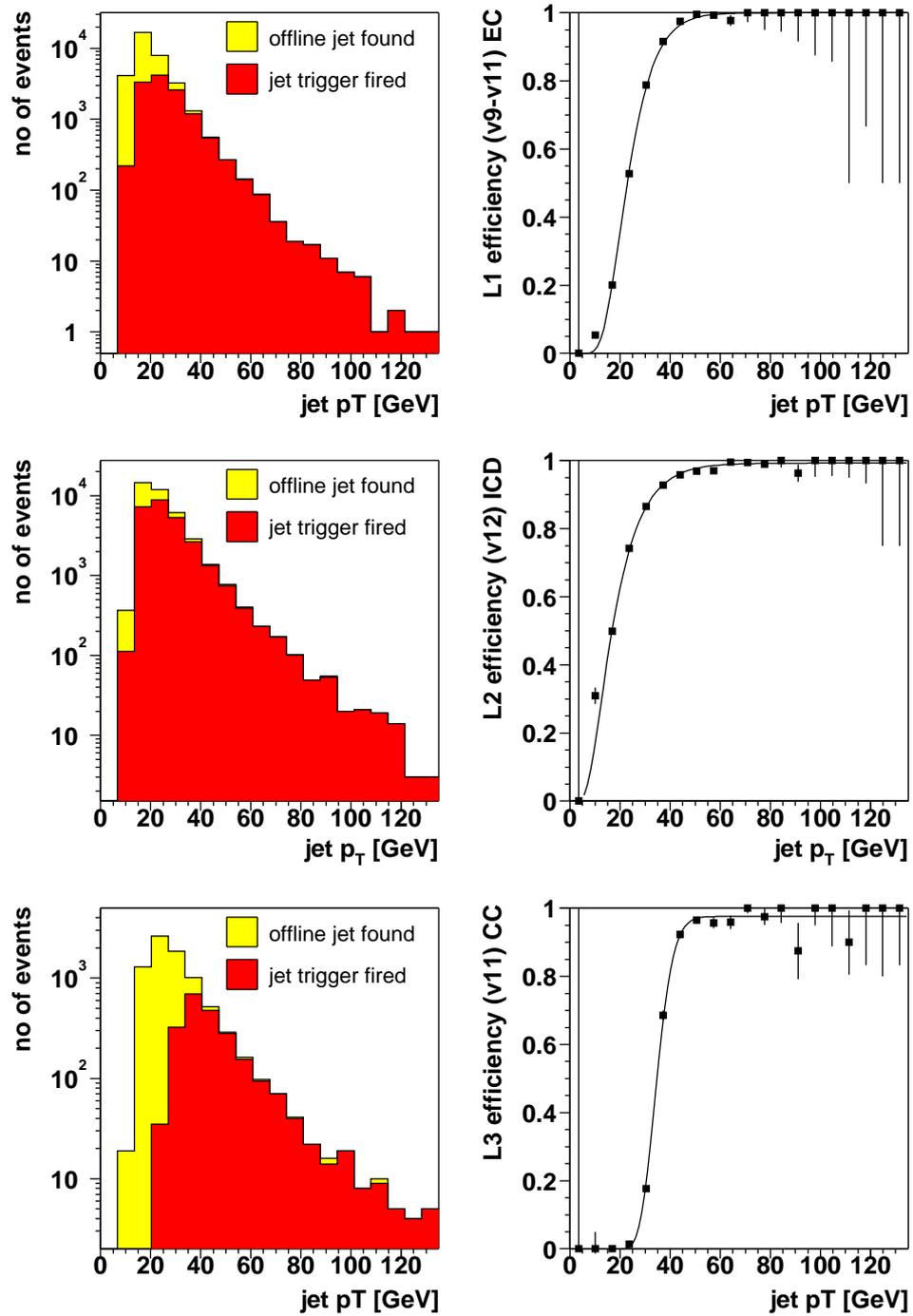


Figure 5.2: Jet trigger efficiency as a function of offline jet  $p_T$  for the Level 1 condition CJT(1,5) (top), for the Level 2 condition JET(1,10) given that the jet already satisfies the Level 1 condition (middle), and the Level 3 condition JET(1,20) given that the jet already satisfies the Level 1 and the Level 2 conditions (bottom).

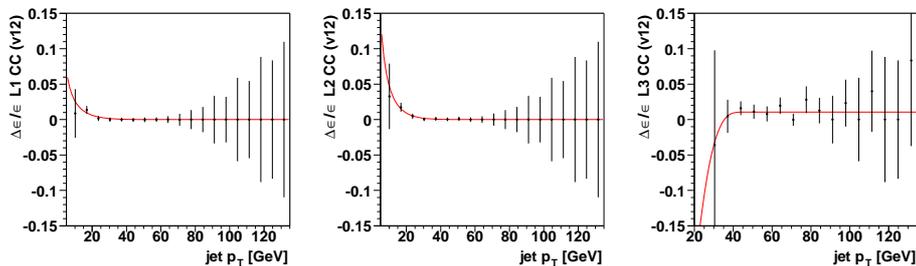


Figure 5.3: Relative difference in the jet trigger efficiency between the two measurements for jets with and without track-match as a function of offline jet  $p_T$  for the Level 1 condition CJT(1,5) (left), for the Level 2 condition JET(1,10) given that the jet already satisfies the Level 1 condition (middle), and the Level 3 condition JET(1,20) given that the jet already satisfies the Level 1 and the Level 2 conditions (right); for trigger version 12, and jets in the CC region.

correction (i.e. not correcting for the energy invisible to the calorimeter) for the jets containing muons, no difference is found in the resulting jet trigger efficiency turn-on curves for muonic and non-muonic jets.

The sample selected with a muon trigger and even more the jets containing a muon are enhanced in heavy flavor. The fact that the trigger efficiency is independent of the sample is a cross check that the trigger efficiency is independent of the jet flavor.

### 5.1.2.1 Systematic Uncertainties

The jet turn-on curves are measured with a large data sample which results in a small statistical uncertainty on the trigger efficiency measurements. Possible systematic effects associated to the method used for measuring the trigger efficiencies could, however, be of the same size or even larger than the statistical uncertainties. In order to quantify possible effects due to the jet quality used in measuring the efficiencies, all the trigger turn-on curves are remeasured requiring, this time, a track-match for every jet and thereby improving the purity of the jet samples, see Fig. 5.3. The relative systematic uncertainties derived from this method go up to approximately 10 % for low  $p_T$  jets and decrease to a few % for high  $p_T$  jets.

### 5.1.3 Event Trigger Efficiency

The probability of an event to satisfy the  $\mu$ +jets triggers is obtained based on the assumptions made above:

$$P_{\mu+\text{jets}} = P_{\text{MU}} \cdot P_{\text{JET}} \quad (5.6)$$

where  $P_{\text{MU}}$  ( $P_{\text{JET}}$ ) represents the probability for the event to satisfy the muon-type (jet-type) conditions. These probabilities can be further subdivided into individual trigger requirements such that,

$$\begin{aligned}
 \text{v8 - v11 : } P_{\text{MU}} &= P_{\text{mu1ptxatxx}} \cdot P_{\text{MUON}(1,\text{med})} \\
 P_{\text{JET}} &= P_{\text{CJT}(1,5)} \cdot P_{\text{JET}(1,20)} \\
 \text{v12 : } P_{\text{MU}} &= P_{\text{mu1ptxatxx}} \cdot P_{\text{MUON}(1,\text{med})} \\
 P_{\text{JET}} &= P_{\text{CJT}(1,3)} \cdot P_{\text{JET}(1,10)} \cdot P_{\text{JET}(1,25)}
 \end{aligned}$$

More details about the triggering can be found in Ref. [141].

## Chapter 6

# Data Sample

The two analyses presented (*topo* and *btag*) make use of slightly different data samples. The *btag* data sample, recorded between August 2002 and September 2003, is a subsample of the data used for the *topo* analysis, recorded between August 2002 and March 2004.

The data samples are derived from the 1MUloose CSG skim [142] and further skimmed requiring one ‘loose’ muon with  $p_T > 15.0$  GeV and  $\Delta R(\mu, jet) > 0.5$  and one jet with  $p_T > 15.0$  GeV.

In the following the data quality selection (Sect. 6.1) and the luminosity of the data sample are discussed (Sect. 6.2).

### 6.1 Data Quality Selection

Of the data delivered by the Tevatron, only a fraction is recorded by the DØ detector. Out of this data only the fraction for which all detector systems are functioning well is used for the data analyses. Finally, only the fraction of data which is reconstructed by the reconstruction software package *reco* is actually used for the data analyses presented. This breakdown of data, quantified in terms of luminosity is presented in Fig. 6.1. Typically, these analyses use about 80 % of the delivered luminosity.

The data quality monitoring is performed on two levels, online and offline. It is crucial for a high data taking efficiency to catch the malfunction of detector components, of the readout or the triggering as early as possible. Online data quality monitoring guarantees to be able to react immediately and thus to maximize the data quality online.

However, there are data quality issues which are not recognized online. The remaining deficient data is eliminated by offline data quality monitoring, as initially introduced by [143]. The idea is to compare basic distributions of physics objects or other variables that describe the detector performance of the data taken with canonical distributions. In a second step a list of possible data corruption methods is identified, quality measures are defined

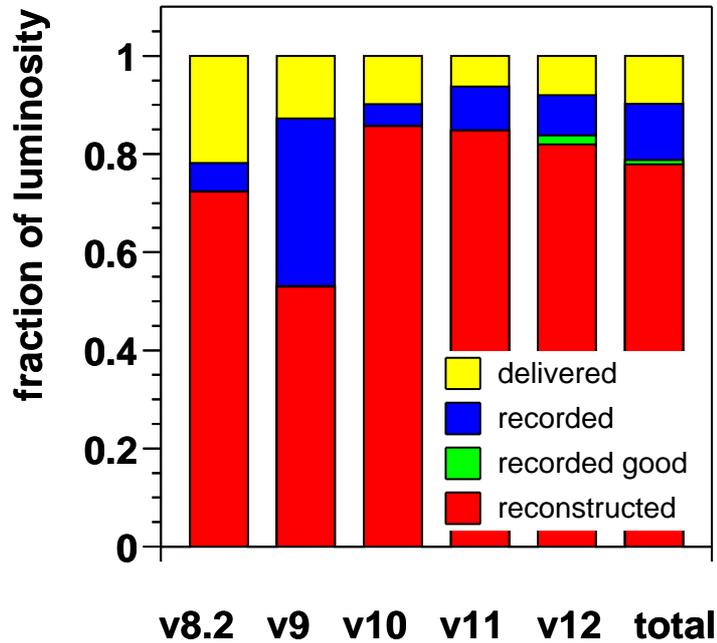


Figure 6.1: Fraction of luminosity that is delivered, recorded, passing the data quality criteria, and reconstructed as a function of the trigger list version (for the *topo* analysis.).

with a maximum discrimination power between usable and unusable data and criteria are defined to classify the data quality. Since the corruption of the data can occur on time scales much smaller than the length of a whole run, the classification is preferably done on the smallest possible units of data.

The data quality selection required for the data events analyzed is broken down in a run based, a luminosity block number (LBN) based, corresponding to approximately one minute of data taking and several thousand events recorded, and an event-by-event based selection.

**Run quality selection:** The good run selection is based on information stored in DØ's Run Quality Database [144]. The requirements are summarized in Table 6.1. It is guaranteed that no hardware failures are known and that all readout crates are included in the readout for the muon system [145], the SMT, the CFT and the calorimeter. The tracking quality is particularly important for the identification of leptons originating from the PV, it is crucial for the *b*-tagging, and tracking is used to build the kinematic variables.

**Luminosity block quality selection:**

- **Ring of Fire:** A grounding problem in the calorimeter, for the most part resolved, can cause a  $\phi$ -ring of energy. This leads to a large missing energy signature. Events which show the characteristic pattern for the *ring-of-fire* activity are eliminated from the data sample. The selection is based on the luminosity block number (LBN).
- **Bad JET/MET LBN List:** The following requirements on the average  $\cancel{E}_T$  and average scalar  $E_T$  are made:
  - shift in average  $\cancel{E}_T$ :  $\sqrt{\langle \cancel{E}_{Tx} \rangle^2 + \langle \cancel{E}_{Ty} \rangle^2} < 6$  GeV,
  - average RMS of  $\cancel{E}_T$ :  $\sqrt{RMS(\cancel{E}_{Tx})^2 + RMS(\cancel{E}_{Ty})^2} < 20$  GeV,
  - average scalar  $E_T$ :  $\langle \text{scalar } E_T \rangle > 60$  GeV.

Files consist of groups of approximately 20 consecutive LBN's. To have a sufficient number of events the data quality selection is based on these files.

**Event quality selection:** The selection of  $t\bar{t}$  events involves requirements of jets and  $\cancel{E}_T$ , both of which are highly susceptible to backgrounds from calorimeter readout malfunctions. Although the signatures for these problems tend to be fairly obvious when looking at event displays, care must be taken in devising an event-wide quality variable which is immune to variations in an event's real (as opposed to instrumentally originating) scalar  $E_T$  or number of jets. This is particularly true for  $t\bar{t}$  events in which large scalar transverse energies can arise. A study to arrive at such an event quality selection is described in [146]. If an event shows significant differences in energy deposition in the precision readout and the Level 1 readout then the event is suspect. A similar quantity is already used for the jet identification, see Sect. 4.5.3. In order to be immune against events with many jets or large energy scalar transverse energy the comparison between the L1 readout and the precision readout is carried out only for trigger towers with Level 1 energy of less than 2 GeV.

An event is rejected if it does not pass the requirement on the Level 1 and the precision readout information [146], and if it is flagged as *coherent noise*, a flag based on precision readout occupancy and RMS within ADC cards [147].

## 6.2 Integrated Luminosity of the Data Samples

Only good luminosity blocks, i.e. luminosity blocks for which the luminosity can be calculated, are used in the analyses. Table 6.2 summarizes the integrated luminosity for the different trigger lists used for the *topo* and the

MUON	Reasonable
SMT	Not Bad
CFT	Not Bad
CAL	Not Bad

Table 6.1: Run quality requirements.

*btag* analyses. The measurement of the luminosity is discussed in Sect. 2.5, the total uncertainty on the integrated luminosity is 6.5% [85].

Trigger List	<i>topo</i>	<i>btag</i> $\mu$ +jets	<i>btag</i> $e$ +jets
$\geq$ v8.2	21.3	21.6	21.0
v9	21.4	22.0	31.7
v10	15.7	16.3	15.9
v11	58.3	60.5	58.0
v12	112.4	38.1	42.2
total	229.1	158.4	168.8

Table 6.2: Breakdown of integrated luminosities,  $\int \mathcal{L}$  [ $\text{pb}^{-1}$ ], by trigger list version. Differences in the luminosity between the *topo* and the *btag* analyses for v8.2-v11 are due to refinements of the data quality selection or data that had not yet been reconstructed for the *btag* analyses (in particular for v12).

## Chapter 7

# The Topological $t\bar{t}$ Analysis

In this chapter the measurement of the  $t\bar{t}$  production cross-section at  $\sqrt{s} = 1.96$  TeV with the DØ detector at the Fermilab Tevatron  $p\bar{p}$  collider, based on the application of a multivariate topological method is presented. The analysis can be subdivided in two major steps. The first one consists of a preselection, based on the selection of the decay products of the leptonically decaying  $W$  boson and jets, that serves to enhance the signal fraction. In the second step the number of  $t\bar{t}$  events is determined by estimating the remaining backgrounds and by comparing the topology of the data sample to the canonical topology of the signal and the backgrounds and fitting the relative number of events.

Section 7.1 discusses the individual preselection cuts and their efficiencies. The contributions to the preselected data set are specified in Sect. 7.2. The multivariate topological discriminant technique and the cross-section extraction procedure are presented in Sect. 7.3. The result of the  $t\bar{t}$  production cross-section measurement is presented in Sect. 7.4, followed by the discussion of the systematic uncertainties in Sect. 7.5.

### 7.1 $t\bar{t}$ Event Preselection

The signature of  $t\bar{t}$  events is discussed in Sect. 2.3.1; the preselection serves to select the individual objects in the final state with high efficiency and purity. The preselection can be summarized as a selection of a leptonically decaying  $W$  boson,  $W \rightarrow \mu\nu$ , in association with four or more jets.

To reject QCD-multijets events, where the muon originates from a heavy flavor decay, the muon is required to have a high transverse momentum (see Fig. 4.6), to come from the primary vertex (see Fig. 4.3) and to be isolated from hadronic activity (see Fig. 4.9). The neutrino is indirectly identified by requiring a high transverse missing energy and, in order to further reject background from QCD-multijets events, the missing energy has to be separated from the muon in  $\phi$ . Only events with four or more jets

are considered. Events with a second lepton with high transverse momentum are analyzed in the  $t\bar{t}$  dilepton analyses [66] and are explicitly vetoed here.

All cuts used to define the preselected sample and the corresponding efficiencies are listed in Table 7.12. The acceptance, reconstruction efficiencies and the survival efficiencies for the preselection cuts for  $t\bar{t}$  events ( $\varepsilon_{pre\,sel}$ ) are measured in a combination of data and Monte Carlo.  $\varepsilon_{pre\,sel}$  is measured in the  $t\bar{t}$  simulation with respect to  $t\bar{t} \rightarrow Wq\bar{q}'b\bar{b}$ , where the following  $W$  boson decays are allowed:  $W \rightarrow \mu\nu_\mu$  or  $W \rightarrow \tau\nu_\tau$  with  $\tau \rightarrow \mu\nu_\mu\nu_\tau$ . Possible inaccuracies in the MC to describe individual object identification or selection efficiencies are corrected by defining MC-to-data correction factors which are typically derived on a control sample where the respective efficiency can be extracted both on MC and data, as detailed in the following. The MC-to-data correction factor is defined as the efficiency measured in data divided by the efficiency measured in MC and can be determined as a function of the relevant quantities. These MC-to-data correction factors are referenced as  $\kappa$  with a descriptive subscript.

In the following Sect. 7.1.1-7.1.16 the factorization of the preselection efficiency ( $\varepsilon_{pre\,sel}$ ) is discussed and the relative efficiencies with respect to the previous cuts are given.

### 7.1.1 Jet Multiplicity Cut

Events with four or more central jets with high  $p_T$  are selected from the hadronic decay of the  $W$  boson and the two  $b$  jets from the top quark decay, see Fig. 2.19. The efficiency to select events with  $\geq 4$  jets with  $p_T > 20$  GeV and  $|\eta| < 2.5$  in  $t\bar{t} \rightarrow \mu + jets$  events is measured in the simulation and found to be:

$$\varepsilon_{\geq 4jets} = 46.45 \pm 0.36\% . \quad (7.1)$$

The jet reconstruction and identification efficiencies in the MC do not reproduce the corresponding efficiencies measured in data. For the efficiency number quoted in Eq. 7.1, the simulation has been corrected by a MC-to-data correction factor as discussed in the following.

#### 7.1.1.1 Jet Reconstruction and Identification Efficiency

A MC-to-data correction factor that accounts for the differences in the reconstruction and identification efficiency for jets is determined in QCD Compton events, where a photon ( $\gamma$ ) is reconstructed back-to-back to a jet in the azimuthal angle. A clean sample of these  $\gamma$ +jet events is recorded with a single EM trigger and selected by the following photon requirements.

- $EMF > 0.9$ ,
- $f_{iso} < 0.15$ ,

- $H\text{-matrix}7 < 20$ ,
- in fiducial (reconstructed away from a calorimeter cell boundary),
- $|\eta_{det}| < 1.1$ ,
- the algorithm which matches the calorimeter cluster to the tracks from the central tracking system did not converge:  $\chi^2 = -1.0$ ,
- no track with  $\Delta R(\gamma, \text{track}) < 0.5$ .

The  $p_T$  of the jet is extracted from the  $p_T$  of the photon which has to balance the jet momentum in the transverse plane and is back-to-back in the azimuthal angle to the jet. The  $\eta$  of the jet cannot be derived from the photon since the photon-jet system is subject to a boost in the  $z$  direction. However, one can reconstruct the jet with the central tracking system and use the spatial information of this track-jet (see Sect. 4.7.1.1) to get  $\eta$ . Thus, the additional selection criteria are:

- exactly one track-jet with  $p_T > 1$  GeV,
- $\Delta\phi(\gamma, \text{track-jet}) > 3.0$ .

These additional selection cuts are designed to reject background events with additional jets from gluon radiation. Additional jets are both vetoed by requiring exactly one track-jet with high  $p_T$  and by making a tight requirement that the photon and the track-jet are back-to-back in the azimuthal angle. The  $p_T$  dependent efficiency in  $\gamma$ +jet events is shown in Fig. 7.1 (left) for jets in the CC ( $|\eta| < 0.8$ ), ICD ( $0.8 < |\eta| < 1.5$ ) and EC ( $1.5 < |\eta| < 2.5$ ) for data and for Monte Carlo. The corresponding MC-to-data correction factor is shown on the right in Fig. 7.1 and fitted with the function

$$f(p_T) = 0.5 \cdot A2 \cdot \left( 1 + \frac{2}{\sqrt{\pi}} \cdot \int_0^{\frac{p_T - A0}{\sqrt{p_T} \cdot A1}} \exp(-t^2) dt \right), \quad (7.2)$$

with the fit values given in Table 7.1.

	A0	A1	A2
CC	9.8±2.7	3.4±0.7	1.01±0.01
ICD	15.0±2.3	2.0±0.7	0.93±0.02
EC	13.5±3.3	2.2±1.0	0.98±0.02

Table 7.1: Fit results for the jet reconstruction  $\times$  identification MC-to-data correction factor using the fit function given in Eq. 7.2.

The Monte Carlo simulation is corrected by randomly removing jets according to the estimated MC-to-data correction factor in order to reproduce the jet reconstruction  $\times$  identification efficiency in data. The  $p_T$  dependent Monte Carlo efficiency after removing jets according to the MC-to-data correction factor is also shown in Fig. 7.1 (left). Effects on the event reconstruction (jet counting,  $\cancel{E}_T$ ) are properly taken into account.

### Systematic Uncertainties

The  $p_T$  dependent MC-to-data correction factor suffers from a number of uncertainties arising from the method used. First of all the  $\gamma$ +jet sample is statistics limited and thus provides a  $\pm 1\sigma$  band for the MC-to-data correction factor versus  $p_T$ , shown in Fig. 7.2. The QCD Compton sample is known to be contaminated with dijet events, see Fig. 4.12, since the latter has a cross section orders of magnitude larger and the rate of electromagnetically fluctuating jets faking a photon is finite. The effect is estimated by varying the selection criteria on the photon to allow for a higher or lower jet contamination. It is found that the resulting variation of the MC-to-data correction factor is contained within the  $\pm 1\sigma$  band. The MC-to-data correction factor for the jet reconstruction and identification can also be derived using a  $Z$ +jet sample, see Fig. 7.3, with the same properties as the  $\gamma$ +jet sample but replacing the  $\gamma$  by a  $Z$  boson decaying either to muons or electrons. The MC-to-data correction factor obtained in this way has a large statistical uncertainty but is – within errors – consistent with the MC-to-data correction factor derived on the  $\gamma$ +jet sample.

By construction the MC-to-data correction factor derived on the  $\gamma$ +jet sample is a function of the  $p_T$  of the photon rather than of the jet. However, the correction is applied to the MC by evaluating the MC-to-data correction factor for the  $p_T$  of the present jets. This is not satisfying since the  $p_T$  resolution is different for jets and photons. The potential size of this effect on the MC-to-data correction factor is estimated by smearing the  $p_T$  of the photon with the energy resolution for jets. The resulting smeared MC-to-data correction factor is well within the  $\pm 1\sigma$  band resulting from the limited statistics in the  $\gamma$ +jet sample. In conclusion, the systematic uncertainty on the MC-to-data correction factor for the jet reconstruction  $\times$  identification is estimated by the  $\pm 1\sigma$  variation given by the statistical uncertainty, as shown in Fig. 7.2.

#### 7.1.2 Muon Acceptance, Reconstruction Efficiency and Cosmic Veto

The efficiency to reconstruct a muon with the standard Muon ID requirements as defined in Sect. 4.3.1:

- $|nseg| = 3$  medium,

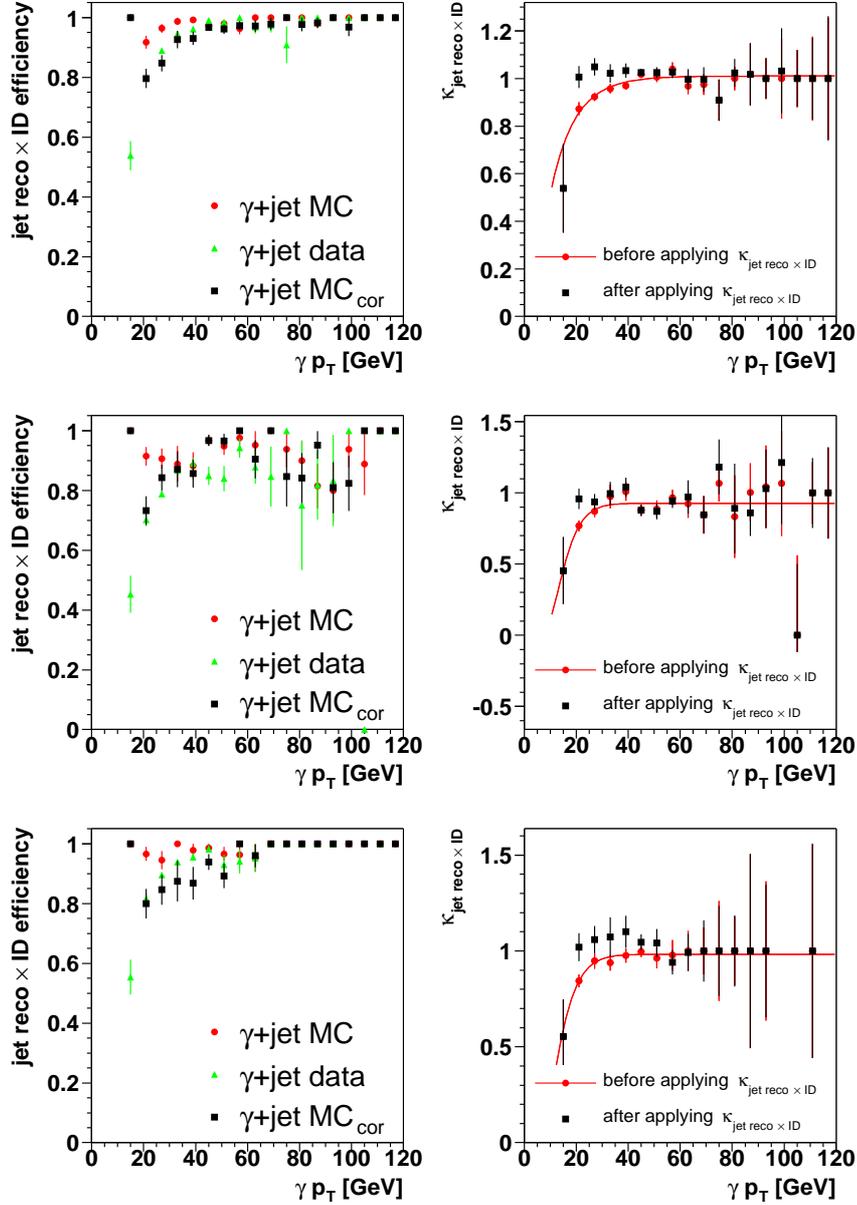


Figure 7.1: Left: Jet reconstruction  $\times$  identification efficiency as a function of the photon  $p_T$  for  $\gamma$ +jet events measured in data, in Monte Carlo and in the corrected MC, where the difference in efficiency is simulated in Monte Carlo. Right: MC-to-data correction factor between Monte Carlo and data before and after correcting the Monte Carlo. Top: For jets in the CC, Middle: For jets in the ICD, Bottom: For jets in the EC.

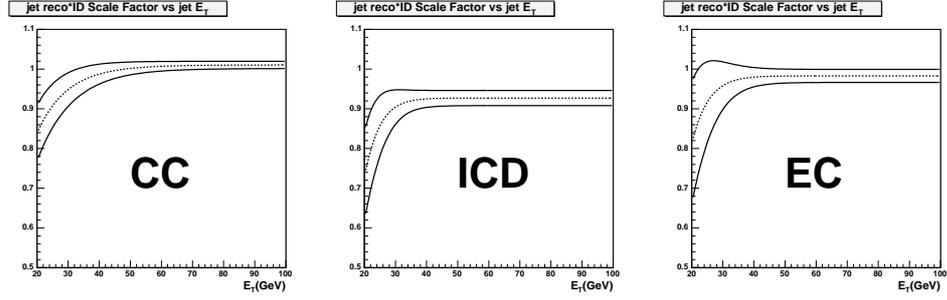


Figure 7.2: MC-to-data correction factor for jet reconstruction  $\times$  identification efficiency in CC, ICD and EC with the associated  $\pm 1\sigma$  band used for the systematic uncertainty on the MC-to-data correction factor.

- $|t_A| < 10$  ns and  $|t_{BC}| < 10$  ns

in  $t\bar{t} \rightarrow \mu + jets$  events is derived from the simulation and is found to be:

$$\varepsilon_{\mu ID} \times acc \times cosmic\ veto = 76.01 \pm 0.45\% .$$

A MC-to-data correction factor that accounts for the possible difference in the muon reconstruction efficiency between data and Monte Carlo is determined on  $Z \rightarrow \mu\mu$  events. To measure the efficiency of the muon reconstruction in data, the tag-and-probe method (see Sect. 4.5.3) is used on  $Z \rightarrow \mu\mu$  events which can be identified by one fully reconstructed muon (tag muon  $\mu_{tag}$ ) and a second muon only reconstructed in the tracking system (probe muon  $\mu_{probe}$ ). To ensure that the reconstructed track belongs to a muon from a  $Z$  decay, the invariant mass of the tag and the probe muon has to be compatible with the  $Z$  mass:  $80 \text{ GeV} < m_{\mu\mu} < 100 \text{ GeV}$ . The existing trigger bias is removed by selecting events recorded with a single muon trigger and matching the tag muon with the muon trigger object. The rate at which the probe muon is matched to a reconstructed muon is a measure of the local muon reconstruction and track-muon matching efficiency. It is found that the efficiency is not dependent on the details of the matching procedure. Fake muons and cosmic muons are reduced by the following additional cuts on the tag muon:

- $|nseg| = 3$  medium,
- $p_T > 20$  GeV,
- $|\eta_{det}| < 2$ ,
- $\Delta R(\mu_{tag}, jet) > 0.5$ ,
- $Rat11 < 0.08$ ,

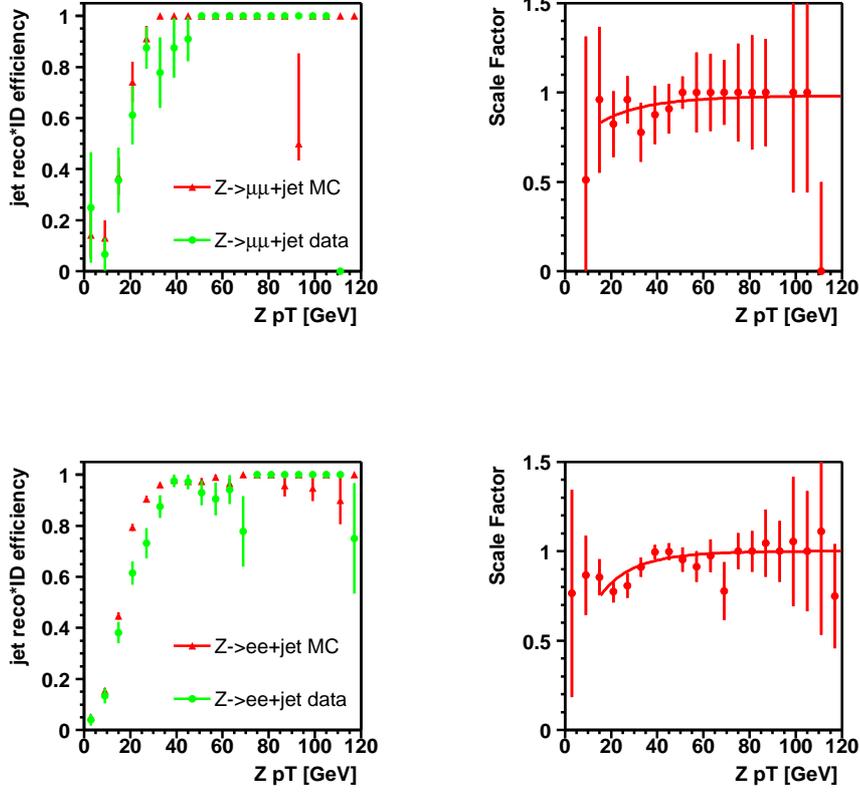


Figure 7.3: Left: Jet reconstruction  $\times$  identification efficiency as a function of the  $Z$   $p_T$  for  $Z + jet$  events measured in data and in Monte Carlo. Right: Correction factor between Monte Carlo and data. Top: For  $Z \rightarrow \mu\mu + jet$  events, Bottom: For  $Z \rightarrow ee + jet$  events.

- $R_{attrk} < 0.06$ ,
- $|t_A| < 10$  ns and  $|t_{BC}| < 10$  ns,
- $|dca|/\sigma_{dca} < 3$ ,
- $\chi_{trk}^2 < 4$ ,
- $|\Delta z(\mu, PV)| < 1$  cm,

on the probe muon (track):

- $p_T > 20$  GeV,
- $|\eta_{det}| < 2$ ,

- $\Delta R(\mu_{probe}, jet) > 0.5$ ,
- $R_{attrk} < 0.06$  GeV,
- $\chi_{trk}^2 < 3$ ,
- $|dca| < 30$   $\mu\text{m}$ ,
- $|dca|/\sigma_{dca} < 2$ ,
- number of SMT hits  $\geq 1$ ,

and on both muons:

- $\Delta\phi(\mu_{tag}, \mu_{probe}) > 2.9$
- acolinearity ( $= \Delta R(\mu_{tag}, \mu_{probe}^{reflected})$ , where  $\mu_{probe}^{reflected}$  is  $\mu_{probe}$  reflected over the PV)  $> 0.1$ ,
- $|\Delta z(\mu_{tag}, \mu_{probe})| < 0.5$  cm,
- opposite charge of tag and probe muon.

After this selection the purity of the probe muons in data is estimated to be on the order of 99 % or higher.

The same measurement is done in Monte Carlo with  $Z \rightarrow \mu\mu$  events and the differential efficiencies as a function of detector  $\eta$ ,  $\phi$  and  $p_T$  are compared. The ratio of the measured efficiencies in data and Monte Carlo is found to depend only on the muon  $p_T$  and the absolute value of detector  $\eta$  and is parameterized using a polynomial of first and of third order respectively. The dependence is assumed to be factorizable, so that  $\kappa_{\mu ID \times acc \times cosmic veto}(p_T, |\eta_{det}|) \equiv C \times \kappa_{\mu ID \times acc \times cosmic veto}(p_T) \times \kappa_{\mu ID \times acc \times cosmic veto}(|\eta_{det}|)$ . The normalization factor  $C$  is calculated so that the total number of predicted muons obtained as the sum over all reconstructed muons in the MC weighted with  $\kappa_{\mu ID \times acc \times cosmic veto}(p_T, |\eta_{det}|)$  is equal to the total number of reconstructed muons in the data. The  $p_T$  and  $|\eta_{det}|$  dependences are fitted with a polynomial of first and of third order, respectively; the fit values are given in Table 7.2, the resulting parameterizations are shown in Fig 7.4. The average MC-to-data correction factor for  $t\bar{t} \rightarrow \mu + jets$  events is:

$$\kappa_{\mu ID \times acc \times cosmic veto} = 1.005 \pm 0.022 \text{ (syst)}.$$

The systematic uncertainty on the measured  $\kappa_{\mu ID \times acc \times cosmic veto}$  as a function of muon  $p_T$  and  $|\eta_{det}|$  is obtained by error propagation of the statistical uncertainty of each fitted parameter.

	p0	p1	p2	p3
$p_T$	$1.12 \pm 0.03$	$-0.0024 \pm 0.0008$	—	—
$ \eta_{det} $	$1.00 \pm 0.03$	$0.1 \pm 0.1$	$-0.01 \pm 0.1$	$-0.03 \pm 0.05$

Table 7.2: Fit results for  $\kappa_{\mu ID} \times acc \times cosmic veto(p_T)$  and  $\kappa_{\mu ID} \times acc \times cosmic veto(|\eta_{det}|)$ , using a polynomial of first and of third order, respectively; p0, p1, p2 and p3 are the polynomial coefficients.

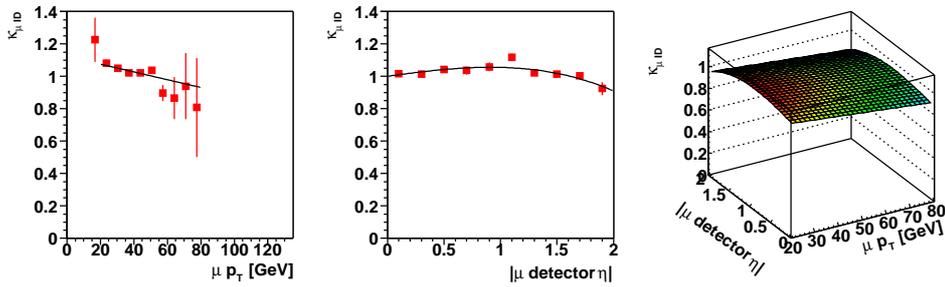


Figure 7.4: MC-to-data correction factor for the muon reconstruction efficiency,  $\kappa_{\mu ID}$ , as a function of muon  $p_T$  (left), as a function of muon  $|\eta_{det}|$  (middle) and the convolution of the two parameterizations (right).

### 7.1.3 Muon Tracking Efficiency

The efficiency to reconstruct a track to the corresponding muon in  $t\bar{t} \rightarrow \mu + jets$  events is measured in the  $t\bar{t}$  simulation and found to be

$$\varepsilon_{\mu track} = 97.21 \pm 0.20\% .$$

The determination of a MC-to-data correction factor that accounts for the difference in tracking efficiency between data and Monte Carlo is described in the following. The efficiency measurement is done by counting the fraction of muons identified in the muon chambers and calorimeter that have a matching central track. Muons used are required to have:

- At least two wire hits in the A segment,
- at least three wire hits in the BC segment,
- at least one scintillator hit in the A segment,
- at least one scintillator hit in the BC segment,
- $|t_A| < 5$  ns and  $|t_{BC}| < 5$  ns,
- track fit in muon chambers converges ( $\chi_{trk}^2 > 0$ ),

- $p_T$  of local muon  $> 2.0$  GeV,
- $|z|$  of local muon track  $< 60$  cm,
- $xy$  impact parameter of local muon track  $< 20$  cm,
- a matching calorimeter MIP trace (so-called MTC), with  $\Delta R(\mu, MTC) < 0.3$ ,
- $\text{Halo}(0.1, 0.4) < 2.5$  GeV,
- $\Delta R(\mu, jet) > 1.0$ .

Once the muon samples have been defined, the central track reconstruction efficiency may be measured by noting the probability of there being a central track pointing toward the muon. To do this, one must first define the size of the region around the muon in which to search for a track. Empirically, the smallest window that includes well over 99% of all matching tracks is found to be  $\Delta\phi \times \Delta\eta = \pm 0.2 \times \pm 0.3$ , centered on the muon direction.

It is possible that other charged particles in the event lie within the muon matching window. After measuring the probability  $\varepsilon_r$  of a random overlap using a control window away from the muon, the true efficiency  $\varepsilon_t$  is related to the measured efficiency  $\varepsilon_m$  by:

$$1 - \varepsilon_m = (1 - \varepsilon_r) \cdot (1 - \varepsilon_t), \quad (7.3)$$

so that

$$\varepsilon_t = \frac{\varepsilon_m - \varepsilon_r}{1 - \varepsilon_r}. \quad (7.4)$$

The data sample used is the entire 2MU CSG skim [142], split into data taken before and after a CFT timing change in Spring 2003. Only events recorded during good runs are considered. The Monte Carlo sample used is a PYTHIA  $Z \rightarrow \mu\mu$  sample.

The track reconstruction efficiency is binned in a two-dimensional grid in  $\eta$  and  $\phi$ . Averaging over all bins, one finds the following relative MC-to-data correction factor:

- Period I (run numbers  $< 175757$ ):  $\kappa_{\mu track} = 98.1\%$ .
- Period II (run numbers  $\geq 175757$ , before November 2003):  $\kappa_{\mu track} = 99.3\%$ .
- Period II (after November 2003):  $\kappa_{\mu track} = 98.5\%$ .

### 7.1.3.1 Systematic Uncertainties

The systematic uncertainties on the measurement are summarized in Table 7.3 and arise from the following sources:

- **Variations in  $\eta$  and  $\phi$ :** Since the  $\eta$  distribution for a given physics channel may be different from that of the control sample, used to derive the above numbers, one needs to account for the variations in the relative efficiency vs.  $\eta$  and  $\phi$ . For Period I the RMS spread is 2.5%, for Period II it's 2.2%, and for Period III it's 3.4%.
- **Variations in  $p_T$ :** The above calculations assume that the efficiency is flat in  $p_T$ . This assumption is very good for the MC sample, but less so for the data. Considering the efficiency in different  $p_T$  regions separately, one finds a variation of  $\kappa_{\mu track}$  of 0.6% in Period I data, 1.2% in Period II, and 1.1% in Period III.
- **QCD contribution to sample:** While all muons in the MC sample are truly isolated, some fraction of the data sample is composed of muons produced in jets (either via heavy quark decay or meson decay in flight). Using a QCD MC sample gives an efficiency consistent with the  $Z \rightarrow \mu\mu$  MC, with an uncertainty of 0.5%.
- **Cosmic rays:** Loosening the timing cuts by a factor of two changes the measured efficiency in data by 0.5% in Period I, 0.2% in Period II, and 0.5% in Period III.
- **Random overlaps:** Doubling the size of the matching window in both  $\eta$  and  $\phi$  changes the measured  $\kappa_{\mu track}$  by 0.1% in Period I, 0.04% in Period II, and 0.1% in Period III.
- **Control sample statistics:** Statistics are plentiful in both, the data and MC samples, so the statistical error on the relative efficiency is 0.1%.

The luminosity weighted average MC-to-data correction factor with systematic uncertainty is:

$$\kappa_{\mu track} = 0.988 \pm 0.030 \text{ (syst)}.$$

### 7.1.4 Efficiency of the Muon $\eta$ Cut

The efficiency to select events with  $|\eta_{\mu}| < 2.0$  in  $t\bar{t} \rightarrow \mu + jets$  events is measured in the simulation and is found to be:

$$\epsilon_{|\eta_{\mu}| < 2.0} = 99.85 \pm 0.05\%.$$

Source	Magnitude (%)		
	Period I	Period II	Period III
$\eta, \phi$ variations	2.5	2.2	3.4
$p_T$ variations	0.6	1.2	1.1
Muons from QCD	0.5	0.5	0.5
Cosmic rays	0.5	0.2	0.5
Random overlaps	0.1	0.04	0.1
Control sample statistics	0.1	0.1	0.1
Total	2.7	2.6	3.6

Table 7.3: Summary of systematic uncertainties on the muon tracking MC-to-data correction factor.

### 7.1.5 Efficiency of the Muon Transverse Momentum Cut

The efficiency to select events with  $p_{T\mu} > 20$  GeV in  $t\bar{t} \rightarrow \mu + jets$  events is measured in the simulation and is found to be:

$$\varepsilon_{p_{T\mu} > 20 \text{ GeV}} = 71.97 \pm 0.56\%.$$

### 7.1.6 Efficiency of the Muon Separation from Jets

The efficiency to select events with  $\Delta R(\mu, jet) > 0.5$  in  $t\bar{t} \rightarrow \mu + jets$  events is measured in the simulation and is found to be:

$$\varepsilon_{\Delta R(\mu, jet) > 0.5} = 78.42 \pm 0.60\%.$$

### 7.1.7 Efficiency of the Track $\chi_{trk}^2$ Cut

A quality cut on the track associated to the muon is applied in order to reduce mismeasured tracks. The  $\chi_{trk}^2 < 4$  efficiency for  $t\bar{t} \rightarrow \mu + jets$  events is measured in the simulation and is found to be:

$$\varepsilon_{\chi_{trk}^2 < 4} = 99.84 \pm 0.07\%.$$

The measured efficiencies for  $\chi_{trk}^2 < 4$  in  $Z \rightarrow \mu\mu$  data and Monte Carlo events and the ratio of the efficiencies are shown in Table 7.4 as a function of the jet multiplicity. A MC-to-data correction factor is determined from a constant fit to the 4 jet multiplicity bins:

$$\kappa_{\chi_{trk}^2 < 4} = 0.986 \pm 0.003.$$

	$\varepsilon_{\chi_{trk}^2 < 4}^{data}$ [%]	$\varepsilon_{\chi_{trk}^2 < 4}^{MC}$ [%]	$\kappa_{\chi_{trk}^2 < 4}$
$N_{jet} = 1$	$98.6 \pm 0.3$	$99.9 \pm 0.0$	$0.987 \pm 0.003$
$N_{jet} = 2$	$96.4 \pm 1.3$	$99.9 \pm 0.0$	$0.966 \pm 0.013$
$N_{jet} = 3$	$100.0 \pm 0.0$	$99.9 \pm 0.1$	$1.001 \pm 0.016$
$N_{jet} \geq 4$	-	$100.0 \pm 0.0$	-

Table 7.4: Efficiencies for  $\chi_{trk}^2 < 4$  in  $Z \rightarrow \mu\mu$  data and Monte Carlo events and MC-to-data correction factor  $\kappa_{\chi_{trk}^2 < 4}$  for events with 1, 2, 3 and  $\geq 4$  jets in the event.

### 7.1.8 Efficiency of the Muon $dca$ Significance Cut

The cut on the  $dca$  significance of the track matched to a muon  $|dca|/\sigma_{dca} < 3$  is used to reject muons from semileptonic decays, while most of the muons from  $W$  decays are kept. The efficiency to select events with  $|dca|/\sigma_{dca} < 3$  in  $t\bar{t} \rightarrow \mu + jets$  events is measured in the simulation and is found to be:

$$\varepsilon_{|dca|/\sigma_{dca} < 3} = 94.69 \pm 0.37\%.$$

The efficiency for the  $|dca|/\sigma_{dca}$  cut is measured both, in  $Z \rightarrow \mu\mu$  data and Monte Carlo events, and is found to be in very good agreement, as shown in Table 7.5 as a function of the jet multiplicity. The ratio of the measured efficiencies in data and Monte Carlo is also shown in Table 7.5, a MC-to-data correction factor is determined from a constant fit to the 4 jet multiplicity bins:

$$\kappa_{|dca|/\sigma_{dca} < 3} = 0.995 \pm 0.003.$$

	$\varepsilon_{ dca /\sigma_{dca} < 3}^{data}$ [%]	$\varepsilon_{ dca /\sigma_{dca} < 3}^{MC}$ [%]	$\kappa_{ dca /\sigma_{dca} < 3}$
$N_{jet} = 1$	$98.4 \pm 0.3$	$98.9 \pm 0.1$	$0.994 \pm 0.004$
$N_{jet} = 2$	$98.9 \pm 0.7$	$99.6 \pm 0.1$	$0.994 \pm 0.007$
$N_{jet} = 3$	$100.0 \pm 0.0$	$99.5 \pm 0.2$	$1.005 \pm 0.016$
$N_{jet} \geq 4$	-	$100.0 \pm 0.0$	-

Table 7.5: Efficiencies for  $|dca|/\sigma_{dca} < 3$  in  $Z \rightarrow \mu\mu$  data and Monte Carlo events and MC-to-data correction factor  $\kappa_{|dca|/\sigma_{dca} < 3}$  for events with 1, 2, 3 and  $\geq 4$  jets in the event.

### 7.1.9 Tight Muon Isolation Efficiency

The tight muon isolation as defined in Sect. 4.3.4:

- $\text{Rat11} < 0.08$ ,
- $\text{Rattrk} < 0.06$ ,

is dependent on the topology of the event, in particular on the jet in the vicinity of the high- $p_T$  muon, and therefore has to be measured in the  $t\bar{t} \rightarrow \mu + jets$  simulation:

$$\varepsilon_{\text{tight isolation}} = 85.41 \pm 0.60\%.$$

Differences between the data and the simulation are quantified using  $Z \rightarrow \mu\mu$  events. The isolation efficiency is measured as a function of jet multiplicity both in simulated and in data events, as shown in Table 7.6. The ratio of the measured efficiencies in data and Monte Carlo is also shown in Table 7.6, a MC-to-data correction factor is determined from a constant fit to the 4 jet multiplicity bins:

$$\kappa_{\text{tight isolation}} = 0.982 \pm 0.009.$$

The jets in  $t\bar{t}$  events are more central than in  $Z$ -plus-jets events, so that the

	$\varepsilon_{\text{tight isolation}}^{\text{data}}$ [%]	$\varepsilon_{\text{tight isolation}}^{\text{MC}}$ [%]	$\kappa_{\text{tight isolation}}$
$N_{\text{jet}} = 1$	$89.5 \pm 0.8$	$91.2 \pm 0.3$	$0.982 \pm 0.009$
$N_{\text{jet}} = 2$	$88.0 \pm 2.2$	$89.6 \pm 0.4$	$0.981 \pm 0.027$
$N_{\text{jet}} = 3$	$85.7 \pm 5.9$	$87.0 \pm 1.1$	$0.985 \pm 0.080$
$N_{\text{jet}} \geq 4$	-	-	-

Table 7.6: Tight isolation efficiencies ( $\text{Rat11} < 0.08$  and  $\text{Rattrk} < 0.06$ ) in  $Z \rightarrow \mu\mu$  data and Monte Carlo events and MC-to-data correction factor  $\kappa_{\text{tight isolation}}$  for events with 1, 2, 3 and  $\geq 4$  jets in the event.

jets are on average closer to the muon in  $t\bar{t}$  events. However, in Fig 7.5 it is shown that the MC-to-data correction factor does not depend on the  $\Delta R(\mu, \text{closest jet})$ .

### 7.1.10 Efficiencies for the Missing Transverse Energy Cut

A cut on the missing transverse energy is applied in order to reject background from QCD-multijets events. However, there are QCD heavy flavor events where neutrinos are produced resulting in real  $\cancel{E}_T$  in the event. The different angular distribution of the missing transverse energy and the muon (or the leading jet) allows to further reject QCD without suffering from a large efficiency loss.

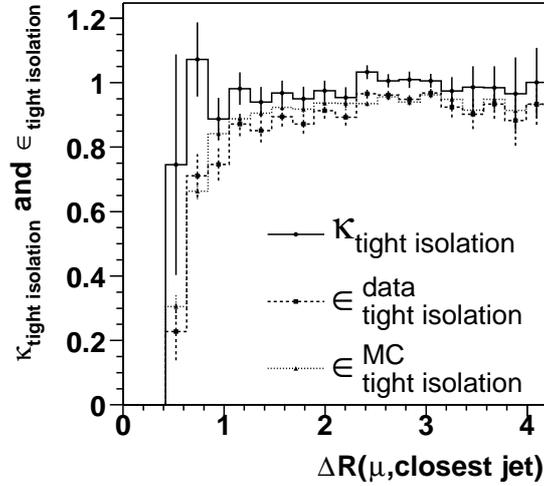


Figure 7.5: Tight muon isolation efficiency for  $Z \rightarrow \mu\mu$  data and Monte Carlo events and the MC-to-data correction factor as a function of the  $\Delta R$  of the muon to the closest jet. No significant dependence is found.

- QCD is found to have the missing transverse energy mainly back-to-back with the leading jet, where the missing transverse energy in these events is generated by the mismeasurement of the leading jet energy.
- QCD is also found to have the missing transverse energy either in the direction of the muon or back-to-back to it. This can be explained by  $b\bar{b}$  production where one or both of the  $b$  hadrons decay semimuonically, which implies the presence of a neutrino and a muon in the jet, however, the jet might not be reconstructed in the calorimeter, or the muon transverse momentum is mismeasured.

In a grid search the cuts on the missing transverse energy and the  $\Delta\phi$  between the missing transverse energy and the muon, respectively the leading jet are varied. A set of cuts is chosen which maximizes the  $t\bar{t}$  efficiency  $\times$  purity, see Fig. 7.6. The optimum set of cuts is found to be:

- $\cancel{E}_T > 20 \text{ GeV}$
- $\Delta\phi(\mu, \cancel{E}_T) > 0.1 \cdot \pi - 0.1 \cdot \pi \cdot \cancel{E}_T [\text{GeV}]/50$
- $\Delta\phi(\mu, \cancel{E}_T) < 0.8 \cdot \pi + 0.2 \cdot \pi \cdot \cancel{E}_T [\text{GeV}]/30$

which is illustrated in Fig. 7.7. The efficiency for this set of cuts is determined with simulated  $t\bar{t}$  events and is found to be:

$$\varepsilon_{\cancel{E}_T - \Delta\phi \text{ cut}} = 89.19 \pm 0.54\%.$$

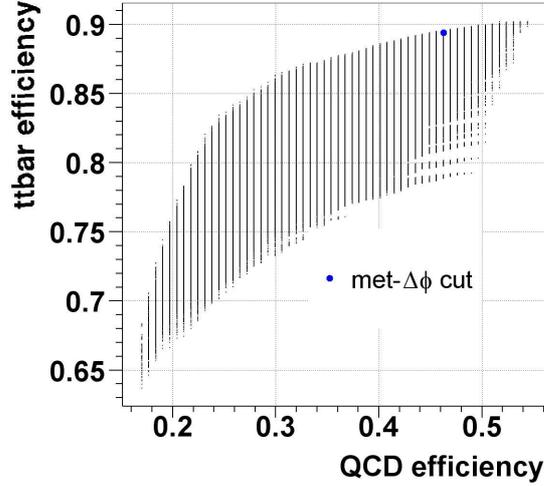


Figure 7.6:  $t\bar{t}$  signal efficiency versus the survival efficiency for the QCD background (QCD-multijets data sample). The cut which maximizes efficiency  $\times$  purity is highlighted by the full point.

### 7.1.11 Efficiency of the Second Lepton Veto (SLV)

Events with a second lepton with high transverse momentum are analyzed in Ref. [66] and are explicitly vetoed here. The corresponding cut efficiency for  $t\bar{t} \rightarrow \mu + jets$  events is measured in the simulation. Events are rejected if a second muon is found which satisfies:

- $p_T > 15$  GeV,
- $|nseg| = 3$  medium,
- central track match,
- $Rat11 < 0.12$ ,
- $Rattrk < 0.12$ .

The efficiency for the second muon veto (SMV) is found to be:

$$\varepsilon_{SMV} = 99.92 \pm 0.05\%.$$

Events are also rejected if an electron is found which satisfies:

- $p_T > 15$  GeV,
- $EMF > 0.9$ ,
- $f_{iso} < 0.15$ ,

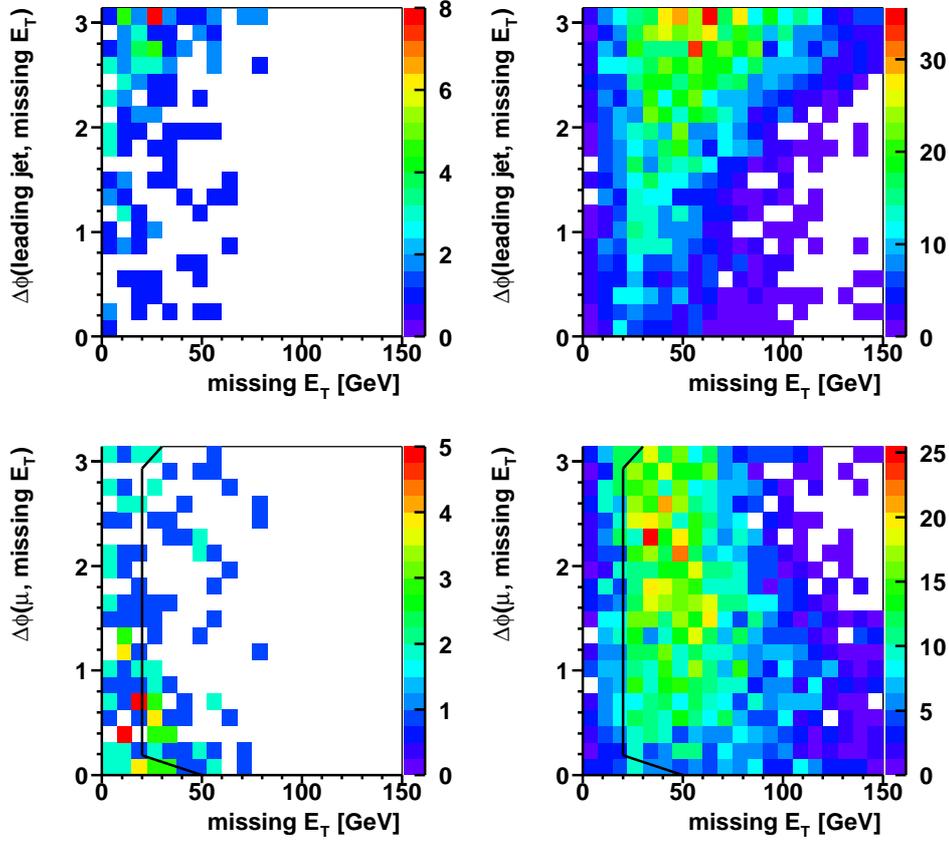


Figure 7.7: Top:  $\Delta\phi(\text{leading jet}, \cancel{E}_T)$  versus  $\cancel{E}_T$ . Bottom:  $\Delta\phi(\mu, \cancel{E}_T)$  versus  $\cancel{E}_T$ . Left: QCD-multijets without applying the  $\cancel{E}_T$  cut. Right:  $t\bar{t}$  MC without applying the  $\cancel{E}_T$  cut. The black lines represent the cut borders.

- $H\text{-matrix7} < 50$ ,
- central track match:  $\chi_{trk}^2 \neq -1.0$ ,
- electron likelihood  $> 0.85$
- $|\eta_{det}| < 1.1$  or  $1.5 < |\eta_{det}| < 2.5$

The efficiency for the electron veto (EV) is found to be:

$$\varepsilon_{EV} = 99.89 \pm 0.06\%.$$

### 7.1.12 Efficiency of the $Z \rightarrow \mu\mu$ Veto

The second muon veto (see Sect. 7.1.11) serves the construction of orthogonal samples for the various final states. However, it is not very efficient in

rejecting  $Z$ -plus-jets events which survive the preselection due to the finite  $\cancel{E}_T$  resolution in events with four or more jets or due to the fact that muons, not fulfilling the medium  $|nseg| = 3$  requirement, are not used for the  $\cancel{E}_T$  correction. Figure 7.8 shows the invariant mass of the selected isolated high  $p_T$  muon and the additional highest  $p_T$  muon which fulfills the following criteria:

- ‘loose’ quality,
- central track match,
- isolated from a jet:  $\Delta R(\mu, jet) > 0.5$ .

This selection picks up a large fraction (‘loose’) of all additional reconstructed isolated muons ( $\Delta R(\mu, jet) > 0.5$ ), with a good  $p_T$  resolution (central track match) and thus, a good invariant mass ( $m_{\mu\mu}$ ) resolution. Soft muons from a semileptonic  $b$ -decay which are expected to be reconstructed in jets are not taken into consideration in order to keep a high efficiency for  $t\bar{t}$  events (where two  $b$  jets are in the final state).

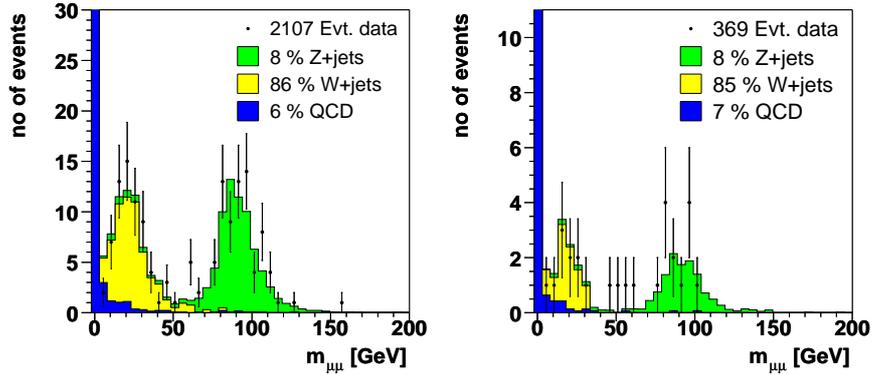


Figure 7.8: Invariant mass of the selected isolated high  $p_T$  muon and any second loose muon with track match and  $\Delta R(\mu, jet) > 0.5$ . The QCD fraction is estimated by the matrix method. The  $Z \rightarrow \mu\mu$ -plus-jet fraction is estimated as given in Eq. 7.5. Left: events with 2 jets reconstructed. Right: events with 3 jets reconstructed.  $m_{\mu\mu}$  is set to zero for events with no second muon.

The number of background events from  $Z$ -plus-jets events is estimated using the relative preselection efficiency and cross section for  $Z \rightarrow \mu\mu$  plus two jets ( $Zjj$ ) with respect to  $W \rightarrow \mu\nu$  plus two jets ( $Wjj$ ) from ALPGEN. Two jets are required (instead of  $\geq$  four jets), which is a good approximation of the relative preselection efficiency for events with  $\geq 4$  jets, since the production mechanisms for additional jets are the same for  $Z$ -plus-jets and

$W$ -plus-jets events.

$$\begin{aligned}
 N_{Wjj} &= \sigma_{Wjj} \cdot BR_{W \rightarrow \mu\nu} \cdot \varepsilon_{Wjj} \cdot \mathcal{L} \\
 N_{Zjj} &= \sigma_{Zjj} \cdot BR_{Z \rightarrow \mu\mu} \cdot \varepsilon_{Zjj} \cdot \mathcal{L} \\
 N_{Zjj} &= \frac{\sigma_{Zjj} \cdot BR_{Z \rightarrow \mu\mu}}{\sigma_{Wjj} \cdot BR_{W \rightarrow \mu\nu}} \cdot \frac{\varepsilon_{Zjj}}{\varepsilon_{Wjj}} \cdot N_{Wjj}.
 \end{aligned} \tag{7.5}$$

The input numbers for Eq. 7.5 are given in Table 7.7. The relative fraction

$\frac{\sigma_{Zjj} \cdot BR_{Z \rightarrow \mu\mu}}{\sigma_{Wjj} \cdot BR_{W \rightarrow \mu\nu}}$	$\varepsilon_{Wjj}$	$\varepsilon_{Zjj}$
0.1	3.12%	2.98%

Table 7.7: Relative  $W$ jj to  $Z$ jj cross section  $\times$  branching ratio from ALPGEN and the preselection efficiencies without applying the  $Z \rightarrow \mu\mu$  veto.

of selected  $W$ -plus-jets and  $Z$ -plus-jet events is estimated to be  $\frac{N_{Wjj}}{N_{Zjj}} = 10.5$  which reproduces very well the number of  $Z$ -plus-jets events seen in the data (see  $m_{\mu\mu}$  around the  $Z$  peak in Fig. 7.8).

A cut on the invariant dimuon mass,  $70 \text{ GeV} < m_{\mu\mu} < 110 \text{ GeV}$ , rejects 35 % of the preselected  $Z$ -plus-jets background. The remaining  $Z$ -plus-jets background cannot be rejected since no second muon is reconstructed mainly for acceptance reasons. The  $t\bar{t}$  efficiency for this cut is estimated in the  $t\bar{t}$  simulation and found to be:

$$\varepsilon_{Z \rightarrow \mu\mu} = 99.82 \pm 0.07\%.$$

### 7.1.13 PV Reconstruction Efficiency

The PV reconstruction efficiency is evaluated using two methods:

- The first method makes use of the preselected signal samples. The primary vertex efficiency is measured after applying all preselection cuts:

$$\varepsilon_{PV}^{data} = 96.5_{-2.6}^{+1.5}\%.$$
 \tag{7.6}

It is presented as a function of jet multiplicity in Table 7.8.

- In the second method the primary vertex cut efficiencies are determined in the  $t\bar{t}$  MC:

$$\varepsilon_{PV}^{t\bar{t}} = 98.0 \pm 0.3\%.$$

A MC-to-data correction factor, derived from  $Z \rightarrow \mu\mu$  events in data and MC, is presented in Table 7.9. Due to the lack of  $Z \rightarrow \mu\mu$  statistics

in data with three or more jets  $\kappa_{PV}$  is determined from a constant fit to the 4 jet multiplicity bins:

$$\kappa_{PV} = 0.994 \pm 0.004.$$

Table 7.10 summarizes the primary vertex efficiencies in  $t\bar{t}$  and  $W$ -plus-jets MC as a function of jet multiplicity.

	$\varepsilon_{PV}^{data}$ [%]
$N_{jet} = 1$	$95.0^{+0.2}_{-0.3}$
$N_{jet} = 2$	$95.0^{+0.4}_{-0.5}$
$N_{jet} = 3$	$97.1^{+0.7}_{-1.0}$
$N_{jet} \geq 4$	$96.5^{+1.5}_{-2.6}$

Table 7.8: Primary vertex cuts efficiencies (in %) as measured in the preselected signal sample as a function of jet multiplicity.

	$\varepsilon_{PV}^{Z \rightarrow \mu\mu data}$ [%]	$\varepsilon_{PV}^{Z \rightarrow \mu\mu MC}$ [%]	$\kappa_{PV}$
$N_{jet} = 1$	$98.6^{+0.4}_{-0.5}$	$99.4 \pm 0.1$	$0.994 \pm 0.004$
$N_{jet} = 2$	$99.0^{+0.6}_{-1.7}$	$99.5^{+0.1}_{-0.2}$	$0.995^{+0.006}_{-0.017}$
$N_{jet} = 3$	$100.0^{+0.}_{-6.3}$	-	-
$N_{jet} \geq 4$	-	-	-

Table 7.9: Primary vertex cut efficiencies (in %) in  $Z \rightarrow \mu\mu$  data and Monte Carlo and MC-to-data correction factor as a function of jet multiplicity.

	$\varepsilon_{PV}^{t\bar{t}}$ [%]	$\kappa_{PV} \cdot \varepsilon_{PV}^{t\bar{t}}$ [%]	$\varepsilon_{PV}^W$ [%]	$\kappa_{PV} \cdot \varepsilon_{PV}^W$ [%]
$N_{jet} = 1$	$96.3^{+1.4}_{-2.3}$	$95.7^{+1.4}_{-2.3}$	$98.5^{+0.3}_{-0.3}$	$97.9^{+0.3}_{-0.3}$
$N_{jet} = 2$	$98.4^{+0.4}_{-0.5}$	$97.8^{+0.4}_{-0.5}$	$97.9 \pm 0.1$	$97.3 \pm 0.1$
$N_{jet} = 3$	$98.4^{+0.2}_{-0.3}$	$97.8^{+0.2}_{-0.3}$	$98.0^{+0.6}_{-0.8}$	$97.4^{+0.6}_{-0.8}$
$N_{jet} \geq 4$	$98.0^{+0.3}_{-0.3}$	$97.4^{+0.3}_{-0.3}$	$98.6^{+0.2}_{-0.3}$	$98.0^{+0.2}_{-0.3}$

Table 7.10: Primary vertex cut efficiencies (in %) in  $t\bar{t}$  and  $W$ -plus-jets MC. For  $\kappa_{PV}$  see Table 7.9.

Both methods (Eq. 7.6 and Eq. 7.7) are in agreement and the PV efficiency quoted is

$$\begin{aligned}\varepsilon_{PV}^{t\bar{t}} \times \kappa_{PV} &= (98.0 \pm 0.3)\% \times (0.994 \pm 0.004) \\ &= 97.4 \pm 0.5 \text{ (stat)} \pm 0.9 \text{ (syst)}\%.\end{aligned}\quad (7.7)$$

The systematic uncertainty on the primary vertex selection efficiency is taken from the observed difference in efficiency between the two methods (Eq. 7.6 and Eq. 7.7).

#### 7.1.14 Efficiency of the $|\Delta z(\mu, PV)|$ Cut

The cut  $|\Delta z(\mu, PV)| < 1$  cm is used to reject both muons from semileptonic decays and cosmic muons, while most of the muons from  $W$  decays are kept. The  $|\Delta z(\mu, PV)| < 1$  cm efficiency for  $t\bar{t} \rightarrow \mu + jets$  events is measured in the simulation and is found to be:

$$\varepsilon_{|\Delta z(\mu, PV)| < 1 \text{ cm}} = 99.92 \pm 0.05\%.$$

The measured efficiencies in  $Z \rightarrow \mu\mu$  data and Monte Carlo events and the ratio of efficiencies are shown in Table 7.11 as a function of jet multiplicity. A MC-to-data correction factor is determined from a constant fit to the 4 jet multiplicity bins:

$$\kappa_{|\Delta z(\mu, PV)| < 1 \text{ cm}} = 0.999 \pm 0.001.$$

	$\varepsilon_{ \Delta z(\mu, PV)  < 1 \text{ cm}}^{data}$ [%]	$\varepsilon_{ \Delta z(\mu, PV)  < 1 \text{ cm}}^{MC}$ [%]	$\kappa_{ \Delta z(\mu, PV)  < 1 \text{ cm}}$
$N_{jet} = 1$	$99.9 \pm 0.1$	$100.0 \pm 0.0$	$0.999 \pm 0.001$
$N_{jet} = 2$	$100.0 \pm 0.0$	$100.0 \pm 0.0$	$1.000 \pm 0.003$
$N_{jet} = 3$	$100.0 \pm 0.0$	$100.0 \pm 0.0$	$1.000 \pm 0.016$
$N_{jet} \geq 4$	-	-	-

Table 7.11: Efficiencies for  $|\Delta z(\mu, PV)| < 1$  cm in  $Z \rightarrow \mu\mu$  data and Monte Carlo events and MC-to-data correction factor  $\kappa_{|\Delta z(\mu, PV)| < 1 \text{ cm}}$  for events with 1, 2, 3 and  $\geq 4$  jets in the event.

#### 7.1.15 Jet Trigger Efficiencies

The determination of the trigger efficiency per jet and the parameterization as a function of the relevant variables is described in Sect. 5.1.2. The jet trigger efficiency for  $t\bar{t} \rightarrow \mu + jets$  events is measured in the simulation by folding in the trigger parameterization per jet, given by Eq. 5.3, and deriving

the event trigger efficiency separately for the L1, the L2 and the L3 trigger requirements:

$$\begin{aligned}\varepsilon_{\text{L1 jet trigger}} &= 99.99 \pm 0.01 \text{ (stat)}_{-0.0}^{+0.0} \text{ (syst)}\%, \\ \varepsilon_{\text{L2 jet trigger}} &= 100.00 \pm 0.00 \text{ (stat)}_{-0.0}^{+0.0} \text{ (syst)}\%, \\ \varepsilon_{\text{L3 jet trigger}} &= 99.76 \pm 0.08 \text{ (stat)}_{-0.1}^{+0.1} \text{ (syst)}\%.\end{aligned}$$

The origin of the systematic uncertainty on the trigger efficiency is discussed in Sect. 5.1.2.

### 7.1.16 Muon Trigger Efficiencies

The determination of the trigger efficiency per medium muon with  $|n_{seg}| = 3$  and  $|t_A| < 10$  ns and  $|t_{BC}| < 10$  ns, and the parameterization as a function of the relevant variable  $\eta$  is described in Sect. 5.1.1. The muon trigger efficiency for  $t\bar{t} \rightarrow \mu + jets$  events is evaluated in the simulation by folding in the trigger parameterization per muon, given by Eq. 5.3, and deriving the event trigger efficiency separately for the L1 and the L2 trigger requirements:

$$\begin{aligned}\varepsilon_{\text{L1 muon trigger}} &= 98.68 \pm 0.23 \text{ (stat)}_{-1.7}^{+1.3} \text{ (syst)}\%, \\ \varepsilon_{\text{L2 muon trigger}} &= 96.95 \pm 0.35 \text{ (stat)}_{-4.4}^{+3.2} \text{ (syst)}\%.\end{aligned}$$

The origin of the systematic uncertainty on the trigger efficiency is discussed in Sect. 5.1.1.

### 7.1.17 Total $t\bar{t}$ Preselection Efficiency

The details of the preselection efficiency measurement are given in the previous sections. The total preselection efficiency is determined by applying all cuts discussed in Sect. 7.1.1-7.1.16 and the corresponding MC-to-data correction factors, and is found to be:

$$\varepsilon_{\text{presel}} = 12.32 \pm 0.23 \text{ (stat)}\%.$$

A summary of the individual efficiencies and MC-to-data correction factors is listed in Table 7.12. The efficiencies are given with respect to  $t\bar{t} \rightarrow Wq\bar{q}'b\bar{b}$ , where the following  $W$  boson decays are allowed:  $W \rightarrow \mu\nu_\mu$  or  $W \rightarrow \tau\nu_\tau$  with  $\tau \rightarrow \mu\nu_\mu\nu_\tau$ .

cut or $\kappa$	exclusive cut efficiency	cumulative cut efficiency
$N_{\text{jet}, p_T > 20 \text{ GeV},  \eta  < 2.5} \geq 4$	$46.45 \pm 0.36$	$46.45 \pm 0.36$
MuonID, acceptance, cosmic veto	$76.01 \pm 0.45$	$35.31 \pm 0.35$
Muon track match	$97.21 \pm 0.20$	$34.33 \pm 0.34$
$ \eta_\mu  < 2.0$	$99.85 \pm 0.05$	$34.28 \pm 0.34$
$p_{T_\mu} > 20 \text{ GeV}$	$71.97 \pm 0.56$	$24.67 \pm 0.31$
$\Delta R(\mu, \text{jet}) > 0.5$	$78.42 \pm 0.60$	$19.34 \pm 0.29$
Muon track $\chi_{\text{trk}}^2 < 4$	$99.84 \pm 0.07$	$19.31 \pm 0.29$
$ dca /\sigma_{dca} < 3$	$94.69 \pm 0.37$	$18.29 \pm 0.28$
Tight muon isolation	$85.41 \pm 0.60$	$15.62 \pm 0.26$
$\cancel{E}_T - \Delta\phi$ cut	$89.19 \pm 0.54$	$13.93 \pm 0.25$
Second muon veto	$99.92 \pm 0.05$	$13.92 \pm 0.25$
$Z \rightarrow \mu\mu$ Veto	$99.82 \pm 0.07$	$13.89 \pm 0.25$
Electron veto	$99.89 \pm 0.06$	$13.87 \pm 0.25$
PV found	$100 \pm 0$	$13.87 \pm 0.25$
$ z_{PV}  < 60 \text{ cm}$	$98.71 \pm 0.23$	$13.71 \pm 0.25$
tracks attached to PV $> 2$	$99.85 \pm 0.08$	$13.68 \pm 0.25$
$\Delta z(\text{D}\cancel{O}\text{recoPV}, \text{D}\cancel{O}\text{recoPV})$	$99.42 \pm 0.15$	$13.61 \pm 0.25$
$ \Delta z(\mu, PV)  < 1 \text{ cm}$	$99.92 \pm 0.05$	$13.60 \pm 0.25$
Trigger efficiency	$95.43 \pm 0.41$	$12.97 \pm 0.24$
$\kappa_{\mu \text{ ID} \times \text{acc} \times \text{cosmic veto}}$	1.005	$13.03 \pm 0.24$
$\kappa_{\mu \text{ track}}$	0.988	$12.88 \pm 0.24$
$\kappa_{\chi_{\text{trk}}^2 < 4}$	0.986	$12.70 \pm 0.24$
$\kappa_{ dca /\sigma_{dca} < 3}$	0.995	$12.63 \pm 0.24$
$\kappa_{\text{Rat11} < 0.08 \text{ and Rattrk} < 0.06}$	0.982	$12.41 \pm 0.23$
$\kappa_{PV}$	0.994	$12.33 \pm 0.23$
$\kappa_{ \Delta z(\mu, PV)  < 1 \text{ cm}}$	0.999	$12.32 \pm 0.23$
$\epsilon_{\text{presel}}$	-	$12.32 \pm 0.23$

Table 7.12: Summary of the  $t\bar{t} \rightarrow \mu + \text{jets}$  event preselection efficiencies and the corresponding MC-to-data correction factors ( $\kappa$ ), if applicable. Only the statistical uncertainties are quoted.

## 7.2 Contributions to the Data Set

The preselection, as described in Sect. 7.1, serves to enhance the signal fraction in the selected data sample, however, not only  $t\bar{t}$  events survive the preselection. The expected number of  $t\bar{t}$  signal events and the remaining backgrounds are discussed in this section. The backgrounds can be subdivided into two components:

- **Physics background:** Physics processes with the same objects in the final state: an isolated high  $p_T$  muon,  $\cancel{E}_T$ , and  $\geq 4$  jets. The dominant contribution comes from the electroweak  $W$  production, accompanied by four or more jets ( $W$ -plus-jets). Minor contributions arise from the electroweak  $Z$  production, accompanied by four or more jets ( $Z$ -plus-jets), where one muon is not in the acceptance region or not reconstructed, leading to  $\cancel{E}_T$ , and  $t\bar{t}$  dilepton events with a muon from the leptonic decay of one  $W$  boson, and the lepton from the other  $W$  boson decay not in the acceptance region or not reconstructed, see Sect. 2.3.
- **Instrumental background:** QCD-multijets background, strong production of four or more jets, with fake  $\cancel{E}_T$  reconstruction and fake muon isolation. The QCD-multijets background is essentially due to heavy flavor QCD events where the muon originates from a semi-leptonic heavy quark decay and the associated jet is not reconstructed, which leads to the fake isolation of the muon. It also leads to a mismeasurement of the  $\cancel{E}_T$ , since the energy deposition of the unreconstructed jet in the calorimeter is not corrected for the JES as it should be. A mismeasurement of the muon momentum can also lead to fake  $\cancel{E}_T$ . There is also a neutrino in the semi-muonic heavy quark decay, which is reconstructed as real  $\cancel{E}_T$ .

### 7.2.1 $t\bar{t}$ Signal

The expected number of  $t\bar{t}$  signal events, for an assumed  $t\bar{t}$  cross section of 7 pb (see also Table 2.2), is given by:

$$N_{\mu+jets}^{expected} = \sigma_{t\bar{t}} \cdot \epsilon_{\mu+jets} \cdot BR_{\mu+jets} \cdot L, \quad (7.8)$$

where the branching ratio is given in Table 2.6, the integrated luminosity  $L$  is given in Table 6.2, and the preselection efficiency is given in Table 7.13 as a function of the jet multiplicity. The number of expected  $t\bar{t}$  events are also given in Table 7.13 as a function of the jet multiplicity.

	$\varepsilon_{\mu+jets}$ [%]	$\varepsilon_{dilepton}$ [%]	$N_{\mu+jets}^{expected}$	$N_{dilepton}^{expected}$
$N_{jet} = 1$	0.41	3.29	1.12	3.36
$N_{jet} = 2$	4.00	10.71	10.93	10.93
$N_{jet} = 3$	11.70	5.05	31.97	5.15
$N_{jet} \geq 4$	12.32	0.98	33.66	1.00

Table 7.13: Preselection efficiencies for the  $t\bar{t}$  signal ( $\mu$ -plus-jets) and  $t\bar{t}$  dilepton as a function of the jet multiplicity in %. The resulting expected number of events are also given, calculated according to Eq. 7.8 and Eq. 7.9, assuming a cross section of 7 pb. The  $t\bar{t}$  all-jets channel does not lead to a measurable contribution to the preselected sample.

### 7.2.2 $t\bar{t}$ Dilepton Background

The expected number of events from the  $t\bar{t}$  dilepton channel,  $t\bar{t} \rightarrow ll'\nu_l\nu_{l'}\bar{b}\bar{b}$ , where both  $W$  bosons decay leptonically, is given by:

$$N_{dilepton}^{expected} = \sigma_{t\bar{t}} \cdot \varepsilon_{dilepton} \cdot BR_{dilepton} \cdot L. \quad (7.9)$$

The preselection efficiency, given in Table 7.13 as a function of jet multiplicity, is determined with respect to  $t\bar{t} \rightarrow WWb\bar{b}$ , where both  $W$  bosons decay leptonically and at least one of them decays via  $W \rightarrow \mu\nu_\mu$  or  $W \rightarrow \tau\nu_\tau$  with  $\tau \rightarrow \mu\nu_\mu\nu_\tau$ . The corresponding branching ratio is calculated to be  $BR = 6.4$  %. The integrated luminosity  $L$  is given in Table 6.2 and the  $t\bar{t}$  cross section is assumed to be 7 pb (see also Table 2.2). The number of expected  $t\bar{t}$  dilepton events are also given in Table 7.13 as a function of jet multiplicity.

The estimated ratio of number of expected events,  $N_{dilepton}^{expected}/N_{\mu+jets}^{expected} = 3.0$  %.

### 7.2.3 Matrix Method - Evaluation of the QCD Background

The expected number of QCD background events is entirely derived from data by defining two samples of events, a loose and a tight set, the latter being a subset of the first. The tight sample ( $N_t$ ) corresponds to the preselection sample, in the loose sample ( $N_\ell$ ) the tight muon isolation requirement is removed, but the loose muon isolation cut ( $\Delta R(\mu, jet) > 0.5$ ) is kept. The number of events in which the muon isolation is real, originating from either a  $t\bar{t}$ , a  $W$ -plus-jets, a  $Z$ -plus-jets or a dilepton event, is denoted by  $N^{sig}$ , the number of events where the muon isolation is fake is denoted by  $N^{QCD}$ . Then  $N_\ell$  and  $N_t$  can be written as:

$$\begin{aligned} N_\ell &= N_\ell^{sig} + N_\ell^{QCD} \\ N_t &= \varepsilon_{sig} N_\ell^{sig} + \varepsilon_{QCD} N_\ell^{QCD} \end{aligned} \quad (7.10)$$

where  $\varepsilon_{sig}$  is the efficiency for a loosely isolated muon from a  $W$  or  $Z$  boson decay to pass also the tight isolation,  $\varepsilon_{QCD}$  is the (fake) rate at which a loosely isolated muon in QCD events appears to be tightly isolated. Solving the linear system in Eq. 7.10 for  $N_\ell^{sig}$  and  $N_\ell^{QCD}$  yields:

$$N_\ell^{sig} = \frac{N_t - \varepsilon_{QCD}N_\ell}{\varepsilon_{sig} - \varepsilon_{QCD}} \quad \text{and} \quad N_\ell^{QCD} = \frac{\varepsilon_{sig}N_\ell - N_t}{\varepsilon_{sig} - \varepsilon_{QCD}}, \quad (7.11)$$

and, multiplying with the respective efficiencies,

$$N_t^{sig} = \varepsilon_{sig} \frac{N_t - \varepsilon_{QCD}N_\ell}{\varepsilon_{sig} - \varepsilon_{QCD}} \quad \text{and} \quad N_t^{QCD} = \varepsilon_{QCD} \frac{\varepsilon_{sig}N_\ell - N_t}{\varepsilon_{sig} - \varepsilon_{QCD}}. \quad (7.12)$$

With  $N_\ell$  and  $N_t$  measured in the data sample,  $N_t^{QCD}$  can be determined by measuring  $\varepsilon_{sig}$  and  $\varepsilon_{QCD}$ .

$\varepsilon_{QCD}$  is measured in the low- $\cancel{E}_T$ -QCD data sample, which has the same requirements as for the preselection but without applying the  $\cancel{E}_T$ -related set of cuts explained in Sect. 7.1.10. Instead, the following two cuts are applied

- $\cancel{E}_T < 10$  GeV,
- $\cancel{E}_{TCAL}$  ( $=\cancel{E}_T$  without muon correction)  $< 25$  GeV,

in order to veto events with real  $\cancel{E}_T$ , in particular  $W$ -plus-jets events (with  $W \rightarrow \mu\nu$ ) for which  $\cancel{E}_{TCAL}$  represents the  $p_T$  of the  $W$  boson. The statistics of this sample is very limited. In order to increase the data set also jets between 15 GeV and 20 GeV are considered.  $\varepsilon_{QCD}$  is found to be constant as a function of  $\cancel{E}_T$  for  $0 \text{ GeV} < \cancel{E}_T < 10$  GeV, shown in Fig. 7.9. The jet

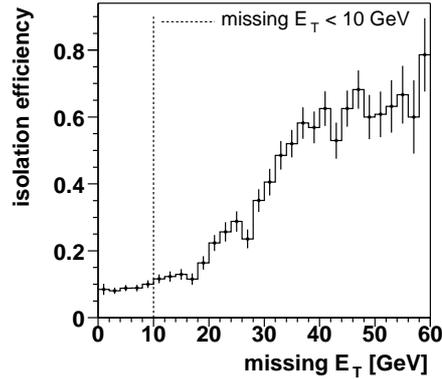


Figure 7.9: Isolation efficiency as a function of  $\cancel{E}_T$  after applying all preselection cuts but the  $\cancel{E}_T$  cut, and requiring  $\cancel{E}_{TCAL} < 25$  GeV.  $\varepsilon_{QCD}$  is determined from the integrated efficiency for  $\cancel{E}_T < 10$  GeV.

multiplicity dependence of  $\varepsilon_{QCD}$  is given in Table 7.14.  $\varepsilon_{QCD}$  varies as a

function of  $|\eta_\mu|$ , as shown in Fig. 7.10, which is due to the  $|\eta|$  dependent performance of the calorimeter and the tracking system to measure the isolation. Since the effect of this variation on the measured cross section is found to be small it is not considered, instead a systematic uncertainty on  $\varepsilon_{QCD}$  is determined from this variation. The  $\pm 1\sigma$  values are given in the last two lines of Table 7.14 by a constant fit to  $\varepsilon_{QCD}$  for  $0.8 < |\eta_\mu| < 1.3$  and for ( $|\eta_\mu| < 0.8$  or  $|\eta_\mu| > 1.3$ ). The central value, derived from a constant fit to the whole  $|\eta_\mu|$  distribution is found to be:

$$\varepsilon_{QCD} = 8.5_{-3.0}^{+3.4} \text{ (syst)\%}.$$

$\varepsilon_{sig}$  is dependent on the muon  $p_T$ , as shown in Fig. 7.10.  $t\bar{t}$  has a higher muon  $p_T$  spectrum than  $W$ -plus-jets, resulting in a higher inclusive tight isolation efficiency. The jet multiplicity dependence of  $\varepsilon_{sig}$  is shown in Table 7.14 for  $t\bar{t}$ , dilepton and  $W$ -plus-jets events, the corresponding efficiency for  $Z$ -plus-jets events is given in Table 7.6. The efficiency decreases with higher jet multiplicities and is very similar for all samples considered. The dilepton and  $Z$ -plus-jets contributions can be neglected as shown in Sect. 7.2.5 and Sect. 7.2.2. In the jet multiplicity bins one to three the  $t\bar{t}$  fraction can be neglected as well, in the fourth bin the expected  $t\bar{t}$  fraction from Eq. 7.8 is used to determine  $\varepsilon_{sig}$ . The systematic uncertainty on  $\varepsilon_{sig}$  for  $\geq 4$  jets is derived by setting the  $t\bar{t}$  fraction to 0 and to 1:

$$\varepsilon_{sig} = 81.0_{-1.7}^{+2.1} \text{ (syst)\%}.$$

$N_l$ ,  $N_t$ ,  $N_t^{sig}$  and  $N_t^{QCD}$  are given in Table 7.15. Figure 7.11 shows the  $W$  transverse mass distribution of the events in the tight sample for each jet multiplicity bin.

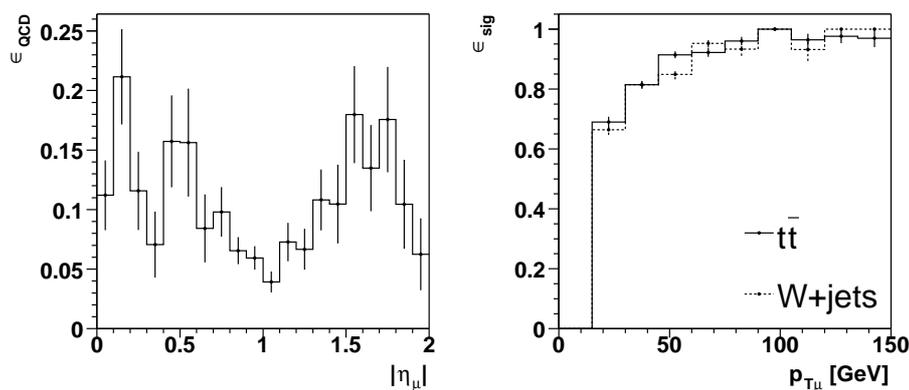


Figure 7.10:  $\varepsilon_{QCD}$  as a function of  $|\eta_\mu|$  and  $\varepsilon_{sig}$  as a function of  $p_{T\mu}$ .

	$\varepsilon_W$ [%]	$\varepsilon_{t\bar{t}}$ [%]	$\varepsilon_{sig}^{uncor}$ [%]	$\varepsilon_{sig}$ [%]	$\varepsilon_{QCD}$ [%]
$N_{jet} = 1$	$90.0 \pm 0.7$	-	$90.0 \pm 0.7$	$88.4 \pm 0.7$	$8.3 \pm 0.6$
$N_{jet} = 2$	$88.6 \pm 0.3$	$87.5 \pm 0.5$	$88.6 \pm 0.3$	$87.0 \pm 0.3$	$9.8 \pm 1.0$
$N_{jet} = 3$	$85.6 \pm 1.6$	-	$85.6 \pm 1.6$	$84.1 \pm 1.6$	$7.1 \pm 2.0$
$N_{jet} \geq 4$	$80.8 \pm 0.9$	$84.6 \pm 0.7$	$82.5 \pm 0.8$	$81.0 \pm 0.8$	$8.6 \pm 4.7$
$0.0 <  \eta_\mu  < 2.0$	-	-	-	-	$8.5 \pm 0.5$
$0.8 <  \eta_\mu  < 1.3$	-	-	-	-	$5.5 \pm 0.5$
$ \eta_\mu  < 0.8$ or $ \eta_\mu  > 1.3$	-	-	-	-	$11.9 \pm 0.9$

Table 7.14: Tight muon isolation efficiency measured for the signal,  $W$ -plus-jets and  $t\bar{t}$  ( $t\bar{t} \rightarrow l\nu_l\nu_l b\bar{b}$  for  $N_{jet} = 2$  and  $t\bar{t} \rightarrow l\nu_l q\bar{q}' b\bar{b}$  for  $N_{jet} \geq 4$ ), in the respective Monte Carlo samples, as a function of jet multiplicity.  $\varepsilon_{sig}^{uncor}$  and  $\varepsilon_{sig}$ , to data corrected by multiplication with  $\kappa_{tight\ isolation}$ , see Sect. 7.1.9. Also the tight muon isolation efficiency for the QCD background in the low- $\cancel{E}_T$ -QCD data sample,  $\varepsilon_{QCD}$ , is given, and the fit result of a constant polynomial for different  $|\eta_\mu|$  regions is shown in the last three rows. Only the statistical uncertainties are shown.

	$N_l$	$N_t$	$N_t^{sig}$	$N_t^{QCD}$
$N_{jet} = 1$	11712	7195	$6859.0 \pm 87.4$	$336.0 \pm 16.2$
$N_{jet} = 2$	3717	2041	$1911.8 \pm 46.3$	$129.2 \pm 6.6$
$N_{jet} = 3$	711	356	$328.8 \pm 19.4$	$27.2 \pm 2.7$
$N_{jet} \geq 4$	149	80	$75.2 \pm 9.2$	$4.8 \pm 1.0$

Table 7.15: Number of loose and tight (preselected) events and the estimated number of non-QCD ( $N_t^{sig}$ ) and QCD events from the matrix method 7.2.3.

## 7.2.4 $W$ -plus-jets Background

The matrix method, discussed in Sect. 7.2.3 not only provides the number of QCD-multijets background events, but trivially also the number of non-QCD events, see Eq. 7.12. These consist of  $W$ -plus-jets,  $Z$ -plus-jets,  $t\bar{t}$  signal and  $t\bar{t}$  dilepton events, as mentioned above. The  $t\bar{t}$  cross section can be calculated theoretically very precisely, see Sect. 2.1.3.3, and the number of expected  $t\bar{t}$  signal and  $t\bar{t}$  dilepton events are discussed in Sect. 7.2.1 and Sect. 7.2.2, respectively.

The  $W$ -plus-jets cross section as a function of jet multiplicity is not known precisely. For this reason the expected number of  $W$ -plus-jets events is determined using the data as explained in Sect. 7.3.

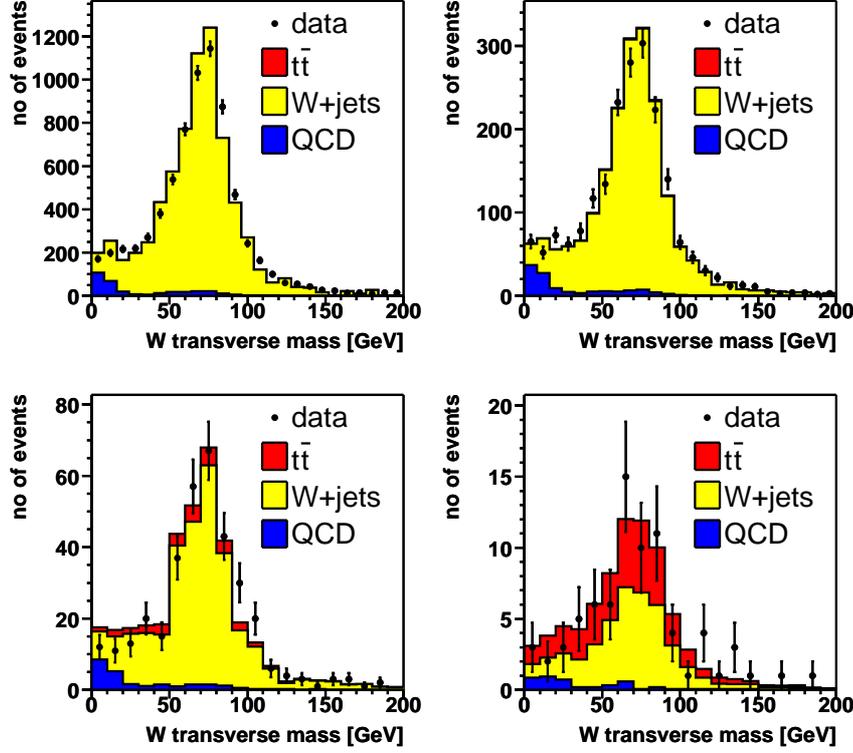


Figure 7.11: The  $W$  transverse mass distribution of the events in the tight sample for  $N_{jet} = 1$ ,  $N_{jet} = 2$ ,  $N_{jet} = 3$ ,  $N_{jet} \geq 4$  (from top left to bottom right), with  $t\bar{t}$  normalized assuming a  $t\bar{t}$  cross section of 7 pb.

### 7.2.5 $Z$ -plus-jets Background

As mentioned in Sect. 7.2.4, the absolute cross section of  $W$ -plus-jets events as a function of jet multiplicity is not known precisely, the same is true for  $Z$ -plus-jets. However, the relative cross sections for  $W$ -plus-jets and  $Z$ -plus-jets production is known very well, see Table 7.16.

$\frac{\sigma_{Zjj} \cdot BR_{Z \rightarrow \mu\mu}}{\sigma_{Wjj} \cdot BR_{W \rightarrow \mu\nu}}$	$\varepsilon_{Wjj}$	$\varepsilon_{Zjj}$
0.1	3.1%	1.9%

Table 7.16: Relative  $W_{jj}$  to  $Z_{jj}$  cross section  $\times$  branching ratio from ALPGEN and the total preselection efficiencies.

Using Eq. 7.5 and the numbers from Table 7.16, the estimated ratio of

number of expected events,  $Z$ -plus-jets /  $W$ -plus-jets = 6.2 %.

### 7.3 Topological Multivariate Analysis

The preselection, as presented in Sect. 7.1, is applied to extract the  $W \rightarrow \mu\nu$ + jets events with maximum efficiency while rejecting the QCD-multijets background. In Sect. 7.2 the contributions of various physics processes to the data set are discussed, and some of the backgrounds are already estimated. However, the contribution from the largest background source, the electroweak production of a  $W$  boson in association with four or more jets remains to be evaluated using the data, as mentioned in Sect. 7.2.4.

The technique of choice is to define a multivariate variable with a maximum discrimination power between the signal and the backgrounds. The number of selected events from the signal and the various background contributions are extracted by fitting their relative contributions by a shape comparison of the discriminating variable.

The discrimination of the  $t\bar{t}$  signal and the backgrounds has two main reasons:

- Due to the large top quark mass of 178 GeV, its decay products are high-energetic, as compared to additional jets produced in  $W$ -plus-jets,  $Z$ -plus-jets or QCD-multijets background events, which originate from QCD bremsstrahlung, and thus tend to be very low-energetic, see Fig. 2.19.
- At the Tevatron with a  $p\bar{p}$  center-of-mass energy of 1.96 TeV, the  $t\bar{t}$  quark pair is mainly produced just above the kinematically allowed threshold, see Sect. 2.1.3.3 and Fig. 2.11. Thus, the  $t\bar{t}$  quark pair is mainly produced at rest, resulting in an isotropic and central distribution of the decay products, whereas additional jet production for the  $W$ -plus-jets,  $Z$ -plus-jets and QCD-multijets backgrounds originates from QCD bremsstrahlung which is collinear divergent and therefore peaks in the forward direction, see Fig. 2.19, Sect. 2.3.2 and Sect. 2.3.3.

A third class of discrimination between the  $t\bar{t}$  signal and the backgrounds uses the fact that jets in a  $t\bar{t}$  event originate mainly from quarks, jets in a background event originate mainly from gluons. The resulting shape of the jets in the calorimeter can be used as a discriminator. However, studies [148] have shown that the MC description of the jet shapes is not sufficient, and these variables are not used.

The modeling of the contributing samples to the data set is discussed in Sect. 7.3.1, the topological variables considered are presented in Sect. 7.3.2, the construction of the multivariate discriminant is described in Sect. 7.3.3,

the cross-section extraction procedure is illustrated in Sect. 7.3.4, the optimization of the discriminant is explained in Sect. 7.3.5 and its result is given in Sect. 7.3.6.

### 7.3.1 Modeling of Signal and Backgrounds

For the shape comparison of the discriminating variable, discussed above, all the contributions to the data set have to be modeled in order to extract the shape of the discriminating variable. The following samples are used:

**$t\bar{t}$  signal:** The  $t\bar{t}$  signal is modeled by the MC simulation, described in Sect. 2.4.1 and by incorporating all the MC-to-data correction factors discussed in Sect. 7.1 as weight factors.

**$W$ -plus-jets background:** The  $W$ -plus-jets background is modeled by the  $Wjjjj$  MC simulation with the factorization scale  $Q^2 = M_W^2 + \sum p_{T_j}^2$  and including the Tune A underlying event description, as described in Table 2.8 and in Sect. 2.4.2, and by incorporating all the MC-to-data correction factors discussed in Sect. 7.1 as weight factors.

**$Z$ -plus-jets background:** The  $Z$ -plus-jets background is modeled by the  $Zjjjj$  MC simulation, described in Sect. 2.4.3 and by incorporating all the MC-to-data correction factors discussed in Sect. 7.1 as weight factors.

**$t\bar{t}$  dilepton:** The  $t\bar{t}$  dilepton contribution is modeled by the MC simulation, described in Sect. 2.4.1 and by incorporating all the MC-to-data correction factors discussed in Sect. 7.1 as weight factors.

**QCD-multijets background:** The QCD-multijets background is described by a data sample selected by applying the preselection (Sect. 7.1), but with inverted muon isolation cuts:

- $\text{Rat11} \geq 0.08$  or  $\text{Rattrk} \geq 0.06$ .

### 7.3.2 Topological Variables

Topological variables are defined aimed at the discrimination of  $t\bar{t}$  and background events, as discussed above. For the calculation of these variables the four-vectors of the final state objects and their correlations are used.

More than two dozens of such variables are constructed, they are summarized in Table 7.17. Only the four leading jets are used to construct the topological variables. This does not reduce the statistical separation power, but it reduces the dependence on systematic effects on the modeling of soft radiation (e.g. underlying event, initial and final state radiation via parton shower). Optionally, also the four-vectors of the muon, the neutrino

Variable	Definition
$H_T$	Scalar sum of transverse energies of the four leading jets.
$H_T^3$	Scalar sum of transverse energies of the third and fourth leading jet.
$N_{JW}$	Number of jets averaged over a range of $E_T$ threshold (15-55 GeV), and weighted by the $E_T$ threshold. This parameterizes the number of jets, taking their hardness into account.
$M_{jj}^{min}$	Minimum di-jet invariant mass.
$E_T^1$	$E_T$ of the jet with highest $E_T$ .
$E_T^2$	$E_T$ of the jet with the next to highest $E_T$ .
$\Delta R_{jj}^{min} E_T^{min}$	Product of minimum di-jet separation in $R$ and $E_T$ of the less energetic jet of that pair.
$M^j$	Invariant mass of the four leading jets.
$M_T^j$	Transverse mass of the four leading jets.
$\cancel{E}_T$	Missing transverse energy.
$p_T^\mu$	$p_T$ of the muon.
$\eta^\mu$	$\eta$ of the muon.
Aplanarity $\mathcal{A}$	$3/2\lambda_3$ .
Sphericity $\mathcal{S}$	$3/2(\lambda_2 + \lambda_3)$ .
Centrality $\mathcal{C}$	$H_T/H$ , the scalar sum of transverse energies divided by the scalar sum of energies of the four leading jets.
$H_T'$	$H_T/H_z$ , the scalar sum of transverse energies of the four leading jets divided by the scalar sum of the longitudinal energies of the four leading jets, the muon and an hypothesis for the neutrino.
$H_T^{2'}$	$H_T^2/H_z$ , the scalar sum of transverse energies of the four leading jets, except the first one, divided by the scalar sum of the longitudinal energies of the four leading jets, the muon and an hypothesis for the neutrino.
$K_T^{min'}$	$\Delta R_{jj}^{min} E_T^A/E_T^W$ , product of minimum di-jet separation in $R$ and $E_T$ of the less energetic jet of that pair, divided by the scalar sum of the transverse energies of the muon and the $\cancel{E}_T$ .
$\Delta R_{jj}^{min}$	Minimum di-jet separation in $R$ .
$\Delta\phi(j^1, \cancel{E}_T)$	Azimuthal opening angle between the leading jet and the $\cancel{E}_T$ .
$\Delta\phi(\mu, \cancel{E}_T)$	Azimuthal opening angle between the muon and the $\cancel{E}_T$ .
$\eta^{max}$	$ \eta $ of the jet with maximum $ \eta $ .
$\eta^1$	$ \eta $ of the jet with maximum $E_T$ .
$\langle\eta^2\rangle^2$	Average $\eta^2$ of the two leading jets.
$\langle\eta^2\rangle^3$	Average $\eta^2$ of the three leading jets.
$\langle\eta^2\rangle^4$	Average $\eta^2$ of the four leading jets.

Table 7.17: Definition of topological variables considered. The normalized momentum tensor is defined in Eq. 7.25 and the three eigenvalues are ordered such that  $\lambda_1 \geq \lambda_2 \geq \lambda_3$ , with  $\lambda_1 + \lambda_2 + \lambda_3 = 1$ . The sets of variables correspond to variables proportional to the energy present in the event, muon kinematic variables, event shape variables, ratios of energy dependent variables and angular variables.

or the  $W$  boson reconstructed from the leptons with a hypothesis of the neutrino  $p_z$ , are used. Additionally, the minimum  $E_T$  and the maximum  $|\eta|$  of the objects considered to define the topological variable are varied. The definition is chosen such that the discrimination between  $t\bar{t}$  and the dominant background,  $W$ -plus-jets, is maximal. The performance plot of the survival efficiencies for a varying straight cut on the topological variable is used as a measure of the quality of the definition; this is demonstrated for the aplanarity in Fig. 7.12.

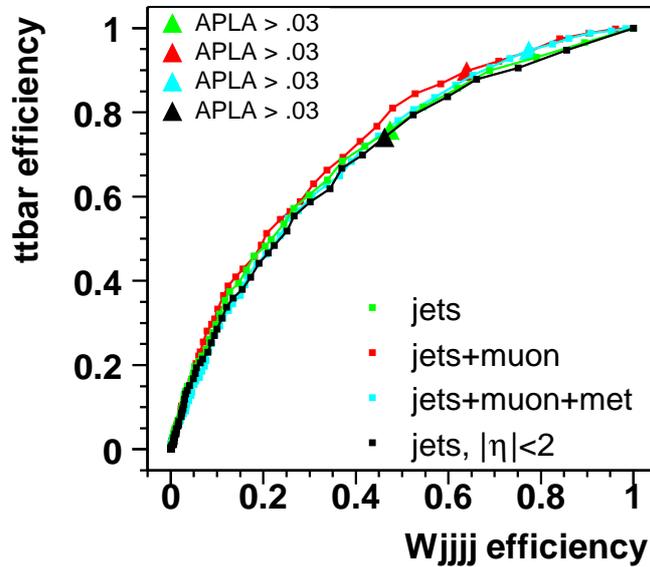


Figure 7.12: Aplanarity performance curves: The survival efficiency for the background ( $W$ -plus-jets) is plotted vs. the survival efficiency for the signal ( $t\bar{t}$ ) for different cut values on the aplanarity. Four different aplanarity definitions are compared, using four different sets of final state objects for their calculation, all jets, all jets plus the muon, all jets plus the muon plus the  $\cancel{E}_T$ , and all jets with  $|\eta| < 2$  (see legend). For convenience the efficiencies for  $\mathcal{A} > 0.03$  are highlighted for all four aplanarity definitions. Including the jets and the  $\mu$  from the  $W$  boson decay in the sum leads to the best separation power.

The variables are not only tested for their discrimination power, but also for their correlation and their sensitivity to systematic uncertainties, consid-

ering the dominant systematic uncertainties originating from the jet energy scale and from the  $W$ -plus-jets Monte Carlo modeling. The systematic uncertainty of the latter is estimated by exchanging the default background MC sample with the factorization scale  $Q^2 = M_W^2 + \sum p_{T_j}^2$  by one with  $Q^2 = \langle p_{T_j} \rangle^2$  (see Sect. 2.4.2). The criteria for the choice of the topological variables can be summarized as:

- Use variables with a strong separation power between signal and background.
- Instead of requiring each single variable to be insensitive against systematic uncertainties, the set of variables is chosen such that the combination of the variables (see Sect. 7.3.3) is robust against the systematic variations. This allows to maintain high statistical sensitivity while reducing systematics, in particular from the JES uncertainty.
- If possible use variables which are not correlated or which have a small correlation. Correlated variables add only new information when the correlation is different for signal and background. However, variables are subject to statistical fluctuations and they enlarge the statistical uncertainties even if they add no new information.

The topological variables are subdivided into two categories, see also Table 7.17:

- Variables which are proportional to the energy present in the event, e.g. the scalar sum of transverse energies of the four leading jets  $H_T = \Sigma E_T$ . These variables have a very good separation power since one peculiarity of the  $t\bar{t}$  events is the large energy which is needed to produce a  $t\bar{t}$  quark pair. However, these variables are also very sensitive to uncertainties in the jet energy scale, which translates into a large systematic uncertainty in the cross-section measurement.
- The second class of variables is characterized by the property of having reduced sensitivity to JES uncertainties. The variables are either defined exclusively from angular qualities, e.g. event shape variables, or by ratios of energy dependent variables where the JES systematic uncertainty largely cancels out.

### 7.3.3 The Topological Likelihood Discriminant

None of the topological variables has enough discrimination power which would allow to place a straight cut on either of them in order to separate the  $t\bar{t}$  signal from the background, see e.g. Fig. 7.12. Therefore the event information contained in the topological variables is combined in a likelihood discriminant. Alternatively, the combination in a neural network, which

takes into account correlations of the variables, was studied. The neural network did not show any improvement in performance over the likelihood and therefore it was not adopted. This section only describes the likelihood discriminant.

A 2-class likelihood with  $t\bar{t}$  as signal and  $W$ -plus-jets as background is chosen. Additional classes are not defined since

- the  $W$ -plus-jets background is found to be the dominant one (80-90%),
- the topological properties for the  $W$ -plus-jets, the  $Z$ -plus-jets and the QCD-multijets background are found to be very similar,
- the  $t\bar{t}$  dilepton background is negligible.

The likelihood discriminant for a set of topological variables, referred to as likelihood input distributions, is built in the following way:

- In order to avoid overtraining of the likelihood only a fraction of the events is used to build the likelihood. It is found that one third of the  $t\bar{t}$  signal and  $W$ -plus-jets background samples is sufficient. However, the full sample is used to evaluate the likelihood discriminant in order to minimize the statistical uncertainty and to guarantee that the same events are used both for the central value and for the systematically varied samples.
- The topological variables are transformed using functions in order to be less sensitive to statistical fluctuations in regions of rapidly varying event counts.
- It is made sure that bins are adequately populated. Bins are combined so that the minimum number of entries per bin is larger than 25.
- Data in the tail of a distribution is consolidated including over- and underflow bins.
- The distributions are normalized to unity and the ratio of these probability density functions for signal over background is built for each distribution:  $S_i/B_i$ .
- The logarithm of the ratios is built and fitted with a polynomial. The fit to the likelihood reduces the sensitivity to single events and further dilutes the statistical dependence between training and evaluation of the likelihood. Fitting the logarithm also simplifies the fit function and symmetrizes the errors on the points.
- A reasonable fit range is chosen. If the abscissa lies beyond this range the fit is evaluated at the range limit.

- The likelihood discriminant definition is motivated by the desire to approximate the optimum discriminant, which in case of  $N$  variables, would be given by:

$$\mathcal{L}_{\text{optimum}} = \frac{S(x_1, \dots, x_N)}{S(x_1, \dots, x_N) + B(x_1, \dots, x_N)}, \quad (7.13)$$

where  $S$  and  $B$  are the signal and background  $N$ -dimensional probability density functions, respectively. Due to the finite statistics available to construct the discriminant, the simplification is made that the  $N$  discriminant variables are uncorrelated and therefore each multidimensional probability density function can be expressed as a product of one-dimensional probability density functions:

$$\begin{aligned} \mathcal{L} &= \frac{\prod_{i=1}^N S_i(x_i)}{\prod_{i=1}^N S_i(x_i) + \prod_{i=1}^N B_i(x_i)} = \frac{\prod_{i=1}^N S_i(x_i)/B_i(x_i)}{\prod_{i=1}^N S_i(x_i)/B_i(x_i) + 1} \\ &= \frac{\exp\left(\sum_{i=1}^N \ln \frac{S_i(x_i)}{B_i(x_i)}\right)}{\exp\left(\sum_{i=1}^N \ln \frac{S_i(x_i)}{B_i(x_i)}\right) + 1}. \end{aligned} \quad (7.14)$$

Denoting the fit to logarithm of the ratios as  $(\ln \frac{S}{B})_i^{\text{fitted}}(x_i)$ , the likelihood discriminant is given by:

$$\mathcal{L} = \frac{\exp\left(\sum_{i=1}^N (\ln \frac{S}{B})_i^{\text{fitted}}(x_i)\right)}{\exp\left(\sum_{i=1}^N (\ln \frac{S}{B})_i^{\text{fitted}}(x_i)\right) + 1}. \quad (7.15)$$

Figure 7.13 demonstrates that the performance curve of various likelihood discriminants is better than the performance curve of a single variable as shown in Fig. 7.12 for the aplanarity. It is also shown that the likelihood discriminants built from the two different categories of topological variables, defined in Sect. 7.3.2, have a different performance in separation power.

### 7.3.4 Cross-Section Extraction Procedure

The number of  $t\bar{t}$  events in the preselected data sample is extracted by describing the distribution of the likelihood discriminant (see Sect. 7.3.3), measured in the preselected data sample, by a linear combination of  $t\bar{t}$ ,  $W$ -plus-jets and QCD-multijets events. The effect of the incorporation of the small  $Z$ -plus-jets and  $t\bar{t}$  dilepton backgrounds is discussed in Sect. 7.5.2.2 and Sect. 7.5.4, respectively. The sensitivity of the discriminating variable to distinguish QCD-multijets and  $W$ -plus-jets events is very small. However, the matrix method (see Sect. 7.2.3) allows to constrain the number of QCD-multijets events.

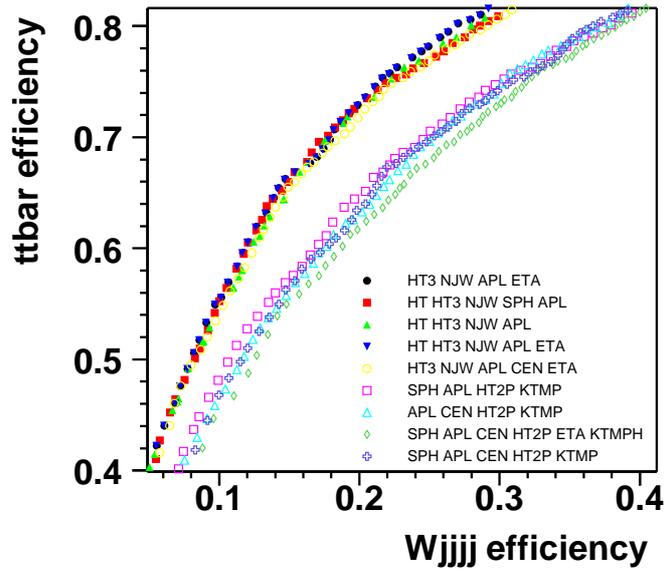


Figure 7.13: Likelihood discriminant performance curves: The survival efficiency for the background ( $W$ -plus-jets) is plotted vs. the survival efficiency for the signal ( $t\bar{t}$ ) for different cut values on the different likelihood discriminants. The discriminants are built from variables given in the legend. The abbreviated names HT, HT3, NJW, APL, SPH, CEN, HT2P, KTMP and ETA correspond to  $H_T$ ,  $H_T^3$ ,  $NJW$ ,  $\mathcal{A}$ ,  $\mathcal{S}$ ,  $\mathcal{C}$ ,  $H_T^{2'}$ ,  $K_T^{min'}$  and  $\eta^\mu$ , respectively. Likelihood discriminants built from any of the variables which are proportional to the energy in the event (HT, HT3 and NJW) show a better performance (see set of curves with lower background efficiency given the same signal efficiency). APL, SPH, CEN, HT2P, KTMP and ETA belong to the category of variables which are relatively immune against JES uncertainties, but the performance is reduced.

A binned likelihood fit (10 bins) to the distribution of the discriminating variable, together with an additional constraint which incorporates the matrix method, is used to derive the most likely number of  $t\bar{t}$  ( $N_t^{t\bar{t}}$ ),  $W$ -plus-jets ( $N_t^W$ ) and QCD-multijets ( $N_t^{QCD}$ ) events contributing to the preselected data sample by maximizing:

$$L(N_t^{t\bar{t}}, N_t^W, N_t^{QCD}) = \left[ \prod_{i=1}^{10} P(n_i^{obs}, \mu_i) \right] \cdot P(N_{\ell-t}^{obs}, N_{\ell-t}), \quad (7.16)$$

where  $P(n, \mu) = \frac{\mu^n e^{-\mu}}{n!}$  generically denotes the Poisson probability density function for  $n$  observed events given an expectation of  $\mu$ .

Since the logarithm is a monotone function, the solution that maximizes the expression in Eq. 7.16 also minimizes:

$$-\ln L(N_t^{t\bar{t}}, N_t^W, N_t^{QCD}) = \left[ \sum_{i=1}^{10} -n_i^{obs} \ln \mu_i + \mu_i \right] - N_{\ell-t}^{obs} \ln N_{\ell-t} + N_{\ell-t}, \quad (7.17)$$

where the terms  $\ln(n_i^{obs}!)$  in the sum and  $\ln(N_{\ell-t}^{obs}!)$  are omitted for simplicity, since they do not change the location of the minimum but lead only to a constant offset.

The number of observed preselected data events that populate the  $i$ -th bin in the likelihood discriminant distribution is denoted by  $n_i^{obs}$ , the corresponding expected number of events is expressed as a function of  $N_t^{t\bar{t}}$ ,  $N_t^W$  and  $N_t^{QCD}$  and is given by:

$$\mu_i(N_t^{t\bar{t}}, N_t^W, N_t^{QCD}) = P_i^{t\bar{t}} N_t^{t\bar{t}} + P_i^W N_t^W + P_i^{QCD} N_t^{QCD}, \quad (7.18)$$

where  $P_i^{t\bar{t}}$ ,  $P_i^W$ , and  $P_i^{QCD}$ , respectively, are the probability density functions for the  $t\bar{t}$ ,  $W$ -plus-jets and QCD-multijets likelihood discriminant distributions evaluated for the  $i$ -th bin.

The second term of Eq. 7.17 is a Poisson constraint on the observed number of events in the ‘‘loose-tight’’ preselected sample and effectively completes the incorporation of the matrix method in the likelihood, and thus, implicitly, the Poisson constraint on the number of QCD-multijets events. By introducing  $N_{\ell-t} = N_{\ell} - N_t$  one makes sure that  $N_{\ell-t}$  and  $N_t$  (implicitly present in the sum over the likelihood discriminant bins) are indeed uncorrelated by construction.

$N_{\ell-t}^{obs} = N_{\ell}^{obs} - N_t^{obs}$  is the number of observed data events after the loose preselection minus the number of observed preselected data events (referred to as ‘‘tight’’ here). The expected number of loose minus tight events,  $N_{\ell-t}$ , can be expressed in terms of  $N_t^{t\bar{t}}$ ,  $N_t^W$  and  $N_t^{QCD}$  by rewriting Eq. 7.10 in the following way:

$$\begin{aligned} N_{\ell} &= N_{\ell}^{t\bar{t}} + N_{\ell}^W + N_{\ell}^{QCD} \\ N_t &= \varepsilon_{sig} N_t^{t\bar{t}} + \varepsilon_{sig} N_t^W + \varepsilon_{QCD} N_t^{QCD}. \end{aligned} \quad (7.19)$$

By substituting  $N_\ell^{t\bar{t}} = \frac{1}{\varepsilon_{sig}} N_t^{t\bar{t}}$ ,  $N_\ell^W = \frac{1}{\varepsilon_{sig}} N_t^W$ , and  $N_\ell^{QCD} = \frac{1}{\varepsilon_{QCD}} N_t^{QCD}$ , and by subtracting the second line from the first line in Eq. 7.19 one obtains:

$$N_{\ell-t} = \frac{1 - \varepsilon_{sig}}{\varepsilon_{sig}} N_t^{t\bar{t}} + \frac{1 - \varepsilon_{sig}}{\varepsilon_{sig}} N_t^W + \frac{1 - \varepsilon_{QCD}}{\varepsilon_{QCD}} N_t^{QCD}. \quad (7.20)$$

The number of preselected  $t\bar{t}$  events,  $N_t^{t\bar{t}}$ , is extracted by inserting Eq. 7.18 and Eq. 7.20 in Eq. 7.17 and by minimizing  $-\ln L(N_t^{t\bar{t}}, N_t^W, N_t^{QCD})$  in Eq. 7.17. The  $t\bar{t}$  production cross-section is then given by:

$$\sigma = \frac{N_t^{t\bar{t}}}{BR \cdot L \cdot \varepsilon_{presel}}, \quad (7.21)$$

where  $BR$  is the branching fraction for the final state considered,  $L$  is the integrated luminosity and  $\varepsilon_{presel}$  is the preselection efficiency.

### 7.3.5 Topological Optimization

The optimization procedure for the choice of the topological variables (see Sect. 7.3.2) used for the construction of the likelihood discriminant (see Sect. 7.3.3) is performed using pseudo-experiments which allow to estimate the expected statistical and dominant systematic uncertainties on the cross-section by the application of the cross-section extraction procedure described in Sect. 7.3.4.

A pseudo-experiment is a representation of the selected data sample, consisting of the same number of events found in the data after the preselection,  $N_t$ , as given in Table 7.15. It is built from the samples used to model the signal and the backgrounds, as described in Sect. 7.3.1, according to the expected contributions. The number of expected  $t\bar{t}$  events is given in Table 7.13, assuming a cross section of 7 pb, the number of expected QCD-multijets events is given by the matrix method given in Table 7.15, and the number of expected  $W$ -plus-jets events is given by  $N_{expected}^W = N_t - N_{expected}^{t\bar{t}} - N_{expected}^{QCD}$ . The  $Z$ -plus-jets background is effectively incorporated by the  $W$ -plus-jets background and the  $t\bar{t}$  dilepton contribution is so small that it has no impact on the optimization.

In each bin of all three (one signal + two backgrounds) likelihood templates, the fluctuations in the number of entries are simulated independently according to the expected number of entries ( $\mu$ ) and the corresponding Poisson distribution  $P(n, \mu)$ . 1000 such pseudo-experiments are simulated. The events for one pseudo-experiment (also called “ensemble”) are drawn randomly from the lists of signal and background events, then the events are put back. Some events will be picked more than once.

The weighted Monte Carlo events are globally reweighted so that the mean of the distribution of the weight factors is unity. This guarantees an optimum usage of the low MC statistics.

The optimization corresponds to the minimization of the figure of merit, which is the expected total uncertainty ( $\sigma_{total}$ ) comprising the expected statistical uncertainty ( $\sigma_{stat}$ ) and the two dominant systematic uncertainties originating from the JES ( $\sigma_{JES}$ ) and the  $W$ -plus-jets Monte Carlo modeling ( $\sigma_{W MC modeling}$ ):

$$\sigma_{total} = \sqrt{(\sigma_{stat}/\sqrt{2})^2 + \sigma_{JES}^2 + \sigma_{W MC modeling}^2}. \quad (7.22)$$

The optimization is performed for a combination of the  $t\bar{t}$  cross-section measurements in the  $\mu$ -plus-jets and electron-plus-jets [149] data samples. Thus, the expected statistical uncertainty for the combined lepton-plus-jets data sample, assumed to have approximately twice the statistics of the  $\mu$ -plus-jets channel alone, is modeled by  $\sigma_{stat}/\sqrt{2}$ . The mean of the statistical error distribution for the 1000 pseudo-experiments, given by the likelihood fit, is used as a measure for the expected statistical uncertainty. The difference in the mean of the distribution of the fitted cross section values using pseudo-experiments with and without systematic variations is used as a measure for the expected systematic uncertainties.

In principal, the optimum likelihood discriminant does not only depend on the likelihood input variables, but also on all preselection cuts, discussed in Sect. 7.1, which can change the expected sample composition and thus the optimum. In particular a variation of the jet multiplicity cut, discussed in Sect. 7.1.1, has a large impact on the expected fraction of signal and backgrounds. As a consequence, both the jet preselection and the topological likelihood input variables are optimized simultaneously. 18 different jet preselections are considered:

$$N_{jet}(p_T > 15 \text{ GeV and } |\eta| < 2.5) \geq 4, \text{ and in addition} \quad (7.23)$$

$$N_{jet}(p_T > p_{Tmin} \text{ and } |\eta| < \eta_{max}) \geq N_j, \quad (7.24)$$

with  $p_{Tmin} = [15 \text{ GeV}, 20 \text{ GeV}, 25 \text{ GeV}]$ ,  $\eta_{max} = [2.5, 2.0, 1.8]$  and  $N_j = [3, 4]$ .

For each of the 18 preselections a topological likelihood discriminant is built from seven likelihood input variables from Table 7.17. Also the topological likelihood discriminants are considered which use any four, five or six out of the seven likelihood inputs, resulting in 64 discriminants per set of seven likelihood input variables. This procedure is repeated for roughly two dozens different sets from Table 7.17, representing all the “reasonable” combinations of likelihood input variables, where “reasonable” refers mainly to the correlation among the variables.

### 7.3.6 Result of the Topological Optimization

The optimum preselection is found to require four jets with  $p_T > 20 \text{ GeV}$  and  $|\eta| < 2.5$ , as given in Sect. 7.1.1. Figure 7.14 shows the expected

systematic uncertainty versus the expected statistical uncertainty for the combined electron and muon sample for one set of seven topological inputs. 64 points are shown, each representing a topological likelihood discriminant built from any four, five, six or seven out of the seven likelihood inputs. The optimum topological likelihood discriminant is represented by the full circle. It is found to be built from the following six variables:

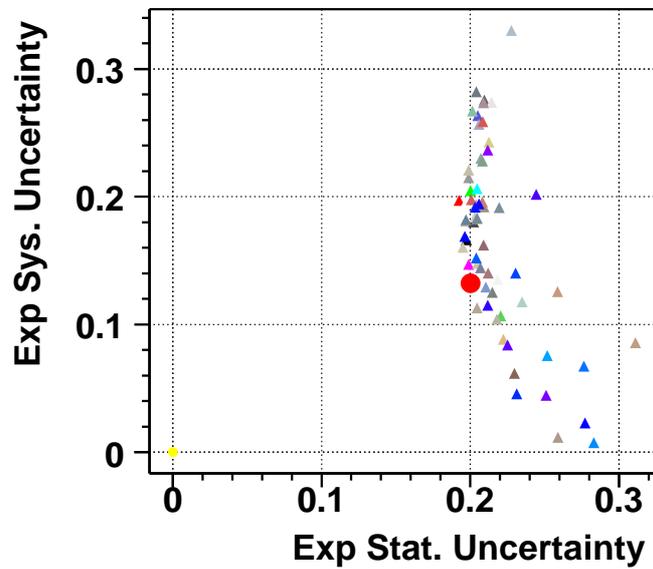


Figure 7.14: Expected systematic uncertainty versus the expected statistical uncertainty (representing the  $\mu$ -plus-jets data statistics) from the likelihood fit for one set of seven topological inputs. One can see 64 points for 64 topological likelihood discriminants. The large point represents the found optimum discriminant.

**Centrality:** The event centrality  $\mathcal{C} = H_T/H$ .  $H_T$  is given by the scalar sum of the  $p_T$  of the four leading jets.  $H$  is the scalar sum of the energy of the four leading jets.

**Aplanarity:** The normalized momentum tensor  $\mathcal{M}$  is defined as:

$$\mathcal{M}_{ij} = \frac{\sum_o p_i^o p_j^o}{\sum_o |\vec{p}^o|^2}, \quad (7.25)$$

where  $p^{\vec{b}}$  is the momentum-vector of a reconstructed object  $o$ ,  $i$  and  $j$  are Cartesian coordinates. By standard diagonalization of  $\mathcal{M}_{ij}$  one may find three eigenvalues  $\lambda_1 \geq \lambda_2 \geq \lambda_3$ , with  $\lambda_1 + \lambda_2 + \lambda_3 = 1$ . The aplanarity  $\mathcal{A}$ , a measure of the flatness of the event, is defined as  $\mathcal{A} = \frac{3}{2}\lambda_3$ , where  $\lambda_3$  is the smallest eigenvalue of the normalized momentum tensor  $\mathcal{M}$  (see Eq. 7.25). Therefore, it is defined in the range  $0 \leq \mathcal{A} \leq 0.5$ . The objects included in the sum are the four leading jets and the  $\mu$  from the  $W$  decay, which have the best discrimination power between  $t\bar{t}$  and  $W$ -plus-jets, see Fig. 7.12. Large values of  $\mathcal{A}$  are indicative of spherical events, whereas small values correspond to more planar events.  $t\bar{t}$  events are quite spherical as is typical for the decay of a heavy object.  $W$ -plus-jets and QCD-multijets events are more planar, primarily due to the fact that the jets in these events arise mainly from initial and final state radiation, and due to the color flow between these jets, defining a plane in which the hadronic activity is localized for the largest part.

**DphiMUMET:**  $\Delta\phi(\mu, \cancel{E}_T)$  is the azimuthal opening angle between the muon and the missing transverse energy.

**Sphericity:** The sphericity  $\mathcal{S}$  of the event is defined as

$$\mathcal{S} = \frac{3}{2}(\lambda_2 + \lambda_3), \quad (7.26)$$

where  $\lambda_2$  and  $\lambda_3$  are the smallest eigenvalues of the normalized momentum tensor  $\mathcal{M}$  (see Eq. 7.25), so that  $0 \leq \mathcal{S} \leq 1$ . Only the four leading jets are included in the sum. Sphericity is essentially a measure of the summed  $p_T^2$  with respect to the event axis; a 2-jet event corresponds to  $\mathcal{S} \approx 0$  and an isotropic event to  $\mathcal{S} \approx 1$ .  $t\bar{t}$  events are quite isotropic as is typical for the decay of a heavy object.  $W$ -plus-jets and QCD-multijets events are less isotropic, primarily due to the fact that the jets in these events arise mainly from QCD bremsstrahlung.

**HT:**  $H_T$  is given by the scalar sum of the  $p_T$  of the four leading jets.

**Ktminp:**  $K'_{Tmin} = \Delta R_{jj}^{min} E_T^{min} / E_T^W$ : provides a measure of the minimum jet  $p_T$  relative to another. In order to reduce the correlation of this variable with the jet energy scale, it is divided by another JES-sensitive variable:  $E_T^W = E_T^{lepton} + \cancel{E}_T$ . Only the four leading jets in an event are considered in the definition of this variable.  $\Delta R_{jj}^{min}$  corresponds to the minimum separation in  $\eta - \phi$  space between a pair of jets and  $E_T^{min}$  is the  $E_T$  of the less energetic jet of that pair.

The probability density functions of the six topological variables are shown in Fig. 7.15 for the  $t\bar{t}$  signal and the  $W$ -plus-jets background. The functions used to transform the likelihood input distributions in order to be less sensitive to statistical fluctuations are:

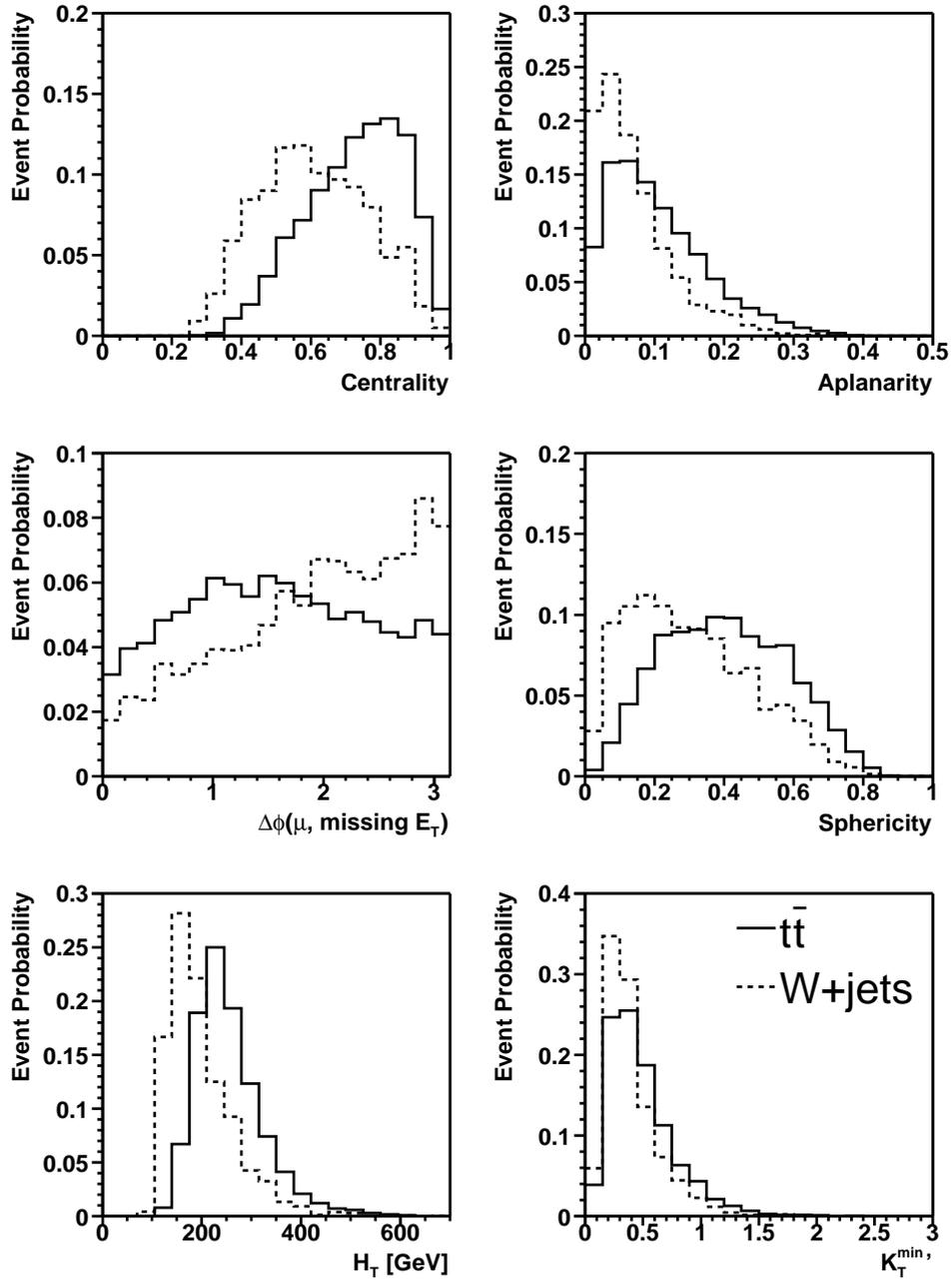
- $\ln(\mathcal{C})$
- $\exp(-11 \cdot \mathcal{A})$
- $\Delta\phi(\mu, \cancel{E}_T)$
- $\ln(\mathcal{S})$
- $\ln(H_T)$
- $\ln(K'_{Tmin})$

The correlation factors among the six variables are summarized in Table 7.18. The logarithm of the ratio of the signal and background prob-

variables	$t\bar{t}$	$W$	QCD	data	prediction
$\exp(-11 \cdot \mathcal{A}), \ln(\mathcal{C})$	-0.29	-0.49	-0.37	-0.40	-0.47
$\exp(-11 \cdot \mathcal{A}), \ln(K'_{Tmin})$	-0.16	-0.12	-0.05	-0.29	-0.16
$\exp(-11 \cdot \mathcal{A}), \ln(\mathcal{S})$	-0.56	-0.61	-0.54	-0.49	-0.61
$\exp(-11 \cdot \mathcal{A}), \Delta\phi(\mu, \cancel{E}_T)$	0.05	0.04	-0.03	0.38	0.06
$\exp(-11 \cdot \mathcal{A}), \ln(H_T)$	0.11	0.09	0.29	-0.18	0.00
$\ln(\mathcal{C}), \ln(K'_{Tmin})$	-0.11	-0.14	-0.27	-0.09	-0.08
$\ln(\mathcal{C}), \ln(\mathcal{S})$	0.42	0.63	0.57	0.62	0.61
$\ln(\mathcal{C}), \Delta\phi(\mu, \cancel{E}_T)$	-0.09	-0.04	-0.05	-0.08	-0.10
$\ln(\mathcal{C}), \ln(H_T)$	0.34	0.23	0.32	0.27	0.36
$\ln(K'_{Tmin}), \ln(\mathcal{S})$	0.09	0.00	-0.07	-0.03	0.06
$\ln(K'_{Tmin}), \Delta\phi(\mu, \cancel{E}_T)$	0.12	0.08	-0.16	-0.18	0.04
$\ln(K'_{Tmin}), \ln(H_T)$	0.07	0.13	-0.01	0.20	0.15
$\ln(\mathcal{S}), \Delta\phi(\mu, \cancel{E}_T)$	-0.01	-0.01	-0.02	-0.11	-0.05
$\ln(\mathcal{S}), \ln(H_T)$	-0.12	0.00	0.04	0.02	0.08
$\Delta\phi(\mu, \cancel{E}_T), \ln(H_T)$	-0.14	-0.22	-0.01	-0.27	-0.22

Table 7.18: Correlation factors for the six topological variables for  $t\bar{t}$ ,  $W$ -plus-jets, QCD-multijets, data and the prediction from the fit to the likelihood distribution (modeled by the corresponding combination of  $t\bar{t}$ ,  $W$ -plus-jets and QCD-multijets).

ability density functions for each of the six likelihood input variables and a polynomial fit to the ratio is shown in Fig.7.16. These fits are the inputs to build the likelihood function according to Eq. 7.15. This likelihood function is then evaluated for the  $t\bar{t}$  signal and all the backgrounds. The resulting normalized likelihood discriminant distributions, referred to as “templates”, represent the probability density for an event to be signal- or background-like. The likelihood discriminant templates are shown in Fig. 7.17 for  $t\bar{t}$  and  $W$ -plus-jets.

Figure 7.15: Likelihood input distributions for  $t\bar{t}$  and  $W$ -plus-jets.

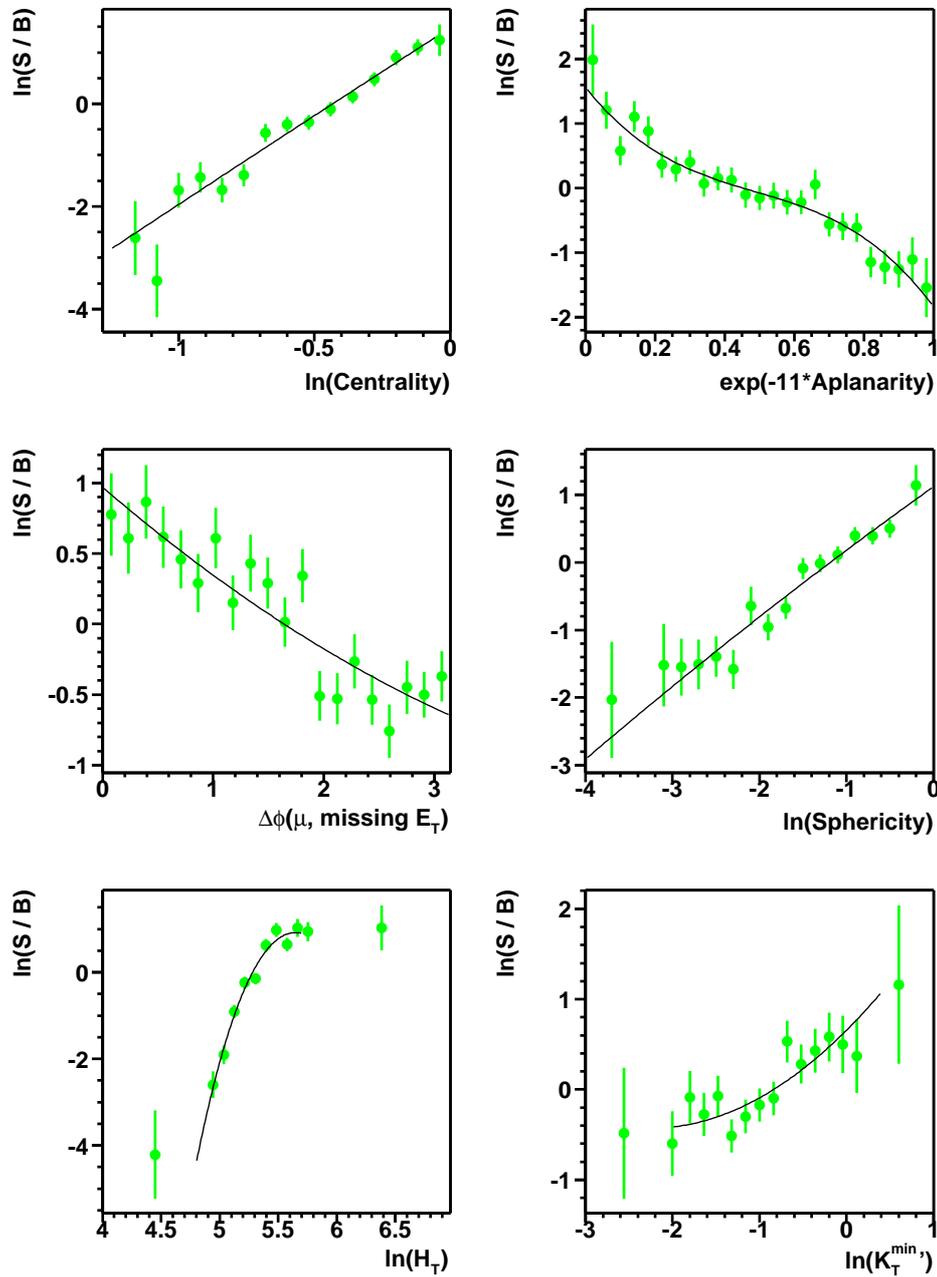


Figure 7.16: Polynomial fits to the logarithm of the ratio of the signal ( $t\bar{t}$ ) over background ( $W$ -plus-jets) probability density functions for each of the six likelihood input variables.

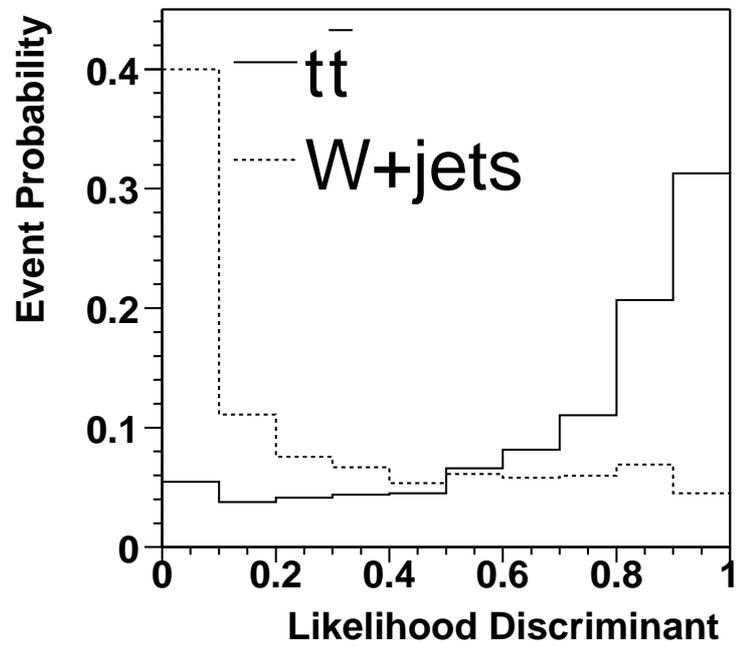


Figure 7.17: Likelihood discriminant templates for  $t\bar{t}$  and  $W$ -plus-jets.

## 7.4 The $t\bar{t}$ Production Cross-Section

The  $t\bar{t}$  production cross-section is determined according to Eq. 7.21, where  $N_t^{t\bar{t}}$  is the number of fitted  $t\bar{t}$  events,  $BR$  is the branching fraction for  $t\bar{t} \rightarrow Wq\bar{q}'b\bar{b}$ , where the  $W$  boson is allowed to decay according to  $W \rightarrow \mu\nu_\mu$  or  $W \rightarrow \tau\nu_\tau$  with  $\tau \rightarrow \mu\nu_\mu\nu_\tau$ ,  $L$  is the integrated luminosity  $\int \mathcal{L}$  and  $\varepsilon_{preselect}$  is the preselection efficiency.

The input values used for the likelihood fit are summarized in Table 7.19.

$N_{\bar{t}}$	$N_t$	$BR$	$L$	$\varepsilon_{preselect}$	$\varepsilon_{sig}$	$\varepsilon_{QCD}$
149	80	17.0%	229.1 pb <sup>-1</sup>	12.3%	81.0%	8.5%

Table 7.19: Inputs for the likelihood fit.

The result of the likelihood fit is shown in Fig. 7.18. The Kolmogorov-Smirnov matching probabilities (KS) for the agreement between the observation in data and the prediction, constrained between 0 and 1, are given in the upper left corner of each plot here for the likelihood discriminant and in the following distributions, based on  $10^4$  bins per variable; the fine binning is needed for the Kolmogorov-Smirnov test to give adequate results. Values close to one represent a good agreement, values close to 0 represent a disagreement between data and the prediction.

The number of fitted events from the likelihood fit and the measured cross section are given in the first row of Table 7.20.

	$N_t^{t\bar{t}}$	$N_t^W$	$N_t^{QCD}$	$\sigma_{p\bar{p} \rightarrow t\bar{t}+X}$
uncorrected QCD-multijets template	$24.3_{-7.6}^{+8.5}$	$51.6_{-9.4}^{+10.3}$	$4.7_{-0.8}^{+0.8}$	$5.06_{-1.58}^{+1.76}$ pb
corrected QCD-multijets template	$24.7_{-7.6}^{+8.5}$	$51.5_{-9.4}^{+10.3}$	$4.7_{-0.8}^{+0.8}$	$5.13_{-1.57}^{+1.76}$ pb

Table 7.20: Fitted number of events from the likelihood fit, and measured  $t\bar{t}$  cross section using the uncorrected ( $N_{\ell-t}$ ) and the corrected ( $N_{\ell-t}^{QCD}$ ) QCD-multijets template.

### 7.4.1 $W$ -plus-jets and $t\bar{t}$ Contamination in QCD-multijets Sample

The shape of the likelihood discriminant for the QCD-multijets background is taken from the “loose-tight” data sample, requiring the full preselection

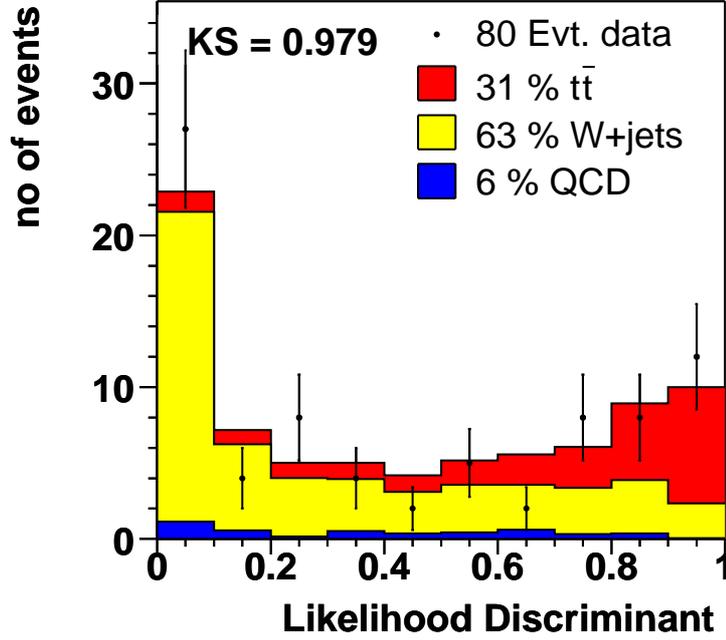


Figure 7.18: Likelihood discriminant distribution for data overlaid with the result from the fit of  $t\bar{t}$ ,  $W$ -plus-jets and QCD-multijets.

but inverting the tight muon isolation (see Sect. 7.3.1). Rewriting Eq. 7.20, one can see that the “loose-tight” ( $N_{\ell-t}$ ) data sample not only consists of QCD-multijets events but is contaminated with  $W$ -plus-jets and  $t\bar{t}$  events:

$$N_{\ell-t} = \frac{1 - \varepsilon_{sig}}{\varepsilon_{sig}} N_t^{t\bar{t}} + \frac{1 - \varepsilon_{sig}}{\varepsilon_{sig}} N_t^W + N_{\ell-t}^{QCD}. \quad (7.27)$$

However, the  $W$ -plus-jets and  $t\bar{t}$  contamination can be subtracted from  $N_{\ell-t}$  to get  $N_{\ell-t}^{QCD}$ , using  $\varepsilon_{sig}$ ,  $N_t^{t\bar{t}}$  and  $N_t^W$  from Table 7.14. Out of 69 “loose-tight” events 5.8 events are expected to be  $t\bar{t}$  and 12.4 events are expected to be  $W$ -plus-jets. This correction procedure can be done in bins of the likelihood discriminant in order to get a new QCD-multijets template, where the non-QCD contamination is removed, as shown in Fig. 7.19. The likelihood fit is repeated with the new QCD-multijets template ( $N_{\ell-t}^{QCD}$ ). The effect on the fitted  $t\bar{t}$  cross section is +1.3 %, as shown in Table 7.20. For the cross section determination the corrected QCD-multijets template is utilized.

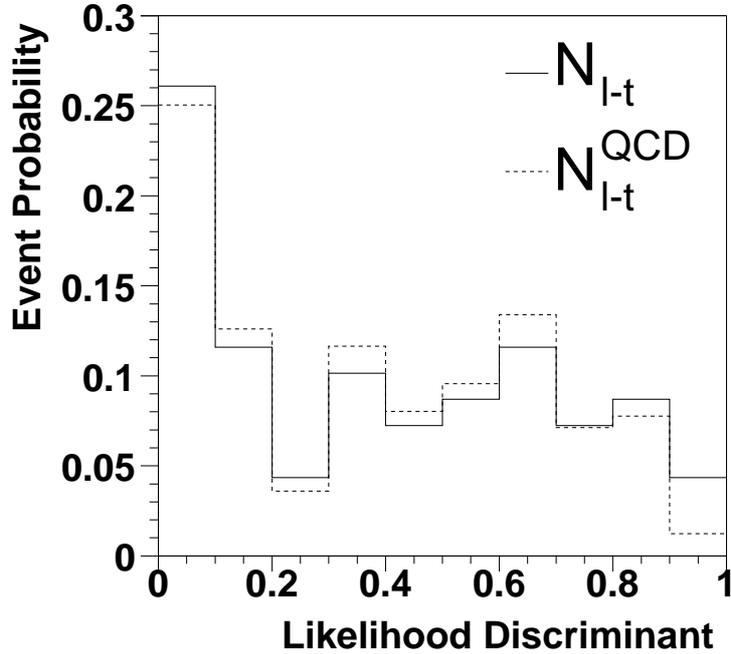


Figure 7.19: QCD-multijets likelihood discriminant template before ( $N_{l-t}$ ) and after subtracting the  $W$ -plus-jets and  $t\bar{t}$  contamination ( $N_{l-t}^{QCD}$ ), using the likelihood discriminant templates shown in Fig. 7.17.

#### 7.4.2 Result of the $t\bar{t}$ Production Cross-Section Measurement

The result of the  $t\bar{t}$  production cross-section measurement in the  $\mu$ -plus-jets channel at  $\sqrt{s} = 1.96$  TeV for a top mass of 175 GeV yields:

$$\sigma_{p\bar{p} \rightarrow t\bar{t}+X} = 5.13_{-1.57}^{+1.76} \text{ (stat) pb.}$$

The likelihood input distributions and various typical distributions are overlaid for the preselected data and the  $t\bar{t}$ ,  $W$ -plus-jets and QCD-multijets samples, normalized to the result of the likelihood fit in Fig. 7.20- 7.24. Further systematic checks for the MC modeling, the JES uncertainty and the uncertainty originating from the correction of the jet reconstruction  $\times$  identification efficiency in the simulation as discussed in Sect. 7.1.1.1, referred to as “JID”, are shown in Appendix A.

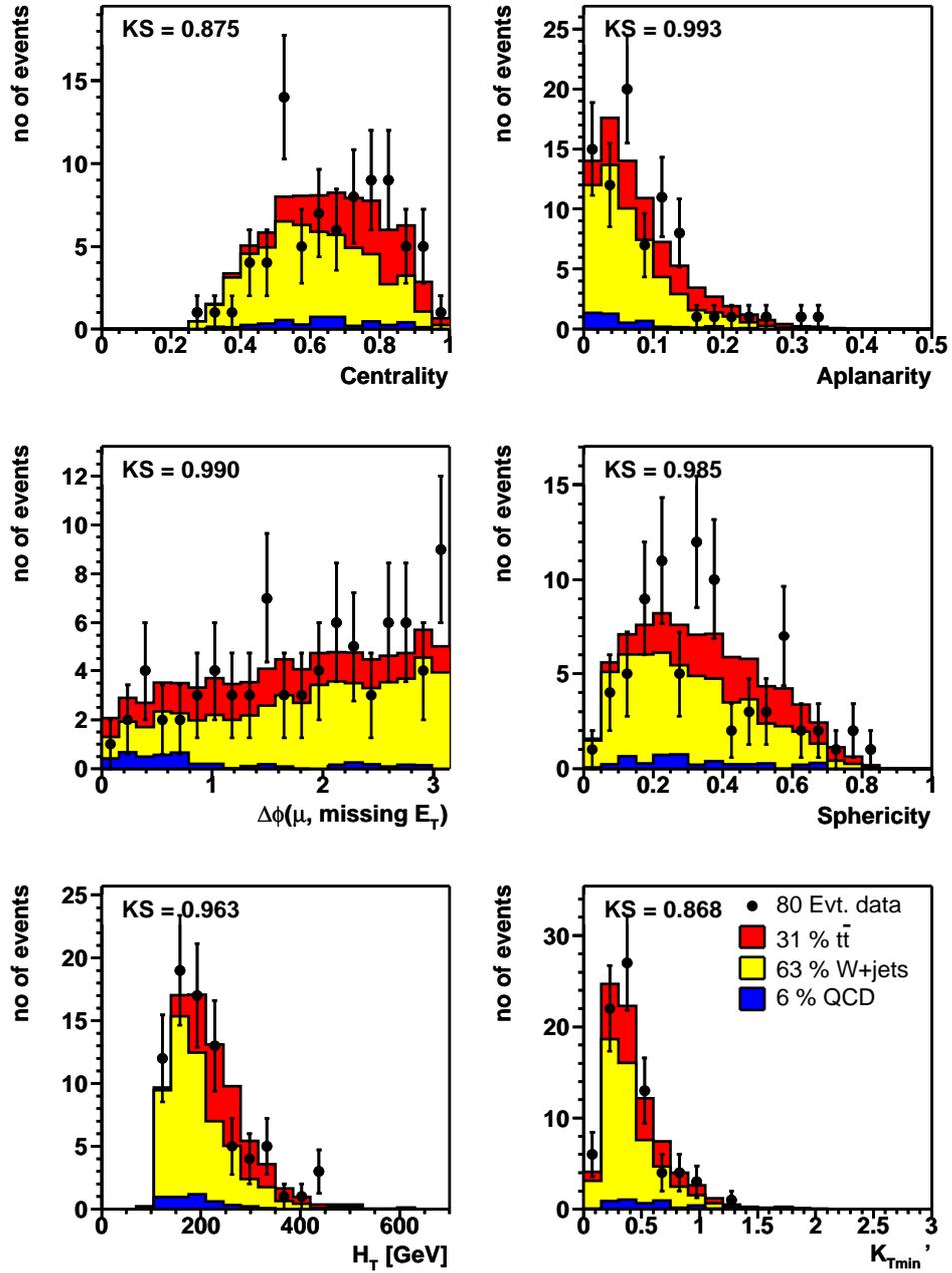


Figure 7.20: Likelihood discriminant input distributions for the preselected data overlaid with the result from the likelihood fit of  $t\bar{t}$ ,  $W$ -plus-jets and QCD-multijets.

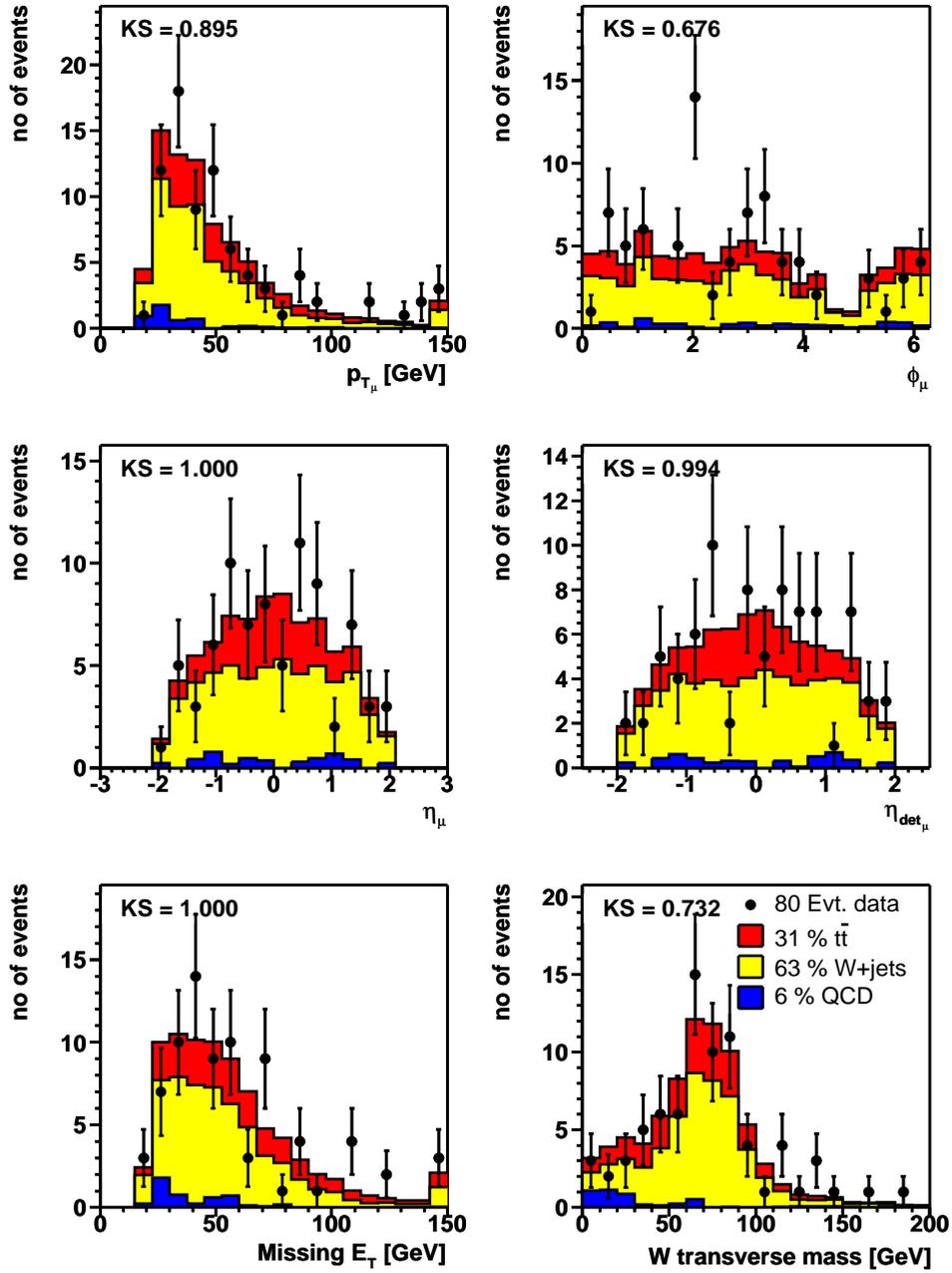


Figure 7.21:  $p_T$ ,  $\phi$ ,  $\eta$ ,  $\eta_{det}$  of the muon,  $\cancel{E}_T$  and  $W$  transverse mass for the preselected data overlaid with the result from the likelihood fit of  $t\bar{t}$ ,  $W$ -plus-jets and QCD-multijets.

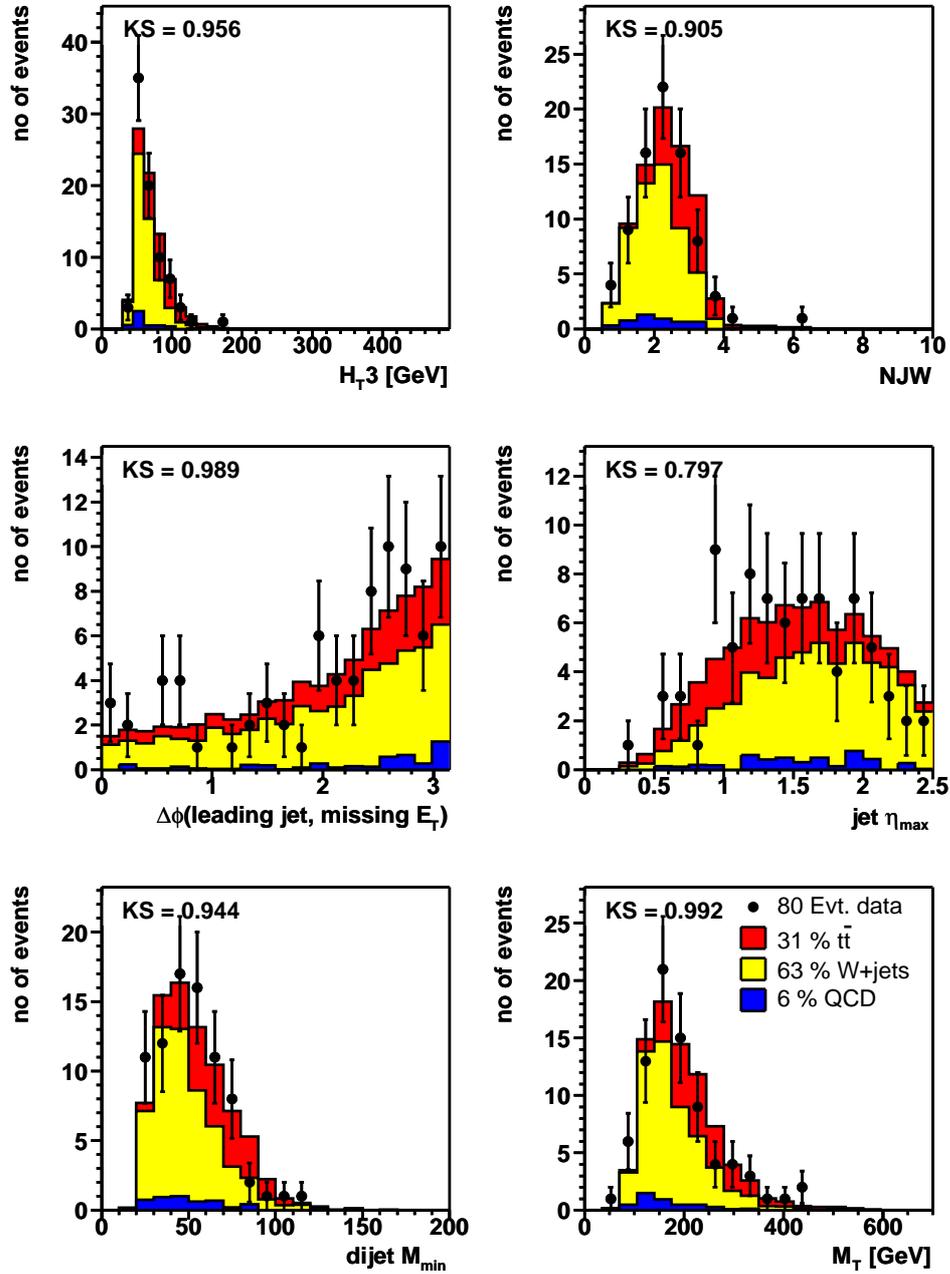


Figure 7.22: Scalar sum of the  $p_T$  of the third and fourth leading jets ( $H_{T3}$ ), NJW is a topological variable built from the  $p_T$  and  $\eta$  of the four leading jets,  $\Delta\phi$  between the leading jet and  $\cancel{E}_T$ , maximum jet  $\eta$ , the minimum dijet mass, transverse mass of the 4 leading jets ( $M_T$ ) for the preselected data overlaid with the result from the likelihood fit of  $t\bar{t}$ ,  $W$ -plus-jets and QCD-multijets.

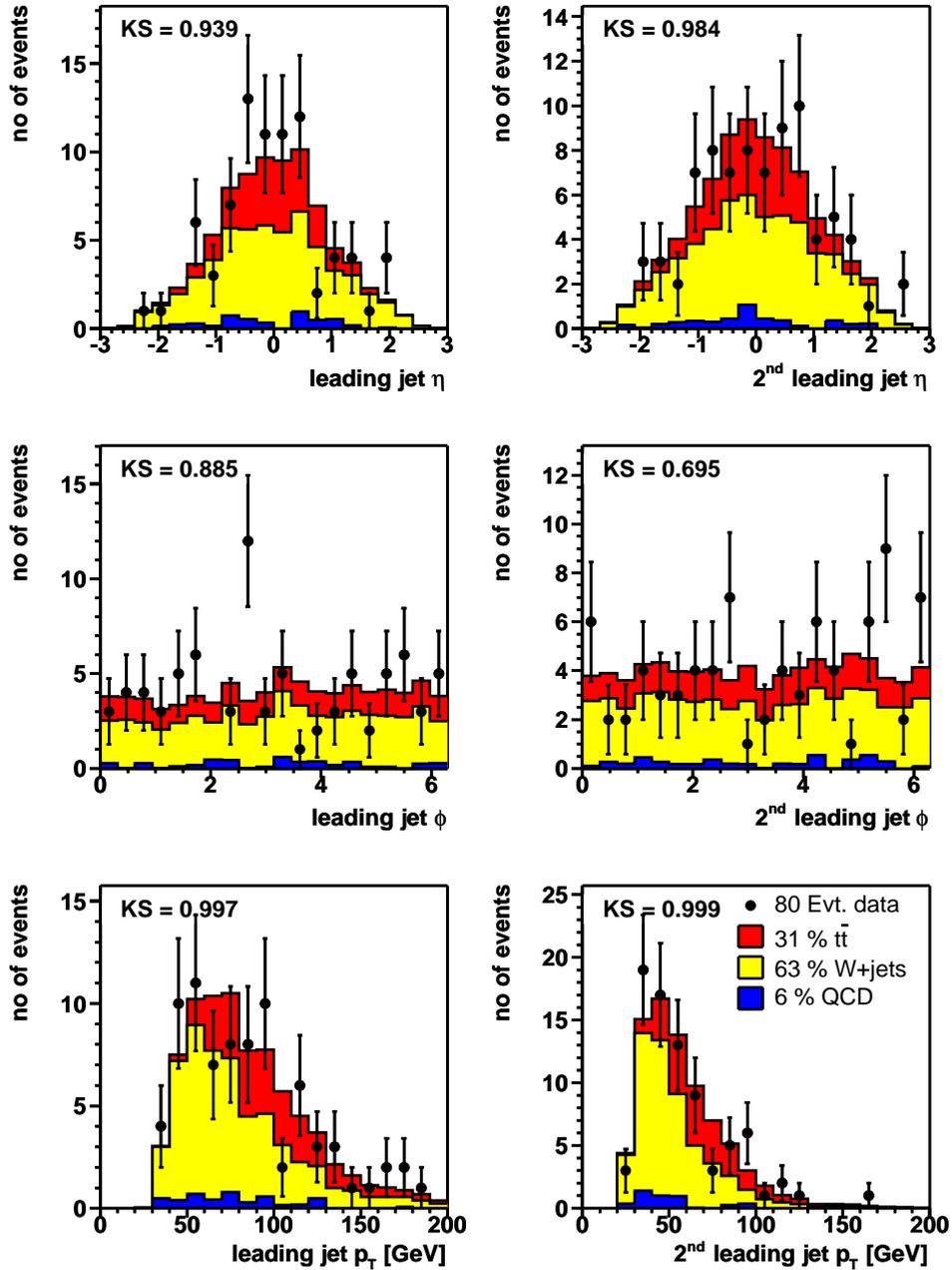


Figure 7.23:  $\eta$ ,  $\phi$  and  $p_T$  of the leading and second leading jet for the preselected data overlaid with the result from the likelihood fit of  $t\bar{t}$ ,  $W$ -plus-jets and QCD-multijets.

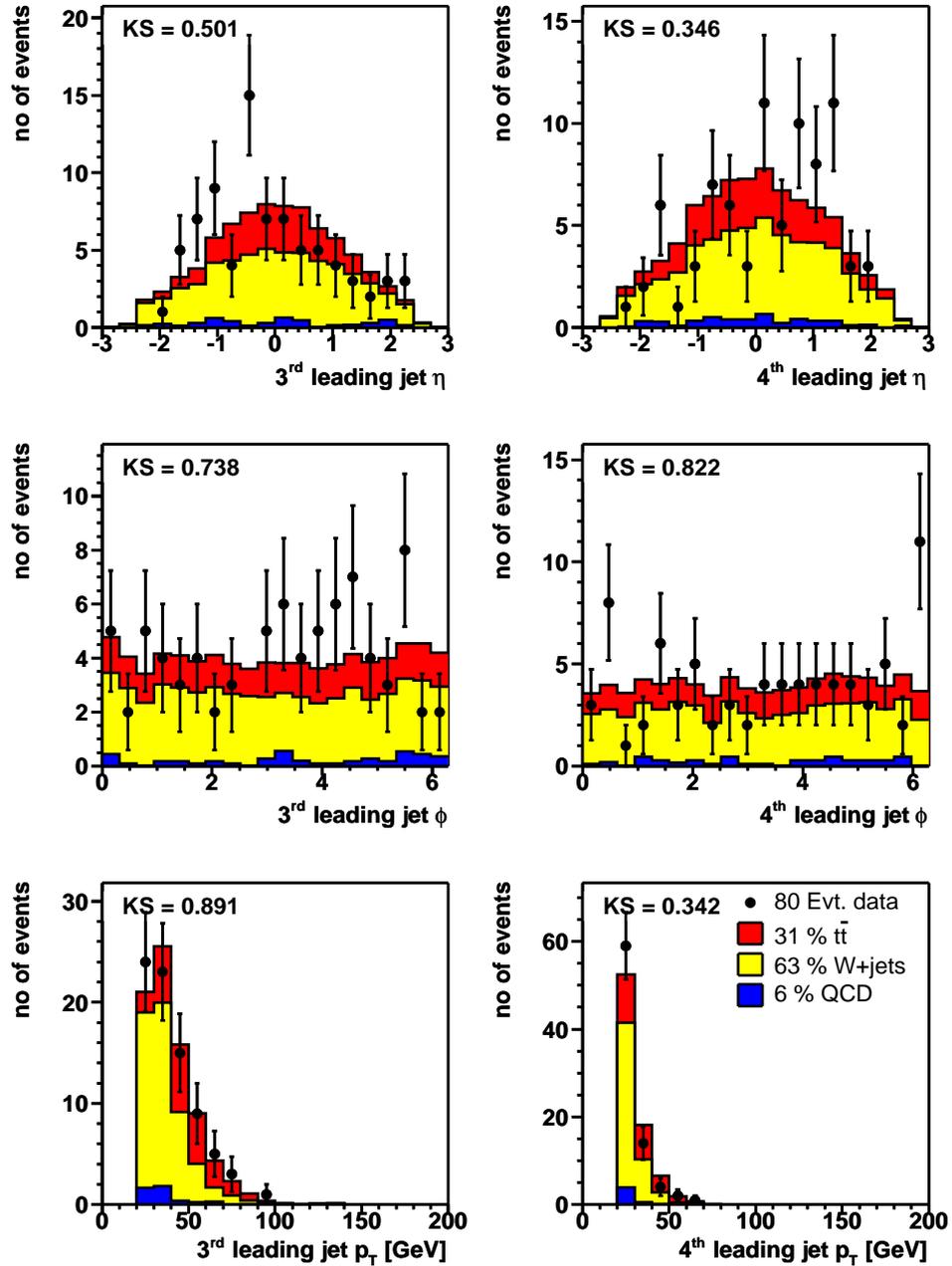


Figure 7.24:  $\eta$ ,  $\phi$  and  $p_T$  of the third leading and fourth leading jet for the preselected data overlaid with the result from the likelihood fit of  $t\bar{t}$ ,  $W$ -plus-jets and QCD-multijets.

## 7.5 Systematic Uncertainties

The systematic uncertainty on the  $t\bar{t}$  cross section measurement is comprised of the systematic uncertainties on all measurements, used to calculate the cross section, as given in Eq. 7.21. This includes the systematic uncertainties on the preselection efficiency (see Sect. 7.5.1), the estimated number of selected  $t\bar{t}$  signal events (see Sect. 7.5.2), the integrated luminosity and the branching ratio (see Sect. 7.5.3).

### 7.5.1 Uncertainties on the Preselection Efficiency

The systematic uncertainty assigned to the preselection efficiency is comprised of the individual systematic uncertainties associated with each preselection cut, the systematic uncertainties on the modeling of the trigger efficiency for  $t\bar{t}$  events, as discussed in Sect. 7.1, and the systematic uncertainties associated with the corrections applied to the MC in order to reproduce the data. The dominant among the latter are the JES correction, discussed in Sect. 4.5.5 and the assigned systematic uncertainty is given by Eq. 4.4, and the correction of the jet reconstruction  $\times$  identification efficiency in the simulation, discussed in Sect. 7.1.1.1 and referred to as “JID”. The systematic uncertainty assigned to the additional energy resolution smearing (JER) of jets in the simulation, discussed in Sect. 4.5.6, has a minor influence on the result. Additionally, a systematic uncertainty is assigned due to the available MC statistics used to derive the preselection efficiency.

A breakdown of the systematic uncertainties on the preselection efficiency is given in Table 7.21.

### 7.5.2 Uncertainties on the Fitted Number of $t\bar{t}$ Events

Systematic uncertainties on the fitted number of  $t\bar{t}$  events,  $N_t^{t\bar{t}}$ , are evaluated by varying the source of the systematic uncertainty, rederiving the likelihood templates for the signal and the backgrounds and repeating the likelihood fit to the data. The dependence of the  $t\bar{t}$  and  $W$ -plus-jets likelihood discriminant templates on the JES uncertainty and the jet reconstruction  $\times$  identification efficiency uncertainty is shown in Fig. 7.25.

All the sources of systematic uncertainties specified for the systematic uncertainty assigned to the preselection efficiency in Sect. 7.5.1 are also considered for the systematic uncertainties on  $N_t^{t\bar{t}}$ .

The uncertainty originating from the available MC statistics is evaluated by repeating the likelihood fit to the data 1000 times, each time allowing for Poisson fluctuations in the  $t\bar{t}$ ,  $W$ -plus-jets and QCD-multijets templates according to the available statistics used to build the template. The fluctuations are realized by replacing the actual number of events, used to populate the  $i$ -th bin, ( $\mu$ ) by  $n$ , drawn each time randomly according to the Poisson

source	systematic uncertainty		
	relative on $\varepsilon_{presel}$	relative on $N_t^{t\bar{t}}$	absolute on $\sigma_{t\bar{t}}$
Primary vertex	$\pm 0.9\%$	—	$\pm 0.09$ pb
Muon track match	$\pm 3\%$	—	$\pm 0.16$ pb
Tight muon isolation	$\pm 0.8\%$	—	$\pm 0.04$ pb
$\mu$ $\sigma_{dca}$	$\pm 0.3\%$	—	$\pm 0.02$ pb
$\Delta z(\mu, PV)$	$\pm 0.1\%$	—	$\pm 0.01$ pb
Muon track $\chi^2$	$\pm 0.3\%$	—	$\pm 0.02$ pb
$\mu$ ID	$\pm 2.2\%$	—	$\pm 0.11$ pb
L1 $\mu$ trigger	+1.3% -1.7%	+0.4% -0.6%	-0.04 pb +0.06 pb
L2 $\mu$ trigger	+3.2% -4.4%	+0.3% -0.5%	-0.14 pb +0.21 pb
JES	+11.2% -12.3%	-4.8% -3.1%	-0.74 pb +0.54 pb
JID	+5.0% -9.3%	+0.7% -8.6%	-0.21 pb +0.02 pb
JER	+0.2% -1.7%	-1.3% -1.6%	-0.08 pb +0.01 pb
L1 jet trigger	—	—	—
L2 jet trigger	—	—	—
L3 jet trigger	+0.1% -0.1%	+0.8% -0.9%	+0.04 pb -0.05 pb
MC statistics	$\pm 1.9\%$		$\pm 0.10$ pb
W MC modeling		+13.2% -13.2%	+0.68 pb -0.68 pb
$\varepsilon_{sig}$		-0.2% +0.1%	-0.01 pb +0.01 pb
$\varepsilon_{QCD}$		-1.1% +0.9%	-0.06 pb +0.05 pb
Template statistics		$\pm 4.5\%$	$\pm 0.23$ pb
total	+13.5% -16.8%	+14.0% -17.2%	+0.96 pb -1.10 pb

Table 7.21: Summary of the relative systematic uncertainties on the  $t\bar{t}$  pre-selection efficiency, the relative systematic uncertainties on the number of fitted  $t\bar{t}$  events (from variations of the templates in the fit), both in %, and the absolute systematic uncertainties on the  $t\bar{t}$  production cross section in pb. The  $^{+1\sigma}$  and  $_{-1\sigma}$  values represent the resulting systematic uncertainties on  $\varepsilon_{presel}$ ,  $N_t^{t\bar{t}}$ , and  $\sigma_{t\bar{t}}$ , respectively, when the corresponding source of the systematic uncertainty is varied by  $^{+1\sigma}$  and  $_{-1\sigma}$ , respectively. Systematic uncertainties are labeled as “—” when the effect is smaller than the precision quoted, and they are left blank when they do not apply.

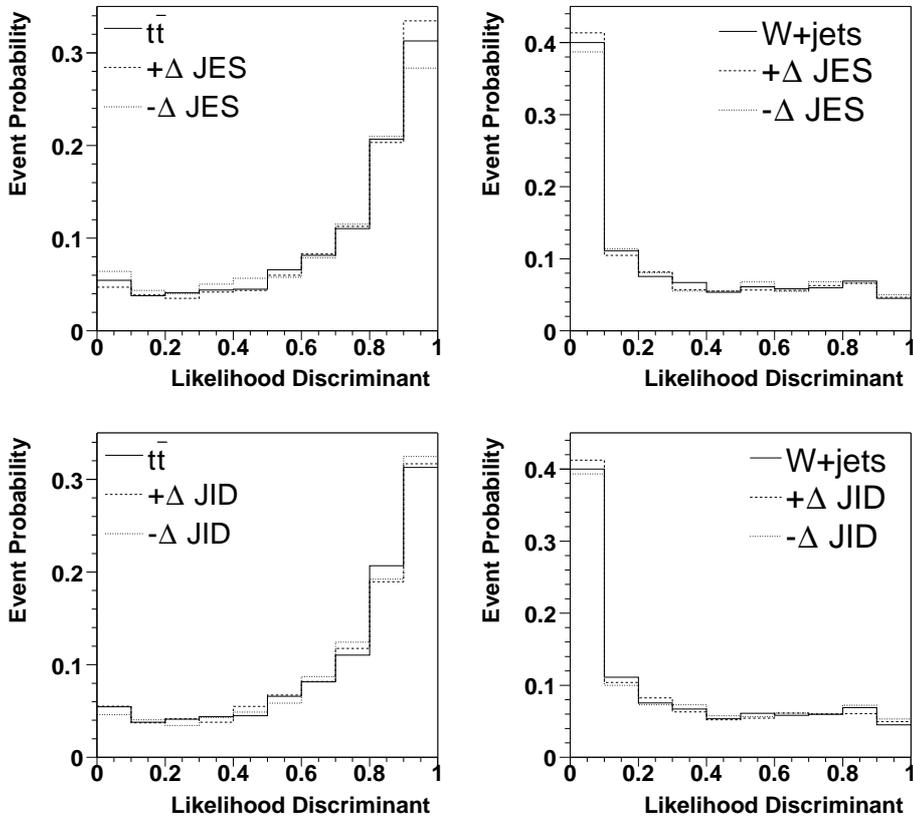


Figure 7.25: Likelihood discriminant distributions. Top: impact of the jet energy scale uncertainty on the  $t\bar{t}$  template (left) and on the  $W$ -plus-jets template (right). Bottom: impact of the jet reconstruction  $\times$  identification efficiency uncertainty on the  $t\bar{t}$  template (left) and on the  $W$ -plus-jets template (right).

distribution  $P(n, \mu)$ . The RMS of the distribution of the fitted number of  $t\bar{t}$  events is an estimate of the systematic uncertainty due to the available statistics used to build the likelihood templates.

There are additional sources of systematic uncertainties not present for the preselection efficiency, originating from the composition of the background and the background modeling:

**Background composition:** The amount of QCD-multijets background in the preselected data sample is determined by  $\varepsilon_{sig}$  and  $\varepsilon_{QCD}$  (see Sect. 7.3.4). The associated uncertainties on  $\varepsilon_{sig}$  (see Eq. 7.13) and on  $\varepsilon_{QCD}$  (see Eq. 7.13) are propagated to the uncertainty on  $N_t^{t\bar{t}}$ .

**$W$ -plus-jets background MC modeling:** Two sources of systematic uncertainties can be distinguished for the  $W$ -plus-jets background MC

modeling:

- The parton-level cuts applied for the MC generation, as given in Table 2.8, are chosen to be far away from the detector response thresholds and acceptance cuts, so that the omission of the contribution removed by the parton-level cuts does not lead to a measurable effect on the  $W$ -plus-jets likelihood template.
- The choice of factorization scale is found to have a measurable impact on the  $W$ -plus-jets likelihood template, as shown in Fig. 7.26. Replacing the  $W$ -plus-jets background Monte Carlo sample generated with the default scale  $Q^2 = M_W^2 + \sum p_{T_j}^2$  by one generated with  $Q^2 = \langle p_{T_j} \rangle^2$  and repeating the likelihood fit to the data leads to a systematic shift in  $N_t^{t\bar{t}}$  of 13.2 %. A comparison of parton-level distributions for the leading parton, the fourth leading parton and the scalar sum of the  $p_T$  of the four leading partons is shown in Fig. 7.27 for  $W$ -plus-jets MC generated with different choices of factorization scale. Changing the factorization scale from the default  $Q^2 = M_W^2 + \sum p_{T_j}^2$  to  $Q^2 = \langle p_{T_j} \rangle^2$  leads to softer  $p_T$  spectra, whereas the other two factorization scales lead to harder  $p_T$  spectra; the distributions for  $Q^2 = M_W^2 + \sum p_{T_j}^2$  are well centered between the other choices. Therefore it is a good approximation to symmetrize the one-sided error determined from the choice of  $Q^2 = \langle p_{T_j} \rangle^2$  as factorization scale.

A breakdown of the systematic uncertainties on  $N_t^{t\bar{t}}$  is given in Table 7.21. A comparison of the observed and the expected (see Sect. 7.3.5) systematic uncertainties on the fitted  $N_t^{t\bar{t}}$  is shown in Fig. 7.28, for the JES, JID and  $W$  MC modeling systematic uncertainties.

In the following, potential sources of systematic uncertainties are discussed which, however, are found to have no impact on  $N_t^{t\bar{t}}$ .

### 7.5.2.1 $W$ -plus-jets Heavy Flavor Composition

The  $W$ -plus-jets background is modeled in MC by the production of  $W$  boson in association with four light jets, as described in Sect. 7.3.1. The dependence of the  $W$ -plus-jets likelihood discriminant template on the heavy flavor content, quantified in Table 2.9, is shown in Fig. 7.29. The difference in the templates is compatible with the statistical uncertainty due to the limited statistics used to build the likelihood templates, and is therefore not quoted as a systematic uncertainty.

### 7.5.2.2 $Z$ -plus-jets Contamination

The likelihood fit to extract  $N_t^{t\bar{t}}$  is repeated replacing the  $W$ -plus-jets template by the template comprising  $W$ -plus-jets and  $Z$ -plus-jets in the expected

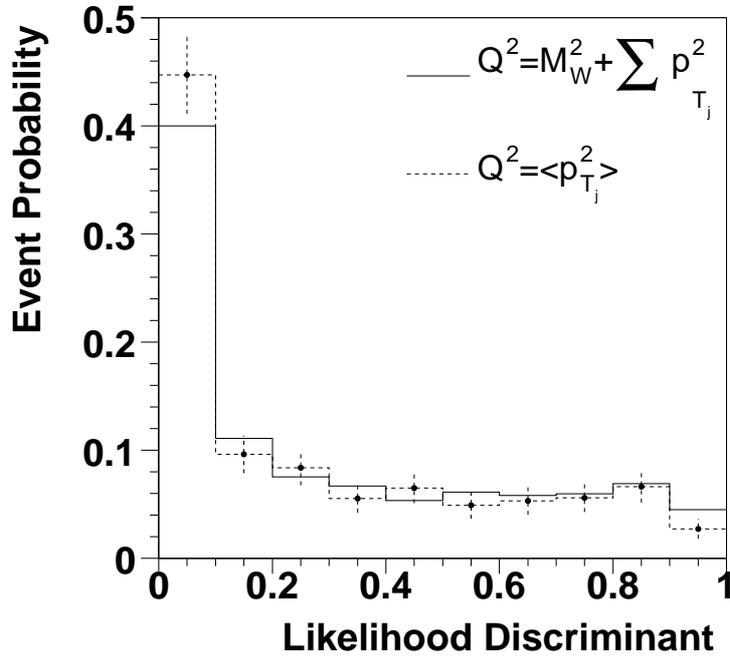


Figure 7.26:  $W$ -plus-jets likelihood discriminant template using two different factorization scales, the default  $Q^2 = M_W^2 + \sum p_{T_j}^2$ , and alternatively  $Q^2 = \langle p_{T_j} \rangle^2$ .

proportions,  $Z/W = 6.2\%$ , as discussed in Sect. 7.1.12 and Sect. 7.2.5. Both templates are shown in Fig. 7.30. The relative change in  $N_t^{t\bar{t}}$  is found to be  $-0.2\%$ . The change is compatible with the statistical uncertainty due to the limited statistics used to build the likelihood templates and is not quoted as a systematic uncertainty.

### 7.5.2.3 Likelihood Overtraining

The likelihood discriminant is built from 1/3 of the  $t\bar{t}$  and  $W$ -plus-jets MC statistics available. The likelihood template is then evaluated using the full MC statistics available. In order to show that no measurable statistical dependence is present due to the usage of 1/3 of the events both, in the building and the evaluation of the likelihood, “new”  $t\bar{t}$  and  $W$ -plus-jets templates are built, using only the 2/3 of the samples which are not used for the discriminant building.

Repeating the likelihood fit with the “new” templates leads to an observed relative change in  $N_t^{t\bar{t}}$  of  $-1.3\%$ . This shift is compatible with the expected shift originating from the statistical independence of the “new” and the default templates, as illustrated in the following. From pseudo-

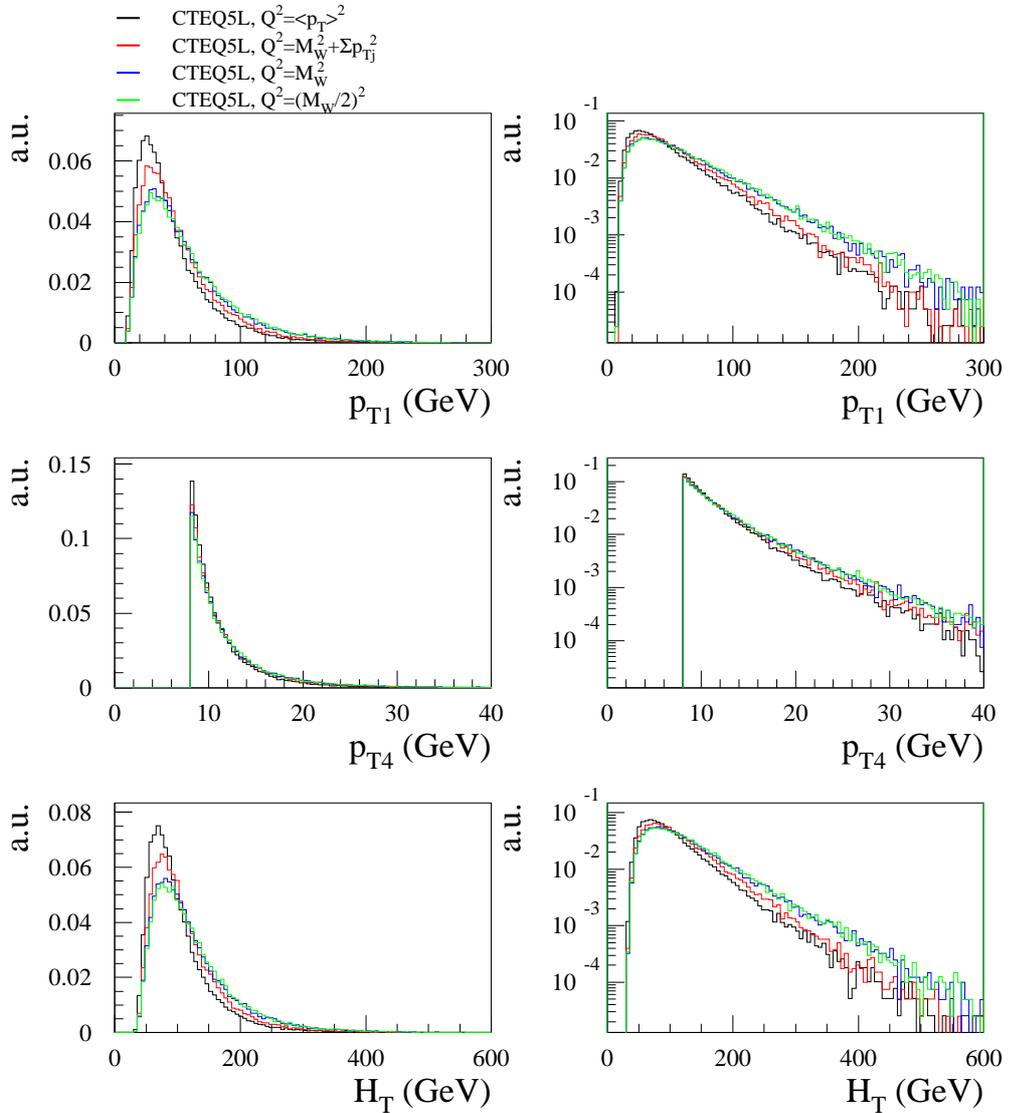


Figure 7.27: Comparison of parton-level distributions: Leading parton ( $p_{T1}$ ), fourth leading parton ( $p_{T4}$ ) and the scalar sum of the  $p_T$  of the four leading partons ( $H_T$ ) for  $W$ -plus-jets MC generated with different choices of factorization scale. Only partons with  $p_T > 8$  GeV and  $|\eta| < 2.5$  are considered.

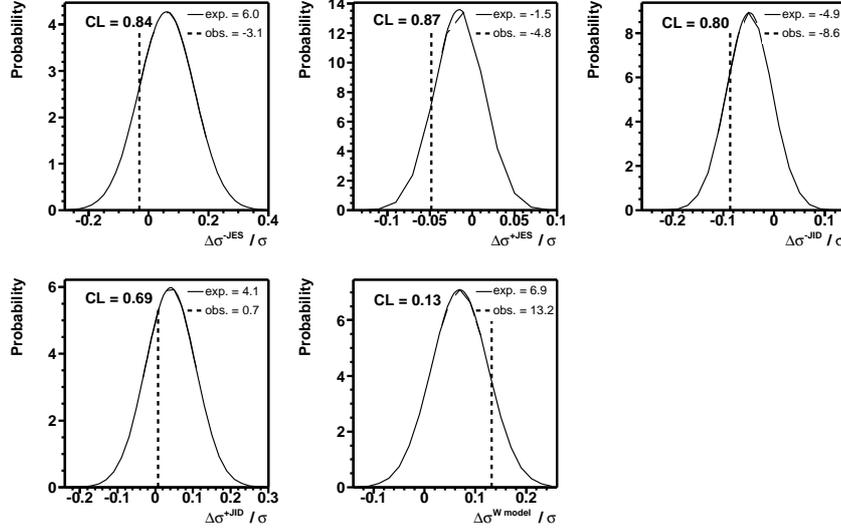


Figure 7.28: A comparison of the observed and the expected systematic uncertainties on the fitted  $N_t^{t\bar{t}}$  for the JES (top left and middle), JID (top right and bottom left) and  $W$  MC modeling (bottom right) systematic uncertainties. “CL” given in the plots denotes the confidence level for finding the observed systematic shift, given the shown expected probability density for the systematic variation, approximated by a Gaussian distribution.

experiments, allowing Poisson fluctuations in 1/3 of the statistics in the  $t\bar{t}$  and  $W$ -plus-jets template, one expects a statistical uncertainty of 2.9 %, in good agreement with the expected statistical uncertainty, allowing Poisson fluctuations in the full sample and dividing by  $\sqrt{3}$ ,  $4.5\%/\sqrt{3} = 2.6$  %.

The observed  $-1.3$  % are very well covered by the expected uncertainty originating from the statistical independence. This suggests that no bias from overtraining seems to be introduced, or if so it has no effect on the fitted  $N_t^{t\bar{t}}$ .

### 7.5.3 Summary of the Systematic Uncertainties

The systematic uncertainties on the  $t\bar{t}$  production cross section ( $\Delta\sigma_{t\bar{t}}$ ), originating from the uncertainties on the preselection efficiency ( $\Delta\varepsilon_{pre\,sel}$ ) and on  $N_t^{t\bar{t}}$  ( $\Delta N_t^{t\bar{t}}$ , from variations of the templates in the fit) are derived, taking into account their correlations, i.e. the systematic uncertainty from a given source can affect the preselection efficiency and the shape of the likelihood template at the same time:

$$\sigma_{t\bar{t}} \pm \Delta\sigma_{t\bar{t}} = \frac{N_t^{t\bar{t}} \pm \Delta N_t^{t\bar{t}}}{BR \cdot L \cdot (\varepsilon_{pre\,sel} \pm \Delta\varepsilon_{pre\,sel})}. \quad (7.28)$$

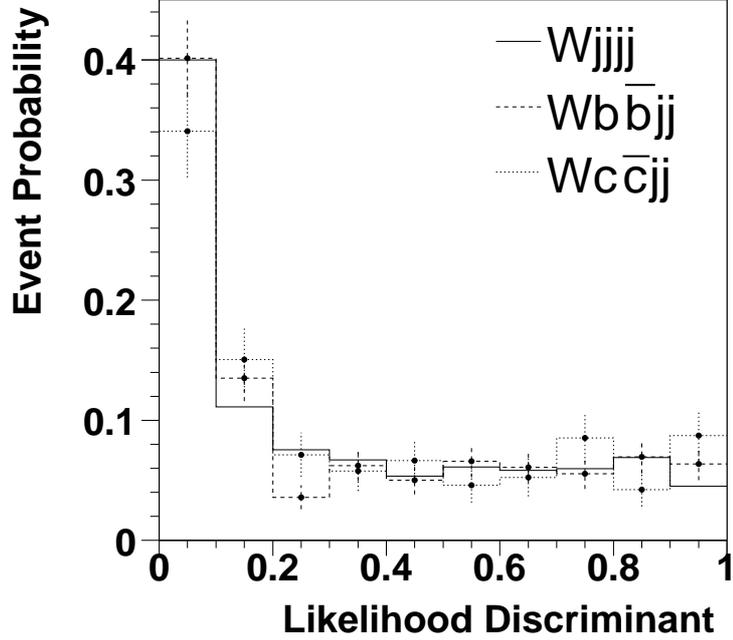


Figure 7.29: Dependence of the likelihood discriminant template on the heavy flavor composition of the  $W$ -plus-jets sample.

A summary of the corresponding systematics is given in Table 7.21.

The relative systematic uncertainty originating from the uncertainty on the luminosity measurement is 6.5 % [85].

The relative systematic uncertainty originating from the uncertainty on the branching ratio, given in Table 2.6, is 0.6 % [39].

#### 7.5.4 Result of the $t\bar{t}$ Production Cross-Section Measurement in the $\mu$ +Jets Channel

The result of the  $t\bar{t}$  production cross-section measurement in the  $\mu$ -plus-jets channel at  $\sqrt{s}=1.96$  TeV for a top mass of 175 GeV yields:

$$\sigma_{p\bar{p}\rightarrow t\bar{t}+X} = 5.13_{-1.57}^{+1.76} (\text{stat}) \pm 0.96_{-1.10} (\text{syst}) \pm 0.33 (\text{lumi}) \text{ pb.}$$

The top quark mass dependence of this result is not included in the systematic uncertainties, but rather discussed in the following. Figure 7.31 shows the top mass dependence of the  $t\bar{t}$  preselection efficiency and the top mass dependence of the fitted  $N_t^{t\bar{t}}$ , using the  $t\bar{t}$  MC samples described in Sect. 2.4.1. The resulting top mass dependence of the  $t\bar{t}$  production cross section, derived according to Eq. 7.28, is shown in Fig. 7.32, together with

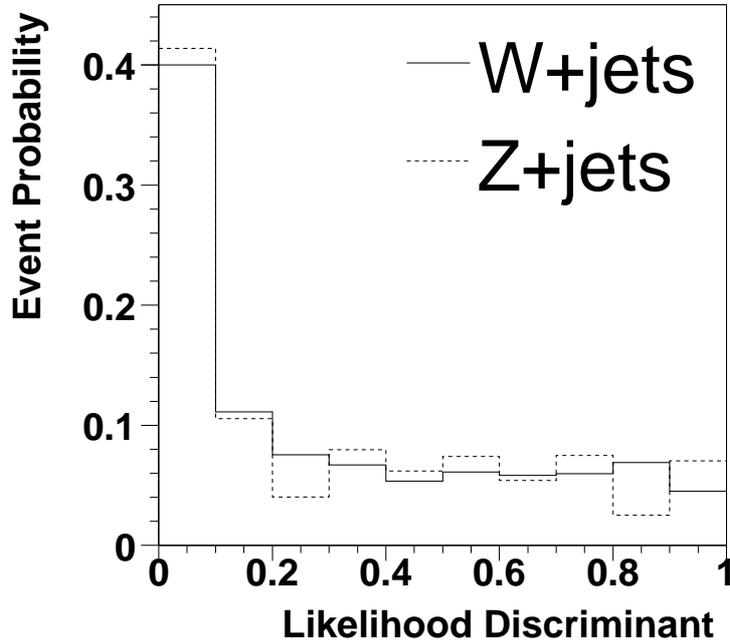


Figure 7.30:  $W$ -plus-jets template compared to the  $Z$ -plus-jets template.

the theoretically expected dependence, whose uncertainty is dominated by the kinematics ambiguity [150].

The presence of a small contamination from  $t\bar{t}_{dilepton}$  events in the preselected sample has an impact on the cross section measurement. The  $t\bar{t}_{dilepton}$  cross section  $\times$  branching ratio is not measured in the analysis presented. Thus, the dependence of the  $t\bar{t}$  production cross section, measured in the  $\mu$ -plus-jets channel, on the assumed  $\sigma_{dilepton} \times BR_{dilepton}$ , is shown in Fig. 7.33. This dependence is derived by treating the  $t\bar{t}_{dilepton}$  contribution as an additional background and by repeating the likelihood fit for varying contributions from  $t\bar{t}_{dilepton}$  events, expressed in multiples of the theoretical *Standard Model* expectation,  $\sigma_{dilepton}^{SM} \times BR_{dilepton}^{SM}$  (between zero and ten).

A comparison of the  $t\bar{t}$  and the  $t\bar{t}_{dilepton}$  template is shown in Fig. 7.33. For the *Standard Model* expectation of  $\sigma_{dilepton} \times BR_{dilepton}$ , the relative change in the measured cross section is found to be  $-1.7\%$ .

### 7.5.5 $t\bar{t}$ Production Cross-Section Measurement in the $e$ +Jets Channel

The topological  $t\bar{t}$  production cross section analysis was optimized in the  $\mu$ -plus-jets channel, as detailed in Sect. 7.3.5 and Sect. 7.3.6. Then the analysis is applied to the electron-plus-jets channel, replacing the isolated high

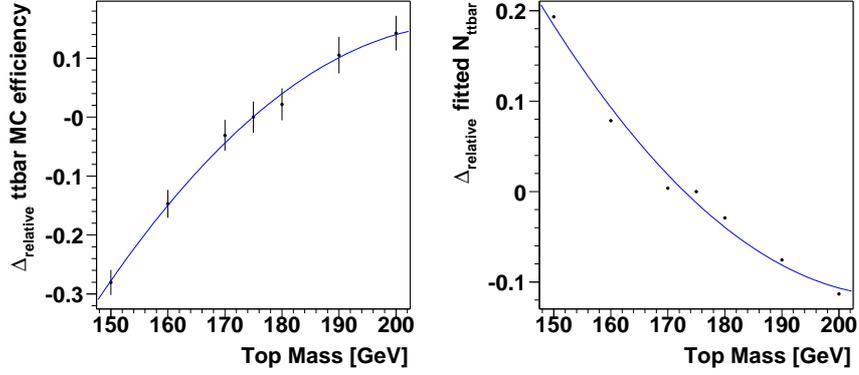


Figure 7.31: Left: Top quark mass dependence of the preselection efficiency. Right: Top quark mass dependence of the fitted number of  $t\bar{t}$  events, resulting from changes in the template distributions.

$p_T$  muon by an isolated high  $p_T$  electron and leaving the analysis method unchanged otherwise. Identically to the  $\mu$ -plus-jets channel, the two main backgrounds are the electroweak  $W$ -plus-jets and the QCD-multijets production, discussed in Sect. 7.2 for the  $\mu$ -plus-jets channel. However, the electron-plus-jets channel has a different source of fake electrons as compared to the  $\mu$ -plus-jets channel, where muons are real but their isolation is fake. Electrons are mostly faked by jets which fluctuate to have a large electromagnetic fraction. The QCD-multijets background is considerably larger in the electron-plus-jets channel as compared to the  $\mu$ -plus-jets channel due to two reasons. Firstly, the cross section of the process which fakes the electrons is larger and secondly, the selection criteria to reject the fake electrons are less efficient.

The numbers relevant for the extraction of the  $t\bar{t}$  cross section are summarized in Table 7.22. The number of fitted events from the likelihood fit are given in Table 7.23. A summary of the systematic uncertainties on the  $t\bar{t}$  cross section is given in Table 7.24.

$N_l$	$N_t$	$BR$	$L$	$\epsilon_{presel}$	$\epsilon_{sig}$	$\epsilon_{QCD}$
230	87	17.1%	226.3 pb $^{-1}$	11.6%	81.7%	16.0%

Table 7.22: Inputs for the likelihood fit [149].

The result of the  $t\bar{t}$  production cross-section measurement in the electron-plus-jets channel at  $\sqrt{s}=1.96$  TeV for a top mass of 175 GeV yields [149]:

$$\sigma_{p\bar{p}\rightarrow t\bar{t}+X} = 8.39_{-1.99}^{+2.17} \text{ (stat)} \quad {}_{-1.41}^{+1.84} \text{ (syst)} \pm 0.55 \text{ (lumi) pb.}$$

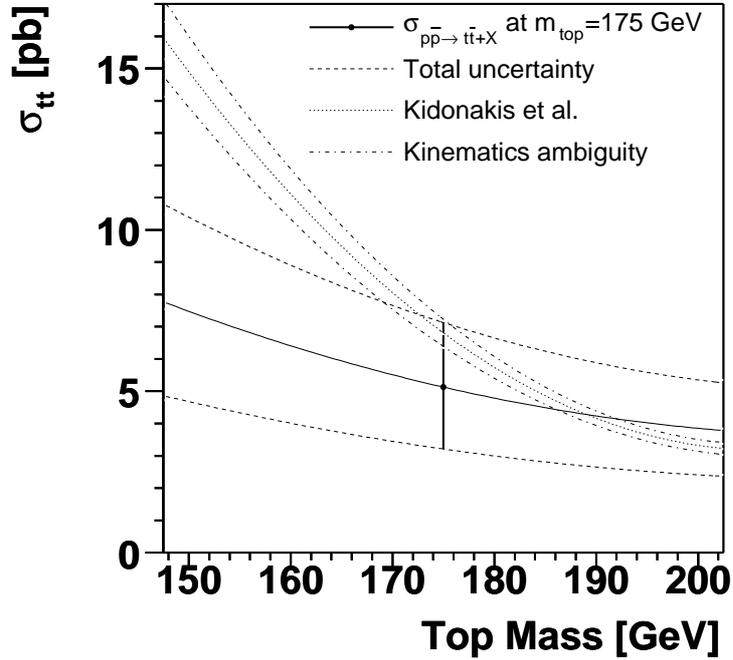


Figure 7.32: Observed and theoretically predicted top mass dependence of the  $t\bar{t}$  production cross section.

$N_t^{t\bar{t}}$	$N_t^W$	$N_t^{QCD}$
$37.7^{+9.8}_{-9.0}$	$22.9^{+10.3}_{-9.4}$	$24.9^{+2.5}_{-2.4}$

Table 7.23: Fitted number of events from the likelihood fit [149].

### 7.5.6 $t\bar{t}$ Production Cross-Section Measurement in the Lepton+Jets Channel

A summary of the systematic uncertainties on both the individual  $\mu$ -plus-jets and electron-plus-jets, as well as on the combined lepton-plus-jets  $t\bar{t}$  cross section is given in Table 7.24. The result of the  $t\bar{t}$  production cross-section measurement in the combined lepton-plus-jets channel at  $\sqrt{s}=1.96$  TeV for a top mass of 175 GeV yields [151]:

$$\sigma_{p\bar{p}\rightarrow t\bar{t}+X} = 6.60^{+1.37}_{-1.28} \text{ (stat)} \quad {}^{+1.25}_{-1.11} \text{ (syst)} \pm 0.43 \text{ (lumi) pb.}$$

source	$\mu$ +jets	$e$ +jets	combined
Primary vertex	$\pm 0.09$ pb	$\pm 0.16$ pb	$\pm 0.12$ pb
Muon track match	$\pm 0.16$ pb		$\pm 0.12$ pb
Tight muon isolation	$\pm 0.04$ pb		$\pm 0.03$ pb
$\mu \sigma_{dca}$	$\pm 0.02$ pb		$\pm 0.01$ pb
$\Delta z(\mu, PV)$	$\pm 0.01$ pb		$\pm 0.01$ pb
Muon track $\chi^2$	$\pm 0.02$ pb		$\pm 0.01$ pb
$\mu$ ID	$\pm 0.11$ pb		$\pm 0.09$ pb
L1 $\mu$ trigger	$-0.04$ pb $+0.06$ pb		$-0.04$ pb $+0.05$ pb
L2 $\mu$ trigger	$-0.14$ pb $+0.21$ pb		$-0.12$ pb $+0.17$ pb
EM tracking + likelihood		$-0.19$ pb $+0.20$ pb	$-0.06$ pb $+0.06$ pb
EM ID		$-0.20$ pb $+0.21$ pb	$-0.06$ pb $+0.06$ pb
L1 EM trigger		—	—
L3 EM trigger		$-0.06$ pb $+0.06$ pb	$-0.02$ pb $+0.02$ pb
JES	$-0.74$ pb $+0.54$ pb	$-1.23$ pb $+1.67$ pb	$-0.97$ pb $+1.12$ pb
JID	$-0.21$ pb $+0.02$ pb	$-0.23$ pb $+0.12$ pb	$-0.17$ pb $+0.09$ pb
JER	$-0.08$ pb $+0.01$ pb	$+0.05$ pb $+0.19$ pb	$+0.04$ pb $+0.12$ pb
L1 jet trigger	—	—	—
L2 jet trigger	—	—	—
L3 jet trigger	$+0.04$ pb $-0.05$ pb	$+0.01$ pb $-0.01$ pb	$+0.02$ pb $-0.03$ pb
W MC modeling	$+0.68$ pb $-0.68$ pb	$+0.01$ pb $-0.01$ pb	$+0.33$ pb $-0.33$ pb
$\varepsilon_{sig}(\mu$ +jets)	$-0.01$ pb $+0.01$ pb		—
$\varepsilon_{QCD}(\mu$ +jets)	$-0.06$ pb $+0.05$ pb		$-0.01$ pb $+0.01$ pb
$\varepsilon_{sig}(e$ +jets)		$-0.01$ pb $+0.01$ pb	—
$\varepsilon_{QCD}(e$ +jets)		$-0.27$ pb $+0.26$ pb	$-0.14$ pb $+0.14$ pb
MC statistics	$\pm 0.10$ pb	$\pm 0.15$ pb	$\pm 0.09$ pb
Template statistics	$\pm 0.23$ pb	$\pm 0.46$ pb	$\pm 0.23$ pb
total	$+0.96$ pb $-1.10$ pb	$+1.84$ pb $-1.41$ pb	$+1.25$ pb $-1.11$ pb

Table 7.24: Summary of the absolute systematic uncertainties on the  $\mu$ +jets, the  $e$ +jets [149] and the combined lepton+jets [151]  $t\bar{t}$  production cross section. The uncertainties in the upper part are correlated between the analysis channels, the uncertainties in the lower part are uncorrelated between the analysis channels. Systematic uncertainties are labeled as “—” when the effect is smaller than the precision quoted, and they are left blank when they do not apply.

A comparison of the  $t\bar{t}$  production cross-section measurements in the  $\mu$ -plus-jets, electron-plus-jets and the combined lepton-plus-jets channels is shown in Fig. 7.34.

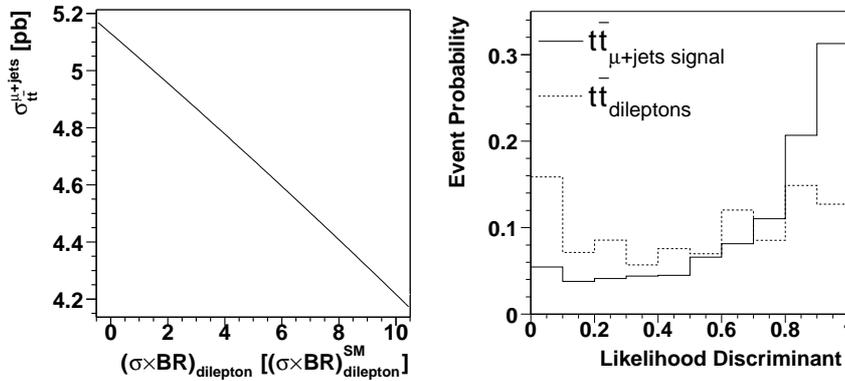


Figure 7.33: Left:  $t\bar{t}$  production cross section, measured in the  $\mu$ -plus-jets channel vs. the assumed  $\sigma_{\text{dilepton}} \times BR_{\text{dilepton}}$ , expressed in multiples of the theoretical *Standard Model* expectation,  $\sigma_{\text{dilepton}}^{\text{SM}} \times BR_{\text{dilepton}}^{\text{SM}}$ , i.e.  $1 \equiv 7 \text{ pb} \times BR_{\text{dilepton}}^{\text{SM}}$ . Right:  $t\bar{t}$  template compared to the  $t\bar{t}_{\text{dilepton}}$  template.

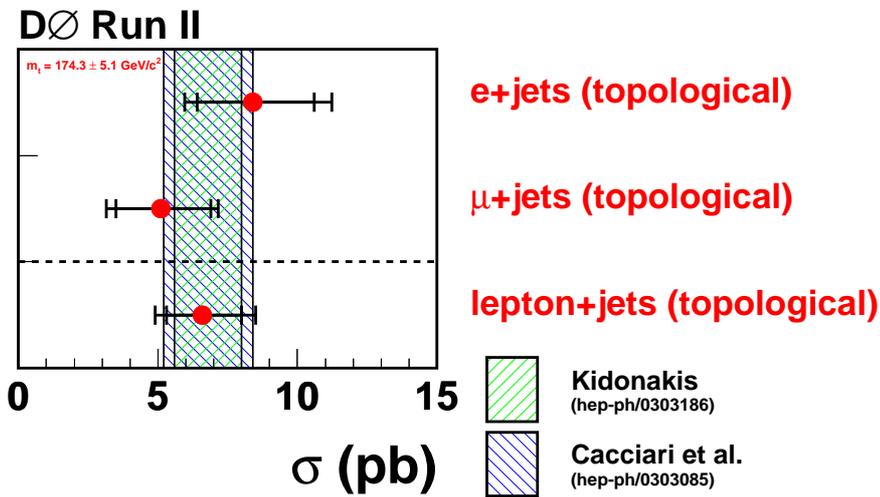


Figure 7.34: A comparison of the  $t\bar{t}$  production cross-section measurements in the  $\mu$ -plus-jets, electron-plus-jets and the combined lepton-plus-jets channel; also shown the theoretical expectation [151].



## Chapter 8

# The $b$ -Tagging Analysis

In this chapter the measurement of the  $t\bar{t}$  production cross-section at  $\sqrt{s} = 1.96$  TeV with the DØ detector at the Fermilab Tevatron, based on the application of lifetime  $b$ -tagging by explicitly reconstructing secondary vertices is presented. The outline of the analysis is the following. In a first step a preselection is applied, based on the selection of the decay products of the leptonically decaying  $W$  boson and jets, that serves to enhance the signal fraction. Most of the non- $t\bar{t}$  processes found in the preselected sample do not contain heavy flavor quarks in the final state. Requiring that one or more of the jets in the event be tagged as defined in Sect. 4.7 keeps more than half of the  $t\bar{t}$  events while removing approximately 95 % of the background. An excess of observed events over the expected background in the tagged (with one or more  $b$ -tags) preselected sample with three or  $\geq$  four jets, where the  $t\bar{t}$  signal is expected to contribute significantly, is assumed to be entirely due to  $t\bar{t}$  production. The  $t\bar{t}$  contribution is negligible in the tagged preselected sample with one or two jets. The observed results with those jet multiplicities serve as a check of the background prediction. In brief, the  $t\bar{t}$  production cross section is extracted by counting the number of tagged events,  $N_{observed}^{tag}$ , and subtracting the estimated number of tagged background events,  $N_{background}^{tag}$ , according to:

$$\sigma = \frac{N_{observed}^{tag} - N_{background}^{tag}}{BR \cdot L \cdot \varepsilon_{pre sel} \cdot \bar{P}^{tag}}, \quad (8.1)$$

where  $BR$  is the branching ratio of the considered final state,  $L$  is the integrated luminosity,  $\varepsilon_{pre sel}$  is the  $t\bar{t}$  preselection efficiency and  $\bar{P}^{tag}$  is the probability for a  $t\bar{t}$  event to have one or more jets identified as  $b$  jets (also called “event tagging probability”).

Section 8.1 discusses the individual preselection cuts and their efficiencies. The details on the determination and parameterization of the  $b$ -tagging efficiencies are given in Sect. 8.2. The application of  $b$ -tagging in the analysis and the contributions to the tagged preselected data set are

specified in Sect. 8.3. The cross-section extraction procedure is illustrated in Sect. 8.4 followed by the discussion of the systematic uncertainties in Sect. 8.5. The result of the  $t\bar{t}$  production cross-section measurement is presented in Sect. 8.6,

## 8.1 $t\bar{t}$ Event Preselection

The preselection for the  $\mu$ +jets channel is identical to the one discussed in Sect. 7.1 with the exception that the jet  $p_T$  cut is lowered from 20 GeV (in the *topo* analysis) to 15 GeV. This leads to an increase in the preselection efficiency by approximately 20 %, however, with a higher background contamination. The latter is affordable since the *b*-tagging application leads to an enormous  $t\bar{t}$  signal enhancement.

The event trigger for the  $e$ +jets channel requires an electron and a jet both at Level 1, 2 and 3 (apart from trigger list v12), the equivalent to Table 5.1 (for the  $\mu$ +jets channel) is given in Table 8.1 for the  $e$ +jets channel; more details on the trigger selection and parameterization of the trigger efficiencies can be found in [141]. The preselection for the  $e$ +jets channel

Trigger List	v8.2-v11	v12
Trigger Name	EM15_2JT15	E1_SHT15_2J20
L1	CEM(1,10)_CJT(2,5)	CEM(1,11)
L2	EM(.85,10)_JET(2,10)	---
L3	ELE_LOOSE_SH.T(1,15)_JET(2,15)	ELE_NLV_SHT(1,15)_JET(2,20)

Table 8.1: Summary of triggers used in trigger lists v8 to v12 for the  $e$ +jets channel.

is discussed in [152] and [13]. The PV and jet selection is identical to the one in the  $\mu$ +jets channel and a similar  $\cancel{E}_T$  cut is applied. Instead of a muon, the reconstruction of an electron is required with the identification criteria listed in Sect. 4.4. A breakdown of the preselection efficiencies for the two jet multiplicities relevant for the cross section determination is given in Table 8.2.

As described above, all jet multiplicities are considered in this analysis, the selection efficiencies per jet multiplicity are given in Table 8.3 both for the  $\mu$ +jets and the  $e$ +jets channel.

All MC samples used in the following are calibrated to the data by using the MC-to-data correction factors  $\kappa$  for the selection and identification efficiencies, given in Table 7.12 for the  $\mu$ +jets channel and in Table 8.2 for the  $e$ +jets channel.

	exclusive efficiency		cumulative efficiency	
	=3 jets	$\geq 4$ jets	=3 jets	$\geq 4$ jets
PV selection	0.955 $\pm$ 0.004	0.953 $\pm$ 0.008	0.955 $\pm$ 0.004	0.953 $\pm$ 0.008
Reconstructed EM	0.570 $\pm$ 0.004		0.545 $\pm$ 0.004	0.543 $\pm$ 0.006
$\kappa_{EM}^{reco}$	0.984 $\pm$ 0.004		0.536 $\pm$ 0.005	0.535 $\pm$ 0.006
$\kappa_{EM}^{selection}$	1.014 $\pm$ 0.001		0.543 $\pm$ 0.005	0.542 $\pm$ 0.006
Second lepton veto	0.973 $\pm$ 0.002		0.529 $\pm$ 0.005	0.527 $\pm$ 0.006
EM $p_T > 20$ GeV	0.931 $\pm$ 0.003		0.492 $\pm$ 0.005	0.491 $\pm$ 0.006
EM track match	0.930 $\pm$ 0.003		0.458 $\pm$ 0.005	0.457 $\pm$ 0.006
$\kappa_{EM}^{track\ match}$	0.945 $\pm$ 0.004		0.432 $\pm$ 0.005	0.432 $\pm$ 0.006
$\cancel{E}_T - \Delta\phi$ cut	0.792 $\pm$ 0.004		0.343 $\pm$ 0.004	0.342 $\pm$ 0.005
$\Delta z(\text{EM,PV})$ cut	0.964 $\pm$ 0.001		0.330 $\pm$ 0.004	0.330 $\pm$ 0.005
EM likelihood cut	0.942 $\pm$ 0.003		0.311 $\pm$ 0.004	0.310 $\pm$ 0.004
$\kappa_{EM}^{likelihood}$	0.931 $\pm$ 0.004		0.290 $\pm$ 0.004	0.289 $\pm$ 0.004
$N_{jet}$	0.336 $\pm$ 0.006	0.565 $\pm$ 0.006	0.097 $\pm$ 0.002	0.163 $\pm$ 0.003
Trigger efficiency	0.930 $\pm$ 0.006	0.931 $\pm$ 0.004	0.090 $\pm$ 0.002	0.152 $\pm$ 0.003

Table 8.2: Breakdown of the  $t\bar{t}$  preselection efficiencies in the  $e$ +jets channel,  $EM$  stands for electron.

	=1 jet	=2 jets	=3 jets	$\geq 4$ jets
$\mu$ +jets	0.086 $\pm$ 0.020	1.64 $\pm$ 0.11	8.64 $\pm$ 0.41	15.91 $\pm$ 0.74
$e$ +jets	0.182 $\pm$ 0.029	2.42 $\pm$ 0.11	9.04 $\pm$ 0.21	15.20 $\pm$ 0.29

Table 8.3: Summary of  $t\bar{t}$  preselection efficiencies in the  $\mu$ +jets and  $e$ +jets channel (%).

## 8.2 The Jet Tagging Efficiencies

The identification of  $b$  jets is defined in Sect. 4.7. It is tuned to identify  $b$  jets with a high efficiency, referred to as  $b$ -tagging efficiency, while keeping the probability low to tag a light jet (from a  $u$ ,  $d$ ,  $s$  quark or gluon), referred to as mis-tagging efficiency. The efficiency to tag a  $c$  jet is referred to as  $c$ -tagging efficiency. The definitions of the jet flavor in the MC are given in Sect. 2.4.2.1. Jets are required to be reconstructed also in the central tracking system, referred to as taggability, in order to be potentially tagged.

Present versions of the  $D\bar{O}$  Monte Carlo simulation do not reproduce the  $b$ -identification efficiencies observed in data, both taggability,  $b$ -,  $c$ - and mis-tagging efficiency. Reasons for this deficiency are dead detector material and noise in the central tracking system being inadequately described by the Monte Carlo simulation, resulting in an overestimated tracking efficiency, in particular within jets.

For this reason, straightforward MC based calculation will not give a correct result, instead the following method is utilized to evaluate the jet tagging efficiencies:

- If possible the efficiencies (taggability,  $b$ -,  $c$ - and mis-tagging efficiency) are derived directly from data. Due to potential inabilities to select pure samples of  $b$ ,  $c$  or light jets in data, corrections which are typically small are determined using Monte Carlo information. These corrections or scale factors are referenced as  $SF$  with a descriptive subscript.
- The jet tagging efficiencies (taggability,  $b$ -,  $c$ - and mis-tagging efficiency) are parameterized as a function of the relevant jet quantities. The preferred variables are tracking related quantities, such as the number of tracks, satisfying certain quality criteria, in a jet. However, the mentioned inability of the MC simulation to describe accurately tracking related quantities, excludes these variables. Instead, the  $p_T$  and the  $\eta$  of a jet is found to describe all dependencies of the jet tagging efficiencies sufficiently.
- Parameterizations are obtained for the  $p_T$  and the  $\eta$  dependence separately. A two-dimensional parameterization is derived by assuming that the dependence is factorizable, so that  $\varepsilon(p_T, \eta) = C\varepsilon(p_T)\varepsilon(\eta)$ . The normalization factor  $C$  is calculated so that the total number of predicted taggable, respectively tagged, jets obtained as the sum over all reconstructed jets weighted with  $\varepsilon(p_T, \eta)$  according to their  $p_T$  and  $\eta$  is equal to the total number of observed taggable, respectively tagged, jets.  
The factorizability is verified by comparing the number of predicted and the number of observed taggable, respectively tagged, jets as a function of jet  $p_T$  and  $\eta$ . This verification is called “closure test”. Statistical uncertainties of the fits to derive the parameterizations are assigned to the taggability, respectively the tagging efficiencies.
- The probability for a given jet of flavor  $\alpha$  ( $b$ ,  $c$ , light) to be tagged is the product of the taggability ( $\varepsilon_{tagg(\alpha)}(p_T, \eta)$ ) and the tagging efficiency ( $\varepsilon_\alpha(p_T, \eta)$ ):

$$\mathcal{P}_\alpha(p_T, \eta) = \varepsilon_{tagg(\alpha)}(p_T, \eta) \cdot \varepsilon_\alpha(p_T, \eta), \quad (8.2)$$

where also the taggability can depend on the flavor of the jet.

The techniques to measure and parameterize these jet tagging efficiencies are presented in the following.

### 8.2.1 Taggability

The jet tagging proceeds in two steps, first by requiring the taggability and then by applying the actual tagging algorithm. This breakdown is made in order to decouple the tagging efficiency from issues related to tracking

inefficiencies and calorimeter noise problems, which are therefore absorbed into the taggability. A jet is taggable if it is matched within  $\Delta R \leq 0.5$  to a track-jet (see also Sect. 4.7.1).

The taggability is measured in the inclusive ( $N_{\text{jet}} \geq 1$ ) preselected data samples, separately for the  $\mu$ +jets and the  $e$ +jets sample. The taggability distribution and the one-dimensional functional parameterizations are shown as a function of jet  $p_T$  and  $\eta$  in Fig. 8.1.

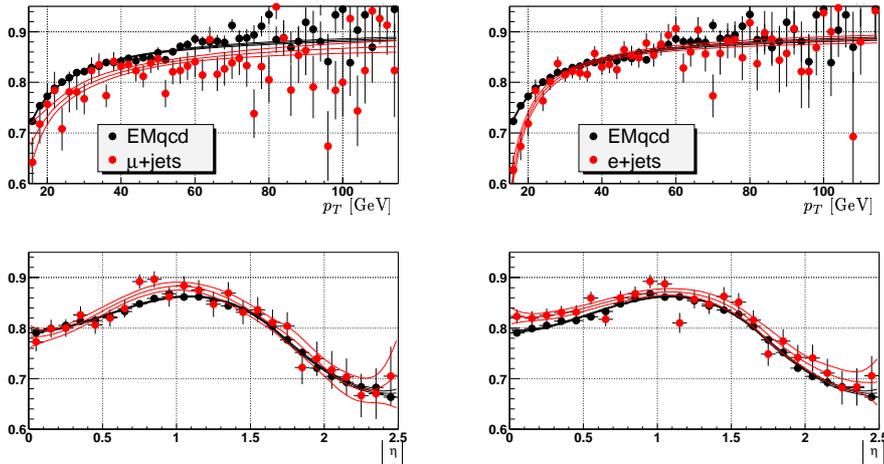


Figure 8.1: Jet taggability as a function of jet  $p_T$  (top) and  $\eta$  (bottom) for the first jet multiplicity bin of the preselected  $e$ +jets (left) and  $\mu$ +jets (right) signal samples and compared to the corresponding taggability in EMqcd. Curves indicate the fit and its  $1\sigma$  error band.

The systematic uncertainty, assigned to the taggability measurement, comprises the statistical uncertainty of the fit to derive the parameterizations and the sample dependence of the taggability, evaluated by replacing the  $\mu$ +jets, respectively the  $e$ +jets sample, by the EMqcd data sample and rederiving the taggability parameterizations. The EMqcd data sample, obtained from the 1EMloose and EM1TRK CSG skims [142, 122] and also shown in Fig. 8.1, is selected requiring:

- $e$ +jets event trigger,
- $\geq 1$  loose electron with  $E_T > 20$  GeV, defined by
  - $EMF > 0.9$ ,
  - $f_{\text{iso}} < 0.15$  and
  - $H\text{-matrix7} < 75$ ,

- $N_{\text{jet}} \geq 1$ ,
- $\cancel{E}_T < 10 \text{ GeV}$ ,

representing a general QCD-multijets sample with an electron in the final state.

### 8.2.1.1 Jet Flavor Dependence of Taggability

The taggability, measured in data, represents an inclusive jet flavor taggability. However, it is dominated by the light jet contribution since the heavy flavor content in the low jet multiplicity bins is small, and thus, it is used to describe the taggability of light flavor jets.

The harder fragmentation of heavy flavor jets leads both to a harder  $p_T$  spectrum of tracks associated to the jet and to an average track multiplicity larger than for light flavor jets, as shown in Fig. 8.2. This leads to an increased taggability, given the track-jet based taggability definition in Sect. 4.7.1 and Sect. 4.7.1.1. The ratios of  $b$  to light and  $c$  to light taggabilities are shown in Fig. 8.3. The difference increases for low  $p_T$  jets up to approximately 10 %, the  $\eta$  dependence is relatively soft. The fits to the ratios are used as flavor dependent correction factors to the taggability. The dependence of the taggability on the description of the underlying event (“Tune A”) is used to assign a systematic uncertainty on the flavor dependent correction.

### 8.2.2 $b$ -Tagging Efficiency

The measurement of the  $b$ -tagging efficiency in data requires the knowledge of the number of taggable  $b$  jets before and after applying the tagging algorithm. Optimally, the measurement is performed on a pure sample of taggable  $b$  jets which fulfills this requirement by definition. However, there is currently no way to select a pure sample of taggable  $b$  jets in the data in an unbiased way. Nevertheless it is preferred to measure the  $b$ -tagging efficiency on a sample enriched in taggable  $b$  jets. The data sample of choice is the muon-in-jet sample, derived from the 1MUloose CSG skim [142], which is selected requiring:

- $N_{\text{jet}} \geq 2$ ,
- $\geq 1$  medium muon with  $\Delta R(\mu, jet) < 0.5$ ,

representing a sample enriched in semimuonic heavy flavor quark decays, as indicated in Sect. 4.7.

Two methods are used to measure the  $b$ -tagging efficiency in data. Both of them measure the efficiency solely for taggable  $b$  jets with a semimuonic  $b$  decay ( $b \rightarrow \mu$ ) where  $\Delta R(\mu, jet) < 0.5$ .

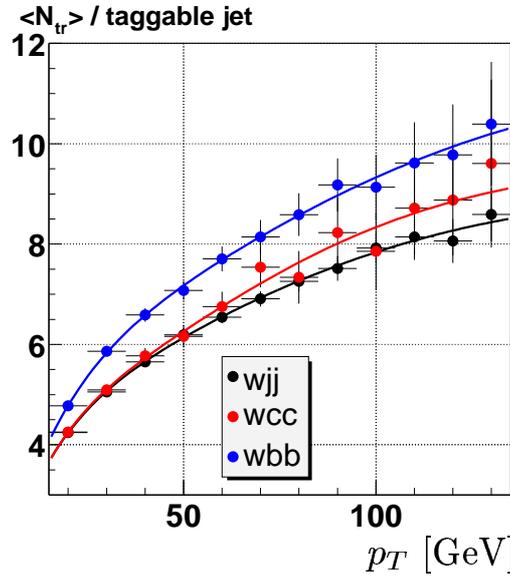


Figure 8.2: Average track multiplicity in a taggable jet as a function of jet  $p_T$  for  $b$ ,  $c$  and light jets in the  $Wb\bar{b}$ ,  $Wc\bar{c}$  and  $Wjj$  Monte Carlo samples.  $b$  jets have the highest and light jets the lowest track multiplicity.

### 8.2.2.1 $p_T^{rel}$ Method

An unbiased discriminating variable is used to evaluate the number of taggable  $b$  jets before ( $N_{b \rightarrow \mu}$ ) and after ( $N_{b \rightarrow \mu}^{tag}$ ) applying the tagging algorithm. This is done analogous to the evaluation of the number of  $t\bar{t}$  events in the *topo* analysis, discussed in Sect. 7.3.4. The fraction of the muon momentum transverse to the momentum of the jet-muon system, referred to as  $p_T^{rel}$  (illustrated in the schematic drawing in Fig. 8.4), discriminates between  $b$  jets and non- $b$  jets, due to the large mass of the  $b$  quark which leads to high values of  $p_T^{rel}$ . The number of taggable  $b$  jets is extracted by describing the distribution of the  $p_T^{rel}$  measured in data by a linear combination of taggable  $b$ ,  $c$  and light jets. The templates for the  $p_T^{rel}$  distributions of the heavy flavor jets is taken from MC, the light jets template is taken from data, by picking tracks randomly within a jet. Then the  $b$ -tagging efficiency is given by:

$$\varepsilon_{b \rightarrow \mu}(p_T, \eta) = \frac{N_{b \rightarrow \mu}^{tag}(p_T, \eta)}{N_{b \rightarrow \mu}(p_T, \eta)}. \quad (8.3)$$

The templates and the fit results are shown in Fig. 8.5.

Two data samples with different  $b$  jet purity are used for the measurement, both of them representing  $b\bar{b}$  production:

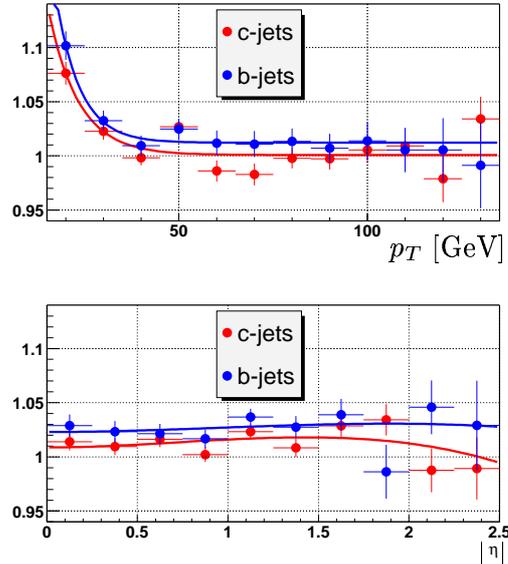


Figure 8.3: Ratio of the  $b$  to light and  $c$  to light jet taggability, measured in a QCD Monte Carlo sample. The ratio is larger for the  $b$  jets.

- the muon-in-jet sample,
- the muon-in-jet sample, where the jet opposite to the jet containing the muon (so-called “away-jet”), is tagged, referred to as muon-in-jet-away-jet-tagged sample. This enriches the sample in heavy flavor due to the contribution of a gluon splitting to a  $b\bar{b}$  quark pair.

### 8.2.2.2 System8 Method

This method allows to calculate the  $b$ -tagging efficiency solely from data. Generally, the method [153] exploits the possibility to measure selection efficiencies for a signal and one or several background classes by measuring the inclusive selection efficiencies for one or more samples with varying signal and backgrounds compositions. The problem can be represented by a system of equations. The system is clearly solvable if

- the samples have different fractions of signal and backgrounds,
- the selection criteria have different efficiencies for these signal and backgrounds,
- the different selection criteria have to be decorrelated to allow the factorization of the efficiencies,

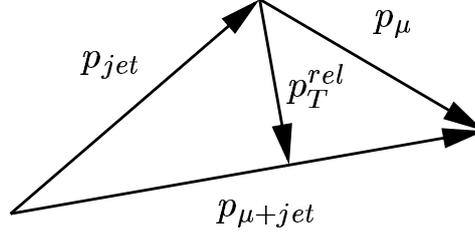


Figure 8.4: Schematic drawing to illustrate the definition of  $p_T^{rel}$ .

- the number of equations (constraints)  $\geq$  the number of unknowns.

Specifically for the measurement of the  $b$ -tagging efficiency, two data samples are used, the muon-in-jet ( $n$ ) and the muon-in-jet-away-jet-tagged sample ( $p$ ). The signal class represents the  $b$  jets and there is one background class representing the non- $b$  jets ( $c$  and light), the selection criteria are the SVT  $b$ -tagging criteria, subject of the measurement, and the SLT (see Sect. 4.7) criteria, requiring that the present muon with  $\Delta R(\mu, jet) < 0.5$  also fulfills  $p_T^{rel} > 1$  GeV. Eight equations are identified:

$$\begin{aligned}
 n &= n_b + n_{non-b} \\
 p &= p_b + p_{non-b} \\
 n^{SVT} &= \varepsilon_b^{SVT} n_b + \varepsilon_{non-b}^{SVT} n_{non-b} \\
 p^{SVT} &= \varepsilon_b^{SVT} p_b + \varepsilon_{non-b}^{SVT} p_{non-b} \\
 n^{SLT} &= \varepsilon_b^{SLT} n_b + \varepsilon_{non-b}^{SLT} n_{non-b} \\
 p^{SLT} &= \varepsilon_b^{SLT} p_b + \varepsilon_{non-b}^{SLT} p_{non-b} \\
 n^{SVT,SLT} &= \varepsilon_b^{SVT} \varepsilon_b^{SLT} n_b + \varepsilon_{non-b}^{SVT} \varepsilon_{non-b}^{SLT} n_{non-b} \\
 p^{SVT,SLT} &= \varepsilon_b^{SVT} \varepsilon_b^{SLT} p_b + \varepsilon_{non-b}^{SVT} \varepsilon_{non-b}^{SLT} p_{non-b}
 \end{aligned}$$

where both the  $(p_T, \eta)$  dependence of each term is omitted and the subscripts  $b \rightarrow \mu$  is abbreviated by  $b$  for reasons of clarity. The terms on the left hand side of the equations are measured in data, and the eight unknowns on the right hand side of the equations consist of the number of taggable  $b$  and non- $b$  jets in the two samples ( $n_b, n_{non-b}, p_b, p_{non-b}$ ) and the tagging efficiencies for taggable  $b$  and non- $b$  jets for the two tagging algorithms SVT and SLT. Instead of an analytical solution, which exists, the system is solved by minimizing a likelihood expression for each  $(p_T, \eta)$  bin considered.

Potential biases of the  $b$ -tagging efficiency measurement originate mainly in

- the uncertainties due to the assumption of the decorrelation between

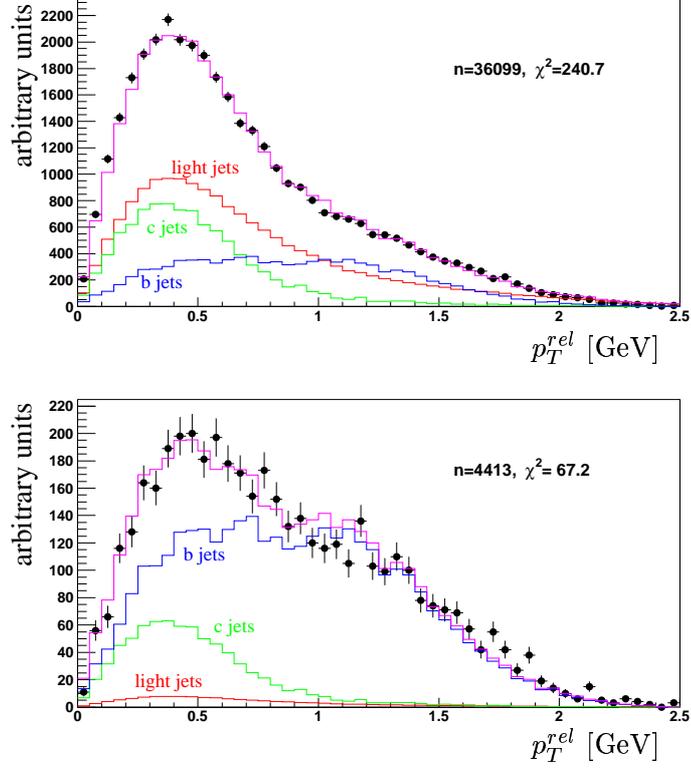


Figure 8.5: The  $p_T^{rel}$  templates for  $b$  jets,  $c$  jets, light jets and the linear combination fitted to the muon-in-jet data sample (solid points) shown before (top) and after (bottom) applying the tagging algorithm.

the SVT and the SLT algorithm

$$\kappa_b = \frac{\varepsilon_b^{SVT,SLT}}{\varepsilon_b^{SVT} \varepsilon_b^{SLT}},$$

- the sample dependence of the SVT tagging efficiency for  $b$  jets

$$\beta = \frac{\varepsilon_b^{SVT}(\text{measured in the muon-in-jet-away-jet-tagged sample})}{\varepsilon_b^{SVT}(\text{measured in the muon-in-jet sample})},$$

- the sample dependence of the SVT tagging efficiency for non- $b$  jets, particularly originating from the assumption that the non- $b$  background composition (i.e. the fraction of  $c$  and light jets) of the two samples is the same

$$\alpha = \frac{\varepsilon_{non-b}^{SVT}(\text{measured in the muon-in-jet-away-jet-tagged sample})}{\varepsilon_{non-b}^{SVT}(\text{measured in the muon-in-jet sample})}.$$

Systematic uncertainties are associated to these biases from measurements in the MC ( $Z \rightarrow b\bar{b} \rightarrow \mu$  is used). Both  $\kappa_b$  and  $\beta$  are found to be compatible with unity.  $\alpha$  is larger than unity, however,  $\varepsilon_b^{SVT}$  is found to be independent on the value of  $\alpha$ .

The average jet  $b$ -tagging efficiencies are summarized in Table 8.4 and are found to be in good agreement with each other. The System8 method is chosen for the determination of the  $b$ -tagging efficiency in the analysis presented, shown in Fig. 8.6, since the associated systematic uncertainties are studied in detail and are well under control.

Method	System8		$p_T^{rel}$
Sample	muon-in-jet	muon-in-jet	muon-in-jet-away-jet-tagged
$\varepsilon_{b \rightarrow \mu}$ [%]	$32.2 \pm 1.2$	$29.8 \pm 0.6$	$30.4 \pm 0.6$

Table 8.4: Average  $b$ -tagging efficiencies (%), obtained by the System8 and the  $p_T^{rel}$  methods in the muon-in-jet and the muon-in-jet-away-jet-tagged sample. Errors are statistical only.

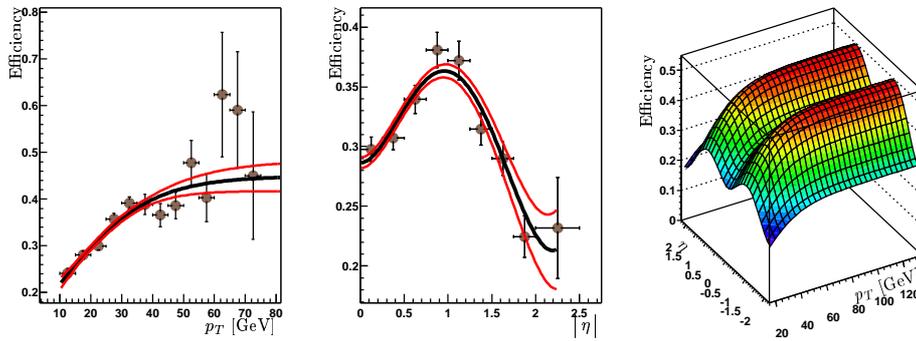


Figure 8.6: Semimuonic  $b$ -tagging efficiency as a function of jet  $p_T$  (left) and jet  $\eta$  (middle) measured in data with the System8 method along with the fit and  $\pm\sigma$  error band. The right plot shows the combined two-dimensional parameterization.

### 8.2.2.3 Inclusive $b$ -Tagging Efficiency

As mentioned above both methods allow only to measure the  $b$ -tagging efficiency for semimuonic  $b$  decays ( $\varepsilon_{b \rightarrow \mu}$ ) which is not necessarily identical to the  $b$ -tagging efficiency for inclusive  $b$  decays ( $\varepsilon_b$ ). A correction factor is

determined in the MC assuming

$$SF_{b \rightarrow \mu}(p_T, \eta) \equiv \frac{\varepsilon_{b \rightarrow \mu}^{data}(p_T, \eta)}{\varepsilon_{b \rightarrow \mu}^{MC}(p_T, \eta)} = \frac{\varepsilon_b^{data}(p_T, \eta)}{\varepsilon_b^{MC}(p_T, \eta)} \equiv SF_b(p_T, \eta). \quad (8.4)$$

Then the tagging efficiency for inclusive  $b$ -decays can be written as:

$$\varepsilon_b(p_T, \eta) \equiv \varepsilon_b^{data}(p_T, \eta) = \varepsilon_b^{MC}(p_T, \eta) SF_b(p_T, \eta). \quad (8.5)$$

The inclusive ( $\varepsilon_b^{MC}$ ) and the semimuonic ( $\varepsilon_{b \rightarrow \mu}^{MC}$ )  $b$ -tagging efficiencies, the latter implicitly contained in  $SF_b(p_T, \eta)$ , are shown in Fig. 8.7.

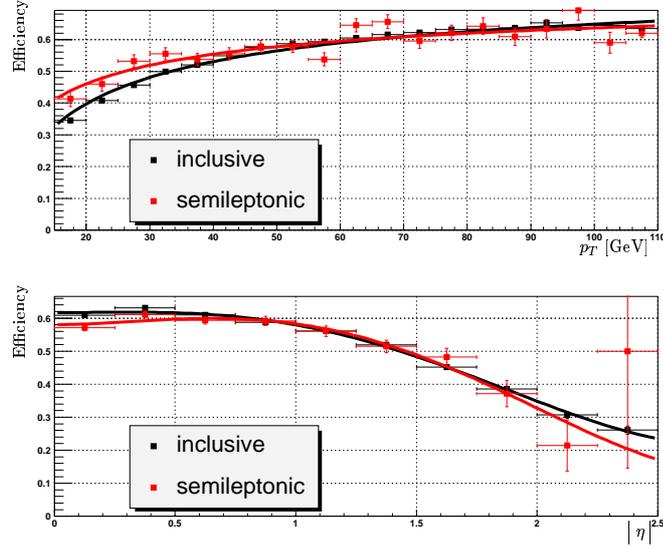


Figure 8.7: Inclusive and semimuonic  $b$ -tagging efficiency measured in the  $t\bar{t}$  Monte Carlo as a function of jet  $p_T$  (top) and  $\eta$  (bottom).

A systematic uncertainty due to the heavy flavor decay model implemented in the MC simulation is associated to  $\varepsilon_b^{MC}$  and  $\varepsilon_{b \rightarrow \mu}^{MC}$  by exchanging EVTGEN by PYTHIA and rederiving the parameterization of  $\varepsilon_b^{MC}(p_T, \eta)$  and  $\varepsilon_{b \rightarrow \mu}^{MC}(p_T, \eta)$ .

### 8.2.3 $c$ -Tagging Efficiency

There is currently no way to measure the  $c$ -tagging efficiency in data, and therefore the  $c$ -tagging scale factor  $SF_c$ . Assuming that  $SF_c(p_T, \eta) = SF_b(p_T, \eta)$  the  $c$ -tagging efficiency can be written as:

$$\varepsilon_c(p_T, \eta) = \varepsilon_c^{MC}(p_T, \eta) SF_c(p_T, \eta), \quad (8.6)$$

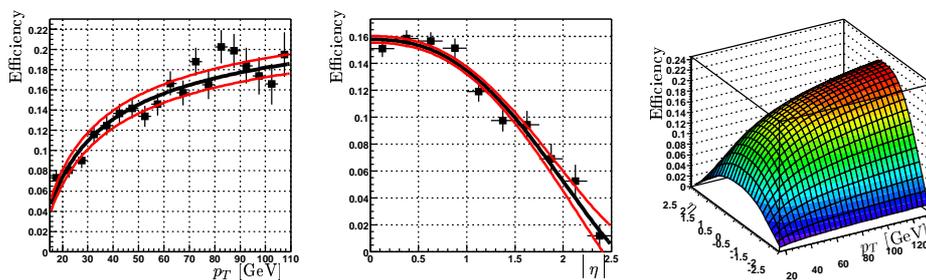


Figure 8.8: Inclusive  $c$ -tagging efficiency measured in the  $t\bar{t}$  Monte Carlo as a function of jet  $p_T$  (left) and  $\eta$  (middle). The right plot shows the combined two-dimensional parameterization.

where  $\varepsilon_c^{MC}(p_T, \eta)$  is shown in Fig. 8.8.

Systematic uncertainties are evaluated due to the assumption that  $SF_c(p_T, \eta) = SF_b(p_T, \eta)$ , by making the replacement:

$$SF_{b/c}(p_T, \eta) \equiv \frac{\varepsilon_{b/c}^{data}(p_T, \eta)}{\varepsilon_{b/c}^{MC}(p_T, \eta)} \rightarrow SF_{b/c}^{rtr}(p_T, \eta) \equiv \frac{\varepsilon_{b/c}^{MC, rtr}(p_T, \eta)}{\varepsilon_{b/c}^{MC}(p_T, \eta)}, \quad (8.7)$$

where  $\varepsilon_{b/c}^{MC, rtr}(p_T, \eta)$  represents the  $b/c$ -tagging efficiency in the MC calibrated to data by a random track removal (rtr) technique [11]. Discrepancies up to 16% are found:  $SF_c^{rtr} \sim 0.84 \cdot SF_b^{rtr}$  and assigned as a systematic uncertainty.

Additionally, a systematic uncertainty due to the heavy flavor decay model implemented in the MC simulation is associated to  $\varepsilon_c^{MC}$  by exchanging EVTGEN by PYTHIA and rederiving the parameterization of  $\varepsilon_c^{MC}(p_T, \eta)$ .

### 8.2.4 Mis-Tagging Efficiency

Mistags, defined as tagged light flavor jets, are caused by random overlap of tracks which are displaced from the primary vertex due to tracking errors and resolution effects. In addition there are contributions from  $K_s^0$  and  $\Lambda^0$  decays, photon conversion ( $\gamma \rightarrow e^+e^-$ ) and nuclear interactions, not completely removed by the  $V^0$  filter, discussed in Sect. 4.7.2.

### Negative Tagging Efficiency

A schematic drawing illustrating the definition of the impact parameter and its sign with respect to a reconstructed track-jet is given in Fig. 4.21. Also shown are the impact parameter significance distributions for tracks from light and  $b$  jets from Monte Carlo. Light jets occur approximately at the

same rate for  $dca/\sigma(dca) > 0$  and for  $dca/\sigma(dca) < 0$ . Since the SVT algorithm is not only symmetric in its treatment of the impact parameter but also the  $L_{xy}$  significance, the tracking related mistags should occur at the same rate for  $L_{xy} > 0$  (defined as a tag or “positive tag” in Sect. 4.7.2) and  $L_{xy} < 0$ , referred to as “negative tag”. Therefore a good estimate for the (positive) mis-tagging efficiency can be obtained from the negative tagging efficiency measured in data,  $\varepsilon_-^{data}(p_T, \eta)$ .

The negative tagging efficiency is measured in the EMqcd data sample and is shown in Fig. 8.9, however, it does not exactly represent the mis-tagging efficiency for the following two reasons.

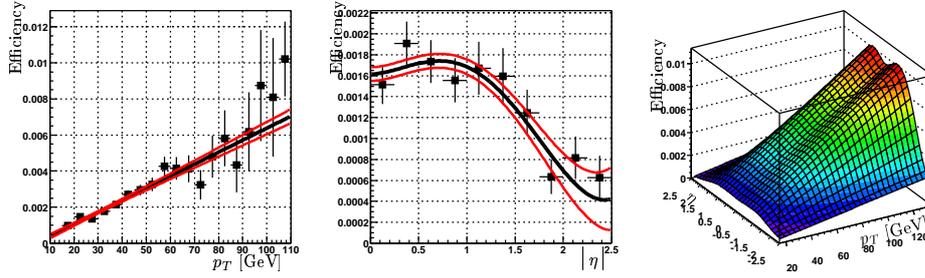


Figure 8.9: Negative tagging efficiency as a function of jet  $p_T$  (left) and jet  $\eta$  (middle). The right plot shows the two-dimensional parameterization.

**Heavy flavor (hf) contamination:**  $\varepsilon_-^{data}(p_T, \eta)$  is determined on the inclusive EMqcd data sample, consisting predominantly of light jets but also of  $b$  and  $c$  jets (2%, respectively 4% as predicted by the PYTHIA QCD-multijets MC simulation). The heavy flavor jets have a higher negative tagging efficiency than light jets. Since a pure sample of light jets cannot be selected in data, a correction factor  $SF_{hf}$  is derived from the PYTHIA QCD-multijets MC simulation:

$$SF_{hf}(p_T, \eta) = \frac{\varepsilon_-^{light}(p_T, \eta)}{\varepsilon_-^{inclusive}(p_T, \eta)}, \quad (8.8)$$

shown in Fig. 8.10, so that

$$\varepsilon_-^{light}(p_T, \eta) = \varepsilon_-^{data}(p_T, \eta) SF_{hf}(p_T, \eta) \quad (8.9)$$

represents the negative mis-tagging efficiency.

**Long lived (ll) particles:** The long lived particles ( $K_s^0$ ,  $\Lambda^0$  decays and interactions with material) present in light jets lead to a larger positive than negative tagging efficiency since they have real lifetime. A

correction factor  $SF_{ll}$  is derived from the  $Wjjjj$  MC simulation, representing the data sample after the preselection:

$$SF_{ll}(p_T, \eta) = \frac{\varepsilon_+^{light}(p_T, \eta)}{\varepsilon_-^{light}(p_T, \eta)} \quad (8.10)$$

also shown in Fig. 8.10, so that

$$\varepsilon_+^{light}(p_T, \eta) = \varepsilon_-^{light}(p_T, \eta) SF_{ll}(p_T, \eta) \quad (8.11)$$

finally represents the positive mis-tagging efficiency.

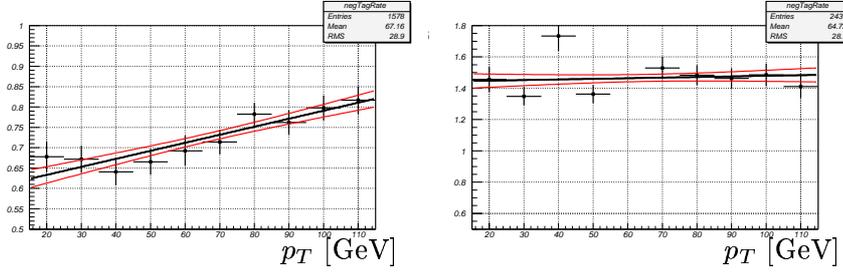


Figure 8.10: Correction factors for contribution of heavy flavor in the negative tag rate ( $SF_{hf}$ , left plot) and contribution to the mistag rate from long lived light particles ( $SF_{ll}$ , right plot).

Both corrections are found to be independent of the jet  $\eta$ , so that the (positive) mis-tagging efficiency is given by

$$\varepsilon_+^{light}(p_T, \eta) = \varepsilon_-^{data}(p_T, \eta) SF_{hf}(p_T) SF_{ll}(p_T). \quad (8.12)$$

The systematic uncertainty associated to the mis-tagging efficiency comprises the statistical uncertainties of the fits to derive the parameterizations for  $\varepsilon_-^{data}(p_T, \eta)$ ,  $SF_{hf}(p_T)$  and  $SF_{ll}(p_T)$  and the sample dependence of  $\varepsilon_-^{data}(p_T, \eta)$ , derived by exchanging the EMqcd sample by the jettrig sample (selected requiring any of the jet triggers); the latter leads to an 8 % uncertainty.

### 8.3 Contributions to the Data Set

The described measurement of the  $t\bar{t}$  cross section requires the knowledge of the probability to identify one or more jets in a  $t\bar{t}$  event as  $b$  jets, referred to as  $\bar{P}^{tag}$  in Eq. 8.1. The probability  $\bar{P}_{background}^{tag}$  is also required for the background prediction, implicitly contained in  $N_{background}^{tag}$  in Eq. 8.1. Since a pure sample both of  $t\bar{t}$  events and most of the background events cannot

be extracted with high statistics in data, MC simulation is used to calculate the event tagging probabilities.

The following method is utilized to evaluate the expected number of tagged signal and background events from the MC simulation:

- The jet tagging probabilities  $\mathcal{P}_\alpha(p_T, \eta)$ , derived in the previous Sect. 8.2, are applied to the preselected MC in order to derive the event tagging probabilities  $P_{event}$  by weighting each reconstructed jet in the MC event by the jet tagging probability  $\mathcal{P}_\alpha(p_T, \eta)$  according to its flavor  $\alpha$ , its  $p_T$  and its  $\eta$ . The probability to have at least one tag in a given event is given by the complement of the probability that none of the jets is tagged:

$$P_{event}^{tag}(\geq 1 \text{ tag}) = 1 - P_{event}^{tag}(0 \text{ tags}) \quad (8.13)$$

$$= 1 - \prod_{j=1}^{N_{jets}} (1 - \mathcal{P}_{\alpha_j}(p_{T_j}, \eta_j)). \quad (8.14)$$

The probabilities to have exactly one and to have two or more tags in an event are used in the analysis presented and are:

$$P_{event}^{tag}(1 \text{ tag}) = \sum_{j=1}^{N_{jets}} \mathcal{P}_{\alpha_j}(p_{T_j}, \eta_j) \prod_{i \neq j} (1 - \mathcal{P}_{\alpha_i}(p_{T_i}, \eta_i)), \quad (8.15)$$

and

$$P_{event}^{tag}(\geq 2 \text{ tags}) = P_{event}^{tag}(\geq 1 \text{ tag}) - P_{event}^{tag}(1 \text{ tag}), \quad (8.16)$$

respectively. The average event tagging probability can be obtained by averaging over all preselected Monte Carlo events.

- The trigger can significantly distort the jet  $p_T$  and also the  $\eta$  spectrum, particularly for the low jet multiplicity bins, and therefore bias the estimate of the average event tagging probability. The probability of an event to satisfy the  $\mu$ +jets trigger is given in Sect. 5, particularly by Eq. 5.6, and can be found for the  $e$ +jets trigger in [141]. The event tagging probability is corrected for this trigger bias by weighting every event with the trigger probability  $P_{event}^{trigg}$ . The average event tagging probability, taking into account the trigger bias, is then computed as:

$$\bar{P}_{event}^{tag} = \frac{\langle P_{event}^{tag} P_{event}^{trigg} \rangle}{\langle P_{event}^{trigg} \rangle}. \quad (8.17)$$

- The expected number of tagged events is obtained by multiplying the expected number of preselected events with the average event tagging probability:

$$N_{event}^{tag} = N_{event}^{presel} \bar{P}_{event}^{tag}. \quad (8.18)$$

### 8.3.1 $t\bar{t}$ Signal

According to Eq. 8.18 and Eq. 7.8 the expected number of tagged  $t\bar{t}$  signal events is given by:

$$N_{l+jets}^{tag,expected} = \sigma_{t\bar{t}} \cdot \varepsilon_{l+jets} \cdot BR_{l+jets} \cdot L_{l+jets} \cdot \bar{P}_{l+jets}^{tag}, \quad (8.19)$$

for the lepton+jets channels and in the same way for the dilepton channels, where the branching ratios are given in Table 2.6, the integrated luminosities  $L_{l+jets}$  are given in Table 6.2, the expected  $t\bar{t}$  cross section is assumed to be 7 pb (see also Table 2.2) and the preselection efficiencies,  $\varepsilon_{l+jets}$ , are given in Table 8.3 as a function of the jet multiplicity (see also Table 8.5 and Table 8.6 for a summary of all preselection efficiencies in the  $\mu$ +jets and  $e$ +jets channel).

	=1 jet	=2 jets	=3 jets	$\geq 4$ jets
$t\bar{t} \rightarrow l+jets$	0.086 $\pm$ 0.020	1.64 $\pm$ 0.11	8.64 $\pm$ 0.41	15.91 $\pm$ 0.74
$t\bar{t} \rightarrow ll$	1.55 $\pm$ 0.09	7.60 $\pm$ 0.34	5.58 $\pm$ 0.26	1.96 $\pm$ 0.11
$tb$	3.06 $\pm$ 0.15	12.18 $\pm$ 0.52	4.34 $\pm$ 0.21	0.951 $\pm$ 0.066
$tqb$	3.10 $\pm$ 0.16	9.25 $\pm$ 0.40	5.61 $\pm$ 0.26	2.28 $\pm$ 0.13
$WW \rightarrow l\nu jj$	4.36 $\pm$ 0.28	12.80 $\pm$ 0.61	2.12 $\pm$ 0.17	0.271 $\pm$ 0.055
$WZ \rightarrow l\nu jj$	4.12 $\pm$ 0.27	13.27 $\pm$ 0.63	2.56 $\pm$ 0.19	0.302 $\pm$ 0.056
$WZ \rightarrow jjll$	3.49 $\pm$ 0.22	9.42 $\pm$ 0.46	2.07 $\pm$ 0.15	0.247 $\pm$ 0.046
$ZZ \rightarrow jjll$	2.83 $\pm$ 0.21	9.96 $\pm$ 0.50	2.80 $\pm$ 0.20	0.363 $\pm$ 0.064
$Z \rightarrow \tau^+\tau^-$	0.170 $\pm$ 0.020	0.118 $\pm$ 0.014	0.044 $\pm$ 0.007	0.005 $\pm$ 0.002

Table 8.5: Summary of preselection efficiencies in the  $\mu$ +jets channel (%). Errors are statistical only.

	=1 jet	=2 jets	=3 jets	$\geq 4$ jets
$t\bar{t} \rightarrow l+jets$	0.182 $\pm$ 0.029	2.42 $\pm$ 0.11	9.04 $\pm$ 0.21	15.20 $\pm$ 0.29
$t\bar{t} \rightarrow ll$	1.99 $\pm$ 0.08	9.17 $\pm$ 0.17	5.86 $\pm$ 0.13	2.06 $\pm$ 0.08
$tb$	3.78 $\pm$ 0.10	14.11 $\pm$ 0.21	4.83 $\pm$ 0.12	1.08 $\pm$ 0.05
$tqb$	3.66 $\pm$ 0.10	11.03 $\pm$ 0.18	6.25 $\pm$ 0.13	2.40 $\pm$ 0.08
$WW \rightarrow l\nu jj$	5.05 $\pm$ 0.23	12.98 $\pm$ 0.34	2.04 $\pm$ 0.14	0.252 $\pm$ 0.049
$WZ \rightarrow l\nu jj$	4.57 $\pm$ 0.22	12.46 $\pm$ 0.33	2.13 $\pm$ 0.14	0.278 $\pm$ 0.051
$WZ \rightarrow jjll$	0.330 $\pm$ 0.052	0.950 $\pm$ 0.085	0.530 $\pm$ 0.064	0.140 $\pm$ 0.033
$ZZ \rightarrow jjll$	0.362 $\pm$ 0.051	1.24 $\pm$ 0.09	0.700 $\pm$ 0.067	0.153 $\pm$ 0.032
$Z \rightarrow \tau^+\tau^-$	0.088 $\pm$ 0.016	0.067 $\pm$ 0.009	0.009 $\pm$ 0.003	0.007 $\pm$ 0.003

Table 8.6: Summary of preselection efficiencies in the  $e$ +jets channel (%). Errors are statistical only.

	=1 jet	=2 jets	=3 jets	$\geq 4$ jets
$W$ +light	$0.36\pm 0.01$	$0.61\pm 0.01$	$0.89\pm 0.01$	$1.18\pm 0.01$
$W(c\bar{c})$	$8.7\pm 0.1$	$7.6\pm 0.1$	$7.6\pm 0.2$	$6.6\pm 1.0$
$W(b\bar{b})$	$35.3\pm 0.2$	$31.8\pm 0.2$	$31.8\pm 0.3$	$30.0\pm 2.3$
$Wc$	$8.7\pm 0.1$	$8.2\pm 0.1$	$8.2\pm 0.1$	$8.7\pm 0.2$
$Wc\bar{c}$		$14.2\pm 0.3$	$14.0\pm 0.3$	$14.1\pm 0.6$
$Wb\bar{b}$		$43.5\pm 0.2$	$42.5\pm 0.3$	$42.2\pm 0.3$
$W$ +jets	$1.06\pm 0.01$	$2.13\pm 0.01$	$2.98\pm 0.02$	$3.75\pm 0.06$
$t\bar{t} \rightarrow l$ +jets	$27.5\pm 3.9$	$37.9\pm 0.6$	$43.0\pm 0.2$	$44.7\pm 0.1$
$t\bar{t} \rightarrow ll$	$38.1\pm 0.3$	$44.7\pm 0.1$	$45.0\pm 0.1$	$44.9\pm 0.2$
$tb$	$36.6\pm 0.2$	$44.5\pm 0.1$	$44.7\pm 0.1$	$44.5\pm 0.3$
$tqb$	$30.2\pm 0.4$	$35.2\pm 0.2$	$38.5\pm 0.2$	$40.6\pm 0.3$
$WW \rightarrow l\nu jj$	$3.28\pm 0.21$	$4.09\pm 0.11$	$4.45\pm 0.30$	$4.01\pm 0.68$
$WZ \rightarrow l\nu jj$	$12.1\pm 0.7$	$11.6\pm 0.5$	$9.3\pm 0.9$	$13.4\pm 3.2$
$WZ \rightarrow jjll$	$2.57\pm 0.18$	$3.87\pm 0.11$	$4.22\pm 0.28$	$2.95\pm 0.56$
$ZZ \rightarrow jjll$	$13.6\pm 1.0$	$13.1\pm 0.6$	$14.0\pm 1.1$	$13.9\pm 3.1$
$Z \rightarrow \tau^+ \tau^-$	$1.40\pm 0.35$	$1.37\pm 0.34$	$2.27\pm 0.92$	$2.59\pm 0.99$

Table 8.7: Event tagging probabilities (%) for 1 tag in the  $\mu$ +jets channel for the  $t\bar{t}$  signal and all backgrounds determined from MC simulation. Errors are statistical only.

The event tagging probabilities,  $\bar{P}_{l+jets}^{tag}$ , are given in Table 8.7 and Table 8.8 for exactly one tag and in Table 8.9 and Table 8.10 for  $\geq 2$  tags for the  $\mu$ +jets, and the  $e$ +jets channel, respectively. The number of expected  $t\bar{t}$  events are given in Table 8.11 and Table 8.12 for exactly one tag and in Table 8.13 and Table 8.14 for  $\geq 2$  tags for the  $\mu$ +jets, and the  $e$ +jets channel, respectively.

### 8.3.2 Matrix Method - Evaluation of the QCD Background

As for the *topo* analysis, the QCD-multijets background is purely determined from data using the matrix method by defining a loose and a tight set, the latter representing the full preselection and being a subset of the first, described in Sect. 7.2.3. For the  $\mu$ +jets sample the loose sample is obtained by dropping the tight isolation requirement, for the  $e$ +jets sample the electron likelihood requirement (see Sect. 4.4) is dropped.

The QCD-multijets background can be evaluated in two ways, either by applying the matrix method after the  $b$ -tagging, or by applying the matrix method before the  $b$ -tagging and by evaluating the QCD-multijets event tagging probability:

	=1 jet	=2 jets	=3 jets	$\geq 4$ jets
$W$ +light	$0.30\pm 0.01$	$0.53\pm 0.01$	$0.81\pm 0.02$	$1.14\pm 0.01$
$W(c\bar{c})$	$8.2\pm 0.1$	$7.4\pm 0.1$	$7.3\pm 0.2$	$6.8\pm 1.3$
$W(b\bar{b})$	$34.1\pm 0.2$	$31.3\pm 0.2$	$31.5\pm 0.3$	$33.4\pm 1.9$
$Wc$	$8.2\pm 0.1$	$7.8\pm 0.1$	$8.4\pm 0.2$	$8.9\pm 0.2$
$Wc\bar{c}$		$13.5\pm 0.3$	$13.4\pm 0.3$	$13.5\pm 0.7$
$Wb\bar{b}$		$43.2\pm 0.2$	$42.2\pm 0.3$	$42.3\pm 0.4$
$W$ +jets	$0.97\pm 0.01$	$2.05\pm 0.01$	$2.96\pm 0.02$	$3.81\pm 0.07$
$t\bar{t} \rightarrow l$ +jets	$23.0\pm 3.0$	$39.3\pm 0.5$	$43.1\pm 0.2$	$45.1\pm 0.1$
$t\bar{t} \rightarrow ll$	$38.4\pm 0.3$	$44.9\pm 0.1$	$45.2\pm 0.1$	$45.3\pm 0.2$
$tb$	$36.8\pm 0.2$	$44.8\pm 0.1$	$44.7\pm 0.1$	$44.6\pm 0.3$
$tqb$	$30.5\pm 0.4$	$35.8\pm 0.2$	$38.8\pm 0.2$	$41.1\pm 0.2$
$WW \rightarrow l\nu jj$	$2.56\pm 0.16$	$4.25\pm 0.12$	$4.13\pm 0.28$	$4.26\pm 0.89$
$WZ \rightarrow l\nu jj$	$12.5\pm 0.7$	$14.0\pm 0.5$	$14.7\pm 1.2$	$9.2\pm 2.8$
$WZ \rightarrow jjll$	$3.50\pm 0.60$	$4.40\pm 0.44$	$4.30\pm 0.50$	$6.0\pm 1.6$
$ZZ \rightarrow jjll$	$13.4\pm 2.3$	$17.3\pm 1.4$	$15.2\pm 1.8$	$13.7\pm 3.8$
$Z \rightarrow \tau^+\tau^-$	$1.35\pm 0.43$	$1.82\pm 0.85$	$6.0\pm 4.3$	$0.94\pm 0.17$

Table 8.8: Event tagging probabilities (%) for 1 tag in the  $e$ +jets channel for the  $t\bar{t}$  signal and all backgrounds determined from MC simulation. Errors are statistical only.

- 1. Applying the matrix method on the tagged data sample:** The loose and the tight samples, originally defined on the preselected sample in Sect. 7.2.3, can be defined in the same way for the tagged samples (1 tag and  $\geq 2$  tags). Solving the linear system of equations in Eq. 7.10 for the tagged samples

$$\begin{aligned}
N_\ell^{tag} &= N_\ell^{tag,sig} + N_\ell^{tag,QCD} \\
N_t^{tag} &= \varepsilon_{sig}^{tag} N_\ell^{tag,sig} + \varepsilon_{QCD}^{tag} N_\ell^{tag,QCD}
\end{aligned} \tag{8.20}$$

allows to directly extract  $N_t^{tag,QCD}$ , according to Eq. 7.12. Both  $\varepsilon_{sig}^{tag}$  and  $\varepsilon_{QCD}^{tag}$  are measured in the same way as described in Sect. 7.2.3. In contrast to  $\varepsilon_{sig}^{tag}$ ,  $\varepsilon_{QCD}^{tag}$  is expected to be possibly different from the untagged  $\varepsilon_{QCD}$ , i.e.  $\varepsilon_{QCD}$  might depend on the heavy flavor composition, since the jets are the source of both the fake electron and the muons with fake isolation. However, given the limited statistics in the tagged samples, no discrepancies could be found and the untagged values are used, i.e.  $\varepsilon_{sig}^{tag} = \varepsilon_{sig}$  and  $\varepsilon_{QCD}^{tag} = \varepsilon_{QCD}$ .

- 2. Applying the matrix method on the untagged data sample:** The number of preselected (untagged) QCD-multijets events is deter-

	=2 jets	=3 jets	$\geq 4$ jets
$W$ +light	$< 0.01$	$< 0.01$	$< 0.01$
$W(c\bar{c})$	$0.036\pm 0.003$	$0.06\pm 0.01$	$0.04\pm 0.01$
$W(b\bar{b})$	$0.47\pm 0.05$	$0.89\pm 0.08$	$1.18\pm 0.78$
$Wc$	$0.020\pm 0.001$	$0.041\pm 0.002$	$0.062\pm 0.003$
$Wc\bar{c}$	$0.60\pm 0.03$	$0.65\pm 0.04$	$0.67\pm 0.07$
$Wb\bar{b}$	$10.3\pm 0.2$	$10.0\pm 0.2$	$9.8\pm 0.3$
$W$ +jets	$0.098\pm 0.002$	$0.188\pm 0.004$	$0.31\pm 0.02$
$t\bar{t} \rightarrow l$ +jets	$4.52\pm 0.32$	$10.3\pm 0.2$	$13.5\pm 0.1$
$t\bar{t} \rightarrow ll$	$10.9\pm 0.1$	$11.6\pm 0.1$	$12.0\pm 0.2$
$tb$	$11.1\pm 0.1$	$11.5\pm 0.1$	$11.9\pm 0.3$
$tqb$	$1.29\pm 0.06$	$5.1\pm 0.1$	$7.6\pm 0.2$
$WW \rightarrow l\nu jj$	$< 0.01$	$0.03\pm 0.01$	$0.03\pm 0.01$
$WZ \rightarrow l\nu jj$	$2.29\pm 0.11$	$1.70\pm 0.21$	$2.68\pm 0.77$
$WZ \rightarrow jjll$	$< 0.01$	$0.03\pm 0.01$	$0.02\pm 0.01$
$ZZ \rightarrow jjll$	$2.45\pm 0.13$	$3.01\pm 0.30$	$2.36\pm 0.63$
$Z \rightarrow \tau^+\tau^-$	$< 0.01$	$0.22\pm 0.19$	$0.02\pm 0.01$

Table 8.9: Event tagging probabilities (%) for  $\geq 2$  tags in the  $\mu$ +jets channel for the  $t\bar{t}$  signal and all backgrounds determined from MC simulation. Errors are statistical only.

mined according to Eq. 7.12 by solving the matrix method as given in Eq. 7.10. In order to extract the number of tagged QCD-multijets events, according to Eq. 8.18, requires the knowledge of the average event tagging probability for a QCD-multijets event,  $\bar{P}_{QCD}^{tag}$ . The latter is measured on the “loose-tight” preselected data sample, consisting of events passing the loose but not the tight selection requirement, known to be dominated by QCD-multijets events (see Sect. 7.4.1).

The results of the matrix method for the preselected (untagged) and the tagged (1 tag) data samples are shown in Table 8.15 and Table 8.16 for the  $\mu$ +jets, and the  $e$ +jets channel, respectively. Also shown are the measurements of  $\varepsilon_{sig}$  and  $\varepsilon_{QCD}$ .  $N_t^{tag, QCD}$  represents the evaluated QCD-multijets background according to the first method.

The second method assumes that the heavy flavor composition in the “loose-tight” data sample, where the event tagging probability is derived from, is identical to the heavy flavor composition in the preselected (tight) sample. In the  $e$ +jets channel this assumption applies, since the instrumental background mainly originates from electromagnetically fluctuating jets misreconstructed as electrons. Thus, the heavy flavor fraction of the remaining jets in the event is not expected to depend on the quality re-

	=2 jets	=3 jets	$\geq 4$ jets
$W$ +light	< 0.01	< 0.01	< 0.01
$W(c\bar{c})$	$0.033\pm 0.003$	$0.07\pm 0.01$	$0.04\pm 0.02$
$W(b\bar{b})$	$0.42\pm 0.05$	$0.70\pm 0.07$	$1.08\pm 0.42$
$Wc$	$0.019\pm 0.001$	$0.042\pm 0.002$	$0.065\pm 0.004$
$Wc\bar{c}$	$0.54\pm 0.03$	$0.57\pm 0.03$	$0.63\pm 0.07$
$Wb\bar{b}$	$10.1\pm 0.2$	$9.7\pm 0.3$	$10.2\pm 0.4$
$W$ +jets	$0.101\pm 0.002$	$0.162\pm 0.004$	$0.28\pm 0.01$
$t\bar{t} \rightarrow l$ +jets	$5.3\pm 0.3$	$10.2\pm 0.2$	$13.9\pm 0.1$
$t\bar{t} \rightarrow ll$	$11.0\pm 0.1$	$11.9\pm 0.1$	$12.5\pm 0.2$
$tb$	$11.5\pm 0.1$	$11.8\pm 0.1$	$11.7\pm 0.3$
$tqb$	$1.36\pm 0.05$	$5.4\pm 0.1$	$7.9\pm 0.2$
$WW \rightarrow l\nu jj$	$0.014\pm 0.003$	$0.017\pm 0.004$	$0.07\pm 0.04$
$WZ \rightarrow l\nu jj$	$2.76\pm 0.12$	$3.05\pm 0.33$	$1.58\pm 0.61$
$WZ \rightarrow jjll$	$0.011\pm 0.001$	$0.025\pm 0.004$	$0.04\pm 0.01$
$ZZ \rightarrow jjll$	$2.97\pm 0.37$	$2.91\pm 0.45$	$3.27\pm 1.26$
$Z \rightarrow \tau^+\tau^-$	$0.36\pm 0.26$	$1.54\pm 1.47$	< 0.01

Table 8.10: Event tagging probabilities (%) for  $\geq 2$  tags in the  $e$ +jets channel for the  $t\bar{t}$  signal and all backgrounds determined from MC simulation. Errors are statistical only.

quirements on this jet (i.e. applying the electron likelihood cut or inverting it). The event tagging probability and the predicted number of tagged QCD-multijets events,  $N_t^{QCD} \bar{P}_{QCD}^{tag}$ , is shown in Table 8.16 and is found to be in very good agreement with the result obtained by the other method ( $N_t^{tag, QCD}$ ). In the  $\mu$ +jets channel the instrumental background originates mainly from semimuonically decaying  $b$  quarks. The heavy flavor fraction is enriched when the isolation criteria is inverted, leading to a higher event tagging probability and therefore the second method cannot be applied to the  $\mu$ +jets channel.

Contributions from QCD-multijets background with  $\geq 2$  tags are estimated to be negligible and are therefore not considered.

### 8.3.3 Small Backgrounds

There is a number of low rate electroweak physics backgrounds with heavy flavor in the preselected sample. The contributions from the following processes:

- diboson production:  $WW \rightarrow l + jets$ ,  $WZ \rightarrow l + jets$ ,  $WZ \rightarrow jjl\bar{l}$ ,  $ZZ \rightarrow lljj$ ,

	=1 jet	=2 jets	=3 jets	$\geq 4$ jets
$W$ +light	$16.3\pm 0.4$	$10.3\pm 0.3$	$3.33\pm 0.18$	$0.76\pm 0.11$
$W(c\bar{c})$	$4.39\pm 0.09$	$2.76\pm 0.08$	$0.95\pm 0.06$	$0.14\pm 0.03$
$W(b\bar{b})$	$11.3\pm 0.2$	$7.2\pm 0.2$	$2.47\pm 0.13$	$0.43\pm 0.07$
$Wc$	$20.0\pm 0.4$	$11.0\pm 0.3$	$2.28\pm 0.13$	$0.36\pm 0.05$
$Wc\bar{c}$		$2.72\pm 0.09$	$1.21\pm 0.07$	$0.36\pm 0.05$
$Wb\bar{b}$		$6.8\pm 0.2$	$2.86\pm 0.15$	$0.83\pm 0.12$
$W$ +jets	$52.0\pm 0.6$	$40.7\pm 0.5$	$13.1\pm 0.3$	$2.88\pm 0.20$
QCD	$5.4\pm 0.9$	$3.75\pm 0.71$	$1.75\pm 0.51$	$0.75\pm 0.30$
$t\bar{t} \rightarrow ll$	$0.32\pm 0.01$	$1.83\pm 0.03$	$1.35\pm 0.03$	$0.46\pm 0.02$
$tb$	$0.20\pm 0.01$	$0.95\pm 0.01$	$0.34\pm 0.01$	$0.074\pm 0.004$
$tqb$	$0.37\pm 0.01$	$1.28\pm 0.02$	$0.85\pm 0.02$	$0.36\pm 0.01$
$WW \rightarrow l\nu jj$	$0.24\pm 0.02$	$0.87\pm 0.03$	$0.16\pm 0.01$	$0.018\pm 0.004$
$WZ \rightarrow l\nu jj$	$0.26\pm 0.02$	$0.79\pm 0.04$	$0.12\pm 0.01$	$0.02\pm 0.01$
$WZ \rightarrow jjll$	$0.015\pm 0.001$	$0.062\pm 0.002$	$0.015\pm 0.001$	$< 0.01$
$ZZ \rightarrow jjll$	$0.05\pm 0.01$	$0.19\pm 0.01$	$0.06\pm 0.01$	$< 0.01$
$Z \rightarrow \tau^+\tau^-$	$0.32\pm 0.08$	$0.22\pm 0.06$	$0.14\pm 0.06$	$0.02\pm 0.01$
background	$59.2\pm 1.0$	$50.7\pm 0.9$	$17.9\pm 0.6$	$4.60\pm 0.36$
syst.	$+5.46$ $-6.80$	$+3.82$ $-4.64$	$+1.24$ $-1.40$	$+0.29$ $-0.34$
$t\bar{t} \rightarrow l$ +jets	$0.04\pm 0.01$	$1.17\pm 0.06$	$7.0\pm 0.2$	$13.5\pm 0.2$
total	$59.2\pm 1.0$	$51.8\pm 0.9$	$24.9\pm 0.6$	$18.1\pm 0.4$
syst.	$+5.47$ $-6.78$	$+3.85$ $-4.59$	$+1.47$ $-1.62$	$+1.55$ $-2.37$
# tags	55	55	39	13

Table 8.11: Summary of observed and predicted number of events with 1 tag in the  $\mu$ +jets channel. Errors are statistical only. Systematic uncertainties are given on the background and the total prediction.

- single top production in the  $s$ - and  $t$ -channel,
- $Z \rightarrow \tau\tau \rightarrow l + jets$

are evaluated in the same way as for the expected  $t\bar{t}$  signal using Eq. 8.19 with the preselection efficiencies given in Table 8.5 and Table 8.6, and the event tagging probabilities given in Table 8.7 and Table 8.8 for exactly one tag and in Table 8.9 and Table 8.10 for  $\geq 2$  tags for the  $\mu$ +jets, and the  $e$ +jets channel, respectively. The expected cross sections and the corresponding branching ratios for the  $\mu$ +jets and the  $e$ +jets channels are given in Table 8.17.

The resulting expected numbers of events are given in Table 8.11 and Table 8.12 for exactly one tag and in Table 8.13 and Table 8.14 for  $\geq 2$  tags

	=1 jet	=2 jets	=3 jets	$\geq 4$ jets
$W$ +light	$17.7\pm 0.4$	$10.3\pm 0.3$	$3.45\pm 0.18$	$1.18\pm 0.13$
$W(c\bar{c})$	$5.7\pm 0.1$	$3.13\pm 0.09$	$0.99\pm 0.06$	$0.29\pm 0.06$
$W(b\bar{b})$	$14.7\pm 0.2$	$8.5\pm 0.2$	$3.17\pm 0.16$	$0.94\pm 0.12$
$Wc$	$23.0\pm 0.4$	$11.7\pm 0.3$	$3.03\pm 0.16$	$0.58\pm 0.07$
$Wc\bar{c}$		$3.32\pm 0.10$	$1.42\pm 0.08$	$0.58\pm 0.07$
$Wb\bar{b}$		$8.3\pm 0.2$	$2.84\pm 0.14$	$1.16\pm 0.13$
$W$ +jets	$61.1\pm 0.6$	$45.3\pm 0.5$	$14.9\pm 0.3$	$4.73\pm 0.24$
QCD	$1.72\pm 0.35$	$3.86\pm 0.57$	$2.52\pm 0.50$	$0.93\pm 0.28$
$t\bar{t} \rightarrow ll$	$0.43\pm 0.02$	$2.35\pm 0.04$	$1.52\pm 0.03$	$0.52\pm 0.02$
$tb$	$0.27\pm 0.01$	$1.19\pm 0.01$	$0.41\pm 0.01$	$0.090\pm 0.004$
$tqb$	$0.48\pm 0.01$	$1.67\pm 0.02$	$1.03\pm 0.02$	$0.42\pm 0.01$
$WW \rightarrow l\nu jj$	$0.23\pm 0.02$	$0.98\pm 0.04$	$0.15\pm 0.01$	$0.02\pm 0.01$
$WZ \rightarrow l\nu jj$	$0.31\pm 0.02$	$0.95\pm 0.04$	$0.17\pm 0.02$	$0.014\pm 0.005$
$WZ \rightarrow jjll$	$< 0.01$	$< 0.01$	$< 0.01$	$< 0.01$
$ZZ \rightarrow jjll$	$< 0.01$	$0.033\pm 0.003$	$0.016\pm 0.002$	$< 0.01$
$Z \rightarrow \tau^+\tau^-$	$0.17\pm 0.06$	$0.18\pm 0.09$	$0.07\pm 0.06$	$< 0.01$
background	$64.7\pm 0.7$	$56.5\pm 0.8$	$20.8\pm 0.6$	$6.7\pm 0.4$
syst.	$+6.83$ $-8.24$	$+4.49$ $-5.14$	$+1.53$ $-1.71$	$+0.54$ $-0.45$
$t\bar{t} \rightarrow l$ +jets	$0.08\pm 0.02$	$1.92\pm 0.08$	$7.9\pm 0.2$	$13.9\pm 0.2$
total	$64.8\pm 0.7$	$58.4\pm 0.8$	$28.7\pm 0.6$	$20.6\pm 0.4$
syst.	$+6.82$ $-8.24$	$+4.49$ $-5.11$	$+1.79$ $-1.89$	$+1.95$ $-2.55$
# tags	64	65	29	30

Table 8.12: Summary of observed and predicted number of events with 1 tag in the  $e$ +jets channel. Errors are statistical only. Systematic uncertainties are given on the background and the total prediction.

for the  $\mu$ +jets, respectively the  $e$ +jets channel.

### 8.3.4 $W$ -plus-jets and $Z$ -plus-jets Background

As in the *topo* analysis, the largest background originates from the electroweak  $W$  boson production in association with jets. In the *topo* analysis the  $Z$ -plus-jets background is found to be a small fraction ( $\lesssim 10\%$ ) of the  $W$ -plus-jets events. The jet production mechanisms are very similar for the  $W$ -plus-jets and the  $Z$ -plus-jets processes, therefore the  $W$ -plus-jets MC samples are used to represent both  $W$ -plus-jets and  $Z$ -plus-jets events. The MC samples do not include the Tune A underlying event description and

	=2 jets	=3 jets	$\geq 4$ jets
$W$ +light	$0.017\pm 0.001$	$0.011\pm 0.001$	$< 0.01$
$W(c\bar{c})$	$0.013\pm 0.001$	$< 0.01$	$< 0.01$
$W(b\bar{b})$	$0.11\pm 0.01$	$0.07\pm 0.01$	$0.02\pm 0.01$
$Wc$	$0.027\pm 0.001$	$0.011\pm 0.001$	$< 0.01$
$Wc\bar{c}$	$0.11\pm 0.01$	$0.056\pm 0.004$	$0.017\pm 0.003$
$Wb\bar{b}$	$1.59\pm 0.05$	$0.67\pm 0.04$	$0.19\pm 0.03$
$W$ +jets	$1.87\pm 0.05$	$0.83\pm 0.04$	$0.23\pm 0.03$
QCD	$< 0.01$	$< 0.01$	$< 0.01$
$t\bar{t} \rightarrow ll$	$0.44\pm 0.01$	$0.35\pm 0.01$	$0.123\pm 0.005$
$tb$	$0.236\pm 0.003$	$0.087\pm 0.002$	$0.020\pm 0.001$
$tqb$	$0.047\pm 0.002$	$0.113\pm 0.004$	$0.068\pm 0.003$
$WW \rightarrow l\nu jj$	$< 0.01$	$< 0.01$	$< 0.01$
$WZ \rightarrow l\nu jj$	$0.15\pm 0.01$	$0.022\pm 0.003$	$< 0.01$
$WZ \rightarrow jjll$	$< 0.01$	$< 0.01$	$< 0.01$
$ZZ \rightarrow jjll$	$0.035\pm 0.002$	$0.012\pm 0.001$	$< 0.01$
$Z \rightarrow \tau^+\tau^-$	$< 0.01$	$0.01\pm 0.01$	$< 0.01$
background	$2.72\pm 0.07$	$1.37\pm 0.05$	$0.40\pm 0.04$
syst.	$+0.54$ $-0.49$	$+0.27$ $-0.26$	$+0.09$ $-0.11$
$t\bar{t} \rightarrow l$ +jets	$0.14\pm 0.01$	$1.68\pm 0.05$	$4.07\pm 0.07$
total	$2.86\pm 0.07$	$3.05\pm 0.07$	$4.47\pm 0.08$
syst.	$+0.56$ $-0.51$	$+0.58$ $-0.54$	$+0.98$ $-1.08$
# tags	4	3	3

Table 8.13: Summary of observed and predicted number of events with  $\geq 2$  tags in the  $\mu$ +jets channel. Errors are statistical only. Systematic uncertainties are given on the background and the total prediction.

the scale is set to  $Q^2 = \left(\frac{M_W}{2}\right)^2$ , as described in Table 2.8 and in Sect. 2.4.2.

As already mentioned for the *topo* analysis in Sect. 7.2.4, the  $W$ -plus-jets cross section as a function of jet multiplicity is not known precisely. Therefore Eq. 8.19 cannot be used to determine the expected number of tagged  $W$ -plus-jets events. The expected number of preselected  $W$ -plus-jets events, however, can be obtained from the matrix method which allows to evaluate the expected number of non-QCD events, referred to as  $N_t^{sig}$ , which comprises the  $t\bar{t}$  signal, the  $W$ -plus-jets background and the small backgrounds listed in Sect. 8.3.3. Thus, the  $W$ -plus-jets contribution can be

	=2 jets	=3 jets	$\geq 4$ jets
$W$ +light	$0.015\pm 0.001$	$0.010\pm 0.001$	$< 0.01$
$W(c\bar{c})$	$0.014\pm 0.001$	$< 0.01$	$< 0.01$
$W(b\bar{b})$	$0.11\pm 0.01$	$0.07\pm 0.01$	$0.03\pm 0.01$
$Wc$	$0.029\pm 0.002$	$0.015\pm 0.001$	$< 0.01$
$Wc\bar{c}$	$0.13\pm 0.01$	$0.060\pm 0.005$	$0.027\pm 0.004$
$Wb\bar{b}$	$1.93\pm 0.06$	$0.65\pm 0.04$	$0.28\pm 0.03$
$W$ +jets	$2.24\pm 0.06$	$0.82\pm 0.04$	$0.35\pm 0.03$
QCD	$< 0.01$	$< 0.01$	$< 0.01$
$t\bar{t} \rightarrow ll$	$0.57\pm 0.01$	$0.40\pm 0.01$	$0.14\pm 0.01$
$tb$	$0.305\pm 0.004$	$0.107\pm 0.003$	$0.024\pm 0.001$
$tqb$	$0.064\pm 0.003$	$0.144\pm 0.004$	$0.081\pm 0.003$
$WW \rightarrow l\nu jj$	$< 0.01$	$< 0.01$	$< 0.01$
$WZ \rightarrow l\nu jj$	$0.19\pm 0.01$	$0.035\pm 0.004$	$< 0.01$
$WZ \rightarrow jjll$	$< 0.01$	$< 0.01$	$< 0.01$
$ZZ \rightarrow jjll$	$< 0.01$	$< 0.01$	$< 0.01$
$Z \rightarrow \tau^+\tau^-$	$0.04\pm 0.03$	$0.02\pm 0.02$	$< 0.01$
background	$3.41\pm 0.07$	$1.53\pm 0.04$	$0.60\pm 0.04$
syst.	$+0.64$ $-0.59$	$+0.29$ $-0.28$	$+0.14$ $-0.12$
$t\bar{t} \rightarrow l$ +jets	$0.26\pm 0.02$	$1.87\pm 0.05$	$4.27\pm 0.07$
total	$3.67\pm 0.07$	$3.40\pm 0.07$	$4.87\pm 0.08$
syst.	$+0.69$ $-0.63$	$+0.65$ $-0.59$	$+1.11$ $-1.17$
# tags	4	5	3

Table 8.14: Summary of observed and predicted number of events with  $\geq 2$  tags in the  $e$ +jets channel. Errors are statistical only. Systematic uncertainties are given on the background and the total prediction.

extracted as:

$$N_{(W \rightarrow l\nu)+nj}^{presel} = N_t^{sig} - N_{t\bar{t} \rightarrow l+jets}^{presel} - N_{t\bar{t} \rightarrow ll}^{presel} - \sum_{bgr\ i} N_{bgr\ i}^{presel}, \quad (8.21)$$

where  $i$  loops over the mentioned small backgrounds. For the cross section determination (see Sect. 8.6) the  $t\bar{t}$  cross section is allowed to float both in Eq. 8.21 and Eq. 8.19. For the Tables 8.11- 8.14 a cross section of 7 pb is assumed to evaluate the expected number of  $t\bar{t}$  events.

In order to evaluate the number of tagged  $W$ -plus-jets events, it remains to determine the event tagging probability and employ Eq. 8.18. For this the MC information is used. It is crucial for the evaluation of the event tagging probability to properly describe the flavor composition of the  $W$ -plus-jets

	=1 jet	=2 jets	=3 jets	$\geq 4$ jets
$N_\ell$	7766	3506	1011	248
$N_t$	5134	2077	510	119
$N_t^{sig}$	$4933.3 \pm 73.9$	$1970.0 \pm 46.8$	$473.6 \pm 23.2$	$110.5 \pm 11.2$
$N_t^{QCD}$	$200.7 \pm 10.3$	$106.9 \pm 5.7$	$36.4 \pm 3.0$	$8.5 \pm 1.2$
$N_\ell^{tag}$	113	102	66	25
$N_t^{tag}$	55	55	39	13
$N_t^{tag,sig}$	$49.6 \pm 7.6$	$51.2 \pm 7.6$	$37.2 \pm 6.4$	$12.2 \pm 3.7$
$N_t^{tag,QCD}$	$5.4 \pm 0.9$	$3.8 \pm 0.7$	$1.8 \pm 0.5$	$0.8 \pm 0.3$
$\varepsilon_{sig}$	$0.866 \pm 0.008$	$0.850 \pm 0.007$	$0.835 \pm 0.011$	$0.819 \pm 0.018$
$\varepsilon_{QCD}$	$0.097 \pm 0.007$	$0.090 \pm 0.008$	$0.082 \pm 0.015$	$0.075 \pm 0.023$

Table 8.15: Number of loose and tight events in the preselected (first part) and the tagged, requiring exactly one tag, data sample (second part) together with the estimated number of non-QCD (*sig*) and QCD events. Also shown the efficiencies  $\varepsilon_{sig}$  and  $\varepsilon_{QCD}$  used in the matrix method.

events after applying the preselection. While the absolute  $W$ -plus-jets cross section as a function of jet multiplicity is not known precisely, the relative cross sections of  $W$ -plus-jets with different flavored jets in the final state for a given jet multiplicity is expected to be described well by the MC simulation. However, first measurements of the relative cross sections [155, 156] and NLO predictions [157] suggest that the LO heavy flavor cross sections have to be scaled up by a factor of approximately 1.5. The exclusive  $W$ -plus-jets MC samples which are considered are discussed in Sect. 2.4.2, and their cross sections are summarized in Table 2.9. The necessity for a parton-jet matching and its implementation is detailed in Sect. 2.4.2.1. The fraction  $F$  of each flavor configuration  $\Phi$  per jet multiplicity bin  $n$ , for events that pass the preselection, is obtained by:

$$F_{\Phi,n} = \frac{\sigma_{\Phi,n}^{eff}}{\sum_{\Phi} \sigma_{\Phi,n}^{eff}}, \quad (8.22)$$

where  $\sigma_{\Phi,n}^{eff} = \sigma_{\Phi,n} \varepsilon_{\Phi,n}^{presel,match}$  is the effective cross-section, and  $\varepsilon_{\Phi,n}^{presel,match}$  is the preselection and matching efficiency. According to the prescription discussed in Sect. 2.4.2.1 the flavor configurations shown in Table 8.18 are identified. One technical problem arises from the fact that no  $W+5$ jets MC samples are available in order to determine the effective cross-sections for the  $W(b\bar{b})jjj$  and the  $W(c\bar{c})jjj$  flavor configurations (see Table 8.18) from the migration from the five jet bin down to the fourth bin, where the heavy quark pair is reconstructed as one jet.

	=1 jet	=2 jets	=3 jets	$\geq 4$ jets
$N_\ell$	9010	4061	1278	397
$N_t$	6452	2387	595	176
$N_t^{sig}$	$6300.5 \pm 81.8$	$2268.3 \pm 49.7$	$534.8 \pm 25.0$	$158.9 \pm 13.5$
$N_t^{QCD}$	$151.5 \pm 8.1$	$118.7 \pm 4.6$	$60.2 \pm 3.1$	$17.1 \pm 1.5$
$N_\ell^{tag}$	92	118	59	44
$N_t^{tag}$	64	65	29	30
$N_t^{tag,sig}$	$62.2 \pm 8.1$	$61.2 \pm 8.2$	$26.1 \pm 5.5$	$29.2 \pm 5.6$
$N_t^{tag,QCD}$	$1.8 \pm 0.5$	$3.8 \pm 0.7$	$2.9 \pm 0.7$	$0.8 \pm 0.3$
$\varepsilon_{sig}$	$0.891 \pm 0.009$	$0.891 \pm 0.009$	$0.891 \pm 0.009$	$0.891 \pm 0.009$
$\varepsilon_{QCD}^{v8.2-v11}$	$0.078 \pm 0.001$	$0.074 \pm 0.002$	$0.080 \pm 0.004$	$0.074 \pm 0.010$
$\varepsilon_{QCD}^{v12}$	$0.127 \pm 0.005$	$0.120 \pm 0.006$	$0.139 \pm 0.016$	$0.108 \pm 0.034$
$\bar{P}_{QCD}^{tag}$	$1.1 \pm 0.2$	$3.3 \pm 0.5$	$4.2 \pm 0.8$	$5.4 \pm 1.6$
$N_t^{QCD} \bar{P}_{QCD}^{tag}$	$1.7 \pm 0.4$	$3.8 \pm 0.6$	$2.5 \pm 0.5$	$0.9 \pm 0.3$

Table 8.16: Number of loose and tight events in the preselected (first part) and the tagged, requiring exactly one tag, data sample (second part) together with the estimated number of non-QCD (*sig*) and QCD events. Also shown the efficiencies  $\varepsilon_{sig}$  and  $\varepsilon_{QCD}$  used in the matrix method ( $\varepsilon_{QCD}$  is found to be dependent on the trigger requirement, different in v8.2-v11 and in v12). The bottom part shows the event tagging probability measured in the “loose-tight” sample and the expected number of QCD-multijets events using the second method (see text).

The method to evaluate the relative flavor decomposition for the jet multiplicity bin with  $N_{jets} \geq 4$  is described in the following. The relative fractions of the available flavor configurations ( $Wjjjj$ ,  $Wcjjj$ ,  $Wc\bar{c}Jj$ ,  $Wb\bar{b}Jj$ ) is determined as described in Eq. 8.22. Their absolute normalization is determined from the  $Wjjjj$  fraction, by fitting the fraction of  $W$ +light jets in the first three jet multiplicity bins, using an  $a + b/\sqrt{N_{jets}}$  dependence, and extrapolating it to the  $N_{jets} \geq 4$  bin (see Figure 8.11). The ratio of fractions for  $W(c\bar{c})$  and  $W(b\bar{b})$ , found to be constant and  $W(c\bar{c})/W(b\bar{b}) \approx 3/2$ , is extrapolated to the  $N_{jets} \geq 4$  bin. This constraint and the fact that all fractions have to add up to unity allows to evaluate  $F_{W(c\bar{c})jjj}$  and  $F_{W(b\bar{b})jjj}$  by solving the following system of equations:

$$\begin{aligned} \frac{F_{W(c\bar{c})jjj}}{F_{W(b\bar{b})jjj}} &= \frac{3}{2}, \\ F_{Wjjjj} + F_{Wcjjj} + F_{Wc\bar{c}Jj} + F_{Wb\bar{b}Jj} + F_{W(c\bar{c})jjj} + F_{W(b\bar{b})jjj} &= 1. \end{aligned}$$

process	NLO $\sigma$ [pb]	$BR_{\mu+jets}$	$BR_{e+jets}$
$tb \rightarrow \ell\nu bb$	0.88	0.1253	0.1259
$tbq \rightarrow \ell\nu bbj$	1.98	0.1253	0.1259
$WW \rightarrow \ell\nu jj$	2.67	0.3912	0.3928
$WZ \rightarrow \ell\nu jj$	0.82	0.3912	0.3928
$WZ \rightarrow jj\ell\ell$	0.24	0.4390	0.4417
$ZZ \rightarrow jj\ell\ell$	0.20	0.4390	0.4417
$Z/\gamma^* \rightarrow \tau\tau (M_{\tau\tau} > 15 \text{ GeV})$	$432.6 \pm 19.1$	0.3171	0.3250

Table 8.17: NLO cross sections and branching ratios for the  $\mu$ +jets and the  $e$ +jets channels for small physics background processes, the  $Z/\gamma^* \rightarrow \tau\tau (M_{\tau\tau} > 15 \text{ GeV})$  cross section is measured by DØ [154].

$W+(=1jet)$	$W+(=2jets)$	$W+(=3jets)$	$W+(\geq 4jets)$
$Wj$	$Wjj$	$Wjjj$	$Wjjjj$
$W(c\bar{c})$	$W(c\bar{c})j$	$W(c\bar{c})jj$	$W(c\bar{c})jjj$
$W(b\bar{b})$	$W(b\bar{b})j$	$W(b\bar{b})jj$	$W(b\bar{b})jjj$
$Wc$	$Wcj$	$Wcjj$	$Wcjjj$
	$Wc\bar{c}$	$Wc\bar{c}J$	$Wc\bar{c}Jj$
	$Wb\bar{b}$	$Wb\bar{b}J$	$Wb\bar{b}Jj$

Table 8.18: Classification of  $W$ -plus-jets flavor configurations per jet multiplicity bin;  $j$  is any of  $u, d, s, g$  and  $J$  is any of  $u, d, s, g, c$  partons,  $(b\bar{b})$  and  $(c\bar{c})$  denote heavy quark pairs reconstructed as one jet.

The summary of the flavor compositions of the  $W$ +jets background is shown in Table 8.19 and Table 8.20 for the  $\mu$ +jets and the  $e$ +jets channel, respectively, where only statistical uncertainties due to finite Monte Carlo statistics are quoted.

### 8.3.4.1 Systematic Uncertainties

The systematic uncertainties, associated with the evaluation of the  $W$ -plus-jets background, are the following:

**Uncertainty due to  $W$ -plus-jets flavor fractions:** Systematic uncertainties arise both from the limited Monte Carlo statistics in the determination of  $\epsilon_{\Phi, n}^{presel, match}$  for the effective cross sections, and from the extrapolation procedure to derive the flavor fractions in the jet multiplicity bin with  $\geq 4$  jets. The latter comprises two uncertainties. The uncertainty on the fraction of  $Wjjjj$  in the  $N_{jets} \geq 4$  bin is set to the difference between the fit evaluated for  $N_{jets} = 3$  and evaluated for

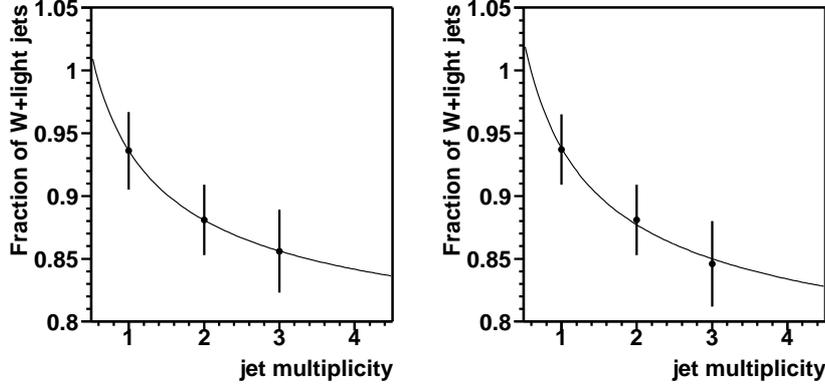


Figure 8.11: Fit to the  $W$ +light fractions in the  $\mu$ +jets (left) and  $e$ +jets (right) channels.

Contribution	$W$ +(=1jet)	$W$ +(=2jets)	$W$ +(=3jets)	$W$ +( $\geq 4$ jets)
$W$ +light	$93.6 \pm 3.1$	$88.1 \pm 2.8$	$85.6 \pm 3.3$	$84.1 \pm 2.3$
$W(c\bar{c})$	$1.03 \pm 0.05$	$1.91 \pm 0.07$	$2.84 \pm 0.18$	$2.8 \pm 1.1$
$W(b\bar{b})$	$0.65 \pm 0.03$	$1.19 \pm 0.04$	$1.77 \pm 0.07$	$1.87 \pm 0.74$
$Wc$	$4.71 \pm 0.19$	$7.01 \pm 0.27$	$6.32 \pm 0.31$	$5.31 \pm 0.33$
$Wc\bar{c}$		$1.00 \pm 0.08$	$1.97 \pm 0.14$	$3.33 \pm 0.43$
$Wb\bar{b}$		$0.81 \pm 0.04$	$1.53 \pm 0.09$	$2.56 \pm 0.17$

Table 8.19: Fraction of the different  $W$ +jets flavor subprocesses (in %) contributing to each exclusive jet multiplicity bin in the  $\mu$ +jets channel. Parton-jet matching and preselection are required. The quoted uncertainty results only from limited Monte Carlo statistics.

$N_{\text{jets}} = 4$ . The default value of  $W(c\bar{c})jj/W(b\bar{b})jj \approx 3/2$  is replaced by the actual result of the constant fit to the ratio  $W(c\bar{c})/W(b\bar{b})$  in the first three jet multiplicity bins, found to be  $1.56 \pm 0.04$ , to quote a systematic uncertainty.

**Uncertainty due to  $Wb\bar{b}$ ,  $W(b\bar{b})$ ,  $Wc\bar{c}$ ,  $W(c\bar{c})$  cross sections:** The samples with two heavy flavor quarks in the final state are subdivided in those where both heavy flavor quarks are reconstructed separately as jets ( $Wb\bar{b}$  and  $Wc\bar{c}$ ) and those where both heavy flavor quarks are reconstructed in one jet ( $W(b\bar{b})$  and  $W(c\bar{c})$ ). The relative cross sections are determined from the LO matrix element generator ALPGEN. However, first measurements of the relative cross sections [155, 156] and NLO predictions [157] suggest that the LO heavy flavor cross

Contribution	$W+(=1\text{jet})$	$W+(=2\text{jets})$	$W+(=3\text{jets})$	$W+(\geq 4\text{jets})$
$W+\text{light}$	$93.7 \pm 2.8$	$88.1 \pm 2.8$	$84.6 \pm 3.4$	$83.4 \pm 2.5$
$W(c\bar{c})$	$1.10 \pm 0.05$	$1.91 \pm 0.07$	$2.71 \pm 0.18$	$3.4 \pm 1.2$
$W(b\bar{b})$	$0.69 \pm 0.02$	$1.23 \pm 0.04$	$2.00 \pm 0.08$	$2.27 \pm 0.82$
$Wc$	$4.50 \pm 0.17$	$6.80 \pm 0.27$	$7.21 \pm 0.36$	$5.30 \pm 0.35$
$Wc\bar{c}$		$1.11 \pm 0.08$	$2.11 \pm 0.16$	$3.43 \pm 0.46$
$Wb\bar{b}$		$0.87 \pm 0.05$	$1.34 \pm 0.09$	$2.20 \pm 0.17$

Table 8.20: Fraction of the different  $W$ +jets flavor subprocesses (in %) contributing to each exclusive jet multiplicity bin in the  $e$ +jets channel. Parton-jet matching and preselection are required. The quoted uncertainty results only from limited Monte Carlo statistics.

sections have to be scaled up by a factor of approximately 1.5. Therefore the uncertainty is obtained by varying the  $Wb\bar{b}$ ,  $W(b\bar{b})$ ,  $Wc\bar{c}$  and  $W(c\bar{c})$  cross sections in ALPGEN by  $\pm 50\%$ . Variations are fully correlated between  $Wb\bar{b}$  and  $W(b\bar{b})$ , and between  $Wc\bar{c}$  and  $W(c\bar{c})$ , while they are uncorrelated between  $Wb\bar{b}$  and  $Wc\bar{c}$ .

**Uncertainty due to the choice of the generation parameters in the  $W$ -plus-jets Monte Carlo:** As mentioned in Sect. 8.3.4, the underlying event is not properly described in the MC samples used due to a small bug. The possible impact on this analysis is estimated as follows: A  $Wjjjj$  MC sample with Tune A, that properly describes the underlying event, is produced and a systematic uncertainty on the  $Wjjjj$  event tagging probability is assigned as the difference between the sample with and without Tune A, found to be very small, 2.8 %. In addition, the  $W$ -plus-jets MC samples used in the *topo* analysis with the generation parameters described in Table 2.8 are used to assign a systematic uncertainty on the event tagging probabilities (the default and the *topo* MC samples differ by the PDF, the factorization scale, the usage of Tune A and by the parton generation cuts). A list of the available samples, the event tagging probabilities and the observed differences are shown in Table 8.21. Due to the limited availability of MC samples the differences are only evaluated for the electron-plus-jets channel and are assumed to be the same for the  $\mu$ -plus-jets channel. The effect of the choice of the generation parameters is found to be small for the heavy flavor components but non-negligible for the light flavor components as shown in Table 8.21. The observed systematic shifts in the jet multiplicity bins  $N_{\text{jets}} = 2$  and  $N_{\text{jets}} \geq 4$  are extrapolated and interpolated to the  $N_{\text{jets}} = 1$  and  $N_{\text{jets}} = 3$  bins, respectively, assuming a linear dependence, leading to the assigned

systematic uncertainties of 6, 12, 18 and 24 % for  $N_{\text{jets}} = 1, 2, 3, \geq 4$ .

As mentioned in Sect. 2.4, the usage of CTEQ 6.1M [41], derived in NLO, is not adequate in conjunction with leading order matrix element generators. However, a comparison of parton-level distributions for the  $Wjjjj$  sample for the leading parton, the fourth leading parton and the scalar sum of the  $p_T$  of the four leading partons shows no measurable dependence on the choice of the parton distribution function (CTEQ 5L [43] or CTEQ 6.1M [41]).

	$\bar{P}_{topo}$	$\bar{P}_{btag}$	$\frac{\bar{P}_{topo} - \bar{P}_{btag}}{\bar{P}_{btag}}$
$Wjj$	0.47	0.53	-12.0 %
$Wcj$	7.6	7.8	-2.7 %
$Wc\bar{c}$	12.4	13.4	-7.5 %
$Wb\bar{b}$	41.8	42.2	-0.9 %
$Wjjjj$	0.86	1.14	-24.5 %

Table 8.21: Event tagging probabilities (in %) in the  $e$ +jets channel for samples with the *topo* and the *btag* generation parameters, and the relative difference, see Table 2.8 for details.

## 8.4 Cross-Section Extraction Procedure

The lepton+jets  $t\bar{t}$  cross sections are extracted using a binned maximum likelihood fit, with four bins per channel ( $\mu$ +jets or  $e$ +jets), considering separately events with 3 or  $\geq 4$  jets and with 1 or  $\geq 2$  tags, i.e. in total eight bins for the combination of both channels:

$$L(\sigma_{t\bar{t}}) = \prod_{i=1}^8 P(n_i^{obs}, \mu_i(\sigma_{t\bar{t}})) \quad (8.23)$$

$$= \prod_{\mu+\text{jets}, e+\text{jets}} \prod_{3, \geq 4 \text{ jets}} \prod_{1, \geq 2 \text{ tags}} P(n_i^{obs}, \mu_i(\sigma_{t\bar{t}})). \quad (8.24)$$

where  $P(n, \mu) = \frac{\mu^n e^{-\mu}}{n!}$  generically denotes the Poisson probability density function for  $n$  observed events, given an expectation of  $\mu$ .

The expectation  $\mu(\sigma_{t\bar{t}})$  per bin is given by:

$$\begin{aligned}
\mu(\sigma_{t\bar{t}}) &= N_{t\bar{t}}^{tag}(\sigma_{t\bar{t}}) + \left[ N_{(W \rightarrow l\nu)+nj}^{tag}(\sigma_{t\bar{t}}) \right] + N_{QCD}^{tag} + \sum_{bgr\ i} N_{bgr\ i}^{tag} \quad (8.25) \\
&= N_{t\bar{t}}^{presel}(\sigma_{t\bar{t}}) \bar{P}_{t\bar{t}}^{tag} + \left[ N_{(W \rightarrow l\nu)+nj}^{presel}(\sigma_{t\bar{t}}) \bar{P}_{(W \rightarrow l\nu)+nj}^{tag} \right] + \\
&\quad N_{QCD}^{tag} + \sum_{bgr\ i} N_{bgr\ i}^{presel} \bar{P}_{bgr\ i}^{tag} \\
&= \sigma_{t\bar{t}} \varepsilon_{t\bar{t}} BR_{t\bar{t}} L \bar{P}_{t\bar{t}}^{tag} \\
&\quad + \left[ \left( N_t^{sig} - N_{t\bar{t}}^{presel}(\sigma_{t\bar{t}}) - \sum_{bgr\ i} N_{bgr\ i}^{presel} \right) \bar{P}_{(W \rightarrow l\nu)+nj}^{tag} \right] \\
&\quad + N_{QCD}^{tag} + \sum_{bgr\ i} \sigma_{bgr\ i} \varepsilon_{bgr\ i} BR_{bgr\ i} L \bar{P}_{bgr\ i}^{tag} \\
&= \sigma_{t\bar{t}} \varepsilon_{t\bar{t}} BR_{t\bar{t}} L (\bar{P}_{t\bar{t}}^{tag} - \bar{P}_{(W \rightarrow l\nu)+nj}^{tag}) + N_t^{sig} + N_{QCD}^{tag} \\
&\quad + \sum_{bgr\ i} \sigma_{bgr\ i} \varepsilon_{bgr\ i} BR_{bgr\ i} L (\bar{P}_{bgr\ i}^{tag} - \bar{P}_{(W \rightarrow l\nu)+nj}^{tag}),
\end{aligned}$$

where  $N_{(W \rightarrow l\nu)+nj}^{presel}$  is expressed as given in Eq. 8.21, and  $i$  in the sum loops over all the small backgrounds listed in Sect. 8.3.3.

## 8.5 Systematic Uncertainties

The systematic uncertainties on the cross section are obtained for each independent source of uncertainty by varying the source by one standard deviation up and down, and propagating the variation into both background estimates and signal efficiencies. A new likelihood function and a new cross section is derived for each such variation. These variations in the central value of the cross section are then summed quadratically to obtain the total systematic uncertainty.

The summary of the systematic uncertainties on the  $t\bar{t}$  cross sections ( $\mu$ +jets,  $e$ +jets and combined) is given in Table 8.22, and the correlations among the two analysis channels are indicated. The uncorrelated systematic uncertainties are usually of statistical origin in either Monte Carlo or data. All systematic uncertainties, besides the ones which originate from the limited MC statistics, are correlated among the two jet multiplicity bins ( $N_{jet} = 3$  and  $N_{jet} \geq 4$ ). All systematic uncertainties are fully correlated between the single tagged and the double tagged subsamples.

source	$\mu$ +jets	$e$ +jets	combined
Muon preselections	+0.39 -0.35		+0.22 -0.21
Electron preselections		+0.22 -0.21	+0.10 -0.10
Muon triggers	+0.39 -0.26		+0.21 -0.15
Electron triggers		+0.29 -0.22	+0.13 -0.10
QCD tagging probability	+0.00 -0.00	+0.23 -0.23	+0.13 -0.13
Matrix method $\epsilon_{QCD}$	+0.05 -0.05	+0.01 -0.01	+0.02 -0.02
Matrix method $\epsilon_{sig}$	+0.01 -0.01	+0.00 -0.00	+0.01 -0.01
Event statistics for matrix method	+0.24 -0.25	+0.19 -0.19	+0.16 -0.15
MC statistics	+0.10 -0.11	+0.12 -0.12	+0.08 -0.08
Jet triggers	+0.01 -0.08	+0.02 -0.01	+0.01 -0.05
Jet energy scale	+1.00 -0.81	+1.07 -0.91	+1.02 -0.84
Jet energy resolution	+0.08 -0.04	+0.01 -0.07	+0.07 -0.02
Jet reco and jet ID	+0.73 -0.00	+0.66 -0.00	+0.68 -0.00
Top mass	+0.25 -0.24	+0.23 -0.22	+0.24 -0.23
Taggability in data	+0.01 -0.01	+0.03 -0.03	+0.02 -0.02
Flavor dependence of taggability	+0.01 -0.01	+0.01 -0.01	+0.01 -0.01
Inclusive $b$ -tag efficiency in MC	+0.08 -0.07	+0.09 -0.09	+0.08 -0.08
Inclusive $c$ -tag efficiency in MC	+0.12 -0.12	+0.11 -0.11	+0.12 -0.12
Semileptonic $b$ -tag efficiency in MC	+0.38 -0.41	+0.43 -0.46	+0.41 -0.44
Semileptonic $b$ -tag efficiency in data	+0.84 -0.72	+0.97 -0.84	+0.91 -0.78
Negative tagging efficiency	+0.12 -0.12	+0.11 -0.11	+0.12 -0.12
$SF_{hf}$ and $SF_{ll}$	+0.05 -0.05	+0.04 -0.04	+0.05 -0.05
$b$ -decay model dependence	+0.32 -0.30	+0.36 -0.34	+0.34 -0.32
$W$ flavor fractions	+0.10 -0.10	+0.11 -0.11	+0.08 -0.08
$Wb\bar{b}$ , $W(b\bar{b})$ , $Wc\bar{c}$ , $W(c\bar{c})$ cross sections	+0.87 -0.97	+0.76 -0.81	+0.81 -0.89
Generation parameters	+0.28 -0.00	+0.23 -0.00	+0.27 -0.00
total	+1.93 -1.68	+1.98 -1.68	+1.89 -1.63

Table 8.22: Summary of the absolute systematic uncertainties in units of pb on the  $\mu$ +jets, the  $e$ +jets and the combined lepton+jets  $t\bar{t}$  production cross section. The uncertainties in the upper part, up to and including “MC statistics”, are uncorrelated between the analysis channels, the uncertainties in the lower part are correlated between the analysis channels.

### 8.5.1 Non-*b*-Tag Related Sources of Systematic Uncertainties

The sources of systematic uncertainties associated with the preselection are described in Sect. 7.1 and Sect. 7.5 for the  $\mu$ +jets channel and are discussed in [152, 13] for the  $e$ +jets channel. These systematic uncertainties correspond to the lepton, jet, matrix method and MC statistics uncertainties in Table 8.22. Minor differences exist with respect to the *topo* analysis which are discussed in the following.

The calculation of the  $W$ -plus-jets and QCD-multijets backgrounds relies on the matrix method. To take into account the Poisson uncertainties on the number of events in the loose and tight samples while accounting for their correlation, the sample is subdivided into four disjoint parts per jet multiplicity bin:

- “loose-tight”, loose and not tight,
- tight and non-tagged,
- tight with 1 tag,
- tight with  $\geq 2$  tag.

The number of events in the first two classes are varied according to the Poisson distributions. The resulting error on the cross section due to statistical fluctuations of the number of events entering the matrix method is classified as a systematic uncertainty.

The MC-to-data correction factor that accounts for the differences in the reconstruction and identification efficiency for jets as discussed in Sect. 7.1.1.1 was not ready in time when this analysis was performed, as a consequence it is not applied for the central value. However, the application of the MC-to-data correction factor is quoted as a systematic uncertainty which only has one sign (one-sided systematic uncertainty).

The uncertainty on the top quark mass is treated as a systematic uncertainty. The central value is quoted for a top quark mass of 175 GeV. Parameterizations of the preselection efficiencies and the event tagging probabilities as a function of the top mass are derived, similarly to what is shown in Fig. 7.31, evaluated at 170 GeV and 180 GeV and propagated to the cross section giving an estimate of the uncertainty.

### 8.5.2 *b*-Tag Related Sources of Systematic Uncertainties

The *b*-tag related systematic uncertainties on the  $t\bar{t}$  cross-section originate mainly from the systematic uncertainties on the tagging efficiencies both in MC and data, as well as on the scale factors and the systematic uncertainties due to the heavy flavor decay model (evaluated by exchanging

EVTGEN by PYTHIA and rederiving the parameterization for all tagging efficiencies), as discussed in Sect. 8.2. Systematic uncertainties due to limited MC statistics to derive the tagging related parameterizations are also considered. Additional uncertainties are associated to the  $W$ -plus-jets background determination as detailed in Sect. 8.3.4.1.

All tagging related systematic uncertainties, but the ones with statistical origin, are treated as correlated between the two channels ( $\mu$ +jets and  $e$ +jets), and between the two jet multiplicity bins ( $N_{\text{jets}} = 3$  and  $N_{\text{jets}} \geq 4$ ), and are summarized in Table 8.22.

## 8.6 The $t\bar{t}$ Production Cross-Section

The expected number of signal and background events (including their statistical and systematic uncertainties), assuming a  $t\bar{t}$  cross section of 7 pb, and the number of observed tagged events are summarized in Table 8.11 and Table 8.13 for the  $\mu$ -plus-jets channel, and in Table 8.12 and Table 8.14 for the electron-plus-jets channel. The actual number of tagged events is in good agreement with the prediction, both in the control bins ( $N_{\text{jet}} = 1, 2$ ) showing that the background is accurately predicted, and in the signal bins ( $N_{\text{jet}} = 3, \geq 4$ ), used to extract the  $t\bar{t}$  cross section. Figure 8.12 visualizes the agreement between prediction and observation for the combined lepton-plus-jets channel. In Fig. B.1- B.8 in Appendix B it is shown that

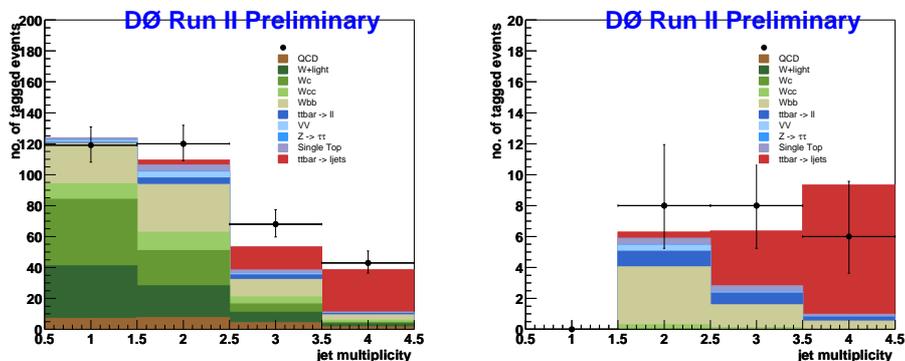


Figure 8.12: Summary plot of the observed and predicted number of tagged events in the combined lepton-plus-jets channel. Left: events with 1 tag. Right: events with  $\geq 2$  tags.

also typical kinematic and topological distributions of the tagged events are reproduced by the prediction of the signal plus the backgrounds.

The result of the  $t\bar{t}$  production cross-section measurement in the  $e$ +jets and  $\mu$ +jets channels and the combination of the two analysis channels for

the SVT lifetime  $b$ -tagging algorithm at  $\sqrt{s} = 1.96$  TeV yields:

$$\begin{aligned} e + \text{jets} & : \quad \sigma_{p\bar{p} \rightarrow t\bar{t}+X} = 9.41_{-1.76}^{+1.93} \text{ (stat)} \quad {}_{-1.68}^{+1.98} \text{ (syst)} \pm 0.61 \text{ (lumi) pb} \\ \mu + \text{jets} & : \quad \sigma_{p\bar{p} \rightarrow t\bar{t}+X} = 6.86_{-1.70}^{+1.90} \text{ (stat)} \quad {}_{-1.68}^{+1.93} \text{ (syst)} \pm 0.45 \text{ (lumi) pb} \\ \text{lepton} + \text{jets} & : \quad \sigma_{p\bar{p} \rightarrow t\bar{t}+X} = 8.24_{-1.25}^{+1.34} \text{ (stat)} \quad {}_{-1.63}^{+1.89} \text{ (syst)} \pm 0.54 \text{ (lumi) pb.} \end{aligned}$$

The results are shown in Fig. 8.13.

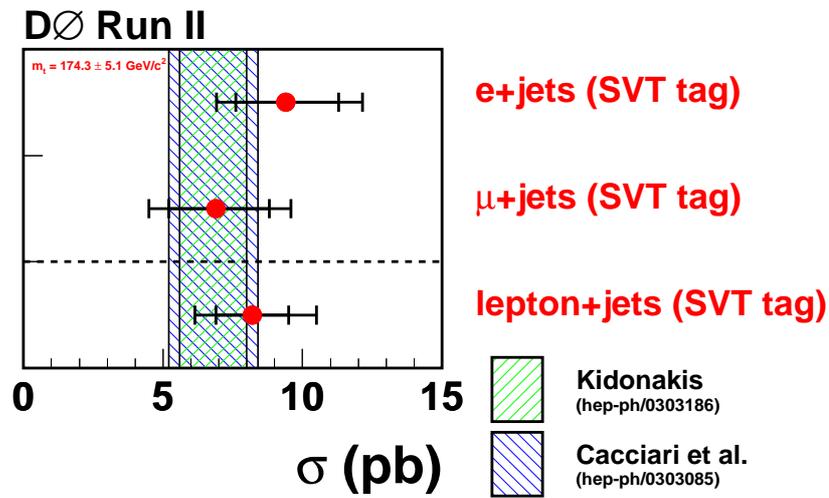


Figure 8.13: A comparison of the  $t\bar{t}$  production cross-section measurements in the  $\mu$ -plus-jets, electron-plus-jets and the combined lepton-plus-jets channel; also shown the theoretical expectation [151].

## Chapter 9

# Summary and Conclusion

Two measurements of the  $t\bar{t}$  production cross section in  $p\bar{p}$  collisions at a center-of-mass energy of 1.96 TeV are presented in events with a charged lepton, a neutrino and  $\geq 4$  jets. After a common preselection of the objects in the final state, two different features of the  $t\bar{t}$  events are exploited in order to extract the  $t\bar{t}$  cross section. The two analyses are complementary and very competitive with each other.

In the *topo* analysis, a likelihood discriminant is built from kinematical variables that describe the topology of an event and have a maximum separation power between the  $t\bar{t}$  signal and the backgrounds along with the property of being preferably insensitive to systematic variations. The  $t\bar{t}$  production cross section in the muon+jets channel is measured to be:

$$\sigma_{p\bar{p} \rightarrow t\bar{t}+X} = 5.13_{-1.57}^{+1.76} \text{ (stat)} \quad {}_{-1.10}^{+0.96} \text{ (syst)} \pm 0.33 \text{ (lumi) pb},$$

using 229.1 pb<sup>-1</sup>. The  $t\bar{t}$  production cross section in the combined lepton+jets channel is measured to be [151]:

$$\sigma_{p\bar{p} \rightarrow t\bar{t}+X} = 6.60_{-1.28}^{+1.37} \text{ (stat)} \quad {}_{-1.11}^{+1.25} \text{ (syst)} \pm 0.43 \text{ (lumi) pb},$$

using 229.1 pb<sup>-1</sup> for the  $\mu$ -plus-jets and 226.3 pb<sup>-1</sup> for the electron-plus-jets channel. In Fig. 9.1 the topological  $t\bar{t}$  production cross-section measurements are shown as a function of the center-of-mass energy and compared with the Run I result.

The *btag* analysis is based on the application of a lifetime *b*-tagging algorithm which explicitly reconstructs secondary vertices, removing approximately 95 % of the background while keeping 60 % of the  $t\bar{t}$  signal. The measurement combines the muon+jets and the electron+jets channel, using 158.4 pb<sup>-1</sup>, and 168.8 pb<sup>-1</sup>, respectively:

$$\sigma_{p\bar{p} \rightarrow t\bar{t}+X} = 8.24_{-1.25}^{+1.34} \text{ (stat)} \quad {}_{-1.63}^{+1.89} \text{ (syst)} \pm 0.54 \text{ (lumi) pb}.$$

In Fig. 9.2 the  $t\bar{t}$  production cross-section measurements using lifetime *b*-tagging are shown as a function of the center-of-mass energy and compared with the Run I result.

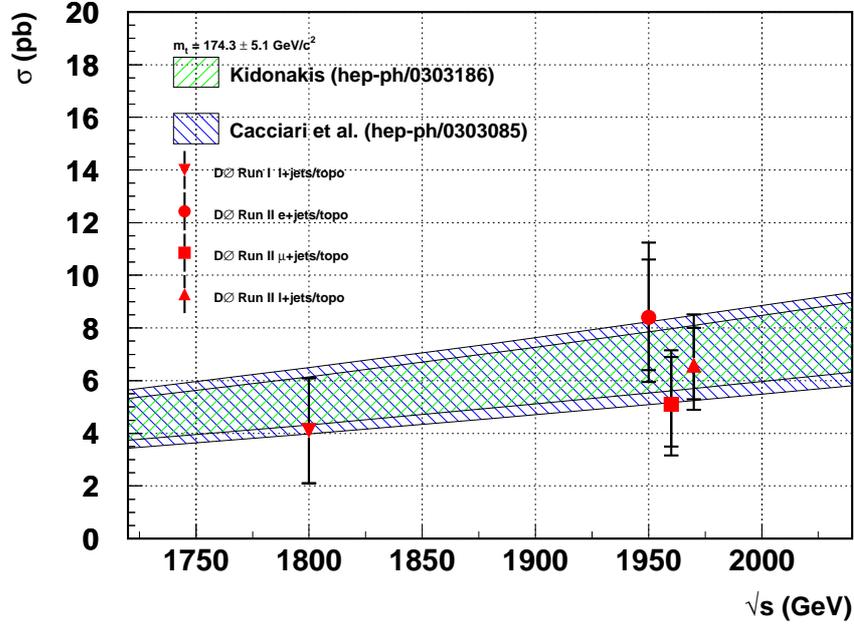


Figure 9.1: Topological  $t\bar{t}$  production cross-section measurements in the  $\mu$ -plus-jets, electron-plus-jets and the combined lepton-plus-jets channel as a function of the center-of-mass energy, compared with the Run I result and the theoretical expectation [151].

All current  $t\bar{t}$  production cross-section measurements from  $D\bar{O}$  are summarized in Fig. 9.3, showing that the measurements in the lepton-plus-jets channels, presented in this thesis, have the highest sensitivity. A comparison of all current  $t\bar{t}$  production cross-section measurements from  $D\bar{O}$  and CDF is shown in Fig. 9.4. The results from both the various analysis channels, the various analysis techniques and the two experiments, CDF and  $D\bar{O}$ , are in good agreement with each other and with the theoretical expectation. There are strong correlations among the channels for CDF and among the various analysis techniques for both experiment which make a combination of the measurements very difficult. The precision of the measurements is currently not high enough to make QCD precision tests, however, with increasing luminosity both the combination and the QCD tests will be done.

With increasing statistics, both analyses will be soon solely limited by the systematic uncertainty, however, most of the systematic uncertainties will be reduced with increasing statistics. The dominant source of systematic uncertainty for both analyses is the uncertainty on the jet energy scale. The

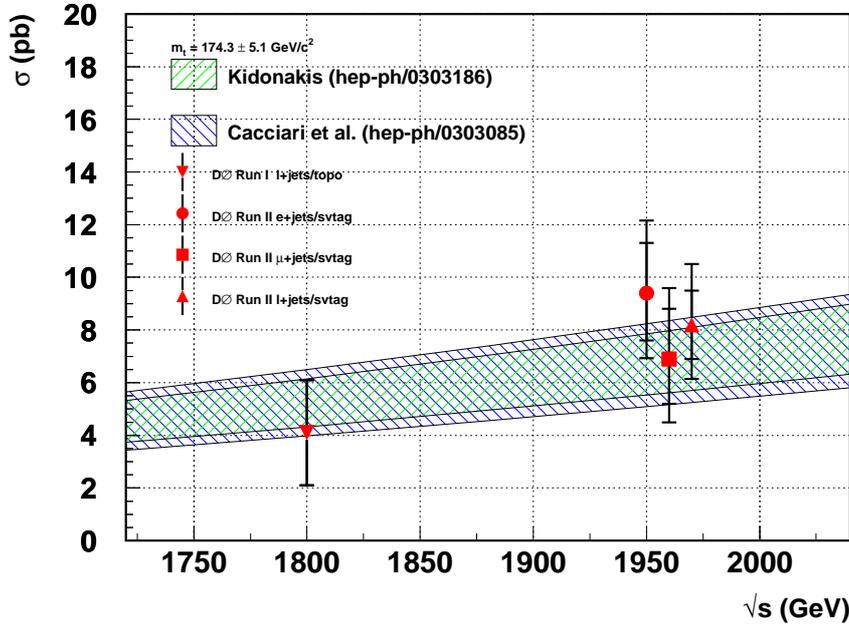


Figure 9.2:  $t\bar{t}$  production cross-section measurements using lifetime  $b$ -tagging in the  $\mu$ -plus-jets, electron-plus-jets and the combined lepton-plus-jets channel as a function of the center-of-mass energy, compared with the Run I result and the theoretical expectation [151].

statistical component will be reduced by collecting more data, the systematic component is at present slightly overestimated, since the systematic effects in data and in MC are treated as uncorrelated. However, many of the systematics are correlated between data and MC and do not affect the cross section measurement. The evaluation of the correlations will decrease the jet energy scale uncertainty dramatically.

The other dominant systematic uncertainty on the cross section in the *topo* analysis originates from the uncertainty on the modeling of the  $W$ -plus-jets background. With increasing data statistics the modeling can be verified in reference samples, vetoing the contribution from  $t\bar{t}$  events by rejecting events with  $b$ -tags or by selecting events with a lower jet multiplicity.

The two other dominant systematic uncertainties on the cross section in the *btag* analysis originate from the uncertainty on the cross sections of the electroweak  $W$  boson production in association with heavy flavor quarks. A discriminant variable, the invariant mass of tracks with a large  $dca$  significance, is defined to disentangle  $b$ ,  $c$  and light jets, which allows to measure the heavy flavor fractions in events with two jets in the final state. This mea-

surement [155, 156] suggests that the LO  $Wb\bar{b}$  and  $Wc\bar{c}$  cross sections have to be scaled up by a factor of approximately 1.5, which is in agreement with NLO predictions, evaluated using the program MCFM [157]. The systematic uncertainties on the  $Wb\bar{b}$  and  $Wc\bar{c}$  cross sections are assigned to cover the factor 1.5. The other limiting systematic uncertainty originates from the measurement of the  $b$ -tagging efficiency in data, which is statistically limited and will decrease with increasing statistics.

Both measurements are in good agreement with the *Standard Model* expectation of approximately 7 pb, and both analyses are in the process of being published in refereed journals (Phys. Rev. Lett.).

As an outlook, the top quark production studies at the Tevatron lay the ground for the future Large Hadron Collider physics project which will be a top factory. There the  $t\bar{t}$  production will be the process of choice to calibrate and understand the detectors, i.e. for the measurement of the jet energy scale, the study of the lepton identification and the study of the  $\cancel{E}_T$ . Moreover, the top quark pair production will be the dominant background for *New Physics* searches beyond the *Standard Model*, and the Higgs boson production.

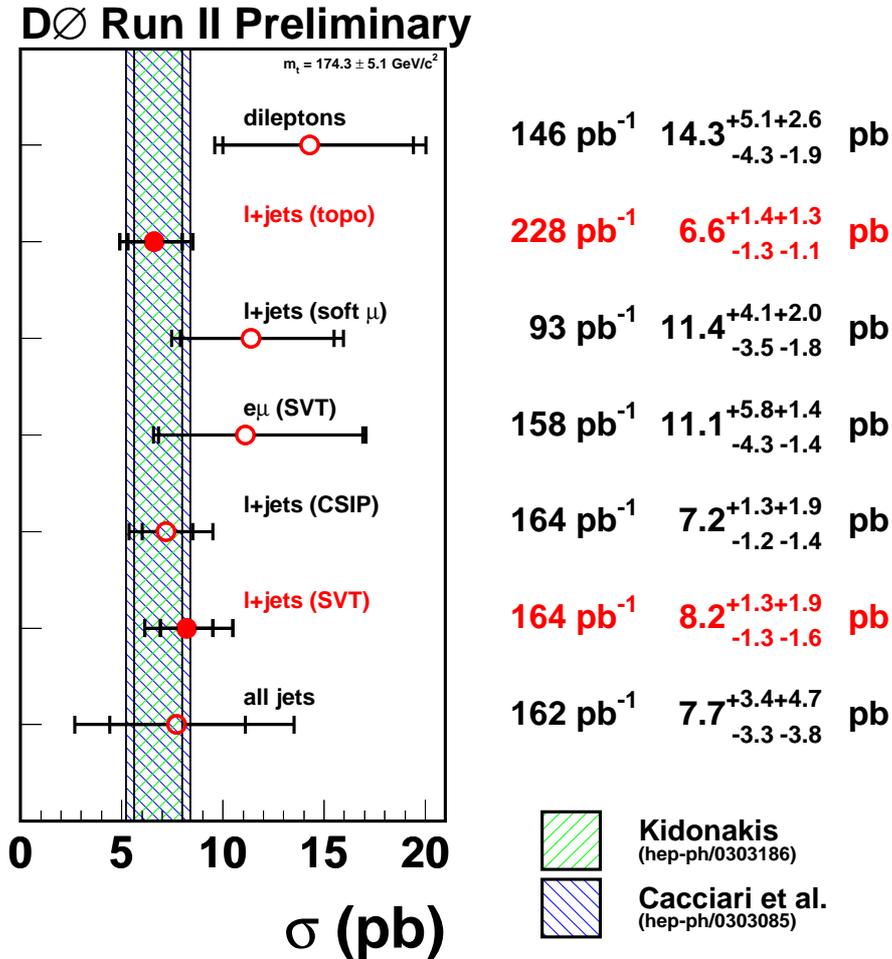


Figure 9.3: Comparison of all current  $t\bar{t}$  production cross-section measurements from DØ in good agreement with the theoretical expectation [151]. The results of the analyses presented are highlighted and are given by the solid points. The integrated luminosities used and the corresponding  $t\bar{t}$  cross-section measurements are quoted. From top to bottom: combined dilepton ( $ee$ ,  $\mu\mu$  and  $e\mu$ ) [158], lepton-plus-jets *topo* as presented in this thesis, lepton-plus-jets using soft muon  $b$ -tagging [159],  $e\mu$  using the SVT lifetime  $b$ -tagging [160], lepton-plus-jets using the CSIP lifetime  $b$ -tagging [13], lepton-plus-jets using the SVT lifetime  $b$ -tagging as presented in this thesis, and all-jets channel [67].

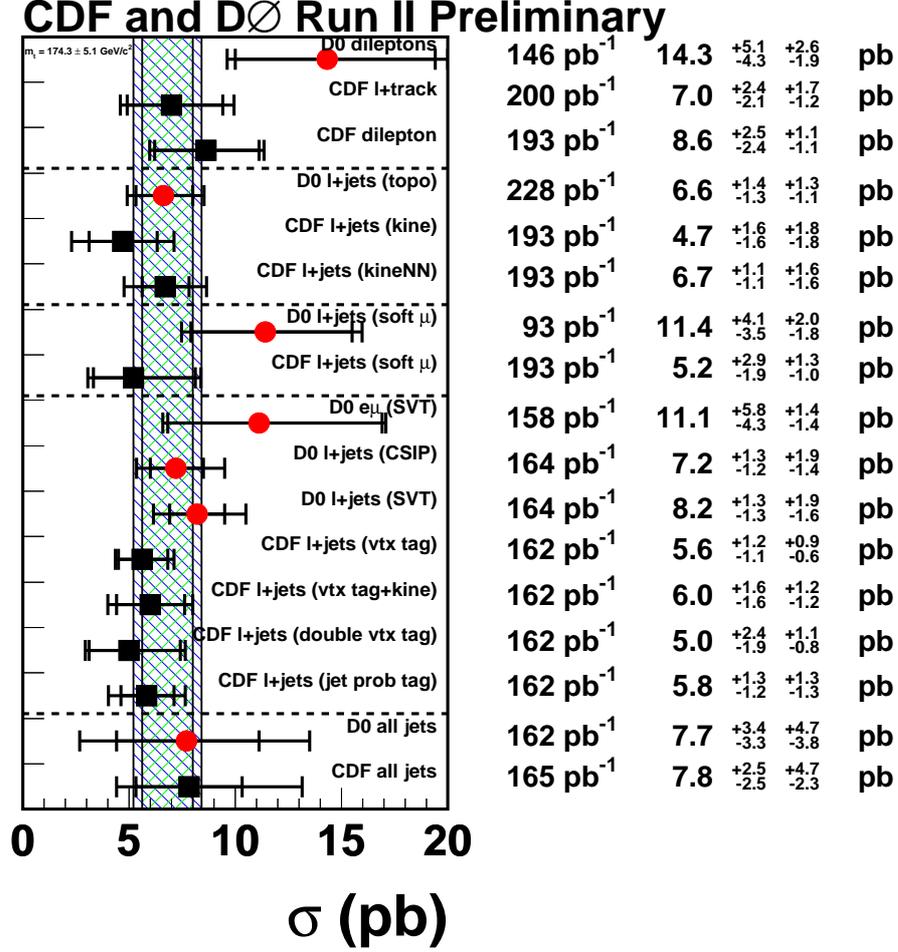


Figure 9.4: Comparison of all current  $t\bar{t}$  production cross-section measurements from DØ and CDF in good agreement with the theoretical expectation [151]. The integrated luminosities used and the corresponding  $t\bar{t}$  cross-section measurements are quoted. For the DØ references see caption of Fig. 9.3. CDF from top to bottom: lepton+track [161], dilepton [162], lepton-plus-jets using  $H_T$  and a neural network output as discriminants [163], lepton-plus-jets using soft muon  $b$ -tagging [164], lepton-plus-jets using SVT lifetime  $b$ -tagging: all tags [165], kinematic fitting of  $b$ -tagged lepton-plus-jets events [166], double tags [165], lepton-plus-jets using jet probability lifetime  $b$ -tagging [167], and all-jets channel [168].

## Appendix A

# Sanity Checks for the Topological Analysis

As a sanity check in the  $\mu$ -plus-jets channel, three additional control samples are defined which are depleted in  $t\bar{t}$  and serve to verify that the  $W$ -plus-jets background is well modeled by the simulation:

**Zero-tag:** Apply the full preselection and require in addition that none of the jets is identified as a  $b$  jet, neither by the SVT lifetime tagging algorithm nor by the soft lepton tagger.  $t\bar{t}$  events contain two  $b$  jets in the event whereas the  $W$ -plus-jets and the QCD-multijets backgrounds mainly consist of light flavor jets. Thus, the  $b$  jet veto leads to a depletion of  $t\bar{t}$  events.

**3-jet-bin:** Apply the full preselection but require that exactly three jets are reconstructed instead of  $\geq$  four.

**2-jet-bin:** Apply the full preselection but require that exactly two jets are reconstructed instead of  $\geq$  four.

Sanity plots for these three samples are shown in the following. Additionally, for the signal selection the default  $W$  MC sample ( $Q^2 = M_W^2 + \sum p_{T_j}^2$ ) is exchanged by the one with  $Q^2 = \langle p_{T_j} \rangle^2$ , and all sanity plots are also shown with the result from the likelihood fit in Fig. A.2- A.6 (called *W MC modeling*). For the *Zero-tag* selection the data is overlaid with the result from the likelihood fit of  $t\bar{t}$ ,  $W$ -plus-jets and QCD-multijets. For the *3-jet-bin* and the *2-jet-bin* samples the expected number of QCD-multijets ( $N_t^{QCD}$ ) and the non-QCD ( $N_t^{sig}$ ) is taken from the matrix method (see Sect. 7.2.3 and Table. 7.15). The number of  $t\bar{t}$  events is shown for an expected cross section of 7 pb. In Fig. A.1 the likelihood discriminant distributions are shown. Figures A.7- A.11 show the sanity plots for the *Zero-tag* selection. Figures A.12- A.15 show the sanity plots for the *3-jet-bin* selection and, Fig. A.16- A.19 show the sanity plots for the *2-jet-bin* selection. The Kolmogorov-Smirnov

matching probabilities (KS) for the agreement between the observation in data and the prediction, constrained between 0 and 1, are given in the upper left corner of each plot.

Data to MC comparison plots including the  $\pm 1\sigma$  uncertainty band for the JES and the JID systematic uncertainty are shown for all sanity plots in Fig. A.20-Fig. A.29.

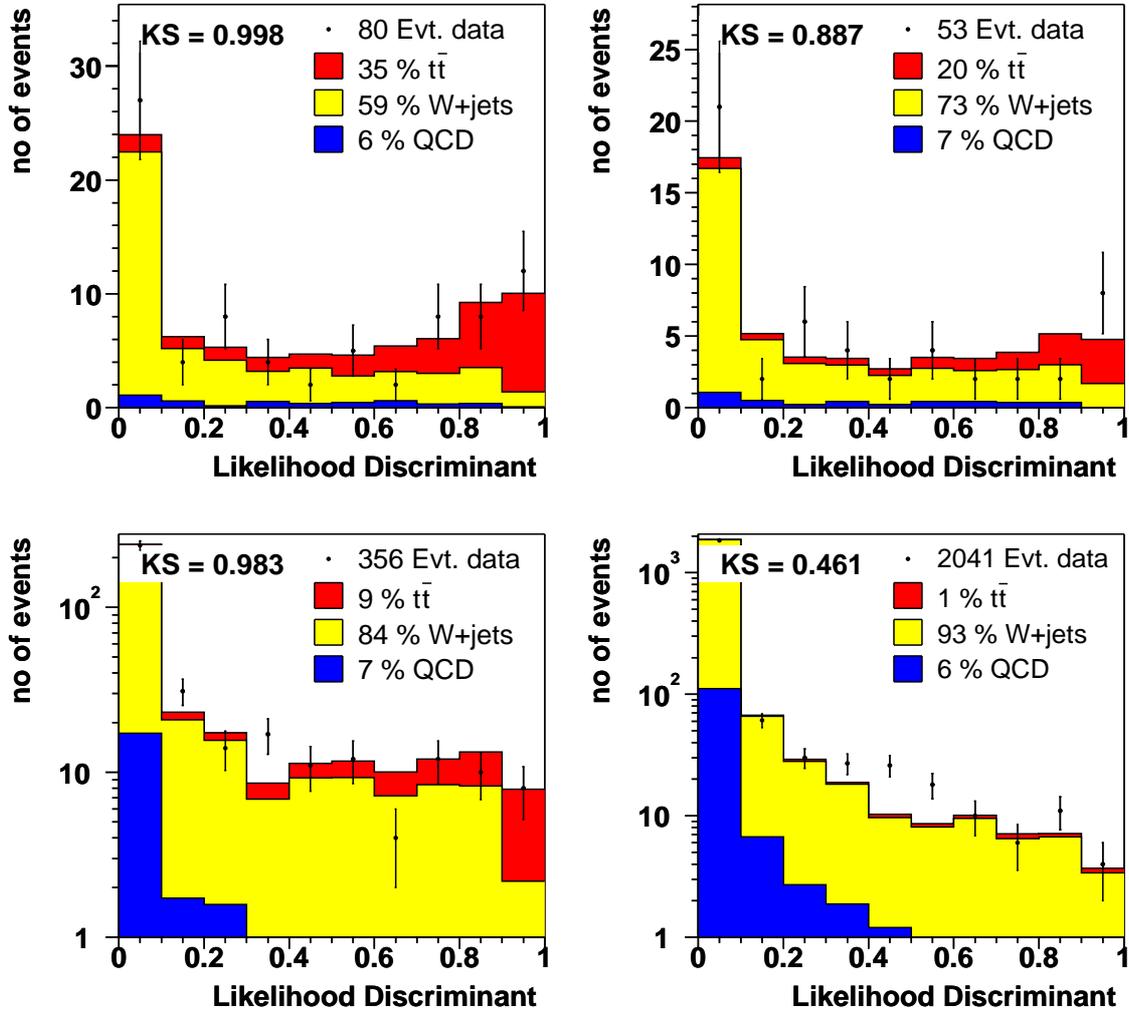


Figure A.1: Likelihood discriminant distribution for data overlaid with  $t\bar{t}$ ,  $W$ -plus-jets and QCD-multijets. Top left:  $W$  MC modeling sample preselected data using  $W$  MC sample with  $Q^2 = \langle p_{T_j} \rangle^2$ ; top right: (*Zero-tag sample*): preselected data with SLV and requiring zero SVT tags in order to reject  $t\bar{t}$  ( $t\bar{t}$ ,  $W$ -plus-jets and QCD-multijets are from the likelihood fit); bottom left: (*3-jet-bin*): preselected data requiring 3 jets reconstructed (expected background and signal shown); bottom right: (*2-jet-bin*): preselected data requiring 2 jets reconstructed (expected background and signal shown).

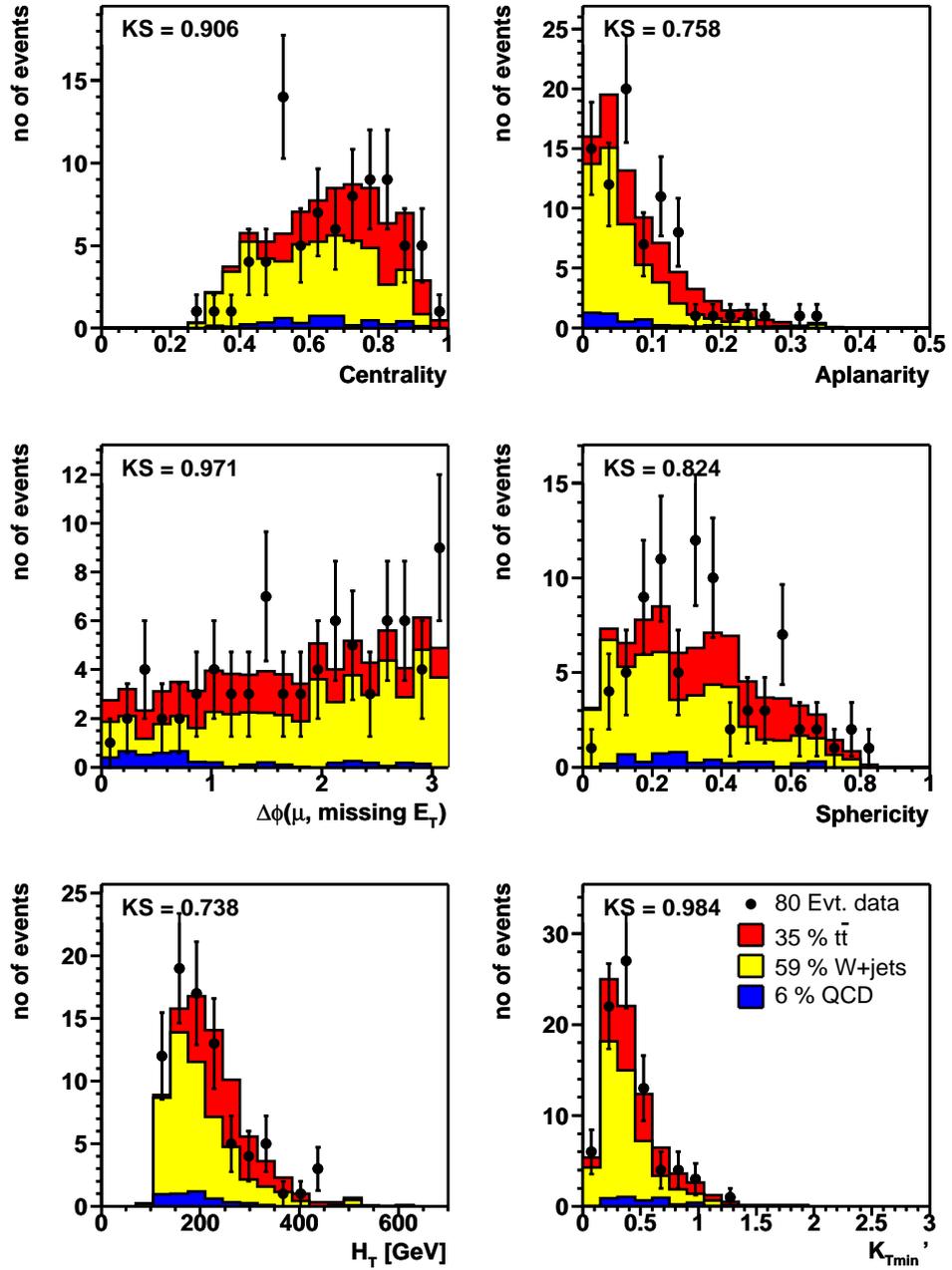


Figure A.2: *W* MC modeling sample: Likelihood discriminant input distributions for the data overlaid with the result from the likelihood fit of  $t\bar{t}$ , *W*-plus-jets and QCD-multijets.

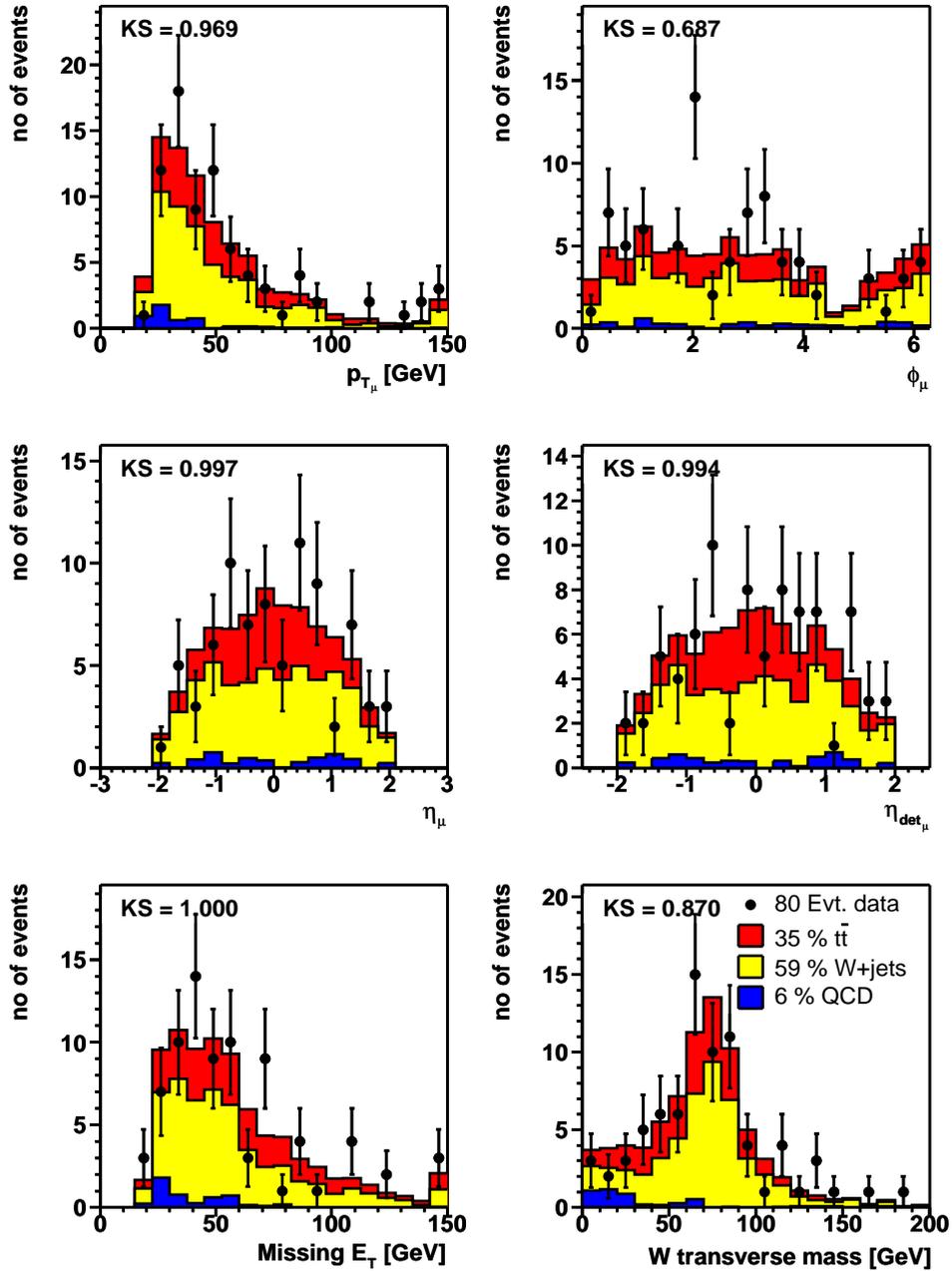


Figure A.3:  $W$  MC modeling sample:  $p_T$ ,  $\phi$ ,  $\eta$ ,  $\eta_{det}$  of the muon,  $\cancel{E}_T$  and  $W$  transverse mass for the data overlaid with the result from the likelihood fit of  $t\bar{t}$ ,  $W$ -plus-jets and QCD-multijets.

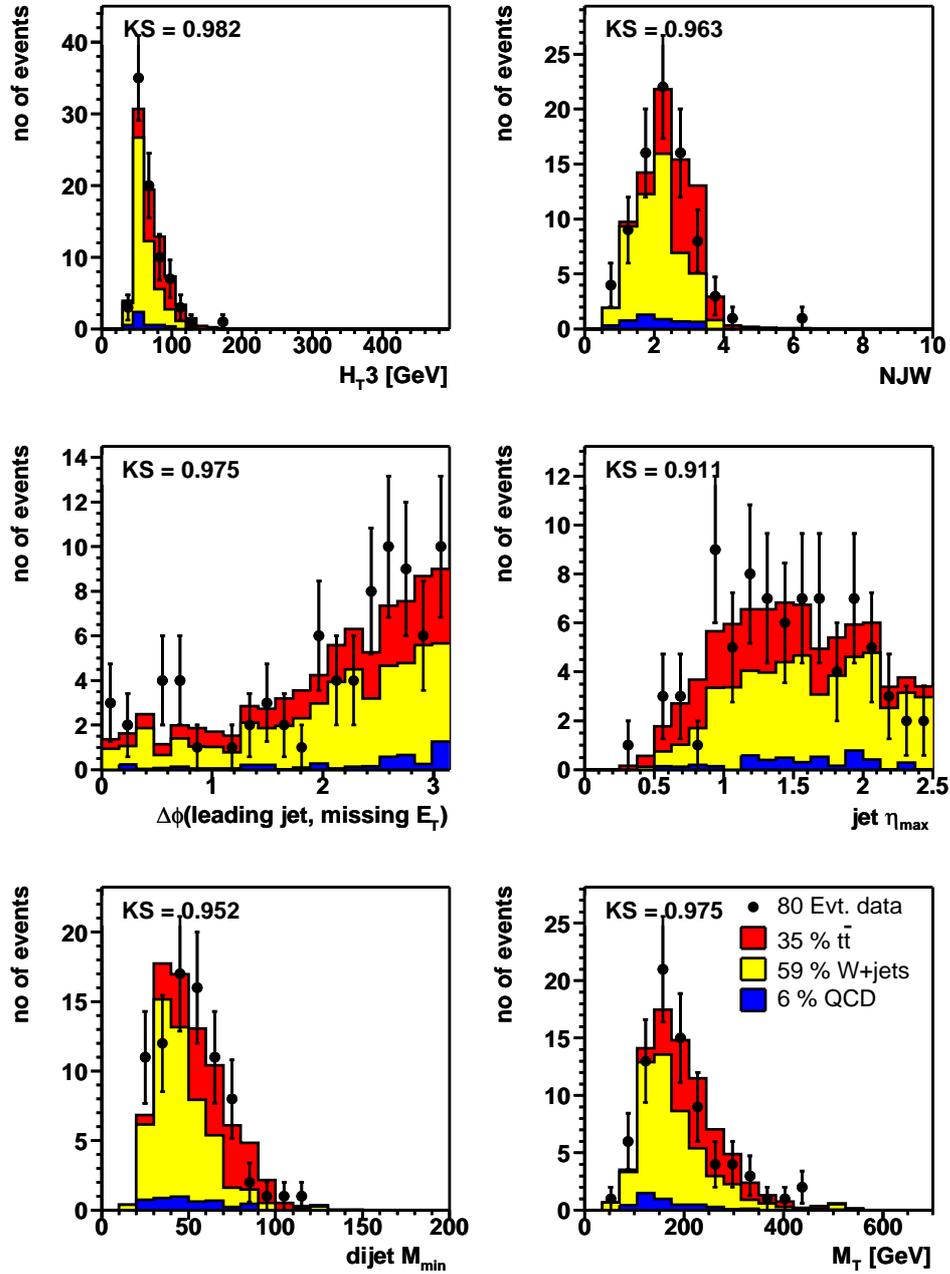


Figure A.4:  $W$  MC modeling sample: Scalar sum of the  $p_T$  of the third and fourth leading jets ( $H_{T3}$ ), NJW is a topological variable built from the  $p_T$  and  $\eta$  of the four leading jets,  $\Delta\phi$  between the leading jet and  $\cancel{E}_T$ , maximum jet  $\eta$ , the minimum dijet mass, transverse mass of the 4 leading jets ( $M_T$ ) for the data overlaid with the result from the likelihood fit of  $t\bar{t}$ ,  $W$ -plus-jets and QCD-multijets.

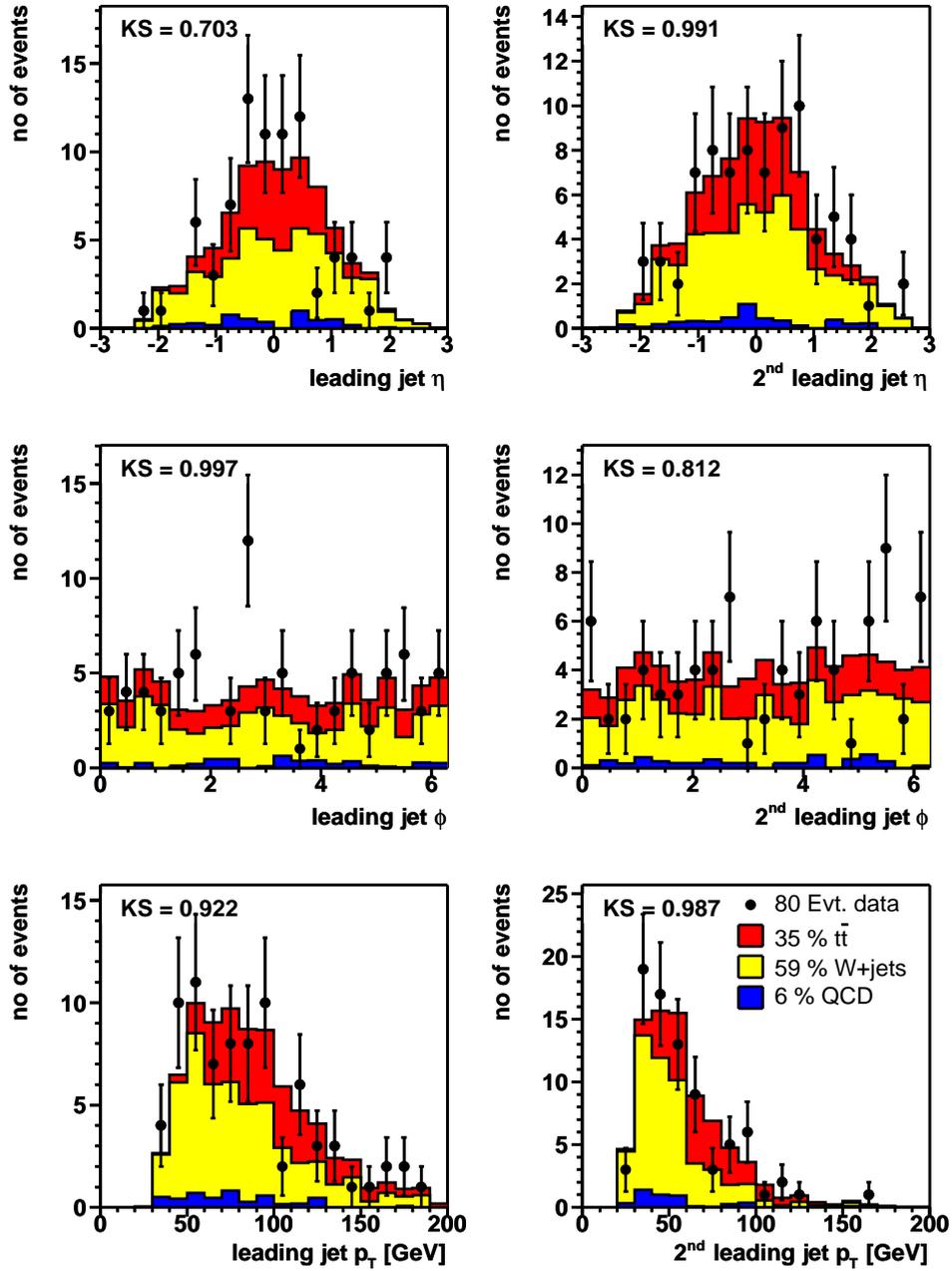


Figure A.5:  $W$  MC modeling sample:  $\eta$ ,  $\phi$  and  $p_T$  of the leading and second leading jet for the data overlaid with the result from the likelihood fit of  $t\bar{t}$ ,  $W$ -plus-jets and QCD-multijets.

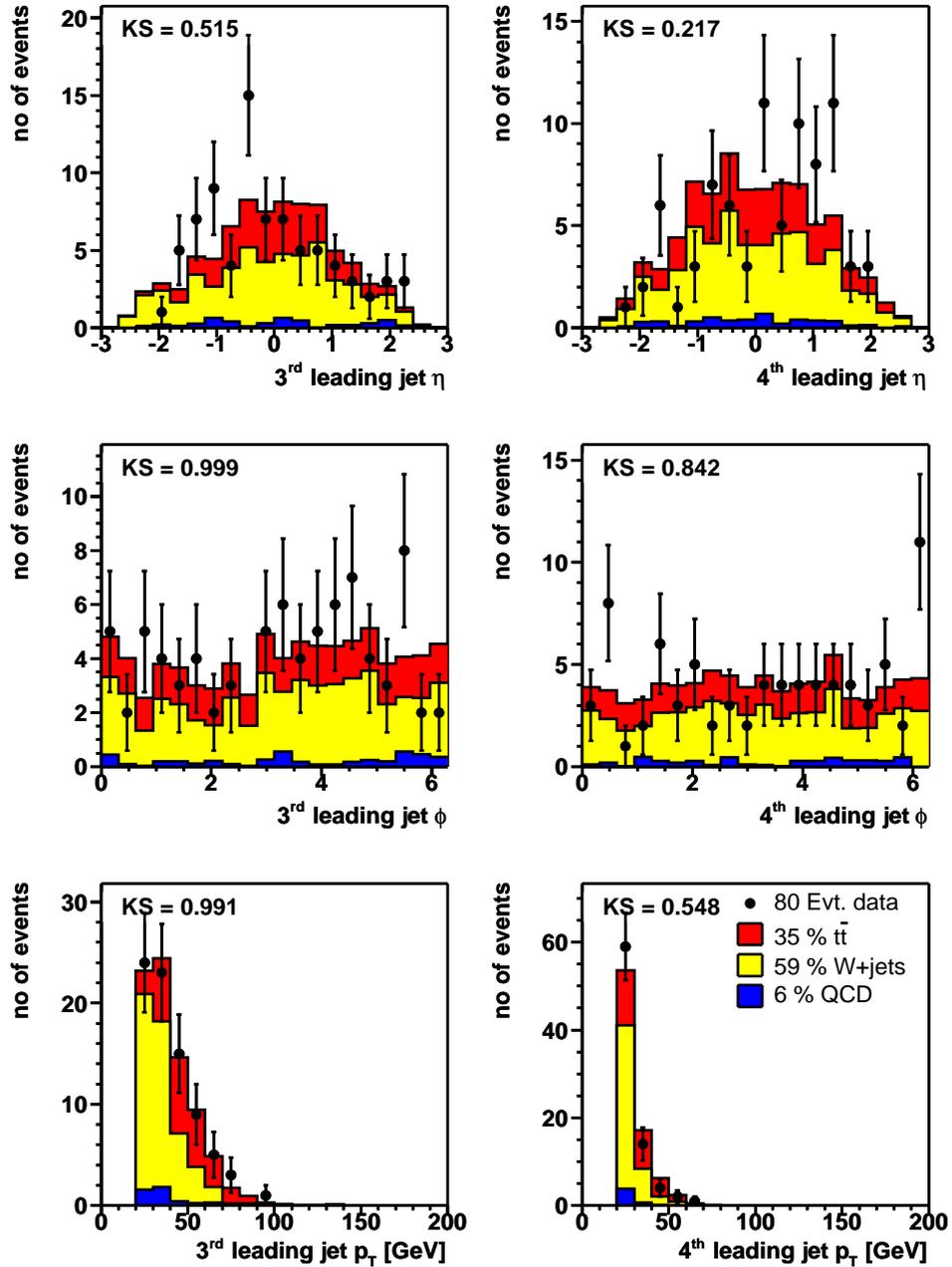


Figure A.6:  $W$  MC modeling sample:  $\eta$ ,  $\phi$  and  $p_T$  of the third leading and fourth leading jet for the preselected data overlaid with the result from the likelihood fit of  $t\bar{t}$ ,  $W$ -plus-jets and QCD-multijets.

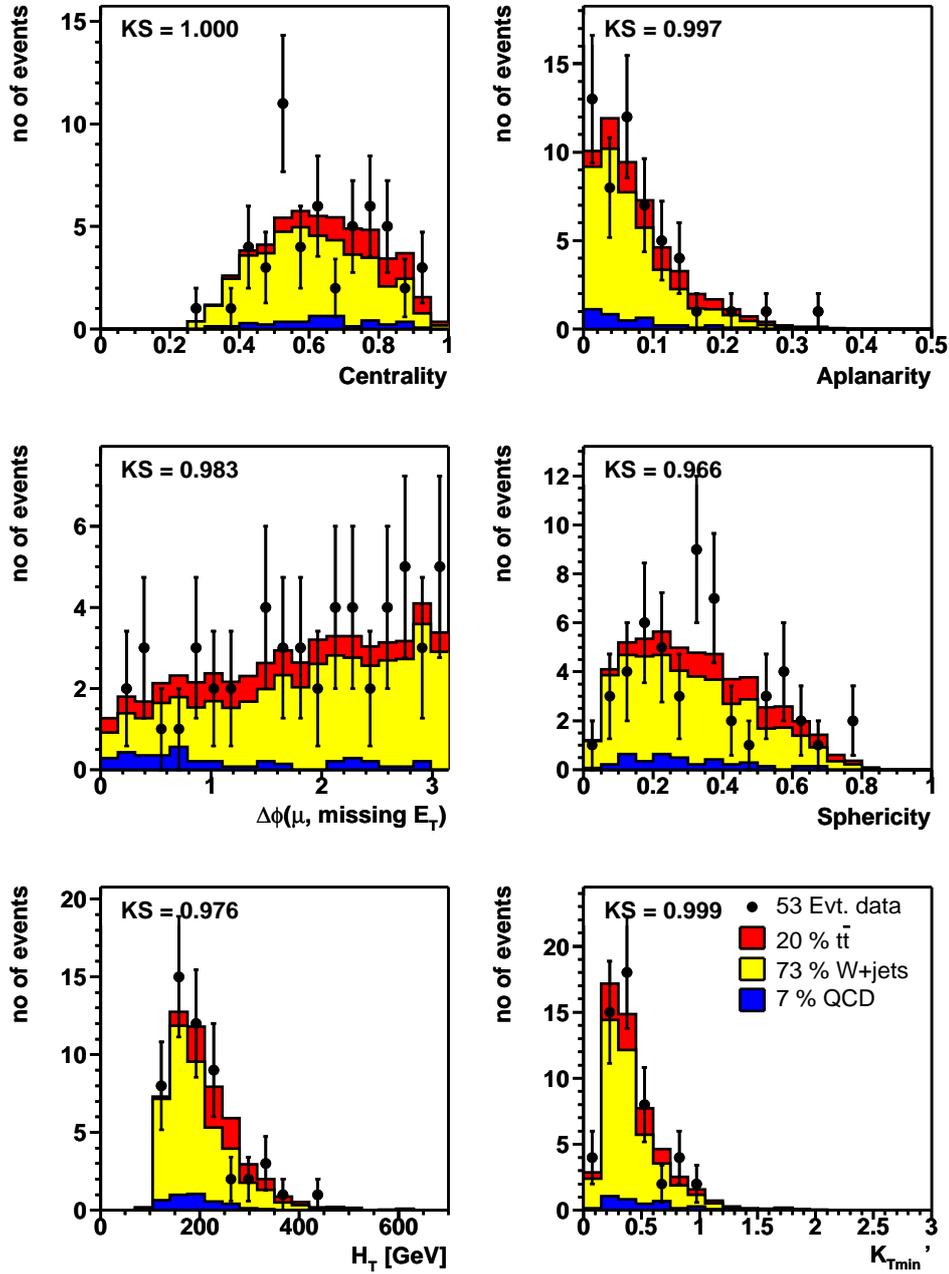


Figure A.7: *Zero-tag sample*: Likelihood discriminant input distributions for the data overlaid with the result from the likelihood fit of  $t\bar{t}$ ,  $W$ -plus-jets and QCD-multijets.

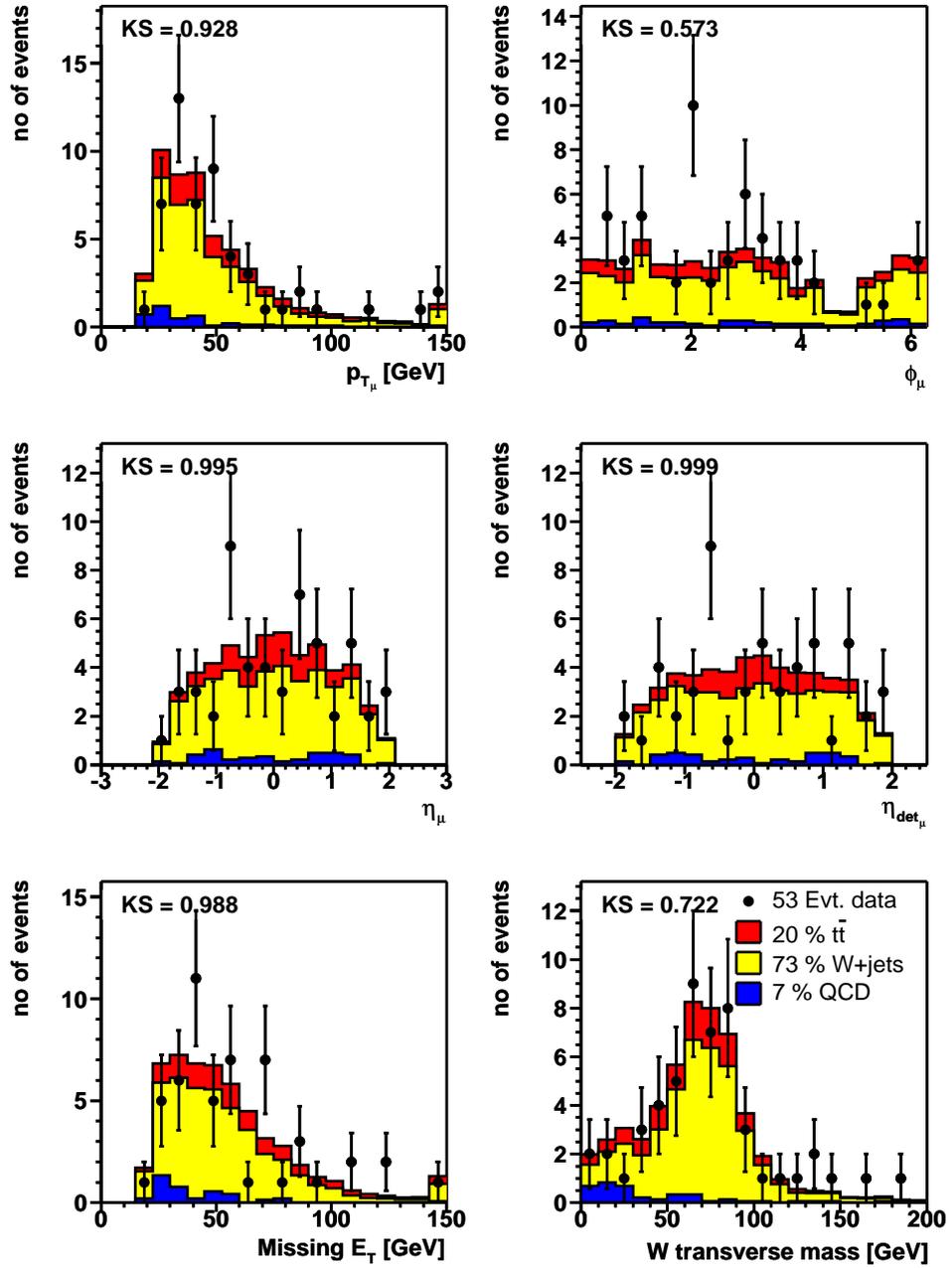


Figure A.8: Zero-tag sample:  $p_T$ ,  $\phi$ ,  $\eta$ ,  $\eta_{det}$  of the muon,  $\cancel{E}_T$  and W transverse mass for the data overlaid with the result from the likelihood fit of  $t\bar{t}$ , W-plus-jets and QCD-multijets.

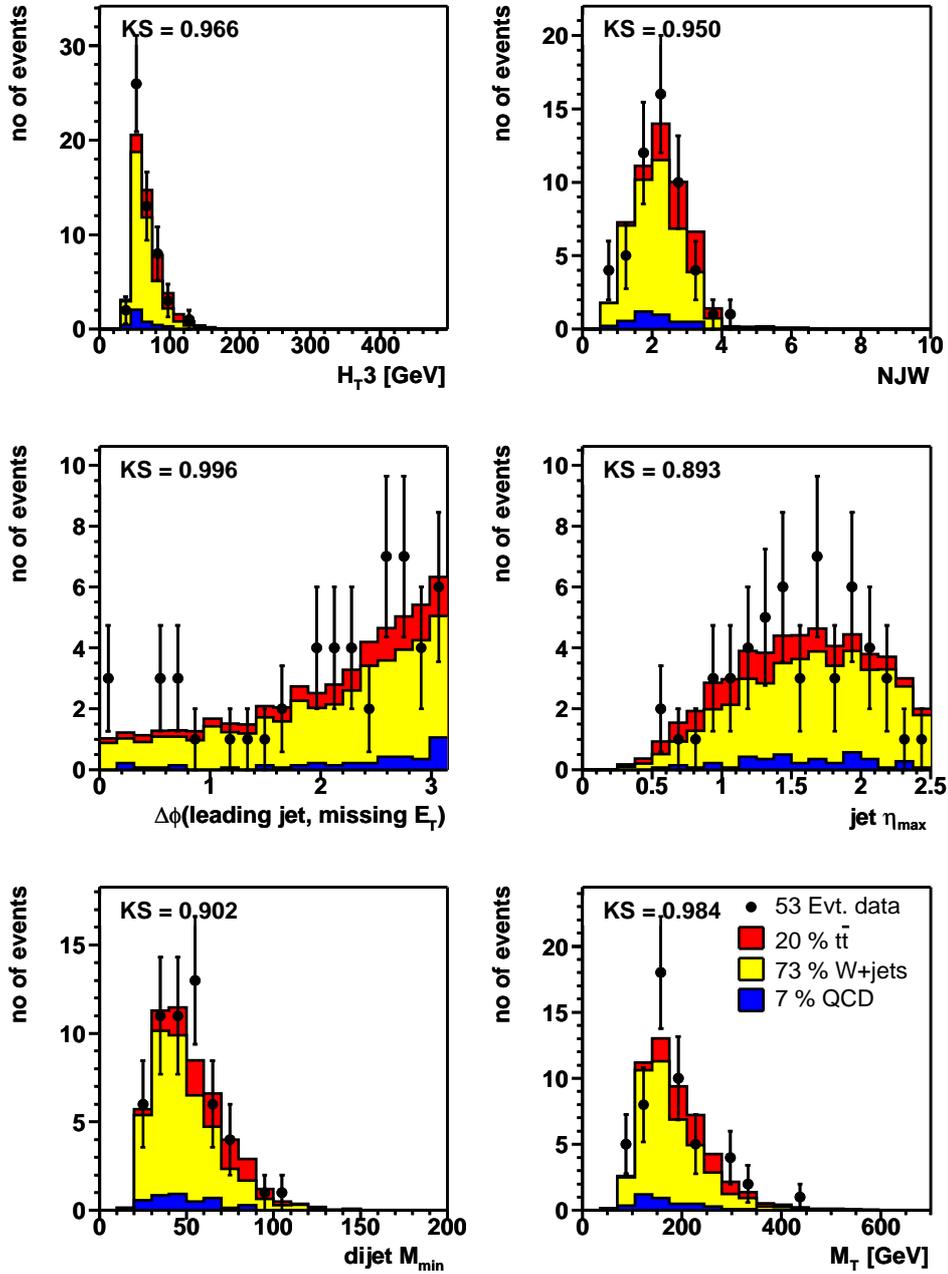


Figure A.9: *Zero-tag sample*: Scalar sum of the  $p_T$  of the third and fourth leading jets ( $H_{T3}$ ), NJW is a topological variable built from the  $p_T$  and  $\eta$  of the four leading jets,  $\Delta\phi$  between the leading jet and  $\cancel{E}_T$ , maximum jet  $\eta$ , the minimum dijet mass, transverse mass of the 4 leading jets ( $M_T$ ) for the data overlaid with the result from the likelihood fit of  $t\bar{t}$ ,  $W$ -plus-jets and QCD-multijets.

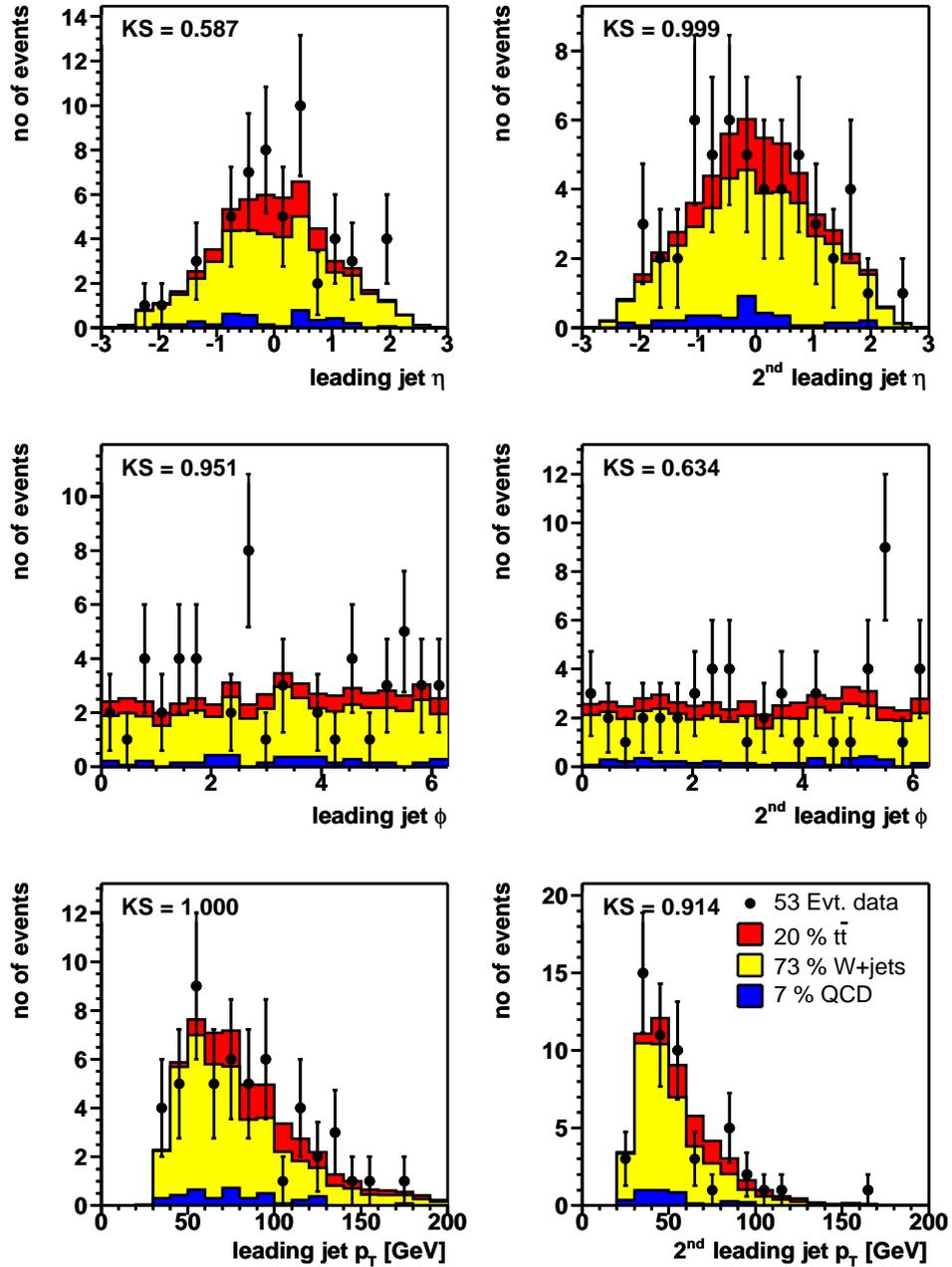


Figure A.10: *Zero-tag sample*:  $\eta$ ,  $\phi$  and  $p_T$  of the leading and second leading jet for the data overlaid with the result from the likelihood fit of  $t\bar{t}$ ,  $W$ -plus-jets and QCD-multijets.

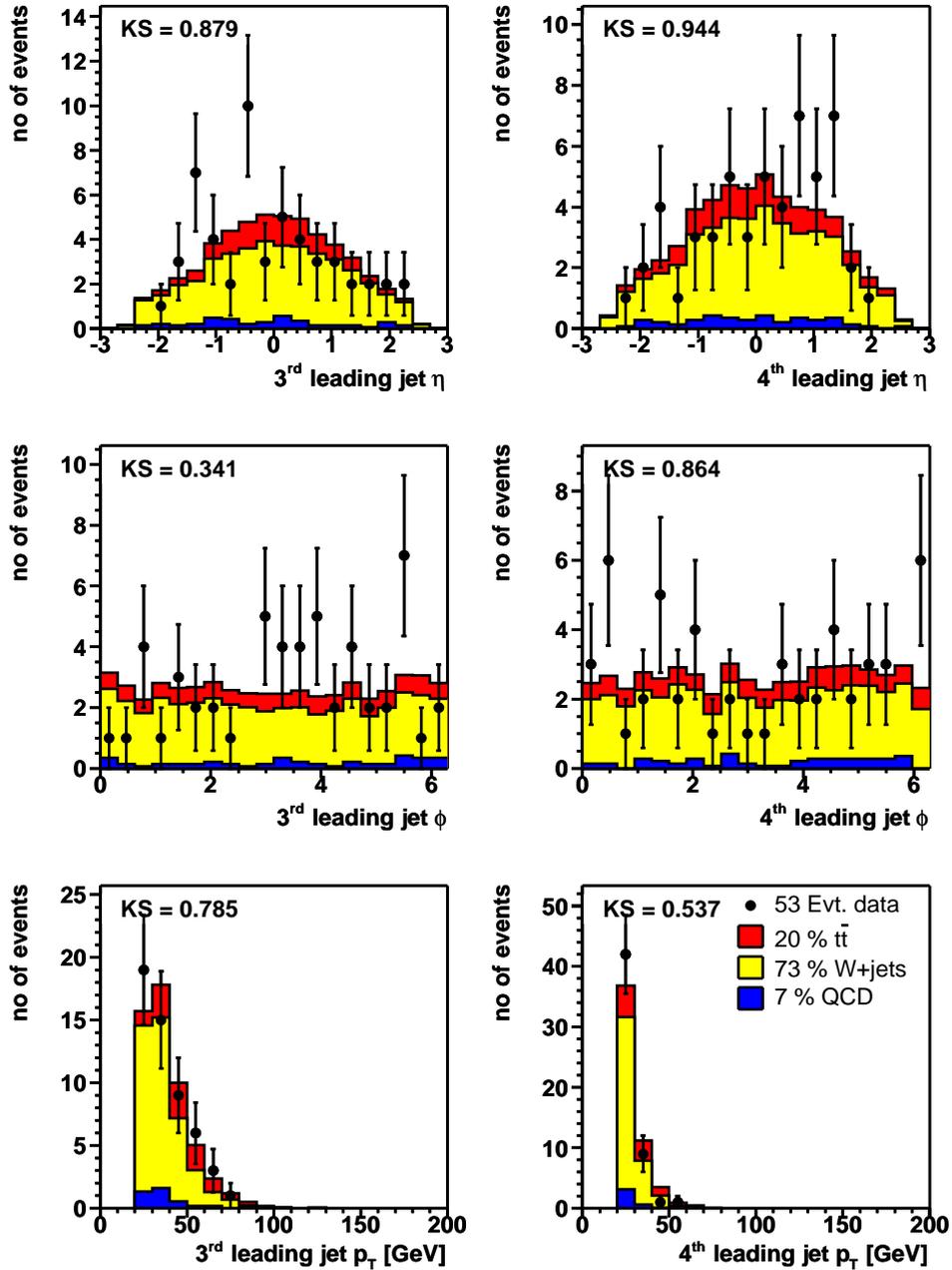


Figure A.11: *Zero-tag sample*:  $\eta$ ,  $\phi$  and  $p_T$  of the third leading and fourth leading jet for the preselected data overlaid with the result from the likelihood fit of  $t\bar{t}$ ,  $W$ -plus-jets and QCD-multijets.

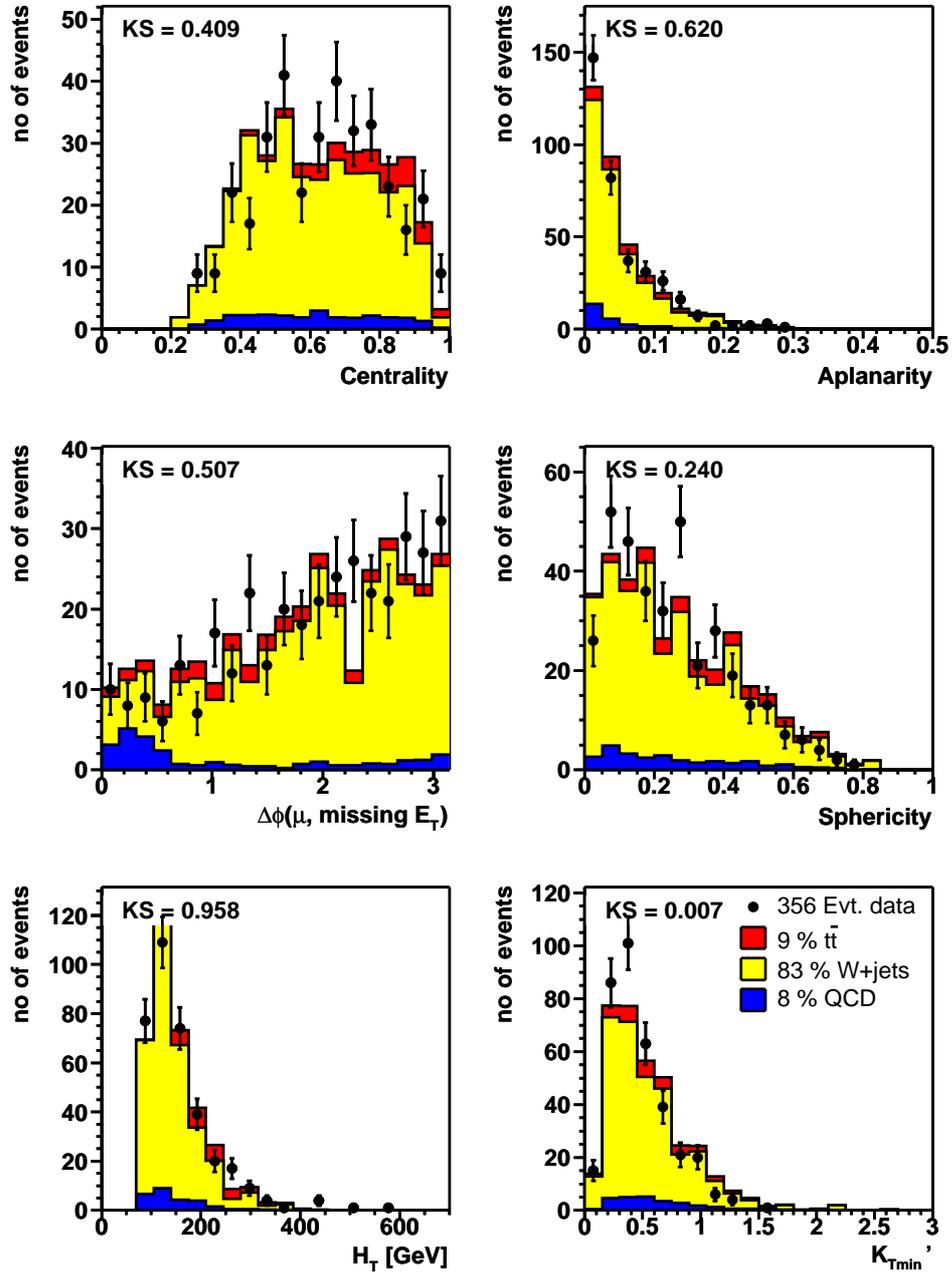


Figure A.12: *3-jet-bin sample*: Likelihood discriminant input distributions for the data overlaid with the expected number of  $t\bar{t}$ ,  $W$ -plus-jets and QCD-multijets.

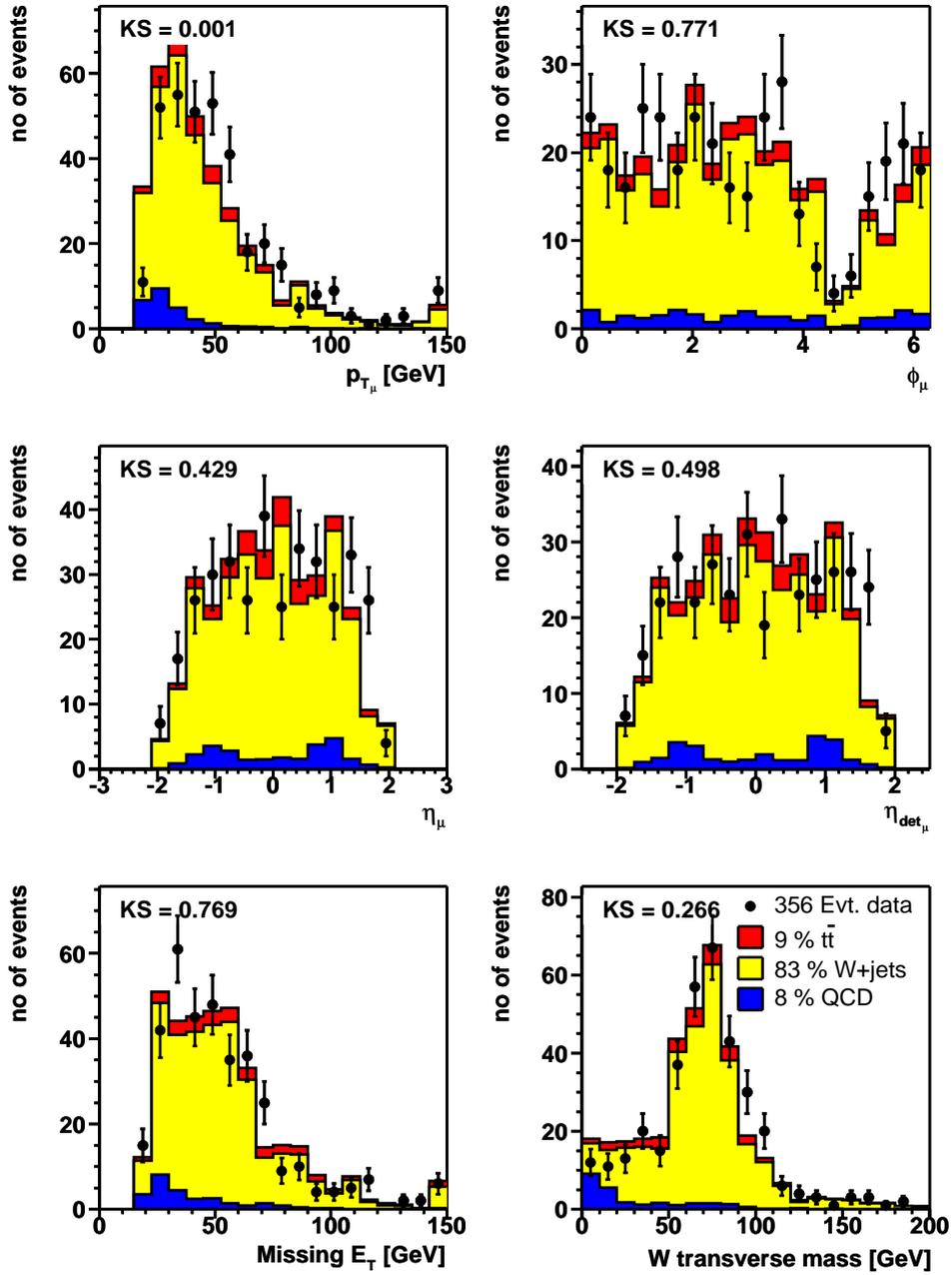


Figure A.13:  $3$ -jet-bin sample:  $p_T$ ,  $\phi$ ,  $\eta$ ,  $\eta_{det}$  of the muon,  $\cancel{E}_T$  and  $W$  transverse mass for the data overlaid with the expected number of  $t\bar{t}$ ,  $W$ -plus-jets and QCD-multijets.

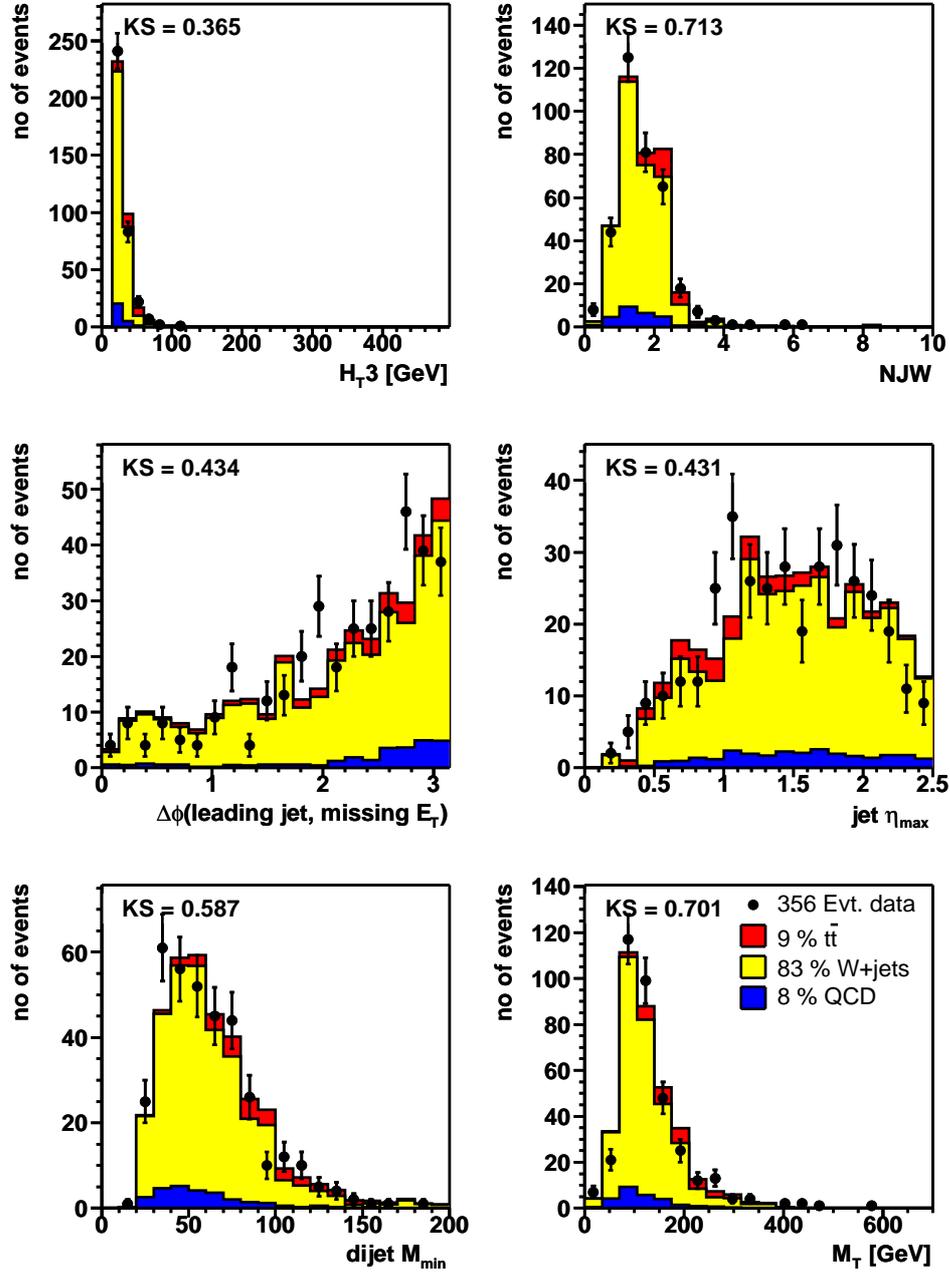


Figure A.14: *3-jet-bin sample*: Scalar sum of the  $p_T$  of the third and fourth leading jets ( $H_{T3}$ ), NJW is a topological variable built from the  $p_T$  and  $\eta$  of the four leading jets,  $\Delta\phi$  between the leading jet and  $\cancel{E}_T$ , maximum jet  $\eta$ , the minimum dijet mass, transverse mass of the 4 leading jets ( $M_T$ ) for the data overlaid with the expected number of  $t\bar{t}$ ,  $W$ -plus-jets and QCD-multijets.

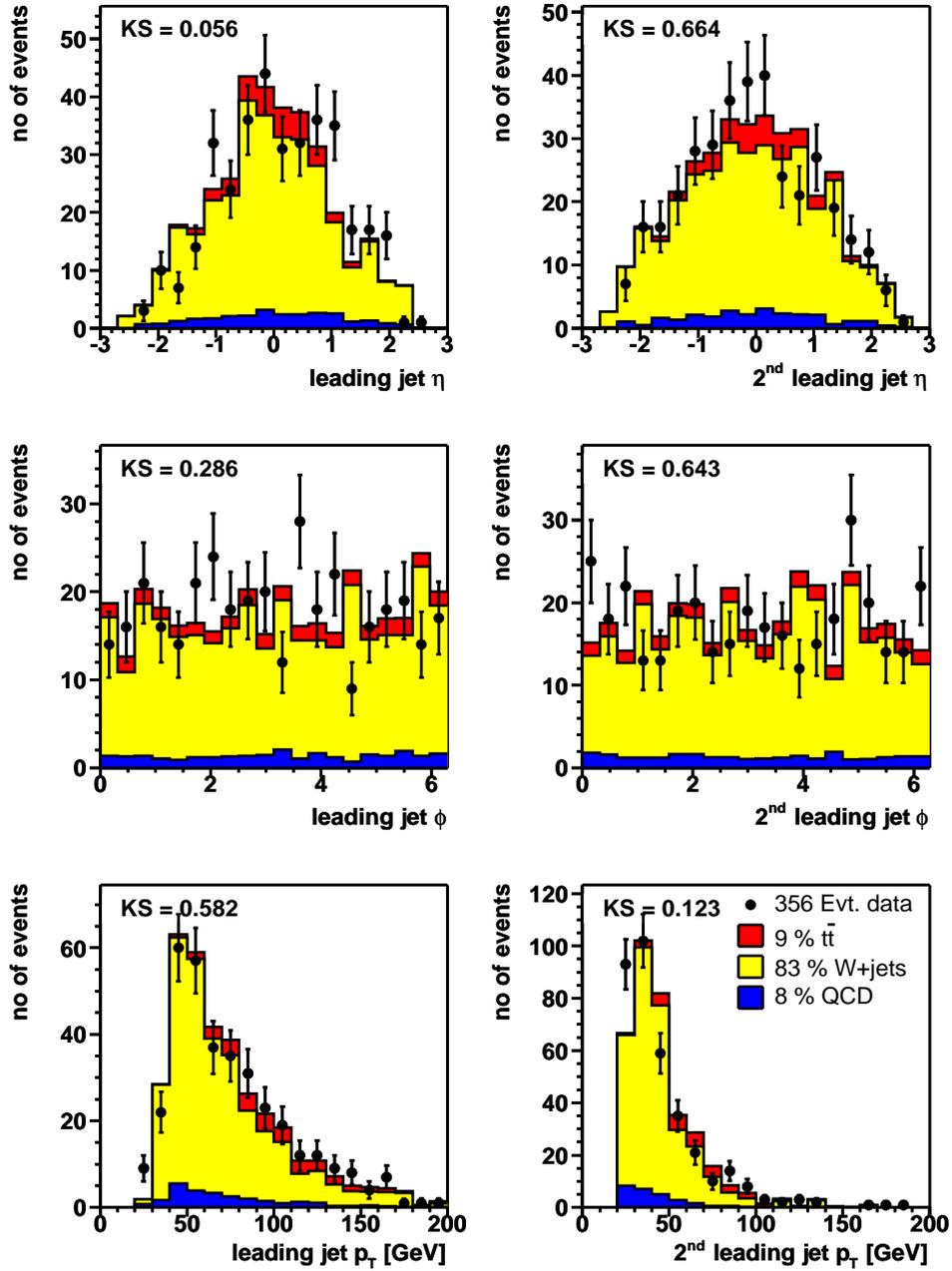


Figure A.15: 3-jet-bin sample:  $\eta$ ,  $\phi$  and  $p_T$  of the leading and second leading jet for the data overlaid with the expected number of  $t\bar{t}$ ,  $W$ -plus-jets and QCD-multijets.

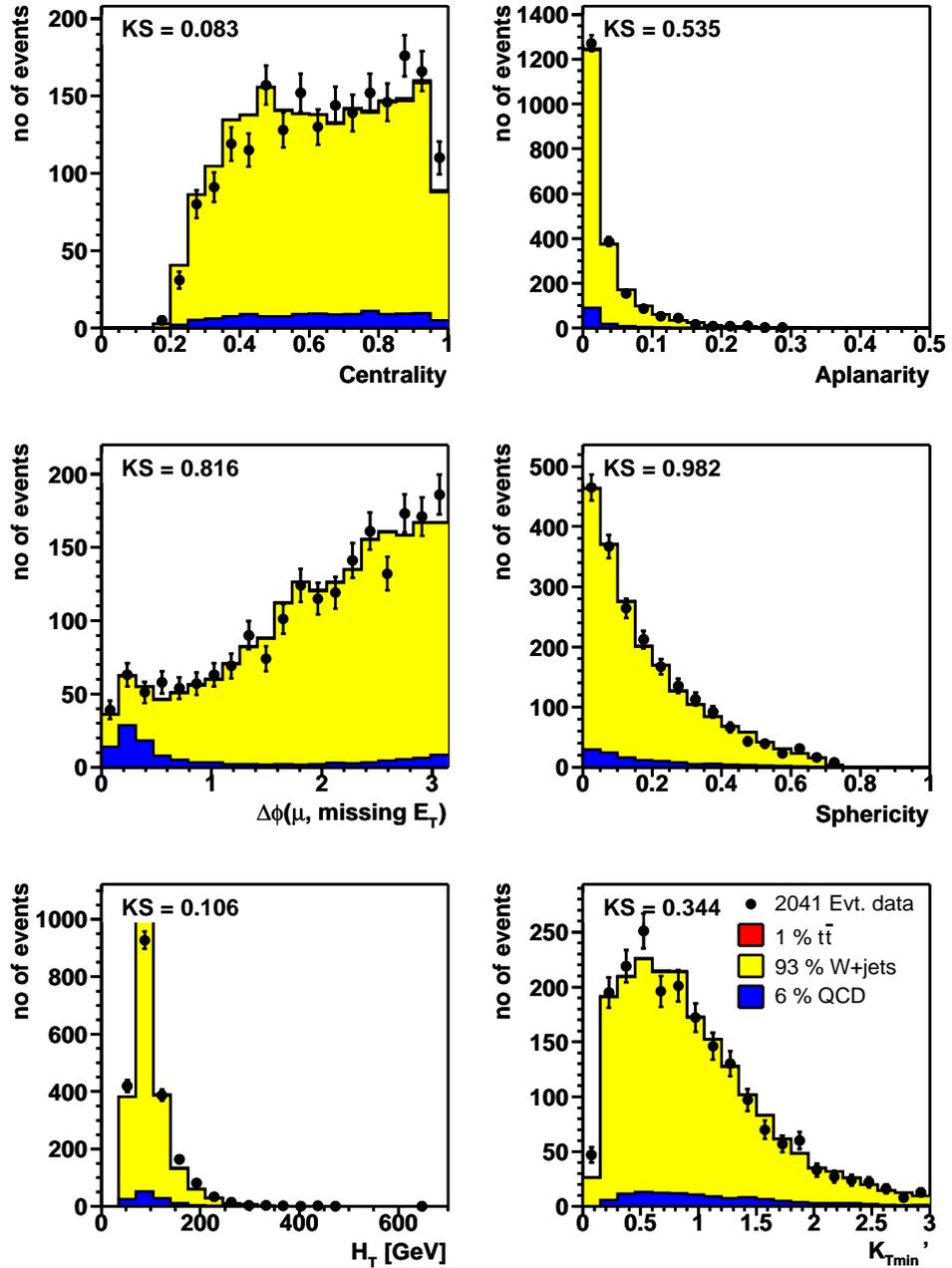


Figure A.16: *2-jet-bin sample*: Likelihood discriminant input distributions for the data overlaid with the expected number of  $t\bar{t}$ ,  $W$ -plus-jets and QCD-multijets.

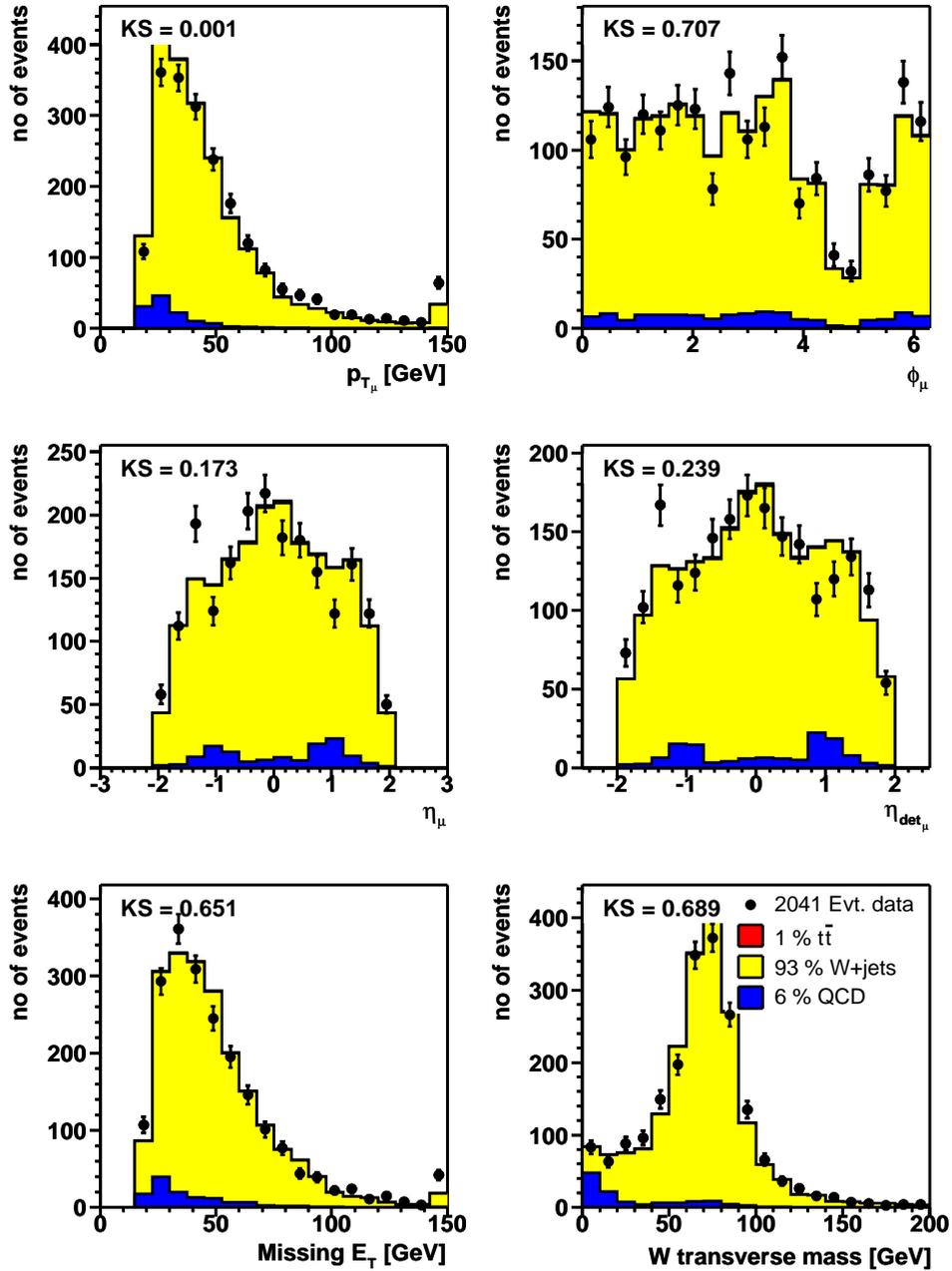


Figure A.17:  $2$ -jet-bin sample:  $p_T$ ,  $\phi$ ,  $\eta$ ,  $\eta_{det}$  of the muon,  $\cancel{E}_T$  and  $W$  transverse mass for the data overlaid with the expected number of  $t\bar{t}$ ,  $W$ -plus-jets and QCD-multijets.

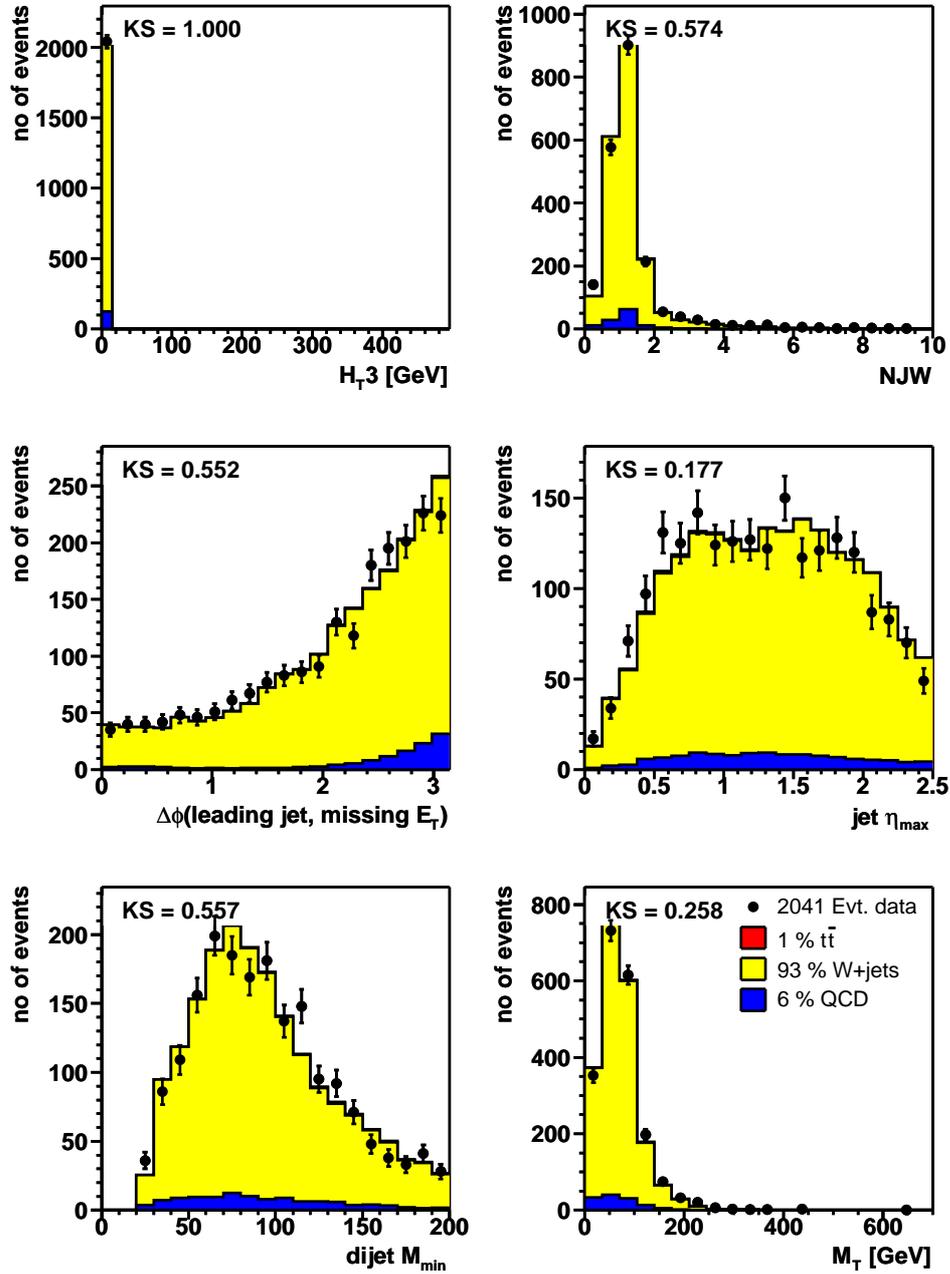


Figure A.18: *2-jet-bin sample*: Scalar sum of the  $p_T$  of the third and fourth leading jets ( $H_{T3}$ ), NJW is a topological variable built from the  $p_T$  and  $\eta$  of the four leading jets,  $\Delta\phi$  between the leading jet and  $\cancel{E}_T$ , maximum jet  $\eta$ , the minimum dijet mass, transverse mass of the 4 leading jets ( $M_T$ ) for the data overlaid with the expected number of  $t\bar{t}$ ,  $W$ -plus-jets and QCD-multijets.

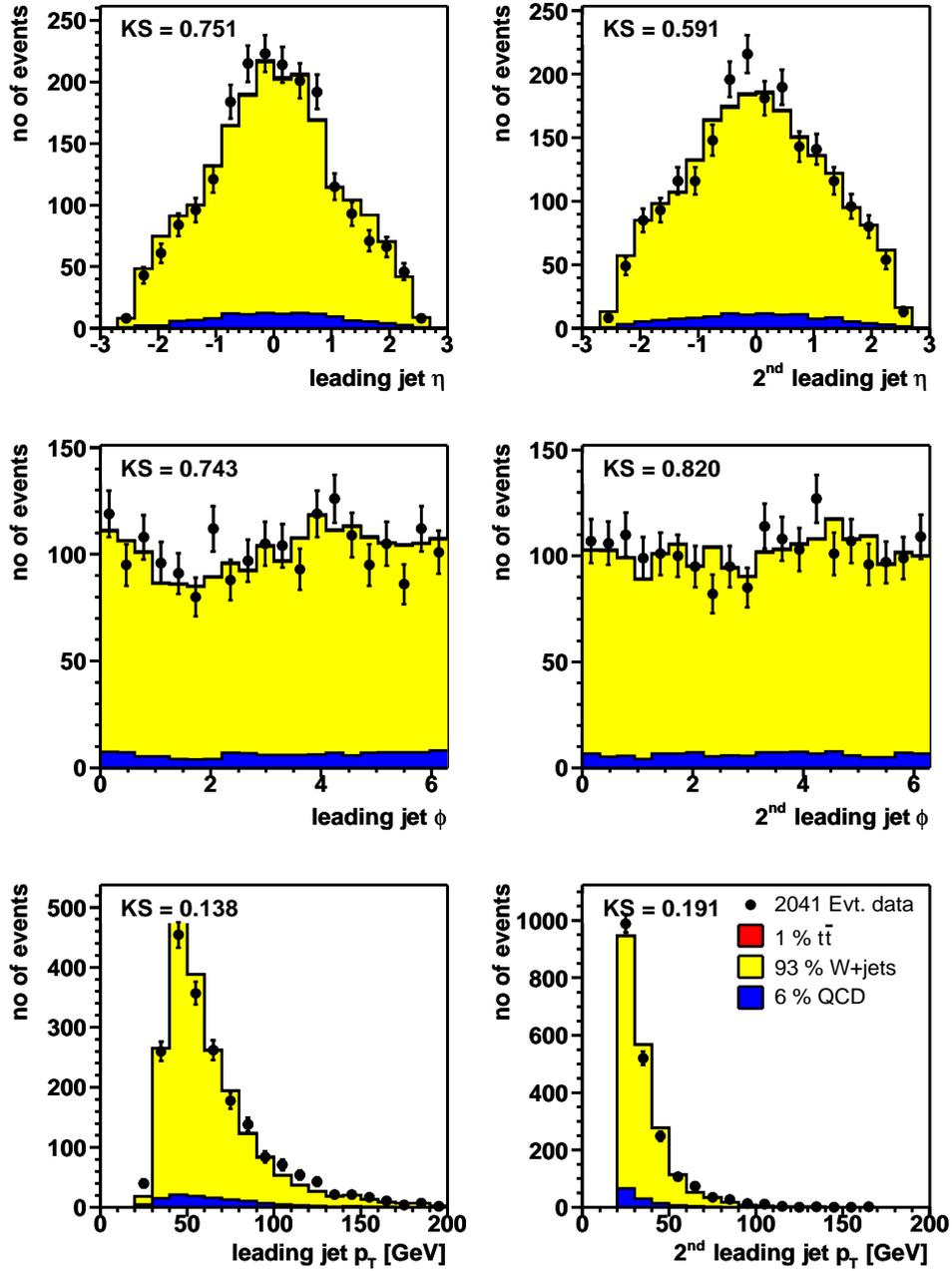


Figure A.19: 2-jet-bin sample:  $\eta$ ,  $\phi$  and  $p_T$  of the leading and second leading jet for the data overlaid with the expected number of  $t\bar{t}$ ,  $W$ -plus-jets and QCD-multijets.

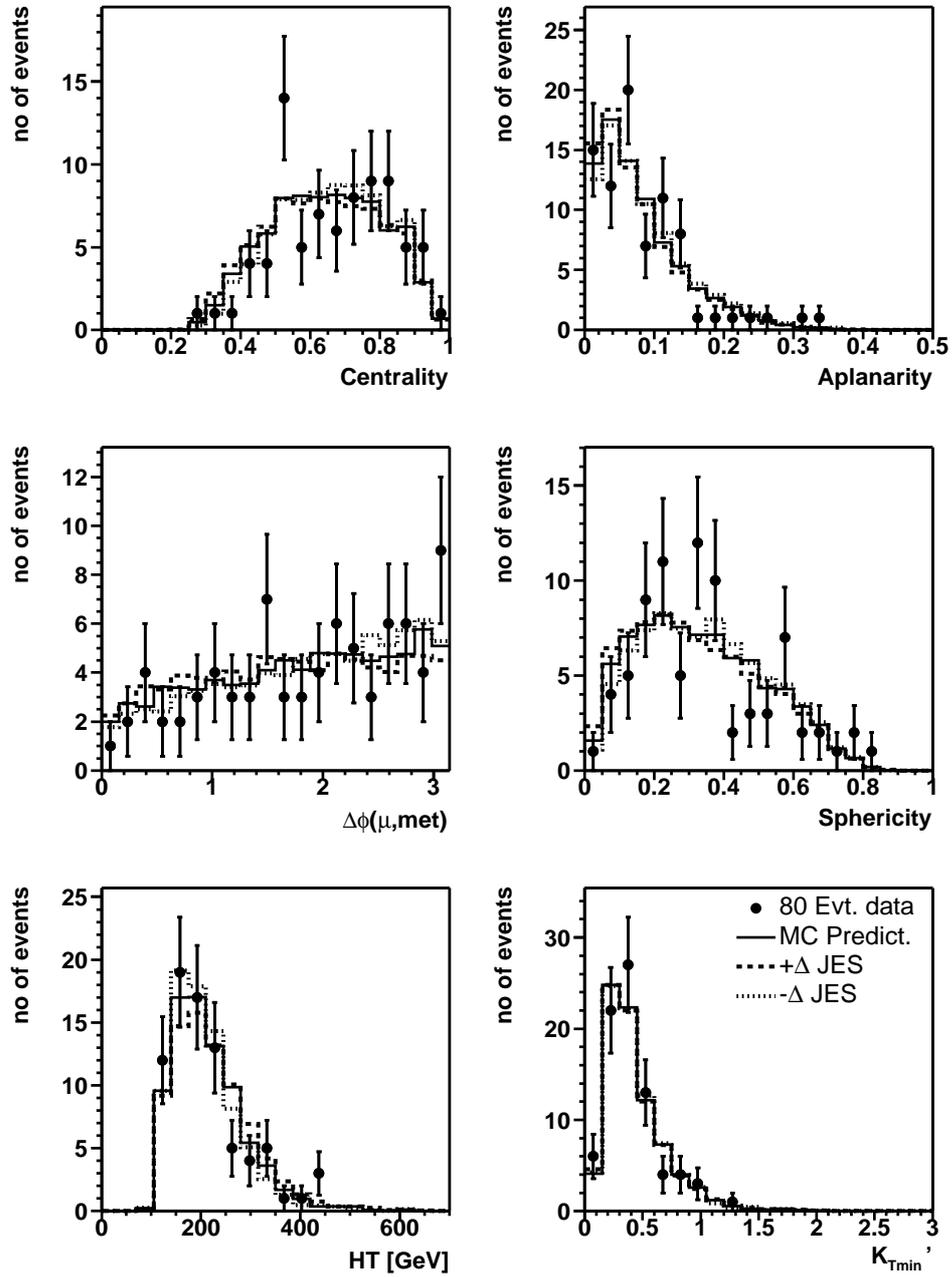


Figure A.20: Data to MC comparison including the  $\pm 1\sigma$  JES uncertainty band for the likelihood discriminant input distributions for the preselected data.

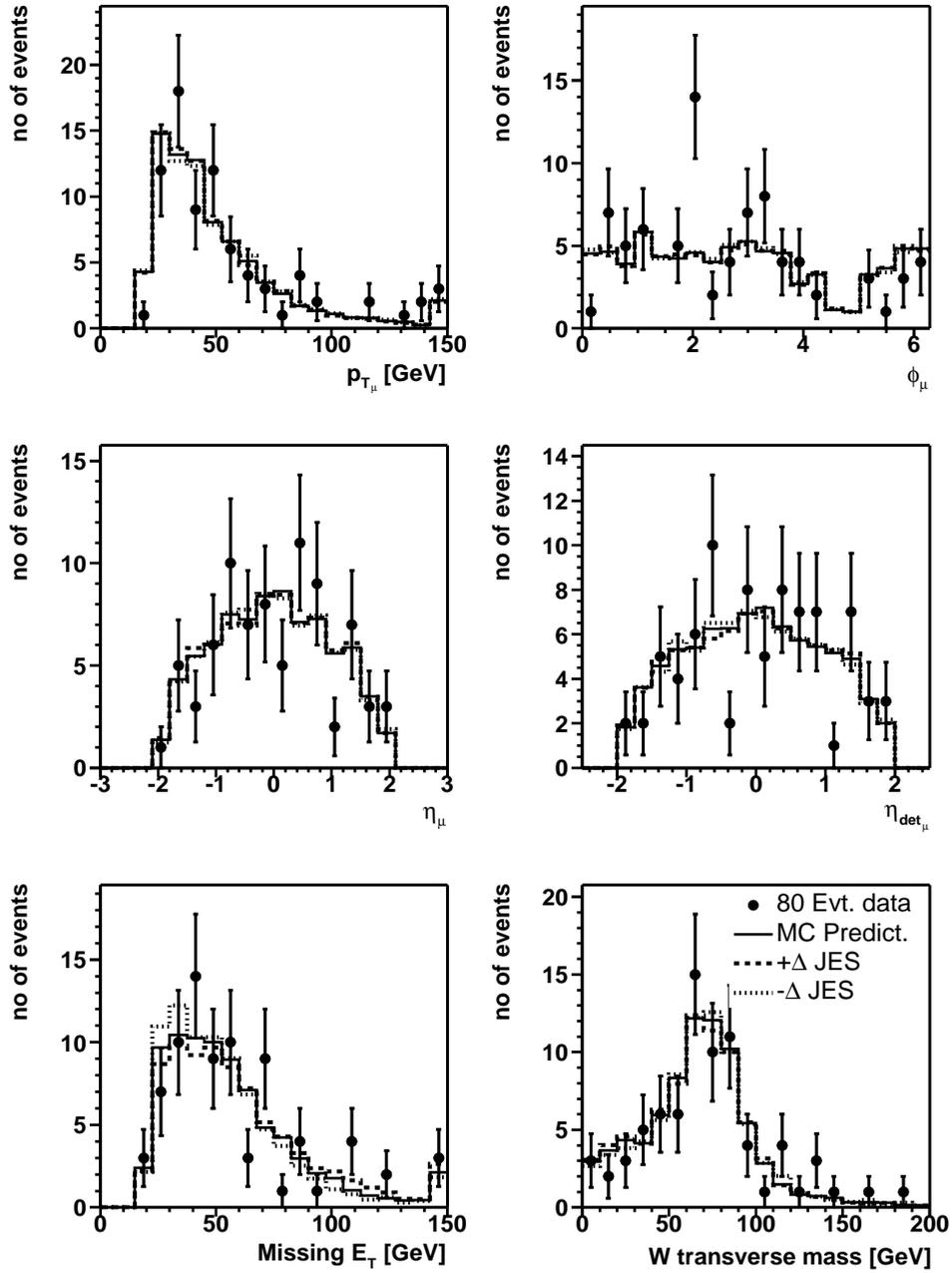


Figure A.21: Data to MC comparison including the  $\pm 1\sigma$  JES uncertainty band for  $p_T$ ,  $\phi$ ,  $\eta$ ,  $\eta_{det}$  of the muon,  $\cancel{E}_T$  and W transverse mass for the preselected data.

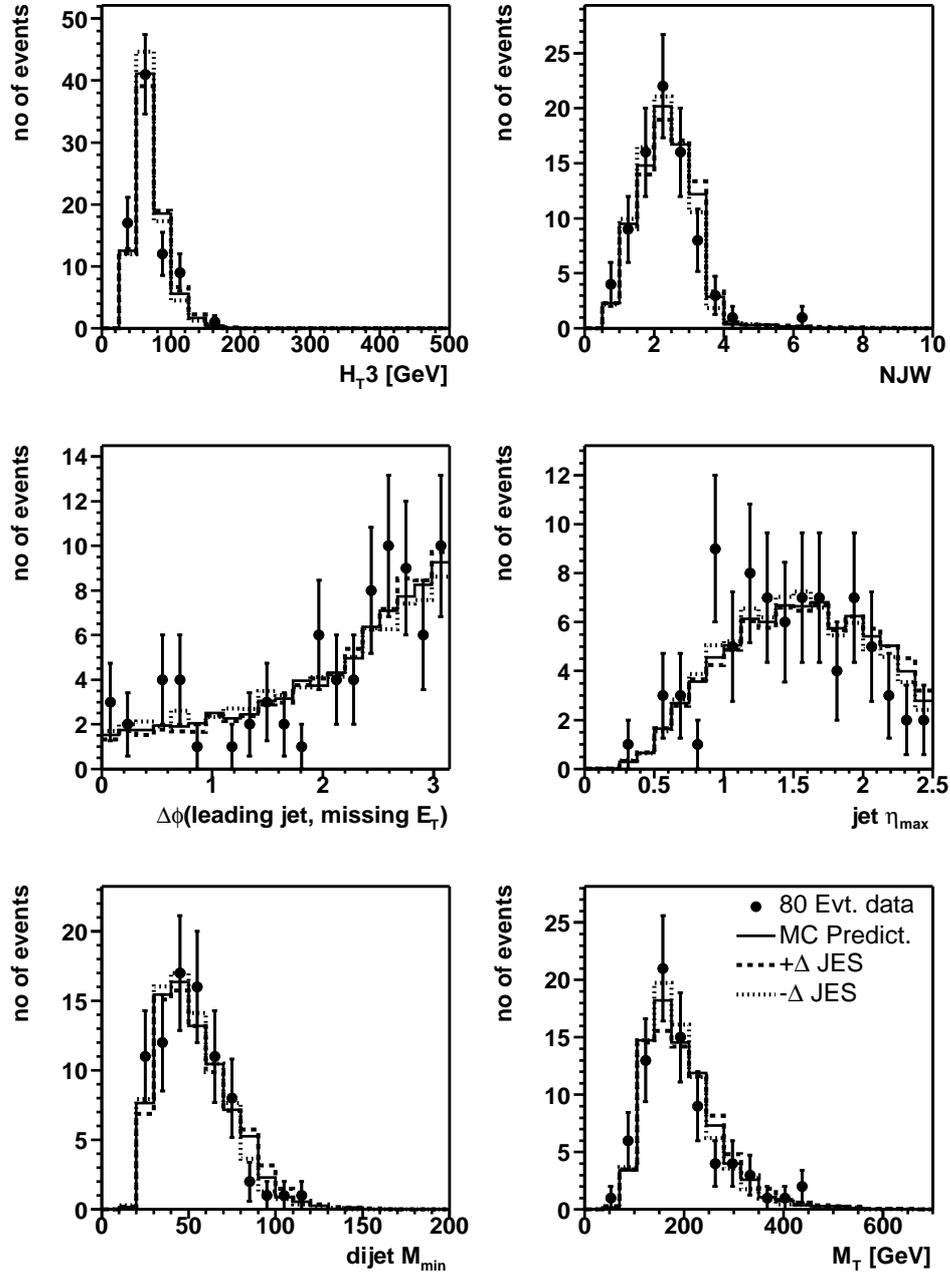


Figure A.22: Data to MC comparison including the  $\pm 1\sigma$  JES uncertainty band for the scalar sum of the  $p_T$  of the third and fourth leading jets ( $H_{T3}$ ), NJW is a topological variable built from the  $p_T$  and  $\eta$  of the four leading jets,  $\Delta\phi$  between the leading jet and  $\cancel{E}_T$ , maximum jet  $\eta$ , the minimum dijet mass, transverse mass of the 4 leading jets ( $M_T$ ) for the preselected data.

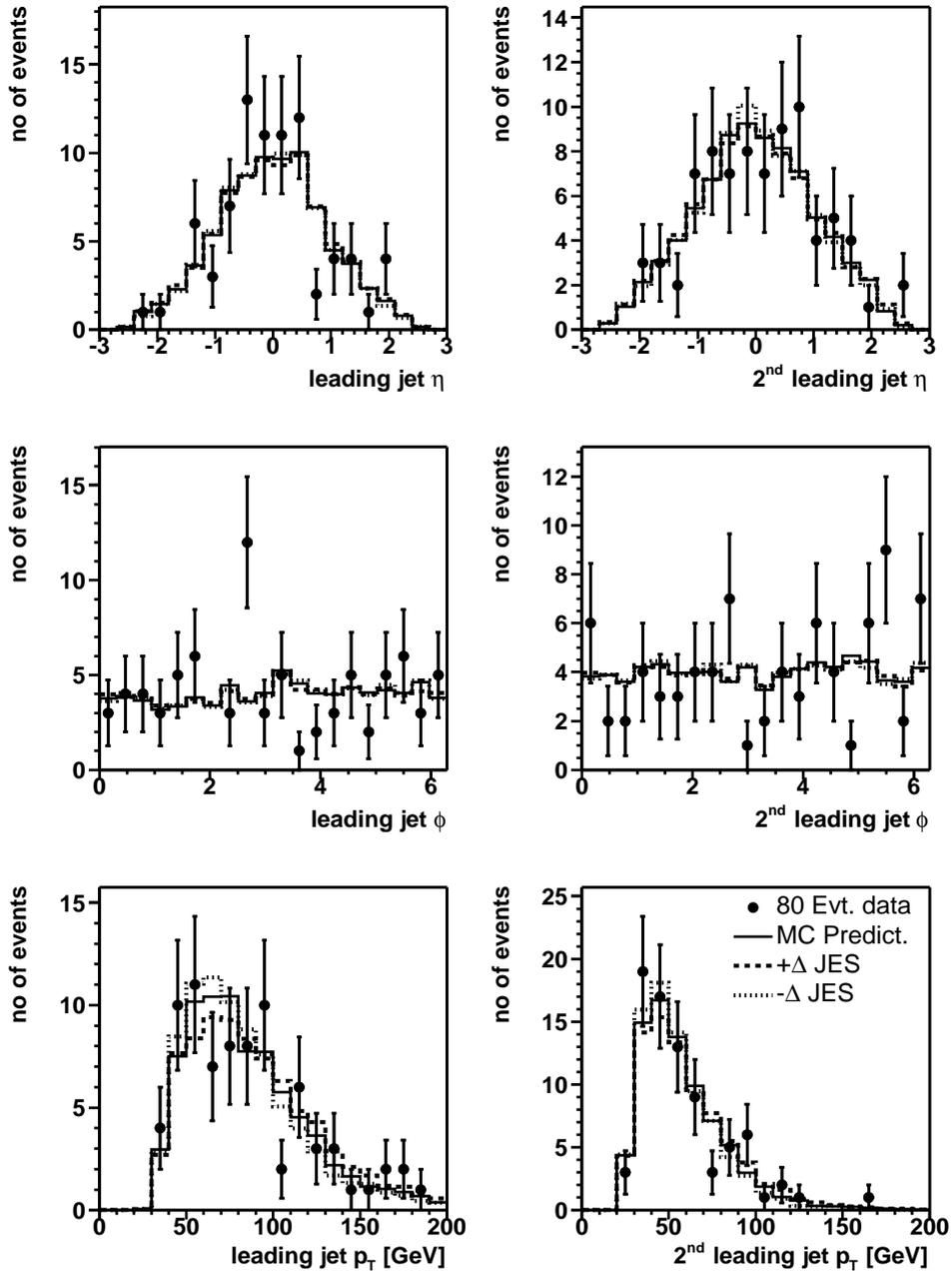


Figure A.23: Data to MC comparison including the  $\pm 1\sigma$  JES uncertainty band for  $\eta$ ,  $\phi$  and  $p_T$  of the leading and second leading jet for the preselected data.

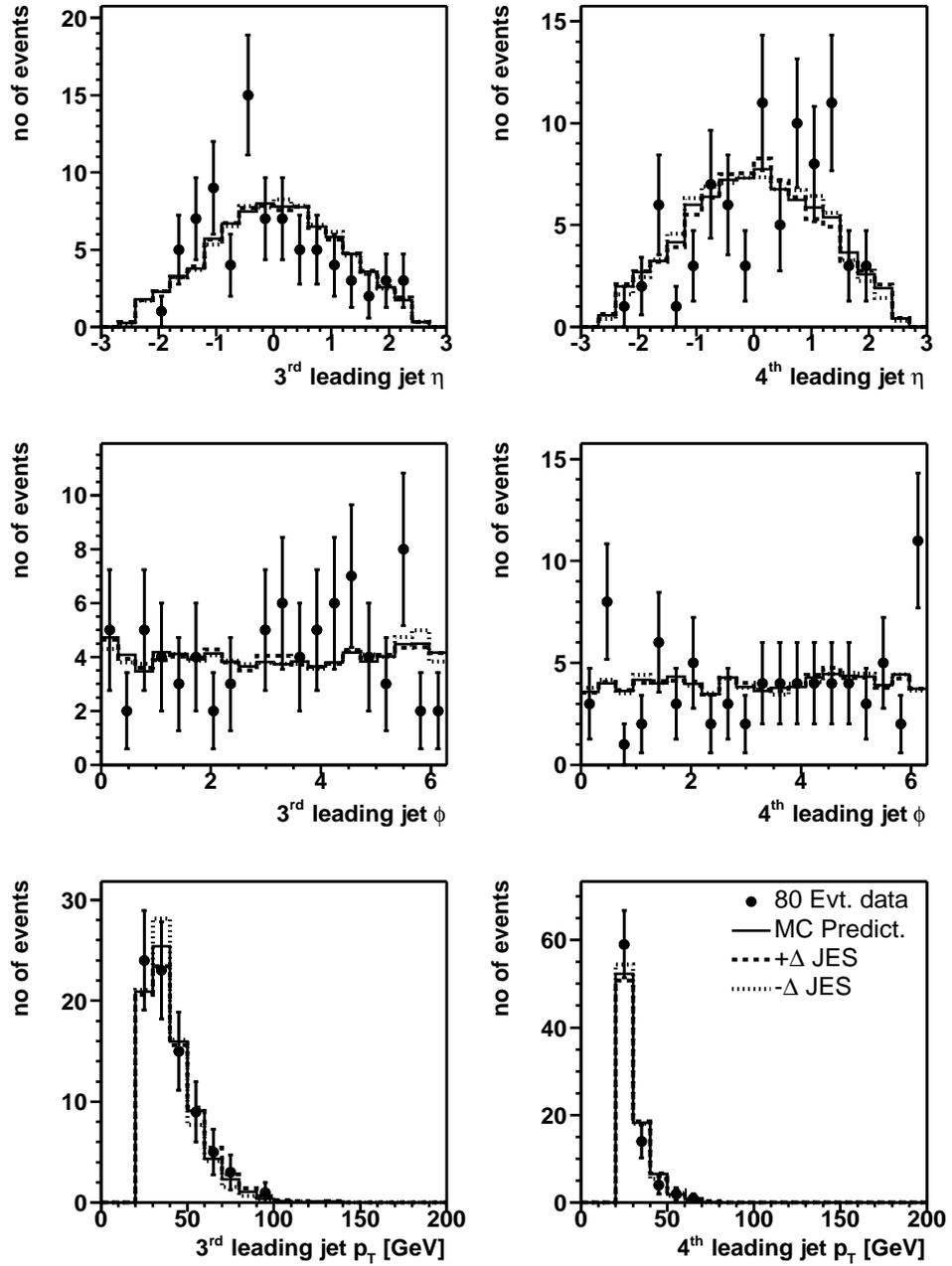


Figure A.24: Data to MC comparison including the  $\pm 1\sigma$  JES uncertainty band for  $\eta$ ,  $\phi$  and  $p_T$  of the third leading and fourth leading jet for the preselected data.

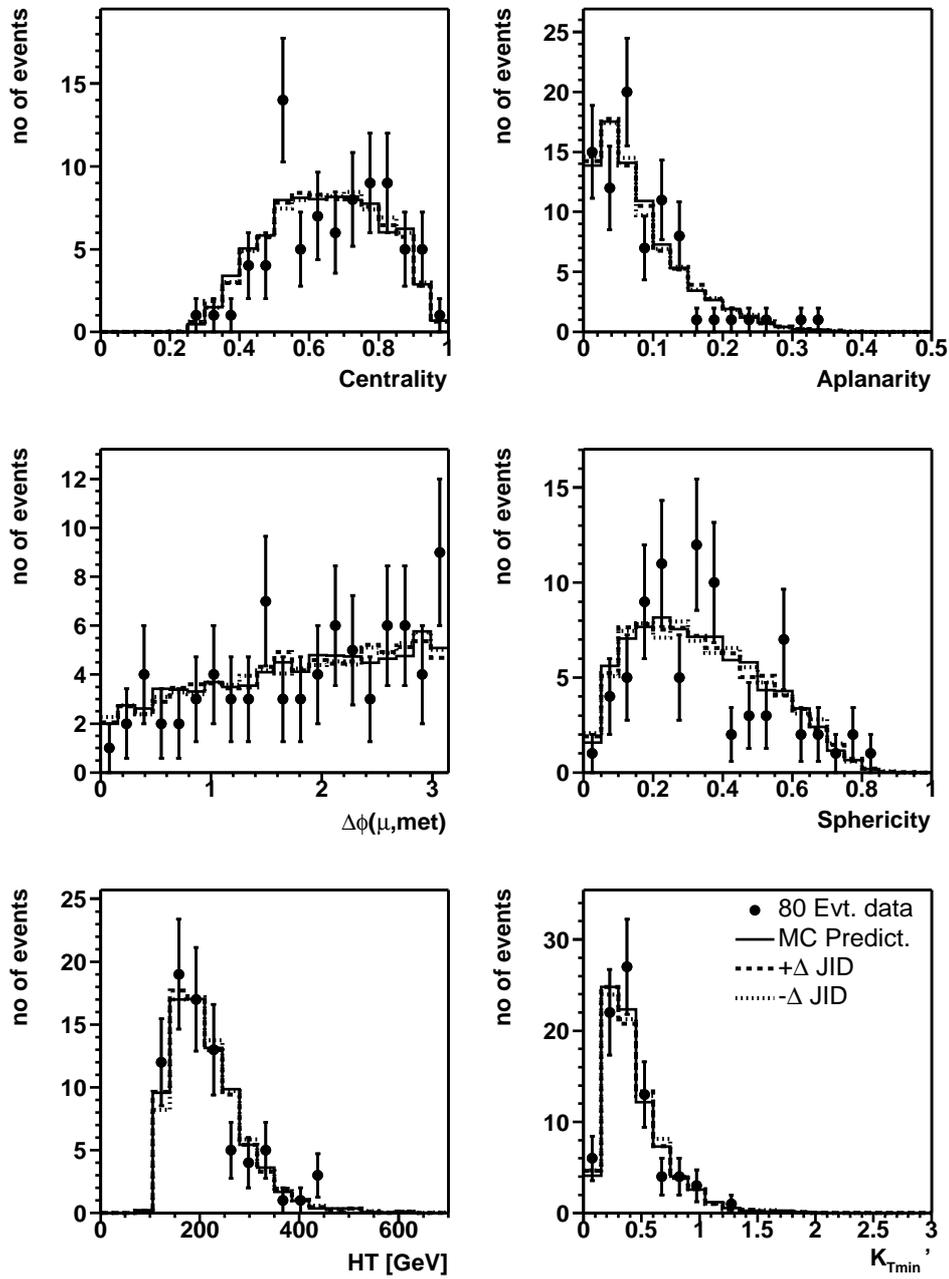


Figure A.25: Data to MC comparison including the  $\pm 1\sigma$  JID uncertainty band for the likelihood discriminant input distributions for the preselected data.

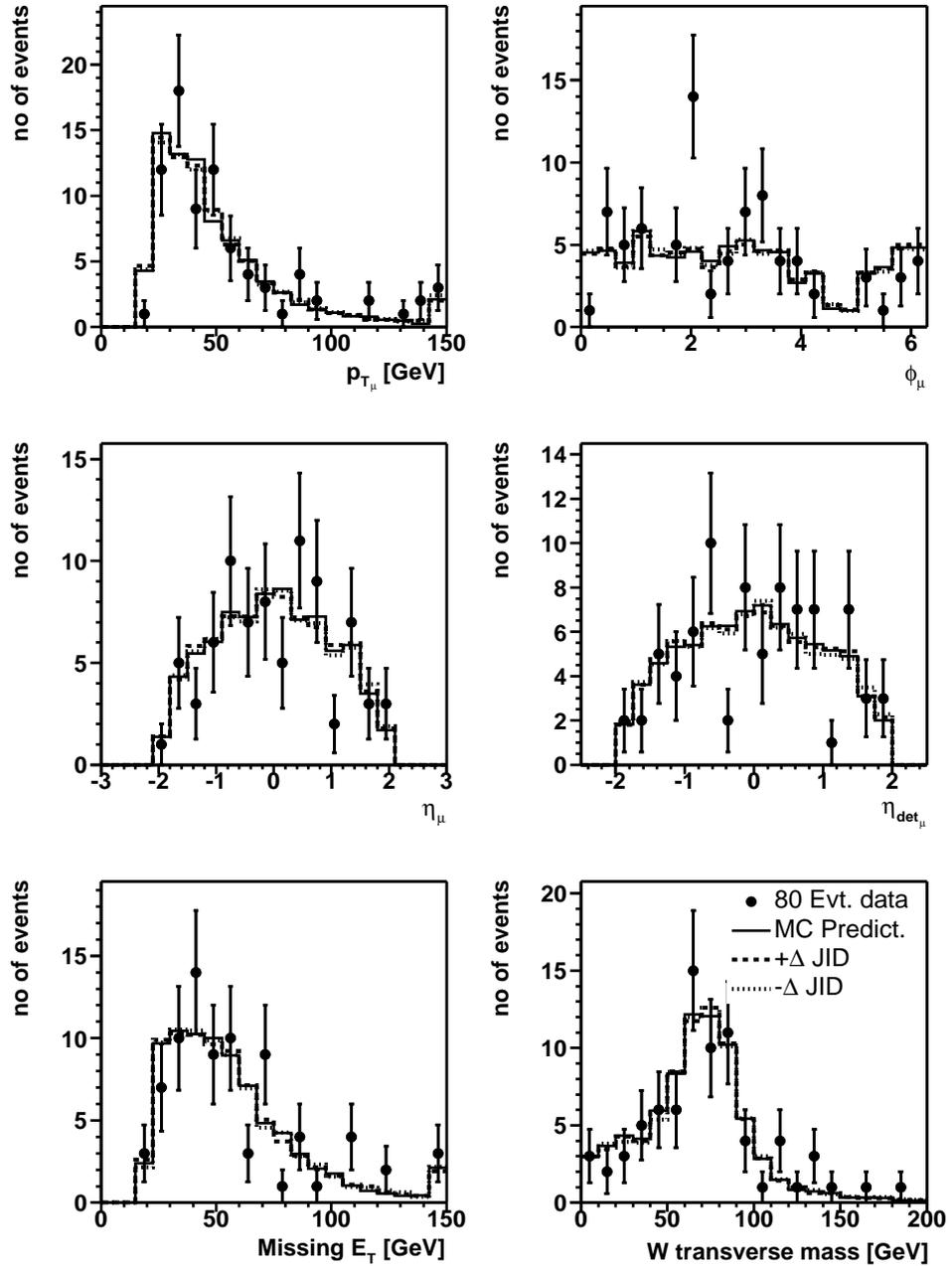


Figure A.26: Data to MC comparison including the  $\pm 1\sigma$  JID uncertainty band for  $p_T$ ,  $\phi$ ,  $\eta$ ,  $\eta_{det}$  of the muon,  $\cancel{E}_T$  and  $W$  transverse mass for the preselected data.

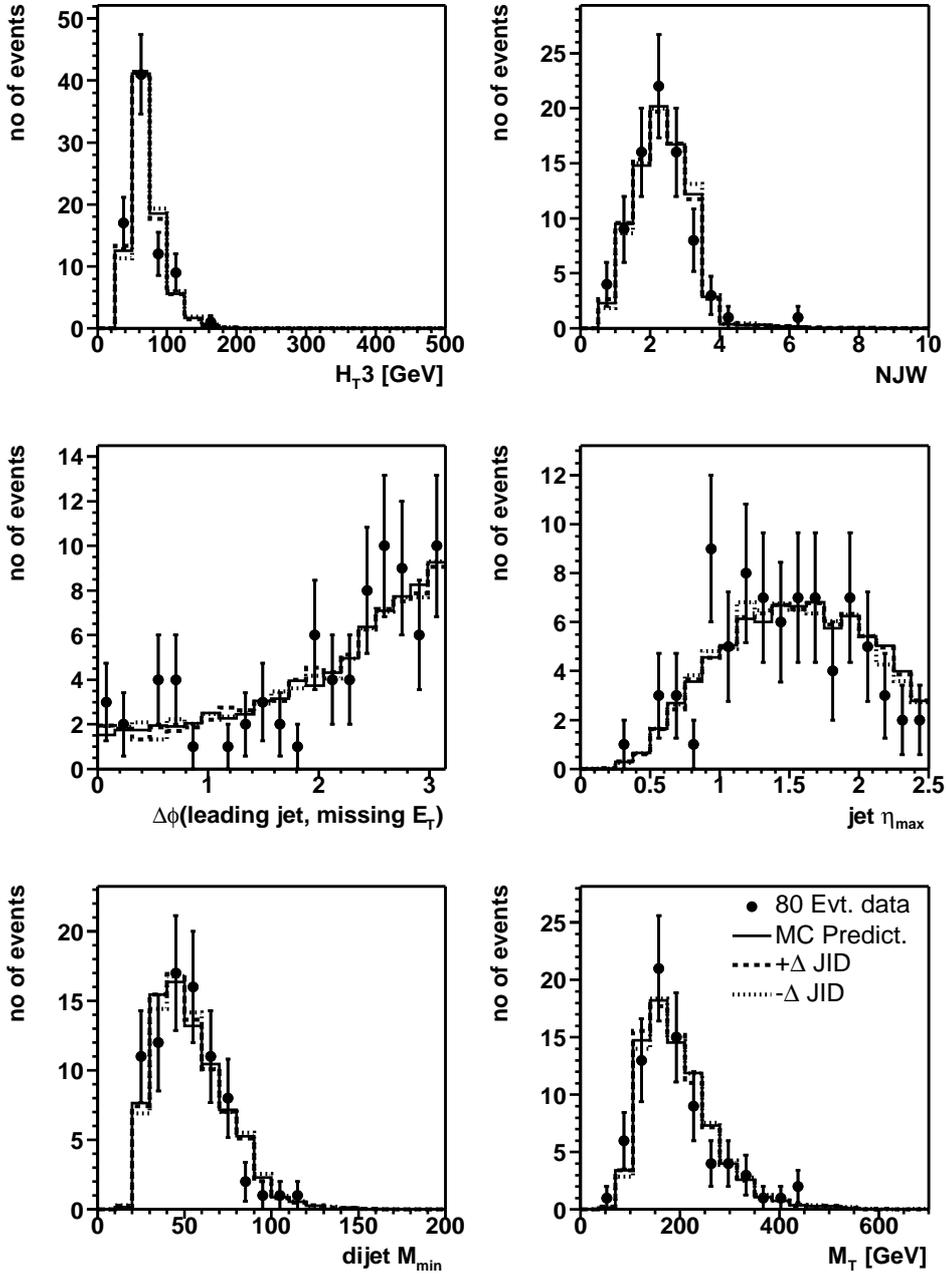


Figure A.27: Data to MC comparison including the  $\pm 1\sigma$  JID uncertainty band for the scalar sum of the  $p_T$  of the third and fourth leading jets ( $H_{T3}$ ), NJW is a topological variable built from the  $p_T$  and  $\eta$  of the four leading jets,  $\Delta\phi$  between the leading jet and  $\cancel{E}_T$ , maximum jet  $\eta$ , the minimum dijet mass, transverse mass of the 4 leading jets ( $M_T$ ) for the preselected data.

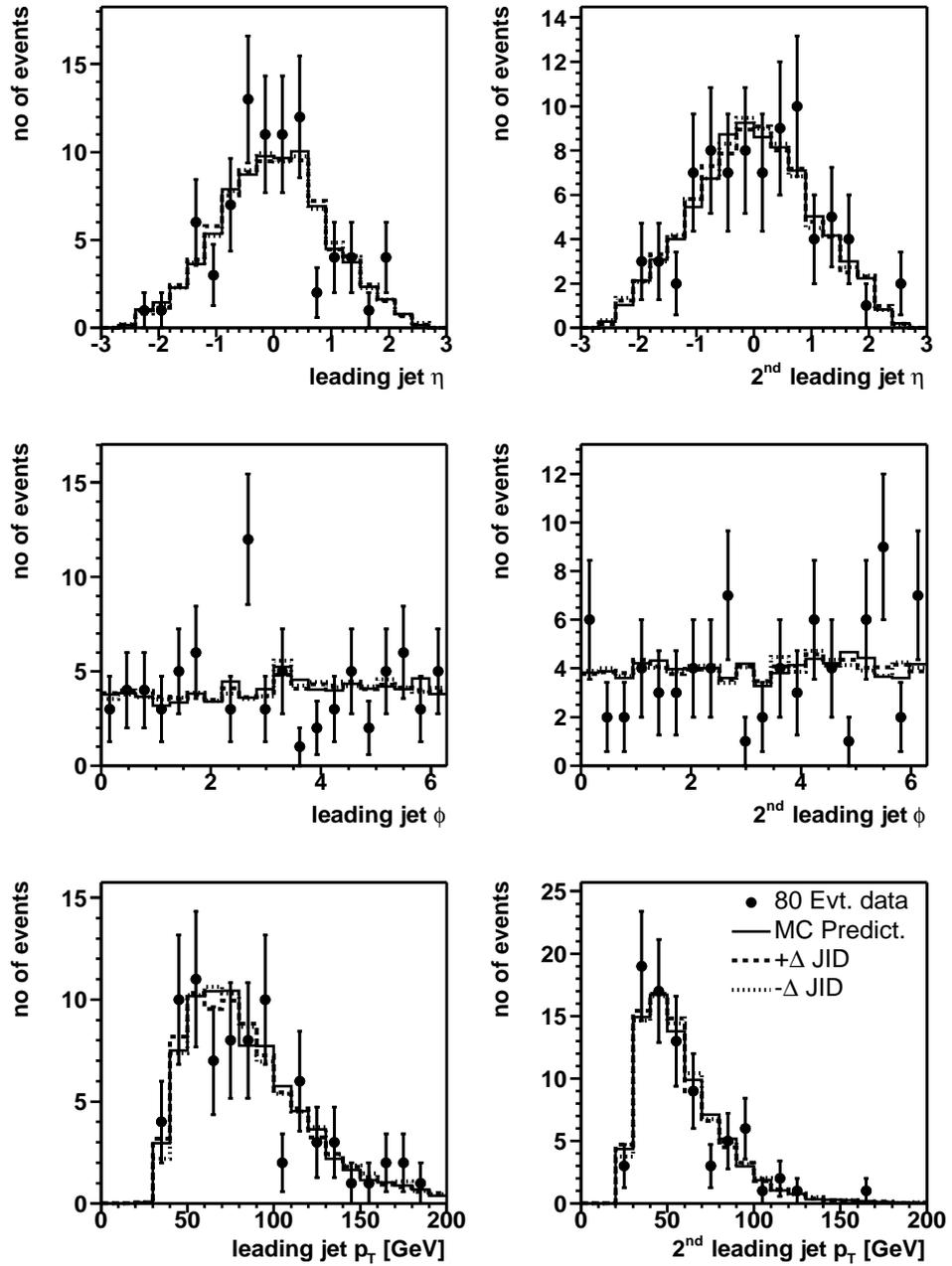


Figure A.28: Data to MC comparison including the  $\pm 1\sigma$  JID uncertainty band for  $\eta$ ,  $\phi$  and  $p_T$  of the leading and second leading jet for the preselected data.

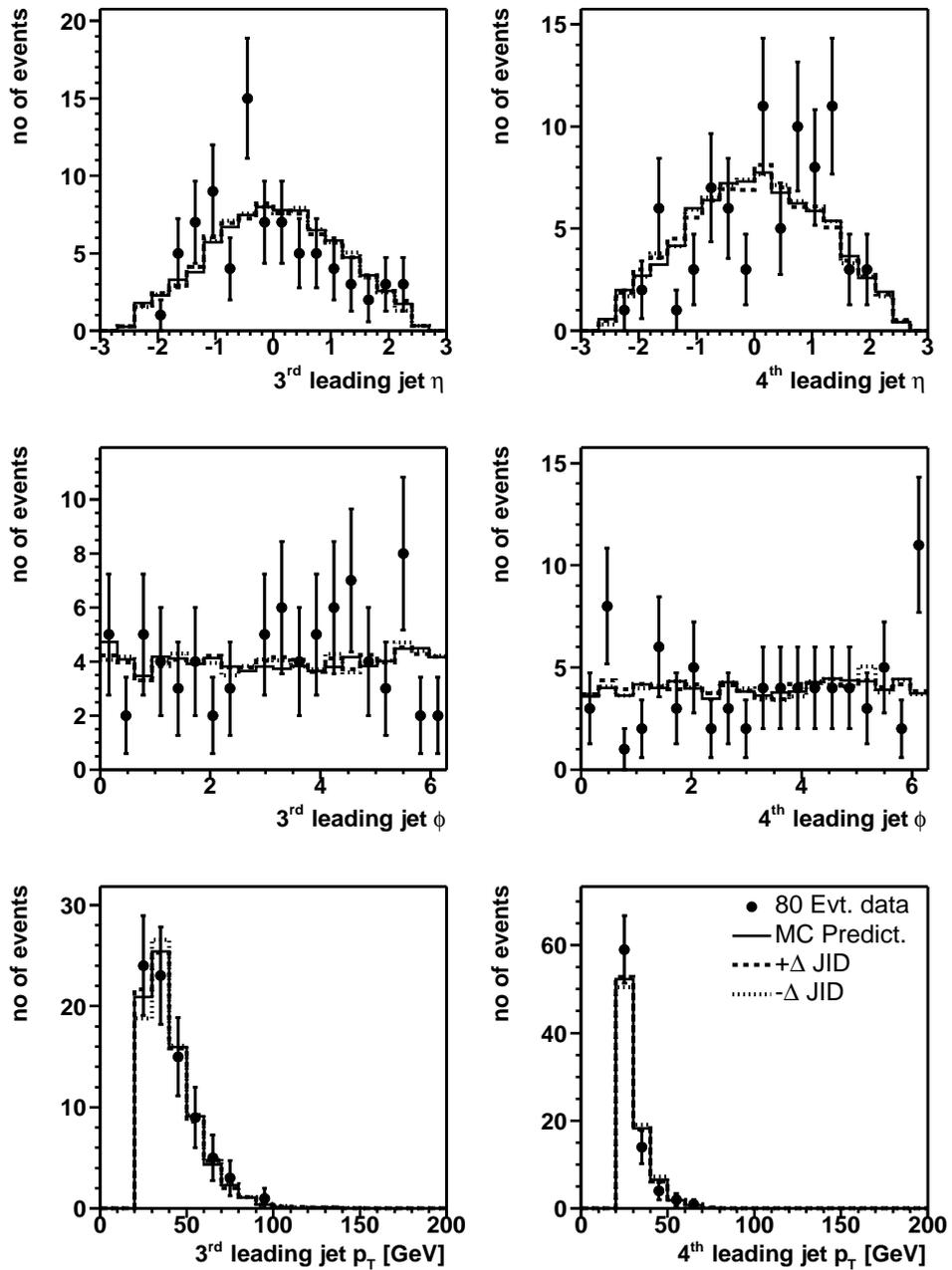


Figure A.29: Data to MC comparison including the  $\pm 1\sigma$  JID uncertainty band for  $\eta$ ,  $\phi$  and  $p_T$  of the third leading and fourth leading jet for the preselected data.



## Appendix B

# Sanity Checks for the *b*-Tagging Analysis

Typical kinematic and topological distributions are shown in Fig. B.1- B.8 for the tagged data events with one tag overlaid with the predicted signal ( $\sigma_{t\bar{t}} = 7$  pb) and backgrounds. The plots show the distributions for the  $W$  transverse mass,  $H_T$ , aplanarity, sphericity,  $\cancel{E}_T$ ,  $H_T^2$ ,  $K_T^{min}$ ,  $\Delta\phi(\text{lepton}, \text{missing } E_T)$ , lepton  $p_T$ ,  $z_{PV}$ , leading jet  $p_T$ , and  $\eta$  of the lepton, for the  $\mu$ -plus-jets and the electron-plus-jets channel, respectively, and for  $N_{\text{jet}} = 1, 2, 3, \geq 4$ . The shapes of all distributions in data are well described by the signal and background predictions

270 APPENDIX B. SANITY CHECKS FOR THE B-TAGGING ANALYSIS

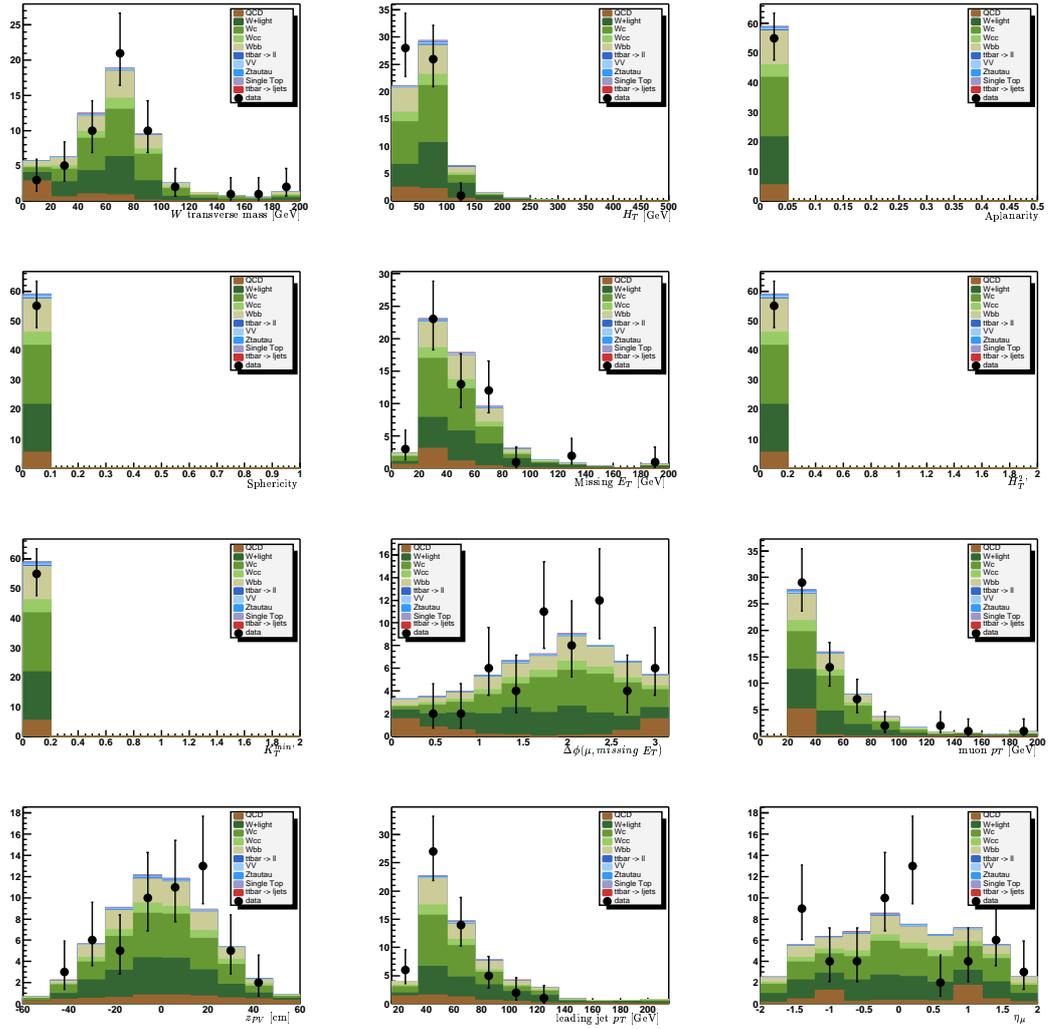
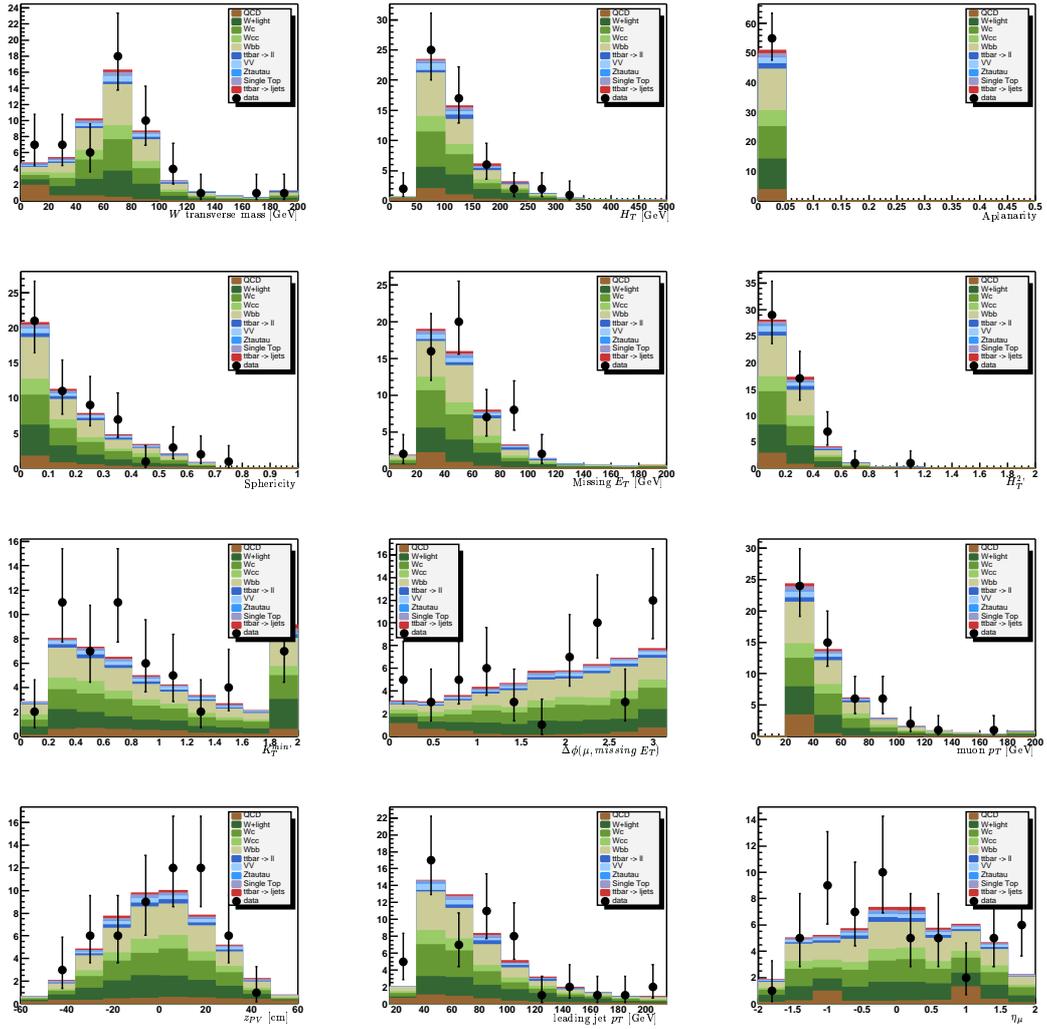


Figure B.1: Sanity checks for the  $\mu+1$  jet.

Figure B.2: Sanity checks for the  $\mu+2$  jets.

272 APPENDIX B. SANITY CHECKS FOR THE B-TAGGING ANALYSIS

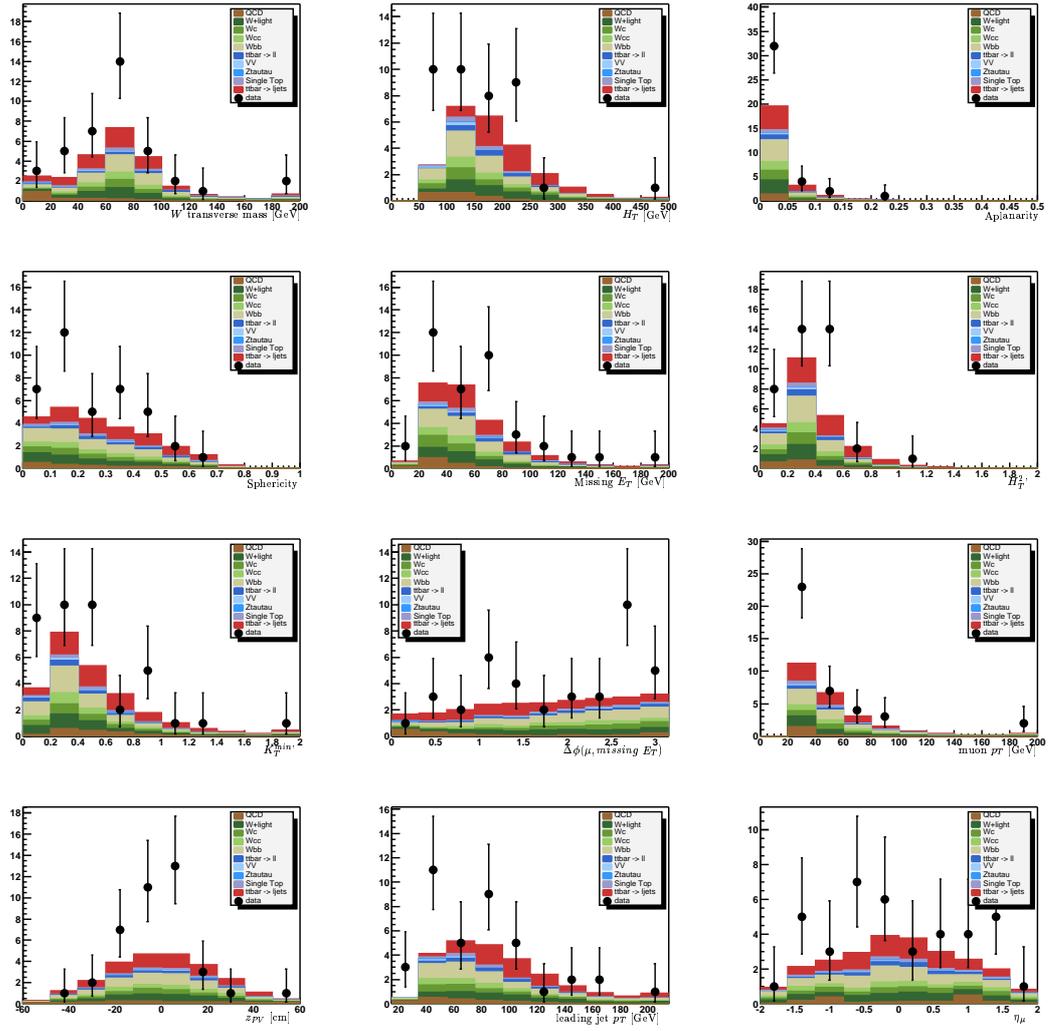
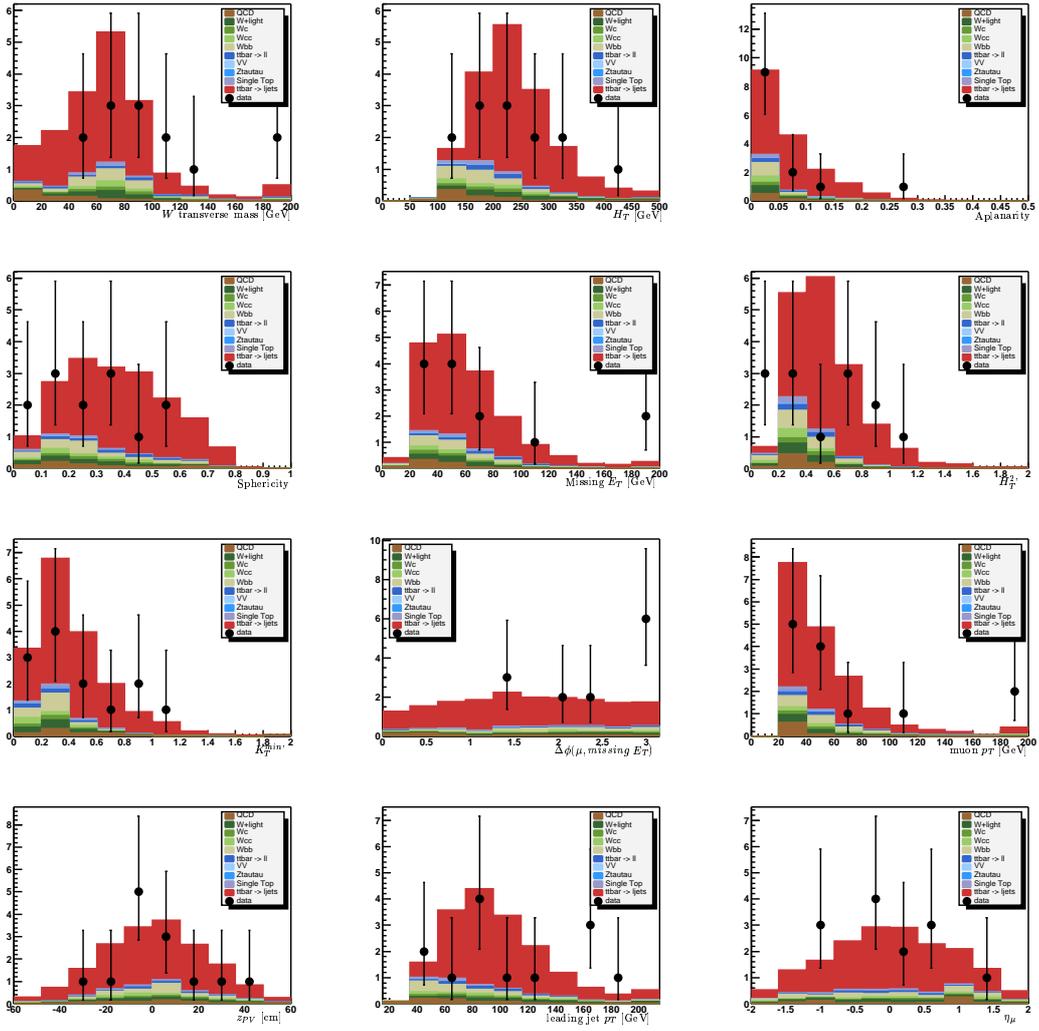


Figure B.3: Sanity checks for the  $\mu+3$  jets.

Figure B.4: Sanity checks for the  $\mu + \geq 4$  jets.

274 APPENDIX B. SANITY CHECKS FOR THE B-TAGGING ANALYSIS

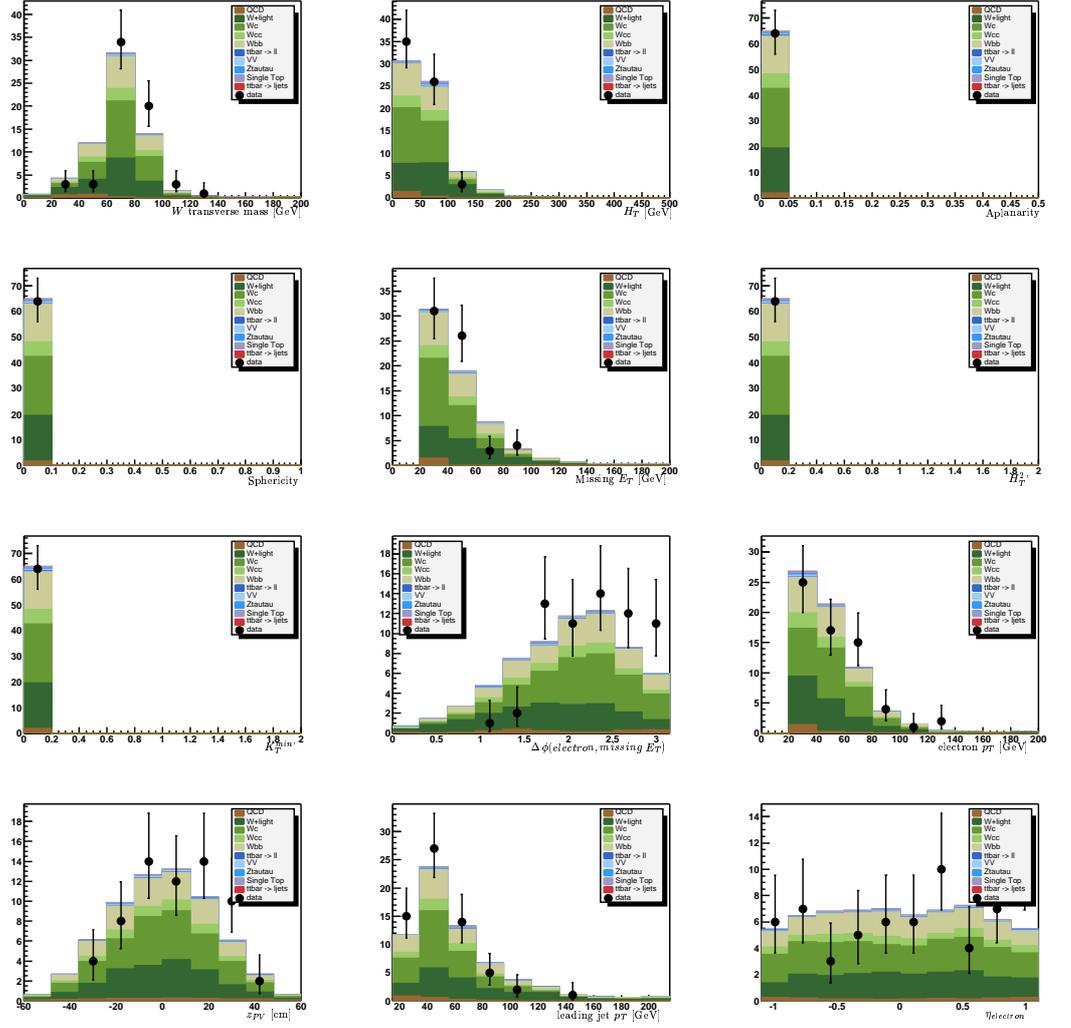
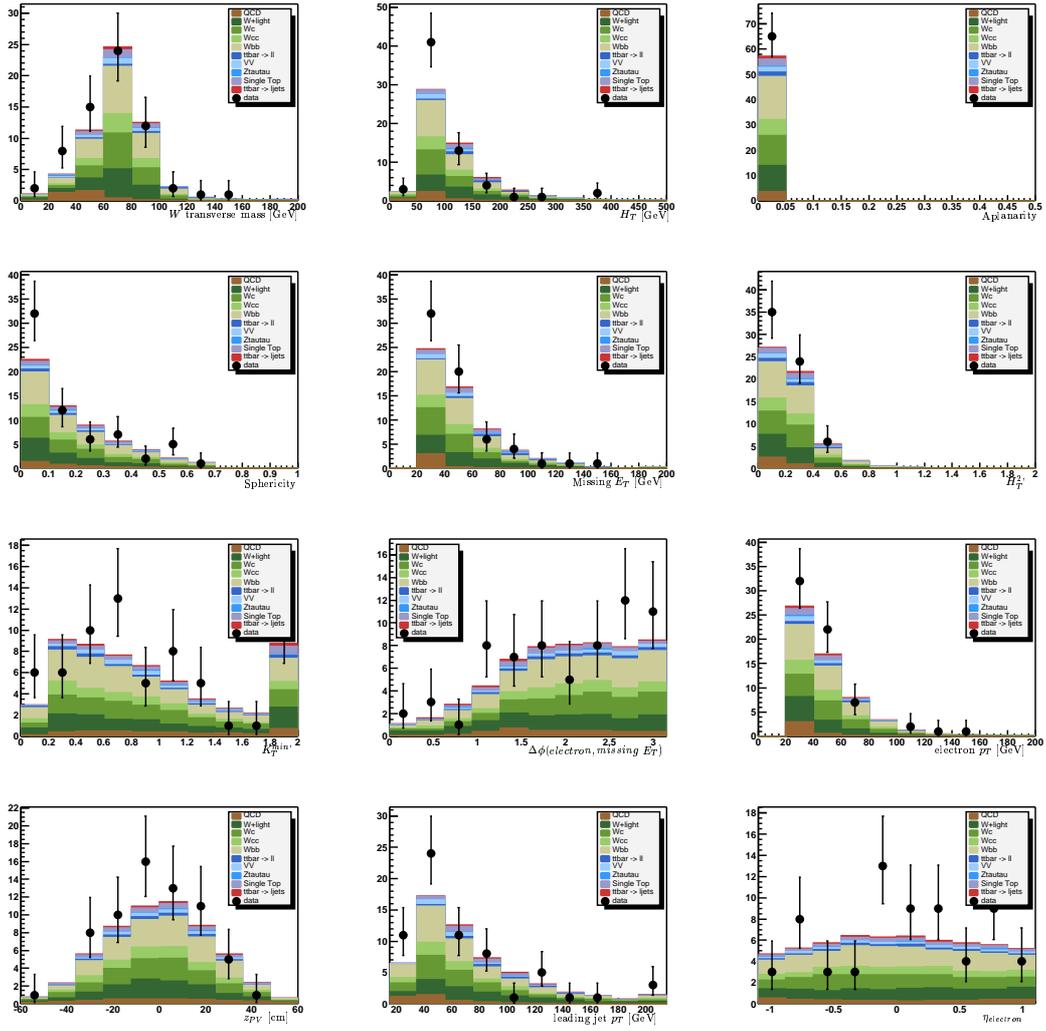


Figure B.5: Sanity checks for the  $e+1$  jet.

Figure B.6: Sanity checks for the  $e+2$  jets.

276 APPENDIX B. SANITY CHECKS FOR THE B-TAGGING ANALYSIS

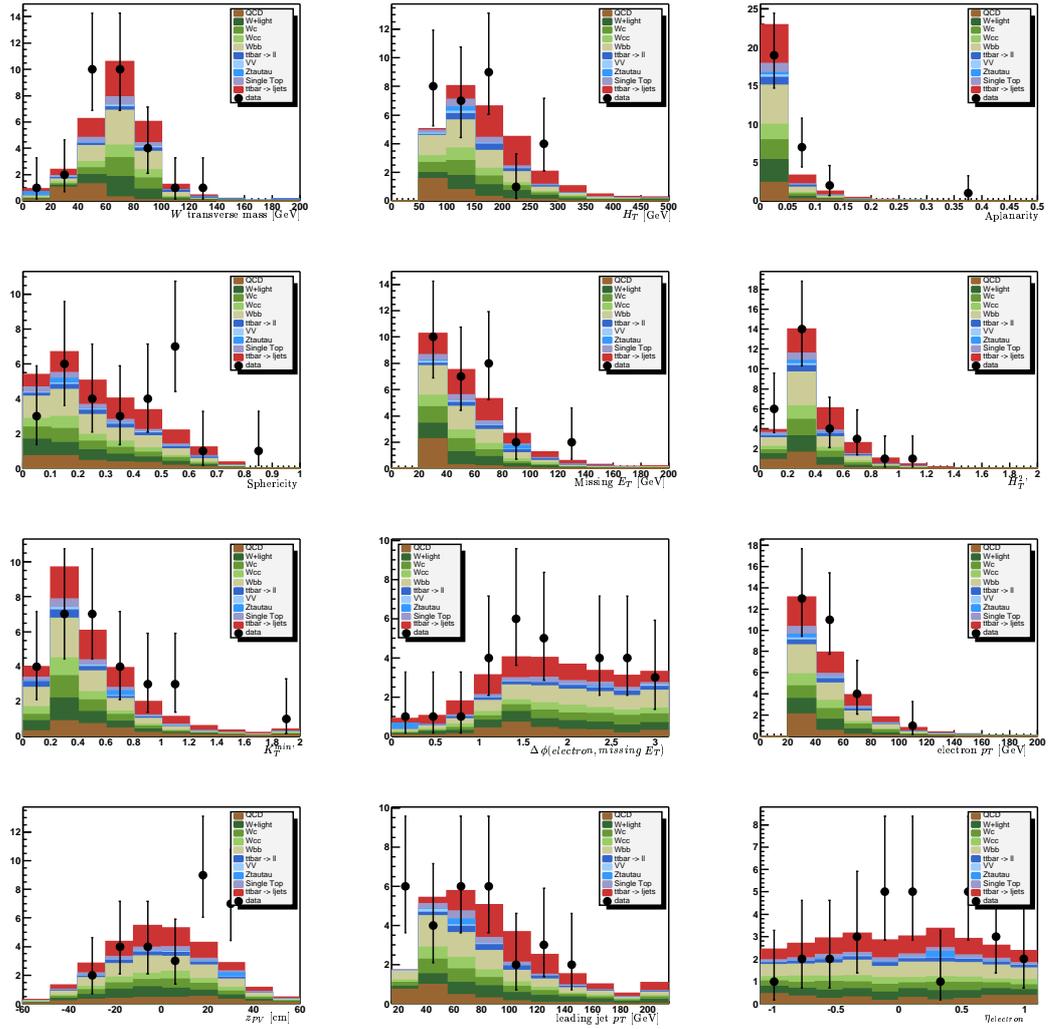
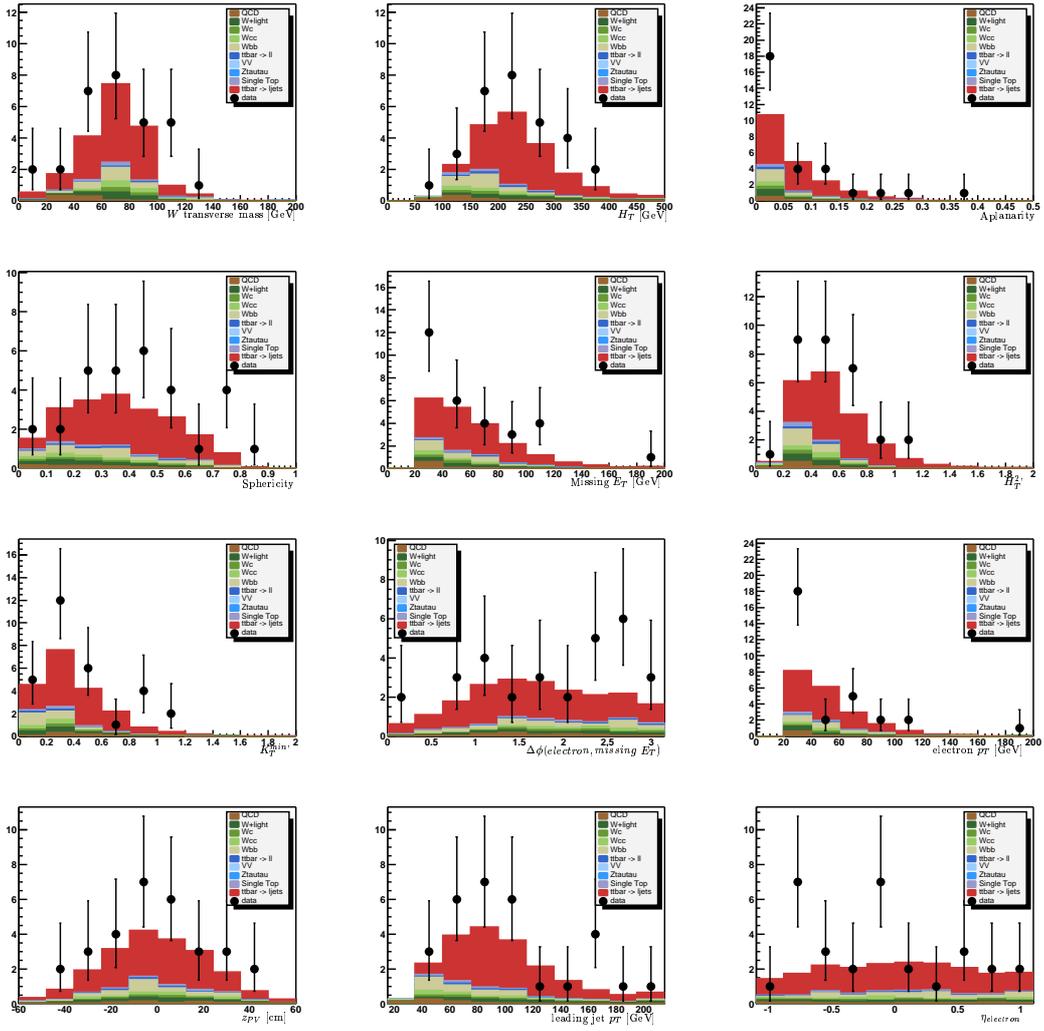


Figure B.7: Sanity checks for the  $e+3$  jets.

Figure B.8: Sanity checks for the  $e+\geq 4$  jets.



# Bibliography

- [1] J. W. Goethe, *Faust 1*, Erstausgabe, Tübingen (1808).
- [2] F. Abe *et al.* [CDF Collaboration], *Observation of the Top Quark in  $p\bar{p}$  Collisions with the Collider Detector at Fermilab*, Phys. Rev. Lett. **74**, 2626 (1995), hep-ex/9503002.
- [3] S. Abachi *et al.* [DØ Collaboration], *Observation of the Top Quark*, Phys. Rev. Lett. **74**, 2632 (1995), hep-ex/9503003.
- [4] F. Abe *et al.* [CDF Collaboration], *Study of  $t\bar{t}$  Production in  $p\bar{p}$  Collisions Using Total Transverse Energy*, Phys. Rev. Lett. **75**, 3997 (1995), hep-ex/9506006.
- [5] D. J. Gross and F. Wilczek, Phys. Rev. D **8** (1973) 3633;  
H. D. Politzer, Phys. Rept. **14** (1974) 129.
- [6] S. Abachi *et al.* [DØ Collaboration], *Update of the Measurement of the  $t\bar{t}$  cross section at  $\sqrt{s}=1.96$  TeV*, FERMI-Conf-03/248-E, July 2003
- [7] M. Klute, *A measurement of the  $t\bar{t}$  production cross section in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96$  TeV with the DØ detector at the Tevatron using final states with a muon and jets*, PhD thesis, Bonn-IR-2004-02, February 2004.
- [8] F. A. Berends, W. T. Giele, H. Kuif, R. Kleiss and W. J. Stirling, Phys. Lett. **B224**, 237 (1989).
- [9] [DØ Collaboration], *Measurement of the  $t\bar{t}$  Production Cross-Section at  $\sqrt{s} = 1.96$  TeV in the Lepton+Jets Final States using a Topological Method*, Conference DØ-Note 4423, April 2004.
- [10] E. Shabalina [For CDF and DØ Collaborations], Eur. Phys. J. C **33** (2004) S472.
- [11] T. Golling *et al.* *Measurement of the  $t\bar{t}$  production cross-section at  $\sqrt{s} = 1.96$  TeV using lifetime tagging*, DØ-Note 4141, April 2003.

- [12] [DØ Collaboration], *Measurement of the  $t\bar{t}$  production cross-section at  $\sqrt{s} = 1.96$  TeV using lifetime tagging*, Conference DØ-Note 4565, August 2004.
- [13] T. Golling *et al.*, *Measurement of the  $t\bar{t}$  production cross-section at  $\sqrt{s} = 1.96$  TeV using lifetime tagging*, DØ-Note 4625, October 2004.
- [14] M. J. G. Veltman, *Physica* **29** (1963) 186;  
G. 't Hooft, *Gauge Field Theory*, in Proceedings of the Adriatic Meeting, Rovinj (1973), ed. M. Martinis *et al.*, North Holland / Am. Elsevier, p.321.
- [15] M. K. Gaillard, P. Grannis and F. J. Sciulli, *The Standard Model of Particle Physics*, *Rev. Mod. Phys.* **71**, 96-111, (1999)
- [16] K. Kodama *et al.* [DONUT Collaboration], *Observation of tau neutrino interactions*, *Physics Letters B* 504 (2001) 218-224.
- [17] D. Griffiths, *Introduction to Elementary Particles*, Wiley, (1987).
- [18] V. D. Barger, R J. N. Phillips, *Collider Physics*, Addison-Wesley, Menlo Park, California, (1987).
- [19] S. Weinberg, *Phys. Rev. Lett.* **19** (1967) 1264;  
S. L. Glashow, *Nucl. Phys.* **22** (1961) 579.
- [20] [LEP Collaborations], *Phys. Lett. B* **276** (1992) 247.
- [21] T. Affolder *et al.* [CDF Collaboration], *Phys. Rev. Lett.* **85** (2000) 2062 [arXiv:hep-ex/0003005].
- [22] T. Affolder *et al.* [CDF Collaboration], *Phys. Rev. Lett.* **87** (2001) 102001.
- [23] V. M. Abazov *et al.*, *Phys. Rev. Lett.* **93** (2004) 011801 [arXiv:hep-ex/0404028].
- [24] S. Abachi *et al.* [DØ Collaboration], *Spin Correlation in  $t\bar{t}$  Production from  $p\bar{p}$  Collisions at  $\sqrt{s} = 1.8$  TeV*, *Phys. Rev. Lett.* 85:256 (2000).
- [25] F. Abe *et al.* [CDF Collaboration], *Kinematics of  $t\bar{t}$  events at CDF*, *Phys. Rev.* **D59**, 092001 (1999).
- [26] By CDF Collaboration and DØ Collaboration and Tevatron Electroweak Working Group, arXiv:hep-ex/0404010 (2004).
- [27] [DØ Collaboration], *Measurement of the Top Quark Mass in lepton+jets Events*, Conference DØ-Note 4574, (2004).

- [28] A. Castro [CDF and DØ Collaborations], FERMILAB-CONF-01-095-E *Prepared for 13th Italian Workshop on LEP Physics (LEP 2001), Rome, Italy, 18-20 Apr 2001.*
- [29] “ATLAS: Detector and physics performance technical design report. Volume 1,” CERN-LHCC-99-14 (1999);  
“ATLAS detector and physics performance. Technical design report. Vol. 2,” CERN-LHCC-99-15 (1999).
- [30] P. W. Higgs, *Phys. Lett.* 12 (1964) 132; *Phys. Rev. Lett.* 12 (1964) 508; *Phys. Rev.* 145 (1966) 1156; F. Englert and R. Brout, *Phys. Rev. Lett.* 13 (1964) 321; G. S. Guralnik, C. R. Hagen and T. W. B. Kibble, *Phys. Rev. Lett.* 13 (1964) 585.
- [31] D. Chakraborty, J. Konigsberg, D. Rainwater, *Review of Top Quark Physics*, hep-ph/0303092 (2003);  
S. Willenbrock, *The Standard Model and the Top Quark*, hep-ph/0211076 (2002).
- [32] T. Affolder *et al.* [CDF Collaboration], *Phys. Rev. Lett.* **84** (2000) 216 [arXiv:hep-ex/9909042].
- [33] V. M. Abazov *et al.* [DØ Collaboration], arXiv:hep-ex/0404040.
- [34] [DØ Collaboration], *Measurement of the W helicity in  $t\bar{t}$  decays at  $\sqrt{s} = 1.96$  TeV in the Lepton+jets Final States using a lifetime tag*, Conference DØ-Note 4545, (2004).
- [35] [LEP Collaboration], arXiv:hep-ex/0312023.
- [36] T. van Ritbergen and R. G. Stuart, *Phys. Rev. Lett.* **82** (1999) 488 [arXiv:hep-ph/9808283].
- [37] <http://lepewwg.web.cern.ch/LEPEWWG/plots/summer2004/>.
- [38] S. Willenbrock, “Hadron Colliders, the Standard Model and Beyond”, hep-ph/0212032 (2002).
- [39] S. Eidelman *et al.* [Particle Data Group Collaboration], *Phys. Lett. B* **592** (2004) 1.
- [40] J .C. Collins and D. E. Soper, *Ann. Rev. Nucl. Part. Sci.* **37** (1987) 383.
- [41] J. Huston *et al.*, *JHEP* **0207** (2002) 012.
- [42] A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, *Phys. Lett. B* **604** (2004) 61 [arXiv:hep-ph/0410230].

- [43] H. L. Lai *et al.*, *Eur. Phys. J. C* **12** (2000) 375.
- [44] A. M. Cooper-Sarkar, R. C. E. Devenish, A. De Roeck, *Int. J. Mod. Phys. A* **13** (1998) 3385-3586.
- [45] R. K. Ellis, W. J. Sterling and B. R. Webber, *QCD and Collider Physics*, Cambridge University Press, (1996).
- [46] P. Nason, S. Dawson and R. K. Ellis, *Nucl. Phys.* **B303** (1988) 607.
- [47] W. Beenakker, H. Kuijf, W. L. van Neerven and J. Smith, *Phys. Rev. D* **40** (1989) 54.
- [48] S. Catani, M. L. Mangano, P. Nason and L. Trentadue *Nucl. Phys.* **B478** (1996) 273-310.  
S. Catani, M. L. Mangano, P. Nason and L. Trentadue *Phys. Lett.* **B378** (1996) 329.
- [49] R. Bonciani, S. Catani, M. L. Mangano and P. Nason *Nucl. Phys.* **B529** (1998) 424-450.
- [50] A. D. Martin *et al.*, arXiv:hep-ph/0211080 (2002).
- [51] M. Cacciari, S. Frixione, G. Ridolfi, M. Mangano, P. Nason, *JHEP* **404**, 68 (2004).
- [52] N. Kidonakis and R. Vogt, *Phys. Rev. D* **68**, 114014 (2003).  
N. Kidonakis and R. Vogt, *Eur. Phys. J. C* **33**, s466 (2004).
- [53] N. Kidonakis, E. Laenen, S. Moch and R. Vogt, *Phys. Rev. D* **64** (2001) 114001 [arXiv:hep-ph/0105041].
- [54] S. Catani, arXiv:hep-ph/0005233.
- [55] V. M. Abazov *et al.* [DØ Collaboration], *Phys. Lett. B* **517** (2001) 282 [arXiv:hep-ex/0106059].
- [56] D. Acosta *et al.* [CDF Collaboration], *Phys. Rev. D* **65** (2002) 091102 [arXiv:hep-ex/0110067].
- [57] R. Schwienhorst [DØ Collaboration], arXiv:hep-ex/0411039.
- [58] D. Acosta *et al.* [CDF Collaboration], arXiv:hep-ex/0410058.
- [59] B. W. Harris, E. Laenen, L. Phaf, Z. Sullivan and S. Weinzierl, *Phys. Rev. D* **66** (2002) 054024 [arXiv:hep-ph/0207055].
- [60] A. Czarnecki and K. Melnikov, *Nucl. Phys. B* **544** (1999) 520 [arXiv:hep-ph/9806244].

- [61] T. Affolder *et al.* [CDF Collaboration], Phys. Rev. Lett. **86** (2001) 3233 [arXiv:hep-ex/0012029].
- [62] [DØ Collaboration], *Measurement of  $BR(t \rightarrow Wb)/BR(t \rightarrow Wq)$* , Conference DØ-Note 4586, (2004).
- [63] P. Merkel [for the CDF Run II collaboration], FERMILAB-CONF-04-229-E *Prepared for 6th International Conference on Hyperons, Charm and Beauty Hadrons (BEACH 2004), Chicago, Illinois, 27 Jun - 3 Jul 2004.*
- [64] F. Abe *et al.* [CDF Collaboration], *Search for Flavor-Changing Neutral Current Decays of the Top Quark in  $p\bar{p}$  Collisions at  $\sqrt{s} = 1.8$  TeV*, Phys. Rev. Lett. **80**, 2525 (1998).
- [65] S. Abachi *et al.* [DØ Collaboration], *Direct Search for Charged Higgs Bosons in Decays of Top Quarks*, Phys. Rev. Lett. **88**, 151803 (2002).
- [66] The Top Production Subgroup, *Measurement of the  $t\bar{t}$  production cross section at  $\sqrt{s} = 1.96$  TeV in dilepton final states*, (2004), linked from [http://www-d0.fnal.gov/Run2Physics/top/private/internal\\_PubsReview.html](http://www-d0.fnal.gov/Run2Physics/top/private/internal_PubsReview.html).
- [67] F. Blekman *et al.*, *Measurement of the  $t\bar{t}$  production cross section in the all-jets channel, using Secondary Vertex Tagging*, DØ-Note 4346, (2004).
- [68] S. I. Alekhin, arXiv:hep-ph/0307219.
- [69] V. M. Abazov *et al.* [DØ Collaboration], arXiv:hep-ex/0410066.
- [70] J. M. Campbell and R. K. Ellis, Phys. Rev. D **60** (1999) 113006 [arXiv:hep-ph/9905386].
- [71] M. L. Mangano *et al.*, *ALPGEN, a Generator for Hard Multiparton Processes in Hadronic Collisions*, CERN-TH-2002-129, FTN-T-2002-06, hep-ph/0206293 (2002).
- [72] T. Affolder *et al.* [CDF Collaboration], Phys. Rev. D **65** (2002) 092002.
- [73] T. Sjöstrand, L. Lonnblad and S. Mrenna, *PYTHIA 6.2: PHYSICS AND MANUAL*, LU-TP-01-21, hep-ph/0108264 (2001).
- [74] G. Marchesini *et al.*, *HERWIG: A Monte Carlo event generator for simulating hadron emission reactions with interfering gluons. Version 5.1 - April 1991*, Comp. Phys. Comm. **67** (1992).
- [75] R. Brun *et al.*, *Simulation program for particle physics experiments, GEANT: user guide and reference manual*, CERN DD 78-2 (1978).

- [76] M. Klute, L. Phaf, D. Whiteson, *TopAnalyze - A Framework Analyze Package for Top Group Analyses*, DØ-Note 4122 (2003).
- [77] S. Anderson, E. Busato, M. Klute, S. Kulik, J. Strandberg, *top\_analyze documentation*, webpage, [http://www-d0.fnal.gov/Run2Physics/top/d0\\_private/wg/top\\_analyze/top\\_analyze\\_documentation.html](http://www-d0.fnal.gov/Run2Physics/top/d0_private/wg/top_analyze/top_analyze_documentation.html)
- [78] D. J. Lange, Nucl. Instrum. Meth. A **462** (2001) 152.
- [79] Z. Was and P. Golonka, arXiv:hep-ph/0411377.
- [80] F. Krauss, R. Kuhn and G. Soff, J. Phys. G **26** (2000) L11 [arXiv:hep-ph/9904274].
- [81] S. Catani, F. Krauss, R. Kuhn and B. R. Webber, JHEP **0111** (2001) 063 [arXiv:hep-ph/0109231].
- [82] M. Mangano. Talk given at the Matrix Element and Monte Carlo Tuning Workshop, Fermilab, Nov 16, 2002.  
<http://cepa.fnal.gov/personal/mreenna/tuning/nov2002/mlm.pdf>
- [83] A. Pukhov *et al.*, arXiv:hep-ph/9908288.
- [84] S. Klimenko, J. Konigsberg, T. M. Liss, FERMILAB-FN-0741 (2003).
- [85] T. Edwards *et al.*, FERMILAB-TM-2278-EB (2004).
- [86] W. T. Giele and S. A. Keller, arXiv:hep-ph/0104053.
- [87] [DØ Collaboration], *Measurement of  $Z \rightarrow ee$  and  $W \rightarrow e\nu$  Production Cross Sections with  $|\eta| < 2.3$* , Conference DØ-Note 4403, (2004).
- [88] [DØ Collaboration], *Measurement of the Cross section for Inclusive Z Production in Di-muon Final States at  $\sqrt{s} = 1.96$  TeV*, Conference DØ-Note 4573, (2004).
- [89] Fermilab Beams division, “RunII Handbook”, Fermilab Beams division RunII webpage, <http://www-bd.fnal.gov/runII/index.html>
- [90] Fermilab Beams division, Fermilab Recycler Ring Technical Design Report, Fermilab-TM-1991, (1996).
- [91] S. Abachi *et al.* [DØ Collaboration], *The DØ Detector*, NIM, A338 (1994) 185.
- [92] S. Abachi *et al.* [DØ Collaboration], *The DØ Upgrade: The Detector and its Physics*, FERMILAB-PUB-96/357-E (1996).
- [93] V. Abazov *et al.* [DØ Collaboration], to be submitted to Nucl. Instrum. Meth. A.  
T. LeCompte and H. T. Diehl, Ann. Rev. Nucl. Part. Sci. **50**, 71 (2000).

- [94] C-C.Miao, FERMILAB-CONF-98-395-E (1998).
- [95] [DØ Collaboration] *The DØ Silicon Microstrip Vertex Detector*, to be submitted to Nucl. Instrum. Meth.
- [96] E. Kajfasz, *DØ Silicon Microstrip Tracker for RunIIa*, hep-ex/0112014 (2001).
- [97] T. Zimmermann *et al.*, *The SVXII Readout chip*, IEEE Trans. Nucl. Sci. **NS42** (1995) 803.
- [98] J. Brzenziak *et al.*, FERMILAB-TM-1886 (1994).
- [99] R. J. Hooper and G. Landsberg, *Search for Large Extra Spatial Dimensions in the Di-Muon Channel with 100 pb<sup>-1</sup> of RunII data*, DØ-Note 4230, August 2003.
- [100] A. Khanov, private communication.
- [101] P. Baringer *et al.* [DØ Collaboration], Nucl. Instrum. Meth. A **469** (2001) 295 [arXiv:hep-ex/0007026].
- [102] M. Adams *et al.*, *Design Report for the Central Preshower Detector for the DØ Upgrade*, DØ-Note 3014, January 1996.
- [103] A. Gordeev *et al.*, *Design Report for the Forward Preshower Detector for the DØ Upgrade*, DØ-Note 3445, May 1998.
- [104] L. Groers, *DØ Calorimeter Upgrades for Tevatron Run II*, DØ-Note 4240, Proceedings for the IXth International Conference on Calorimetry in Particle Physics, Annecy, France, Oct 9-14, 2000.
- [105] B. Abbott *et al.* [DØ Collaboration], Nucl. Instrum. Meth. A **424** (1999) 352 [arXiv:hep-ex/9805009].
- [106] S. Abachi *et al.* [DØ Collaboration], Nucl. Instrum. Meth. A **324** (1993) 53, FERMILAB-PUB-92/162-E.
- [107] B. Abbott *et al.* [DØ Collaboration], Phys. Rev. D **58** (1998) 092003 [arXiv:hep-ex/9712029].
- [108] J. Stark, private communication.
- [109] T. Diehl *et al.*, *Technical Design of the Central Muon System*, DØ-Note 3365, March 1998.
- [110] T. Diehl *et al.*, *Technical Design for the DØ Forward Muon Tracking Detector Based on Mini-Drift Tubes*, DØ-Note 3366, December 1997.

- [111] T. Diehl *et al.*, *Technical Design Report for the  $D\bar{O}$  forward trigger scintillator counters*, D $\bar{O}$ -Note 3237, November 1987.
- [112] J. Buttler, *Local Muon Momentum Resolution*, D $\bar{O}$ -Note 4002, July 2002.
- [113] [D $\bar{O}$  Collaboration], *Reconstruction of B Hadron Signals at D $\bar{O}$* , Conference D $\bar{O}$ -Note 4481, (2004).
- [114] A. Brandt *et al.*, Fermilab-Pub-97-377 (1997).
- [115] U. Amaldi *et al.*, Phys. Lett. B **43** (1973) 231.
- [116] J. Barreto *The D $\bar{O}$  Forward Proton Detector (FPD) Status*, D $\bar{O}$ -Note 4285, October 2003.
- [117] R. E. Kalman, J. Bas. Eng. **82D** 35 (1960);  
R. E. Kalman and R. S. Brucy, J. Bas. Eng. **83D** 95 (1961);  
P. Billoir, Nucl. Instrum. Meth. A **225** (1984) 352.
- [118] A. Garcia-Bellido *et al.*, *Primary Vertex certification in p14*, D $\bar{O}$ -Note 4320, January 2004.
- [119] A. Schwartzman and M. Narain, *Probabilistic Primary Vertex Selection*, D $\bar{O}$ -Note 4042, November 2002.
- [120] T. Golling *et al.*, *p14 Muon-ID certification note*, D $\bar{O}$ -Note 4350, February 2004.
- [121] Emily Nurse and Paul Telford, *Measurement of cross section times branching ratio for  $Z \rightarrow \mu\mu^-$  in  $p\bar{p}$  collisions at 1.96 TeV*, D $\bar{O}$ -Note 4231, August 2003.
- [122] Top Physics Working Group, *D $\bar{O}$  Top Analyses and Data Sample for the Winter Conference 2004*, D $\bar{O}$ -Note 4419, April 2004.
- [123] Chakravarti, Laha, and Roy, *Handbook of Methods of Applied Statistics*, Volume I, John Wiley and Sons, pp. 392-394, (1967).
- [124] M. Klute, *Measurement of the  $t\bar{t}$  cross section at  $\sqrt{s} = 1.96$  TeV in the muon-plus-jets channel*, D $\bar{O}$ -Note 4185, July 2003.
- [125] D. Whiteson and M. Kado, *Muon Isolation Studies*, D $\bar{O}$ -Note 4070, December 2002.
- [126] S-J. Park *et al.*, *Electron Likelihood in p14*, D $\bar{O}$ -Note 4449, May 2004.
- [127] G. Blazey *et al.*, *Run II Jet Physics*, D $\bar{O}$ -Note 3750, April 2000.

- [128] U. Bassler and G. Bernardi, *Towards a Coherent Treatment of Calorimetric Energies: Missing Transverse Energy, Jets, E.M. Objects and the T42 Algorithm*, DØ-Note 4124, March 2003.
- [129] J-R. Vlimant, U. Bassler, G. Bernardi and S. Trincaz-Duvoid, *Technical description of the T42 algorithm for the calorimeter noise suppression*, DØ-Note 4146, May 2003.
- [130] G. Bernardi, E. Busato and J-R. Vlimant, *Improvements from the T42 Algorithm on Calorimeter Objects Reconstruction*, DØ-Note 4335, January 2004.
- [131] T. Golling, presentation in the CALGO meeting, November 4<sup>th</sup> 2003, <http://www-d0.hef.kun.nl//askArchive.php?base=agenda&category=a03543&id=a03543s1t7/transparencies>.
- [132] S. Anderson *et al.*, *Measurement of the  $t\bar{t}$  cross section at  $\sqrt{s} = 1.96$  TeV*, DØ-Note 4116, July 2003.
- [133] Calgo/jetid certification web page, [http://www-d0.fnal.gov/computing/algorithms/calgo/jet/jetID\\_p14.html](http://www-d0.fnal.gov/computing/algorithms/calgo/jet/jetID_p14.html).
- [134] DØ Jet Energy Scale study group, [http://www-d0.fnal.gov/phys\\_id/jes/d0\\_private/certified/v5.1/links.html](http://www-d0.fnal.gov/phys_id/jes/d0_private/certified/v5.1/links.html).
- [135] A. Kumar *et al.*, *Oversmearing of Missing Transverse Energy in  $Z \rightarrow ee + X$  Monte Carlo Events*, DØ-Note 4551, July 2004.
- [136] A. Schwartzman and M. Narain, *Secondary Vertex Reconstruction using the Kalman Filter*, DØ-Note 3908, September 2001.
- [137] DØ  $b$ -ID web page, [http://www-d0.fnal.gov/phys\\_id/bid/d0\\_private/bid.html](http://www-d0.fnal.gov/phys_id/bid/d0_private/bid.html).
- [138] T. Christiansen, *The D-Zero L2-Muon Trigger Performance for P11*, DØ-Note 4053, November 2003.
- [139] R. Schwienhorst, *Top Trigger Selection and Application of Turn-On Curves to the Monte Carlo*, DØ-Note 4508, July 2004.
- [140] M. Klute and A. Quadt, *Measurements of Level 1 Trigger Efficiencies from DØ Data*, DØ-Note 3949, March 2002.
- [141] M. Agelou *et al.*, *Top Trigger Efficiency Measurement and the top\_trigger package*, DØ-Note 4512, July 2004.
- [142] Common Sample Group (CSG), <http://www-d0.fnal.gov/Run2Physics/cs/index.html>.

- [143] M. Klute *et al.*, *Top group ROOT tuples selection and Data Quality Monitoring*, DØ-Note 3942, February 2002.
- [144] The DØ Run Quality Database,  
<http://d0db-prd.fnal.gov/run/runQuery.html>.
- [145] V. Abazov, T. Diehl and R. McCroskey, *Good and Bad Muon Global Runs Early in Run II*, DØ-Note 3938, April 2002.
- [146] K. Ranjan *et al.*, *Calorimeter Event Quality Using Level 1 Confirmation*, DØ-Note 4554, July 2004.
- [147] Calgo group web page,  
<http://www-d0.fnal.gov/computing/algorithms/calgo/calgo.html>.
- [148] T. Golling *Jet Shape Variables*, talk in the top production meeting, July 19.2004.
- [149] C. Gerber *et al.*, *Measurement of the  $t\bar{t}$  Production Cross-Section at  $\sqrt{s} = 1.96$  TeV in the Electron+Jets Final State using a Topological Method*, DØ-Note 4662, December 2004.
- [150] N. Kidonakis, private communication.
- [151] A. Quadt *et al.* DØ-Note in preparation.
- [152] Top Physics Working Group, *Measurement of the  $t\bar{t}$  production cross-section in the lepton+jets channels at  $\sqrt{s} = 1.96$  TeV (topological)*, DØ-Note 4422, April 2004.
- [153] B. Clément *et al.*, *SystemD or how to get signal, backgrounds and their efficiencies with real data*, DØ-Note 4159, March 2003.
- [154] V. M. Abazov *et al.* [DØ Collaboration], arXiv:hep-ex/0412020, submitted to Phys. Rev. Lett.
- [155] T. Golling *Measurement of the top quark pair production cross section in the lepton+jets channel*, DØ Collaboration Week, Dec.10.2004.  
<http://www-d0.hef.kun.nl/fullAgenda.php?ida=a042001>.
- [156] A. Juste *Parton matching and modeling issues*, DØ Top Quark Mass Workshop, Dec.4.2004.  
<http://www-d0.hef.kun.nl///fullAgenda.php?ida=a041955>.
- [157] J. Campbell and J. Huston, arXiv:hep-ph/0405276.
- [158] S. Anderson *et al.*, *Measurement of the  $t\bar{t}$  Production Cross-section at a center of mass energy of 1.96 TeV in Dilepton Final States*, DØ-Note 4623, October 2004.

- [159] F. Beaudette and J.-F. Grivaz, *Measurement of the  $t\bar{t}$  production cross section in the lepton+jets channels with soft muon tag*, DØ-Note 4206, July 2003.  
E. Barberis *et al.*, *Combined Results for the  $t\bar{t}$  Cross Section Measurement*, DØ-Note 4246, September 2003.
- [160] C. Clément *et al.*, *Measurement of the  $t\bar{t}$  Production Cross-section at  $\sqrt{s} = 1.96$  TeV in the  $e\mu$  Channel Using Secondary Vertex  $b$ -tagging*, DØ-Note 4528, July 2004.
- [161] D. Acosta *et al.* [CDF Collaboration], *Phys. Rev. Lett.* **93** (2004) 142001 [arXiv:hep-ex/0404036].
- [162] [CDF Collaboration], *A global analysis of the high- $p_T$  dilepton sample using 200  $pb^{-1}$  of Run II data*, Conference CDF-Note 7192.
- [163] [CDF Collaboration], *Measurement of the cross section for  $t\bar{t}$  production using event kinematics in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96$  TeV*, Conference CDF-Note 7154.
- [164] [CDF Collaboration], *Measurement of the  $t\bar{t}$  production cross section in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96$  TeV using lepton plus jets events with soft muon  $b$ -tagging*, Conference CDF-Note 7174.
- [165] [CDF Collaboration], *Measurement of the  $t\bar{t}$  production cross section in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96$  TeV using lepton+jets events with secondary vertex  $b$ -tagging*, Conference CDF-Note 7138, Fermilab-Pub-04-275-E, submitted to *Phys. Rev. D*.
- [166] [CDF Collaboration], *Measurement of the  $t\bar{t}$  production cross section in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96$  TeV using kinematic fitting of  $b$ -tagged lepton+jets events*, Conference CDF-Note 7087, submitted to *Phys. Rev. D*.
- [167] [CDF Collaboration], *Measurement of the  $t\bar{t}$  production cross section in the jet probability tagged sample in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96$  TeV*, Conference CDF-Note 7236.
- [168] [CDF Collaboration], *Measurement of the  $t\bar{t}$  production cross section in the in the all-hadronic channel*, Conference CDF-Note 7075.



# Acknowledgments

This page is devoted to all who made my dissertation possible, who supported and helped me arriving finally and happily at this last page.

I am very thankful to my supervisor Prof. Dr. Norbert Wermes, for giving me the opportunity to work within the DØ collaboration at FERMILAB, for his reliability, for giving me “plenty of rope” and at the same time being there whenever it was necessary.

I would like to thank Prof. Dr. Erwin Hilger for being so kind to take over the *burden* of coreferee, reading and correcting.

I would like to express my gratitude to Dr. Arnulf Quadt who made sure that the past three years have been a great and intense experience for me - it has been a pleasure and an honor and I am looking forward to collaborating (*work hard, play hard*) in the near and the long future with you. Not to forget, I owe a debt of gratitude for steady proofreading, inspiring conversations and sagacious advice.

I am indebted to Dr. Aurelio Juste who opened my eyes over and over again to the very subtleties of physics without losing the view of the big picture, thank you for shaping me and my respect for the complete physicist you are.

I am deeply grateful to all being interested in my analyses, sharing their vast experience with me and inspiring me, in particular Prof. Dr. Tom Ferbel and Prof. Dr. Mark Strovink, and the convenors of the top groups Prof. Dr. Ela Barberis, Dr. Christophe Clement, Prof. Dr. Cecilia Gerber, Dr. Marumi Kado, Prof. Dr. Bob Kehoe, Dr. Flera Rizatdinova, Dr. Lisa Shabalina and Dr. Chris Tully.

Many thanks to all who contributed to make this dissertation a very pleasurable experience - not only in terms of physics: the indescribable Prof. Dr. Gregorio Bernardi *et son savoir vivre*, Dr. Catherine Biscarat *pour le refuge*, Prof. Dr. Martin Erdmann for the initiation and the advice, my co-DJane Dr. *BeatE* Heinemann, Dr. Markus Klute for sharing some long night, Dr. Christian Schwanenberger to kick off my back-up career, Dr. Jan Stark for his challenging critical viewpoint (*der Geist, der stets verneint*), Jonas Strandberg for the smooth collaboration, meine *Bespaßerin* Dr. Julia Thom, and the DØ-Bonn group Dr. Jochen Cammin, Kevin Kröniger, Jörg Meyer, Teresa Negrini, Su-Jung Park, Markus Warsinsky, and Dr. Marc-

Andre Pleier.

I am very grateful to the Graduiertenförderung des Landes Nordrhein Westfalen, the Heinrich Hertz-Stiftung and the German academic exchange service (DAAD) for the financial support during my dissertation.

I would like to thank my family, particularly my parents for providing me with the essential *Urvertrauen* and for reminding me, together with my brother, my sister and all my friends who have not forgotten me, of the bearable lightness of being whenever I needed it.

Finally, to all the people I did not mention: "I have not forgotten you, vielen Dank, thanks a lot, merci beaucoup, muchas gracias, spasibo!"