Measurement of the $t\bar{t}$ production cross-section in the $\tau + jets$ final state at $\sqrt{s} = 7$ TeV with the ATLAS detector at the LHC

Maria Chiara Conidi

Institut de Física d Altes Energies
Universitat Autònoma de Barcelona
Departament de Física
Facultat de Ciències
Edifici Cn E-08193 Bellaterra (Barcelona)

January 2013

Supervised by: Martine Bosman
Institut de Física d Altes Energies
Edifici Cn E-08193 Bellaterra (Barcelona)

Tutored by: Maria Pilar Casado
Universitat Autònoma de Barcelona

Introduction

The top quark was the last elementary particle of the Standard Model to be discovered, before the recent Higgs boson candidate. The top quark discovery in 1995 at Tevatron $p\bar{p}$ collider at Fermilab allowed to complete the third generation of quarks. From the beginning it was clear that the top quark played a very special role in the Standard Model theory, due to its large mass of ∼173.5 GeV [1]. Its interest resides in its potential link to the Electroweak Symmetry Breaking and the mechanism that generates the mass of elementary particles. The top quark has a very short lifetime and decays before it hadronizes. It thus offer also the unique possibility to study the properties of a bare quark, including its spin. For these and other reasons the measurement of its properties is crucial for the understanding and verification of the Standard Model predictions. Moreover, it offers the possibility to investigate new physics through its decay (for example to a beyond the Standard Model Charged Higgs boson), or through its production via new particles decaying predominantly to top quarks (like a new generation of quarks (t') and supersymmetric particles).

The pp collider LHC at CERN is actually a *top factory*, which can produce as much as a top-pair per second. It opens a new era in top physics. The data collected in the first two years of data taking in 2010-2011 at $\sqrt{s} = 7$ TeV (about 800 000 quark pairs produced per LHC experiment) and those collected at $\sqrt{s} = 8$ TeV in 2012, may contain the key to answer some of the open questions of the Standard Model and provide evidence to support new physics theories.

The top quark decays in about 100% of the cases into a W boson and a b quark. Three $t\bar{t}$ decay topology are defined according to the decay mode of the two W which determine the different final states: all hadronic channel, when both the W bosons decay into hadronic jets; dileptonic channel, when both W bosons decay into leptons and semi-leptonic channel when one W decays hadronically and the other into leptons. The decay topology treated in this thesis is the semi-leptonic decay with τ lepton in the final state observed in its hadronic decay mode: $t \rightarrow b(W \rightarrow \tau\nu_\tau) \rightarrow b(\tau \rightarrow had\bar{\nu}_\tau)\nu_\tau$. This represents one of the most challenging experimental final state, due to the difficulty of reconstructing and identifying the hadronically decaying τ and due to the presence of more than one neutrino as the source of missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$). The final state contains various additional jets, two of them originating from a b-quark. Jet reconstruction with precise jet energy scale estimation and efficient and well calibrated tagging of b-quark jets constitute other challenging experimental issues.

The aim of this thesis is to measure the $t\bar{t}$ production cross-section observed in the semileptonic tau final state described above. The measurement is done using 2.05 fb^{-1} of LHC data produced at the center-of-mass energy $\sqrt{s} = 7$ TeV, collected by the AT-

LAS detector during 2011. It provides a test of the Standard Model probing the QCD production of $t\bar{t}$ pairs in proton-proton collisions and the weak decay of the top quark.

The main processes which constitute a background to the $t\bar{t} \rightarrow \tau + jets$ signal are the production of W boson in association with jets and multi-jets QCD production. Other smaller backgrounds comprise Z boson production in association with jets, single top quark production and diboson production (WW, WZ and ZZ). The analysis strategy has been designed to take maximum advantage of the statistics of signal events, separating events in various categories depending on the type of hadronic tau decay mode, the multiplicity of jets and the multiplicity of b-quark jets. These categories feature different signal-to-background ratios. Their characteristics will be used in a simultaneous likelihood fit to separate signal from backgrounds and to constrain with data some of the systematic uncertainties affecting the measurement, improving thus the precision of the measurement. These uncertainties are of experimental nature or are related to the physics modeling of the processes. The fit incorporates also $t\bar{t}$ candidates selected in the semileptonic electron channel ($t\bar{t} \rightarrow e + jets$ events) improving the statistical power of the data in constraining some systematic uncertainties. The measured cross-section is then compared to other experimental results and to the Standard Model prediction.

This thesis is organized as follows: in Chapter 1 the bases of the Standard Model of particle physics are presented; in Chapter 2 some insight into the physics of quark top is given; Chapter 3 contains a description of the main characteristics of the ATLAS detector at the LHC; in Chapter 4 the physics object reconstruction in the ATLAS detector is presented. Chapter 5 is dedicated to the ATLAS trigger system and in particular to the trigger used in this analysis. In Chapter 6 the strategy of selection of $t\bar{t} \rightarrow \tau + jets$ events and of the cross-section measurement is explained. Finally the results of this analysis are discussed in Chapter 7 as well as the prospects for the future.

Contents

List of Figures

List of Tables

Chapter 1

The Standard Model

1.1 Introduction

The Standard Model is a quantum field theory, formulated in the 60 s by S.L.Glashow [2], A. Salam [3] and S. Weinberg [4]. It successfully describes all the known fundamental matter constituents and their interactions at smallest scales (10^{-18}m) and at the highest energies achieved until now. According to the Standard Model two kind of elementary particles exist: the fermions, the ordinary matter constituents, and the gauge bosons that are the mediators of the interaction between particles. There are 12 known fermions (listed in Table 1.1); all of them are spin-$\frac{1}{2}$ particles, and for each of them there exists a corresponding antiparticle with equal mass and opposite additive quantum numbers.

Fermions	Generation			Electric charge	Interactions
	I	II	III	(e)	
Quark	u	c	t	$+\frac{2}{3}$	Strong
					Weak
	d	s	b	$-\frac{1}{3}$	EM
Lepton	ν_e	ν_μ	ν_τ	0	Weak
	e^-	μ^-	τ^-	-1	Electroweak

Table 1.1: Fundamental Standard Model fermions.

The fermions are grouped in *families* or *generations*: until now the existence of only three generations has been assumed, and all their fermions have been experimentally discovered. Each generation consists of doublets of particles of increasing mass, classified into *quarks* and *leptons* depending on the force they are subject to. The two members of leptons and quarks doublets differ by one unit of electric charge. Under the assumption that the quantum numbers and the electric charge have the same sign, the electric charge of each quark can be derived from the Gell-Mann and Nijsihima formula:

$$Q = T_3 + \frac{B}{2} + \frac{s + c + b + t}{2} \tag{1.1}$$

where T_3 is the third component of the weak isospin, B is the baryonic number and s, c, b, t are the quantum numbers related to the *flavor* of the quark.

The range of fermion masses is at least 11 order of magnitude wide. The top quark is the heaviest fundamental particle with a mass of 173.5 GeV [1] [CHECK PDG VALUE], and is probably related with the mass generation process. The flavor symmetry breaking mechanism and the mass hierarchy are part of Standard Model theory, although they are not yet fully understood.

There is a *mixing* mechanism between the three quark generations, parametrized by the Cabibbo-Kobayashi-Maskawa (CKM) matrix, that relates the flavor eigenstates of the weak interaction to the mass eigenstates. The second component of the quark doublets is not a physical state but rather a mixture of three *down* physical states, each of them weighted by the corresponding CKM matrix element. The origin of CKM matrix is not yet explained in the Standard Model theory.

In the hypothesis that neutrinos are not massless, another unitary matrix analogous to CKM is introduced, to explain the neutrino oscillation between flavor states observed in recent experiments [1]. If mixing of charged lepton exists, the weak interaction mediated by the W boson that decays into lepton+neutrino leads to the possibility of coupling between each charged lepton mass eigenstates and each neutrino mass eigenstates. Each neutrino flavor eigenstates $(\nu_e, \nu_\mu, \nu_\tau)$ would then be a linear combination of the eigenstates of mass (ν_1, ν_2, ν_3), determined by the mixing matrix.

As anticipated before, the Standard Model is a quantum field theory, which means that particles are treated as excitations of quantum oscillators of the corresponding fields. In other words, elementary particles are the *quanta* of the elementary fields described by the theory. The last actors of the play are the forces associated to the fields. They are interpreted as the exchange of virtual particles, named the force carriers or mediators. The Standard Model is also a *gauge* theory: the Lagrangian is invariant under some symmetry (or *gauge*) transformations. Ensuring invariance of the Lagrangian under local transformations (depending on space-time coordinates) requires the introduction of new gauge vector fields and give rise to interactions mediated by the associated quanta the *gauge* bosons. The underlying symmetry in the Standard Model is $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$. In particular, the interactions described by the Standard Model are divided in:

- Electroweak interaction, with the corresponding symmetry group $SU(2)_L \otimes U(1)_Y$;
- Strong interaction, with the corresponding symmetry group $SU(3)_C$.

More details about these theories are reported in the next sections.

1.2 Quantum Electrodynamics

The Quantum Electrodynamic (QED) is an *abelian gauge theory*[1], described by the symmetry $U(1)_Q$, where Q represents the electric charge. The QED coupling is called the *fine-structure constant*, even if the name is not really appropriate since it is not actually a constant, but its value varies noticeably in function of the momentum transferred in the

---

[1] A gauge theory is said *abelian* when its generators commute. One of the consequences of this is the fact that the photons cannot interact with themselves.

Interaction	Acts on	Particles experiencing	Particles mediating	Electric charge (GeV)	Mass	Spin
Electromagnetic	Electric Charge	All charged particles	Photon (γ)	0	0	1
Weak	Flavor	Quarks, $W^{\pm}$	$W^{\pm}$	± 1	80.403 ± 0.029	1
		Leptons, Z^0	Z^0	0	91.1876 ± 0.0021	1
Strong	Color charge	Quarks, Gluons	Gluons (g)	0	0	1
Gravitation	Mass and energy	All All	Graviton (not observed yet)	0	0	2

Table 1.2: Fundamental interactions and their proprieties. Gravitation is separately shown for completeness even if is nos part of the Standard Model.

interaction (figure 1.3), as happens to the other coupling of the theories that compose the Standard Model. For this reason, they all are otherwise called *running couplings*. Anyway, the *fine-structure constant* can be defined as:

$$\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c} \simeq \frac{1}{137} \qquad , \tag{1.2}$$

where: $e = 1,602176 \times 10^{19}\ C$ is the electron charge; $\epsilon_0 = 8.854187 \times 10^{12}\ Fm^{-1}$ is the vacuum dielectric constant; $\hbar \equiv \frac{h}{2\pi} = 1.054571 \times 10^{-34}\ Js$, being h the Planck constant, and $c = 299792458\ ms^{-1}$ is the velocity of light in the vacuum. As a consequence of the zero mass of its carriers, the photons, the electromagnetic force has infinite range of action.

1.3 Weak interactions

Historically, the discovery of the first weak interaction process has been the neutron *beta* decay, defined as *charged current*. It has been introduced by Fermi, in a purely vectorial version. The intensity of the weak interaction is parametrized by the Fermi constant G_F, which value can be expressed as:

$$\frac{G_F}{(\hbar c)^3} = \sqrt{2}\frac{g^2}{8m_W^2} = 1.16 \times 10^{-5} GeV^{-2} \qquad , \tag{1.3}$$

where m_W is the W boson mass and g one of the electroweak couplings that will be introduced in next section. Nowadays is known that the weak interaction is mediated by charged bosons of very high mass (around 80 GeV) and that only the *left-handed* components of the fermions are subject to this interaction. The unification of the weak and the electromagnetic interactions theory, lead to the additional introduction of weak *neutral currents* mediated by a neutral boson of mass around 90 GeV, as explained in the next chapter.

1.4 Electroweak unification

The electroweak interaction theory (EW) describes the unification of the electromagnetic and weak interactions, through the introduction of the gauge group of the weak *left-handed* isospin and the hypercharge, being the Lagrangian symmetric under $SU(2)_L \times U(1)_Y$ transformation. The Lagrangian contains a vectorial minus vector-axial term (V-A), that is not symmetric under parity transformations[2]: that s the origin of the experimental fact that all the electroweak interactions violate Parity. As V-A interaction couples only the *left-handed* state of the particles, the fermionic fields Ψ are classified in *left-handed* and *right-handed* fields, of the form:

$$\psi_L = \frac{1}{2}(1-\gamma_5)\psi \qquad and \qquad \psi_R = \frac{1}{2}(1+\gamma_5)\psi, \tag{1.4}$$

where γ_5 is given by the product of the *gamma* (or Dirac) matrices times the i number: $\gamma_5 = i\gamma_0\gamma_1\gamma_2\gamma_3$. The fermionic fields are arranged into doublets ($T = \frac{1}{2}$) and singlets ($T = 0$) of weak isospin: leptons are shown in table 1.3 while quarks are shown in table 1.4, together with their characteristic quantum numbers.

Lepton	T	T_3	Q	Y
ν_e	$\frac{1}{2}$	$\frac{1}{2}$	0	-1
e^-_L	$\frac{1}{2}$	-$\frac{1}{2}$	-1	-1
e^-_R	0	0	-1	-2

Table 1.3: Weak isospin and hypercharge quantum numbers for leptons.

In the doublets, the neutrinos and the *up*-type quarks (u, c, t) present a value of weak isospin of $T_3 = +\frac{1}{2}$, while the charged leptons and the *down*-type quarks (d, s, b) have weak isospin $T_3 = -\frac{1}{2}$. The weak hypercharge Y, defined by the relation $Q = T_3 + Y/2$, for the member of doublets is $Y = -1$ for leptons and $Y = \frac{1}{3}$ for quarks.
The singlets of weak isospin instead are *right-handed* fermions with hypercharge $Y = -2$ as regards charged leptons, $Y = 4/3$ and $Y = -2/3$ as regards *up*-type and *down*-type

Quark	T	T_3	Q	Y
u_L	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{1}{3}$
d_L	$\frac{1}{2}$	-$\frac{1}{2}$	-$\frac{1}{3}$	$\frac{1}{3}$
u_R	0	0	$\frac{2}{3}$	$\frac{4}{3}$
d_R	0	0	-$\frac{1}{3}$	-$\frac{2}{3}$

Table 1.4: Weak isospin and hypercharge quantum numbers for quarks.

---

[2] The Parity operator is defined as spatial inversion around the origin. The invariance under parity transformation can be translated into the fact that applying the parity operator twice, the original wave-function remains unchanged.

quarks, respectively. Charged current in Electroweak theory are those that transform each member of the doublets into the other, while the singlets are transformed back into themselves. Neutral current instead transmute singlets as well as the upper and the lower term of the doublets into themselves.

The two couplings of the EW theory, corresponding to the symmetry groups $SU(2)_L$ and $U(1)_Y$, are related by the formula:

$$e = gsin\theta_W = g'cos\theta_W, \tag{1.5}$$

where θ_W is the Weinberg angle, which value has been determined in various experiments: from $\nu - e$ scattering; through the electroweak interference in scattering processes of e^+e^-; from the study of the Z boson decay width and from the relation between W and Z masses. Combining all these experiments results the obtained value is:

$$sin^2\theta_W = 0.23113 \pm 0.0005. \tag{1.6}$$

The Weinberg angle introduces a mixing, between the neutral boson W^3 of $SU(2)_L$ group and the neutral boson B of $U(1)_Y$ group. So, the physics states A_μ (the photon) and Z_μ (the Z boson), responsible of neutral electroweak currents, are a linear combination of them:

$$A_\mu = B_\mu cos\theta_W + W^3_\mu sin\theta_W, \qquad Z_\mu = -B_\mu sin\theta_W + W^3_\mu cos\theta_W. \tag{1.7}$$

On the other hand, the charged bosons $W^\pm$, are a liner combination of the W_1 and W_2 states:

$$W^\pm_\mu = \frac{W^1_\mu \mp iW^2_\mu}{\sqrt{2}}. \tag{1.8}$$

The charged W boson couples similarly to quarks as well as to leptons with same chirality, while the Z coupling depends strongly on the particle. Actually, the strength of the Z coupling to a fermion is given by:

$$g_Z(f) = \frac{g}{cos\theta_W}(T_3 - Q_f \sin^2\theta_W) \tag{1.9}$$

where Q_f is the fermion electric charge in unity of elementary charge e; while T_3 is the third component of the weak isospin.

It is also possible to describe the coupling of fermions with the neutral weak field in terms of *right-handed* and *left-handed* currents, by introducing the following couplings:

$$g_L = (T_3 - Qsin^2\theta_W) \tag{1.10}$$

$$g_R = -Qsin^2\theta_W \tag{1.11}$$

	ν_e, ν_μ, ν_τ	e, μ, τ	u, c, t	d', s', b'
g_L	$\frac{1}{2}$	$-\frac{1}{2} + sin^2\theta_W$	$\frac{1}{2} - \frac{2}{3}sin^2\theta_W$	$-\frac{1}{2} + \frac{2}{3}sin^2\theta_W$
g_R	0	$+sin^2\theta_W$	$-\frac{2}{3}sin^2\theta_W$	$\frac{1}{3}sin^2\theta_W$

Table 1.5: Couplings of fermions with the Z neutral boson.

There is *mixing* between the generations of the quarks, which means that the quark doublets in nature occur as:

$$(u_L, d'_L) \qquad (c_L, s'_L) \qquad (t_L, b'_L) \qquad , \tag{1.12}$$

where the *flavor* states d'_L, s'_L and b'_L are actually mixtures of the *mass* states d_L, s_L and b_L. The W couplings allow any charge changing transition; in such a way are possible decays as for example $c \to s$ or $s \to u$, with the associated production of a virtual W boson that materializes decaying into light quarks or leptons. The *mixing* between the flavor and the mass states of quarks has been parametrized using a 3×3 mixing matrix known as the Cabibbo-Kobayashi-Maskawa (CKM) matrix represented as:

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{1.13}$$

in such a way that:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}$$

The V_{ud} matrix element for example, parametrizes the coupling of the interaction between an u-quark with a W boson resulting in a d-quark. The CKM matrix is mostly diagonal, with some small off-diagonal terms, especially for mixings between light quarks. One possible parametrization expresses each matrix element through the combination of four parameters: the cosine and sine of the three mixing angles between the quarks generation, $\theta_{i,j}$ with $i < j=$ 1,2,3. The fourth parameter is actually a phase, δ, which value slightly different from unity accounts for the CP[3] violation observed in some electroweak interactions. If CP were an exact symmetry the laws of nature would be the same in matter and anti-matter, but interestingly it turns out that CP asymmetry is about 0.3%. Generation mixing between leptons can only occur in the case that neutrinos have masses: this would generate the phenomena of neutrino oscillation, investigated by the last generation of neutrino experiments.

The electroweak theory formulated by Glashow-Salam and Weinberg cannot be described by an exact symmetry in $SU(2)_L \times U(1)_Y$ due to the fact that the gauge bosons are not all massless: the only massless mediator boson is actually the photon, while the W and Z boson are very massive instead. Their large masses explain the short range of the electroweak interaction, but the introduction of a corresponding mass term in the Lagrangian

---

[3] CP is the combined operation of charge conjugation and parity. For example,CP turns a left-handed electron in a right-handed positron.

breaks its gauge invariance. Moreover, a mass term introduced by hand implies that the diagrams with loops taken into account in the matrix element amplitude calculation of physics processes would diverge, leaving the theory no more renormalizable[4].

1.5 EW symmetry breaking and the Higgs mechanism

For the reasons explained in the previous section, in order to explain the masses of the Z and W gauge bosons exchanged in electroweak interactions, the introduction by hand of a term of mass in the Lagrangian is not a solution. A spontaneous symmetry breaking is needed, in a way that would retain the renormalizability of the theory. In addition, it is also not possible to include in the Langrangian a classical term of mass for fermions, because a term of the form:

$$m(\psi_L^\dagger \psi_R + \psi_R^\dagger \psi_L) \tag{1.14}$$

is not gauge invariant under $SU(2)_L$.
The Higgs mechanism has been proposed by Peter Higgs in 1964, as a mechanism inside the Standard Model able to break spontaneously the electroweak symmetry at a certain energy scale. At high transferred momentum, $q^2 >> M_Z^2, M_W^2$, the masses of fermions and bosons becomes irrelevant and the symmetry is restored. In the simplest formulation of the Higgs model, the mass of the quarks, leptons and weak vector bosons are all interpreted as the result of the interaction with a single Higgs field. As usual, there is a particle associated to each quantum field: from the EW symmetry breaking mechanism the existence of at least one scalar boson, called Higgs boson, arises. The new particle observed at CERN by both ATLAS and CMS experiments seems to be a good candidate for the Higgs boson. It has been observed, in two different decay channels: H into four leptons and H into two photons final states, as announced on 4th of July of 2012. The Higgs mechanism assumes the existence of a fundamental scalar neutral field which Lagrangian is given by

$$\mathcal{L}_{Higgs} = (D_\mu \phi)^\dagger (D_\mu \phi) - V(\phi^\dagger \phi), \tag{1.15}$$

where D_μ is the covariant derivative and ϕ is a doublet of weak isospin of scalar complex fields, with $Y = 1$, of the form:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 & \\ \phi_3 + i\phi_4 & , \end{pmatrix} \tag{1.16}$$

the first component having charge $Q = 1$, in units of elementary charge e, and the second being neutral.

The Higgs potential $V(\phi^\dagger \phi)$ is given by:

$$V(\phi^\dagger \phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 \tag{1.17}$$

where μ and λ are arbitrary parameters. When μ^2 and λ are both positive, the potential has the familiar parabolic shape, with a minimum at zero value of the field (see figure 1.1,

---

[4] Renormalizability is a sort of re-parametrization of the Feynman diagrams amplitudes of a process in a way that the infinities of some of the higher order terms cancel.

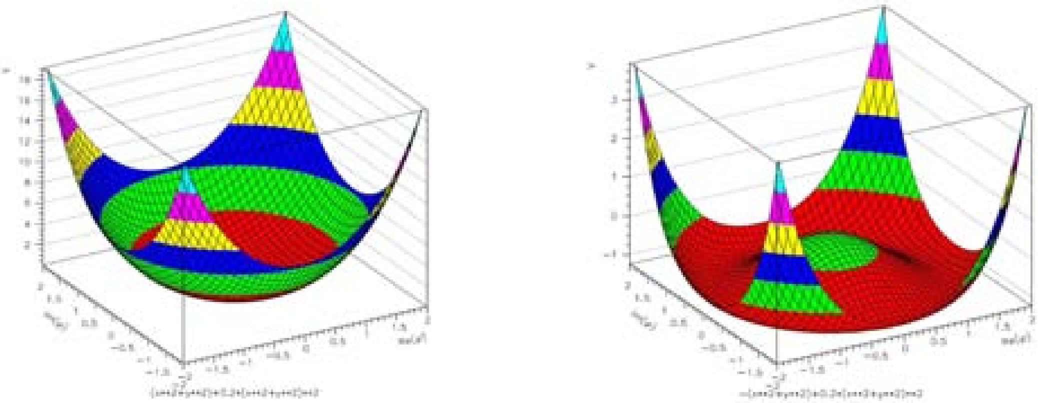


Figure 1.1: Graphical representation of $V(\phi^\dagger\phi) = \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2$ potential for $\mu^2 > 0$ (left) and $\mu^2 < 0$ (right).

left). While if $\mu^2 < 0$ and $\lambda > 0$ the potential assumes the so called *Mexican hat* or *bottle bottom* shape (see figure 1.1, right). Instead of having a single minimum, this potential has a ring of minima in the complex plane, *i.e.* each time that

$$|\phi|^2 = -\frac{\mu^2}{2\lambda} = \frac{v^2}{2}, \tag{1.18}$$

where v indicates the vacuum expectation value (VEV), that is the amount of the Higgs field in the empty space. This means that minimizing the energy now requires non-zero field through out all space, and the potential is perfectly symmetric with respect to phase rotations: the probability for the field to assume any of these values is the same. By choosing exactly one of these values at the minimum energy, a spontaneous breaking of the symmetry occurs. A particular convenient choice is: $\langle\phi_1\rangle = \langle\phi_2\rangle = \langle\phi_4\rangle = 0$ and $\langle\phi_3\rangle = v$. In this way it is possible to expand the field around the minimum value, thanks to the introduction of four scalar fields θ_1, θ_2, θ_3, and H[7]:

$$\phi(x) = e^{i\tau\cdot\theta(x)/v}\left(\begin{array}{c} 0 \\ \frac{v+H(x)}{\sqrt{2}} \end{array}\right). \tag{1.19}$$

The Lagrangian results locally $SU(2)_L$ invariant: this justifies the choice of a particular *gauge* to remove the three fields $\theta_i(x)$ that represent non-physical degrees of freedom (Goldstone bosons); the symmetry breaking allows to absorb such degrees of freedom as longitudinal modes of W and Z bosons while they acquire mass.

The Higgs doublet becomes then:

$$\phi = \frac{1}{\sqrt{2}}\left(\begin{array}{c} 0 \\ v + H(x) \end{array}\right). \tag{1.20}$$

The choice of ϕ_0 as vacuum state with $T = \frac{1}{2}$, $T_3 = -\frac{1}{2}$ and $Y = 1$ breaks the $SU(2)_L$ and $U(1)_Y$ symmetries; but the $U(1)_Q$ symmetry is preserved because the field ϕ_0 have zero electric charge.

By substituting $\phi(x)$ in the Lagrangian, the terms of mass for the $SU(2)_L$ and $U(1)_Y$ gauge field can be obtained. The masses of the gauge bosons comes out from the definitions of their physics states:

$$m_{W^\pm} = \frac{1}{2}vg; \tag{1.21}$$

$$m_Z = \frac{1}{2}v\sqrt{g^2 + g'^2}; \tag{1.22}$$

$$m_A = 0. \tag{1.23}$$

The A_μ field, related to the $U(1)_Q$ symmetry, remain massless, in agreement with the fact that it represents the photon; while Z and W bosons acquire masses. The physical meaning of the substitution illustrated above is that the Z and W interact with the Higgs field: they travel through what would be otherwise an empty space, and collide with the Higgs field. The mass comes to represent the *inertia* of the particles: so that heaviest particles go slower than light ones given equal pushes, because the light particles interact less (bump less often) with the Higgs field. Actually, the γ doesn t interact at all with the Higgs field and go straight with the speed of light. The Lagrangian contains also self-interaction of Higgs boson terms and the mass of the Higgs boson itself is given by $m = \sqrt{2\lambda v^2}$.
Also fermions, as the bosons, acquire masses interacting with the Higgs field. Their masses are generated introducing in the Lagrangian terms invariant under $SU(2)$ transformations. For example, in case of electrons:

$$-G_e\left[(\nu_e, \overline{e})_L \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} e_R + \overline{e}_R(\phi^-, \overline{\phi^0}) \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L\right].$$

The Higgs doublet has the exact quantum numbers of $SU(2)_L \otimes U(1)_Y$ that allow the coupling with the *left* and *right* component of the fermions. Substituting the ϕ doublet by the form:

$$\phi(x) = \sqrt{\frac{1}{2}}\begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \tag{1.24}$$

the following terms are obtained :

$$-\frac{G_e}{\sqrt{2}}v(\overline{e}_L e_R + \overline{e}_R e_L) - \frac{G_e}{\sqrt{2}}(\overline{e}_L e_R + \overline{e}_R e_L)H = -\frac{G_e v}{\sqrt{2}}\overline{e}e - \frac{G_e}{\sqrt{2}}\overline{e}eH. \tag{1.25}$$

The first term is the mass term for fermion e (electron in this case) while the second term represents the interaction of the fermion with Higgs boson. The mass of the fermion is $m_e = \frac{G_e v}{\sqrt{2}}$, where G_e is just a parameter to be determined by the experiment: in fact the Higgs mechanism can be used to generate masses of fermions but cannot predict the actual value of the masses. For the leptons the procedure is analogous, while to generate the quarks masses, the conjugated doublet ϕ_C of the fields has to be added in the Lagrangian. After the spontaneous symmetry breaking it assumes the form:

$$\phi_C = -\sqrt{\frac{1}{2}}\begin{pmatrix} v + H \\ 0 \end{pmatrix}. \tag{1.26}$$

This field allows the coupling of the doublet *left* of quarks with the *right* component of the *up*-type quark.

The Higgs mechanism provide also an explanation to the CKM matrix generation (1.13) and of the flavor changing charged currents that change the flavor of quarks and mix the three quark generation allowing transitions between them.

Beyond the Standard Model many extensions of this simplest version of the Higgs sector have been developed. One of them, called the Minimal Supersymmetric Model, performs a scenario with two complex Higgs doublets. At *tree-level* this model depends on just two parameters: one common choice for them is the mass of the Higgs with even value of CP (m_A), and the ratio between the vacuum expectation values of the two Higgs doublet ($\tan\beta$).

1.6 Quantum Chromodynamics

In the Standard Model the strong interactions are described in Quantum Chromodynamics (QCD) theory, where the mediators of the strong interaction are the gluons, at the same way that in QED the photons are mediators of the electromagnetic interaction. The *charge* exchanged in QCD is called *color charge*. One of the main differences between QED and QCD is that there are 3 states of color charge, usually called *red*, *green* and *blue*. Quarks can change their color charge through QCD processes, while leptons don t carry any color charge and so don t take part in QCD. Since color charge must be conserved in interaction vertices, gluons have themselves to carry a color charge.
In terms of symmetries, the QCD Lagrangian is gauge invariant under $SU_C(3)$ group transformations, *i.e.* under the mixing of three fermion fields, that are three independent copies of the same quark, distinguished by color quantum number. A general $SU_C(3)$ gauge transformation can be written as:

$$\psi \rightarrow \psi' = e^{i\frac{g}{2}\alpha\cdot\lambda}\psi, \tag{1.27}$$

where:

- g represent the gauge coupling of the theory;
- the components of λ are the $SU_C(3)$ generators, represented by eight unitary 3×3 matrices;
- the components of α are eight free parameters, required to be $\alpha_i = \alpha(x)_i$ in order to make this a local gauge transformation.

As the α_i transformation generators don t commute, the QCD is a non-Abelian theory. For this fact, the eight massless gluons predicted by QCD can interact directly between themselves, unlike photons. There are two interesting implications of this feature of gluons that make the strong interaction very different from QED. The self-interaction amongst gluons makes the force appear stronger at longer distances (lower energies), causing the *confinement* of quarks in hadron: free quarks cannot be observed, they always form bound states of *mesons* ($q\bar{q}$) or *hadrons* ($q_iq_jq_k$). The observed states are actually singlets of color, that means states with no net color quantum number. The only quark that doesn t form bound states is the top quark: due to its large mass it decays more quickly compared to the typical hadronization time ($\Gamma_{top} \gg \Lambda_{QCD}$). Therefore the top decay offers a unic possibility to study the proprieties of *bare* quarks.

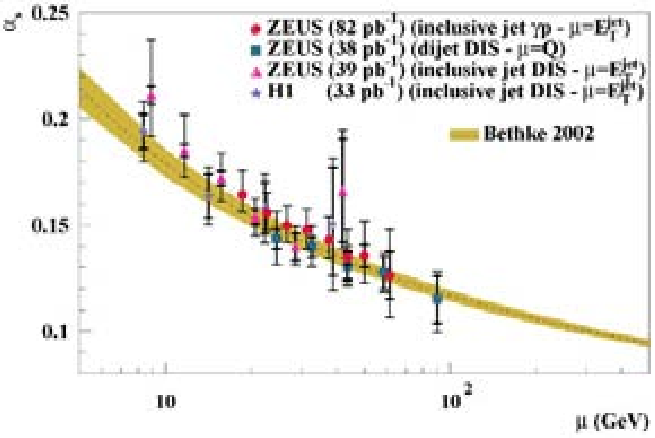


Figure 1.2: Results from the HERA experiments H1 and ZEUS which directly confirm the asymptotic freedom: the strength of the strong force described by the strong coupling α_S, decreses with increasing energy scale (2004 Nobel Prize in Physics).

However, on short distance scales (or equivalently at high energy scales) quarks behave as free particles, and this characteristic is known as *asymptotic freedom*. This behavior is in practice the opposite of the electric charge screening phenomenon. In fact in QED, the measured charge depends on the distance of the probe charge: supposing that the charge to be measured is positive and the other negative, the former will be surrounded by electrons, and the coupling α comes to represent the strength of the repulsion between the probe-electron and the sum of the screening charge. Its asymptotic value is $\sim 1/137$, but actually depends on the distance (energy): smaller the distance (greater the energy), greater is the strength of the electric interaction. On the contrary in QCD, each quark is surrounded by gluons that carry the same color charge, in a way that the coupling becomes arbitrarily feeble at arbitrary short distance (high momentum), while it s is greater as far as the distance becomes greater. So, at very small distances quarks behave as non interacting particles.
It is very interesting to study inter-quark processes at LHC, because only at these energy scales it is possible to do high precision tests using perturbative calculations as in QED.

At first perturbative order, the coupling of strong interaction would be given by:

$$\alpha_S(\mu^2) = \frac{12\pi}{(33-2n_f)\cdot ln(\frac{\mu^2}{\Lambda^2})} \tag{1.28}$$

where n_f is the number of fermions, μ is the momentum transferred in the interaction and Λ is a free arbitrary parameter. From the comparison between the theoretical predictions and the experiment, it results $\Lambda \approx 250\ MeV/c$. A perturbative expansion in QCD holds only if $\alpha_S << 1$, condition verified for $\mu^2 >> \Lambda^2 \approx 0.06\ GeV^2/c^2$.

The dependence of α_S on μ^2 corresponds to its dependence on the separation amongst quarks. The asymptotic behavior can be expressed in terms of μ: in the limit of $\mu^2 \to \infty$ the quarks can be considered as free particles. On the contrary, at great separation (small

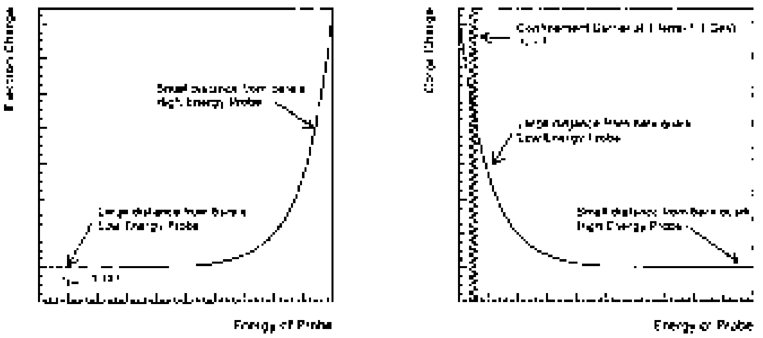


Figure 1.3: Sketch of the different behavior of electromagnetic and strong coupling α and α_S: the resulting properties are known as charge screening for the former (left) and as asymptotic freedom for the latter (right).

μ^2) the *confinement* occurs. The attempt of extract the quark from an hadron produce a hadronic jet through the formation of gluons and quark-antiquark pairs (*fragmentation*). At high energy scales, the fragmentation is described by perturnative QCD, as said before, but at low energy scales the *soft gluons* radiation has to be described using phenomenological models (non-perturbative QCD). The experimental study of such *soft* hadronic processes allowed the achievement of some progress in non-perturbative QCD predictions.

1.7 Limits of the Standard Model

The Standard Model theory gave very satisfactory results in the past years and we can say that the present time is one of the most interesting moments for Standard Model, because of the so long-awaited experimental detection of the Higgs boson. The discovery of a very likely candidate for the Higgs has become reality thanks to the challenging energies and luminosity achieved by the LHC and the optimal performance of ATLAS and CMS detectors and the efficiency of their collaborations. The evolution of technology in recent times allowed to do more and more accurate measurements also in astrophysics and cosmology. However, some issues outside of the Standard Model but related to particle physics, should be mentioned: they suggest the need to go beyond the present theory and extend it in order to describe new phenomena, as those listed here below.

1. The Dark Matter issue, which the Standard Model cannot provide any solution to. One of the most renowned theories explains the presence of dark matter in the Universe assuming the existence of a neutral stable and very massive particle (at least of the Z boson mass order of magnitude) that interact very feebly with matter.

2. The matter-antimatter asymmetry observed in the Universe requires both the barionic number and CP violation. The former is not expected in the Standard Model, while the latter is expected through a phase of the CKM matrix. Anyway the measured value cannot explain the level of asymmetry observed.

3. The existence of a non-zero mass for neutrinos is required by the evidence of their

flavor oscillation during their propagation. But in the Standard Model neutrinos have to be massless except if the existence of the *right-handed* neutrino is admitted.

In addition to this incompatibilities, some aspects of the Standard Model need to be better understood, in such a way to build a more general theory:

- The large number number of degrees of freedom[5]: the masses of fermions and bosons, the values of the couplings, the coefficients of the mixing electroweak matrix (19 parameters in total). The dependence on too much arbitrary variables is maybe an indication that it is not a fundamental theory.
- With the renormalization theory the couplings of the interactions described in the Standard Model are expressed in function of the energy scale. By extrapolating the coupling values for very high energies it seems that they get closer: this suggests the possibility of a Great Unification Theory (GUT) of all the fundamental interactions. Nevertheless at high energy scale the couplings seem not to converge.
- Gravity is not described in the Standard Model: on one hand at the low energy scale at which the Standard Model has been validated by the experiment, gravity can be neglected, but on the other hand an universal quantum theory should incorporate it too.
- The Standard Model cannot explain neither the limit at three on the number of generation or families of leptons and quarks, nor the observed differences between their masses.
- The hierarchy problem: in the Electroweak symmetry breaking mechanism the term μ^2 is required to be negative in order to obtain a ring of minima for the vacuum expectation value of the Higgs field. Any how this is true at tree-level, but at higher orders the renormalization is needed in order to re-absorb the loop divergences and make the theory valid at arbitrarily high energies. At the Plank scale the value of Λ^2 parameter is of the order of 10^{38} GeV^2: assuming an Higgs mass lower than 1 TeV (as the LHC experiments seem to show), the value of $-\mu^2$ has to be extremely small compared to the scale. In other phenomenology theories verified over a certain threshold, like GUT, the contribution of the various loops is of the order of the threshold value. The only way to obtain this in the Standard Model, is to apply a very *fine-tuning* of the parameters.

In summary, the Standard Model looks like a sort of effective theory, able to accurately describe the phenomenology at the electroweak symmetry breaking scale (up to 1 TeV); beyond that scale seems to be more adequate to elaborate an new theory rather than an extension of the model. The unsolved problems of the Standard Model seem to indicate that the Standard Model is rather a low energy limit of a more fundamental theory. It is possible that we are now in a case analogous to the situation in the 60 s, when the electroweak theory has been built from the unification of electromagnetism and weak interaction theory. For this reason the idea of a more general theory that would include also gravity, could be a likely possibility.

---

[5] Degrees of freedom are the parameter that cannot be extracted by the theory and have to be obtained by the experiment

Chapter 2

Top physics at LHC

The top quark ([15], [16]), the most massive elementary particle ever observed, has been experimentally discovered in 1995 at Fermilab , and in such a way the third generation of quarks of the Standard Model has been completed. Due to the scale of mass of the top quark, its production is only accessible at very high center-of-mass energies ($\sqrt{s}$), so far by hadron colliders only. At Tevatron, the Fermilab proton-antiproton ($p\bar{p}$) collider, characterized by an center mass energy of 1.8 TeV in the first run and 1.96 TeV in Run II, using CDF and Dff experiments simultaneously, the production of $t\bar{t}$ pairs through strong interaction has been studied. In addition, the kinematic properties and the spin correlations have been studied and top mass and production cross-section have been measured. The Large Hadron Collider at CERN has been defined as a *top factory*: a top-pair can be produced every second. The properties studied at the Tevatron, have been carefully re-examined at LHC with the first 5 fb^{-1} of data collected in 2011 at $\sqrt{s} = 7$ TeV.

The strong interest in the top quark is justified because of its large mass and its link to the Electroweak Symmetry Breaking (EWBS) and to the Higgs mechanism that plays an important role in the origin of the mass of elementary particles. An accurate mass determination is crucial, together with that of the W and Higgs, because they intervene in the calculation of radiative corrections of the main Standard Model observables (see Figure 2.1). On one hand the study of top-pairs production is a very good channel to probe the current understanding of the strong interaction and the predictions of perturbative QCD calculations; on the other hand the single top quark production provides a possibility to investigate the weak interaction. In addition, the detailed study of the top quark properties could give indications of new physics signals, through the measurement of anomalous couplings in top quark production or decays. Events with top quarks decaying into charged leptons and neutrinos constitute a background of many new physics searches. The understanding of experimental signature of the top events, involving the ATLAS detector in many of its components, plays therefore an important role for the estimation of the potential discovery of new physics. In particular the $lepton + jets\ t\bar{t}$ decay channel has the advantage of large cross section and high signal to noise ratio (S/B), that allows to have high purity samples with appreciable statistics in relatively short data taking periods.

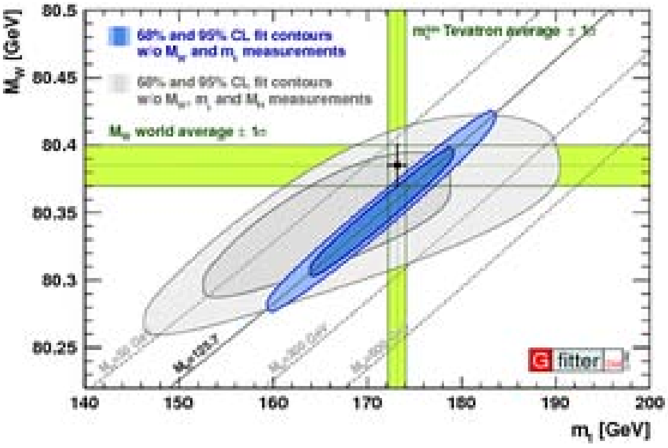


Figure 2.1: CL contours obtained from scans of fits with fixed values of M$_W$ and m$_{top}$. The narrower blue and larger grey allowed regions are the results of the fit including and excluding the M$_H$ measurements, respectively. The green bands indicate the 1σ regions of the M$_W$ and m$_{top}$ measurements (world averages) [17].

2.1 $t\bar{t}$ pairs production

In the context of the Standard Model, the top quark can be produced in quark-antiquark pairs predominantly via strong interaction, while single tops are produced via weak interaction. On one hand the hadron colliders allow to reach higher center-of-mass energies, but on the other hand they have to deal with the complicated theoretical predictions of hadronic interactions, in the complicated scenario of composite particles collisions. To simplify the calculations of composite hadrons, the *QCD factorization theorem* has been formulated, according to which there is a factorization scale μ_F^2 that allows to represent the process as the incoherent sum of interactions between partons at energy scales greater than μ_F^2 (perturbatively calculated). The contributions of the different partons contained in the proton are weighted with the Parton Distribution Functions (PDF): they provide the probability for a given parton to carry a certain fraction of the longitudinal momentum of the proton. The parton distribution function takes also into account all the processes that affect the parton interaction at energy scales lower than μ_F^2 scale. Hadronic collisions can hence be divided in long distance (low momentum transfer, Q^2) interactions and short distance (high momentum transfer) interactions with associated production of high mass and high transverse momentum particles. With this approach the proton itself can be described as a mixture of partons (quarks, antiquarks and gluons) interacting among them at low energy scale $\Lambda_{QCD} < 1\ GeV$ vias *soft* interactions, while the *hard* processes characterized by high transverse momentum $\geq O(100\ \text{GeV})$ happen between the partons that constitute the proton. As a consequence of the factorization theorem, the partons ($a,\ b$) taking part in the elementary process can be considered almost free and the cross-section can be calculated using perturbative calculations that not depend on the hadron to which the partons belong. In order to regularize the divergences that occur in higher orders calculations, another energy scale μ_R^2 together with the running coupling $\alpha_S(\mu_R^2)$ have to be introduced, in such a way that: $\widehat{\sigma}_{(a+b)\to X} = \widehat{\sigma}_{(a+b)\to X}(\widehat{s}, \alpha_S(\mu_R^2), \mu_R^2)$. The distribution

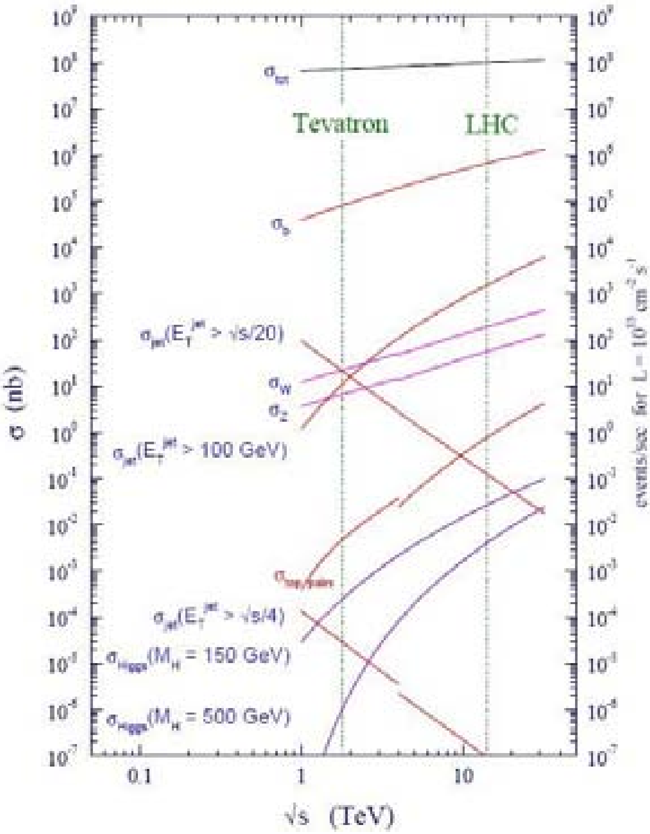


Figure 2.2: Total production cross-section and production cross-sections of the main Standard Model processes as a function of the center-of-mass energy. Dotted lines correspond to Tevatron Run II center-of-mass energy ($\sqrt{s} = 1.96TeV$) and designed LHC center-of-mass energy ($\sqrt{s} = 14TeV$).

of the proton longitudinal momentum between its partons is then described by a PDF, $f_{a/A}(x, \mu_F^2)$; it represents the probability to find a given parton a into the hadron A, with a momentum fraction x, when it is explored at the energy scale μ_F^2. If we consider, in addition to the factorization theorem, the parton interactions as elastic scatterings (which is true in the limit of such short times that any interaction with other partons can occur) and assume negligible mass for the partons, the fraction of momentum x carried by the parton (Bjorken scale variable) can be defined as:

$$x = \frac{Q^2}{M\nu}, \tag{2.1}$$

where Q^2 is the transferred four-momentum, M the mass of the proton and ν the transferred energy. The PDF also takes into account the singularities arising in perturbative calculations of the partonic cross section. Thanks to the factorization theorem it is then possible to calculate the $t\bar{t}$ production cross-section by integrating the cross-section of the elementary process with the partonic distribution function of the proton:

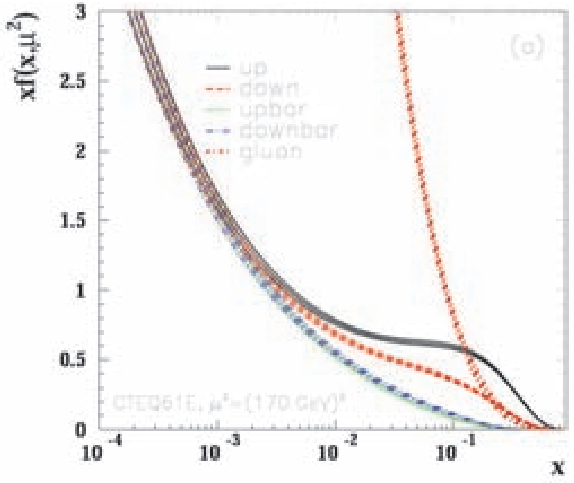


Figure 2.3: CTEQ61 parton distribution functions with their uncertainty bands, for the m_{top} mass scale ($Q^2 = (170\text{ GeV})^2$) for (anti-) up quarks, (anti-) down quarks and gluons in the proton.

$$\sigma_{A+B\to t\bar{t}}(\sqrt{s}, m_t) = \sum_{a,b=g,q,\bar{q}} \int \sigma_{a+b\to t\bar{t}}(s, \alpha_s(\mu_R^2), \mu_R^2, \mu_F^2, m_t) f_{a/A}(x_a, \mu_F^2) f_{b/B}(x_b, \mu_F^2) dx_a dx_b, \tag{2.2}$$

where $\sigma_{A+B\to t\bar{t}}(\sqrt{s}, m_t)$ is the $t\bar{t}$ production cross-section from the collision of hadrons A and B, that at LHC are both protons; s is the collider center-of-mass energy; m_t is the mass of top quark; $f_{a/A(b/B)}(x_{a(b)}, \mu_F^2)$ is the PDF of parton a (b) and proton A (B); $\sigma_{a+b\to t\bar{t}}$ is the partonic cross-section; μ_R^2 is the renormalization scale and μ_F^2 is the factorization scale. One should notice that in the formula above the physics cross-section $\sigma_{A+B\to t\bar{t}}$ does not depend on μ_R^2 and μ_F^2 scales introduced for the calculation, while the parton distribution functions and parton cross sections depend on them and consequently also the results of calculation at each finite order. This dependence decreases with the inclusion in the calculation of higher order terms. In practical applications, both scales are supposed equal to the typical momentum scale of the elementary scattering process, to the transverse momentum of the produced particles or to the mass of the produced particle. In case in which $t\bar{t}$ pairs are produced it is usually required $\mu_F = \mu_R = \mu = m_t$. The PDFs have been determined experimentally, for example with deep inelastic scattering of leptons on nucleons: they have been extracted from the measured cross-section, calculating in perturbative series the partonic cross-section. Since the PDFs are universal and does not depend on the process which they are derived from, they can be used to predict the cross sections of other elementary scattering processes, whenever they are calculated at the same perturbative order and same renormalization scheme. Figure 2.3 shows the main parton distributions function in the proton for the $t\bar{t}$ production at Tevatron or LHC. All PDFs vanish for large momentum fractions x and the gluon density starts to dominante over the valence quark density around $x = 0.13$. There is no flavor symmetry between the distributions of quarks up or down, neither for valence quark nor

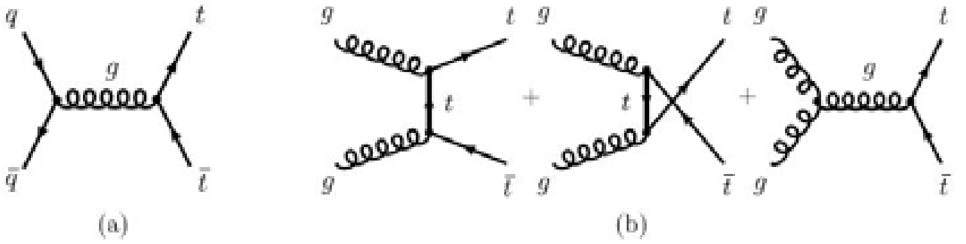


Figure 2.4: Feynman diagrams at LO for $t\bar{t}$ pairs production: (a) quark-antiquark annihilation; (b) gluon-gluon fusion.

for the *sea* quarks[1] (the contribution of the latter becoming more important at high Q^2). For values of x smaller than 0.1 the typical PDF uncertainties of the valence quarks and gluons are around 5%. At high values of x instead they increase drastically, especially the PDFs of gluons. The center-of-mass energy available for the parton collision $\sqrt{s}$ is related to the total center-of-mass energy $\sqrt{s}$ (given by the sum of the energies carried by the two protons) through the Bjorken scale variables by the relation:

$$\widehat{s} = x_a x_b s, \tag{2.3}$$

where x_1 (x_2) represents the fraction of the four-momentum carried by parton 1 (2). To produce a top quark pair, the square root of the center-of-mass energy at the parton level must be at least equal to the kinematic threshold: $\sqrt{s} = (2m_t)^2$. The Feynman diagrams at leading order for the $t\bar{t}$ pairs production by proton-proton collisions are shown in Figure 2.4: the quark-antiquark scattering (a) and the gluon-gluon scattering (b) respectively. The relative importance of the two amplitudes depends on the center-of-mass energy of the collisions: for the Tevatron $p\bar{p}$ collisions ($\sqrt{s} \simeq 2\ TeV$) the production of top quark pairs is kinematically limited to the region dominated by quarks. Since fractions of momentum close to the kinematic threshold are needed, the $t\bar{t}$ pair production in the $p\bar{p}$ collider is dominated by the process of quarks and anti-quark annihilation (about 85% during Run II). On the contrary, at the LHC the dominant process is gluon-gluon fusion (with a contribution of about 85% at $\sqrt{s}$= 7 TeV). , since at its center-of-mass energy low fractions of momentum are enough to produce $t\bar{t}$ pairs. This justifies the choice of pp collisions instead of $p\bar{p}$ collisions without any loss in production cross section and avoiding technological complication related to the production of an intense antiprotons beam. Consequently, the total production cross-section is also different amongst the two colliders: at LHC at design conditions the cross-section will be 100 times greater than at the Tevatron. Complete perturbative calculations for heavy quarks production cross-section have been performed up to the next-to-next-to-leading QCD perturbative order (NNLO).

---

[1] Sea quarks are those coming from momentary materialization of gluons which form quark-antiquarks pairs.

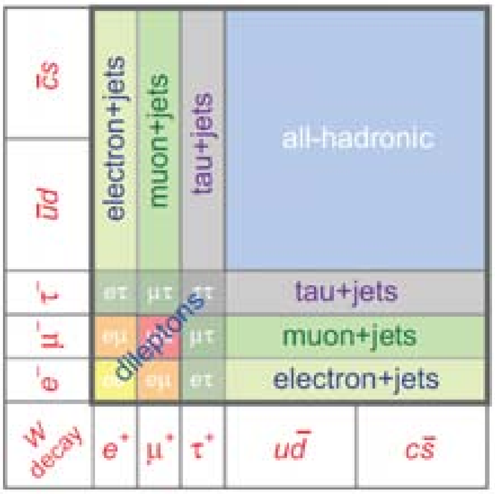


Figure 2.5: Diagram of the branching ratios of the $t\bar{t}$ decay channels.

2.2 $t\bar{t}$ pairs decay

Since the mass of the top quark is above threshold for the two-body decay $t \rightarrow Wq$, where q denotes a quark *down*-type (d, s or b), it represents the dominant decay. The contribution to the total decay width from each flavor of quark is proportional to the square of the corresponding CKM matrix element V_{qb}. Using the unitarity of CKM matrix and assuming three quark generations, at the 95% confidence level the matrix elements can be constrain as follows:

$$\begin{aligned} |V_{td}| &= 0.00862^{+0.00026}_{-0.00020}; \\ |V_{ts}| &= 0.0407^{0.0011}_{-0.0007}; \\ |V_{tb}| &= 0.999152^{+0.000030}_{-0.000045}. \end{aligned} \tag{2.4}$$

The top quark therefore decays in the almost totality of cases to Wb [1]. The limits on the CKM matrix elements shown in Eq. 2.4 would change dramatically if there were more than three generations of quarks. The V_{td} and V_{ts} matrix elements cannot be measured with the top decay at the tree level, but in the context of the Standard Model they can be derived from the mixing of B mesons. The direct measurement of V_{tb} matrix element without assuming three generations of quarks and the unitarity of the CKM matrix it is only feasible through single top production processes, since the production rate in each channel is proportional to $|V_{tb}|^2$. A way to estimate the relative size of $|V_{tb}|$ with respect to $|V_{td}|$ and $|V_{ts}|$ is to measure the ratio R between the top decay modes, which can be expressed through the elements of the CKM matrix, as:

$$R = \frac{B(t \rightarrow Wb)}{B(t \rightarrow Wq)} = \frac{|V_{tb}|^2}{|V_{tb}|^2 + |V_{ts}|^2 + |V_{td}|^2}. \tag{2.5}$$

Assuming the CKM matrix unitarity, the denominator of the previous expression is 1 and limits can be provided to $|V_{tb}|$.

The top decay width according to the Standard Model, taking into account QCD corrections at the first perturbative order, can be expressed in the following way:

$$\Gamma_t = |V_{tb}|^2 \frac{G_F m_t^3}{8\pi\sqrt{2}} \left(1 - \frac{m_W^2}{m_t^2}\right)^2 \left(1 + 2\frac{m_W^2}{m_t^2}\right) \left[1 - \frac{2\alpha_s}{3\pi}\left(\frac{2\pi^2}{3} - \frac{5}{2}\right)\right]. \tag{2.6}$$

In the formula above, it is assumed that $\frac{m_W^2}{m_t^2} \to 0$, $m_t^2 \gg m_W^2$. It expresses the top decay width within an accuracy of 2%. The other decays allowed by the Standard Model ($t \to Wd$, $t \to Ws$) constitute only a negligible contribution to the total decay width ($\Gamma_t = \sum_q \Gamma_{tq}$) proportional to $|V_{td}|^2$ and $|V_{ts}|^2$.

The decay width increases with the top mass: using $\alpha(m_Z) = 0.1176$ and $G_F = 1.16637 \times 10^{-5}(GeV/c^2)^{-2}$, Γ_t results to be 1.02/1.26/1.54 GeV/c^2 for top masses of 160/170/180 GeV/c^2 respectively. The resulting mean life of top quark is therefore $\tau_t = \Gamma^{-1} \approx (1.3/GeV/c^2)^{-1}$ that is about $5 \times 10^{-25}s$; significantly smaller than hadronization time $\tau_{had} = \Lambda_{QCD}^{-1} \approx 3 \times 10^{-24}s$. Consequently the top quark decays before that hadrons can be produced and no $t\bar{t}$ bound states do exist (the so-called *toponium*, analogously to other quarks resonances).

As mentioned before, the top quark decays almost exclusively as $t \to Wb$, producing a signature quite different from the collimated hadronic jets that represent the signature of lighter quarks. The W boson decays 67.6% of the time in $u\bar{d}$ or $c\bar{s}$, or else in the conjugate decays according to the sign of the W charge. The remaining times it decays into a charged lepton l and the corresponding neutrino ν_l. The abundance of any particular pair of quarks is determined by the corresponding element of the CKM matrix. In particular, the production of b-quarks is suppressed by a factor $\mid V_{cb} \mid^2 \simeq 1.7 \cdot 10^{-3}$. For this reason, the W boson decay can be considered a clean source of light quarks.

The decay of top quarks pairs leads to a three channel scenario, depending on how each W boson decays. The decay channels are classified as detailed here below and the branching fractions are sketched in Figure 2.5.

- **Dileptonic decay channel:** about 10.3% of all the $t\bar{t}$ decays both W bosons decay leptonically, resulting in final states with two leptons, each produced together with a neutrino, and two b-jets. The identification of this decay mode is done through the request of two leptons with high transverse momentum (p_T) and the presence of missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$); this allows to get a quite pure sample of top events. Usually, events with taus are not considered in this signature, reducing the branching fraction to ∼ 5%.

- **All hadronic decay channel:** it represents 46.2% of all $t\bar{t}$ decays. As both W bosons decay into jets, an all hadronic $t\bar{t}$ sample consist of events with six jets: 4 of them belonging to W decays, to which are added the two b-jets products together with the W bosons. In this case, there isn t any high p_{T} lepton to be triggered on. The signature is very similar to multi-jets QCD events, which are expected to be more abundant as the signal. The other challenge of this channel is the presence of a high combinatorial background in the reconstruction of top mass.

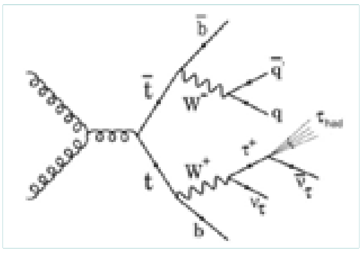


Figure 2.6: Scheme of the $t\bar{t}$ semi-leptonic decay in the $\tau + jets$ channel.

- **Semi-leptonic decay channel:** it accounts for 43.5% of all $t\bar{t}$ decays. The presence of a single isolated lepton with high transverse momentum provides a convenient signal to trigger and for the suppression of QCD and $W + jets$ background events. This signature, offers a good compromise between the purity of leptonic W decays and the statistics offered by the hadronic W branching fraction. Usually, people refers to semi-leptonic $t\bar{t}$ decays meaning just the events where the lepton is an electron or a muon, while the events involving tau leptons are handled separately. In the $e + jets$ and $\mu + jets$ channels, the neutrino transverse momentum can be reconstructed because it is the only source of transverse missing energy, which makes these channels much easier to deal with.

  The semi-leptonic $t\bar{t}$ with a hadronic tau in the final state which event topology corresponds to 10% of all $t\bar{t}$ decays, are the events of interest of this thesis. They are characterized by the presence of two neutrinos, as shown in Figure 2.6: the missing transverse energy is larger than in the electron and muon channels and it is impossible to separate the contribution due to each neutrino.

  Moreover, taus are the most difficult leptons to identify at hadron colliders, due to the multiple signature they leave in the detector. The branching ratio of tau leptons to one or more charged and/or neutral hadrons and a tau neutrino is BR($\tau \rightarrow hadrons + \nu_\tau$) $\sim$ 65%. Hadronically decaying taus appear as narrow jets; this signature is easily mimicked by hadronic jets or electrons. The decays of tau lepton into lighter leptons has a lower branching ratio: BR($\tau \rightarrow l\nu_l\nu_\tau, l = e, \mu$) $\sim$ 35%. In addition, electrons and muons from the tau leptonic decay can hardly be discriminated from electrons or muons produced from W decays. For these reasons, tau identification algorithms thus address only the hadronic tau decays. The branching ratio for hadronic tau decays into one charge hadron together with other neutral hadrons and neutrinos (*1-prong* decay topology) is $\sim$ 47%; while the decays into three charged hadrons + neutrinos (*3-prong* decay topology) constitute 15% of the total hadronic decays. Other modes contribute 3%. A typical way to collect top events with taus in the final state is achieved by requiring large missing transverse energy, several jets, at least one of which is identified as a b-jet, and vetoing the presence of electrons or muons. Usually leptonic top events remain as a background for the tauonic signal. More details of the analysis of this $t\bar{t}$ decay topology are given in Chapter 6.

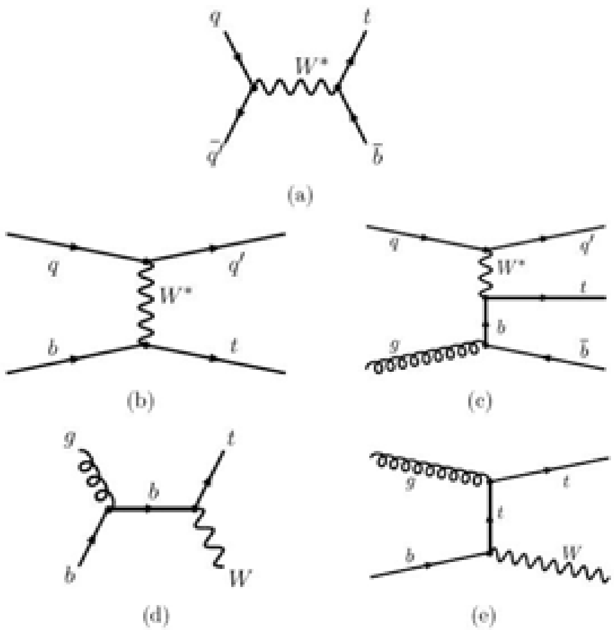


Figure 2.7: Feynman diagrams representing the single top electroweak production processes: (a) s-channel; (b,c) t-channel; (d,e) associated production.

2.3 Single top production

In addition to the production of top pairs via strong interaction processes discussed in the previous section, the top quark can also be produced singularly via weak interaction through the Wtb vertex, while Wts and Wtd vertices are strongly CKM suppressed. According to the Standard Model, there are three mechanisms for the single top production (the respective Feynmann diagrams are illustrated in Figure 2.7):

1. **t-channel production**: two cases can be distinguished. A first case, describes the coupling of a virtual W boson with a b-quark of the proton sea and the consequent production of a top quark: $qb \rightarrow q't$ (where q and q' represent light quarks). The second t-channel process, also known as Wg fusion, is a contribution at higher order of $\mathcal{O}$ (α_S) in which the splitting of a gluon into a pair $b\bar{b}$ occurs, followed by the interaction of the b-quark with the W boson: $qg \rightarrow tq'\bar{b}$;

2. **Associate production of a top quark and a W boson**: this process can be expressed as $gb \rightarrow tW$, where a real W is produced in association with the top quark from a gluon and a b-quark ($q^2 = m_W^2$);

3. **s-channel production**: Top quark production in this channel is due to the exchange of a charged W: the diagram corresponding to this process is achieved by rotating Wb scattering diagram in the t-channel.

The processes that have been described above have to be considered together with the respective processes obtained by charge conjugation. The three processes differ amongst them both in initial and final state and they are classified by referring to the final state: the s-channel is denoted tb; the t-channel as tq , tqb and the associated production with Wt. Correspondingly, their signatures can be used to discriminate between the different modes of production: the s-channel is characterized by a additional b-quark product together with the top; the t-channel from a light quark forward and the associated production (Wt) from the W and the top decay products. Since the associated production involves the decay of an additional W boson produced in association with the top quark, this signature differs slightly from the t-channel and s-channel production. The W decay provides either an additional jet or an extra charged lepton and $E_{\mathrm{T}}^{\mathrm{miss}}$ compared to the other single top production channels. This signature is only relevant at the LHC because the production cross-section at the Tevatron is negligibly small. Because of the b quark and the incoming gluon, the t-channel and Wt channel rates depend in no negligible way on their respective PDFs: therefore the measurements of cross sections in these channels constitute the opportunity to learn more about the PDFs of the gluon and the b-quark.

At LHC, the t-channel is the dominant process. Both the t-channel and s-channel processes have no dileptonic final state and the lepton+jets channel provides the cleanest event signature. Although it is a weak process, the cross-section for production of single top is of the same order of magnitude as the production of $t\bar{t}$ pairs, (actually amounts to about half of that of $t\bar{t}$ production according to approximate (N)NLO calculations [11]).

The measurement of single top production makes accessible the study of weak interactions of top quark: it is possible to investigate directly the $|V_{tb}|$ CKM matrix element, that is one of the less accurately known quantities related to the top quark. $|V_{tb}|$ determines the coupling strength at the Wtb vertex in addition to the universal electroweak coupling constant. Since the cross-section in all single top production modes is proportional to $|V_{tb}|^2$, under the assumption of CKM unitarity, its measurement constitutes the only way to determine directly the absolute value of $|V_{tb}|^2$.

Due to its short lifetime, the top quark keeps unchanged the polarization acquired at the production time, so that highly polarized top-quarks are expected in single top production. Detailed observations of the top-quark polarization can hence test the $V-A$ structure of the top quark charged-current weak interaction by studying the angular correlations between the decay products, or to reveal new production mechanisms through depolarization observations ([14]). Various extensions of the Standard Model predict eventual deviations from the standard Wtb interaction; both single top cross-section and polarization measurements are expected to be sensitive probes for the search of anomalous top quark couplings and to look for new physics beyond the Standard Model. The s-channel is sensitive to the existence of an exotic charged boson, which could be a W' boson or a charged Higgs, coupled to the $top - bottom$ weak isospin doublet, which could be detected through a eventual increase of the observed cross-section. The production rate in t-channel could eventually increase due to flavor changing neutral current (FCNC) processes, suppressed by the Standard Model, affecting couplings between *up*-type quarks and a neutral boson (a Higgs boson, a gluon, a photon or a Z). Couplings of this type would be difficult to be experimentally observed in the s-channel because the absence of the b quark, which is essential for the reconstruction of their final state. Finally, the Wt channel is the only way that would allow a more direct verification of the Wtb Standard Model vertex, since it contains a non-virtual W boson.

Although the rate of production is not so different from the production of $t\bar{t}$ pairs, the signature of the single top production is more difficult to separate from the background processes, which is why the first steps of this signal were made only very recently. While the t-channel process has been measured separately at the Tevatron, the Wt process has never been observed before. LHC publications and preliminary results of almost all these measurement are already available with the statistics collected at 7 TeV and also some at 8 TeV of center-of-mass energy ([11], [13], [12]).

2.4 Measurement of top properties at LHC

2.4.1 $t\bar{t}$ cross-section

At the LHC, in pp collisions with $\sqrt{s} = 7$ TeV, an approximate production cross-section has been calculated at NNLO using CTEQ6.6 parton distribution functions with an uncertainty of less than 10%. At these energies about 80% of the production is from gluon-gluon fusion. The resulting theoretical prediction of the top quark cross-section is $\sigma_{t\bar{t}} = 165^{+11}_{-16}$ pb, assuming a top quark mass of 172.5 GeV/c^2 [27]. Measurements of the top quark production cross-section are good tests of perturbative QCD. For these reasons, measurements of the top-quark production cross sections are important components of the LHC and Tevatron physics programs. The $t\bar{t}$ cross-section has been measured in many different channels with the data of both accelerators. The last update on such measurements as a function of center-of-mass energy is summarized in the plot shown in Figure 2.8, which takes into account also the most recent 8 TeV ATLAS result.

The aim of this thesis is to provide a measurement of the $t\bar{t}$ cross-section in the $\tau + jets$ channel, which only recently became feasible because of the need of a significant amount of data. The measurement presented here has been done for a LHC center-of-mass energy of $\sqrt{s} = 7$ TeV and for an integrated luminosity of about 2.05 fb^{-1} of data collected in 2011. Details of this analysis are given in Chapter 6.

2.4.2 Top mass

The top quark mass (m_{top}) is a very interesting property to be studied, first because the large difference of its value with respect to other quarks masses (it s roughly forty times greater than b-quark mass, and over 10 000 times more than the up quark mass) points out its not trivial role in the Standard Model mechanisms. Since in the Standard Model the top mass is a free fundamental parameter, it must be experimentally determined. The top quark mass contributes to electroweak radiative corrections Δr related to higher order diagrams with loops, with terms proportional to m_{top}^2/m_Z^2. This implies that its contribution is greater than the Higgs mass contribution, proportional to $\log(m_H/m_Z)$. Furthermore, Δr is related to W boson mass through the relation:

$$m_W^2 = \frac{\pi\alpha}{\sqrt{2}G_F \sin^2_{\Theta_W}}(1 + \Delta r). \tag{2.7}$$

For these reasons, the top mass is used together with other electroweak observables to infer the Higgs boson mass in both SM and non-SM scenarios. In addition, the precise

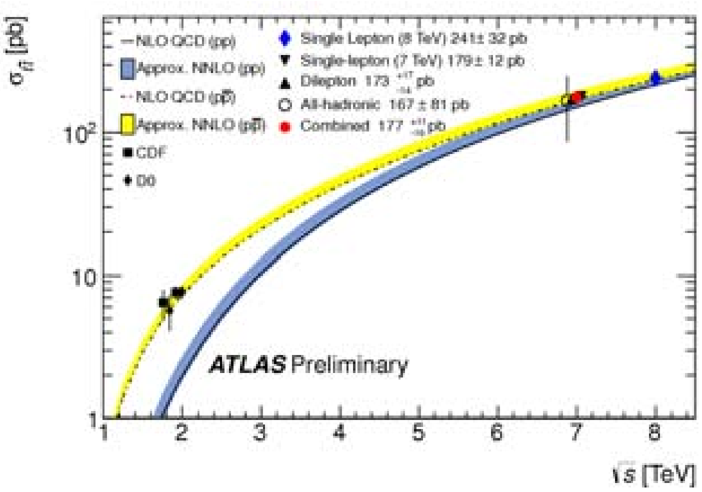


Figure 2.8: Dependence of top production cross section on the centre of mass energy $\sqrt{s}$ from theoretical predictions based on a top-quark mass of 172.5 GeV together with the dilepton, single-lepton, and all-hadronic measurements from ATLAS (LHC), as well as the combined measurement. Error bars represent measurement uncertainties (including statistical, systematic, and luminosity). Selected results obtained by Tevatron experiments are also shown.

measurement of this parameter has crucial importance in the determination of many other properties of the top quark. As an example, the dependence of the theoretical computation of $t\bar{t}$ cross-section from the top quark mass is $\sim 3\%/(GeV/c^2)$. The main methodology used to determine the top mass at hadron colliders consists of measuring the invariant mass of the decay products of the top quark candidates and deducing m_{top} using sophisticated analysis methods. The most precise measurements of this type use semileptonic the $t\bar{t} \to lepton + jets$ channel. Indirect measurements analysis are also ongoing, as for example those that parametrize the $t\bar{t}$ cross-section as a function of m_{top} and extract its values using a likelihood fit. A summary of the ATLAS results in measurements of top mass in the different channels until June 2012 (ICHEP Conference) is shown in Figure 2.9.

Due to the short top quark lifetime, the top quark is the only quark that can be studied before hadronization occurs; providing an unique opportunity to test the CPT symmetry conservation by measuring directly the mass difference between a quark and its antiquark. The advantage in this measurement is that almost all systematics affecting the m_{top} analysis cancel out in the Δm_{top} determination as they affect the measurement of m_{top} and $m_{anti-top}$ as well. The CDF, Dff and CMS collaborations have measured this difference to be in agreement with the Standard Model prediction of no difference, to a precision up to $\Delta m_{top}/m_{top} = 0.7\%$ [16].

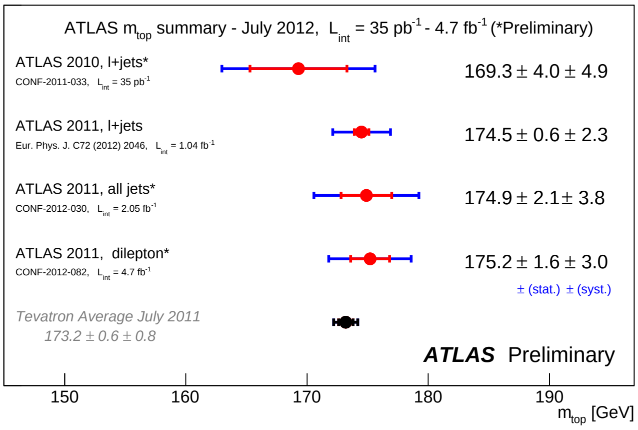


Figure 2.9: Summary of measurements of the top mass as of 30 Jun 2012 (ICHEP)

2.4.3 Other properties

Width. In the Standard Model, the total decay width of the top quark Γ_t described in Section 2.2, is dominated by the partial decay width $\Gamma(t \rightarrow Wb)$, which has been computed to be:

$$\Gamma_t = |V_{tb}|^2 \frac{G_F m_t^3}{8\pi\sqrt{2}}. \tag{2.8}$$

The total width ranges from 1.26 to 1.4 depending on the value used for top mass (170-175). Deviations from this value could be induced by decays of top quarks to non-SM particles such as scalar top partners, charged Higgs, or FCNC decays. The decay width of an unstable particle can be measured from its mass spectrum, or altenatively can be obtained from the equation $\Gamma_t = \Gamma(t \rightarrow Wb)/BR(t \rightarrow Wb)$, where $\Gamma(t \rightarrow Wb)$ is measured throught the single top cross-section and $BR(t \rightarrow Wb)$ is measured in top quark pairs events. The most precise measurement has been done by Dff using this latter method resulting in a Γ_t value of 0.90 ± 0.04, which deviates of $\sim 2.5\ \sigma$ from the Standard Model prediction.

W **boson helicity.** Another opportunity offered by the fact that top quark decays before hadronization, is the possibility to study directly an electroweak interaction vertex. Due to the V-A structure of Wtb vertex and the large mass of the W boson, the Standard Model predicts that 69.8% of the W produced in top decays are longitudinally polarized, while 30.1% have left-handed helicity and 0.041% are right-handed. The W longitudinal, right-handed and left-handed helicity fractions are extracted from the $cos\theta^*$ distribution, where θ^* is defined by the angle between the down type quark from the W decay and the opposite of the top direction in the W rest frame. Measurements of W helicity fractions have been done at Tevatron and also at LHC by ATLAS collaboration in both dilepton and lepton + jets channel with 0.7 fb^{-1} of integrated luminosity. Furthermore, ATLAS uses the same data to extract limits on anomalous coupling at the Wtb vertex.

Spin correlations. Although the $t\bar{t}$ pairs in the Standard Model are produced non-polarized, theirs spins are predicted to be significantly correlated. The spin of the top quarks cannot be measured directly in hadron colliders, but the angular distribution of their decay products depends on their spin polarization, since the decay occurs via electroweak interaction before hadronization. Hence, the consistency of Standard Model prediction can be probed through angular distribution measurements. If decorrelation would be measured, it could indicate that the top spins flipped before decay or else that a non-SM decay occurred ($t \rightarrow H^+b$, for example), in which the spin orientation information doesn t propagate. Complete decorrelation has been excluded by Tevatron measurements, even if the measurement is affected by large systematic and statistical uncertainties. At LHC the strength of the spin correlation is expected to be different: the predicted value at NLO calculation differs for $q\bar{q}$ annihilation and gg fusion processes. The preliminary results from LHC are not consistent with SM prediction [16].

Charge asymmetry. At leading order, the SM predicts $t\bar{t}$ production to be symmetric under charge conjugation. However, interferences from gluon emissions in initial and final states from higher-order $q\bar{q}$ annihilation diagrams (as $q\bar{q} \rightarrow t\bar{t}g$ and $qg \rightarrow t\bar{t}q$) result in a small asymmetry. At Tevatron charge asymmetry (A_C) can be measured by the detection of a forward-backwards asymmetry in the final states belonging to top and anti-top: it is predicted that the top quark will be emitted preferentially in the direction of the incoming quark and the anti-top in the direction of the anti-quark . Both CDF and Dff observed a non-zero forward-backwards asymmetry (A_{FB}), at least 2σ above the value predicted by the SM. A confirmation that such asymmetry is not compatible with SM prediction, it could indicate the existence of some resonant state with mass larger than twice the top quark mass, even if it doesn t appear in the $m_{t\bar{t}}$ spectrum due to kinematic limits of the accelerators. On the other hand non-Standard Model t-channel processes would also be observed as a forward-backward asymmetry. The symmetric nature of the initial state (pp instead of $p\bar{p}$) at the LHC means that there is no natural way of defining a *forward* or *backwards* direction, increasing the challenge of pursuing A_{FB} measurements. Another complication at LHC is the fact that the gluon-gluon fusion is the dominant process to produce $t\bar{t}$ pairs at $\sqrt{s}$ =7 TeV. Nevertheless, QCD predicts a small excess of anti-top quarks at central pseudorapidity values while top quarks are produced, on average, at higher absolute rapidities. This arises from the fact that because $t\bar{t}$ production via $q\bar{q}$ annihilation is dominated by initial valence quarks with large momentum fractions while anti-quarks coming from the proton sea have smaller momentum fractions. While in the $t\bar{t}$ rest frame the top are preferentially emitted in the direction of the initial quarks, the boost into the laboratory frame moves the top quark mainly in the forward or backward directions, while anti-top quarks are kept more in the central region. One of the definitions used for the charge asymmetry A_C is [26]:

$$A_C = \frac{N(\Delta|y| > 0) - N(\Delta|y| < 0)}{N(\Delta|y| > 0) + N(\Delta|y| < 0)}, \tag{2.9}$$

where $\Delta|y| \equiv |y_t| - |y_{\bar{t}}|$ is the difference between the absolute values of the top and anti-top rapidities and N($\Delta|y| > 0$) (N($\Delta|y| < 0$)) is the number of events with positive or negative $\Delta|y|$. ATLAS already provided A_C measurements with 7 TeV data in both lepton+jets [25] and dilepton [24] channels, all in agreement with Standard Model prediction of $A_C =$ 0.006.

2.5 Summary of new physics implications of top analysis

In order to summarize, the primary focus of top quark physics nowadays is the search for some evidence of physics beyond the Standard Model, in particular new physics that would help to explain the top quark s singular differences in comparison with the rest of the quark sector. Strategies in the search for new physics associated with the top quark can be categorized into the following two groups. First there are the direct searches for new physics associated with top quark production or decay, meaning for example the searches of new heavy resonances decaying into top quarks or searches of new particles produced in top quark decays. The other strategy category is to measure properties of the top quark predicted within the Standard Model, such as its production cross section, as well as its decay branching fractions, looking for deviations compared to the Standard Model predictions. For example, the strategy adopted in this thesis to measure the $t\bar{t}$ cross-section assuming the Standard Model branching ratios, could be used in the future in charged Higgs boson searches (see 7.3).

Furthermore, a fourth generation of heavy quarks it is not excluded by the SM fit to the existing precision measurement of electroweak observables, and would be an input to the understanding of CP violation in the Universe. These exotic quarks would appear in detectors very similarly to events with SM top quark production. Finally the supersymmetric theory (SUSY) which suggests that the supersymmetric partners of the third generation quarks could be the lightest SUSY squarks. The production and decay of stop quarks would appear kinematically similar to SM top quark production.

2.6 Signal and background modeling

There are many Monte Carlo programs which simulate top quark production. The ATLAS collaboration uses the Next-to-Leading-Order (NLO) Monte Carlo program MC@NLO. The parton distribution functions (PDF) set employed is the CTEQ one. In this thesis the next-to-leading order (NLO) generator MC@NLO v.3.41 is used to produce the $t\bar{t}$ signal sample, assuming a top mass of 172.5 GeV and with the NLO parton density function (PDF) set CTEQ66. To simulate parton shower all the tree-level computations are passed to HERWIG, which also account for hadronization; while the subroutine JIMMY describes the underlying events. The events with tau leptons are passed to TAUOLA package to simulate the tau decays. A full detector response simulation is the last step for the decayed particles and it is provided by the GEANT program. Finally pile-up events are also added to the primary collision.

With respect to the backgrounds, the main contribution in this analysis involving $t\bar{t}$ events decaying into one hadronic tau and jets comes from the W boson produced in association with jets. To simulate $W+jets$ production, and $Z+jets$ as well, the ALPGEN (v2.13) model is used. ALPGEN implements the exact LO matrix element for final states with up to 6 partons. $W+jets$ events with up to 5 partons and $Z+jets$ events with up to 5 partons and with the dilepton invariant mass $m_{ll} > 40$ GeV are generated using the LO PDF set CTEQ6L1. Separate samples of $W+jets$ events are generated to include $b\bar{b}$ and $c\bar{c}$ quark pair production at the matrix element level. In addition a sample of $W+c+jets$ events is produced. An overlap removal procedure is applied to avoid double counting of events between $W+light \quad jets$ and $W+b\bar{b}$ samples. Other smaller contributions, apart the already mentioned $Z+jets$, are the single top and dibosons events. The single top events are generated with MC@NLO as the $t\bar{t}$ making use of the diagram removal scheme to remove overlaps between the single top and the $t\bar{t}$ final states. Finally diboson events as of $WW+jets$, $WZ+jets$ and $ZZ+jets$ modeled with HERWIG. For all backgrounds the parton shower and hadronization is simulated with HERWIG and interfaced with JIMMY for the underlying event model.

The most difficult background to deal with in this analysis is nonetheless the QCD multi-jet background. Part of the complexity of its treatment is the difficulty to model these events with the Monte Carlo. For this reason, data driven methods are used to extract such background from specific control region of the data.

To estimate the systematic uncertainty related to the signal modeling, other $t\bar{t}$ Monte Carlo samples produced with different generators and interfaces have been used in this thesis, in particular have been considered: $t\bar{t}$ events produced with POWHEG generator interfaced with PYTHIA to estimate the uncertainty on the parton shower modeling; $t\bar{t}$

events produced with ALPGEN to estimate the uncertainty on the fragmentation simulation and specific AcerMC $t\bar{t}$ events to evaluate the effect of the initial and final state radiation (ISR and FSR) in the cross-section measurement. The problem of merging NLO calculations with parton shower simulations is the overcounting of events. The MC@NLO was the first proposed generator which provides an acceptable solution to that. In the current version of the MC@NLO code, the MC subtraction terms have been computed for HERWIG. Due to the subtraction procedure, MC@NLO can generate events with negative weights, (although that doesn t imply a negative cross-section), POWHEG was proposed as alternative solution. In the POWHEG method the hardest radiation is generated first, with a technique that yields only positive-weighted events [29]. Finally, AcerMC generator is dedicated for the simulation of the specific Standard Model background (like W and Z boson production in association with jets) characterized by the presence of the heavy flavor jets and multiple isolated leptons in the final state [30].

Chapter 3

The ATLAS experiment at the LHC

3.1 The Large Hadron Collider

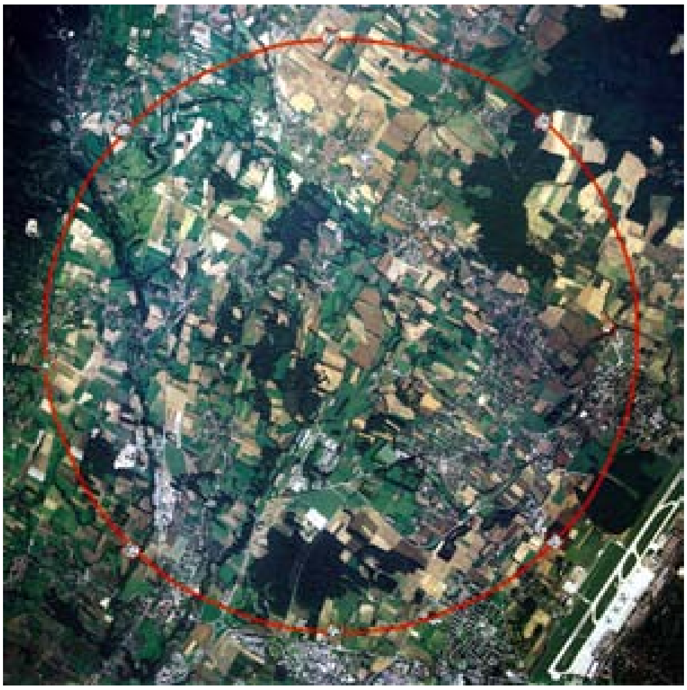


Figure 3.1: Areal view of the area where the tunnel of LHC can be found, at ~ 100 m underground (represented by the red circle). The ATLAS site is roughly at the 2 o clock position.

The Large Hadron Collider (LHC) is the gigantic particle accelerator placed at CERN laboratories at about 100 m underground, in the same tunnel that hosted in the past the LEP (Large Electron-Positron collider). The LHC project dates back to 1991 and was designed to extend research on the Standard Model, in particular, the Higgs boson searches for masses above the experimental limits fixed by LEP and Tevatron experiments. The first beam injection occurred on September 2008. In November 2010 started the data taken of the first collision data. In July 4, 2012 at CERN has been observed a particle consistent with the Higgs boson within 5 sigmas. Further searches are in progress to explore the proprieties of the new particle and to determine whether it is compatible with the Standard Model Higgs or what else.

LHC produces collisions between protons and heavy ions with a center of mass energy never achieved before. After the 2011 data taken at 7 TeV, which this thesis is based on, at the moment has been reached a center of mass energy of 8 TeV in pp interactions and it is expected to achieve the 14 TeV in 2014. While for the heavy ion collisions the energy that is expected to be reached is about 5.5 TeV per nucleon. The LEP/LHC tunnel circumference is about 27 km long and it hosts 1232 dipoles superconductors that create a magnetic field of 8.4 Tesla necessary to keep the protons of 7 TeV circulating on the orbit of $\sim$4.3 km radius. Before being injected in the main ring the two beams are accelerated by 4 smaller accelerators, which progressively increase the beam energy (see Figure 3.3). The first system is the linear accelerator LINAC, which accelerates protons up to a few tens of MeV. Subsequently, the protons pass to the Proton Synchrotron Booster (PSB) and then they are accelerated up to about 1 GeV by the Proton Synchrotron (PS).

Figure 3.2: First dipole installed in the LHC tunnel in 2005.

Finally they are injected into the Super Proton Synchrotron (SPS), where are accelerated up to 450 GeV before being introduced in the main ring. The packages of protons collide in the different collision points at intervals of 50 ns, that will be reduced to 25 ns during next year data taking. At such collision point are set the main detectors: ATLAS, CMS, LHCb and ALICE (see Figure 3.4). ATLAS and CMS detectors are dedicated to the study of the physics of the Standard Model and beyond; LHCb is aimed at the study of the physics of the Standard Model s third generation of particles, particularly the b-quark contained in B mesons; while ALICE experiment is dedicated to the physics of heavy ions interactions. Only the LHCb experiment uses one of the high energy proton beams in a fixed-target set-up.

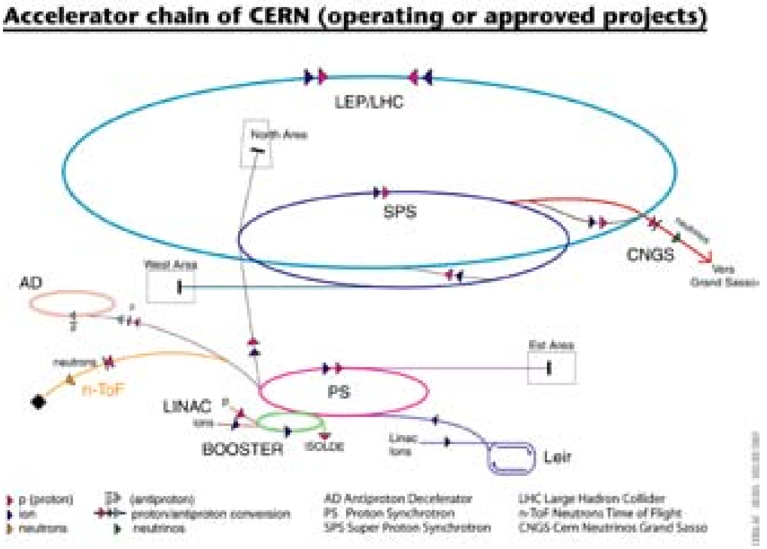


Figure 3.3: Map of the accelerators complex at CERN site.

After the accelerator commissioning, the first phase of LHC activity is taking place with a luminosity[1] of $10^{33} cm^{-2} s^{-1}$ (low luminosity phase), that have been gradually reached from a luminosity of $10^{31} cm^{-2} s^{-1}$. The project luminosity of $10^{34} cm^{-2} s^{-1}$ (high luminosity phase) is expected to be achieved in 2014 runs. The adoption of a hadron accelerator rather than a leptonic one is motivated by large emission of energy due to synchrotron radiation in the case of e^+e^- collisions. Indeed the energy lost by charged particles moving along circular orbits, according to the Eq.

$$\frac{dE}{dt} \propto \frac{E^4}{m^4 R} \tag{3.3}$$

This implies that, at a given energy and radius, the electrons lose a fraction of energy $(m_p/m_e)^4 \sim 10^{12}$ times greater than that lost by a proton beam having the same characteristics. The use of electron beams with the desired performance would require either a circular accelerator with a radius much greater than LEP, or a linear accelerator: both solutions much more expensive than the one adopted. Hadron accelerators present much more challenges with respect to lepton accelerators, due to the fact that protons are particles with an internal structure. Proton-proton collisions can be classified into two categories: *soft* and *hard* collisions. A soft collision acn be modelled as a proton interacting as a single block. These interactions are characterized by low moment transferred, then the particles

---

[1] The luminosity is defined as

$$\mathcal{L} = f \frac{n_1 n_2}{4\pi \sigma_x \sigma_y} \tag{3.1}$$

where σ_x and σ_y are the horizontal and vertical beam size (the particles in the beam are supposed gaussian distributed); f is the collision rate and n the number of particles per beam The expected number of events for a certain process at luminosity $\mathcal{L}$ is found as:

$$N = \mathcal{L} \times \sigma \tag{3.2}$$

where σ is the cross-section of the process itself.

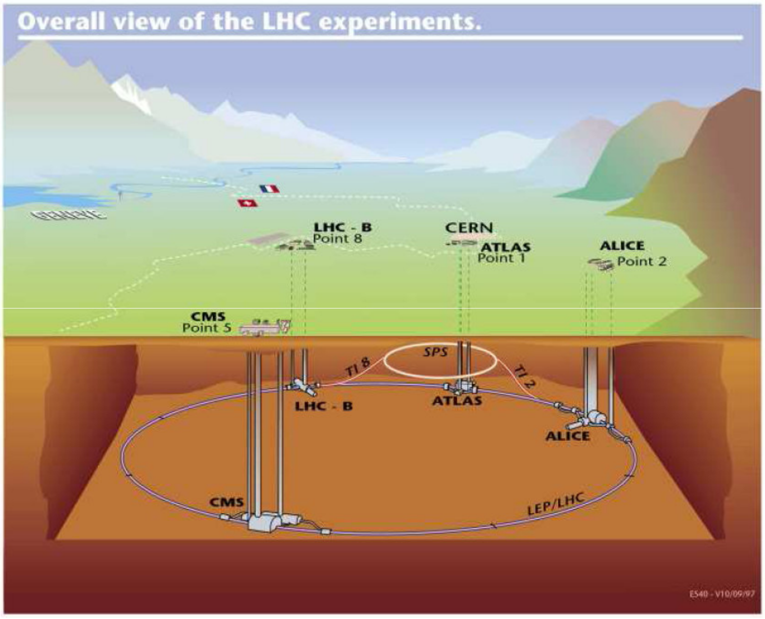


Figure 3.4: Map of the main experiments at the LHC: ATLAS, CMS, LHCb and ALICE.

outgoing from the interaction point have low transverse momentum ($< p_T > \simeq 500$ MeV). The events generated by soft interaction are defined as *minimum bias events* and constitute the majority of the events. In hard collisions instead the internal structure of the proton plays an important role in the interaction: the collision is considered as elementary QCD process between two constituents partons. These interactions are characterized by high transferred momentum allowing the production of new particles. One limitation of hadron accelerators is due to the fact that, since the fraction of the proton momentum carried by each parton is unknown, the center of mass energy of a single partonic interaction is not known as well. A second problem is that the hard collisions, significant for the discovery of new physics, are characterized by a cross section extremely low compared to soft processes cross-section. Furthermore, at high luminosity for each hard interaction are produced in addition ~25 soft collisions per bunch crossing, that constitute the so called *pile-up* events. For these reasons, the detectors and the data acquisition system are designed in a way to be able to discriminate interesting events from background, requiring the identification of experimental signatures such as missing transverse energy or secondary vertices. Identifying such final states for interesting processes imposes further requirements on the integrated luminosity needed, and on the particle-identification capabilities of the detector. These benchmark physics goals are reflected into a set of general requirements for the LHC detectors.

3.1.1 LHC detector requirements

1. The high rate of overlapping events in the detector require a fast and sophisticate electronic system, capable of discriminate in time the events minimizing the effect of pile-up.

2. To spatially discriminate events handling with the particle fluxes, a fine granularity is needed, in order to proceed to reduce pile-up.

3. The detectors are designed to cover as much solid angle possible around the beams interaction point: ATLAS acceptance in the azimuth angle (ϕ) is almost total and the coverage in terms of pseudorapidity[2] ($|\eta| < 5$).

4. An efficient trigger system is crucial to select interesting events ensuring the required rejection factor, to allow to select the interactions of interest, reducing the initial event rate of 40 MHz to a 200 Hz rate of recorded events, as explained in Chapter 5.

5. The ability of identificate particles is one of the crucial points at the LHC: this determines the characteristics of the sub-detectors. For the identification of b-jets, for example, is needed an excellent resolution of the tracker detector to determine the secondary vertices position. On the other hand, the muon chambers must ensure the muon identification and a good resolution on their transverse momentum.

6. Data acquisition capability to properly record the large amount of data produced at the high luminosity of LHC, is another fundamental requirement.

3.2 The ATLAS detector

3.2.1 Detector overview

The work presented in this thesis is part of the ATLAS collaboration project, consisting of about 3000 people from 38 different countries working on the experiment. ATLAS (A Toroidal Lhc ApparatuS) is a general purpose detector, which main interest is the search of the origin of the mass at the electroweak scale. The detector has been optimized to have as much sensitivity as possible in the expected Higgs mass range. In this regard, the measures of electrons, photons and muons are done at high resolution, such as the determination of the missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) and energy of the jets in the calorimeters. Particular accuracy is dedicated to the b quarks and tau leptons identification. The ATLAS detector has a forward-backward symmetric cylindrical geometry with respect to the interaction point, being the longitudinal axis coincident with the direction of the beams. Like most

---

[2] The pseudorapidity is defined as:

$$\eta = -\ln \tan(\frac{\theta}{2}), \tag{3.4}$$

where θ is the polar angle measured from the beam axis.

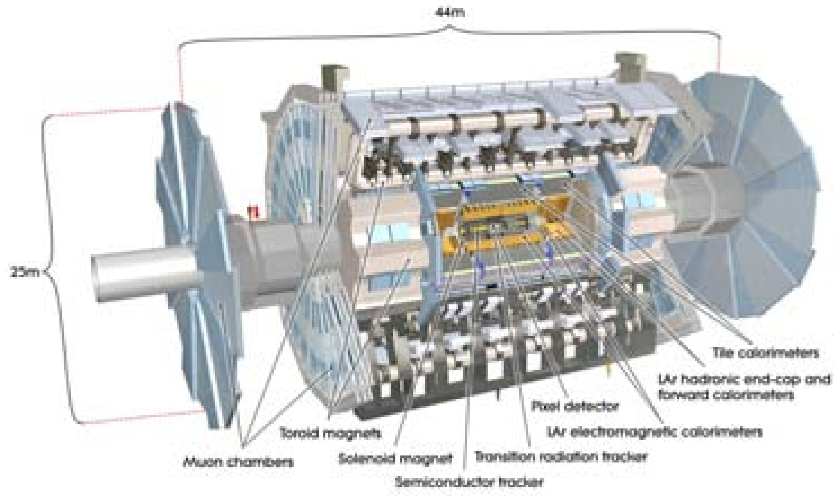


Figure 3.5: The ATLAS detector at LHC: sub-detectors overview.

of the detectors associated with modern colliders, ATLAS is composed by several sub-detectors, each of which has been designed and optimized for a specific task within the general objective: the required performance of the detector is therefore achieved thanks to the correct performance of each single part. In Figure 3.5 is represented a cut-view of the detector, in which sub-detectors are pointed out. They can be classified in three blocks:

1. **Inner tracking system:** The inner tracking system (Inner Detector) is placed close to the beam-pipe and includes multiple detectors: pixels and silicon microstrips in the inner part and straw-tube trackers with the capability of detect transition radiation in the outer part. This set of detectors is contained in a cylinder of 1.15 m of radius and 6.2 m large, which covers a range in pseudorapidity of $|\ \eta\ |< 2.5$. These detectors are characterized by an excellent spatial resolution that allows to identify primary and secondary interaction vertices. The Inner Detector is immersed in a 2 T solenoidal magnetic field, which allows to determine the moment of the charged particles, through the measurement of their radius of curvature.

2. **Calorimetric system:** The system consists of an high granularity liquid-argon (LAr) electromagnetic sampling calorimeter and a hadronic scintillator. The electromagnetic calorimeter is made up of the whole barrel + end-cap for the interval $|\ \eta\ |< 3.2$ and by a forward calorimeter for the range $3.2 <|\ \eta\ |< 4.9$. The hadronic calorimeter consists of a central cylinder and two extensions on both sides consisting of modules of scintillating material. Also the two hadron calorimeter end-caps are made with the liquid argon technology, as the electromagnetic calorimeter.

3. **Muon spectrometer:** The muon detection system is immersed in the magnetic field generated by an external toroid and is composed by a long barrel and two inserted end-cap magnets. The barrel ($|\ \eta\ |< 1$) consists of three concentric cylindric wire chambers of radius of 5, 7.5 and 10 m respectively, while the end-cap ($1 <|\ \eta\ |< 2.7$) consists of 4 concentric disks at a distance of 7, 10.8, 14 and 21.5 m respectively from

the interaction point. The outer radius of the spectrometer coincides with ATLAS maximum radius and is equal to 11 m; the outer end-cap disk is situated at 23 m from the interaction point.

In the next sections, the structure and the function of each class of sub-detectors is exposed in more detail.

3.2.2 The magnets

As mentioned above, the ATLAS detector employs two different magnetic fields: one internal for the inner detector and the other external for the muon chambers. The magnetic fields system is produced by an internal solenoid and eight external superconducting toroids.

- **Central solenoid:** the internal solenoid is constituted by a winding of superconducting material on a cylindrical support of internal (external) diameter 2.46 m (2.56) and length of 5.8 m. It is placed between the inner detector and the electromagnetic calorimeter and is able to generate a magnetic field of 2 T. To achieve the desired calorimeter performance, the solenoid layout has been optimize to keep the material thickness in front of the calorimeter as low as possible: the solenoid windings have been inserted inside the cryostat of the LAr calorimeter; in this way the thickness of the solenoid and the cryostat is only 0.66 radiation lengths (X_0).
- **External toroids:** the external magnetic field, needed by the muon spectrometer, is generated in the barrel by a toroid consisting in a system of eight superconducting coils which sizes are 25.3 m in lenght, 9.4 m of internal diameter and 20.1 m of outer diameter. The coils are arranged along the beams direction and each of them is provided by its own cryogenic system. Two end-cap toroids generate the magnetic field required for optimizing the bending power in the end-cap regions of the muon spectrometer system. The windings of the end-cap toroids are inserted in a single cryostat each. The overall magnetic system, sketched in Figure 3.6, provides a toroidal magnetic field of 0.5 T in the central region and 1 T in the end-cap.

3.2.3 The inner tracker

Contained in the central solenoid that produces a magnetic field of 2 T, the ATLAS Inner Detector (ID)is designed to provide hermetic and robust pattern recognition,excellent momentum resolution and both primary and secondary vertex measurement for charged tracks above a given p_{T} threshold. The Inner Detector (ID), which layout is shown in Figure 3.8, is contained within a cylindrical envelope of length ~3.5 m and of radius ~1.15 m. It shares with the central solenoid a single cryostat. The coil of the solenoid is designed to be as thin as possible, without sacrificing its reliability. Due to the large number of tracks that is produced in LHC high luminosity collisions, the resolution requested for the vertices and the momentum can be achieved thanks to the high granularity that characterizes the sub-detectors system. The semiconductor trackers, *i.e.* the silicon microstrip (SCT) and silicon pixels, which reconstruct the particle tracks using concentric detection layers, fulfill these characteristics. In particular, the finest granularity is offered by the

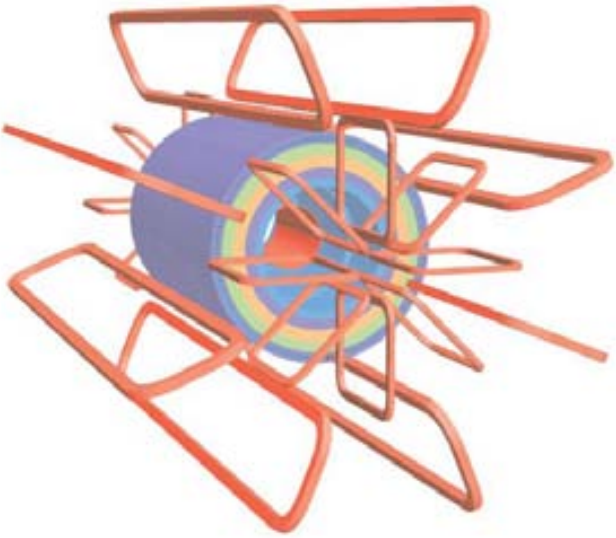


Figure 3.6: Sketch of the geometry of the magnet windings. The solenoid winding lies inside the calorimeter volume.

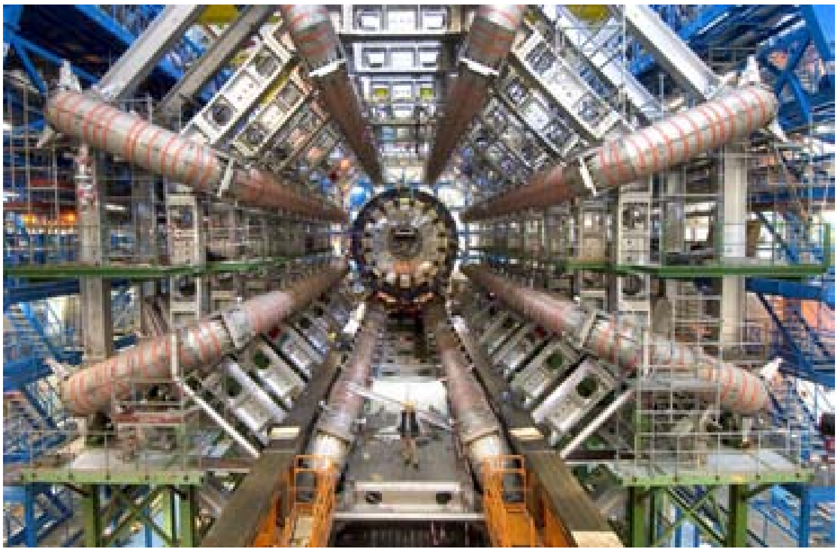


Figure 3.7: Magnetic toroids system installed in the ATLAS cavern.

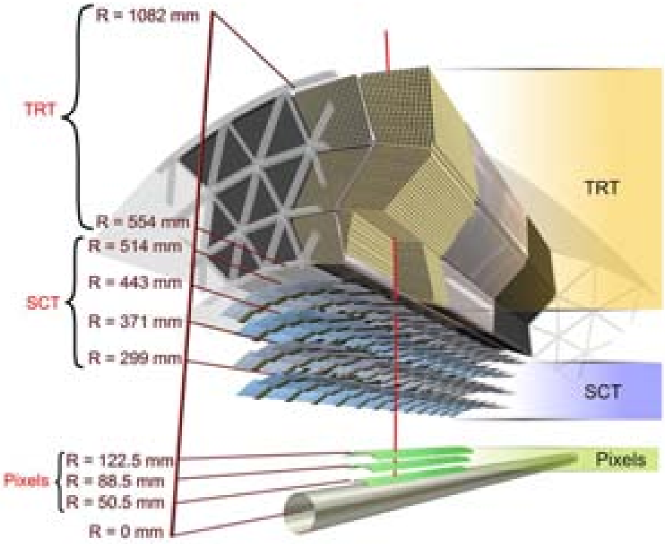


Figure 3.8: The ATLAS Inner Detector scheme.

pixel detector, positioned in the region immediately close to the interaction point. The high-radiation environment imposes stringent conditions on the inner detector sensors, but also on the electronic devices and the mechanical structures. The number of layers in which a precision measurement is performed should be limited both by the huge amount of material involved and for the high cost of so sophisticated detectors. The overall structure of the three ID sub-detectors is projected in a way to allow a continuous tracking, crossing small thickness of material and low costs. A particle track first crosses some pixels and microstrip layers, and then a large number of measurements and longer measured track length are provided by the drift tubes of the Transition Radiation Tracker (TRT). The combination of these two approaches provides a good resolution in tracks reconstruction. The TRT information contributes significantly to the measurement of the momentum, even if the accuracy on the single point position is lower with respect that measured in the silicon, because this is compensated by the large number of points provided. The relative accuracy of each point is well balanced, so that none of them predominates over the others in the momentum resolution. This increases the robustness of the performance. The positioning of the individual elements must be known with an accuracy greater than that given by the intrinsic resolution of the detector in order to not introduce an error greater than 20% on the tracking parameters. What leads to the requirement for the pixel detector to be aligned in R-ϕ within $\simeq 7\mu m$ and the detector SCT within $\simeq 12\mu m$.

Pixel detector

The pixel detector (Figure **??**) is designed to provide for each track three high-resolution points as close as possible to the interaction point. It is composed by a cylindric barrel

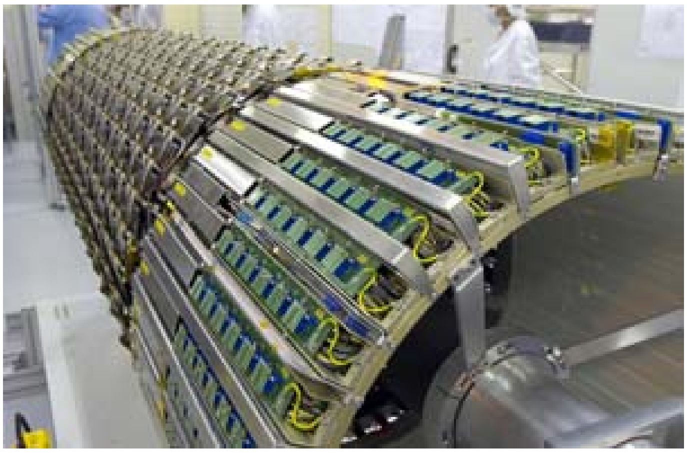


Figure 3.9: View of the silicon microstrips of the semiconductor tracker detector (SCT).

and two end-caps. The barrel is made up of three concentric layers with radius of 5.05 cm (B-Layer), 8.85 cm (Layer 1) and 12.25 cm (Layer 2) respectively, with an average length of 80 cm. The disks are also composed of three layers placed at increasing distance from the center of the barrel. The first disk is placed at $z = 49.5$ cm, the second one at $z = 58.0$ cm and the third at $z = 65.0$ cm. All pixel sensors are identical and have a minimum dimension per pixels in $R - \phi \times z$ of $50 \times 400 \mu m^2$. On the capability of this detector mainly depends on the tracking and the subsequent reconstruction of secondary vertices of short lifetime particles, as B mesons and τ leptons. Since the pixel is the closest detector to the interaction point, it is also required to be highly resistant to radiation.

SCT (Semiconductor Tracker)

The SCT is designed to provide eight track points at intermediate distance from the interaction point, in order to contribute to the momentum and the impact parameter measurements. The system, based on silicon microstrips (see Figure 3.9), has a surface area of 61 m^2 and is almost two orders of magnitude larger than the previous generation of semiconductor detectors such as those used in LEP; also the capability to resist to radiation is much higher, although this implies limits to the spatial resolution. The SCT barrel is constituted by four double layers of microstrip with radius between 30 and 60 cm; each double layer consists of strips aligned along the azimuthal direction and strips rotated by 40 mrad with respect to the first set, in order to provide measures of the coordinates $R - \phi$ and z. The angular coverage of this system is limited to $|\eta| \leq 1.4$ but in the end-cap regions the detector presents a set of strips running radially and a set of stereo strips at an angle of 40 mrad that increase the angular coverage to $|\eta| \leq 2.5$. The strips have a width of 80 μm, are 12 cm long and have a spatial resolution of 17 μm in $R - \phi$ and 580 μm in z. A chain of electronic amplification, filtering, triggering and signal transport is associated to the strips. The SCT, as the pixel detector, also requires a cooling system to disperse the heat generated by the electronics.

TRT (Transition Radiation Tracker)

The TRT structure is based on the use of straw tubes detectors, which consent the identification of electrons through the transition radiation. The straw tubes consist of thin metal wires, of diameter of 30 μm each, inserted into a barrel of 4 mm in diameter which contains a mixture of non-flammable gases (70% Xe, 20% CO_2 and 10% CF_4) for a total volume of 3 m^3. Passing through one of the drift tubes, a charged particle ionizes the gas producing a discharge that is collected by the wire thanks to the electric field applied. The time between the instant at which the particle crosses the tube and the instant at which the signal is detected (*drift time*) allows to measure the coordinates of the point where the particle has passed through the tube. Typically 36 points are given for each track. The total number of TRT readout channels is about 351,000, each of which provides a measure of the time of drift giving a spatial resolution of 130 microns. Each sensor is isolated from the others; in such a way it is able to give an accurate response also in the extremely high rate of particles provided by the LHC. Other important advantages of this technique are the natural radiation resistance and the relatively low costs.

The combination of precision trackers (pixel and SCT detectors) at small radii with the TRT at a larger radius gives very robust pattern recognition and high precision in both $R-\phi$ and z coordinates. In Table 3.1 are shown the resolutions in $R-\phi$ and z of the various Inner Detector sub-systems.

System	Resolution (μm)	
	$R-\phi$	z
Pixel	~ 10	~ 120
SCT	17	580
TRT	130	-

Table 3.1: Inner Detector sub-systems spatial resolution. [32].

3.2.4 Calorimeters

The ATLAS calorimeter system follows the classical scheme of large composite detectors, which consists in an electromagnetic calorimeter followed (in the radial coordinate) by a hadronic calorimeter (see Figure 3.10). The purpose of the calorimeters is to measure the energy and the position of electrons, photons and hadronic jets coming from the interaction point. To do this, the initial position of the particle is first identified and then the spatial development of the relative shower is traced. The high luminosity working conditions and the signatures of the processes of physics interest require an excellent performance of the calorimeter, mainly in terms of granularity, response time and energy resolution. In the following, are presented in more detail the different class of calorimeters.

LAr electromagnetic calorimeter

The main features of an electromagnetic calorimeter, which plays an important role in its performance in the reconstruction of electrons and photons, are the energy resolution and the granularity. The granularity is defined in terms of calorimeter cell size in the $\Delta\eta \times \Delta\phi$

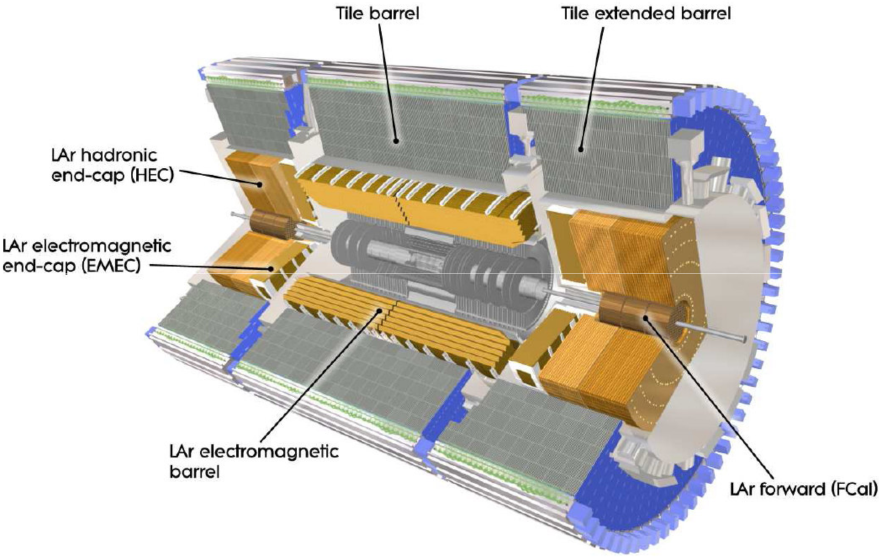


Figure 3.10: 3-D scheme of the ATLAS calorimeter system.

plane. The electromagnetic calorimeter resolution is given by the equation:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \tag{3.5}$$

where E is the energy of the particle expressed in GeV; the symbol $\oplus$ indicates summation in quadrature; a is related to the shower fluctuations; b is related to electronic noise and c is the so-called *constant term* due to non-uniformity of the detector or calibration systematic errors. The technology chosen for the electromagnetic calorimeter consists in liquid argon ionization chambers, which features are suitable to the LHC experimental environment: good performance in terms of electronic noise (around 400 MeV), high radiation resistance and technological possibility to realize the granularity request of $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ that ensures the achievement of the desired level of energy resolution and high performance in terms of tracking. The liquid argon ionization chamber has however a disadvantage: the signals that it produces have a time size (about 400 ns) greater than the time range between two successive collisions and this contributes to the pile-up. The electrodes of the electromagnetic calorimeter are characterized by a peculiar geometry. In standard ionization chambers the plan where the electrodes stand are placed perpendicularly to the direction of incidence of particle and thus in a direction parallel to the axis of the detector. This configuration would implies that the areas of the calorimeter used for the passage of read-out channels would not be useful for detection. In addition this technology would not allow the fine granularity required for the ATLAS calorimeter. For these reasons, has been studied a geometry of the electrodes and the detector said *accordion* (see Figures 3.11

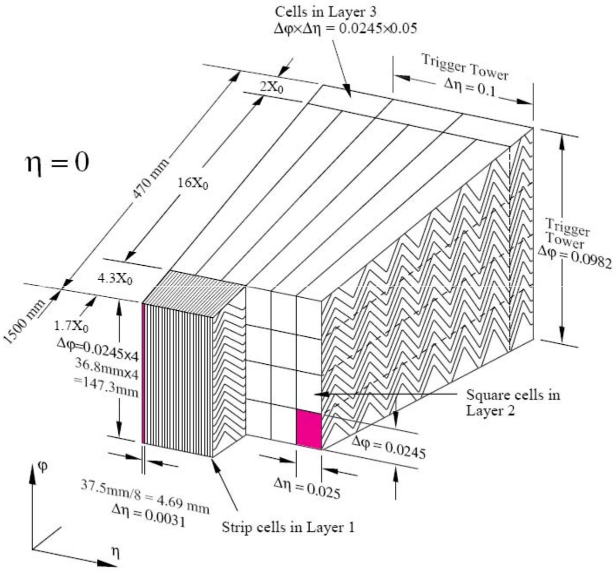


Figure 3.11: Sketch of the *accordion* geometry of the EM calorimeter.

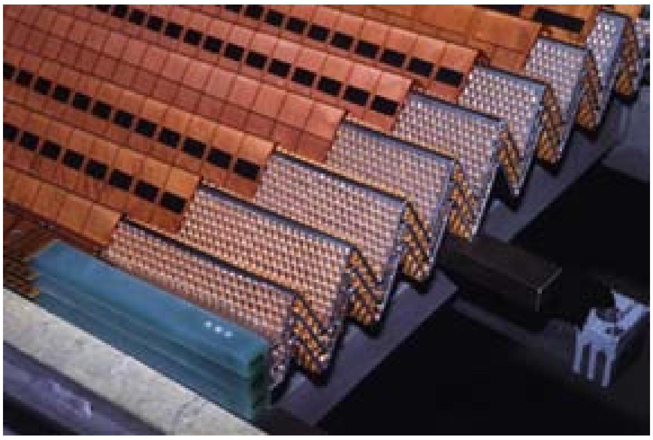


Figure 3.12: Detail of the *accordion* structure of the EM calorimeter.

and 3.11): readout electrodes and the absorber elements are arranged in radial direction and folded like an accordion. This geometry allows to achieve a complete symmetry in the ϕ direction avoiding the introduction of dead space. Thanks to this fact, to follow the shower radially it is not necessary to link the different planes of electrodes using cables, since the accordion electrodes themselves operate as transmission channels. In this way, a particle couldn t pass through the calorimeter without being detected. The curvature of the electrodes varies with the distance from the center: this allows to keep as constant as possible the distance between electrodes and absorbers that otherwise would tend to increase. The calorimeter is segmented into η and ϕ for a total of 190,000 cells suitable for the measurement of electromagnetic showers. The segmentation in the ϕ coordinate is obtained by connecting four electrodes to a single read-out channel: the coverage is guaranteed by 256 cells that lead to a granularity of $\frac{2\pi}{256} \simeq 0.025$ rad. With regard to the coverage in η, the division in cells is accomplished by photolithographic incision on the electrodes.
From the structural point of view, two parts compose the electromagnetic calorimeter: the barrel, in pseudorapidity region $|\eta| < 1.4$, and the end-cap, for $1.4 < |\eta| < 3.2$. Both calorimeter components are basically built with the same technology, liquid argon and accordion geometry, but differ in the spatial arrangement of the elements, according to the different η region which they are optimized for. The calorimeter barrel occupies a hollow cylinder with half length equal to ±6.65 m, inner radius of 1.15 m and outer radius of 2.25 m. Four regions can be distinguished.

The Presampler. The electrodes of the calorimeter itself are preceded by a module, said *Presampler*, made with the technology of argon liquid too. It is placed before the cryostat and therefore also inside the solenoid. The presampler consists of a liquid argon layer of thickness between 1.1 and 0.5 cm, without absorbers. The fine granularity of the presampler allows, during off-line reconstruction and through a proper calibration, to recover the energy loss due to material placed in front of the calorimeter.

The Strips (Sampling 1). This module terminates at a distance of $6X_0$ from the beam interaction point (including the contribution of all material that precedes it) and is characterized by a fine granularity in η.

The Middle (Sampling 2). This sampling covers the radiation length range between 6 and $22X_0$. Its goal is to completely contain the electromagnetic showers related to γ/e with energies up to 50 GeV. Therefore, the maximum deposition of electromagnetic energy is expected to be found in this module.

The Back (Sampling 3). This module consitute an addition of $2X_0$ in units of radiation lenght at the end of the calorimeter: it can be used together with the hadronic calorimeter for the identification and characterization of hadronic jets.

The importance of this segmentation is that allows the study of the longitudinal development of the shower and increases the ability of the detector to identify electrons, photons and pions. The end-caps are electrode systems constituted by two concentric wheels, one for $1.4 < |\eta| < 2.5$ and the other for $2.5 < |\eta| < 3.2$. In the electromagnetic calorimeter geometry there are three pseudorapidity regions, called *cracks*, in which the detector response is subject to a certain deterioration:

- at $\eta = 0$, the gap between the two semi-cylinder of the barrel;
- at $1.37 <\mid \eta \mid< 1.52$, the transition region between barrel and end-caps, where the inner detector cables are placed;
- at $\eta = 2.5$, the region between the two wheels of the end-cap.

In Table 3.2, the granularity in the barrel and end-cap regions and for each sampler is shown.

	Barrel	**Endcap**			
eta range	0 - 1.475	1.375 - 1.8	1.8 - 2.0	2.0 - 2.5	2.5 - 3.2
Presampler	0.025×0.1	0.025×0.1			
Sampling 1	0.003×0.1	0.003×0.1	0.004×0.1	0.006×0.1	0.1×0.1
Sampling 2	0.025×0.025	0.025×0.025	0.025×0.025	0.025×0.025	0.1×0.1
Sampling 3	0.05×0.025	0.05×0.025	0.05×0.025	0.05×0.025	

Table 3.2: Granularity expressed in $(\Delta\eta \times \Delta\phi)$ of the electromagnetic calorimeter, sampler by sampler.

The hadronic calorimeter

The main function of the hadron calorimeter is the contribution to energy reconstruction of the hadronic jets produced in the interactions. It is composed of a central barrel, two extended barrel and a liquid argon end-cap for each side. The central barrel is 5.8 m long and has inner radius of 2.28 m and outer radius of 4.25 m; it covers the pseudorapidity region of $\mid \eta \mid< 1$. At the two extremity are placed two cylindric extended barrels of the same radial dimensions and length of 2.6 m, extending the acceptance up to $\mid \eta \mid= 1.7$. Both the central barrel and the two extended barrel consist of pre-assembled modules composed by alternating plans of absorbing material (steel) and plans of scintillating material. The hadron calorimeter end-caps consist of liquid argon ionization chambers and cover the region $1.5 <\mid \eta \mid< 3.2$. They are constituted by plane copper electrodes. The advantage of this technology stands in its high resistance to radiation and in the relatively low cost. Moreover, it allows fine-granularity (0.1×0.1) at high η region, which fact offers an optimal pile-up reduction.

The forward calorimeter

The forward calorimeter is designed to cover the pseudorapidity region of $3.2 <\mid \eta \mid< 4.9$. In this region extremely high doses of radiation as well as considerable multiplicity of tracks are expected. The calorimeter consists in three modules for each end-cap: the first in z direction is made of copper and it is optimized for electromagnetic measurement, while the other two, made of tungsten, measure predominantly the energy of hadronic interactions. The calorimeter technique is again the liquid argon technique, but the the electrodes structure is completely different from the rest, consisting of concentric rods and tubes parallel to the beam axis. The LAr in the gap between the rod and the tube is the sensitive medium.

	η coverage	**Number of samplings**	**granularity ($\Delta\eta$ x $\Delta\phi$)**
Hadronic tile			
Barrel	$\lvert\eta\rvert < 1.0$	3	sampling 1= 0.1×0.1 sampling 2= 0.1×0.1 sampling 3= 0.2×0.1
Extended barrel	$0.8 < \lvert\eta\rvert < 1.7$	3	sampling 1= 0.1×0.1 sampling 2= 0.1×0.1 sampling 3= 0.2×0.1
Hadronic LAr			
	$1.5 < \lvert\eta\rvert < 3.2$	4 4	$1.5 < \lvert\eta\rvert < 2.5 = 0.1 \times 0.1$ $2.5 < \lvert\eta\rvert < 3.2 = 0.2 \times 0.2$
Forward			
	$3.1 < \lvert\eta\rvert < 4.9$	3	0.2×0.2

Table 3.3: Geometry and granularity ($\Delta\eta \times \Delta\phi$) of the different modulus of the hadronic and forward calorimeters.

3.2.5 Muon spectrometer

Muons are the only particles of those that cross the ATLAS detector, which penetrate enough to pass through the calorimeters and reach the external spectrometer. Their momentum is measured by the curvature of their trajectory in the magnetic field produced by the toroidal magnet system. Given the high interaction rate, the spectrometer must fulfill strict constraints in terms of response time and tolerance to radiation. It also has to be able to operate in difficult background conditions, determined by both penetrating

Figure 3.13: Installation of a part of one of the muon spectrometer wheels.

particles produced in the primary collision, and the background radiation mainly due to neutrons and photons in the MeV energy range coming from secondary interactions. The spectrometer for muons hosts in a hollow cylinder of radial dimensions of 5 m (inner radius) and 10 m (outer radius), and its total length is ± 23 m from the interaction point. It is designed to detect charged particles with $|\eta|$ <2.7. The spectrometer is immersed in the magnetic field generated by the external toroid that, in the central cylindrical region, is oriented in a way that the Lorentz force is along the z axis (the beams direction). The detector chambers are arranged on three cylinders, concentric with the axis of the beam. In the external disks region instead, the Lorentz force is along R and the chambers are arranged in four disks concentric with the beam axis. On almost the entire range of rapidity of the spectrometer, the Monitored Drift Tubes (MDTs) provide an accurate measurement along the curvature direction. They ensure a spatial resolution of about 80 μm. At large values of η and in the nearby the interaction point as well the MDTs are replaced by multiwire proportional chambers with cathodes segmented into strips (Cathode Strip Chambers, CSCs). Characterized by higher granularity then MDTs, the CSCs are more suitable to high counting rates and to the difficult underlying conditions that characterize these regions. In the central cylinder, the trigger system is composed of resistive plate chambers (RPCs), while in the external disks multiwire proportional chambers are used, denominated Thin Gap Chambers (TGCs) because they are characterized by a small separation between the anode and the cathode. The trigger system provides bunch-crossing identification and, therefore, the time resolution is required to be better than the time spacing between bunches (25 ns, at project conditions). Physical events used to determine the trigger are selected through well-defined p_{T} cuts, optimized to reject the background. The trigger system is also able to measure the muon coordinate in the direction orthogonal to that reconstructed by the MDTs, with typical resolutions of 5-10 mm. From test-beam data, the resolution of the toroidal muon spectrometer has been determined as explained in [32] :

$$\frac{\sigma_{p_t}}{p_t} < 10\% \qquad \text{per } p_t < 1TeV/c. \tag{3.6}$$

Chapter 4

Object reconstruction and identification in the ATLAS detector

In order measure the $t\bar{t}$ production cross-section in the $\tau + jets$ decay channel, we have to be able to select events with, in the final state:

- at least b-jets;
- 1 hadronic decaying tau lepton;
- a certain quantity of $E_{\mathrm{T}}^{\mathrm{miss}}$ that account for the two neutrinos produced in the W and tau decays.

The ATLAS detector provides a powerful reconstruction of such objects, thanks to the combined performance of its multiple components: the Calorimeters, the Inner Detector and the Muon Chambers, which in turn are composed of different sub-detectors (see Chapter 3). The information provided by these sub-detectors depend on their characteristics and on the particle that pass through it. For example, both electrons and photons lose energy passing through the electromagnetic calorimeter. But the electrons, that are charged particles, have also an associated track in the subsystems devoted to the particle tracking due to ionization, while the photons don t ionize and don t leave any track. So, the combination of calorimeter and tracker information let to distinguish between electrons and photons. The muons and the jets produced by gluons and quarks have a well defined signature, while the heavy flavor quarks (b, c) present a secondary vertex in their signature. Tau leptons have a secondary vertex too, but they are in general reconstructed from their decay product: leptons (electrons or muons) or jets. In the present work when we refer to tau we mean hadronic tau (τ_{had}), *i.e.* tau leptons decaying into hadronic jets. Starting from the information provided by the sub-detectors, the ATLAS software make the event reconstruction in three phases. The first step is the initialization, in which the information of the detector geometry and the magnetic field are uploaded. Then, the signals from each sub-detector is reconstructed according to the following scheme:

- In the Inner Detector, the signals from pixels and SCT are combined to determine the three-dimensional coordinates of the track of the particles.

- In the TRT, the points where the particle pass through are determined from the energy deposition position and the drift time of the charge. The reconstruction software is designed to identify the tracks, that are the helicoidal paths compatibles with the measured points reconstructed by each single plane of the detectors and yield to the optimal value of their parameters.

- In the calorimeters, the cells that provide a significant amount of energy are labeled *seeds* of the clusters reconstructed by the software and candidate to be tagged as electrons, photons, hadronic jets and tau jets. The clusters are then calibrated to provide the better energy estimation.

- In the muon spectrometer, analogously to the Inner Detector, each point position is calculated in different tracking planes and then combined.

- Finally, the information provided by the different sub-detectors is combined together.

In this chapter are described the techniques of reconstruction and identification of each object involved in the event selection we are interested in. In particular, the object reconstruction described in this chapter follows this order: tracks in Section 4.1; vertices in Section 4.2; muons in Section 4.3; electrons in Section 4.4; missing transverse energy in Section 4.5; jets in Section 4.6. Section 4.7 describes briefly the b-tagging techniques. Finally the last Section (**??**) is devoted to the hadronic tau reconstruction and identification; the identification efficiency and the energy calibration techniques are also described.

4.1 Track reconstruction

The track reconstruction is part of the reconstruction of the majority of the particles that cross the detector; in particular it is crucial for leptons and jet reconstruction. The tracks are reconstructed using the Inner Detector and the Muon spectrometer when it is needed. The Inner Detector reconstruction methods consists of three steps [36]:

- Pre-processing: *Space-points* are built from pixel and SCT hits and *drift circles* are reconstructed from TRT hits.

- Track-finding: the high granularity of the pixel detector is exploited, looking for prompt tracks originated from the interaction point. The track candidate is obtained combining the space-points from the three pixel layers with those from SCT; then the candidates are fitted. Using quality cuts on the clusters, the outliers are removed, cluster-to-track association ambiguities solved and mis-reconstructed tracks removed. The drift circle informations is used to solve the left-right ambiguities due to the extrapolation of the SCT track seed to include the TRT hits. Then the tracks are re-fitted with the information of the three inner sub-detectors.

- Post-processing: using the reconstructed tracks, primary and secondary vertices are reconstructed (see next Section)and photon conversions are identified.

4.2 Vertex reconstruction

A dedicated vertex finder is used to reconstruct the primary vertices, followed by algorithms dedicated to the secondary vertices reconstruction. In a collision in ATLAS, several vertices are produced, originating from the hard-scattering or other soft interactions. To distinguish the interesting event from the pile-up and minimum bias events, a primary vertex is searched between all these vertices. During the first run of LHC, the probability of pp collisions to pile-up in the same bunch-crossing was estimated to be 10^{-5}. Consequently, the primary vertex finder was configured to reconstruct exactly one primary vertex at each bunch-crossing. For data collected at 7 TeV, the amount of pile-up was estimated as 10^{-3}, which was no longer negligible. A new reconstruction strategy, based on iterative vertex finder and adaptive vertex fitter was used. The transverse momentum requirement for charged particle tracks was lowered to $p_{\mathrm{T}} > 100$ MeV. First, exactly one vertex is fitted from all the pre-selected tracks. Then, tracks incompatible by more than 7σ with the initial estimate are used to seed and reconstruct a new vertex candidate. The process is repeated until all available tracks are used or no new vertex seed can be created. The beam-spot information was used to constrain the vertex fit. The pile-up events were identified as triggered bunch-crossing where at least one additional primary vertex with at least 4 fitted tracks is reconstructed [39].

4.3 Muons

The muon identification starts from a track reconstructed in the muon spectrometer. If such track matches to an associated track into the Inner Detector, the muon is defined as combined muon , otherwise stand-alone muon . There are also the so called tagged muons , whose track is first reconstructed in the Inner Detector and then extrapolated to the Muon spectrometer. For each identification type exist two reconstruction algorithm: `MuID` or `Staco`. The former, combines from the beginning the track reconstruction provided by the Inner Detector and by the Muon Spectrometer; the latter, combines successively the information provided separately from the inner Detector and the Muon Spectrometer. For the scope of this thesis, we define the muon objects just to use them as a veto for taus and electrons and to apply the overlap removal to identified objects. As recommended for Release 16 analysis at $\sqrt{s}$= 7 TeV, have been used combined muons reconstructed with `MuID` algorithm.

4.4 Electrons

The reconstruction of electrons in ATLAS relies on the signal produced by two subdetectors the inner detector tracker and the electromagnetic calorimeter. The algorithm, which reconstructs clusters in the calorimeter and associates them to reconstructed tracks of charge particles in the inner detector, allows an optimal reconstruction of the four-momentum for the full momentum and pseudorapidity range and for all luminosities. Information from both detectors is used to identify electrons with the lowest possible amount of background, keeping in mind that the optimum between identification efficiency and background rejection depends on the analysis. Electron reconstruction starts with the creation of a preliminary set of *seed* clusters, built using a *sliding-window* algorithm. It consist of mapping the calorimeter cells into 3×5 in $\eta \times \phi$ cell units of dimensions 0.025 × 0.025. The fixed cell window is moved across the tower grid until a local maximum is found. A cluster seed is found when the energy in the window is greater than 2.5 GeV, summing the energy in all longitudinal layers. Duplicated clusters are removed from nearby seed clusters. In the region of tracker detector ($\eta| < 2.5$), an electron is defined by the existence of one or more reconstructed tracks matched to a seed cluster. The central part of the electron reconstruction consists then in the track-to-cluster matching. Reconstructed tracks are matched to seed clusters by extrapolating them from their last measurement point to the second layer of the calorimeter. The impact point coordinates (η , ϕ) are compared to the corresponding coordinates of the seed cluster: if the difference is below a certain threshold then the track is considered matched to the cluster. To take into account the Bremsstrahlung losses, the $\Delta\phi$ window is taken larger in the part where the track bends due to the magnetic field. In case of tracks that not contains silicon hits, since the η accuracy in TRT is not so good, only a matching restricted to the ϕ coordinate is required. In case that more than one track matches the same seed cluster (for example in the case of electromagnetic showers), all the tracks are retained ad ordered according to the quality of the matching, and the information from all the tracks is used in the identification step. The track with the smallest difference ΔR between its impact point on the electromagnetic calorimeter and the seed cluster is considered as the best match. In the collection of electron at this stage there is a significant contamination of non-prompt electron from photon conversion because all objects that have tracks matched to seed clusters are treated as electron. Then the particle identification criteria distinguish prompt electrons from converted photons. The final electromagnetic cluster size is defined using a window of 3×7 in the barrel and 5×5 cell units in the end-cap. In order to contain more energy, the cluster is taken larger in the ϕ direction, where the magnetic field curves the particle trajectory, as said above. In the end-caps, the cluster is taken symmetric in (η, ϕ) because the magnetic field is weaker. The reconstruction efficiency of low-p_{T} electrons (a few GeV) is extrapolated from the track to the electromagnetic calorimeter and a cluster is built using the track impact point as a seed. The electron identification step, consists on the definition of discriminant variables and apply cuts on them that provide good separation between separated electrons and jets faking electrons. These variables include calorimeter, tracker and combined calorimeter/tracker informations. Three reference sets of cuts are defined with increasing background rejection power: loose, medium and tight.

4.5 Missing transverse energy

In a collision experiment, like ATLAS, since the total momentum is expected to conserve, the missing transverse energy (E_T^{miss}) is defined as the momentum imbalance in the transverse plane. It is obtained from the negative vector sum of the momenta of all particles detected in a *pp* collision in the transverse plane. The missing quantity of transverse momentum can signal the presence of unseen known particles, such as neutrinos, or else the presence of new physics, such as stable weakly interacting supersymmetric particles. In the analysis presented here, large amount of missing transverse energy is expected in the final state, due to the presence of two neutrinos. A precise measurement of the missing transverse energy is then crucial also to distinguish the signal final state from other eventual processes that LHC is looking for, as the Higgs boson in $H \rightarrow \tau\tau$ decay channel. An important requirement of the measurement of $E_{\mathrm{T}}^{\mathrm{miss}}$is the minimization of the impact of limited detector coverage, finite detector resolution, dead material regions, noisy readout channels. The fake component of $E_{\mathrm{T}}^{\mathrm{miss}}$due to these sources has to be suppressed by severe cuts. For example in searches of Higgs boson decaying into two leptons, the final state can be mimicked by $Z \rightarrow ll + jets$ processes with a overestimated amount of $E_{\mathrm{T}}^{\mathrm{miss}}$. For this reason one of the goals to achieve is a $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction the most accurate than possible. In ATLAS, the $E_{\mathrm{T}}^{\mathrm{miss}}$reconstruction combine the contributions from energy deposits in the calorimeters with muons reconstructed in muon spectrometer, such that:

$$E_{x(y)}^{miss} = E_{x(y)}^{miss,calo} + E_{x(y)}^{miss,\mu}. \tag{4.1}$$

The calculation of each term will be described in the next two sections. Low-p_T tracks are used to recover low p_{T}particles which are missed in the calorimeters, while muons reconstructed from the inner detector are used to recover muons in regions not covered by the muon spectrometer. The value of $E_{\mathrm{T}}^{\mathrm{miss}}$and its azimuthal angle (ϕ^{miss}) are then calculated as:

$$\begin{aligned} E_{x(y)}^{miss} &= \sqrt{(E_x^{miss})^2 + (E_y^{miss})^2}, \\ \phi^{miss} &= \arctan(E_y^{miss}, E_x^{miss}). \end{aligned} \tag{4.2}$$

4.5.1 $E_{\mathrm{T}}^{\mathrm{miss}}$calorimeter term calculation

The $E_{\mathrm{T}}^{\mathrm{miss}}$reconstruction uses calorimeter cells calibrated according to the reconstructed physics object to which they are associated. Calorimeter cells are associated with a reconstructed and identified high-p_T parent object in a chosen order: electrons, photons, hadronically decaying τ-leptons, jets and muons. Cells not associated with any such object are also taken into account in $E_{\mathrm{T}}^{\mathrm{miss}}$calculation: their contribution, named $E_T^{miss,CellOut}$, is important in the $E_{\mathrm{T}}^{\mathrm{miss}}$resolution. The contribution of each physics object instead is calculated from the negative sum of calibrated cell energies associated to it, as:

$$\begin{aligned} E_x^{miss,obj} &= -\sum_{i=1}^{N_{cell}^{obj}} E_i \sin\theta_i \cos\phi_i, \\ E_y^{miss,obj} &= -\sum_{i=1}^{N_{cell}^{obj}} E_i \sin\theta_i \sin\phi_i, \end{aligned} \tag{4.3}$$

where E_i, θ_i and ϕ_i are the energy, the polar angle and the azimuthal angle of the cell. The summation is done on all cells in the pseudorapidity range $|\eta| < 4.5$, where data are well described by the MC simulation. Hence the total calorimeter contribution to $E_{\mathrm{T}}^{\mathrm{miss}}$ is calculated as:

$$\begin{aligned} E_{x(y)}^{miss,calo} = {} & E_{x(y)}^{miss,e} + E_{x(y)}^{miss,\gamma} + E_{x(y)}^{miss,\tau} + E_{x(y)}^{miss,jets} + E_{x(y)}^{miss,SoftJets} + \\ & E_{x(y)}^{miss,CellOut} + (E_{x(y)}^{miss,calo,\mu}), \end{aligned} \tag{4.4}$$

where, in particular:

- $E_{x(y)}^{miss,jets}$ is reconstructed from cells in clusters associated to jets with calibrated p_{T}> 20 GeV, while $E_{x(y)}^{miss,SoftJets}$ is reconstructed from cells in clusters associated to jets with 7 GeV < p_{T}< 20 GeV;
- $E_{x(y)}^{miss,CellOut}$ is the contribution of cells in topoclusters not associated to any reconstructed object;
- $E_{x(y)}^{miss,calo,\mu}$ is the contribution originating from the energy lost by muons in the calorimeter and it is not always included (see next section).

Because of the high granularity of the calorimeter, it is crucial to suppress noise contributions and to use only the cells containing a significant signal: this is achieved using topo-clusters, with the exception of electrons and photons for which a different clustering algorithm is used.

4.5.2 $E_{\mathrm{T}}^{\mathrm{miss}}$muon term calculation

The $E_{\mathrm{T}}^{\mathrm{miss}}$muon term is calculated as the summation of the momenta of the tracks associated to all the selected muons, reconstructed with $\eta < 2.7$:

$$E_{x(y)}^{miss,\mu} = - \sum^{muons} p_{x(y)}^{\mu}. \tag{4.5}$$

In the region $\eta < 2.5$, only well-reconstructed muons in the muon spectrometer with a matched track in the inner detector are considered (combined muons). This matching requirement considerably reduces contributions from fake muons (reconstructed muons not corresponding to true muons), sometimes created from very energetic jets. To deal appropriately with the energy deposited by the muon in the calorimeters, $E_{x(y)}^{miss,calo,\mu}$, the muon term of $E_{\mathrm{T}}^{\mathrm{miss}}$is calculated differently for isolated and non-isolated muons (those within a distance $\Delta R < 0.3$ of a reconstructed jet in the event):

- The p_{T} of an isolated muon is determined from the combined measurement of the inner detector and the muon spectrometer, taking into account the energy deposited in the calorimeters. In this case the energy lost by the muon the calorimeters ($E_{x(y)}^{miss,calo,\mu}$) is not added in the calorimeter term (Eq. 4.4) to avoid double counting of energy.
- For non-isolated muons, the energy deposited by the muon in the calorimeters cannot be resolved. Therefore the muon spectrometer measurement of the muon momentum after the energy loss in the calorimeter is used and the term $E_{x(y)}^{miss,calo,\mu}$ is added to the calorimeter contribution to the $E_{\mathrm{T}}^{\mathrm{miss}}$.

For the region of pseudorapidity outside the fiducial volume of the inner detector (2.5 < η <2.7), there is no matched track requirement and the muon spectrometer p_{T} alone is used for both isolated and non-isolated muons. Muons outside the acceptance of the muon spectrometer ($|\eta|$ >2.7) or in small inactive regions of the spectrometer (around $|\eta| = 0$ and $|\eta| \sim 1.2$) are lost. The muons reconstructed by segments matched to inner detector tracks extrapolated to the muon spectrometer are used to recover their contribution to

$E_{\mathrm{T}}^{\mathrm{miss}}$in the $|\eta| \sim 1.2$ region. Although the core of the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution is not much affected by the muon term, any muons which are not reconstructed, badly measured or fake, can be source of fake $E_{\mathrm{T}}^{\mathrm{miss}}$.

4.6 Jets

In a hard interaction, a jet is defined as the outgoing partons resulting in a shower of collimated particles, due to soft and collinear showering and hadronization. A jet algorithm is needed to match the hadronic and the partonic level jet; and the requirements to the algorithms are to be infrared safe (the reconstructed jet doesn t change with arbitrary soft radiation) and collinear safe (not sensitive to collinear splitting of the particle). Jets are reconstructed in ATLAS using the anti-k_T algorithm with distance parameter R= 0.4 or R=0.6. The excellent performance of such algorithm is due to this characteristic: soft particles tend to cluster with hard ones before they cluster among themselves. This fact ensures that soft particles don t change the shape of the jet, while the hard particles do it. At this point, we have to distinguish between hadron level jets and detector level jets. The hadron level jets in Monte Carlo simulation (MC truth jets), are built from the collection of event generator particles coming from the main hard scattering, before their interaction with the detector. On the other hand, the detector level jets are reconstructed for both Monte Carlo and Data from energy deposits in the calorimeter in form of topological cluster or calorimeter towers. In both hadron level and detector level, the four-momentum of the jet is the sum of the of its constituents four-momenta (detector level constituents are considered massless).

4.6.1 Topo-clusters formation algorithm

The clustering algorithm groups the cells with positive energy deposits and determine the energy released by the original parton. The topological clustering in three dimensions algorithm is designed to follow the shower development of a single particle interacting with the calorimeter, taking advantage of its fine granularity. The calorimeter noise is efficiently suppressed by this algorithm, because of its significance based criteria of clustering. The clustering starts from a *seed* cell with signal to noise ratio (S/N) above 4. The signal is defined as the absolute value of the energy deposited in the cell, while the noise is the quadratic sum of the electronic noise due to read-out (measured in empty bunch crossings), and the noise due to pile-up. Then, the cluster group criteria implies the inclusion of the seed neighboring cells (in three dimension) with S/N >2; then the new cells are iteratively used as seed too, until no more medium significance neighboring cells are found. Finally, all neighboring cells in the cluster perimeter are added to contain the shower tails. A splitting algorithm is used to separate showers from two distinct particles and double counting is avoid. At the end, the final topo-cluster have the following characteristics:

- variable number of constituent cells;
- energy equal to the sum of the cell contained in it;
- zero mass;

- reconstructed direction calculated from the energy weighted average of pseudorapidities and azimuthal angles of the constituent cells.

4.6.2 Jet energy scale calibration

In ATLAS the jets are reconstructed at the electromagnetic scale, that accounts correctly the energy deposited in the calorimeter by electromagnetic showers, but not the hadronic ones. For this reason, the EM+JES calibration is used to correct the energy and momentum of jets measured in the calorimeter to those of jet at the hadronic scale. The detector effects affecting the energy measurement and corrected by the calibration are summarized here below:

- calorimeter non-compensation: partial measurement of the energy deposited by hadrons (ATLAS calorimeter response is lower for hadron than for electrons);
- dead material: energy lost in non-sensitive regions of the detector;
- calorimeter leakage: energy deposits from particles not contained in the calorimeter;
- out-of-cone shower: energy deposits from particles in hadronic level jet not reconstructed in the detector;
- calorimeter clustering and jet reconstruction inefficiencies.

The EM+JES calibration scheme consists in the following subsequent steps:

1. **Pile-up correction.** Part of the jet energy is comes from multiple proton-proton interactions within the same bunch crossing of the event in interest (in-time pile-up). A correction for pile-up is derived from minimum bias data as a function of the number of reconstructed primary vertices, N_{PV} he jet pseudorapidity η and the bunch spacing.
2. **Jet origin correction.** Calorimeter jet directions are reconstructed using the geometrical center of ATLAS detector as reference. The jet four-momentum is corrected for each event such that the direction of each topo-cluster points back to the primary hard-scattering vertex. The improvement in angular resolution is smaller than 1% and the energy is unaffected.
3. **Jet energy final correction.** The final step of ATLAS EM+JES calibration consists of a jet-by-jet correction depending on the jet $p_{\mathrm{T}}$and η based on MC di-jet simulation. The goal is to restore the reconstructed jet energy to the hadron level jet energy (MC truth). First, the reconstructed jets are matched to truth jets within ΔR=0.3. Both are required to be isolated, i.e. not to have other calorimeter (truth) jet with EM-scale (truth) $p_{\mathrm{T}}>$ 7 GeV within ΔR=2.5R, where R is the distance parameter of the jet algorithm. The EM-scale response, defined as $\mathcal{R} = E_{EM,calo}/E_{truth}$, is measured in bins of p_T^{truth} and detector pseudorapidity η_{det} for each calorimeter-truth pair of jets. For each (E_{truth}, η_{det})-bin , the measured EM-scale energy response $\langle\mathcal{R}\rangle$ is defined as the peak of a Gaussian fit to the

$E_{EM,calo}/E_{truth}$ distribution, and the average $\langle E_{EM,calo}\rangle$ is determined. For each η_{det}-bin, the $(\langle E_{EM,calo}\rangle, \langle\mathcal{R}\rangle)$ points are fitted, having parametrized the response as

$$\mathcal{R}(E^{EM}_{calo}, \eta_{det}) = \sum_{i=0}^{N_{max}} \frac{a_i}{(lnE^{EM}_{calo})^i}, \tag{4.6}$$

where a_i are free parameters and N_{max} is chosen between 1 and 6 depending on the goodness of the fit. Finally, the EM reconstructed energy is corrected like this:

$$E^{EM+JES}_{calo} = \frac{E^{EM}_{calo}}{\mathcal{R}(E^{EM}_{calo}, \eta_{det})}. \tag{4.7}$$

Since the correction factor is applied to all components of the four-momentum, the jet direction is unchanged.

4.6.3 Jet energy scale uncertainties

The sources of JES systematic uncertainties can be summarized in the categories listed here below.

- **Uncertainty due to the JES calibrated method.** At low p_{T}, is observed a non-closure deviation in the jet energy and p_{T} response after the JES calibration in the Monte Carlo sample. This implies that the kinematic observables are not restored to the energy of the corresponding truth jet. The systematic uncertainty due to non-closure of the nominal JES calibration is taken as the larger deviation in either energy or p_{T} from unity.

- **Uncertainty due to the calorimeter response.** The calorimeter response uncertainty is estimated as a function of jet η and p_{T} from the propagation of single constituent particle uncertainties to the jet. *In-situ* measurements of particle response reduces the uncertainty due to the limited knowledge of the detector geometry due to dead material and of the exact modeling of the particle interaction in the detector.

- **Uncertainty due to the detector simulation.** Differences between simulated noise and real noise in the calorimeter cell can lead to discrepancy in the top-cluster shape and to the presence of fake clusters. While the noise in data continuously changes, the noise in MC is fixed to the RMS of the energy distribution in a cell at the simulation time and that leads to biases. Possible biases to the energy scale are due also to the simulation of the detector material. The uncertainty contribution due to additional material in the inner detector and overall additional dead material are estimated usind specific Monte Carlo samples produced with distorted geometries.

- **Uncertainty due to the physics model and the parameters employed in the Monte Carlo event generator.** The contribution of the JES uncertainty due to the modeling of the fragmentation and underlying event and other parameters of the Monte Carlo event generators are obtained comparing the combination of different generators and tune interfaces.

- **Uncertainty due to the relative calibration for jets in the end-cap.** The JES uncertainty is evaluated in six different eta bins. For all jets with $\eta > 0.8$ the uncertainty is derived from the uncertainty of the central barrel region ($0.3 < \eta < 0.8$) and then corrected with a relative calibration derived in di-jet events through a Matrix Method technique in data and in several MC event samples. The uncertainty from the relative inter-calibration is taken as the RMS deviation of the MC prediction from the data and is added in quadrature to the baseline uncertainty (the total JES uncertainty measured in the central barrel region).

4.6.4 JES uncertainties for heavy quark jets

Heavy flavor jets, in particular those induced by bottom quark (b-jets) play a very important role in this analysis, because of the presence of at least two of them in the $t\bar{t}$ process. The uncertainty in the calorimeter response to b-jets is evaluated using single hadron response measurement in samples of inclusive di-jet and $b\bar{b}$ di-jets events. The uncertainty arising from the modeling of the b-quark production mechanism and the b-quark fragmentation can be determined from systematics variations of Monte Carlo simulation. The b-jet JES uncertainty is obtained adding the calorimeter response uncertainty and the uncertainties from the systematic Monte Carlo variations in quadrature. The validation of the calorimeter p_T^{jet} measurement is done comparing it to the one from tracks associated to the jet for inclusive jets and identified b-jets. From the comparison of data to Monte Carlo simulation the b-jet energy scale uncertainty relative to the inclusive jet sample is estimated.

4.6.5 Jet energy resolution

The jet energy resolution have been measured with two different methods:

- **The di-jet balance method.** This method is based on momentum conservation in the transverse plane and is derived from the measurement of the asymmetry resolution. The asymmetry between the transverse momenta of the two leading jets A($p_{T,1}, p_{T,2}$) is defined as:

  $$A(p_{T,1}, p_{T,2}) = \frac{p_{T,1} - p_{T,2}}{p_{T,1} + p_{T,2}}. \tag{4.8}$$

  The fitted Gaussian σ_A is related to the relative jet resolution accordingly to:

  $$\sigma_A = \frac{\sqrt{(\sigma_{p_{T,1}})^2 + (\sigma_{p_{T,1}})^2}}{\langle p_{T,1} + p_{T,2} \rangle}. \tag{4.9}$$

  The assumption of transverse momentum balance implies $\langle p_{T,1} \rangle = \langle p_{T,2} \rangle \equiv p_T$, while requiring the jets to be in the same rapidity region signifies $\sigma_{p_{T,1}} = \sigma_{p_{T,2}} = \sigma_{p_T}$. Then the equation 4.9 can be reduced to $\sigma_A \simeq \frac{\sigma_{p_T}}{\sqrt{2} p_T}$ and the relative jet energy resolution is given by:

  $$\frac{\sigma_{p_T}}{p_T} = \sqrt{2}\sigma_A. \tag{4.10}$$

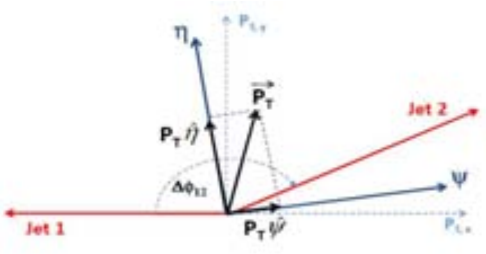


Figure 4.1: A scheme to define the variables of the bi-sector technique for JER. The η-axis corresponds to the azimuthal angular bi-sector of the di-jet system and the Ψ-axis is defined as orthogonal to the η-axis

- **The bi-sector technique.** This other method is based on the definition of an imbalance (transverse) vector, $\overrightarrow{P_T}$, as the vector sum of the two leading jets in the di-jet event. Such vector is projected along an orthogonal coordinate system in the transverse plane (Ψ, η), where η is chosen in the direction that bisects the angle formed by $\overrightarrow{P_{T,1}}$ and $\overrightarrow{P_{T,2}}$, as illustrated in Fig. (fig 4 nota JER). For a perfectly balanced di-jet event, $\overrightarrow{P_T} = 0$. Fluctuations due to different sources make the variance of each component of $\overrightarrow{P_T}$ to be non-zero. The basic assumption of the bi-sector method is that the fluctuations in the Ψ and η components at the particle level are equal: $\sigma_{\Psi}^{2part} = \sigma_{\eta}^{2part}$. At calorimeter level, $\overrightarrow{P_T^{calo}}$ will further differ from zero because of instrumentation effects, which measurement is the goal of this method. If both jets belong to the same rapidity region, they have the same average jet resolution and with this assumption it can be shown that:

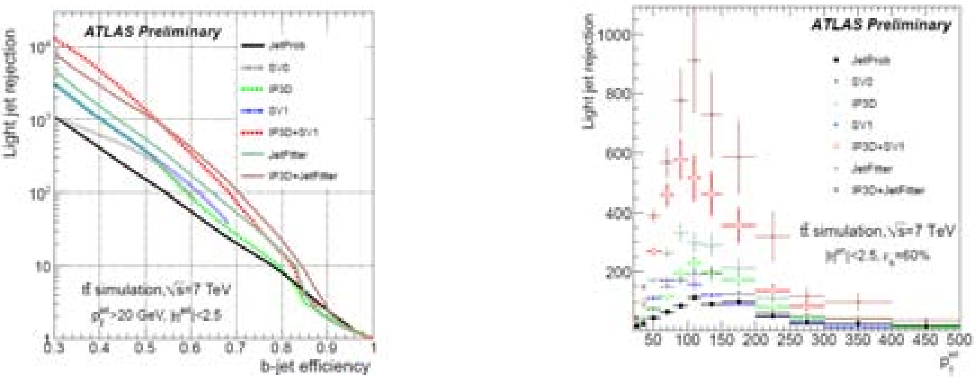
$$\frac{\sigma_{P_T}}{\langle P_T \rangle} = \frac{\sqrt{\sigma_{\Psi}^{2\,calo} - \sigma_{\eta}^{2\,calo}}}{\sqrt{2}\langle P_T \rangle |\cos \Delta\phi_{12}|}, \tag{4.11}$$

where $\Delta\phi_{12}$ is the angle between the two leading jet p_{T}.

The dispersions σ_{Ψ} and σ_{η} are extracted from a Gaussian fits to the $P_{T,\Psi}$ and $P_{T,\eta}$ distributions in bins of p_{T}. To be noticed that the resolution is expressed in terms of calorimeter observables only.

4.6.6 Jet reconstruction efficiency

The jet reconstruction efficiency is determined in the Monte Carlo simulation by counting in how many cases the calorimeter jet can be matched to a truth jet. Reconstructed jets are matched to truth jets, if their axes are within a $\Delta R = 0.4$. Since track jets and calorimeter jets are reconstructed by independent ATLAS sub-detectors, the ability of the Monte Carlo simulation to correctly reproduce the jet reconstruction in the data is tested using track jets. A tag-and-probe method is implemented to measure *in situ* the jet reconstruction efficiency for track jets. The jet reconstruction efficiency is hence measured in a sample of



(a) (b)

Figure 4.2: (a) Light-jet rejection as a function of the $b$jet tagging efficiency for the early tagging algorithms (JetProb and SV0) and for the high-performance algorithms, based on simulated $t\bar{t}$ events. (b) Light-jet rejection as a function of the jet transverse momentum p_{T}, for operating points of the various tagging algorithms leading to the same $\epsilon_b^{t\bar{t}}$=60%, based on simulated $t\bar{t}$ events.

minimum bias events and it is compared to a minimum bias Monte Carlo simulation: the matching efficiency in data and Monte Carlo shows a good overall agreement except at low p_T^{jet} ($<$ 25 GeV). The systematic uncertainty of the *in situ* determination is estimated varying certain selection requirements to both data and MC and it results larger than the shift between data and Monte Carlo simulation. A systematic uncertainty of 2% for jets with $p_{\mathrm{T}}<$ 30 GeV is assigned and negligible for higher p_{T}.

4.7 B-tagging

The ability to identify jets containing b-hadrons is important for the high-p_{T} physics of a general-purpose experiment at the LHC such as ATLAS. In particular, in $t\bar{t}$ processes, where at least two b-jets are expected, the performance of the b-tagging could affect the precision in the results. This is especially true for the method used in this analysis, as explained in Section 6.2: b-tagging is a delicate point because the distinction into different b-tagged jets bins allows to separate signal enriched regions from $W + jets$ background enriched regions and to know where to extract the signal cross-section. The key objects for b-tagging purposes are tracks, which reconstruction has been described in Section 4.1. Tracks are required to fulfill certain b-tagging quality requirements and to be associated to a reconstructed jet. The most critical track parameters for b-tagging are the transverse and longitudinal impact parameters. The transverse impact parameter d_0, is the distance of closest approach of the track to the primary vertex point in the r-ϕ projection. The z coordinate of the track at this point of closest approach is referred as z_0 or longitudinal impact parameter. The knowledge of the position of the primary interaction point of the proton-proton collision is important since it defines the reference point: some details about vertex reconstruction can be found in Section 4.2. To determine the vertex position with enough good resolution, the primary vertex must be reconstructed from at least five tracks. On the basis that the decay point of the b-hadron must lie along its flight path, the impact parameter is signed to further discriminate the tracks from b-hadron decays from tracks originated from the primary vertex. The b-tagging algorithms can be classified in the following way.

- **Impact parameter based algorithms.** These algorithms are based on the combination of the impact parameter significance (d_0/σ_{d_0}) of all the tracks in the jet. This can be done in a simple way, like JetProb tagging algorithm does, or with more sophisticated techniques as a likelihood ratio in which input variable have been previously compared with pre-defined and smoothed distributions for both b- and light-jet hypothesis (IP3D algorithm).
- **Secondary vertex based algorithms.** Including informations from the vertex formed by the decay products of the b-hadron helps to increase the discrimination between b-jets and light jets. All two-tracks pairs that form a good vertex, using only tracks associated to the jet and far enough from the primary vertex are built. Vertices compatible with V^0 or material interaction are rejected. The decay length significance $L_{3D}/\sigma_{L_{3D}}$ measured in 3D and signed with respect to the jet direction can be used as a discriminating variable between b-jets and light jets: this is the principle of SV0 tagger. To increase the discrimination power, the high-performance tagging algorithm SV1 takes advantage of three of the vertex properties: the invariant mass of all tracks associated to the vertex, the ratio of the sum of the energies of the tracks in the vertex to the sum of the energies of all tracks in the jet, and the number of two-tracks vertices. These variables are combined using a likelihood ratio technique.
- **JetFitter algorithm.** This algorithm is able to exploit the topology of weak b- and c-hadrons inside the jet. A Kalmar filter is used to find a common line on which the primary vertex and the b- and c-vertices lie, as well as their position on this line, giving an approximated path of the b-hadron. With this approach b- and c-hadrons vertices are not necessarily merged. The discrimination between b-, c- and light jets is based on a likelihood using similar variables as in SV1 algorithm and additional variables as flight length significance of the vertices.

In this analysis, the combination of JetFitter and IP3D jet algorithm (JetFitter-CombNN tagger) has been employed at its operational point that provide 70% of efficiency. A combination of IP3D and SV1 algorithms is also possible.

The b-tag efficiency is defined as the fraction of reconstructed jets originated from b-quarks that are tagged by the b-tagging algorithm. To label the true flavor of the jets, is used the following procedure: a jet is labeled as a b-quark if a b-quark with $p_{\mathrm{T}}> 5$ GeV is found in a cone of size $\Delta R = 0.3$ around the jet direction. The various labeling hypothesis are tried in this order: b quark, c quark and τ lepton. When none of these hypothesis is satisfied, the jet is labeled as a light jet. The light rejection (or mistag rate) is the reciprocal of the fraction of jets that are labeled as light jets and are actually tagged incorrectly by the algorithm. In Figure **??** (a) shows the light-jet rejection as a function of b-tag efficiency for the various ATLAS b-tagging algorithms. It is obtained varying continuously the operating point of each tagger, *i.e.* the cut on its output discriminating variable. The jets are from $t\bar{t}$ simulated events and satisfy the following cuts: $p_{\mathrm{T}}> 20$

GeV and $\eta < 2.5$. For the same tagging efficiency high-performance tagging algorithms as that used in this analysis, achieve much lower mis-tagging rate even at relatively high efficiencies with respect to the algorithm previously used: at 70% b-tagging efficiency the IP3D+JetFitter algorithm achieves a mis-tag rate lower than 1%. In Figure **??** (b) is shown that the light rejection depends strongly of the kinematics and the variation is specific of each tagging algorithm. It is plotted the light rejection in function of the jet p_{T} for the various algorithms operating in such a way that they all lead to the same efficiency $\epsilon_b^{tt} = 60\%$. The optimal performance for each algorithm is reached at $p_{\mathrm{T}}\sim$ 100 GeV: below this threshold, tracks in jets are relatively soft and therefore multiple scattering is compromising the resolution on the impact parameter, while above the threshold several other effects reduce the performance. The tagging rate is defined as the fraction of jets that are tagged out of those that could be tagged by a given algorithm, for a specific choice of its operating point. In Fig. 4.3 the tagging rates of IP3D+JetFitter for the inclusive jet sample and for the one enriched in heavy-flavor jets are shown, with algorithm configuration such that $\epsilon_b^{tt} = 70\%$. The tagging rate predicted by the simulation agrees with experimental data to within 20%.

4.8 Taus

4.8.1 Reconstruction

Tau leptons decay in the 65% of the cases in hadrons.

The hadronic tau reconstruction starts considering the set of calorimeter jets reconstructed with the anti-k_T algorithm, from topological clusters of calorimeter cells, using a distance parameter R = 0.4. Each of these jets, calibrated with the Local Hadron Calibration (LC) is referred as a *seed* for the tau reconstruction algorithm, if the jet p_{T} is more than 10 GeV and its pseudorapidity $fabs\eta < 2.5$ (the coverage of the ATLAS tracking system). Then, the four momentum of the all tau visible decay products (not the neutrino), is reconstructed, defined in term of p_{T}, η and ϕ. The psedorapidity η and the azimuthal angle ϕ are taken from the sum of the four-vectors of the topological clusters associated to the seed, assuming zero mass for each cluster. Tau candidates are represented by massless 4-vectors, so the transverse momentum and the transverse energy are the same. Track are then associated to each tau candidate if their distance from the tau candidate axis is $\Delta R < 0.2$ (*core cone*) and if they pass the quality criteria listed here below:

- $p_{\mathrm{T}} >$ 1 GeV;
- number of pixel hits ≥ 2;
- number of pixel hits + number of SCT hits ≥ 7;
- $fabsd_0 < 1.0$ mm;
- $fabsz_0 sin\theta < 1.5$ mm;

where d_0 is the distance of the closest approach of the track to the reconstructed primary vertex in the transverse plane and z_0 is the longitudinal distance of the closest approach, as said in the previous section. Tau candidates are classificated as single or multi-prong

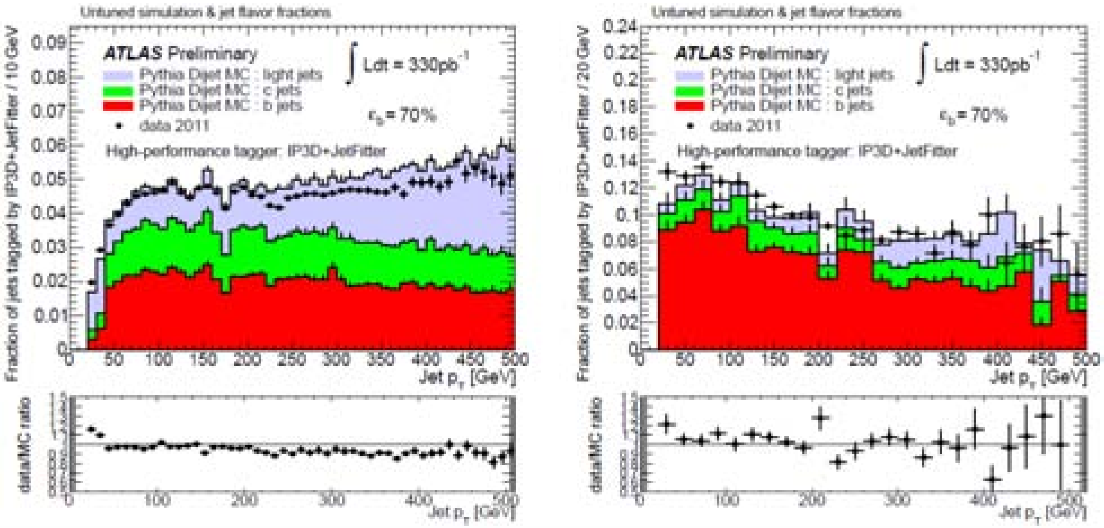


Figure 4.3: Distribution of the tagging rate for the IP3D+JetFitter tagger at the operational point $\epsilon_b \approx 70\%$ for data (solid black points) and for MC (filled histograms for the various flavours) versus the jet p_{T} for an inclusive light jet sample (left) and an enriched heavy-flavour jet samole (right). The ratio data/MC is shown at the bottom of each plot.

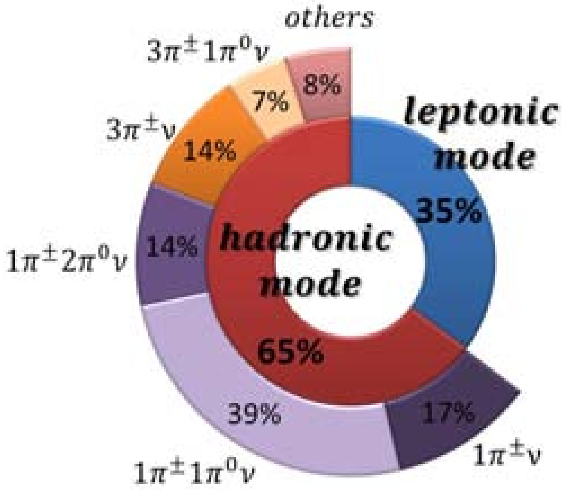
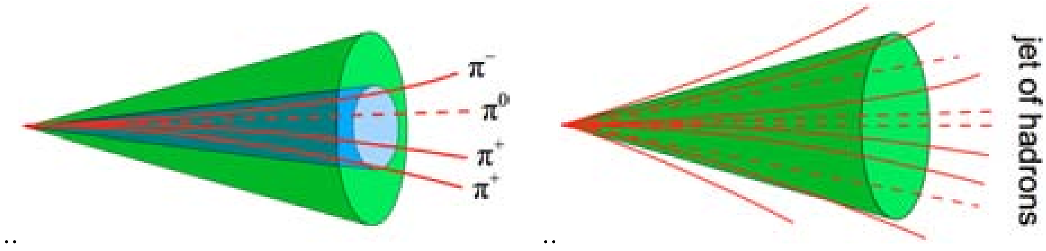


Figure 4.4: Graphical representation of the different decay modes of the tau lepton and the relative branching ratios.



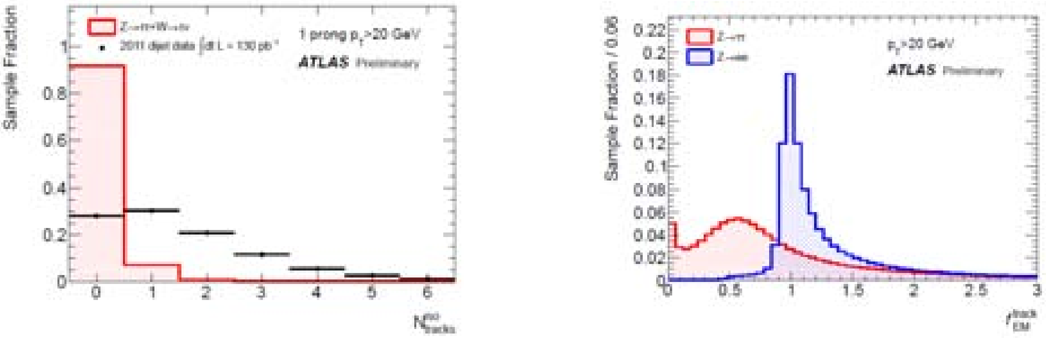
(a) (b)

Figure 4.5: (a) Schematich view of a hadronic tau decay. (b) Schematich view of a jet.

depending on the number of tracks counted in the *core cone*. Tracks within the *isolation annulus*, defined as $0.2 < \Delta R < 0.4$ from the tau candidate axis, are also used in the calculation of discriminating variables and are required to satisfy the same track quality criteria.

4.8.2 Identification

The step that follow the reconstruction is the identification, as described in this section. A list of discriminating identification variables based on calorimeter and tracking information calculated in the reconstruction step, are combined into multivariate discriminants to reject QCD jets and electrons. The complete list can be found in [35], some of them introduced for the new optimization to take into account the pile-up condition of 2011 data. For example a new pile-up insensitive variable for jet discrimination is the number of track in the isolation annulus (N^{iso}_{track}), represented in Fig. for signal and jet background. On the other hand, an example of variable used to distinguish taus from electron background, is the electromagnetic track fraction, f^{track}_{EM}, i.e. the ratio of the transverse energy deposited



(a) (b)

Figure 4.6: (a) Example of a jet discriminating variable: number of track in the isolation annulus for MC simulated $Z \rightarrow \tau\tau$ and $W \rightarrow \tau\nu$ signal sample and a di-jet background sample selected from 2011 data. (b) Example of an electron discriminating variable: electromagnetic track fraction for MC simulated $Z \rightarrow \tau\tau$ signal and $Z \rightarrow ee$ background events. The distributions are normalized to unity [35].

in the electromagnetic calorimeter over the transverse momentum of the leading track:

$$f_{EM}^{track} = \frac{\Sigma_{i \in EM}^{\Delta R_i < 0.4} E_{T,i}}{p_{T,1}^{track}}, \tag{4.12}$$

where i runs over all cells in the EM calorimeter within the *wide cone* ($\Delta R < 0.4$).

Jet discrimination

There are three methods used in the tau identification to accept true hadronically decaying tau leptons and reject reconstructed candidate from QCD jets: cut-based, likelihood function and boosted decision tree discrimination. All of them use a discriminant built with various identification variables and for each discriminant three working points are defined and optimized: loose, medium and tight, which yield signal efficiency of approximately 60%, 45% and 30%, respectively. The optimization of the discriminants have been done using PYTHIA $W \rightarrow \tau\nu$, $Z \rightarrow \tau\tau$ and $Z' \rightarrow \tau\tau$ Monte Carlo samples for signal and a QCD jet background sample taken from a selection of di-jet events in ATLAS data, collected in 2011, corresponding to a integrated luminosity of 130 pb^{-1} (for optimization details see [35]). For the scope of this thesis, have been used the likelihood discrimination method. The likelihood function, $L_{S(B)}$, for signal (background) is defined as the product of the one-dimensional probability density function, $p_i^{S(B)}$, of each identification variable considered, x_i:

$$L_{S(B)} = \prod_{1=1}^{N} p_i^{S(B)}(x_i). \tag{4.13}$$

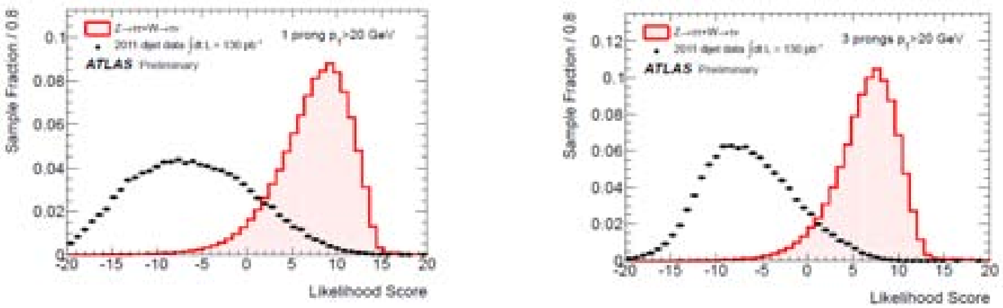


Figure 4.7: The log-likelihood-ratio for 1-prong (left) and 3-prong (right) tau candidates [35].

The discriminant used by the likelihood based tau identification is defined as the log-likelihood-ratio between signal and background:

$$d = ln\left(\frac{L_S}{L_B}\right) = \sum_{i=1}^{N} ln\left(\frac{p_i^S(x_i)}{p_i^B(x_i)}\right). \tag{4.14}$$

Each probability density function is calculated as the fraction of events per bin in a histogram of the x_i distribution. To maximize the discrimination power of the method, both the signal and background samples are split into categories (p_{T} bins, number of prongs and number of reconstructed vertices to reduces the dependence of pile-up conditions) and the likelihood is trained separately for each category. Also tau polarization is taken into account, but is found negligible. The log-likelihood ratio (Likelihood Score) for 1-prong and 3-prong separately is shown in Fig.

Loose, medium and tight selections on the log-likelihood score have been defined which yield on average 60%, 45% and 30% signal efficiency. The efficiency are approximately independent in p_{T}as the cuts on the likelihood are calculated in bins of p_{T}.

The boosted decision tree jet discriminant output (BDTJetScore), is also used in the present analysis. The Boosted Decision Tree approach used for tau identification is a multivariate algorithm with a continuous discriminant output. The difference of using this kind of approach instead of simply applying a given cut on a given variable is that an event that fails a single cut continue to be considered in the decision tree. This method is used to better identify taus and separate signal from background as it reduces the misidentification. In particular, we took into account the continuous discriminant outputs and cut directly on them: in general, closer is the output score to 1, more probably the candidate object is tau-like. Hence on one side, the `BDTJetScore` is useful to discriminate taus from jets; in this analysis a cut on the BDTJetScore is applied to improve the rejection of multi-jet QCD background (see Section 6.1.5). On the other side, the electron discriminant output `BDTEleScore` can be used as electron veto as explain in next section, because it separates electrons from taus. In Fig. 4.8 is shown the jet BDT score distribution compared between $W \to \tau\nu$, $W \to \tau e$ and $W \to \tau\mu$ Monte Carlo simulated events.

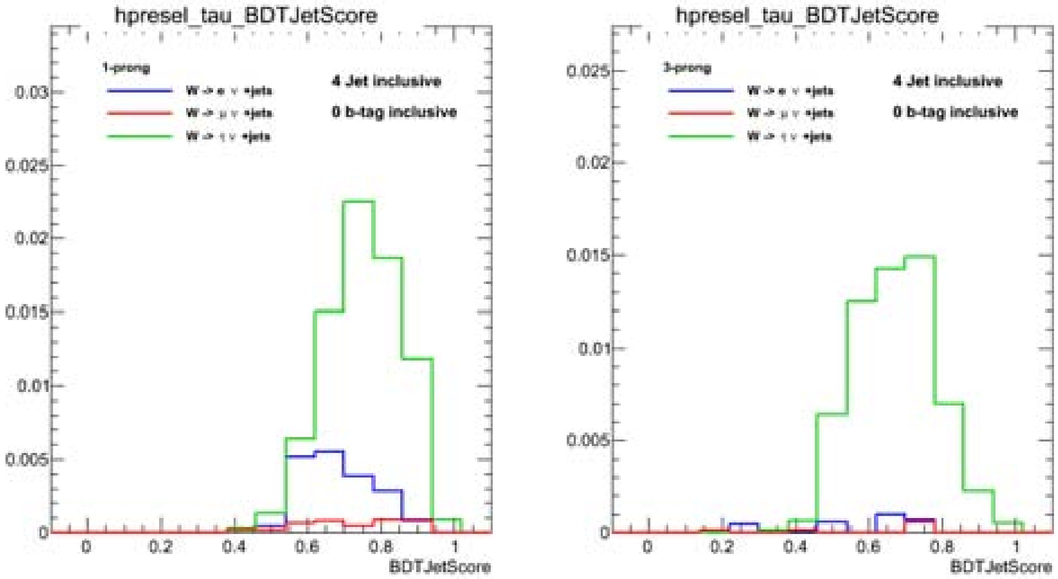


Figure 4.8: The electron BDT score distribution (BDTJetScore). Comparison between $W \rightarrow \tau\nu$, $W \rightarrow e\nu$ and $W \rightarrow \mu\nu$ Monte Carlo simulated events. The plots are obtained with the baseline selection of this analysis (see Chapter 6), except the cut on the BDTJetScore. At least four jets are required, without any b-tagging requirement. At left (right), 1-prong (3-prong) tau events only are considered.

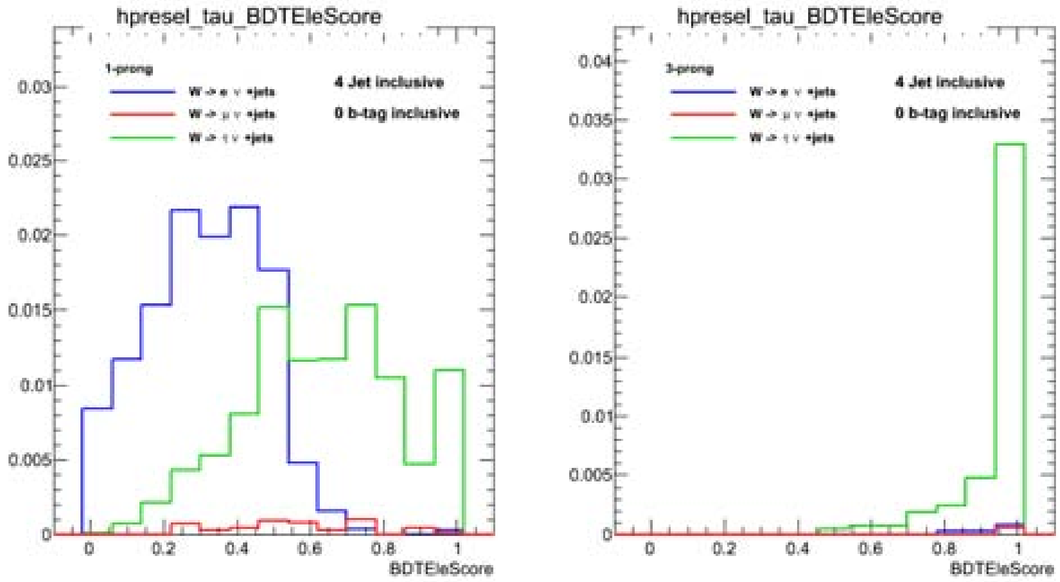


Figure 4.9: The electron BDT score distribution (BDTEleScore). Comparison between $W \to \tau\nu$, $W \to e\nu$ and $W \to \mu\nu$ Monte Carlo simulated events. The plots are obtained with the baseline selection of this analysis (see Chapter 6), except the cut on the BDTEleScore. At least four jets are required, without any b-tagging requirement. At left (right), 1-prong (3-prong) tau events only are considered.

Electron discrimination

The electron signature in the detector can be easily confused with the similar signature of taus decaying hadronically in 1-prong: in fact after the kinematic cuts that reject the jet related backgrounds a significant part of the remaining background is due to electrons. The properties that better help to distinguish between the two signatures are the difference in the shower produced by a tau lepton, wider and longer than electron shower and the emission of transition radiation of the electron track. Discriminants to reject electron misidentified as tau are built using these and other properties through the variables defined in the reconstruction. The most common discrimination methods are the cut-based and the boosted decision tree based discriminants. In this analysis is used the electron BDT discriminant as electron veto, applying a cut on the discriminating variable itself, named BDTEleScore, without the use of defined working points. The score of BDT electron discriminant distribution is shown in Fig. 4.9 for different leptonic decay channels of the W: in particular there is a clear separation between electrons from $W \to e\nu$ and hadronically decaying tau from $W \to \tau\nu$ decays.

The performance of the identification algorithms is evaluated from data using the tag-and-probe method. The tag-and-probe crucial point consists in the ability to select a high purity sample of signal candidates (probes) from data without applying identification to probe itself. This is possible by tagging events from a process that have a very distinct signature, which also contains a real signal candidate. In general, strict selection is applied to the tag, while selection on the probe is as looser as possible, to minimize bias on the identification variable of the probe. After the signal sample is selected, tau identification methods are applied to the probe candidates. The identification efficiency is extracted by estimating the number of real signal candidates in the sample before and after the identification. If the background contribution is non-negligible has to be subtracted from the sample. The same selection is then applied to MC simulated signal samples and the ratio of the efficiency calculated in data (ϵ^{ID}_{DATA}) to the MC efficiency (ϵ^{ID}_{MC}) are calculated as correction factors:

$$C^{ID}_{MC} = \frac{\epsilon^{ID}_{DATA}}{\epsilon^{ID}_{MC}}. \tag{4.15}$$

The tau identification efficiency have been measured with tag-and probe methods with both $Z \rightarrow \tau\tau$ events in 800 pb^{-1}and with $W \rightarrow \tau\nu$ events selected from 1.37 fb^{-1}of data collected with ATLAS in 2011. The measurement of the efficiency with $Z \rightarrow \tau\tau$ have been done together with $Z \rightarrow \tau\tau \rightarrow l\tau_{had}$ analysis [35]. The muon channel only have been considered and the muon have been used as the tag. For the measurement with $W \rightarrow \tau\nu$ events instead, the events have been tagged with a large amount of $E_{\mathrm{T}}^{\mathrm{miss}}$and the number of real hadronically decaying taus is estimated by fitting the number of tracks distribution. Three templates have been used in the fit: one for real taus, one for misidentified electrons and one for misidentified QCD jets. The templates for taus and electrons are taken from MC, while the jet template is extracted from data in a control region.

Both measurements yield to scale factors for MC very close to unity for all the identification discriminant and working point: so the prescription is to not apply any correction to the MC sample. The uncertainty varies depending on the method and the discriminant: in the case of likelihood discriminant and at the tight working point and for tau $p_{\mathrm{T}} > 30 GeV$, the prescription is to use 9.3% of uncertainty in analysis that take into account both 1-prong and 3-prong contributions and 7.8% for 1-prong only analysis [51].

4.8.3 Tau energy scale

Because 1-prong and 3-prong hadronic tau decays each consist of a specific mix of charged and neutral pions, the energy scale needs an additional calibration for taus with respect to jet calibration. The calorimeter clusters associated to the jet seed are first calibrated using the local hadron calibration (LC), but an additional correction is applied to restore the tau energy to the true value. The initial direction is taken from the four-vector sum of the clusters associated to the seed. All clusters within a cone of $\Delta R < 0.2$ around the initial direction contribute to the calculation of tau energy. The smaller cone radius reduces the pile-up effects and provides a good energy resolution. At this point, an additional correction factor, determined from $Z \rightarrow \tau\tau$ and $Z' \rightarrow \tau\tau$ simulated samples, is applied to the four-momentum of the tau candidate at LC scale. Reconstructed tau candidates are required to match to a true hadronically decaying tau within $\Delta R < 0.2$ and to pass the loose cut identification; no other kinematical requirements are applied. The tau response is defined as the ratio of reconstructed tau energy at the LC scale divided by the true tau visible energy. Then the response is binned in true visible energy, reconstructed $|\eta|$ and for 1-prong and multi-prong candidates. In each bin, the response distribution is fitted with a Gaussian and the response is taken as the mean of the Gaussian fit. A response curve is built in function of LC energy scale, in bins of η and N_{track}. The uncertainty on the energy scale is also evaluated using MC simulation of $Z \to \tau\tau$ and $Z' \to \tau\tau$ processes. The method consists in comparing the nominal results of the reconstruction of p_{T} of tau candidates passing loose identifications, with Monte Carlo simulations with alternative configurations. A similar method was used to measure the ATLAS jet energy scale uncertainty. The systematic uncertainty is defined as the difference from the nominal configuration of the quantity f_s, the relative scale difference between reconstructed and true visible momentum:

$$f_s = \frac{p_T^{reco} - p_T^{truth}}{p_T^{truth}}, \tag{4.16}$$

where:

- p_T^{reco} is the reconstructed transverse momentum after the tau energy scale correction and required to match to a true hadronically decaying tau within $\Delta R < 0.2$;
- p_T^{truth} is the true visible transverse momentum.

The uncertainties are calculated in the barrel, end-cap and in transition regions, for 1-prong and multi-prong candidates.

In each region are exploited six different p_{T} bins, ranging from 15 to 110 GeV. The considered sources of uncertainties are: Monte Carlo generator and underlying event model, hadronic shower model, amount of detector material, EM energy scale, topological clustering noise thresholds, pile-up and non-closure. The total uncertainty on the tau energy scale in each bin is taken by adding the individual uncertainties in quadrature.

Chapter 5

The trigger

The selection of hadronically decaying tau events is challenging due to the high background rate at the LHC. Achieving an efficient selection of tau events implies the increase of the discovery potential in new physics channels, concerning Higgs boson and Supersymmetry. In this chapter after an introduction about the ATLAS trigger system, the tau trigger selection technique and its performance with 2011 data is briefly presented.

5.1 ATLAS trigger system

The aim of the trigger and data acquisition (TDAQ) systems is to be able to select events of interest in order to reduce the initial rate of ~1 GHz of events in proton-proton interaction (40 MHz bunch-crossing) to a final rate of about 200-600 Hz to be used in physics analysis. ATLAS demonstrated to be able to achieve this goal thanks to its three level system that reduces the rate in three consecutive steps:

- Level 1 trigger (hardware based): reduces the rate from 40 MHz to 75 KHz;
- Level 2 trigger (software based): reduces the rate from 75 KHz to 3-5 KHz;
- Event Filter (software based): reduces the rate from 3-5 KHz to 200-600 Hz.

The first level trigger (L1) is an hardware system based on muon trigger detectors and calorimeter information with a granularity of about $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$. It aims at identifying high transverse energy objects: electromagnetic, like electrons and photons, or hadronic, as jet and hadronic tau, as well as events with large missing energy or missing transverse energy. At L1 a Region of Interest (RoI) is defined for each energy deposition cluster in the calorimeter: the RoI is passed to the next level of trigger where it is used as a *seed* for the reconstruction algorithm. The *"L1 accept"* decision is taken within 2.5 μs, based on the multiplicity of the muon and calorimeter objects combined in the Central Trigger Processor (CTP). The decision of the CTP depends on the combination of the different objects received matching any of the required L1 *trigger items*, stored in a L1 *trigger menu*. The items in the menu can be programmed and associated to *prescale* factors, to reduce the fraction of the accepted events to be recorded.

The second level trigger (L2) and the Event Filter (EF) constitute the High Level Trigger (HLT) and consist of software algorithms. The HLT analyzes data from all ATLAS sub-detectors, but only in the RoI provided by the L1: this fact and the use of fast reconstruction algorithms speed up substantially the execution process. The Event Filter (EF) uses algorithms very similar to those employed in the offline object reconstruction to select events and reduce the rate. The EF has at its disposal more sophisticated calibration and alignment tools with respect to L2 and can access to the entire event, such that the calculation of overall event variable like $E_{\mathrm{T}}^{\mathrm{miss}}$ is available. In addition, at HLT topological requirements are considered, such as mass cuts or b-tagging. The decision time of HLT is 40 ms (L2) + 4 s (EF). The *HLT trigger menu* is composed of a list of physics signatures, each of which is the result of a *chain* of processing through the L2 and EF originating from an accepted L1 item. Each chain (or *Slice*) has the goal of identifying a different particle or signature. The full trigger menu contains electron, photons, muon, tau, jet, b-jet, missing energy, total energy, jet energy and B-physics triggers. The chain execution proceeds up to the step where a signature is not satisfied; at that point the chain is deactivated for the current event. An event is registered if it is selected by at least one of these chains.

To deal with the high rate and to optimize the efficiency of important physics channels, different triggers are often used in combination with various setting of thresholds. The analysis presented here, for example, relies on events passing a combined tau and transverse missing energy trigger: `tau29_medium_xe35_noMu`. This trigger has a p_{T} threshold at 29 GeV for the tau object and a $E_{\mathrm{T}}^{\mathrm{miss}}$ threshold at 35 GeV. The same trigger is used also in $H^{\pm} \rightarrow \tau^{\pm}\nu$ analysis, having similar final state.

5.2 Tau trigger

The tau trigger is designed to select hadronic tau decays, which signature is characterized by one or three tracks due to charged hadrons (1-prong or 3-prong). In addition, a neutrino is produced and often also one or more neutral pions. The trigger steps for the tau are detailed here below.

Level 1. At L1, electromagnetic (EM) and hadronic (HAD) calorimeter trigger towers are used to calculate the energy inside a core region and an isolated region around the core. The core region is defined as the tower of dimensions $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ while the isolation region is defined as the ring of towers within $\Delta\eta \times \Delta\phi = 0.4 \times 0.4$ surronding the core region. The energy threshold and the identification criteria were adjusted as the luminosity increased during 2011: each trigger item was tuned to maximize the physics acceptance while keeping the rate under control.

Level 2. L2 algorithms use the L1 information as a seed and refine the L1 RoI position. Both tracking and calorimeter information are taken into account to discriminate taus from multi-jet background. The discrimination is based on shape variables calculated in all layers of the calorimeters that highlight the characteristics of the tau jet compared to the QCD jets, as the narrowness and the track multiplicity. One of the key variable used in the L2 selection is the EM Radius (R_{EM}), defined in the same way as in the offline reconstruction:

$$R_{EM} = \frac{\sum_{Cells} E_{Cells} \cdot R_{Cells}}{\sum_{Cells} E_{Cells}}, \tag{5.1}$$

where E_{Cells} is the energy calculated in a cell and R_{Cells} the distance of each cell to the L2 candidate position. Note that the sum is extended to all cells in the first three samplings of the calorimeter, while in 2010 only the second layer of EM calorimeter was taken into account and the R_{Cells} was squared in the R_{EM} calculation. The R_{EM} provides an estimation of the lateral size of the cluster in the EM calorimeter. Other useful variables are the number of tracks and the ratio of the scalar sum of p_{T} of all tracks in an isolation region to the scalar sum of the p_{T} of all tracks in a core region. The isolation region at L2 is defined as the annulus between the cones of radius 0.1 and 0.3 around the direction of the highest p_{T} track, while the core region is defined as the cone of radius 0.1 around the same direction point. Figure 5.2 shows the comparison of the R_{EM} variable used for the L2 selection between $Z \rightarrow \tau\tau$ signal events and di-jet background, as simulated in Monte Carlo.

Event Filter The EF, exploits the same characteristics as L2 but with more refined algorithms, taking advantage of the possibility of a larger processing time. Actually, the EF selection algorithms are very similar to those used in the offline reconstruction and a detector level calibration is also applied. The selection is based on the tau E_T, the number of associated tracks and shape variables. Different selection cuts are applied to 1-prong or multi-prong taus. The shape variables used are:

- R_{EM}, defined as in Eq. 5.1;
- R_{Track}, that accounts for the lateral size of the track system and is defined as

$$R_{Track} = \frac{\sum p_{T,track} \cdot \Delta R_{track}}{\sum p_{T,track}}; \tag{5.2}$$

  where ΔR_{track} is the distance between the track and the EF tau candidate direction;
- f_{track}, which is the fraction of the total transverse energy of the tau candidate carried by the highest p_{T} track.

In Figure 5.2, the distributions of R_{EM} for 1-prong and 3-prong, comparing $Z \rightarrow \tau\tau$ events and di-jet background are shown: there is good discrimination power and the good correlation with the corresponding offline variables. A less precise energy calibration applied at the EF with respect to the offline causes the small shift in the respective distributions.

5.3 $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger

The offline reconstruction of $E_{\mathrm{T}}^{\mathrm{miss}}$ (see Chapter 4) uses an algorithm which combines the energy information provided by calorimeter cells and by the muon spectrometer. This algorithm implies very long time to access to the stored information, reject the noisy cells and calculate the $E_{\mathrm{T}}^{\mathrm{miss}}$. The algorithm used in HLT trigger balances the algorithm performance and speed. The goal of $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger is to trace interesting events and at the same time distinguishing fake sources of $E_{\mathrm{T}}^{\mathrm{miss}}$. The trigger thresholds are driven by background, mainly di-jet events which ideally would have no missing transverse energy, but

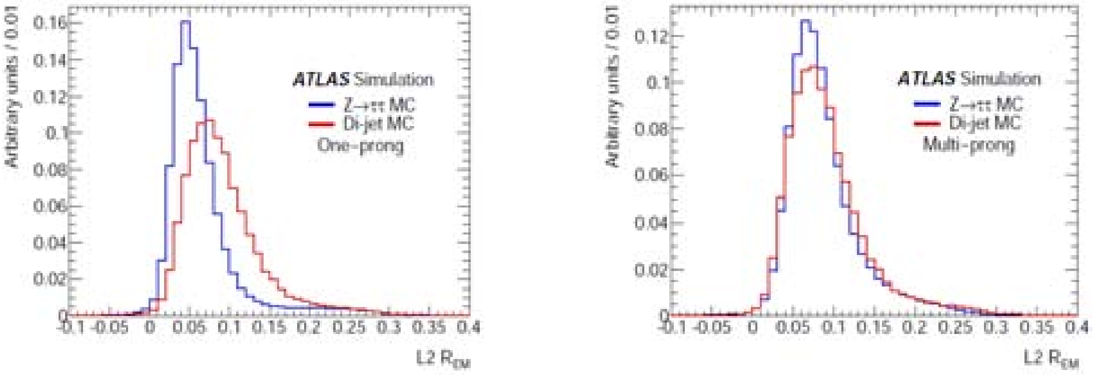


Figure 5.1: Comparison of L2 EM Radius distribution for 1-prong (left) and 3-prong (right) between $Z \to \tau\tau$ signal and di-jet background from MC simulation.

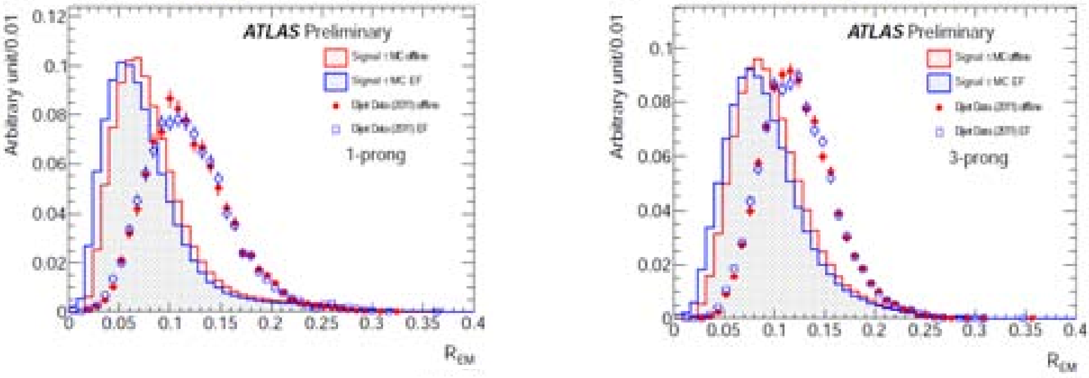


Figure 5.2: Comparison of EF EM Radius distribution for 1-prong (left) and 3-prong (right) between $Z \to \tau\tau$ MC simulated signal and di-jet 2011 data. The same quantity calculated offline in the same samples is shown too.

due to the calorimeter resolution an asymmetry can be measured. Other algorithms are used to found fake $E_{\mathrm{T}}^{\mathrm{miss}}$ sources such as calorimeter cracks and dead cells, or beam-halo events. In the 2011 data analysis, minimum bias events dominated by pile-up constitute the main challenge of $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers. Due to pile-up collisions large transverse energy are deposited in calorimeter cells, increasing the probability of measurement fluctuations giving large $E_{\mathrm{T}}^{\mathrm{miss}}$. Sophisticated triggers have been developed that either use better methods of calculating $E_{\mathrm{T}}^{\mathrm{miss}}$ or use additional information of the quantity to be triggered on. For example, triggers based on $E_{\mathrm{T}}^{\mathrm{miss}}$ significance (labeled with `_xs_` suffix) parameterize the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution as a function of the sum of the energy deposited in calorimeters ($\sum E_T$) and set a threshold for the ratio of $E_{\mathrm{T}}^{\mathrm{miss}}$ to $E_{\mathrm{T}}^{\mathrm{miss}}$ significanceare used [47].

5.4 Trigger efficiency measurements

The data collected in 2011 provided a sufficiently clean sample of hadronic taus that has been used to estimate the performance of the single tau triggers and measure their efficiencies. The trigger efficiency is defined as the fraction of tau trigger objects that pass the trigger decision in a trigger chain with respect to the off-line tight BDT tau candidate [49]. The off-line candidate is required to be matched within $\Delta R < 0.2$ to an EF tau candidate that passed the trigger chain. As all the single tau triggers are prescaled in data, the decision is recalculated off-line using the L1 and HLT selections on the trigger object.

The trigger used in this analysis for the selection of events with a tau candidate is a combined $\tau + E_{\mathrm{T}}^{\mathrm{miss}}$ trigger, `EF_tau29_medium_xe35_noMu`: it requires both a τ jet identified with medium quality criteria with a transverse momentum of at least 29 GeV and at least 35 GeV of missing transverse energy calculated without applying a correction for muon objects. The trigger efficiency for the inclusive tau with respect to the off-line tau candidates passing all identification requirements as a function of p_{T} is measured in data using a sample of $Z \to \tau\tau$ events obtained by triggering on a single-lepton trigger, and the Monte Carlo is directly weighted by the measured efficiency ([48]). In Figure 5.3 the efficiency for single tau trigger `EF_tau29_medium` as a function of tau p_{T} is shown, measured with 3.6 fb^{-1} of 2011 data at $\sqrt{s} = 7$ TeV. The lower canvas shows the ratio between the efficiency calculated in data and the one calculated in the Monte Carlo, which represents the scale factor in function of tau p_{T}.

In the context of search for a charged Higgs boson in the $\tau + jets$ final state in $t\bar{t}$ decays [50], the efficiency of the `EF_tau29_medium_xe35_noMu` trigger has been estimate using a data-driven method based on the selection of $t\bar{t}$ decays in the $\mu + \tau$ channel, treated with the tag-and-probe technique. On one hand, such events are expected to be very similar to $t\bar{t}$ events with τ+jets final state, apart from the higher quantity of $E_{\mathrm{T}}^{\mathrm{miss}}$ due to the presence of an additional neutrino associated to the muon. On the other hand, the advantage is the possibility of use the muon trigger to tag the $\mu+\tau$ final state. The trigger efficiency can be deduced from the distributions in Figure 5.4, which illustrates the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution before and after the additional application of the `EF_tau29_medium_xe35_noMu` trigger. From the ratio of these efficiencies between data and Monte Carlo, scale factors for the $\tau + E_{\mathrm{T}}^{\mathrm{miss}}$ trigger are obtained. Due to the low statistic available the data are binned into four $E_{\mathrm{T}}^{\mathrm{miss}}$ and tau p_{T} bins. For the p_{T} binning, are defined the following

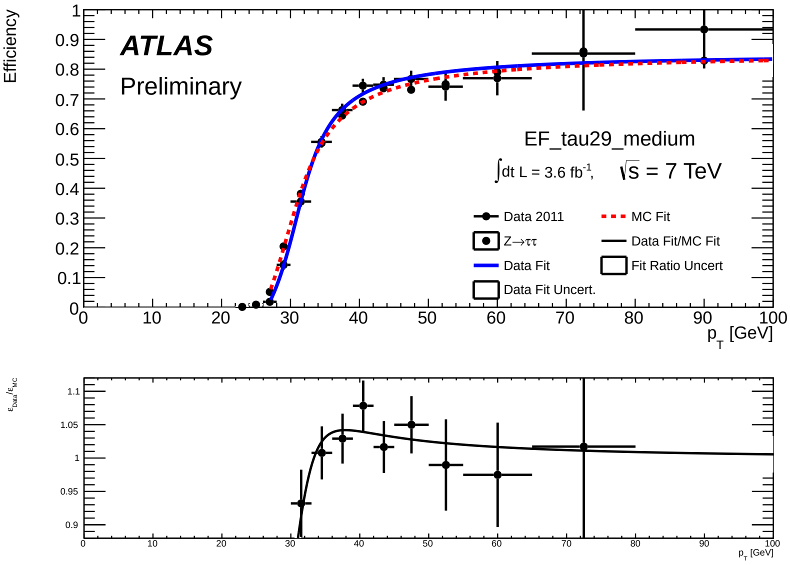


Figure 5.3: (Up) Efficiency of single tau trigger `EF_tau29_medium` measured with 2011 $Z \to \tau\tau$ data events. (Down) Ratio of the efficiency measured in data to the efficiency measured in MC.

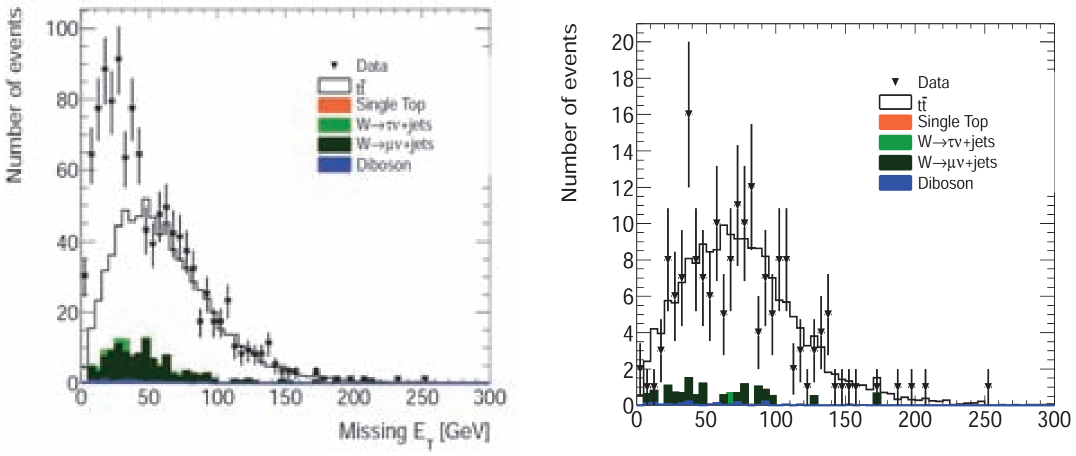


Figure 5.4: $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution for events in $\mu + \tau$ control sample ([50]), before (left) and after (right) additionally applying the `EF_tau29_medium_xe35_noMu` trigger. B-K data period of 7 TeV data are used, for a total integrated luminosity of 2.05 fb^{-1}.

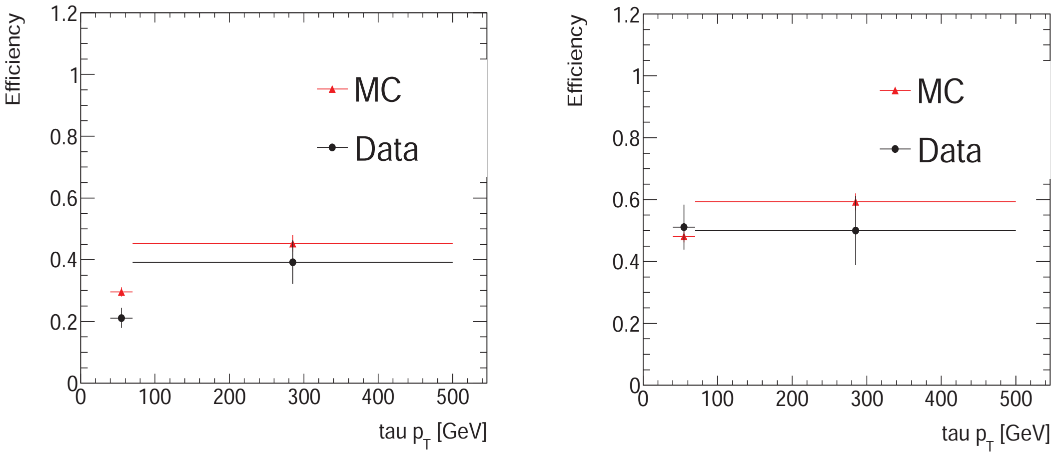


Figure 5.5: Efficiencies of the `EF_tau29_medium_xe35_noMu` trigger in data and Monte Carlo in the $\mu+\tau$ control sample studied in [50], in 65 GeV $\leq E_{\mathrm{T}}^{\mathrm{miss}} <$ 100 GeV (left) and 100 GeV $\leq E_{\mathrm{T}}^{\mathrm{miss}} <$ 500 GeV regions. B-K data period of 7 TeV data are used.

regions: 40 GeV $\leq p_{\mathrm{T}} <$ 70 GeV and 70 GeV $\leq p_{\mathrm{T}} <$ 500. The $E_{\mathrm{T}}^{\mathrm{miss}}$ binning is defined as two regions of 65 GeV $\leq E_{\mathrm{T}}^{\mathrm{miss}} <$ 100 GeV and 100 GeV $\leq E_{\mathrm{T}}^{\mathrm{miss}} <$ 500 GeV. The scales factors corresponding to these bins are reported in table 5.1.

	p_{T}= [40,70) GeV	p_{T}= [70,500) GeV
$E_{\mathrm{T}}^{\mathrm{miss}}$= [65,100) GeV	0.76 ± 0.09	0.86 ± 0.13
$E_{\mathrm{T}}^{\mathrm{miss}}$= [100,500) GeV	1.10 ± 0.12	0.91 ± 0.14

Table 5.1: Trigger scale factors for `EF_tau29_medium_xe35_noMu` as calculated in [50].

In Figure 5.5 the efficiencies in the $E_{\mathrm{T}}^{\mathrm{miss}}$ bins are shown. More details about the scale factor used in this analysis and its error, are provided in Chapter 6. In Appendix B.1 some checks of the $\tau+E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency in the context of this analysis are reported.

Chapter 6

Measurement of the $t\bar{t}$ cross-section in the $\tau + jets$ final state

The aim of this study is the measurement of the $t\bar{t}$ cross-section in the channel with $\tau + jets$ in the final state. This channel is the less studied so far, because of the challenge that the tau lepton reconstruction and identification represents. To cope with the high rate of fake taus and be able to measure the cross-section with a good accuracy, much more data are needed than in the analogous semi-leptonic analysis with other leptons (one electron or one muon) in the final state.

More precisely, the signal is characterized by the decay of one of the top quarks into a W boson and a b quark, with the subsequent W boson decay into a tau that in turns decays hadronically: $t \to b(W \to \tau\nu_\tau) \to b(\tau \to had\bar{\nu_\tau})\nu_\tau$. The W from the other top decays hadronically: $t \to bW(q\bar{q} \to jets)$. The final state is then characterized by the presence of:

- at least four jets, of which at least two are from the b quarks, decay products of the top quarks;
- one τ-jet from τ lepton in the hadronic decay mode;
- missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) due to two tauonic neutrinos (one neutrino and one anti-neutrino) which are produced in the decays of the W and the τ lepton.

The reconstruction of this challenging final state depends strongly on the reconstruction and identification of the τ object and the correct understanding of the $E_{\mathrm{T}}^{\mathrm{miss}}$; on the other hand also the jet reconstruction and energy scale, as well as the b-tagging are crucial for this measurement. The analysis strategy has been designed in order to take advantage of the maximum amount of signal events. A likelihood fit method is used to charactwrize the selected data sample in terms of signal and backgrounds and to extract the $t\bar{t}$ cross-section. The strategy is inspired the $t\bar{t}$ cross-section measurement in the e and μ semi-leptonic channels [65] but adapted to deal with the specific problematic of tau leptons. The strategy is presented in Section 6.2, after describing the reconstructed physics objects used in the analysis (Section 6.1). In Section 6.2 the event selection and optimization are then detailed; in Section 6.3.1 the technique to estimate the QCD background is explained, while in Section 6.5 the likelihood fit method used to extract the cross-section from the data is presented. The systematic uncertainties considered in this analysis are presented in Section 6.6. Finally, the results of the cross-section measurement are presented in the next Chapter.

This analysis uses 2.05 $\text{fb}^{-1}$of data collected by ATLAS at the LHC in *pp* collisions at a centre of mass energy of 7 TeV between March and August 2011. The data taking is divided into data periods regrouping LHC fills (or runs) in which the trigger conditions and the LHC operation mode remained stable. The data periods used in this thesis are those labeled as *Periods B-K*. Segments of individual ATLAS data runs, named *luminosity blocks*, are included if they were collected during LHC stable colliding beams runs, all the ATLAS sub-systems were fully-operational and the data quality has been estimated sufficient to be suitable for physics analysis. The total integrated luminosity considered in this analysis includes only luminosity blocks labeled as *good* in this sense, which constitute the 90% of the data taken. The uncertainty in the luminosity estimate is 3.7% [52]. Monte Carlo simulated samples are used to optimize the selection procedure, to calculate the acceptance for $t\bar{t}$ events and to evaluate the contributions from some background processes. The Monte Carlo simulated samples used in this thesis, have been generated as described in Chapter 2.6.

At the instantaneous luminosity of 1.3×10^{33}, the peak luminosity reached by the LHC during the data taking considered, the effect of pile-up of additional collisions becomes important. The effect of pile-up is considered in the Monte Carlo simulation: on average 6 extra event are added and a small uncertainty is taken into account to restore the mismatch in the number of observed vertices between data and Monte Carlo. The complete list of the datasets used in this analysis can be found in Appendix [?].

6.1 Objects definition

As mentioned before, to measure the cross-section in the $t\bar{t} \rightarrow \tau + jets$ decay channel we are interested in the reconstruction of all the objects that characterize the final state: hadronic jets, missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) that accounts for the neutrinos, and the τ lepton. We also identify electron and muon objects to use them as a veto to the tau lepton. In addition, the electron play an important role in the strategy used for the cross-section measurement (see next section). The definition of these objects is the following.

6.1.1 Electrons

Electron candidates are defined as electromagnetic clusters reconstructed with the cluster based or the track based algorithm (`ElectronAODCollection` with `author==1` or 3), consistent with the energy deposition of an electron in the calorimeter and associated with a well measured track, matched to the cluster. Most of the quality requirements are contained in the `ElectronTight` definition optimized to provide good efficiency (75%) for prompt electrons and good background rejection. Electrons passing the ID cut are required to have $E_{\mathrm{T}} \geq 25GeV$ (where E_{T} is defined as $E_{cluster}/\cosh(\eta_{track})$) and $|\eta_{\mathrm{cluster}}| < 2.47$ (where $\eta_{cluster}$ is the pseudorapidity of the cluster in the calorimeter associated with the candidate). The candidates in the crack region between barrel and calorimeter ($1.37 < |\eta_{cluster}| < 1.52$) are rejected. To exclude the background from photon conversion, we require the track to have an associated hit in the first pixel layer. The isolation requirement requires the transverse energy deposited in a cone of $\Delta R < 0.2$ (corrected to take into account the energy leak of the electron) to be less than 4 GeV.

6.1.2 Muons

The candidates are reconstructed combining the information of the tracks in muon chambers and the tracks in the inner detector. We consider muons contained in the `MuidMuonCollection` and we explicitly look for `Tight` and `Combined` muons (`author==MuonParameters::MuidCo`): We select muons with $p_{\mathrm{T}}> 10$ GeV and $|\eta| < 2.5$. For the track quality, we apply the following hits requirements:

- number of hits in the B layer > 0 if the track does not cross a dead region;
- number of pixel hits plus number of crossed dead pixel sensors > 1;
- number of SCT hits plus number of crossed dead SCT sensors ≤ 6;
- number of pixel holes plus number of SCT holes < 2;
- sum of TRT hits and TRT outliers > 5 and TRT outliers/(TRT hits + TRT outliers) < 0.9 for $|\eta| < 0.9$;
- TRT outliers/(TRT hits + TRT outliers) < 0.9 if TRT hits > 5 for $|\eta| < 0.9$.

To comply with the isolation requirements the calorimeter energy in a cone of $\Delta R = 0.3$ should be less than 4 GeV and the sum of track transverse momenta in a cone of $\Delta R = 0.3$ should be less than 4 GeV. Moreover, the distance from any jets with $p_{\mathrm{T}}> 20$ GeV is required to be $\Delta R > 0.4$, to suppress muons from heavy flavor decays inside jets.

6.1.3 Jets

Jets are reconstructed with the anti-k_T algorithm with distance parameter $\Delta R = 0.4$ starting from topological clusters of energy deposits in calorimeters. The energy is calibrated at the electromagnetic scale appropriate for the energy deposited by electrons or photons. Jets are calibrated with Monte Carlo based p_{T} and η dependent correction factors to restore the full hadronic energy scale (see Chapter 4). The jet is removed from the collection if the distance with the closest selected electron is smaller than $\Delta R = 0.2$; to avoid double-counting of electrons. The jet energy scale uncertainty varies from 2% to 7% as a function of jet p_{T} and η. The jet energy resolution and the reconstruction efficiency applied to the Monte Carlo samples have been measured in data. The jet is considered as a b-jet if the JetFitterCOMBNN tagger returns a value under 0.35, which corresponds to about 70% tagging efficiency and it is characterized at the same time by a light jet rejection fraction of 99 [42] which corresponds to a mistag rate <1%, as mentioned in Chapter 4.

6.1.4 Missing transverse energy

The missing transverse energy definition is `MET_RefFinal_em_tight` and is constructed from the vector sum of calorimeter energy deposits and muons reconstructed in the muon system. Cells are calibrated according to the reconstructed physics objects they are associated to: muons, electrons with $p_{\mathrm{T}} > 10$ GeV, jets and soft jets at the EM scale (consistently with the object definitions described above). Remaining calorimeter cells are also taken into account.

6.1.5 Taus

As mentioned before the most difficult object to identify is the τ lepton. In the present analysis, the definition of a τ lepton is based on the following requirements.

1. The tau candidate has to be found by the calorimeter-based algorithm (the `author` to be 1 or 3);
2. Only taus with 1 or 3 associated tracks are considered. Because of their different characteristics and purity, we decided to consider separately the contribution of tau events with 1 or 3 associated tracks (respectively, 1-prong and 3-prongs tau events);
3. As an identification criteria we used the `tau_SafeLlhTight` , whose efficiency of selecting τ events is $\sim 30\%$;
4. The overlap removal is performed in the following way: first the overlap between selected taus and selected electrons and between selected taus and selected muons is considered. Then we consider overlap between the taus that satisfy the above selection and with jets, both objects are required to have $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$. When an electron or a muon overlaps with a tau, we reject the tau, while when a jet overlaps with the tau we reject the jet. We define as overlapping two objects found within a relative distance ΔR smaller than 0.2;
5. Electron veto based on Boosted Decision Tree continuous discriminant output: `BDTEleScore` > 0.5. The ATLAS $t\bar{t} \rightarrow \tau + e/\mu$ analysis note [62] shows that in a $Z \rightarrow \tau\tau$ sample 85% of the reconstructed τ leptons decaying hadronically satisfy the requirement `BDTEleScore` > 0.5 while the rejection factor for electron is about 60.
6. Discrimination of taus from QCD jets, requiring `BDTJetScore` > 0.6. In [62] is shown that a fit of `BDTJetScore` discriminant lead to the choice of a cut at 0.7 to select a tau sample with good purity τs; in this analysis we chose a cut at 0.6 as a good compromise between signal acceptance and background rejection.
7. The tau candidate selected with the requirements described above is accepted if it matches with the `EF_tau29_medium` tau trigger object within $\Delta R < 0.2$.

6.2 $t\bar{t} \rightarrow \tau + jets$ events selection and analysis strategy

At the root of the event selection there is the choice of the trigger: we use the τ+$E_{\mathrm{T}}^{\mathrm{miss}}$ trigger with a threshold of 29 GeV on the tau p_{T} and 35 GeV on calorimeter-based

$E_{\mathrm{T}}^{\mathrm{miss}}$(`EF_tau29_medium_xe35_noMu`). Such trigger selection requires to apply the corresponding kinematical cuts in the analysis. We select events with an off-line reconstructed and identified tau object with $p_{\mathrm{T}} > 35$ GeV and reconstructed $E_{\mathrm{T}}^{\mathrm{miss}} > 60$ GeV, in order to be in the plateau of the trigger efficiency turn-on curves for the respective variables. Regarding the tau object, we also require that the reconstructed tau candidate matches the tau object triggered in the Event Filter. Events are vetoed if they contain additional leptons (electrons and muons as defined in the previous section).

The contributions of events characterized by 1-prong or 3-prong hadronic tau decays (we refers here to them as *1-prong channel* and *3-prong channel* for conciseness) are treated separately as the fake tau contributions and the signal-to-background ratio (S/B) differ.

	1-prong tau		**3-prong tau**	
	= 3 jets	≥ 4 jets	= 3 jets	≥ 4 jets
$t\bar{t}$	108.8 ± 2.3	189.0 ±3.1	56.9 ± 1.7	94.6 ±2.2
W+jets	501.1 ± 23.5	172.3 ±11.1	208.2 ± 12.4	99.2 ±8.6
Z+jets	60.8 ± 6.2	38.4 ±5.0	34.8 ± 4.8	18.4 ±3.3
Single top	20.6 ± 1.96	12.2 ±1.4	11.2 ± 1.3	6.3 ±1.0
Dibosons	2.9 ± 0.9	1.2 ±0.5	0.8 ± 0.5	0.0 ±0.0
QCD	271.9 ± 16.5	213.0 ±14.6	518.0 ± 22.7	331.4 ± 18.2
Total prediction	970.5 ±27.9	633.1 ±16.9	790 ±18	550.4 ±13.2
Data	848.0	569.0	489.0	375.0

Table 6.1: 1-prong and 3-prong tau yields after the event selection, without applying b-tagging.

The $t\bar{t}$ final state contains in addition to the hadronically decaying tau at least four jets, two of them originated by b quarks. Jets are required in the analysis to have p_{T} above 25 GeV, as done in most ATLAS top-related analysis based on criteria related to jet energy resolution, energy scale and b-tagging. Table 6.1 gives the yields of events with exactly 3 jets and events with at least 4 jets. The rates observed in data are compared to the Monte Carlo expectation for signal and backgrounds . Note that $t\bar{t}$ events populate mostly the 4 jets inclusive category, but that a non negligible part of the signal falls into the 3 jets exclusive category.

The data rates are affected by a filter applied on the data stream resulting from jet, tau, $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers. This filter is applied when producing the reduced data sets (D3PDs) for ATLAS top related analysis, The filter was designed to provide high efficiency for top events while reducing overall significantly the volume of data in that stream. The filter requires at least 4 jets with p_{T} of more than 20 GeV, two of them with at least 40 GeV, or 5 jets with p_{T} of more than 20 GeV. Note that hadronically decaying taus are counted as jets for the purpose of the filter. The filter is also applied on the Monte Carlo events, for consistency. The efficiency of this filter for $t\bar{t}$ events with a tau with $p_{\mathrm{T}} > 35$ GeV and reconstructed $E_{\mathrm{T}}^{\mathrm{miss}} > 60$ GeV is of 99% both for events with 4 jets or more with $p_{\mathrm{T}} > 25$ GeV and for events with 3 jets $p_{\mathrm{T}} > 25$ GeV. The table shows that the most relevant background contributions come from the $W + jets$ and QCD multi-jets processes. The latter is the most critical and difficult background to estimate in this analysis, specially in the 3-prong mode. Note that the errors quoted on the expected rates in the table are purely statistical. More details on the normalization of the backgrounds and the associated systematic uncertainties are given later in this chapter.

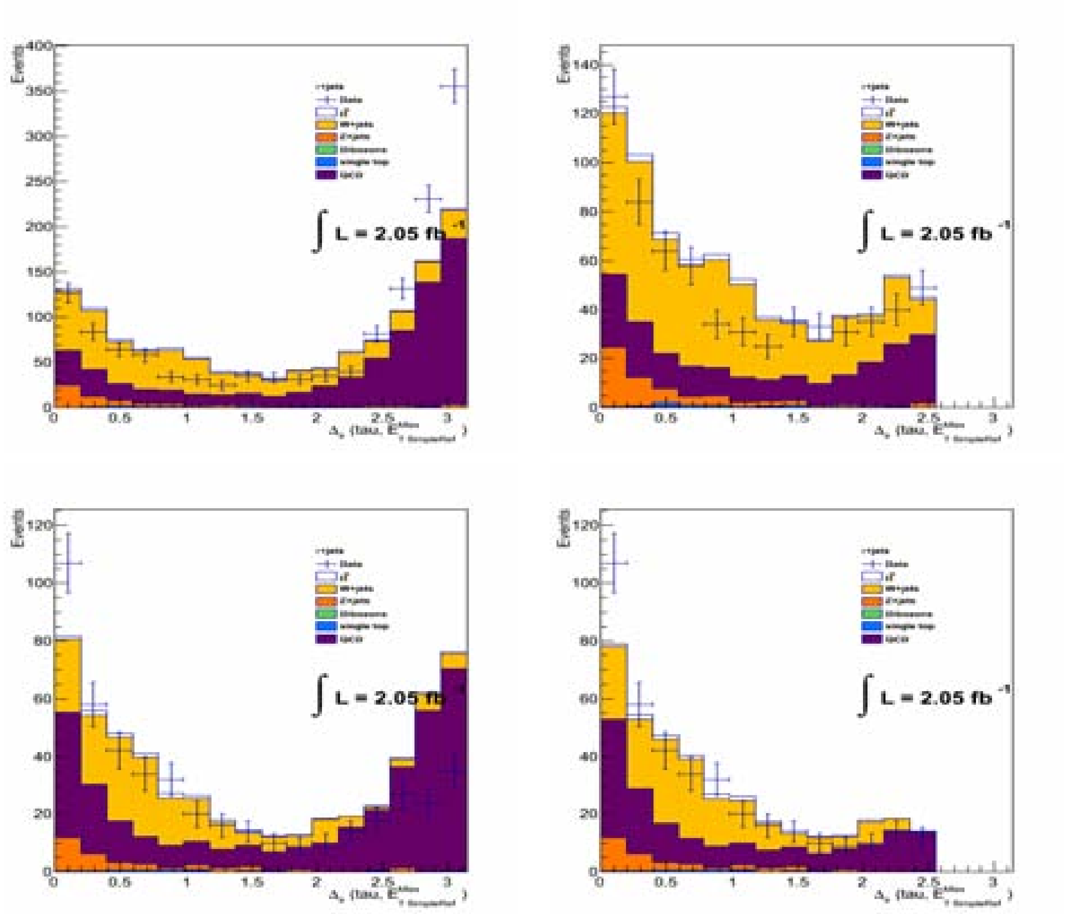


Figure 6.1: Comparison of the $\Delta\phi(\tau,\ E_{\mathrm{T}}^{\mathrm{miss}})$ distribution in data and Monte Carlo simulated samples, before (left) and after (right) applying the cut at 2.5. Exactly three jets and exactly zero b-tagged jets are required. 1-prong (3-prong) tau channel is shown at the top (bottom).

Selection cuts have been optimized for maximum QCD multi-jet events rejection while keeping good acceptance for the signal. This was done for the choice of the tau identification criteria as discussed in the previous section. On the other hand, several kinematical variables have been investigated for their potential discrimination against QCD multi-jet events. The $E_{\mathrm{T}}^{\mathrm{miss}}$ provides a good handle and the cut has been set at 60 GeV. Another interesting variable is the the angle $\Delta\phi$ between the tau and the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector. As shown in Figure 6.1, there is a large contribution of QCD at large $\Delta\phi$ values. When the $E_{\mathrm{T}}^{\mathrm{miss}}$ in the event is generated mostly by a mis-reconstructed jet corresponding to the tau candidate, the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector is back-to-back to the tau. Consequently, these events will also have a

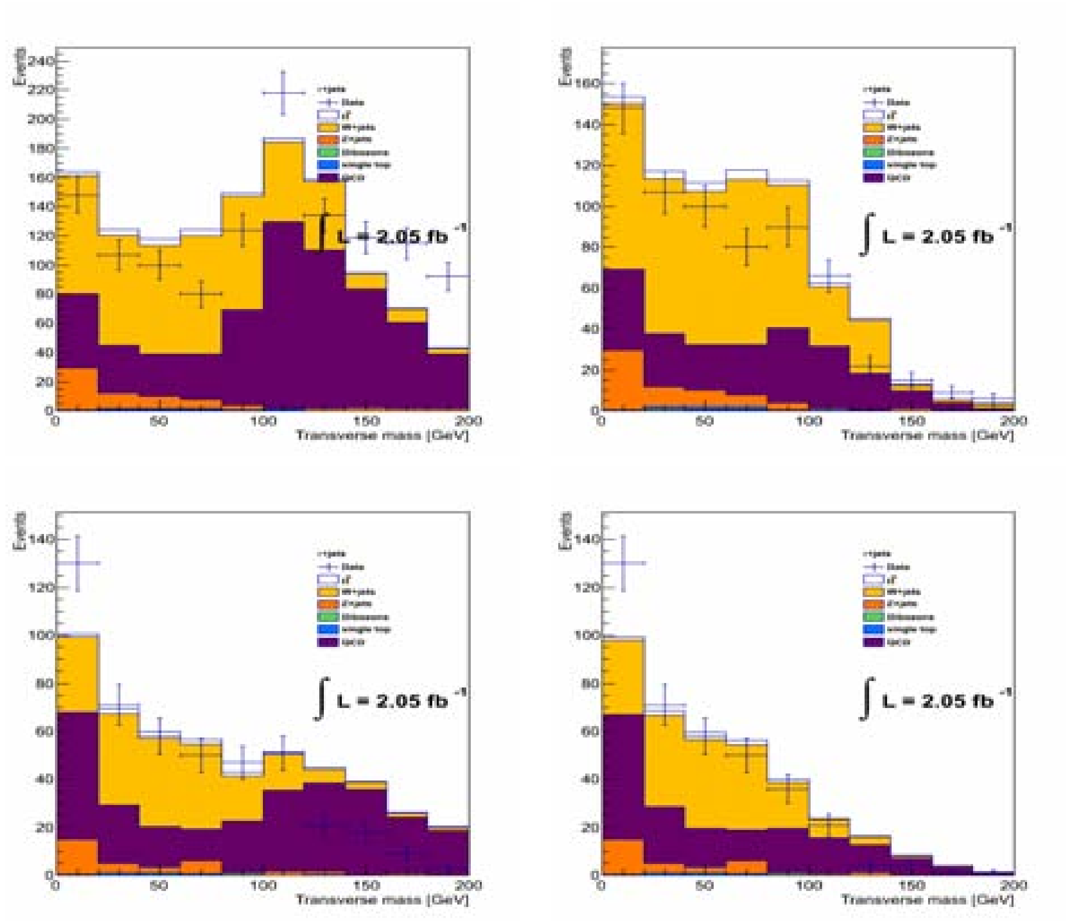


Figure 6.2: Comparison of the W transverse mass distribution in data and Monte Carlo simulated samples, before (left) and after (right) applying the cut on $\Delta\phi(\tau,\ E_{\mathrm{T}}^{\mathrm{miss}})$ at 2.5. Exactly three jets and exactly zero b-tagged jets are required. 1-prong (3-prong) tau channel is shown at the top (bottom).

high reconstructed W transverse mass as seen in the left column of Figure 6.2. There is a clear accumulation of QCD events above 100 GeV that is observed. There is a threshold effect on the transverse mass of such events, which value corresponds approximately to the scalar sum of $E_{\mathrm{T}}^{\mathrm{miss}}$ and tau p_{T} and is thus shaped by the cuts applied on these variables at trigger and analysis level. A cut on $\Delta\phi < 2.5$ reduces significantly this source of fake tau events and the effect on the W transverse mass distribution is clearly seen on the right of Figure 6.2.

The other important handle against the QCD background is the identification of the jets originating from b-quarks. In this analysis, the JetFitterCombNN tagger has been used at its working point which corresponds to 70% of tagging efficiency. The selected events can be classified into three categories in terms of the number of jets tagged as b-jets: events with exactly 0 b-jets; events with exactly 1 b-jet and events with at least 2 b-jets. Table 6.2 gives the yields for each b-tag region.

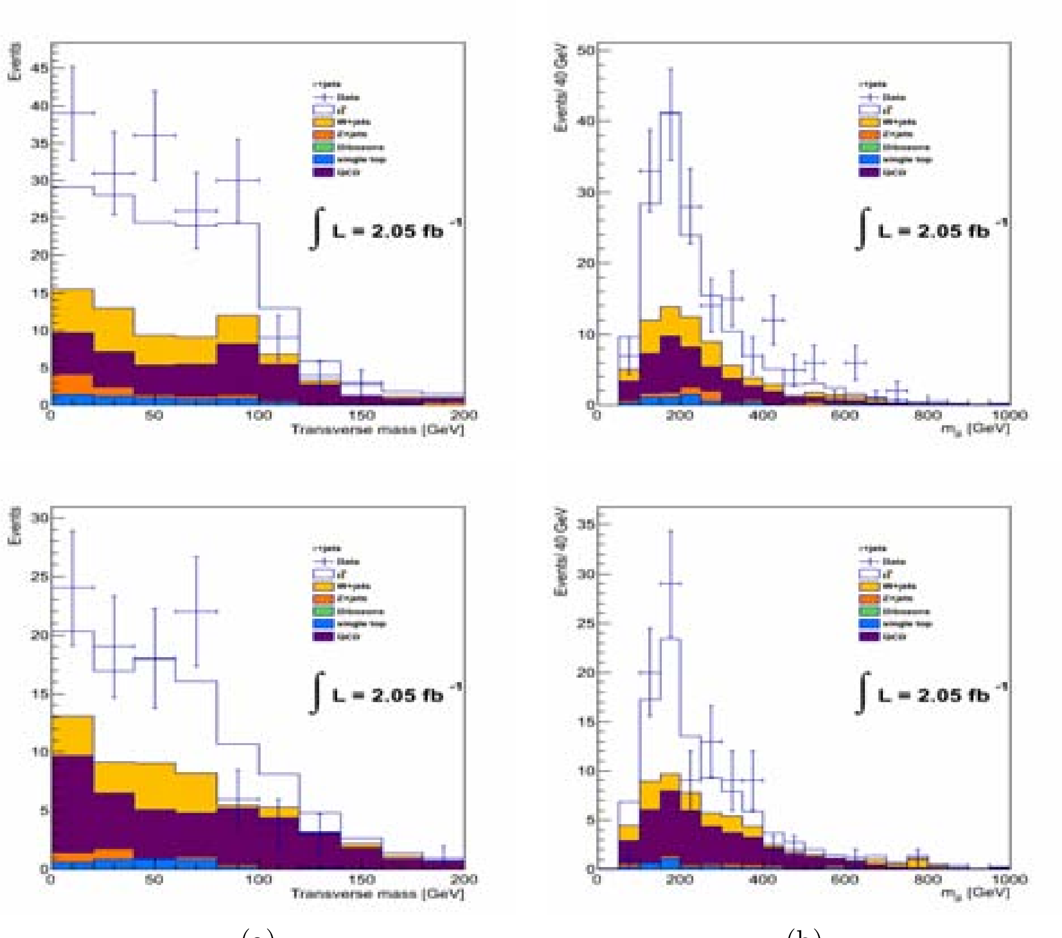
	1-prong		3-prong	
	= 3 jets	≥ 4 jets	= 3 jets	≥ 4 jets
0 *b*-tag exclusive				
$t\bar{t}$	22.4 ± 1.1	24.8 ±1.2	11.2 ± 0.8	13.6 ±0.9
W+jets	451.9 ± 22.6	140.3 ±9.6	178.0 ± 10.9	80.4 ±7.4
Z+jets	57.0 ± 6.0	32.3 ±4.6	28.6 ± 4.3	15.3 ±3.0
Single top	5.9 ± 1.0	3.5 ±0.8	3.7 ± 0.7	1.6 ±0.5
Dibosons	2.5 ± 0.8	1.1 ±0.5	0.7 ± 0.5	0.0 ±0.0
QCD	224.5 ± 11.6	156.7 ±8.53	432.0 ± 11.9	285.9 ±9.4
Total prediction	764.1 ±26.2	358.6 ±13.7	654.2 ±16.7	397.0 ±12.3
Data	646.0	296.0	378.0	221.0
1 *b*-tag exclusive				
$t\bar{t}$	55.3 ± 1.6	82.4 ±2.1	29.9 ± 1.2	42.5 ±1.4
W+jets	44.4 ± 6.0	27.0 ±5.2	25.3 ± 5.2	16.0 ±3.9
Z+jets	4.1 ± 1.8	5.8 ±2.1	6.4 ± 2.3	2.4 ±1.2
Single top	10.4 ± 1.4	6.0 ±0.9	5.3 ± 0.9	3.6 ±0.7
Dibosons	0.3 ± 0.2	0.1 ±0.1	0.0 ± 0.0	0.0 ±0.0
QCD	41.4 ± 5.2	35.4 ±5.4	40.0 ± 1.9	37.1 ±1.9
Total prediction	155.9 ±8.5	156.8 ±8.1	106.9 ±6.2	101.7 ±4.8
Data	156.0	179.0	90.0	97.0
2 *b*-tags inclusive				
$t\bar{t}$	31.3 ± 1.2	81.7 ±2.0	15.9 ± 0.9	38.5 ±1.4
W+jets	4.8 ± 2.0	4.6 ±2.3	5.2 ± 2.8	2.5 ±1.8
Z+jets	0.0 ± 0.0	0.5 ±0.5	0.0 ± 0.0	1.0 ±1.0
Single top	4.3 ± 0.9	2.9 ±0.7	2.1 ± 0.6	1.1 ±0.5
Dibosons	0.1 ± 0.1	0.0 ±0.0	0.1 ± 0.1	0.0 ±0.0
QCD	6.0 ± 2.7	20.9 ±4.0	6.0 ± 0.8	8.3 ±1.1
Total prediction	46.5 ±3.6	110.6 ±5.1	29.3 ±3.1	51.4 ±2.8
Data	46.0	94.0	21.0	57.0

Table 6.2: Yields after the event selection, for events with exactly 0 jet, 1 jet or at least 2 jets tagged as *b*-jets. The 1-prong and 3-prong tau event yields are shown separately.

The following observations can be made:

- in the 0 *b*-tag exclusive region, the *W*+jets and QCD background events dominate;
- in the 1 *b*-tag exclusive region, the signal-to-background ratio (S/B) is around 1:2 for 3 jets exclusive and around 1:1 for 4 jets inclusive events;
- in the 2 *b*-tag inclusive region, the signal-to-background ratio (S/B) is around 2:1 for 3 jets exclusive and around 3:1 for 4 jets inclusive events.

By splitting events according to the jet multiplicity and the number of b-tagged jets, we obtain six different regions characterized by different contributions of the signal and the main backgrounds. We apply this procedure separately for the 1-prong and 3-prong channels. The analysis strategy based on a likelihood fit with templates for the signal and backgrounds will make use of the full information such as to profit from the maximum amount of signal events and be able to characterize as precisely as possible the contribution of the most relevant backgrounds.



(a) (b)

Figure 6.3: (a) W transverse mass distribution in data and Monte Carlo simulation or 4 jets inclusive events with one b-tagged jet in 1-prong (top) and 3-prong (bottom) channels. (b) Reconstructed mass of the hadronic top in data and Monte Carlo simulation in 1-prong (top) and 3-prong (bottom) channels for the same events.

We investigated which variable is most appropriate to be used for the fit, among them the W transverse mass and the reconstructed mass of the candidate hadronic top defined as the combination of 3 jets having the highest vector sum p_{T}(labelled m_{jjj}, [65]). The Figure 6.3 shows examples of both distribution for 4 jets inclusive events with one b-tagged jet for the 1-prong and the 3-prong channels. The W transverse mass discriminates between the QCD multi-jet background and W to tau decays from the $t\bar{t}$ signal or $W + jets$ events, while m_{jjj} differentiates the resonant $t\bar{t}$ signal from the smooth shape of $W + jets$ and QCD. We choose to use the W transverse mass for its more efficient QCD discrimination. Another reason for that choice is the potential extension of this analysis to the charged Higgs searches with taus in the final state: the mass of the Higgs boson could be observed as a characteristic shoulder in the transverse mass distribution.

In the analysis, we will also make use of an auxiliary sample of events corresponding to the semileptonic decay mode in the electron channel. The experience from the $t\bar{t}$ cross section measurement in the semileptonic electron and muon channels [65] has shown that these channels are characterized by higher statistics of $t\bar{t}$ and $W + jets$ events and are contaminated by much less QCD background. This sample will provide valuable information on various aspects of the contributing processes, notably on $W + jets$.

To select a sample of $t\bar{t} \rightarrow e + jets$ events, we use the same baseline cut-flow used for the $t\bar{t} \rightarrow \tau + jets$ events, except for the following requirements:

- the trigger: single electron trigger `EF_e20_medium` or `EF_e22_medium`, according to the period for data and randomly for the MC;
- exactly one electron with $p_{\mathrm{T}} > 35$ GeV;
- other leptons (taus and muons) in the event are vetoed.

The yields corresponding to the electron channel selection are reported in table 6.2.

Electron channel		
	= 3 jets	≥ 4 jets
$t\bar{t}$	2582.5 ± 13.1	4761.9 ±18.0
W+jets	9830.9 ± 119.6	4133.2 ±65.4
Z+jets	433.6 ± 19.4	289.3 ±15.3
Single top	433.4 ± 10.2	315.4 ±8.0
Dibosons	169.6 ± 8.0	58.8 ±4.7
QCD	109.0 ± 19.3	54.5 ±16.1
Total prediction	13559.0 ±124.0	9613.1 ±72.0
Data	13304.0	10241.0

Table 6.3: Electron channel yields at the end of the selection, without any b-tagging requirement.

6.3 Inputs to the cross-section measurement

Before going into the details of the cross-section measurement, it is useful to describe some aspects related to the determination of the background processes and others related to the acceptance for signal and background.

6.3.1 Background estimation

As seen in Table 6.2 , the main backgrounds to the $t\bar{t} \rightarrow \tau + jets$ signal process are the W-boson production in association with jets and the QCD multi-jet processes. The understanding and the correct estimation of these backgrounds is crucial for a precise measurement of the $t\bar{t}$ cross-section. Other contributions to the background come from the Z boson produced in association with jets, single top processes and diboson productions like WW, WZ and ZZ. In this thesis, all the backgrounds are simulated by Monte Carlo, except for the QCD which is estimated using two complementary data-driven methods, the Matrix Method and the so-called TRF Method .

$W + jets$ estimation and uncertainties

In this section, we describe the *a priori* normalization of the $W + jets$ process used in the analysis and their related uncertainties.

The total rate of the $W + jets$ processes with 3- and $\geq$ 4-jets are poorly predicted by LO generators. To reduce that uncertainty , one can use the ratio of $W + n + 1$ jets events to $W + n$ jets events that is expected to be constant, the so-called Berends scaling [56]. The number of events in each bin multiplicity can then be extrapolated from the 2 jets bin multiplying by such ratio. The cross-section in the 2 jets bin has been obtained from a measurement of the W+jets asymmetry [57]. The corresponding scale factors, that are used as *a priori* normalization, are 0.833 for the 3 jets bin and 0.952 for the 4 jets inclusive bin. At a later stage in the analysis, we will extract from the fit the absolute normalization taking advantage of the constraint from data. Another issue is heavy flavor fraction in the $W + jets$ samples. In addition to the inclusive W+jets MC samples we use additional samples which simulate $W + b\bar{b}$, $W + c\bar{c}$ and $W + c$ processes at LO. Since these final states could be contained also in the inclusive sample as part of the parton shower, one needs to apply an overlap removal procedure to avoid partial double counting. This is implemented in the HFOR tool [58]. The W+jets heavy flavor fraction has been measured in W+$2jets$ data events studying the properties of secondary vertices. From this studies scale factors have been extracted to apply to the MC predictions: $W + b\bar{b}$ and $W + c\bar{c}$ samples are scaled by 1.63 ± 0.76 while $W + c$ sample is scaled by a factor 1.11 ± 0.35. The sum of the inclusive events and the heavy flavor events is kept constant, according to the normalization described above. However the scaling affects the composition of $W + jets$ background. The impact of the uncertainty on the fraction can be estimated by means of modified templates where the fractions of $W + b\bar{b}$, $W + c\bar{c}$ and $W + c$ are modified up and down by their uncertainty.

Matrix Method for QCD multi-jets background estimation

The Matrix Method (MM), developed at the D0 experiment [59], allows to estimate the contribution of mis-identified taus events in QCD multi-jets events from data, by defining two different samples that differ in the tau identification criteria used and describing the relation between them by a system of equations. The solution of the system of equations constitutes the weight to be applied to data in order to obtain the QCD contamination in the selected sample. We ll refer to the first sample as the *tight* and to the second as the *loose* selection. The tight sample is a subset of the loose sample and corresponds to

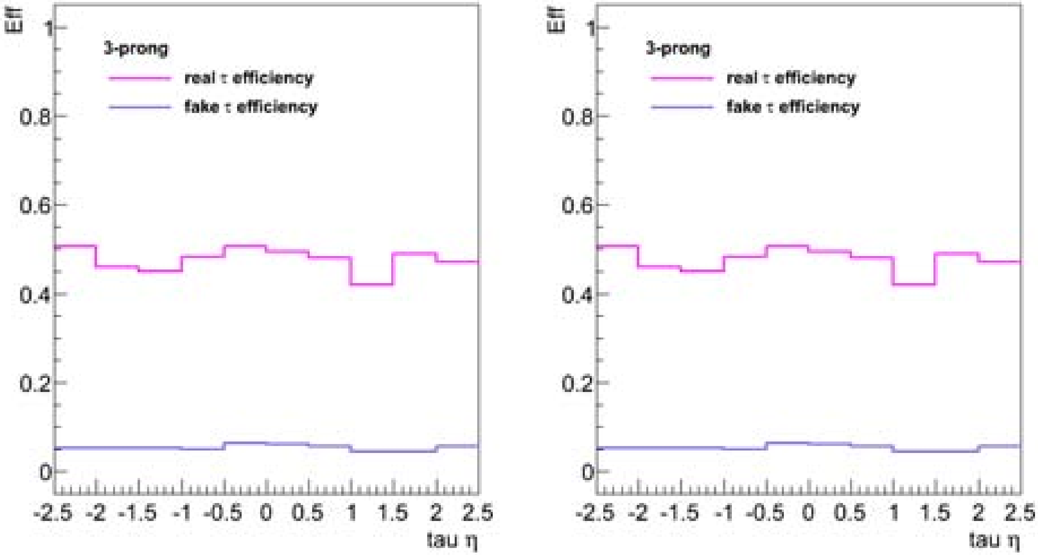


Figure 6.4: Efficiencies of select real and fake taus used to calculate the event weights of the Matrix Method. At left (right), the 1-prong (3-prong) channel is shown.

the selected sample for the analysis. In both samples the tau is defined as described in Section 6.1 except for two selection criteria:

1. the tau ID requirement: to fulfill the loose (tight) selection we require the event to contain a tau identified according to the `SafeLlhLoose` (`SafeLlhTight`) operational point;

2. the tau-jet discrimination cut: no cut is applied on `BDTJetScore` in the loose selection.

We consider that each sample is composed of real and misidentified taus, such that it is licit to write the following formulas:

$$N^{loose} = N^{loose}_{real} + N^{loose}_{fake}; \tag{6.1}$$

$$N^{tight} = N^{tight}_{real} + N^{tight}_{fake}. \tag{6.2}$$

Therefore the efficiency for a loose tau to also pass the tight selection can be defined as:

$$\epsilon^{real} = \frac{N^{tight}_{real}}{N^{loose}_{real}} \quad and \quad \epsilon^{fake} = \frac{N^{tight}_{fake}}{N^{loose}_{fake}}. \tag{6.3}$$

And the Eq. 6.2 can be rewritten as:

$$N^{tight} = \epsilon_{real} \cdot N^{loose}_{real} + \epsilon_{fake} \cdot N^{loose}_{fake}. \tag{6.4}$$

We thus have a linear system of equations (Eq. 6.1 and 6.4) whose solutions are the following:

$$N_{real}^{tight} = \frac{\epsilon_{real}}{\epsilon_{real} - \epsilon_{fake}}(N^{tight} - \epsilon_{fake} N^{loose}) \tag{6.5}$$

$$N_{fake}^{tight} = \frac{\epsilon_{fake}}{\epsilon_{real} - \epsilon_{fake}}(\epsilon_{real} N^{loose} - N^{tight}) \tag{6.6}$$

In this way we can estimate the number of events in the tight selection (the selection used in the cross-section measurement) that arise from the real and misidentified taus, respectively. In order to have a reliable estimation with this method, the following conditions have to be satisfied:

1. to have enough statistical precision, the efficiencies real and fake have to be sufficiently different because they enter in the denominator as a difference: $\epsilon_{real} - \epsilon_{fake}$;
2. both efficiencies have to be as much as possible independent of the topological features of the sample, so that they can be calculated in control samples and then applied in the region of interest for the analysis;
3. if there is any significant dependence of the efficiencies on the kinematics or topology of the sample, it must be parameterized.

The method described above can be implemented as a tool to estimate the contribution of QCD multi-jets weighting event-by-event the data sample. Events that satisfy the loose selection will be weighted by:

$$w_{loose} = \frac{\epsilon_{fake} \cdot \epsilon_{real}}{\epsilon_{real} - \epsilon_{fake}}, \tag{6.7}$$

while an event that satisfies the tight selection will be weighted by:

$$w_{tight} = \frac{\epsilon_{fake}(\epsilon_{real} - 1)}{\epsilon_{real} - \epsilon_{fake}}. \tag{6.8}$$

The result is an estimation of the contribution of QCD events to the loose or tight sample respectively.

Ideally the efficiencies are measured in the data in a signal dominated or background dominated sample. Z boson decays to leptonic pairs, in the case of electrons and muons, provide a high-statistics and high-purity sample of leptons. The case of taus is less favorable, with less statistics and worse signal to background ratio. We opted to measure the ϵ_{real} from $t\bar{t}$ Monte Carlo sample. To do that we required in the event:

- exactly 1 tau with $p_T > 35$ GeV, matched to a truth tau (within $\Delta R < 0.2$ of a true visible tau decay);
- matching to the triggered tau;
- events with additional leptons are vetoed;
- at least 1 jet with $p_T > 25$ GeV without any b-tagging requirement;
- $E_{\mathrm{T}}^{\mathrm{miss}} > 35$ GeV;

- $\Delta\phi(\tau, E_T^{miss}) < 2.5$.

On the other hand, to measure the ϵ_{fake} we select in the JetTauEtMiss data stream the events with the following requirements:

- exactly 1 tau with $p_T > 35$ GeV;
- matching to the triggered tau;
- events with additional leptons are vetoed;
- at least 1 jet with $p_T > 25$ GeV without any b-tagging requirement;
- $40GeV < E_{\mathrm{T}}^{\mathrm{miss}} < 60GeV$;
- $\Delta\phi(\tau, E_T^{miss}) < 2.5$.

The cut on the missing transverse energy has been set between 40 and 60 GeV because 40 GeV is the lower limit imposed by the trigger and 60 GeV is the cut that separates the signal dominated region (with a reasonable S/B ratio for the analysis) from the QCD dominated region. The residual contribution of true taus, has been estimated in $t\bar{t}$ and W+jets Monte Carlo events, selected with the same cuts used to define the QCD dominated region in data. Such contribution which amount to the 19% (0.06%) for tight (loose) events in the 1-prong channel and to 11% (0.02%) in the 3-prong channel, has been subtracted from the data sample used to calculate the fake efficiency. The dependence of the efficiency on the tau pseudorapidity is taken into account when the weights per event are calculated. In Figure 6.4 shows the real and fake efficiencies as a function of the tau η, in both 1-prong and 3-prong channels. The dependence of the efficiency fake on other variables (tau p_{T}, $E_{\mathrm{T}}^{\mathrm{miss}}$) has been studied but their effect are small. The largest source of uncertainty comes from the truth subtraction method. A systematic uncertainty of 50% is assigned. It is set to 100% for samples containing a b-tagged jet. The above efficiencies are used to calculate the *a priori* rates. Later in the analysis, we adjust the level of QCD background to the data.

TRF method for QCD estimation

The Matrix Method provides potentially a quite accurate estimation, however it may be in some cases limited by statistics, in particular for samples with two jets tagged as b-jets. In such cases, the W transverse mass distribution may present spikes. In order to obtain smoother shapes more reliable for the fit, we applied in addition another data-driven method: the Tagging Rate Fraction (TRF) technique. The main idea is to use as QCD shape template for the fit the untagged distribution and apply a per-jet tagging weight, building in this way the 1 b-tag exclusive and the 2 b-tag inclusive distributions. The procedure is the following.

- We select a *loose-not-tight* sample of data in a QCD enriched region (control region) defined as the events with at least 1 jet, none of them tagged, and missing transverse energy between 40 and 60 GeV. The *loose-not-tight* selection of tau events consists in requiring to be satisfied the `SafeLlhLoose` tau identification criteria and to exclude the events with taus that satisfy also the `SafeLlhTight` criteria. The other cuts are the same as the tight selection, but the cut on `BDTJetScore` that is not applied.

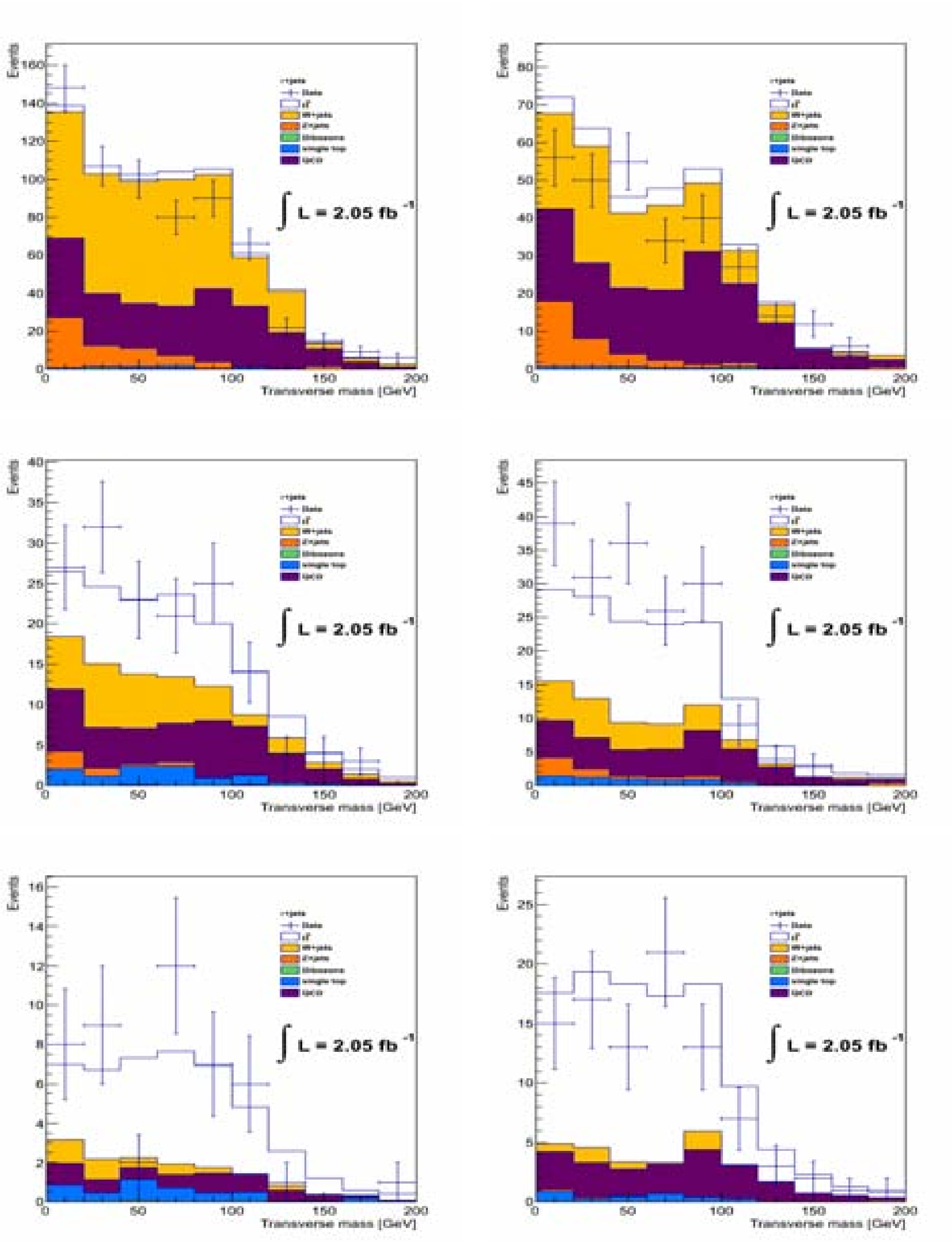


Figure 6.5: W transverse mass distribution in the 0- (top), 1- (medium) and $\geq$ 2- (bottom) b-jet multiplicity bins in the 1-prong tau channel. The data (dots with error bars) are compared to the expectation.

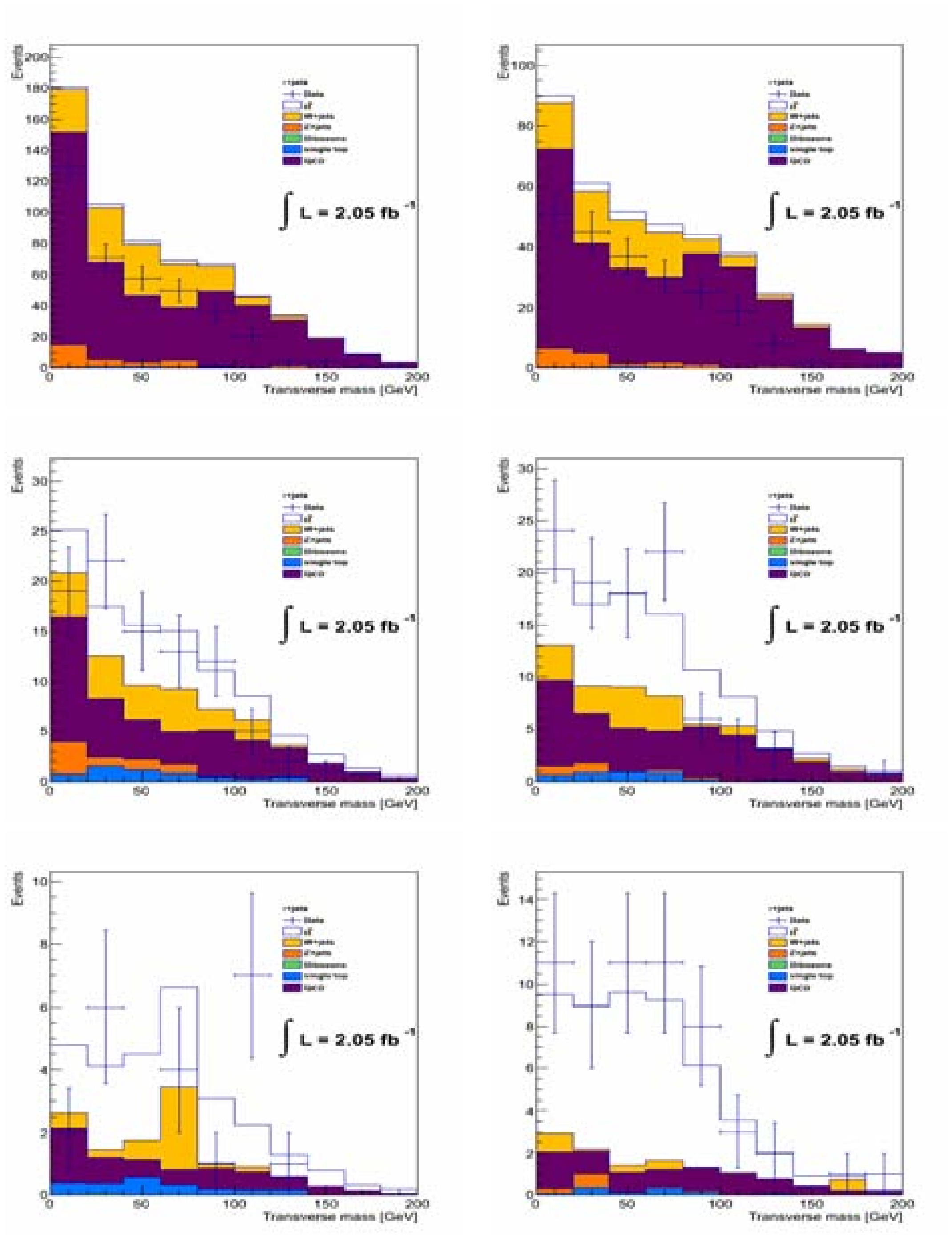


Figure 6.6: W transverse mass distribution in the 0- (top), 1- (medium) and $\geq$ 2- (bottom) b-jet multiplicity bins in the 3-prong tau channel. The data (dots with error bars) are compared to the expectation.

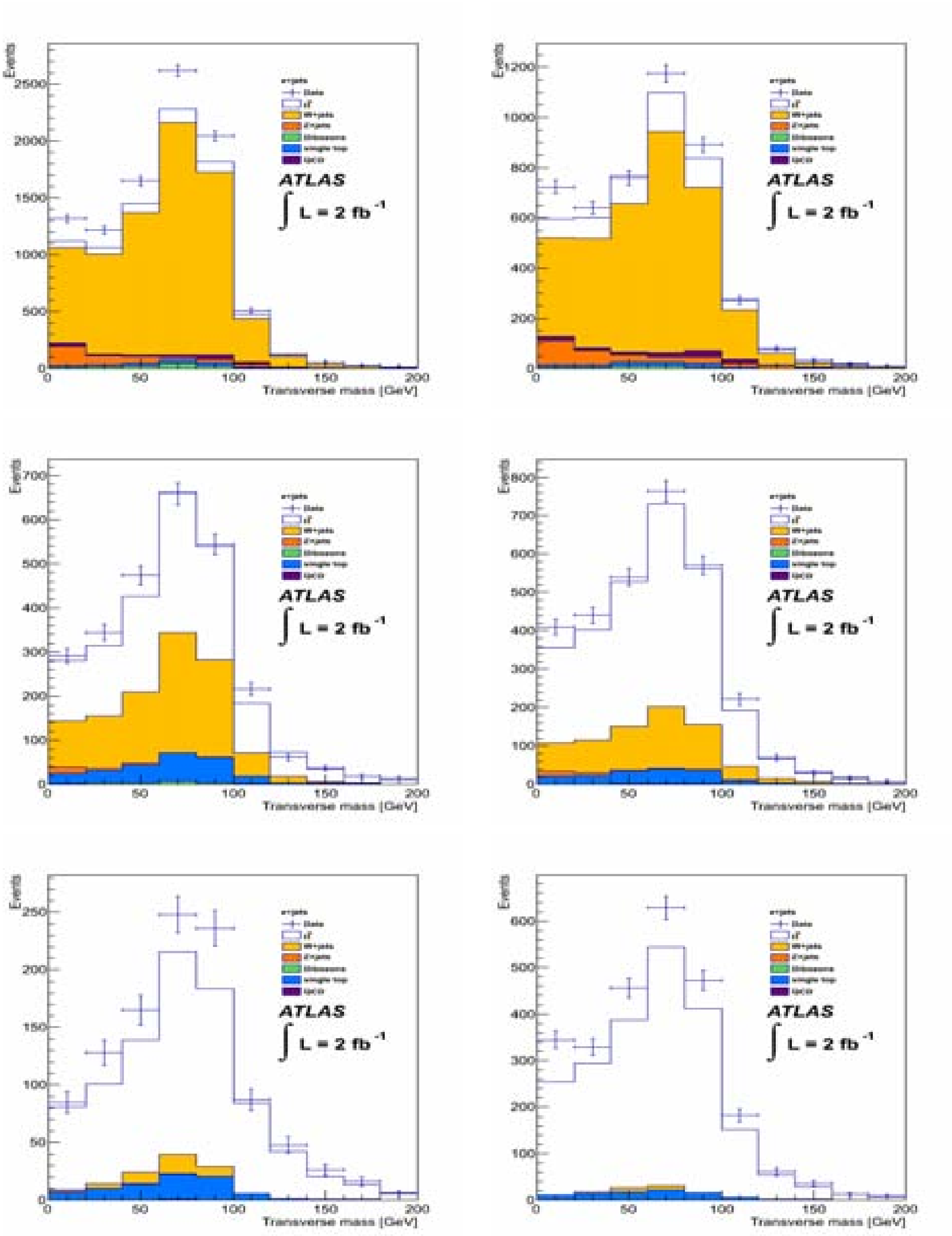


Figure 6.7: W transverse mass distribution in the 0- (top), 1- (medium) and $\geq$ 2- (bottom) b-jet multiplicity bins in the electron channel. The data (dots with error bars) are compared to the expectation.

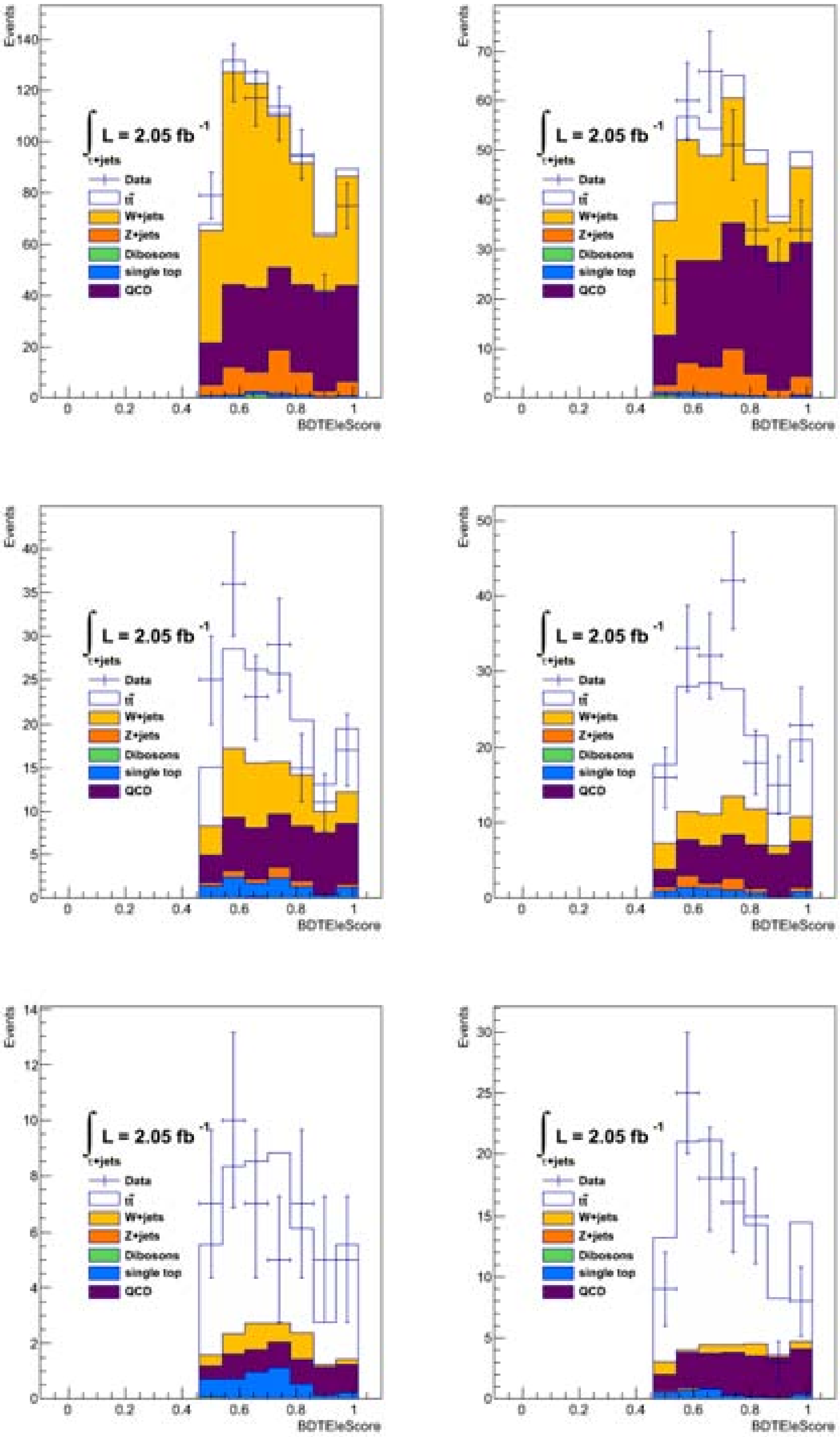


Figure 6.8: Data-MC comparison: BDTEleScore in the 1-prong tau channel. A cut at 0.5 is applied.

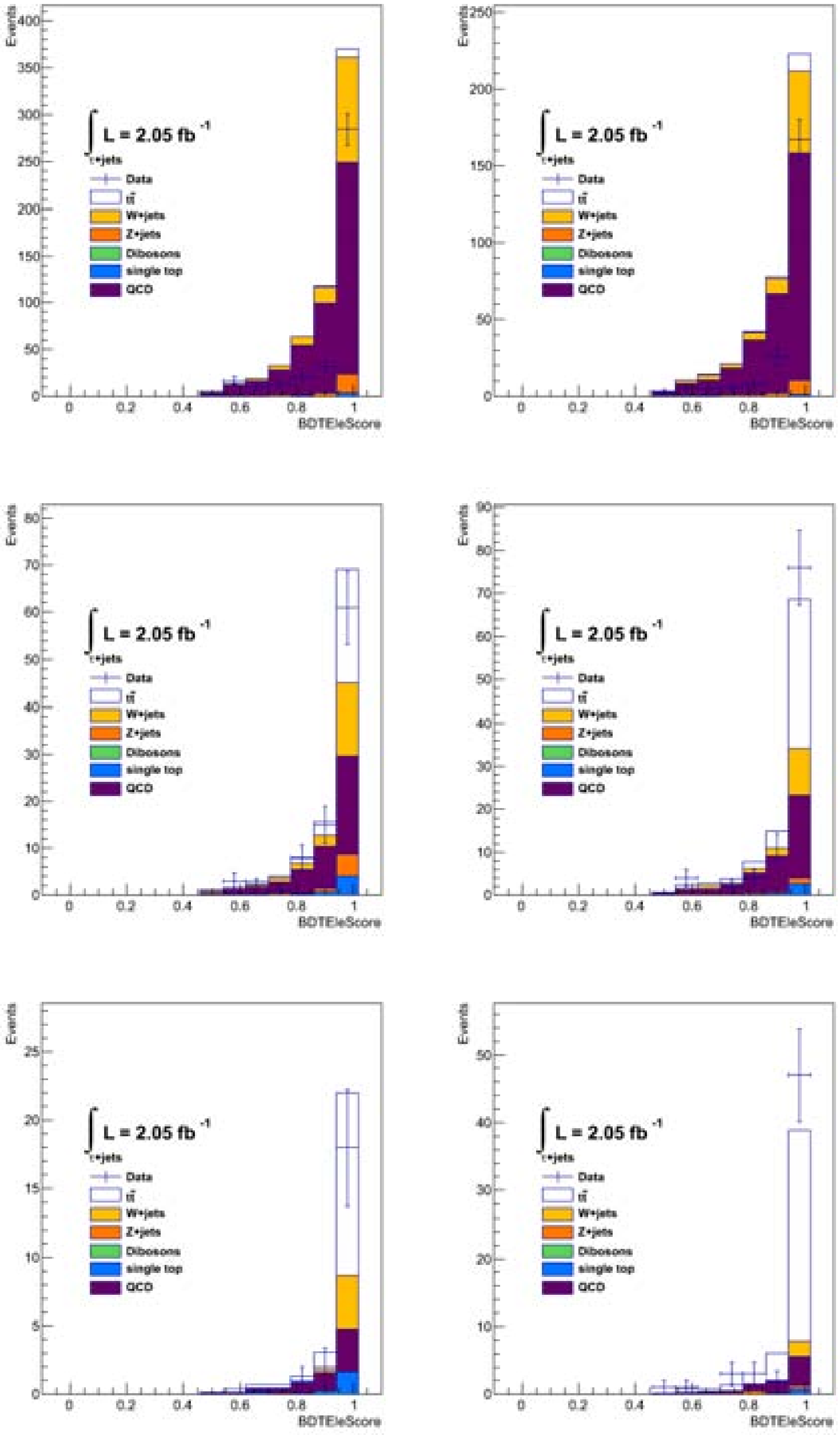


Figure 6.9: Data-MC comparison: BDTEleScore in the 3-prong tau channel. A cut at 0.5 is applied.

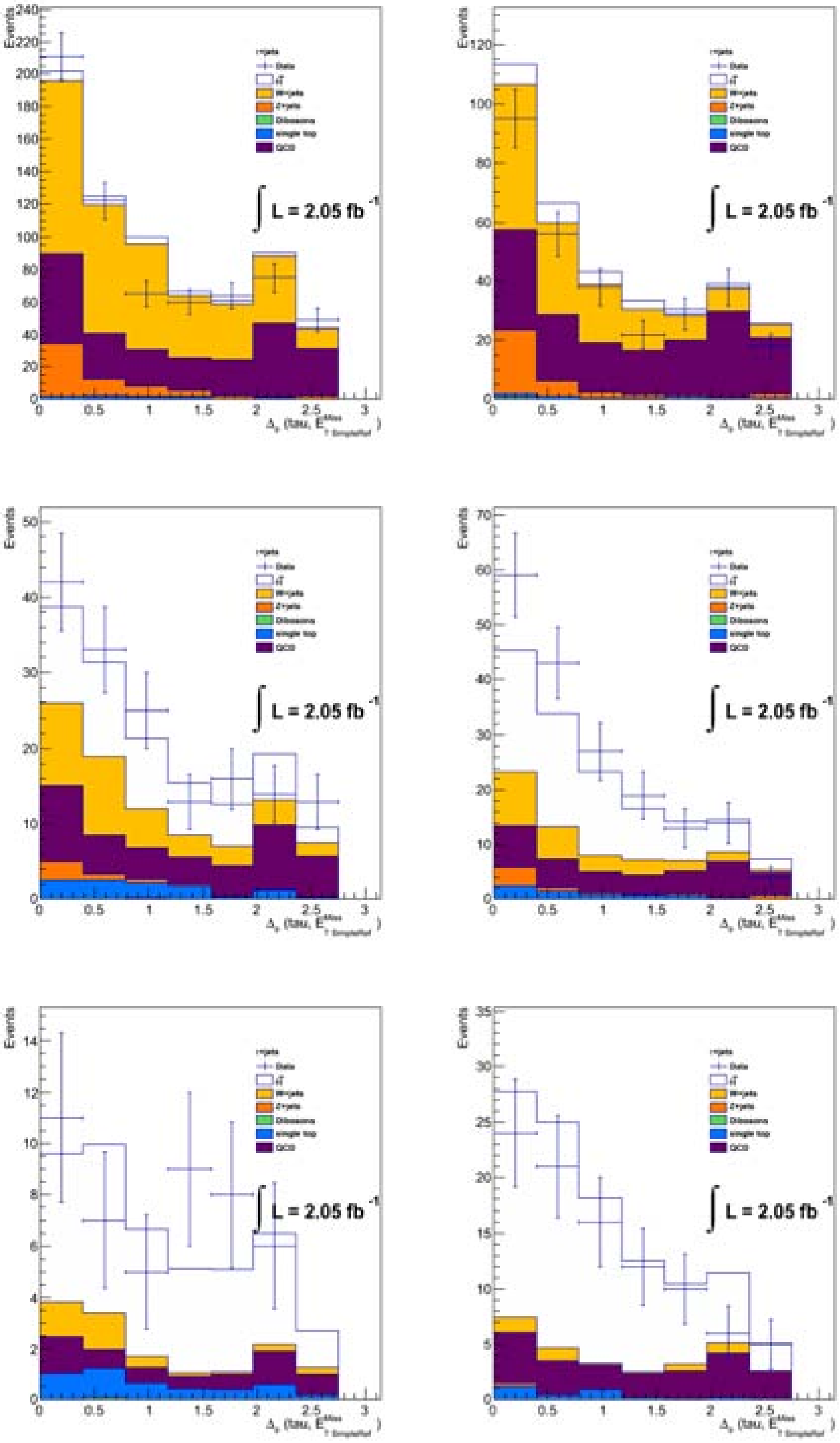


Figure 6.10: Data-MC comparison: $\Delta\phi$ (τ , $E_{\mathrm{T}}^{\mathrm{miss}}$) in the 1-prong tau channel. A cut at 2.5 is applied.

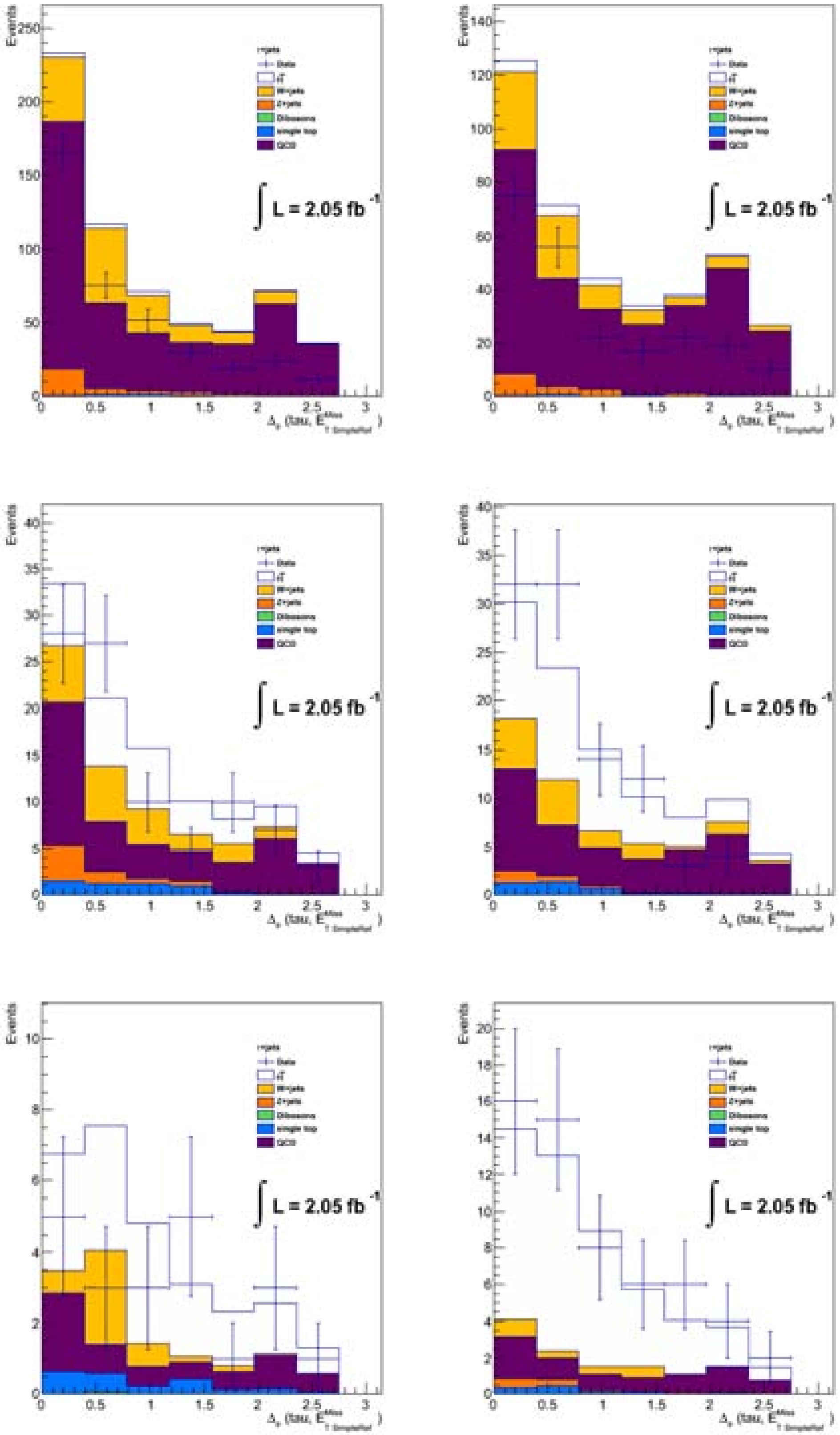


Figure 6.11: Data-MC comparison: $\Delta\phi$ (τ , $E_{\mathrm{T}}^{\mathrm{miss}}$) in the 3-prong tau channel. A cut at 2.5 is applied.

- We produce then for this sample of data a map of the efficiency for a jet to be tagged as a function of the jet η and jet p_T. A weight is then applied to each event. The weight represents the average probability to have a certain number of jets tagged. Such weight is calculated as a function of the jet η and jet p_T (per-jet tagging rate), in the way shown in the Eq. 6.9. In this formula, TRF is the tagging efficiency in function on the η and p_T of the jet.

$$w_{trf} = 1 - \prod_{i}^{njets} (1 - TRF(\eta_i, p_{Ti})), \tag{6.9}$$

Combining the precision of MM and the improvement in the shapes due to applying the TRF method, our final QCD estimation is obtained using the TRF study normalized to the number of entries of the MM distributions.

TRF method for Monte Carlo samples

The TRF method is also useful in the case of other background processes with limited statistics of simulated samples. Therefore we calculated per event weight based on the jet MC tagging probability and applied them to the untagged distribution of each MC simulated sample, similarly to what we did for QCD. In this case, the jet tagging probability calculation is based on the MC truth level information on the nature of the quark that originate the jet. Each distribution is then rescaled to the number of events of the unweighted distribution.

6.4 Comparison between data and Monte Carlo prediction

Figures 6.5, 6.6 and 6.7 respectively for 1-prong, 3-prong and electron channel, show the W transverse mass distribution in data and Monte Carlo samples obtained applying the baseline selection and the background estimation tools described before. These distributions are the templates that we fitted to extract the $t\bar{t}$ cross-section.

The comparison between data and Monte Carlo in another discriminant variables are also shown in this Section. The `BDTEleScore` distribution is shown in Figures 6.8 and 6.9 for 1-prong and 3-prong respectively. The background contribution to `BDTEleScore` and `BDTJetScore` discriminant in the region selected by the event selection cuts give an idea of the composition of the selected sample. The `BDTEleScore` is cut at 0.5 to veto those electrons which otherwise fulfill the tau selection requirements. The discrimination power of this variable has more impact in the 1-prong tau channel where the contamination from electrons is more important.

Concerning the tau kinematics, the distribution of the angle between tau p_{T} and the $E_{\mathrm{T}}^{\mathrm{miss}}$ ($\Delta\phi(\tau, E_{\mathrm{T}}^{\mathrm{miss}})$) is shown in Figures 6.10 and 6.9, for 1-prong and 3-prong respectively. This variable plays an important role in this analysis, as has been shown in Section 6.2, because it is involved in the calculation of the W transverse mass. In addition, the $\Delta\phi(\tau, E_{\mathrm{T}}^{\mathrm{miss}})$ allows to visualize the QCD multi-jets contamination, which fact convinced us to optimize the selection of signal events applying a cut on this variable. In all the cases shown, the data are in reasonably agreement with the sum of simulated signal and background.

Analogous comparison plots are reported in Appendix B. The `BDTJetScore` is again a discriminating variable: the cut at 0.6 accepts tau-like object and rejects jet-like objects; for this reason it is more significative in the 3-prong channel, where the main source of fake tau is represented by QCD multi-jets. The kinematics of the tau lepton is described in data and Monte Carlo by the distribution of tau p_{T}, tau pseudorapidity and the distribution of the angle between the tau p_{T} and the leading jet p_{T} ($\Delta\phi(\tau$,leading jet $p_{\mathrm{T}})$). Other informations of the event topology are contained in the distribution of $E_{\mathrm{T}}^{\mathrm{miss}}$ and of the reconstructed mass of the hadronic top though the sum of the p_{T} of the three jets more energetic.

The TRF method provides in general smooth shapes although the lack of statistics also in Monte Carlo, makes the histograms with two identified b-jets not sufficiently populated.

6.5 Profile likelihood fit method

In order to extract the cross section we use a profile likelihood fit method. We perform a likelihood fit of the transverse mass of the $W(\tau\nu)$ decay in the selected data to a weighted sum of templates corresponding to the $t\bar{t}$ signal ($\mathcal{D}_{t\bar{t}}$) and the background of $W + jets$ ($\mathcal{D}_W$), QCD ($\mathcal{D}_{QCD}$), and the rest of physics backgrounds, single top, Z+jets, and diboson ($\mathcal{D}_{other}$), such that:

$$\mathcal{D}_{data} = k_{t\bar{t}} \times \mathcal{D}_{t\bar{t}} + k_W \times \mathcal{D}_W + k_{QCD} \times \mathcal{D}_{QCD} + \sum k_{other} \times \mathcal{D}_{other}. \tag{6.10}$$

The term $k_{t\bar{t}}$ is the factor that multiplies the $t\bar{t}$ nominal cross-section: a scaling factor of one would correspond to the nominal SM prediction 164.6 pb [54]. The other k factors provide the respective size of the various backgrounds. We fit simultaneously $k_{t\bar{t}}$ (the signal) and k_W (the most abundant background). The factor k_{QCD} is set such as to reproduce the amount of multi-jet events predicted by the data-driven method discussed in Section 6.3.1, while k_{other} is set to one such as to reproduce the Standard Model expectation for the other backgrounds. All template shapes are provided by the Monte Carlo simulation applying the TRF method, except for QCD template shapes that are extracted from the data (see Section 6.3.1). The latter are thus not affected by physics modeling or uncertainties related to reconstructed objects.

In general, for a continuous variable, the likelihood is defined as the product of the probability that each component assumes the assigned value: We use the MINUIT package [66] to perform the likelihood fit. In practice, the fit minimizes minus the logarithm of the likelihood in order to deal with sums rather than products making it easier to compute the derivatives used in the minimization process:

$$-2\ln L(k_{t\bar{t}}, k_W, \overrightarrow{\alpha}) \propto -2 \sum_{i=0}^{i=N_{bins}} n_i \ln(\mu_i) - \mu_i + \sum_{j=0}^{j=N_{syst}} \alpha_j^2, \tag{6.11}$$

with n_i are the observed number of events in each bin and μ_i the expected number with $\mu_i = \mu_i(k_{t\bar{t}}, k_W, \overrightarrow{\alpha})$. The sum runs over all bins of the set of fitted histograms (N_{bins}). We use a profile likelihood fit, which allows that systematic variations due to external factors are not simply used to test the bias of the fit due to these factors, but enter directly in the minimization process of the fit [67]. Compared to a standard likelihood function, the profile likelihood adds to the fitting function a set of N_{syst} nuisance parameters α_j, which control the size of the corresponding systematic uncertainties. Each nuisance parameter is assumed to be Gaussian-distributed around the nominal value of zero and with width one. A fitted value of ± 1 corresponds to the $\pm 1\sigma$ variation given in the input for the corresponding systematic uncertainty. The addition of nuisance parameters to the fit allows to effectively check and constrain *in situ* the systematic uncertainties using the data themselves. The uncertainty on the fitted values of the nuisance parameters ($\Delta\alpha$) defines the 68% confidence level (CL) range in which the variation of the systematic uncertainty is compatible with data. In this approach all nuisance parameters are fitted together and the total uncertainty is determined taking into account the correlation between them.

6.6 Treatment of the systematic uncertainties

To extract the $t\bar{t}$ cross-section in the channel with $\tau + jets$ in the final state, we fit the $W(\tau\nu)$ decay transverse mass distributions in the 1-prong and 3-prong decay modes for the various jet and b-tagged jet multiplicity shown in Figures 6.5, 6.6. We fit simultaneously the $W(e\nu)$ decay transverse mass of the auxiliary electron sample shown in 6.7. In the fit, we will assume that the normalization of the $W + jets$ background process, and the rest of physics backgrounds, single top, Z+jets, and diboson, is common to the τ and the electron final state, e.g. that they are described by the SM. In the case of the $t\bar{t}$ sample, we will not constrain *a priori* the relative τ and the electron final states normalization. On the other hand, in all cases we will assume the same physics modeling of the $t\bar{t}$ production process.

In this section we describe the systematic uncertainties affecting the measured cross section and their implementation in the fit. There are various categories of systematic uncertainties: the ones affecting the reconstructed objects, the ones affecting the modeling of the signal and background physics processes and the ones related to the data-driven QCD background. All systematic uncertainties are included as nuisance parameters in the fit, except the ones predicted to be small (at the percent level) that are treated separately.

The systematics uncertainties affecting jets and $E_{\mathrm{T}}^{\mathrm{miss}}$ are listed below:

- **Jet energy scale.** To estimate the impact of the Jet Energy Scale (JES) uncertainty, we use the `JetUncertainties-00-03-04-01` version of the JetProvider package that provides the individual sources for JES uncertainty as explained in [38]. We use the envelope jet energy scale systematic p_{T} and η dependent. The JES systematic uncertainty is of the order of 4% for jets with p_{T} of 20 GeV and 2.5% at 100 GeV. After applying the factor for the *up* variation or *down* variation, the selection cuts and the full analysis are applied. An additional JES uncertainty is applied to b-jets. The ATLAS Jet-ETMiss group provides a set of scale factors to be applied according to jet p_{T} to all true b-jets in Monte Carlo. The value goes from 2.5% to 0.7%, for jets with p_{T} between 20 GeV and > 600 GeV. The uncertainty is then added in quadrature to the JES uncertainty for inclusive jets [60].

- **Jet energy resolution.** The jet resolution (JER) is calculated with the JERUncertaintyProvider tool, which provides the smearing factor to be applied to the nominal samples that take into account the difference in resolution between data and Monte Carlo [69]. Such factor is the quadratic difference between the jet resolution in data and MC, where the expected fractional jet p_{T} resolution is given as a function of its p_{T} and η. Only an *up* variation is defined for this uncertainty, because the jet smearing cannot be under-estimated in MC.

- **Jet reconstruction efficiency.** The efficiency of reconstruction for a jet in the calorimeter with respect to the tracked jets, presents differences between data and MC. This efficiency is estimated using the tag-and-probe method in QCD di-jet events. A sample where the 0.4% of jets with p_{T} under 35 GeV are randomly rejected is used. The jet reconstruction uncertainty is one-side; the *down* variation is obtaining through a symmetrization of the *up* variation.

- **BC-tagging and light-tagging.** It is very important to estimate the contribution of the uncertainty of heavy flavor tagging of jets, as the signal purity depends critically on the number of b-tagged jets. For the JetFitterCombNN tagger used at the working point corresponding to a 70% of b-tagging efficiency, the recommended scale factors ranges between 0.94 and 0.99 depending on the jet p_{T} [61]. This uncertainty is considered 100% correlated between for b- and c- quark jets. The uncertainty on the light-quark tagging probability is provided separately. It is not included as nuisance parameter in the fit since its overall contribution is less than 1%.

- **Missing transverse energy.** This uncertainty is calculated using the TopMETTool that takes into account the contribution of the pile-up, soft-jet and cell-out terms. These are additional contributions to the ones coming from the JES uncertainty of all jets in the vent. The uncertainty due to the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation uncertainty has also been removed from the fit since its contribution is less than 1%.

- **LAr hole.** Part of the data collected in 2011 suffers of a leakage in the acceptance region due to an hardware problem related with some FEBs (front end boards) of the Liquid Argon calorimeter. For the data interested by this issue (just the runs between 180614 and 185352) have been removed the events if a jet is found near that region, requiring a jet $p_{\mathrm{T}}>$ 20 GeV. In data, the lost of energy due to the dead Front-End-Boards is taken into account in the determination of the p_{T}, while the jet p_{T} in the MC is not affected. The event removal is performed through the MET_Cleaning_Utils tool, while the systematic uncertainty is obtained running the tool varying the jet p_{T} threshold by ± 4 GeV. After the evaluation of the contribution of this systematic uncertainty, it has been removed from the fit as it was less than 1%.

The systematic uncertainties related specifically to the τ and the electron are the following:

- **Tau identification and trigger scale factors.** Data-driven measurements of the tau identification efficiency have been performed using $Z \to \tau\tau$ and $W \to \tau\nu$ tag and probe analysis. The uncertainty related to the `tau_SafeLlhTight` ID working point and a tau p_{T} greater that 30 GeV is 9.3%. For 1-prong only samples, the calculated uncertainty is 7.8% [51]. In general, the relative uncertainties are in the range 7-12% in the majority of p_{T} range. At this level of precision no significant discrepancies between the efficiencies in data and Monte Carlo were found [35], so that no scale factor has been applied for the nominal distributions. For what concerns the trigger efficiency, we applied a global scale factor of 0.76 as measured in a dilepton $t\bar{t}$ sample (as described in Chapter 5) with a 4% associated uncertainty. We use the value obtained for events with 40 GeV $\leq p_{\mathrm{T}} <$ 70 GeV and 65 GeV $\leq E_{\mathrm{T}}^{\mathrm{miss}} <$ 100 GeV, the interval of p_{T} and $E_{\mathrm{T}}^{\mathrm{miss}}$ where most of our events are included (see table 5.1). Complementary studies of the trigger efficiencies obtained with the electron semi-leptonic sample and described in Appendix B.1 did not show evidence of dependence of the scale factor on the value of $E_{\mathrm{T}}^{\mathrm{miss}}$.

We thus combined both ID and trigger sources and assign an uncertainty of 20% on the scale factor to be profiled by the fit. This *a priori* uncertainty covers the statistical and systematic uncertainties of the measurement and the value of the trigger scale factor measured in the higher $E_{\mathrm{T}}^{\mathrm{miss}}$ bin.

- **Tau energy scale.** Because 1-prong and 3-prong tau decay each consist of a specific mix of charged and neutral pions, the tau energy scale calibration differs in the final step from the general hadronic calibration, as described in Chapter **??**. Tau energy scale uncertainties are provided as a function of the tau p_{T}, the tau η and the number of tracks, irrespectively of the working point used for the identification [51]. For taus with $p_{\mathrm{T}} >$ 35 GeV as we used in the analysis, the uncertainty in the 1-prong case is 3.5%, 5% and 4.5% for η ranges: $|\eta| < 1.3$, $1.3 < |\eta| < 1.6$ and $|\eta| > 1.6$ respectively. In the multi-prong case instead, an uncertainty of 5.5% is applied to multi-prong taus with p_{T} between 30 and 40 GeV (for all η values), while the uncertainty is 4.5% (5%) in η region $|\eta| < 1.3$ ($|\eta| > 1.3$) for tau $p_{\mathrm{T}} >$ 40 GeV.

- **Electron reconstruction, identification and trigger.** A scale factor is applied to the electron in the MC simulation to correct the difference between data and Monte Carlo due to the reconstruction, the identification and the trigger efficiency. The quadratic sum of the statistical and the systematic uncertainties on the scale factor is taken as systematic uncertainty and is of the order of 2.5%. This scale factor is considered as normalization factor only with an associated uncertainty.

The background processes estimated with Monte Carlo samples are affected by the uncertainties of the simulation model. In general it could affect both the shape and the normalization of the sample. Here below the different backgrounds and the treatment of the systematic uncertainties related with them are listed:

- **Z+jets, single top and diboson background normalization.** We did the same also for the Z+jets, single top and diboson MC samples: varying their normalization by, respectively, 100%, 10% and 5% according to [52]. We consider the uncertainty of each jet bin fully correlated with the others, but we consider uncorrelated the uncertainties between different channels (1-prong tau, 3-prong tau and electron channel). The uncertainty on the shape is not taken into account because of the very few statistics of these backgrounds.

- ***W*+jets modeling.** On one hand the systematic uncertainty associated to the Berends scaling (see Section 6.3.1) is the spread around the central value of the ratio between $W + n\ jets$ events and $W + (n+1)jets$, which amounts to 24%. The uncertainties related to the estimation are detailed in Section 6.3.1 as well. Moreover another uncorrelated uncertainty of 25% is added to take into account the error in the extrapolation of the scale factors from the 2 jets bin to the higher multiplicity bins. The uncertainty on $W + b\bar{b}$, $W + c\bar{c}$ are treated as fully correlated between them, but uncorrelated with $W + c$ fraction.

The $t\bar{t}$ signal sample used in this analysis has been generated with MC@NLO generator model, which differs from other models because they account for different phase space of events, even if some regions of the phase space overlap. The systematic uncertainty related to the $t\bar{t}$ modeling is treated as nuisance parameter as well as the other uncertainties. The validity of this method has been demonstrated [70] under the these conditions:

- the analysis exploit jet-multiplicity dependence: our analysis strategy is based on the classification of the signal between events with 3 jet and at least 4 jet;
- the shape of the observables doesn t depend on the $t\bar{t}$ model used;
- consider enough number of models to extrapolate the acceptance in the full phase space: we used 5 different $t\bar{t}$ models, the same that have been used to probe the validation of the method.

In particular the following models have been considered:

AcerMC ISR and FSR. ISR and FRS samples produced with AcerMC generator are compared to MC@NLO to take into account the effect of the initial and final state radiation. The variations with respect to MC@NLO are parametrized with nuisance parameters.

POWHEG and ALPGEN. To estimate the uncertainty associated to the signal generator MC@NLO, the latter is compared to two other samples generated with POWHEG and ALPGEN, interfaced with HERWIG/JIMMY to simulate the hadronization. The variations with respect to MC@NLO are parametrized with nuisance parameters.

POWHEG+HERWIG vs POWHEG+PYTHIA. To estimated the error on the fragmentation calculation, we estimate the relative difference between the hadronization simulation provided firstly by HERWIG/JIMMY and then by PYTHIA, both in association with POWHEG events generator. Such difference is then applied to the MC@NLO sample and the variation with respect to the nominal MC@NLO is parametrized with a nuisance parameter.

A systematic error of 1.4% associated to this method of profile the $t\bar{t}$ modeling has been estimated [70], which has to be summed in quadrature after the fit.

An uncertainty on the parton distribution functions is also added out the profile. Following the recommendation of the PDF4LHC working group to evaluate the PDFs uncertainty, the estimated value of 1.7%, obtained in the 700 pb^{-1} $t\bar{t}$ semi-leptonic analysis with electrons and muons [68] is summed in quadrature separately.

The QCD multijet background is extracted via a data-driven method. Uncertainties affecting the normalization, as well as the shapes are taken into account, as described below:

- **Multi-jet background normalization.** The normalization of the QCD multi-jet background is fixed into the fit to the value obtained by the Matrix Method (see Section 6.3.1). In order to let the normalization to adjust itself accordingly to data we let the normalization vary by its uncertainty that is 50 % for the QCD multi-jet background in the untagged case and 100 % when at least one jet is tagged as b-jet.

- **Multi-jet background shape.** An uncertainty is associated to the determination of QCD background shape with the TRF method as described in Section 6.3.1. In order to take into account such uncertainty, another QCD template with different shape is passed to the fit. This varied template is derived from the shape of the QCD sample obtained with the Matrix Method in the inclusive pre-tagged selection (mostly 3 or more jets), which does not suffer the lack of statistics. Actually this template is built by scaling bin-per-bin the QCD distributions obtained with the *loose-not-tight* event selection used in the TRF method in each jet multiplicity and b-tagging categories. The weight is calculated as the ratio of the Matrix Method distribution and the TRF distribution both in a 1 jet inclusive pre-tagged sample. The shape variation between the varied template and the nominal QCD template, shown in Appendix [**?**], is taken as *a priori* uncertainty.

Finally, a nuisance parameter (called *k_tt_e* in the following) is assigned as scale factor that takes into account the $t\bar{t}$ normalization in the electron channel. A nominal uncertainty of 100% is adopted, in such a way the normalization of the $t\bar{t}$ sample in the electron channel becomes uncorrelated with respect to that of 1-prong and 3-prong tau channel. The value of the *k_tt_e* nuisance parameter obtained with the profile fit represents the factor to be multiplied to the *k_ttbar* and to the nominal cross-section in order to obtain the $t\bar{t} \rightarrow e+jets$ cross-section.

The effect of the main systematic uncertainties can be seen in Appendix E: the up and down variations related to some of the systematic uncertainties are shown with respect to the nominal templates as an example.

In the next Chapter the results of the fit and the measurement of the $t\bar{t}$ cross-section are discussed.

Chapter 7

Results and Outlook

In this chapter, we will describe the result of the cross-section measurement, present a series of validation tests, compare our results with previous measurements and finally evaluate the prospects for a future measurement based on a similar strategy but with a higher statistics sample.

7.1 Results of the profile likelihood fit

As described in the previous chapter, we fit the data for the fraction of $t\bar{t}$ and W+jets events, and the magnitude of all the nuisance parameters. Table 7.1 gives the results of the fit. The distributions rescaled and reshaped according to the result of the combined fit of the three channels (1-prong tau, 3-prong tau and electron) are shown in Figures 7.1, 7.2 and 7.3, respectively. In general, good agreement is observed.

The data and Monte Carlo yields calculated after the combined fit are reported in tables 7.2, 7.3 and 7.4, for 1-prong, 3-prong and electron channel respectively.

We can see in Table 7.1 that the $t\bar{t}$ cross-section in the channel with $\tau\ + jets$ in the final state (k_{ttbar}) is measured with a total relative error of the order of 16%. The ratio of $t\bar{t}$ cross-section in the electron and the $\tau\ + jets$ channels (k_tt_e) is constrained slightly more, at the level of 12%, as some of the nuisance parameters like the luminosity do not affect the ratio. The W cross-section (k_{Wjets}) is measured more precisely ($\sim$ 6%) as it relies mainly on the information from the electron channel. The global τ ID and τ+$E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency ($TauIdSF$) is constrained to 40% of the pre-assigned uncertainty of 20%. The contraining power of the data comes from the measured ratio of the $W \to \tau$ to $W \to e$ that is constrained by the ratio of W branching fractions. This also helps constraining the Tau Energy Scale. The QCD background normalizations and shapes in the two τ channels are both significantly constrained with respect to the large *a priori* assigned uncertainties. On the other hand, jet related uncertainties and b-tagging uncertainties are constrained mainly by the electron channel information, as the electron channel is statistically much more powerful. The same is true for the $W + jets$ and the $t\bar{t}$ modeling uncertainties,

k_{ttbar}	1.245	0.212/-0.183
k_{Wjets}	1.064	0.063/-0.060
Systematic	α	$\Delta\alpha$
k_tt_e	-0.137	0.129/-0.112
Lumi	-0.025	1.005/-0.988
QCDnorm_corr_tau_1p	-0.666	0.303/-0.295
QCDnorm_corr_tau_btag_1p	0.014	0.473/-0.448
QCDnorm_corr_tau_3p	-1.240	0.098/-0.094
QCDnorm_corr_tau_btag_3p	0.278	0.513/-0.497
QCDnorm_corr_ele	-0.943	0.906/-0.885
QCDnorm_corr_ele_btag	-0.001	1.000/-1.000
TES	0.295	0.304/-0.332
TauIdSF	-0.348	0.444/-0.435
JES	0.195	0.270/-0.285
bJES	-0.790	0.904/-0.611
JER	0.253	0.193/-0.196
JetRecoEff	-0.008	0.851/-0.898
BCtag	0.302	0.186/-0.173
Wjets_Berends	-0.552	0.244/-0.259
Wjets_HFQQ	-0.167	0.325/-0.347
Wjets_HFQQ_3ex	-0.209	0.485/-0.478
Wjets_HFQQ_4in	0.280	0.681/-0.736
Wjets_HFC	-0.421	0.521/-0.454
Wjets_HFC_3ex	-0.286	0.552/-0.503
Wjets_HFC_4in	-0.156	0.859/-0.770
ttbar_ISR	-0.002	0.134/-0.130
ttbar_FSR	0.125	0.114/-0.121
ttbar_Frag	0.017	0.194/-0.202
ttbar_NLO_1	-0.145	0.182/-0.196
ttbar_NLO_2	0.097	0.134/-0.130
Xsect_singleTop	0.068	0.989/-0.991
Xsect_Zjets	0.789	0.196/-0.192
Xsect_DB	-0.026	1.000/-1.000
ElectronFactors	0.034	0.991/-0.993
qcdShape1p	0.291	0.248/-0.247
qcdShape3p	-0.791	0.235/-0.209

Table 7.1: Results of the fit to the 1-prong tau, 3-prong tau and electron distributions. The column α gives the fitted value of the nuisance parameters, while $\Delta\alpha$ provides the corresponding errors expressed in units of the a-priori 1 σ uncertainty.

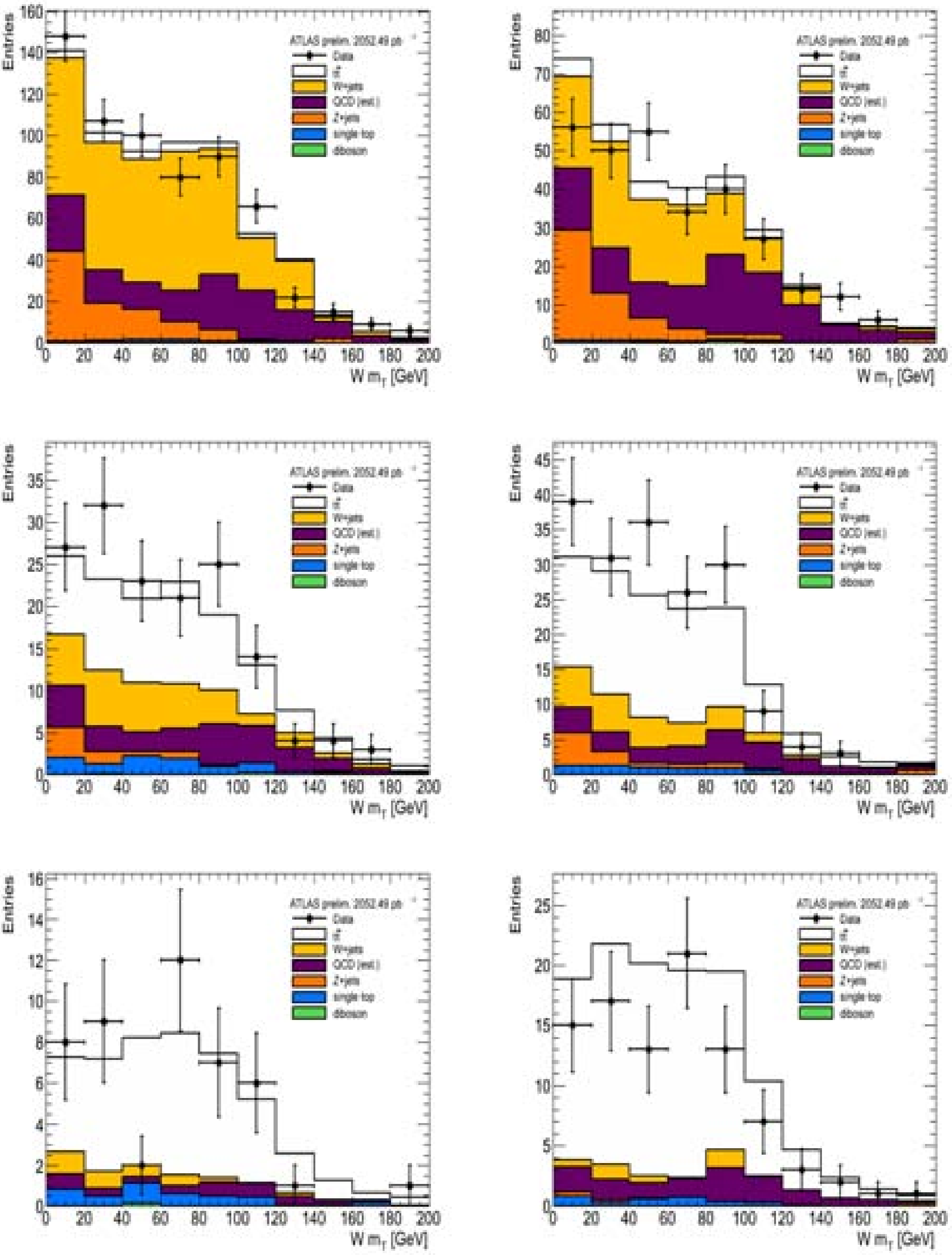


Figure 7.1: W transverse mass distribution after the combined fit in the 0- (top), 1- (medium) and ≥ 2- (bottom) b-jet multiplicity bins in the tau 1-prong channel.

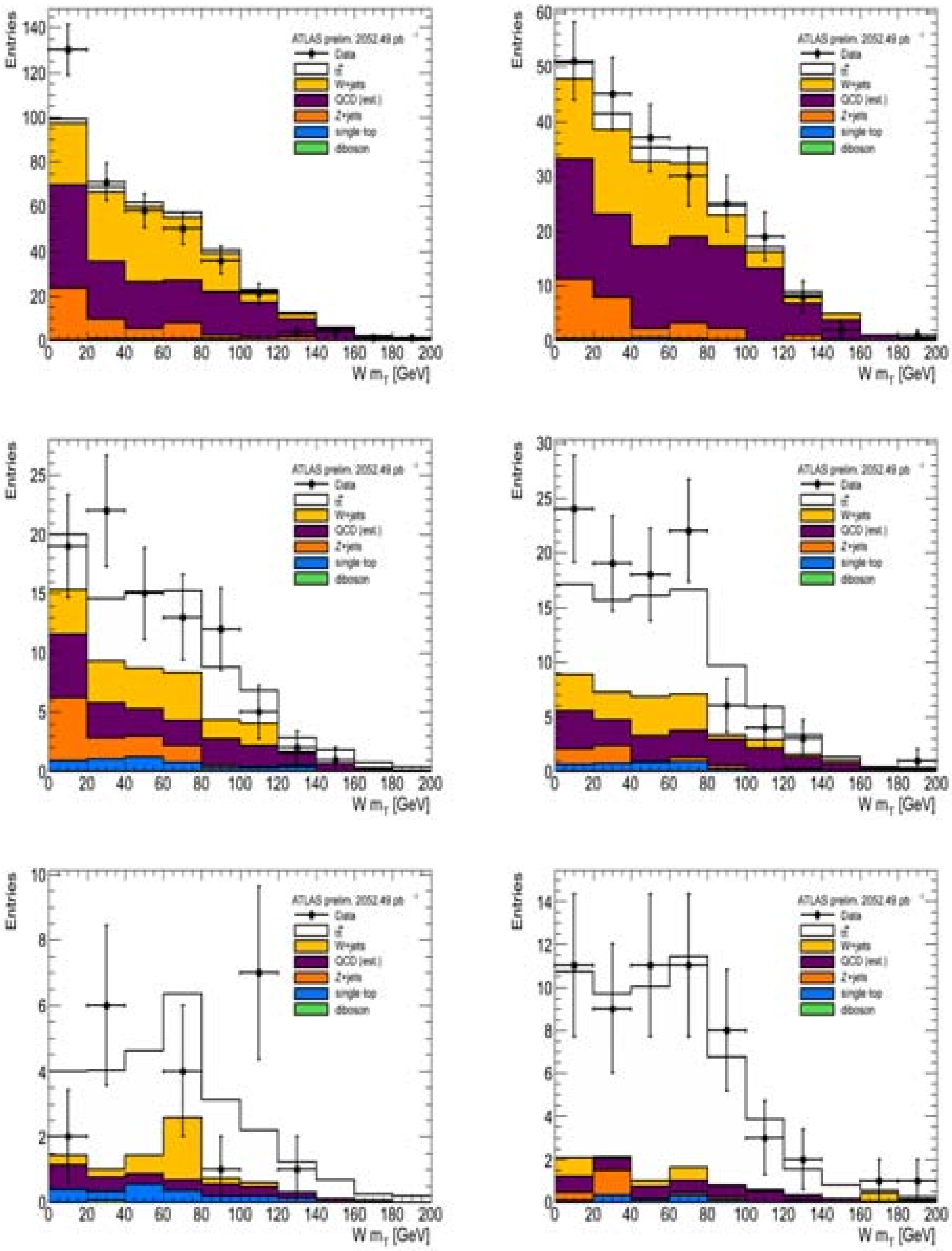


Figure 7.2: W transverse mass distribution after the combined fit in the 0- (top), 1- (medium) and $\geq$ 2- (bottom) b-jet multiplicity bins in the tau 3-prong channel.

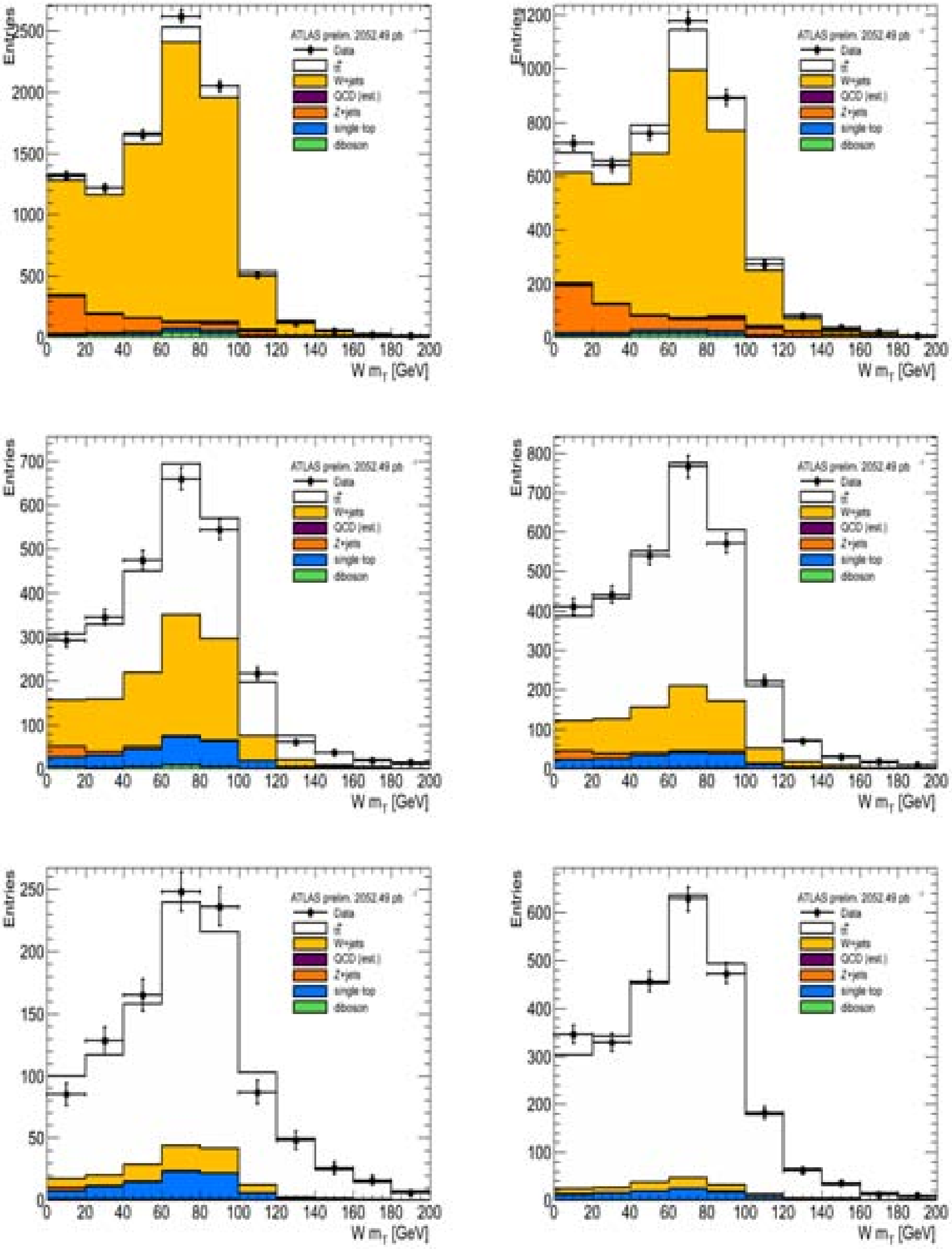


Figure 7.3: W transverse mass distribution after the combined fit in the 0- (top), 1- (medium) and $\geq$ 2- (bottom) b-jet multiplicity bins in the electron channel.

0btagex	3 jets	$\geq$ 4 jets
$t\bar{t}$	24.0± 5.0	25.8± 5.6
Wjets	369.2± 45.5	125.9± 18.1
QCD	148.7± 33.7	103.4± 23.3
Zjets	94.4± 13.9	54.5± 7.9
Single Top	5.0± 0.9	3.4± 0.6
Diboson	2.3± 0.2	1.0± 0.1
Prediction	643.5± 58.5	314.1± 31.0
Data	643.0	294.0

1btagex	3 jets	$\geq$ 4 jets
$t\bar{t}$	62.7± 12.8	93.6± 18.8
Wjets	32.3± 6.0	24.7± 4.8
QCD	27.7± 13.1	23.6± 10.8
Zjets	6.9± 1.0	9.8± 1.4
Single Top	10.0± 1.4	6.1± 0.9
Diboson	0.3± 0.0	0.2± 0.0
Prediction	139.9± 19.3	158.1± 22.3
Data	153.0	178.0

2btagin	3 jets	$\geq$ 4 jets
$t\bar{t}$	37.0± 7.4	97.2± 19.7
Wjets	3.5± 0.9	4.1± 1.2
QCD	4.0± 1.9	14.0± 6.3
Zjets	0.0± 0.0	0.9± 0.1
Single Top	4.1± 0.6	3.2± 0.5
Diboson	0.1± 0.0	0.0± 0.0
Prediction	48.7± 7.7	119.5± 20.7
Data	46.0	93.0

Table 7.2: Tau 1 prong post fit yields.

0btagex	3 jets	≥ 4 jets
$t\bar{t}$	12.3± 2.6	14.4± 3.4
Wjets	145.2± 18.2	70.1± 10.7
QCD	162.1± 20.9	106.7± 13.4
Zjets	47.8± 6.8	26.5± 4.0
Single Top	3.3± 0.5	1.5± 0.2
Diboson	0.6± 0.1	0.1± 0.1
Prediction	371.3± 28.6	219.3± 17.9
Data	376.0	218.0

1btagex	3 jets	≥ 4 jets
$t\bar{t}$	33.7± 7.1	46.7± 9.2
Wjets	18.5± 3.9	14.6± 3.0
QCD	18.6± 6.9	17.1± 6.6
Zjets	10.8± 1.6	4.1± 0.6
Single Top	4.9± 0.7	3.5± 0.5
Diboson	0.0± 0.0	0.0± 0.0
Prediction	86.5± 10.8	86.1± 11.7
Data	89.0	97.0

2btagin	3 jets	≥ 4 jets
$t\bar{t}$	18.3± 3.7	46.2± 9.3
Wjets	3.4± 1.2	2.5± 0.7
QCD	2.8± 1.0	3.8± 1.5
Zjets	0.0± 0.0	1.8± 0.3
Single Top	2.1± 0.3	1.1± 0.1
Diboson	0.0± 0.0	0.0± 0.0
Prediction	26.6± 4.0	55.5± 9.4
Data	21.0	57.0

Table 7.3: Tau 3 prong post fit yields.

0btagex	3 jets	$\geq$ 4 jets
$t\bar{t}$	486.3± 117.8	602.2± 141.4
Wjets	8009.8± 620.2	3338.5± 442.3
QCD	74.5± 63.3	57.2± 47.3
Zjets	736.0± 84.3	477.1± 54.6
Single Top	112.0± 13.0	63.9± 7.8
Diboson	146.4± 10.2	54.1± 5.3
Prediction	9564.9± 640.2	4593.0± 470.0
Data	9566.0	4605.0

1btagex	3 jets	$\geq$ 4 jets
$t\bar{t}$	1394.2± 311.6	2220.7± 496.4
Wjets	979.1± 138.5	629.4± 115.6
QCD	0.1± 0.2	0.1± 0.1
Zjets	46.3± 5.4	57.9± 6.6
Single Top	240.6± 25.9	168.6± 19.4
Diboson	23.5± 1.7	9.4± 1.0
Prediction	2684.0± 342.0	3086.0± 510.1
Data	2660.0	3073.0

2btagin	3 jets	$\geq$ 4 jets
$t\bar{t}$	858.7± 191.8	2335.4± 524.7
Wjets	77.3± 16.2	88.7± 23.2
QCD	0.1± 0.1	0.1± 0.2
Zjets	5.2± 0.6	1.9± 0.2
Single Top	82.8± 9.6	96.3± 10.5
Diboson	2.8± 0.2	0.7± 0.1
Prediction	1027.0± 192.7	2523.0± 525.3
Data	1045.0	2532.0

Table 7.4: Electron post fit yields.

The nuisance parameters of the systematic uncertainties are all fitted together in the minimization process with the $t\bar{t}$ and $W + jets$ normalizations, taking into account the correlation among them. The matrix of correlations resulting from the fit is graphically shown in Figure 7.4. The parameters are mostly uncorrelated except for a few cases. Here we discuss a few examples. The parameter k_{ttbar} normalizing the $t\bar{t}$ cross-section is anti-correlated with $QCDnorm\_corr\_tau\_btag\_1p$, the normalization of the QCD background in the 1-prong tau decay channel with a b-tagged jet, where most of the signal is and QCD is an important background. The parameter k_{ttbar} is also anti-correlated with $TauIDSF$, the normalization of the τ ID and τ+$E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency as the latter enters directly in the signal acceptance calculation. The parameter $TauIDSF$ is in turn anti-correlated with $QCDnorm\_corr\_tau\_1p$ and $QCDnorm\_corr\_tau\_3p$, as $W + jets$ and QCD background are competing contributors to the category of events without b-tagged jet. The higher the share of $W + jets$, the higher will be the efficiency for tau identification and trigger extracted from the simultaneous fit with the electron channel.

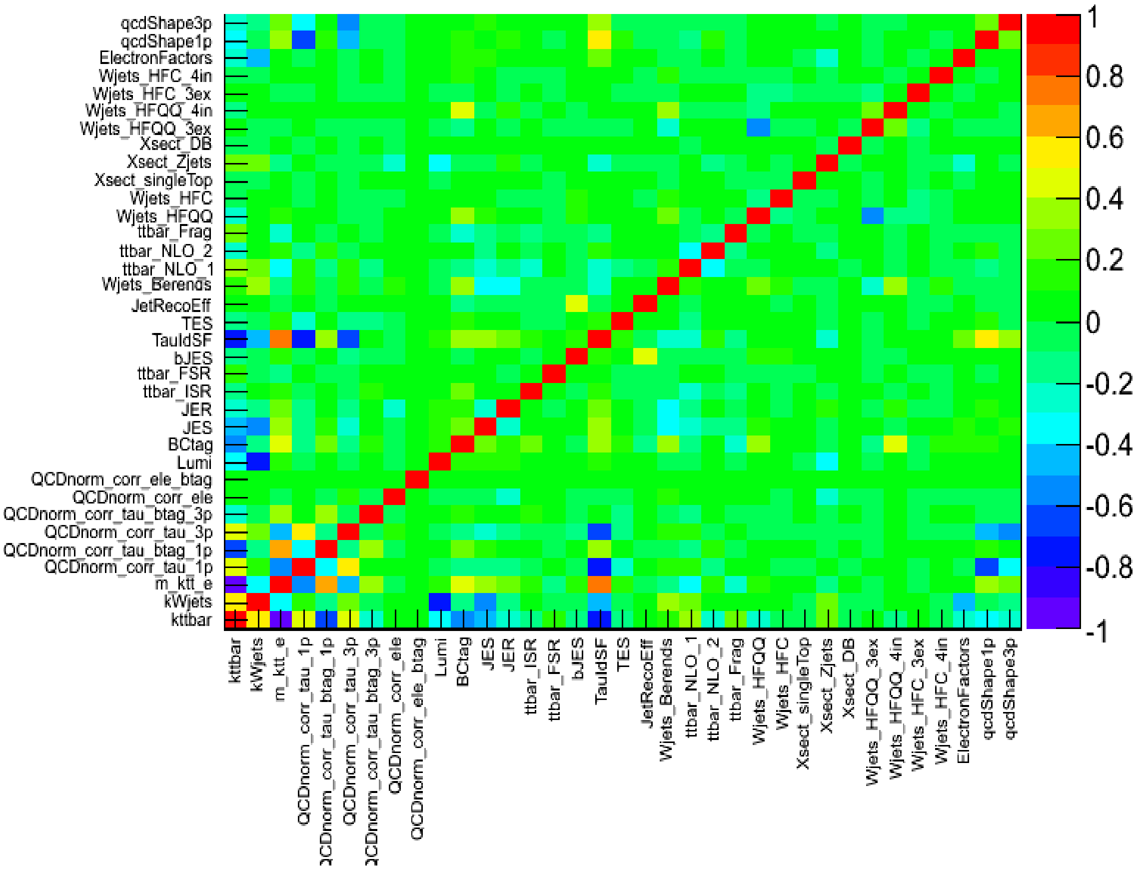


Figure 7.4: Correlations of the fitted nuisance parameters and k-factors.

To test that the fit is well-behaved, we perform the fit on a series of pseudo-experiments generated by varying the content of expected content of each bin predicted according to

Poisson statistics. The result of 100 pseudo-experiments is shown in Table 7.5. We see that the measured errors are in good agreement with the expectation.

As mentioned above, the profile likelihood provides the total uncertainty on the fitted quantities, but the contribution of each single uncertainty is not well defined because the fit takes into account the correlation among the nuisance parameters in the minimization process. We classified the systematic uncertainties into categories and then we estimated the contribution to the cross-section of each category. To do that, we iterate the fit with pseudo-experiments, removing one group at the time and calculating again the errors on the $k_{t\bar{t}}$ as the quadratic difference with the nominal value obtained with all nuisance parameters together. The result of the fit with pseudo-experiments, together with the estimated contribution of each category of systematic uncertainties, are shown in Table 7.6. The statistical error is 7%. It is estimated by removing all nuisance parameters from the fit, except for the ratio of $t\bar{t}$ cross-section in the electron and the $\tau + jets$ channels (k_tt_e). The largest contribution to the systematic error comes from the uncertainty on the τ ID and τ+$E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency, and the tau energy scale. This uncertainty is partially constrained via the W+jets process, but the power of that constrain is limited by the statistics of the sample and the QCD background contribution. The second largest contribution comes from QCD, with the normalization contributing most. Jet related and BC-tagging uncertainties contribute 4% each, while signal and W+jets modeling contribute another 5% each. These uncertainties have been significantly constrained by the fit to the electron sample, and their impact has thus been reduced in a relevant way.

We have also fitted the electron sample only. The result is shown in Table 7.7 . the results are compatible with the ones obtained for the full fit for all the relevant parameters. The constraining power of jet related and BC-tagging uncertainties, $t\bar{t}$ and W+jets modeling is slightly reduced compared to the full fit.

7.2 The $t\bar{t}\rightarrow \tau+jets$ cross-section and comparison with other results

The $t\bar{t}\rightarrow \tau + jets$ cross-section measured in this analysis is

$$\mathbf{205 \pm 11(stat) \pm 39(syst)pb}.$$

This result can be compared with other ATLAS measurements at $\sqrt{s}$ =7 TeV shown in Figure 7.5. In particular, we can compare this result to the other measurement of the $t\bar{t}\rightarrow \tau + jets$ cross-section at $\sqrt{s}$ =7 TeV with a sample of 1.7 fb^{-1}that results in a value of

$$\mathbf{194 \pm 18(stat) \pm 46(syst)pb}$$

[73] . In that case, events with at least five jets are selected, first at trigger-level and then at analysis-level, where two of the jets are identified as having originated from b-quarks. After identifying the two jets likely to come from the hadronic decay of one of the top quarks, one of the remaining jets is selected as the hadronic tau candidate from the other top quark. The tau contribution is then separated from quark- or gluon initiated jets with a one-dimensional fit to the distribution of the number of tracks associated with the tau candidate. So, we see that the two measurements involve different triggers and a different

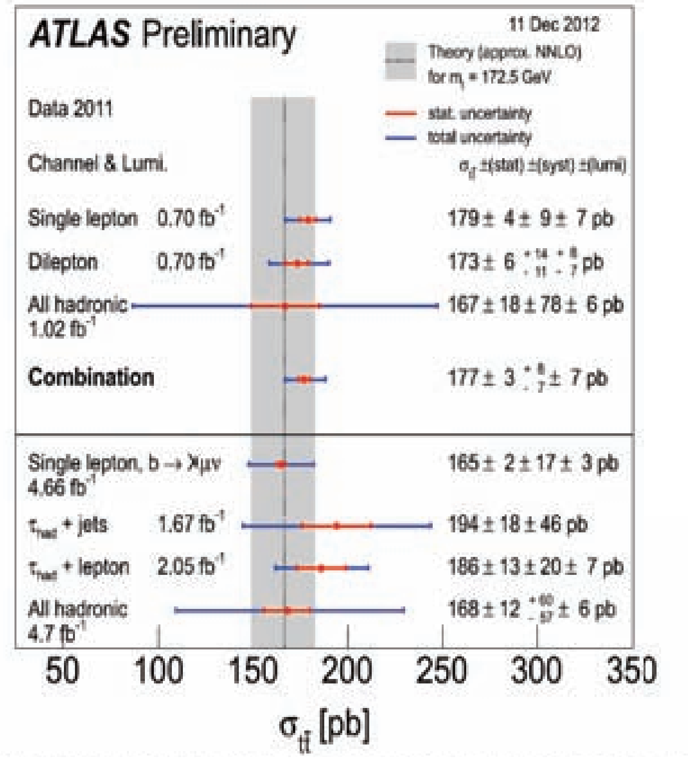


Figure 7.5: Summary of measurements of the top-pair production cross-section compared to the corresponding theoretical expectation based on an approximate NNLO calculation (Hathor 1.2). The lower part shows additional newer measurements not included in the combination.

way of identifying taus and are thus quite independent. The measured cross-sections are compatible. The total error of our measurement is about 30% smaller. We can compare also to the measurement of the $t\bar{t}\rightarrow \tau + lepton$ cross-section at $\sqrt{s}$ =7 TeV with a sample of 2 fb^{-1} [62] of

$$\mathbf{186 \pm 13(stat) \pm 21(syst) pb}$$

compatible with the results obtained in the semileptonic channel. That measurement relies on a different trigger, the single electron or muon trigger, but uses similar identification criteria of the tau hadronic decay as in this study.

All the measurements involving taus, although slightly higher, are compatible within errors with the combined result of ATLAS obtained with electrons and muons of

$$\mathbf{177 \pm 3(stat) \pm 10(syst) pb}$$

and the approximate NNLO calculation of HATHOR [19] with $\sigma_{t\bar{t}} = 167^{+17}_{-18}$ pb.

We have measured the $t\bar{t}$ cross-section in the semileptonic electron channel resulting in a value of

$$\mathbf{178 \pm 14(stat + syst) pb}.$$

in good agreement with the ATLAS combined measurement. We measure also directly the ratio of $t\bar{t}$ cross-section in the electron and the $\tau\ + jets$ channels. Some of the systematic uncertainties cancel in the ratio which is measured to be $0.86^{+0.13}_{-0.11}$ compatible with 1 as expected in the Standard Model.

7.3 Discussion and prospects for a higher statistics measurement

In this section, we discuss the prospects of performing the $t\bar{t}\rightarrow \tau + jets$ cross-section measurement with a significantly higher statistics sample like the $\sim$ 25 fb^{-1}of data at $\sqrt{s}$ =8 TeV.

As we have seen in the previous section, the largest systematic errors are related to the τ ID and τ+$E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiencies. With a high statistics data sample, the τ ID performance can be measured rather precisely with $Z \rightarrow \tau\tau$ decays. The precise measurement of the τ+$E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency is more difficult. The measurement that we have done with W+jets events would improve with more statistics but the precision would still be limited in comparison with the improved statistical precision of the measurement. In addition, the efficiency should preferably be measured in bins of p_{T} and $E_{\mathrm{T}}^{\mathrm{miss}}$. So one should use an additional sample, like $t\bar{t}$ decays in the $\mu + \tau$ and apply a tag-and-probe method to extract the efficiency for the trigger as reported in Section 5.4, [50]. This method could be extended to the electron channel to maximize the statistics.

As the statistical power of the data increases, one should ensure that the implementation of nuisance parameters is done with the required granularity of the sources of uncertainties to avoid overconstraining them. For example the JES and BC-tagging nuisance parameters should be divided in more components. With more statistics, one could also divide the sample in events with exactly 3 jets, exactly 4 jets or 5 or more jets. This would improve the sensitivity to the physics process modeling and jet related experimental uncertainties. A careful optimization of the cuts to minimize the contribution from QCD should be done. Yet another possible improvement consists in fitting more that one variable. In addition to the fit of the W transverse mass, we could fit for example the invariant mass of the candidate hadronic top decay. The latter brings a different separation power between the signal and the two main backgrounds: W+jets and QCD.

Ultimately, one can envisage to perform a simultaneous measurement of the semileptonic and dileptonic decays modes such has to have a fully consistent definition of the signal, treatment of the systematic errors and their correlation. The fit could be done under the hypothesis of the Standard Model, or including templates for the contribution of new physics like the existence of a Charged Higgs.

k_{ttbar}	1.020	0.194/-0.169
k_{Wjets}	1.005	0.065/-0.060

Systematic	α	$\Delta\alpha$
k_tt_e	0.014	0.181/-0.150
Lumi	-0.038	1.000/-0.998
QCDnorm_corr_tau_1p	-0.002	0.334/-0.329
QCDnorm_corr_tau_btag_1p	-0.019	0.344/-0.330
QCDnorm_corr_tau_3p	0.023	0.117/-0.115
QCDnorm_corr_tau_btag_3p	-0.007	0.240/-0.233
QCDnorm_corr_ele	-0.053	0.886/-0.885
QCDnorm_corr_ele_btag	-0.001	0.997/-0.997
TES	0.072	0.545/-0.709
TauIdSF	0.010	0.533/-0.522
JES	-0.021	0.317/-0.325
bJES	-0.058	0.866/-0.817
JER	-0.028	0.202/-0.205
JetRecoEff	0.123	0.809/-0.840
BCtag	0.001	0.241/-0.234
Wjets_Berends	0.007	0.261/-0.276
Wjets_HFQQ	-0.030	0.410/-0.424
Wjets_HFQQ_3ex	0.007	0.621/-0.664
Wjets_HFQQ_4in	-0.007	0.727/-0.711
Wjets_HFC	-0.030	0.664/-0.629
Wjets_HFC_3ex	-0.063	0.789/-0.701
Wjets_HFC_4in	0.003	0.849/-0.826
ttbar_ISR	0.081	0.253/-0.221
ttbar_FSR	0.031	0.186/-0.191
ttbar_Frag	0.033	0.229/-0.244
ttbar_NLO_1	-0.039	0.226/-0.237
ttbar_NLO_2	0.015	0.160/-0.160
Xsect_singleTop	0.041	0.989/-0.990
Xsect_Zjets	0.003	0.176/-0.172
Xsect_DB	-0.002	0.999/-0.998
ElectronFactors	-0.031	0.995/-0.992
qcdShape1p	-0.013	0.273/-0.270
qcdShape3p	-0.115	0.385/-0.380

Table 7.5: Fit results and constraints on the nuisance parameters obtained with pseudo-data.

Systematic contributions (100 pseudo-experiments)		
	$k_{t\bar{t}}$	
Fitted values	**1.0**	
Statistical Error (%)	**+7**	**-7**
Tau related uncertainties (%)	+12	-8
Jet related uncertainties (%)	+4	-4
BC tagging (%)	+4	-4
Electron SF (%)	<1	<1
Signal modeling (%)	+6	-6
QCD modeling (normal.) (%)	+11	-9
QCD modeling (shape) (%)	+6	-4
W+jets modeling (%)	+5	-5
MC background rate (%)	<1	<1
Luminosity (%)	+2	-3
Total Systematic (%)	**+18**	**-15**
PDF and other small errors(%)	+2	-2
Total Error (%)	**+19**	**-17**

Table 7.6: Fit result and relative contributions of systematic uncertainties obtained with 100 pseudo-experiments.

k_{ttbar}	1.080	0.086/-0.080
k_{Wjets}	1.104	0.064/-0.060
Systematic	α	$\Delta\alpha$
Lumi	-0.025	0.996/-0.996
QCDnorm_corr_ele	-0.845	0.924/-0.912
QCDnorm_corr_ele_btag	-0.002	1.000/-1.000
JES	0.157	0.297/-0.323
bJES	-1.112	1.164/-0.666
JER	0.177	0.233/-0.224
JetRecoEff	0.368	0.932/-0.917
BCtag	0.374	0.248/-0.238
Wjets_Berends	-0.376	0.271/-0.291
Wjets_HFQQ	-0.214	0.417/-0.436
Wjets_HFQQ_3ex	-0.308	0.605/-0.596
Wjets_HFQQ_4in	0.442	0.753/-0.889
Wjets_HFC	-0.514	0.699/-0.650
Wjets_HFC_3ex	-0.409	0.693/-0.628
Wjets_HFC_4in	0.075	0.948/-0.946
ttbar_ISR	-0.014	0.261/-0.244
ttbar_FSR	-0.216	0.266/-0.225
ttbar_Frag	0.182	0.303/-0.304
ttbar_NLO_1	-0.190	0.268/-0.262
ttbar_NLO_2	-0.103	0.199/-0.195
Xsect_singleTop	0.088	0.997/-0.991
Xsect_Zjets	0.670	0.215/-0.209
Xsect_DB	-0.027	1.000/-1.000
ElectronFactors	-0.018	1.000/-0.997

Table 7.7: Results of the fit to the electron distributions. The column α gives the fitted value of the nuisance parameters, while $\Delta\alpha$ provides the corresponding errors expressed in units of the a-priori 1 σ uncertainty.

Chapter 8

Conclusions

This thesis contributes to the area of top physics with the measurement of the $t\bar{t}$ production cross-section via strong interaction in pp collisions at the LHC. The measurement is done in the semileptonic decay channel with a hadronically decaying tau in the final state ($t\bar{t} \to \tau + jets$), which is challenging experimentally. This analysis is done using the 2.05 fb^{-1} of data collected by ATLAS in 2011 at 7 TeV center-of-mass energy. The cross-section is extracted with a profile likelihood fit of the transverse mass of the leptonically decaying W, combining the information of the 1-prong tau, 3-prong tau and electron channels. The cross-section in the tau and electron channel can vary independently. Systematic uncertainties are implemented as nuisance parameters in the fit and are constrained by the data improving the precision of the measurement. The result of the $t\bar{t} \to \tau + jets$ cross-section measurement is:

$$\mathbf{205 \pm 11(stat) \pm 39(syst)pb}.$$

It is compatible with the result of another measurement of the $t\bar{t}\to \tau + jets$ cross-section at $\sqrt{s}$ =7 TeV done with a sample of 1.7 fb^{-1} and based on a very different technique. Our measurement achieves a better precision with a relative error about 30% smaller. Both measurements are compatible with the Standard Model cross-section calculated at NNLO [19].

We measured also the cross-section in the $t\bar{t} \to e + jets$ channel, which resulted in the value of

$$\mathbf{178 \pm 14(stat + syst)pb}$$

in good agreement with the ATLAS combined measurement in the electron and muon channels. The ratio of the $t\bar{t}$ cross-section in the electron and the τ $+ jets$ channels is measured to be $0.86^{+0.13}_{-0.11}$ compatible with 1 as expected in the Standard Model.

As future perspective, we discussed possible improvements of the method developed in this thesis that could be achieved with a higher statistics sample like the 25fb^{-1} of 8 TeV data.

Bibliography

[1] J. Beringer et al., *Particle Data Group*, Phys. Rev. D86, 010001, 2012.

[2] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. 22, 579-588, 1961.

[3] A. Salam, N. Svartholm. ed. Elementary Particle Physics: Relativistic Groups and Analyticity. Eighth Nobel Symposium. Stockholm: Almquvist and Wiksell, 367, 1968.

[4] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. 19, 1264-1266, 1967.

[5] M. E. Peskin, D. V. Schroeder, *Quantum field theory*, Addison-Wesley, 1995.

[6] D. Green, *High P_T Physics at Hadron Colliders*, Cambdrige University Press, 2005.

[7] F. Halzen, A. D. Martin, *Quarks and Leptons*, John Wiley & Sons,1984.

[8] V. D. Barger, R. J. N. Phillips, *Collider Physics*, Updated Edition, Frontiers in Physics Edition,1996.

[9] R. K. Ellis, W. J. Stirling, B. R. Webber, *QCD and Collider Physics*, Cambdrige University Press, 2003.

[10] A. Pomarol, *Beyond the Standard Model*, arXiv:1202.1391v1 [hep-ph], February 7, 2012.

[11] A. Lleres, A. Lucotte, C. Monini, *Search for single-top production in the s-channel at* $\sqrt{s} = 7$ *TeV with the ATLAS detector*, ATLAS NOTE (Not reviewed, for internal circulation only), November 12, 2012.

[12] K. Becker *et al. Measurement of t-Channel Single Top-Quar1 k Production in pp Collisions at* $\sqrt{s} = 8$ *TeV using Neural Networks*, ATLAS NOTE (Not reviewed, for internal circulation only), September 3, 2012.

[13] The ATLAS Collaboration, *Search for the Wt production in the lepton+jets channel with ATLAS*, ATLAS NOTE (Not reviewed, for internal circulation only), March 29, 2012.

[14] A. Lleres and X. Sun, *Measurement of the Single Top-Quark Polarization with 4.7* fb^1, ATLAS NOTE (Not reviewed, for internal circulation only), September 17, 2012.

[15] The ATLAS Collaboration, *Top quark physics at ATLAS*, ATL-PHYS-CSC-2007-000, April 22, 2008.

[16] K. Lannon, F. Margaroli, C. Neu, *Measurements of the Production, Decay and Properties of the Top Quark: A Review*, arXiv:1201.5873v1 [hep-ph], January 27, 2012.

[17] M. Baak *et al.*, *The Electroweak Fit of the Standard Model after the Discovery of a New Boson at the LHC*, arXiv:1209.2716v2 [hep-ph], September 25, 2012.

[18] The ATLAS Collaboration, *Top quark properities*, ATL-PHYS-CSC-2007-000, April 18, 2008.

[19] M. Aliev *et al.*, *HATHOR - HAdronic Top and Heavy quarks cross section calculatoR*, Comput. Phys. Commun.182 (2011) 1034-1046, arXiv:1007:1327 [hep-ph] .

[20] The ATLAS Collaboration, *Top quark mass measurements with ATLAS*, ATL-PHYS-CSC-2007-000, April 8, 2008.

[21] The ATLAS Collaboration, *Determination of Top pair production cross-section in ATLAS*, ATL-PHYS-CSC-2007-000, April 22, 2008.

[22] E. P. Cortezon, *Top Mass Measurements at LHC and Tevatron (slides)*, Hadron Collider Physics Symposium, Kyoto, November 12-16, 2012.

[23] The ATLAS Collaboration, *Measurement of the Top Quark Mass with the Template Method in the* $t\bar{t} \rightarrow lepton + jets$ *Channel using ATLAS Data*, arXiv:1203.5755v2 [hep-ex], June 12, 2012.

[24] The ATLAS Collaboration, *Measurement of the charge asymmetry in dileptonic decays of top quark pairs in pp collisions at* $\sqrt{s} = 7$ *TeV using the ATLAS detector*, ATLAS-CONF-2012-057, June 4, 2012.

[25] The ATLAS Collaboration, *Measurement of the charge asymmetry in top quark pair production in pp collisions at* $\sqrt{s} = 7$ *TeV using the ATLAS detector*, ATLAS-CONF-2011-106, August 4, 2011.

[26] The ATLAS Collaboration, *Measurement of the charge asymmetry in top quark pair production in pp collisions at* $\sqrt{s} = 7$ *TeV using the ATLAS detector*, CERN-PH-EP-2012-015, arXiv:1203.4211v2 [hep-ex], Jun 28, 2012.

[27] W.-M. Yao *et al.*, *2007 Review of Particle Physics*, J. Phys. G 33, 1 (2006) and 2007 partial update for the 2008 edition.

[28] ATLAS TWiki, *https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/TopMC10For2011Data#MC10_*

[29] S. Frixione, P. Nason, C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, arXiv:0709.2092v1 [hep-ph], September 13, 2007.

[30] B. P. Kersevan and E. Richter-Was, *The Monte Carlo Event Generator AcerMC versions 2.0 to 3.8 with interfaces to PYTHIA 6.4, HERWIG 6.5 and ARIADNE 4.1*, arXiv:hep-ph/0405247v3, November 12, 2012.

[31] The ATLAS Collaboration, G. Aad et al.,*The ATLAS Experiment at the CERN Large Hadron Collider*, 2008 JINST 3 S08003.

[32] G. Aad et al., *Journal of Instrumentation* 3 (2008) S08003.

[33] ATLAS Collaboration, Expected Performance of the ATLAS Experiment, Detector, Trigger and Physics, CERN-OPEN-2008-020, December 2008.

[34] The ATLAS Collaboration, *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment*, ATLAS-CONF-2011-077, August 1,2011.

[35] The ATLAS Collaboration, *Performance of the Reconstruction and Identification of Hadronic Tau Decays with ATLAS*, ATLAS-CONF-2011-152, November 13,2011.

[36] The ATLAS Collaboration, *Expected Performance of the ATLAS Detector*, CERN-OPEN-2008-020, December 2008.

[37] The ATLAS Collaboration, *Jet energy measurement with the ATLAS detector in proton-proton collisions at* $\sqrt{s} = 7$ *TeV*, CERN-PH-EP-2011-191, December 30,2011.

[38] The ATLAS Collaboration, *Jet energy scale and its systematic uncertainty in proton-proton collisions at* $\sqrt{s} = 7$ *TeV with ATLAS 2010 data*, ATLAS-CONF-2011-032, March 22, 2011.

[39] K. Prokofiev on behalf of The ATLAS Collaboration, *Reconstruction of primary vertices in pp collisions at energies of 900 GeV and 7 TeV with the ATLAS detector*, SNSN-323-63, October 8, 2010.

[40] ATLAS Collaboration, Performance of Missing Transverse Momentum Reconstruction in Proton-Proton Collisions at $\sqrt{s} = 7$ TeV with ATLAS, CERN-PH-EP-2011-114, December 6, 2011.

[41] The ATLAS Collaboration, *Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data*, ATLAS-CONF-2011-102, July 20, 2011.

[42] ATLAS TWiki, *ihttps://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/BTaggingBenchmarks#b_t*

[43] O. Arnaez on behalf of the ATLAS Collaboration, *Electron reconstruction and identification with the ATLAS detector*, ATL-PHYS-PROC-2009-118, October 1, 2009.

[44] The ATLAS Collaboration, *Expected electron performance in the ATLAS experiment*, ATL-PHYS-INT-2010-126, November 26, 2010.

[45] The ATLAS Collaboration, *Jet energy resolution and selection efficiency relative to track jets from in-situ techniques with the ATLAS detector using proton-proton collisions at a center of mass energy* $\sqrt{s} = 7$ *TeV*, ATLAS-CONF-2010-054, July 20, 2010.

[46] D. Berge et al., *The Configuration System of the ATLAS Trigger*, ATL-DAQ-PROC-2009-047, December 20, 2009.

[47] ATLAS TWiki, *https://twiki.cern.ch/twiki/bin/viewauth/Atlas/TAPMMissingEt.*

[48] The ATLAS Collaboration, *$Z \to \tau\tau$ cross section measurement in proton-proton collisions at 7 TeV with the ATLAS experiment*, ATLAS-CONF-2012-006, February 26, 2012.

[49] The ATLAS Collaboration, *Performance of the ATLAS tau trigger in 2011*, ATLAS-COM-CONF-2012-054 (not published yet), September 26, 2012.

[50] P.Assamagan, *et. al.*, *Search for charged Higgs boson in the τ + jets final state in $t\bar{t}$ decays with 4.6 fb^{-1} of pp collision data recorded at $\sqrt{s} = 7$ TeV with the ATLAS experiment*, Supporting note (for internal circulation only), April 24, 2012.

[51] ATLAS TWiki, *https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/TauSystematicsSummerConf201*

[52] ATLAS TWiki, *https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/TopSystematicUncertainties2011*

[53] ATLAS TWiki,*https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/TopCommonObjects2011rel16.*

[54] ATLAS TWiki,*https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/TopMC10For2011Data#MC10_C*

[55] The ATLAS Collaboration, *Search for charged Higgs bosons decaying via $H^{\pm} \to \tau\nu$ in $t\bar{t}$ events using 4.6 fb^{-1} of pp collision data at $\sqrt{s} = 7$ TeV with the ATLAS detector*, ATLAS-CONF-2012-011 March 2, 2012.

[56] F.A. Berends *et al.*, *Multijet production in W, Z events at $p\bar{p}$ colliders*,Physics Letters B224 (1989) 237, March 1989.

[57] ATLAS TWiki, *https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/WplusJetsBackgroundsforTopAn*

[58] ATLAS TWiki, *https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/HforTool#ATLAS_notes.*

[59] The D0 Collaboration, *Measurement of the $t\bar{t}$ production cross-section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ using kinematic characteristics of lepton + jets events*, arXiv:0705.2788v1 [hep-ex] May 19, 2007.

[60] ATLAS TWiki, *https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/TopJetLiaisonR16Recommendat*

[61] ATLAS TWiki, *https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/Analysis16#Calibrations_to_be_u*

[62] The ATLAS Collaboration, *Measurement of the top quark pair production cross section with ATLAS in pp collisions at $\sqrt{s} = 7$ TeV using final states with an electron or a muon and a hadronically decaying τ lepton* , arXiv:1205.2067v1 [hep-ex] May 9, 2012.

[63] The ATLAS Collaboration, *Measurement of the $t\bar{t}$ production cross section in the final state with a hadronically decaying tau lepton and jets using the ATLAS detector*, ATLAS-CONF-2012-032, March 12, 2012.

[64] M.Bosman, *et al.* , *Simultaneous measurement of the $t\bar{t}$production cross- section and R_b in the lepton plus jets final state at $\sqrt{s} = 7$ TeV*, ATLAS NOTE, 2012.

[65] J.Nadal, *Simultaneous measurement of the top quark pair production cross-section and R_b in the ATLAS experiment*, July 2012.

[66] F. James, *MINUIT, Function Minimization and Error Analysis, Reference Manual* CERN Version 94.1, CERN Program Library Long Writeup D506, CERN Geneva, Switzerland.

[67] W. Verkerke, *Guide to parametrized likelihood analysis* (Version 1.07), April 19, 2012.

[68] The ATLAS Collaboration, *Measurement of the $t\bar{t}$ production cross-section in pp collisions at $\sqrt{s}$ = 7 TeV using kinematic information of lepton+jets events*, ATLAS-CONF-2011-121, August 21, 2011.

[69] ATLAS TWiki, *https://twiki.cern.ch/twiki/bin/view/Main/JetEnergyResolutionProviderPrescription* .

[70] The ATLAS Collaboration, *Simultaneous measurement of the top quark pair production cross-section and R_b in the ATLAS experiment*, ATLAS NOTE (not yet published).

[71] T. Aaltonen *et. al.*, *Measurements of the Top-quark Mass and the $t\bar{t}$ Cross Section in the Hadronic $\tau+$ Jets Decay Channel at $\sqrt{s} = 1.96$ TeV* arxiv:1208.5720v4 [hep-ex], September 25, 2012.

[72] V.M. Abazov, *et. al.*, *Measurement of $t\bar{t}$ production in the $\tau+$ jets topology using $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV* arxiv:1008.4284v2 [hep-ex], October 27, 2010.

[73] The ATLAS Collaboration, *Measurement of the $t\bar{t}$ production cross section in the tau + jets channel using the ATLAS detector* , arXiv:1211.7205v1 [hep-ex], November 30, 2012.

Appendix A

Monte Carlo samples

Table 1: MC10 samples used in this analysis.

Sample
105200.T1_McAtNlo_Jimmy.merge.NTUP_TOP.e598_s933_s946_r2302_r2300_p572
107650.AlpgenJimmyZeeNp0_pt20.merge.NTUP_TOP.e737_s933_s946_r2302_r2300_p572
107651.AlpgenJimmyZeeNp1_pt20.merge.NTUP_TOP.e737_s933_s946_r2302_r2300_p572
107652.AlpgenJimmyZeeNp2_pt20.merge.NTUP_TOP.e737_s933_s946_r2302_r2300_p572
107653.AlpgenJimmyZeeNp3_pt20.merge.NTUP_TOP.e737_s933_s946_r2302_r2300_p572
107654.AlpgenJimmyZeeNp4_pt20.merge.NTUP_TOP.e737_s933_s946_r2302_r2300_p572
107655.AlpgenJimmyZeeNp5_pt20.merge.NTUP_TOP.e737_s933_s946_r2302_r2300_p572
107660.AlpgenJimmyZmumuNp0_pt20.merge.NTUP_TOP.e737_s933_s946_r2302_r2300_p572
107661.AlpgenJimmyZmumuNp1_pt20.merge.NTUP_TOP.e737_s933_s946_r2302_r2300_p572
107662.AlpgenJimmyZmumuNp2_pt20.merge.NTUP_TOP.e737_s933_s946_r2302_r2300_p572
107663.AlpgenJimmyZmumuNp3_pt20.merge.NTUP_TOP.e737_s933_s946_r2302_r2300_p572
107664.AlpgenJimmyZmumuNp4_pt20.merge.NTUP_TOP.e737_s933_s946_r2302_r2300_p572
107665.AlpgenJimmyZmumuNp5_pt20.merge.NTUP_TOP.e737_s933_s946_r2302_r2300_p572
107670.AlpgenJimmyZtautauNp0_pt20.merge.NTUP_TOP.e737_s933_s946_r2302_r2300_p572
107671.AlpgenJimmyZtautauNp1_pt20.merge.NTUP_TOP.e737_s933_s946_r2302_r2300_p572
107672.AlpgenJimmyZtautauNp2_pt20.merge.NTUP_TOP.e737_s933_s946_r2302_r2300_p572
107673.AlpgenJimmyZtautauNp3_pt20.merge.NTUP_TOP.e737_s933_s946_r2302_r2300_p572
107674.AlpgenJimmyZtautauNp4_pt20.merge.NTUP_TOP.e737_s933_s946_r2302_r2300_p572
107675.AlpgenJimmyZtautauNp5_pt20.merge.NTUP_TOP.e737_s933_s946_r2302_r2300_p572
107680.AlpgenJimmyWenuNp0_pt20.merge.NTUP_TOP.e600_s933_s946_r2302_r2300_p572
107681.AlpgenJimmyWenuNp1_pt20.merge.NTUP_TOP.e600_s933_s946_r2302_r2300_p572
107682.AlpgenJimmyWenuNp2_pt20.merge.NTUP_TOP.e760_s933_s946_r2302_r2300_p572
107683.AlpgenJimmyWenuNp3_pt20.merge.NTUP_TOP.e760_s933_s946_r2302_r2300_p572
107684.AlpgenJimmyWenuNp4_pt20.merge.NTUP_TOP.e760_s933_s946_r2302_r2300_p572
107685.AlpgenJimmyWenuNp5_pt20.merge.NTUP_TOP.e760_s933_s946_r2302_r2300_p572
107690.AlpgenJimmyWmunuNp0_pt20.merge.NTUP_TOP.e600_s933_s946_r2302_r2300_p572
107691.AlpgenJimmyWmunuNp1_pt20.merge.NTUP_TOP.e600_s933_s946_r2302_r2300_p572
107692.AlpgenJimmyWmunuNp2_pt20.merge.NTUP_TOP.e760_s933_s946_r2302_r2300_p572
107693.AlpgenJimmyWmunuNp3_pt20.merge.NTUP_TOP.e760_s933_s946_r2302_r2300_p572
107694.AlpgenJimmyWmunuNp4_pt20.merge.NTUP_TOP.e760_s933_s946_r2302_r2300_p572
107695.AlpgenJimmyWmunuNp5_pt20.merge.NTUP_TOP.e760_s933_s946_r2302_r2300_p572
107700.AlpgenJimmyWtaunuNp0_pt20.merge.NTUP_TOP.e600_s933_s946_r2302_r2300_p572
107701.AlpgenJimmyWtaunuNp1_pt20.merge.NTUP_TOP.e600_s933_s946_r2302_r2300_p572
107702.AlpgenJimmyWtaunuNp2_pt20.merge.NTUP_TOP.e760_s933_s946_r2302_r2300_p572
107703.AlpgenJimmyWtaunuNp3_pt20.merge.NTUP_TOP.e760_s933_s946_r2302_r2300_p572
107704.AlpgenJimmyWtaunuNp4_pt20.merge.NTUP_TOP.e760_s933_s946_r2302_r2300_p572
107705.AlpgenJimmyWtaunuNp5_pt20.merge.NTUP_TOP.e760_s933_s946_r2302_r2300_p572
105985.WW_Herwig.merge.NTUP_TOP.e598_s933_s946_r2302_r2300_p572
105986.ZZ_Herwig.merge.NTUP_TOP.e598_s933_s946_r2302_r2300_p572
105987.WZ_Herwig.merge.NTUP_TOP.e598_s933_s946_r2302_r2300_p572
108340.st_tchan_enu_McAtNlo_Jimmy.merge.NTUP_TOP.e598_s933_s946_r2302_r2300_p572
108341.st_tchan_munu_McAtNlo_Jimmy.merge.NTUP_TOP.e598_s933_s946_r2302_r2300_p572
108342.st_tchan_taunu_McAtNlo_Jimmy.merge.NTUP_TOP.e598_s933_s946_r2302_r2300_p572
108343.st_schan_enu_McAtNlo_Jimmy.merge.NTUP_TOP.e598_s933_s946_r2302_r2300_p572
108344.st_schan_munu_McAtNlo_Jimmy.merge.NTUP_TOP.e598_s933_s946_r2302_r2300_p572
108345.st_schan_taunu_McAtNlo_Jimmy.merge.NTUP_TOP.e598_s933_s946_r2302_r2300_p572
108346.st_Wt_McAtNlo_Jimmy.merge.NTUP_TOP.e598_s933_s946_r2302_r2300_p572

Sample
107280.AlpgenJimmyWbbFullNp0_pt20.merge.NTUP_TOP.e786_s933_s946_r2302_r2300_p572
107281.AlpgenJimmyWbbFullNp1_pt20.merge.NTUP_TOP.e786_s933_s946_r2302_r2300_p572
107282.AlpgenJimmyWbbFullNp2_pt20.merge.NTUP_TOP.e786_s933_s946_r2302_r2300_p572
107283.AlpgenJimmyWbbFullNp3_pt20.merge.NTUP_TOP.e786_s933_s946_r2302_r2300_p572
117284.AlpgenWccFullNp0_pt20.merge.NTUP_TOP.e786_s933_s946_r2302_r2300_p572
117285.AlpgenWccFullNp1_pt20.merge.NTUP_TOP.e786_s933_s946_r2302_r2300_p572
117286.AlpgenWccFullNp2_pt20.merge.NTUP_TOP.e786_s933_s946_r2302_r2300_p572
117287.AlpgenWccFullNp3_pt20.merge.NTUP_TOP.e786_s933_s946_r2302_r2300_p572
117293.AlpgenWcNp0_pt20.merge.NTUP_TOP.e786_s933_s946_r2302_r2300_p595
117294.AlpgenWcNp1_pt20.merge.NTUP_TOP.e786_s933_s946_r2302_r2300_p595
117295.AlpgenWcNp2_pt20.merge.NTUP_TOP.e786_s933_s946_r2302_r2300_p595
117296.AlpgenWcNp3_pt20.merge.NTUP_TOP.e786_s933_s946_r2302_r2300_p595
117297.AlpgenWcNp4_pt20.merge.NTUP_TOP.e786_s933_s946_r2302_r2300_p595
105205.AcerMCttbar.merge.NTUP_TOP.e574_s933_s946_r2302_r2300_p572
105860.TTbar_PowHeg_Jimmy.merge.NTUP_TOP.e871_s933_s946_r2302_r2300_p572
105861.TTbar_PowHeg_Pythia.merge.NTUP_TOP.e871_s933_s946_r2302_r2300_p572
117255.AcerMCttbar_isr_down.merge.NTUP_TOP.e574_s933_s946_r2302_r2300_p572
117256.AcerMCttbar_isr_up.merge.NTUP_TOP.e574_s933_s946_r2302_r2300_p572
117257.AcerMCttbar_fsr_down.merge.NTUP_TOP.e574_s933_s946_r2302_r2300_p572
117258.AcerMCttbar_fsr_up.merge.NTUP_TOP.e574_s933_s946_r2302_r2300_p572
117259.AcerMCttbar_isr_down_fsr_down.merge.NTUP_TOP.e598_s933_s946_r2302_r2300_p572
117260.AcerMCttbar_isr_up_fsr_up.merge.NTUP_TOP.e598_s933_s946_r2302_r2300_p572
116108.AlpgenJimmyttbbinclNp0_baseline.merge.NTUP_TOP.e791_s933_s946_r2302_r2300_p572
117899.AlpgenJimmyttbarlnlnNp5_baseline.merge.NTUP_TOP.e804_s933_s946_r2302_r2300_p572
117898.AlpgenJimmyttbarlnlnNp4_baseline.merge.NTUP_TOP.e804_s933_s946_r2302_r2300_p572
117897.AlpgenJimmyttbarlnlnNp3_baseline.merge.NTUP_TOP.e804_s933_s946_r2302_r2300_p572
117889.AlpgenJimmyttbarlnqqNp5_baseline.merge.NTUP_TOP.e804_s933_s946_r2302_r2300_p572
117888.AlpgenJimmyttbarlnqqNp4_baseline.merge.NTUP_TOP.e804_s933_s946_r2302_r2300_p572
117887.AlpgenJimmyttbarlnqqNp3_baseline.merge.NTUP_TOP.e804_s933_s946_r2302_r2300_p572
105896.AlpgenJimmyttbarlnqqNp2_baseline.merge.NTUP_TOP.e600_s933_s946_r2302_r2300_p572
105895.AlpgenJimmyttbarlnqqNp1_baseline.merge.NTUP_TOP.e600_s933_s946_r2302_r2300_p572
105894.AlpgenJimmyttbarlnqqNp0_baseline.merge.NTUP_TOP.e600_s933_s946_r2302_r2300_p572
105892.AlpgenJimmyttbarlnlnNp2_baseline.merge.NTUP_TOP.e600_s933_s946_r2302_r2300_p572
105891.AlpgenJimmyttbarlnlnNp1_baseline.merge.NTUP_TOP.e600_s933_s946_r2302_r2300_p572
105890.AlpgenJimmyttbarlnlnNp0_baseline.merge.NTUP_TOP.e600_s933_s946_r2302_r2300_p572

Appendix B

Data-MC comparison distributions

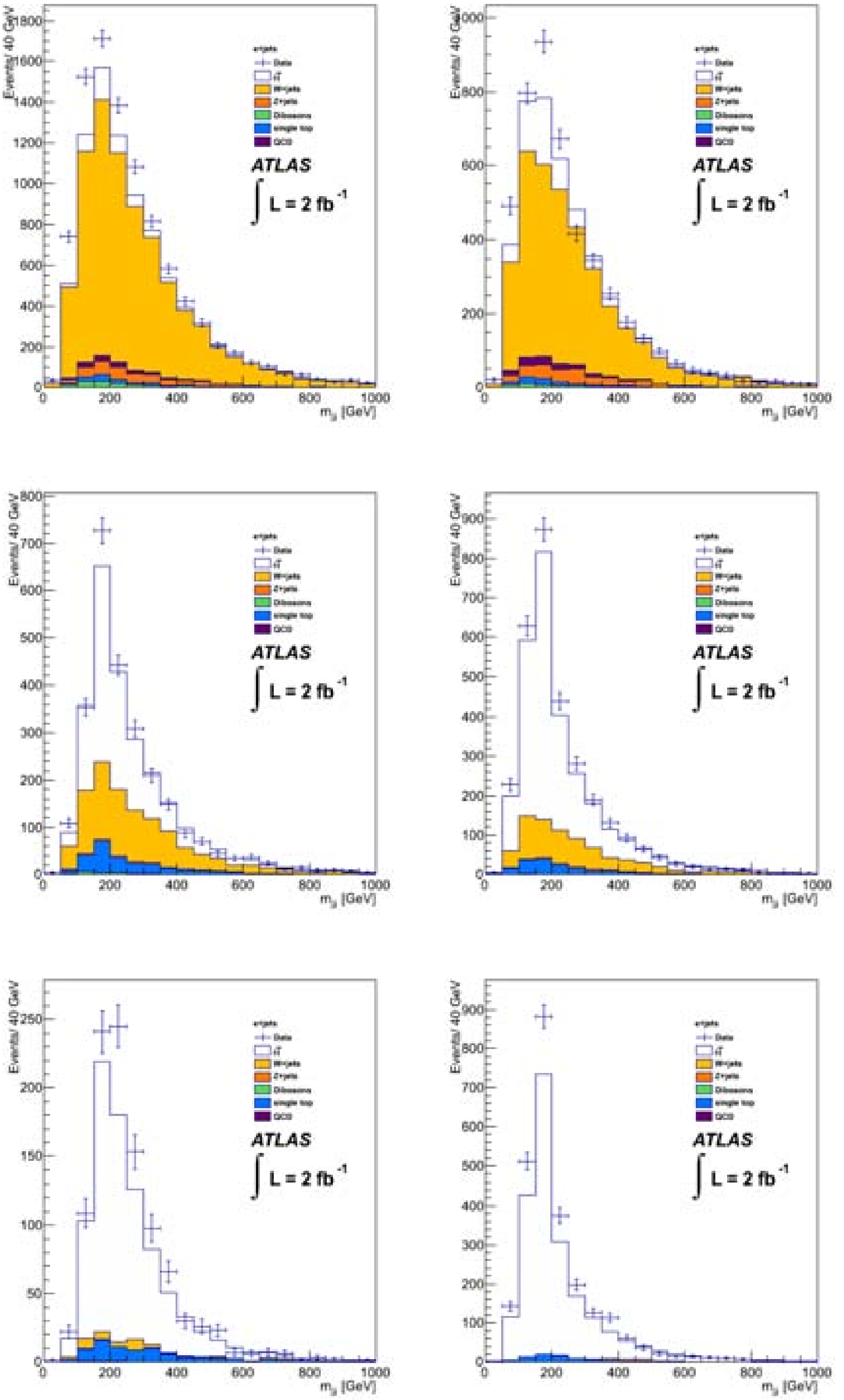


Figure 1: Data-MC comparison: hadronic top mass in the electron channel.

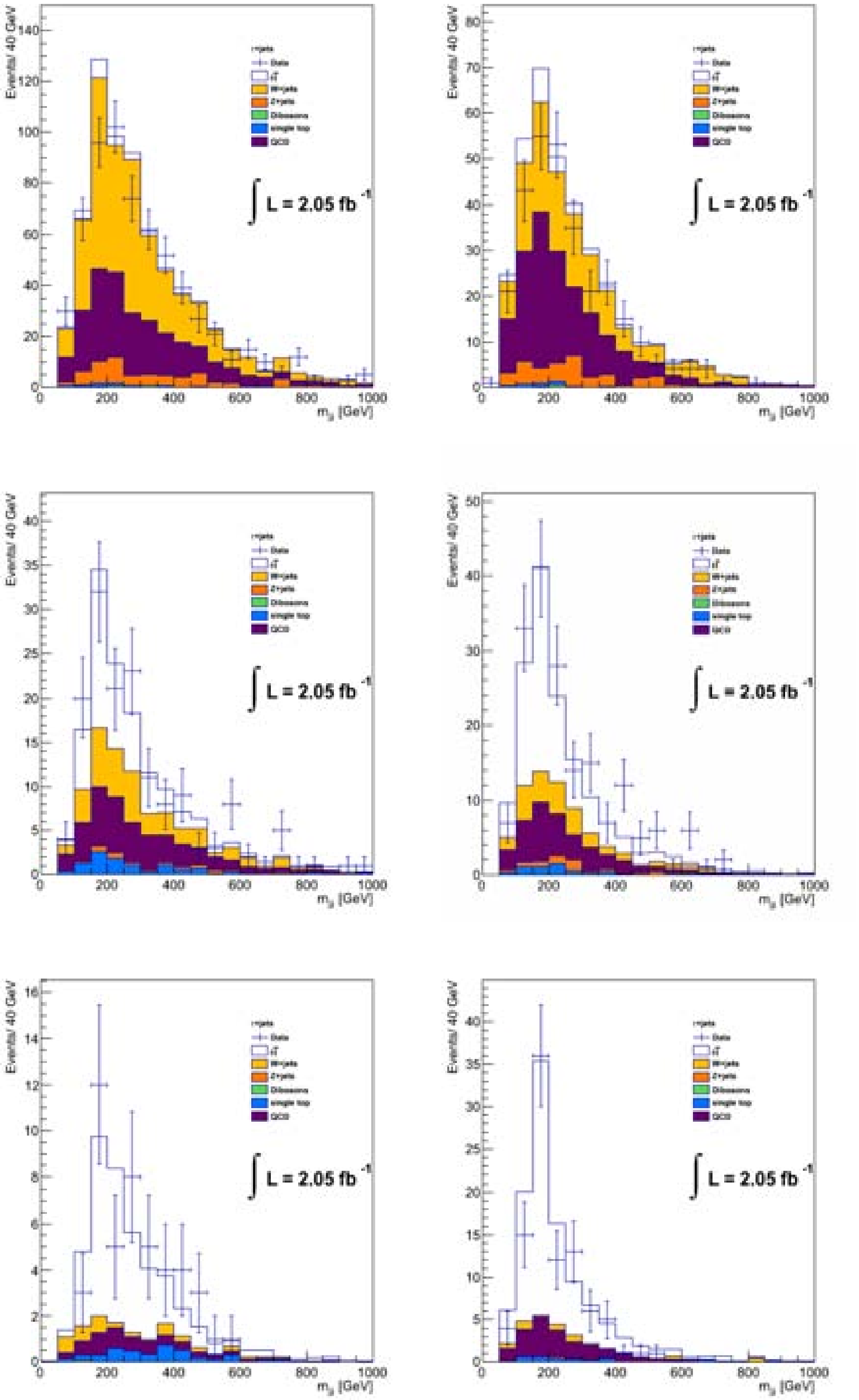


Figure 2: Data-MC comparison: hadronic top mass in the 1 prong tau channel.

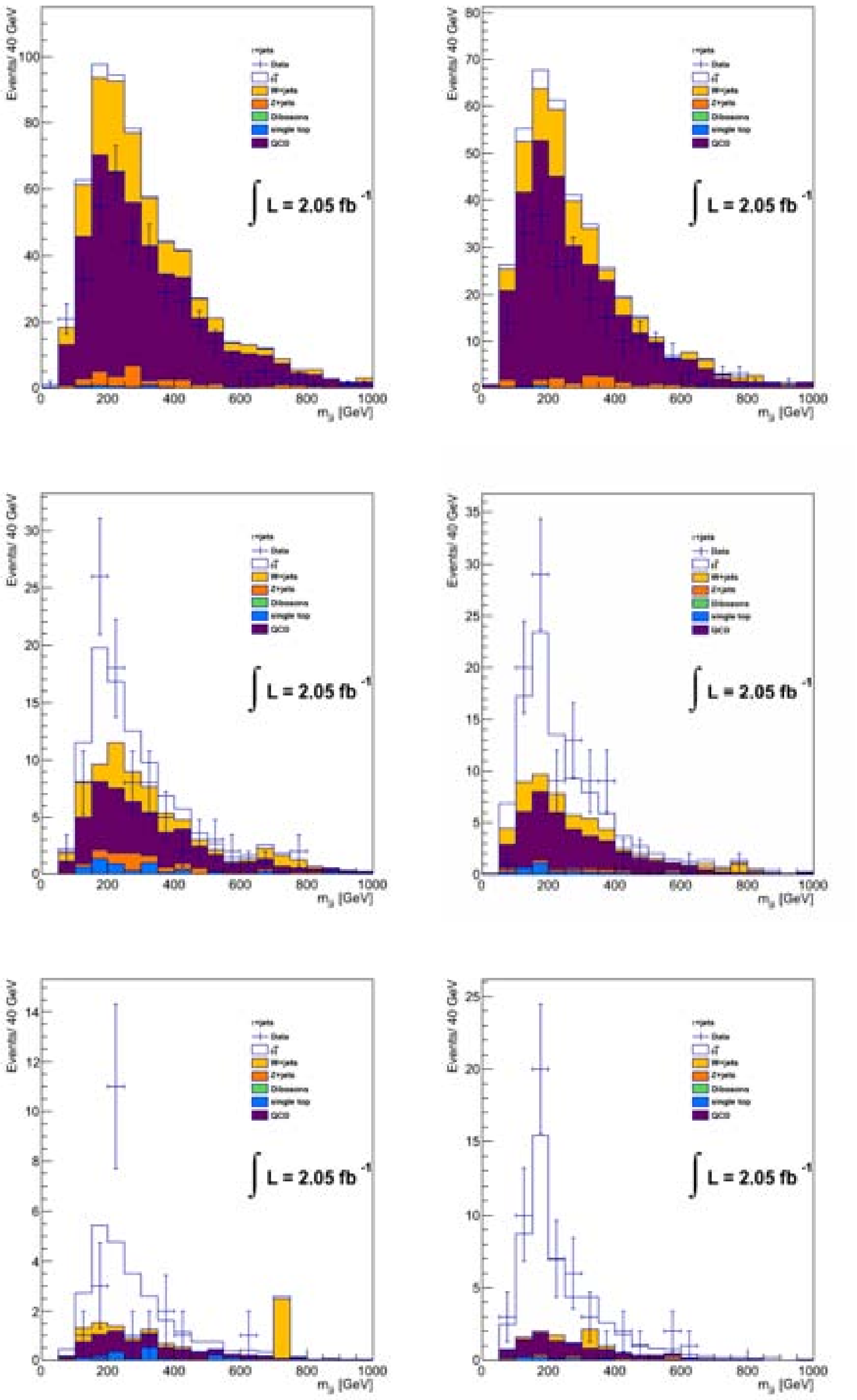


Figure 3: Data-MC comparison: hadronic top mass in the 3 prong tau channel.

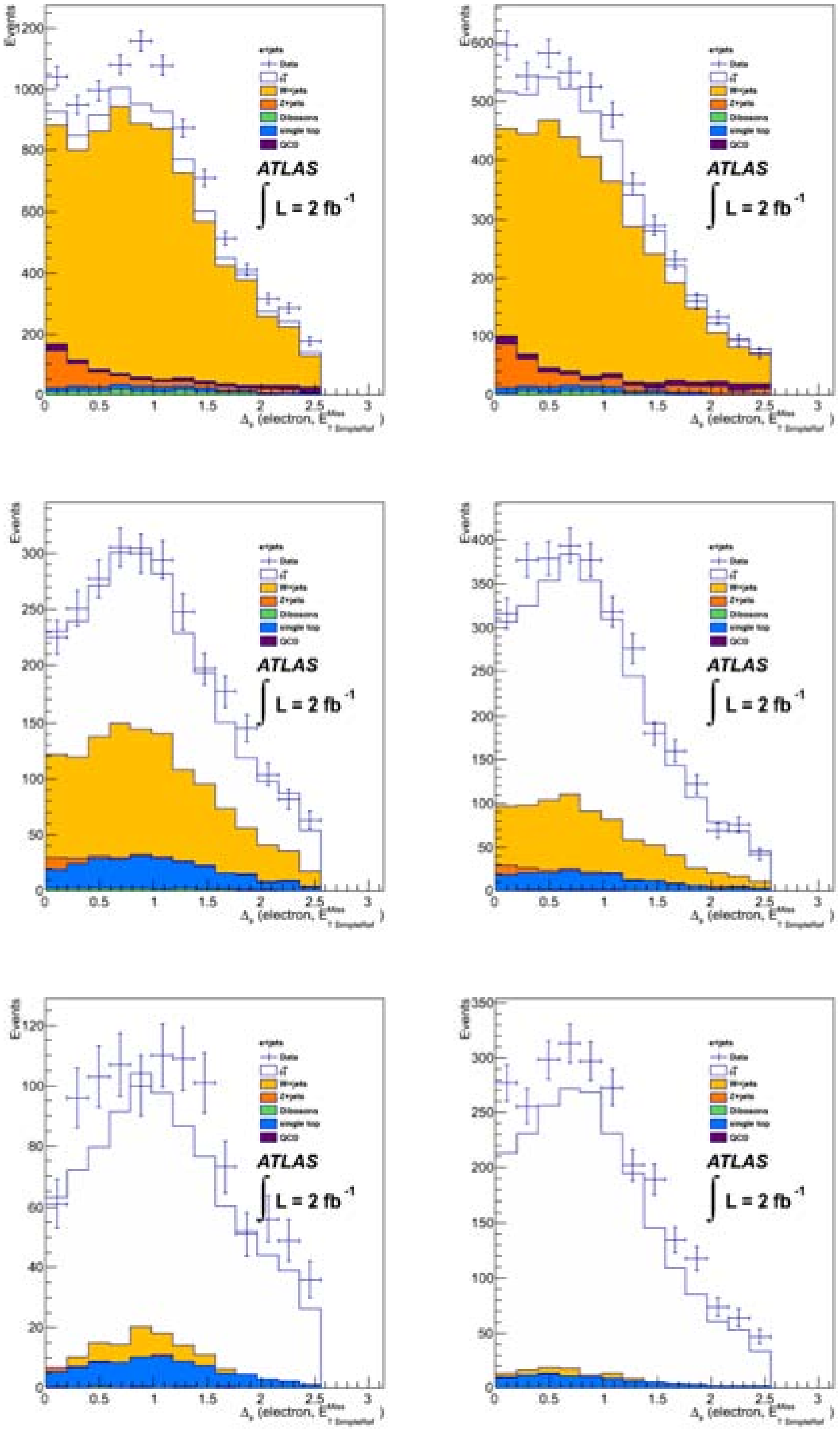


Figure 4: Data-MC comparison: $\Delta\phi$ (e , $E_{\mathrm{T}}^{\mathrm{miss}}$) in the electron channel.

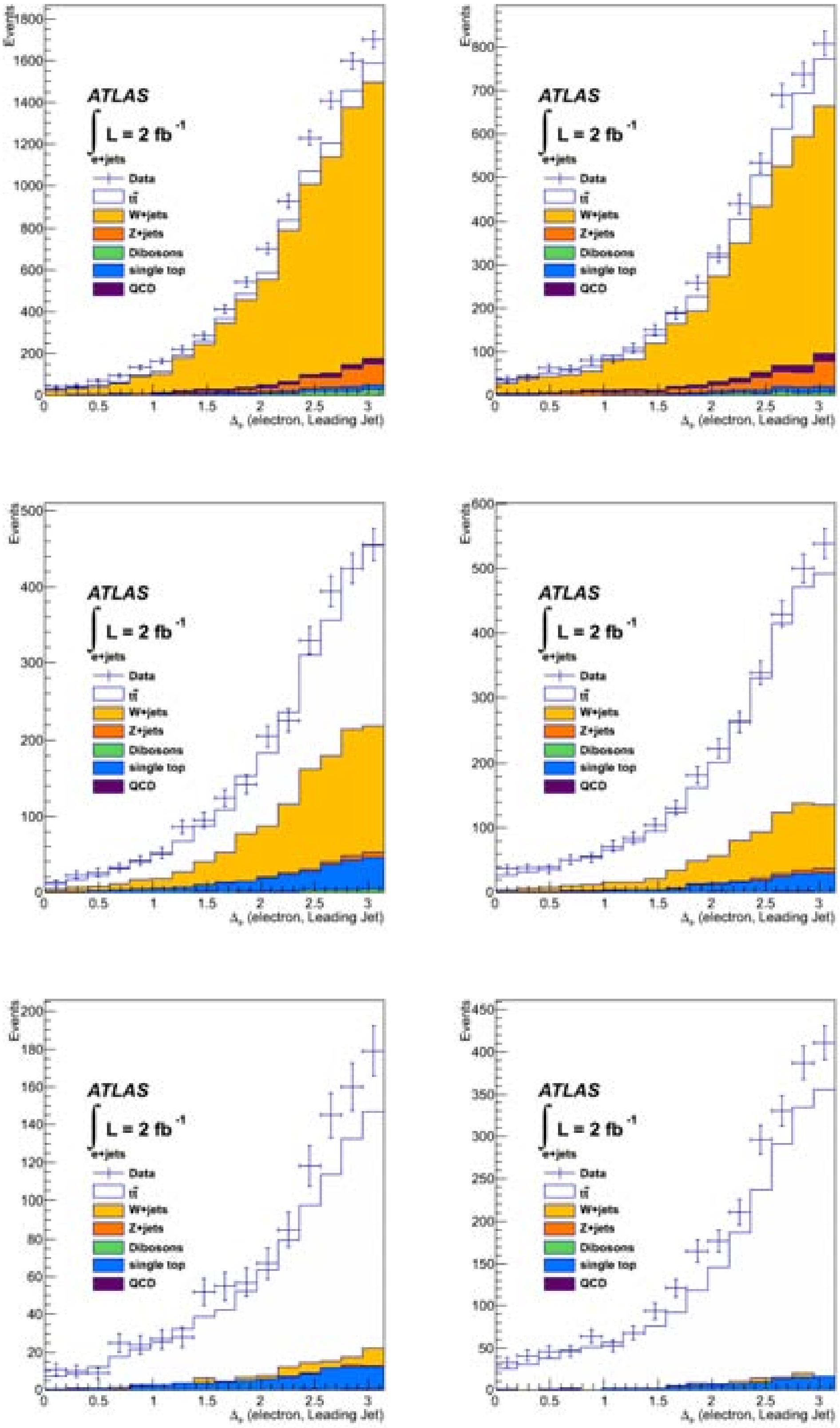


Figure 5: Data-MC comparison: $\Delta\phi$ (e , leading jet p_{T}) in the electron tau channel.

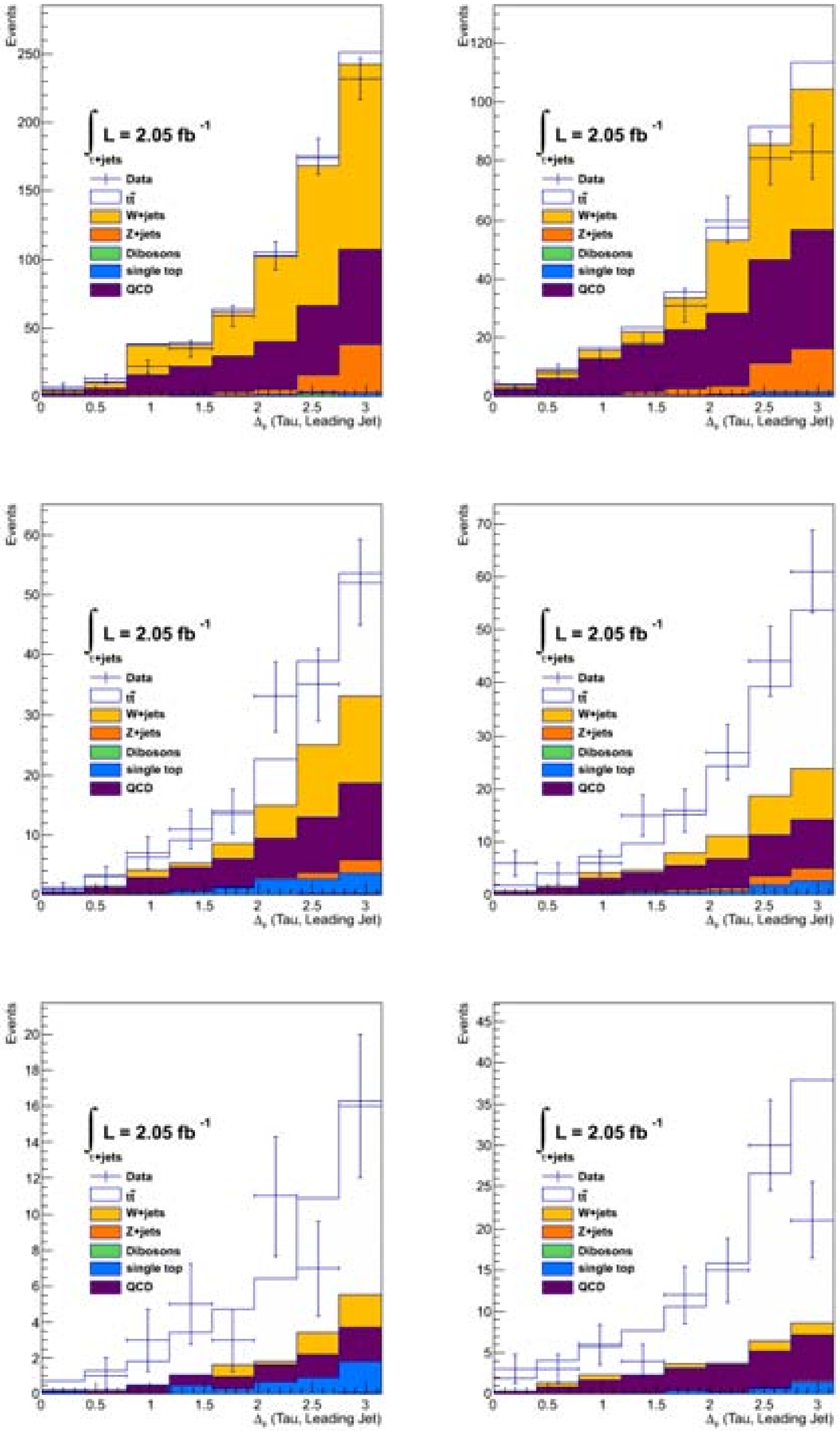


Figure 6: Data-MC comparison: $\Delta\phi$ (τ , leading jet p_{T}) in the 1 prong tau channel.

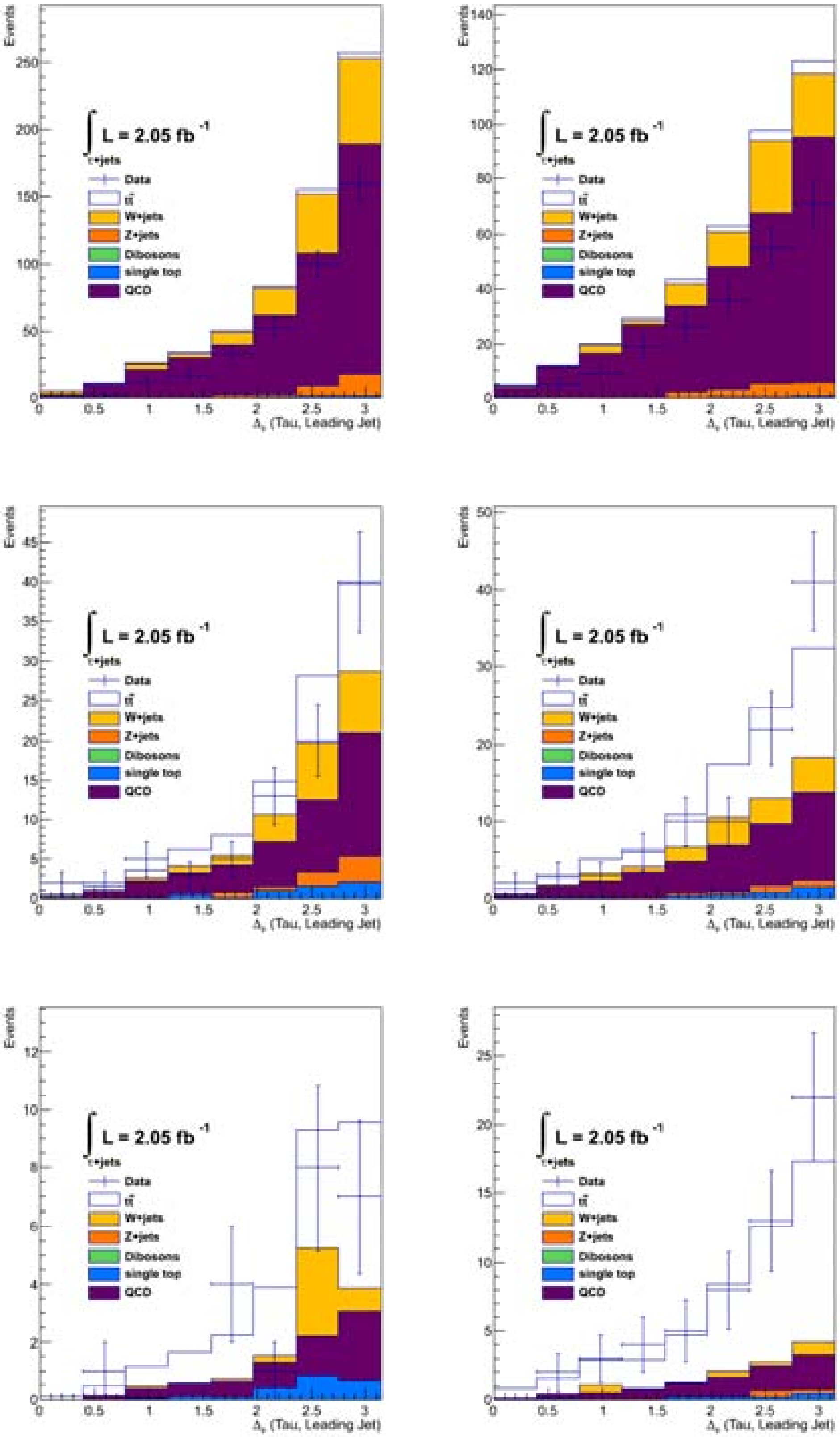


Figure 7: Data-MC comparison: $\Delta\phi$ (τ , leading jet p_{T}) in the 3 prong tau channel.

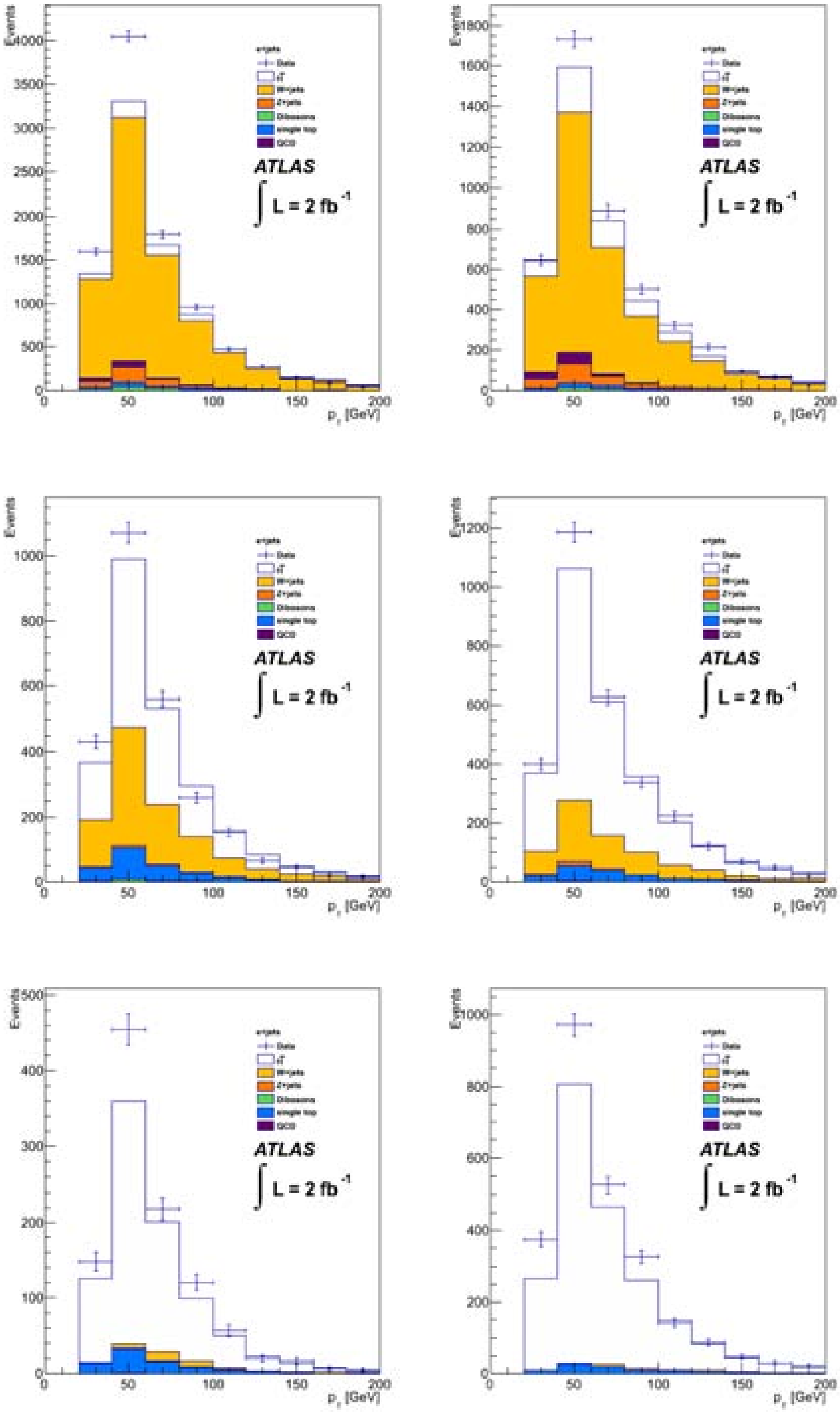


Figure 8: Data-MC comparison: lepton (electron) p_T in the electron channel.

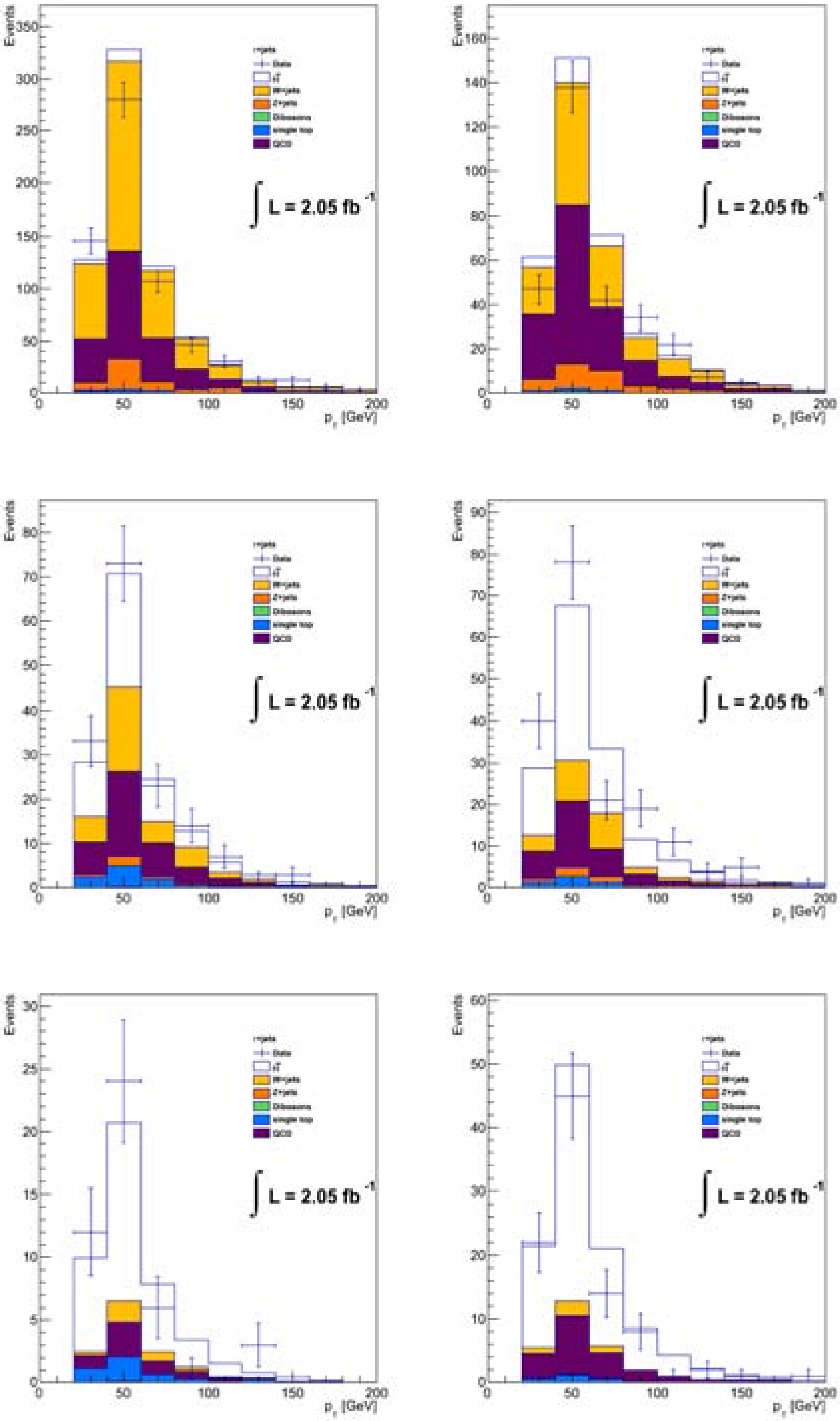


Figure 9: Data-MC comparison: tau p_{T} in the 1 prong tau channel.

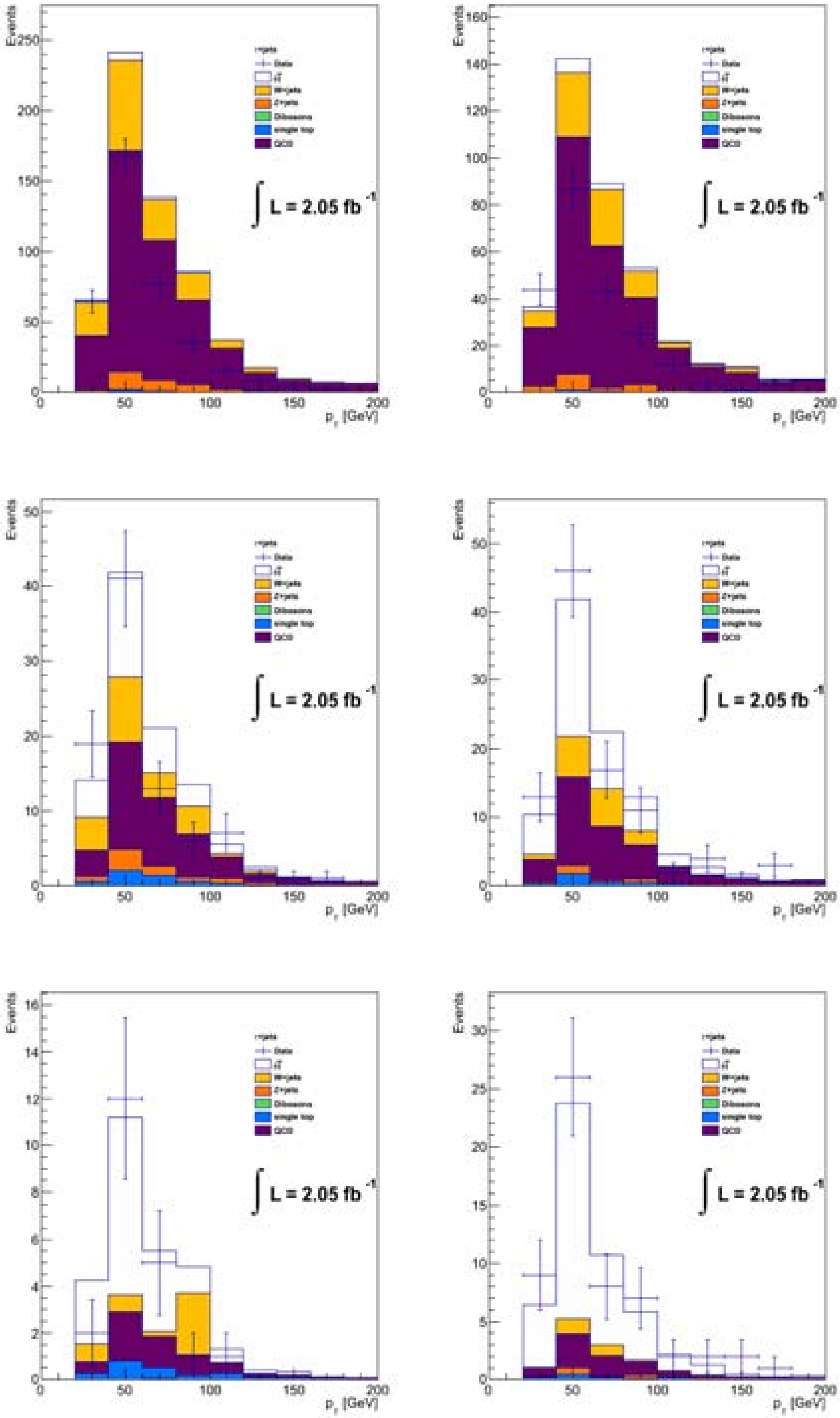


Figure 10: Data-MC comparison: tau p_{T} in the 3 prong tau channel.

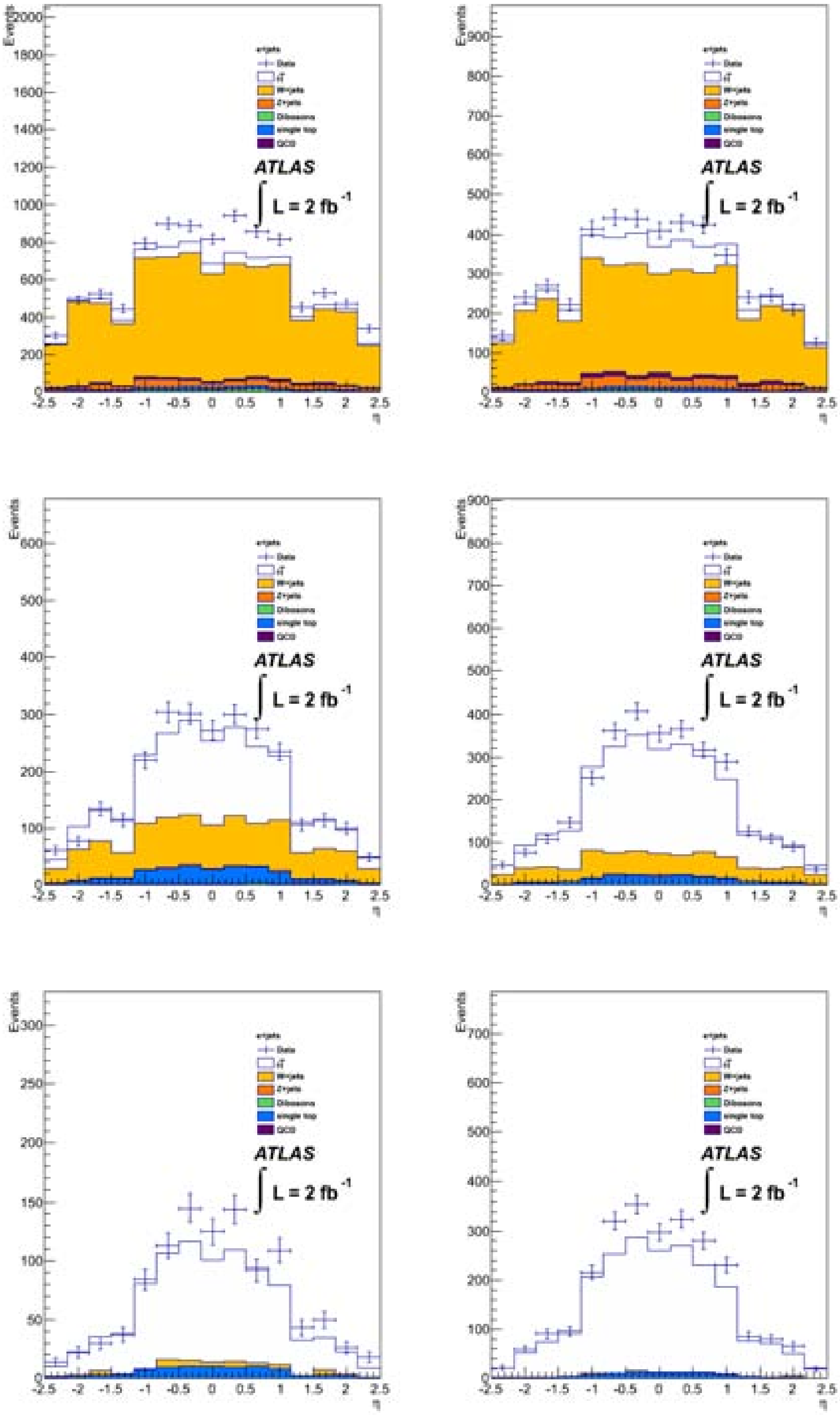


Figure 11: Data-MC comparison: lepton (electron) η in the electron channel.

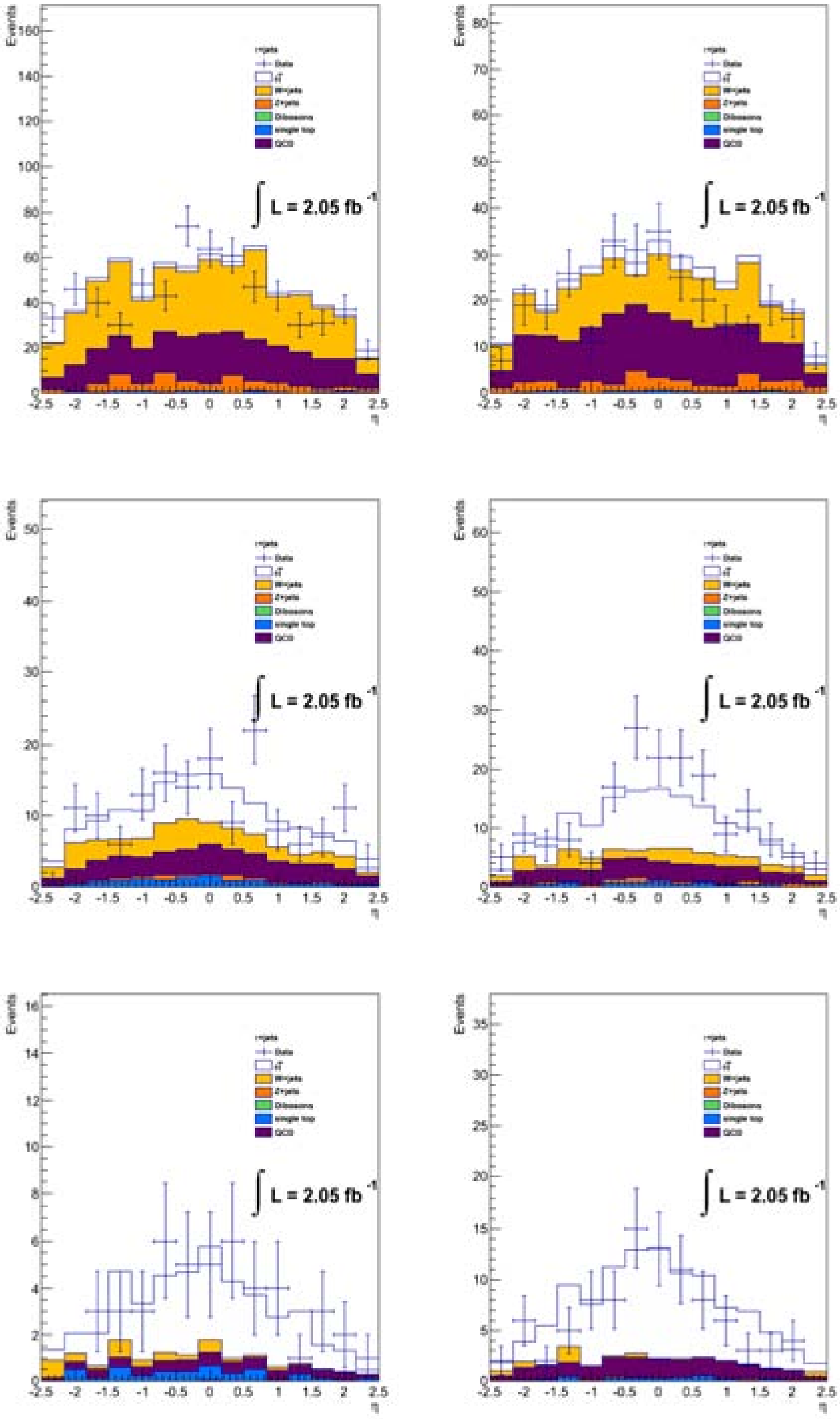


Figure 12: Data-MC comparison: tau η in the 1 prong tau channel.

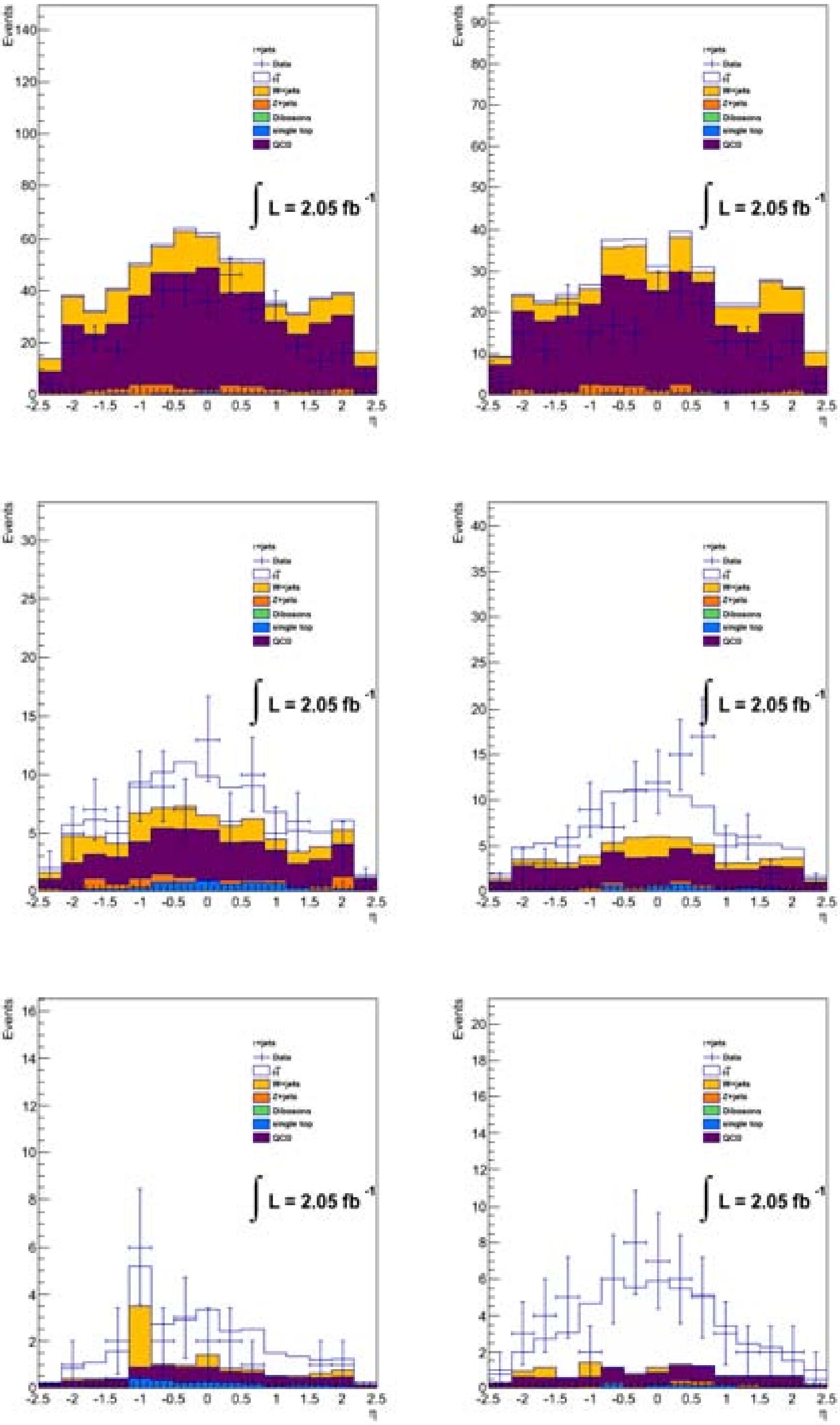


Figure 13: Data-MC comparison: tau η in the 3 prong tau channel.

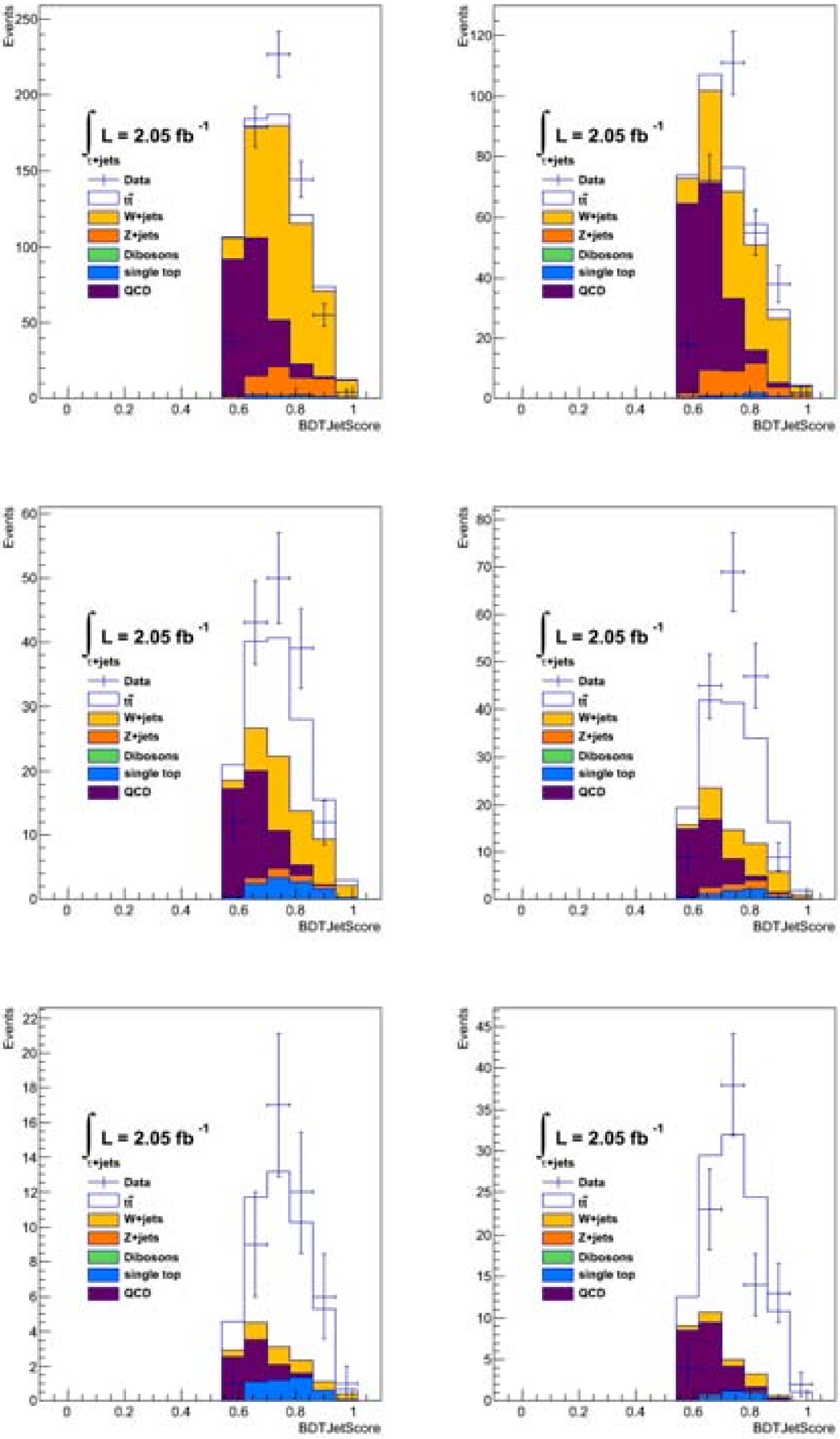


Figure 14: Data-MC comparison: BDTJetScore in the 1 prong tau channel.

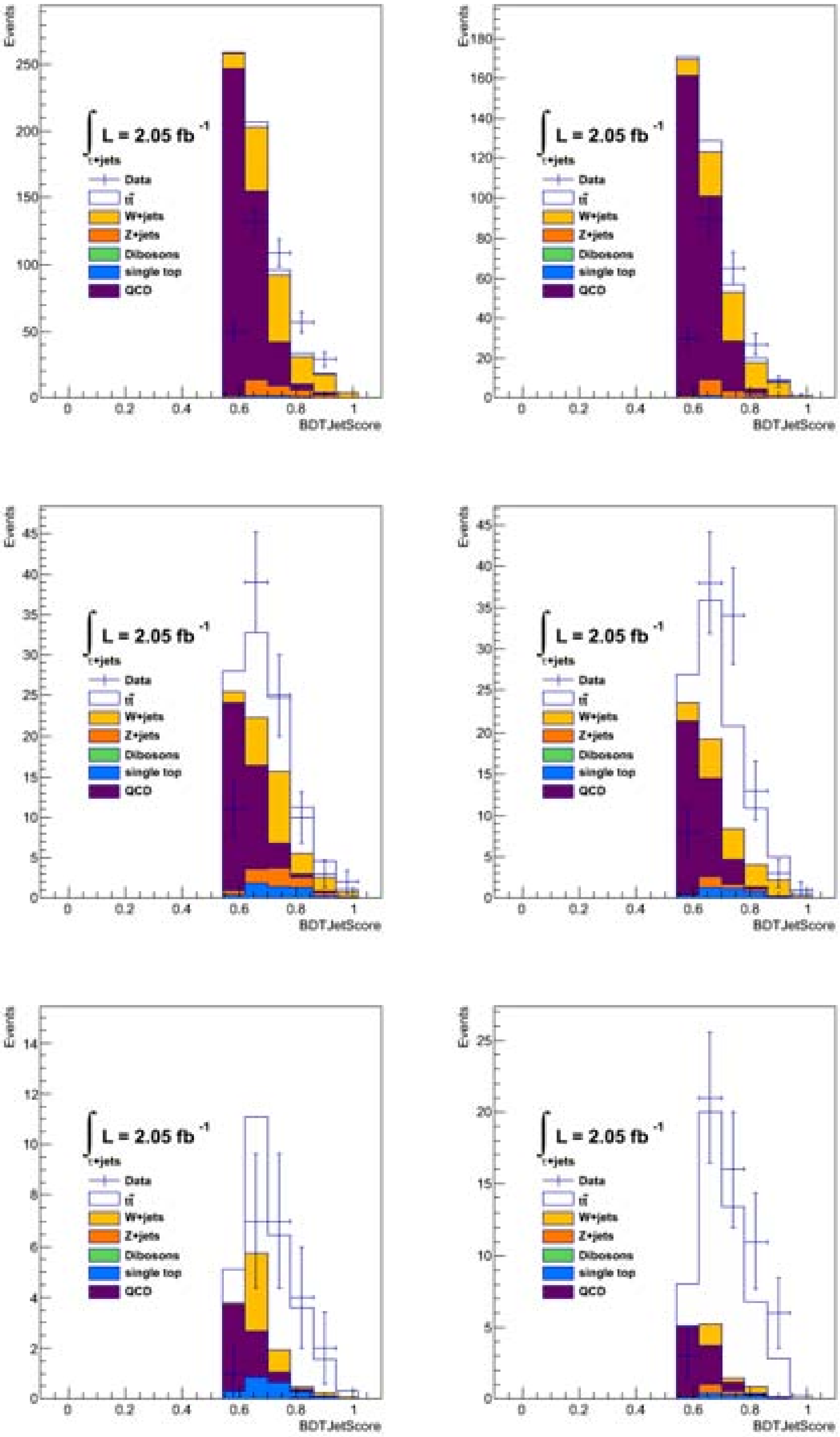


Figure 15: Data-MC comparison: BDTJetScore in the 3 prong tau channel.

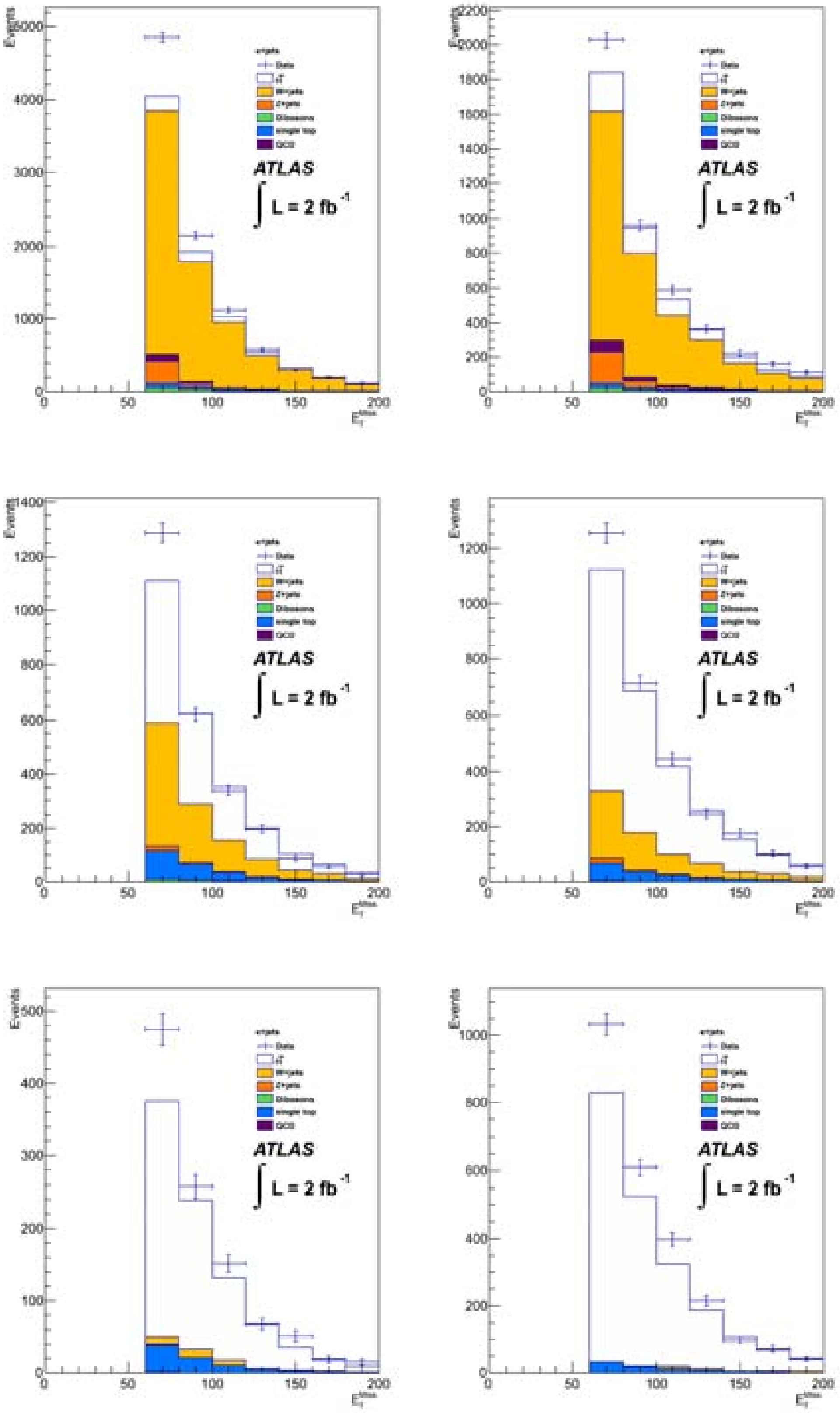


Figure 16: Data-MC comparison: missing transverse energy in the electron channel.

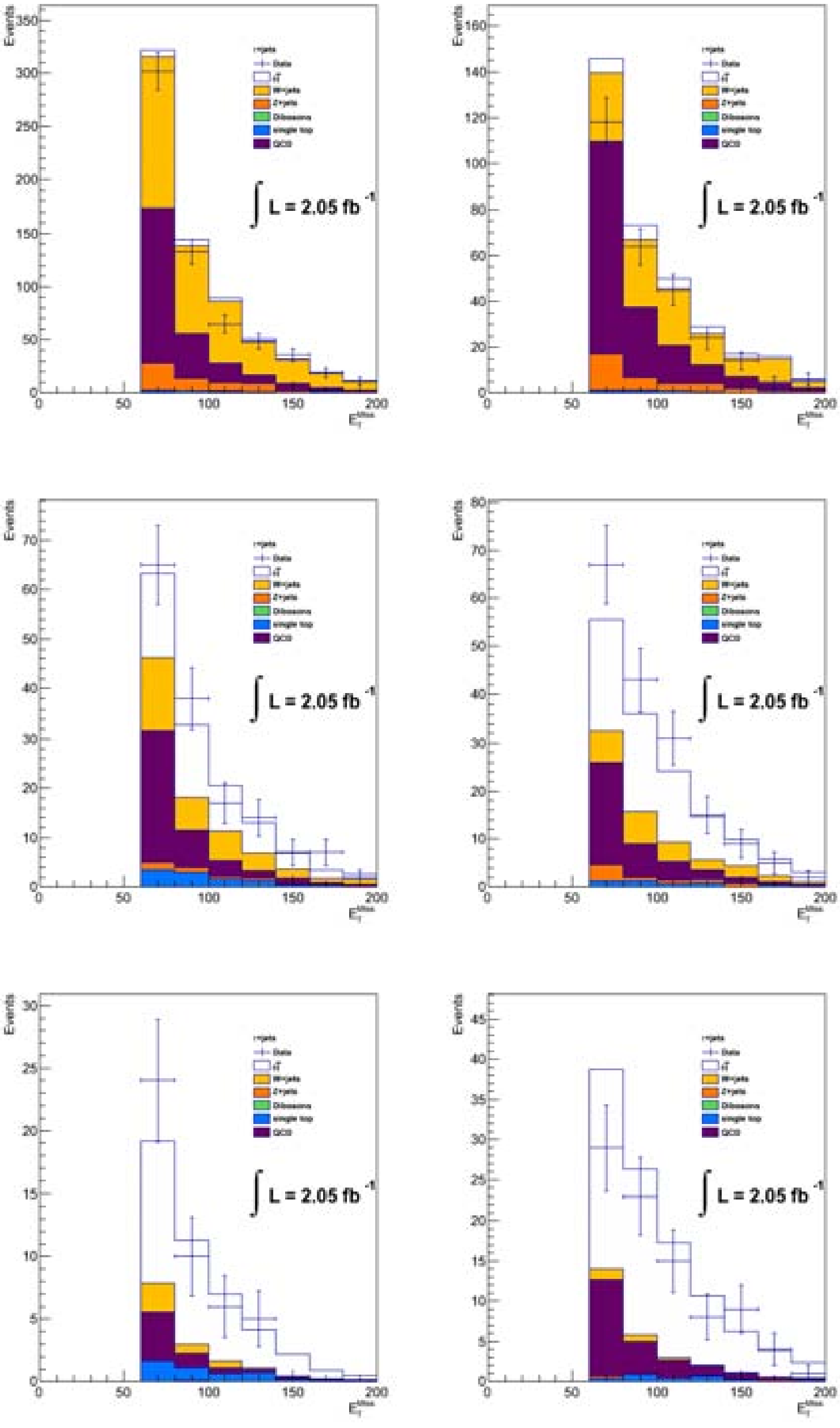


Figure 17: Data-MC comparison: missing transverse energy in the 1 prong tau channel.

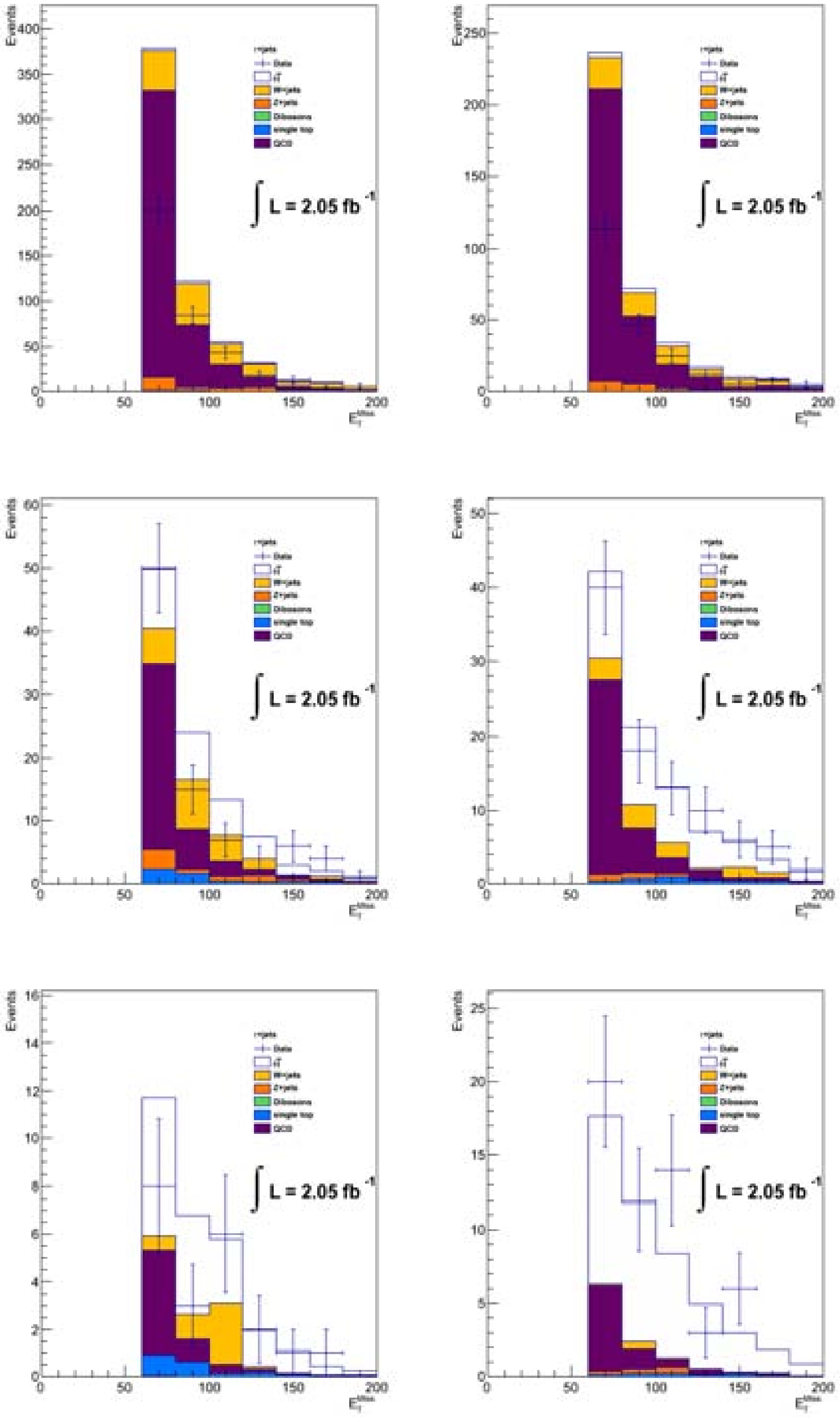


Figure 18: Data-MC comparison: missing transverse energy in the 3 prong tau channel.

B.1 Trigger Studies

In the hypothesis that the single tau trigger efficiency and the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency can be considered in first approximation uncorrelated, one can estimate and apply the efficiencies separately, assuming they factorize.

The inclusive tau trigger `EF_tau29_medium` efficiency has been measured in $Z \to \tau\tau$ enriched samples in the context of the cross-section measurement through tag-and-probe methods [48]. Using the ATLAS TauTriggerCorrections tool, based on the efficiencies obtained with these studies, we extract the scale factor related to `EF_tau29_medium` trigger efficiency and its *up* and *down* variation as a function of tau p_{T}. This scale factor has been applied to each MC simulated sample and the $\pm 1\sigma$ varied templates passed to the fit to profile the uncertainty.

To estimate the $E_{\mathrm{T}}^{\mathrm{miss}}$ component of the trigger, we performed a measurement of the trigger efficiency with data. To do this study we applied the 1-prong tau selection (except for the electron veto which is not applied) to a sample that fulfilled the single electron trigger (EF_e20_medium or EF_e22_medium, depending on the data period) both in data (Egamma stream) and MC. This is a high statistic sample with negligible QCD background. This sample is dominated by true electrons. On the other hand, this measurement is not sensitive to aspects of the hadronic tau decays, like shower shape, etc.: for this reason we considered it as a measure of the efficiency of the part of the trigger relative to the $E_{\mathrm{T}}^{\mathrm{miss}}$ cut. The electrons that pass the 1-prong tau selection apart from the BDT based electron veto, are fake taus neither suppressed by kinematic cuts nor by the tau identification criteria. Therefore, to apply the electron trigger ensure us to deal with lepton-like objects from semi-leptonic $t\bar{t}$ decay that in some case would pass the tau selection. Then, we select the events we are interested on requiring in addition the trigger used in the baseline selection of this analysis: the `EF_tau29_medium_xe35_noMu` trigger. The ratio between the distributions obtained in data and Monte Carlo provides an estimation of the efficiency of the $E_{\mathrm{T}}^{\mathrm{miss}}$component of the $\tau + E_{\mathrm{T}}^{\mathrm{miss}}$ trigger. In Figure B.1.1 is shown the ratio of the trigger efficiency between data and the sum of the Monte Carlo samples in function of $E_{\mathrm{T}}^{\mathrm{miss}}$, `BDTEleScore` and tau p_{T}. The linear fit of the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution provides the scale factor that account for the efficiency of the $E_{\mathrm{T}}^{\mathrm{miss}}$ component.

According to the previous considerations, no significant dependence can be noticed neither for scale factor related to the tau in all the p_{T} range nor for the scale factor related to the $E_{\mathrm{T}}^{\mathrm{miss}}$ part in all $E_{\mathrm{T}}^{\mathrm{miss}}$ range. This justify the use of a single scale factor as a global systematic to the nominal templates of the fit.

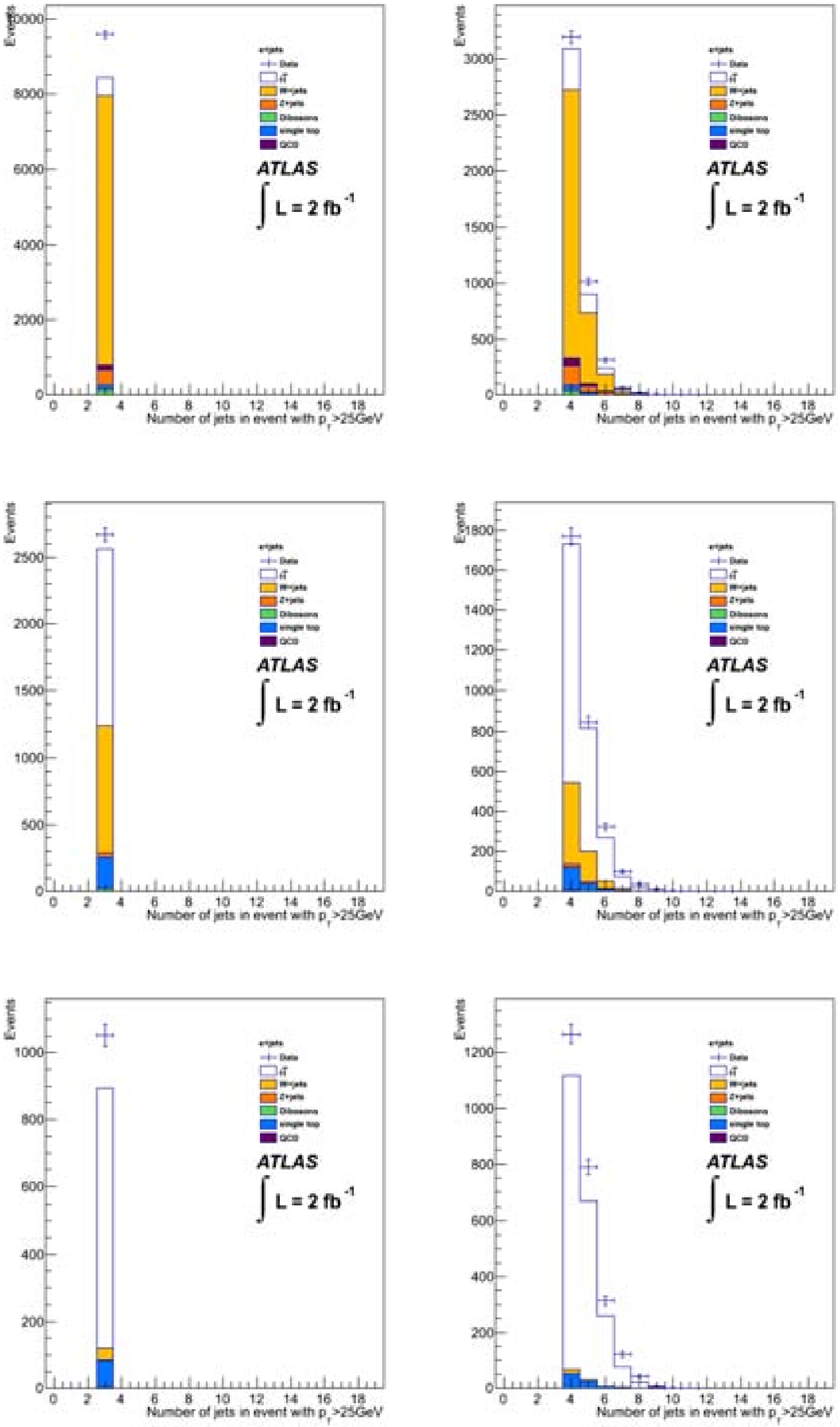


Figure 19: Data-MC comparison: number of jets with $p_{\mathrm{T}} > 25$ GeV per event, in the electron channel.

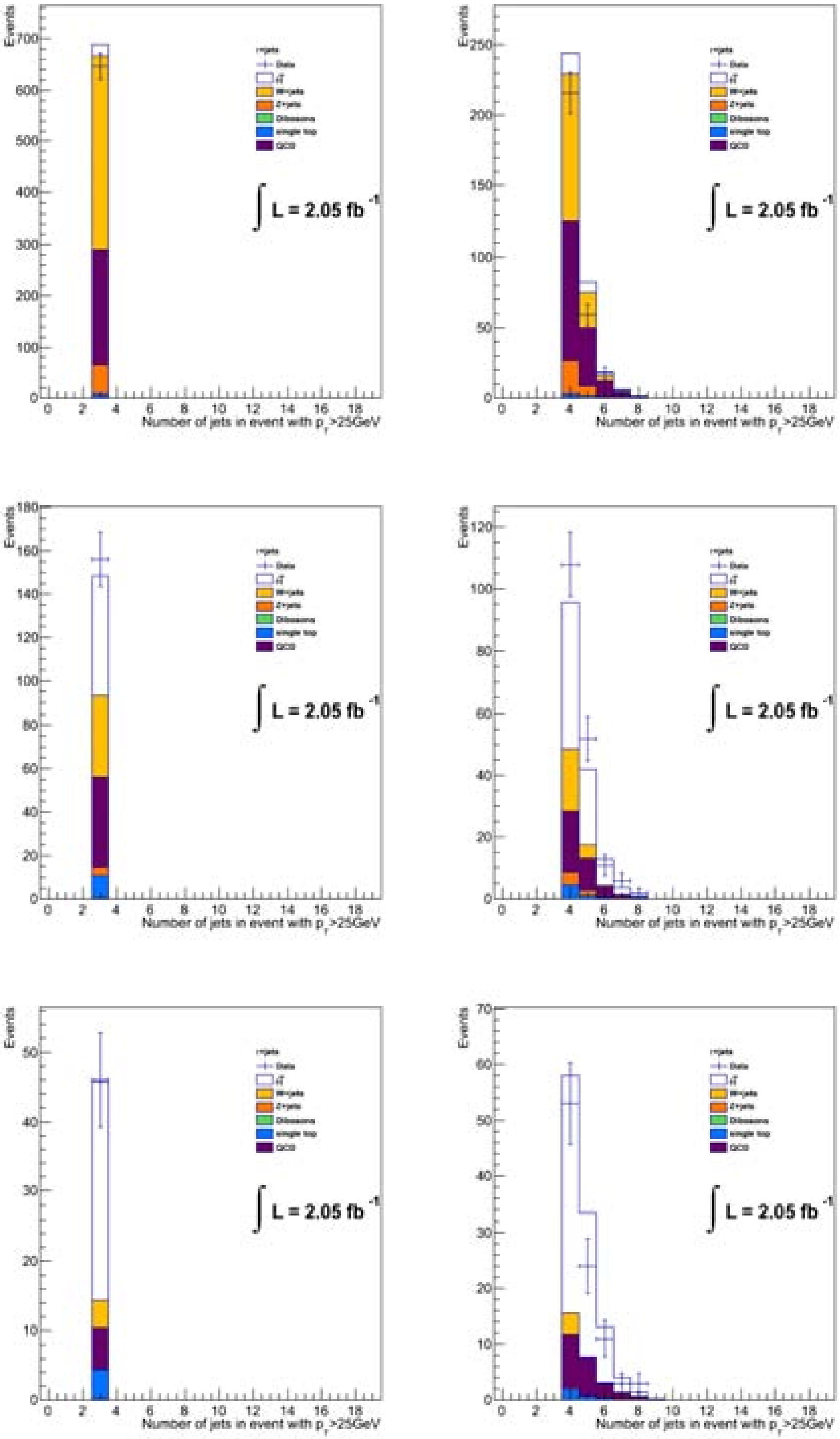


Figure 20: Data-MC comparison: number of jets with $p_{\mathrm{T}} > 25$ GeV per event, in the 1 prong tau channel.

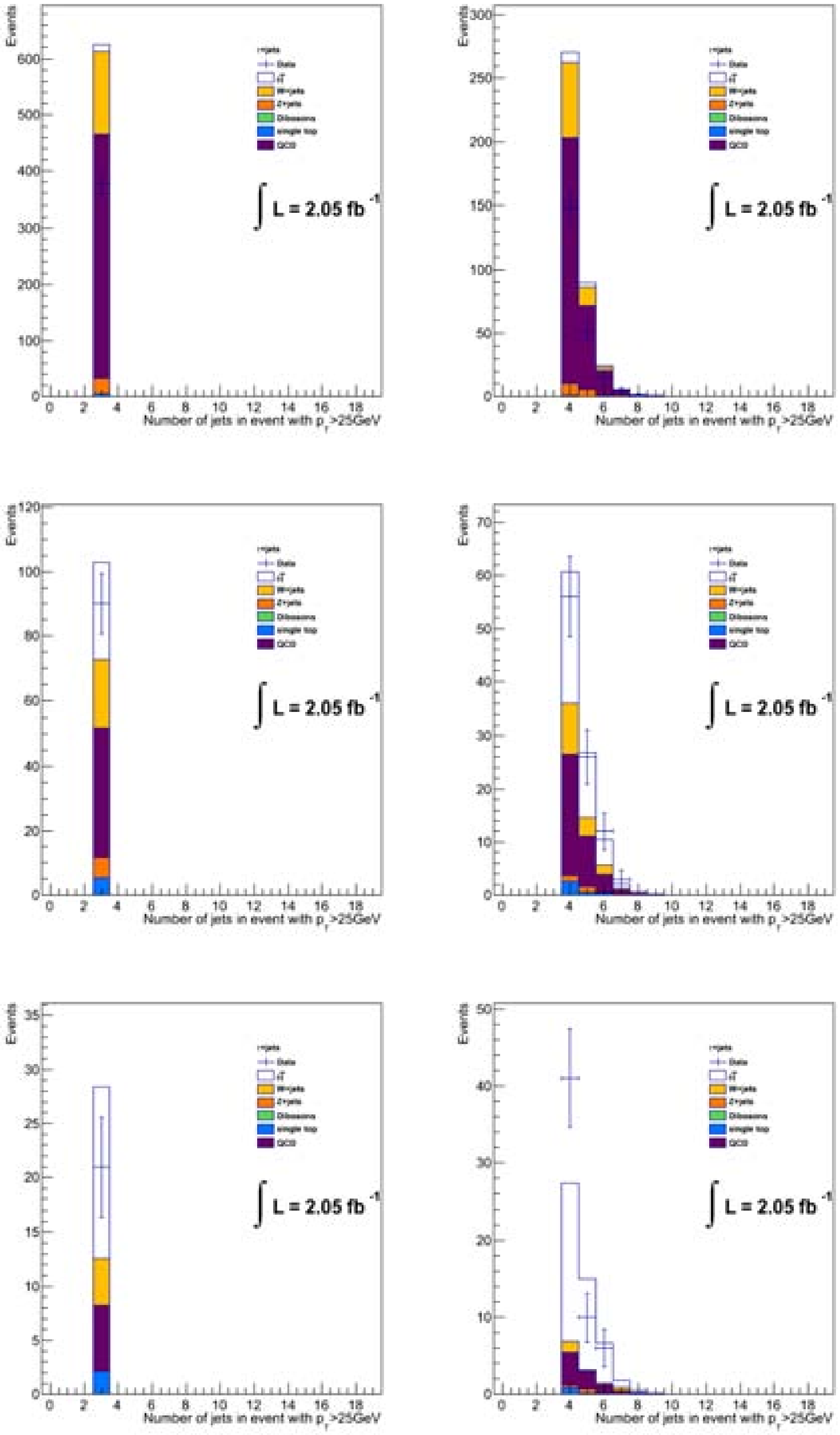
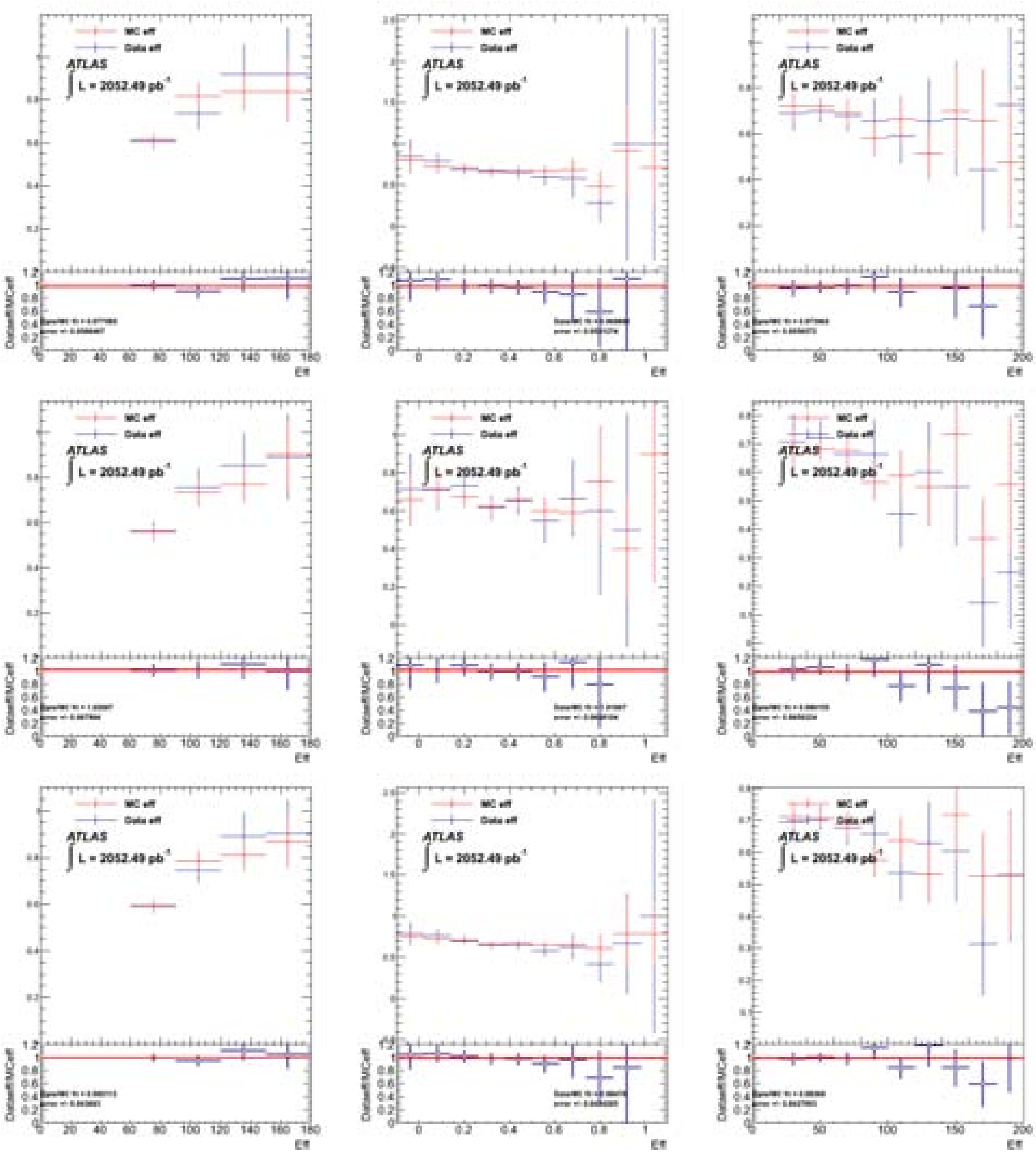


Figure 21: Data-MC comparison: number of jets with $p_{\mathrm{T}} > 25$ GeV per event, in the 3 prong tau channel.



$E_{\mathrm{T}}^{\mathrm{miss}}$ BDTEleScore tau p$_T$

Figure B.1.1: Ratio of the tau+MET trigger efficiency versus $E_{\mathrm{T}}^{\mathrm{miss}}$ (left), BDTEleScore (center) or tau p_T (right) between data and sum of MC in the 0 b-tag inclusive bin and for exactly 3 jets (top), at least 4 jets (medium) and at least 3 jets (bottom). No electron veto (BDTEleScore) applied; 1 prong contribution only is considered.

Appendix C

Shape comparison

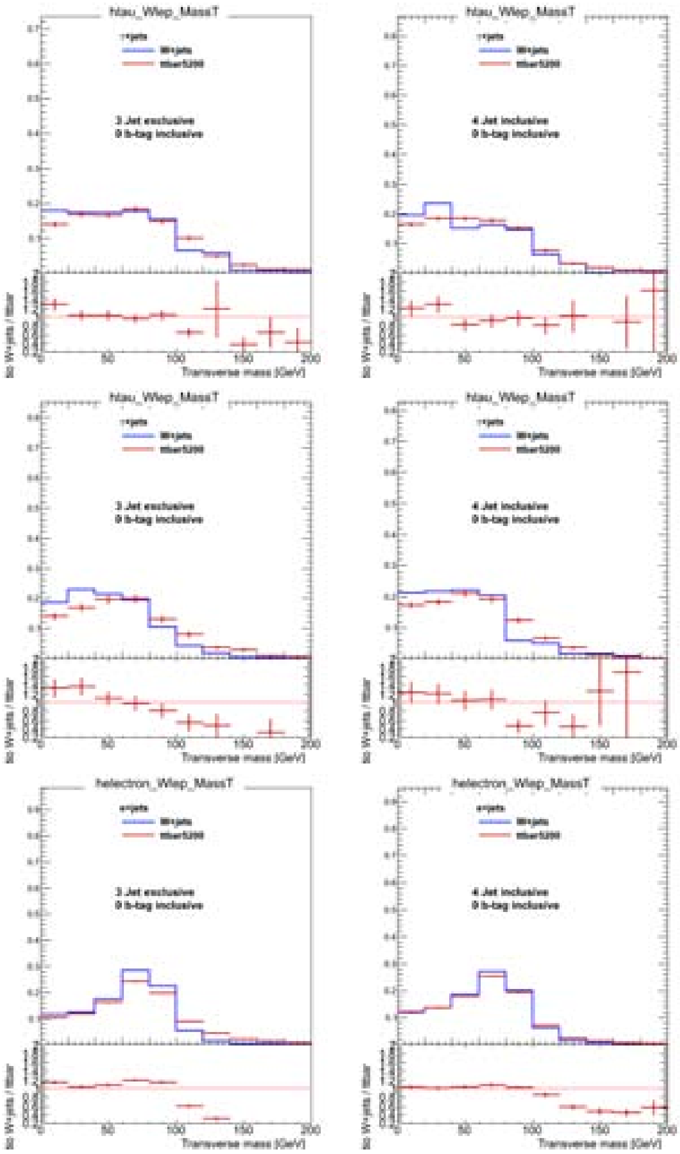


Figure 1: W m_T shape comparison between W+jets and $t\bar{t}$ in the untagged sample, for 1-prong (top), 3-prong (medium) and electron (bottom) channel. At left (right): 3 jet exclusive (4-jet inclusive).

Appendix D

QCD shape variation

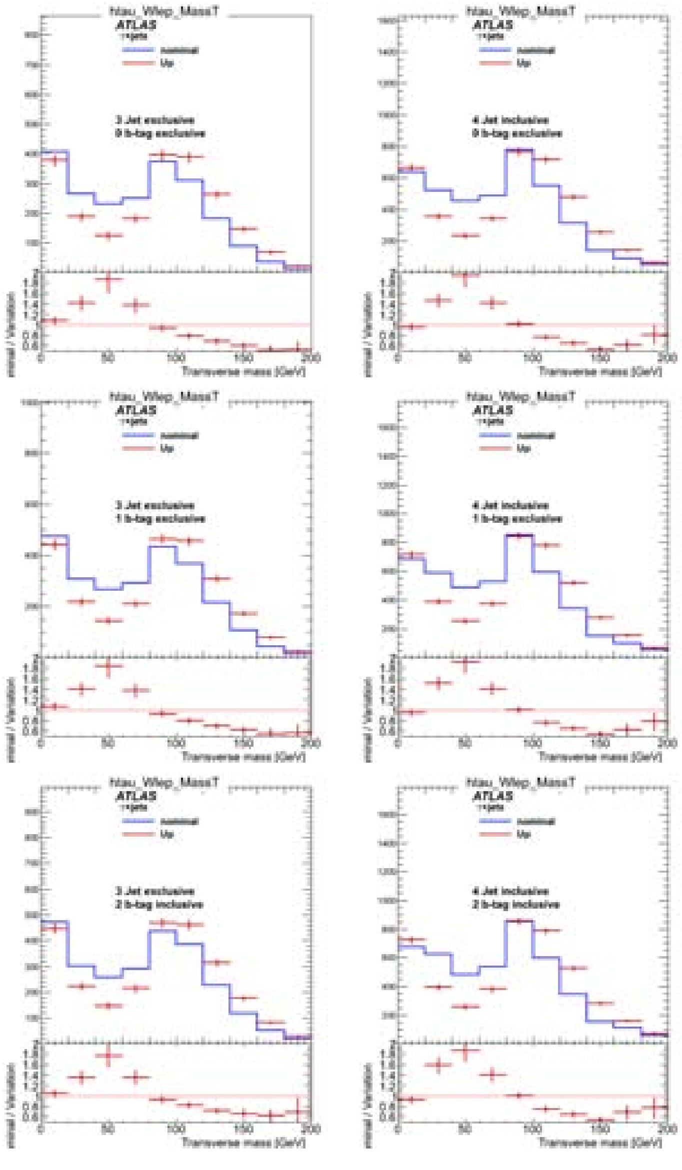


Figure 1: W m_T shape comparison between the nominal QCD template and the varied template based on Matrix Method, in the 1-prong channel. The variation is used as *a priori* uncertainty on the QCD shape in the fit. At left (right): 3 jet exclusive (4-jet inclusive). From top to bottom: 0 b-tag exclusive; 1 b-tag exclusive and 2 b-tag inclusive region of the selected events.

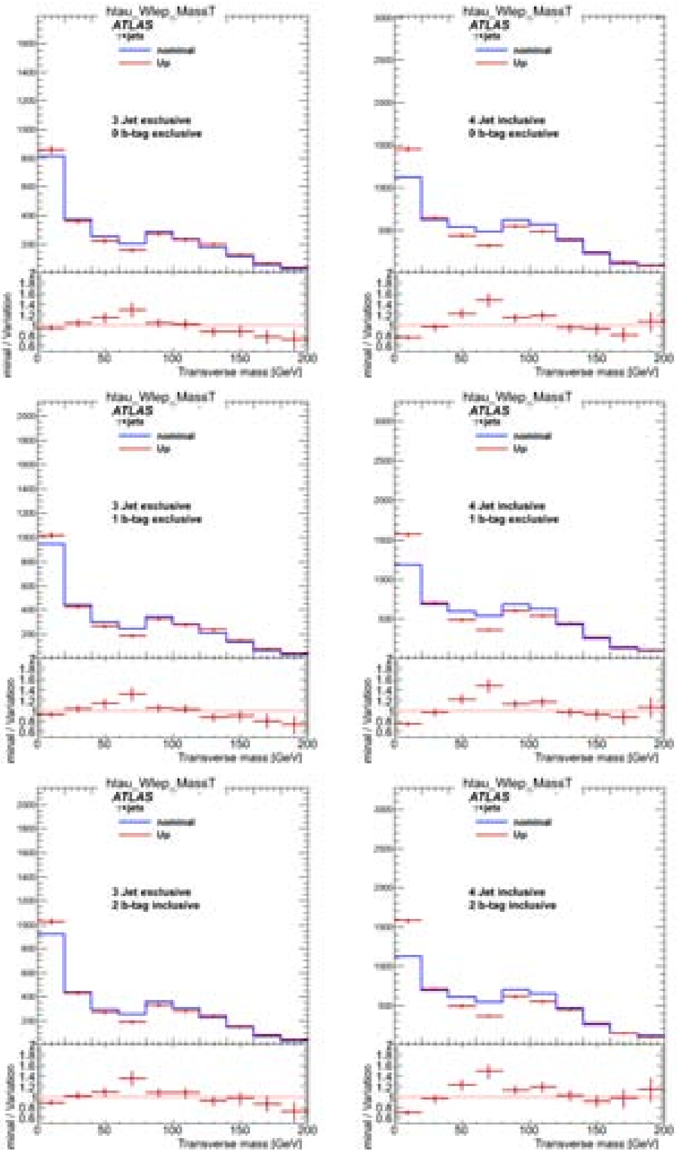
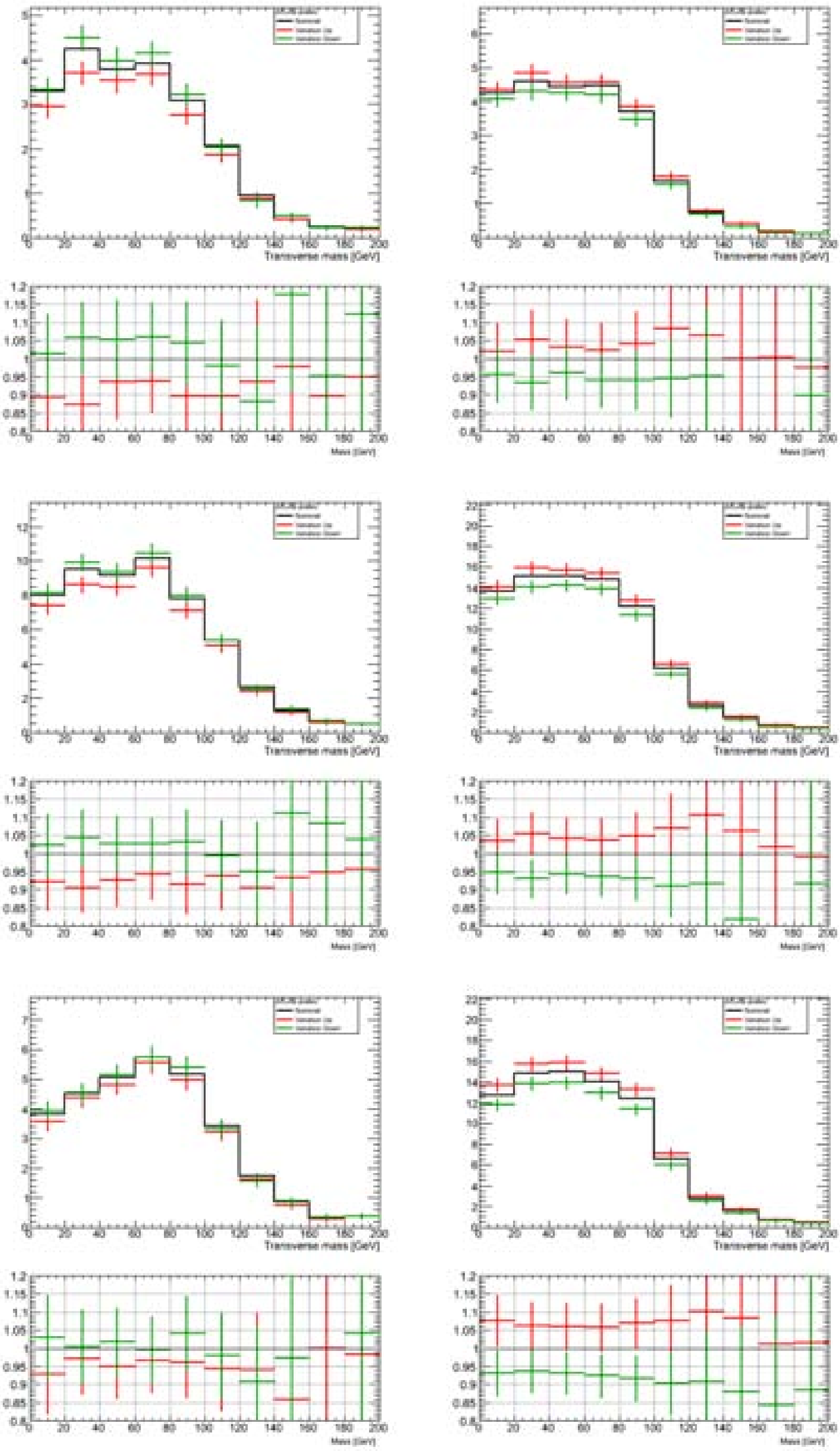


Figure 2: W m$_T$ shape comparison between the nominal QCD template and the varied template based on Matrix Method, in the 3-prong channel. The variation is used as *a priori* uncertainty on the QCD shape in the fit. At left (right): 3 jet exclusive (4-jet inclusive). From top to bottom: 0 b-tag exclusive; 1 b-tag exclusive and 2 b-tag inclusive region of the selected events.

Appendix E

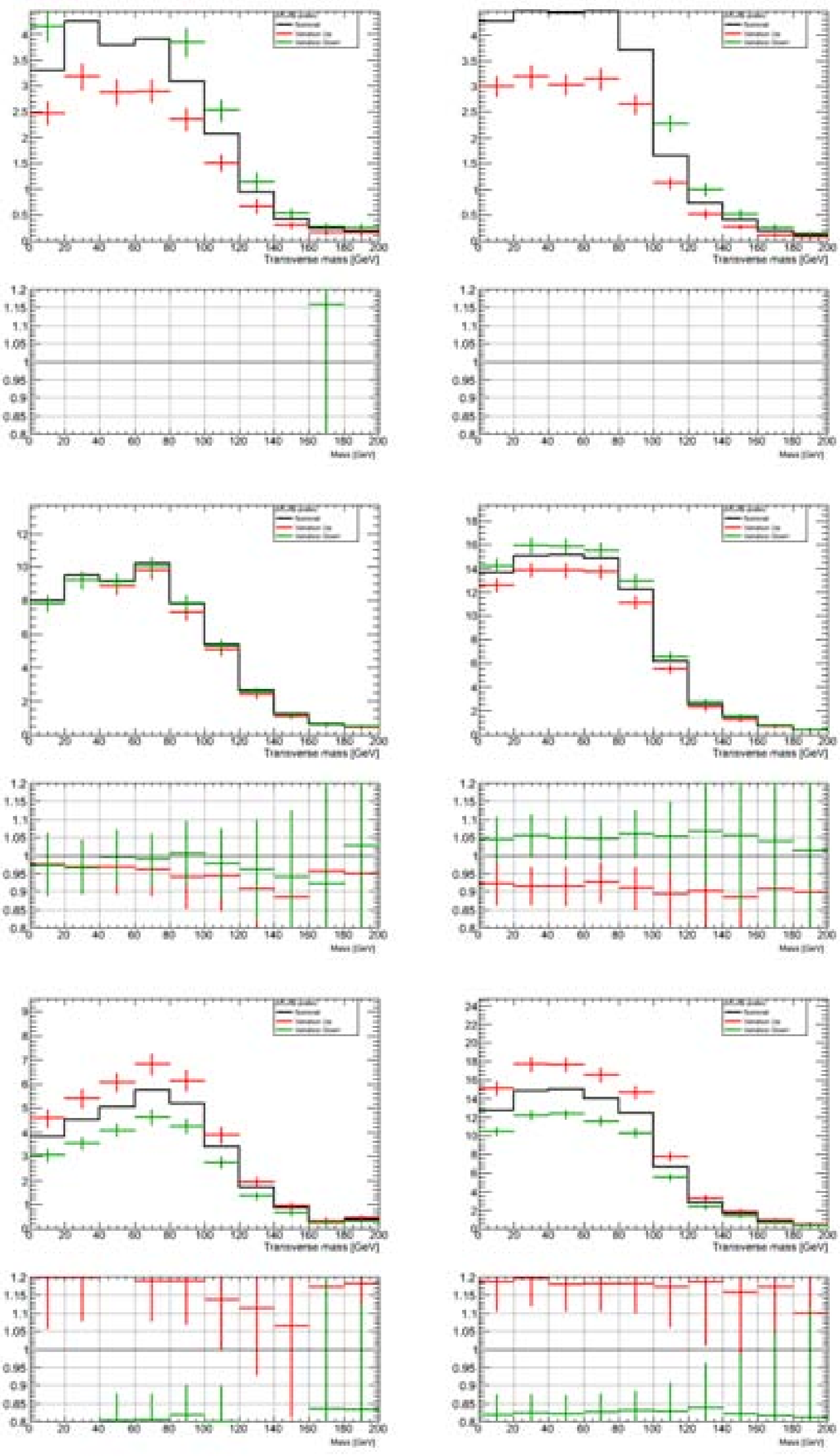
Systematic uncertainties validations

E.1 Systematic variations in $t\bar{t}$ with respect to the nominal value in 1-prong channel



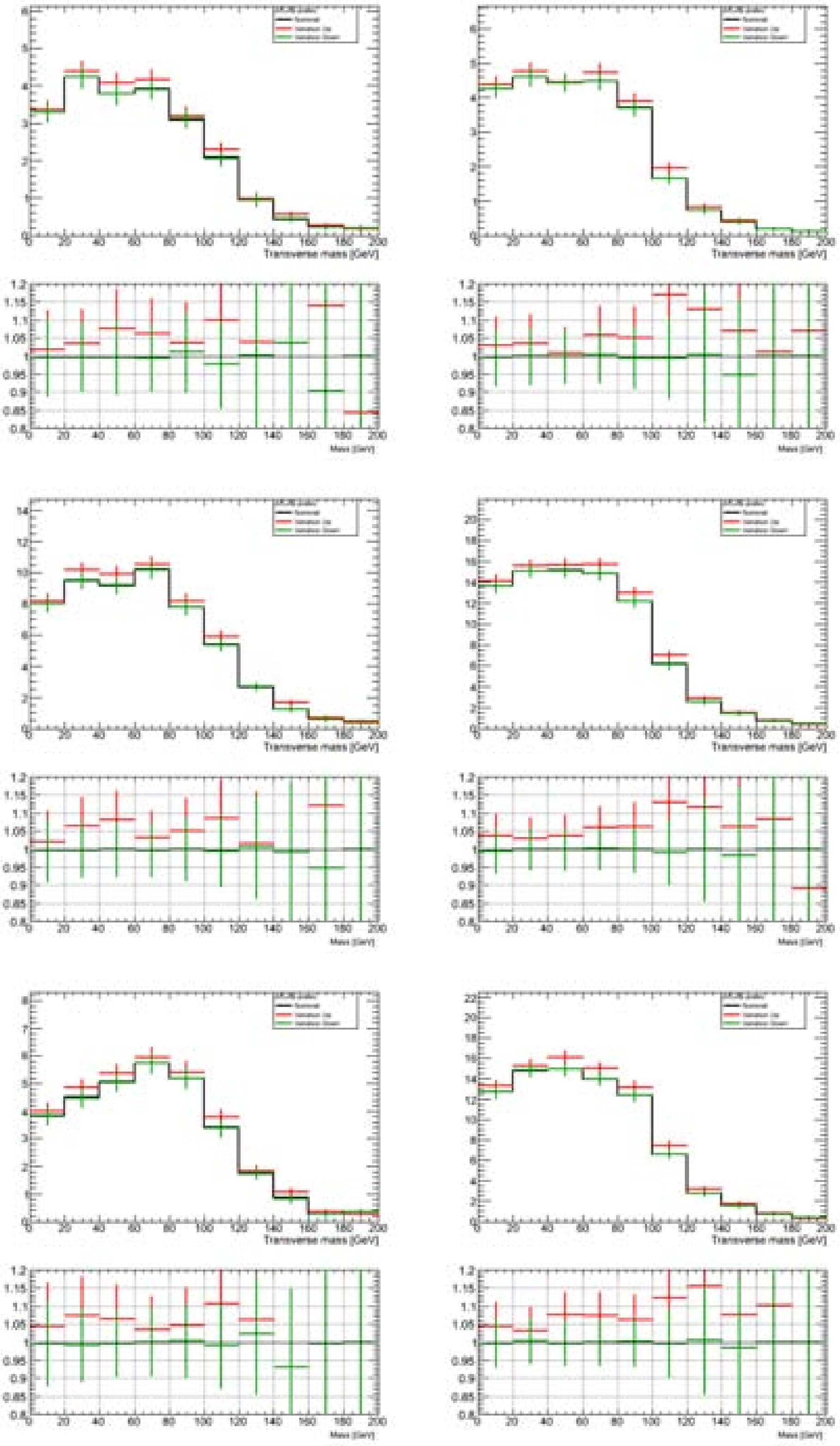
3-jet ≥ 4-jet

Figure E.1.1: Up and down variations from the nominal value due to JES systematics uncertainty in the W transverse mass in the 0- (top), 1- (medium) and ≥ 2- (bottom) *b*-jet multiplicity bins in 1 prong tau channel, using ttbar MC sample.



3-jet ≥ 4-jet

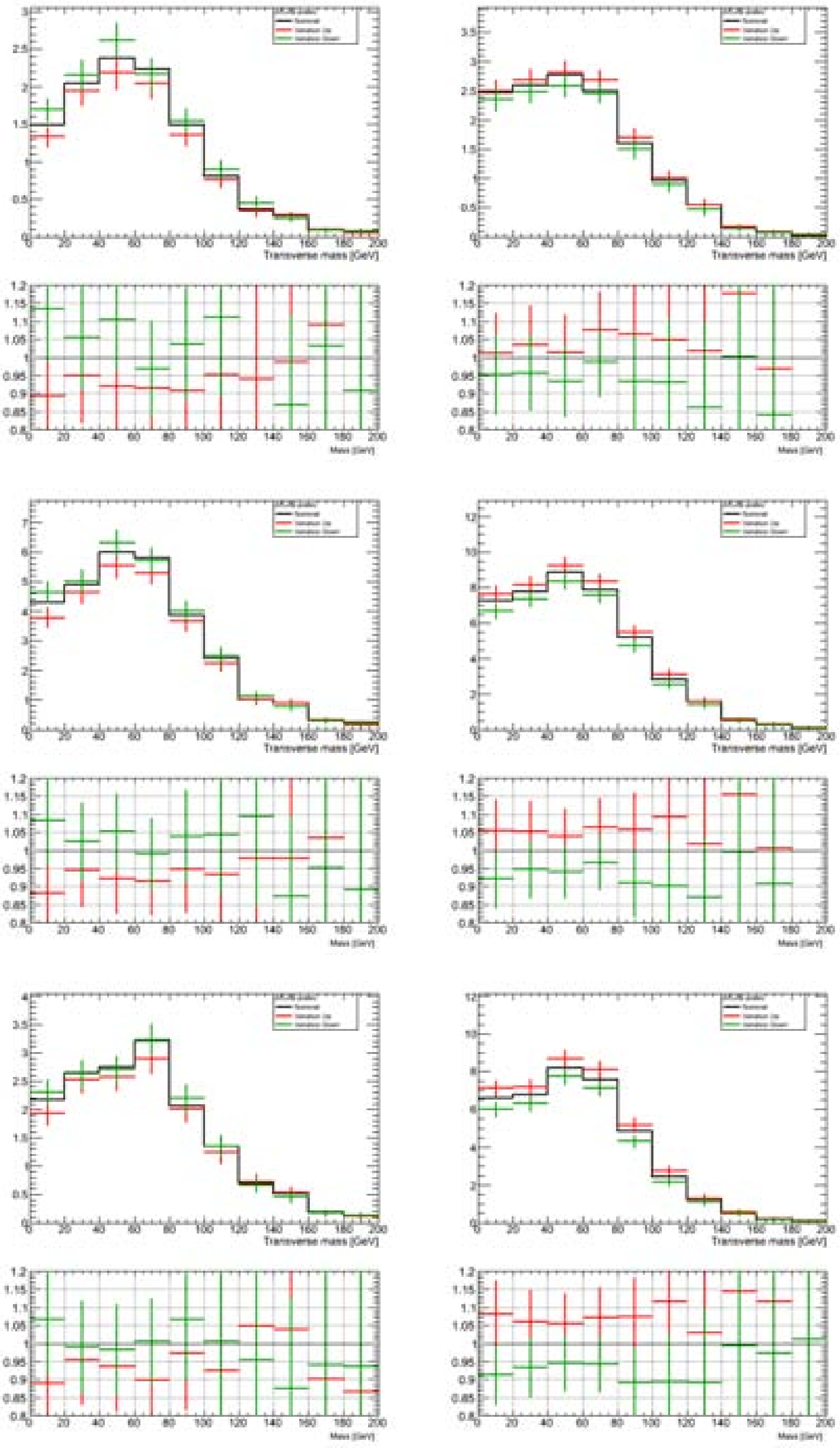
Figure E.1.2: Up and down variations from the nominal value due to BCtag systematics uncertainty in the W transverse mass in the 0- (top), 1- (medium) and ≥ 2- (bottom) b-jet multiplicity bins in 1 prong tau channel, using ttbar MC sample.



3-jet ≥ 4-jet

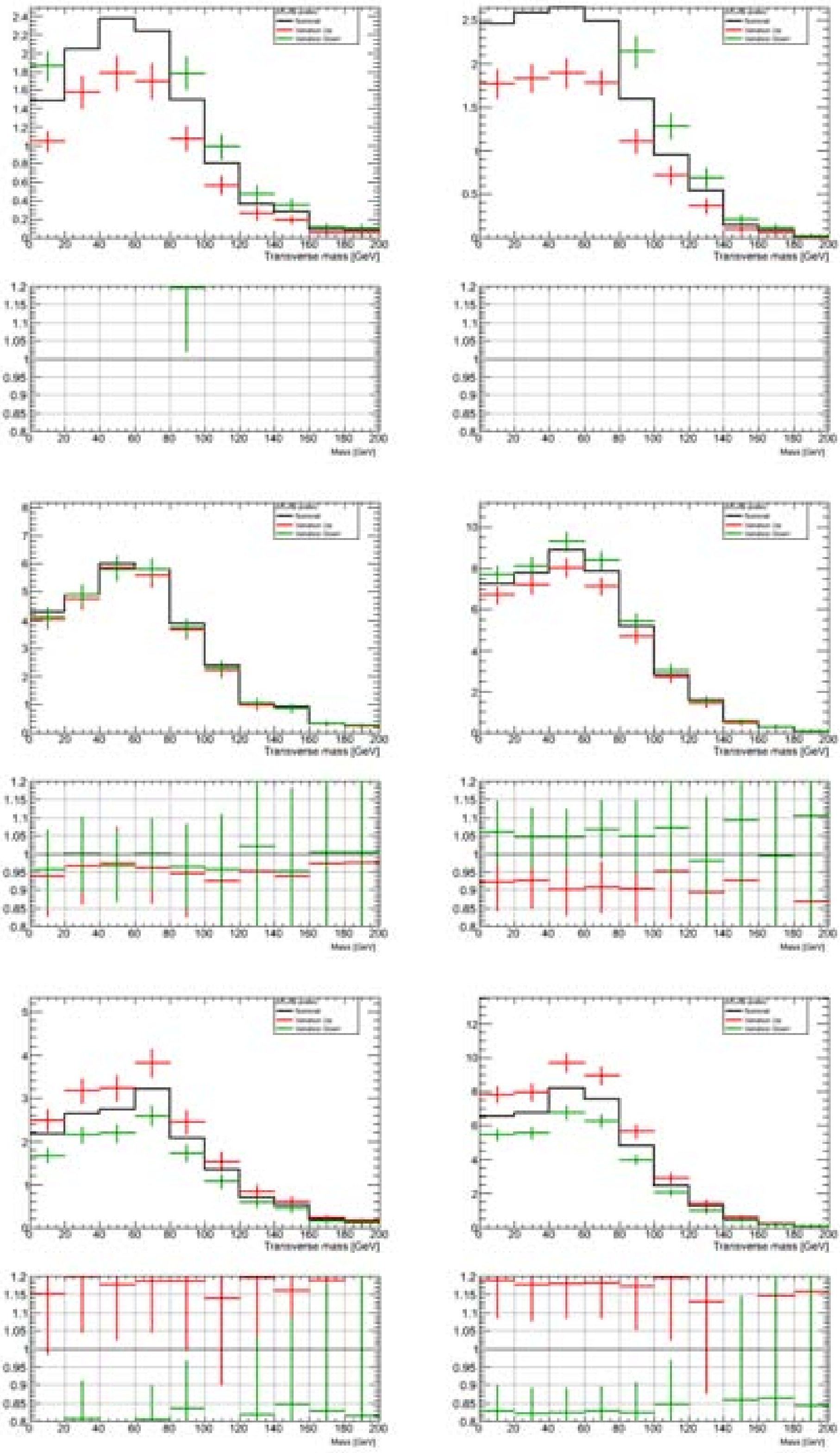
Figure E.1.3: Up and down variations from the nominal value due to TES systematics uncertainty in the W transverse mass in the 0- (top), 1- (medium) and ≥ 2- (bottom) *b*-jet multiplicity bins in 1 prong tau channel, using ttbar MC sample.

E.2 Systematic variations in $t\bar{t}$ with respect to the nominal value in 3-prong channel



3-jet ≥ 4-jet

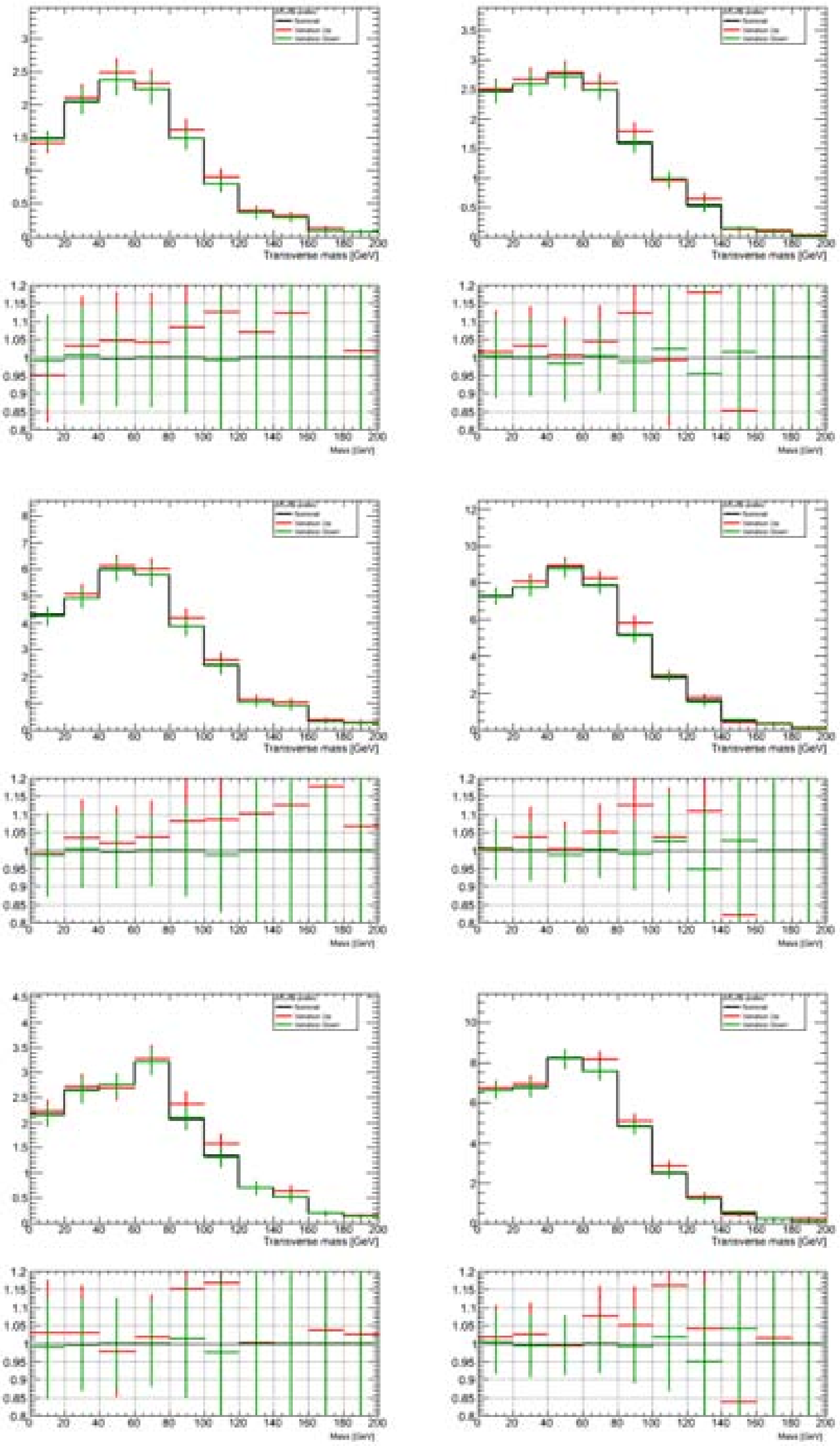
Figure E.2.1: Up and down variations from the nominal value due to JES systematics uncertainty in the W transverse mass in the 0- (top), 1- (medium) and ≥ 2- (bottom) *b*-jet multiplicity bins in 3 prongs tau channel, using ttbar MC sample.



3-jet

≥ 4-jet

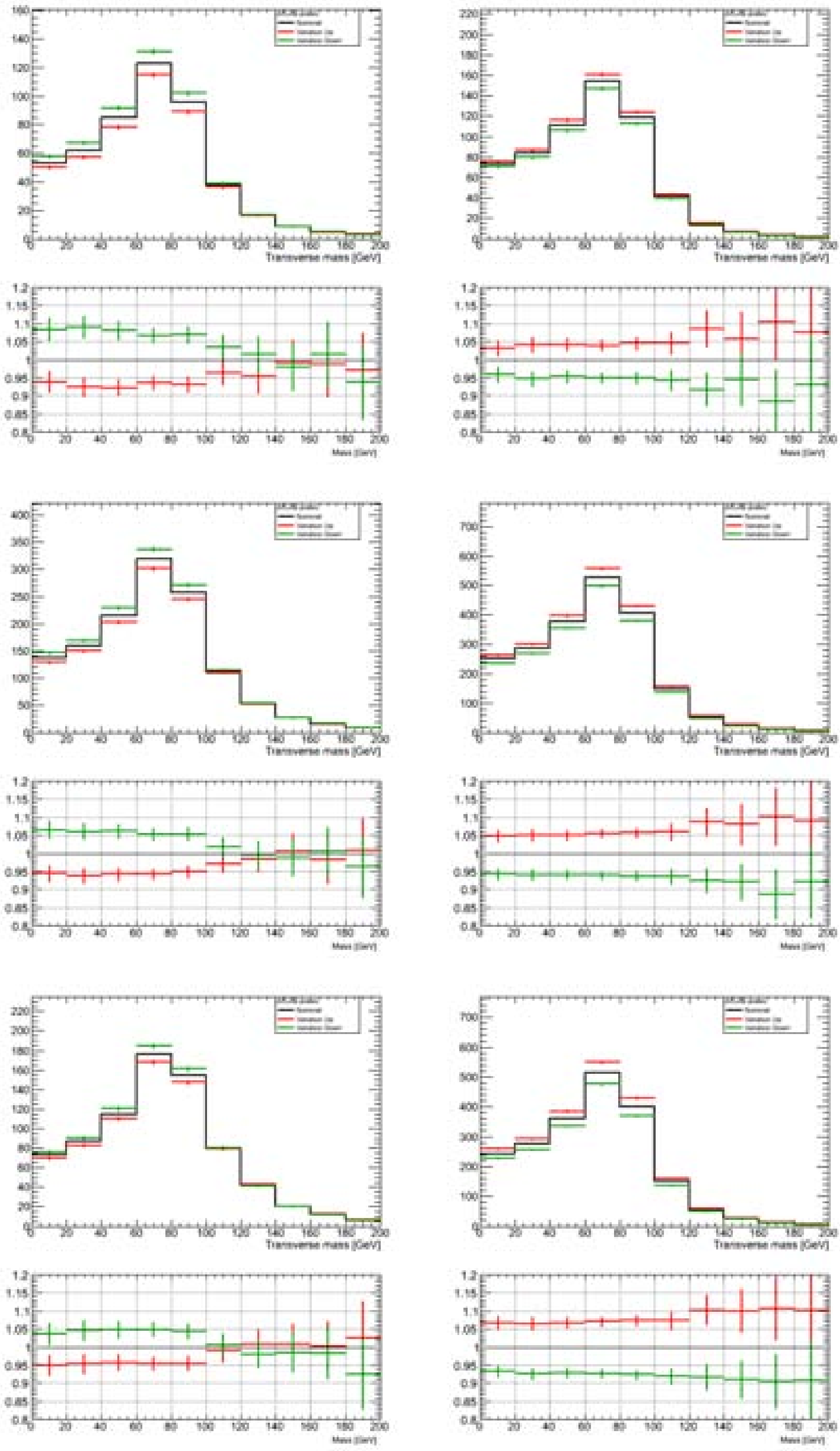
Figure E.2.2: Up and down variations from the nominal value due to BCtag systematics uncertainty in the W transverse mass in the 0- (top), 1- (medium) and ≥ 2- (bottom) *b*-jet multiplicity bins in 3 prongs tau channel, using ttbar MC sample.



3-jet ≥ 4-jet

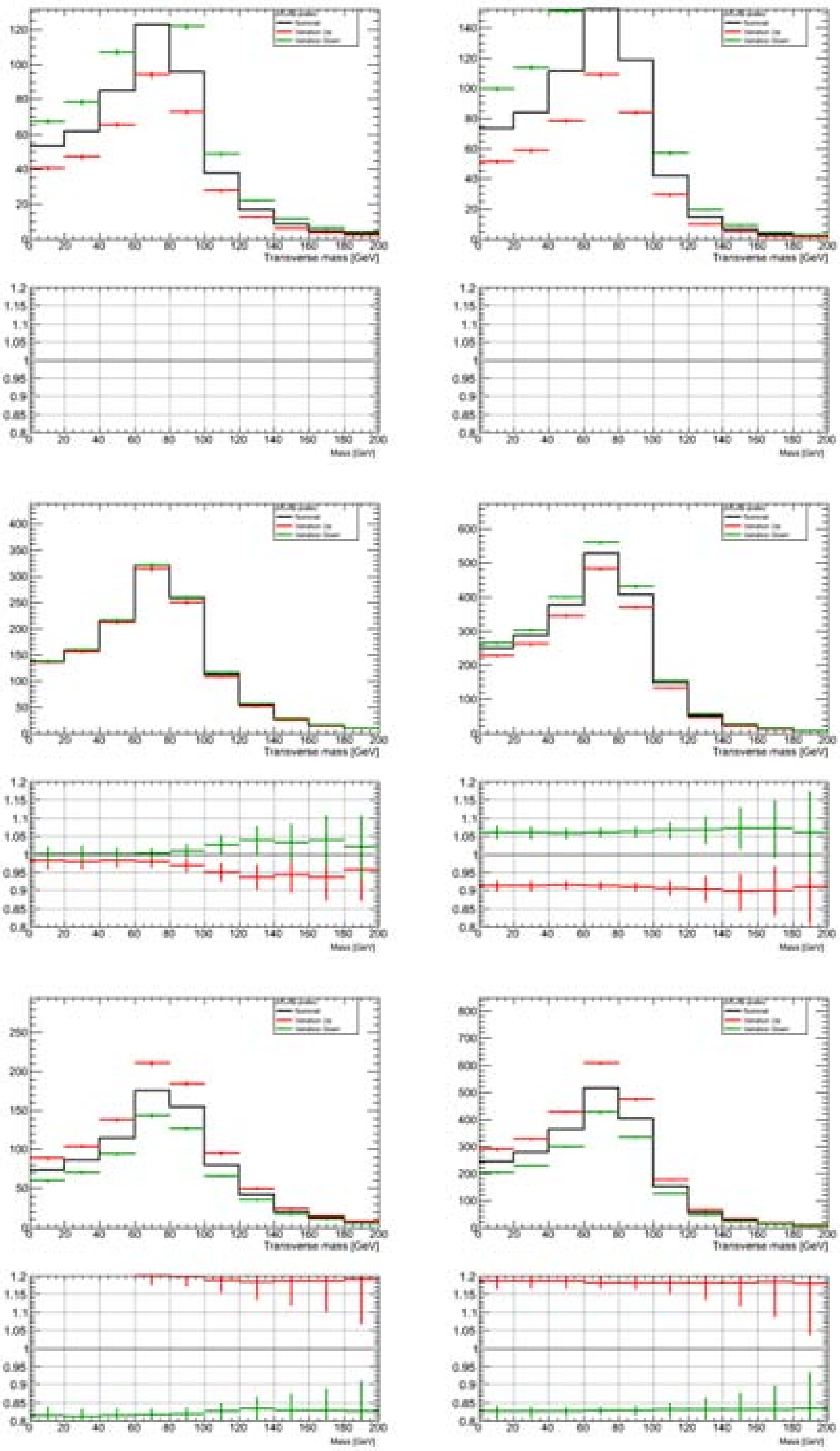
Figure E.2.3: Up and down variations from the nominal value due to TES systematics uncertainty in the W transverse mass in the 0- (top), 1- (medium) and ≥ 2- (bottom) *b*-jet multiplicity bins in 3 prongs tau channel, using ttbar MC sample.

E.3 Systematic variations in $t\bar{t}$ with respect to the nominal value in electron channel



3-jet ≥ 4-jet

Figure E.3.1: Up and down variations from the nominal value due to JES systematics uncertainty in the W transverse mass in the 0- (top), 1- (medium) and ≥ 2- (bottom) *b*-jet multiplicity bins in electron channel, using ttbar MC sample.



3-jet $\geq$ 4-jet

Figure E.3.2: Up and down variations from the nominal value due to BCtag systematics uncertainty in the W transverse mass in the 0- (top), 1- (medium) and $\geq$ 2- (bottom) *b*-jet multiplicity bins in electron channel, using ttbar MC sample.