$Z \to \tau_h\tau_l$ Cross Section Measurement and $\tau\tau$ Mass Reconstruction with the ATLAS Detector at the LHC

Dissertation

zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades
„Doctor rerum naturalium “
der Georg-August-Universität Göttingen

vorgelegt von
Despoina Evangelakou
aus Athen, Griechenland

Göttingen, 2012

Referent: Prof. Dr. Arnulf Quadt
Korreferentin: Prof. Dr. Ariane Frey

Tag der mündlichen Prüfung: 18.07.2012

$Z \to \tau_h\tau_l$ Cross Section Measurement and $\tau\tau$ Mass Reconstruction with the ATLAS Detector at the LHC

by
Despoina Evangelakou

In this dissertation, the $Z \to \tau\tau$ decay where one τ decays leptonically and the other one hadronically is presented. The process is studied with data collected from the ATLAS experiment in pp collisions at $\sqrt{s} = 7$ TeV at the Large Hadron Collider (LHC). A first observation of the process and an initial measurement of its cross section is done with 36 $\text{pb}^{-1}$of data recorded during the 2010 run. The measured cross section value in the mass region 66-116 GeV is $857.6 \pm 84.3(\text{stat}) \pm 124.7(\text{syst}) \pm 29.2(\text{lumi}) \pm 2.8(\text{theo})$ pb for the muon channel and $1142 \pm 138.6(\text{stat}) \pm 197.7(\text{syst}) \pm 38.9(\text{lumi}) \pm 2.6(\text{theo})$ pb for the electron channel. The measurement is updated with 1.55-1.34 $\text{fb}^{-1}$of data taken in 2011 and is $912.4 \pm 15.0(\text{stat}) \pm 94.7(\text{syst}) \pm 33.7(\text{lumi})$ pb for the muon channel and $998.1 \pm 23.7(\text{stat}) \pm 131.9(\text{syst}) \pm 36.9(\text{lumi})$ pb for the electron channel. Finally, the clean $Z \to \tau\tau$ sample is used to study ditau ($\tau\tau$) mass reconstruction methods. The considered methods are effective mass, collinear approximation, missing mass calculator, 'true' transverse mass and bound mass. All methods are found to have advantages and disadvantages and their applicability depends on the needs of the analysis, such as whether it is for discovery or mass measurement purposes.

Post address:
Friedrich-Hund-Platz 1
37077 Göttingen
Germany

II.Physik-UniGö-Diss-2012/06
II. Physikalisches Institut
Georg-August-Universität Göttingen
July 2012

Contents

Chapter 1

Introduction

"All of us at some time have paused to wonder at our existence. As children we asked our parents embarrassing questions about where we came from and, in retrospect, probably received some embarrassing answers. In later years we may have asked this question in a more mature form, either in accepting or rejecting some form of religion. Scientists that dedicate themselves to pure research have never stopped asking this question." taken from [1].

Physicists, nowadays, are closer to having a scientific answer to this question more than ever before. Cosmological theories, like the *Big Bang*, are attempting to explain the beginning and the evolution of the universe. It is believed that after the Big Bang an inconceivable amount of energy concentrated in an infinitesimal volume was set free. The pressure this energy applied to itself forced the system to expand and particles appeared. These were interacting with each other at higher energies and more often than it would ever be possible in the universe nowadays. For a brief instance, this expansion was happening with an enormous speed, that allowed the universe to double in size. This period is called *inflation*, it lasted from 10^{-36} seconds to $10^{-33} - 10^{-32}$ seconds, and it allowed the universe to grow by a factor of 10^{75}! At the end of the inflation the particles had all become energy. This energy was not stable and it started collapsing back into matter. The matter, that is the elementary particles that we know today, interacted with each other and the produced energy forced the universe to expand, but with slower rate than during the inflation period. Through continuous and steady interactions, the planets and the galaxies that we now observe were created.

The estimated age of the universe is 15 billion years. However, although we have an idea of how it was created, we do not know what it is made of. Astronomical observations of the gravitational distortions of the space and measurements of the energy relic from the Big Bang with the Hubble telescope indicate a greater amount of matter and energy than the one initially estimated. It is believed that the known or visible matter accounts only for about 5% of the total matter/energy of the universe. Apart from the visible matter, there is the so-called *dark matter*, that interacts weakly with the visible matter and hence, it has not been observed, yet. This dark matter makes up 25% of the universe and according to some theories it might have been produced during the period of inflation. Finally, the largest part of the universe (70%) is considered to be filled with *dark energy*, an energy of yet unknown properties that is assumed to explain the accelerating expansion of the universe.

Parallel to the development of the cosmological theories, particle physics through theoretical models and experiments provides a unique chance to catch a glimpse into those early moments and verify or reject some of the ideas. The two fields are, eventually, merging in an effort to explain "where are we coming from". From the particle physics point of view, at high energies, like the ones that existed in the universe after the period of inflation, matter consists of

twelve elementary particles and their anti-particles that interact with each other via fundamental forces. The theory that summarises how these particles come about and how they interact is a mathematical model called the Standard Model of particle physics. It was developed in the early 1970s by Sheldon Glashow, Steven Weinberg and Abdus Salam and since then it has withstood the challenges of many experimental tests.

The matter particles of the Standard Model can be divided into two groups; *leptons* and *quarks*. The motivation for this classification comes from the way the constituents of the groups 'feel' the fundamental forces. There are six quarks: *up*, *down*, *charm*, *strange*, *bottom* and *top*, ordered in ascending mass values. The first two quarks, up and down, are the main comprising particles of the proton and the neutron, which form the nucleus of an atom and thus are the only ones naturally occurring today. The rest of the quarks can only be produced in laboratories. Nonetheless, that was not the case in the early universe when significantly more energy was available and the four heavier quarks were also present and were interacting.

There are, also, six leptons: *electron*, *electron neutrino*, *muon*, *muon neutrino*, *tau* (τ) and *tau neutrino*. Most people are familiar with the electron as a constituent of the atom and as carrier of the electric energy. The muon and the τ leptons are the heavier siblings of the electron; i.e. all of their intrinsic properties, such as charge and spin, are identical to those of the electron. The neutrinos are like little brothers to the corresponding charged leptons; they themselves have no charge and have very small masses, but they can still interact with the "heavier" brothers.

Up to this point, the word "interaction" has been mentioned many times. When it comes to humans it is quite clear what an interaction is, but what do physicists mean when they speak of interactions? The answer is the four fundamental forces of nature: gravity, electromagnetism, weak and strong. The gravitation and the electromagnetic forces are the most widely known ones. The first one is an attraction felt by all massive bodies and the second one, either attractive or repulsive, affects charged particles. The less known forces are the weak and the strong. The weak has the special capability of changing the type of the interacting particle. For example, one type of weak interaction is the change of an electron into an electron neutrino, or an up-quark to a down-quark. For the leptons this change can only happen between the charged particle and its neutral partner. On the other hand, the only constraint for the quarks is the presence of enough energy/mass for the lighter ones to transform into heavier ones. Finally, the strong force is experienced exclusively by the quarks. That is, the quarks inside the proton should be repelling or attracting each other due to the electromagnetic charge they carry, instead, due to the strong force, they are bound together. This property prevents us from observing single quarks in nature or experiments. Except for gravity, the other three forces play an important role in the interactions of the elementary particles, at least at the currently accessible energy range, and are described by the Standard Model.

The interactions between two particles happen with the exchange of a *mediator*. A mediator is a particle that "carries" the force, as for example in electromagnetism the photon, γ, mediates the interaction between two charged particles. The weak force has three such particles: two charged, $W^{\pm}$, and one neutral Z^0, and the strong force has the gluon, g. The mediators, or *bosons*, have completely different properties than the matter particles. In particular, the photon and the gluon are massless, while the $W^{\pm}$ and Z^0 are quite massive, $m_W \approx 80$ GeV and $m_Z \approx 91$ GeV . The most favourable explanation for why the W and Z bosons are massive, while the photons and gluons are massless, in the Standard Model is given via the introduction of one more particle, the *Higgs boson*. Interactions of leptons and quarks with the latter can also explain their masses. The discovery of the Higgs boson would be the final verification of the success of the Standard Model, since it is the only predicted particle that has not yet been

observed[1].

The experimental device that physicists use to look into the properties and the interactions of elementary particles is a particle accelerator. An accelerator brings particles to velocities close to the speed of light and then lets them collide. At the collision point a detector is positioned, which serves as a giant camera that captures snapshots of the collision and its evolution. The newest and most powerful accelerator is currently operating at CERN in Geneva, CH, namely the Large Hadron Collider (LHC). LHC is built not only for finding or excluding the Higgs boson, but also for searching for new physics beyond the Standard Model predictions. The detector used for obtaining the data for this work is the ATLAS detector.

The main focus of this research has been the Z boson, as it is an irreducible background for the search for Higgs boson production. The chosen decay channel is the one where the boson decays to two τ leptons and their neutrinos. The τ leptons, being the heaviest of the leptons, are more probable to couple with the Higgs boson and, hence, a potential observation of the latter has a higher probability. The challenge that is posed, though, is that the τ leptons have a very short lifetime and decay further to lighter leptons and quarks. In order to identify them a smart mechanism needs to be developed to separate them from primary particles, e.g. electrons and muons. Once the τ leptons are well identified, then the Z boson is reconstructed. A good knowledge of the latter will allow searches for the Higgs boson in the $\tau\tau$ decay channel.

This dissertation has been organised as follows: the Standard Model formalism and the necessity of the Higgs boson are explained in Chapter 2. Following, in Chapter 3, the experimental setup, i.e. the LHC and the ATLAS detector, is described. Chapter 4 is concerned with the tools used to merge the theoretical calculations and the experimental results, namely Monte Carlo simulations. The particle identification and reconstruction in ATLAS are the subject of Chapter 5. Chapter 6 is focusing on the analysis strategy followed, in order to observe Z bosons decaying to two τ leptons with the ATLAS detector. The analysis is done first with the data collected during the 2010 run. These results are statistically limited resulting in unwanted uncertainties; therefore, the analysis is repeated with the data collected during spring and summer 2011. The latter amount of data is sufficient for decreasing the statistical uncertainties, but the measurement is still dominated by systematical ones. The cross section of the Z is measured with both dataset, see Chapter 8. Furthermore, the larger statistical sample of the 2011 data run allowed for a study of different mass reconstruction methods of the $Z \to \tau\tau$ process, which is presented in detail in Chapter 9. The final results and some thoughts for future improvements are given in Chapter 10.

[1]Only a week after this thesis was submitted, CERN announced the observation of a new particle at 125 GeV that seems to have all the properties of the long-sought Higgs boson. For details see [2, 3].

Chapter 2

The Standard Model

The Standard Model of particle physics is the theoretical model that best describes the current understanding of the world of elementary particles. It incorporates the electroweak theory (EW) theory and Quantum Chromodynamics (QCD), within which all fundamental particles and interactions among them are described. In this chapter, the mathematical formulation of the theories that comprise the Standard Model is explained.

2.1 Particles and Force Mediators

The picture of the world of the very small has changed completely since the discovery of the first particle; the electron in 1897. According to the Standard Model there are twelve elementary particles and four types of force carrying particles. The first ones, also called matter particles, are divided into two categories; leptons and quarks. The matter particles carry half-integer spin and hence they follow Fermi-Dirac statistics, giving them the name fermions. The force carriers, on the other hand, have an integer spin and obey Bose-Einstein statistics, hence called bosons.

A Table of the fermions and bosons grouped per type and per generation is shown in Fig. 2.1. The fermions are further divided into three generations. Particles of the same generation share certain properties and they have different masses. There are six leptons; the electron (e) and its neutrino (ν_{e}), the muon (μ) and its neutrino (ν_μ) and the tau (τ) and its neutrino (ν_τ). The leptons and their electromagnetic charge and leptonic number are shown in Table 2.1. The quarks are six, too, and are similarly divided into three generations. The quarks of the subsequent generation are heavier than the antecedent ones. The quark flavours are: up (u), down (d), charm (c), strange (s), top (t) and bottom (b). These are also shown in Table 2.1.

The great success of the Standard Model is that all predicted particles have been experimentally observed with the exact properties predicted. The last particles to be discovered were the top quark, being the heaviest of the fermions, and the ν_τ, being the most difficult to detect. The top quark was observed at the TeVatron at Fermilab, USA, in 1995 and the ν_τ was seen by the DONUT collaboration (Direct Observation of the NU Tau) at Fermilab as well, in the year 2000.

The elementary particles interact with each other through three fundamental forces; the electromagnetic, the weak and the strong force[1]. For each force there is a mediator: the photon (γ) for the electromagnetic, the $W^\pm$ and the Z^0 for the weak and the gluon (g) for the strong

[1]In principle, there are four fundamental forces: electromagnetism, weak, strong and gravity. But the last one is very small compared to the other interactions at the current experimental energies. If one goes to energies up to the Planck scale ($\sim 10^{19}$ GeV) then gravity is quantised and plays a significant role. Nonetheless, there is no theory yet that describes satisfactorily interactions at such high energies.

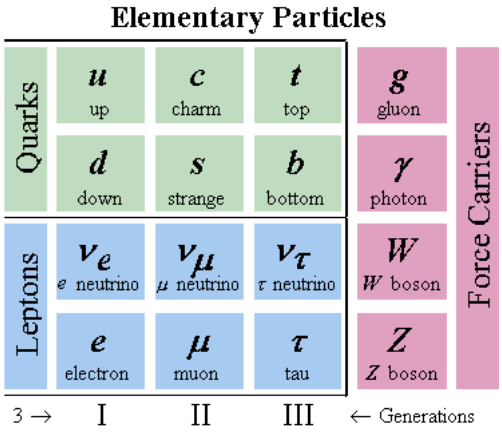


Figure 2.1: Table of the Standard Model particles.

Table 2.1: Most important quantum numbers of leptons and quarks. At the top half of the table the quantum numbers of the leptons are shown, while at the bottom half the ones for quarks.

fermion generations	1	2	3	EM charge
leptons	e	μ	τ	-1
	ν_e	ν_μ	ν_τ	0
lepton nr	$L_e = 1$	$L_\mu = 1$	$L_\tau = 1$	
quarks	u	c	t	+2/3
	d	s	b	-1/3

force. Not all bosons have the same intrinsic properties, however they are all characterised by the same spin value, spin 1. For example, photons and gluons are massless, but the W and the Z bosons are exceptionally massive

$$m_W = 80.385 \pm 0.015 \text{ GeV} \quad \text{and} \quad m_Z = 91.1876 \pm 0.0021 \text{ GeV}$$

[4]. The masses of the W and Z bosons come about from the interaction with a postulated field known as the *Higgs field*. The Higgs field gives rise to a massive scalar particle with spin 0, the *Higgs boson*. The Higgs field and boson are explained in detail in Section 2.2.3. The W boson carries electromagnetic charge, while the photon, the Z boson and the gluon are neutral. Nonetheless the gluons carry a different kind of charge, colour charge. Colour is only carried by quarks and gluons and it is a property of the strong interactions. The fact that gluons interact not only with quarks but also with each other adds complexity to the description of the strong force. The significance of this last property of gluons will become more obvious in Section 2.2.1.

2.2 Field Theories

The particles are described within the Standard Model as field quanta, i.e. as excited states of an underlying field. Global symmetry transformations give rise to conserved quantities, according to Noether's theorem. By requiring that the theory is invariant under local symmetry transformations, called local gauge invariance, the fundamental forces of nature arise. In the next sections the theories that are used to describe the three kind of interactions are explained.

2.2.1 Quantum Chromodynamics

Although the strong force was known since the 1930s, a consistent gauge formulation of it was only achieved in the 1970s. The theory of the strong interactions is Quantum Chromodynamics (QCD) and it is a non-Abelian SU(3) gauge theory. The Lagrangian is of the following form [5]

$$\mathcal{L}_{\text{classical}} + \mathcal{L}_{\text{gauge-fixing}} + \mathcal{L}_{\text{ghost}}. \tag{2.1}$$

The first term is the classical part

$$\mathcal{L}_{\text{classical}} = -\frac{1}{4} F^A_{\alpha\beta} F_A^{\alpha\beta} + \sum_{\text{flavours}} \bar{q}_a (i\not{D} - m)_{ab} q_b, \tag{2.2}$$

where $F^A_{\alpha\beta}$ is the field strength tensor coming from a gluon field $\mathcal{A}^A_\alpha$, q_i the quark field and $\not{D} = \gamma_\mu(\partial^\mu + ieA_\mu)$. The classical Lagrangian describes the interaction of spin-$\frac{1}{2}$ particles of mass m, the quarks, and massless spin-1 particles, the gluons. The field strength tensor is derived through

$$F^A_{\alpha\beta} = \left[\partial_\alpha \mathcal{A}^A_\beta - \partial_\beta \mathcal{A}^A_\alpha - g f^{ABC} \mathcal{A}^B_\beta \mathcal{A}^C_\beta\right], \tag{2.3}$$

where the indices A, B, C run over the eight degrees of freedom of the gluon fields. Thus each gluon carries two different colour charges, in order to be able to couple to quarks. The third term of this equation gives rise to triplet and quartic gluon self-interactions.

In order to perform perturbation theory, a gauge fixing term is necessary. The second term of eq.(2.1) fixes the class for covariant gauges to a parameter λ

$$\mathcal{L}_{\text{gauge-fixing}} = -\frac{1}{2\lambda} (\partial^\alpha \mathcal{A}^A_\alpha)^2. \tag{2.4}$$

In non-Abelian theories, like QCD, the gauge-fixing term needs to be followed by a ghost term given by

$$\mathcal{L}_{\text{ghost}} = \partial_\alpha \eta^{A\dagger} (D^\alpha_{AB} \eta^B). \tag{2.5}$$

Ghost fields are necessary for cancelling unphysical degrees of freedom. Then, from eqs. (2.2-2.5) one can derive the Feynman rules for the quark-gluon interactions from the operator

$$S = i \int \mathcal{L} d^4x = S_0 + S_I, \tag{2.6}$$

where S_0 is calculated by the free part of the Lagrangian and S_I by the interaction term. The inverse propagator of fermions can be obtained by inverting the $-S_0$, while S_I is used for deriving the Feynman rules of the interaction which are treated as perturbations.

The quark and gluon propagators are found to be

$$\Gamma^{(2)}_{ab}(p) = -i\delta_{ab}(\not{p} - m) \quad \text{inverse quark propagator,} \tag{2.7}$$

$$\Gamma^{(2)}_{AB,\alpha\beta}(p) = -i\delta_{AB}\left[p^2 g_{\alpha\beta} - (1 - \frac{1}{\lambda}) p_\alpha p_\beta\right] \quad \text{inverse gluon propagator} \tag{2.8}$$

in momentum space.

The self-interaction term of the Lagrangian leads to an important property of the strong interactions, the asymptotic freedom. That means that in a non-Abelian theory, like SU(3), the interaction among particles becomes weaker, the smaller the distance between them, or respectively the larger the momentum transfer q^2. In such a case, the coupling constant of the interaction becomes smaller for higher energies. This property explains why at high energies quarks and gluons behave as almost free particles, while at low energies are bound and therefore it is impossible to be observed as colour singlets.

2.2.2 Electroweak Theory

In the 1920s, a quantum field theory that describes the electromagnetic interactions was developed, Quantum Electrodynamics (QED). The QED Lagrangian is

$$\mathcal{L} = \frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \bar{\psi}(i\not{D} - m_e)\psi, \tag{2.9}$$

where $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$ the field tensor and $\not{D} = \gamma_\mu(\partial^\mu + ieA_\mu)$ the covariant derivative. QED is invariant under local gauge transformations of the U(1) group and it predicts a massless boson, the photon, as a mediator. The theory was proven very successful since all its predictions were confirmed experimentally.

A few decades later, in the 1960s, the theory was extended to include the weak interactions by S. Glashow, S. Weinberg and A. Salam. The electroweak (EW) or GWS theory is based on a local SU(2)⊗U(1) gauge symmetry whose generators are the weak isospin operator T_3 for the SU(2) and the hypercharge Y for the U(1). Hypercharge is defined as

$$Y = 2(Q - T_3),$$

where Q is the electric charge. In the SU(2) representation the fermions are grouped in left–handed doublets and right–handed singlets. The U(1) group assigns the hypercharge to the fermions.

The gauge invariance of the EW Lagrangian is ensured via the introduction of three massless bosons $W^i, i = 1, 2, 3$, coming from the SU(2) and one massless from the U(1), B. The EW Lagrangian is given by

$$\mathcal{L} = -\frac{1}{4}W^{i\,\mu\nu}W^i_{\mu\nu} - \frac{1}{4}B^{\mu\nu}B_{\mu\nu} \tag{2.10}$$

The $W^i, i = 1, 2, 3$, and B are the field tensors for the SU(2) and U(1) fields, respectively,

$$W^i_{\mu\nu} = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu - g_W \epsilon^{ijk} W^j_\mu W^k_\nu \tag{2.11}$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu, \tag{2.12}$$

where g_W being the SU(2) gauge coupling. The Lagrangian, thus, describes four massless vector bosons forming a singlet (B) and a triplet (W^i) under weak isospin.

This is an unphysical model as in nature, three massive bosons, $W^\pm$, Z^0, have been observed and only one massless, the *photon*. Therefore, the SU(2)⊗U(1) symmetry has to be broken. That is possible with the introduction of the *Higgs mechanism*.

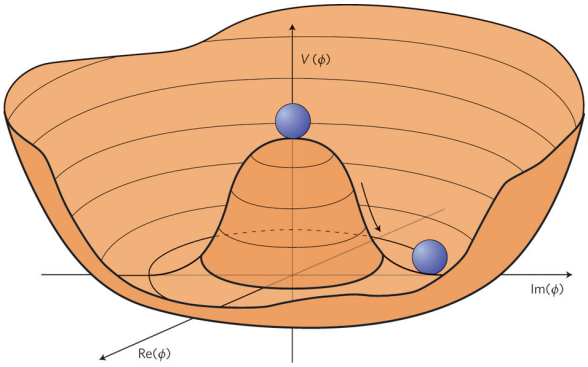


Figure 2.2: Higgs potential. The minimum of the potential is not at $\phi = 0$, but instead in a circle at $|\phi| = \frac{v}{\sqrt{2}}$.

2.2.3 The Higgs Field

In order to be able to get massive fermions and gauge bosons, complex scalar fields are introduced. For example, an isospin doublet ϕ that transforms under the SU(2) group is

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \sqrt{\frac{1}{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{2.13}$$

The EW Lagrangian becomes

$$\begin{aligned} \mathcal{L} = (\partial^\mu \phi^\dagger + i g_w W^\mu \cdot T\phi^\dagger + \frac{1}{2} i g'_w B^\mu \phi^\dagger) \times \\ (\partial^\mu \phi + i g_w W^\mu \cdot T\phi + \frac{1}{2} i g'_w B^\mu \phi) - V(\phi^\dagger \phi), \end{aligned} \tag{2.14}$$

where $V(\phi^\dagger\phi)$ is the Higgs potential that is chosen to be

$$V(\phi^\dagger \phi) = \lambda(\phi^\dagger \phi)^2 - \mu^2 \phi^\dagger \phi. \tag{2.15}$$

This potential, illustrated in Fig. 2.2, has a circle of degenerate minima at

$$|\phi| = \sqrt{\frac{\mu^2}{2\lambda}} = \frac{v}{\sqrt{2}}. \tag{2.16}$$

If a particular vacuum expectation value (VEV) is chosen for the field ϕ, such as

$$< \phi >_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \tag{2.17}$$

the SU(2)⊗U(1) symmetry of the vacuum state is broken. The fluctuations around the new minimum can be parametrised through a field $H(x)$ which will be the Higgs field

$$< \phi >_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}. \tag{2.18}$$

Under gauge transformations the unphysical degrees of freedom disappear and in their place massive gauge bosons appear. The Higgs boson is the only dynamical remaining field.

The mass of gauge bosons can be obtained by redefining the fields W^3 and B through new fields A_μ and Z_μ which propagate independently

$$\begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_w & \sin\theta_w \\ -\sin\theta_w & \cos\theta_w \end{pmatrix} \begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix}, \tag{2.19}$$

where the electroweak mixing angle θ_w is given by

$$\cos\theta_w = \frac{g_w}{\sqrt{g_w^2 + g_w'^2}}, \ \sin\theta_w = \frac{g_w'}{\sqrt{g_w^2 + g_w'^2}}. \tag{2.20}$$

From this rotation the W and Z bosons acquire their masses. That is

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2) \quad \text{with mass} \quad m_W = g_w\frac{\upsilon}{2}, \tag{2.21}$$

$$Z_\mu^0 = \frac{1}{\sqrt{g_w^2 + g_w'^2}}(g_w W_\mu^3 - g_w' B_\mu) \quad \text{with mass} \quad m_Z = \sqrt{g_w^2 + g_w'^2}\frac{\upsilon}{2}. \tag{2.22}$$

The photon remains massless

$$A_\mu = \frac{1}{\sqrt{g_w^2 + g_w'^2}}(g_w' A_\mu^3 + g_w Z_\mu) \quad \text{with mass} \quad m_A = 0. \tag{2.23}$$

The self-couplings of the Higgs boson are given by

$$\mathcal{L}_{Higgs} = \frac{1}{2}\partial_\mu H \partial^\mu H - \mu^2 H^2 - \lambda\upsilon H^3 - \frac{1}{4}\lambda H^4, \tag{2.24}$$

hence its mass is

$$m_H = \sqrt{2}\mu = \sqrt{2\lambda}\upsilon \tag{2.25}$$

[6]. The mass of the Higgs boson is the only free parameter in the Standard Model in the Higgs sector and it is yet to be measured. In Fig. 2.3 (left) the estimated production cross sections for the Higgs boson for pp collisions at 7 TeV centre-of-mass energy are shown. Up to now, the Higgs boson has not been observed, but experimental and theoretical limits constrain the value of its mass. The latest ATLAS limits are shown in Fig. 2.4 and exclude masses in the ranges

$$112.9 \ \text{GeV} < m_H < 115.5 \ \text{GeV}, \ 131.0 \ \text{GeV} < m_H < 238.0 \ \text{GeV}$$

and

$$251.0 \ \text{GeV} < m_H < 466.0 \ \text{GeV}$$

at 95%CL [8]. It is, also, worth mentioning that a 2σ excess has been observed by both LHC experiments and the TeVatron at $\sim$126 GeV [9, 10][2]. Finally, the mass of the fermions is the result of Yukawa interactions between the Higgs boson and fermion fields:

$$m_f = g_f\frac{\upsilon}{\sqrt{2}}.$$

The Yukawa coupling, g_f, is stronger the heavier the fermion is. Hence, interactions of the Higgs boson with third generation quarks and leptons is favoured. In Fig. 2.3 (right), the decays of the Standard Model Higgs boson to fermions and bosons is shown.

[2]On the 4th July, CERN announced the observation of a new particle at 125 GeV that seems to have all the properties of the long-sought Higgs boson. For details see [2, 3].

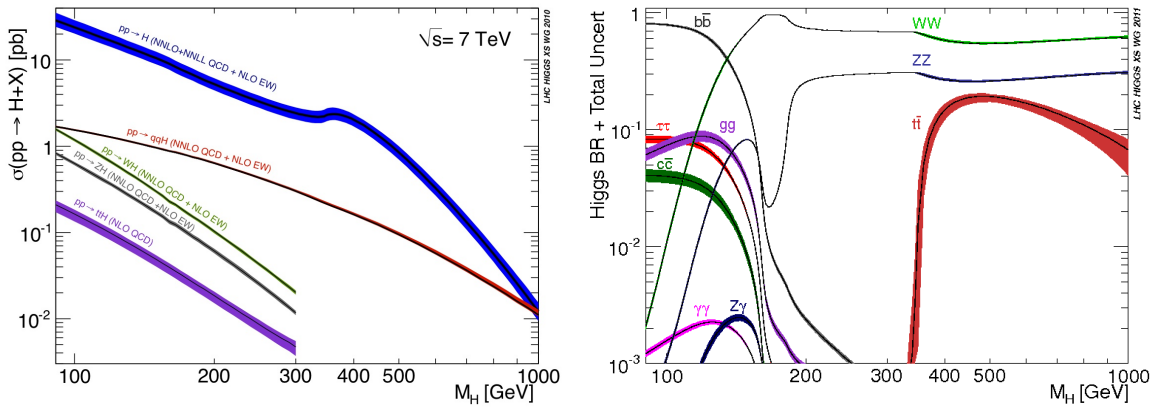


Figure 2.3: Left: Production cross sections for the Standard Model Higgs boson for pp collisions at 7 TeV. Right: Decay rates for the Standard Model Higgs boson. Both plots taken from [7].

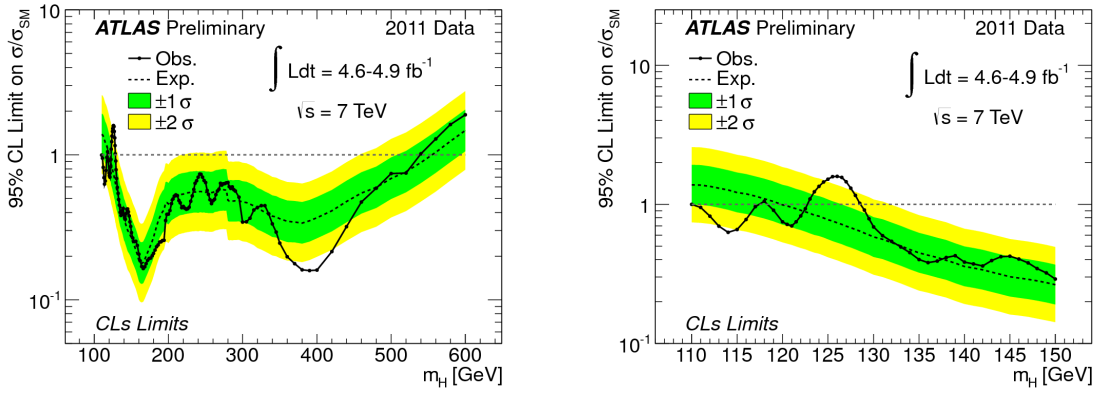


Figure 2.4: The observed (full line) and expected (dashed line) 95% CL combined upper limits on the SM Higgs boson production cross section divided by the Standard Model expectation as a function of m_H in the full mass range considered in this analysis (left) and in the mass range $110 < m_H < 150$ GeV (right). The dotted curves show the median expected limit in the absence of a signal and the green and yellow bands indicate the corresponding 68% and 95% intervals. Taken from [8].

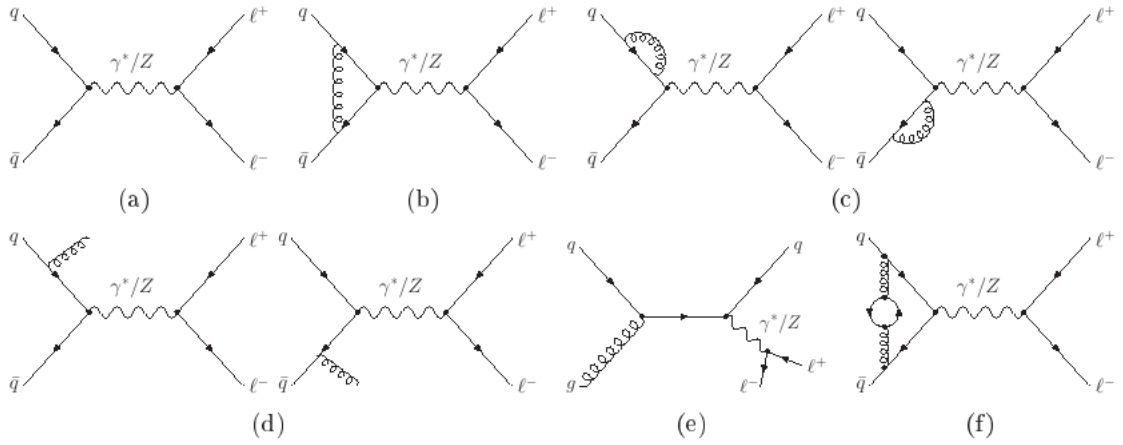


Figure 2.5: Feynman diagrams of the production and decay channels of the Z boson in a hadron collider.

2.2.4 The Z boson

The weak interaction is mediated via two kind of vector bosons, a charged $W^{\pm}$ and a neutral Z. The W was known already from its role in nuclear decays, being a mediator of the interaction for beta decay. The Z being neutral and very similar to the photon was more difficult to observe. First evidence came from neutrino elastic scattering in 1973 at the Gargamelle bubble chamber at CERN. A phenomenon that was explained as the interaction of neutrinos with electrons with the exchange of a Z boson. The actual discovery of the particle was achieved at the SPS $p\bar{p}$ collider at CERN in the early 1980s. The current world average of the mass and the decay width are

$$m_Z = 91.1876 \pm 0.0021 \text{ GeV} \quad \text{and} \quad \Gamma_Z = 2.4952 \pm 0.0023 \text{ GeV}, \tag{2.26}$$

respectively [4].

The main production mechanism is annihilation of a quark-antiquark pair. The (anti)quark can be a valence or a sea (anti)quark depending on the colliding particles. In hadron colliders, it is impossible to distinguish between Z boson and an off-shell photon production, γ^*, thus henceforth they are considered as one. The Feynman diagram of the leading order process[3] is shown in Fig. 2.5(a). In the Feynman diagram a quark–antiquark pair annihilates via γ^*/Z to a lepton pair. Higher order corrections come from QCD initial and final state radiation or QCD Compton scattering ($q + g \rightarrow \gamma^*/Z + q$), shown in Fig. 2.5(b)-(f). The total cross section is calculated from the sum of all possible processes up to next-to-next-to-leading order diagrams. The cross sections for leading order (LO), next-to-leading order (NLO) and next-to-next-to-leading order (NNLO) accuracy are shown in Table 2.2. The Z boson decays to leptons ($ee, \mu\mu, \tau\tau, \nu\nu$) and hadrons. The partial decay widths are shown in Table 2.3.

The NNLO QCD corrections of the theoretical prediction for the cross section times branching ratio for pp collisions at $\sqrt{s} = 7$ TeV is estimated at [12]. The value is found to be

$$\sigma^{NNLO}_{\gamma^*/Z \rightarrow ll} = 0.96 \pm 0.05 \,\text{nb} \quad \text{for } 66 < m_{ll} < 116 \text{ GeV}.$$

[3]The cross section of a process is the sum of the contributing Feynman diagrams. If only tree-level diagrams are considered the calculation is referred to as leading order, while if one-loop diagrams are taken into account it is called next-to-leading order, for two-loop diagrams next-to-next-to-leading order and so on.

Table 2.2: Central values for the production cross sections of $\sigma \times \mathrm{BR}(\gamma^*/Z \to ll)$ process at leading order (LO), next-to-leading order (NLO) and next-to-next-to-leading order (NNLO) accuracy [11].

order in α_s	$\sigma(\gamma^*/Z \to ll)$ [nb]
LO	0.758
NLO	0.938
NNLO	0.964

Table 2.3: Main decay modes of the Z boson [4].

decay mode	Branching Fraction (Γ_i/Γ) [%]
e^+e^-	3.363 ± 0.004
$\mu^+\mu^-$	3.366 ± 0.007
$\tau^+\tau^-$	3.367 ± 0.008
invisible	20.00 ± 0.06
hadrons	69.91 ± 0.06

Chapter 3

Experimental Setup

The experimental setup that physicists use to look into the world of the elementary particles is a particle accelerator combined with a simpler or more complicated detector. That is exactly what it is done for the discovery of the last missing particle in the Standard Model, the *Higgs boson*. The Large Electron Positron (LEP) collider at CERN in the late eighties and nineties and the TeVatron at Fermilab from 1995 till 2011 have tried to search for the elusive particle, although without success. In their place a much stronger machine is built; the Large Hadron Collider (LHC) at CERN. LHC is designed to accelerate protons at centre-of-mass energy of 14 TeV and lead ions at total centre-of-mass energy of 1150 TeV. Collisions take place at four interaction points where four detectors have been installed.

The design energy has not yet been reached, due to technical problems. Instead the machine operated at 7 TeV for protons for two years (2010-2011) with great success. A fact that provided with technical knowledge and confidence to upgrade the run to 8 TeV for 2012. At the end of 2012, LHC will shut down for one year to enable an upgrade of the accelerator to the design energy and improvements on detector parts for the experiments. Further improvements of the LHC have already been planned for 2016 and 2020 and involve luminosity upgrades and, in many cases, complete sub-detector replacements. In this chapter the LHC and the setup of the experiment ATLAS, whose data are used for this thesis, are described.

3.1 Large Hadron Collider

The LHC is a ring accelerator of 26.7-km circumference. The tunnel is on average 100 m below the surface and extends from the Jura mountains till the Léman lake. The accelerator is designed for circulating proton-proton (pp) beams in two independent rings [13]. Additionally, lead-ion beams (Pb^{+2}) have been integrated in the physics program, giving the opportunity for studies in the QCD transition region and quark-gluon plasma physics. The trajectories of the protons or Pb ions is steered with the help of super-conducting magnets that operate at 1.9 K.

The two beams meet at four interaction points underground, where detectors have been placed. The four detectors, i.e. four experiments, are ALICE, ATLAS, CMS and LHCb. ATLAS and CMS are general purpose detectors and their physics interests vary from Higgs searches to studies of rare B meson decays or forward physics. ALICE and LHCb, on the other hand, are specialised experiments in relativistic heavy ion physics the former and B meson physics and CP-violation the latter. Two more detectors are placed in the LHC ring, the LHCf and TOTEM. LHCf, situated on either side of the ATLAS detector, measures neutral pions in the forward direction accumulating data that will help to get closer to understanding the ultrahigh-energy

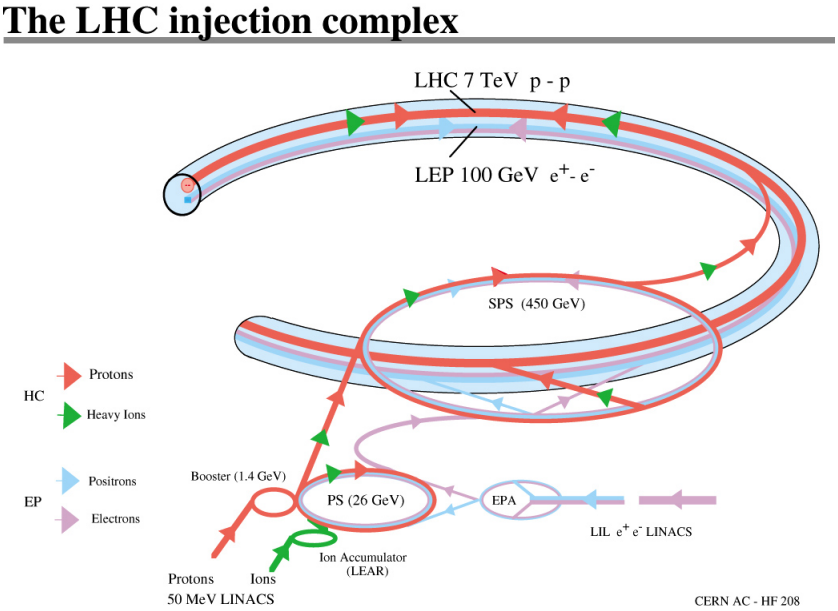


Figure 3.1: LHC acceleration chain. In red are shown the injection point and the pre-acceleration rings for protons and in green for lead ions.

cosmic ray events. TOTEM shares the cavern with CMS and, also, detects forward escaping particles from the collision point to study the structure of the proton and to measure the LHC luminosity.

In order to achieve the high energies necessary, the proton (*Pb-ion*) beams are going through a series of pre-accelerators. From the source of protons (*Pb-ion*) the particles are sent to the Linac2 (*Linac3*) for the first acceleration, then they continue to the Proton Synchrotron Booster (PSB) (*Low Energy Ion Ring (LEIR)*), to the Proton Synchrotron (PS), then to the Super Proton Synchrotron (SPS) to end up in LHC. A schematic of the acceleration chain is shown in Fig. 3.1. The source of protons is hydrogen atoms which are stripped of their electrons. The protons are accelerated in the Linac2 up to 50 MeV before they are injected into the PSB. The Booster is composed of four, superimposed rings and gets the protons to energies up to 1.4 GeV[1]. Then, the PS takes over and pushes the energy of the protons to 26 GeV (*5.9 GeV/nucleon*). SPS follows with a ring of 6.9 km in circumference and maximum energy 450 GeV for protons (*177 GeV/nucleon*). In the LHC ring the proton beams are accelerated to their maximum energy; 3.5 TeV in 2010, 2011 runs and 4.0 TeV in 2012. The design energy per beam is 7 TeV (*1.38 TeV/nucleon*) and it is planned to be achieved after a machine upgrade in 2014.

3.1.1 Luminosity

In collider experiments, a very important variable is the integrated luminosity ($\mathcal{L}$), which describes the flux of particles produced per unit area per collision. The luminosity accounts for the amount of data delivered and is used to estimate the amount of events of a process of interest

[1]The production of lead ions is more complicated and out of the scope of this thesis. Some information is provided in the ALICE blog page [14].

produced per collision. The number of events for the process under study is given by

$$N_{\text{event}} = \int \mathcal{L}\sigma_{\text{event}} dt, \tag{3.1}$$

where σ_{event} is the cross section of the particular process. The luminosity depends on the beam parameters according to the following relation

$$\mathcal{L} = \frac{N_b^2 n_b f_{\text{rev}} \gamma_r}{4\pi\varepsilon_n \beta^*} F, \tag{3.2}$$

where N_b is the number of particles per bunch, n_b the number of bunches per beam, f_{rev} the revolution frequency, γ_r the relativistic γ factor, ε_n the normalised transverse beam emittance, β^* the β function at the collision point and F the geometric luminosity reduction factor due to crossing angle at the interaction point[2].

ATLAS and CMS are designed to make full use of the delivered luminosity, $\mathcal{L} = 10^{34}\,\text{cm}^{-2}\text{s}^{-1}$, for proton runs, while for LHCb the maximum necessary luminosity for its physics program is $\mathcal{L} = 10^{32}\,\text{cm}^{-2}\text{s}^{-1}$. For heavy ion runs the luminosity is lower aiming at peak values of $\mathcal{L} = 10^{27}\,\text{cm}^{-2}\text{s}^{-1}$.

Since LHC is not running yet at the design centre-of-mass energy of 14 TeV some of the parameters are different than the design ones given in [13, 15, 16]. During the 2010 data run that started in March and ended in November of the year, protons were accelerated to 3.5 TeV per beam [17]. The total instantaneous luminosity was $0.1 - 1.2 \times 10^{28}\text{cm}^{-2}\text{s}^{-1}$ achieved with 10^{10} protons per bunch with 2(3) bunches in each beam out of which 1(2) pairs collided. In 2011, the beam energy and the protons per bunch remained the same, but the number of colliding bunches from 1-2 reached 1331 in one fill with corresponding bunch spacing from 75 ns in the beginning to 25 ns at the end of the run period. The parameter values for the beams at the ATLAS interaction point for $\sqrt{s} = 7$ and 14 TeV are given in Table 3.1. The total delivered integrated luminosity by the LHC and the recorded luminosity by ATLAS for the 2010 and 2011 data taking periods are shown in Fig. 3.2. Out of the total delivered luminosity, 36 $\text{pb}^{-1}$from the 2010 run and 1.3-1.5 $\text{fb}^{-1}$for the 2011 run are used for this analysis. The systematic uncertainty on these measurements is estimated with van der Meer scans to be 3.4% in the 2010 runs [18] and 3.7% in the 2011 runs [19].

3.2 ATLAS detector

In Fig. 3.3 the ATLAS detector is shown. ATLAS is one of the two multi-purpose experiments at LHC with dimensions $44 \times 21 \times 15$ m^3 (length $\times$ width $\times$ height). The main physics aim of the experiment is the discovery or exclusion of the elusive *Higgs* boson, hence the design of the sub-systems is based on achieving this goal. A fine-resolution calorimeter system for both charged leptons and hadrons is constructed to be able to separate signal from QCD and Standard Model processes. The good particle identification is obtained with the fine and multi-layer inner tracking detector and the muon spectrometer. Finally, a fast rejection of uninteresting events is accomplished with a three-level trigger system. In the following sections a more detailed description of the ATLAS sub-systems is given.

[2] $F = \left(1 + \left(\frac{\theta_c \sigma_z}{2\sigma^*}\right)^2\right)$, where θ_c is the full crossing angle at the interaction point (IP), σ_z the RMS bunch length and σ^* the transverse RMS beam size at the IP. It is assumed that the two beams have equal parameters and $\sigma_z << \beta$.

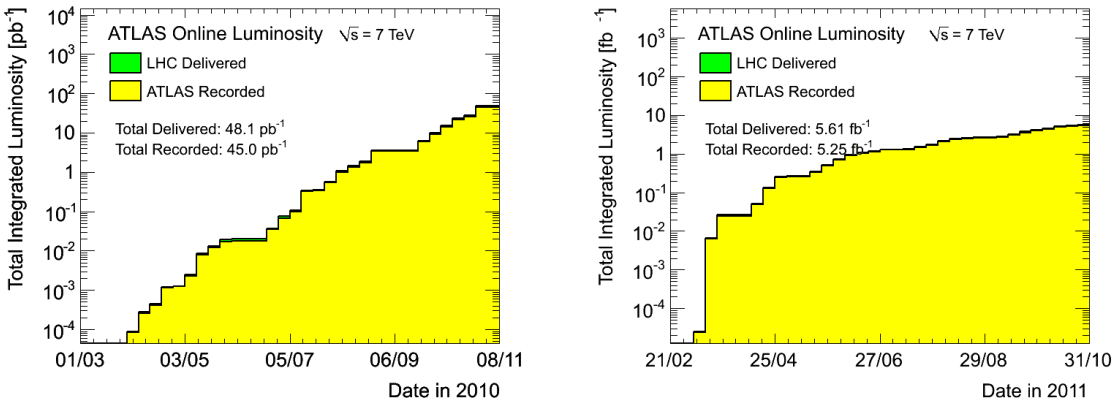


Figure 3.2: Total integrated luminosity delivered by the LHC and recorded by ATLAS per week for 2010 (left) and 2011 (right) data taking periods.

Table 3.1: Beam parameters at the ATLAS interaction point. Under 14 TeV the design values are stated and under 7 TeV the values for the 2010 (2011) run [17]. The parameters in the Table are explained in Section 3.1.1.

$\sqrt{s}$ [TeV]	14 (design)	7 in 2010 (2011)
f_{rev} (Hz)	11245	11245
n_c	2808	1-2 (200 - 1331)
N_b ($10^{10}\,p$/bunch)	11.5	1.1
$\varepsilon_{N_x,y}$ (μm-rad)	3.75	1.5 - 10
$\beta^*_{x,y}$ at IP (m)	0.55	11 - 2
$\sigma^*_{x,y}$ (μm)	16.6	100 - 40
σ_z (mm)	75	35 - 60
L ($cm^{-2}s^{-1}$)	1.0×10^{34}	$0.1 - 1.2 \times 10^{28}$ (0.30×10^{28})

3.2.1 Coordinate System

The ATLAS detector is a complicated system where several reference systems could be defined. The preferred one starts at the interaction point in the "heart" of the ATLAS detector and the z-axis is identified with the beampipe. Naturally, the $x - y$ plane is transverse to the beam. The positive x-axis is pointing towards the centre of the LHC ring, the positive y-axis upwards and the positive z-axis towards the A-side of the detector. More meaningful coordinates for measuring physics quantities are the azimuthal and polar angles and the pseudorapidity. The azimuthal angle ϕ is defined around the beam axis and the polar angle θ is the angular distance from the beampipe. Finally, for approximately massless particles, the pseudorapidity,

$$\eta = -\ln \tan \frac{\theta}{2},$$

is used. In the case where the mass of the particle cannot be ignored, such as jets, instead of the pseudorapidity the rapidity is used

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z}.$$

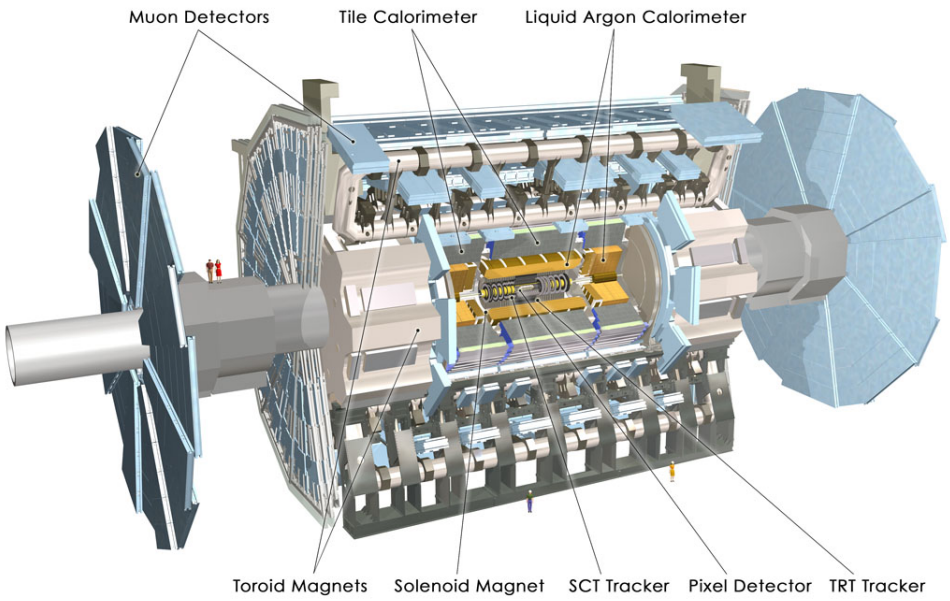


Figure 3.3: ATLAS detector. The main sub-detectors and the magnets are indicated with arrows.

The transverse momentum is, finally, defined in the $x - y$ plane as

$$p_T = \sqrt{p_x^2 + p_y^2}.$$

3.2.2 Inner Detector

In the innermost part of ATLAS the Inner Detector (ID) is located [20, 21]. Its aim is to track charged particles that fly away from the collision point. As a charged particle crosses the detector, electron-hole pairs are created. The latter drift in opposite directions when electrical field is applied. The signal is created only by migrating charges and is proportional to the square of the absolute charge of the particle. If magnetic field is applied around the detector, the path of the traversing particles is curved and its momentum can be estimated. At the LHC, about 1000 particles are produced per collision. Thus, a three-level tracking system is built and submerged in a strong solenoid magnetic field. The one closest to the collision point is the Silicon pixel detector, which is surrounded by the Silicon Strip (SCT) and the Transition Radiation Tracker (TRT). The pseudorapidity coverage of the Inner Detector is $|\eta| < 2.5$. The solenoid magnet operates at nominal value of 2 T.

The pixel detector, the inner most detector shown in Fig 3.4, comprises three barrels at radii 5.0, 9.0 and 12.0 cm and three disks on each side between radii 9.0 and 15.0 cm. The three-layer structure provides three precision measurements over the full acceptance, $|\eta| < 2.5$, of the detector. Its main aim is to determine the resolution of the impact parameter and the secondary vertex of the interaction. The latter is important for B meson and τ lepton studies. The innermost layer of the pixel detector, called b-layer, is the one that provides the secondary vertex measurement. In total the pixel detector consists of 1744 modules each 62.4 mm long and 21.4 mm wide with 46080 pixel elements. Due to its proximity to the collision point, the

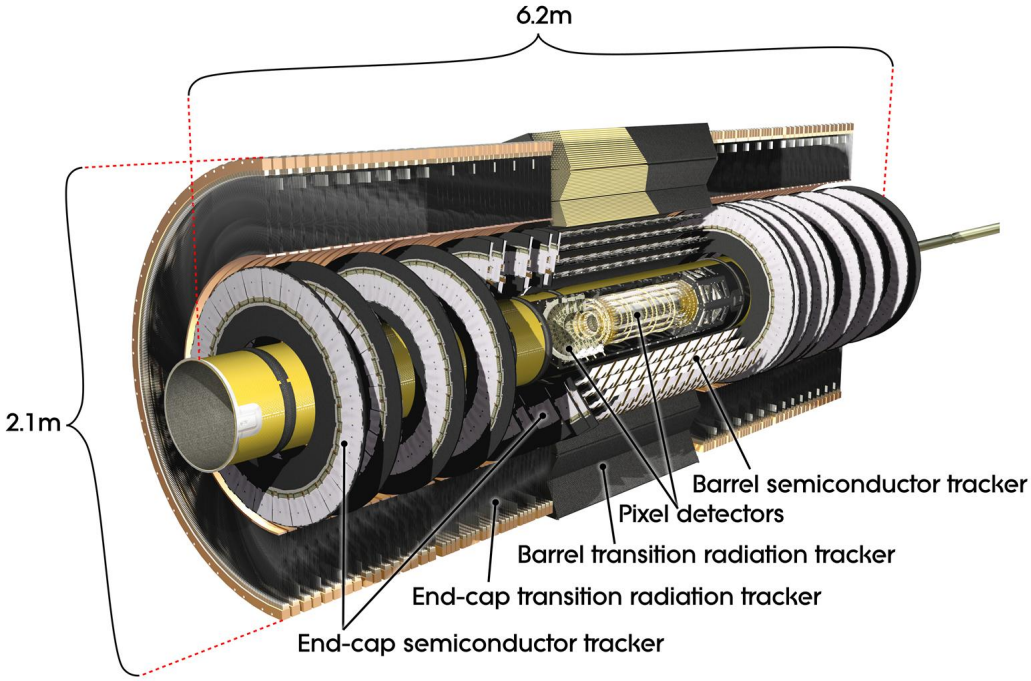


Figure 3.4: Inner Detector (ID) and its main parts. The pixel, the semiconductor (SCT) and the transition radiation (TRT) tracker.

pixel detector is exposed to over 300 kGy of ionising radiation and hence needs to be replaced after only three years of operation at maximal luminosity. At the moment, an upgrade of the pixel detector is scheduled for the shut-down phase of the LHC in 2013. During this upgrade an extra layer will be inserted closer to the collision point, called the Insertable B-layer (IBL). The whole pixel detector will be completely replaced in 2016.

The SCT (Semiconductor Tracker), envelops the pixel, as shown in Fig. 3.4. It has eight layers of Silicon micro-strip detectors, with dimensions 6.36×6.40 cm^2 each. The SCT provides precision measurements in the r-ϕ and z coordinates, hence contributing to the determination of the momentum, the impact parameter and the vertex position in the full pseudorapidity range.

The TRT (Transition Radiation Tracker) is the last layer of the Inner Detector and its geometrical coverage is $|\eta| < 2.0$. 50000 straws in the barrel and 320000 in the endcaps are surrounded by 3 m^3 of gas, which is a mixture of Xe(70%), CO_2(27%) and O_2(3%). The TRT is designed to provide a good electron-photon separation, thanks to the Xenon gas.

3.2.3 Calorimeters

The energy of the particles that are traversing the detector is measured in the calorimeters; the electromagnetic (EM) for charged light particles and the hadronic for hadrons. The physics processes that the EM calorimeter is based on is bremsstrahlung and pair production. Charged particles, mainly electrons, traversing the calorimeter react with the material of the detector and emit photons via bremsstrahlung radiation. These photons via pair-production produce electron-

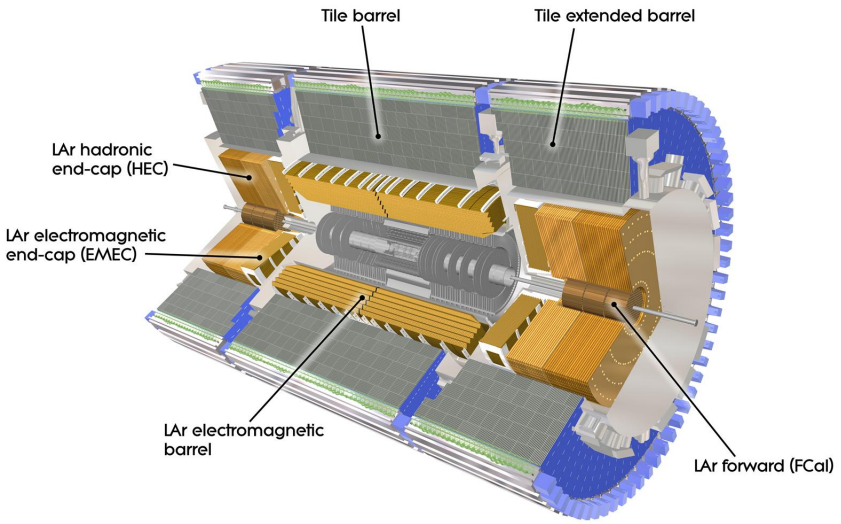


Figure 3.5: ATLAS electromagnetic (EM) and hadronic calorimeters and the forward calorimeter (FCal).

positron pairs that lose energy again via bremsstrahlung. This process, called electromagnetic shower, continues till the electrons have energy lower than 1 MeV, when they ionise and can be detected as an electromagnetic pulse. Similarly, hadrons going through matter produce a hadronic shower. Only that, in this case, the cause of the energy loss is not bremsstrahlung but the strong interaction and the splitting of the initial hadron produces new particles (quarks and gluons) with lower energy. Again, the lower energy particles are detected via ionisation. The distance travelled in the detector is called radiation length (X_0) for electrons/photons and interaction length (λ) for hadrons. In general, the hadrons are more penetrating than electrons, i.e. they need more material to lose all their energy and ionise, which explains why the hadronic calorimeter has a lot more material and it is situated outside the electromagnetic one.

The ATLAS calorimeter is shown in Fig. 3.5 and it is extensively described in [20, 21]. The EM and the hadronic calorimeters are cylindrical around the beam pipe providing a full ϕ symmetry. In η, they extend out to $|\eta| = 4.9$ with the help of a forward calorimeter. Different technologies are used in order to achieve high performance resolution in the total pseudorapidity range.

The EM calorimeter consists of multiple layers of Kapton electrodes and lead absorber plates put into an accordion-shape configuration to ensure full ϕ symmetry. This construction is placed on a cryostat wall and the gaps in between are filled with liquid Argon (LAr). A presampler behind the cryostat wall is necessary for correcting for the energy lost in the material upstream of the calorimeter. The lead thickness in the absorber plates varies depending on the pseudorapidity region to maintain high performance in energy resolution. The total thickness of the calorimeter is more than 24 radiation lengths in the barrel and more than 26 in the endcaps. In the barrel, the EM calorimeter is comprised of three longitudinal sections. The first section is important for particle identification, such as separating electrons and photons from pions, and it provides a precise position measurement in pseudorapidity. The whole EM calorimeter is positioned inside a 0.5-1 T magnetic field, created by a toroidal magnet. The total

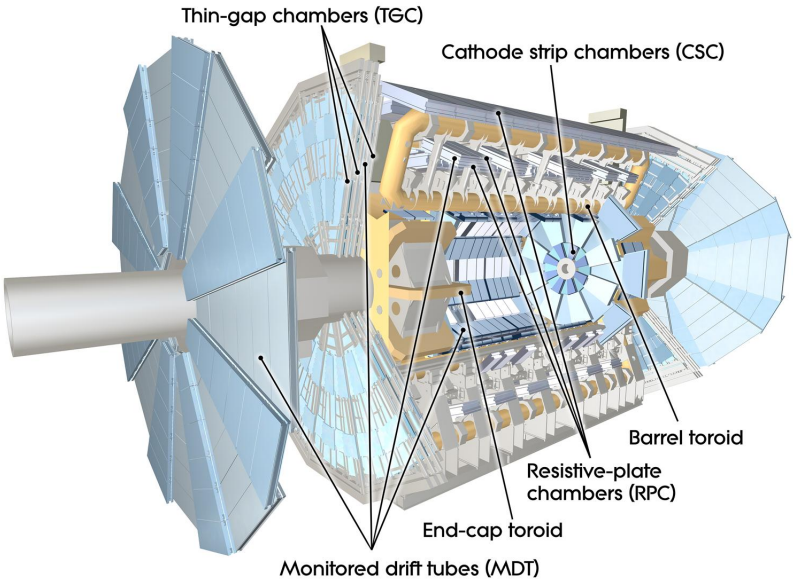


Figure 3.6: Muon spectrometer (MS) and the different parts is made of.

pseudorapidity coverage of the EM calorimeter is $|\eta| < 3.2$.

The Hadronic calorimeter extends to $|\eta| < 3.2$ using different techniques: tile in the barrel, liquid Argon at the endcaps (HEC) region. The tile calorimeter is located in the barrel and extended barrel in $|\eta| < 1.7$. It is a sampling calorimeter with steel as the absorber and scintillator as the active medium providing a radial length of 7.4 interaction lengths. At the endcap region, the hadronic calorimeter uses liquid Argon as medium because it has higher resistance to radiation. The HEC consists of two independent wheels made of copper plates and reaches $|\eta| = 3.2$.

The FCal detects particles flying in the forward direction, i.e. very close to the beam pipe, hence the need for a material that can withstand high radiation, thus using liquid Argon. The FCal is supported by the endcap cryostat wall and is comprised of three sectors. The first is made out of copper and it measures electrons and photons. The other two are made out of tungsten and detect hadronic decays. The FCal pseudorapidity coverage is $3.1 < |\eta| < 4.9$.

3.2.4 Muon Spectrometer

The muons traverse the Inner Detector and the calorimeters interacting only lightly with them, before they reach the spectrometer, where the muon event is triggered and the track of the particle is reconstructed.

The ATLAS muon spectrometer (MS) is shown in Fig. 3.6 together with the toroid magnet that envelops it. The MS covers the pseudorapidity range $|\eta| < 2.7$ and full ϕ coverage. It can provide momentum resolution of about 10% for 1 TeV tracks [20, 21, 22]. Its standalone momentum measurement capability extends from a few GeV, for muons that lost energy in the calorimeters, up to a few TeV. The tracks are bent by a superconducting air-core toroid magnet. The magnetic field is almost always perpendicular to the muon trajectories to avoid multiple scattering and to maintain good track resolution.

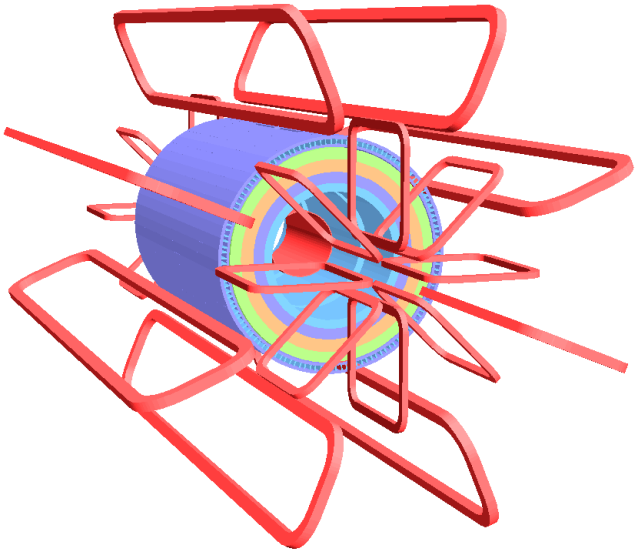


Figure 3.7: Barrel solenoid and toroid magnets of ATLAS.

The tracking and hence, the momentum measurement of the muons is done with drift tubes and cathode strip chambers, In the barrel, $|\eta| < 2.0$, the Monitor Drift Tubes (MDT) are arranged in three cylindrical layers around the beam axis. The Cathode Strip Chambers (CSC) are positioned at $2.0 < |\eta| < 2.7$ and are used for their high granularity and resistance against radiation. At $|\eta| = 0$ a gap for services to the inner systems is left.

Apart from measuring the momentum of the muons the MS is equipped with a muon trigger system that can deliver fast track information. Resistive Plate Chambers (RPC) are installed in $|\eta| < 1.05$ and Thin Gap Chambers (TGC) in the endcaps ($1.05 < |\eta| < 2.4$). Both type of chambers have a response time spanning from 15 to 25 ns and, thus, they are able to tag the beam-crossing.

3.2.5 Magnets

In Fig. 3.7 the solenoid and the toroid magnets of ATLAS are shown [20, 21]. The solenoid in the barrel provides a 2-Tesla magnetic field for the Inner Detector. The toroid creates a 3.9-Tesla field in the barrel and a 4.1-Tesla field in the endcaps for the muon spectrometer. The magnets are cooled down to 4.5 K by forced flow of helium.

3.3 Trigger and Data Acquisition Systems (TDAQ)

The main challenge for the ATLAS experiment is the high event rate, 10^9 pp collisions per second. In order to be able to store and analyse this amount of data, an online selection of "interesting" events is developed, namely the trigger system. The not-interesting events are rejected in three stages, called Level 1 (L1), Level 2 (L2) triggers and Event Filter (EF). The bunch crossing rate at design luminosity (10^{34} cm^{-2}s^{-1}) is 40 MHz, while the front-end systems can record events at a rate of 75 kHz (upgradable to 100 kHz). The Level 2 and EF, usually together referred to as High Level Trigger (HLT), reduce the event rate down to 100 Hz and final event size of $\sim$ 1.5 Mbyte. The design of the data acquisition system is such that it allows

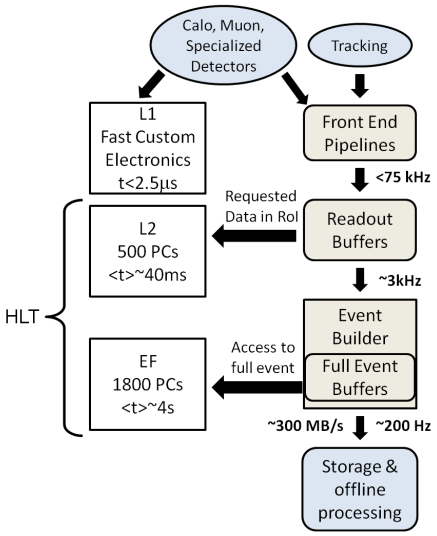


Figure 3.8: ATLAS trigger and data acquisition system layout.

continuous monitoring of the data from when an event is triggered till it is stored for physics analysis. An outline of the TDAQ (trigger and data acquisition) system is shown in Fig. 3.8.

3.3.1 Level 1 Trigger

The Level 1 (L1) trigger is hardware-based and discussed in detail in [21, 23]. It is installed on the calorimeters and the MS to reduce material occurrence from cables and read-out electronics. High p_{T} muons are triggered by the RPCs in the barrel and the TGCs in the endcaps. High p_{T} electrons and photons, jets and hadronically decaying τ leptons and missing and total transverse energy are identified in the calorimeters (EM, hadronic and forward) using reduced-granularity information. The decisions are based on simple combinations of objects which are required to coincide or be vetoed. Examples of such decisions are a single isolated muon with $p_{\mathrm{T}} > 20$ GeV (`L1_MU20I`), EM clusters with $E_T > 30$ GeV (`L1_EM30`), or a jet with $E_T > 100$ GeV and $E_{\mathrm{T}}^{\mathrm{miss}} > 100$ GeV (`L1_J100_XE100`).

As shown in Fig. 3.8, the detector signals are stored in the front-end pipelines waiting for the L1 trigger decision with a latency of less than 2.5 μs. Once the decision is taken the data are read out into readout buffers (ROBs). The output event rate from L1 is 75 kHz for luminosity 10^{34} cm^{-2}s^{-1} and for bunch spacing 25 ns. During the 2010 and 2011 runs, the trigger system was run at moderate rates. In 2010, at the average luminosities were $10^{27} - 10^{32}$ cm^{-2}s^{-1} and the bunch spacing was 35-45 ns. The resulting event rate was 30 kHz [24]. In 2011, when the luminosity was increased to $\mathcal{L} = 10^{32} - 5 \cdot 10^{33}$ cm^{-2}s^{-1} and the bunch crossing was reduced to 75-25 ns the event rate was ~ 60 kHz. Regions of Interest (RoIs) are defined at this point to be further investigated by the HLT.

3.3.2 High Level Trigger

The Level 2 (L2) and Event Filter (EF) triggers are both software based and they use sophisticated algorithms to reject events. The computer farm used for the HLT consists of 800 nodes configurable as either L2 or EF and 300 dedicated EF nodes. More details on the architecture and design can be found in [23, 21].

The L2 trigger looks into the regions of interest (RoIs) defined at L1 to verify whether a triggered object is valid. Afterwards, for all the objects except for the jets extra information is retrieved from the Inner Detector. That can be the transverse momentum from the TRT or the η and ϕ coordinates from the SCT. At this point, L2 RoIs are defined. L2 trigger decisions can be, for example, sharper p_{T} thresholds than at L1, calorimeter isolation requirements on muons or EM clusters or cluster-track matching using the Inner Detector information. In the end, the L2 trigger further reduces the rate to ~ 3 kHz with average processing time of ~ 40 ms/event.

In the EF the full event is reconstructed and offline algorithms are used to reject events. At this final stage the event rate is reduced to $\sim$ 200 Hz with an average processing time ~ 4 seconds/event. An event is reconstructed at the Event Builder (EB) from the information stored at ROBs using the full granularity of the detector. The EF algorithms can be simple p_{T} thresholds and isolation as in L2 or more complicated methods, like likelihood methods and boosted decision trees.

After the EF decision, the event is stored for offline analysis. The data are written to streams depending on the type of trigger that was fired: *Egamma* for EM triggers, *Muons* for muon triggers, *JetTauEtmiss* for jet, hadronic τ and $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers, *MinBias* for the minimum bias triggers. There are also several calibration streams for specific studies. About 10% of the data is written to an *express* stream which is used for online calibration and Data Quality (DQ) checks.

3.4 Data Quality Monitoring

Given the complexity of the ATLAS detector and the trigger and data acquisition system, constant and systematic monitoring is necessary. The data quality framework is complex so that it allows for independent operation of each sub-detector for calibration and debugging or, during data-taking, for exclusion of one or more subsystems if necessary. The online monitoring system consists of the Information Service (IS), the Online Histogramming Service (OHS) and the Data Quality Monitoring (DQM). These three services facilitate the information sharing and make the communication between the different environments possible. They use network and CPU load minimisation algorithms, allowing for the monitoring of functional parameters published by different hardware and software components. The whole data quality monitoring chain of ATLAS is shown in the schema in Fig. 3.9.

The Information Service (IS) is used for sharing information between applications in a distributed environment [25]. The main component of the IS is the IS-repository, which stores and manipulates any information coming from the related applications. The IS is using the tree-level object model, i.e. objects, classes and meta-classes. In the meta-type level the information is described in XML format and includes definition of type attributes, such as name, type and optional text explanation for each attribute. In the second level, the IS meta-types generate automatically programming language declarations. The IS information in the XML files is mapped to classes in C++ or Java. The third level contains the actual information as instances of the classes created at the second level. These instances are used to put or get information to/from the IS repository.

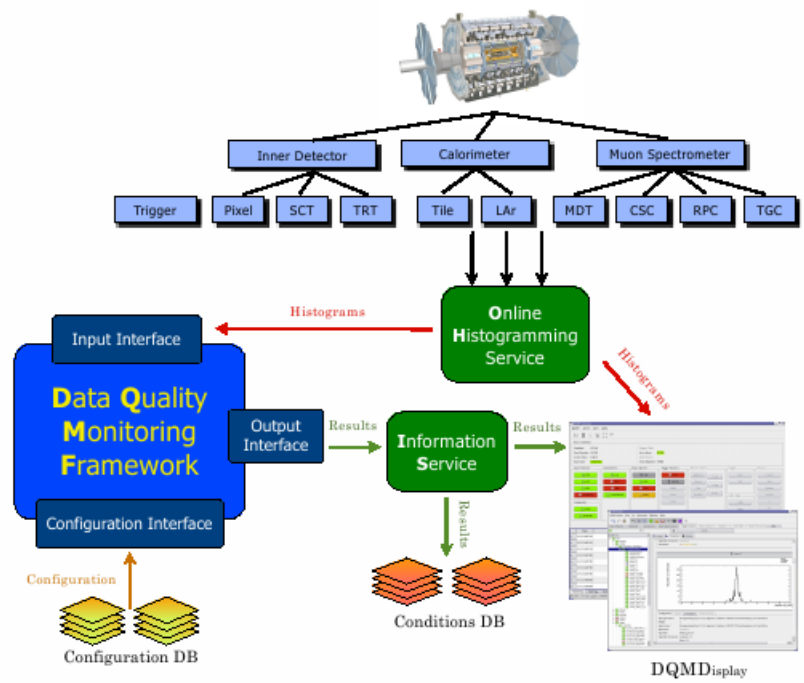


Figure 3.9: Layout of the Data Quality Monitoring Framework in ATLAS.

The Online Histogramming Service (OHS) is based on the IS repository and it permits user-defined histogram-tasks to retrieve histograms from different Histogram Providers. The providers might be handling monitoring, analysis tasks, or even histograms produced in read-out drivers (RODs). OHS is not responsible for booking, filling, storing or presenting the histograms. In order to receive a histogram from the OHS, the user must have a receiver interface which translates the histograms into an understandable format. For publishing histograms, the user needs a provider interface which can translate the histograms into a format which is understood by the OHS [26, 27].

Online Histogramming Presenter (OHP) is used for displaying the histograms that are produced by the different monitoring tasks. That is, histograms produced by monitoring applications, sub-detectors, detector control systems, or TDAQ, and published on the OHS server.

3.4.1 Data Quality Monitoring (DQM)

The Data Quality Monitoring system is a complex system that gets information from the detector, the Detector Control System (DCS) and the TDAQ and makes automatic data quality assessments. This information can be histograms, counters, voltages etc. stored in the OHS. After the DQ analysis has been performed the DQ status is sent to the IS for storage and to the ATLAS Control Room (ACR) for real-time monitoring.

Data Quality Monitoring Framework (DQMF) is an online framework for performing data quality assessments during a run through user-defined algorithms [28]. DQMF is organised in a tree-like structure in which each sub-detector and physics object is represented as a `DQRegion`. Quality assessments are performed through pre-defined algorithms, `DQAlgorithms`,

and parameters, DQParameters, and the final decision is assigned to DQFlags. Each DQRegion may contain one or more sub-regions which in their turn may be further split. The DQAlgorithm defines which algorithm should be used to perform the data quality tests. The DQParameter specifies which input histogram(s) should be used for the specific DQAlgorithm. A DQFlag could be red, yellow, green or undefined and is published at the DQM Output. During the 2011 run, the published flags are more specific, attributing specific name strings, called *defects*, if there is a problem. That allows an easier and faster assessment of errors and as a result quicker and more specific action. In the end, all the flags are collected and summarised by the IS. The results of the DQAlgorithms are published with the DQMDisplay to the ACR shifter. The DQMD is a graphic user interface that summarises the DQFlags for all the detector systems and the physics objects. All the results are published and stored in the conditions database (COOL) from where they can be accessed by the ATLAS offline framework for analysis.

3.5 Computing Model - The Grid

The ATLAS and, in general, the LHC computing model has been designed to provide quick access to all data for all members of the collaboration through Grid computing. Data recorded by the HLT are processed 'on site' at the CERN computing facilities, called Tier-0. Then, the processed, raw, data are copied at computing centres around the world, the Tier-1 centres. Datasets with reduced information are distributed to Tier-2 centres for physics analyses. A general software framework has been developed in ATLAS to meet the needs of the experiment ranging from the HLT until the end-user, ATHENA. The computing and software models are described in the following paragraphs.

The Grid connects computers, supercomputers, storage systems, or laptops around the world to provide almost unlimited computer resources [29]. A user can use the Grid to perform complicated analyses or tasks that demand a lot of CPU. The LHC experiments are making use of the Grid technology with 140 computing centres in 35 countries. The Worldwide LHC Computing Grid (WLCG) project has been created in order to build and maintain data storage and infrastructure for all the high energy experiments at LHC. The software necessary for accessing the Grid has been developed independently from each experiment. In ATLAS there are two such applications; PanDA (Production ANd Distributed Analysis) and GANGA (GAUDI, ATHENA and Grid Alliance). GANGA is the first middleware developed for ATLAS and it is based on GAUDI, an LHCb data processing framework [30]. It is a user-centred interface for ATHENA job assembly, submission, monitoring and control. The second middleware software is PanDA which was initially developed by the US ATLAS community for production and analysis. In 2007, it was expanded for use by the whole collaboration. Its advantages are scalability, robustness, efficient resource utilisation, minimal operations' manpower and tight integration of data management with processing workflow [31, 32].

The Tiers are the hardware, computational facilities for data processing and storage. The Tier centres are organised in three layers; Tier-0 is the computer farm at CERN, Tier-1s are large computer farms around the world and Tier-2s are regional computer farms in several universities and laboratories. As the data are being recorded a first processing is happening at the Tier-0 facility at CERN. This data is archived at CERN and copied over to the Tier-1 centres around the world. At the Tier-1 facilities the data are being reprocessed and various processed versions are being produced. The latter are then copied to the Tier-2 facilities for physics analyses.

At the Tier-2 centres the simulated data are also produced and stored. The analyses jobs are exclusively hosted by Tier-2 centres.

The data types produced in ATLAS are RAW data, ESDs, AODs, TAG and DPDs.

RAW data is the format in which the events are produced as the output of the event filter (EF). Each event is 1.6 MB arriving at an event rate of 200 Hz. The size of the file that is transferred from the EF to the Tier-0 is at most 2 GB, that is each file contains events of a single run. The RAW data are kept on disks for about 1 year.

ESD stands for Event Summary Data and is the data format after reconstruction. The target size is 500 kB/event and is stored in POOL ROOT files. The ESDs are kept for a period of 6 weeks for the `Egamma`, `Muons` and `JetTauEtmiss` streams. The ESDs for debug streams and calibration are kept longer.

AOD, Analysis Object Data, is derived from ESDs and the event size is smaller, ~ 100 kB/event. The AODs contain physics objects and information important for analysis. The file format is again POOL ROOT files.

TAG data are event-level metadata used for identification and selection of events or datasets of interest.

DPD, Derived Physics Data, is the format mainly used by physics analysers. It is n-tuple style representation of the data that a user can easily access and display them via the standard analysis tools, like ROOT.

3.5.1 Offline Software

The offline software of the ATLAS experiment, ATHENA, is serving multiple purposes: processing of RAW events, delivery of the processed results to the user and provision of tools for analysis [30]. It is object-oriented using primarily C++ and less FORTRAN and Java. ATHENA, is designed and developed with the following properties:

- abstract interfaces that allow for uniformity and flexibility among the different environments, e.g. HLT, offline and detector description;
- extensive dynamic libraries;
- separation between data and algorithms, that allows for changes at run time, for example of the track-finding algorithm without having to recompile the particle identification algorithms;
- clear separation between algorithmic code and the technology used to store the data, since the latter might change over the lifetime of the experiment; and
- potential for accommodating different types of developers, i.e more or less capable programmers.

The configuration within the ATHENA framework is done with scripts written in Python. ATHENA is under constant development and a new release is produced once a year. In this thesis release 16 has been used.

An analysis within the ATHENA framework requires to run over AOD files. But the AOD files are large making the analysis slow and less flexible. Hence, a different framework is used, called SFrame [33]. SFrame is based on ROOT trees (n-tuples as an output of DPD files) and makes use of the fact that a physics analysis is performed in cycles, e.g. calculating new

quantities for each event. All the basic framework needed, such as creating I/O trees, I/O of histograms, looping over events is provided by the framework. The user should still write his/her own selection/optimisation steps and should provide meta-data describing the cycle, for example luminosity, Monte Carlo weights, etc. Finally, one needs to compile the code only once and all variables can be changed through XML files at run-time.

GEANT4 is the toolkit used by ATLAS to simulate the detector itself [34, 30]. It provides functions for easy geometry description and browsing, for propagating particles through detector components or dead material and description of the physical processes. Since 2003, the GEANT4 packages have been embedded in ATHENA allowing for improvements on the description of physics processes, Monte Carlo truth simulation and the usage of ATLAS databases for the output.

Chapter 4

Phenomenology of pp collisions

In high-energy proton-proton collisions the actual constituents that interact are not the protons, but their composing parts, quarks and gluons, simply referred to as partons. In fact, at least one parton from one proton will collide with at least one parton from the other proton. This process is referred to as the hard-scattering. The two-parton interaction will produce a particle, which in turn will decay further to other particles. The proton remnants, also, produce secondary interactions which create low energy QCD jets in the detector. The collision and its evolution is shown in Fig. 4.1 and it is explained in some detail in the next sections. At the end of this chapter the tools used in high energy physics for simulating these processes are described.

4.1 Description of a pp collision

The frequency with which a specific process happens is measured with the production cross section and it is given by

$$\sigma_{pp} = \int dx_a dx_b f_a(x_a, Q^2) f_b(x_b, Q^2) \hat{\sigma}_{ab\to X}, \tag{4.1}$$

where $ab \to X$ is a specific process in which a parton a from one proton and a parton b from the other proton interact and produce a resonance X, $\hat{\sigma}_{ab\to X}$ is the cross section of this process and Q^2 the momentum transfer from the partons to the hadronic system. $f_i(x_i, Q^2)$ express the probability that a parton of type i with momentum fraction x_i is found inside a proton. These probabilities are called parton distribution functions (PDF) and cannot be computed in perturbative QCD but are measured experimentally.

The cross section of the specific process, $\hat{\sigma}_{ab\to X}$, is calculated with the use of Feynman rules. That means, that all contributing tree-level Feynman diagrams are drawn and the matrix element, $\mathcal{M}$, is written down. The cross section is the square of the matrix element of the hard-process integrated over the allowed phase-space.

At high energies and multiplicities the tree-level description of a parton interaction is not sufficient, because virtual corrections and real emissions become important in the initial and final state of the interaction. These are included in the matrix element via NLO and NNLO terms. The perturbative expansion results in a finite cross section estimation only if the coupling constant is less than one. But, in the case of the strong coupling, α_S, its value is close to one at low energies, $\Lambda_{QCD} \approx 300$ MeV. The divergence can be avoided with the use of approximate methods, namely hadronisation.

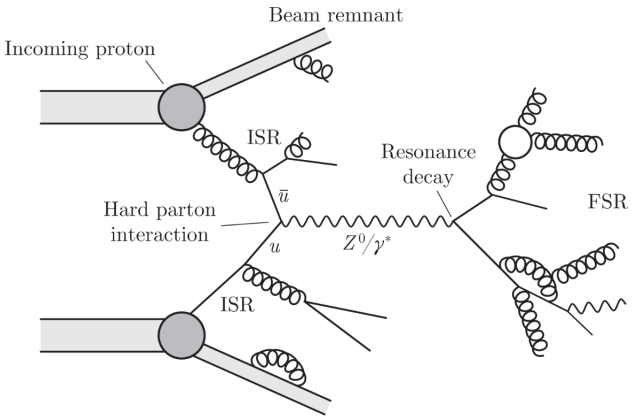


Figure 4.1: Schematic representation of a pp collision, including initial and final state radiation. Taken from [35].

Parton Distribution Functions (PDF). The PDFs describe the fraction of the momentum x of the protons that is carried away by the struck partons. They can only be estimated through global fits to data from deep inelastic scattering experiments (DIS), Drell-Yan (W/Z) and jet production measurements. There are two big groups that perform the fits to data and provide the PDF sets; CTEQ [36] and MSTW [37] collaborations. Several other smaller groups exist as well that perform specific fits, like the HERA collaboration and NNPDF [38]. The obtained PDF sets differ in the input data and in the initial parametrisation of the parton densities at a reference scale Q_0. As an example, the PDF sets provided by the MSTW group for the MSTW2008 are shown in Fig. 4.2.

Uncertainties on the PDFs cannot be evaluated by simply using a different PDF set, since both the CTEQ and MSTW are using similar methods to extract the distributions, instead the Hessian technique is used [40]. By diagonalising a matrix with all the free parameters of the fit, 20×20 for the CTEQ PDFs, the eigenvector directions provide the basis for the determination of the PDF error for any cross section. The experimental uncertainties are propagated linearly through the PDFs.

Parton Shower. Higher order corrections to the cross section arise predominantly from the emission of soft and collinear radiation. The initial and final partons of the hard-process are allowed to split into lower energy partons through the parton shower until the hadronisation scale. In numerical calculations, the corrections are added to the coloured lines via Sudakov form factors [41].

Hadronisation. Individual partons are not observable, therefore, all partons of the interaction are combined into colour-singlets with a process called hadronisation. There are two techniques used to hadronise quarks and gluons; the string (Lund) model and the cluster model. The string model assumes that colour strings connect parton pairs by creating certain tension. When the distance between the strings increases, the tension of the string increases too, until it breaks. A new quark-antiquark pair appears after the breaking which combined with the initial partons creates colourless mesons or baryons. In the cluster model all initial gluons are split into a quark-antiquark pair. Then, the neighbouring quarks are combined to form colourless clusters.

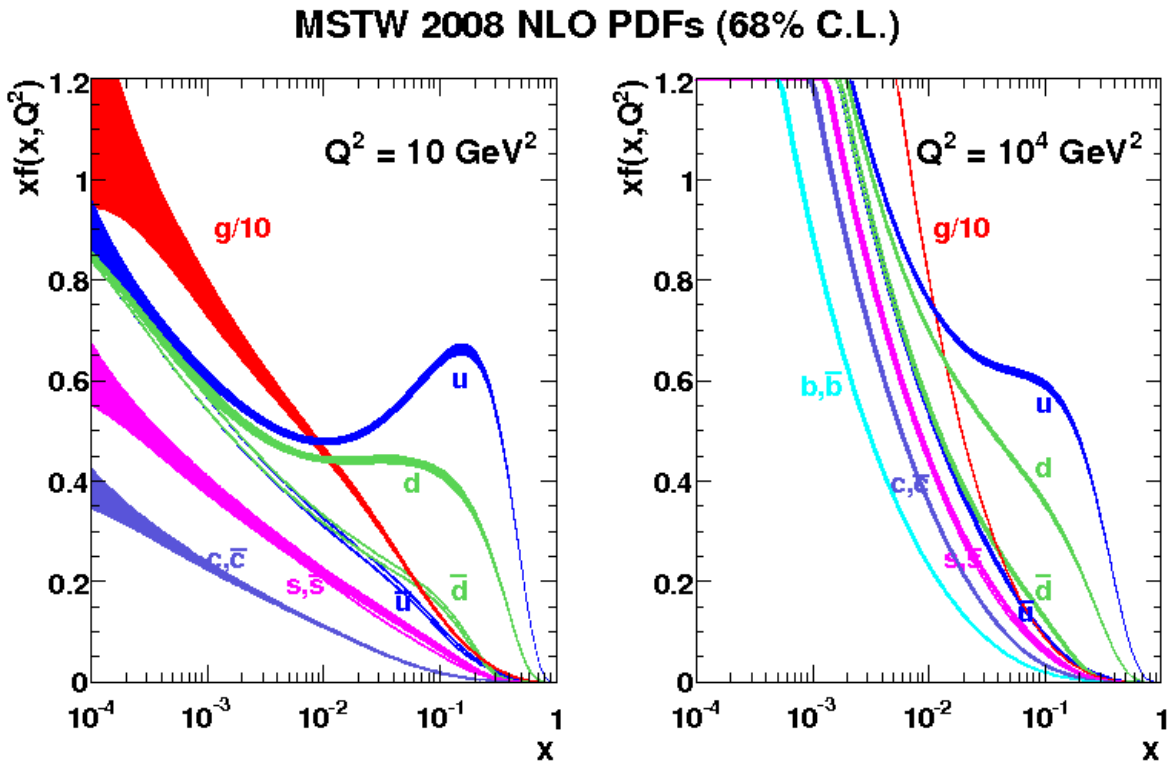


Figure 4.2: PDF sets at $Q^2 = 10$ GeV^2 (left) and at $Q^2 = 10000$ GeV^2 (right) for quarks and gluons estimated by MSTW2008 at NLO accuracy [39].

The clusters are converted into hadrons depending on their mass.

Underlying Event. Soft processes that takes place at the same time as the hard scattering are generally called underlying event. The underlying event originates from the spectator quarks of the interaction or the hadronisation process. These processes cannot be computed and a parametrisation based on data is used.

4.2 Monte Carlo Generators

The previously described processes, which take place in a pp collision, are modelled with Monte Carlo generators (MC). A variety of generators is available, because they utilise different approaches for the calculation of the hard scattering cross section, the PDFs, the hadronisation and parton shower models. In ATLAS, different Monte Carlo generators are used depending on the process of interest. The ones relevant for this thesis are AlpGEN, PYTHIA, HERWIG++ and MC@NLO. The specialised generators TAUOLA and PHOTOS are interfaced with the other generators for the simulation of τ lepton and photon processes, respectively.

AlpGEN is a tree-level matrix element plus parton shower generator (ME+PS). It is designed for generating Standard Model processes with high jet multiplicity at the final state in hadronic

collisions [42, 43]. The computation is done using the exact leading order calculation of a partonic matrix element, including b and t quark masses and t and gauge boson helicities. The generated events are interfaced with HERWIG [44] for the hadronisation and JIMMY [45] for the underlying event or with PYTHIA [46]. The default set of parton distribution functions used is CTEQ6.1, but in ATLAS the MRST LO** is used. The simulated processes relevant for this analysis are

- $W\rightarrow$ leptons, light quarks plus N jets, $N \leq 6$,
- γ^* / $Z\rightarrow$ leptons, light quarks or neutrinos plus N jets, $N \leq 6$.

The AlpGEN samples are used only in the analysis of the 2011 data, because in the 2010 Monte Carlo production the τ lepton polarisation was not correctly included in the AlpGEN dataset.

PYTHIA generates hadronic events in pp, $p\bar{p}$, ee and ep collisions. In ATLAS, PYTHIA6.4 is used which is written in FORTRAN [46]. A large number of processes is implemented at leading order in PYTHIA, amongst which are QCD, light and heavy boson EW decays, SUSY and exotics. The hard-process is simulated with a LO matrix element and it includes initial and final state radiation. The underlying event is also an intrinsic feature of the code covering from charged particle multiplicity to long-range correlations to minijets. For hadronisation in PYTHIA is done with the string model. In this analysis, PYTHIA is used for the simulation of the electroweak processes, W/Z decays, for the 2010 data analysis.

MC@NLO is a full next-to-leading order event generator. It includes diboson, heavy quark, Higgs boson, $W^{\pm}, Z, \gamma^*$ and dilepton processes. Mass effects and spin correlations are treated appropriately from the code. The showering and hadronisation are done with HERWIG [44, 47]. To avoid double counting of partons and emissions from the parton shower and the hard scattering, negative weights have been implemented in the code. These only occur for a small fraction of the events. MC@NLO is used to simulate $t\bar{t}$ events and for studying the systematic uncertainties on the modelling of the τ lepton decays.

HERWIG/Jimmy is a general purpose Monte Carlo event generator, that allows for simulations of lepton-lepton, lepton-hadron and hadron-hadron collisions. In HERWIG the parton shower is angular-ordered and the hadronisation is done with the cluster model. It is interfaced with the add-on generator Jimmy for including multiple-parton interactions. It is written in FORTRAN and can be interfaced to TAUOLA for τ lepton decays. It has a large "menu" of hard processes, such as QCD $2\rightarrow2$, electroweak boson production and several SUSY processes. For the PDFs in ATLAS, the MRST LO** is used and for the underlying event the AUET1 tunings [48].

HERWIG++ is a new event generator written in C++ [49]. HERWIG++ has an improved description of the parton shower and the hadronisation model. The version used in ATLAS is HERWIG++ 2.4.2 and 2.5.1 for the simulation of QCD multijet and diboson events.

TAUOLA is a specialised package for generating τ leptons including the spin polarisation. The code is written in FORTRAN and there is a newer C++ version. It describes mainly the weak current with first order QED corrections for leptonic decays, but one can also tune it for non standard interactions, τ neutrino mass and others. More information and the code can be found in [50, 51].

PHOTOS is a Monte Carlo algorithm for implementing QED radiative corrections to the event generator programs [52].

Chapter 5

Object Reconstruction

The data recorded from the trigger system cannot be directly used for physics analyses. The events that have passed the high-level trigger are reprocessed within the ATHENA framework and saved in AODs, see Section 3.5. In ATHENA, physics objects and event variables are built, such as muons, electrons, jets, etc and their performance is studied. In the following, the methods used for the object reconstruction and identification are presented. Additionally, the results of the reconstruction and identification studies for the 2010 and 2011 data taking are discussed. The efficiency studies are used for the calibration of the data and scale factors for the Monte Carlo are provided for the physics analyses.

5.1 Electrons

5.1.1 Electron Reconstruction

The electron reconstruction is based on clusters originating from the EM calorimeter which are then associated to tracks reconstructed in the Inner Detector. A preliminary seed cluster is defined with minimum energy of 2.5 GeV using the *sliding-window* algorithm[1]. In the case of a true electron this initial cluster has dimensions 3×5 in units 0.025×0.025 in the η/ϕ space of the middle layer of the calorimeter. In order to avoid duplicate events, the energies of neighbouring clusters are compared and the identical ones are removed. From Monte Carlo studies the cluster reconstruction efficiency is expected to be 100% for electrons with $E_{\mathrm{T}} > 15$ GeV [53]. Henceforth only electrons lying in the barrel region of the detector ($|\eta| < 2.5$) are discussed, since the ones in the forward region ($2.5 < |\eta| < 4.9$) are of no relevance to this analysis.

An electron is defined whenever tracks from the Inner Detector are associated to a seed cluster. The track is, then, extrapolated and its η and ϕ coordinates are compared to the coordinates of the cluster. If the difference is below a certain threshold, the track is matched to the cluster. After the track matching is done a final electron cluster is defined using 3×7 longitudinal towers of cells. The energy of the electron is given by the energy of the cluster. The η and ϕ coordinates are taken from the corresponding cluster parameters, except for when the track has only TRT hits, then both quantities are taken from the track. At this point the four-vector of an electron can be defined. Currently, prompt electrons and converted photons are

[1] The *sliding-window* algorithm forms a fixed-size rectangular cluster such that the energy of the cluster becomes maximum. Then, it checks all the neighbouring cells until the energy of the cluster drops below a threshold (2.5 GeV) [53]. The size of the cluster differs depending on the type of particle (electron/photon) and the region of the calorimeter (barrel/endcaps).

both reconstructed as electrons and the separation is done with further identification criteria. In this way a high electron reconstruction efficiency is achieved.

5.1.2 Electron Identification

The electron baseline identification variables in ATLAS are chosen so that they provide a good separation between electrons and jets that fake electrons. According to the background rejection efficiency, three sets of cuts are defined: *loose*, *medium* and *tight*.

The *loose* set is purely calorimeter based and is determined by two types of variables, two shower shape variables of the second layer of the EM calorimeter and the hadronic leakage:

- the ratio in η of cell energies in 3×7 to 7×7 (R_η);
- the lateral width of the shower ($w_{\eta 2}$); and
- the ratio of E_{T} in the first layer of (whole of) the hadronic calorimeter to the E_{T} of the EM cluster, R_{had1} (R_{had}), over a range $|\eta| > 0.8$ and $|\eta| < 1.37$.

The *medium* cuts include the *loose* and additionally, the following tighter calorimeter and track-based ones

- the shower shape variables of the first layer of the EM calorimeter, i.e. the ratio of the energy difference associated with the largest and second largest energy deposit over the sum of these energies (E_{ratio}) and the total shower width (w_{stot});
- a good track is required to have more than one hits in the pixel detector, the sum of the hits in the pixel and the SCT to be more than 7 and the transverse impact parameter (d_0) less than 5 mm; and
- for track matching the difference in pseudorapidity ($\Delta\eta_1$) between the cluster and the track needs to be smaller than 0.01.

The track-based cuts are tested with the first data and are found to be quite robust. The last set of cuts, the *tight*, is the one used in this analysis, since it provides a good separation between real and fake electrons. The *tight* cuts are defined as the *medium* ones and furthermore the following criteria:

- a b-layer cut (number of b-layer hits > 1);
- a tighter transverse impact parameter cut ($d_0 < 1$ mm);
- extra track matching criteria; the $\Delta\phi$ between the cluster and the track to be smaller than 0.02, a tighter $\Delta\eta < 0.005$ and the ratio of the cluster energy to the track momentum E/p from the TRT provides particle identification;
- the total number of TRT hits and the ratio of high-threshold hits to the total number of hits in the TRT; and
- check for photon conversions and reject them.

The jet rejection, defined as

$$R_{jet} = \frac{N_{\mathrm{truthjets}}}{N_{\mathrm{fakejets}}}$$

is measured in data for the three sets of cuts and it is found to be 500 for loose, 5000 for medium and 50000 for tight electrons with $E_{\mathrm{T}} > 20$ GeV [53, 54].

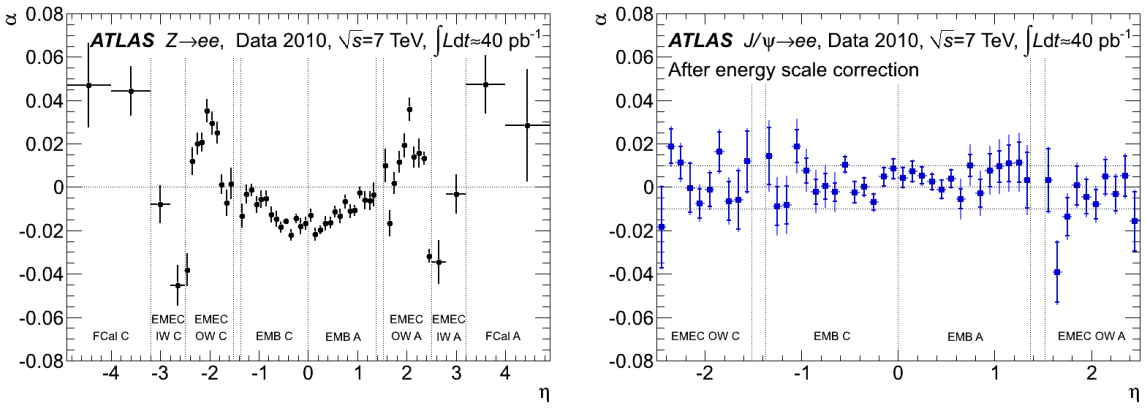


Figure 5.1: Energy scale correction factors in pseudorapidity bins derived from $Z \to e^+e^-$ (left) and $J/\psi \to e^+e^-$(right) data fits. The uncertainties shown for the $Z \to e^+e^-$ analysis are statistical only, while the $J/\psi \to e^+e^-$ are statistical and systematical. Taken from [53].

5.1.3 Energy Scale Calibration

For the electron energy scale and the alignment of the EM calorimeter $Z \to e^+e^-$ and $J/\psi \to e^+e^-$ decays are mainly used. An alternative method that uses the ratio of the energy measured in the EM calorimeter over the momentum measured in the Inner Detector, E/p, in $W \to e\nu$ events has been studied, as well. This method gives compatible results with the $Z \to e^+e^-$ study, but the systematics are larger and thus is used only as a cross check. More details are given in [53].

The energy calibration is done in three steps:

1. the actual detector hits are converted into deposited energy via the electronic calibration of the EM calorimeter,

2. energy corrections based on Monte Carlo studies are applied to the energy deposits of the reconstructed clusters and, finally,

3. using $Z \to e^+e^-$ decays ($J/\psi \to e^+e^-$ for low p_{T} electrons), an in-situ calibration permits the determination of the energy scale and alignment for the whole calorimeter coverage.

The mis-calibration between the measured and true energy for a given region i is given by

$$E^{\mathrm{meas}} = E^{\mathrm{true}}(1 + \alpha_i), \tag{5.1}$$

where E^{meas} is the energy measured at the calorimeter after Monte Carlo-based energy scale corrections, E^{true} is the true electron energy and α_i is the energy correction factor. In Fig. 5.1, the α_i values are given in several η bins to account for the different detector regions. The correction factors are of the order of 2% in the barrel region and 5% in the forward one. Several systematic uncertainties have been evaluated and the dominant one is found to be the additional material in front of the calorimeter and on the presampler which vary from -2% to +1.2%. The correction factors are applied on data events when defining electrons in the analysis.

Table 5.1: Effective constant term of the calorimeter energy resolution as measured from $Z \to e^+e^-$ 2010 data. The values are binned in η-regions. The detector region initials in the table stand for: EMB for EM calorimeter barrel, EMEC for EM calorimeter endcaps, OW/IW for outer/inner wheel. Taken from [53].

subsystem	η-range	c_{data} (%)
EMB	$\eta < 1.37$	1.2 ± 0.1 (stat) $\pm^{0.5}_{0.6}$ (syst)
EMEC-OW	$1.52 < \eta < 2.47$	1.8 ± 0.4 (stat) $\pm$ 0.4 (syst)
EMEC-IW	$2.5 < \eta < 3.2$	3.3 ± 0.2 (stat) $\pm$ 1.1 (syst)

5.1.4 Energy Resolution

The energy resolution in the calorimeter is given by

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c, \tag{5.2}$$

where a is the sampling, b the noise and c the constant term. All three terms are parametrised in bins of η. The first two terms of eq. (5.2) are taken from Monte Carlo and the constant term, c, is estimated from data with the following formula

$$c_{\text{data}} = \sqrt{2\left(\left(\frac{\sigma}{m_Z}\right)^2_{\text{data}} - \left(\frac{\sigma}{m_Z}\right)^2_{MC}\right) + c^2_{MC}}, \tag{5.3}$$

where c_{MC} is the constant term in the Monte Carlo simulation and its value is 0.5%, m_Z is the Z mass and σ the Gaussian fit value of the experimental resolution of the $Z \to e^+e^-$ distribution. The measured values of c_{data} are shown in Table 5.1. The dominant systematic uncertainty is coming from the use of the Monte Carlo value for the sampling term, 0.4% for the barrel region.

5.1.5 Total Efficiency

There are four steps in the electron selection: trigger, reconstruction, identification and isolation. The efficiencies of these steps are estimated with a tag-and-probe method. *Probe* electrons are selected with certain selection cuts, *tag* requirements, in order to acquire a clean and unbiased sample. The efficiency of a selection can be estimated by applying the specific cut to the probe sample. Any background contamination is subtracted from the probe sample using the appropriate data-driven technique. The final value is given from the ratio of probe electrons at a selection cut to the initial number of probe electrons. The total efficiency correction of the electron is

$$c = \epsilon_{\text{event}} \cdot \alpha_{\text{reco}} \cdot \epsilon_{\text{ID}} \cdot \epsilon_{\text{trig}} \cdot \epsilon_{\text{isol}}, \tag{5.4}$$

where

- ϵ_{event} is the efficiency of the event preselection, such as primary vertex requirement and event cleaning,
- α_{reco} is the basic reconstruction efficiency to find an electromagnetic cluster and match it to particle cuts,

Table 5.2: *Medium* and *tight* identification efficiencies as measured in $W \to e\nu$ and $Z \to e^+e^-$ channels for $|\eta| < 2.47$ excluding the transition region and endcaps and $20 < E_T < 50$ GeV. For $15 < E_T < 20$ GeV in $|\eta| < 0.8$ $W \to e\nu$ and $J/\psi \to e^+e^-$ are used. Taken from [53].

ID selection	channel	data (%)	MC (%)	ratio
medium	$W \to e\nu$	$94.1 \pm 0.2 \pm 0.6$	96.9	$0.971 \pm 0.002 \pm 0.007$
	$Z \to e^+e^-$	$94.7 \pm 0.4 \pm 1.5$	96.3	$0.984 \pm 0.004 \pm 0.015$
tight	$W \to e\nu$	$78.1 \pm 0.2 \pm 0.6$	77.5	$1.009 \pm 0.003 \pm 0.007$
	$Z \to e^+e^-$	$80.7 \pm 0.5 \pm 1.5$	78.5	$1.028 \pm 0.006 \pm 0.016$
medium	$W \to e\nu$	$75.8 \pm 8.8 \pm 8.1$	94.9	$0.80 \pm 0.09 \pm 0.07$
	$J/\psi \to e^+e^-$	$80.0 \pm 7.3 \pm 10.2$	81.9	$0.98 \pm 0.09 \pm 0.14$
tight	$W \to e\nu$	$61.9 \pm 6.0 \pm 7.0$	78.3	$0.79 \pm 0.08 \pm 0.09$
	$J/\psi \to e^+e^-$	$68.1 \pm 7.3 \pm 9.0$	69.1	$0.99 \pm 0.11 \pm 0.15$

Table 5.3: The reconstruction efficiencies given with and without the requirement for a Silicon hit integrated over $|\eta| < 2.47$ excluding the transition region and endcaps and over $20 < E_T < 50$ GeV. Taken from [53].

Reconstruction selection	data (%)	MC (%)	ratio
nominal reconstruction	$98.7 \pm 0.1 \pm 0.2$	98.3	$1.005 \pm 0.001 \pm 0.002$
track Silicon hit requirement	$94.3 \pm 0.2 \pm 0.8$	93.1	$1.013 \pm 0.002 \pm 0.008$

- ϵ_{ID} is the identification cuts efficiency,
- ϵ_{trig} is the trigger efficiency and
- ϵ_{isol} is the efficiency when isolation requirements applied to the electron candidate.

The identification efficiency is calculated by applying the *medium* and *tight* identification criteria on a probe sample. The efficiencies derived from data are compared to the ones derived from Monte Carlo and their ratio is extracted. The efficiencies and the data-Monte Carlo ratio are given in Table 5.2. Separate values are given for low energy electrons, $15 < E_T < 20$ GeV, using $J/\psi \to e^+e^-$ decays instead of $Z \to e^+e^-$. That is due to the low statistics in this E_{T} region, which necessitates the use of smaller part of the detector for this calculation.

The reconstruction efficiency is studied with respect to the sliding-window clusters in the EM calorimeter using the $Z \to e^+e^-$ decays. The resulting efficiencies are calculated in $|\eta| < 2.47$ for *medium* and *tight* electrons and are given in Table 5.3.

The trigger efficiency is estimated as the fraction of the offline electrons that pass a given trigger. By design, the triggers have a 100% efficiency in the E_{T} plateau region, if one compares them to *medium* or *tight* offline electrons. For the 2010 electron channel the `EF_e15_medium` trigger is used and for the 2011 a combined τ+e trigger, `EF_tau16_loose_e15_medium`. The first trigger is a single trigger firing whenever an electron with *medium* identification criteria and $E_{\text{T}} > 15$ GeV passes through the detector. The second trigger combines the former electron

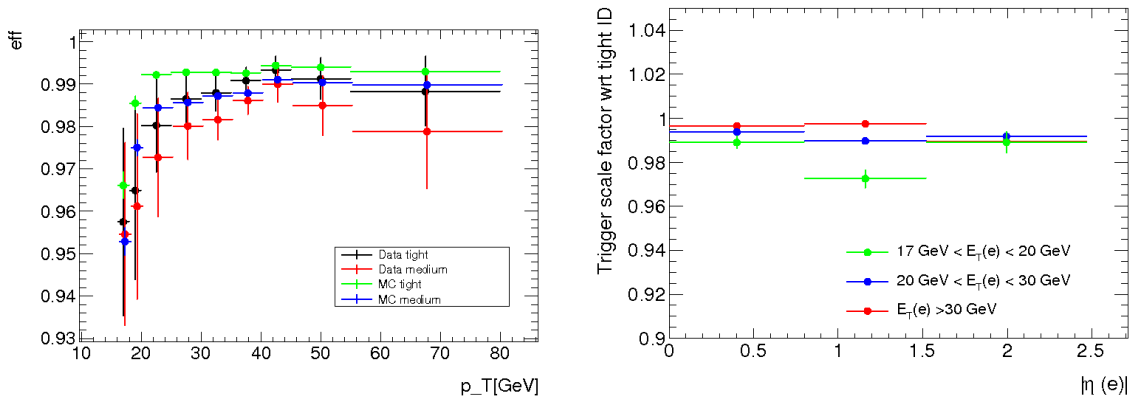


Figure 5.2: Scale factors for the electron trigger for the 2010 (left) and 2011 (right) $Z \to \tau\tau$ analysis. Taken from [55, 56], respectively.

Table 5.4: Electron trigger efficiency measured with respect to offline selected electrons for the 2010 analysis. Taken from [55].

trigger	EF_e15_medium
16 - 18 GeV	95.8 ± 2.2 (stat) ± 0.6 (syst)
18 - 20 GeV	96.5 ± 2.1 (stat) ± 0.4 (syst)
> 20 GeV	99.05 ± 0.22 (stat) ± 0.08 (syst)

trigger with a hadronic τ trigger. The latter triggers on hadronic τ candidates that pass the *loose* identification criteria, see Section 5.4.2, and have $E_{\mathrm{T}} > 16$ GeV. The trigger efficiencies in the 2010 data are calculated in $W \to e\nu$ decays with respect to offline electrons with $E_{\mathrm{T}} > 15$ GeV. In 2011, the electron trigger is part of the combined trigger, nonetheless its efficiency is measured independently in $Z \to e^+e^-$ decays and is later combined with the τ trigger efficiency [55, 56]. The scale factors are shown in Fig. 5.2 and in Table 5.4.

The isolation variables are tracker- and calorimeter- based. The former is defined as the scalar sum of the transverse momentum of tracks in a hollow cone of radius X = 0.2, 0.3, 0.4 about the seed lepton ($I^X_{p_{\mathrm{T}}}$). Similarly, the latter is calculated as the sum of the energy deposit on the EM calorimeter in a hollow cone about the seed lepton ($I^X_{E_{\mathrm{T}}}$). Often the isolation cone parameter is divided by the p_{T} or E_{T} of the candidate lepton to reduce the dependence on the momentum of the lepton. The isolation efficiency is strongly analysis dependent and thus it is estimated case by case. The isolation variables for the electron in the $Z \to \tau\tau$ analysis are $I^{0.3}_{E_{\mathrm{T}}}/E_{\mathrm{T}}$ for the 2010 analysis, $I^{0.4}_{E_{\mathrm{T}}}/E_{\mathrm{T}}$ for the 2011 and $I^{0.4}_{p_{\mathrm{T}}}/p_{\mathrm{T}}$. Scale factors are estimated with respect to *tight* electrons and are parametrised in E_{T} and η bins [55, 56]. The final values are shown in Fig. 5.3.

5.2 Muons

Muons are identified in the muon spectrometer, which provides a wide pseudorapidity coverage, $|\eta| < 2.7$, and a large momentum range, 3 GeV $< p_T <$ 1 TeV. An almost clean signal of

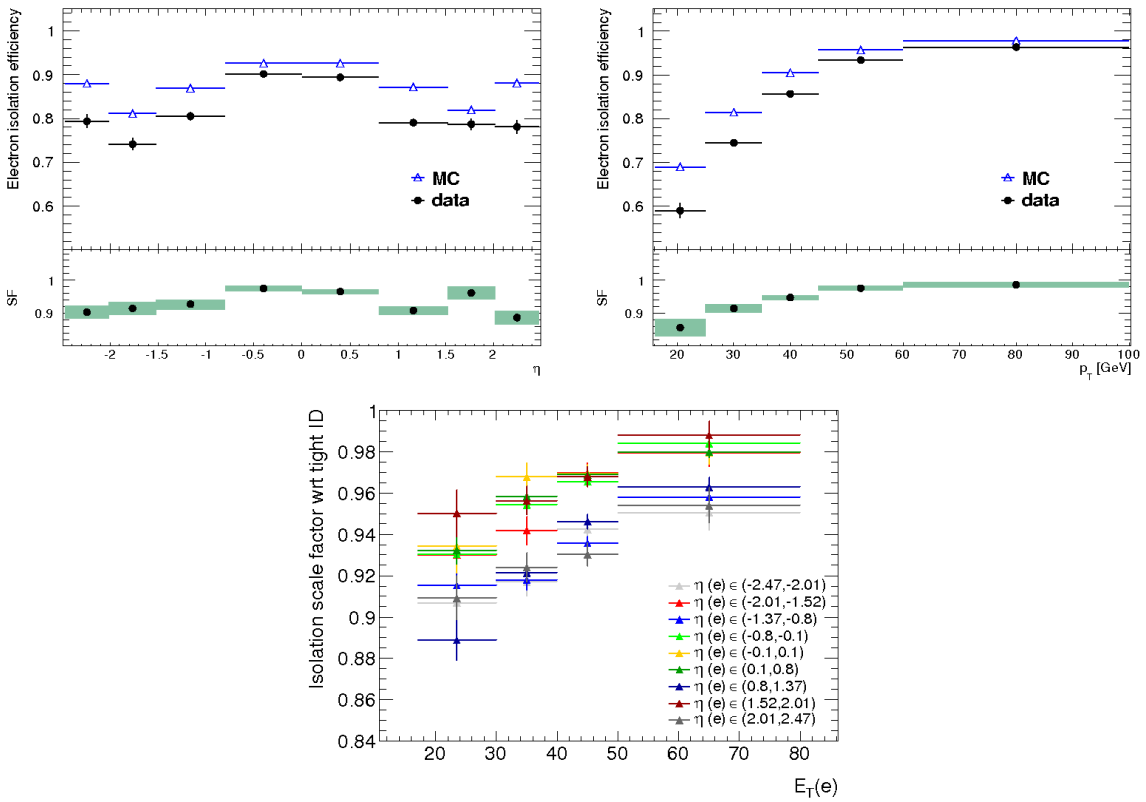


Figure 5.3: Isolation scale factors in bins of E_{T} and η for the 2010 (top) $Z \to \tau\tau$ analysis. In the bottom plot, the isolation scale factors for the analysis of the 2011 data with respect to E_{T} for different η ranges. Taken from [55, 56], respectively.

muons is detected in the spectrometer since electrons, photons and hadrons are absorbed by the calorimeter. A secondary tracking measurement is given by the Inner Detector for $|\eta| < 2.5$ and 30 GeV $< p_T <$ 200 GeV. Extra information is provided by the calorimeter whenever a muon looses energy in it.

In ATLAS, two algorithm chains are used for the identification and reconstruction of muons, i.e. STACO (chain 1) [57] and MuId (chain 2) [58]. For STACO muons, the momentum of the muon candidate needs to be measured in both the Inner Detector and the Muon Spectrometer and the final momentum is the weighted average of the two measurements. On the other hand, for the MuId muons a combined track fit is performed for all related hits in the Inner Detector and the Muon Spectrometer.

5.2.1 Muon Reconstruction and Identification

For the identification and reconstruction of muons three different algorithms are used: *standalone, combined* and *segmented tagged*. Here a brief summary of them is given and for more detailed description see [59, 60, 61].

The **standalone** algorithm uses tracks reconstructed in the spectrometer and extrapolates those to the beampipe. The tracks are reconstructed separately in each of the three segments of the spectrometer and then they are combined to form one track. The final track is corrected for energy losses in the calorimeters. The extrapolation to the beampipe slightly differs depending

on the chain algorithm, MuonBoy [57, 62] for STACO or Moore for MuId [63]. The standalone algorithm reconstructs muons at $|\eta| < 2.7$ and p_T more than a few GeV.

In the case of the **combined** algorithm (CB), muons reconstructed in the Muon Spectrometer are then combined with Inner Detector tracks to provide the final muon object. The Inner Detector tracks are identified in the pixel and SCT sub-detectors in a pseudorapidity range of $|\eta| < 2.5$. The difference between STACO and MuId algorithms appears when calculating the combined track vector. The first one performs a statistical combination of the inner and outer track vectors, while the latter produces a partial refit. In the second case, the algorithm starts from the inner track vector and adds the measurement from the outer track. The advantage of this method is that it allows for corrections due to multiple scattering and energy loss in the detector's material and magnetic field.

The last category of reconstructed muons are the **segmented tagged** (ST). They are Inner Detector tracks extrapolated to the spectrometer and matched to the closest segment in the spectrometer. In ATLAS there are two such algorithms, MuTag for STACO [57] and MuGirl for MuId [64]. As its name reveals, the MuTag defines a tag chi-square between the predicted extrapolated track and the near-by segments. Neither the tracks nor the segments should have been reconstructed by the STACO algorithm for this case. On the other hand, the MuGirl makes use of neural networks to define a discriminant. In contrast to MuTag, MuGirl finds and reconstructs all possible muons.

The combined muons are reconstructed with the highest purity of all three, since they form independently spectrometer tracks varying in η and ϕ. Their superiority is obvious in the $\eta \sim 0$ and $|\eta| \sim 1.2$ regions. The first is the region with almost no muon chambers leaving space for Inner Detector and calorimeter services. The second region is the transition region between barrel and endcaps, where there is only one chamber available. Nonetheless, the segmented muons are still very useful because they provide high efficiency when a muon candidate has not crossed enough chambers for precision measurements or has low momentum and hence, it cannot cross to the outer detector layer. Apart from the extreme cases where one or the other muon type performs better, most commonly a muon is reconstructed by both, the combined and standalone. Then, standalone muons are a separate category only when they are not matched to Inner Detector tracks. The segmented tagged and combined ones are different by construction for the STACO chain, while overlap removing is done in the case of MuId.

For this analysis, performance tests between STACO and MuId have been conducted with the 2010 data. No difference was found between the two chains thus the STACO muons are chosen for the final measurement.

5.2.2 Muon Efficiency and Resolution

The reconstruction efficiency for muons is calculated as a product of the efficiencies of the Inner Detector, the Muon Spectrometer and the matching between the two tracks

$$\epsilon^{reco}(CB + ST) = \epsilon^{ID}\epsilon^{comb}\epsilon^{MS}.$$

For the estimation of the individual efficiencies the tag-and-probe method is used. In $Z \to \mu^+\mu^-$ and $J/\psi \to \mu^+\mu^-$ decays, one track is selected as a combined muon, *tag muon*, and the other track, *probe muon*, must be either a standalone muon or an Inner Detector track. The first is used for the measurement of the Inner Detector efficiency and the second for the Muon Spectrometer and the track matching efficiencies. The $Z \to \mu^+\mu^-$ is used for measurements of muons with $p_T > 20$ GeV, while the $J/\psi \to \mu^+\mu^-$ for low p_{T} muons with Inner Detector $p > 3$ GeV. The efficiencies are binned in five pseudorapidity regions:

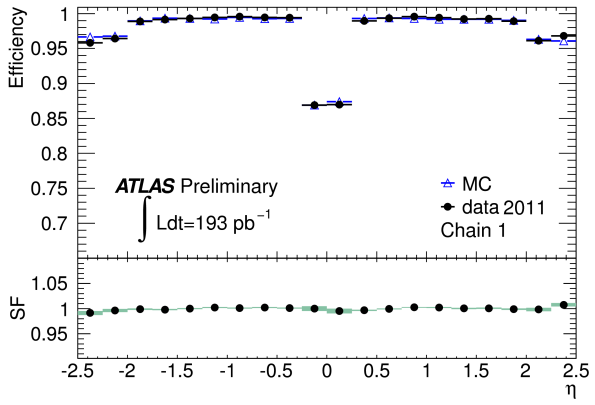


Figure 5.4: Combined muon reconstruction efficiency with respect to the inner tracking efficiency as a function of the pseudorapidity for muons with $p_T > 20$ GeV. The lower part of the plot shows the ratio between data measurements and Monte Carlo predictions. Taken from [55].

- $|\eta| < 0.1$, crack region,
- $0.1 < |\eta| < 1.1$, barrel region,
- $1.1 < |\eta| < 1.3$, transition region between barrel and endcaps,
- $1.3 < |\eta| < 2.0$, endcap region and
- $2.0 < |\eta| < 2.5$, forward region.

The resulting efficiencies and scale factors are shown in Fig. 5.4 with respect to η bins. In p_{T} bins the efficiency is 100% and the scale factor 1.0.

The momentum resolution of muons is parametrised in p_{T}, η and ϕ [61]. The Inner Detector resolution is calculated as the quadratic sum of a term describing the multiple scattering contribution ($a_{ID}(\eta,\phi)$) and a term for the intrinsic resolution given by the magnetic field in the Inner Detector, the spatial resolution of the detector components and the residual misalignment of them ($b_{ID}(\eta,\phi)$):

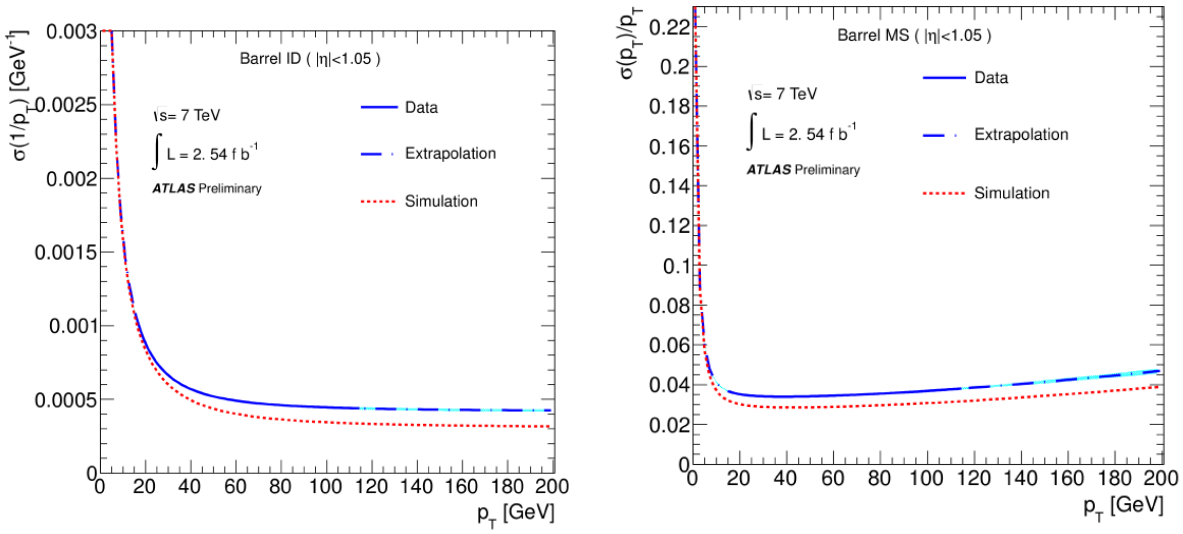
$$\frac{\sigma_{ID}(p_T)}{p_T} = a_{ID}(\eta,\phi) \oplus b_{ID}(\eta,\phi) \cdot p_T. \tag{5.5}$$

The momentum resolution for the standalone muons is given by

$$\frac{\sigma_{SA}(p_T)}{p_T} = a_{MS}(\eta,\phi) \oplus b_{MS}(\eta,\phi) \cdot p_T \oplus \frac{c(\eta,\phi)}{p_T}, \tag{5.6}$$

where the first two terms are similar to eq. (5.5) but for the Muon Spectrometer and the third term accounts for the fluctuations of the muon energy loss in the calorimeters which needs to be added to the Muon Spectrometer momentum to get the muon momentum at the pp interaction point.

The $a_{ID}(\eta,\phi)$ term is constrained by measurements and agrees with the Monte Carlo prediction within 5% in the barrel and 10% in the endcaps. The $b_{ID}(\eta,\phi)$ term is taken by the resolution of the $Z \to \mu^+\mu^-$ di-muon mass distribution. The b term for the Muon Spectrometer is estimated from a special dataset where the toroidal magnetic field is turned off. The resulting value for $b_{MS}(\eta,\phi)$ is 0.2 TeV^{-1} in the barrel and the MDT endcap and 0.4 TeV^{-1} in



(a) ID barrel region

(b) MS barrel region

Figure 5.5: Muon Momentum resolution as estimated from a template fit to the $Z \to \mu^+\mu^-$ invariant mass. In the left is shown the resolution in the barrel region for the Inner Detector and in the right the Muon Spectrometer. For the specific plots the MuID algorithm is used but the results are similar for the STACO algorithm. Taken from [65].

the CSC endcap region. The energy loss term of eq. (5.6) is significant only for low transverse momenta ($\lesssim$ 15 GeV) hence cosmic-ray data and $J/\psi \to \mu^+\mu^-$ decays are used to estimate this term. Comparisons with Monte Carlo predictions show good agreement for the $c(\eta, \phi)$ term with 10-20% systematic uncertainties.

Concluding the total muon momentum resolution is found to be about 10% worse in data than in the Monte Carlo, due to mis-alignments of the Inner Detector and the Muon Spectrometer as it is shown as an example in Fig.5.5 for the data collected in 2011.

5.2.3 Muon Trigger and Isolation Efficiencies

The muon trigger efficiencies are evaluated with the tag-and-probe method in $Z \to \mu^+\mu^-$ decays. For the 2010 analysis three triggers are used depending on the run conditions: `EF_mu10_MG`, `EF_mu13_MG` and `EF_mu13_MG_tight`. The prefix `EF` in the trigger names means that it is an EF trigger. The suffix `MG` refers to the reconstruction algorithm for the triggered muons which is Muon Girl. A muon fires the first trigger if it has $p_{\mathrm{T}} > 10$ GeV and the last two if it has $p_{\mathrm{T}} > 13$ GeV. In all cases the trigger efficiency is found to be close to 95% in the endcap region and 80% in the barrel for offline muon $p_{\mathrm{T}} > 20$ GeV. The efficiency plots are similar for all three triggers, hence the one for `EF_mu13_MG` trigger is given as representative in Fig 5.6 [55].

In 2011, the `EF_mu15i` and `EF_mu15i_medium` triggers are used. The triggers fire when an isolated muon with $p_{\mathrm{T}} > 15$ GeV traverses the detector. The isolation in the trigger is done at the Level2 of the trigger system to ensure low Event Filter rates. There is track and a calorimeter isolation requirement. The track parameter is defined in a hollow cone around the lepton track of $\Delta R = 0.2$ and it is divided by the p_{T} of the track, $I_{p_{\mathrm{T}}}^{0.2}/p_{\mathrm{T}}$. A cut is implemented at $I_{p_{\mathrm{T}}}^{0.2}/p_{\mathrm{T}} < 0.05$. The calorimeter isolation is defined as the sum of the energy in a hollow

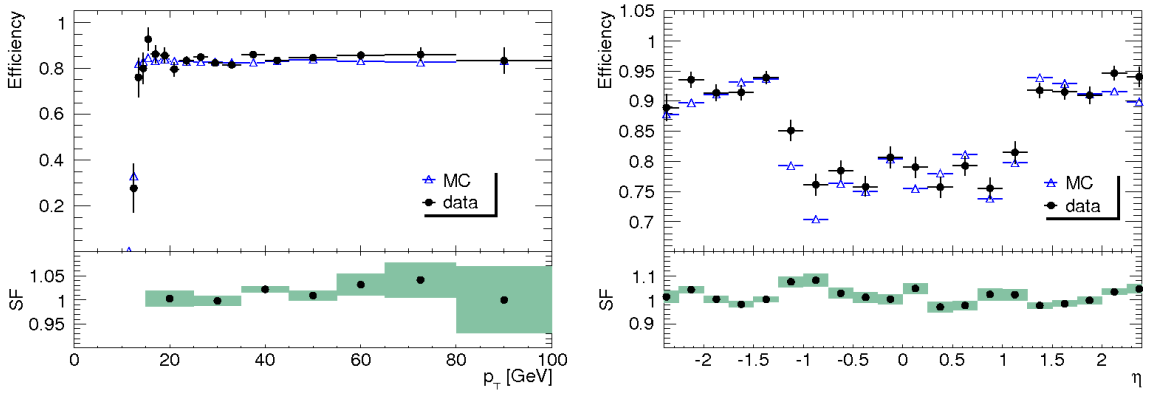


Figure 5.6: Efficiency of the EF_mu13_MG as a function of p_{T} and η and the corresponding scale factors for the 2010 analysis. Taken from [55].

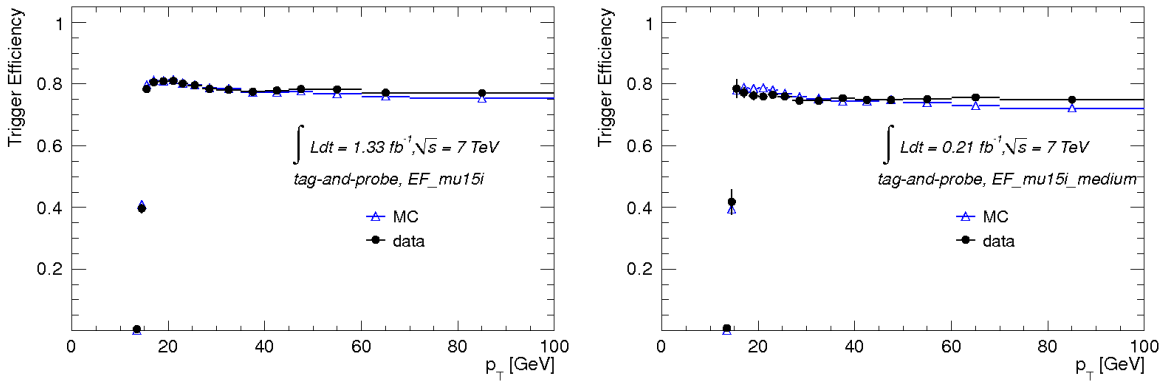


Figure 5.7: Trigger efficiency of the EF_mu15i (left) and EF_mu15i_medium (right) with respect to the (p_{T}) of the probe muon. Taken from [56].

cone around the lepton tower of $\Delta R = 0.4$, $I_{E_{\mathrm{T}}}^{0.4}$. A cut at $I_{E_{\mathrm{T}}}^{0.4} < 0.05$ is applied. The difference between the two triggers is the L1 seed; the first has a p_{T} threshold of 10 GeV (`L1_MU10`) and the second of 11 GeV (`L1_MU11`). In Monte Carlo the `EF_mu15i` is correctly simulated but the `EF_mu15i_medium` is absent. To simulate the latter trigger `EF_mu15i` and `L1_MU11` are combined. Trigger efficiencies are determined using the $Z \to \mu^+\mu^-$ tag-and-probe method for the two triggers, the `EF_mu15i` and `EF_mu15i_medium`. An small $\eta - \phi$ dependence is found in the scale factors, specially in the barrel-endcap transition region. The final scale factors used for rescaling the Monte Carlo are p_T dependant and close to unity as shown in Fig. 5.7 [56].

The muon isolation variables are defined in the same way as the electron ones. For the $Z \to \tau\tau$ analysis, the $I_{E_{\mathrm{T}}}^{0.3}/p_{\mathrm{T}}$ and $I_{p_{\mathrm{T}}}^{0.4}/p_{\mathrm{T}}$ are used. The scale factors are estimated in $Z \to \mu^+\mu^-$ decays and are shown in Fig. 5.8.

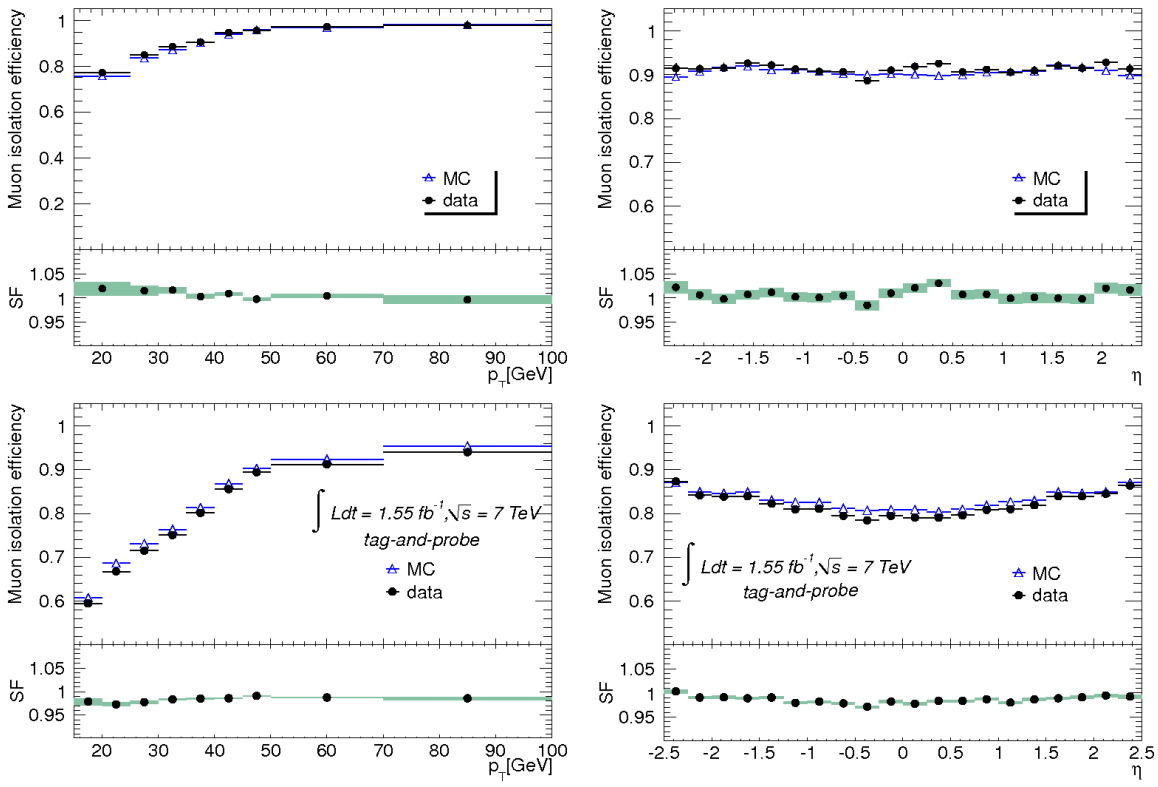


Figure 5.8: Isolation efficiency for the 2010 data (top) and 2011 (bottom) parametrised in p_{T} and η. Taken from [55, 56], respectively.

5.3 Jets and Missing Transverse Energy

5.3.1 Jets

A hadronic jet is reconstructed by a jet algorithm that combines the energy depositions in the electromagnetic and the hadronic calorimeters [59, 66]. The energy reconstruction is done at the electromagnetic energy scale (EM), i.e. the energy deposited by the particles of an electromagnetic shower in the calorimeter. The four-momentum of a jet is reconstructed from the corrected energy and angles with respect to the primary event vertex.

Jet Calibration

The energy measurement is calibrated to associate the jet energy measurement on the calorimeter to the true energy of the stable particles that created the jet. The calibration aims at correcting for detector effects, such as:

- not full measurement of the energy deposited by hadrons on the calorimeter;
- energy losses in inactive regions of the detector;
- energy leakage outside the calorimeters;
- energy deposits of true jets falling outside the reconstruction jet cone; and

- noise thresholds and reconstruction efficiency.

The calibration formula used for the first data is

$$E^{\text{jet}}_{\text{calib}} = \frac{E^{\text{jet}}_{\text{meas}}}{\mathcal{F}_{\text{calib}}(E^{\text{jet}}_{\text{meas}})}, \quad \text{where} \quad E^{\text{jet}}_{\text{meas}} = E^{\text{jet}}_{\text{EM}} - \mathcal{O}(N_{PV}). \tag{5.7}$$

The $E^{\text{jet}}_{\text{calib}}$, $E^{\text{jet}}_{\text{meas}}$, $E^{\text{jet}}_{\text{EM}}$ are the calibrated, measured and at EM scale jet energies, respectively, and $\mathcal{O}(N_{PV})$ is the correction for energy coming from multiple pp interactions depending on the number of primary vertices (N_{PV}). In principle, the jet energy scale (JES) corrections are applied on the reconstructed jet energy that is measured at the EM calorimeter (EM+JES calibration scheme). In this way the systematic uncertainty is evaluated from a single hadron response measurement and the systematic variations are taken from Monte Carlo simulations. A more refined method is used for the calibration of hadronic τ candidates, that is the local cluster weighting (LCW) [67]. In LCW method, the calorimeter cells, which are topologically connected, are clustered together based on single pion Monte Carlo events. The clusters are then corrected for all the detector effects without associating them to a jet algorithm, hence the naming *local*. Finally the jets are built from the calibrated clusters using a jet algorithm.

The anti-k_t algorithm [68] with distance parameter[2] $R = 0.4$ or 0.6 is used for the reconstruction of the jets. Topological clusters (*topoclusters*) or calorimeter *towers* are the input for the jet algorithm. A *topocluster* is a group of calorimeter cells that follows the shower development looking for local maxima. The calorimeter *towers*, on the other hand, have constant dimensions $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ and follow the cell real dimensions.

Jet Quality Selection Criteria

Before any hadronic jet is reconstructed and calibrated, quality cuts are applied to suppress non-collision background events [66, 69]. These cuts are given in Table 5.5 and the variables are explained here for the *medium* selection, since it is the one used for the $Z \to \tau\tau$ analysis. An event is rejected if:

1. ($f_{\text{HEC}} > 0.5$ AND $|f_{\text{HECquality}}| > 0.5$) OR $|E_{\text{neg}}| > 60$ GeV: the fraction of the jet energy deposited in the HEC calorimeter is greater than 50% and the fraction of the HEC calorimeter contribution to the jet energy from cells with abnormal signal shape is greater than 50%; or the jet has negative energy greater than 60 GeV; or

2. $f_{\text{HEC}} > 1 - |Q_{\text{HEC}}|$: the fraction of the jet energy deposited in the HEC calorimeter is greater than the fraction of the HEC calorimeter contribution to the jet energy from cells with normal signal shape;

3. $f_{\text{EM}} > 0.9$ AND $|f_{\text{quality}}| > 0.8$ AND $|\eta| < 2.8$: the jet is central and more than 95% of the jet energy is deposited in the EM calorimeter and the fraction of the total jet energy coming from cells with an abnormal signal shape is greater than 80%;

4. $f_{\text{EM}} < 0.05 \text{AND} |\eta| \geq 2$: the jet is not central and at most 5% of the jet energy is deposited in the EM calorimeter; or

[2]The main principle upon the jet algorithms are based is that partons are produced in a hard process and combine into collinear pairs of partons, as discussed already in Chapter 4. A jet algorithm decides whether two proto-jets originate from the same parton leaving the hard process. This decision is taken by defining a collinearity measure in distance space, R, and the transverse momentum of one proto-jet with respect to another or to the beam axis. The anti-k_t algorithm uses the following relation: $y_{ij} = \frac{\Delta R_{ij}}{D} min(p_{T,i}^{-1}, p_{T,j}^{-1}) \quad y_{iB} = p_{T,i}^{-1}$, where D is a measure of the angular size of the jet and can take values between 0.4 and 1.5.

Table 5.5: *Loose* and *Medium* jet quality selection criteria used to reject fake jets and non-collision events. Taken from [66].

	loose	*medium*
HEC spikes	($f_{HEC} > 0.5$ and $\|f_{HECquality}\| > 0.5$) or $\|E_{neg}\| > 60$ GeV	Loose or $f_{HEC} > 1 - \|f_{HECquality}\|$
Coherent EM noise	$f_{EM} > 0.95$ and $f_{\text{quality}} > 0.8$ and $\|\eta\| < 2.8$	Loose or $f_{EM} > 0.9$ and $f_{\text{quality}} > 0.8$ and $\|\eta\| < 2.8$
Non-collision background	$\|t_{\text{jet}}\| > 25$ ns ($f_{EM} < 0.05$ and $f_{ch} < 0.05$ and $\|\eta\| < 2$) or ($f_{EM} < 0.05$ and $\|\eta\| \geq 2$) or ($f_{max} > 0.99$ and $\|\eta\| < 2$)	Loose or $\|t_{\text{jet}}\| > 10$ ns or ($f_{EM} < 0.05$ and $f_{ch} < 0.1$ and $\|\eta\| < 2.$) or ($f_{EM} > 0.95$ and $f_{ch} < 0.05$ and $\|\eta\| < 2$)

5. $|f_{\text{max}}| > 0.99$ AND $|\eta| < 2$: the jet is central and the fraction of the maximum total energy in a single calorimeter layer is higher than 99%;

6. $|t_{\text{jet}}| < 10$ ns: the jet primarily contains cells with energy deposited less than 10 ns before or after the nominal proton-proton collision time; or

7. $|f_{\text{ch}}| < 0.1$ AND $f_{\text{EM}} < 0.05$ AND $|\eta| < 2$: the sum of the transverse momentum of the tracks associated to a central jet is less than 10% of the total jet energy and at most 5% of the jet energy is deposited in the EM calorimeter; or

8. $|f_{\text{ch}}| < 0.05$ AND $f_{\text{EM}} > 0.95$ AND $|\eta| < 2$: the sum of the transverse momentum of the tracks associated to a central jet is less than 5% of the total jet energy and at least 5% of the jet energy is deposited in the EM calorimeter.

Jet Energy Calibration (EM+JES scheme)

As mentioned above the EM+JES calibration scheme allows to apply corrections to the energy measured as a function of jet p_{T} and η at the EM scale. The first step is to correct for pile-up effects; that is energy deposits from multiple interactions per bunch crossing (in-time pile-up). For the 2010 data run, the average additional energy due to pile-up is subtracted from the calorimeter-measured energy using constants that are estimated with *in situ* measurements in the calorimeters. In 2011, this correction is not applied, because the out-of-time pile-up (energy deposits from multiple interactions from previous bunches) becomes important. In this case, the pile-up contribution is estimated with event re-weighting of the Monte Carlo to the observed average interaction vertices per event. The next step is to correct the direction of the jet. That is that the reconstructed jet is coming from the primary vertex of the interaction and not the geometrical centre of the detector. Finally, the jet four-vector is reconstructed and corrected with constants derived from Monte Carlo truth studies. The systematic uncertainties evaluated for the EM+JES calibration scheme with 2010 data is shown in Table 5.6. The uncertainties are given in bins of p_{T} and η. Amongst the systematic uncertainty sources that are studied are Monte Carlo modelling, out-of-time pile-up, JES calibration and detector effects, for details see [66].

The JES measurement in 2011 data is checked with the p_{T} balance method [70]. Events with exactly one Z boson and one jet should have their momentum balanced in the transverse plane. The p_{T} balance is used for the jet energy calibration. If p_{T}^{jet} and p_{T}^{Z} the transverse momenta of

Table 5.6: Maximum EM+JES jet energy scale uncertainties for different p_{T} and η regions from Monte Carlo studies on anti-k_t jets with $R = 0.4$. Taken from [66].

η region	maximum fractional JES uncertainty (%)		
	$p_T^{jet} = 20$ GeV	200 GeV	1.5 TeV
$0.0 < \|\eta\| < 0.3$	4.1	2.3	3.1
$0.3 < \|\eta\| < 0.8$	4.3	2.4	3.3
$0.8 < \|\eta\| < 1.2$	4.4	2.5	3.4
$1.2 < \|\eta\| < 2.1$	5.3	2.6	3.5
$2.1 < \|\eta\| < 2.8$	7.4	2.7	-
$2.8 < \|\eta\| < 3.2$	9.0	3.3	-
$3.2 < \|\eta\| < 3.6$	9.3	3.5	-
$3.6 < \|\eta\| < 4.5$	13.4	4.9	-

the jet and the Z respectively, then their ratio, $p_{\mathrm{T}}^{jet}/p_{\mathrm{T}}^{Z}$, is used for the energy calibration. In Fig. 5.9, the final calibration results for part of the 2011 data is shown. The data is compared to PYTHIA Monte Carlo and within uncertainties they agree well.

Jet Energy Resolution (JER)

The width of the Gaussian fit of the jet p_{T} distribution is used to estimate the jet energy resolution (JER). JER is studied using two methods: the di-jet balance and the bi-sector technique [71]. The first one takes advantage of the fact that in a di-jet event the momentum of one jet recoils against the momentum of the second one. The bi-sector technique, on the other hand, assumes that the fluctuation in the ψ and η is uniform at particle level, where (ψ, η) is the transverse plane in a di-jet event and η is the direction that bisects the angle between the two jets, see Fig. 5.10. For either method, the fractional jet transverse resolution is given by

$$\frac{\sigma_{p_{\mathrm{T}}}}{p_{\mathrm{T}}} = \frac{N}{p_{\mathrm{T}}} \oplus \frac{S}{\sqrt{p_{\mathrm{T}}}} \oplus \mathcal{C}, \tag{5.8}$$

where N, S and $\mathcal{C}$ are the noise, stochastic and constant terms, respectively. The Monte Carlo results are fitted to the equation and compared to data. In Fig. 5.11 a good agreement between data (2010(a) and 2011(b)) and Monte Carlo predictions is found for both methods.

5.3.2 Missing Transverse Energy ($E_{\mathrm{T}}^{\mathrm{miss}}$)

The missing energy in ATLAS [59, 72] is defined as the sum of the calorimetric component ($E^{\mathrm{miss,calo}}$) and the muon one ($E^{\mathrm{miss},\mu}$)

$$E_{x(y)}^{\mathrm{miss}} = E_{x(y)}^{\mathrm{miss,calo}} + E_{x(y)}^{\mathrm{miss},\mu}. \tag{5.9}$$

From the x- and y- components, the missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) and its azimuthal angle (ϕ^{miss}) are calculated as

$$E_T^{\mathrm{miss}} = \sqrt{(E_x^{\mathrm{miss}})^2 + (E_y^{\mathrm{miss}})^2}, \tag{5.10}$$

$$\phi^{\mathrm{miss}} = \arctan\left(E_y^{\mathrm{miss}}, E_x^{\mathrm{miss}}\right). \tag{5.11}$$

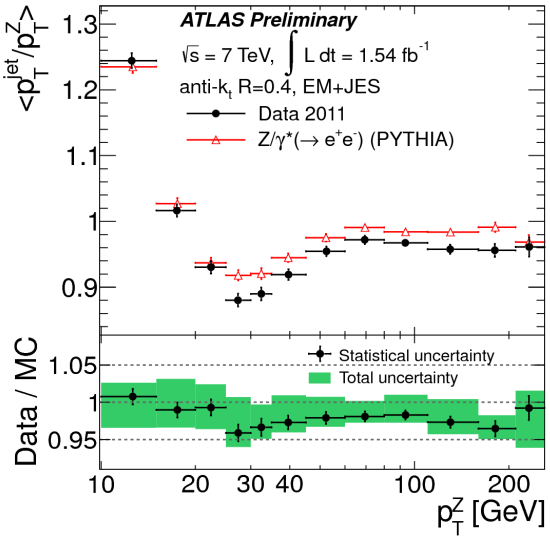


Figure 5.9: Mean p_{T} balance for jets and data-to-MC ratio with statistical and systematic uncertainties. Taken from [70].

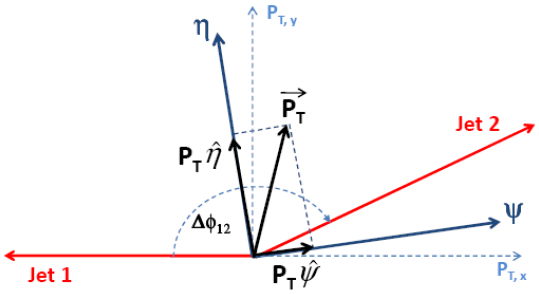


Figure 5.10: Sketch of the coordinates used in the bi-sector technique. Taken from [71].

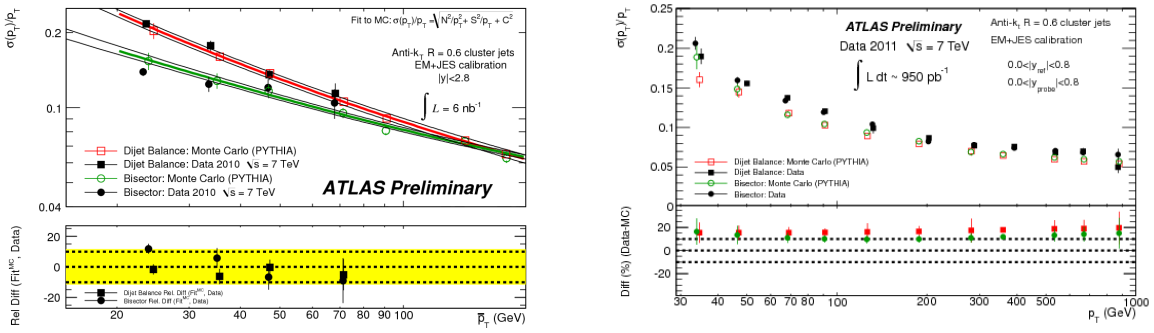
The calorimetric part of the missing energy consists of the calorimeter cells associated to the reconstructed objects, electrons, photons, hadronic τ candidates, jets and muons. The cells that are not associated to any object are summed in the $E_T^{\mathrm{miss,CellOut}}$ term, which is important for the missing energy resolution estimation. Each term of the missing energy is given by

$$E_x^{\mathrm{miss,term}} = -\sum_{i=1}^{N_{\mathrm{cell}}^{\mathrm{term}}} E_i \sin\theta_i \cos\phi_i, \quad E_y^{\mathrm{miss,term}} = -\sum_{i=1}^{N_{\mathrm{cell}}^{\mathrm{term}}} E_i \sin\theta_i \sin\phi_i, \tag{5.12}$$

where E_i, θ_i and ϕ_i are the energy, the polar and azimuthal angle, respectively. The total $E_{x(y)}^{\mathrm{miss,calo}}$ is calculated as

$$E_{x(y)}^{\mathrm{miss,calo}} = E_{x(y)}^{\mathrm{miss},e} + E_{x(y)}^{\mathrm{miss},\gamma} + E_{x(y)}^{\mathrm{miss},\tau} + E_{x(y)}^{\mathrm{miss,jets}} + E_{x(y)}^{\mathrm{miss,softjets}} + (E_{x(y)}^{\mathrm{miss,calo},\mu}) + E_{x(y)}^{\mathrm{miss,CellOut}}. \tag{5.13}$$

Each term of eq. (5.13) is reconstructed as follows:



(a) (b)

Figure 5.11: Jet energy resolution for the di-jet balance and bi-sector methods as a function of the average p_{T} of a jet. Lower plot shows the relative difference between data and Monte Carlo and the uncertainty on each point. In (a) the results with 2010 data, taken from [71], and in (b) with the 2011 data, taken from [70].

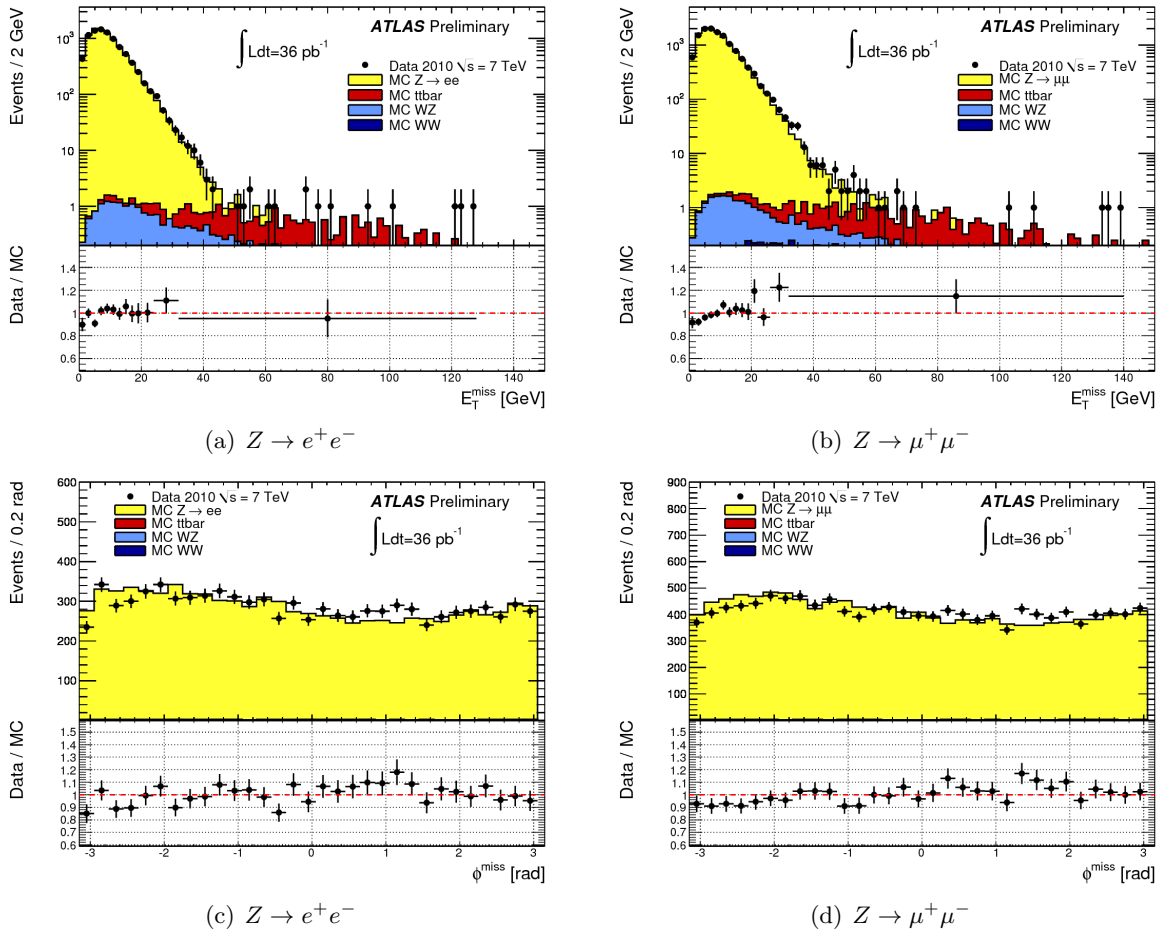
- $E_{x(y)}^{\mathrm{miss},e}$, $E_{x(y)}^{\mathrm{miss},\gamma}$ and $E_{x(y)}^{\mathrm{miss},\tau}$ from three-dimensional topoclusters of electrons, photons and τ leptons, respectively;
- $E_{x(y)}^{\mathrm{miss,jets}}$ and $E_{x(y)}^{\mathrm{miss,softjets}}$ are taken from cells in clusters associated to jets with calibrated $p_T > 20$ GeV or $7 < p_T < 20$ GeV, respectively;
- $E_{x(y)}^{\mathrm{miss,calo},\mu}$ is the energy loss of the muons in the calorimeter;
- $E_{x(y)}^{\mathrm{miss,CellOut}}$ is taken from cells of topoclusters which are not attributed to any of the reconstructed objects.

For the calibration, the default scheme for each object is used and it is compared to the default selection cuts, defined as optimal in each case [72].

The muon part of the missing energy is given by

$$E_{x(y)}^{\mathrm{miss},\mu} = -\sum_{\mathrm{muons}} p_{x(y)}^{\mu} \tag{5.14}$$

in the pseudorapidity range of $|\eta| < 2.5$, using only muons with matched tracks in the Muon Spectrometer and the Inner Detector. Well-reconstructed muons reduce the number of fake muons considered. To avoid double counting of minimal ionising particles in the calorimeters, isolated and non-isolated muons are treated differently in the calculation of the muon missing energy term. Non-isolated muons are muons lying within a cone of $\Delta R < 0.3$ about a reconstructed jet. In the case of an isolated muon, the $E_{x(y)}^{\mathrm{miss,calo},\mu}$ is not added in eq. (5.13), since the p_{T} of the muons is measured in the Inner Detector and the Muon Spectrometer taking into account the energy deposited in the calorimeters [59]. If the muon is non-isolated then the energy deposited in the calorimeter cannot be distinguished from the calorimetric energy depositions of the particles in the jet. Then, the measurement in the spectrometer is used to estimate the energy loss in the calorimeter and the term $E_{x(y)}^{\mathrm{miss,calo},\mu}$ is added in eq. (5.13). For measurements outside the fiducial volume of the spectrometer and at the transition regions the segmented muons are used.



(a) $Z \to e^+e^-$ (b) $Z \to \mu^+\mu^-$

(c) $Z \to e^+e^-$ (d) $Z \to \mu^+\mu^-$

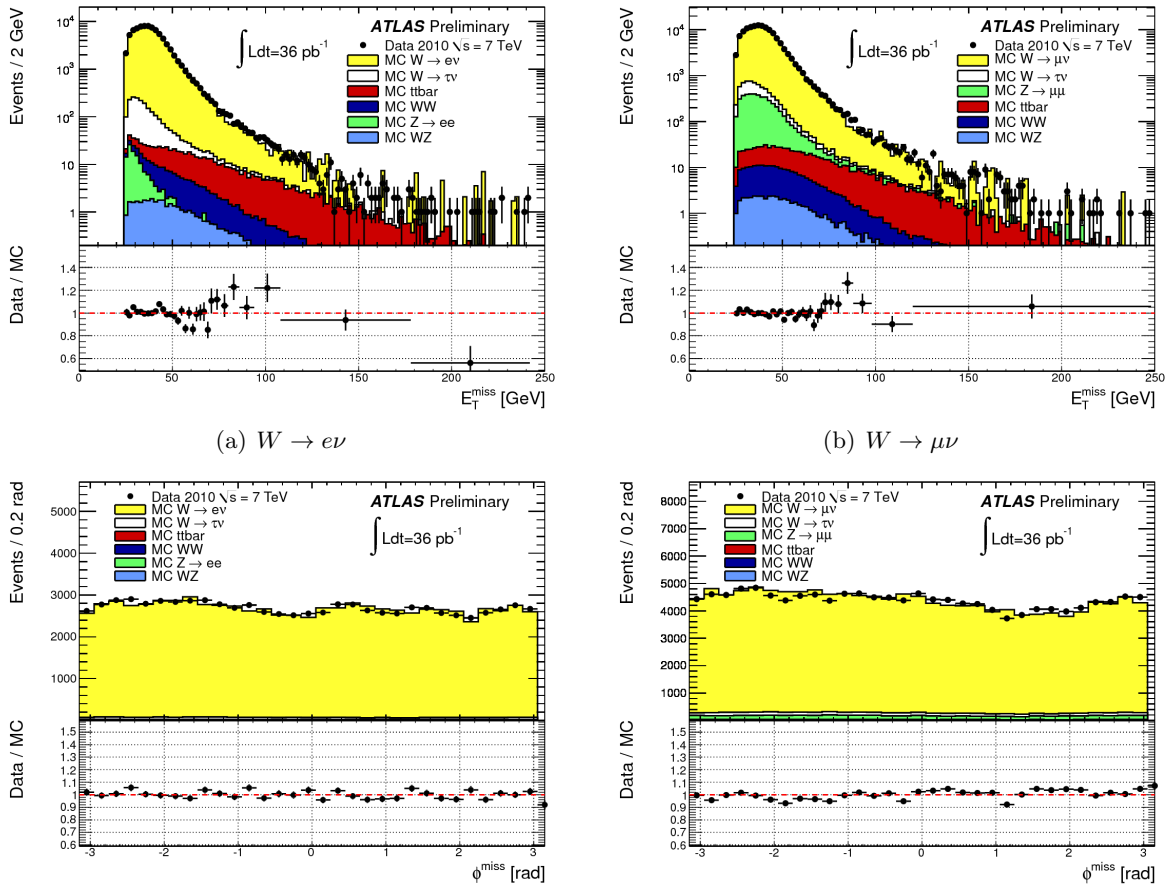
Figure 5.12: Distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$ (top) and ϕ^{miss} (bottom) as measured in $Z \to e^+e^-$ (left) and $Z \to \mu^+\mu^-$ (right) 2010 data. Monte Carlo expectations are normalised to the cross section and luminosity of the processes studied. Taken from [73].

Missing Energy Performance and Resolution

The performance of the missing energy is checked in minimum bias, dijet and $Z \to \ell\ell$ events where no real missing energy is expected. In all cases a relatively good agreement between data and Monte Carlo is seen except for the tails of the E_x^{miss} and E_y^{miss} distributions, see Fig. 5.12. The mis-matching in the tails is mainly due to poor Monte Carlo statistics and fake muon events that are rejected with tighter cuts at analysis level [72].

The scale of the $E_{\mathrm{T}}^{\mathrm{miss}}$ is investigated in $W \to e\nu$ and $W \to \mu\nu$ events, where a lot of missing energy is expected due to the neutrinos of the processes. The $E_{\mathrm{T}}^{\mathrm{miss}}$ and ϕ^{miss} distributions for the $W \to \ell\nu$ events are shown in Fig. 5.13, where a good agreement between data and Monte Carlo predictions is observed. The overall systematic uncertainty is found to be on average 2.6% in $W \to \ell\nu$ events with the value increasing for high $\sum E_T$. The linearity of the $E_{\mathrm{T}}^{\mathrm{miss}}$ is given by the ratio

$$\frac{E_{\mathrm{T}}^{\mathrm{miss}} - E_T^{\mathrm{miss,True}}}{E_T^{\mathrm{miss,True}}}.$$



(a) $W \to e\nu$ (b) $W \to \mu\nu$

(c) $W \to e\nu$ (d) $W \to \mu\nu$

Figure 5.13: Distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$ (top) and ϕ^{miss} (bottom) as measured in $W \to e\nu$ (left) and $W \to \mu\nu$ (right) data. Monte Carlo expectations are normalised to the cross section and luminosity of the processes studied. Taken from [73].

The mean value is expected to be zero for the case of the perfect alignment. In Fig. 5.14 (left) a deviation at low $E_{\mathrm{T}}^{\mathrm{miss}}$ values is observed, due to the finite resolution of the $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement. The bias in muon decays in contrast to electron ones is probably due to an underestimation of the $E_T^{\mathrm{miss,calo},\mu}$ term [73].

The resolution of the $E_{\mathrm{T}}^{\mathrm{miss}}$ is evaluated in events with no real missing energy, as in the case of the performance studies. The E_x^{miss} and E_y^{miss} components are equal to zero and the resolution is taken from the width of the combined distribution of E_x^{miss} and E_y^{miss} in bins of $\sum E_T$. Fits of the distribution for the different processes result in a resolution given by the formula

$$\sigma = k \cdot \sqrt{\sum E_T}, \tag{5.15}$$

where the parameter k is about 0.5 $\mathrm{GeV}^{1/2}$. The result is shown in Fig. 5.14 (right) for studies with the 2010 data.

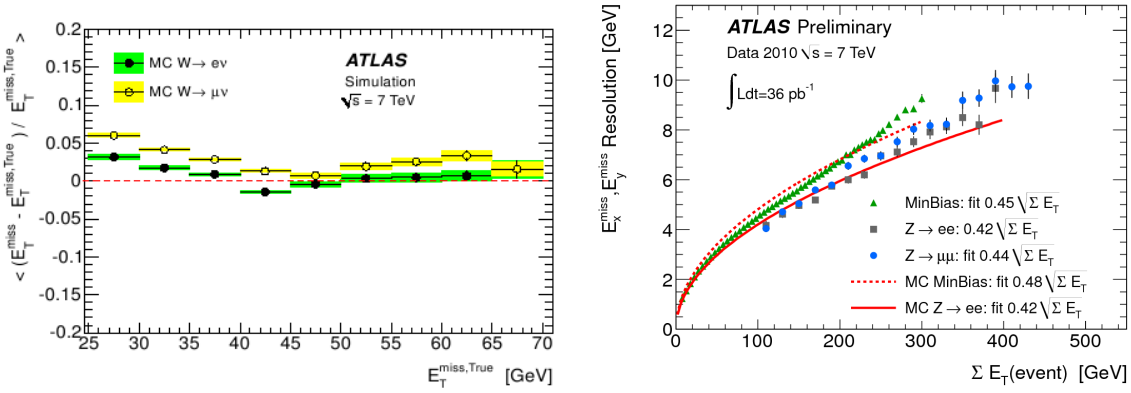


Figure 5.14: Left: Linearity of $E_{\mathrm{T}}^{\mathrm{miss}}$ with respect to true $E_{\mathrm{T}}^{\mathrm{miss}}$ as observed in $W \to e\nu$ and $W \to \mu\nu$ events [72]. Right: E_x^{miss} and E_y^{miss} resolution as a function of the total transverse energy in the event for data collected in the 2010 run. The resolution is fitted with eq. (5.15). Taken from [73].

5.4 Tau Leptons

The τ leptons differ from the electrons and the muons because they decay further to light leptons or hadrons. Their mass is 1776.82 ± 0.16 MeV and their mean life time 2.9×10^{-13} seconds [4]. That means that they decay before they leave the beam pipe. τ leptons decay to hadrons with a BR = 64.7% and to lighter leptons with a BR = 35.3%. In Table 5.7 the detailed decay modes and their branching ratios are given.

The fact that the τ lepton has many decay modes makes it very difficult to identify, specially in hadronic collisions, like at LHC. In particular, the leptonic decays are impossible to distinguish from primary electrons and muons, hence, only the hadronic modes are considered. From the hadronic decays mainly the ones with one or three charged pions at their final state are reconstructed and are usually referred to as one- and three-prong τ candidates. Moreover, the hadronic τ candidates do not provide a clean signature as they can easily be misinterpreted for QCD multijet production. In order to discriminate the hadronic τ signal from the multijet background, variables based on the shape of the calorimetric shower and the track multiplicity are used in either a cut-based method or more sophisticated ones, like likelihood or boosted decision trees.

A hadronic τ event needs to be well isolated both in the Inner Detector and the calorimeters. In the Inner Detector the single- or three-prong charged tracks are reconstructed. The single track should not be associated to any muon segment nor carry any of the electron characteristics. The three-track system, on the other hand, needs to be well collimated in (η, ϕ) and the invariant mass should be well below the τ lepton mass. The charge of the τ lepton is estimated from the charged tracks.

In the calorimeter, the energy of the visible decay products, i.e. all decay products but the neutrinos, is deposited. The shower produced is rather narrow and almost all deposited in the electromagnetic calorimeter. About 55% of the energy comes from π^0s. Each cluster should be matched to tracks close to the impact point of the latter and the mass of the cluster should be well below the mass of the τ lepton.

Table 5.7: Decay modes of the τ lepton [4].

Decay Modes	Branching Ratio [%]
Leptonic Decay Modes	
$\mu\,\nu_\mu\,\nu_\tau$	17.4
$e\,\nu_e\,\nu_\tau$	17.9
Hadronic Decay Modes	
$\pi^\pm\,\pi^0\,\nu_\tau$	25.5
$\pi^\pm\,\nu_\tau$	10.9
$\pi^\pm\,\pi^0\,\pi^0\,\nu_\tau$	9.3
$\pi^\pm\,\pi^0\,\nu_\tau$	9.3
$\pi^\pm\,\pi^\pm\,\pi^\mp\,\nu_\tau$	1.0
$\pi^\pm\,\pi^0\,\pi^0\,\pi^0\,\nu_\tau$	4.6
$\pi^\pm\,\pi^\pm\,\pi^\mp\,\pi^0\,\nu_\tau$	1.5
5 charged pions	< 0.1
$K^\pm\,(nK^0_s)\,(n\pi^0)\,\nu_\tau$	1.5

5.4.1 Hadronic Tau Reconstruction

For the hadronic τ reconstruction and identification two algorithms have been developed; a calorimeter-based and a track-based one. The calorimeter-based algorithm starts by finding clusters in both the hadronic and electromagnetic calorimeter and then, creates the identification variables based on tracker and calorimeter information. The track-based algorithm starts from low multiplicity, high quality, collimated tracks around a leading track. Then, the energy is calculated with an energy-flow algorithm using the tracks and the deposit in the EM calorimeter. Again, all identification variables are built with information from both the tracker and the calorimeter. During the 2010-2011 data taking only the calorimeter-based algorithm was used.

The τ reconstruction starts with calorimetric jets reconstructed with the anti-k_t algorithm, [68]. As input for the jet algorithm a distance parameter of $R = 0.4$ is used from topoclusters of calorimetric cells [74, 75]. The clusters are calibrated using a local hadron calibration (LC) [76]. The algorithm is run over all jets, referred to as *seeds*, with $p_T > 10$ GeV and $|\eta| < 2.5$.

Afterwards, the kinematic variables of all τ candidates are calculated. The momentum four-vector is built as a massless vector of the p_T, η and ϕ coordinates. Then, the tracks are associated to a τ candidate if they lie within a cone of $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.2$ around the leading track. The tracks are considered as tracks coming from τ decays when they fulfil the following quality criteria:

- $p_T > 1$ GeV,
- number of pixel hits ≥ 2,
- number of pixel hits + number of SCT hits ≥ 7,
- $|d_0| < 1\,\text{mm}$,
- $|z_0 \sin\theta| < 1.5\,\text{mm}$,

where d_0 is the distance of closest approach of the track to the reconstructed primary vertex in the transverse plane, while z_0 is the longitudinal distance of closest approach. In the end, the τ

candidate is classified as one- or three-prong depending on the number of the tracks inside the cone core.

5.4.2 Hadronic Tau Identification Variables

For the hadronic τ candidates several track- and calorimeter-based variables are used. In the following a short description of those is given, while more details can be found in [59, 77].

Track Radius (R_{track}) is the track width weighted by p_T

$$R_{\text{track}} = \frac{\sum_i^{\Delta R_i<0.4} p_{T,i}\,\Delta\, R_i}{\sum_i^{\Delta R_i<0.4} p_{T,i}}, \tag{5.16}$$

where i runs over all core and isolation tracks of the τ candidate and $p_{T,i}$ is the track transverse momentum.

Leading Track Momentum Fraction (f_{track}) is the fraction of the transverse momentum of the leading track, $p_{T,1}^{\text{track}}$, over the total transverse momentum, p_T^{τ},

$$f_{\text{track}} = \frac{p_{T,1}^{\text{track}}}{p_T^{\tau}}, \tag{5.17}$$

where the transverse momenta are calibrated at the EM energy scale. In the case of a single track, f_{track} is not equal to 1, since the total transverse momentum can have contributions from calorimeter deposits from neutral particles.

Core Energy Fraction (f_{core}) is the fraction of transverse energy within a small cone ($\Delta R < 0.1$) of the τ candidate, which exploits the collimation of the energy deposition of a τ jet in contrast to a QCD multijet.

$$f_{\text{core}} = \frac{\sum_{i\epsilon\{\text{all}\}}^{\Delta R_i<0.1} E_{T,i}}{\sum_{j\epsilon\{\text{all}\}}^{\Delta R_j<0.4} E_{T,j}}, \tag{5.18}$$

where i runs over all cells associated to the τ candidate within $\Delta\, R < 0.1$ and j runs over all cells in the wider cone $\Delta\, R < 0.4$.

Number of Isolation Tracks ($N_{\text{track}}^{\text{iso}}$) is the number of tracks in the isolation annulus.

Calorimetric Radius (R_{cal}) is the weighted shower width in the calorimeters by the total transverse energy

$$R_{\text{cal}} = \frac{\sum_{i\epsilon\{\text{all}\}}^{\Delta R_i<0.4} E_{T,i}\,\Delta\, R_i}{\sum_{j\epsilon\{\text{all}\}}^{\Delta R_i<0.4} E_{T,i}}, \tag{5.19}$$

where i runs over all cells in all layers of the EM and hadronic calorimeters.

Ring Isolation (f_{iso})

$$f_{\text{iso}} = \frac{\sum_{i\epsilon\{\text{EM }0-2\}}^{0.1<\Delta R<0.2} E_{T,i}}{\sum_{j\epsilon\{\text{EM }0-2\}}^{\Delta R<0.4} E_{T,j}}, \tag{5.20}$$

where i runs over cells in the first three layers of the EM calorimeter in the annulus $0.1 < \Delta R < 0.2$ around the τ candidate axis and j runs over EM cells in the wide cone.

Cluster Mass ($m_{\text{eff.clusters}}$) is the invariant mass of the clusters that consist of the seed jet calibrated at the LC (local cluster) energy scale

$$m_{\text{eff.clusters}} = \sqrt{\left(\sum_{\text{clus}} E\right)^2 - \left(\sum_{\text{clus}} \mathbf{p}\right)^2}, \tag{5.21}$$

where eff. clusters stands for effective clusters, which are the first N leading transverse energy clusters and N is

$$N = \frac{(\sum_i E_{T,i})^2}{\sum_i E_{T,i}^2},$$

where i runs over all clusters associated to the τ candidate and N is rounded up to the nearest integer.

Track Mass (m_{tracks}) is the invariant mass of both core and isolation tracks

$$m_{\text{tracks}} = \sqrt{\left(\sum_{\text{tracks}} E\right)^2 - \left(\sum_{\text{tracks}} \mathbf{p}\right)^2}. \tag{5.22}$$

Transverse Flight Path Significance (S_T^{flight}) is the significance of the decay length of the secondary vertex for multi-prong τ candidates in the transverse plane

$$S_T^{\text{flight}} = \frac{L_T^{\text{flight}}}{\delta L_T^{\text{flight}}}, \tag{5.23}$$

where L_T^{flight} is the reconstructed signed decay length and $\delta L_T^{\text{flight}}$ is its estimated uncertainty. For the secondary vertex fit, core tracks are used.

Leading Track IP Significance ($S_{\text{leadtrack}}$) is the impact parameter (IP) significance of the leading track of the τ candidate

$$S_{\text{leadtrack}} = \frac{d_0}{\delta d_0}, \tag{5.24}$$

where d_0 is the distance of the closest approach of the track to the reconstructed primary vertex in the transverse plane and the δd_0 is its estimated uncertainty.

First 2(3) Leading Clusters Energy Ratio ($f_{2\,\text{lead clus}}, f_{3\,\text{lead clus}}$) is the ratio of the energy of the first two (three) clusters over the total energy of all clusters associated to the τ candidate. The energy of the clusters is sorted from higher to lower energies.

Maximum $\Delta R\,(\Delta\, R_{\max})$ is the maximal $\Delta\, R$ distance between the core track and the axis of the τ candidate.

Electromagnetic Fraction (f_{EM}) is the fraction of the E_T of the τ candidate deposited in the EM calorimeter

$$f_{\mathrm{EM}} = \frac{\sum_{i\epsilon\{\mathrm{EM}\ 0-2\}}^{\Delta R_i<0.4} E_{T,i}}{\sum_{j\epsilon\{\mathrm{all}\}}^{\Delta R_j<0.4} E_{T,j}}, \tag{5.25}$$

where i runs over all cells of the first three layers of the EM calorimeter and j all layers of the calorimeter.

TRT HT fraction (f_{HT}) is the fraction of the high-threshold hits in the TRT for the leading core track

$$f_{\mathrm{HT}} = \frac{\text{High-Threshold TRT hits}}{\text{Low-Threshold TRT hits}}. \tag{5.26}$$

The purpose of this variable is to distinguish between electrons and one-prong hadronic τs. Electrons being lighter than pions are more probable to produce transition radiation that, in turn, gives rise to high-threshold hits in the TRT.

Hadronic Track fraction $(f_{\mathrm{had}}^{\mathrm{track}})$ is the fraction of the hadronic transverse energy of the leading track

$$f_{\mathrm{had}}^{\mathrm{track}} = \frac{\sum_{i\epsilon\{had\}}^{\Delta R_i<0.4} E_{T,i}}{p_{T,1}^{\mathrm{track}}}, \tag{5.27}$$

where i runs over all cells in the hadronic calorimeter within the wide cone.

Maximum Strip $E_T\,(E_{T,\max}^{\mathrm{strip}})$ is the maximum transverse energy of a cell in the presampler of the EM calorimeter, that is not be associated to the leading track.

Electromagnetic Track Fraction $(f_{\mathrm{EM}}^{\mathrm{track}})$ is the ratio of the transverse energy deposited in the EM calorimeter over the transverse momentum of the leading track

$$f_{\mathrm{EM}}^{\mathrm{track}} = \frac{\sum_{i\epsilon\{EM\}}^{\Delta R_i<0.4} E_{T,i}}{p_{T,1}^{\mathrm{track}}}, \tag{5.28}$$

where i runs over all cells in the electromagnetic calorimeter within the wide cone.

Hadronic Radius (R_{had}) is the weighted shower width in the Hadronic calorimeter

$$R_{\mathrm{had}} = \frac{\sum_{i\epsilon\{had,EM3\}}^{\Delta R_i<0.4} E_{T,i}\,\Delta\, R_i}{\sum_{i\epsilon\{had,EM3\}}^{\Delta R_i<0.4} E_{T,i}}, \tag{5.29}$$

where i runs over all wide-cone, τ-candidate-associated cells in the hadronic calorimeter and the third layer of the EM calorimeter.

Electromagnetic Radius (R_{EM}) is the weighted shower width in the EM calorimeter

$$R_{\mathrm{EM}} = \frac{\sum_{i\epsilon\{\mathrm{EM\ 0-2}\}}^{\Delta R_i<0.4} E_{T,i}\,\Delta\,R_i}{\sum_{i\epsilon\{\mathrm{EM\ 0-2}\}}^{\Delta R_i<0.4} E_{T,i}}, \tag{5.30}$$

where i runs over cells in the first three layers of the EM calorimeter.

Corrected Cluster Isolation Energy ($E^{\mathrm{iso}}_{T,\mathrm{corr}}$) is the transverse energy of the isolated clusters, corrected for pile-up

$$E^{\mathrm{iso}}_{T,\mathrm{corr}} = E^{\mathrm{iso}}_{T} - \delta\,E^{\mathrm{iso}}_{T} = \sum_{i}^{0.2<\Delta R_i<0.4} E_{T,i} - \delta\,E^{\mathrm{iso}}_{T}, \tag{5.31}$$

where i runs over all clusters associated to the τ candidate and ΔR_i is defined between the cluster and the τ candidate axis. $\delta E^{\mathrm{iso}}_{T}$ is the pile-up correction term and is calculated as $(1-\mathrm{JVF})\times\sum p_{T,\mathrm{track}}$. JVF stands for Jet Vertex Fraction and is calculated for each τ candidate as the fraction of the sum of the tracks p_{T} associated to the seed jet that are consistent with the primary vertex. The $p_{T,\mathrm{track}}$ is the sum of the transverse momentum of the tracks associated to this jet.

Combinations of these variables are also used for discriminating hadronic τ candidates from QCD jets and electrons. For the τ-jet separation, there are three algorithms, cut-based, likelihood-based (LLH) and boosted decision trees (BDT), while for the τ-e there are two, cut-based and BDT. Each method uses different combination of the above mentioned variables. That is shown in detail in Table 5.8. Distributions of selective variables for one- and three-prong τ candidates are shown in Fig. 5.15 (2010) and Fig. 5.16 (2011) (for jet background) and Fig. 5.17 (for electron background) [77].

5.4.3 Discriminating from Jets

All three identification algorithms have been designed so that most of the QCD multijets are rejected while most of the τ signal is kept. The discriminants, cuts, likelihood and BDT, are optimised for three working points: *loose, medium* and *tight*, corresponding to 60%, 45% and 30% signal efficiency [77].

In the first year of data taking, the detector is not well understood neither is the behaviour of the algorithms, that is why cut-based techniques are preferred. That is not the case during the 2011 data taking, when the BDT algorithms are used. The parameters used for the cut-based identification in the data collected in 2010 and the parameters used to train the BDTs for the 2011 data are explained in the following section. For completeness, a short description of the likelihood function is, also, given.

The Cut-based algorithm for the 2010 analyses is using three variables: R_{EM}, f_{track} and R_{track}. In 2011, due to higher pile-up events more robust variables are chosen, like R_{track}, $N^{\mathrm{iso}}_{\mathrm{track}}$ and $E^{\mathrm{iso}}_{\mathrm{T,corr}}$. In either case, an optimisation is done on cut values to ensure the expected signal efficiencies.

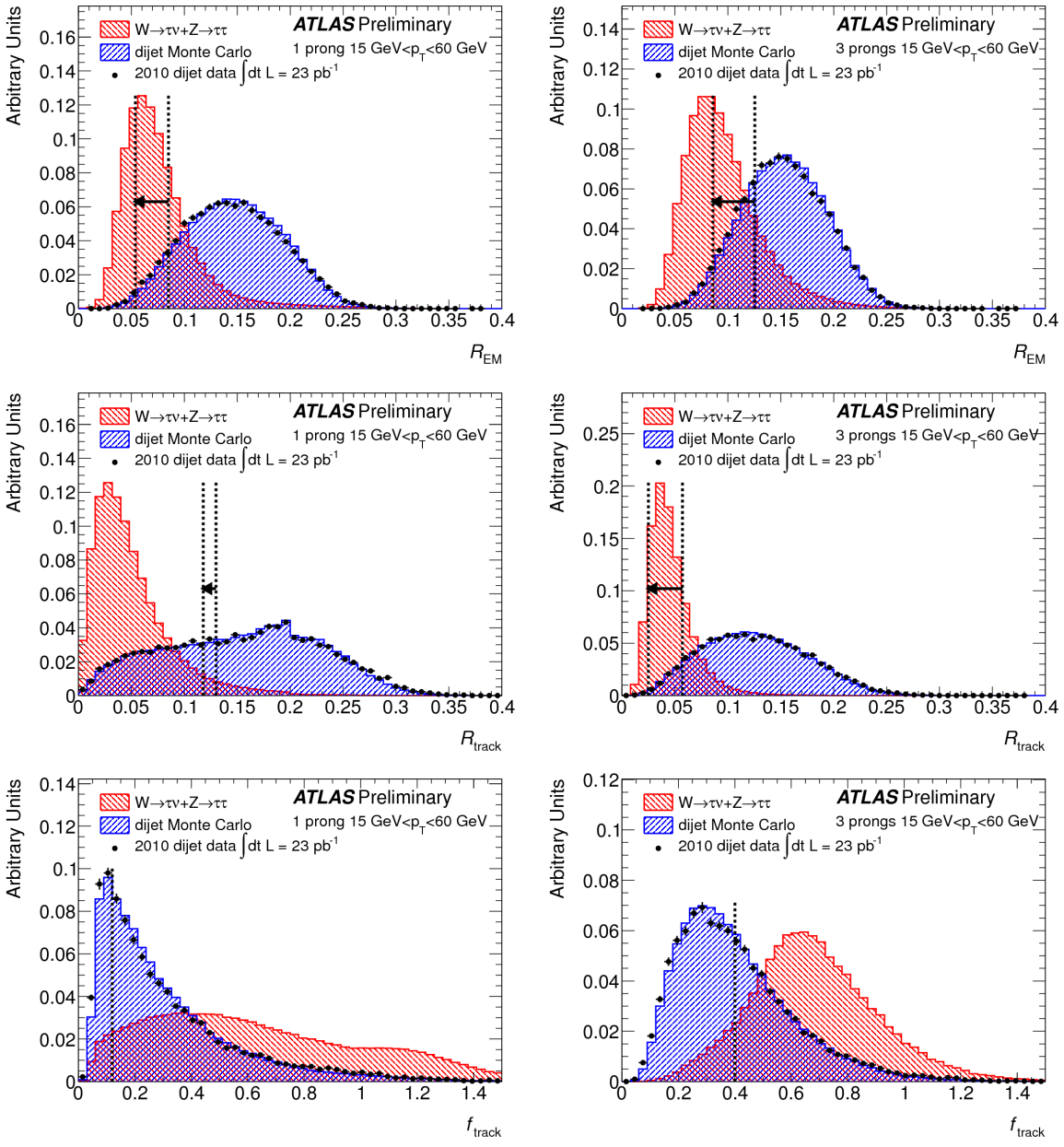


Figure 5.15: Distributions of a selection of jet discriminating variables for MC simulated $Z \to \tau\tau$ and $W \to \tau\nu$ signal samples and a di-jet background sample selected from 2010 data ($\mathcal{L} = 23\,\mathrm{pb}^{-1}$). The dashed lines indicate the cut boundaries for the cut-based identification. The characteristic range of the cut values is demonstrated by showing lines for the cuts for candidates with $p_{\mathrm{T}} = 20$ GeV and then an arrow pointing to the cut for candidates with $p_{\mathrm{T}} = 60$ GeV. Taken from [78].

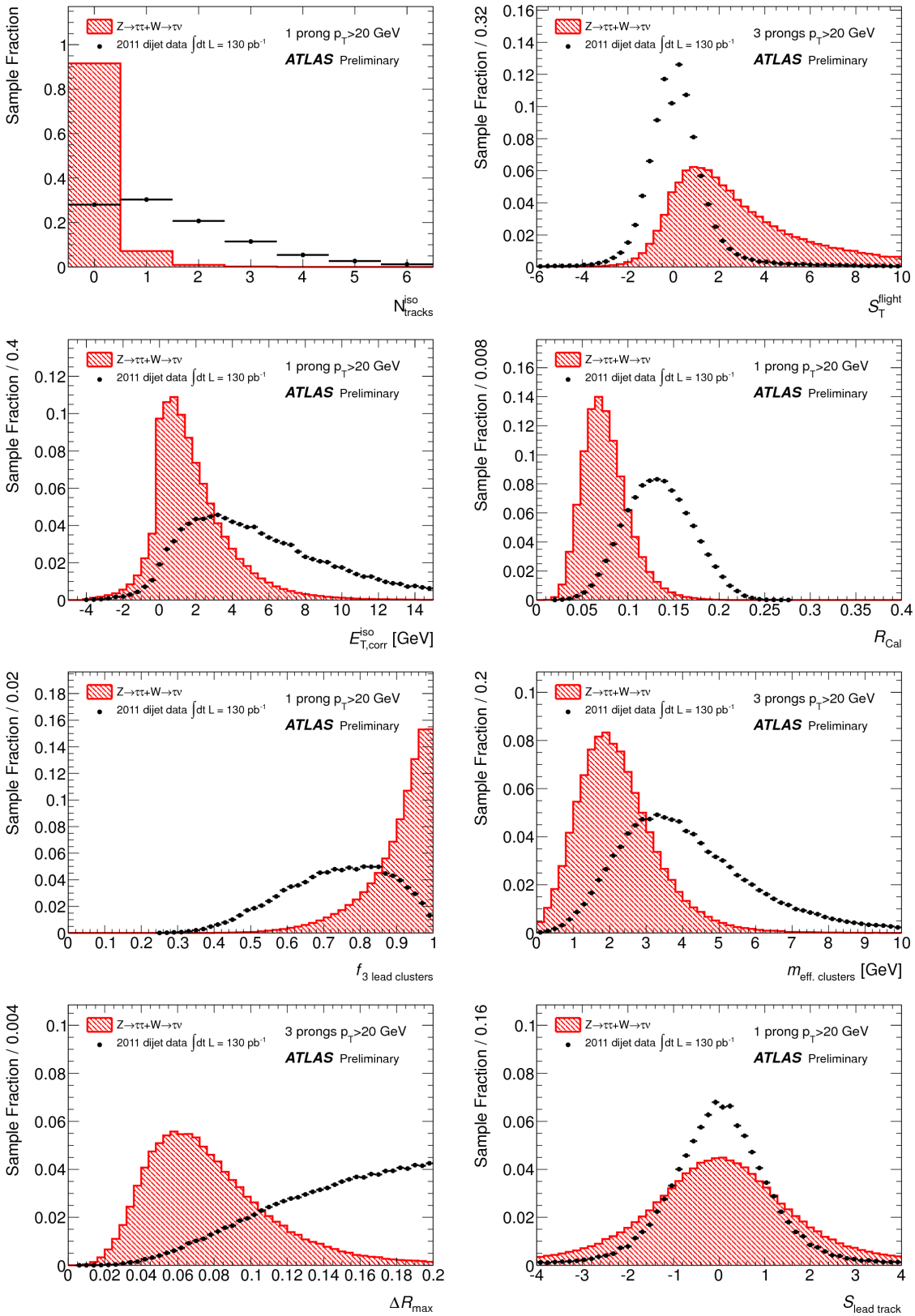


Figure 5.16: Distributions of a selection of jet discriminating variables for MC simulated $Z \to \tau\tau$ and $W \to \tau\nu$ signal samples and a di-jet background sample selected from 2011 data ($\mathcal{L} = 130\,\mathrm{pb}^{-1}$). Taken from [77].

Table 5.8: Complete list of variables used by the cut-based (Cut), likelihood-based (LLH) and boosted decision tree (BDT) jet discriminants, and the cut-based and BDT electron discriminants. The equation numbers refer to the variable definitions in Section 5.4.2. The use of variables for the jet discriminants is defined separately for 1-prong (1) and three-prong (m) candidates. Taken from [77].

Variable	eq.	Jet discriminants						Electron discriminants	
		Cut		LLH		BDT		Cut	BDT
		1	m	1	m	1	m	1	1
R_{track}	5.16	•	•	•	•	•	•		•
f_{track}	5.17	•	•			•	•		•
f_{core}	5.18			•	•	•	•		•
$N_{\text{track}}^{\text{iso}}$		•	•	•		•	•		
R_{Cal}	5.19			•		•	•		
f_{iso}	5.20								•
$m_{\text{eff. clusters}}$	5.21					•	•		
m_{tracks}	5.22				•		•		
$S_{\text{T}}^{\text{flight}}$	5.23		•		•		•		
$S_{\text{lead track}}$	5.24					•	•		
$f_{\text{2 lead clusters}}$				•					
$f_{\text{3 lead clusters}}$						•	•		
ΔR_{max}					•		•		
f_{EM}	5.25								•
f_{HT}	5.26							•	•
$f_{\text{Had}}^{\text{track}}$	5.27							•	•
$E_{\text{T,max}}^{\text{strip}}$								•	•
$f_{\text{EM}}^{\text{track}}$	5.28							•	
R_{Had}	5.29								•
$E_{\text{T,corr}}^{\text{iso}}$	5.31	•	•						

Projective Likelihood Function is the second method used to discriminate τ candidates from jets. The discriminant is the logarithm of the ratio of the likelihoods of signal and background

$$d = ln\left(\frac{L_s}{L_B}\right) = \sum_{i=1}^{N} ln\left(\frac{p_i^s\left(x_i\right)}{p_i^B\left(x_i\right)}\right).$$

where $p_i^{s(B)}\left(x_i\right)$ are the probability density functions (pdfs) of each identification variable x_i for signal (s) and background (B). The pdfs are the fraction of events per bin in a histogram of the x_i distribution. The discriminant is split in the following bins for optimal performance: 1-prong, 3-prong, three p_T bins and three bins depending on the number of reconstructed vertices.

Boosted Decision Trees are trained for one- and three-prong τ candidates. Separate training is, also, performed depending on the average number of primary vertices per collision. Although the BDT score is a continuous line ranging from 0 for background-like to 1 for signal-like distributions, loose, medium and tight selections are defined. The jet BDT score is p_T-dependant. The BDT score distributions for one- and three-prong events is shown in Fig. 5.18.

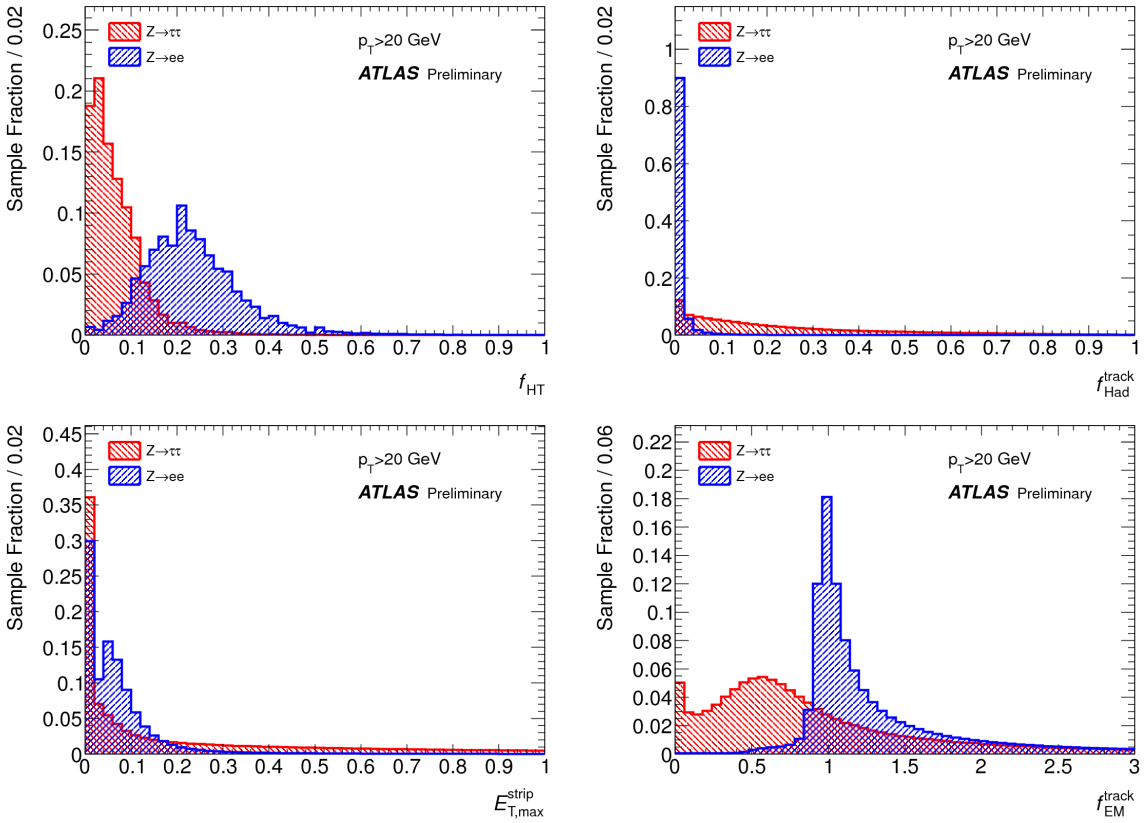


Figure 5.17: Distributions of a selection of identification variables for MC simulated $Z \to \tau\tau$ signal and $Z \to ee$ background events. Taken from [77].

Signal and Background Efficiencies are defined as follows for all of the above identification methods

Signal efficiency:

$$\varepsilon_{\text{sig}}^{n\text{-prong}} = \frac{(\#\text{ of } \tau \text{ candidates with } n \text{ reconstructed tracks, passing ID})}{(\#\text{ of true visible hadronic } \tau \text{ decays with } n \text{ prongs})}. \tag{5.32}$$

Background efficiency:

$$\varepsilon_{\text{bkg}}^{n\text{-prong}} = \frac{(\#\text{ of } \tau \text{ candidates with } n \text{ reconstructed tracks, passing ID})}{(\#\text{ of } \tau \text{ candidates with } n \text{ reconstructed tracks})}. \tag{5.33}$$

All "true" information is taken from the truth-level Monte Carlo generator. The truth-matched τ candidates are those reconstructed candidates within $\Delta R < 0.2$ of the true visible (vis) τ decay and the fiducial region for the true hadronic τs is $|\eta| < 2.5$ and $E_T^{\text{vis}} > 10$ GeV. One- and three-prong candidates are considered for the signal efficiency, while any number of prongs is allowed for the background. For more accurate results the efficiencies are binned in p_T and η. Fig. 5.19 and Fig. 5.20 show the signal efficiency with respect to the background rejection for the three methods for different p_T bins and number of prongs for 2010 and 2011 data. Although in general the BDT performs better than the other two methods, in practice

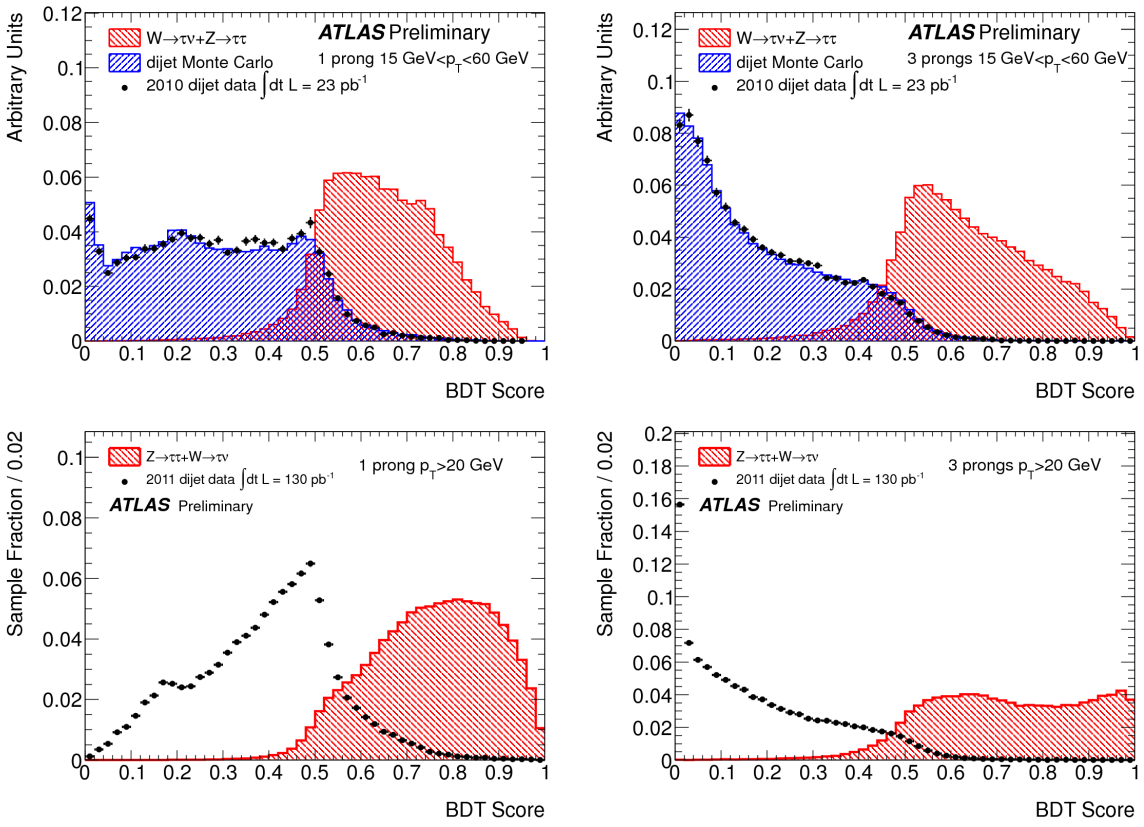


Figure 5.18: The jet BDT score for 1-prong (left) and 3-prong (right) τ candidates. For the 2010 analysis in the top (plots taken from [78]) and for the 2011 in the bottom (plots taken from [77]).

the background rejection is strongly correlated with the specific analysis cuts. After signal to background efficiency studies, the cut-based algorithm is chosen for 2010 analysis and the BDT for the 2011.

5.4.4 Discriminating from Electrons

A significant background for one-prong τ candidates are the electrons. Although the signatures are quite similar, techniques have been developed to distinguish between the two objects. These techniques are taking advantage of several differences such as transition radiation emitted by the electrons or the longer and wider shower of the τ candidates in the calorimeter. In ATLAS two methods exist, one cut-based and one using BDTs. Similarly to the identification methods, three working points with corresponding signal efficiencies are defined; *loose* (95%), *medium* (85%) and *tight* (75%).

Cut-based electron veto uses two shower shape variables, $f_{\mathrm{Had}}^{\mathrm{track}}$ and $E_{\mathrm{T,max}}^{\mathrm{strip}}$, and the pseudorapidity range to categorise the candidates. The electron veto is optimised for one-prong candidates and two pseudorapidity ranges: barrel ($|\eta_1^{\mathrm{track}}| < 1.7$) and endcap ($|\eta_1^{\mathrm{track}}| > 1.7$). The cut selection is applied on two variables for the barrel region, f_{HT} and $f_{\mathrm{EM}}^{\mathrm{track}}$, and only on $f_{\mathrm{EM}}^{\mathrm{track}}$ for the endcaps. The precise categorisation is shown in Table 5.9.

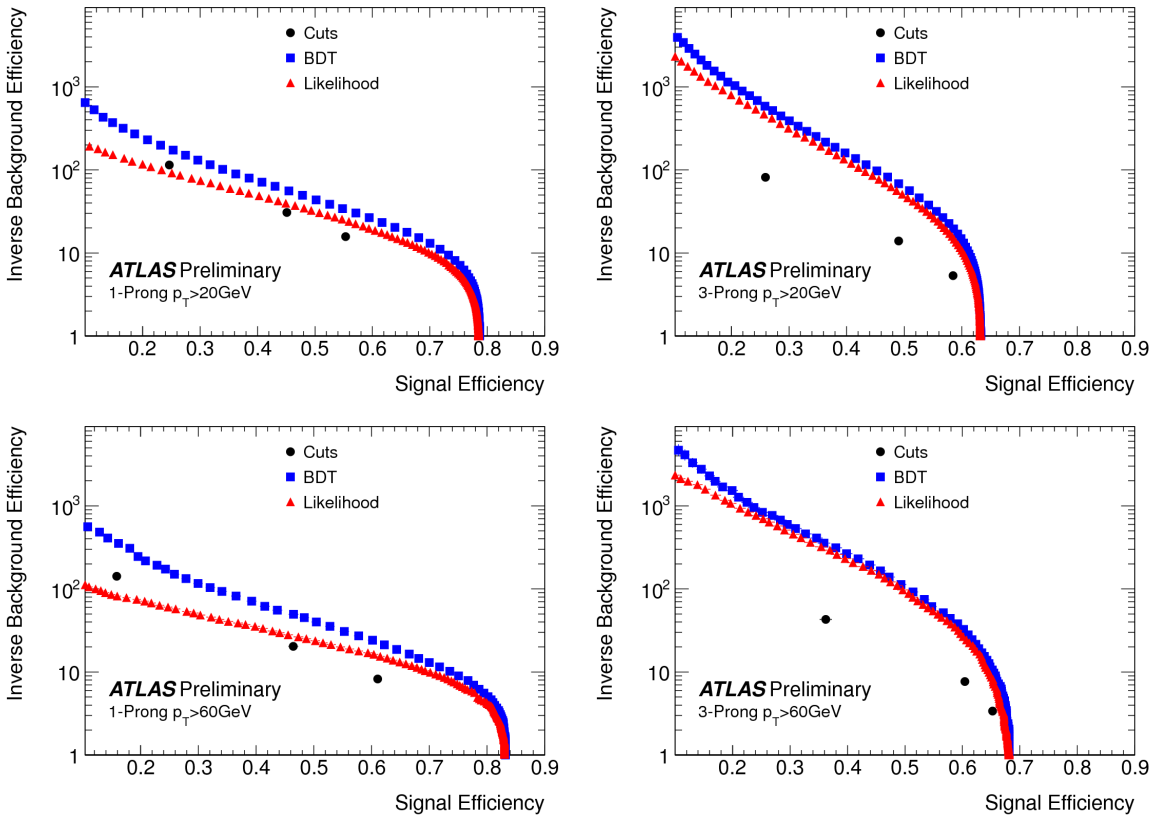


Figure 5.19: Inverse background efficiency as a function of signal efficiency for 1-prong (left) and 3-prong (right) candidates, in low (top) and high (bottom) p_{T} ranges, for all jet discriminants with the 2010 data. Taken from [78].

Table 5.9: Category structure used in the cut-based electron veto. Boundary and cut values depend on the particular selection. Candidates selected by these cuts are considered to be τ leptons. The last two columns show the fraction of τ leptons and electrons in each category before any cut is applied. The boundary values and cuts shown here correspond to the loose electron veto configuration. Taken from [77].

Category definition		Cut	τ fraction	electron fraction
Barrel Candidates ($\lvert\eta\rvert < 1.7$)				
$f_{\mathrm{Had}}^{\mathrm{track}} > 0.04$	-	$f_{\mathrm{HT}} < 0.24$	38%	2.5%
$f_{\mathrm{Had}}^{\mathrm{track}} \leq 0.04$	$E_{\mathrm{T,max}}^{\mathrm{strip}} > 0.25$	$f_{\mathrm{HT}} < 0.18$	20%	6.6%
	$E_{\mathrm{T,max}}^{\mathrm{strip}} \leq 0.25$	$f_{\mathrm{HT}} < 0.10$	15%	64%
		$f_{\mathrm{EM}}^{\mathrm{track}} < 9.8$		
endcap Candidates ($\lvert\eta\rvert > 1.7$)				
$f_{\mathrm{Had}}^{\mathrm{track}} > 0.02$	-	-	10%	0.5%
$f_{\mathrm{Had}}^{\mathrm{track}} \leq 0.02$	-	$f_{\mathrm{EM}}^{\mathrm{track}} \leq 2.7$	17%	26%

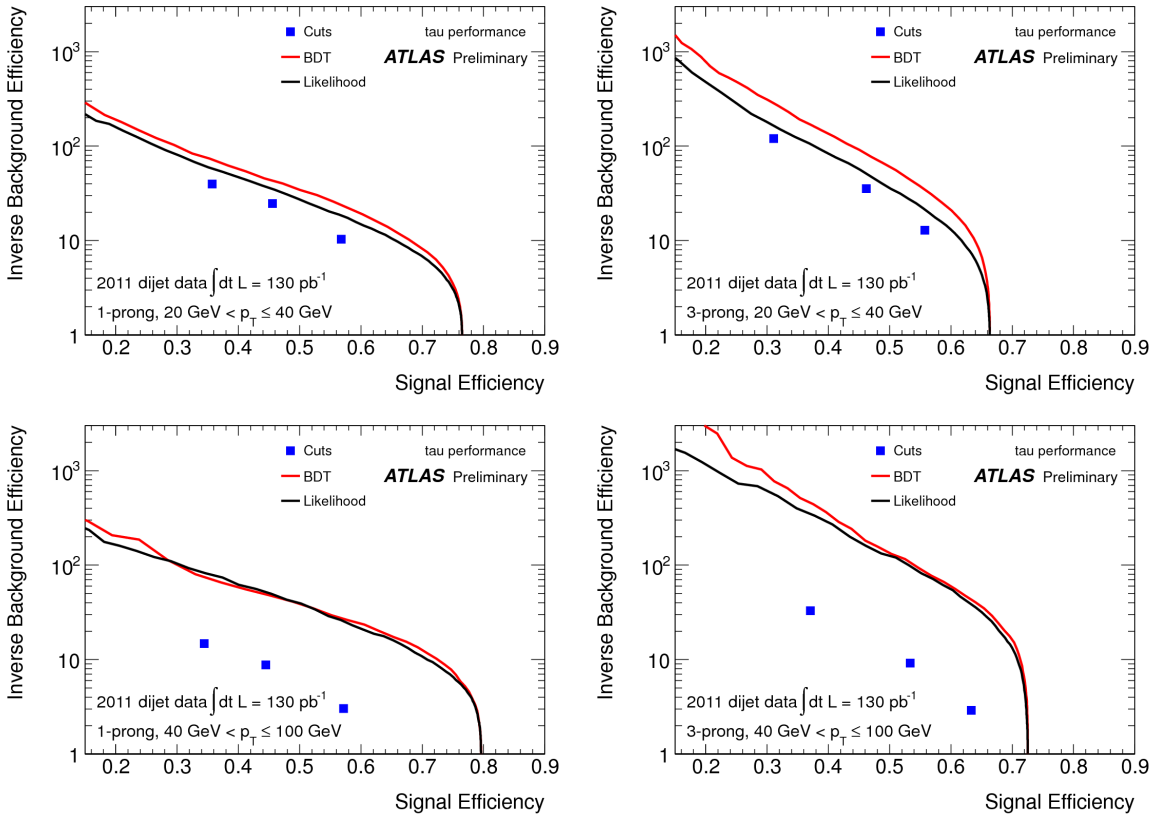


Figure 5.20: Inverse background efficiency as a function of signal efficiency for 1-prong (left) and 3-prong (right) candidates, in low (top) and high (bottom) p_{T} ranges, for all jet discriminants. Taken from [77].

Boosted Decision Tree discriminants are built in four pseudorapidity ranges: ($|\eta| < 1.37$), crack ($1.37 < |\eta| < 1.52$), endcap ($1.52 < |\eta| < 2.0$) and forward endcap ($2.0 < |\eta| < 2.3$). For the training the variables listed in Table 5.8 are used for each η region. A good separation between electrons and τ candidates is shown in Fig. 5.21.

The mis-identification probabiliy of the electron veto methods is shown in Fig. 5.22. The signal and background efficiencies are estimated using Eq. (5.32) and (5.33). Similarly to the jets discrimination, the actual background rejection depends on the particular analysis cuts. For the $Z \to \tau\tau$ analysis a cut-based veto is found to provide better background rejection while keeping most of the signal for both 2010 and 2011 data analysis as seen in Fig. 5.23.

5.4.5 Hadronic Tau Energy Scale

Due to the limited amount of data, the energy scale of the τ candidate is evaluated from Monte Carlo simulations of $Z \to \tau\tau$, $W \to \tau\nu$ and $Z' \to \tau\tau$ processes. The energy of the τ candidate is given by the calorimeter response and is calibrated by applying a correction factor at the EM energy scale. The sum over the energies of the cells within a cone of $\Delta R < 0.4$ of the seed jet axis gives the electromagnetic energy and the correction factor is

$$R(p_T^{EM}) = \frac{p_T^{EM}}{p_T^{gen}}. \tag{5.34}$$

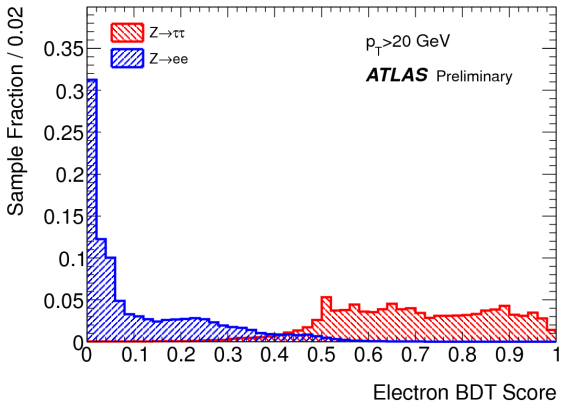


Figure 5.21: Score of the BDT-based electron veto for MC simulated hadronic τ decays and electrons. Taken from [77].

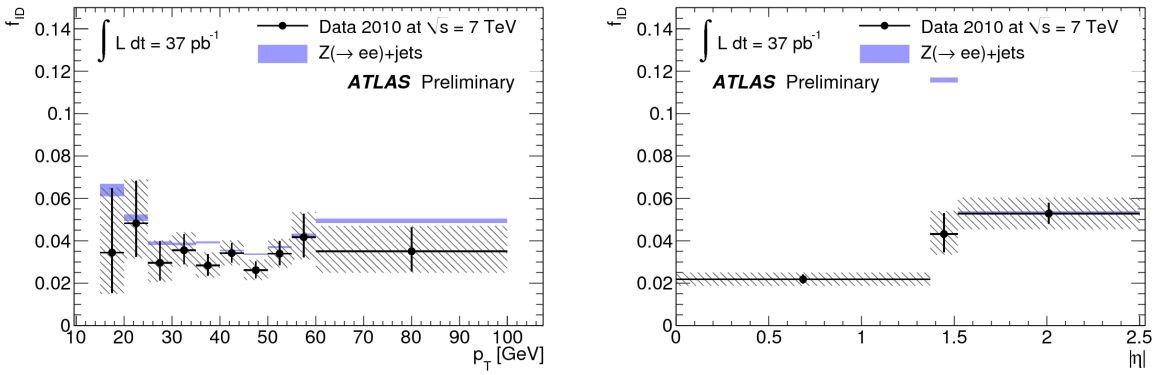


Figure 5.22: The mis-identification probabilities as a function of the tau lepton transverse momentum for the looser working point in bins of p_{T}(left) and η (right). Taken from [79].

p_T^{EM} is the transverse momentum of the τ candidate at EM scale and p_T^{gen} is the p_{T} at truth level. Having estimated the response function the energy is calibrated to the τ energy scale (TES) by scaling the EM scale energy by

$$p_T^{TES} = \frac{1}{R(p_T^{EM})} p_T^{EM}. \tag{5.35}$$

Studies of the reliability of the use of Monte Carlo have been conducted. Agreement for pions and protons in the range of a few hundred MeV to 6 GeV is excellent between data and Monte Carlo, details of the studies are given in [78].

Uncertainties on the TES measurement are estimated by comparing the reconstructed p_{T} of the τ candidate passing the loose identification criteria in nominal Monte Carlo and alternative configurations. In each case the difference is calculated as

$$f_s = \frac{p_T^{reco} - p_T^{truth}}{p_T^{truth}}, \tag{5.36}$$

where p_T^{reco} is the reconstructed transverse momentum after the TES correction is applied and p_T^{truth} is the true visible p_{T}. The f_s is calculated separately for 1- and 3-prong τ candidates,

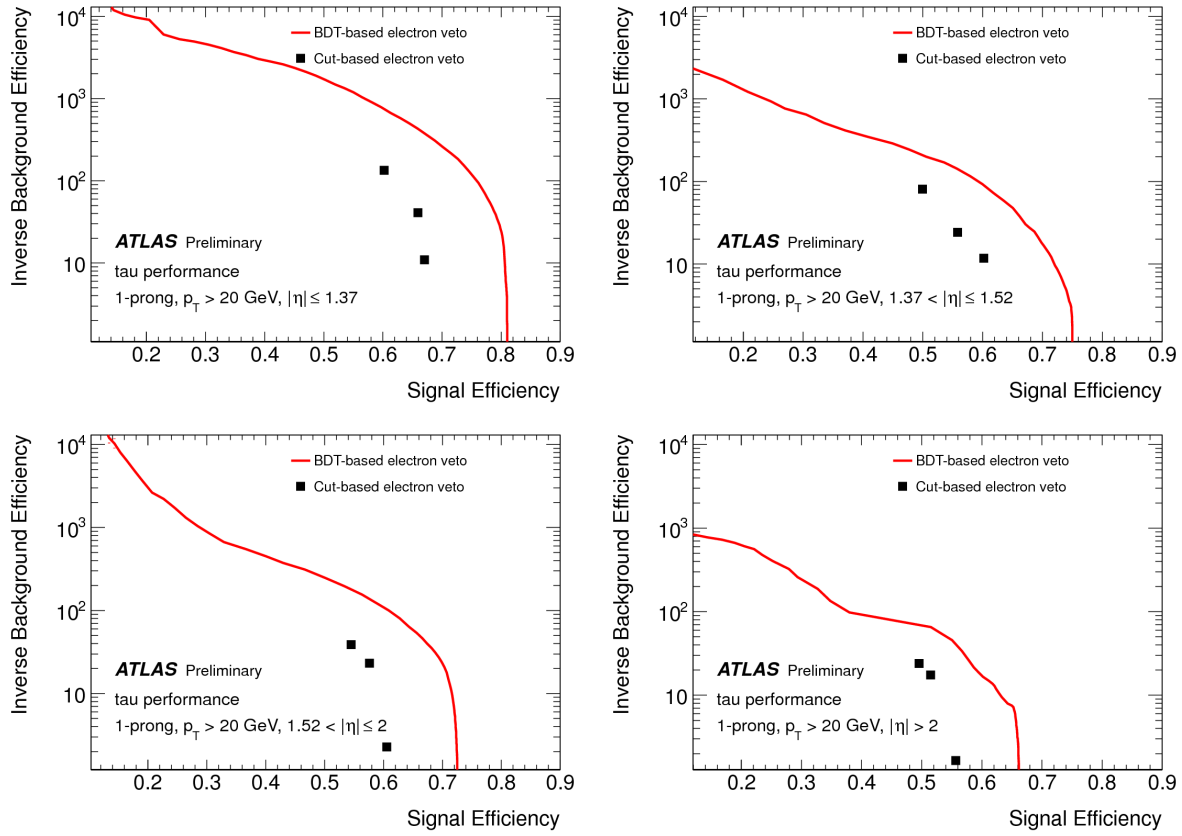


Figure 5.23: Inverse background efficiency as a function of signal efficiency for 1-prong reconstructed τ candidates with $p_{\mathrm{T}} > 20$ GeV, in four regions of $|\eta|$, for both electron discriminants. Taken from [77].

in the barrel, endcap and transition regions and in six p_{T} bins ranging from 15 to 110 GeV. The alternative scenarios evaluated as systematic uncertainties of the TES are: Monte Carlo generator and underlying event model; hadronic shower model; amount of detector material; electromagnetic energy scale; topological clustering noise thresholds; pile-up and non-closure [77, 79]. The effect of all these uncertainties on the TES are shown in Fig. 5.24 for the 2010 data run and in Fig. 5.24 for the 2011 data run. The percentage effect is summarised in Table 5.10. The dominant uncertainty is coming from the hadronic shower modelling and the non-closure.

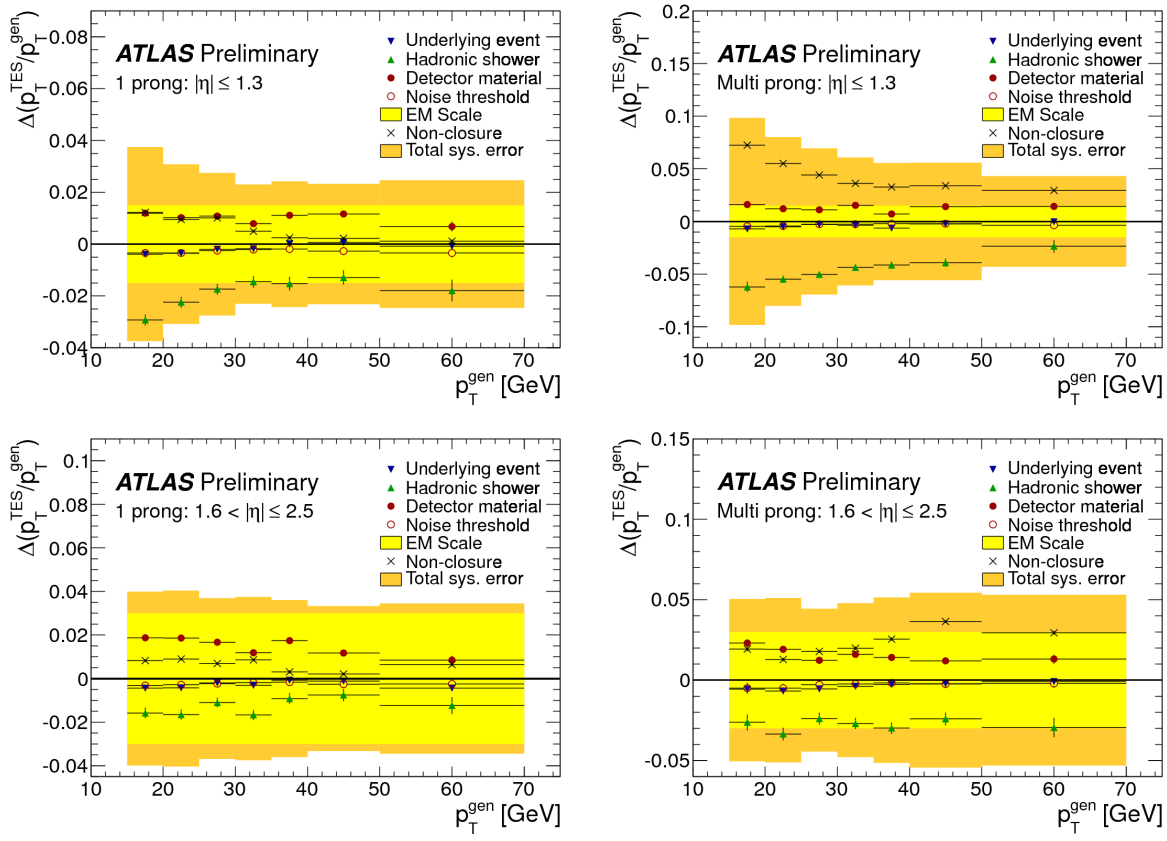


Figure 5.24: Systematic uncertainties on the τ energy scale for 1-prong (left) and three-prong (right) candidates, in the barrel (top), crack (middle) and endcap (bottom) regions for the 2010 data run. Each different marker represents a separate source of uncertainty as indicated in the legend. The yellow band shows the combined uncertainty from all sources. Taken from [78].

Table 5.10: Total systematic uncertainty on the τ energy scale. Taken from [77].

p_{T} range [GeV]	$\lvert\eta\rvert < 1.3$	$1.3 < \lvert\eta\rvert < 1.6$	$\lvert\eta\rvert > 1.6$
	1-prong candidates		
15-20	5.5%	5%	4.5%
20-30	4.5%	5%	4.5%
>30	3.5%	5%	4.5%
	three-prong candidates		
15-20	6.5%	9.5%	6.5%
20-30	6.5%	5.5%	5.5%
30-40	5.5%	5.5%	5.5%
>40	4.5%	5%	5%

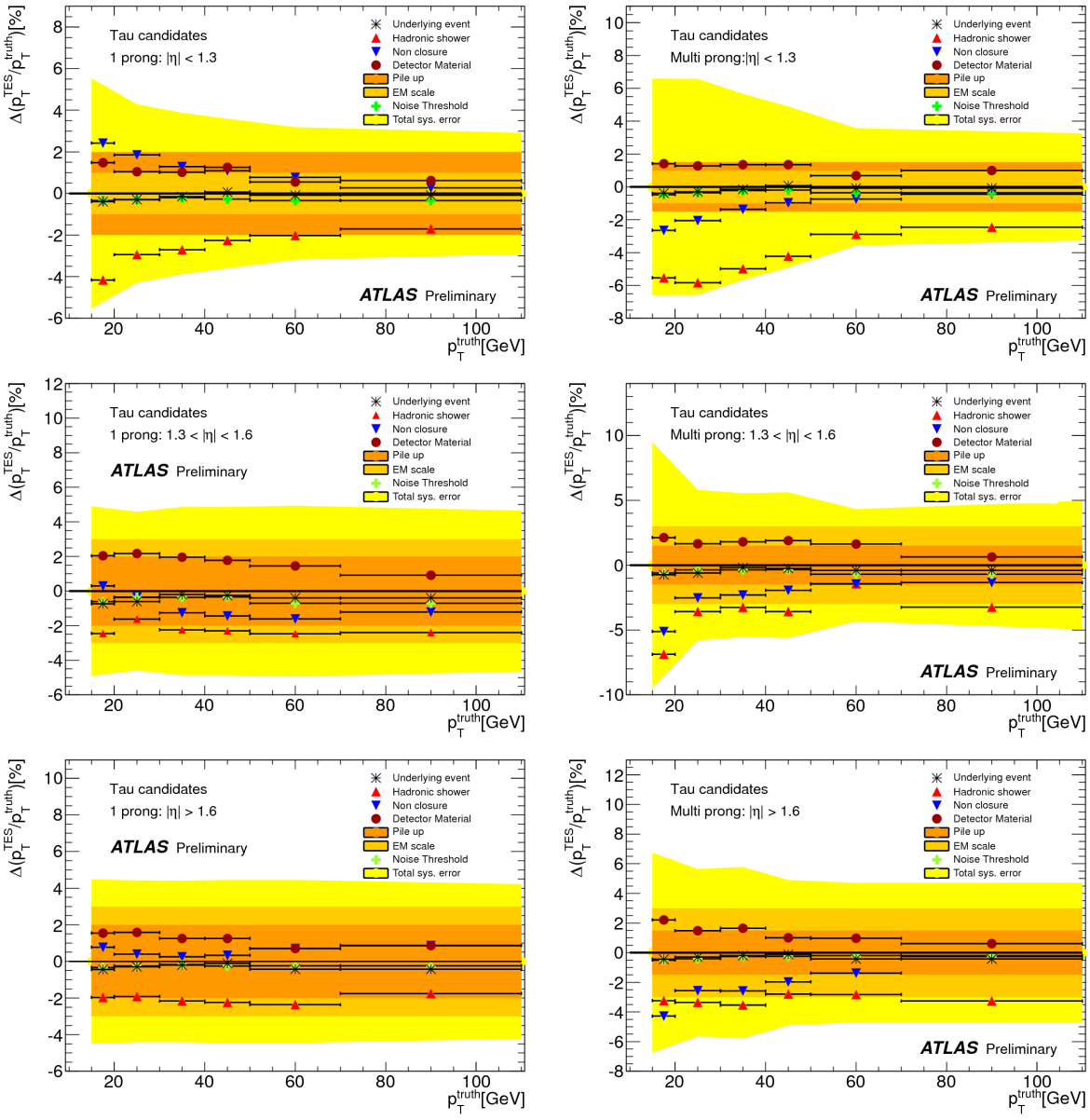


Figure 5.25: Final systematic uncertainty on the τ energy scale for 1-prong (left) and three-prong (right) candidates, in the barrel (top), crack (middle) and endcap (bottom) regions for the 2011 data run. Each different marker represents a separate source of uncertainty as indicated in the legend. The yellow band shows the combined uncertainty from all sources. Taken from [77].

Chapter 6

$\gamma^*/Z \to \tau\tau$ Selection

The production of a Z boson followed by its decay into τ leptons and neutrinos is an irreducible background for Higgs boson searches in the τ lepton channels. A series of requirements has been developed for selecting $Z \to \tau\tau$ events in the semi-leptonic channel, in which one τ decays hadronically and the other to an electron or muon. The measurement is first done with the 2010 data and then repeated with the 2011 data. In the first measurement, the amount of data collected (36 pb^{-1}) resulted in a cross section dominated by statistical uncertainties, that is why the study is repeated with the much larger sample of 2011 run (1.3-1.5 fb^{-1}). In this chapter, the analysis strategy is explained in detail. In most cases the cuts are the same for both datasets, whenever that is not the case it is explicitly stated.

6.1 Monte Carlo samples

For the production of the Monte Carlo samples the full ATLAS simulation [34] is used. The pile-up conditions are modelled by overlaying simulated minimum bias events over the original hard-scattering event. The AUET1 ATLAS specific tune [48] is used. The Monte Carlo samples are from the MC10 production. The W+jets and γ^*/Z background and signal samples are generated with PYTHIA6 [80] in 2010 and AlpGEN [42] in 2011. The cross sections and K-factors[1] for the NNLO calculations are taken from [82]. For the $t\bar{t}$ background the MC@NLO generator is used [83] with CTEQ66 PDFs. The diboson samples are generated at leading order with HERWIG++ [44]. In all samples the τ decays are modelled with the TAUOLA package [50] and all generators are interfaced to PHOTOS [52] to simulate the effect of final state QED radiation. All simulated samples used in this study, the cross sections and the K-factors are listed in the tables in the Appendix A.

6.2 Data Samples

The measurement has been performed with collision data at $\sqrt{s} = 7$ TeV. Good quality of data is ensured via analysis-specific Good Runs List (GRL).

[1] As explained in Section 4.1 the leading order calculation of the cross section of a process is not sufficient to match the experimental accuracy, therefore the parton shower is used. In Monte Carlo generators this technique depicts the shapes of the distributions accurately but not the normalisation. A solution to this problem is given by normalising the total cross section to a K-factor, $K = \frac{\sigma^{NNLO}}{\sigma^{LO}}$ [81].

Table 6.1: The 2010 run periods, the corresponding range of run numbers and the integrated luminosity for the electron and muon channel.

run period 2010	run number	int. luminosity (pb^{-1}) electron channel	int. luminosity (pb^{-1}) muon channel
period E	160387 – 161948	0.8	0.5
period F	162347 – 162882	1.7	1.7
period G	165591 – 166383	5.5	5.5
period H	166466 – 166964	7.0	7.0
period I	167575 – 167844	20.7	20.7
		35.7	35.5

In 2010 the GRL requires for the following flags to be set: `ATLGL` is `ready` for the ATLAS global and for data taking, `ATLSOL`, `ATLTOR` for the solenoid and toroid magnets to be on and `lumi` for reliable luminosity estimation. The following flags need to be "Green": the L1 triggers central (`L1CTP`), calorimeter (`L1CAL`) and the muon endcap (`L1MUE`) and barrel (`L1MUB`); the high level triggers for electrons (`TRELE`), muons (`TRMUO`), jets (`TRJET`) and missing transverse energy (`TRMET`); the combined data quality flags for τ candidates (`cp_tau`), electrons in the forward (`cp_eg_electron_forward`), the endcaps (`cp_eg_electron_endcap`) and the barrel regions (`cp_eg_electron_barrel`), muons triggered by either muon algorithm (`cp_mu_mmuidcb` or `cp_mu_mstaco`), jets in the barrel (`cp_jet_jetb`), in the endcaps (`cp_jet_jetea`, `cp_jet_jetec`), or the forward regions (`cp_jet_jetfa`, `cp_jet_jetfc`) and total $E_{\mathrm{T}}^{\mathrm{miss}}$ (`cp_met`), $E_{\mathrm{T}}^{\mathrm{miss}}$ in the calorimeters (`cp_met_calo`) and the muon component of the $E_{\mathrm{T}}^{\mathrm{miss}}$ (`cp_met_muon`).

In 2011 the GRL requires for the following flags to be set: the `ATLAS` partition to be on, the data-taking tag to be set to `data11_7TeV`, the data quality defects for the various subsystems to be "Green" and the detector status to be set to `DetStatus-v28-pro08-07`. The defects are: global detector status (`global_status`), solenoid magnet (`atlsor`), toroid magnet (`atltor`), muon trigger (`trig_muo`), muon combined for MUID algorithm (`cp_mu_mmuidcb`), muon combined for STACO algorithm (`cp_mu_mstaco`), electron trigger (`trig_ele`), electron combined in the barrel region (`cp_eg_electron_barrel`), electron combined at the endcaps (`cp_eg_electron_endcap`), electron combined in the forward region (`cp_eg_electron_forward`), jet trigger (`trig_jet`), jet combined in the barrel region (`cp_jet_jetb`), jet combined in the endcap A (`cp_jet_jetea`), jet combined in the endcap C (`cp_jet_jetec`), jet combined in the forward A (`dq cp_jet_jetfa`), jet combined in the forward C (`dq cp_jet_jetfc`), $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger (`trig_met_metcalo`), $E_{\mathrm{T}}^{\mathrm{miss}}$ combined (`cp_met`), tau trigger (`trig_tau`), tau combined (`cp_tau`), b-jet (`cp_btag_life`), pixel B-layer (`pix0`), tracking system (`cp_tracking`), inner detector vertex finding algorithms (`idvx`), inner detector good beam spot position (`idbs`), luminosity (`lumi`).

The luminosities for the data periods are calculated with the official ATLAS tool *lumicalc* [84]. The resulting data periods and the luminosities are given respectively in Tables 6.1 and 6.2.

Simulation of Pile-up

A bunch in LHC contains 10^{10} protons and multiple bunches are circulating the accelerator per fill. In 2010, the bunch separation for the first train was 150 ns and for the second 225 ns. During the 2011 data taking, the LHC ran with two different configurations of spacing per bunch

Table 6.2: The 2011 run periods, the corresponding range of run numbers and the integrated luminosity per trigger per channel.

run period 2011	run number	trigger	int. luminosity (pb^{-1})
electron channel			
period E-J	180614 – 186755	`EF_tau16_loose_e15_medium`	1344.0
muon channel			
period B2-I	178044 – 186493	`EF_mu15i`	1529.62
period J	186516 – 186755	`EF_mu15i_medium`	17.38
total			1547.0

Table 6.3: Number of reconstructed vertices and corresponding pileup weights.

Nvertex	Event Weight
1	1.970 ± 0.008
2	1.242 ± 0.004
3	0.853 ± 0.003
4	0.633 ± 0.002
5	0.509 ± 0.003
6	0.427 ± 0.004
7	0.392 ± 0.006
8	0.38 ± 0.01
9	0.39 ± 0.02
10	0.41 ± 0.05
≥ 11	0.89 ± 0.14

train; 75 ns (period B) and 50 ns (periods D-J). As a result, apart from the primary interaction, several additional interactions take place per event. These additional interactions are called pile-up events. Two sources of pile-up arise: **in-time**, interactions from the same bunch crossing, and **out-of-time**, interactions from the previous bunch crossing. In 2010 the pile-up is mainly in-time with the maximum number of primary vertices being thirteen and the average about 2.2. In 2011 both in-time and out-of-time pile-up are important. The maximum number of primary vertices is sixteen with an average number of interactions per bunch crossing being six. In order to account for the additional interactions, the number of reconstructed vertices is used for the 2010 data analysis and the average number of pile-up interactions, $< \mu >$, for 2011. In Monte Carlo samples the pile-up is simulated under given initial conditions and then re-weighted to correctly match the data per run. For re-weighting the Monte Carlo to the 2010 data an event weight is calculated for each number of primary vertices. The event weights are shown in Table 6.3 and the primary vertex distribution before and after the re-weighting in shown in Fig. 6.1. In 2011, a PileupReweighting tool [85] is developed centrally. The tool compares the $< \mu >$ distribution in data and Monte Carlo and calculates a scaling factor, which is then propagated to each Monte Carlo event. In Fig.6.2 the number of pile-up vertices is shown for data and Monte Carlo for period D, as an indication of the pile-up conditions and the effectiveness of the tool.

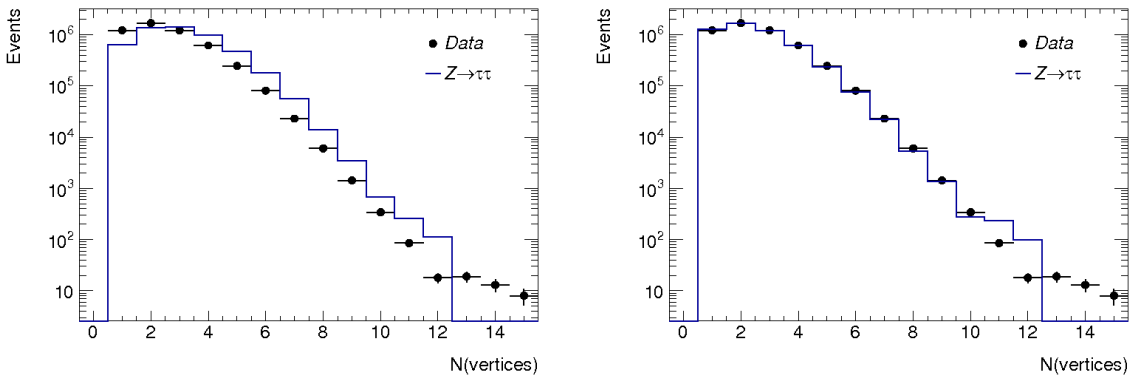


Figure 6.1: The primary vertex distribution before (left) and after (right) pile-up re-weighting for data 2010 and signal Monte Carlo.

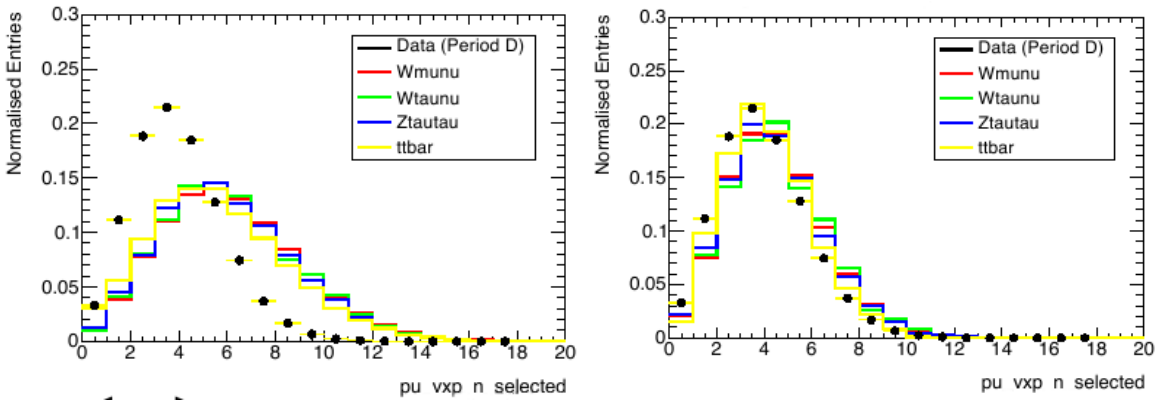


Figure 6.2: The number of pile-up vertices in data and Monte Carlo before re-weighting (left) and after (right) for period D of the 2011 data run.

6.3 Analysis Cuts

The analysis is done in four steps: event preselection, object preselection, object selection and event selection. The event preselection contains cuts that are ATLAS wide and ensures a good quality of data. In the object preselection, loose selection criteria are required to define muon, electrons, hadronic τ candidates and jets. These objects are used for rejecting events that fake the signal. Tighter cuts are used to define "selected" objects which are used for the event selection, i.e. Z boson selection.

6.3.1 Event Preselection

Primary Vertex Requirement

The reconstruction of the interaction or, else, primary vertex is based on the reconstruction of the tracks of charged particles that transverse the Inner Detector, in particular, the pixel and SCT detectors. The tracks are associated to vertex candidates and a fitting algorithm is run to get the correct vertex position and its corresponding error matrix [86]. In the analysis, collision candidate events are ensured by requiring at least one reconstructed primary vertex with at least

Table 6.4: Limits used to calculate the number of objects in the LAr hole in the case of real data. The event is rejected if at least one object inside the LAr hole is found. For the τ candidates the p_T cut is lower for the muon-hadron channel than for the electron-hadron channel, see Section 6.3.2.

Object	p_T [GeV]	η coordinates	ϕ coordinates
electrons	15	$-0.1 < \eta < 1.55$	$-0.888 < \phi < -0.492$
τ candidates	25 (20)	$-0.1 < \eta < 1.55$	$-0.9 < \phi < -0.5$
jets	20	$-0.2 < \eta < 1.6$	$-0.988 < \phi < -0.392$

four reconstructed tracks.

Jet Cleaning

The cuts explained in Section 5.3.1 in Table 5.5 are applied on data, aiming at vetoing events with non-collision jets. The jets should be reconstructed with the anti-k_t algorithm in a cone of $\Delta R = 0.4$, $p_T > 20$ GeV and $|\eta| < 2.5$. The *medium* set is used for the analysis except for the cut on central jets, $|\eta| < 2.0$, with $|f_{\text{track}}| < 0.05 \quad \&\& \quad f_{\text{EM}} > 0.95$, because it rejects large amount of real τ candidates.

LAr hole Cleaning

During part of the 2011 data taking, a part of the LAr calorimeter was not operational leading to an acceptance hole in the EM calorimeter. This hole may affect the reconstruction of electrons, τ candidates or jets if they lie in it. Hence, data events are rejected if any of these objects points towards the hole. For that preselected electrons, τ candidates and jets are checked whether they are lying within certain η and ϕ region of the detector. The exact criteria are shown in Table 6.4. In Monte Carlo the hole is not present and hence, an event weight is assigned. The weight is calculated as the fraction of the data that the hole is present with respect to the whole data taking period. This weight is 0.38 for the electron channel and 0.44 for the muon channel.

Trigger Requirement

There are two options when triggering for semi-leptonic $Z \to \tau\tau$ events, either trigger on the lepton or on the hadronic τ. In the former case, a single lepton trigger is required to fire. Such a trigger reaches an efficiency plateau almost instantaneously, but the p_{T} threshold is relatively high for $Z \to \tau\tau$ events, $p_{\text{T}} > 20$ GeV. A lower threshold, $p_{\text{T}} > 15$ GeV, is obtained by requiring the lepton to be isolated. In the latter case, triggering on τ jets requires sophisticated rejection algorithms to avoid triggering on multijet events. These algorithms follow the offline identification selection: cut-based, likelihood and BDT. In 2010 and 2011 data taking periods the τ triggers are mainly cut-based. The other two algorithms are used for studies, in order to gain experience with them before they are officially used. A third kind of triggers has been implemented in the 2011 data taking, the combined hadronic τ and lepton triggers. Thus, low offline p_{T} thresholds for both leptons and τ candidates is maintained. The disadvantage of the combined trigger is that an efficiency plateau is reached slowly.

For the 2010 analysis the luminosity is lower and single triggers with low p_{T} thresholds can be used. The muon channel is split in three data-taking periods depending on the minimum unprescaled trigger:

- `EF_mu10_MG` for periods E4 to G1, where the trigger fires when a muon reconstructed with the Muon Girl algorithm and $p_{\mathrm{T}} > 10$ GeV;
- `EF_mu13_MG` for periods G2 to run 167576, similarly to the previous trigger but for muons with $p_{\mathrm{T}} > 13$ GeV; and
- `EF_mu13_MG_tight` for run 167607 to end of period I2, for muons which additionally pass the tight identification criteria.

The method to extract the efficiency of the triggers is explained in Section 5.2.3.

For $\tau_e\tau_{\mathrm{h}}$ decays the `EF_e15_medium` trigger is used, which fires when medium electrons with $E_{\mathrm{T}} > 15$ GeV pass. The trigger efficiencies are given in Table 5.4 in Section 5.1.5.

For the 2011 analysis complex high level triggers are chosen; a combined hadronic τ and electron for the electron channel and an isolated muon one for the muon channel.

Specifically, the `EF_tau16_loose_e15_medium` for $\tau_e\tau_{\mathrm{h}}$ fires if there is at least one hadronic τ that passes the loose identification criteria with $p_T > 16$ GeV and at least one electron that satisfies the medium identification criteria with $p_T > 15$ GeV. The turn on curve of the τ part of the trigger reaches the efficiency plateau at 25 GeV, see Fig. 6.3(a,b). The trigger was active during data taking periods F through I ($\mathcal{L} = 1.3\,\mathrm{pb}^{-1}$). In Monte Carlo simulations, on the other hand, this trigger is not present, since an optimisation of the τ trigger hypothesis took place after the production was finished. Hence, on simulated events only the electron part of the trigger is required to have fired (`EF_e15_medium`). To account for trigger inefficiencies in Monte Carlo, a Z tag-and-probe study is conducted for both the electron and the τ parts of the trigger. The resulting efficiencies are shown in Fig. 6.3(a) and (b) for the τ part. In Fig. 6.3(c) for the electron trigger efficiency is given as a reminder, although its estimation is explained in detail in Section 5.1.5. Due to the lack of simulated τ trigger, the study is done only on data and the resulting efficiencies are applied as event weights to the analysis. For the electron part, scale factors are evaluated and applied as well to the analysis [56].

Triggering on $\tau_\mu\tau_{\mathrm{h}}$ events, the `EF_mu15i` and `EF_mu15i_medium` triggers are used. The first one is active from data taking periods B through I and the latter for J, resulting in a total luminosity of $\mathcal{L} = 1.55\,\mathrm{fb}^{-1}$. The triggers and their efficiencies are described in Section 5.2.3.

6.3.2 Object Preselection

The object preselection cuts are different for 2010 and 2011 analyses due to the higher thresholds in the triggers used in 2011.

Muons

For the 2010 dataset: *Combined* muons are reconstructed with the STACO algorithm with $p_{\mathrm{T}} > 15$ GeV and $|\eta| < 2.4$. It is required that the longitudinal distance of the muon track to the primary vertex is less than 10 mm.

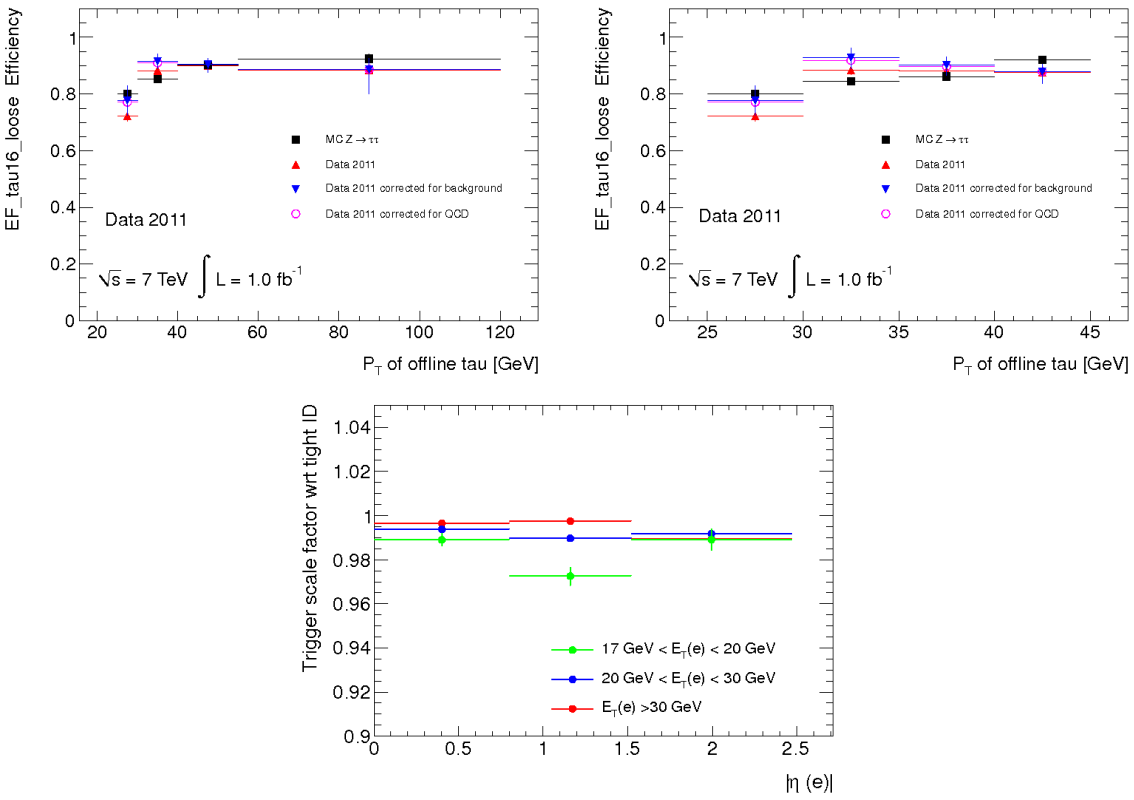


Figure 6.3: Top Left: EF_tau16_loose trigger efficiency with respect to transverse momentum (p_{T}) of probe τ. Top Right: EF_tau16_loose trigger efficiency with respect to transverse momentum (p_{T}) of probe τ, zoomed into the region most relevant for the present study. Bottom: EF_e15_medium trigger scale factor with respect to *tight* electrons. All plots taken from [56].

For the 2011 dataset: A preselected muon is required to have $p_T > 6$ GeV and lie within $|\eta| < 2.7$. It should satisfy the "loose" identification criteria which includes combined, segmented-tagged and stand-alone muons, see Section 5.2.1. The specific low p_{T} threshold is required for rejecting non-leading muons that can fake a τ candidate.

Scale factors accounting for differences between data and Monte Carlo on muon reconstruction and p_{T} resolution effects are applied.

Electrons

For the 2010 dataset: Preselected electrons fulfil the *medium* identification criteria, see Section 5.1.2, have $E_T > 15$ GeV and lie within $|\eta| < 2.47$ excluding the transition region between the barrel and the endcap calorimeters, $1.37 < |\eta| < 1.52$. In order to account for detector regions with problems, the *Object Quality maps* are defined. These are two dimensional (η, ϕ) histograms which map the faulty regions. In case an electron is found to have been reconstructed in a problematic area it is rejected. The same is done in Monte Carlo to ascertain similar efficiencies. For the Monte Carlo the run that represents the most current detector status is used. For this analysis this run is the 167521.

For the 2011 dataset: *Medium* electrons with $E_T > 15$ GeV are defined as preselected electrons. It is required that they lie within $|\eta| < 2.7$ excluding the transition region between the barrel and the endcap calorimeters, $1.37 < |\eta| < 1.52$. Object Quality criteria[2] are also checked and only electrons with quality status "good" are accepted.

All electrons in the data events are corrected for energy mis-calibration according to the factors calculated with the $Z \to e^+e^-$ and $J/\psi \to e^+e^-$ in-situ calibration, see Section 5.1.3. These electrons are used for the LAr hole cleaning, Section 6.3.1, for electron-muon and electron-τ candidate overlap removal, later in this section, and for rejecting electrons that fake $\tau\tau$ events, namely the dilepton veto cut, see Section 6.3.5.

τ Candidates

For the 2010 dataset: τ candidates are defined with $p_{\mathrm{T}} > 20$ GeV and within $|\eta| < 2.47$ excluding the transition region between the barrel and the endcap calorimeters, $1.37 < |\eta| < 1.52$. No identification criterion is required at this level.

For the 2011 dataset: Preselected τ candidates are required to have $p_{\mathrm{T}} > 20(25)$ GeV for $\tau_\mu\tau_{\mathrm{h}}(\tau_e\tau_{\mathrm{h}})$ channel and lie within $|\eta| < 2.47$ excluding the transition region between the barrel and the endcap calorimeters, $1.37 < |\eta| < 1.52$. A different p_{T} cut for the two channels is motivated by the different triggers used. That is, a single muon trigger for the $\tau_\mu\tau_{\mathrm{h}}$ which allows for a lower cut on the τ candidate p_{T}, while in the $\tau_e\tau_{\mathrm{h}}$ the combined τ+e trigger does not. As mentioned in Section 6.3.1 the τ part of the combined trigger does not reach a plateau before 25 GeV. An additional geometrical cut in pseudorapidity is required to reject fake electrons in $|\eta| < 0.03$. The preselected τ candidates are used only for the LAr hole cleaning checks.

Jets

Although not part of the main analysis, jets are defined. The jets are reconstructed with the anti-k_T algorithm with distance parameter $\Delta R = 0.4$. Local calibrated (LC) jets with $p_{\mathrm{T}} > 20$ GeV and within $|\eta| < 4.5$ are used for the jet cleaning cuts (2010 and 2011) and for the LAr hole cleaning checks (2011), see Section 6.3.1.

Overlap Removal

Any of the above objects might be reconstructed from the same localised response in the ATLAS detector, thus an overlap removal procedure is performed. The process starts from muon and electron candidates, since they can be selected with higher purity. Then one checks whether a τ candidate overlaps with either lepton. A jet is removed if it overlaps with any of the other objects. The tolerable cone for lepton-lepton overlapping is $\Delta R > 0.2$ and $\Delta R > 0.4$ for τ-lepton, τ-jet or jet-lepton.

[2]Electrons are required not to have been flagged with certain "defects" during the Data Quality assessment. Some information is given in Section 3.4 and more details can be found in the ATLAS-restricted twiki page [87].

6.3.3 Object Selection

For the main analysis tighter selection cuts are required for the analysis-relevant objects: muons, electrons and τ candidates.

Muons

For the 2010 dataset: Additional cuts are Inner Detector track requirements [88]. A muon is accepted if:

- if it is expected to hit the B layer then it should have at least 1 hit in the B layer, if not there should be no hit at all;
- the sum of the hits and the dead sensors in the pixel detector are larger than 1;
- the total number of SCT hits and dead sensors are more than 5;
- the number of pixel and SCT holes are less than 2 and
- the following TRT outliers prerequisites are true
  - $|\eta| < 1.9$:
    $\frac{\text{nTRT Outliers}}{(\text{nTRT Hits} + \text{nTRT Outliers})} < 0.9$ and (nTRT Hits + nTRT Outliers) > 5
  - $|\eta| >= 1.9$:
    $\left(\frac{\text{nTRT Outliers}}{(\text{nTRT Hits} + \text{nTRT Outliers})} < 0.9 \quad \text{and} \quad (\text{nTRT Hits} + \text{nTRT Outliers}) > 5\right)$
    or (nTRT Hits + nTRT Outliers) < 6.

For the 2011 dataset: Muons with transverse momentum higher that 17 GeV and within $|\eta| < 2.4$ are used for the main analysis. Protection against cosmic-ray or background muons, for example from B meson decays, is taken by requiring the longitudinal distance of the muon track from the primary vertex to be less than 10 mm. Further Inner Detector quality criteria are applied to "loose" muons. These are the same as for 2010 data except for a looser cut in the number of pixel and SCT holes that should be less than 3.

Electrons

Tight electrons with $p_{\mathrm{T}} > 16$ GeV in the 2010 analysis and $p_{\mathrm{T}} > 17$ GeV in the 2011 are required.

τ Candidates

For the 2010 dataset: Cut-based identification criteria are set at the selected τ candidates. One-prong τ candidates are required to pass the *medium* identification criteria and the three-prong τ candidates the *tight* criteria, see Section 5.4.3. Electrons faking τ candidates are vetoed with a cut-based electron veto, see Section 5.4.4. For the analysis the *tighter* working point is used. Scaling factors are used to account for discrepancies between data and Monte Carlo efficiencies for the electron veto and the identification, see Section 5.4.1.

For the 2011 dataset: The "loosely" defined τ candidates at the preselection step are constrained further by requiring the candidates to pass the *medium* BDT identification cuts, see Section 5.4.1. The choice of the identification method and signal efficiency is made based on acceptance studies of the three methods to the full analysis. Electrons faking single-prong τ candidates are suppressed with a cut-based electron veto at *tight* rejection efficiency, see Section 5.4.4. The data-Monte Carlo discrepancies are accounted for with a correction factor applied to τ candidates in Monte Carlo matched to a true electron within $\Delta R = 0.2$.

6.3.4 Lepton Isolation

Electrons or muons coming from leptonic τ decays are isolated compared to leptons from QCD multijets. In ATLAS, tracker- and calorimeter- based variables are used for defining such isolation cuts. The former is defined as the scalar sum of the transverse momentum of tracks in a hollow cone of radius X = 0.2, 0.3, 0.4 about the seed lepton ($I^X_{p_\mathrm{T}}$). Similarly, the latter is calculated as the sum of the energy deposit on the EM calorimeter in a hollow cone about the seed lepton ($I^X_{E_\mathrm{T}}$). Often the isolation cone parameter is divided by the p_T or E_T of the candidate lepton to reduce the dependence on the momentum of the lepton. For the $Z \to \tau\tau$ analysis the following variables are studied for the $\tau_\mu\tau_\mathrm{h}$

- $I^{0.2}_{E_\mathrm{T}}/p_\mathrm{T}$, $I^{0.3}_{E_\mathrm{T}}/p_\mathrm{T}$, $I^{0.4}_{E_\mathrm{T}}/p_\mathrm{T}$ and
- $I^{0.2}_{p_\mathrm{T}}/p_\mathrm{T}$, $I^{0.3}_{p_\mathrm{T}}/p_\mathrm{T}$, $I^{0.4}_{p_\mathrm{T}}/p_\mathrm{T}$

and for the $\tau_e\tau_\mathrm{h}$

- $I^{0.3}_{E_\mathrm{T}}/E_\mathrm{T}$, $I^{0.4}_{E_\mathrm{T}}/E_\mathrm{T}$ in 2010, $I^{0.3}_{E_\mathrm{T}\mathrm{corrected}}/E_\mathrm{T}$, $I^{0.4}_{E_\mathrm{T}\mathrm{corrected}}/E_\mathrm{T}$ in 2011 and
- $I^{0.3}_{p_\mathrm{T}}/p_\mathrm{T}$, $I^{0.4}_{p_\mathrm{T}}/p_\mathrm{T}$.

In 2011, the E_T isolation cone is corrected for pile-up contributions.

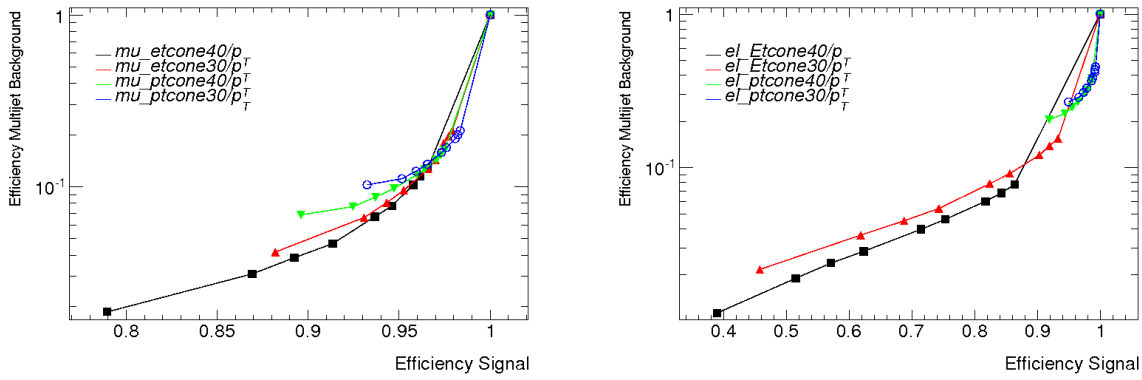
2010 data run The signal efficiency with respect to the QCD multijet efficiency of those variables is shown in Fig. 6.4, from which the optimal working points are chosen to be

- for the $\tau_\mu\tau_\mathrm{h}$: $I^{0.3}_{E_\mathrm{T}}/p_\mathrm{T} < 0.1$ and $I^{0.4}_{p_\mathrm{T}}/p_\mathrm{T} < 0.06$,
- for the $\tau_e\tau_\mathrm{h}$: $I^{0.4}_{E_\mathrm{T}}/E_\mathrm{T} < 0.06$ and $I^{0.4}_{p_\mathrm{T}}/p_\mathrm{T} < 0.06$.

In the 2011 data run the optimal working points are chosen to be

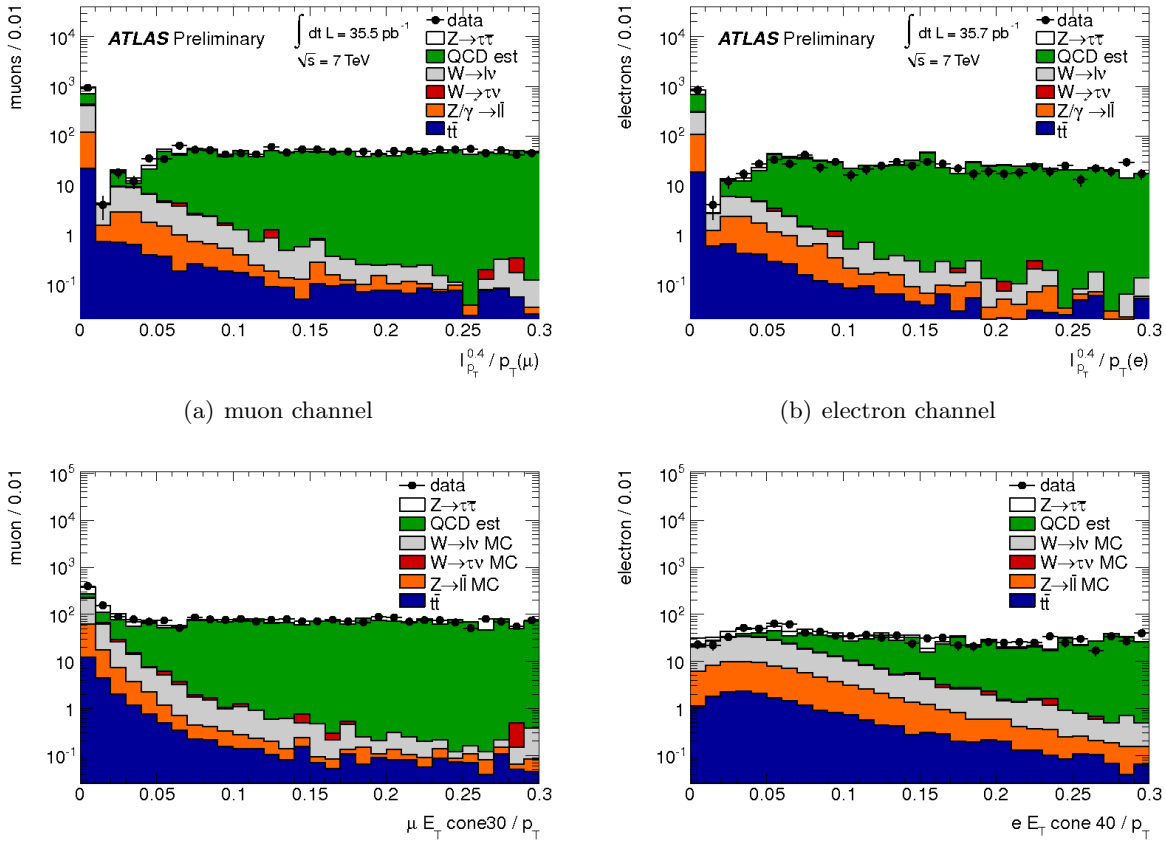
- for the $\tau_\mu\tau_\mathrm{h}$: $I^{0.3}_{E_\mathrm{T}}/p_\mathrm{T} < 0.04$ and $I^{0.4}_{p_\mathrm{T}}/p_\mathrm{T} < 0.03$,
- for the $\tau_e\tau_\mathrm{h}$: $I^{0.4}_{E_\mathrm{T}\mathrm{corrected}}/E_\mathrm{T} < 0.1$ and $I^{0.4}_{p_\mathrm{T}}/p_\mathrm{T} < 0.03$.

Fig. 6.5 and 6.6 show the distributions for the chosen isolation variables for the two channels for 2010 and 2011 analyses. The chosen cuts result in rejecting most of the multijet background while keeping the $Z \to \tau\tau$ signal. The multijet background is estimated from data that the τ candidate and the lepton have the same sign (SS data), Section 6.4.3. The rest of the electroweak background processes are estimated from Monte Carlo.



(a) muon channel (b) electron channel

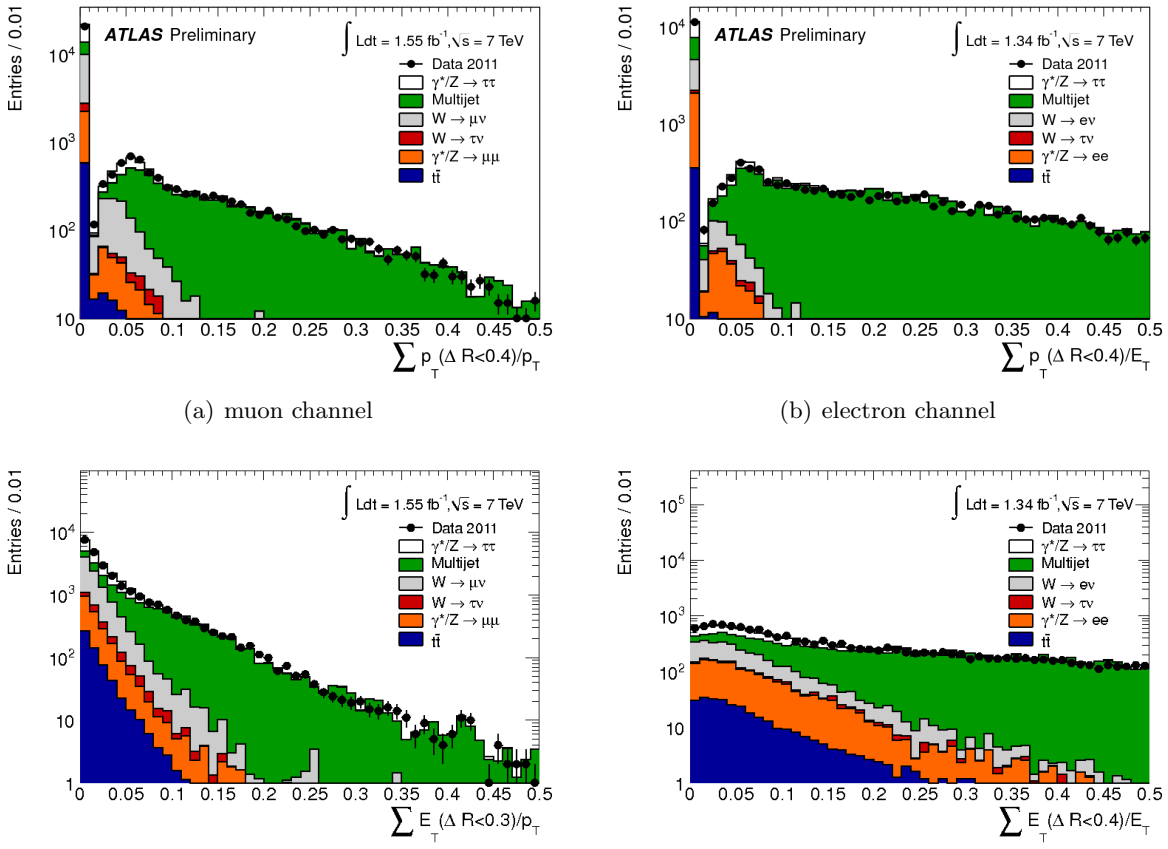
Figure 6.4: Signal versus multijet background efficiency for electrons (left) and muons (right) for the 2010 analysis. The signal is taken from Monte Carlo simulations, while the QCD multijet is estimated from same sign data events.



(a) muon channel (b) electron channel

(c) muon channel (d) electron channel

Figure 6.5: Comparison between data 2010 and Monte Carlo simulated events of the isolation variables for both channels. The distributions are produced requiring opposite sign-same sign $\tau_e\tau_{\rm h}$ and $\tau_\mu\tau_{\rm h}$ events. No further event selection has been made.



(a) muon channel

(b) electron channel

(c) muon channel

(d) electron channel

Figure 6.6: Comparison between data 2011 and Monte Carlo simulated events of the isolation variables for both channels. The distributions are produced requiring opposite sign-same sign $\tau_e\tau_{\mathrm{h}}$ and $\tau_\mu\tau_{\mathrm{h}}$ events. No further event selection has been made.

6.3.5 Event Selection

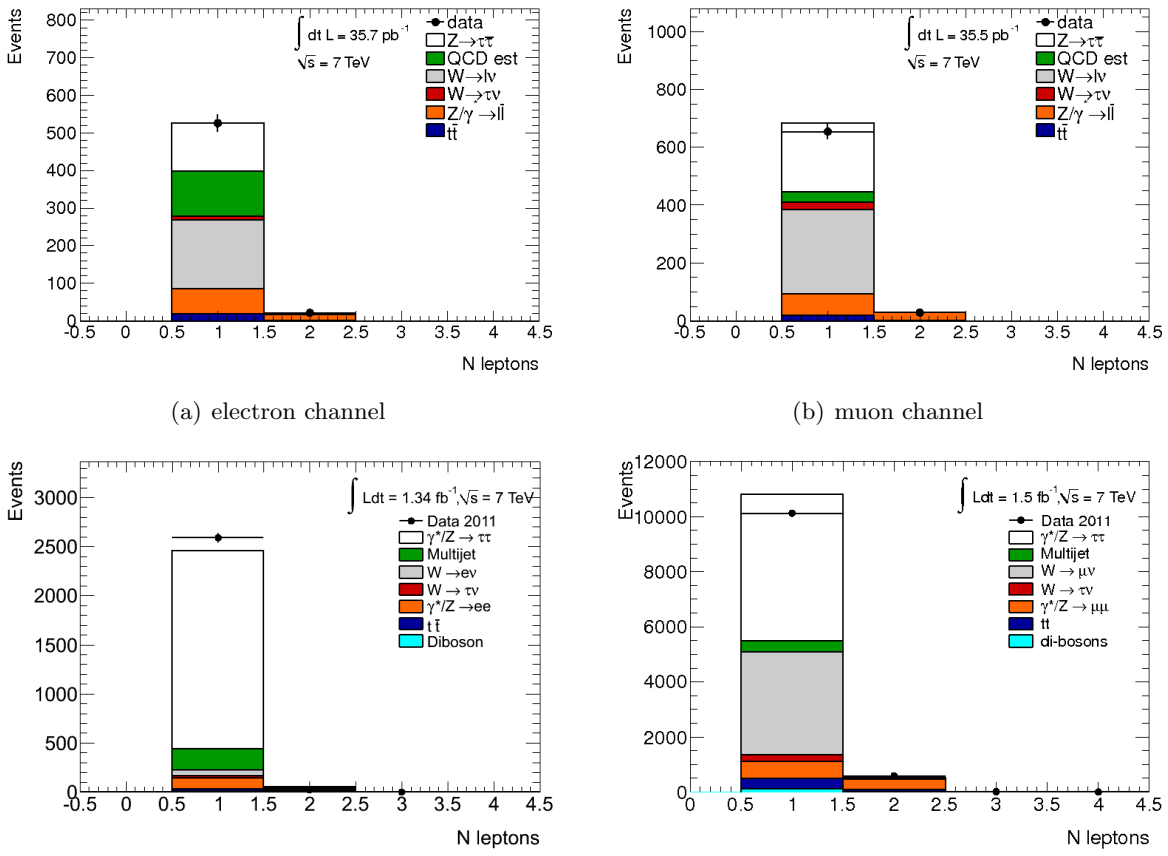
The objects defined earlier (selected and preselected) are used for the $Z \to \tau\tau \to \ell\,\tau_{\mathrm{h}} + 4\nu$ cuts described below.

Dilepton Veto

Events coming from $Z \to \ell\ell +$ jets or $Z \to \tau\tau \to \ell\,\ell + 4\nu$ are vetoed using the preselected electrons and muons. The distribution of the number of preselected leptons can be seen in Fig. 6.7 for the 2010 (top) and 2011 (bottom) analysis.

Opposite Sign Requirement

Further reduction of the multijet background is achieved by requiring the lepton (e or μ) and the τ candidate to be oppositely charged. Naturally, all other electroweak processes in which a QCD jet is faking a τ candidate are suppressed, too.



(a) electron channel

(b) muon channel

(c) electron channel

(d) muon channel

Figure 6.7: Distribution of the number of preselected leptons for the events which have passed all event selection cuts except for the dilepton veto. In the top row, the distributions for the 2010 data is shown and in the bottom row, the ones for the 2011 data.

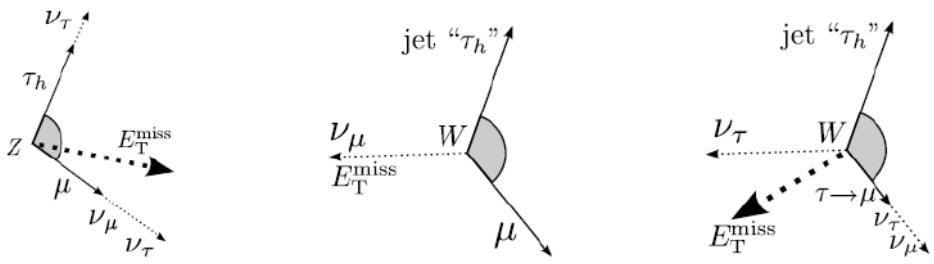
W+jets suppression Cuts

The second larger background process after QCD multijets is W+jets. This decay can fake the signal in two ways: the W decays to a lepton and a jet is misidentified as a τ candidate, or the W decays to a hadronic τ and a multijet is mis-identified for a lepton. A cut on the transverse mass (m_T) of the lepton and the $E_{\mathrm{T}}^{\mathrm{miss}}$ reduces this background because the W boson escapes almost entirely in the transverse plane. The m_T is calculated from

$$m_{\mathrm{T}}(\ell, E_{\mathrm{T}}^{\mathrm{miss}}) = \sqrt{2\, p_{\mathrm{T}}(\ell) \cdot E_{\mathrm{T}}^{\mathrm{miss}} \cdot (1 - \cos\Delta\phi(\ell, E_{\mathrm{T}}^{\mathrm{miss}}))}, \tag{6.1}$$

where $p_{\mathrm{T}}(\ell)$ the transverse momentum of the electron or muon and $\phi(\ell, E_{\mathrm{T}}^{\mathrm{miss}})$ the angle between the decaying lepton and the missing transverse energy.

A secondary cut is applied taking advantage of the topology of the W and Z decays. Z bosons are heavier than the τ leptons, so the latter are produced boosted and their decay products are collimated along the τ lepton direction. In most of the cases the transverse energy of the Z bosons is small and the τ leptons decay back-to-back. Experimentally, the p_{T} of the neutrinos is summed up as one $E_{\mathrm{T}}^{\mathrm{miss}}$ vector that lies within the angle of the two decay leptons,



(a) $Z \to \tau\tau \to \mu\tau_{\mathrm{h}}$ (b) $W \to \mu\nu$ (c) $W \to \tau\nu \to \mu\nu\nu\nu$

Figure 6.8: Left plot shows the sum of the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector lies between the τ lepton decay products in the case of $Z \to \tau\tau$ decay. Middle and right sketches show a $W \to \mu\nu$ and $W \to \tau\nu$ decay, respectively. In this case, the transverse momenta of the mother and daughter particles are balanced and, the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector lies outside the angle that the decay products form.

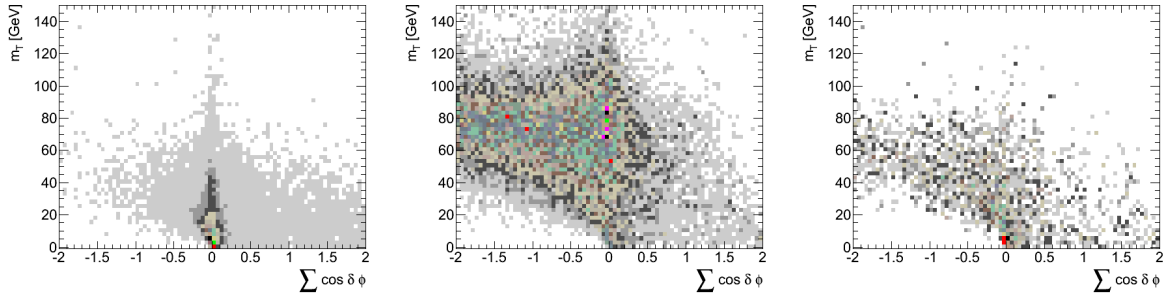


Figure 6.9: Correlation between the $\sum \cos\delta\phi$ and the transverse mass for signal (left), background $W \to e\nu$ (middle) and $W \to \tau\nu$ (right) Monte Carlo events.

see Fig. 6.8(a). In the case of $W \to \ell\nu$ and $W \to \tau\nu$ decays, the lepton, the jet and the neutrino(s) all fly at different directions, therefore the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector tends to point outside the angle that the lepton and the jet or τ candidate form, as shown in Fig. 6.8(b,c). A variable defined to express this is

$$\sum \cos\Delta\phi = \cos\left(\phi(\ell) - \phi(E_{\mathrm{T}}^{\mathrm{miss}})\right) + \cos\left(\phi(\tau_{\mathrm{h}}) - \phi(E_{\mathrm{T}}^{\mathrm{miss}})\right). \tag{6.2}$$

The correlation between the two variables, m_T and $\sum \cos\Delta\phi$, for signal and W background is shown in Fig. 6.9. The optimal cutting points for the analyses are

$$m_T < 50 \text{ GeV} \quad \text{and} \quad \sum \cos\delta\phi > -0.15. \tag{6.3}$$

One can see in Fig. 6.10, 6.11 the $\sum \cos\delta\phi$ and transverse mass distributions for 2010 and 2011 data, respectively. The $Z \to \tau\tau$ events rest at low m_T values and $\sum \cos\delta\phi \sim 0$, while the W+jets on $m_T \sim 80$ GeV and $\sum \cos\delta\phi < 0$.

τ candidate Event Cuts

Further suppression of processes that fake τ candidates is achieved by the following cuts. The τ candidate has to have either one or three associated tracks (1- or 3-prong hadronic τs) and EM charge ± 1. An additional cut is applied to suppress muons faking τ candidates in the 2011 data analysis. That is the fraction of the electromagnetic energy deposited in the calorimeter, f_{EM}, has to be larger than 0.1. The distribution of the number of tracks is shown in Fig. 6.12 for 2010 (top) and 2011 data (bottom) and the one for the f_{EM} in Fig. 6.13.

Visible Mass Reconstruction

After all selection cuts have been applied the invariant mass of the two τ leptons is calculated. The simplest approach for the $\tau\tau$ decays is to calculate the invariant mass of the visible decay products, i.e. the hadronic τ candidate and the lepton (e/μ). In Fig. 6.14 the reconstructed visible mass is plotted for the 2010 (top) and 2011 (bottom) dataset. The distributions peak lower than the real Z mass, since the method does not account for the neutrinos involved in the decay. Other methods that take into account the $E_{\mathrm{T}}^{\mathrm{miss}}$ of the event and the $Z \to \tau\tau \to \ell\,\tau_{\mathrm{h}} + 4\nu$ topology exist and are studied in Chapter 9. The additional cut on the electromagnetic fraction of the energy deposited in the calorimeter that is introduced in the 2011 analysis helped suppressing the $Z \to \mu^+\mu^-$ background significantly, as it can be seen by comparing the distributions in Fig. 6.14 (b) and (d). A similar cut has not been found for the electron channel resulting in a large amount of $Z \to e^+e^-$ background left contaminating the signal, see Fig. 6.14 (a) and (c).

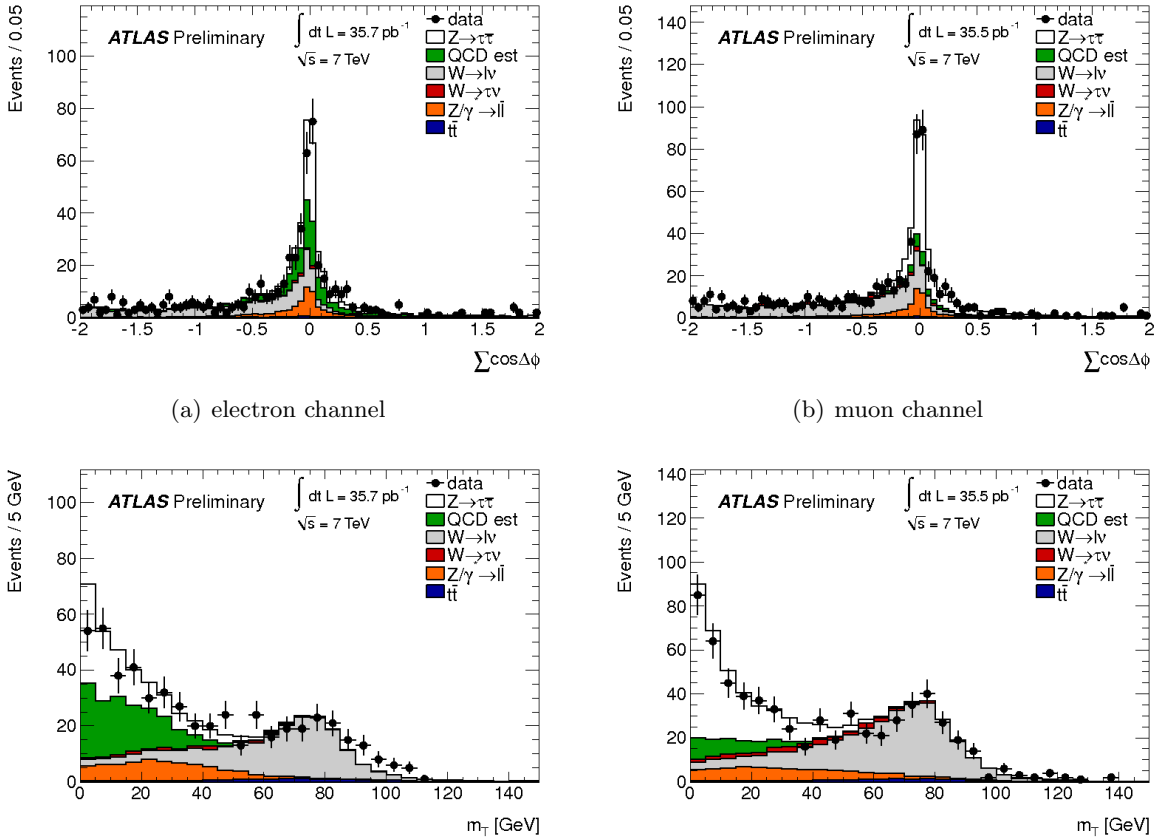
A summary of all selection cuts applied, together with the event yields at each cut step for data, signal Monte Carlo and the various backgrounds is given in Table 6.5 for the $\tau_e\tau_{\mathrm{h}}$ channel, Table 6.6 for the $\tau_\mu\tau_{\mathrm{h}}$ channel for the 2010 analysis and Table 6.7 for the $\tau_\mu\tau_{\mathrm{h}}$ channel and Table 6.8 for the $\tau_e\tau_{\mathrm{h}}$ channel for the 2011 analysis. The event yields of the backgrounds are estimated as described in Section 6.4.

Table 6.5: Numbers of events passing the cumulative event selections for the muon channel. The statistical errors on the least significant digits are given in the parentheses.

	data	$\gamma^*/Z \to \tau\tau$	multijets	$\gamma^*/Z \to \mu\mu$	$W \to \mu\nu$	$W \to \tau\nu$	$t\bar{t}$	Diboson
object selection	1365	261(3)	163(9)	216(2)	649(6)	54(3)	38.9(5)	8.6(1)
dilepton veto	1291	260(3)	162(8)	125(2)	648(6)	54(3)	34.3(5)	7.8(1)
W suppression cuts	462	231(3)	90(4)	58(1)	66(2)	18(2)	7.8(2)	1.34(5)
$m_{\text{vis}} = 35 - 75$ GeV	327	205(2)	71(3)	23.1(9)	23(1)	10(1)	2.4(1)	0.49(3)
$N_{\text{trk}}(\tau_{\text{h}}) = 1$ or 3, $\lvert Q(\tau_{\text{h}})\rvert = 1$	247	187(2)	42(3)	15.3(7)	12.1(8)	5(1)	1.4(1)	0.32(2)
opposite sign	213	186(2)	23(3)	11.1(5)	9.3(7)	3.6(8)	1.3(1)	0.28(2)

Table 6.6: Numbers of events passing the cumulative event selections for the electron channel. The statistical errors on the least significant digits are given in the parentheses.

	data	$\gamma^*/Z \to \tau\tau$	multijets	$\gamma^*/Z \to ee$	$W \to e\nu$	$W \to \tau\nu$	$t\bar{t}$	Diboson
object selection	1203	141(2)	402(12)	164(1)	409(4)	24(2)	33.0(4)	6.4(1)
dilepton veto	1144	140(2)	400(11)	116(1)	409(4)	24(2)	29.1(4)	5.9(1)
W suppression cuts	449	125(2)	159(6)	70(1)	43(1)	10(1)	6.7(2)	0.98(4)
$m_{\text{vis}} = 35 - 75$ GeV	273	107(1)	95(4)	19.2(7)	12.8(7)	3.7(7)	1.7(1)	0.32(2)
$N_{\text{trk}}(\tau_{\text{h}}) = 1$ or 3, $\lvert Q(\tau_{\text{h}})\rvert = 1$	180	98.5(1)	53(4)	11.0(5)	6.7(5)	1.8(5)	1.13(9)	0.21(2)
opposite sign	151	98(1)	25(3)	6.9(5)	4.8(4)	1.5(4)	1.02(8)	0.18(1)



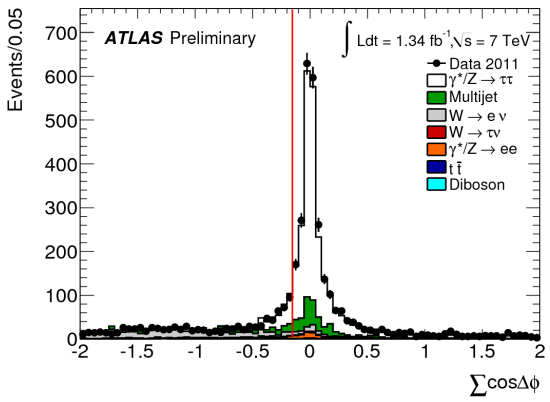
(a) electron channel

(b) muon channel

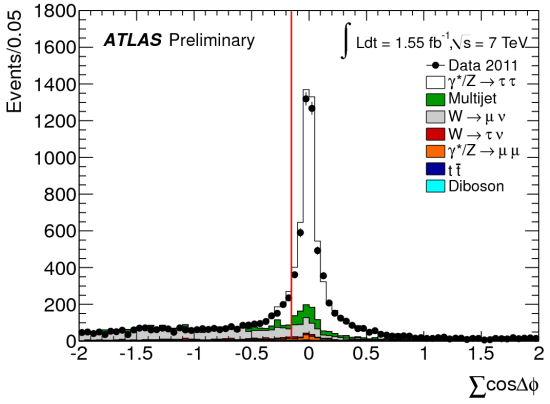
(c) electron channel

(d) muon channel

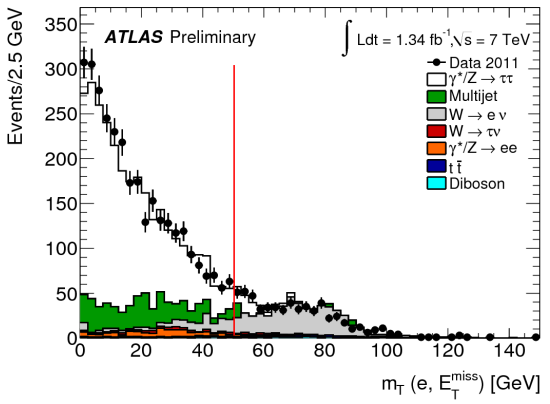
Figure 6.10: Distributions of $\sum \cos\Delta\phi$ (top) and m_{T} (bottom) for 2010 data events which have passed all event selection cuts except the cuts on these two variables for the $\tau_e\tau_{\mathrm{h}}$ (left) and the $\tau_\mu\tau_{\mathrm{h}}$ (right). The red line indicates the value of the corresponding cut.



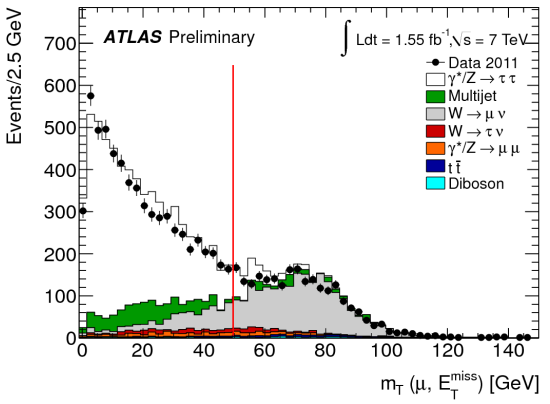
(a) electron channel



(b) muon channel

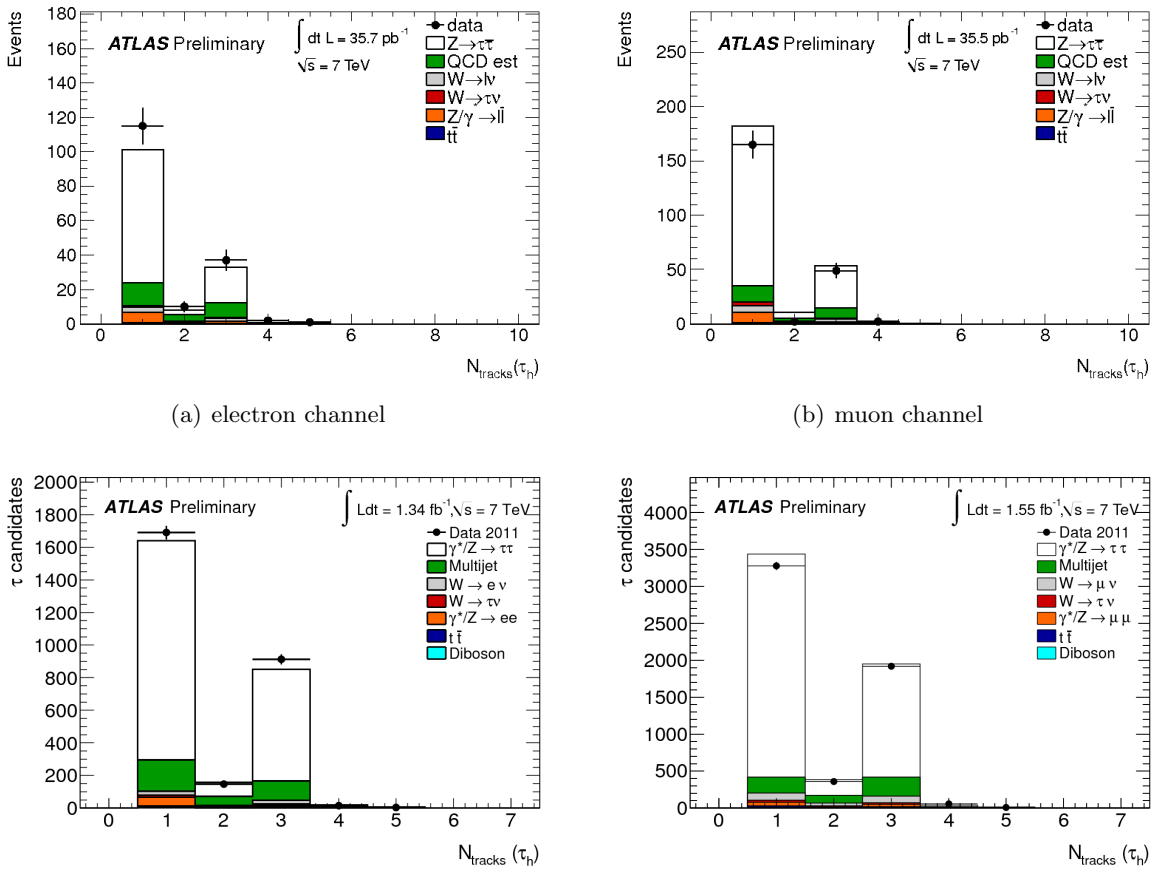


(c) electron channel



(d) muon channel

Figure 6.11: Distributions of $\sum \cos\Delta\phi$ (top) and m_{T} (bottom) for the 2011 data events which have passed all event selection cuts except the cuts on these two variables for the $\tau_e\tau_{\mathrm{h}}$ (left) and the $\tau_\mu\tau_{\mathrm{h}}$ (right).



(a) electron channel

(b) muon channel

(c) electron channel

(d) muon channel

Figure 6.12: Number of tracks associated with the τ candidate, after all cuts of the event selection except for the final τ candidate related cuts for the 2010 (top) and 2011 (bottom) data taking and for the $\tau_e\tau_{\mathrm{h}}$ (left) and the $\tau_\mu\tau_{\mathrm{h}}$ channel (right).

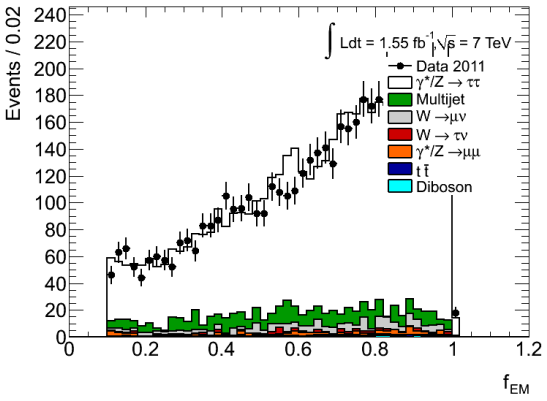
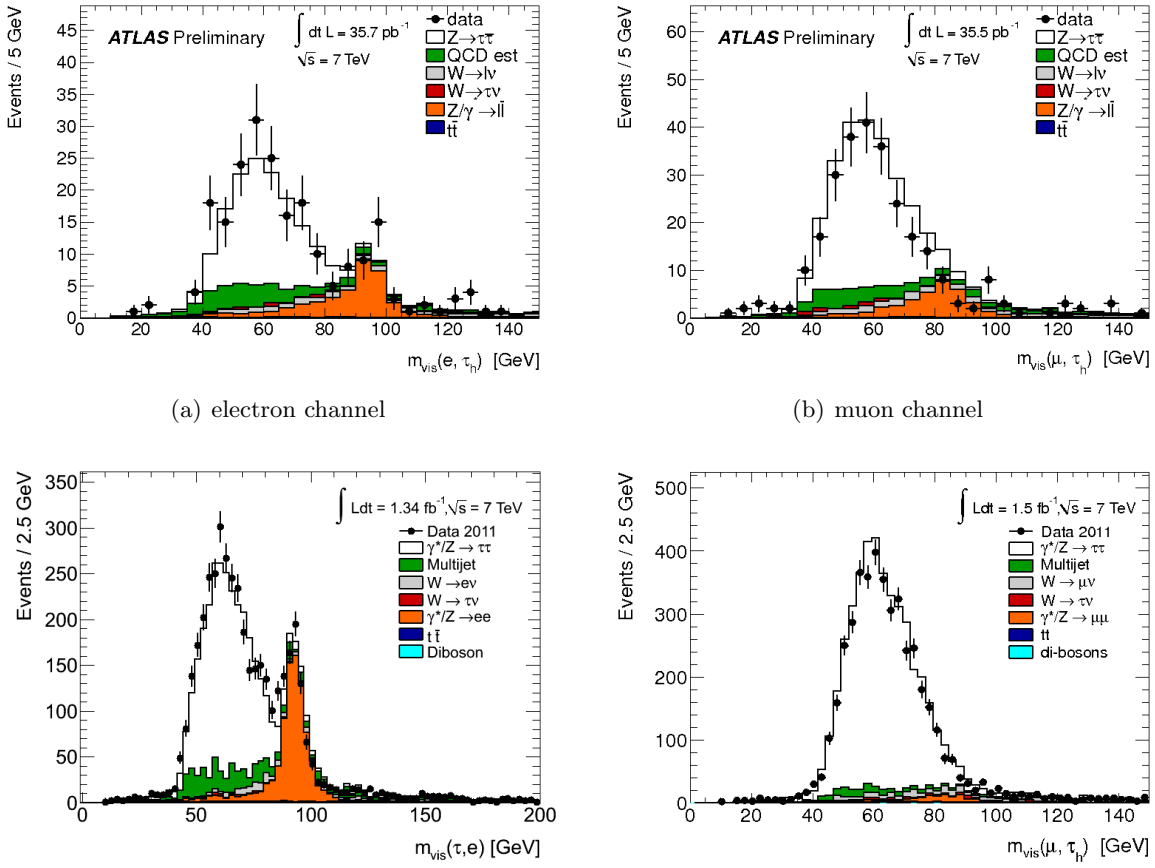


Figure 6.13: Distribution of f_{EM} for the muon channel after the final event selection in the 2011 data.



(a) electron channel

(b) muon channel

(c) electron channel

(d) muon channel

Figure 6.14: Distribution of the visible mass after all event selection cuts for the 2010 (top) and 2011 data (bottom) and for the $\tau_e\tau_{\mathrm{h}}$ (left) and the $\tau_\mu\tau_{\mathrm{h}}$ (right) .

Table 6.7: Cutflow table for $\tau_\mu\tau_\mathrm{h}$ channel. In brackets, the statistical error is shown.

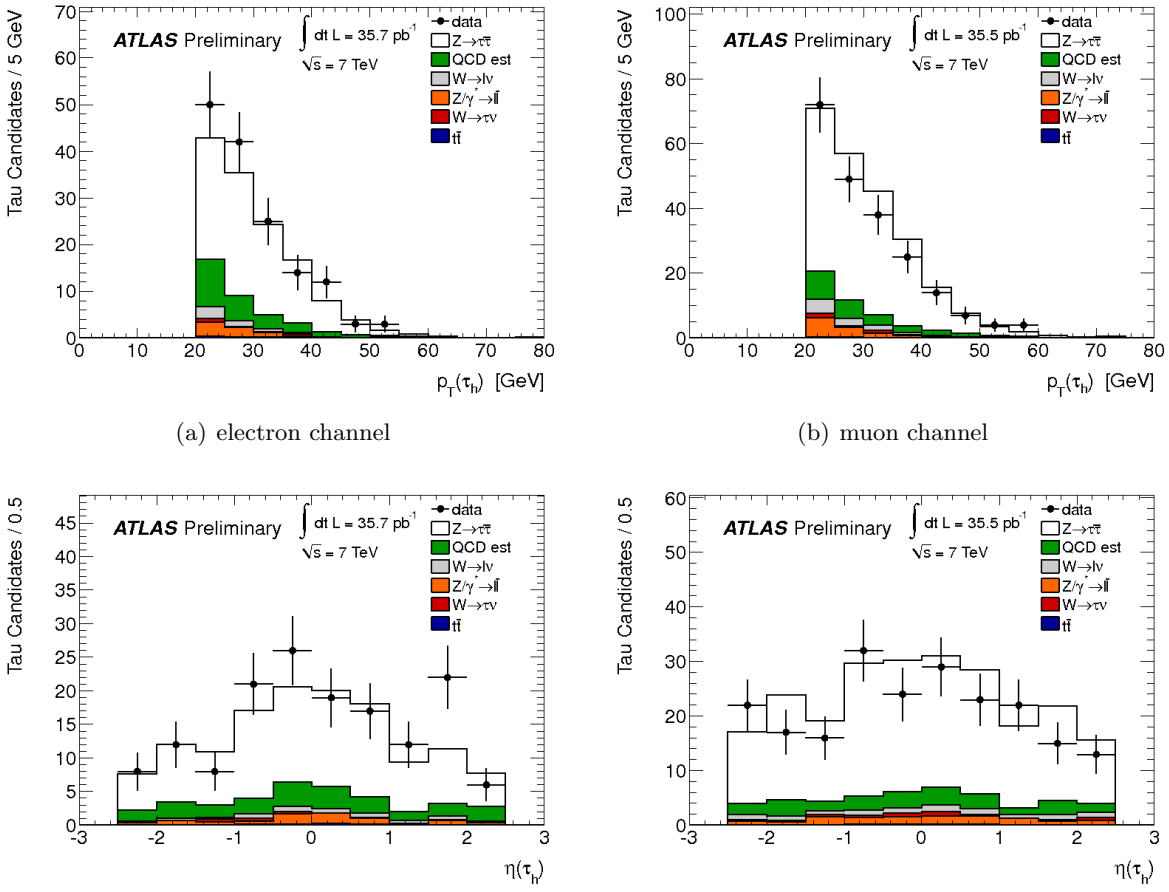
	Data 2011	$Z/\gamma^* \to \tau\tau$	$W \to \mu\nu$	$W \to \tau\nu$	$Z/\gamma^* \to \mu\mu$	$t\bar{t}$	Dibosons	multijets
Object selection	35020(187)	7857(63)	28341(243)	1862(64)	4205(39)	950(6)	352(8)	
Opposite sign	17550(133)	7059(60)	6886(86)	461(23)	1612(25)	577(5)	198(7)	1363(108)
Dilepton veto	16470(128)	7027(60)	6859(86)	459(23)	921(21)	458(5)	181(6)	1265(100)
W cuts	7397(86)	5788(54)	582(22)	129(10)	408(15)	102(2)	32(3)	764(44)
$N_\mathrm{tracks}(\tau_\mathrm{h}) = 1$ or 3	6445(80)	5208(51)	444(19)	103(8)	187(10)	87(2)	27(2)	564(39)
$\lvert charge(\tau_\mathrm{h})\rvert = 1$	6422(80)	5200(51)	442(19)	102(8)	186(10)	87(2)	27(2)	555(39)
$m_\mathrm{vis} = 35 - 75$ GeV	5184(72)	4544(49)	186(13)	49(5)	81(7)	31(1)	15(2)	432(30)

Table 6.8: Cutflow table for $\tau_e\tau_\mathrm{h}$ channel. In brackets, the statistical error is shown.

	Data 2011	$Z/\gamma^* \to \tau\tau$	$W \to e\nu$	$W \to \tau\nu$	$Z/\gamma^* \to ee$	$t\bar{t}$	Dibosons	multijets
Object selection	15200(123)	3393(33)	4660(57)	291(12)	2362(28)	534(4)	174(5)	
Opposite sign	8675(93)	3087(32)	2158(39)	127(7)	1575(24)	340(3)	103(4)	1156(60)
Dilepton veto	8441(92)	3067(31)	2149(39)	127(7)	1450(24)	271(3)	97(4)	1154(58)
W cuts	4649(68)	2570(28)	210(12)	50(4)	900(19)	59(1)	18(2)	726(36)
$N_\mathrm{tracks}(\tau_\mathrm{h}) = 1$ or 3	4358(66)	2456(28)	180(11)	41(4)	879(19)	54(1)	16(1)	593(33)
$\lvert\mathrm{charge}(\tau_\mathrm{h})\rvert = 1$	4351(66)	2453(28)	179(11)	41(4)	878(19)	53(1)	16(1)	584(32)
$m_\mathrm{vis} = 35 - 75$ GeV	2600(51)	2029(25)	45(5)	18(2)	64(4)	17(1)	6(1)	300(21)

6.3.6 Performance Plots data run 2010

After applying the full event selection, a visible mass window cut, 35 GeV $< m_{\text{vis}} <$ 75 GeV, is required. This cut rejects the Z+jets events providing a clean $Z \to \tau\tau$ sample for performance studies and a cross section measurement. Fig. 6.15 - 6.18 show the following distributions for the $\tau_e\tau_{\text{h}}$ channel and the $\tau_\mu\tau_{\text{h}}$ channel after all selection cuts have been applied: p_{T} and η of the selected τ candidates, electrons and muons, $E_{\text{T}}^{\text{miss}}$ and $\Delta\phi$ between the τ and the lepton for the 2010 data analysis.



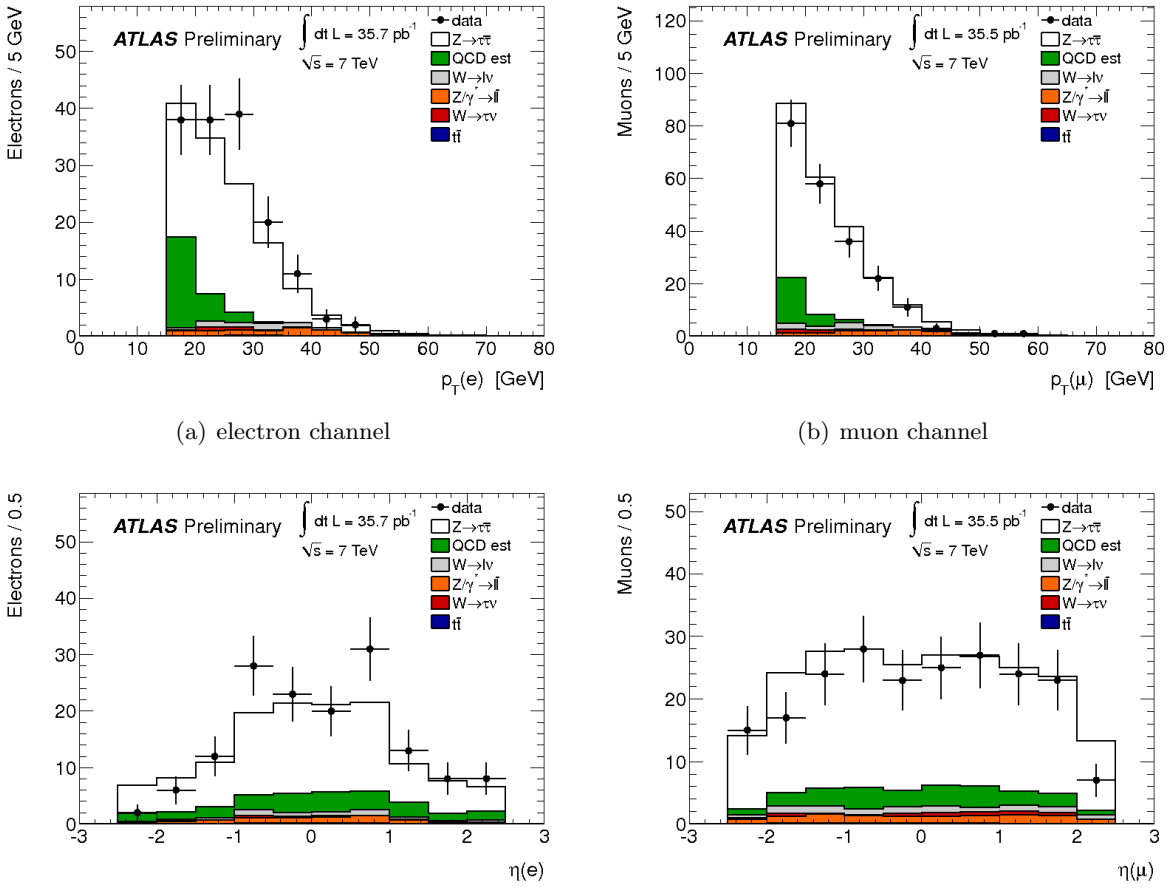
(a) electron channel

(b) muon channel

(c) electron channel

(d) muon channel

Figure 6.15: Distribution of p_{T} and η of the selected τ candidate after the final event selection for the $\tau_e\tau_{\text{h}}$ (left) and the $\tau_\mu\tau_{\text{h}}$ (right) with the 2010 data.



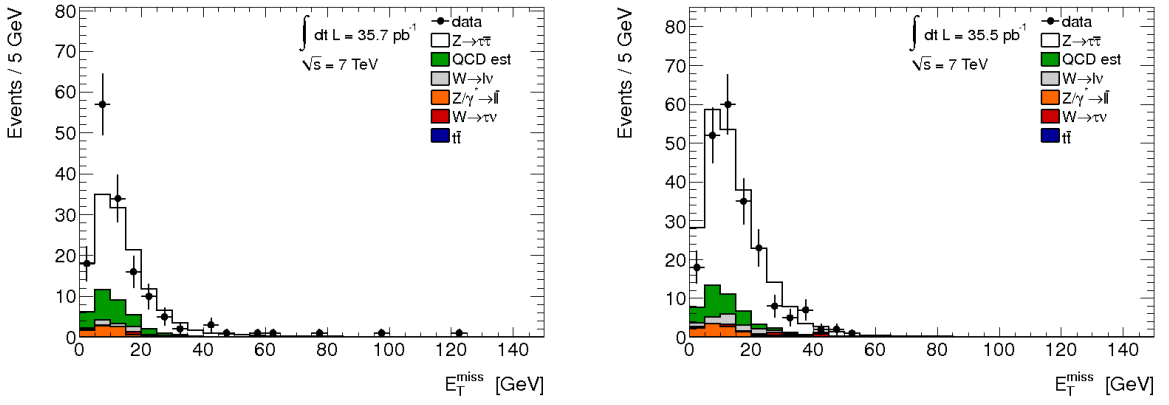
(a) electron channel

(b) muon channel

(c) electron channel

(d) muon channel

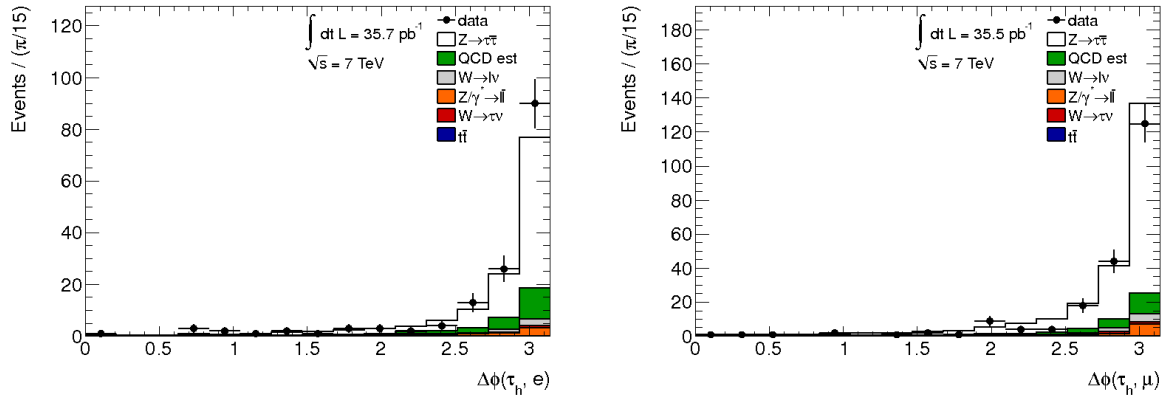
Figure 6.16: Distribution of E_T or p_{T} and η of the selected electron or muon after final event selection for the $\tau_e\tau_{\mathrm{h}}$ (left) and the $\tau_\mu\tau_{\mathrm{h}}$ (right) with the 2010 data.



(a) electron channel

(b) muon channel

Figure 6.17: Distribution of the missing transverse energy after final event selection for the $\tau_e\tau_{\mathrm{h}}$ (left) and the $\tau_\mu\tau_{\mathrm{h}}$ (right) with the 2010 data.



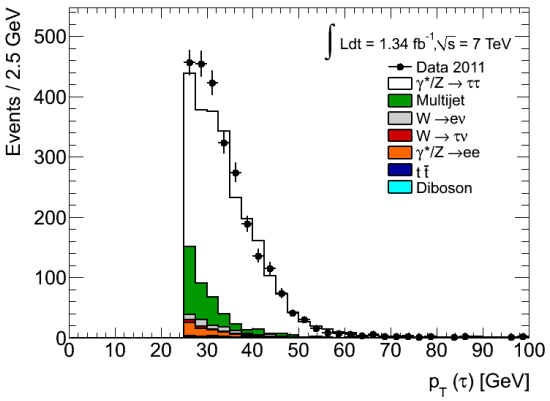
(a) electron channel

(b) muon channel

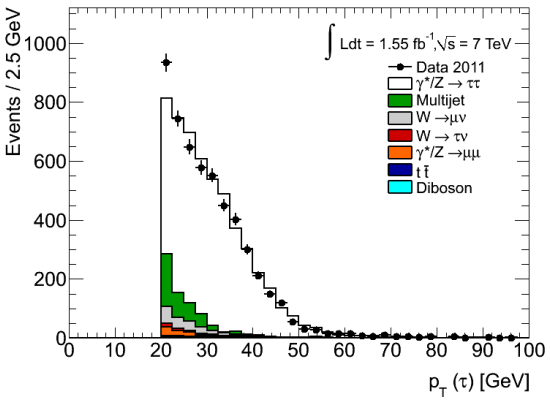
Figure 6.18: Distribution of $\Delta\phi$ between the τ and the lepton after final event selection for the $\tau_e\tau_{\mathrm{h}}$ (left) and the $\tau_\mu\tau_{\mathrm{h}}$ (right) with the 2010 data.

6.3.7 Performance Plots data run 2011

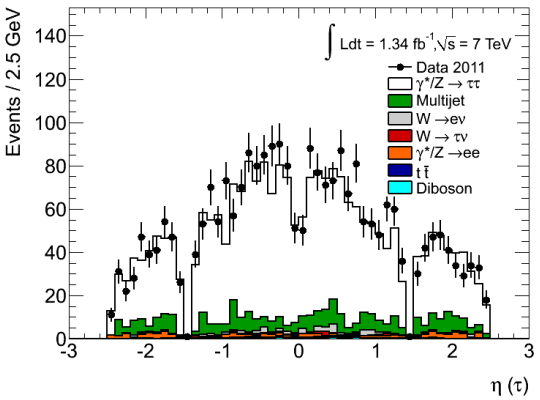
Similarly, Fig. 6.19 - 6.22 show the following distributions for the $\tau_e\tau_h$ channel and the $\tau_\mu\tau_h$ channel for the 2011 data after all cuts and the visible mass window have been applied. The agreement of Monte Carlo predictions with data is worse in the electron channel (compared to the muon one) due to the absence of the exact trigger in the simulations.



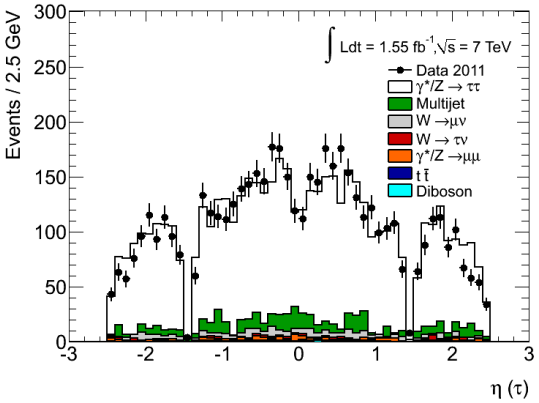
(a) electron channel



(b) muon channel



(c) electron channel



(d) muon channel

Figure 6.19: Distribution of p_T and η of the selected τ candidate after the final event selection for $\tau_e\tau_h$ (left) and $\tau_\mu\tau_h$ (right) with the 2011 data.

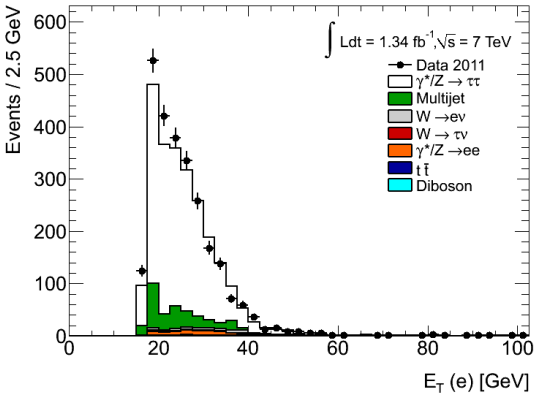
The performance of the BDT τ identification variables is checked at the end of the event selection in the 2011 data. The following variables are plotted showing a good agreement between data and Monte Carlo for the muon channel

- in Fig. 6.23:
  - the calorimetric radius, R_{Cal};
  - the ratio of Transition Radiation Tracker (TRT) hits of the leading track with high threshold to those with low threshold, f_{HT};
  - the invariant mass of the constituent clusters of the seed jet, $m_{cluster}$;

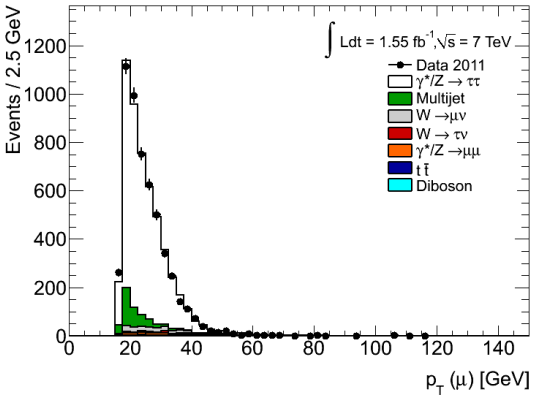
  - the impact parameter significance of the leading track of the τ candidate , $S_{\text{leadtrack}}$; and
  - the finally calculated BDT score before applying the τ identification BDT medium and after;

- in Fig. 6.24:
  - the core energy fraction, f_{core};
  - the invariant mass of the track system, m_{tracks};
  - the track radius, R_{track}; and
  - the transverse flight path significance, $S_{\text{T}}^{\text{flight}}$.

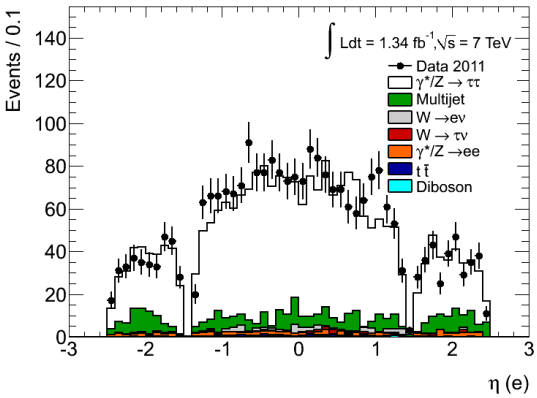
These variables are explained in Section 5.4.2 and in more detail in [77]. In the $\tau_e\tau_{\text{h}}$ channel, due to the use of a hadronic τ trigger a bias on these distributions is expected.



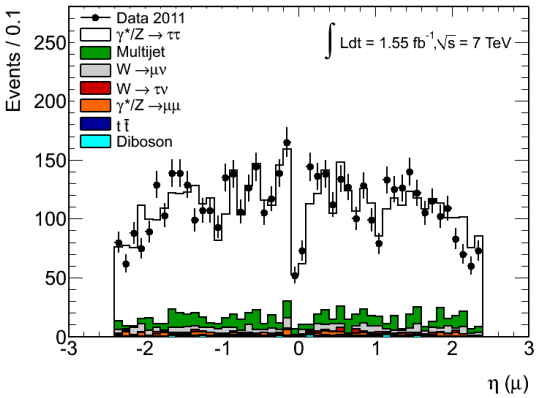
(a) electron channel



(b) muon channel

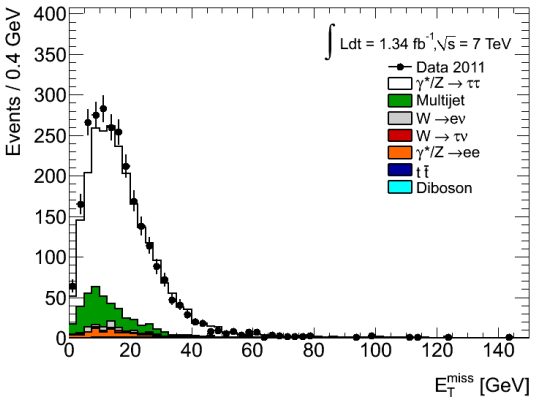


(c) electron channel

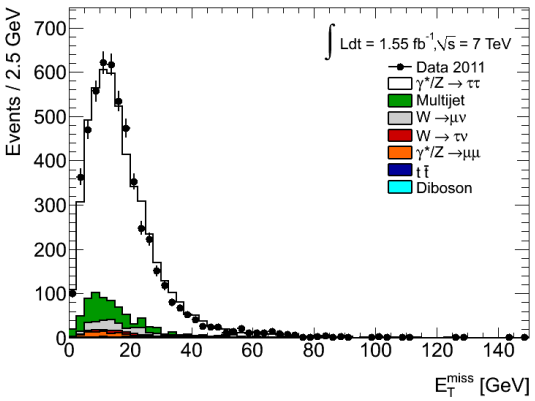


(d) muon channel

Figure 6.20: Distribution of E_T or p_{T} and η of the selected electron or muon after final event selection for $\tau_e\tau_{\mathrm{h}}$ (left) and $\tau_\mu\tau_{\mathrm{h}}$ (right) with the 2011 data.

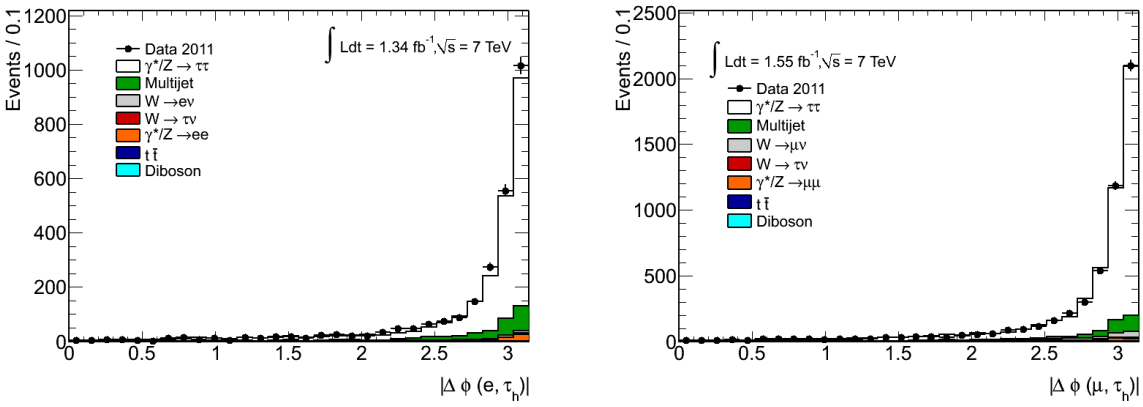


(a) electron channel



(b) muon channel

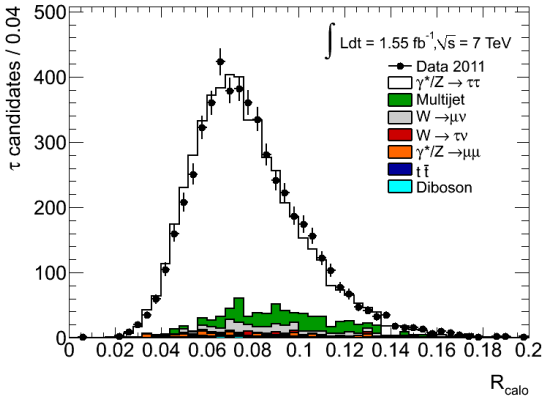
Figure 6.21: Distribution of the missing transverse energy after final event selection for $\tau_e\tau_{\mathrm{h}}$ (left) and $\tau_\mu\tau_{\mathrm{h}}$ (right) with the 2011 data.



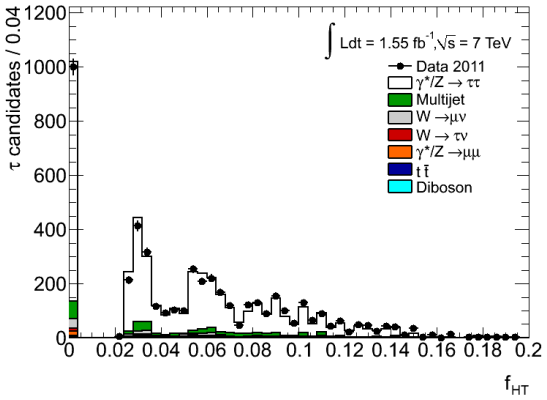
(a) electron channel

(b) muon channel

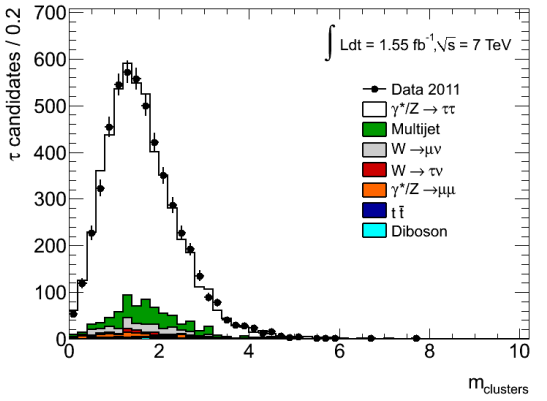
Figure 6.22: Distribution of $\Delta\phi$ between the τ and the lepton after final event selection for $\tau_e\tau_{\mathrm{h}}$ (left) and $\tau_\mu\tau_{\mathrm{h}}$ (right) with the 2011 data.



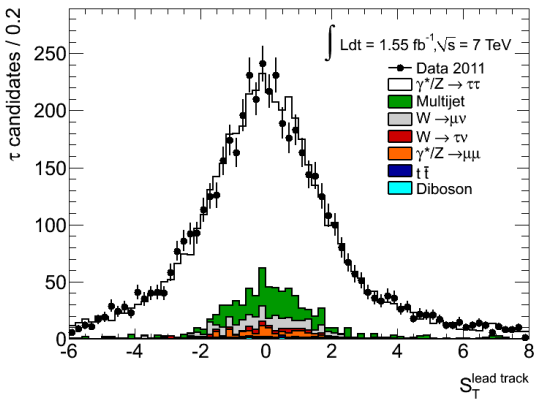
(a) The calorimetric radius.



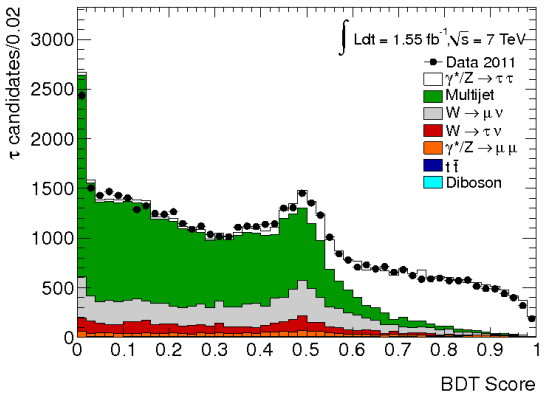
(b) The TRT HT fraction.



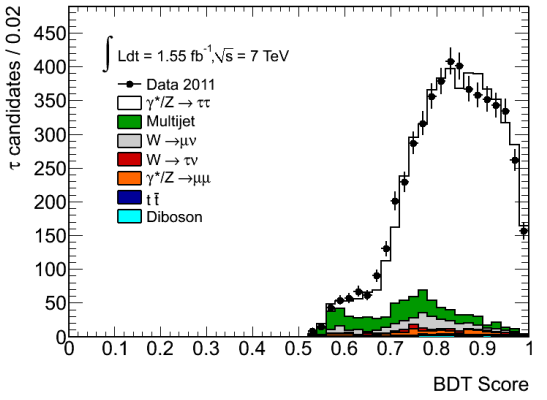
(c) The cluster mass.



(d) The leading track impact parameter significance.

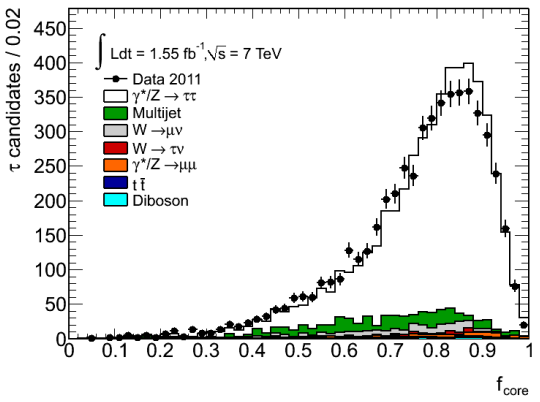


(e) The BDT score before applying the BDT medium identification.

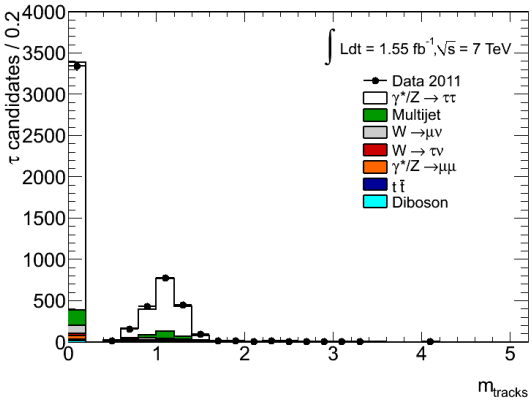


(f) The BDT score after applying the BDT medium identification.

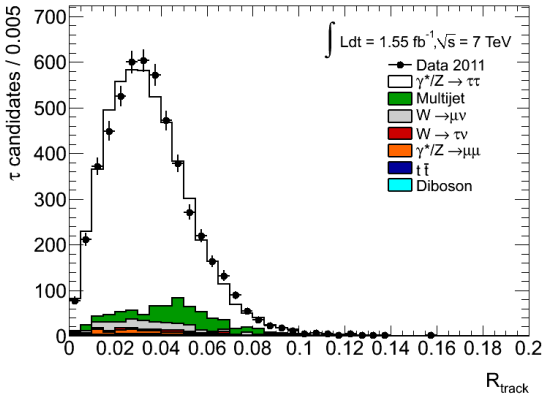
Figure 6.23: Distributions of the τ identification for events passing the full event selection of the $\tau_\mu\tau_\mathrm{h}$ channel.



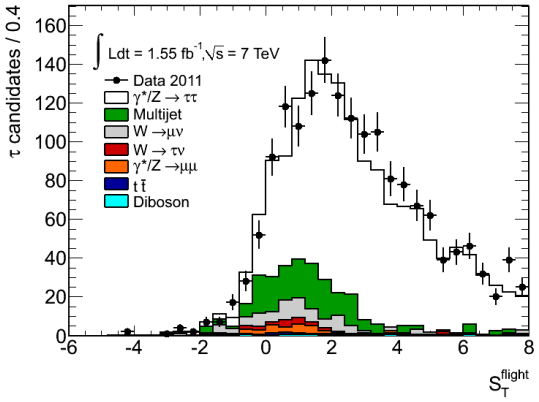
(a) The core energy fraction.



(b) The track mass.



(c) The track radius.



(d) The transverse flight path significance.

Figure 6.24: Further distributions of the τ identification for events passing the full event selection of the $\tau_\mu\tau_\mathrm{h}$ channel.

6.4 Background Estimation

The main background processes for the semi-leptonic $Z \to \tau\tau$ decay are QCD multijets, Z+jets, W+jets, $t\bar{t}$ and diboson (WW, WZ, ZZ) decays. The $Z \to \tau\tau \to \ell\,\ell$+$4\nu$ decays are a negligible background thanks to the dilepton veto and the τ candidate related requirements, Section 6.3.5. For comparing the data with the expectation values, all the electroweak background processes are simulated with Monte Carlo generators and are normalised to the measured luminosity and the theoretical NNLO cross section, given in tables in the Appendix A. In addition, Z+jets, W+jets Monte Carlo are normalised to collision data with a scale factor to account for hadronic τ misidentification. The multijet background is derived by a fully data-driven method. In the following sections, the methods used to estimate W+jets (Section 6.4.1), Z+jets (Section 6.4.2) and QCD multijets (Section 6.4.3) backgrounds are explained.

6.4.1 W+jets Background

The W+jets background can fake $Z \to \tau\tau \to \ell\,\tau_{\rm h} + 4\nu$ decays in two ways; $W \to \ell\nu$ were the lepton is real and a jet from the event is misidentified for a τ candidate, or $W \to \tau\nu$ were the τ candidate is real and a soft QCD jet is wrongly identified as a lepton. The shape of both processes is described satisfactorily by Monte Carlo simulations, but the number of events after applying the τ identification cut is overestimated. Therefore a scaling factor, k_W, is applied.

The scale factor for the W backgrounds is estimated in a W-rich control region. That is ensured by inverting the relative cuts, i.e. $m_T < 50$ GeV and $\sum \cos\delta\phi > -0.15$, see Section 6.3.5. The τ candidate $p_{\rm T}$ distribution in the W control region is shown in Fig. 6.25. In the figure one sees that Monte Carlo overestimates the data, when there is a τ identification requirement. It is, also, shown that the control region is mainly dominated by W+jets and the electroweak contributions from other processes is small. The multijets background is completely suppressed here.

Having obtained a rather clean control region the k_W factor is estimated from the following formula

$$N_W^{\rm WCR} \to k_W N_W^{\rm WCR} = N_{\rm data}^{\rm WCR} - N_{Z\to\ell\ell}^{\rm WCR} - N_{t\bar{t},diboson}^{\rm WCR} \tag{6.4}$$

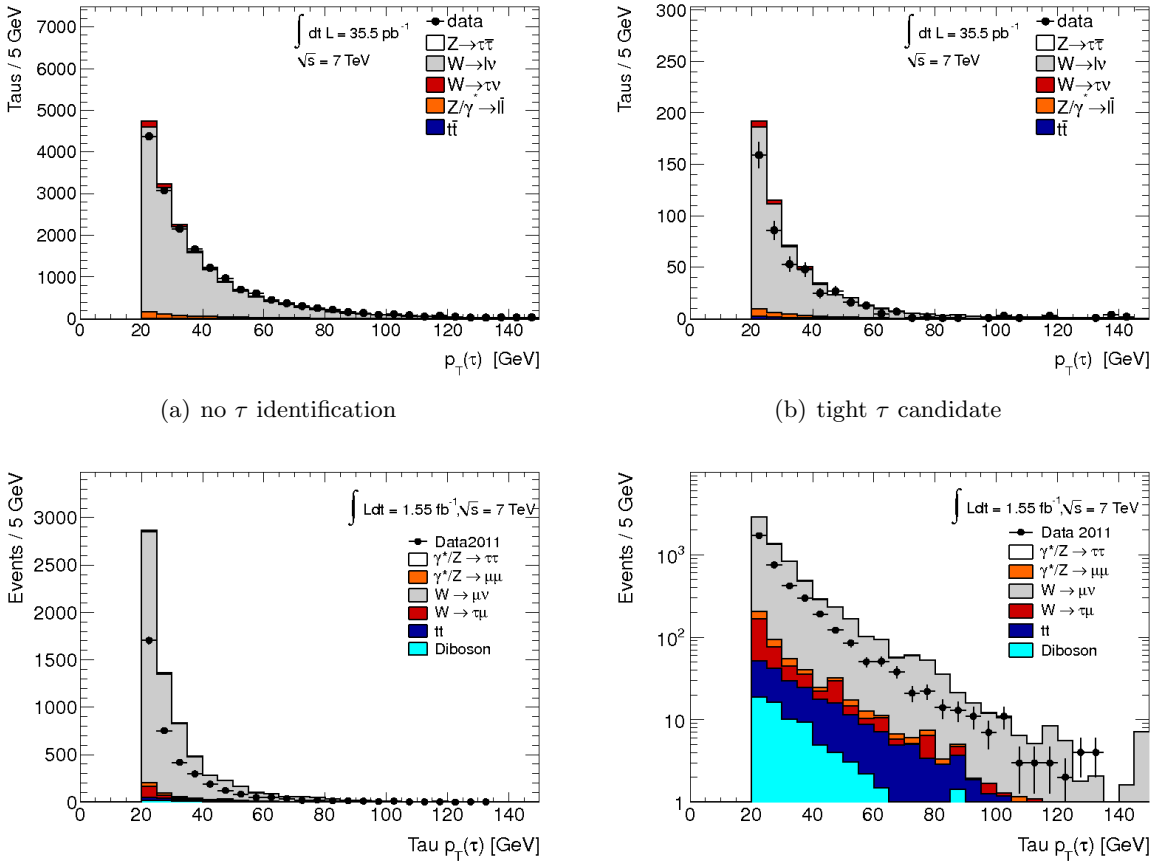
Two k_W factors are calculated; one for opposite sign events and one for same sign. For the 2010 dataset the k_W factors are:

$$k_W = \begin{cases} 0.63 \pm 0.07\ (\text{stat.}) & \tau_{\rm e}\tau_{\rm h}\ \text{channel, opposite sign}\,, \\ 0.83 \pm 0.15\ (\text{stat.}) & \tau_{\rm e}\tau_{\rm h}\ \text{channel, same sign}\,, \\ 0.73 \pm 0.06\ (\text{stat.}) & \tau_{\mu}\tau_{\rm h}\ \text{channel, opposite sign}\,, \\ 0.94 \pm 0.13\ (\text{stat.}) & \tau_{\mu}\tau_{\rm h}\ \text{channel, same sign}\,. \end{cases} \tag{6.5}$$

Similarly, for the 2011 dataset the derived factors are:

$$k_W = \begin{cases} 0.44 \pm 0.02\ (\text{stat.}) & \tau_{\rm e}\tau_{\rm h}\ \text{channel, opposite sign}\,, \\ 0.56 \pm 0.04\ (\text{stat.}) & \tau_{\rm e}\tau_{\rm h}\ \text{channel, same sign}\,, \\ 0.54 \pm 0.01\ (\text{stat.}) & \tau_{\mu}\tau_{\rm h}\ \text{channel, opposite sign}\,, \\ 0.74 \pm 0.03\ (\text{stat.}) & \tau_{\mu}\tau_{\rm h}\ \text{channel, same sign}\,. \end{cases} \tag{6.6}$$

The necessity for the normalisation factor comes from the fact that the Monte Carlo generators cannot predict accurately the fraction of the misidentification rate of τ candidates from jets depending on whether the jet originates from quarks or from gluons. This rate is different



(a) no τ identification

(b) tight τ candidate

(c) no τ identification

(d) tight τ candidate

Figure 6.25: Distribution of the p_{T} of the τ candidate in the W control region, with (left) and without (right) the τ identification criterion applied for the 2010 data (top) and 2011 data (bottom).

because the gluons are blind to the electromagnetic charge. Therefore, the W+g process shows no charge preference. In contrast, the quarks have charge and favour the production of positively charged quark jets in a pp collision. The different rate between quark and gluon jets is reflected by the smaller k_W factor in the case of opposite sign events. In the 2011 analysis an additional difference is observed between electron and muon channel, eq. (6.6). This difference is attributed to the bias from the τ-trigger in the $\tau_e\tau_{\mathrm{h}}$ channel.

6.4.2 Z+jets Background

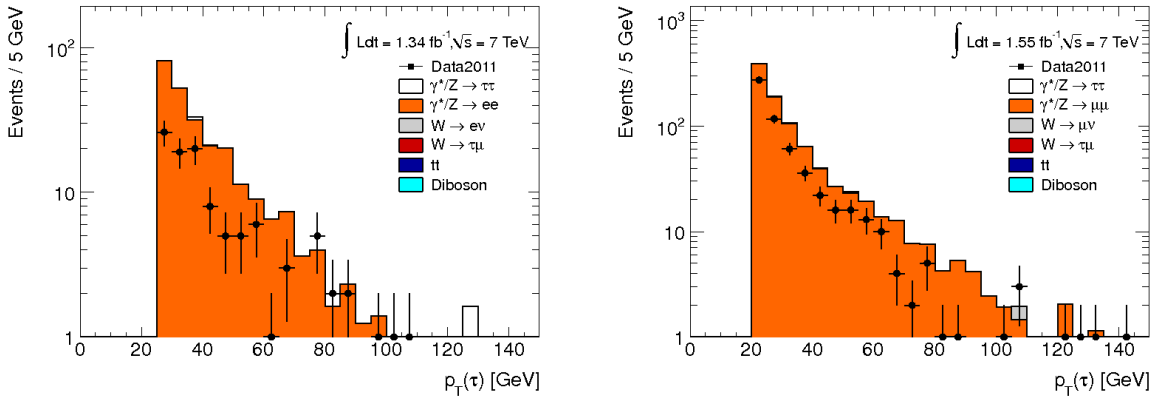
After suppressing the W+jets background, the Z+jets dominates. A $Z \to \ell\ell$ decay can be wrongly reconstructed as a $Z \to \tau\tau$ for two reasons; either one of the leptons is misidentified for a τ candidate, or a jet is misidentified for a τ candidate. The first case affects mainly the $\tau_e\tau_{\mathrm{h}}$ channel, since it is more probable for an electron to fake a τ candidate than for a muon. This background is well described by Monte Carlo, so no correction factor is needed. The second process is similar to the W+jets background. In 2010 analysis a detailed study is conducted [79] and the event weights, shown in Table 6.9, are applied to the analysis.

Table 6.9: Scale factors for tau fake rate obtained in Z+jets events in 2010 data run. The given uncertainties incorporate statistics and systematics.

number of vertices	1-prong medium tau	3-prong tight tau	>= 2-prong tight tau
$1, 2$	0.949 ± 0.220	0.855 ± 0.280	0.929 ± 0.160
> 2	0.626 ± 0.240	1.151 ± 0.436	0.741 ± 0.181

In the 2011 analysis, a Z-rich control region is defined by inverting the dilepton veto, i.e. ask for two leptons, and by requiring the invariant mass of the two leptons to lie within the Z mass peak ($66 < m_{ll} < 116$ GeV). From Fig. 6.26, one can see that the control region contains only $Z \to \ell\ell$ decays. The scale factor calculated from this region is

$$k_Z = \begin{cases} 0.39 \pm 0.05 \text{ (stat.)} & \tau_e\tau_h \text{ channel}, \\ 0.57 \pm 0.04 \text{ (stat.)} & \tau_\mu\tau_h \text{ channel}. \end{cases} \tag{6.7}$$



(a) electron channel

(b) muon channel

Figure 6.26: Transverse momentum distributions of τ candidates in the Z-enriched control region. Monte Carlo simulation overestimates data after the identification of τ candidates for $\tau_e\tau_h$ (left) and $\tau_\mu\tau_h$ (right) channels.

6.4.3 QCD multijets Background

The QCD jets are very likely to fake either the τ candidate or the lepton from the $Z \to \tau\tau \to \ell\,\tau_h + 4\nu$ decay, even after applying the isolation cuts on the leptons. For that reason, a method describing this background at every stage of the analysis is sought. The use of Monte Carlo simulations fails due to the lack of sufficient statistics. The next option is to create control regions in data that are dominated by jets and use those for estimating the amount and shape of the multijet events.

In fact, four such regions are defined:

A. a region with one isolated lepton and opposite sign events;

B. a region with one isolated lepton and same sign events;

C. a region with one non-isolated lepton and opposite sign events; and

D. a region with one non-isolated lepton and same sign events.

The first one in the list (A), where the exact analysis cuts described in Section 6.3.5 are applied, is the signal region. The last one (D), where all analysis cuts are inverted, is dominated by QCD jets. The two intermediate (B,C) contain mainly QCD events but the contamination with signal and EW backgrounds is high. The main assumption of the method is that, for a given cut-stage of the analysis, the ratio between opposite sign (OS) and same sign (SS) multijet events is the same for isolated and non-isolated leptons, i.e.

$$R_{\mathrm{OS/SS}} = \frac{N^{A}_{\mathrm{multijet}}}{N^{B}_{\mathrm{multijet}}} = \frac{N^{C}_{\mathrm{multijet}}}{N^{D}_{\mathrm{multijet}}}, \tag{6.8}$$

where $N^{i}_{\mathrm{multijet}}$ is the number of events in each region and it is calculated as

$$N^{i}_{\mathrm{multijet}} = N^{i}_{Data} - N^{i}_{Z\to\tau\tau} - N^{i}_{Z\to\ell\ell} - N^{i}_{t\bar{t},diboson} - N^{i}_{W\to\ell\nu} - N^{i}_{W\to\tau\nu}, \quad \text{for } i = \mathrm{B,C,D}. \tag{6.9}$$

The signal and electroweak backgrounds are subtracted from the data events from Monte Carlo predictions.

In Tables 6.10 and 6.11, the number of events per control region are given for 2010 and 2011 data, respectively. The resulting OS/SS ratio for 2010 analysis is

$$R_{\mathrm{OS/SS}} = \begin{cases} 1.07 \pm 0.04 \text{ (stat.)} & \tau_{\mathrm{e}}\tau_{\mathrm{h}} \text{ channel,} \\ 1.07 \pm 0.07 \text{ (stat.)} & \tau_{\mu}\tau_{\mathrm{h}} \text{ channel.} \end{cases} \tag{6.10}$$

and for 2011 analysis is

$$R_{\mathrm{OS/SS}} = \begin{cases} 1.06 \pm 0.03 \text{ (stat.)} & \tau_{\mathrm{e}}\tau_{\mathrm{h}} \text{ channel,} \\ 1.13 \pm 0.04 \text{ (stat.)} & \tau_{\mu}\tau_{\mathrm{h}} \text{ channel.} \end{cases} \tag{6.11}$$

Any systematic uncertainties entering the calculation of the ratios are evaluated in Section 7.7.

Table 6.10: Number of events in the ABCD control regions used for the multijet background estimation in the 2010 data. Signal and other background processes are estimated by Monte Carlo, while the number of multijet events is estimated from data after subtracting the EW processes (eq. (6.9)).

		muon channel		electron channel	
		isolated lepton	non-isolated lepton	isolated lepton	non-isolated lepton
		region A	region B	region A	region B
	data	213(15)	1521(39)	151(12)	398(20)
	$Z \to \tau\tau$	185(2)	8.4(4)	97(1)	3.2(2)
	$\gamma \to \tau\tau$	0.7(3)	0.05(5)	0.3(2)	0(0)
OS Events	QCD	24(6)	1511(39)	23(6)	394(20)
+ final selection	$W \to \ell\nu$	9.3(7)	0.3(1)	4.8(4)	0.2(1)
	$W \to \tau\nu$	3.6(8)	0.08(8)	1.5(4)	0.04(4)
	$Z \to \ell\ell$	8.7(3)	0.33(6)	4.9(2)	0.12(3)
	$\gamma \to \ell\ell$	2.4(4)	0.16(8)	2.0(3)	0.03(3)
	$t\bar{t}$	1.3(1)	0.99(8)	1.02(8)	0.11(3)
	Di-Boson	0.28(2)	0.052(8)	0.18(1)	0.009(3)
		region C	region D	region C	region D
	data	34(6)	1415(38)	29(5)	367(19)
	$Z \to \tau\tau$	1.3(2)	0.3(8)	1.0(1)	0.23(7)
	$\gamma \to \tau\tau$	0.06(6)	0.09(9)	0.2(1)	0(0)
SS Events	QCD	22(6)	1413(38)	21(5)	367(19)
+ final selection	$W \to \ell\nu$	3.7(5)	0.09(6)	2.3(3)	0(0)
	$W \to \tau\nu$	2.1(7)	0.2(2)	0.3(3)	0(0)
	$Z \to \ell\ell$	1.9(1)	0.11(3)	2.7(3)	0.05(2)
	$\gamma \to \ell\ell$	2.5(4)	0.11(8)	1.3(3)	0.13(11)
	$t\bar{t}$	0.21(4)	0.61(6)	0.1(3)	0.06(18)
	Di-Boson	0.044(7)	0.021(4)	0.029(5)	0.005(3)

Table 6.11: Number of events in the ABCD control regions used for the multijet background estimation in the 2011 data. Signal and other background processes are estimated by Monte Carlo, while the number of multijet events is estimated from data after subtracting the EW processes (eq. (6.9)).

		muon channel		electron channel	
		isolated lepton	non-isolated lepton	isolated lepton	non-isolated lepton
		region A	region B	region A	region B
	data	5184	577	2600	353
	$\gamma/Z \to \tau\tau$	4544	44	2029	19
OS Events	$\gamma/Z \to \ell\ell$	81	57	64	29
+ final selection	$W \to \ell\nu$	186	64	45	15
	$W \to \tau\nu$	49	22	18	5
	$t\bar{t}$	31	5	17	2
	Dibosons	15	1	6	1
	multijet	-	384	-	282
		region C	region D	region C	region D
	data	1728	1352	2626	2403
	$\gamma/Z \to \tau\tau$	196	1	71	4
SS Events	$\gamma/Z \to \ell\ell$	2	4	3	-
+ final selection	$W \to \ell\nu$	12	1	2	-
	$W \to \tau\nu$	2	-	1	-
	$t\bar{t}$	2	-	3	1
	Dibosons	-	-	-	-
	multijet	1514	1345	2546	2397

Chapter 7

Evaluation of Systematic Uncertainties

Except for the statistical uncertainties, several systematic ones enter in estimation of the final number of predicted events for the $Z \to \tau\tau \to \ell\,\tau_{\rm h} + 4\nu$ process. These uncertainties may contribute indirectly through the estimated scale factors, event weights and efficiencies, or directly through the background estimation methods. The strategy followed to estimate them is similar in the two datasets, however whenever that is not the case the differences are highlighted and explained.

7.1 Pile-up

The pile-up conditions are different in the two datasets. For the 2010 data run, event weights per primary vertex have been calculated, see Section 6.2. The systematic uncertainty is estimated by varying the weights by one standard deviation, $\pm\,1\sigma$, and it is on average $\pm 0.35\%$ for the different processes ($Z \to \tau\tau$, $Z \to \ell\ell$, $t\bar{t}$).

For the 2011 data, the pile-up is both in-time and out-of-time. The re-weighting in this case is done using the average number of pile-up interactions per event, see Section 6.2. It is assumed that no systematic uncertainty enters the re weighting.

7.2 Muon Trigger, Reconstruction and Isolation Efficiency

The systematic uncertainties of the muon scale factors are evaluated by varying the parameters of the tag-and-probe methods used for their estimation. There are three scale factors applied on muons: trigger, reconstruction and isolation.

The uncertainties on the reconstruction scale factors are studied by the muon combined performance group and they are propagated through the analysis. The uncertainty on the resolution of the muons is considered by varying $\pm\,1\sigma$ the resolutions of the Inner Detector and the Muon Spectrometer separately. The total effect on the final selection is $<1\%$ (2010 data) and 0.2% (2011 data).

The trigger scale factors are calculated with respect to reconstructed and isolated muons and they are given in regions of the (η, ϕ) plane, since they are constant in $p_{\rm T}$. The systematic uncertainty is estimated to be 3% (2010 data) and 1% (2011 data).

The systematic uncertainty on the measurement of the isolation efficiency is estimated by varying the isolation cut on the selected probe. Scale factors are calculated with a loose isolation

Table 7.1: Correction factors for the case that a jet is faking a τ candidate estimated in Z+jets events in 2010 data analysis [79].

$\lvert\eta\rvert < 1.37$	$1.37 < \lvert\eta\rvert < 1.52$	$\lvert\eta\rvert < 1.52$
1.2 (0.4)	0.2 (0.1)	1.2 (0.5)

cut and with no isolation cuts. It is found that the scale factors are constant and the systematic uncertainty is 1% (2010 data) and 1.3% (2011 data).

7.3 Electron Trigger, Reconstruction and Isolation Efficiency

Similarly, the electron systematic uncertainty on the trigger, reconstruction and isolation scale factors are estimated by varying the parameters in the tag-and-probe methods.

For the reconstruction and identification uncertainties, the measurement is provided by the electron combined performance group and the result is propagated through the analysis. The systematic uncertainty is found to be $\pm(1-17)\%$ in 2010 data. The uncertainty on the trigger scale factor is 2% in 2010 data and the isolation one $\sim 3\%$. For the 2011 data, the electron scale factors are considered to be correlated. The systematic uncertainty is evaluated as the product of the trigger, reconstruction, identification and isolation scale factors, see Tables 8.6 (2010) and 8.7 (2011).

7.4 Tau Trigger, Identification and Electron-Misidentification Efficiency

The τ trigger efficiency is only relevant in the 2011 measurement when the combined $\tau + e$ trigger is required. It is measured in $Z \to \tau_e\tau_{\mathrm{h}}$ events with the tag-and-probe method. The efficiency is only measured in data because the specific trigger is not present in Monte Carlo. Also, because the trigger decisions themselves are different between periods for this particular trigger, only 1 fb^{-1} of data (periods F to H) is used for the estimation of the trigger efficiency. In the tag-and-probe method, the tag electron is an electron passing the `EF_e15_medium` trigger and the probe is the hadronic τ candidate [56]. The dominant uncertainty arises from the QCD multijet contamination, which is estimated by the ratio of the uncertainty on same-sign events to opposite-sign events and is taken to be $\sim 5\%$.

The systematic uncertainty on the identification efficiency for cut-based medium (tight) τ candidates is $\pm 9\%(12\%)$ (2010 data) and for BDT-based medium τ candidates is 5.1% (2011 data).

In 2010 data, two types objects are considered to be faking the τ candidates; electrons and jets. The correction factors for jets misidentification is given in Table 7.2 and the uncertainty on them is found to be $\pm(17-38)\%$ depending on the number of vertices and the number of associated tracks to the τ candidate. The electron fake rate is given in Table 7.1 and the systematic uncertainty on the final number of events is found to be $\pm 50\%$. The uncertainties are so large due to the small number of Z+jets events used. In the 2011 analysis, where the statistics are not an issue, the estimated uncertainty is of the order of a few per-mille.

Table 7.2: Correction factors for the case that an electron is faking a τ candidate in Monte Carlo simulations for 2010 data analysis [79].

number of vertices	1-prong "medium" τ candidate	multi-prong "tight" τ candidate
1,2	0.949 ± 0.220	0.929 ± 0.160
> 2	0.626 ± 0.240	0.741 ± 0.181

Table 7.3: Factors used to scale the τ candidates energy both up and down in Monte Carlo for 2010 analysis [78].

Relative uncertainty	1-prong energy scale systematics			multi-prong energy scale systematics		
Detector Region	$\lvert\eta\rvert < 1.3$	$1.3 < \lvert\eta\rvert < 1.6$	$\lvert\eta\rvert > 1.6$	$\lvert\eta\rvert < 1.3$	$1.3 < \lvert\eta\rvert < 1.6$	$\lvert\eta\rvert > 1.6$
$15 < p_{\mathrm{T}} < 20$ GeV	4.5%	7.0%	4.5%	14.0%	7.0%	6.5%
$20 < p_{\mathrm{T}} < 30$ GeV	3.5%	7.0%	4.5%	10.0%	7.0%	6.5%
$p_{\mathrm{T}} > 30$ GeV	2.5%	7.0%	4.5%	8.0%	7.0%	6.5%

7.5 Energy Scale

The dominant systematic uncertainty is coming from the estimation of the energy scale of the τ candidate, the electron and the $E_{\mathrm{T}}^{\mathrm{miss}}$. These three uncertainties are treated as fully correlated, since it is difficult to disentangle the effects.

In the electron channel, the hadronic τ, the electron and the $E_{\mathrm{T}}^{\mathrm{miss}}$ energy scales are varied simultaneously according to the recommendations of the performance groups. The τ energy scale uncertainty is provided by the τ combined performance group in bins of p_{T} and η separately for one- and multi-prong τ candidates. The systematic uncertainty is shown in Table 7.3 for 2010 and Table 7.4 for 2011 data.

The energy scale uncertainty for the $E_{\mathrm{T}}^{\mathrm{miss}}$ is estimated in a slightly different way for the 2010 and 2011 data. In 2010 analysis, the muon and electron related topoclusters are subtracted from the $E_{\mathrm{T}}^{\mathrm{miss}}$ estimation, since electron and muon energy scale variation is propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation. In 2011 analysis, the opposite approach is preferred. That is, the electron energy scale is not propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation, only the muon one and all the $E_{\mathrm{T}}^{\mathrm{miss}}$ topoclusters are varied at the same time. The shift of the $E_{\mathrm{T}}^{\mathrm{miss}}$ is following a p_{T}-dependant

Table 7.4: Factors used to scale the τ candidates energy both up and down in Monte Carlo for 2011 analysis [77].

Relative uncertainty	1-prong energy scale systematics			multi-prong energy scale systematics		
Detector Region	$\lvert\eta\rvert < 1.3$	$1.3 < \lvert\eta\rvert < 1.6$	$\lvert\eta\rvert > 1.6$	$\lvert\eta\rvert < 1.3$	$1.3 < \lvert\eta\rvert < 1.6$	$\lvert\eta\rvert > 1.6$
$15 < p_{\mathrm{T}} < 20$ GeV	5.5%	5.0%	4.5%	6.5%	9.5%	6.5%
$20 < p_{\mathrm{T}} < 30$ GeV	4.5%	5.0%	4.5%	6.5%	5.5%	5.5%
$p_{\mathrm{T}} > 30$ GeV	3.5%	5.0%	4.5%	–	–	–
$30 < p_{\mathrm{T}} < 40$ GeV	–	–	–	5.5%	5.5%	5.5%
$p_{\mathrm{T}} > 40$ GeV	–	–	–	4.5%	5.0%	5.0%

Table 7.5: The effect of shifting the cluster energies along with shifting the τ energy on the final event yield in the 2010 data. For the electron channel, the electron energy is also scaled up or down depending on the sign of a.

Shift	$Z \to \tau\tau(e)$	$Z \to ee$	W+Jets	$t\bar{t}$
a=0.07, N-1=0.93	9.39	29	21.1	10.6
a=0.06, N-1=1.17	9.21	27	20	11
a=0.05, N-1=1.50	9.16	24	18.6	10.2
a=-0.05, N-1=1.50	-9.1	-4.11	-8.74	-12.5
a=-0.06, N-1=1.17	-9.26	-4.58	-9.07	-12.9
a=-0.07, N-1=0.93	-9.44	-5.7	-10.1	-13
cluster $\|\eta\| > 3.2$				
a=1.1, N-1=0	-0.154	1.88	-2.15	1.63
a=0.9, N-1=0	-0.246	-0.803	-3.36	-0.0744
Shift	$Z \to \tau\tau(\mu)$	$Z \to \mu\mu$	W+Jets	$t\bar{t}$
a=0.07, N-1=0.93	7.54	24.3	18.1	17.8
a=0.06, N-1=1.17	7.5	24.1	17.9	17.6
a=0.05, N-1=1.50	7.4	23.6	18.1	17.9
a=-0.05, N-1=1.50	-8.24	-17.9	-18.5	-13.5
a=-0.06, N-1=1.17	-8.42	-19.2	-18.4	-14
a=-0.07, N-1=0.93	-8.55	-20.1	-18.4	-14.7
cluster $\|\eta\| > 3.2$				
a=1.1, N-1=0	0.0517	-0.04	-0.0895	0.129
a=0.9, N-1=0	-0.0501	0.617	-3.71	1.5

relation [12, 73]

$$1 + a\left(1 + \frac{N-1}{p_{\mathrm{T}}}\right), \tag{7.1}$$

where a is the uncertainty on the cluster energy scale and it is derived by E/p studies on single isolated hadrons [89, 90]. The $(N-1)$ corrects for the change in the clustering efficiency and scale in a non-isolated environment, such as noise threshold effects from near-by clusters. It is estimated by comparing the E_{cell}/p measurement to the total E/p. For this analysis, the a and $(N-1)$ are varied independently in two η ranges, $|\eta| < 3.2$ and $|\eta| \geq 3.2$. The effect on the $E_{\mathrm{T}}^{\mathrm{miss}}$ energy scale is shown in Table 7.5.

The maximal uncertainty arises for $|\eta| < 3.2$ and $a = -0.07$, $N-1 = 0.93$ GeV and for $|\eta| > 3.2$ and $a = 0.09$, $N-1 = 0$ GeV. These values for $|\eta| < 3.2$ and $a = \pm 0.07$, $N-1 = 0.93$ GeV are used for the evaluation of the systematic uncertainty on the 2011 analysis.

For the muon channel, apart from the τ candidate and the $E_{\mathrm{T}}^{\mathrm{miss}}$ energy scale uncertainty, the muon one is varied as well, and then, propagated into the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation, as explained above. The final effect of the energy scale shift in the measurement is given in Tables 7.5 for the 2010 and 8.7 for the 2011 measurement.

7.6 W and Z Background Estimation

The W+jets and Z+jets backgrounds are estimated from Monte Carlo and normalised to data using normalisation factors, as described in Sections 6.4.1 and 6.4.2. The statistical error on the

calculation of the k_W and k_Z factors is assigned as a systematic uncertainty on the background estimation. All other systematics affecting Monte Carlo data that are described in the previous paragraphs are evaluated and found to be smaller than the statistical uncertainty on these factors.

7.7 QCD multijets background Estimation

The QCD multijets background is estimated with the data-driven method described in Section 6.4.3. Apart from the statistical uncertainty in the calculation of the OS/SS ratio, the stability of the ratio against the assumptions made is checked. That is, the assumption that the $R_{\mathrm{OS/SS}}$ ratio is independent of the isolation cut and the subtraction of Monte Carlo predicted events in the control regions. All effects are added in quadrature to obtain the final systematic uncertainty on this background.

Specifically, in the 2010 data analysis, the stability of the $R_{\mathrm{OS/SS}}$ against the isolation requirement is checked by inverting the τ candidate identification criterion. In the new control regions

- the τ candidate is required to be *loose* but *not tight* and the lepton-τ_{had} pair to have same charge (SS) and
- the τ candidate is required to be *loose* but *not tight* and the lepton-τ_{had} pair to have opposite charge (OS).

The $R_{\mathrm{OS/SS}}$ is plotted against the calorimeter isolation variable $I^{0.4}_{E_{\mathrm{T}}}/E_{\mathrm{T}}$ for muons and $I^{0.3}_{E_{\mathrm{T}}}/E_{\mathrm{T}}$ for electrons in Fig. 7.1(a,b). The ratio is constant against the isolation variables within statistics, except for the first bin in which a large contamination of electroweak background is observed. A better modelling of this region is not possible due to the low Monte Carlo statistics. The stability of the $R_{\mathrm{OS/SS}}$ ratio is also checked against the τ identification criteria; *loose*, *medium* and *tight*. This is plotted in Fig. 7.1(c,d) and it is found to be independent within statistics.

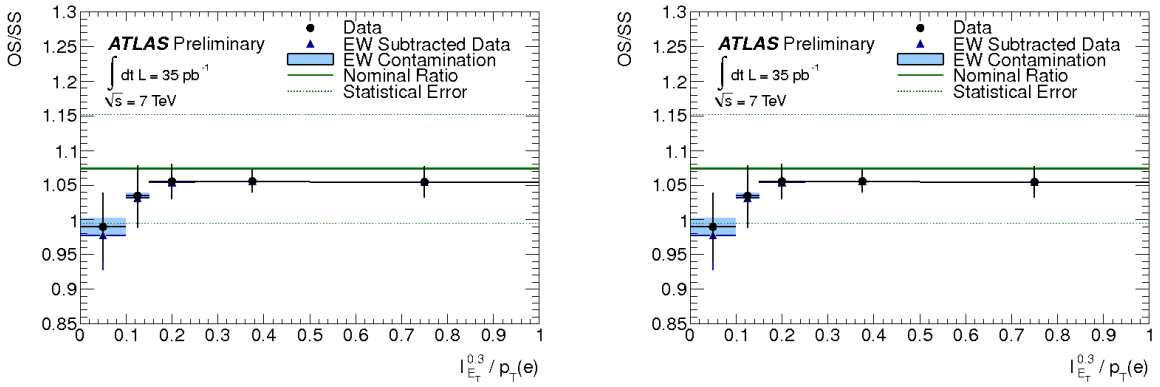
For the 2011 data, the $R_{\mathrm{OS/SS}}$ stability is checked against both the calorimeter and the tracking isolation for the muon channel. The same control regions as in the 2010 analysis are defined. The new $R_{\mathrm{OS/SS}}$ ratio is shown in Fig. 7.2(a,b) for calorimeter and track isolation bins, respectively. The maximal deviation from the $R_{\mathrm{OS/SS}}$ value is 3%.

In the electron channel, due to the use of the τ trigger, the identification criteria cannot be inverted in the offline selection. Therefore, the stability of the $R_{\mathrm{OS/SS}}$ ratio is tested against relaxed/tighter isolation cuts. The effect is always smaller than the statistical error. Nevertheless, the same systematic uncertainty of 3% that is estimated in the muon channel is assigned here, too. The stability of the ratio is also checked throughout the cutflow and a deviation of 4% is seen only in the $\tau_e\tau_{\mathrm{h}}$ channel, see Fig. 7.2(c).

The systematic uncertainty due to Monte Carlo background subtraction in the signal and control regions is estimated by varying the cross section up and down by one standard deviation. The result for 2010 data is $\pm 14.8\%$ for the $\tau_\mu\tau_{\mathrm{h}}$ and $\pm 12.5\%$ for the $\tau_e\tau_{\mathrm{h}}$, while for 2011 data there is 1% effect in the $\tau_\mu\tau_{\mathrm{h}}$ and almost none in the $\tau_e\tau_{\mathrm{h}}$.

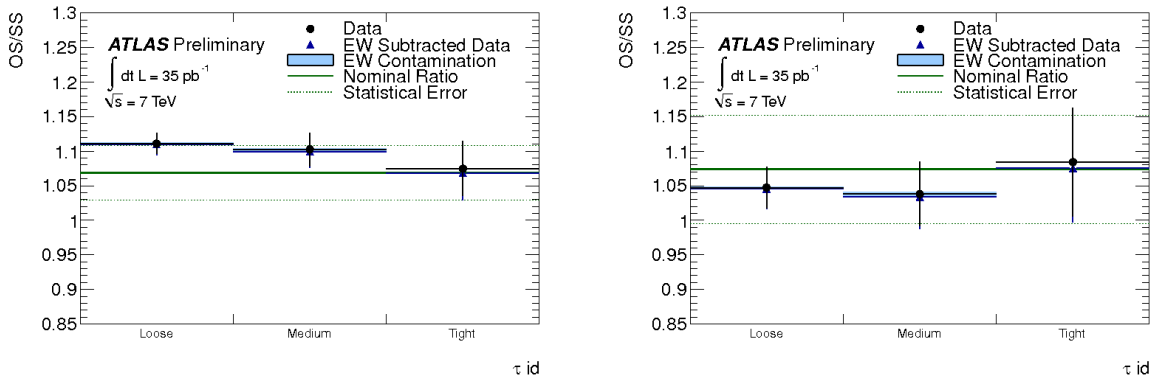
7.8 Summary of Systematic Uncertainties

All the systematic uncertainties on the background prediction are summarised in Table 7.6 for the 2010 measurement and in Table 8.7 for the 2011 measurement. For the W+jets background,



(a) muon channel

(b) electron channel



(c) muon channel

(d) electron channel

Figure 7.1: Stability of the OS vs SS ratio as a function of calorimeter isolation (top) and the τ identification (bottom) for muons (left) and electrons (right) in data.

the statistical and systematic uncertainty is included, as described in Section 7.6. The same is true for the k_Z factor used in 2011 analysis (Section 6.4.2). In the case of the QCD multijet background, except for the systematics on the stability of the $R_{\mathrm{OS/SS}}$ ratio the uncertainties relating to Monte Carlo predictions are taken into account. That is done by shifting the Monte Carlo up and down by the relevant amount simultaneously in the signal and the control regions. The final effect on the signal measurement is calculated by adding all the partial effects in quadrature.

Table 7.6: Summary of the effect of all systematic and statistical uncertainties on the final predicted signal ($N_{\text{obs}} - N_{\text{bg}}$) for both the electron and muon channels on the 2010 data.

	muon channel (% Deviation)				electron channel (% Deviation)			
Systematic	N-Bkgrd	multijet	W+Jets	$\gamma^\star/Z + t\bar{t}$	N-Bkgrd	multijet	W+Jets	$\gamma^\star/Z + t\bar{t}$
lepton efficiency	0.24	1.6	3.0	2.8($Z/\gamma^\star$) 5.9($t\bar{t}$)	0.31	2.7	7.0	5.0(Z) 9.0($\gamma^\star/t\bar{t}$)
lepton resolution(μ energy scale)	0.018	0.092	0.2	0.2	0.012	0.065	0.2	0.2
Problematic calorimeter regions	0.0023	0.016	0	0	0.022	0.14	0.4	0.4
e charge misidentification	0.0023	0.016	0	0	0.014	0.064	0.21	0.21
τ id efficiency	0.078	0.54	0	0	0.11	0.55	0	0
Energy scale lepton and τ	1.7	10	18.8	24.3(Z) 18.0($t\bar{t}$)	1.3	8.8	21.4	29.1(Z) 13.1($t\bar{t}$)
lepton-jet τ fake rate	1.1	5	0	24	0.66	3.9	0	21
pile-up reweighting	0.029	0.17	0.35	0.35	0.02	0.12	0.35	0.35
Jet Cleaning	0.067	0.41	0	1.8	0.066	0.51	0	2.5
k_W factor	0.12	3.6	8.20	0	0.18	2.2	11.1	0
OS/SS Ratio	0.76	5.2	0	0	2	10	0	0
Theoret. cross section	0.2	1.4	0	5.0/8.3	0.12	1.3	0	5.0/8.3
Total systematic uncertainty	2.2	13	21	34	2.5	14	25	35
Statistical uncertainty	9.8	27	8.3	3	12	27	9.9	3.8
Total uncertainty	10	30	22	34	12	30	27	35

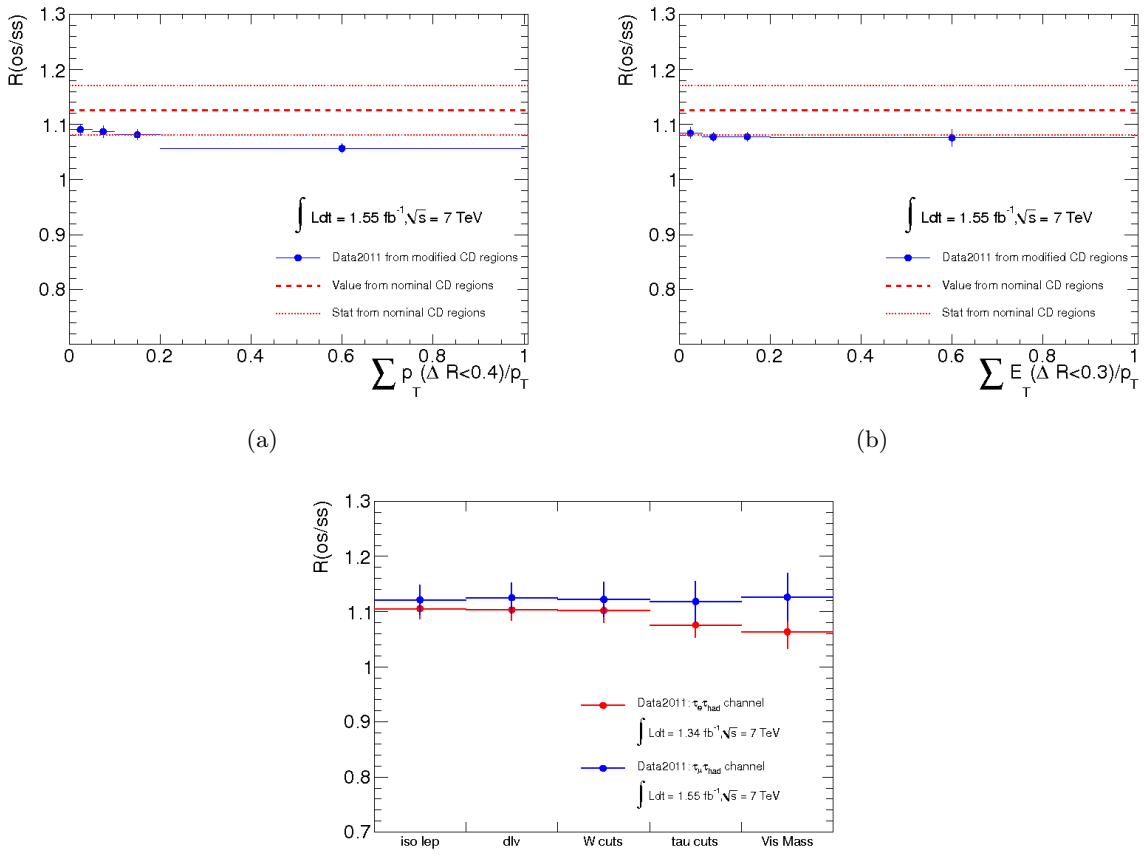


Figure 7.2: Top: In 2011 data, the dependence of $R_{OS/SS}$ as a function of the isolation variables ($I^{0.4}_{p_{\mathrm{T}}}/p_{\mathrm{T}}$ (left) and $I^{0.3}_{E_{\mathrm{T}}}/p_{\mathrm{T}}$ (right)) in the $\tau_\mu\tau_{\mathrm{h}}$ channel, in the multijets-pure region of inverted τ identification requirements. Bottom: In 2011 data, the dependence of $R_{OS/SS}$ as a function of the cutflow for the semi-leptonic channels.

Chapter 8

$\gamma^*/Z \to \tau\tau$ Cross Section Measurement

8.1 Experimental Measurement

The cross section of the $Z \to \tau\tau \to \ell\,\tau_{\rm h} + 4\nu$ decay is estimated using the formula

$$\sigma(Z \to \tau\tau) \times BR(\tau \to l\,\nu\,\nu, \tau \to \tau_{had}\,\nu) = \frac{N_{\rm obs} - N_{\rm bkg}}{A_Z \cdot C_Z \cdot \mathcal{L}} \tag{8.1}$$

where

- $N_{\rm obs}$ is the number of observed events in data
- $N_{\rm bkg}$ is the number of estimated background events
- A_Z denotes the kinematic and geometric acceptance for the signal process. It is calculated by the ratio of the number of events at generator level that fall within the fiducial regions defined below, $N^{\rm gen\,kin}_{\rm dressed}$, over the number of events at generator level whose invariant mass at Born level lies within the mass window $66 < m_Z < 116$ GeV, $N^{\rm gen\,m_{inv}}_{\rm born}$.

$$A_Z = \frac{N^{\rm gen\,kin}_{\rm dressed}}{N^{\rm gen\,m_{inv}}_{\rm born}}. \tag{8.2}$$

  A "dressed" τ decay product is a lepton or a τ candidate that has been associated to photons radiated by either the τ leptons or by the decay products themselves. The photon should be radiated within a cone of $\Delta R < 0.1$ for electrons and muons and $\Delta R < 0.4$ for hadronic decay products. This process allows to perform a partial QED final state radiation correction back to the Born level, although it excludes the radiation at wide angle. The A_Z factor by construction includes a correction for events that migrate from outside the invariant mass window in the fiducial cuts.

  The central values for the A_Z factor were determined using a default PYTHIA Monte Carlo sample generated with the modified LO parton distribution function (PDF) MRSTLO* [91] and the corresponding ATLAS MC10 tune[1] [92]. Although AlpGEN is used in the analysis for signal and background processes, it is avoided in the estimation of the A_Z

[1] Although the combination of a leading order PDF with a NLO plus the matrix element is wrong, we are doing it here for technical reasons.

factor because it underestimates the Z production at low rapidity. The source of the problem is that the leading order PDF (CTEQ6L1) used in AlpGEN cannot describe accurately the distributions from higher order corrections. In contrast a next-to-leading order PDF, like CTEQ6.6, provides a more realistic description of the data. This problem is not expected to affect the reconstruction-level description of event kinematics, but could affect the extrapolation to the total cross section. Therefore, PYTHIA with MRSTLO* PDF is used for the determination of the geometrical acceptance.

The lower bound on the invariant mass of the default sample is 10 GeV and therefore the sample includes a tail of low-mass γ^*/Z events from outside the Z peak that can possibly migrate to the fiducial region. The obtained central values are reported in Tables 8.1 (top) for the 2010 data and (bottom) for the 2011 data. The A_Z value is small compared to the acceptance values at the $Z \to e^+e^-$ and $Z \to \mu^+\mu^-$ processes [12] because the neutrinos, which carry a large fraction of the τ lepton momentum, are not included in the definition of the fiducial regions.

Table 8.1: Central values for the A_Z acceptance factor from PYTHIA ATLAS MC10 Monte Carlo generated with MRSTLO* PDF at generator level and C_Z correction factor determined using the same sample at generator level and after full detector simulation for the 2010 (top) and 2011 (bottom) data analysis.

	Muon channel	Electron channel
A_Z	0.11691 ± 0.00023 (stat.)	0.10073 ± 0.00021 (stat.)
C_Z	0.2045 ± 0.0024 (stat.)	0.1197 ± 0.0017 (stat.)

	Muon channel	Electron channel
A_Z	0.0976 ± 0.0002 (stat.)	0.0687 ± 0.0002 (stat.)
C_Z	0.1417 ± 0.0016 (stat.)	0.1009 ± 0.0013 (stat.)

- C_Z is the correction factor that accounts for the experimental imperfections, such as trigger, reconstruction and identification inefficiencies within the geometrical acceptance. It is defined as

$$C_Z = \frac{N^{\text{reco pass}}}{N^{\text{gen kin}}_{\text{dressed}}} \tag{8.3}$$

  where $N^{\text{reco pass}}$ is the number of signal events that pass the analysis cuts after full simulation corrected with data-driven factors as described in Sections 6.3.2, 6.3.3, 6.3.5. $N^{\text{gen kin}}_{\text{dressed}}$ is the same as for the A_Z numerator. By construction C_Z includes a correction for migrations from outside of the acceptance. C_Z is calculated with AlpGEN using CTEQ6L1.

- $\mathcal{L}$ denotes the integrated luminosity for the channel of interest.

Using eq. (8.1) and the A_Z and C_Z values from Table 8.1 one can calculate the total inclusive cross section for the two channels.

- 2010 data analysis:

$$\sigma(Z \to \tau\tau) \times BR(\tau \to \mu\,\nu\,\nu, \tau \to \tau_{\text{had}}\,\nu) = 192.9 \pm 19.0(\text{stat}) \pm 28.1(\text{syst}) \pm 6.6(\text{lumi})\ \text{pb} \tag{8.4}$$

for the $\tau_\mu\tau_\mathrm{h}$ channel,

$$\sigma(Z \to \tau\tau) \times BR(\tau \to e\,\nu\,\nu, \tau \to \tau_\mathrm{had}\,\nu) = 264.1 \pm 32.1(\mathrm{stat}) \pm 45.7(\mathrm{syst}) \pm 9.0(\mathrm{lumi})\ \mathrm{pb} \tag{8.5}$$

for the $\tau_e\tau_\mathrm{h}$ channel.

- 2011 data analysis:

$$\sigma(Z \to \tau\tau) \times BR(\tau \to \mu\,\nu\,\nu, \tau \to \tau_\mathrm{had}\,\nu) = 205.2 \pm 3.4(\mathrm{stat}) \pm 21.3(\mathrm{syst}) \pm 7.6(\mathrm{lumi})\ \mathrm{pb} \tag{8.6}$$

for the $\tau_\mu\tau_\mathrm{h}$ channel,

$$\sigma(Z \to \tau\tau) \times BR(\tau \to e\,\nu\,\nu, \tau \to \tau_\mathrm{had}\,\nu) = 230.9 \pm 5.5(\mathrm{stat}) \pm 30.5(\mathrm{syst}) \pm 8.5(\mathrm{lumi})\ \mathrm{pb} \tag{8.7}$$

for the $\tau_e\tau_\mathrm{h}$ channel.

If the A_Z is set to unity then the cross section can be calculated for the fiducial regions defined below. The measurement becomes independent of the extrapolation procedure to the full phase space and thus less dependent on the theoretical uncertainties of the model. The fiducial regions are defined by the following cuts (the first number refers to the 2010 selection and in parenthesis the 2011):

$\tau_\mu\tau_\mathrm{h}$ channel:

- Muon: $p_\mathrm{T}>$ 15(17) GeV, $|\eta| < 2.4$
- Tau: $E_\mathrm{T}>$ 20 GeV, $|\eta| < 2.47$, excluding $1.37 < |\eta| < 1.52$
- Event: $\Sigma \cos\Delta\phi > -0.15$, $m_T < 50$ GeV, m_vis within $35 < m_Z < 75$ GeV.

$\tau_e\tau_\mathrm{h}$ channel:

- Electron: $E_\mathrm{T}>$ 16(17) GeV, $|\eta| < 2.47$, excluding $1.37 < |\eta| < 1.52$
- Tau: $E_\mathrm{T}>$ 20(25) GeV, $|\eta| < 2.47$, excluding $1.37 < |\eta| < 1.52$
- Event: $\Sigma \cos\Delta\phi > -0.15$, $m_T < 50$ GeV, m_vis within $35 < m_Z < 75$ GeV.

The fiducial cross section is measured via

$$\sigma^{fid}(Z \to \tau\tau) \times BR(\tau \to l\,\nu\,\nu,\ \tau \to\ \tau_\mathrm{had}\,\nu) = \frac{N_{obs} - N_{bkg}}{C_Z \cdot \mathcal{L}} \tag{8.8}$$

If one plugs in the measured values into eq. (8.8), the final fiducial cross section is

- 2010 data analysis:

$$\sigma^{fid}(Z \to \tau\tau) \times BR(\tau \to \mu\,\nu\,\nu, \tau \to\ \tau_\mathrm{had}\,\nu) = 22.55 \pm 2.22(\mathrm{stat}) \pm 3.21(\mathrm{syst}) \pm 0.77(\mathrm{lumi})\ \mathrm{pb} \tag{8.9}$$

for the $\tau_\mu\tau_\mathrm{h}$ channel and

$$\sigma^{fid}(Z \to \tau\tau) \times BR(\tau \to e\,\nu\,\nu, \tau \to\ \tau_\mathrm{had}\,\nu) = 26.59 \pm 3.23(\mathrm{stat}) \pm 4.54(\mathrm{syst}) \pm 0.91(\mathrm{lumi})\ \mathrm{pb} \tag{8.10}$$

for the $\tau_e\tau_\mathrm{h}$ channel.

- 2011 data analysis:

$$\sigma^{fid}(Z \to \tau\tau) \times BR(\tau \to \mu\,\nu\,\nu,\ \tau \to\ \tau_{\rm had}\,\nu) = 20.0 \pm 0.3(\text{stat}) \pm 2.0(\text{syst}) \pm 0.7(\text{lumi})\ \text{pb} \tag{8.11}$$

for the $\tau_\mu\tau_{\rm h}$ channel and

$$\sigma^{fid}(Z \to \tau\tau) \times BR(\tau \to e\,\nu\,\nu,\ \tau \to\ \tau_{\rm had}\,\nu) = 15.9 \pm 0.4(\text{stat}) \pm 2.0(\text{syst}) \pm 0.6(\text{lumi})\ \text{pb} \tag{8.12}$$

for the $\tau_e\tau_{\rm h}$ channel.

Finally, the inclusive cross section after correcting for the $(\tau \to l\,\nu\,\nu, \tau \to \tau_{\rm had})$ branching ratio, 0.2250 ± 0.0009 for the $\tau_\mu\tau_{\rm h}$ channel and 0.2313 ± 0.0009 for the $\tau_e\tau_{\rm h}$ channel [4], is

- 2010 data analysis:

$$\sigma(Z \to \tau\tau,\ m_{\rm inv} : [66{-}116]\ \text{GeV}) = 857.6{\pm}84.3(\text{stat}){\pm}124.7(\text{syst}){\pm}29.2(\text{lumi}){\pm}2.8(\text{theo})\ \text{pb} \tag{8.13}$$

for the $\tau_\mu\tau_{\rm h}$ channel,

$$\sigma(Z \to \tau\tau,\ m_{\rm inv} : [66{-}116]\ \text{GeV}) = 1142{\pm}138.6(\text{stat}){\pm}197.7(\text{syst}){\pm}38.9(\text{lumi}){\pm}2.6(\text{theo})\ \text{pb} \tag{8.14}$$

for the $\tau_e\tau_{\rm h}$ channel.

- 2011 data analysis:

$$\sigma(Z \to \tau\tau,\ m_{\rm inv} : [66{-}116]\ \text{GeV}) = 912.4{\pm}15.0(\text{stat}){\pm}94.7(\text{syst}){\pm}33.7(\text{lumi})\ \text{pb} \tag{8.15}$$

for the $\tau_\mu\tau_{\rm h}$ channel,

$$\sigma(Z \to \tau\tau,\ m_{\rm inv} :\ [66-116]\ \text{GeV}) = 998.1 \pm 23.7(\text{stat}) \pm 131.9(\text{syst}) \pm 36.9(\text{lumi})\ \text{pb} \tag{8.16}$$

for the $\tau_e\tau_{\rm h}$ channel.

8.2 Systematic Uncertainties

The systematic uncertainties on the cross section measurement are estimated by propagating the effect of the individual uncertainties to the cross section measurement, as described in Section 7, for signal and background events and the uncertainty on the estimation of the A_Z, C_Z factors.

8.2.1 Geometrical Acceptance

The systematic uncertainty on the A_Z factor is mainly coming from the limited knowledge of the proton PDFs and the modelling of the Z boson production at a hadron collider like LHC. Three components are varied to assess the systematic uncertainty:

- Uncertainty within one PDF set is evaluated according to the method explained in Section 4.1. The PDF set is CTEQ6.6 at NLO. The error eigenvector for this PDF set is shifted by 1σ up and down and the corresponding A_Z^i factors are calculated for each entry of the matrix. The uncertainty is derived via

$$\Delta A_Z = \frac{1}{2}\sqrt{\sum (A_Z^{i+} - A_Z^{i-})^2} \tag{8.17}$$

- Deviations between different PDF sets. The ones considered here are MRSTLO*, the nominal one, and CTEQ6.6 and HERAPDF1.0 for cross checks. The maximal deviation is taken as the systematic uncertainty.
- Uncertainty due to the modelling of the parton shower is evaluated using MC@NLO interfaced with HERWIG parton shower. These samples are generated with the CTEQ6.6 PDF sets and the ATLAS tune MC10 for pile-up. HERWIG does not treat correctly τ polarisation effects, so MC@NLO samples are used for that. The correction factor is estimated to be 0.9917 ± 0.0002 for the muon channel and 0.9904 ± 0.0002 for the electron.

Tables 8.2 and 8.3 show the geometric acceptance for the above described variations and the relative uncertainties for the 2010 data analysis. Similarly, in Tables 8.4 and 8.5 are shown the respective geometric acceptance and the relative uncertainties for the 2011 data analysis.

Table 8.2: Central values (PYTHIA MRSTLO*) and variations (others) of the A_Z geometric and kinematic acceptance factor for 2010 analysis.

	Muon channel	Electron channel
PYTHIA MRSTLO*	0.1169	0.1007
PYTHIA CTEQ6.6	0.1191	0.1026
PYTHIA HERAPDF1.0	0.1185	0.1020
PYTHIA CTEQ6.6 $m_{Z/\gamma^*} > 60$ GeV	0.1185	0.1022
MC@NLO CTEQ6.6 $m_{Z/\gamma^*} > 60$ GeV	0.1174	0.1016
MC@NLO CTEQ6.6 $m_{Z/\gamma^*} > 60$ GeV spin effect correction	0.1165	0.1006

Table 8.3: Relative uncertainties on the A_Z factors for the 2010 analysis.

	Muon channel	Electron channel
CTEQ 6.6 eigenvector set	1.2%	1.2%
Different PDF sets	1.9%	1.9%
Model dependence	1.8%	1.6%
Total uncertainty	2.9%	2.8%

Table 8.4: Central values (PYTHIA MRSTLO*) and variations (others) of the A_Z geometric and kinematic acceptance factor for 2011 analysis.

	Muon Channel	Electron Channel
PYTHIA MRSTLO*	0.0976	0.0687
PYTHIA CTEQ6.6	0.0998	0.0699
PYTHIA HERAPDF1.0	0.0992	0.0692
PYTHIA CTEQ6.6 $m_{Z/\gamma^*} > 60$ GeV	0.0994	0.0698
MC@NLO CTEQ6.6 $m_{Z/\gamma^*} > 60$ GeV	0.0973	0.0679

Uncertainties due to QED radiation and the modelling of the τ lepton decays are evaluated in the 2010 analysis and are found negligible. In particular, the QED modelling by PHOTOS

Table 8.5: Relative uncertainties on the A_Z factors for the 2011 analysis.

	Muon Channel	Electron channel
CTEQ 6.6 eigenvector set	0.9%	1.2%
Different PDF sets	2.2%	1.7%
Model dependence	2.1%	2.7%
Total uncertainty	3.1%	3.4%

has an accuracy better than 0.2%, which is smaller than the PDF uncertainty evaluated. For the τ lepton decays, instead of TAUOLA, a SHERPA sample is generated that includes its own libraries to decay τ leptons. The total theoretical systematic uncertainty on the A_Z factor is calculated by adding in quadrature all the different sources and a final 3% uncertainty is assigned on both channels in 2010 analysis and a 3.1% for $\tau_e\tau_{\rm h}$ and a 3.4% for $\tau_\mu\tau_{\rm h}$ in 2011 analysis.

8.2.2 Experimental Acceptance

The systematic uncertainty on experimental acceptance factor, C_Z, is the sum of the uncertainties described in Section 7 on signal Monte Carlo. Each of the uncertainties is varied and a new C_Z factor is calculated, listed in Tables 8.6 (2010) and 8.7 (2011).

Table 8.6: Relative systematic uncertainties in % for C_Z for the 2010 estimate for both semi-leptonic channels [55].

Systematic uncertainty	$\delta C_Z/C_Z$ $\tau_\mu\tau_{\rm h}$	$\delta C_Z/C_Z$ $\tau_e\tau_{\rm h}$
lepton efficiency	3.6%	9.2%
lepton resolution (μ energy scale)	0.2%	0.2%
Problematic regions in the calorimeter	–	0.4%
e charge misidentification	–	0.21%
τ id efficiency	8.6%	8.6%
Energy scale lepton and τ	8.6%	9.4%
Pileup re-weighting	0.4%	0.4%
Jet cleaning	1.8%	1.8%
Total systematic uncertainty	12.8%	15.8%
Statistical uncertainty	1.2%	1.4%
Total uncertainty	12.8%	15.9%

For each of the considered systematic uncertainties, the total cross section is calculated and the result is shown in Tables 8.8 (2010) and 8.9 (2011). The total systematic uncertainty on the cross section measurement is obtained by adding in quadrature all the estimated uncertainties.

Table 8.7: Relative systematic uncertainties in % for C_Z for the 2011 estimate for both semi-leptonic channels [56].

Systematic uncertainty	$\delta C_Z/C_Z$ (%) $\tau_\mu\tau_\mathrm{h}$	$\delta C_Z/C_Z$ (%)$\tau_e\tau_\mathrm{h}$
lepton trigger, reco, ID and isolation efficiency	1.7	4.8
muon resolution	<0.05	-
electron resolution	-	0.2
jet resolution	-	-
LAr hole	-	0.1
τ ID efficiency	5.1	5.1
electron-tau fake rate	-	0.02
e, τ, jet and $E_\mathrm{T}^\mathrm{miss}$ energy scale	7.9	9.5
tau trigger efficiency	-	4.5
Total systematic unc.	9.6	12.6
Statistical uncertainty	1.1	1.5

Table 8.8: Relative systematic uncertainties in % for the total cross-section measurement in 2010 data analysis.

Systematic uncertainty	$\delta\sigma/\sigma$ $\tau_\mu\tau_\mathrm{h}$	$\delta\sigma/\sigma$ $\tau_e\tau_\mathrm{h}$
τ id efficiency	8.6	8.6
e charge misidentification	–	0.14
k_W factor	0.12	0.18
Energy scale lepton and τ	10	11
Jet Cleaning	1.9	1.9
multijet estimate method	0.75	2
Pileup re-weighting	0.38	0.37
Problematic calorimeter regions	–	0.42
Theoret. cross section	0.2	0.12
lepton efficiency	3.8	9.6
lepton resolution(μ energy scale)	0.22	0.21
lepton-jet τ fake rate	1.1	0.66
A_Z systematics	2.9	2.9
Total Systematic	15	17
Statistical uncertainty	9.8	12
Luminosity	3.4	3.4

Table 8.9: Systematic and statistical uncertainties on the total cross section measurement in the 2011 data analysis.

Systematic uncertainty	$\delta\sigma/\sigma$ (%) $\tau_\mu\tau_\mathrm{h}$	$\delta\sigma/\sigma$ (%) $\tau_e\tau_\mathrm{h}$
lepton trigger, reco, ID and isolation efficiency	1.7	5.0
muon resolution	< 0.05	-
electron resolution	-	0.1
jet resolution	-	-
LAr hole	-	0.1
τ ID efficiency	5.2	5.2
electron-tau fake rate	-	0.2
e, τ, jet and $E_\mathrm{T}^\mathrm{miss}$ energy scale	8.2	9.3
tau trigger efficiency	-	4.7
k_W normalisation factor	<0.05	0.04
k_Z normalisation factor	<0.05	<0.05
QCD estimation	0.8	1.3
Background MC normalisation	0.1	0.2
A_Z uncertainties	3.1	3.4
MC statistics	1.2	1.4
Total systematic unc.	10.4	13.2
Statistical uncertainty	1.6	2.4

Chapter 9

Ditau Mass Reconstruction

Due to the presence of the neutrinos in the τ lepton decays, the reconstruction of the full $\tau\tau$ invariant mass poses an experimental challenge. A number of methods have been proposed for that purpose, typically in the context of Z/H separation. In the following section, the performance of several mass reconstruction techniques is evaluated using the relatively pure semileptonic $Z \to \tau\tau$ sample obtained for the cross section measurement. The methods considered are: the effective mass in Section 9.1, the collinear approximation in Section 9.2, the missing mass calculator (MMC) in Section 9.3, the mass bound (m_{bound}) in Section 9.4, the "true" transverse mass (m_T^{true}) in Section 9.5 and the combination of the latter two methods in Section 9.6. The motivation for this combination is given in the last section.

There are two methods that are not considered here: visible mass and transverse mass. Visible mass is the invariant mass of the visible decay products of the two τ leptons, electrons, muons or quarks, and it was used for the measurement of the cross section of the $Z \to \tau\tau$ decay. This reconstruction method is not used here, since it does not include the neutrinos of the τ decay and hence, does not reconstruct the whole event. Transverse mass is defined in Section 6.3.5 in eq. (6.1). It is used to obtain the invariant mass of the decaying particles when the parent decays into one visible lepton and a neutrino. In this case, one assumes that the mass of the decay products is zero and the parent particle is produced with zero recoil. The transverse mass is a perfect method to reconstruct a W boson, where the approximation that the missing transverse momentum is identical to the transverse momentum of the neutrino is valid. Due to the presence of two neutrinos this approximation is not applicable to a Z boson. Hence the method is ignored for the purposes of this chapter.

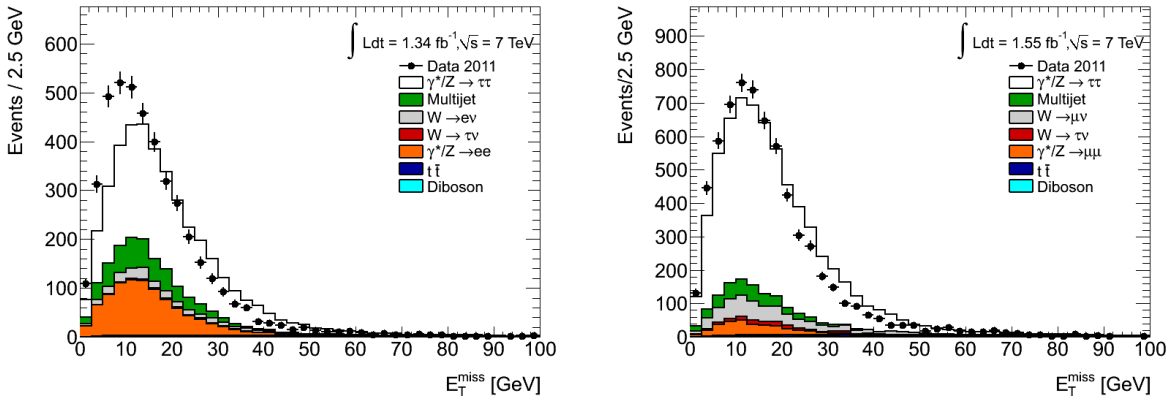
For all considered methods, the invariant mass of the $\tau\tau$ system is reconstructed after all standard selection cuts and an additional $E_{\text{T}}^{\text{miss}}$ requirement are applied. The distribution of the $E_{\text{T}}^{\text{miss}}$ after the standard selection is shown in Fig 9.1. The Monte Carlo does not describe the $E_{\text{T}}^{\text{miss}}$ well below 20 GeV, which is why for the mass reconstruction a requirement for

$$E_{\text{T}}^{\text{miss}} > 20 \text{ GeV} \tag{9.1}$$

has been implemented.

9.1 Effective Mass

The effective mass is the transverse invariant mass of the decay products of a parent particle that decays to one visible and one invisible part. The effective mass, in contrast to the transverse mass, does not require the parent particle to be produced with no recoil. Also, it does not make



(a) electron channel

(b) muon channel

Figure 9.1: Distribution of the $E_{\mathrm{T}}^{\mathrm{miss}}$ for $\tau_e\tau_{\mathrm{h}}$ (left) and $\tau_\mu\tau_{\mathrm{h}}$ (right) after all standard cuts applied.

any assumptions about the masses of the decay products. It is defined as

$$m_{\tau\tau}^{eff} = p_{T\,\tau 1} + p_{T\,\tau 2} + p_T^{miss}, \tag{9.2}$$

where $p_{T\,\tau i}$ are the transverse momenta of the visible τ lepton decay products and p_T^{miss} the missing transverse momentum. The method takes advantage of the well-measured transverse momenta of electrons, muons and hadronic τ candidates. Additionally, it accounts for the neutrinos of the decay by including the transverse missing momentum of the event. The disadvantage of the method is that it does not take into account the τ lepton mass and the angles in the transverse plane. As a result, the invariant mass of the parent particle (Z or H) is not well constrained and thus, an accurate determination of its mass is impossible.

The effective mass for the $\tau_e\tau_{\mathrm{h}}$ and $\tau_\mu\tau_{\mathrm{h}}$ channels are shown in Fig. 9.6(a,b) after all necessary requirements have been applied.

9.2 Collinear Approximation

The collinear approximation is a popular method for the $\tau\tau$ invariant mass reconstruction because it succeeds in accurately measuring the mass of the parent particle by requiring the τ lepton decay products to be produced collinearly with the τ in the laboratory frame [93]. This condition is fulfilled when the mass of the parent particle is significantly larger than that of the τ lepton ($m_{H/Z}/2 >> m_\tau$) and therefore, the τ leptons are produced highly boosted.

The invariant mass of the $\tau\tau$ system is calculated by

$$\begin{aligned} m_{\tau\tau}^2 &= (p_{\tau_1} + p_{\tau_2})^2 \\ &= m_{\tau_1}^2 + m_{\tau_2}^2 + 2\,(E_{\tau_1}E_{\tau_2} - \vec{p_{\tau_1}}\vec{p_{\tau_2}}) \end{aligned} \tag{9.3}$$

where p_{τ_i}, m_{τ_i} and E_{τ_i} are the four-momenta, the mass, and the energy of the two τ_i leptons, $i = 1, 2$. The mass of the τ leptons is considered to be zero, m_{τ_1}, $m_{\tau_2} = 0$, eq. (9.3) becomes

$$\begin{aligned} m_{\tau\tau}^2 &= 2\,(E_{\tau_1}E_{\tau_2} - \vec{p_{\tau_1}}\vec{p_{\tau_2}}) \\ &= 2 \cdot E_{\tau_1}E_{\tau_2}(1 - \cos\theta_{\tau_1\tau_2}), \end{aligned} \tag{9.4}$$

where $\theta_{\tau_1\tau_2}$ the angle between the τ leptons. When the collinearity requirement is considered, the energy of the τ leptons can be written as the sum of the visible decay products, E_{visi}, and the neutrinos, $E_{\nu i}$,

$$E_i = E_{\text{vis}\,\text{i}} + E_{\nu\, i}.$$

$$m^2_{\tau\tau} = 2 \cdot (E_{\text{vis}\,1} + E_{\nu\,1})(E_{\text{vis}\,2} + E_{\nu\,2})(1 - \cos\theta_{\tau_1\tau_2}), \tag{9.5}$$

Then the only unknowns of the system are the fraction of the τ momentum carried away by the neutrinos, which can be calculated by requiring the momenta of the neutrinos to be equal to the x- and y-components of the $E_{\text{T}}^{\text{miss}}$

$$x_1 = \frac{p^1_{\text{T}}}{p^1_{\text{T}} + p^{\text{miss}}_{\text{T}}} = \frac{p^1_x p^2_y - p^1_y p^2_x}{p^1_x p^2_y + E^{miss}_x p^2_y - p^1_y p^2_x - E^{miss}_y p^2_x} \tag{9.6}$$

$$x_2 = \frac{p^2_{\text{T}}}{p^2_{\text{T}} + p^{\text{miss}}_{\text{T}}} = \frac{p^1_x p^2_y - p^1_y p^2_x}{p^1_x p^2_y - E^{miss}_x p^1_y - p^1_y p^2_x + E^{miss}_y p^2_x}. \tag{9.7}$$

By re-writing the invariant mass formula with the help of eqs. (9.6, 9.7), the $\tau\tau$ mass is given by

$$m_{\tau\tau} = \frac{m_{(\text{vis}\,1\text{vis}\,2)}}{\sqrt{x_1 x_2}} \quad \text{for } x_1 x_2 > 0, \tag{9.8}$$

where $m_{(\text{vis}\,1\text{vis}\,2)}$ is the invariant mass of the visible decay products.

The fractions x_i are meaningful when the τ leptons are not produced back-to-back, otherwise eqs. (9.6, 9.7) are linearly dependent and cannot be solved. Hence, a requirement on the angle between the visible decay products is imposed

$$|\Delta\phi(\tau_{\text{vis}\,1}, \tau_{\text{vis}\,2})| < 2.9.$$

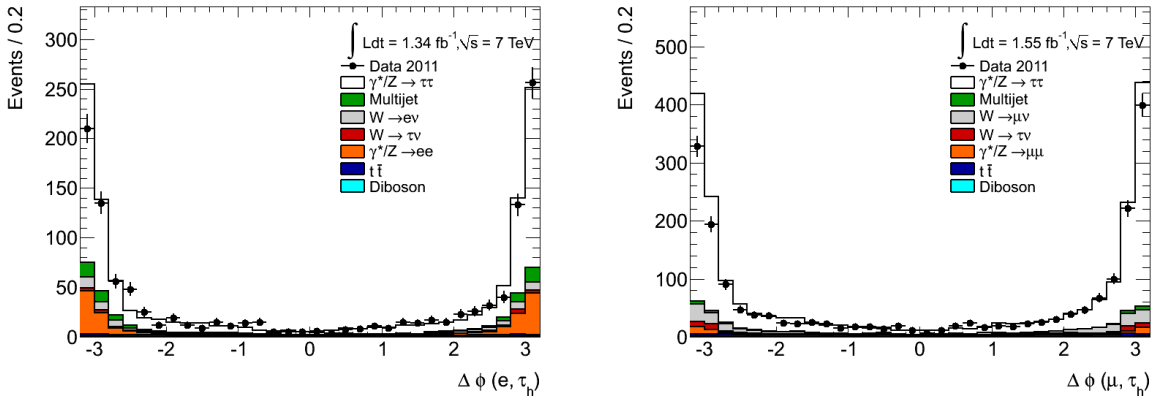
This requirement reduces the acceptance rate of signal events to ∼50%. Nonetheless, with sufficient statistics, one can get a good estimation of the true mass of the Z or H bosons.

The $Z \to \tau\tau$ invariant mass is reconstructed with the collinear approximation for the $\tau_e\tau_{\text{h}}$ and $\tau_\mu\tau_{\text{h}}$ channels after all selection requirements and the $\Delta\phi(\tau_{\text{h}}, \tau_{\text{l}})$ restriction have been applied. The $\Delta\phi(\tau, \text{lepton})$ distribution after the $E_{\text{T}}^{\text{miss}}$ cut is plotted in Fig. 9.2. The mass distributions are shown in Fig. 9.6(c,d) after the $E_{\text{T}}^{\text{miss}}$ cut and Fig. 9.8 after the angular cut. In the mass distributions a preference towards higher masses appears, which is expected due to the constraint imposed by eq. (9.8), i.e. $m_{\tau\tau} > m_{(\text{vis}\,1\text{vis}\,2)}$.

The collinear approximation is sensitive to the mis-measurements of the $E_{\text{T}}^{\text{miss}}$ due to the orientation of the τ leptons. To account for it, a 1σ variation in the calculation of the energy scale of the event, see Section 7, has been performed and propagated through to the reconstructed mass. The result is given in Table 9.3. Varying the energy scale has a small effect on the central value of the distribution, ∼4% for both channels. On the other hand, the width of the Z distribution is more sensitive to the mis-measurements of the energy, especially in the electron channel, ∼12%.

9.3 Missing Mass Calculator

The Missing Mass Calculator (MMC) is a relatively new method introduced in [94] for the reconstruction of the $\tau\tau$ invariant mass. It resembles the collinear approximation except that



(a) electron channel

(b) muon channel

Figure 9.2: Distribution of the $\delta\phi$ for $\tau_e\tau_{\rm h}$ (left) and $\tau_\mu\tau_{\rm h}$ (right) after all standard cuts and $E_{\rm T}^{\rm miss} > 20$ GeV applied.

it avoids the use of the collinearity constraint by imposing constraints on the orientation of the neutrinos and the visible decay products.

Writing the equations for the reconstruction of the full event topology reveals that there are more unknown parameters than there are available equations

$$E_{\rm x}^{\rm miss} = p_1^{\rm miss}\,\sin\theta_1^{\rm miss}\,\cos\phi_1^{\rm miss} + p_2^{\rm miss}\,\sin\theta_2^{\rm miss}\,\cos\phi_2^{\rm miss}, \tag{9.9}$$

$$E_{\rm y}^{\rm miss} = p_1^{\rm miss}\,\sin\theta_1^{\rm miss}\,\sin\phi_1^{\rm miss} + p_2^{\rm miss}\,\sin\theta_2^{\rm miss}\,\sin\phi_2^{\rm miss}, \tag{9.10}$$

$$\begin{aligned}(m_1^\tau)^2 = {} & (m_1^{\rm miss})^2 + (m_1^{\rm vis})^2 + 2\sqrt{(p_1^{\rm vis})^2 + (m_1^{\rm vis})^2}\sqrt{(p_1^{\rm miss})^2 + (m_1^{\rm miss})^2} \\ & - 2\,p_1^{\rm vis}\,p_1^{\rm miss}\,\cos\Delta\theta_1^{\rm vm},\end{aligned} \tag{9.11}$$

$$(m_2^\tau)^2 = (m_2^{\rm vis})^2 + 2\sqrt{(p_2^{\rm vis})^2 + (m_2^{\rm vis})^2}\,p_2^{\rm miss} - 2\,p_2^{\rm vis}\,p_2^{\rm miss}\,\cos\Delta\theta_2^{\rm vm}, \tag{9.12}$$

where $E_{\rm x,y}^{\rm miss}$ are the x and y components of the missing energy, $m_i^{\rm vis}$, $p_i^{\rm vis}$ the mass and momentum of the two visible decay products and m_τ the mass of the τ lepton = 1.777 GeV. These variables are measured experimentally and are considered as known parameters. The angle between the visible and invisible decay products, $\theta_i^{\rm vm}$, can be expressed in terms of the other variables and thus taken as a known quantity. However, the momentum of the τ leptons carried away by the neutrinos, $p_i^{\rm miss}$, and the invariant mass of the neutrinos, $m_i^{\rm miss}$, are unknowns. For a hadronically decaying τ there is only the τ neutrino present and hence $m_i^{\rm miss}$ can be set to zero. For the leptonically decaying one, that is not the case.

The system of equations can be constrained by using kinematic information from the event. For example, the three-dimensional angle between the directions of visible and invisible decay products in $\theta\times\phi$ space, $\Delta\theta_{3D}$ [95]. The angle $\Delta\theta_{3D}$ is evaluated from $Z\to\tau\tau$ events for different values of the τ transverse momentum, $10 < p_{\rm T} < 230$ GeV, and for the three types of τ decays, leptonic, one- and three-prong. In Fig. 9.3, the calculated $\Delta\theta_{3D}$ is shown for a representative $p_{\rm T}$ range, $45 < p_{\rm T} < 50$ GeV. For the hadronically decaying τ, eqs. (9.9-9.12) can be solved for every pair $(\Delta\phi_1, \Delta\phi_2)$. For the leptonic one, the grid needs to contain the neutrino invariant mass $(\Delta\phi_1, \Delta\phi_2, m_1^{miss})$. The expected value of $\Delta\theta_{3D}$ is used as a probability density function in a global event probability to obtain the best estimate for the $m_{\tau\tau}$ for the event

$$\mathcal{P}_{\rm event} = \mathcal{P}(\Delta\theta_1, p_{\tau_1}, \text{decay type}) \times \mathcal{P}(\Delta\theta_2, p_{\tau_2}, \text{decay type}).$$

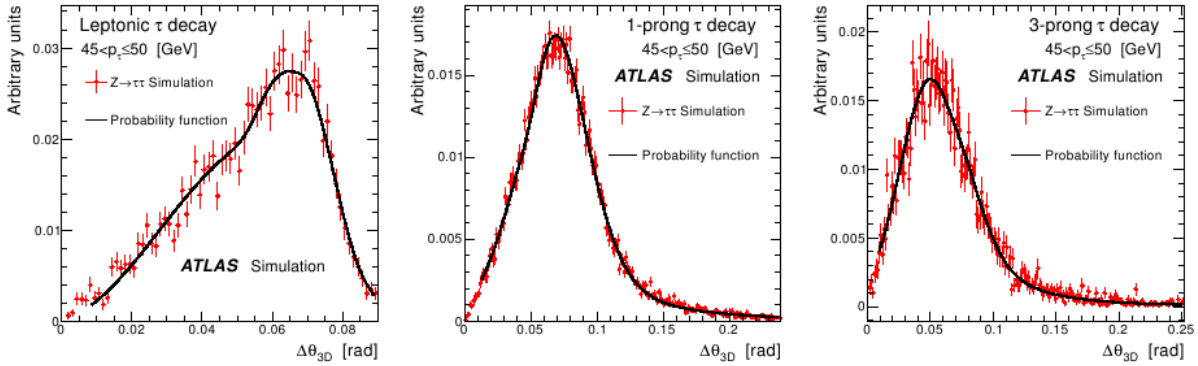


Figure 9.3: Example of the probability distribution functions $\Delta\theta_{3D}$ for a particular τ p_{T} ($45 < p_{\mathrm{T}} \leq 50$ GeV). These functions are used in the calculation of the likelihood for the three τ decay modes: leptonic (left), single-prong (middle) and three-prong (right). Taken from [95].

The main constraint of the method comes from the direction of the missing energy, hence any mis-measurements might affect the performance of the method. To account for it, an extra weight for the resolution of the missing energy in x- and y- directions is introduced. In this case the final event weight is given by

$$\mathcal{P} = \mathcal{P}(\Delta\theta_1, p_{\tau_1}, \text{decay type}) \times \mathcal{P}(\Delta\theta_2, p_{\tau_2}, \text{decay type}) \times \mathcal{P}(\Delta E_x^{miss}) \times \mathcal{P}(\Delta E_y^{miss}).$$

The MMC efficiency is defined as the ratio of the number of the events that are reconstructed with the MMC over the number of events that are given as an input to the algorithm. It is found to be higher than 94% for signal events. The efficiency loss is mostly driven by the fact that some pairs of angles might lie outside the $(\Delta\phi_1, \Delta\phi_2)$ grid. Another reason is the fact that the measured missing energy might fluctuate more than 3σ in resolution and thus exceed the current scan range.

Finally, the systematic uncertainties of the method are of two kinds; τ candidate related ones and non-τ candidate related. In order to account for the former category the τ and $E_{\mathrm{T}}^{\mathrm{miss}}$ energy scale are varied according to the recipe described in Section 7. In Table 9.3 the shift of the peak and the width of the MMC distribution is shown for the two channels for signal Monte Carlo events. The peak position is not affected much by the energy scale variation, 3%, while the width is more sensitive to it, 10%. For the latter kind of uncertainties, the $E_{\mathrm{T}}^{\mathrm{miss}}$ angular resolution and jet energy resolution are varied by $\pm 1\sigma$. The effect on the nominal values for the signal is shown in Table 9.4. In both cases the effect on the peak and the width of the distribution is small, $< 6\%$. The effect on the background is difficult to quantify due to its low statistics.

Since only the angular distribution of the $E_{\mathrm{T}}^{\mathrm{miss}}$ is used for the estimation, the $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement motivated in the introduction of the chapter is of less importance. Nonetheless the $\tau\tau$ invariant mass is calculated with and without it for comparison. The MMC mass for the $\tau_e\tau_{\mathrm{h}}$ and $\tau_\mu\tau_{\mathrm{h}}$ are shown in Fig. 9.5 after all analysis cuts have been applied. Similarly, in Fig. 9.6(e,f) are shown the MMC distributions for the two channels after including the $E_{\mathrm{T}}^{\mathrm{miss}}$ cut.

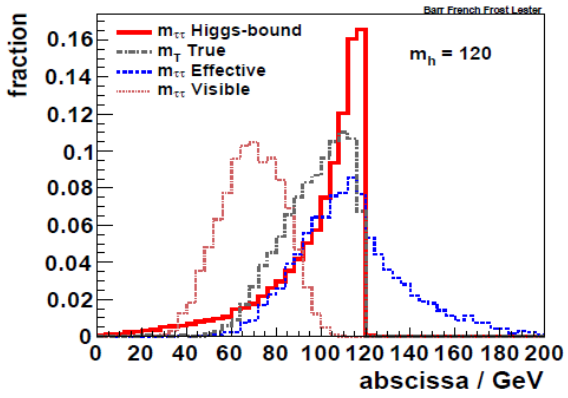


Figure 9.4: Distribution of the m_{bound} variable for truth-level Monte Carlo events for a Higgs signal at 120 GeV. The variable is compared to m_T^{true}, m_{eff} and the visible mass. Taken from [96].

9.4 Mass Bound

The mass bound method was developed to mainly separate between Z and H bosons in the $\tau\tau$ final states [96]. This is achieved by constructing a variable capable of discriminating between the Higgs signal and the Z boson irreducible background. The discriminant is constructed by minimising the mass of the parent particle over all possible values of the phase-space, i.e. all combinations constrained by the mass of the involved particles and the topology of the decay

$$m_{\tau\tau}^{\text{bound}} = \min_{Q_1^\mu, Q_2^\mu | \aleph} \sqrt{H^\mu H_\mu} \tag{9.13}$$

where

$$H^\mu = P_1^\mu + Q_1^\mu + P_2^\mu + Q_2^\mu$$

is the four-momentum of the visible ($P_{1,2}^\mu$) and invisible ($Q_{1,2}^\mu$) τ lepton decay products. In the ideal case, the m_{bound} should provide a sharp cut off, keeping all the Z events below m_Z and the H events above a hypothetic m_H:

$$m_Z \leq m_{\text{bound}} \leq m_H.$$

This is shown at truth level Monte Carlo events in Fig. 9.4.

The minimisation of the m_{bound} is subject to the constraints ($\aleph$) comprising the mass-shell requirement for the $\tau\tau$ system:

$$\begin{aligned} Q_1^\mu Q_{1\mu} &= 0 && (9.14)\\ Q_2^\mu Q_{2\mu} &= 0 && (9.15)\\ (P_1^\mu + Q_1^\mu)(P_{1\mu} + Q_{1\mu}) &= m_\tau^2 && (9.16)\\ (P_2^\mu + Q_2^\mu)(P_{2\mu} + Q_{2\mu}) &= m_\tau^2 && (9.17) \end{aligned}$$

and the missing transverse momentum:

$$\vec{q}_{1T} + \vec{q}_{2T} = \vec{p}_T^{\,\text{miss}}. \tag{9.18}$$

There is a solution to this set of constraints *if and only if* the following additional condition is true

$$m_{T2}(P_1, P_2, \vec{p}_T^{\,\text{miss}}) < m_\tau, \tag{9.19}$$

where m_{T2} is the 'stransverse' mass[1] of the τ decay particles in a decay with multiple invisible decay products, as defined in [97, 98]. The constraint (9.19) is necessary for providing a minimisation region for the m_{bound} mass. In an ideal experiment, τ pairs satisfy this relation *a priori*, but when detector resolution effects are considered that is not the case anymore. In the $Z \to \tau\tau$ study, it is found that about 40% of the signal events are rejected.

The bound mass method uses only the four-momenta of the visible τ decay products and the measured missing energy. Thus no additional systematic uncertainty is introduced in the analysis, apart from the ones investigated in the cross section estimation, Section 7. The shift of the peak and the width of the distribution when varying the energy scale is given in Table 9.3. The reason why only the energy scale uncertainty is quoted is that it is the dominant one and the others are not expected to have a significant impact on the mass reconstruction. Also, the energy scale contains a shift in the $E_{\text{T}}^{\text{miss}}$ which is one of the main inputs of the method.

In Fig. 9.7(a,b), the m_{bound} distributions are plotted for the two semi-leptonic channels. Although, 40-50% of the signal is lost, the method is still valuable, since it succeeds in rejecting most of the Z+jets background. The tails at high masses are due to $E_{\text{T}}^{\text{miss}}$ resolution effects. That is the m_{bound} does not have a steep cut off at the parent particle's mass, as it is designed to, Fig. 9.4. This effect might reduce the discriminating power of the method between the Z boson and a H boson in case it is relatively light (~ 125 GeV).

9.5 "True" Transverse Mass

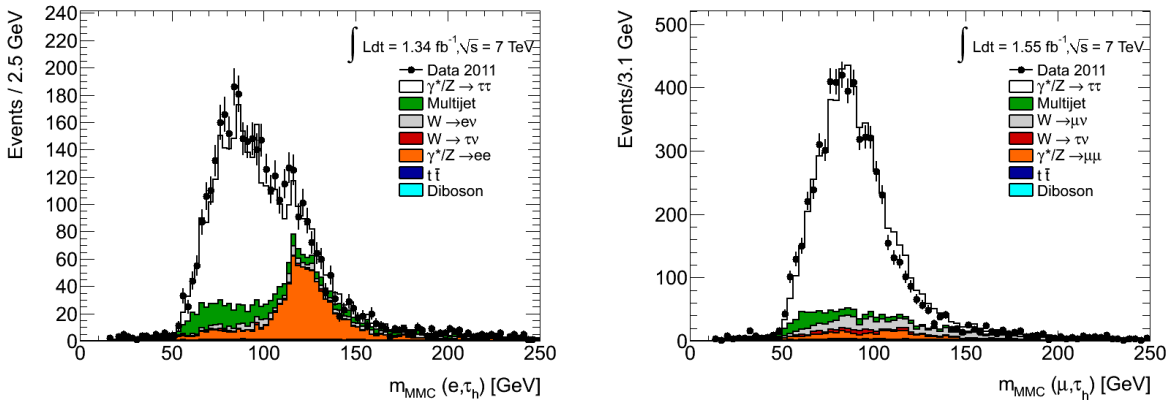
A generalised version of the 'traditional' transverse mass is the "true" transverse mass. In the case of the "true" transverse mass the clustering of the decay products is done as visible decay products and invisible decay products. The invariant mass of the visible particles does not need to give the mass of the parent particle, while the invariant mass of the invisible part has to be zero, since neutrinos are considered to be massless. Hence, the m_T^{true} is similar to m_{bound}, but the τ leptons do not need to be produced on-shell [97, 99]. m_T^{true} is defined as

$$(m_{\text{T}}^{\text{true}})^2 = m_{\text{vis1vis2}}^2 + 2\left(p_T^{miss}\sqrt{(m_{\text{vis1vis2}}^2 + \vec{p}_{\text{T,vis1vis2}}^{\,2})} - \vec{p}_T^{\,\text{miss}} \cdot \vec{p}_{\text{T,vis1vis2}}\right). \tag{9.20}$$

Although m_T^{true} does not have any constraints for the τ system, the motivation for using it is that it can reconstruct the transverse mass of the visible and invisible τ decay products for every event. In Fig 9.7(c,d), the m_T^{true} distribution is shown for $\tau_e\tau_{\text{h}}$ and $\tau_\mu\tau_{\text{h}}$. From the figures it occurs that the tails at high masses are not present in the Z peak, however the Z+jets background peaks at the expected value of a light Higgs boson, $\sim$126 GeV. That means that if the m_T^{true} were to be used for Higgs boson discovery, a good handle on the electroweak background would be necessary.

As in the case of the m_{bound} the systematic uncertainties for the method are related to the experimental measurement of the involved objects. The dominant uncertainty comes from the energy scale measurement. The effect on the peak and the width of the signal from shifting it by 1σ is given in Table 9.3. Equivalently to the other methods the mass width is shifted by maximum 5% and the peak 3%.

[1] The method as well as the name 'stransverse' mass were initially developed in searches for supersymmetrical particles.



(a) electron channel

(b) muon channel

Figure 9.5: MMC distribution for $\tau_e\tau_\mathrm{h}$ (left) and $\tau_\mu\tau_\mathrm{h}$ (right). The distribution is plotted after all standard cuts.

9.6 "Combined" Mass

The m_bound method can be a good discriminator between the Z and the Higgs boson. However, due to constraint (9.19) about 40% of the signal phase space is rejected. On the other hand, the m_T^true method is reconstructed for every event, but it does not constrain the τ system, thus admitting a significant amount of electroweak background. The drawbacks of the two methods can be surpassed if they are combined: the $m_{\tau\tau}$ invariant mass is reconstructed with the m_bound method, unless it fails, in which case the m_T^true is used instead

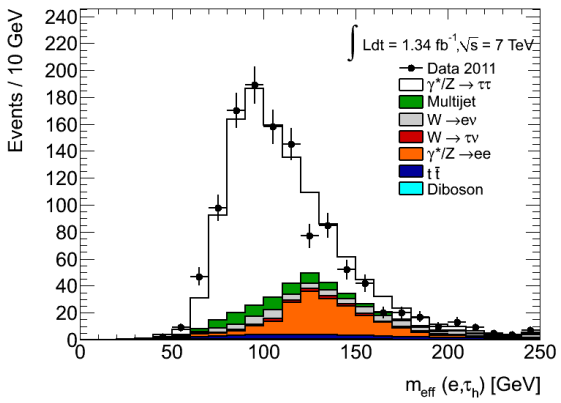
$$m_\mathrm{combi} = m_\mathrm{bound} \text{ or } m_T^\mathrm{true}.$$

In Fig. 9.7(e,f), the m_combi distribution is plotted after the $E_\mathrm{T}^\mathrm{miss}$ cut for $\tau_e\tau_\mathrm{h}$ and $\tau_\mu\tau_\mathrm{h}$ channels. It is shown that the combined distribution has smaller tails at high masses and a smaller Z+jets contamination. Nonetheless, a better handle on the background is necessary, especially in the electron channel.

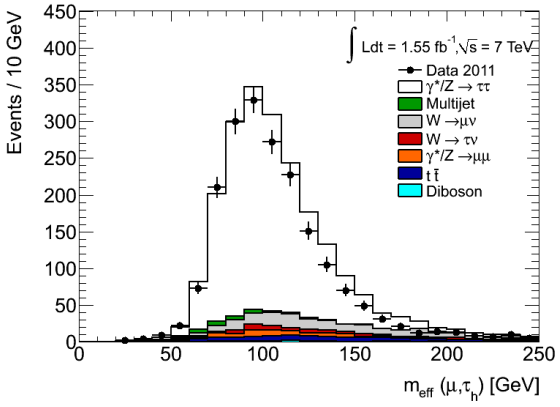
The systematic uncertainty of the energy scale shift on the peak and width of m_combi is given in Table 9.3. The peak position is only shifted by 3(4)% for the $\tau_e\tau_\mathrm{h}(\tau_\mu\tau_\mathrm{h})$ and the width by 5%.

9.7 Performance of the Mass Methods

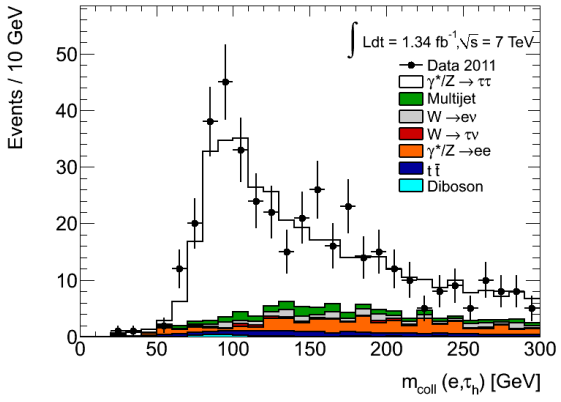
The Z resonance is described by a Breit-Wigner distribution if no detector effects are taken into account. On the other hand, the experimental resolution of the energy $(1/p_\mathrm{T})$ measurement on the electrons (muons) is approximately described by a Gaussian distribution. To account for this, the Z resonance is fitted by a convolution of the Breit-Wigner and the Gauss functions in the evaluation of the performance of the mass methods. For the case that the Z decays to τ leptons, the τ reconstruction effects are dominating and hence, the Z resonance is better described by a Gaussian distribution than by a Breit-Wigner. Nonetheless, a Gaussian and a Breit-Wigner function are fitted to the data and the signal Monte Carlo for all the mass methods after all relevant cuts have been applied. The data are fitted after subtracting the background



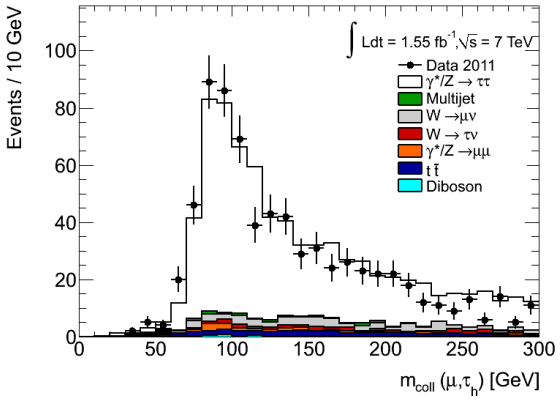
(a) m_{eff} electron channel



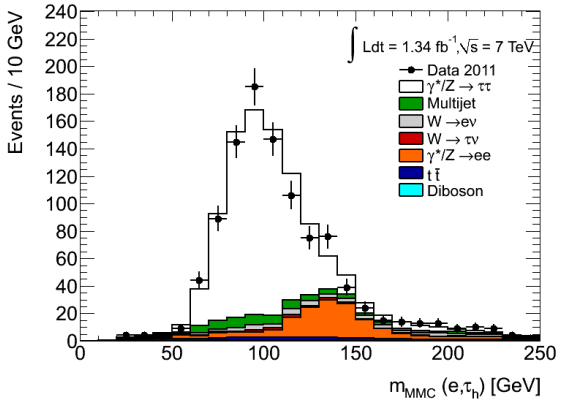
(b) m_{eff} muon channel



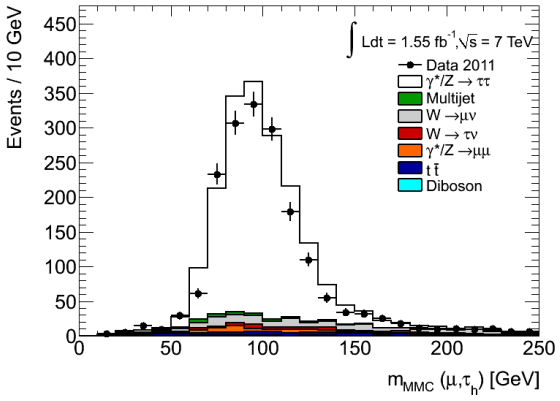
(c) m_{coll} electron channel



(d) m_{coll} muon channel

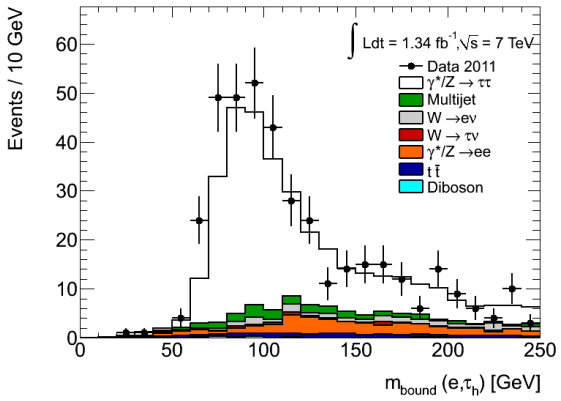


(e) MMC electron channel

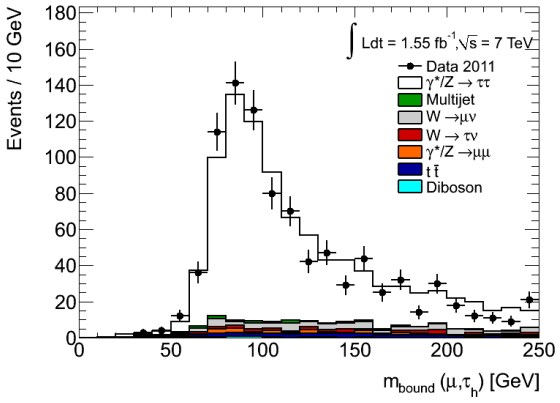


(f) MMC muon channel

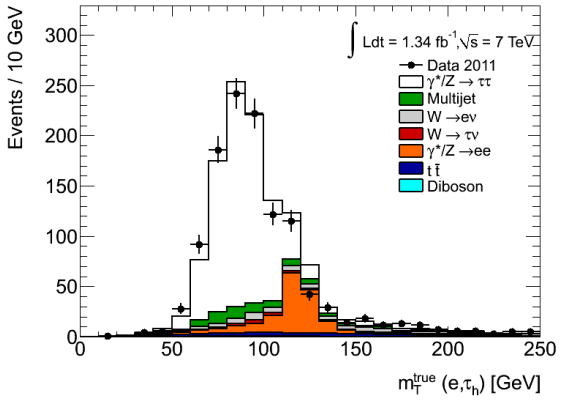
Figure 9.6: Mass distributions of the $\tau\tau$ system for $\tau_e\tau_{\mathrm{h}}$ (left) and $\tau_\mu\tau_{\mathrm{h}}$ (right). The distributions are plotted after all standard cuts including $E_{\mathrm{T}}^{\mathrm{miss}} > 20$ GeV.



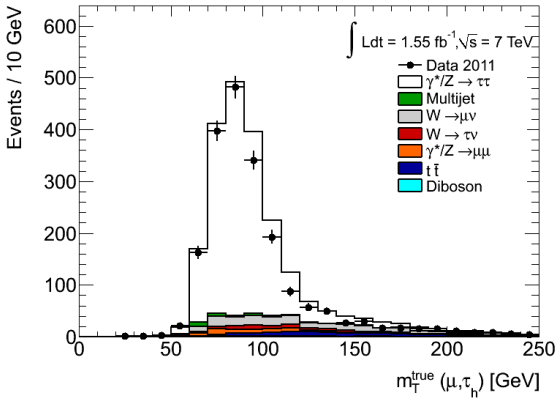
(a) m_{bound} electron channel



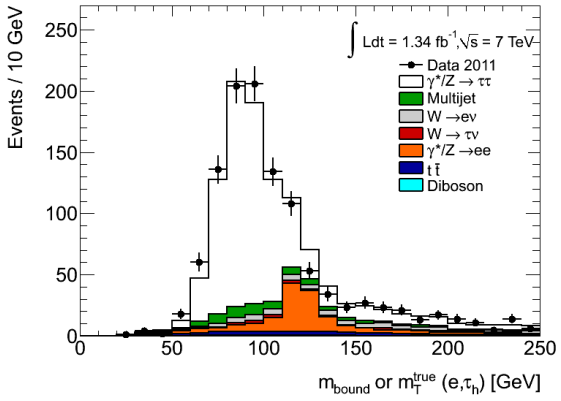
(b) m_{bound} muon channel



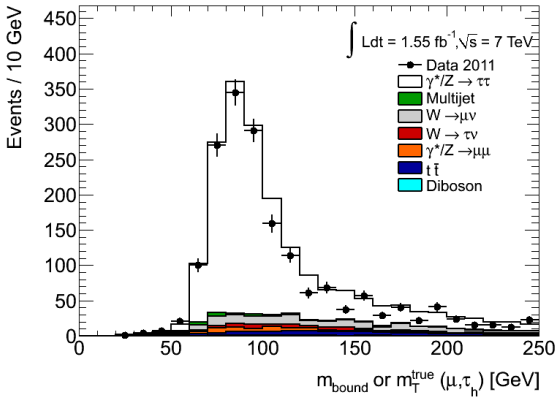
(c) m_T^{true} electron channel



(d) m_T^{true} muon channel

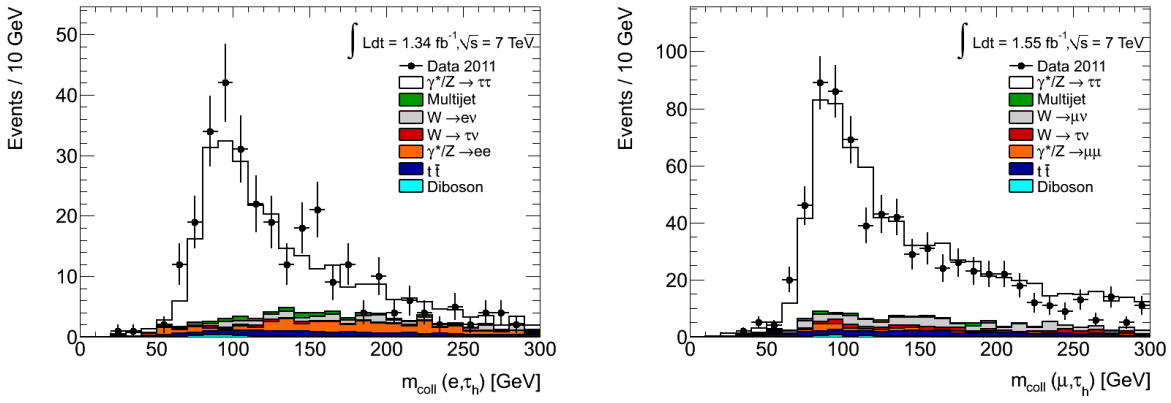


(e) m_{bound} or m_T^{true} electron channel



(f) m_{bound} or m_T^{true} muon channel

Figure 9.7: Mass distributions of the $\tau\tau$ system for $\tau_e\tau_{\text{h}}$ (left) and $\tau_\mu\tau_{\text{h}}$ (right). The distributions are plotted after all standard cuts including $E_{\text{T}}^{\text{miss}} > 20$ GeV.



(a) electron channel

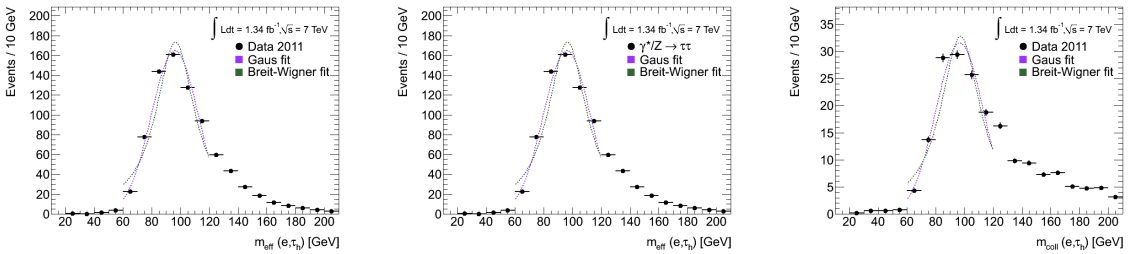
(b) muon channel

Figure 9.8: Collinear approximation for $\tau_e\tau_{\rm h}$ (left) and $\tau_\mu\tau_{\rm h}$ (right). The distribution is plotted after all standard cuts including $E_{\rm T}^{\rm miss} > 20$ GeV and $|\Delta\phi(\tau, \text{lepton})| < 2.9$.

contribution from the Monte Carlo predictions. The resulting curves are shown in Fig. 9.9 and 9.10 and mean values and widths are summarised in Tables 9.1 and 9.2.

To check the stability of the methods, the masses are reconstructed after varying the dominant systematic uncertainty of the process, the energy scale, and the deviation from the nominal value is given in Table 9.3. For the MMC two extra systematic uncertainties are checked, the $E_{\rm T}^{\rm miss}$ and the jet energy resolution. The corresponding effects are presented in Table 9.4. The systematic uncertainties are discussed previously for each method separately.

A final check of the performance of the methods is done by estimating the ratio of the signal over the background contributions, which is summarised in Table 9.5. The muon channel has a higher background rejection efficiency due to the stronger suppression of the Z+jets background. Amongst the studied methods, the ones that constrain the phase space using the τ decay topology, like the MMC and the $m_{\rm bound}$ methods, provide a better separation between signal and background. However, the collinear approximation offers comparable separation power without using statistical methods that might introduce additional systematic uncertainties.



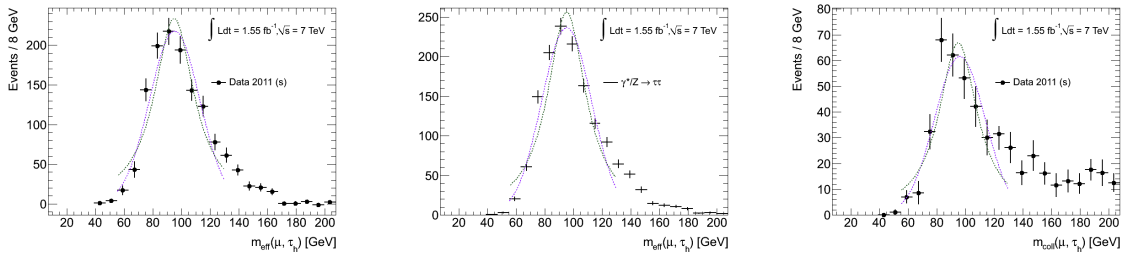
(a) signal data – electron channel (b) signal MC – electron channel (c) signal data – electron channel

(d) signal MC – electron channel (e) signal data – electron channel (f) signal MC – electron channel

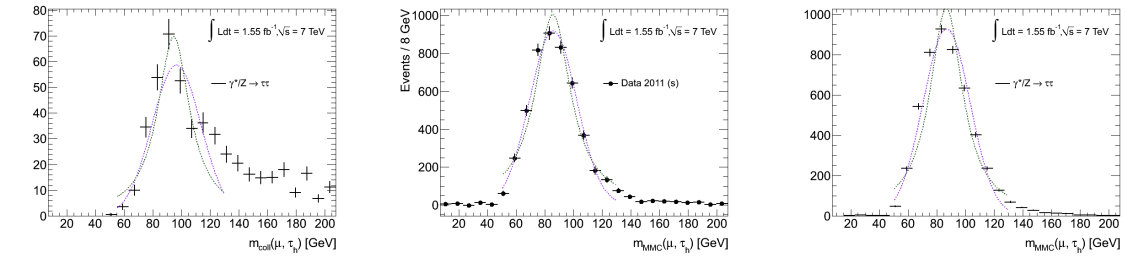
(g) signal data – electron channel (h) signal MC – electron channel (i) signal data – electron channel

(j) signal MC – electron channel (k) signal data – electron channel (l) signal MC – electron channel

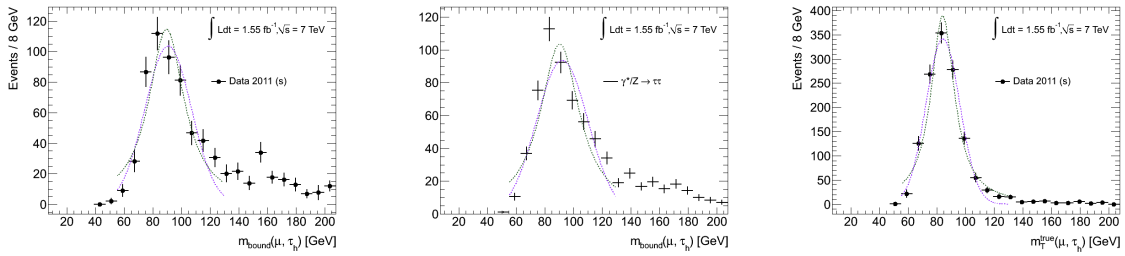
Figure 9.9: Fit all mass distribution for the $\tau_e\tau_\mathrm{h}$ channel after the necessary cuts have been applied, for signal Monte Carlo (right) and data after subtracting the background expectation from Monte Carlo (left).



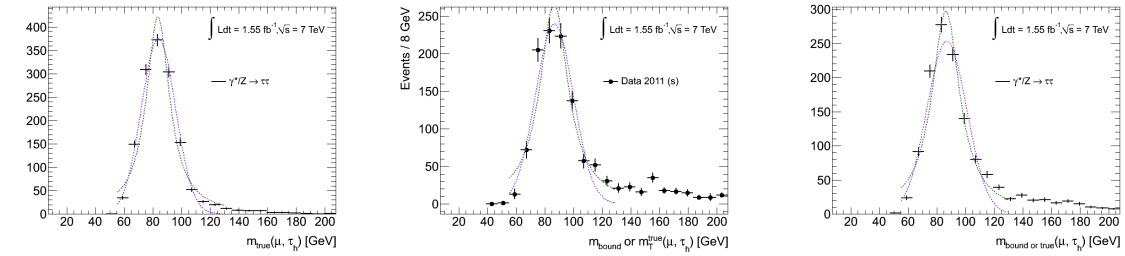
(a) signal data – muon channel (b) signal MC – muon channel (c) signal data – muon channel



(d) signal MC – muon channel (e) signal data – muon channel (f) signal MC – muon channel



(g) signal data – muon channel (h) signal MC – muon channel (i) signal data – muon channel



(j) signal MC – muon channel (k) signal data – muon channel (l) signal MC – muon channel

Figure 9.10: Fit of mass distribution for the $\tau_\mu\tau_\mathrm{h}$ channel after the necessary cuts have been applied.

Table 9.1: The Gaussian (g) and Breit Wigner (BW) mean and width values of the reconstructed masses for signal Monte Carlo.

$\tau_e\tau_h$	μ (Gauss) [GeV]	σ (Gauss) [GeV]	μ (BW) [GeV]	width (BW) [GeV]
m_{vis}	61.1 ± 0.1	10.4 ± 0.1	60.5 ± 0.1	14.7 ± 0.1
m_{eff}	96.0 ± 0.1	16.3 ± 0.1	96.3 ± 0.2	33.0 ± 0.4
m_{coll}	98.2 ± 0.4	17.1 ± 0.3	99.1 ± 0.4	35.1 ± 0.9
m_{MMC}	97.9 ± 0.1	17.8 ± 0.1	98.5 ± 0.1	29.2 ± 0.3
m_{bound}	92.8 ± 0.3	17.0 ± 0.3	92.6 ± 0.3	36.5 ± 1.0
m_T^{true}	87.9 ± 0.1	14.5 ± 0.1	87.5 ± 0.1	28.2 ± 0.3
m_{combi}	90.5 ± 0.1	14.7 ± 0.1	90.0 ± 0.1	28.5 ± 0.3
$\tau_\mu\tau_h$	μ (Gauss) [GeV]	σ (Gauss) [GeV]	μ (BW) [GeV]	width (BW) [GeV]
m_{vis}	59.3 ± 0.5	7.8 ± 0.3	59.4 ± 0.2	12.0 ± 0.5
m_{eff}	95.2 ± 0.4	17.5 ± 0.4	95.2 ± 0.5	32.8 ± 1.1
m_{coll}	96.5 ± 1.1	17.1 ± 0.8	94.5 ± 0.9	2.7 ± 2.1
m_{MMC}	86.8 ± 0.2	16.4 ± 0.1	86.6 ± 0.2	28.3 ± 0.4
m_{bound}	92.1 ± 0.7	18.0 ± 0.7	90.5 ± 0.8	33.7 ± 2.0
m_T^{true}	83.9 ± 0.2	11.9 ± 0.2	83.6 ± 0.3	20.0 ± 0.5
m_{combi}	86.9 ± 0.3	14.0 ± 0.3	86.3 ± 0.3	23.9 ± 0.8

Table 9.2: The Gaussian and Breit Wigner (BW) mean and width values of the reconstructed masses for data.

$\tau_e\tau_h$	μ (Gauss) [GeV]	σ (Gauss) [GeV]	μ (BW) [GeV]	width (BW) [GeV]
m_{eff}	95.4 ± 1.1	16.9 ± 0.8	95.7 ± 1.2	28.8 ± 2.0
m_{coll}	94.7 ± 1.8	15.5 ± 1.4	95.1 ± 2.0	25.6 ± 3.5
m_{MMC}	95.0 ± 0.9	15.8 ± 0.8	96.1 ± 1.0	26.5 ± 1.8
m_{bound}	90.3 ± 1.6	16.4 ± 1.2	91.6 ± 2.0	31.3 ± 3.8
m_T^{true}	86.2 ± 0.7	14.5 ± 0.6	87.0 ± 0.8	26.8 ± 1.7
m_{combi}	90.0 ± 0.8	14.9 ± 0.7	90.3 ± 0.8	26.1 ± 1.9
$\tau_\mu\tau_h$	μ (Gauss) [GeV]	σ (Gauss) [GeV]	μ (BW) [GeV]	width (BW) [GeV]
m_{vis}	59.7 ± 0.8	7.4 ± 0.4	59.2 ± 0.4	11.3 ± 0.8
m_{eff}	95.2 ± 0.8	17.4 ± 0.7	94.9 ± 0.9	33.5 ± 2.2
m_{coll}	96.1 ± 1.6	17.7 ± 1.3	94.9 ± 1.8	32.5 ± 3.9
m_{MMC}	85.8 ± 0.3	16.4 ± 0.3	85.6 ± 0.4	30.3 ± 0.9
m_{bound}	90.7 ± 1.0	16.2 ± 1.0	89.6 ± 1.1	30.2 ± 2.7
m_T^{true}	84.3 ± 0.4	11.8 ± 0.4	84.1 ± 0.5	20.0 ± 0.9
m_{combi}	86.7 ± 0.6	13.2 ± 0.6	86.5 ± 0.6	24.0 ± 1.4

Table 9.3: The shift of the mean and width of the Gauss fit function for the different masses when the energy scale is varied in $\tau_e\tau_h$ and $\tau_\mu\tau_h$ channels.

Energy Scale $[\tau_e\tau_h]$	μ [GeV]	σ [GeV]	Energy Scale $[\tau_\mu\tau_h]$	μ [GeV]	σ [GeV]
m_{eff}	3%	$^{2}_{-7}$%	m_{eff}	3%	4.5%
m_{coll}	4%	12%	m_{coll}	4%	3.5%
m_{MMC}	3%	$^{11}_{-7}$%	m_{MMC}	3%	10%
m_{bound}	$^{3}_{-9}$%	9%	m_{bound}	2%	6%
m_T^{true}	1%	4.9%	m_T^{true}	3%	4.4%
m_{combi}	$^{2}_{-5}$%	9.1%	m_{combi}	2.6%	5.7%

Table 9.4: The shift of the mean and width for the Gauss fit functions for MMC when the $E_{\text{T}}^{\text{miss}}$ resolution (METr) and the jet energy resolution (JER) are varied.

m_{MMC} $[\tau_e\tau_h]$	μ [GeV]	σ [GeV]	m_{MMC} $[\tau_\mu\tau_h]$	μ [GeV]	σ [GeV]
JER	1%	3%	JER	0.1%	0.9%
METr	0.5%	0.2%	METr	0.1%	0.7%

Table 9.5: Signal over background efficiency of the mass methods.

$\tau_e\tau_h$	s/bg	$\tau_\mu\tau_h$	s/bg
m_{vis}	3.16 ± 0.005	m_{vis}	6.37 ± 0.001
m_{eff}	3.12 ± 0.005	m_{eff}	6.20 ± 0.002
m_{coll}	3.56 ± 0.010	m_{coll}	7.09 ± 0.003
m_{MMC}	3.61 ± 0.006	m_{MMC}	6.75 ± 0.002
m_{bound}	3.47 ± 0.014	m_{bound}	7.28 ± 0.003
m_T^{true}	3.13 ± 0.005	m_T^{true}	6.11 ± 0.001
m_{combi}	3.14 ± 0.005	m_{combi}	6.26 ± 0.001

Chapter 10

Conclusions

In this thesis, the measurement of the semi-leptonic $Z \to \tau\tau$ decay was presented with the data collected in 2010 and 2011 by ATLAS. The latter set of data was also used to study $\tau\tau$ mass reconstruction methods.

Studying and understanding the $Z \to \tau\tau$ signal and, as a matter of fact, all Z and W boson decays is a pre-requisite for any searches for the Higgs boson or new physics. That is because the properties and the decays of the Z and the W bosons are known to a great accuracy and can serve as calibration tools for the data and the methods used. Especially, for Higgs boson searches, the $Z \to \tau\tau$ process is an irreducible background lying in the same mass range that a light Higgs boson is expected. Hence, before claiming discovery or exclusion of the latter, it is necessary to prove that the former is well-modelled.

The first step is the measurement of the cross section of the process. For that the semi-leptonic decays were used, where one τ lepton decays to pions and the other to either electrons or muons. The main background contribution comes from the QCD multijets, the suppression of which is achieved with an isolation requirement on the electrons and muons. Moreover, the remaining events are modelled with a two-dimensional template fit, called ABCD method. Further kinematical and identification requirements have been implemented to suppress the electroweak background events.

The final measured inclusive cross section when taking into account the branching fraction of the τ leptons to pions and electrons or muons is:

- with the data collected in 2010:

  $$\sigma(Z \to \tau\tau,\ m_{\text{inv}} : [66{-}116]\ \text{GeV}) = 857.6{\pm}84.3(\text{stat}){\pm}124.7(\text{syst}){\pm}29.2(\text{lumi}){\pm}2.8(\text{theo})\ \text{pb} \tag{10.1}$$

  for the $\tau_\mu\tau_{\text{h}}$ channel,

  $$\sigma(Z \to \tau\tau,\ m_{\text{inv}} : [66{-}116]\ \text{GeV}) = 1142{\pm}138.6(\text{stat}){\pm}197.7(\text{syst}){\pm}38.9(\text{lumi}){\pm}2.6(\text{theo})\ \text{pb} \tag{10.2}$$

  for the $\tau_e\tau_{\text{h}}$ channel.

- with the data collected in 2011:

  $$\sigma(Z \to \tau\tau,\ m_{\text{inv}} : [66{-}116]\ \text{GeV}) = 912.4{\pm}15.0(\text{stat}){\pm}94.7(\text{syst}){\pm}33.7(\text{lumi})\ \text{pb} \tag{10.3}$$

  for the $\tau_\mu\tau_{\text{h}}$ channel,

  $$\sigma(Z \to \tau\tau,\ m_{\text{inv}} :\ [66-116]\ \text{GeV}) = 998.1 \pm 23.7(\text{stat}) \pm 131.9(\text{syst}) \pm 36.9(\text{lumi})\ \text{pb} \tag{10.4}$$

  for the $\tau_e\tau_{\text{h}}$ channel.

The measurements agree within uncertainties with the theoretical NNLO expectation value, 960 $\pm$ 50 pb^{-1}.

After having established the selection criteria and methods for the $Z \to \tau\tau$ decay, the process is used to evaluate the most prominent reconstruction methods for the $\tau\tau$ system. These methods are: effective mass, collinear approximation, missing mass calculator, bound mass, "true" transverse mass and the combination of the last two. The effective and the "true" transverse mass, although they are reconstructing the full event, they are not utilising τ properties and hence, do not constrain the event. Nonetheless, they are able to reconstruct the Z resonance with 100% efficiency. Therefore, they can be useful if one is limited by statistics but has a good knowledge of the background processes. The opposite is true for the collinear approximation and the bound mass, in which case, 40-50% of the events are rejected from the phase-space constraints of the methods, but a good signal over background separation is possible. The combined method seems to overcome the disadvantages of the two separate methods, but, nonetheless, it inherits some of the disadvantages, more notably, the events reconstructed with the "true" transverse mass suffer from low background rejection. Finally, the missing mass calculator constrains the $\tau\tau$ system in almost 95% of the phase-space and thanks to the likelihood method it uses, it succeeds in providing good signal to background separation. One drawback is the sensitivity of the method on the mis-measurements of the angular coordinates between the visible decay products and the missing transverse momentum. Additionally, due to the need for initial tuning of the discriminant, the method is not flexible to any changes in the event selection.

To summarise, a first measurement of the $Z \to \tau\tau$ cross section was done with the data collected in 2010, which was then improved by the measurement in the data collected in 2011. The latter were used to further reconstruct the invariant mass of the $\tau\tau$ system and evaluate different methods. Further improvements are always possible, specially the better understanding of the detector and the tools is acquired. A larger statistical sample is not necessary for the cross section measurement, rather a better estimation of the energy scale uncertainties is necessary. In contrast, the reconstruction of the full invariant mass still suffers from statistics, so an extension of the studies to the full 2011 dataset would be beneficial. Then, it would also be possible to study and evaluate the uncertainties on the background events and hence, more concrete conclusions could be drawn on the discovery power of the methods. Nonetheless, the methods and selection criteria developed have already been used for the evaluation of systematic uncertainties on the identification and reconstruction of the hadronic τ candidates in tag-and-probe studies in data. Also, in Higgs boson searches in the $\tau\tau$ decay mode, several of the studied methods are already used and the others are currently under consideration.

Appendix A

Monte Carlo Samples

All simulated samples used in this study are listed in Table A.1 for 2010 and Tables A.2, A.3, A.4 and A.5 for 2011. In the tables are given the number of generated events, the particular dataset number and reference tags (AMI tag) for the specific ATLAS reconstruction campaign. Additionally, the NNLO cross sections and the K-factors, where applied, are listed.

Table A.1: List of Monte Carlo samples used for this analysis. All samples were generated with PYTHIA, except for $t\bar{t}$ which was generated with MC@NLO and the diboson samples generated with HERWIG.

Dataset	Dataset Number	AMI Tag	Events	Cross-Section / pb
$Z \to ee$ ($m_{\ell\ell} > 60$ GeV)	106046	e574_s933_s946_r1831_r2040	5 M	990
$Z \to \mu\mu$ ($m_{\ell\ell} > 60$ GeV)	106047	e574_s933_s946_r1831_r2040	5 M	990
$Z \to \tau\tau$ ($m_{\ell\ell} > 60$ GeV)	106052	e574_s934_s946_r1833_r2040	2 M	990
$W \to e\nu$	106043	e574_s933_s946_r1831_r2040	7 M	1046
$W \to \mu\nu$	106044	e574_s933_s946_r1831_r2040	7 M	1046
$W \to \tau\nu$	107054	e574_s934_s946_r1833_r1700	2 M	1046
$t\bar{t}$	105200	e598_s933_s946_r1831_r1700	1 M	91.50
$J1e$ (electron filter $p_{\mathrm{T}} > 8$ GeV, $\lvert\eta\rvert < 3$)	109271	e600_s934_s946_r1833_r1700	998 k	8.81×10^5
$J2e$ (electron filter $p_{\mathrm{T}} > 8$ GeV, $\lvert\eta\rvert < 3$)	109272	e574_s934_s946_r1833_r1700	497 k	2.54×10^5
$J3e$ (electron filter $p_{\mathrm{T}} > 8$ GeV, $\lvert\eta\rvert < 3$)	109273	e574_s934_s946_r1833_r1700	499 k	3.72×10^4
$J0mu$ (muon filter $p_{\mathrm{T}} > 8$ GeV, $\lvert\eta\rvert < 3$)	109276	e574_s933_s946_r1831_r1700	967 k	8.48×10^5
$J1mu$ (muon filter $p_{\mathrm{T}} > 8$ GeV, $\lvert\eta\rvert < 3$)	109277	e574_s933_s946_r1831_r1700	997 k	8.14×10^5
$J2mu$ (muon filter $p_{\mathrm{T}} > 8$ GeV, $\lvert\eta\rvert < 3$)	109278	e574_s933_s946_r1831_r1700	495 k	2.21×10^5
$J3mu$ (muon filter $p_{\mathrm{T}} > 8$ GeV, $\lvert\eta\rvert < 3$)	109279	e574_s933_s946_r1831_r1700	499 k	2.85×10^4
$\gamma^*/Z \to \tau\tau$ (10 GeV $< m_{\ell\ell} <$ 60 GeV)	107055	e574_s933_s946_r1831_r2040	190 k	396.7
$\gamma^*/Z \to ee$ (15 GeV $< m_{\ell\ell} <$ 60 GeV)	108320	e574_s933_s946_r1831_r1700	996 k	146.2
$\gamma^*/Z \to \mu\mu$ (15 GeV $< m_{\ell\ell} <$ 60 GeV)	108319	e574_s933_s946_r1831_r1700	999 k	146.2
WW	105985	e598_s933_s946_r1831_r2040	250 k	11
ZZ	105986	e598_s933_s946_r1831_r2040	250 k	1.0
WZ	105987	e598_s933_s946_r1831_r2040	250 k	3.4

Table A.2: List of Monte Carlo samples for the Z+jets production. The samples are generated with AlpGEN and are split per number of initial partons (NpX, X=0, ..., 5). A minimum parton p_T cut at 20 GeV is applied at production.

Process	Dataset Number	AMI Tag	LO Cross Section * NNLO-factor [pb]	Events
$Z \to \tau\tau$ ($m_{\ell\ell} > 40$ GeV) + Np0	107670	e844_s933_s946_r2302_r2300	668.40 * 1.25	6608784
$Z \to \tau\tau$ ($m_{\ell\ell} > 40$ GeV) + Np1	107671	e844_s933_s946_r2302_r2300	134.81 * 1.25	1327672
$Z \to \tau\tau$ ($m_{\ell\ell} > 40$ GeV) + Np2	107672	e844_s933_s946_r2302_r2300	40.36 * 1.25	403864
$Z \to \tau\tau$ ($m_{\ell\ell} > 40$ GeV) + Np3	107673	e844_s933_s946_r2302_r2300	11.25 * 1.25	109947
$Z \to \tau\tau$ ($m_{\ell\ell} > 40$ GeV) + Np4	107674	e844_s933_s946_r2302_r2300	2.79 * 1.25	29977
$Z \to \tau\tau$ ($m_{\ell\ell} > 40$ GeV) + Np5	107675	e844_s933_s946_r2302_r2300	0.77 * 1.25	9990
$Z \to ee$ ($m_{\ell\ell} > 40$ GeV) + Np0	107650	e737_s933_s946_r2302_r2300	668.32 * 1.25	6612265
$Z \to ee$ ($m_{\ell\ell} > 40$ GeV) + Np1	107651	e737_s933_s946_r2302_r2300	134.36 * 1.25	1333745
$Z \to ee$ ($m_{\ell\ell} > 40$ GeV) + Np2	107652	e737_s933_s946_r2302_r2300	40.54 * 1.25	404873
$Z \to ee$ ($m_{\ell\ell} > 40$ GeV) + Np3	107653	e737_s933_s946_r2302_r2300	11.16 * 1.25	109942
$Z \to ee$ ($m_{\ell\ell} > 40$ GeV) + Np4	107654	e737_s933_s946_r2302_r2300	2.88 * 1.25	29992
$Z \to ee$ ($m_{\ell\ell} > 40$ GeV) + Np5	107655	e737_s933_s946_r2302_r2300	0.83 * 1.25	8992
$Z \to \mu\mu$ ($m_{\ell\ell} > 40$ GeV) + Np0	107660	e737_s933_s946_r2302_r2300	668.68 * 1.25	6619010
$Z \to \mu\mu$ ($m_{\ell\ell} > 40$ GeV) + Np1	107661	e737_s933_s946_r2302_r2300	134.14 * 1.25	1334723
$Z \to \mu\mu$ ($m_{\ell\ell} > 40$ GeV) + Np2	107662	e737_s933_s946_r2302_r2300	40.33 * 1.25	403886
$Z \to \mu\mu$ ($m_{\ell\ell} > 40$ GeV) + Np3	107663	e737_s933_s946_r2302_r2300	11.19 * 1.25	109954
$Z \to \mu\mu$ ($m_{\ell\ell} > 40$ GeV) + Np4	107664	e737_s933_s946_r2302_r2300	2.75 * 1.25	29978
$Z \to \mu\mu$ ($m_{\ell\ell} > 40$ GeV) + Np5	107665	e737_s933_s946_r2302_r2300	0.77 * 1.25	9993

Table A.3: List of Monte Carlo samples for W+jets production. The samples are generated with AlpGEN and are split per number of initial partons (NpX, X=0, ..., 5). A minimum parton p_T cut at 20 GeV is applied at production.

Process	Dataset Number	AMI Tag	LO Cross Section * NNLO-factor [pb]	Events
$W \to \tau\nu$ + Np0	107700	e844_s933_s946_r2302_r2300	6918.60 * 1.20	3259564
$W \to \tau\nu$ + Np1	107701	e844_s933_s946_r2302_r2300	1303.20 * 1.20	2496467
$W \to \tau\nu$ + Np2	107702	e844_s933_s946_r2302_r2300	378.18 * 1.20	3764804
$W \to \tau\nu$ + Np3	107703	e844_s933_s946_r2302_r2300	101.51 * 1.20	1008514
$W \to \tau\nu$ + Np4	107704	e844_s933_s946_r2302_r2300	25.64 * 1.20	248864
$W \to \tau\nu$ + Np5	107705	e844_s933_s946_r2302_r2300	7.04 * 1.20	64950
$W \to e\nu$ + Np0	107680	e600_s933_s946_r2302_r2300	6921.60 * 1.20	3455037
$W \to e\nu$ + Np1	107681	e798_s933_s946_r2302_r2300	1304.30 * 1.20	2499513
$W \to e\nu$ + Np2	107682	e760_s933_s946_r2302_r2300	378.29 * 1.20	3768265
$W \to e\nu$ + Np3	107683	e760_s933_s946_r2302_r2300	101.43 * 1.20	1009641
$W \to e\nu$ + Np4	107684	e760_s933_s946_r2302_r2300	25.87 * 1.20	249869
$W \to e\nu$ + Np5	107685	e760_s933_s946_r2302_r2300	7.00 * 1.20	69953
$W \to \mu\nu$ + Np0	107690	e600_s933_s946_r2302_r2300	6919.60 * 1.20	3466523
$W \to \mu\nu$ + Np1	107691	e798_s933_s946_r2302_r2300	1304.20 * 1.20	2499513
$W \to \mu\nu$ + Np2	107692	e760_s933_s946_r2302_r2300	377.83 * 1.20	3768893
$W \to \mu\nu$ + Np3	107693	e760_s933_s946_r2302_r2300	101.88 * 1.20	1009589
$W \to \mu\nu$ + Np4	107694	e760_s933_s946_r2302_r2300	25.75 * 1.20	254879
$W \to \mu\nu$ + Np5	107695	e760_s933_s946_r2302_r2300	6.92 * 1.20	69958

Table A.4: List of Monte Carlo samples for the low mass Z+jets process. The samples are generated with AlpGEN and are split per number of initial partons (NpX, X=0, ..., 5). A minimum parton p_T cut at 20 GeV is applied at production.

Process	Dataset Number	AMI Tag	LO Cross Section * NNLO-factor [pb]	Events
$\gamma^*/Z \to \tau\tau$ (10 GeV $< m_{\ell\ell} <$ 40 GeV) + Np0	116270	e844_s933_s946_r2302_r2300	3055.1 * 1.25	959877
$\gamma^*/Z \to \tau\tau$ (10 GeV $< m_{\ell\ell} <$ 40 GeV) + Np1	116271	e844_s933_s946_r2302_r2300	84.93 * 1.25	296945
$\gamma^*/Z \to \tau\tau$ (10 GeV $< m_{\ell\ell} <$ 40 GeV) + Np2	116272	e844_s933_s946_r2302_r2300	41.47 * 1.25	498804
$\gamma^*/Z \to \tau\tau$ (10 GeV $< m_{\ell\ell} <$ 40 GeV) + Np3	116273	e844_s933_s946_r2302_r2300	8.36 * 1.25	149953
$\gamma^*/Z \to \tau\tau$ (10 GeV $< m_{\ell\ell} <$ 40 GeV) + Np4	116274	e844_s933_s946_r2302_r2300	1.85 * 1.25	39980
$\gamma^*/Z \to \tau\tau$ (10 GeV $< m_{\ell\ell} <$ 40 GeV) + Np5	116275	e844_s933_s946_r2302_r2300	0.46 * 1.25	9995
$\gamma^*/Z \to ee$ (10 GeV $< m_{\ell\ell} <$ 40 GeV) + Np0	116250	e660_s933_s946_r2302_r2300	3055.2 * 1.25	999859
$\gamma^*/Z \to ee$ (10 GeV $< m_{\ell\ell} <$ 40 GeV) + Np1	116251	e660_s933_s946_r2302_r2300	84.92 * 1.25	299940
$\gamma^*/Z \to ee$ (10 GeV $< m_{\ell\ell} <$ 40 GeV) + Np2	116252	e660_s933_s946_r2302_r2300	41.41 * 1.25	499880
$\gamma^*/Z \to ee$ (10 GeV $< m_{\ell\ell} <$ 40 GeV) + Np3	116253	e660_s933_s946_r2302_r2300	8.38 * 1.25	149940
$\gamma^*/Z \to ee$ (10 GeV $< m_{\ell\ell} <$ 40 GeV) + Np4	116254	e660_s933_s946_r2302_r2300	1.85 * 1.25	39973
$\gamma^*/Z \to ee$ (10 GeV $< m_{\ell\ell} <$ 40 GeV) + Np5	116255	e660_s933_s946_r2302_r2300	0.46 * 1.25	9995
$\gamma^*/Z \to \mu\mu$ (10 GeV $< m_{\ell\ell} <$ 40 GeV)+ Np0	116260	e660_s933_s946_r2302_r2300	3054.9 * 1.25	999869
$\gamma^*/Z \to \mu\mu$ (10 GeV $< m_{\ell\ell} <$ 40 GeV)+ Np1	116261	e660_s933_s946_r2302_r2300	84.87 * 1.25	299890
$\gamma^*/Z \to \mu\mu$ (10 GeV $< m_{\ell\ell} <$ 40 GeV)+ Np2	116262	e660_s933_s946_r2302_r2300	41.45 * 1.25	499864
$\gamma^*/Z \to \mu\mu$ (10 GeV $< m_{\ell\ell} <$ 40 GeV)+ Np3	116263	e660_s933_s946_r2302_r2300	8.38 * 1.25	149939
$\gamma^*/Z \to \mu\mu$ (10 GeV $< m_{\ell\ell} <$ 40 GeV)+ Np4	116264	e660_s933_s946_r2302_r2300	1.85 * 1.25	39988
$\gamma^*/Z \to \mu\mu$ (10 GeV $< m_{\ell\ell} <$ 40 GeV)+ Np5	116265	e660_s933_s946_r2302_r2300	0.46 * 1.25	9996

Table A.5: List of Monte Carlo samples for $t\bar{t}$ process generated with MC@NLO and for diboson production generated with HERWIG.

Process	Dataset Number	AMI Tag	Cross Section [pb]	Events
$t\bar{t}$ (no fully hadronic decays)	105200	e844_s933_s946_r2302_r2300	90.15	14845714
WW	105985	e598_s933_s946_r2302_r2300	17.02	2495756
ZZ	105986	e598_s933_s946_r2302_r2300	5.54	249906
WZ	105987	e598_s933_s946_r2302_r2300	1.26	249923

Bibliography

[1] J. Allday, *Quarks, Leptons and the Big Bang*, CRC Press, (2001), ISBN 9780750308069.

[2] The ATLAS Collaboration, G. Aad *et al.*, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys.Lett.B **716(1)** (2012) 1.

[3] The CMS Collaboration, S. Chatrchyan *et al.*, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys.Lett.B (2012), 1207.7235.

[4] The Particle Data Group Collaboration, K. Nakamura *et al.*, *Review of particle physics*, J. Phys. **G37** (2010) 075021.

[5] R. K. Ellis, W. J. Stirling, and B. R. Webber, *QCD and Collider Physics*, Cambridge Monographs on Particle Physics, Nuclear Physics, and Cosmology, Cambridge University Press, (2003), ISBN 9780521545891.

[6] D. J. Griffiths, *Introduction to Elementary Particles*, Physics Textbook, Wiley-VCH, (2008), ISBN 9783527406012.

[7] The LHC Higgs Cross Section Working Group Collaboration, S. Dittmaier *et al.*, *Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables*, 1101.0593.

[8] G. Aad *et al.*, *An update to the combined search for the Standard Model Higgs boson with the ATLAS detector at the LHC using up to 4.9 fb^{-1} of pp collision data at $\sqrt{s}$ = 7 TeV*, Tech. Rep. ATLAS-CONF-2012-019, CERN, Geneva, (Mar 2012).

[9] The ATLAS Collaboration, G. Aad *et al.*, *Combined search for the Standard Model Higgs boson using up to 4.9 fb-1 of pp collision data at sqrt(s) = 7 TeV with the ATLAS detector at the LHC*, Phys.Lett. **B710** (2012) 49, 1202.1408.

[10] The Tevatron New Physics Higgs Working Group, CDF Collaboration, D0 Collaboration Collaboration, C. Group, D. Collaborations, the Tevatron New Physics, and H. Working, *Updated Combination of CDF and D0 Searches for Standard Model Higgs Boson Production with up to 10.0 fb^{-1} of Data*, 1207.0449.

[11] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, *FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order*, Comput.Phys.Commun. **182** (2011) 2388, 1011.3540.

[12] The ATLAS Collaboration, G. Aad *et al.*, *Measurement of the $W \to l\nu$ and $Z/\gamma* \to ll$ production cross sections in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, 1010.2130, CERN-PH-EP-2010-037.

[13] L. Evans and P. Bryant, *LHC Machine*, Journal of Instrumentation **3(08)** (2008) S08001.

[14] A. Dusto, *Making Heavy Ion Beams for the LHC*, Available online, Published on 11 Nov 2011.

[15] O. S. Bruening, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, *LHC Design Report Volume I*, CERN, Geneva, (2004).

[16] M. Benedikt, P. Collier, V. Mertens, J. Poole, and K. Schindl, *LHC Design Report Volume III*, CERN, Geneva, (2004).

[17] G. Aad *et al.*, *Characterization of Interaction-Point Beam Parameters Using the pp Event-Vertex Distribution Reconstructed in the ATLAS Detector at the LHC*, Tech. Rep. ATLAS-CONF-2010-027, CERN, Geneva, (May 2010).

[18] G. Aad *et al.*, *Updated Luminosity Determination in pp Collisions at root(s)=7 TeV using the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2011-011, CERN, Geneva, (Mar 2011).

[19] G. Aad *et al.*, *Luminosity Determination in pp Collisions at sqrt(s) = 7 TeV using the ATLAS Detector in 2011*, Tech. Rep. ATLAS-CONF-2011-116, CERN, Geneva, (Aug 2011).

[20] The ATLAS Collaboration, G. Aad *et al.*, *The ATLAS Experiment at the CERN Large Hadron Collider*, Journal of Instrumentation **3(08)** (2008) S08003.

[21] G. Aad *et al.*, *ATLAS detector and physics performance: Technical Design Report, 1*, Technical Design Report ATLAS, CERN, Geneva, (1999), Electronic version not available.

[22] G. Aad *et al.*, *ATLAS muon spectrometer: Technical Design Report*, Technical Design Report ATLAS, CERN, Geneva, (1997), distribution.

[23] G. Aad *et al.*, *ATLAS level-1 trigger: Technical Design Report*, Technical Design Report ATLAS, CERN, Geneva, (1998).

[24] The ATLAS Collaboration, G. Aad *et al.*, *Performance of the ATLAS Trigger System in 2010*, The European Physical Journal C - Particles and Fields **72** (2012) 1, 1110.1530, 10.1140/epjc/s10052-011-1849-1.

[25] *Information Service page*, http://atlas-onlsw.web.cern.ch/Atlas-onlsw/components/is/.

[26] A. Dotti, P. Adragna, D. C. Cimino, M. Della Pietra, R. Ferrari, G. Gaudio, C. Roda, D. Salvatore, W. R. Vandelli, and P. F. Zema, *OHP: An Online Histogram Presenter for the ATLAS experiment*, Tech. Rep. ATL-DAQ-CONF-2006-006. ATL-COM-DAQ-2006-014, CERN, Geneva, (Mar 2006).

[27] *Online Histogramming page*, http://atlas-onlsw.web.cern.ch/Atlas-onlsw/oh/oh.htm.

[28] A. Corso-Radu, H. Havadand, Y. Ilchenko, S. Kolos, H. Okawa, K. Slagle, and A. Taffard, *Data Quality Monitoring Framework for ATLAS Experiment: Performance Achieved with Colliding Beams at LHC*, Tech. Rep. ATL-DAQ-PROC-2011-007, CERN, Geneva, (Jan 2011).

[29] *Grid Computing page*, Available online.

[30] G. Aad *et al.*, *Atlas Computing: technical design report*, CERN, Geneva, (2005).

[31] T. Maeno, *PanDA: distributed production and distributed analysis system for ATLAS*, Journal of Physics: Conference Series **119(6)** (2008) 062036.

[32] *The PanDA Production and Distributed Analysis System*, https://twiki.cern.ch/twiki/bin/viewauth/Atlas/PanDA, ATLAS internal twiki.

[33] A. Krasznahorkay, D. Berge, J. Haller, S. Ask, N. Berger, T. Eifert, and A. Hoecker, *SFrame ROOT/PROOF analysis framework*, http://sourceforge.net/apps/mediawiki/sframe/.

[34] The The GEANT4 Collaboration, S. Agostinelli *et al.*, *GEANT4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.

[35] *Experimental Particle Physicists at Imperial College London*, http://imperialhep.blogspot.de/2011/08/strangeness-at-lhcb.html, Imperial's LHCb blog.

[36] *CTEQ*, http://www.phys.psu.edu/ cteq/, CTEQ site.

[37] *MRST*, http://mstwpdf.hepforge.org/, MRST site.

[38] *Neural Network Parton Distribution Functions*, http://nnpdf.hepforge.org/, NNPDF site.

[39] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur.Phys.J. **C63** (2009) 189, `0901.0002`.

[40] M. Dobbs, S. Frixione, E. Laenen, K. Tollefson, H. Baer, *et al.*, *Les Houches guidebook to Monte Carlo generators for hadron collider physics*, `hep-ph/0403045`, Compiled by the Working Group on Quantum Chromodynamics and the Standard Model.

[41] M. E. Peskin and D. V. Schroeder, *An Introduction To Quantum Field Theory*, Advanced Book Program, Addison-Wesley Publishing Company, (1995), ISBN 9780201503975.

[42] M. L. Mangano, F. Piccinini, A. D. Polosa, M. Moretti, and R. Pittau, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, Journal of High Energy Physics **2003(07)** (2003) 001.

[43] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. Polosa, *AlpGEN*, http://mlm.home.cern.ch/mlm/alpgen/, AlpGEN homepage.

[44] G. Corcella *et al.*, *HERWIG 6.5*, JHEP **0101** (2001) 010.

[45] J. Butterworth, J. Forshaw, M. Seymour, and R. Walker, *JIMMY*, http://jimmy.hepforge.org/, JIMMY site.

[46] T. Sjostrand, S. Mrenna, and P. Skands, *PYTHIA*, http://home.thep.lu.se/ torbjorn/Pythia.html, PYTHIA site.

[47] S. Moretti, K. Odagiri, P. Richardson, M. Seymour, and B. Webber, *HERWIG*, http://hepwww.rl.ac.uk/theory/seymour/herwig/, HERWIG site.

[48] G. Aad *et al.*, *First tuning of HERWIG/JIMMY to ATLAS data*, ATL-PHYS-PUB-2010-014.

[49] S. Gieseke, A. Ribon, P. Richardson, M. Seymour, B. Webber, and P. Stephens, *HERWIG++*, http://herwig.hepforge.org/, HERWIG++ site.

[50] S. Jadach, Z. Was, R. Decker, and J. H. Kühn, *The τ decay library TAUOLA, version 2.4.*

[51] R. Decker, S. Jadach, M. Jezabek, J. Kuhn, and Z. Was, *TAUOLA*, http://wasm.home.cern.ch/wasm/goodies.html, TAUOLA site.

[52] P. Golonka and Z. Was, *PHOTOS Monte Carlo: A Precision tool for QED corrections in Z and W decays.*, Eur. Phys. J. **C45** (2006) 97.

[53] The ATLAS Collaboration, G. Aad *et al.*, *Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data*, 1110.3174.

[54] G. Aad *et al.*, *Expected electron performance in the ATLAS experiment*, Tech. Rep. ATL-PHYS-PUB-2011-006, CERN, Geneva, (Apr. 2011).

[55] G. Aad *et al.*, *Observation of $Z \to \tau_h\tau_l$ Decays with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2011-010, CERN, Geneva, (Feb 2011).

[56] G. Aad *et al.*, *$Z \to \tau\tau$ cross section measurement in proton-proton collisions at 7 TeV with the ATLAS experiment*, Tech. Rep. ATLAS-CONF-2012-006, CERN, Geneva, (Feb 2012).

[57] M. Verducci, M. Cirilli, O. Kortner, D. Orestano, F. Petrucci, J. Rothberg, N. van Eldik, Z. van Kesteren, and M. Woudstra, *Conditions Database and Calibration Software Framework for ATLAS Monitored Drift Tube Chambers*, in: *Nuclear Science Symposium Conference Record, 2006. IEEE*, vol. 1, (Nov. 2006), 2006 622 –626.

[58] T. Lagouri, D. Adams, K. Assamagan, M. Biglietti, G. Carlino, G. Cataldi, F. Conventi, A. Farilla, Y. Fisyak, S. Goldfarb, E. Gorini, K. Mair, L. Merola, A. Nairz, A. Poppleton, M. Primavera, S. Rosati, J. Shank, S. Spagnolo, L. Spogli, G. Stavropoulos, M. Verducci, and T. Wenaus, *A muon identification and combined reconstruction procedure for the ATLAS detector at the LHC at CERN*, Nuclear Science, IEEE Transactions on **51(6)** (Dec 2004) 3030 .

[59] G. Aad *et al.*, *Expected performance of the ATLAS experiment: detector, trigger and physics*, CERN, Geneva, (2009).

[60] G. Aad *et al.*, *Determination of the muon reconstruction efficiency in ATLAS at the Z resonance in proton-proton collisons at sqrt(s)=7 TeV*, Tech. Rep. ATLAS-CONF-2011-008, CERN, Geneva, (Feb 2011).

[61] O. Kortner, O. Liebig, and E. Moyse, *Muon Reconstruction Efficiency and Momentum Resolution of the ATLAS Experiment in Proton-Proton Collisions at $\sqrt{s} = 7$ TeV in 2010*, Tech. Rep. ATL-COM-PHYS-2012-069, CERN, Geneva, (Jan 2012), ATLAS internal.

[62] R. Nicolaidou, L. Chevalier, S. Hassani, J. F. Laporte, E. L. Menedeu, and A. Ouraou, *Muon identification procedure for the ATLAS detector at the LHC using Muonboy reconstruction package and tests of its performance using cosmic rays and single beam data*, Journal of Physics: Conference Series **219(3)** (2010) 032052.

[63] D. Adams, K. A. Assamagan, M. Biglietti, G. Carlino, G. Cataldi, F. Conventi, A. Farilla, Y. Fisyak, S. Goldfarb, E. Gorini, T. Lagouri, K. Mair, L. Merola, A. Nairz, A. Poppleton, M. Primavera, S. Rosati, J. T. Shank, S. Spagnolo, L. Spogli, G. D. Stavropoulos, M. Verducci, and T. Wenaus, *Track reconstruction in the ATLAS Muon Spectrometer with MOORE 007*, Tech. Rep. ATL-SOFT-2003-007, CERN, Geneva, (May 2003).

[64] S. Tarem, Z. Tarem, N. Panikashvili, and O. Belkind, *MuGirl – Muon identification in the ATLAS detector from the inside out*, in: *Nuclear Science Symposium Conference Record, 2006. IEEE*, vol. 1, (Nov 2006), 2006 617 –621.

[65] O. Kortner, E. Moyse, and A. Salvucci, *Muon resolution plots*, Tech. Rep. ATL-COM-PHYS-2011-1504, CERN, Geneva, (Nov 2011).

[66] The ATLAS Collaboration, G. Aad *et al.*, *Jet energy measurement with the ATLAS detector in proton-proton collisions at sqrt(s) = 7 TeV*, `1112.6426`.

[67] G. Aad *et al.*, *Properties of Jets and Inputs to Jet Reconstruction and Calibration with the ATLAS Detector Using Proton-Proton Collisions at* $\sqrt{s} = 7$ *TeV*, Tech. Rep. ATLAS-CONF-2010-053, CERN, Geneva, (Jul 2010).

[68] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti-k(t) jet clustering algorithm*, JHEP **0804** (2008) 063, `0802.1189`.

[69] G. Aad *et al.*, *Data-Quality Requirements and Event Cleaning for Jets and Missing Transverse Energy Reconstruction with the ATLAS Detector in Proton-Proton Collisions at a Center-of-Mass Energy of sqrts* $= 7$ *TeV*, Tech. Rep. ATLAS-CONF-2010-038, CERN, Geneva, (Jul 2010).

[70] G. Aad *et al.*, *Validating the measurement of jet energies with the ATLAS detector using Z + jet events from proton-proton collisions at sqrt(s) = 7 TeV*, Tech. Rep. ATLAS-CONF-2011-159, CERN, Geneva, (Dec 2011).

[71] G. Aad *et al.*, *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 sqrts = 7 TeV*, Tech. Rep. ATLAS-CONF-2010-054, CERN, Geneva, (Jul 2010).

[72] The ATLAS Collaboration, G. Aad *et al.*, *Performance of Missing Transverse Momentum Reconstruction in Proton-Proton Collisions at 7 TeV with ATLAS*, Eur.Phys.J. **C72** (2012) 1844, `1108.5602`.

[73] G. Aad *et al.*, *Reconstruction and Calibration of Missing Transverse Energy and Performance in Z and W events in ATLAS Proton-Proton Collisions at 7 TeV*, Tech. Rep. ATLAS-CONF-2011-080, CERN, Geneva, (Jun 2011).

[74] M. Cacciari and G. P. Salam, *Dispelling the* N^3 *myth for the* k_t *jet-finder*, Phys.Lett. **B641** (2006) 57, `hep-ph/0512210`.

[75] W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, S. Menke, S. Rajagopalan, D. Rousseau, S. Snyder, and G. Unal, *Calorimeter Clustering Algorithms: Description and Performance*, Tech. Rep. ATL-LARG-PUB-2008-002, ATL-COM-LARG-2008-003, CERN, Geneva, (Apr 2008).

[76] T. Barillari, E. Bergeaas Kuutmann, T. Carli, J. Erdmann, P. Giovannini, K. J. Grahn, C. Issever, A. Jantsch, A. Kiryunin, K. Lohwasser, A. Maslennikov, S. Menke, H. Oberlack, G. Pospelov, E. Rauter, P. Schacht, F. Spano, P. Speckmayer, P. Stavina, and P. Strizenec, *Local Hadronic Calibration*, Tech. Rep. ATL-LARG-PUB-2009-001-2, ATL-COM-LARG-2008-006, CERN, Geneva, (Jun 2008).

[77] G. Aad *et al.*, *Performance of the Reconstruction and Identification of Hadronic Tau Decays with ATLAS*, Tech. Rep. ATLAS-CONF-2011-152, CERN, Geneva, (Nov 2011).

[78] G. Aad *et al.*, *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons*, Tech. Rep. ATLAS-CONF-2011-077, CERN, Geneva, (May 2011).

[79] G. Aad *et al.*, *Measurement of the Mis-identification Probability of Hadronically Decaying Tau Leptons from Hadronic Jets and from Electrons*, Tech. Rep. ATLAS-CONF-2011-113, CERN, Geneva, (Aug 2011).

[80] T. Sjostrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026.

[81] T. Plehn, *Lectures on LHC Physics*, Lect.Notes Phys. **844** (2012) 1, `0910.4182`.

[82] J. Butterworth *et al.*, *Single Boson and Diboson Production Cross Sections in pp Collisions at* $\sqrt{s} = 7$ *TeV*, ATL-COM-PHYS-2010-695.

[83] S. Frixione and B. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP **0206** (2002) 029, `hep-ph/0204244`.

[84] *Luminosity Calculation Tool*, https://atlas-datasummary.cern.ch/lumicalc.

[85] *Extended Pileup Reweighting*, https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/ExtendedPileupReweighting, ATLAS only page.

[86] G. Aad *et al.*, *Performance of primary vertex reconstruction in proton-proton collisions at* $\sqrt{s}$ *=7 TeV in the ATLAS experiment*, Tech. Rep. ATLAS-CONF-2010-069, CERN, Geneva, (Jul 2010).

[87] *LAr hole cleaning*, https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/LArCleaningAndObjectQuality, ATLAS internal twiki.

[88] G. Aad *et al.*, *Muon reconstruction efficiency in reprocessed 2010 LHC proton-proton collision data recorded with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2011-063, CERN, Geneva, (Apr 2011).

[89] G. Aad *et al.*, *In-situ pseudo-rapidity inter-calibration to evaluate jet energy scale uncertainty and calorimeter performance in the forward region*, Tech. Rep. ATLAS-CONF-2010-055, CERN, Geneva, (Jul 2010).

[90] G. Aad *et al.*, *Measurement of jet production in proton-proton collisions at 7 TeV centre-of-mass energy with the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2010-050, CERN, Geneva, (Jul 2010).

[91] A. Sherstnev and R. S. Thorne, *Parton Distributions for LO Generators*, Eur. Phys. J. **C55** (2008) 553.

[92] G. Aad *et al.*, *Charged particle multiplicities in pp interactions at $\sqrt{s}$ = 0.9 and 7 TeV in a difractive limited phase-space measured with the ATLAS detector at the LHC and new PYTHIA6 tune*, Tech. Rep. ATLAS-CONF-2010-031, CERN, Geneva, (Jul 2010).

[93] The ATLAS Collaboration, G. Aad *et al.*, *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, 0901.0512.

[94] The ATLAS Collaboration, A. Elagin, P. Murat, A. Pranko, and A. Safonov, *A New Mass Reconstruction Technique for Resonances Decaying to di-tau*, 1012.4686.

[95] G. Aad *et al.*, *Search for Neutral MSSM Higgs Bosons decaying to $\tau^+\tau^-$ pairs in proton-proton collisions at $\sqrt{7}$ TeV with the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2011-132, CERN, Geneva, (Aug 2011).

[96] A. J. Barr, S. T. French, J. A. Frost, and C. G. Lester, *Speedy Higgs boson discovery in decays to tau lepton pairs: $h \to \tau\tau$*, JHEP **10** (2011) 080, 1106.2322.

[97] A. J. Barr, T. J. Khoo, P. Konar, K. Kong, C. G. Lester, *et al.*, *Guide to transverse projections and mass-constraining variables*, 1105.2977, review of transverse masses.

[98] A. Barr, C. Lester, and P. Stephens, *m(T2) : The Truth behind the glamour*, J. Phys. **G29** (2003) 2343, hep-ph/0304226.

[99] A. J. Barr, B. Gripaios, and C. G. Lester, *Measuring the Higgs boson mass in dileptonic W-boson decays at hadron colliders*, JHEP **0907** (2009) 072, 0902.4864.

Resumé

Contact Information

Despoina Evangelakou
Nationality: Hellenic
Place of birth: Athens, Greece
Office Address: Friedrich-Hund-Platz 1, 37075 Göttingen, DE
cellular phone: +49-(0)1769 99 20 482
e-mail: devangel@gwdg.de

Education

Doctoral Thesis in Particle Physics, Georg-August Universität, Göttingen, DE **2008-2012**

- *Dissertation Topic*: "$Z \to \tau\tau$ cross section measurement and $\tau\tau$ Mass Reconstruction with the ATLAS detector at the LHC"
- *Promotor*: Prof. Dr. Arnulf Quadt *Co-promotor*: Prof. Dr. Ariane Frey

M.Sc. in Particle and Astroparticle Physics, Universiteit van Amsterdam, Amsterdam, NL **2006-2008**

- *Dissertation Topic*: "Particle Identification studies with the TPC and TOF detectors of the ALICE experiment. Application in the Reconstruction of the ϕ meson"
- *Advisor*: Prof. Dr. Raimond Snellings

Diploma in Applied Mathematics and Physics, National Technical University of Athens, Athens, GR **1999-2006**

- *Dissertation Topic*: "Unimodular Gravity"
- *Advisor*: Assistant Prof. Dr. Alexandros Kehagias

Mathematics Department, University of Ioannina, Ioannina, GR **1998-1999**

High-School Diploma, 3rd High-school of Peristeri, GR **1998**

Publications

(Member of the ATLAS author list since October 2010)

- "Observation of $Z \to \tau_l\tau_h$ Decays with the ATLAS Detector", ATLAS-CONF-2011-010
- "Measurement of $Z \to \tau\tau$ production cross section in proton-proton collisions at $\sqrt{7}$ TeV with the ATLAS detector", Phys. Rev. D 84, 112006 (2011)
- "$Z \to \tau^+\tau^-$ cross section measurement in proton-proton collisions at 7 TeV with the ATLAS experiment", ATLAS-CONF-2012-006

Selected Talks

- "Comparison of the $\tau\tau$ mass reconstruction methods in ATLAS",
  $m \to \tau\tau$ *workshop*, Freiburg, DE (Apr 2012)
- "Comparison of the $\tau\tau$ mass reconstruction methods in ATLAS",
  Deutsche Physikalische Gesellschaft Tagung 2012, Göttingen, DE (Feb 2012)
- "Overview of the mass reconstruction techniques for the diTau system",
  $m \to \tau\tau$ *workshop*, Göttingen, DE (Sept 2011)
- "Z+jets Cross Section Measurement",
  ATLAS-D meeting, DESY Hamburg, DE (Nov 2009)

Physics Outreach

The Real Physics Behind The Angels and Demons Book and Movie **May 2009**

Guided Tours for High-School Students at CERN (in Greek) **2010**

Expand Your Horizons (in french and English) **Nov 2011**
Small Physics Experiments for High-School students in the region of Geneva (CH) with purpose the encouragement of women's participation in sciences.

Rent A Scientist (in German) **May 2012**
I helped introducing high-school students to particle physics and LHC using the program "Hands on Particle Physics Masterclasses, ATLAS"

Acknowledgements

I don't think of my PhD time only as a time for learning but also, and some times more, as an opportunity to get to know people and myself better. The reason I 'm saying that is because, during these last three years or, actually, the six years I 've been away from *home*, first in Amsterdam and then in Göttingen and Geneva, I got to meet and interact with lots of people. All of them showed me new ways to think, through their different attitude towards life and work than mine. So, at this point I want to try to remember the most important ones and in a way say 'Thank you!' for having changed me.

First, I need to thank my 'boss' Arnulf. You were there for me when Markus left three and half years ago, and not one moment did you make me feel unwanted in the institute. Later (and every time), when I was disappointed and stressed you believed in me and encouraged me to continue with my work and the PhD. It turned out that you were right and I did manage to finish it! Vielen Dank für deine Vertrauen und Hilfe.

Many thanks to Ariane Frey for being my co-promotor, reading my thesis and helping me to 'keep things simple' when I was preparing for the final exam.

I can't thank enough all my colleagues in the ATLAS $Z \to \tau\tau$ and $H \to \tau\tau$ groups. I want to specially mention Susanne Kühn, Justin Griffiths, Anna Kaczmarska and Stan Lai you helped me a lot when I was starting with the analysis and you gave me a chance to work with the team and learn a lot. A big thanks and congrats to Frank Seifert, Daniele Capriotti, Jana Novakova and Elias Coniavitis for all the hard work we did together, all the support, the constructive discussions, the detailed emails any time of the day, any day of the week till we succeeded with the publication.

Apart from the work part there are all these 'other' hours that friends and colleagues helped to go by easier and in a more pleasant way. First, to my friends in Amsterdam, *Inci, Eleni, Menelaos*, I 've to say that I love you all. You 've been there for me every time I freaked out and I needed a way out in the city or a friend's hug and to be pampered.

Magda and Barbara, you are amazing and my very best friends. You 've been there for me in so many occasions. Many of our trips are memorable, but here I need to mention only one; the spa-weekend in Evian a bit before the stress started for us. UNFORGET-TABLE!!!!! Magda, I need to thank you specially for finding time to proof-read my thesis in the mayhem of moving and publishing. Efharisto!!!!!! bardzo dziekuje!!!!!!

In Göttingen, there are so many people that played their own special role at one point or another; *Matthias G., Jörn, Nina, Anna, Stefan, Jens, Fabian, Olaf, Boris, Beni, Shabnaz* for coffee breaks at Uni-Nord, cinema nights, whisky and Eurovision nights and so many other fun moments; *Eric and Henrike* you were very patient and cool office-mates; *Erik* thanks tons for correcting my english, explaining me the theory and all this without complaining; *Adam* efharisto for supporting me and helping me to adapt in goe and in the PhD life, for all the adventures and, finally, for unconsciously helping me choose a

career; *Ulla*, danke for showing me how to stand on my own feet and 'take care of myself' at a professional level; *Trudi, Karin and Lucie*, thanks for the best welcome in Göttingen I could have wished for and for your good words and warm hugs whenever I needed them.

In Geneva, the crew and the distractions were different: *Leticia and Bruno, Olivier and Caroline, Benjamin, Julian, summies 2006.* Lots coffees at R1 or drinks by the lakes or biking through the french-swiss vignards for wine tasting, talking about physics, whining about our supervisors or dancing. It's been great!

Finally, my last year in Göttingen I gained a bunch of very good friends which helped me pull through at this very difficult period: *Anna and Oleg, Gabi, Natassa, Tamara, Leonid and Kostas. Oleg*, I owe you big time for the last minute corrections and the glass of wine, that night... you know which one! *Gab's*, without your late night visits at the office, bike rides, coffees and so much more, the last year would have been miserable. *Annoula*, I 'm running out of words. You supported me, encouraged me, pushed me to work when I had to and told me to let go when I should. Efharisto! And not to forget: "Guys, hold on, we 'll make it out of here!"

Ooops! I almost forgot. No acknowledgements are complete without a thanks to the family. Μανουλα, Τέρας, Μπιφ, δε θα τα κατάφερνα ποτέ χωρίς την αγάπη και την υποστήριξή σας. Ευχαριστώ!