



HAL
open science

Search for $t\bar{t}$ resonances using the ATLAS detector

Silvestre Marino Romano Saez

► **To cite this version:**

Silvestre Marino Romano Saez. Search for $t\bar{t}$ resonances using the ATLAS detector. Other [cond-mat.other]. Université Blaise Pascal - Clermont-Ferrand II, 2016. English. NNT : 2016CLF22736 . tel-01471591

HAL Id: tel-01471591

<https://theses.hal.science/tel-01471591>

Submitted on 20 Feb 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Numéro d'ordre: DU 2736
EDSF: 879

PCCF T 1608

Université Blaise Pascal de Clermont-Ferrand
U.F.R. Science et Technologies

**ÉCOLE DOCTORALE DES SCIENCES
FONDAMENTALES**

THÈSE

pour obtenir le grade de

**DOCTEUR D'UNIVERSITÉ
SPECIALITÉ: PHYSIQUE DES PARTICULES**

Présentée par
Silvestre Marino ROMANO SAEZ

Search for $t\bar{t}$ resonances using the ATLAS detector.

**Thèse soutenue publiquement le 11 Octobre 2016
devant la commission d'examen:**

Président:	Jean	ORLOFF
Rapporteurs:	Jeremy	ANDREA
	Frederic	DERUE
Examineur:	Tomas	DAVIDEK
Directeurs:	Dominique	PALLIN
	Samuel	CALVET

Abstract

Massive particles decaying into a top-antitop pair ($t\bar{t}$) are predicted by many theoretical models, which are introduced to provide explanations to the various open questions raised by the current formulation of the Standard Model of Particle Physics (SM). A search for new $t\bar{t}$ resonances is presented in this manuscript, using the 2015 dataset (early Run 2) from the proton-proton collisions at Large Hadron Collider (LHC) with a beam energy of 13 TeV at the centre-of-mass and with an integrated luminosity of 3.2 fb^{-1} collected by the ATLAS detector. The final state signature focuses on the lepton-plus-jets channel, which is characterised by the presence of an electron or muon, certain number of jets of hadrons and missing transverse energy. The invariant mass of the $t\bar{t}$ system ($m_{t\bar{t}}$) is the main observable on this search, which is used to test the compatibility of the data with the SM-only hypothesis. For the high $m_{t\bar{t}}$ region, the selected events are dominated by top-quarks with collimated decay products, and jet substructure techniques have to be used to select the $t\bar{t}$ pairs (boosted scenario), while for low $m_{t\bar{t}}$, the selected events are dominated by top-quarks with well separated decay products (resolved scenario).

Preliminary results on the boosted analysis are presented in this manuscript. In the analysed mass spectra, no evidence of the existence of new particles was found. For a topcolor-assisted technicolor Z'_{TC2} boson with a relative width of 1.2%, masses below 2 TeV are excluded, improving the previous limits obtained at Run 1. Perspective studies to include the resolved topologies are also discussed, where the estimation of the QCD multi-jet background using the “matrix method” is exposed in detail. In addition, the in-situ “jet rescaling” method is proposed to improve the $m_{t\bar{t}}$ resolution. The impact after each step of the rescaling procedure is presented using 8 TeV simulations.

Resumé

Le Modèle Standard (MS) de la physique des particules décrit les particules élémentaires et leurs interactions à l'exception de la gravité. Jusqu'à présent, la plupart des résultats expérimentaux sont en accord avec les prédictions du MS. Cependant, il existe encore des questions de nature expérimentale ou théorique qui restent sans réponse. Les particules massives qui se désintègrent en une paire de quarks top-antitop sont prédites par certains des modèles de physique au-delà du Modèle Standard, qui tentent de répondre à ces questions.

Une recherche de nouvelles résonances $t\bar{t}$ est présentée dans ce manuscrit, en utilisant l'ensemble de données 2015 (début des Run 2) à partir des collisions proton-proton produites par le *Large Hadron Collider* (LHC) avec une énergie de faisceau de 13 TeV dans le centre de masse et avec une luminosité intégrée de 3.2 fb^{-1} recueillie par le détecteur ATLAS. La signature de l'état final recherché est caractérisée par la présence d'un électron ou d'un muon, d'un certain nombre de jets de hadrons et de l'énergie transverse manquante. La masse invariante des système $t\bar{t}$ ($m_{t\bar{t}}^{reco}$) est la principale observable sur cette recherche, et est utilisée pour tester la compatibilité des données avec l'hypothèse MS seule. Pour la région à grand $m_{t\bar{t}}$, les quarks top sont produits avec un grande impulsion et leurs produits de désintégration sont collimatés. Dans cette région, la sous-structure des jets doit être utilisée pour identifier les paires $t\bar{t}$ (scénario "boosted"), tandis que pour de basses valeurs de $m_{t\bar{t}}$, les produits de désintégration des quark top sont bien séparés (scénario "resolved").

Les résultats préliminaires sur l'analyse "boosted" sont présentés dans ce manuscrit. Dans les spectres de masse analysés, aucune preuve de l'existence de nouvelles particules n'a été trouvée. Pour une boson Z'_{TC2} "Topcolor assisted technicolor" avec une largeur relative de 1.2 %, des masses inférieures à 2 TeV sont exclues, l'améliorant les limites antérieures obtenues lors du Run 1. Des études de perspective pour inclure les topologies "resolved" sont également discutées, avec en particulier l'estimation du bruit de fond multi-jet en utilisant la "méthode de la matrice". En outre, la méthode d'étalonnage des jets in-situ est proposée pour améliorer la résolution sur $m_{t\bar{t}}$. L'impact après chaque étape de la procédure de mise à l'échelle est présentée en utilisant des simulations à 8 TeV.

Contents

1	Theoretical context	13
1.1	The Standard Model of particle physics	14
1.1.1	Theoretical formulation	15
1.1.2	SM overview, successes and weaknesses	21
1.2	Beyond the Standard Model	25
1.2.1	Topcolor assisted technicolor	25
1.2.2	Extra-dimensions	26
1.2.3	Two Higgs Doublet Model motivation	26
1.3	Top-quark physics at the LHC	27
1.3.1	Top-quark production modes	28
1.3.2	Top-quark decay	29
1.3.3	$t\bar{t}$ signatures	29
1.4	Searching for $t\bar{t}$ resonances	31
1.4.1	Outlook	32
2	Experimental context	35
2.1	The Large Hadron Collider	36
2.1.1	Proton acceleration chain	37
2.1.2	Proton collisions at LHC	38
2.2	The ATLAS detector	40
2.2.1	ATLAS coordinate system	41
2.2.2	Magnet system	42
2.2.3	Inner detector	42
2.2.4	Calorimeters	45
2.2.5	Muon spectrometer	48
2.2.6	Forward detectors	49
2.2.7	Data acquisition system	50
2.2.8	Computing model	53
2.2.9	Data quality in 2015 data-taking	54
2.3	Outlook	55

3	The Tile Calorimeter and calibration by laser in-time runs	57
3.1	Tile calorimeter overview	58
3.2	TileCal calibration systems	61
3.3	The Laser system	63
3.3.1	Laser runs	64
3.3.2	Laser calibration	65
3.3.3	Statistical dependence of C_{laser} for laser in-time runs	66
3.3.4	Compatibility between C_{las} and C_{Cesium}	68
3.3.5	Source of systematic uncertainty of the C_{laser}	69
3.4	Statistical uncertainty studies using pseudo-data	72
3.4.1	Generation of the $R_{i,p}^{MC}$ distribution	73
3.4.2	The statistical dependence of the laser precision	74
3.4.3	Emulation of the Cesium scans	75
	Outlook	76
4	pp collisions: simulation and physic object identification	79
4.1	Simulation of the proton-proton collisions	81
4.1.1	Event simulation	81
4.2	Track and vertex reconstructions	84
4.3	Electrons	85
4.3.1	Electron reconstruction	85
4.3.2	Electron identification	86
4.4	Muon	87
4.4.1	Muon reconstruction	87
4.4.2	Muon identification	88
4.5	Lepton isolation and correction	89
4.5.1	Lepton isolation	89
4.6	Jets of hadrons	91
4.6.1	Jet reconstruction algorithms	92
4.6.2	Jet calibration	93
4.6.3	Track-based pile-up jets suppression	94
4.6.4	Large-R jets and Top-tagging	96
4.6.5	b-tagging algorithm	97
4.7	Overlap removal	99
4.8	Missing transverse energy	100
4.9	Outlook	101
5	Search for $t\bar{t}$ resonances in boosted regimes	103
5.1	Strategy	104
5.2	Event selection	105
5.2.1	$t\bar{t}$ topologies	105

CONTENTS

5.2.2	Event pre-selection	106
5.2.3	Boosted selection	107
5.3	Invariant $t\bar{t}$ mass reconstruction	108
5.4	Signal simulation	110
5.5	Background processes	110
5.6	Data driven backgrounds	112
5.6.1	Non-prompt leptons	112
5.6.2	\mathbf{W} +jets background normalisation	117
5.7	Systematic Uncertainties	117
5.8	Control plots	119
5.9	Compatibility with SM only hypothesis	119
5.10	Upper production cross section limits on $t\bar{t}$ resonances	126
6	Improving the sensitivity at low $m_{t\bar{t}}$	131
6.1	Resolved selection	132
6.1.1	Low mass topology	132
6.1.2	Boosted and resolved combination	133
6.2	Estimation of the QCD multi-jet background	133
6.2.1	Requirements for the QCD multi-jet estimation	134
6.2.2	Changes with respect the previous QCD multi-jet estimation	134
6.2.3	Estimation of the real and fake rates	135
6.2.4	Performance of the QCD multi-jet background estimation	140
6.3	Reconstruction of the invariant mass of the $t\bar{t}$ system	147
6.3.1	χ^2 minimisation algorithm	147
6.4	Improving the $t\bar{t}$ invariant mass resolution	149
6.4.1	Impact of the neutrino reconstruction	149
6.4.2	Jet rescaling method	150
6.4.3	Impact of the rescaling	155

CONTENTS

Introduction

The Standard model of particle physics (SM) introduces a mathematical formalism capable to describe three of the known fundamental interactions. This theory has been under construction since last half century by many of the greatest names in physics, with a successful predictive power. Nevertheless, the SM does not provide a complete picture of the modern understanding of the fundamental interactions, and some pieces of theoretical and experimental orders are still missing. The problem of the un-natural low mass of the Higgs boson, or the role of the top-quark (the heaviest one) in the electro-weak symmetry breaking, have been reinforced by the discovery of the Higgs boson in 2012. The so-called Beyond-the-SM (BSM) theories, are developed to propose explanations of the SM open questions, introducing new symmetries, interactions, or extra-dimensions. The top-quark may be directly linked to new physics processes because of its high mass. Some BSM theories predict new heavy particles decaying into a $t\bar{t}$ resonance if they are heavy enough. The validity or the refutation of these BSM theories can be tested at high energy experiments, such as the Large Hadron Collider (LHC) at CERN.

The analysis exposed in this manuscript develops a model independent method sensitive to the invariant mass of $t\bar{t}$ events produced at LHC. The $t\bar{t}$ events are obtained from the reconstructed physics objects produced at the proton-proton collision of LHC and measured using the ATLAS detector. The $t\bar{t}$ invariant mass is reconstructed with different algorithms implemented on the kinematic regime of the top-quarks. A statistical test allows to determine if the data agrees with the SM prediction, in which case an upper limit is set on the production cross-section of several BSM signal benchmarks. Systematic uncertainties associated to the reconstruction of the physics object and the background modelling, are taken into account in the limit setting procedure.

The Chapter 1 present a brief introduction to the SM formulation and the motivations of the BSM theories which are used as benchmarks in the search. A description of CERN apparatus is presented in Chapter 2, covering the LHC accelerator chain and the ATLAS detector. The Chapter 3 presents the results of a study testing the feasibility of the laser in-time runs to perform the calibration of energy measurements of the Tile calorimeter. In Chapter 4, several techniques to reconstruct and identify the physics objects produced in the LHC collisions are presented. The algorithms and methods used for the identification of the particles coming from the decay products

CONTENTS

of the top-quark are exposed. The results of the analysis searching for $t\bar{t}$ resonances in the boosted kinematic regime is exposed in Chapter 5. This analysis uses 3.2 fb^{-1} of collision data recorded by the ATLAS detector at early Run 2, where the proton beam energy reaches 13 TeV at the centre-of-mass (2015 data-taking). In this manuscript, preliminary results using the 2015 datasets are presented in the boosted kinematic regime, which has been approved for Moriond 2016 conference. In addition, an improvement of the sensitivity at the low mass region could lead to more accurate cross-section limits. A perspective study is presented to search for $t\bar{t}$ resonances in the low mass region in Chapter 6.

Chapter 1

Theoretical contex

Contents

1.1	The Standard Model of particle physics	14
1.1.1	Theoretical formulation	15
1.1.2	SM overview, successes and weaknesses	21
1.2	Beyond the Standard Model	25
1.2.1	Topcolor assisted technicolor	25
1.2.2	Extra-dimensions	26
1.2.3	Two Higgs Doublet Model motivation	26
1.3	Top-quark physics at the LHC	27
1.3.1	Top-quark production modes	28
1.3.2	Top-quark decay	29
1.3.3	$t\bar{t}$ signatures	29
1.4	Searching for $t\bar{t}$ resonances	31
1.4.1	Outlook	32

The best validated model of particle physics so far is the Standard Model (SM), developed in the context of a quantum field theory. The SM provides a beautiful theoretical framework which is able to accommodate the present knowledge on electro-weak and strong interactions. But despite its phenomenological success, the SM leaves many un-answered questions to solve. Beyond the Standard Model (BSM) theories are devoted to providing answers.

A brief overview of the SM formulation is presented in this chapter, followed by an overview of the BSM theories used as benchmarks for the search for $t\bar{t}$ resonances. In addition, the phenomenology of the top-quark physics is exposed.

1.1 The Standard Model of particle physics

The SM [1, 2, 3, 4, 5] aims to describe the known fundamental interactions between the elementary particles, which are classified into fermions and bosons. The fermions are particles with half-integer spin, following the Fermi-Dirac statistics, while the bosons are particles with an integer spin, following the Bose-Einstein statistics. According to the Pauli exclusion principle, only one fermion can be in a particular quantum state. Under the SM framework, the fermions are usually the constituents of matter while the bosons are the particles that transmit interactions (force carriers), or the constituents of radiation.

The fermions are divided in three families or generations (see Figure 1.1). The stable matter is built from fermions of the first generation, and the second and third generations are composed by short-lived particles which can be observed in high energy interactions, decaying into first generation particles. These fermionic families differ only by their mass and their flavour quantum number.

Each family is classified in quarks and leptons. The leptons are sensitive to the weak and electromagnetic interactions, and are subdivided into an electrically charged and electrically neutral leptons. The charged leptons are conformed by the electron, muon, and tau, while the neutral leptons are the associated neutrinos. The neutrinos weakly interact with matter; they oscillate between generations due to the non-zero neutrino mass and neutrino mixing [6]. The quarks are sensitive to the three interactions. They are divided into up- and down-type quarks depending on their fractional electrical charges. The up-type includes the up-, charm-, and top-quarks, while the down-type includes the down-, strange- and bottom-quarks. The colour charge, associated to the strong interaction, is an important property of the quarks being by convention: red, green or blue. The quarks are bound in combinations called hadrons, where three bounded quarks are named “baryons” and quark-antiquark pair are named “mesons”.

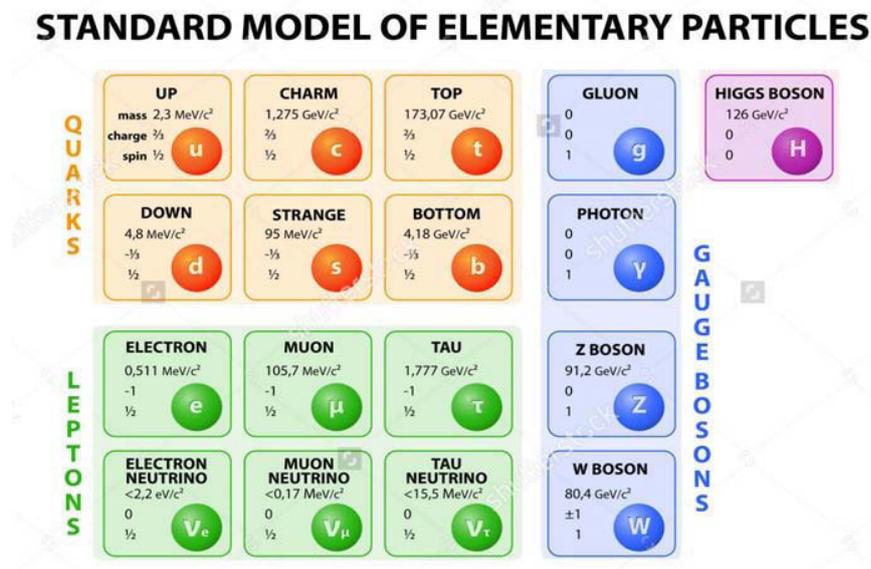


Figure 1.1: Scheme of the elementary particles in the SM of particle physics.

In addition, there is an antiparticle associated to each of the particles mentioned. They carry the same mass than the corresponding particle but with opposite quantum numbers.

The SM interactions are described via the exchange of spin-1 gauge fields. The force carrier of the electro-magnetic interaction is the photon γ , a massive and not self-interacting gauge boson. The weak interaction has three massive gauge mediators called Z^0 and W^\pm bosons with a short range of interaction. Finally, the strong interaction is mediated by eight massless gluons which carry the two color charges.

The Higgs boson is the latest SM particle found [7, 8, 9, 10, 11], which is a remanent of the Spontaneous Symmetry Breaking (SSB) mechanism of the electro-weak interaction that provides a mass to the SM particles. A Higgs boson particle was discovered in July 2012 by ATLAS and CMS Collaborations.

1.1.1 Theoretical formulation

In the context of a quantum field theory, the SM lagrangian density is invariant under a symmetry group of local gauge transformations. The symmetry group is $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$, where $SU(3)_C$ is associated to the strong interaction and $SU(2)_L \otimes U(1)_Y$ to the electro-weak interaction. The C denotes color charge, L the left-handed fields, and Y the hypercharge. The generators of

the symmetry group correspond to the force carriers.

Details about the formulation of the involved theories are summarised in the following sections.

The electro-magnetic interaction

Quantum electrodynamics (QED) is the relativistic quantum field theory which describes the electro-magnetic phenomena based on the local gauge symmetry U(1) [12, 13]. The QED lagrangian density describes the coupling between the charged fermionic fields (Ψ) to the gauge boson A_μ and their corresponding kinematics:

$$\mathcal{L}_{QED} = \bar{\Psi}(i\gamma^\mu D_\mu - m)\Psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \quad (1.1)$$

where γ^μ denotes the Dirac matrices, m is the mass of the fermion and D_μ is the covariant derivate and the electro-magnetic field tensor $F_{\mu\nu}$ which are defined as:

$$D_\mu = \partial_\mu - ieA_\mu, \quad (1.2)$$

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \quad (1.3)$$

where e is the coupling constant (equal to the fundamental electric charge). To ensure the gauge invariance of the lagrangian density under a local phase change of Ψ ($\Psi(x) \rightarrow e^{i\theta(x)}\Psi(x)$), the vector field (A_μ) associated to the photon should transform as:

$$A_\mu(x) \xrightarrow{U(1)} A_\mu(x) - e\partial_\mu\theta(x). \quad (1.4)$$

Notice that a mass term for the photon ($m^2 A_\mu A^\mu$) is forbidden to preserve the gauge invariance.

The gauge invariance leads to a local conservation of the electro-magnetic charge, which is usually written in terms of the dimensionless ratio α_{QED} , known as the fine structure constant ($\alpha_{QED} = \frac{e^2}{4\pi}$)¹. α_{QED} is a fundamental parameter of the theory which represent the strength of the electro-magnetic interaction. In perturbation theory, the virtual particle corrections to the propagators can diverge. The procedure to remove such divergences in the Feynman diagrams calculations is denominated as renormalisation, which re-defines the measurable observables at a given energy scale (μ_R) to include the virtual particle corrections. The value of α_{QED} grows logarithmically with the energy scale, going from 1/137 at low energies, to 1/127 for energies at the GeV order.

¹In units where $\hbar = c = 1$.

Electro-weak theory

The Electro-weak theory is the unification of the electro-magnetic and weak interactions, where $SU(2)_L \otimes U(1)_Y$ is the gauge symmetry group of the interaction [14]. The local invariance leads to four gauge vector bosons W_μ^i ($i = 1, 2, 3$) from $SU(2)$ and B_μ for $U(1)$. The W^\pm , Z^0 bosons and the photon A_μ are mixture of these gauge vector fields.

The electro-weak theory is a chiral theory, where only left-handed fermions interact with the weak interaction. Therefore, the fermions are organised in invariant $SU(2)_L$ doublet fields Ψ_L (left-handed) and $SU(2)_L$ singlet fields Ψ_R (right-handed):

$$\Psi_L = P_L \Psi = \frac{1}{2}(1 - \gamma^5)\Psi, \quad (1.5)$$

$$\Psi_R = P_R \Psi = \frac{1}{2}(1 + \gamma^5)\Psi, \quad (1.6)$$

where $P_{L,R}$ are the chirality operators and $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$ is a product of the Dirac matrices. The Ψ_L and Ψ_R structure for the first generation is:

$$\Psi_L : \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \begin{pmatrix} u \\ d \end{pmatrix}_L ; \quad (1.7)$$

$$\Psi_R : e_R, u_R, d_R, \quad (1.8)$$

and the other generations follows the same scheme. To describe weak interactions, several fermionic flavors and different properties for left- and right-handed fields are taken into account on the electro-weak lagrangian density:

$$\mathcal{L}_{EW} = \bar{\Psi}_L(i\gamma^\mu D_\mu)\Psi_L + \bar{\Psi}_R(i\gamma^\mu D_\mu)\Psi_R - \frac{1}{4}W_{\mu\nu}^i W^{\mu\nu i} - \frac{1}{4}B_{\mu\nu} B^{\mu\nu}, \quad (1.9)$$

where the field strength tensors are defined as:

$$W_{\mu\nu}^i = \partial_\nu W_\mu^i - \partial_\mu W_\nu^i - g\epsilon^{ijk}W_\mu^j W_\nu^k, \quad (1.10)$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu. \quad (1.11)$$

and the covariant derivate is:

$$D_\mu = \partial_\mu + \frac{ig}{2} I^k W_\mu^k + \frac{ig'}{2} Y B_\mu, \quad (1.12)$$

where ϵ^{ijk} is the Levi Civita tensor, I^k are the generators of the $SU(2)_L$ and the hyper charge Y is the generator of $U(1)_Y$, which is connected to the electric charge and weak isospin ($Q = I_3 + Y/2$). The coupling constants g and g' are related by the weak mixing angle (θ_W):

$$\tan \theta_W = \frac{g'}{g}, \quad (1.13)$$

which can be used to find the photons and the massive weak bosons in terms of these fields:

$$A_\mu = \cos \theta_W B_\mu + \sin \theta_W W_\mu^3, \quad (1.14)$$

$$Z_\mu^0 = -\sin \theta_W B_\mu + \cos \theta_W W_\mu^3, \quad (1.15)$$

$$W_\mu^\pm = \frac{1}{2}(W_\mu^1 \mp iW_\mu^2). \quad (1.16)$$

At this point, all the particles introduced in the lagrangian density are massless to preserve the gauge invariance. But this symmetry is broken in the nature, since the gauge mediators of the weak interaction are massive. In the SM, the SSB of the electro-weak interaction is performed via the Brout, Englert and Higgs (“BEH”) mechanism.

BEH Mechanism

Mass terms are forbidden to the force carriers to preserve the $SU(2)_L$ gauge symmetry. But in the 60s, Brout, Englert and Higgs [15, 16] have introduced a mechanism to the electro-weak theory to explain the massive weak Z^0 and W^\pm bosons. It is based in the SSB on the electro-weak interaction:

$$SU(2)_L \otimes U(1)_Y \xrightarrow{SSB} U(1)_{QED}, \quad (1.17)$$

which means that after the electro-weak symmetry breaking, it remains the symmetry of the electromagnetism and massive the weak gauge bosons.

BEH mechanism is based on the introduction of a complex scalar fields with hyper charge $Y = 1$:

$$\phi(x) = \begin{pmatrix} \phi^+(x) \\ \phi^0(x) \end{pmatrix}. \quad (1.18)$$

Let's consider the lagrangian density associated to such scalar field:

$$\mathcal{L}_{Higgs} = \partial_\mu \phi^\dagger \partial^\mu \phi - V(\phi^\dagger \phi), \quad (1.19)$$

where $V(\phi^\dagger \phi)$ is the scalar potential. \mathcal{L} is constructed to be invariant under local phase transformations of the scalar field ($\phi(x) \rightarrow e^{i\theta(x)}\phi(x)$). Then, the scalar potential that contains all possible invariant terms which preserve the gauge symmetry is:

$$V(\phi^\dagger \phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2, \quad (1.20)$$

where the parameters μ and λ are set to find the ground states of $V(\phi)$, which has a minimum at $\phi^\dagger \phi = -\frac{\mu^2}{2\lambda}$ for $\mu^2 < 0$ and $\lambda > 0$. The ground state is infinitely degenerated. So, the symmetry is broken by choosing a particular ground state solution:

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \quad \text{with } v = \sqrt{\frac{-\mu^2}{\lambda}}, \quad (1.21)$$

where v is the vacuum expectation value (~ 246 GeV). Then, the scalar doublet can be parametrised in terms of the Higgs field H :

$$\phi = \exp\left(\frac{i \sigma^a \xi_a(x)}{2v}\right) \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}, \quad (1.22)$$

where the ξ_a with ($a = 1, 2, 3$) are scalar fields (the four initial degrees of freedom are preserved). In the unitary gauge ($\xi_a = 0$), the kinetic term of the scalar lagrangian density takes the form:

$$(D_\mu \phi)^\dagger D^\mu \phi \rightarrow \frac{1}{2} \partial_\mu H \partial^\mu H + (v + H)^2 \left(\frac{g^2}{4} W_\mu^\dagger W^\mu + \frac{g^2}{8 \cos^2 \theta_W} Z_\mu Z^\mu \right). \quad (1.23)$$

Therefore, a relation between the masses of the electro-weak gauge bosons is predicted by the SM:

$$m_Z = \frac{m_W}{\cos \theta_W}. \quad (1.24)$$

Similarly, the BEH mechanism explains the mass of the fermions from the interaction terms with the scalar boson:

$$\mathcal{L}_{mass} = \lambda_d \bar{Q}_L \phi d_R - \lambda_u \bar{Q}_L \tilde{\phi} u_R - \lambda_\ell \bar{\ell} \phi e_R + h.c. \quad (1.25)$$

where $Q_L = (u_L, d_L)$, $\ell = (\nu_L, e_L)$, $\tilde{\phi} = -i\sigma_2 \phi^*$ and λ_i is the Yukawa coupling with the Higgs field. The mass of the fermions becomes directly proportional to their coupling to the Higgs field:

$$m_f = \frac{\lambda_f v}{\sqrt{2}}. \quad (1.26)$$

The strong interaction

Quantum chromodynamics is a theory describing the strong interaction between colored particles with an exchange of gluons. QCD is constructed as a non-Abelian gauge theory, invariant under a local $SU(3)_C$ symmetry [17, 18]. For a quark field of flavor f , the lagrangian density is:

$$\mathcal{L}_{QCD} = \sum_f i \bar{q}_f \gamma^\mu D_\mu q_f - \frac{1}{4} F_{\mu\nu}^A F^{\mu\nu A}, \quad (1.27)$$

where the quarks are represented as colour triplets. The covariant derivate and the tensor field $F_{\mu\nu}$ are defined as follow:

$$D_\mu = \partial_\mu + i g_s t^A G_\mu^A, \quad (1.28)$$

$$F_{\mu\nu}^A = \partial_\mu G_\nu^A - \partial_\nu G_\mu^A + g_s f^{ABC} G_\mu^B G_\nu^C, \quad (1.29)$$

where G_μ^A are the gluon fields with $A = (1, \dots, 8)$, f^{ABC} is the structure constants of $SU(3)$. Each gluon is associated to a generator matrix of the $SU(3)_C$ group, corresponding to the Gell-Mann matrices (t_A) which satisfies the Lie algebra $[t_A, t_B] = i f_{ABC} t_C$. The g_s is the strong coupling constant which can also be expressed in terms of the QCD interaction strength $\alpha_s = \frac{g_s^2}{4\pi}$.

In perturbation theory, α_s depends on the energy scale of the observed coupling (q^2), and can be written at leading order in terms of the renormalisation scale (μ_R) as:

$$\alpha_s(q^2) \propto \frac{1}{\ln(q^2/\mu_R^2)}. \quad (1.30)$$

The most important properties of QCD are the asymptotic freedom and the confinement, which depend on the value of α_s . For high energy scales q , the value of α_s vanishes asymptotically. Therefore, the strong interaction become weaker at high energies and the quarks are described as

free particles in this regime. On the other hand, the α_s starts to grow at low energies. This is known as the confinement, which implies that free quarks cannot be found in the nature at large distances. Only colorless bound states are invariant under transformations of the symmetry group, and can be observed; they are named hadrons. The large number of known mesonic and baryonic states can be described by QCD: mesons as $M = q\bar{q}$ and baryons as $B = qqq$.

Cabibbo-Kobayashi-Maskawa matrix

The weak interaction, mediated by the massive W gauge bosons, is able to change the flavor of the fermions. But the mass eigenstates are not the same as the interaction eigenstates. The unitary transformation which connects both representations for quarks is called Cabibbo-Kobayashi-Maskawa matrix (V_{CKM}). For leptons, such transformation is called Pontecorvo-Maki-Nakagawa-Sakata, but the PMNS matrix that explains the neutrino oscillations is not a part of the SM.

The current representation of the V_{CKM} is parametrised assuming only three fermion generations. Each matrix element represents the probability for a flavor change of a quark via the mediation of a W-boson. The value of the matrix elements are obtained experimentally [39]:

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} 0.97425 & 0.2253 & 0.00413 \\ 0.225 & 0.986 & 0.0412 \\ 0.0084 & 0.0400 & 0.99 \end{pmatrix}. \quad (1.31)$$

Since $|V_{tb}|$ is much larger than $|V_{td}|$ and $|V_{ts}|$, the top-quark has a very large probability to decay with a Wtb vertex (see section 1.3.1).

1.1.2 SM overview, successes and weaknesses

Since the formulation of the SM, it has been tested in multiple high energy experiments and its validity has been confirmed with high precision measurements. The SM has also predicted some of the particles discovered by such experiments. Just to mention some of the most important results: the discovery of the third generation fermions which provide a natural mechanism for CP violation through the complex phase of the CKM matrix [19], the discovery of the massive electro-weak gauge vector bosons at the CERN SPS collider in 1983 [20], the prediction of the top-quark mass and its discovery in 1995 [21, 22] at the Tevatron, the discovery of the Higgs boson at the LHC [10, 11], etc.

As mentioned, the symmetry group of the SM after the electro-weak SSB is $SU(3)_C \times U(1)_{QED}$. Using the formulation exposed, the full lagragian density is:

$$\mathcal{L}_{SM} = \mathcal{L}_{QCD} + \mathcal{L}_{QED} + \mathcal{L}_{Higgs} + \mathcal{L}_{mass}, \quad (1.32)$$

However, there are 19 free parameters in the SM formulation: 9 fermion masses (Yukawa couplings to the Higgs field), 3 mixing angles and a CP violation phase of the unitary CKM matrix, the vacuum expectation value, the Higgs boson mass, 3 coupling constant corresponding to each symmetry in SM and the QCD CP violation phase (no evidence of QCD violated process has been observed yet).

Experimental success

The consistency of the SM is confirmed through the electro-weak fit [23], where the fundamental parameters of the SM formulation can be measured simultaneously using the data from the high energy experiments. The result of the fit is shown in Figure 1.2 in terms of the deviation between the SM prediction and the measured parameter in units of the measurement uncertainty. The fit results are shown for both scenarios, including or not the M_H measurement, since all parameters entering electro-weak precision observables are known for the former case, allowing a full assessment of the consistency of the SM at the electroweak scale. No deviation exceeds the three standard deviations, showing a satisfying consistency with the SM prediction.

The Figure 1.3 shows an impressive agreement of several cross section measurements for the production of SM processes at the LHC, compared to the corresponding theoretical expectations calculated at NLO or higher. The experimental measurements are performed using proton-proton collision data at $\sqrt{s} = 7, 8$ and 13 TeV, recorded by the ATLAS experiment [24].

SM weaknesses

Despite its predictive power and descriptive precision, the SM has limitations of experimental and theoretical nature.

First of all, the theoretical framework of SM does not describe the gravitational force. The general theory of relativity is the current understanding of the gravitational interaction, which is formulated within the “classical physics” framework. A quantum theory of gravity is needed to describe the general relativity theory with the principles of quantum mechanics. The usual prescriptions of quantum field theory can be used to describe the force of gravity as mediated by an exchange of graviton bosons. But such theory is not renormalizable.

Then, the SM does not explain the reason for the large magnitude difference between the weak interaction and the gravitational interaction (hierarchy problem) and why the Higgs boson mass is so small (naturalness issue). To illustrate this point, let's consider the one-loop quantum correction to the Higgs boson mass (δm_H^2):

$$\delta m_H^2 = \frac{\lambda_f^2 \Lambda^2}{8\pi^2} \quad (1.33)$$

where λ_f is the Yukawa coupling with the fermions and Λ is the scale of the process. At orders of the Plank scale, the scale at which the gravitation has to be taken into account (energies of

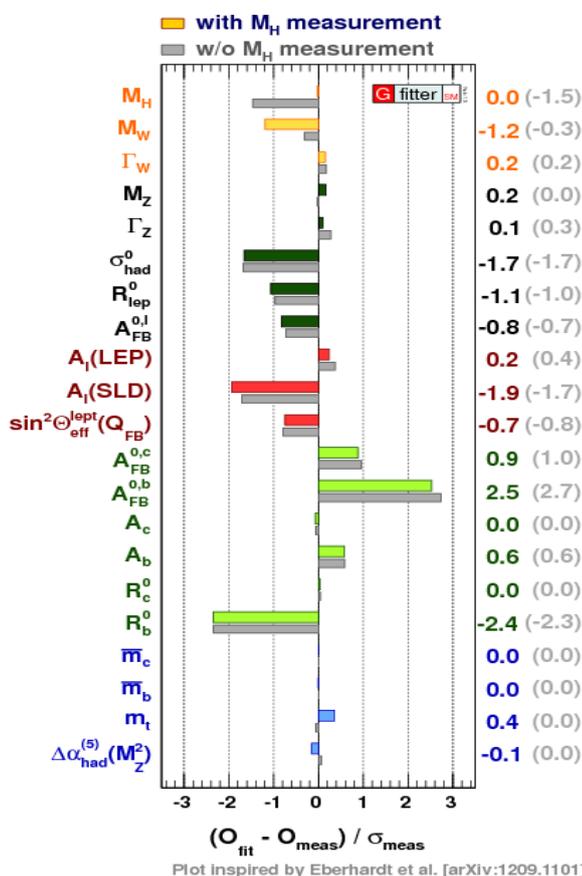


Figure 1.2: Differences between the SM prediction and the measured parameters including the measurement of the Higgs boson mass (coloured) and without (grey). Pseudo experiments are generated for the SM parameters according to Gaussian distributed values around their expected values and with standard deviations equal to the full experimental uncertainty. A χ_{min}^2 distribution is obtained using the toy-data to check the agreement with the idealized distributions [23].

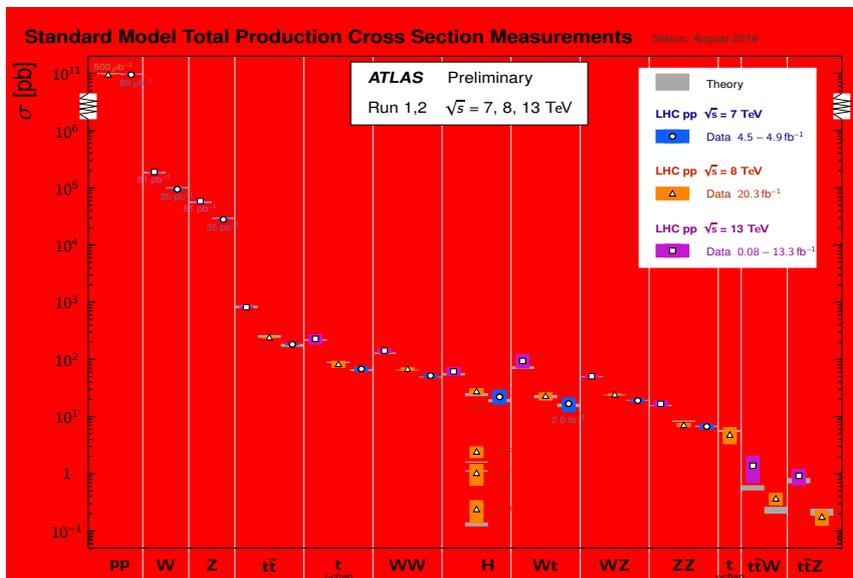


Figure 1.3: Summary of several SM total production cross section measurements performed by the ATLAS experiment [24]. All theoretical expectations were calculated at NLO or higher and the cross section measurements were performed using the 2011, 2012, or 2015 datasets.

about 10^{19} GeV), the correction is very large. Therefore, to ensure $m_H^2 \sim 100$ GeV, there should be an “un-natural” fine-tuning of the parameters that enters in the radiative term. Therefore, the un-natural low Higgs mass requires new physics at the TeV scale energies.

Cosmological observations shows that the SM explains only about 5% of the energy presents in the observable universe. There is about 26% of the energy corresponding to the dark matter, which unlike SM matter, it does not interact via the electro-magnetic force. This means it does not absorb, reflect or emit light. The SM does not propose any fundamental particles for a dark matter candidate.

To summarise a few more open questions: the SM does not predict the number of fermion families observed or their mass hierarchy; in the SM, the neutrinos are massless, but experimental measurements of the neutrino flavor oscillations indicate they have a small mass.²

²Mass terms for neutrinos can be added artificially to the SM, but these lead to new theoretical problems as the origin of the neutrino mass terms.

1.2 Beyond the Standard Model

Physics beyond the Standard Model (BSM) refers to the theoretical proposals that aim to explain the deficiencies of the SM [25, 26, 27]. Some of these BSM models predicts new particles decaying in top-quark pairs. The motivation for the benchmark models historically used by this analysis are presented in this section. For the time-scale of this manuscript, only the signal simulation for a Topcolor assisted technicolor Z' boson was used (see Chapter 5).

1.2.1 Topcolor assisted technicolor

There are models proposing solutions to the naturalness problem by introducing new interactions, such as the “technicolor” force. An alternative mechanism is introduced to generate the masses of the electro-weak gauge bosons through the dynamics of the new gauge interactions. Additional massless fermions sensible to the technicolor interaction (techni-fermions) are introduced, leading to the formation of a techni-fermion condensates. However, this model can not generate the masses of the fermions. This interaction is asymptotically free at very high energies and becomes strong and confining when the energy decreases to the electro-weak scale (246 GeV).

Additional new interactions has to be included to obtain the massive SM fermions, known as “Extended Technicolor” forces (ETC). The energy scale for the SSB of the symmetry associated to ETC is around 100 GeV. This means that this mechanism can not explain the large top-quark mass. The ETC approach is based on the gauge dynamics of fermions only [28].

Alternative models were developed in the early 90s to provide an explanation of the large top-quark mass with the introduction of a new interaction for the third generation quarks, which is called “topcolor”. A large mass $t\bar{t}$ condensate is formed by strong interactions at the energy scale, Λ_t . The resulting low-energy theory simulates the SM if the topcolor scale is very high $\Lambda_t \sim 10^{15}$ GeV $\gg m_t$. Therefore, this topcolor scenario is highly unnatural, requiring a high fine-tuning.

“Topcolor Assisted Technicolor” (TC2) is denoted for the combination of these theories [29, 30]. Under this framework, the electro-weak SSB is performed by the technicolor interactions which are strong near 1 TeV. The light quark and lepton masses are still generated by ETC. The topcolor interaction, also with a scale near 1 TeV, generates the $t\bar{t}$ condensates and the large top-quark mass. Their marriage allows to introduce technicolor interactions which explains the large top mass. The breaking of the gauge symmetries proposed generates a massive neutral gauge boson Z'_{TC2} , which is a color single resonance with spin-1.

Since technicolor models propose an alternative mechanism to explained the masses of the electro-weak gauge bosons, it is strongly disfavoured after the Higgs discovery. Although, technicolor is not completely discarded since there are models introducing a technicolor scalar which can be interpreted as the Higgs boson [31]. Nevertheless, the motivation to use a TC2 model is independent of its validity, since this benchmark has already been used in previous searches by ATLAS and CMS collaborations (see section 1.4), and the consistency of the analysis with respect to previous results can be directly compared.

1.2.2 Extra-dimensions

Some models propose extra dimensions in the Randall-Sundrum formalism (RS) with a single warped extra dimension to explain the hierarchy and naturalness problems. Studies of the phenomenological consequences of the Kaluza-Klein mode of the graviton (G_{KK}) in RS theories have focused primarily on the scenario where all SM particles reside on a four-dimensional slice (denoted as “3-Brane”) [32, 33]. This model introduces an extra dimension (y coordinate) of finite size with two branes placed at $y = 0$ and $y = \pi r_c$, denominated Planck and TeV brane respectively. The metric proposed to solve the Einstein equation of general relativity should describe a flat and static 4-dimensional universe:

$$ds^2 = e^{-k|y|} \eta_{\mu\nu} dx^\mu dx^\nu + dy^2 = g_{ab} dx^a dx^b \quad (1.34)$$

where the warp factor ($e^{-k|y|}$) is written as an exponential for convenience; k determines the curvature of the spacetime; and $\eta_{\mu\nu} = \text{diag}(-1, 1, 1, 1)$ is the 4-dimensional Minkowski metric. In this model, the SM matter is restricted to live inside the TeV brane and the gravitation is spread on the extra-dimension with a weaker contribution in the TeV brane. In that way, the hierarchy problem is understood.

In more realistic models [34, 35, 36], the SM fields are allowed to propagate in the bulk. The particle fields propagating in the extra-dimension exhibit a discrete spectrum of mass states (KK excitations). In addition to the graviton, excited states of the gluons and electro-weak gauge bosons are predicted by these models.

1.2.3 Two Higgs Doublet Model motivation

So far, the SM contains the most simple scalar structure, just one SU(2) doublet. But the scalar sector of the SM can be explored in case there is a richer structure. The most simple generalisation

is the introduction of another SU(2) doublet, the Two Higgs Doublet Model (2HDM) [37].

A good motivation of 2HDM is that it is needed by the supersymmetry. In the Minimal Supersymmetric Standard Model (MSSM), which is a simple supersymmetric model, each SM fermion/boson has a supersymmetric boson/fermion partner: a particle with same mass and quantum numbers but their spin differs by 1/2. By doing so, the radiative correction of Eq. 1.33 cancel out. This means that the scale of validity of the theory is extended. In addition, many supersymmetric theories predicts stable massive particles, which are electrically neutral and that interacts weakly with the SM particles (weakly-interacting massive particle, WIMP). However, supersymmetric particles have not been observed yet. But if they exist, the supersymmetry is broken in nature.

In 2HDM, the electro-weak SSB is performed with an extension of the BEH mechanism. There are eight degrees of freedom which will be used to provide masses to the gauge bosons after the symmetry breaking. In general, the vacuum structure of 2HDMs is very rich, where the most general scalar potential contains 14 parameters. However, most phenomenological studies of 2HDMs are developed using several simplifying assumptions. It is usually assumed that CP is conserved in the Higgs sector, and that CP is not spontaneously broken.

In this particular formulation, three degrees of freedom are associated to the weak gauge bosons. The remaining five degrees of freedom are associated to physical Higgs bosons: two are associated to complex doublets, two real CP-even scalars, one real CP-odd scalar. The most important parameters used in the 2HDMs studies are: $\tan\beta$, which is ratio of the vacuum expectation values of the scalar doublets $\tan\beta = v_1/v_2$; and α , which is the mixing angle of CP-even fields. These two parameters α and β determine the interactions of the various Higgs fields with the SM vector bosons and fermions.

1.3 Top-quark physics at the LHC

The top-quark was predicted before its direct observation in 1995, by the CDF and DØ experiments at the Tevatron accelerator in Fermilab. The top-quark is the heaviest of all known quarks with a mass $m_t = 173.2 \pm 0.9$ GeV which implies a large coupling to the Higgs boson [39]. For this reason, it is often speculated whether the top-quark has a special role in the electro-weak SSB.

1.3.1 Top-quark production modes

Top-quarks are produced at the LHC, either in $t\bar{t}$ pair through the strong interaction, or singly through the weak interaction. In addition, since the mass of the top-quark is larger than Λ_{QCD} , $t\bar{t}$ production at LHC can be successfully described in terms of QCD. The LO contribution of gluon induced $gg \rightarrow t\bar{t}$, and quark induced $q\bar{q} \rightarrow t\bar{t}$ processes can be seen in the Feynman diagrams in Figure 1.4.

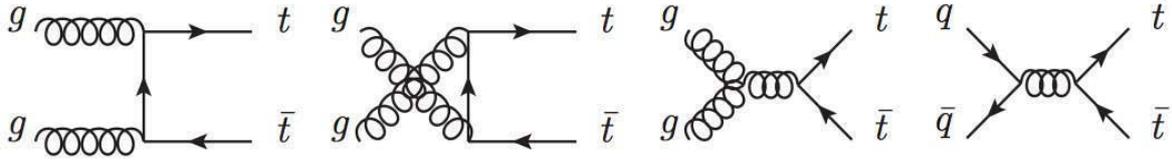


Figure 1.4: Feynman diagrams for $t\bar{t}$ production at leading order QCD.

Single top-quarks are produced through the electro-weak interaction through almost exclusively the Wtb vertex (since $|V_{tb}| \gg |V_{td}|, |V_{ts}|$). There are three different single-top production modes (see Figure 1.5):

In the t-channel mode a space-like W-boson scatters off a b-quark, which is either considered through the b-quark PDF in the proton (flavor excitation) or produced via gluon splitting $g \rightarrow b\bar{b}$.

In the s-channel mode a time-like W-boson is produced from two quarks and decaying into $t\bar{b}$.

In the tW-channel mode which is also called associated production, the top-quark is produced in association with a W-boson.

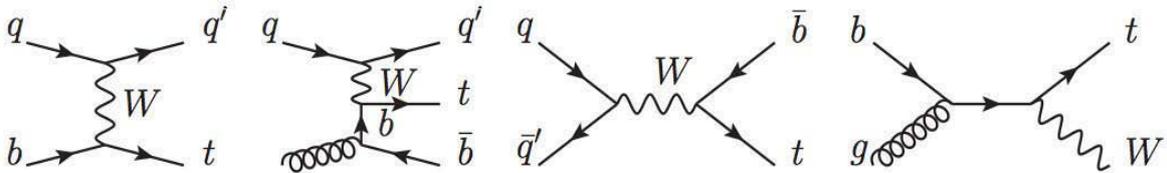


Figure 1.5: Feynman diagrams for single top-quark production at LO QCD. From left to right: t-channel produced via flavor excitation and via W-gluon fusion; s-channel production; tW-channel production.

Single top-quark production is interesting for various reasons. Measuring all three production modes, since they are sensitive to the Wtb vertex. Non-SM couplings would indicate the presence of contributions from new physics. Also, single top-quark production allows to directly measure the CKM matrix element $|V_{tb}|$ [38].

1.3.2 Top-quark decay

Due to the large mass of the top-quark, it decays before hadronizing with a lifetime of about $O(10^{-25})$ s. Therefore, the top-quark has to be reconstructed from its decay products. The SM predicts only three possible decays: $t \rightarrow bW^+$, $t \rightarrow sW^+$, $t \rightarrow dW^+$. The probability of occurrence for these processes are proportional to the corresponding CKM matrix element $|V_{tq}|^2$ with $q=b,s,d$. This means that the branching ratio for the decay modes of the top-quark is dominated by the $t \rightarrow bW^+$ process: $BR(t \rightarrow bW^+) = 0.99$ [39]. Then, the final signature of the top-quark decay depends of the W-boson decay modes: the leptonic W-boson decay mode into a charged lepton and neutrino is about 33%, while the hadronic decay into quark-anti-quark pair is about 67 % [39]. If a tau lepton is produced in the leptonic W-boson decay mode, it will subsequently decay into an electron or muon plus neutrino, or in quark pairs.

1.3.3 $t\bar{t}$ signatures

There are three possible final state signatures of the $t\bar{t}$ system based in the decay mode of the W-bosons (Figure 1.6), with their respective branching ratio [39]:

Di-lepton channel: both W-bosons decay into lepton (electron, muon, tau) and neutrino, $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow \bar{l}\nu_l b l' \bar{\nu}_{l'} \bar{b}$. The branching fraction is about 6%.

Lepton+jets channel: one W-boson decays into lepton and neutrino, the other one into a quark-anti-quark pair, $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow q\bar{q}' b l \bar{\nu}_l \bar{b} + q\bar{l}' \nu_l b q\bar{q}' \bar{b}$. The branching fraction is about 38%.

Hadronic channel: both W-bosons decay into a quark - anti-quark pair, $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow q\bar{q}' b q'' \bar{q}'' \bar{b}$. The branching fraction is about 56%.

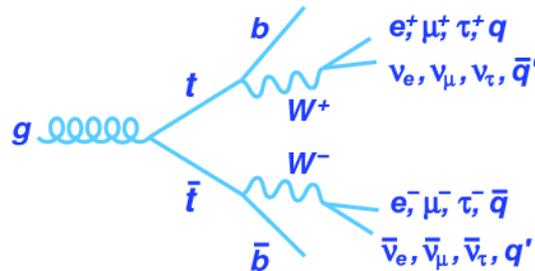


Figure 1.6: Feynman diagrams for $t\bar{t}$ decay modes.

The dominant decay modes are lepton plus jets and hadronic. Although the large branching ratio in the hadronic channel, the background from events with many jets of hadrons are likely to pass the analysis selection. Instead, the multi-jet background are efficiently suppressed by an analysis selecting lepton plus jets events.

Using the QCD parton model (see section 4.1.1), the inclusive production cross section of the process $pp \rightarrow t\bar{t}$ can be computed at NNLO order in QCD. The measured production cross-section of $t\bar{t}$ events by ATLAS and CMS are presented in Figure 1.7, obtained using 2015 p-p collisions at $\sqrt{s} = 13$ TeV. The measurement obtained independently in each decay mode, and the compatibility between measurement and expectation is in good agreement considering the total uncertainties.

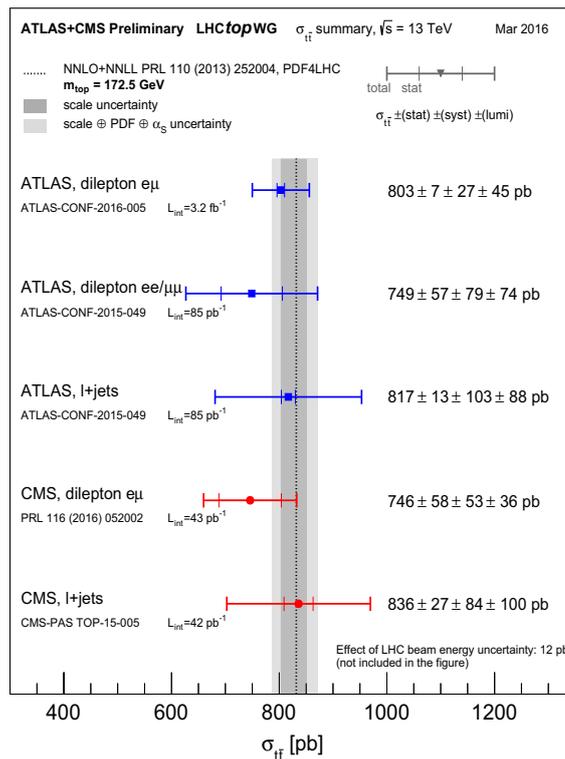


Figure 1.7: Measurements of the top-pair production cross-section at 13 TeV compared to the exact NNLO QCD calculation. The uncertainties bands are obtained considering the renormalisation and factorisation scale, parton density functions and the strong coupling. The measurements and the predictions use $m_t = 172.5$ GeV [38].

1.4 Searching for $t\bar{t}$ resonances

In the analysis exposed in Chapters 5 and 6, the search for $t\bar{t}$ resonances is performed with the 2015 data recorded by the ATLAS detector and limits are set on the Z'_{TC2} benchmark (see Figure 1.8). But in previous iterations of this search performed also by the ATLAS [40], CDF [41], CMS [42] and DØ [43] Collaborations, limits on the production cross-sections have already been set for Z'_{TC2} and other benchmark models. No significant deviation from the total expected background were found.

The search at Run 1 ($\sqrt{s} = 8$ TeV) with the ATLAS detector was performed for various signal benchmark models, upper limits at 95% CL (see definition in section 5.10) are set on the cross-section times branching ratio using a profile likelihood-ratio test. For the Z'_{TC2} benchmark, the limit on the production cross-section is obtained for a Z'_{TC2} of width 1.2% is excluded for masses lower than 1.8 TeV (see Figure 1.9), while masses below 2.0 TeV are expected to be excluded. The mass limits are higher for a Z'_{TC2} width of 2% (3%), reaching 2.0 TeV (2.3 TeV). In this analysis, limits are also set on KK-gluons (with various width) and KK-gravitons predicted in RS models and a scalar $t\bar{t}$ resonance.

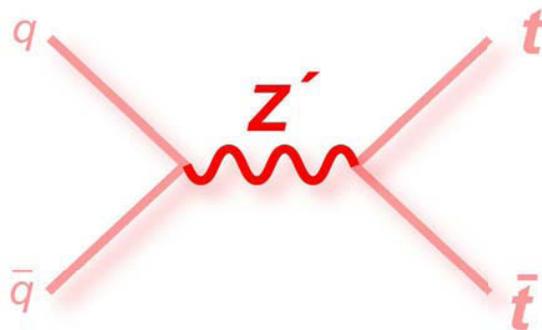


Figure 1.8: Feynman diagram of a topcolor assisted technicolor Z'_{TC2} produced from quark annihilation.

An analysis to search for $t\bar{t}$ resonances in the lepton+jets channel is also performed by the CMS collaboration [42], using 19.6 fb^{-1} of data collected in Run 1 collisions. Even if the strategy is similar to the analysis from ATLAS, both analyses differ in the selection of physics objects, systematic uncertainties and invariant $t\bar{t}$ mass reconstruction. The CMS detector also set limits with a 95 % CL on the production of a Z'_{TC2} boson (Figure 1.10) for several widths: 1.2% and 10%; excluding masses below 2.1 and 2.7 TeV respectively. This analysis also set limits on KK-gluons for various width.

In Run 1, upper limits have been set by the ATLAS collaboration on the production cross-section times branching ratio and the allowed mass range on KK graviton (g_{KK}) and KK gluon (G_{KK}), using 20.3 fb^{-1} [40]. For the g_{KK} , the limits on the production cross-section vary from 4.8 pb for a mass of 0.4 TeV, until 0.09 pb for a mass of 3 TeV. A g_{KK} of width 15.3 % is excluded for masses lower than 2.2 TeV. While for the G_{KK} , the cross-section limit is 2.5 pb for a mass of 0.4 TeV to 0.03 pb for 2.5 TeV, with no mass range excluded. For early Run 2, the signal simulation samples for these benchmarks were not available for the time-scale of this manuscript.

1.4.1 Outlook

The motivations and results from previous searches for $t\bar{t}$ resonances have been presented in this chapter. The ATLAS analysis using 2015 dataset from the p-p collisions is presented in this manuscript, aiming to improve the current upper limits by using the early Run 2 datasets. A increase of \sqrt{s} leads to an increase in the production cross-section of the BSM signals. But in parallel, the cross section of the electro-weak background also increases. Therefore, it is crucial for this analysis be very efficient in the identification of $t\bar{t}$ event candidates in order to improve the sensitivity on the detection of the benchmarks signals.

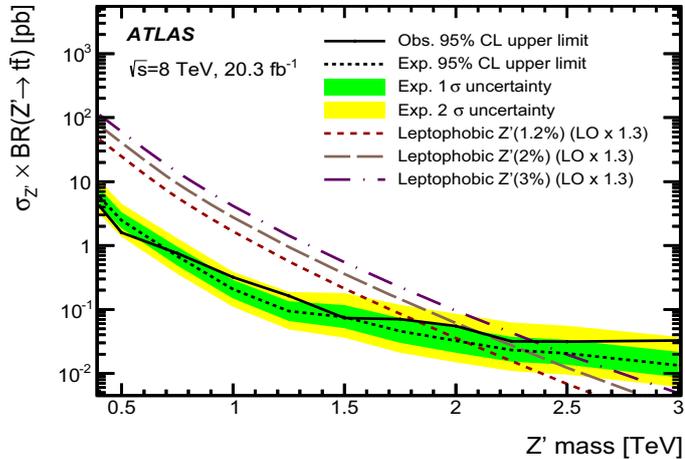


Figure 1.9: Observed and expected upper limits results by the ATLAS analysis at Run 1 using 20.3 fb^{-1} of data collected in p-p collisions at $\sqrt{s} = 8 \text{ TeV}$. The cross-sections times $t\bar{t}$ branching ratio as function of the different mass point for the Z' boson [40].

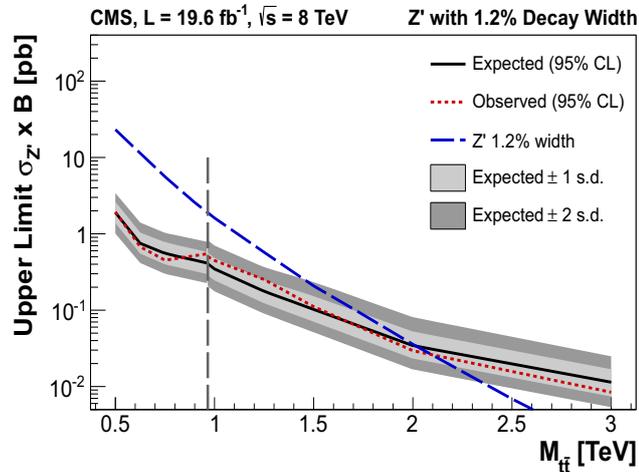


Figure 1.10: Observed and expected upper limits results by the CMS analysis at Run 1 using 19.6 fb^{-1} of data collected in p-p collisions at $\sqrt{s} = 8 \text{ TeV}$. Observed and expected upper limits on the cross-sections times $t\bar{t}$ branching ratio as function of the different mass point for the Z' boson. The vertical dashed line indicates the transition between threshold and boosted analysis, chosen based on the sensitivity of the expected limit [42].

Chapter 2

Experimental context

Contents

2.1	The Large Hadron Collider	36
2.1.1	Proton acceleration chain	37
2.1.2	Proton collisions at LHC	38
2.2	The ATLAS detector	40
2.2.1	ATLAS coordinate system	41
2.2.2	Magnet system	42
2.2.3	Inner detector	42
2.2.4	Calorimeters	45
2.2.5	Muon spectrometer	48
2.2.6	Forward detectors	49
2.2.7	Data acquisition system	50
2.2.8	Computing model	53
2.2.9	Data quality in 2015 data-taking	54
2.3	Outlook	55

The CERN’s Large Hadron Collider (LHC) marks a new era in particle physics by reaching unprecedented energies. The Higgs boson discovery (July 2012) was announced by ATLAS and CMS experiments using 2011-2012 datasets from the proton-proton (pp) collisions of the LHC. The 2013 Nobel prize in physics was awarded jointly by François Englert and Peter Higgs for “the theoretical discovery of a mechanism that contributes to our understanding of the origin of mass of subatomic particles, and which recently was confirmed through the discovery of the predicted fundamental particle, by the ATLAS and CMS experiments at CERN’s Large Hadron Collider” [44].

An increase of the luminosity and the centre-of-mass energy was scheduled for the p-p collisions at LHC in early 2015, to extend the reach of new physics in higher mass scales and to increase the production cross section of interesting processes. The analysis presented in this thesis uses the 2015 datasets to search for new resonances decaying into top-quark pairs, where the physics objects in the final state are identified and reconstructed using the ATLAS sub-detectors. A brief description of the LHC accelerator chain and ATLAS particle detector is presented in the following sections.

2.1 The Large Hadron Collider

The LHC [45] is the largest and most energetic particle collider in the world, starting its first run on September 2008. It is located under the franco-swiss border at CERN (Figure 2.1), in a tunnel approximately 100 meters under the surface. It consists of a 27 kilometre ring of superconducting magnets where two adjacent and parallel beams of protons (or heavy ions) are accelerated in opposite directions. The proton beams intersect each other at four interaction points, where the main particle detectors are located: ALICE [46], ATLAS [47], CMS [48] and LHCb [49].

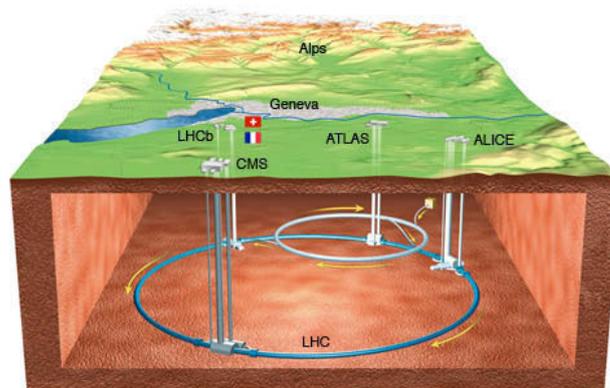


Figure 2.1: Geographical location of the LHC and of the four main experiments.

2.1.1 Proton acceleration chain

Before being injected in the LHC ring, the proton's energy is increased progressively in the accelerator chain (Figure 2.2). The proton beams are obtained from the ionisation of hydrogen atoms with an electric field. Then, the linear accelerator (LINAC) provides an energy up to 50 MeV to the proton beams. Finally a chain of three circular accelerators is used before the injection in the LHC: the Proton Synchrotron Booster (PSB), the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS), where the proton beams reach an energy of 1 GeV (PSB), 26 GeV (PS) and 450 GeV (SPS) respectively. Finally, the protons are transferred to the two beam pipes of the LHC to reach a maximum energy of 6.5 TeV under normal operating conditions.

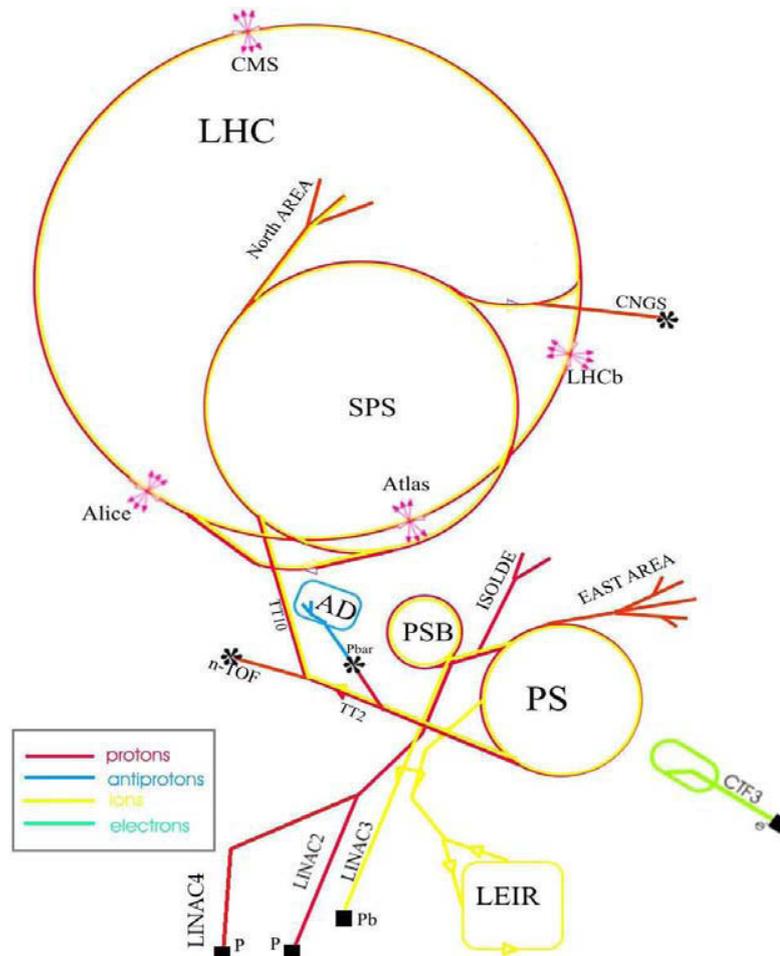


Figure 2.2: Schematic view of the CERN's accelerator chain.

Between the beam injection into LHC and the pp collisions, there are several operational stages. Once the beams are injected into the LHC, the proton bunches circulate through separated vacuum tubes. The LHC use superconducting radio frequency cavities to accelerate the beams, and stabilize the proton bunches in buckets. Each vacuum tube is surrounded by superconducting magnets; 1296 dipole magnets are placed to achieve the bending and 392 quadrupole magnets for the focalisation of the beams (Figure 2.3). The parameters of the machine are optimised for the acceleration of the beams until the energy of collision (ramp). The proton beam are accelerated close to the speed of light (in approximately 20 minutes) through the superconductive elements (Niobium-Titanium) which are cooled until 1.9 K [50][51]. Finally, the beams are squeezed and adjusted for the collisions at the four intersection points. The “stable beams” status is declared, and the data recording can start using the different particle detectors.

The energy at the centre of mass of the pp collision was $\sqrt{s} = 7$ TeV in 2011, reaching $\sqrt{s} = 8$ TeV in 2012. The first period of LHC’s collisions (Run 1) officially ended on February 2013, starting the “long shutdown” period for planned upgrades (2 years). The second period of collisions (Run 2) restarted in June 2015 reaching an energy in the centre-of-mass of $\sqrt{s} = 13$ TeV.

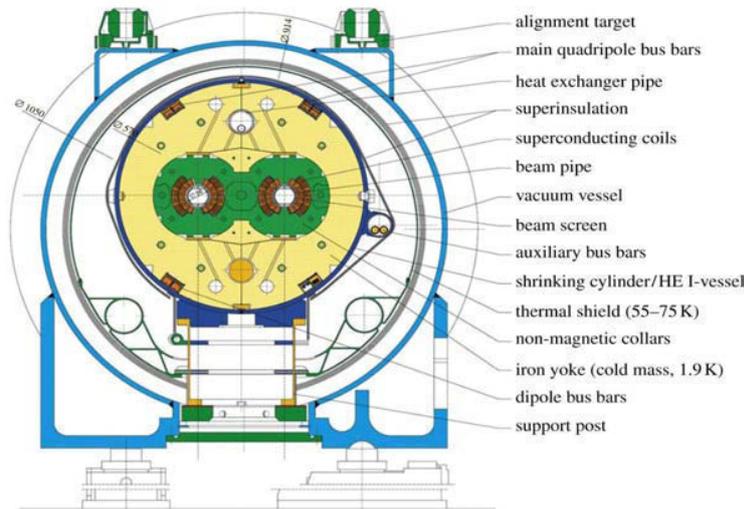


Figure 2.3: Transverse section of the superconducting LHC dipole magnet.

2.1.2 Proton collisions at LHC

The production rate of $pp \rightarrow X$ events is proportional to the instantaneous luminosity \mathcal{L} , which depends only on the intrinsic beam parameters:

$$\frac{dN_{pp \rightarrow X}}{dt} = \mathcal{L} \sigma_{pp \rightarrow X}, \quad (2.1)$$

where $\sigma_{pp \rightarrow X}$ is the production cross section, which is typically small for particles predicted by BSM theories. Therefore, to maximise the number of potential events coming from new phenomena, a significant luminosity is needed. \mathcal{L} can be written as:

$$\mathcal{L} = \frac{N_b^2 n_b f_{rev} \gamma}{4\pi \epsilon_n \beta^*} F, \quad (2.2)$$

where N_b is the number of particles per bunch, n_b is the number of bunches per beam, f_{rev} is the revolution frequency, γ is the relativistic Lorentz factor, ϵ_n is the normalised transverse beam emittance, the β^* function quantises the amplitude of oscillations of the protons at the collision point and F is a geometric luminosity factor due to the crossing angle at the intersection point:

$$F = \left(1 + \left(\frac{\theta_c \sigma_z}{2\sigma_*} \right)^2 \right)^{1/2}, \quad (2.3)$$

where θ_c is the crossing angle of the beams, σ_z and σ_* are the RMS of the longitudinal and transversal bunch length respectively. Figure 2.4 shows the instantaneous luminosity peak's delivered to ATLAS during pp collisions at $\sqrt{s} = 13$ TeV in 2015.

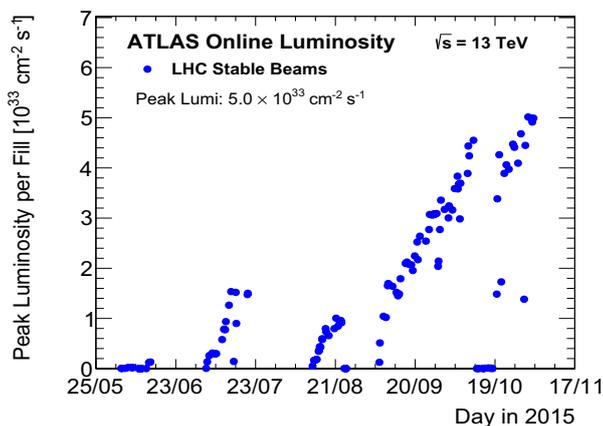


Figure 2.4: Instantaneous luminosity peaks for 2015 data taking.

The integrated luminosity allows to quantify the amount of data delivered by the LHC. It corresponds to the integral of the instantaneous luminosity over a time period. Figure 2.5 shows the time evolution of integrated luminosity delivered by the LHC (green), recorded by ATLAS (yellow)

and the certified runs with good quality data (blue) during pp collisions in 2015.

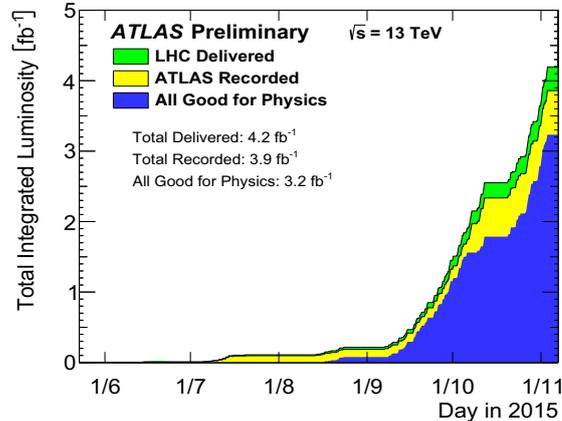


Figure 2.5: Integrated luminosity for 2015 data taking.

The probability to have multiple pp collisions increases linearly with the luminosity. Other pp collisions in addition to the collision of interest, are collectively referred as “pileup” (PU). It presents a serious challenge to physics analyses at LHC since the reconstructed physics object are affected from additional energy contributions. In-time pile-up corresponds to the multiple pp collisions within the same bunch crossing. The experimental observable used as estimator of in-time PU is the multiplicity of reconstructed primary vertex N_{pv} . The out-of-time PU corresponds to pp interactions occurred in a previous bunch crossing. It is the result of long electronic integration times, becoming significant when the spacing decreases between the bunch crossing.

2.2 The ATLAS detector

A Toroidal LHC ApparatuS (ATLAS) is the largest detector located at the LHC ring, with 44 m of length, 25 m of height and 7500 tons of weight. The physics program covered by ATLAS aims to perform precision measurements of the SM observables and to search for particles predicted by different BSM theories. This section describes how ATLAS is composed of several sub-detectors around the interaction point (Figure 2.6). The inner detector, the closest to the beam pipe, is used for the tracking of charged particles. It is followed by the calorimetric system which measures the energy and direction of the particles by total absorption. The muons are penetrating enough to cross through the calorimeters until the muon spectrometer. Moreover a system of magnets provides the magnetic field for the tracking systems and the muon spectrometer. Because of the high cross sections of QCD multi-jet background in pp collisions, a trigger system is used to

select potential interesting events. In addition, ATLAS has also multiple detectors to monitor the luminosity and the beam position.

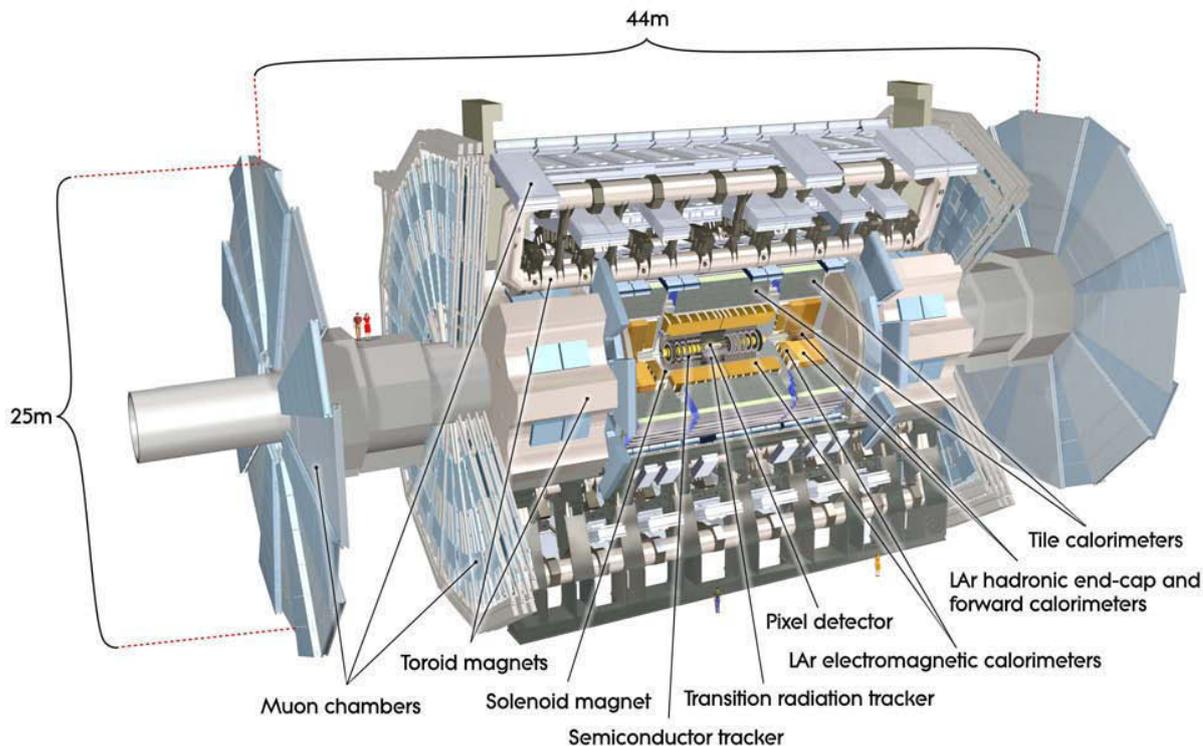


Figure 2.6: Schematic view of the components of the ATLAS detector.

2.2.1 ATLAS coordinate system

To parametrize the trajectories of particles traversing ATLAS, a cartesian right-handed coordinate system is defined with the origin corresponding to the geometrical centre of the detector. The counter-clockwise beam direction defines the positive z -axis and the positive x -axis points to the centre of the LHC ring. The x - y plane defines the transverse plane and the azimuthal angle ϕ is defined with respect to the positive x -axis. The polar angle θ is defined with respect to the longitudinal direction.

Since the partons interacting in pp collisions have an unknown fraction of the proton momentum, boost of Lorentz invariant quantities are preferred to describe the object kinematics. The rapidity is used to replace the polar angle:

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right) \quad (2.4)$$

where E and p_z denotes the energy and the longitudinal momentum component of an object detected by ATLAS. Differences in rapidity are Lorentz invariant under boosts along the longitudinal axis. The rapidity is equivalent to the pseudorapidity in the limit where a particle is travelling close to the speed of light:

$$\eta = -\ln \left(\tan \frac{\theta}{2} \right) \quad (2.5)$$

The angular plane is formed by the variables η and ϕ , on which the distance is defined as:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \quad (2.6)$$

The following kinetic variables are also used since they are invariant under a boost in the longitudinal direction:

$$E_T = \frac{E}{\cosh \eta} = E \sin \theta \quad \text{and} \quad p_T = \sqrt{p_x^2 + p_y^2} \quad (2.7)$$

2.2.2 Magnet system

A system of superconducting magnets [55] provides the magnetic field for bending the charged particle trajectories. It is composed by the central solenoid and the toroidal magnets, which provide the magnetic field for the inner detector and the muon spectrometer respectively (Figure 2.7). The central solenoid is composed by four magnets aligned to the beam axis and provides an axial magnetic field of 2.0 T. The toroidal magnet is composed by a barrel toroid and two endcap toroids, providing a magnetic field reaching up to 2.5 T in central region and to 0.35 T in endcap region.

2.2.3 Inner detector

The inner detector [56] [57] is designed to reconstruct the trajectories of the charged particles, leading to an accurate measurement of the transverse momentum with an expected resolution of $\sigma_{p_T}/p_T = (0.05 \%) p_T \oplus 1 \%$. It also has an important role for the reconstruction of primary and secondary vertices. It is composed by three high precision subsystems: the pixel detector, the semiconductor tracker (SCT) and the transition radiation tracker (TRT), covered by the magnetic field

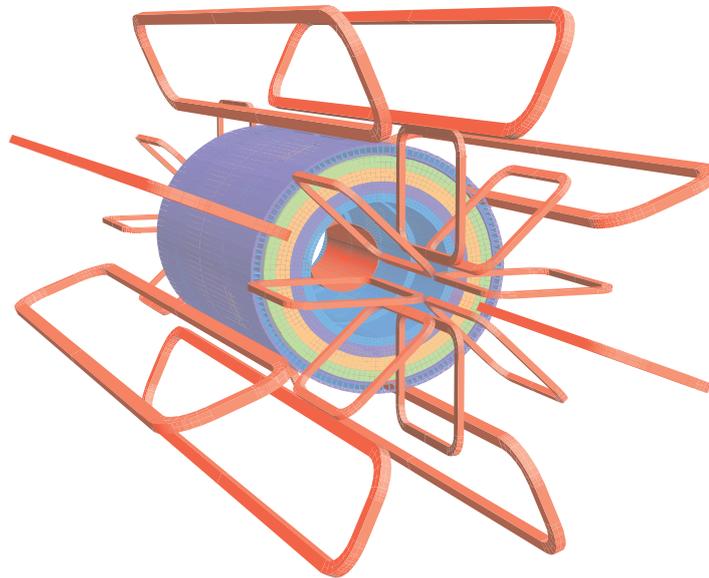


Figure 2.7: ATLAS magnet system, composed by the central solenoid and the toroidal magnets.

from the central solenoid. A charged particle with $p_T > 5.0$ GeV and $|\eta| < 2.0$ passing through the inner detector, typically leaves 4 hits at the pixel detector, 4 hits at SCT and around 30 hits at TRT (Figure 2.8).

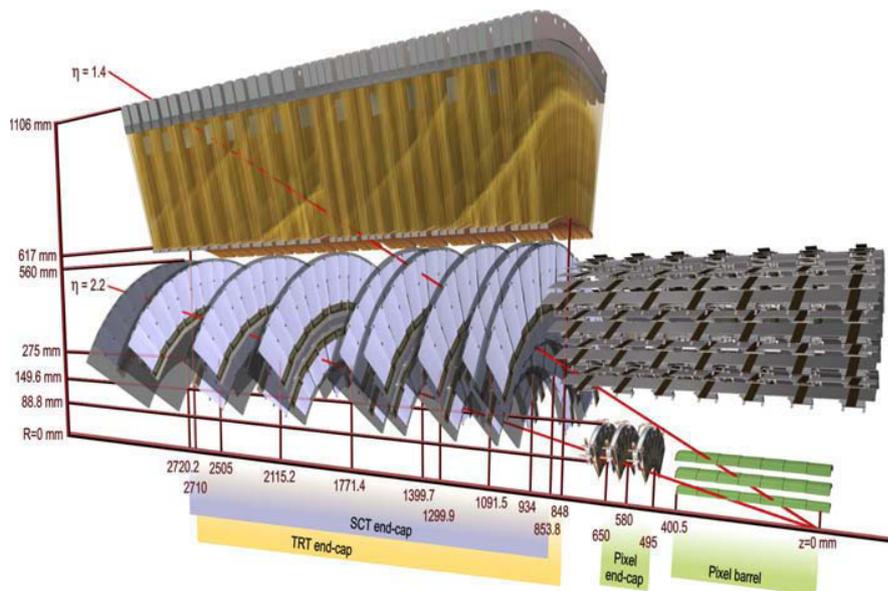


Figure 2.8: Components of the ATLAS Inner detector.

Pixel detector

Since the pixel detector is the innermost tracking subsystem, it is the most important detector for the reconstruction of the secondary vertices, used for the identification of the jets coming from a b-quark. It consists of 1744 modules built of a layer of silicon connected to the read-out electronic, having in each sensor tile about 47000 pixels (nominal pixel size is $50 \mu\text{m} \times 400 \mu\text{m}$). It was up to 2013 composed by three concentric layers around the beam axis and three endcap disks on each side of the interaction point which are perpendicular to the beam axis, having a coverage up to $|\eta| < 2.5$ and a complete ϕ coverage. In 2013, an insertable B-Layer (IBL) [58] has been placed between the beam pipe and the inner pixel layer (with a radial extension between 31 and 40 mm) improving the reconstruction of primary and secondary vertices, increasing the identification efficiency of jets coming from b-quarks and the discrimination of electrons from converted photons.

When charged particles from the interaction point pass through the silicon, a electron-hole pair is created in the semiconductor and a measurable signal is produced thanks to an external electrical field, determining the position of the charged particles. The pixel detector has an experimental resolution of $40 \mu\text{m}$ in the longitudinal direction and $8 \mu\text{m}$ in the $(r - \phi)$ plane [54].

SemiConductor Tracker

Located after the pixel detector, it is the middle component of the inner detector. Having a similar technology than the pixel detector, the SCT is composed of 4 coaxial cylindrical layers in the barrel region and nine endcap disks along the beam axis, providing a coverage up to $|\eta| < 2.5$. It consists of about 16000 silicon strip sensors distributed on 4088 modules. It has an expected resolution of $580 \mu\text{m}$ in the longitudinal direction and $17 \mu\text{m}$ in the $(r - \phi)$ plane.

Transition Radiation Tracker

The outermost component of the inner detector is the TRT. It is composed of polyimide straw tubes, filled with a gaseous mixture (Xe, CO₂ and O₃). Each tube contains a thin wire going through the centre of the tube. The gas atoms are ionised by the charged particles passing through the tubes. Then, the electrons and charged atoms are separated by an electric field existing between the surface of the tube and the wire.

In addition to the TRT, the reconstructed tracks are useful for the discrimination between pions and electrons. A relativistic particle emits photons when passing through medias with different dielectric constants. The total energy radiated is proportional to the Lorentz γ factor, which is significantly larger for an electron than for a pion with the same energy.

2.2.4 Calorimeters

The calorimetric system of ATLAS [61] (Figure 2.9) is located outside the solenoid magnet, providing a coverage region of $|\eta| < 4.9$ and a full ϕ coverage. It is designed to absorb the interacting particles coming from the interaction point (except muons and neutrinos), measuring their energy and position. The incident particle interacts with the calorimeter material through different processes depending of the particle, its kinematics and the material.

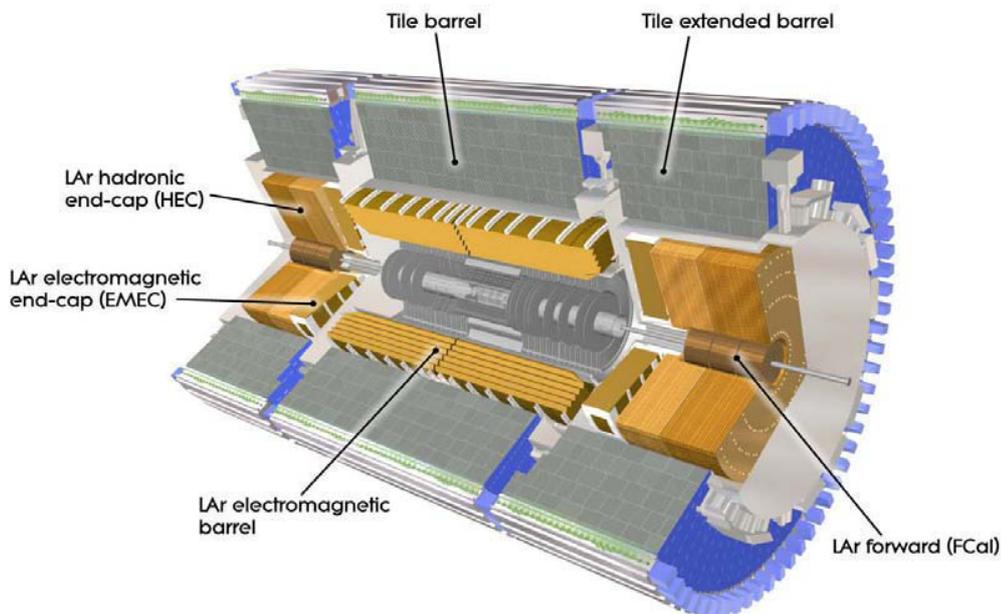


Figure 2.9: Schematic view of the components of ATLAS calorimetric system.

Electromagnetic (EM) calorimeter

The ATLAS EM calorimeter [62, 63] is designed for the identification and measurement of photon and electron energy and position, with a coverage of $|\eta| < 3.2$ (excluding the crack region $|\eta| \in [0, 1.37) \cup (1.52, 2.47]$) and a full coverage in ϕ . The EM is a sampling calorimeter with accordion geometry, which uses liquid argon (LAr) as active medium. The charged particles traversing the calorimeter ionize the LAr, where the electrons are collected as a current with an external electric field. The LAr is kept at a constant temperature of 88 K thanks to cryostats.

The EM calorimeter is longitudinally segmented in three layers in the region up to $|\eta| < 2.5$, having different spacial resolution in the (η, ϕ) plane (Figure 2.10). The first layer acts like a

preshower detector, where only a small fraction of the energy from the incoming particles are deposited there. It is finely segmented in η : $\Delta\eta \times \Delta\phi = 0.0031 \times 0.1$. Most of the energy from the incoming particles are deposited in the second layer, which has a segmentation of $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$. The third layer collects only the shower tails, and has a segmentation of $\Delta\eta \times \Delta\phi = 0.05 \times 0.025$. Such fine segmentation allows precise measurement of electrons and photons with an energy resolution of $\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 0.7\%$.

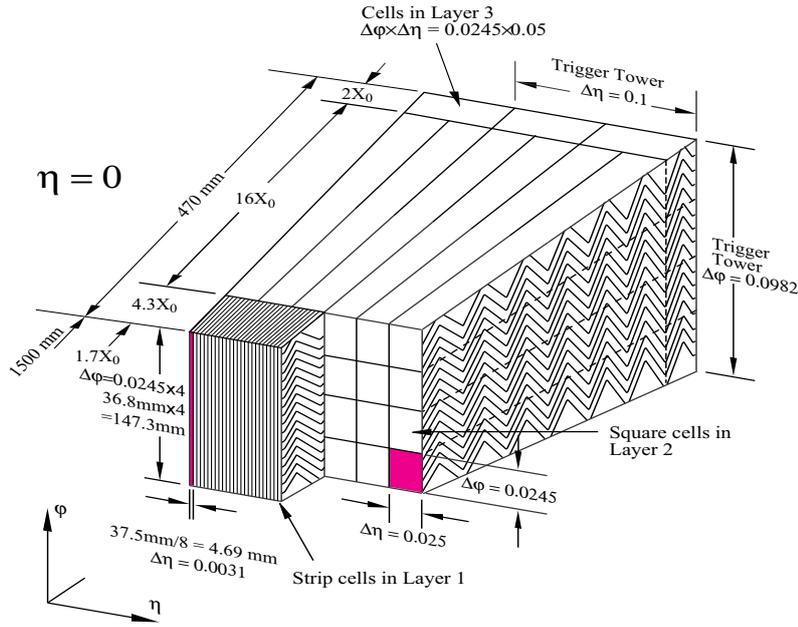


Figure 2.10: Module of the ATLAS electromagnetic calorimeter (central region).

The hadronic calorimeters

The hadronic calorimeters are designed for the reconstruction of the hadronic jets. The technology implemented is different depending on the η region. It is composed by three sub-detectors: the Tile calorimeter and the LAr hadronic endcaps, having a coverage of $|\eta| < 1.7$ and $1.5 < |\eta| < 3.2$ respectively.

The Tile Calorimeter (TileCal) [64, 65] is a sampling calorimeter having scintillating plastic tiles as active medium between layers of steel which plays the role of absorber material (Figure 2.11). When an hadron crosses the absorber material of the TileCal, it produces an hadronic shower which excites the atoms in the scintillator tiles. The de-excitation of such atoms produce light which is collected and transmitted with fibers to the photomultipliers (PMT), where the light is converted into an electrical response. The current generated inside the PMT is proportional

to the energy deposited in the scintillator tiles and the PMT's gain is obtained from the current amplification given by the product of individual *dynode* contributions. The PMTs responses are digitized and transmitted to the ATLAS read-out system in order to measure the deposited energy by the hadron showers. The sampling effect and the full chain lead to an expected energy resolution of $\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\%$, and more details about its structure and granularity are discussed in section 3.1.

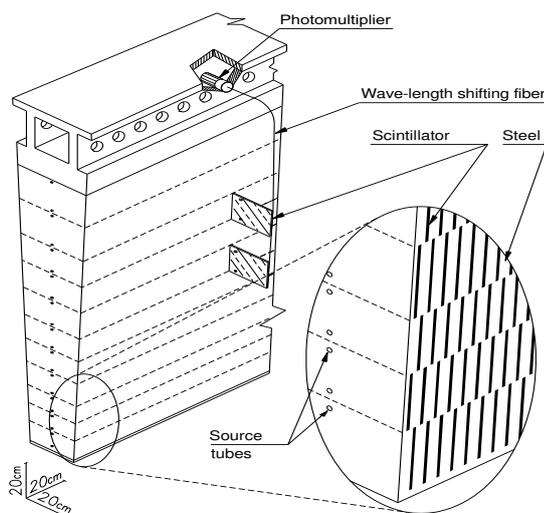


Figure 2.11: Scheme of a module of the tile hadronic calorimeter of ATLAS. It is composed by scintillating tiles and absorbers

The Hadronic Endcap Calorimeter (HEC) is a sampling calorimeter using LAr as active medium and copper as absorber. It extends the hadronic calorimeter to cover higher pseudorapidity regions. The granularity of the HEC is $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ in the region $1.5 < |\eta| < 2.5$, while it is $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ for $2.5 < |\eta| < 3.2$. Its expected energy resolution is $\frac{\sigma_E}{E} = \frac{100\%}{\sqrt{E}} \oplus 10\%$.

The Forward Calorimeter (FCal) [66] ensures the coverage of the calorimetric system in the forward region, having a coverage of $3.1 < |\eta| < 4.9$. This is important to allow a precise determination of the missing transverse energy, although the measurement has to be under the different conditions of intense flux of particles. FCal is a sampling calorimeter having 2 cylindrical endcaps placed approximately at 4.7 m from the interaction point along the beam axis (Figure 2.12). It uses LAr as active medium and copper as absorber. The expected energy resolution for pions is $\frac{\sigma_E}{E} = \frac{94\%}{\sqrt{E}} \oplus 7.5\%$.

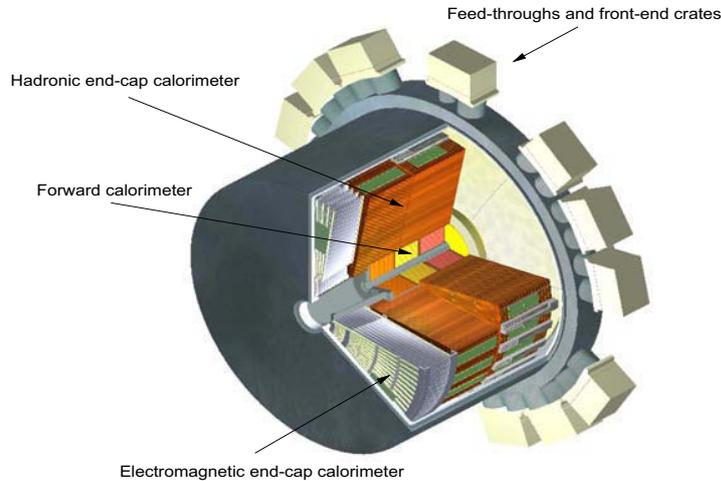


Figure 2.12: Scheme of the Forward calorimeter.

2.2.5 Muon spectrometer

The muon spectrometer (MS) [67] is designed to detect the charged particles passing the calorimetric systems (mainly the muons), and measuring their trajectory and momentum in a coverage range of $|\eta| < 2.7$. The MS has two types of high precision chambers: the Monitored Drift Tube chambers (MDT) and Cathode Strip Chambers (CSC), covering the barrel and forward region respectively. The MDT consists of drift tubes filled with argon/ CO_2 mixture, where the anode wire collects the charge produced by the ionising muons. The CSC system consists of two disks with eight multi-wire chambers each, filled with a gas mixture.

The trigger components of the MS uses two kinds of detector: the Resistive Plate Chambers (RPC) in the barrel region and Thing Gap Chambers (TGC) in the endcaps, covering the pseudorapidity region of $|\eta| < 2.4$ and a full ϕ coverage. The TGC uses the same technology as the RPC, which consists of a gas mixture (CO_2 /n-pentane for TGC and $\text{C}_2\text{H}_2\text{F}_4$ /Iso- C_4H_{10} / SF_6 for RPC) which is enclosed by two resistive plates. A voltage applied between the plates allows to detect the ionising muon, which creates an avalanche along the ionisation tracks towards the anode.

Using the magnetic field provided by the toroidal magnets, the momentum measurement is performed with a resolution of 4 % for muons with p_T around 100 GeV, increasing to 10% for muons with $p_T = 1$ TeV. The track measurements from MS could be combined with those from the inner detector to improve the muon resolution. The MS is instrumented with a separated high precision tracking and muon trigger chambers.

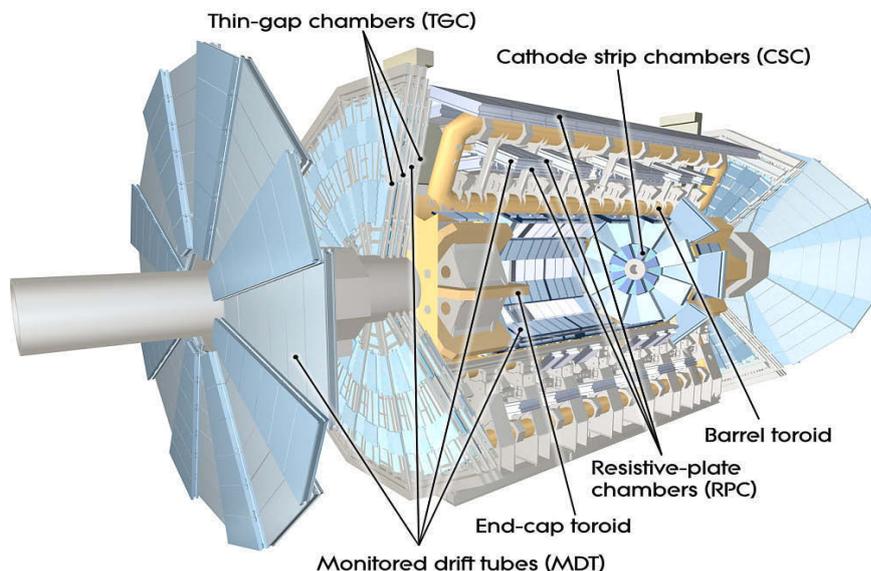


Figure 2.13: ATLAS Muon spectrometer detector.

2.2.6 Forward detectors

Several detector systems perform the measurement of the luminosity and beam position in the forward region of ATLAS. The LUMInosity measurement using Cerenkov Integrating Detector (LUCID) is the main online relative-luminosity monitor of ATLAS. Located at 17 m from the interaction point, it detects inelastic scattering in the forward region. The Minimum Bias Trigger System (MBTS) consists of 32 scintillating paddles organised in two disks, which are on both sides at 3.65 m from the interaction point. The main role of the MBTS is to trigger the data acquisition minimum bias events for the initial running period at low luminosity. The Zero-Degree Calorimeter (ZDC) performs the detection of neutrons and photons from the collisions in a very high pseudorapidity region ($|\eta| > 8.3$). Located at 140 m from the interaction point, it plays a major role in determining the centrality of the heavy ion collisions. The ZDC is a sampling calorimeter, having modules composed of alternating tungsten (absorber material) and quartz plates. The Absolute Luminosity For ATLAS (ALFA) allows the measurement of the proton-proton total cross section and the absolute luminosity. Located at 240 m from the interaction point, ALFA is composed by 8 scintillating fibre trackers located inside roman pots stations designed to approach until 1 mm from the beam.

2.2.7 Data acquisition system

Triggers

The ATLAS trigger system is responsible for deciding whether a given beam crossing is recorded or not. The selection at the trigger-level must provide enough rejection to reduce the event recording rate according to the offline computing power and storage capacity, recording only “interesting events” coming from the 40 MHz rate LHC collisions (Figure 2.14). For Run 2, the ATLAS trigger system is based on two levels of online event selection: the first-level trigger (Level-1 trigger) is implemented in hardware using a subset of the detector information to reduce the accepted rate to 100 kHz, it is followed by a software-based trigger called higher-level trigger (HLT), which reduces the rate of events recorded to 1 kHz.

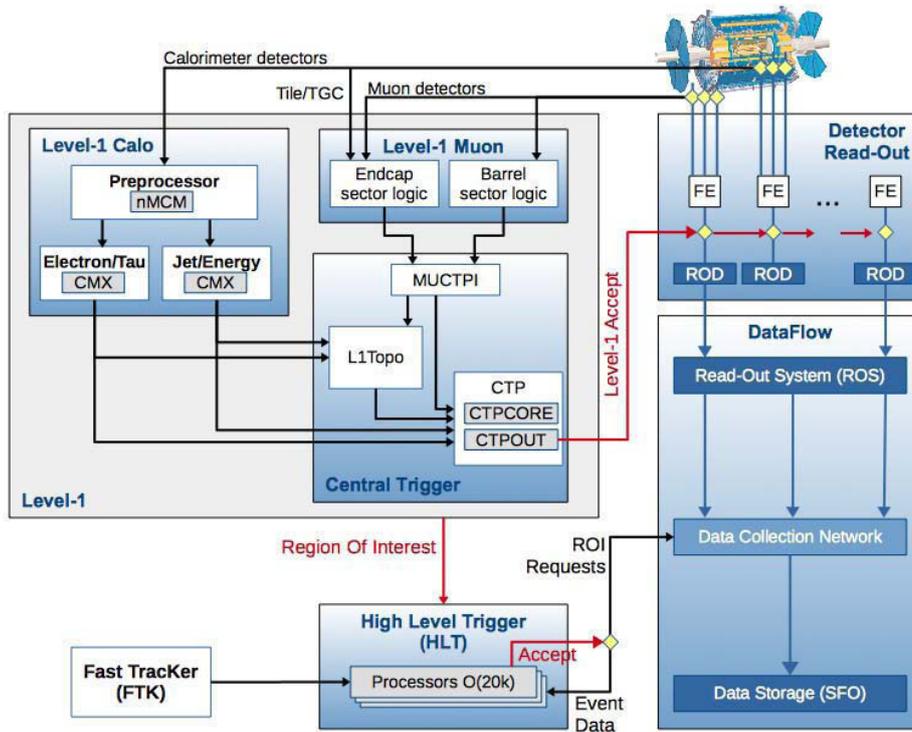


Figure 2.14: The ATLAS Trigger system.

Level-1 trigger

The Level-1 trigger performs an hardware-based online event selection using information of reduced granularity from all the calorimeters and the muon trigger system, to select events with

high- E_T objects (electrons, taus, jets, muons). The Level-1 calorimeter trigger (L1Calo) processes the real-time responses of the electromagnetic and hadronic calorimeters to produce inputs for the Central Trigger Processor (CTP). The initial system counts the physics object multiplicity above the E_T threshold for electrons, taus, jets. L1Calo has access to information of energy deposits from $\Delta\eta \times \Delta\phi$ regions of 0.1×0.1 , called “trigger towers”. The E_T of the object is estimated with the addition of adjacent trigger towers in a sliding window [68]. For different muon p_T thresholds, the Level-1 muon trigger uses the hits in the RPC and TGC to identify muons in the central region and in the endcaps respectively. Both L1Calo and L1Muon define Regions of Interest (ROI) in the plane (η, ϕ) , which correspond to the regions of the detector where the trigger algorithm have identified physics objects.

High-Level Trigger

The event satisfying the ROI multiplicity requirements and the E_T thresholds are passed to the HLT trigger [68]. It was composed by two different trigger farms during the Run 1 that were merged into a single farm for the Run 2. The HLT trigger takes decisions based on complex algorithms, using the full granularity and precision of the ATLAS detector (in association with the ROI defined at the Level-1 trigger). The trigger selections are optimized to minimize differences between the HLT and the offline analysis selections.

The chain of algorithms used to define a trigger selection is called “trigger chain”, and its name follows the convention:

$$[\text{LEVEL}]_{-}[\text{N TYPE}][\text{THRESHOLD}]_{-}[\text{QUALITY}]_{-}[\text{ISOLATION}]$$

where:

LEVEL: corresponds to the trigger level used (L1 or HLT).

N TYPE(S): indicates the type and multiplicity of object candidates.

THRESHOLD: corresponds to the transverse momentum threshold for a trigger selection.

QUALITY: indicates the rigour of requirement in the algorithms.

ISOLATION: indicates the isolation working point implemented.

Let's consider the following trigger chain "HLT_e60_lhmedium" to illustrate the conventions [69]. It selects events with a single electron of LHMedium quality (see section 4.3.2) and $p_T > 60$ GeV using the HLT algorithm, with any isolation requirement.

An important characteristic in the trigger system is the prescaling, which are used to control the rate of events passing a given trigger. For the loose lepton definition in the estimation of the QCD multi-jet background (section 5.6.1), a prescaled muon trigger is used (HLT_mu20_L1MU15) which has a time-constant prescaling factor of 10. Therefore, only one event is accepted every 10 events fulfilling the trigger requirements.

Lepton trigger

The online trigger system reconstructs and identifies electron and muon candidates. For electrons, the level-1 trigger uses the signals in the EM and hadronic calorimeters to calculate the E_T , where a given threshold can be set for different η regions. The E_T is computed within a granularity of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ in the central region. A veto on the hadronic leakage can also be applied at level-1 trigger by requiring a threshold on the fraction of energy measured in the hadronic calorimeter behind the core of the EM cluster.

Muon candidates are first formed using the hits in the tracking chambers at the MS pointing to the beam interaction region. Then, such muon candidate is combined with the track in the inner detector. The p_T measurement is refined with the implementation of weights obtained from the average p_T measured with the MS and the inner detector. In addition, the degree of isolation of the combined muon is obtained from the addition of the p_T of tracks in the inner detector within a cone centred around the muon candidate after its p_T subtraction. The muon triggers have a limited geometric coverage in the central region to allow space for detector supply services.

The lepton candidates at HLT are reconstructed and selected with a likelihood ratio or cut-based identification algorithms, with certain efficiency after accounting for residual background contamination. The trigger efficiency $\epsilon_{trigger}$ is defined as the fraction of events selected by a given trigger against the electrons or muons from the background sources. For illustration, the $\epsilon_{trigger}$ of the combined L1 and HLT e24_lhmedium.L1EM20VH trigger as a function of the E_T of the electron candidates is presented in Figure 2.15. The ratio of the efficiencies in data and simulation is used to correct the simulated samples.

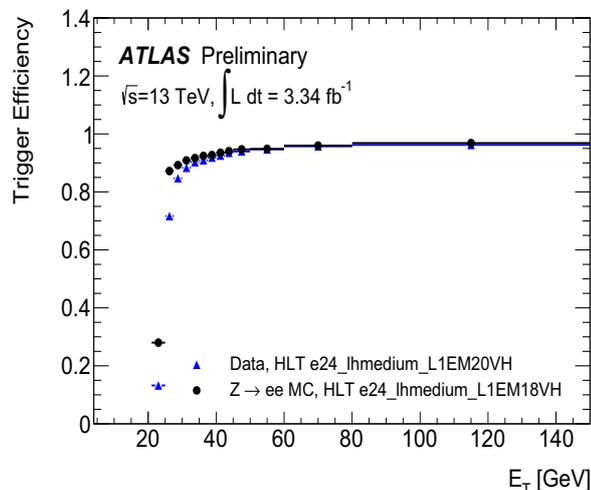


Figure 2.15: Efficiency of the combined L1 and HLT e24_lhmedium_L1EM20VH trigger as a function of the electron E_T . The offline reconstructed electron is required to pass $E_T > 24$ GeV and the LHmedium identification criteria (see section 4.3.2). The efficiency was measured with a tag-and-probe method using $Z \rightarrow ee$ decays and they are compared to simulations. The error bars show the combined statistical and systematic uncertainties [69].

2.2.8 Computing model

The Worldwide LHC Computer Grid (WLCG) is a network connecting many computer centres based at the countries which are members of the ATLAS Collaboration. It provides a computational infrastructure to store and analyse the data recorded by the particle detectors at LHC. The WLCG is structured in different layers, named “Tiers”, each with a specific set of tasks and services. The primary event (or raw data) processing occurs at CERN in the Tier-0 facility. Then, the datasets are copied to several Tier-1 facilities around the world, which have the responsibility to host and provide long-term access to the datasets. Finally, the data are copied to the Tier-2 facilities based at scientific institutes. They have enough computing resources for specific analysis tasks. In addition, there are additional computing resources based in the scientific institutes for data analysis, denoted as Tier-3 resources. In particular, Tier-2 and Tier-3 facilities are present at LPC, Clermont-Ferrand.

The main requirement on the ATLAS computing model is to enable the access to the raw data for monitoring, calibration and alignments activities, and access to the reconstructed data for analysis. The raw data format is transformed into the “Event Summary Data” format, which is based on the information from the ATLAS sub-detectors: calorimeter cells, and tracks in the inner detector and muon chambers. In addition, the data format are also stored in terms of the

reconstructed physics objects (see Chapter 4), named Analysis Object Data (AOD). The AOD are finally reduced in DxAOD, a format of smaller size that contains only the necessarily information for the analysis.

For the physics analyses, the datasets can be processed with ROOT [70], a software framework which provides the functionalities needed for the processing of big datasets, to perform statistical analysis, for data visualisation and storage, etc.

2.2.9 Data quality in 2015 data-taking

It is essential to understand which datasets are good to perform a physics analysis. A data quality (DQ) status flags are determined by the DQ representatives from each sub-detector system, which are used by the combined performance and trigger groups to declare the data as good/flawed/bad for the different physics objects. These flags are set for each luminosity block (approximately 2 minutes long fraction of runs), where the most low-level flags are based on detector control conditions, flagging possible hardware and data-taking problems such as nominal voltages, temperature, humidity, etc. Finally, a “Good Run List” (GRL) are built using all the luminosity blocks for which the sub-detectors useful for a given analysis were operational.

The relative fraction of luminosity associated to data of good quality delivered by the various sub-detector of ATLAS during the 2015 data-taking is shown in Figure 2.16. The total fraction of luminosity which is good for physics analysis correspond to 87.1 % of the total luminosity delivered by the LHC.

ATLAS pp 25ns run: August-November 2015										
Inner Tracker			Calorimeters		Muon Spectrometer				Magnets	
Pixel	SCT	TRT	LAr	Tile	MDT	RPC	CSC	TGC	Solenoid	Toroid
93.5	99.4	98.3	99.4	100	100	100	100	100	100	97.8
All Good for physics: 87.1% (3.2 fb⁻¹)										
<small>Luminosity weighted relative detector uptime and good data quality (DQ) efficiencies (in %) during stable beam in pp collisions with 25ns bunch spacing at $\sqrt{s}=13$ TeV between August-November 2015, corresponding to an integrated luminosity of 3.7 fb⁻¹. The lower DQ efficiency in the Pixel detector is due to the IBL being turned off for two runs, corresponding to 0.2 fb⁻¹. Analyses that don't rely on the IBL can use those runs and thus use 3.4 fb⁻¹ with a corresponding DQ efficiency of 93.1%.</small>										

Figure 2.16: Performance of the ATLAS detector for the 2015 data-taking. Runs with a bunch spacing of 25 ns are taken between August and November 2015, corresponding to a recorded integrated luminosity of 3.2 fb⁻¹ [71].

2.3 Outlook

The ATLAS detector is used for the identification and reconstruction of the particles coming from the pp collisions of LHC. The 2015 pp collisions recorded with ATLAS are used for the search for new resonances decaying into top quark pairs, where the good quality datasets used have an integrated luminosity of 3.2 fb^{-1} (Figure 2.16). The Chapter 3 is going to provide more details about the TileCal calibration and presents a study demonstrating the potential of new calibration technique, while identification and reconstruction of the physics objects used in this analysis are described in Chapter 4.

Chapter 3

The Tile Calorimeter and calibration by laser in-time runs

Contents

3.1	Tile calorimeter overview	58
3.2	TileCal calibration systems	61
3.3	The Laser system	63
3.3.1	Laser runs	64
3.3.2	Laser calibration	65
3.3.3	Statistical dependence of C_{laser} for laser in-time runs	66
3.3.4	Compatibility between C_{las} and C_{Cesium}	68
3.3.5	Source of systematic uncertainty of the C_{laser}	69
3.4	Statistical uncertainty studies using pseudo-data	72
3.4.1	Generation of the $R_{i,p}^{MC}$ distribution	73
3.4.2	The statistical dependence of the laser precision	74
3.4.3	Emulation of the Cesium scans	75
Outlook	76

Several calibration systems provide corrections to the energy measurements of TileCal. In particular, the laser system allows to monitor the linearity and stability of the PMTs and the associated electronic chain. In dedicated calibration runs, a well known laser light is sent to the photocathodes of the PMTs and deviations of the response are used to compute the calibration constants [72]. But the laser pulses can also be sent during the empty gap between the bunch crossings of the LHC (laser in-time runs). The usage of these runs could lead to a laser in-time calibration in the future, which could be done with an higher frequency than the usual laser calibration runs.

The author’s qualification task¹ is presented in this section, where the analysis exposed has converged in an internal ATLAS note [73]. The objective of this study is to obtain an estimate of the uncertainties in the calibration constants using the laser in-time runs, and determine if it is feasible to use such kind of runs to perform the calibration. The 2012 datasets are used for this analysis, which was taken with the first version of the Laser system. The laser system has been upgraded for Run 2, but the laser in-time runs are not used yet for the laser calibration.

3.1 Tile calorimeter overview

The geometry of the TileCal detector consists of a cylindrical steel-scintillator structure composed by a central barrel covering the region $|\eta| < 1$, and two extended barrels which covers the region $0.8 < |\eta| < 1.7$. The hadronic showers cross the scintillator tiles and induce the production of ultraviolet light, carried to the 9852 photomultipliers (PMTs) by wavelength shifting optical fibers.

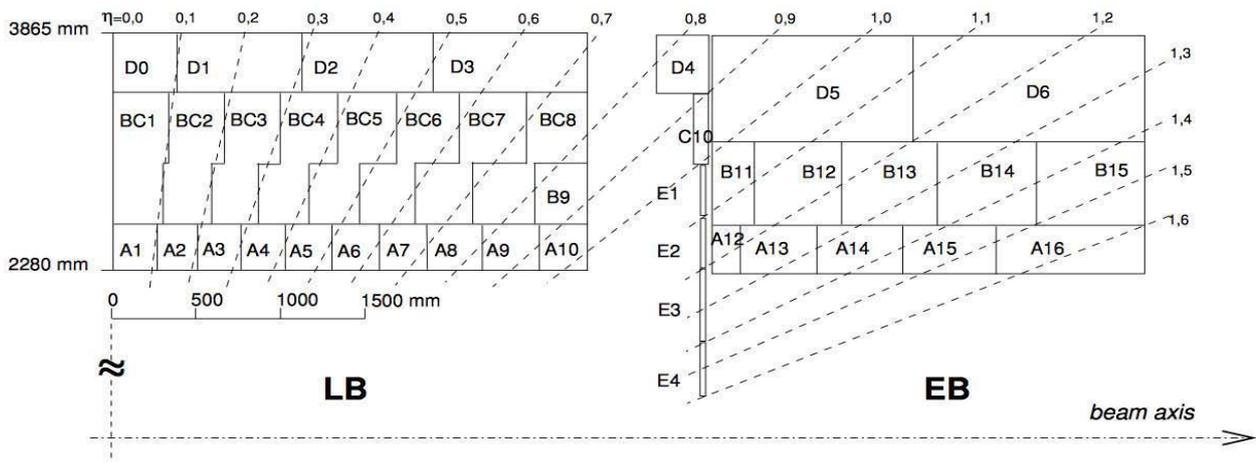


Figure 3.1: Scheme of cell distribution in ATLAS Tile calorimeter.

¹Necessary to be qualified as ATLAS author.

The TileCal is divided in four partitions: LBA and LBC (for side A and C of the long-barrel), and EBA and EBC (for side A and C of the extended barrels). Each partition is divided azimuthally into 64 modules. The modules are composed by a super-drawer where the PMTs and their electronic cards are placed. The modules contain 45 PMTs in the long-barrel partitions and 32 PMTs in extended barrel partitions. TileCal is also longitudinally segmented in three layers, with various cell granularity (Figure 3.1): the long-barrel is segmented in layers A, BC and D and the extended barrels in layers A, B and D. The cell granularity is $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ for the layers A and BC/B, and $\Delta\eta \times \Delta\phi = 0.1 \times 0.2$ for the layer D. In addition, the E cells are placed between the barrels to perform the energy measurements of the particles interacting with the non-instrumented part of the detector.

Optical system As mentioned previously, the TileCal is a sampling calorimeter using steel as absorber and scintillating tiles as active material. The scintillator tiles, made of polyester, are oriented to be perpendicular to the beam axis. The time response of the tiles correspond to the time spent in the de-excitation of the atoms ionised by the incoming particles, which is around 10 ns (smaller than the nominal proton bunch spacing). A total of approximately 460.000 scintillating tiles are installed in TileCal, almost half in the long-barrel, a quarter in each extended barrel and a small number for the cells between the barrels.

The ionising particles crossing the tiles induces the production of light in the UV wavelength range. The light propagates through the edges of the tiles where it is given by wavelength shifting (WLS) *optical fibers* to the PMT where a light mixer optimises the detection uniformity of the photocathode. The shift of the light to a longer wavelength is performed to match the sensitive region of the photocathode.

Photomultiplier The photomultiplier in the *PMT block* (Figure 3.2) transforms the incoming light from the scintillator tiles into an electrical response, which is digitised by the electronics inside the super-drawers. The PMT block is composed by the photomultiplier, a high-voltage divisor and an electronic board (named “*3-in-1*”); and is covered by μ -metal² to avoid the interference with the external magnetic field.

The photomultipliers are composed of a photocathode and several dynodes. The high-voltage divisor connects the PMT to the electronic card, and provides a voltage to each *dynode*. When the photons strike the photocathode, electrons are ejected by the photoelectric effect (photoelectrons), and are multiplied in the *dynodes* via the secondary emission process. The photoelectrons are

²Nickel-iron combination which allows the shielding of sensitive electronic equipment against static or low-frequency magnetic fields.

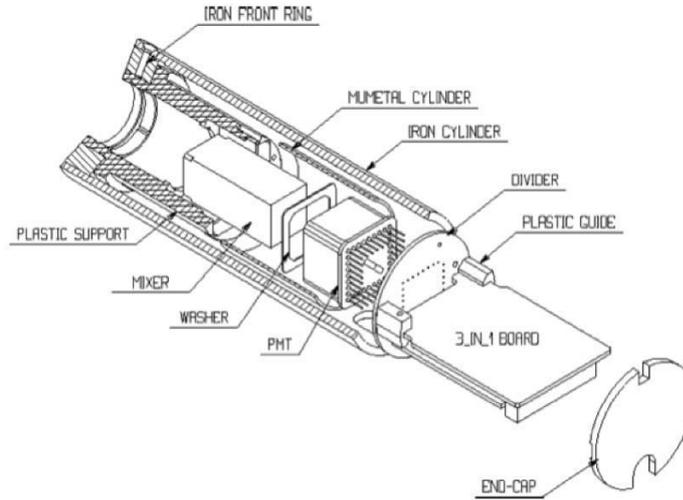


Figure 3.2: Schematic view of a PMT block.

collected by the anode located in the latest dynode. The multiplication factor of the photoelectrons is defined as the PMT’s gain, which should be $O(10^5)$ to ensure an optimal performance (photomultiplier R5900 Hamamatsu).

To ensure the required gain on each PMT, a specific high-voltage should be applied depending on its intrinsic characteristics. A high voltage system (Figure 3.3) supplies the PMTs with the proper voltage. The same high voltage source is used for all the PMTs per module (HV_{in}), having two possible values: $HV_{in} = -830$ V and $HV_{in} = -950$ V. Each module is provided with a HV_{micro} card which controls two other electronic cards called HV_{opto} . A regulation loop is located in the HV_{opto} cards to minimise the difference between the high voltage to be implemented to the PMTs (HV_{out}) and the expected value (HV_{set}). Finally, the PMTs are supplied with a high voltage between $HV_{in} - 360$ V $< HV_{out} < HV_{in} - 1$ V. A good regulation of HV_{out} is crucial for the performance of the energy measurements of the PMTs, as a variation $\Delta HV_{out} = 0.1$ V (drift) introduces a gain variation of 0.1% at a typical voltage of 700 V [76].

Read-out electronic The “3-in-1” electronic board in the PMT block forms the front-end of the electronics read-out chain. These boards are connected perpendicular by to the high voltage divisor (Figure 3.2), and have three main functions: the signal pulse shaping, a charge injection calibration and the integration of the PMTs signals for the Cesium calibration. The 3-in-1 boards have three analogical outputs, where two of them corresponds to outputs for signal amplification (low and high gain) and the other output for the level-1 trigger.

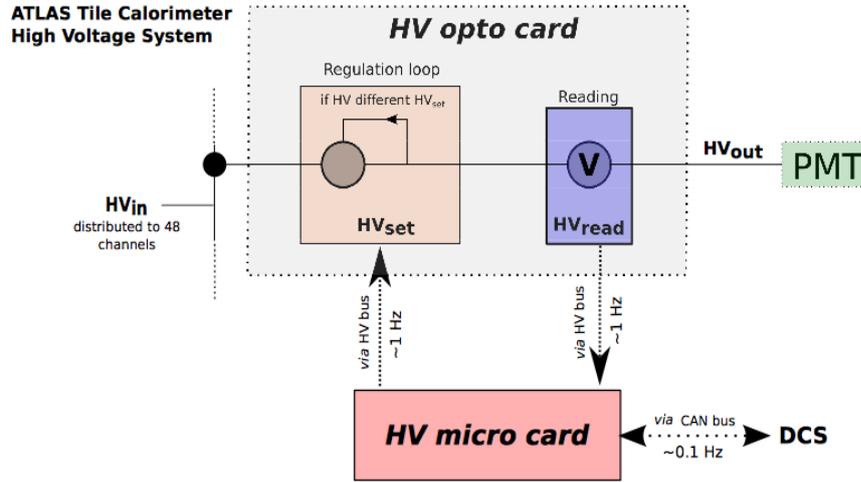


Figure 3.3: Scheme of the HV distribution systems of TileCal.

3.2 TileCal calibration systems

When a particle deposits its energy in the calorimeter, the electric output from the PMTs is converted into digital signals. But some variations of the optical and electronic responses could degrade the energy measurements. The TileCal calibration consists in finding the relation between the energy deposited by the particles and the digital signal value. As this response can change with time, the detector has to be calibrated frequently. TileCal has three calibration systems providing correction constants to the energy measurement covering the full hardware chain (Figure 3.4): the Cesium system, the laser system and the charge injection system (CIS) [74, 75].

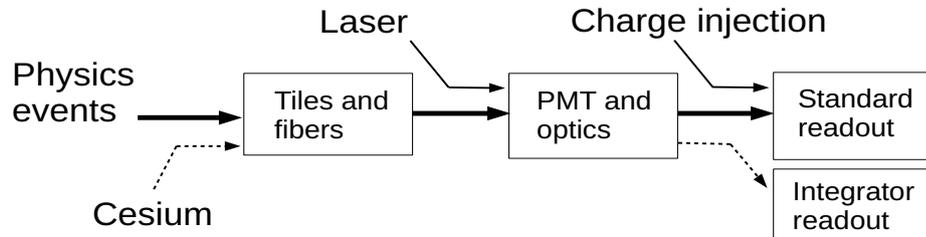


Figure 3.4: Hardware calibration chain of TileCal.

The Cesium calibration system: it allows the absolute calibration since it covers the full energy measurement chain for each channel and the injected energy is well know, but it uses a different readout system. It consists in a movable radioactive Cesium source ^{137}Cs (thanks to an hydraulic system) that irradiates the scintillating tiles with photons of 662 keV. Deviations of a cell response with respect to a reference Cesium scan are interpreted as variations of the channel overall gain, which are translated into calibration constants C_{Cesium} . This calibration is performed typically once per month since each Cesium scan takes long time-periods.

The laser system: the laser calibration provides corrections to the instabilities on the PMT responses and its associated electronic chain. Laser pulses can be sent to the PMTs during the proton-proton collisions or during dedicated laser-calibration runs. The wavelength of the laser light is close to the wavelength of the light coming from the scintillating tiles in typical LHC collisions. Deviations on the PMT response with respect to a reference laser run is used to obtain calibration constants C_{laser} (more details are exposed in section 3.3). The laser calibration is performed relative to a reference laser run, which is chosen to be the closest one to a Cesium scan.

The Charge Injection System: it simulates a physics signal by the injection of a known “charge” into the ADC. It leads to the calibration of the read-out electronics and provides a precise estimation of the electronic noise and linearity. It also provides a quantitative relationship between the analogical physical signals from the PMTs and the electronic response of the read-out system, i.e. the conversion from ADC counts to pC ($C_{ADC \rightarrow pC}$). The calibration constants are updated with a lower frequency than for the other calibration systems, since a global gain variation due to readout problems are not frequent. But the calibration is performed several times per week to monitor drifts in individual problematic channels.

The detector response is converted from ADC counts into energy units (GeV) combining the correction constants for each channel (E_i):

$$E_i = A_i \cdot C_{i,ADC \rightarrow pC} \cdot C_{pC \rightarrow GeV} \cdot C_{i,laser} \cdot C_{i,Cesium}, \quad (3.1)$$

where A_i is the signal amplitude in the channel i and $C_{pC \rightarrow GeV}$ is the conversion factor³ from charge into energy measured from the cell responses to electrons in test beam campaigns [76]. $C_{ADC \rightarrow pC}$ is determined by the CIS calibration with precision of about 0.7% [75]. A statistical precision of 0.3% [77] is estimated for the laser calibration constants C_{laser} using laser runs in the absence of collision with 10^4 pulses per run. Finally, the Cesium calibration provides energy corrections C_{Cesium} with a precision of 0.3% [78].

³Tilecal EM scale factor is 1.050 ± 0.003 pC/GeV.

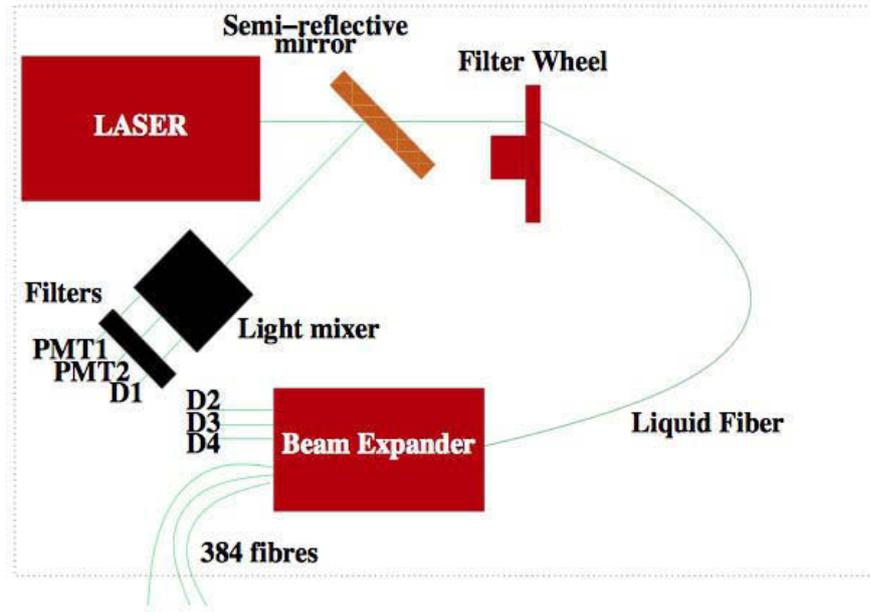


Figure 3.5: Schematic view of the laser system in Run 1.

3.3 The Laser system

The laser system used in Run 1 has been upgraded during the 2013 shutdown. It has been re-designed to improve the main shortcomings, as the inhomogeneity of the light splitting at the level of the distribution box, the inclusion of more photodiodes to measure the laser light at different stages of the optical path and the inclusion of a redundant internal calibration of the photodiodes [80]. The results of the analysis exposed have been obtained using the laser data before the upgrade, so a breve description of the laser system used in Run 1 is presented in this section [81].

In the laser box (Figure 3.5), the light source is a frequency-doubled infrared laser emitting a light beam with a wavelength of 532 nm, which is close to the optimal wavelength (480 nm) of the light transmitted by the WLS fibers. The energy of the laser pulses (few μJ) is enough to saturate the electronic of the PMT. A small fraction of the laser light is sent to the mixing block via a semi-reflecting mirror, to be measured by a Si photodiode⁴ (D^1) for the amplitude calibration and by two PMTs for timing purposes. The laser light which is not reflected by the mirror passes through a filter in the wheel, which contains seven different density filters with attenuation factors from 3 to 1000 to cover the possible energy spectrum. Then, the laser light is collected by a fiber which is linked to the distribution box⁵, splitting the laser light towards all the PMTs using 384

⁴The photodiodes are calibrated using an ^{241}Am source of α particles.

⁵Based on two lenses to expand the diameter of the laser light and a diffuser to prevent effects of the light

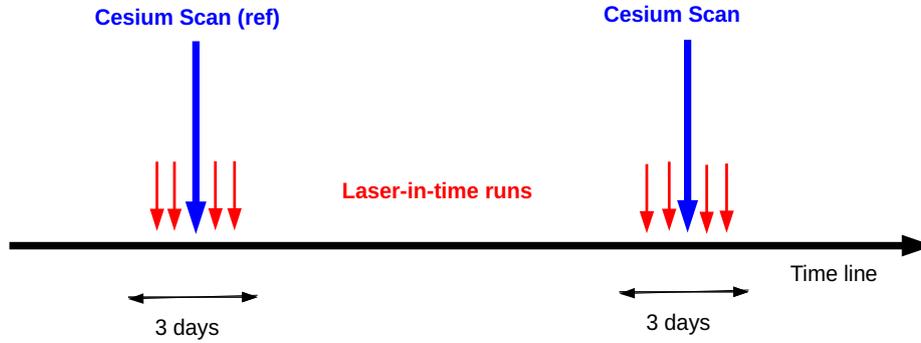


Figure 3.6: Time-line of the calibrations runs of TileCal. The laser-in-time runs are performed during physics runs, which can increase the frequency of the laser calibration.

optical fibers connected to a patch panel. Each module is fed by two fibers, 45 PMTs per fiber in long-barrel and 16 PMTs per fiber in the extended barrels. Other three photodiodes measure the intensity of the laser light after the light splitting in the distribution box.

3.3.1 Laser runs

The laser calibration system can be used in the following modes [81]:

Laser calibration runs: high frequency laser pulses are sent during the calibration periods. The laser runs using a filter with attenuation factor of 3 are considered as low gain read-out (LG) and those using a filter with attenuation factor of 330 as high gain read-out (HG). Each of these runs has 10^4 and 10^5 pulses respectively. The compatibility between the deviations detected using the LG and HG laser runs is demanded to validate the calibration for a channel.

Laser runs in physics data taking: low frequency laser pulses are sent during the empty gaps between the bunch crossings of the LHC. Such kind of runs are denoted as laser in-time runs, they are typically used to measure the timing on the detector responses to perform time calibrations channel per channel. The filter used has an attenuation factor of 100. Since the bunch spacing time and the physics runs are limited, the number of pulses in the laser in-time runs is also limited.

coherence.

Run number	Initial date	Final date	Number of events
207975	2012-08-03 23:53:40	2012-08-04 12:48:39	44160
208015	2012-08-05 02:33:50	2012-08-05 05:29:34	10129
208123	2012-08-06 15:02:20	2012-08-07 00:25:26	32335
208125	2012-08-07 00:46:07	2012-08-07 03:22:06	8998

Table 3.1: Laser in-time runs closest to the Cesium scan of Aug 5th 2012. Selected as reference to get the laser gain variation.

Run number	Initial date	Final date	Number of events
208484	2012-08-12 22:42:52	2012-08-13 00:12:18	5078
208485	2012-08-13 00:32:49	2012-08-13 10:25:18	33462
208629	2012-08-15 03:36:14	2012-08-05 09:02:00	18802
208631	2012-08-15 09:22:33	2012-08-15 11:45:09	8133

Table 3.2: Laser in-time runs closest to the Cesium scan of Aug 14th 2012.

3.3.2 Laser calibration

The “direct method” is used to compute the gain variation with respect to the reference runs for both types of run. After the pulse-by-pulse subtraction of the pedestals, the normalised response $R_{i,p}$ for each channel is computed:

$$R_{i,p} = \frac{E_{i,p}^{PMT}}{D_p^1}, \quad (3.2)$$

where $E_{i,p}^{PMT}$ is the PMTs response for the channel i and the pulse p , and D_p^1 is the response of the photodiode D^1 to the pulse p . The mean value of the $R_{i,p}$ distribution, defined as R_i , is used to compute the gain variation (Δ_i) for each channel i :

$$\Delta_i = \frac{R_i - R_i^{ref}}{R_i^{ref}}, \quad (3.3)$$

where R_i^{ref} is the value of R_i for the collision run chosen as reference.

The light splitting in the distribution box is not homogeneous and varies as function of time. Therefore, the gain variations are corrected by the fiber correction. It is computed using the PMT response of the most distant cells⁶ to the interaction point, assuming that the luminosity has not

⁶The D cells for long-barrel and the corner cells B13-15 and D5,6 for extended barrels.

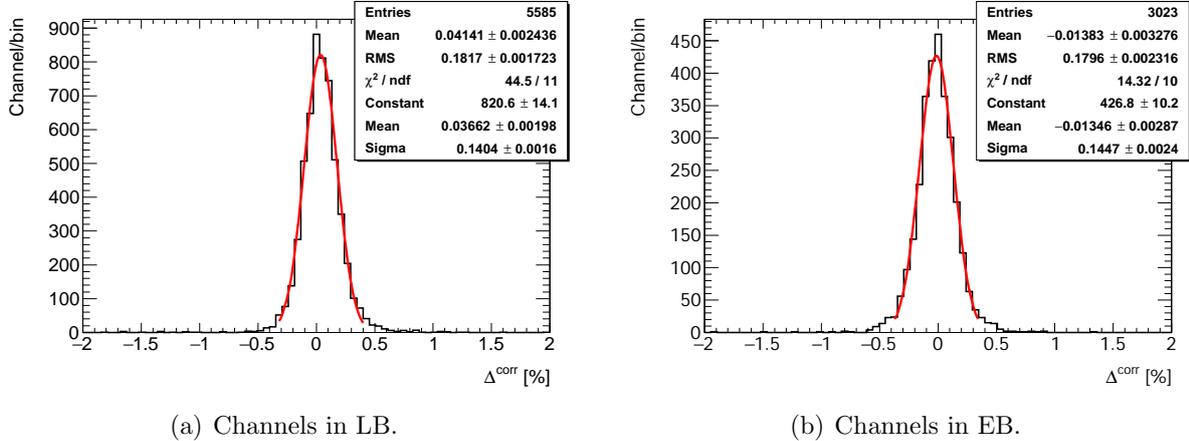


Figure 3.7: Δ^{corr} distributions corresponding to the pair of laser in-time runs.

an important radiation effect on its PMTs. The stability assumption on these channels is not fully valid since some PMTs are unstable in reality (due to HV variations for example). These PMTs are removed from the $\Delta_{f(i)}$ computation with an iterative procedure [72].

$$\Delta_i^{corr} = \Delta_i - \Delta_{f(i)}, \quad (3.4)$$

and finally, the laser calibration constants are computed with the corrected gain variation:

$$C_{laser}^i = \frac{1}{1 + \Delta_i^{corr}}, \quad (3.5)$$

3.3.3 Statistical dependence of C_{laser} for laser in-time runs

As mentioned before, the objective of this study is to determine if the statistical precision of the C_{las} derived with laser-in-time runs is good enough to be used. But there were not many laser in-time runs in the 2012 pp collisions with enough laser pulses. At least 5000 pulses are demanded to choose the laser in-time runs, which represents 50% of the pulses sent in LG and 5% in HG laser runs in calibration periods. The laser-in-time runs summarised in Tables 3.1 and 3.2 were found to be closest to the Cesium scans on Aug 5th (set as reference) and Aug 14th 2012 (Figure 3.6). The laser in-time runs selected are in a three days range around the Cesium scan to check whether the change in the response seeing by the laser in-time is compatible with the one seeing by the Cesium scans.

Statistical uncertainty of laser in-time calibration

The statistical uncertainty of the laser calibration can be computed as follows: for the channel i , the means $R_{i\alpha}$ and standard deviation $RMS_{i\alpha}$ of the normalised response distributions are considered for two laser runs $\alpha = 1, 2$ denoting the reference run as $\alpha = 1$. Then, the statistical uncertainty on $R_{i\alpha}$ depends on the number of pulses N_α (see Eq 3.2):

$$\sigma_{i\alpha} = \frac{RMS_{i\alpha}}{\sqrt{N_\alpha}}. \quad (3.6)$$

The laser runs are independent from each other, so the statistical uncertainty on the gain variation for the channel i (σ_i^{las}) can be estimated from a quadratic sum of the statistical uncertainty for each run (see Eq 3.3):

$$\sigma_i^{las,stat} = \sigma_{i,1} \oplus \sigma_{i,2}, \quad (3.7)$$

The $RMS_{i,1}$ and $RMS_{i,2}$ are of the same order of magnitude, so the previous expression can be approximated as:

$$\sigma_i^{las,stat} \propto \sqrt{\frac{1}{N_1} + \frac{1}{N_2}}. \quad (3.8)$$

If this statistical uncertainty is significant compared to the other sources of systematics, then it should be seen when the dispersion of Δ_i is drawn as function of $\sqrt{\frac{1}{N_1} + \frac{1}{N_2}}$ (as long as it can be assumed a dispersion of the true responses is not large).

The distributions of the corrected gain variation Δ^{corr} (Eq. 3.4) in a time period of two weeks is derived for the selected laser in-time runs (Table 3.1 and 3.2), excluding the pathological channels.⁷ A fit with a Gaussian function is used to obtain the spread on the Δ^{corr} distributions for the channels in long- and extended barrels (Figure 3.7). The fits were performed in a ± 2 RMS window to exclude contributions from the channels with large drifts.

The spread on the Δ^{corr} distribution obtained from the fit results for cells in the seconds layer⁸ (BC- and B) are shown in Figures 3.8 as function of $\sqrt{\frac{1}{N_1} + \frac{1}{N_2}}$. The $\sigma^{las,stat}$ values present a small linearity dependence, which means that some systematic uncertainties and/or the dispersion of the true response dominate over the statistical uncertainty.

⁷The pathological channel list is provided by the data quality group in 2012.

⁸It is shown in section 3.3.4 that splitting the channels by layers leads to a better fit with a gaussian function.

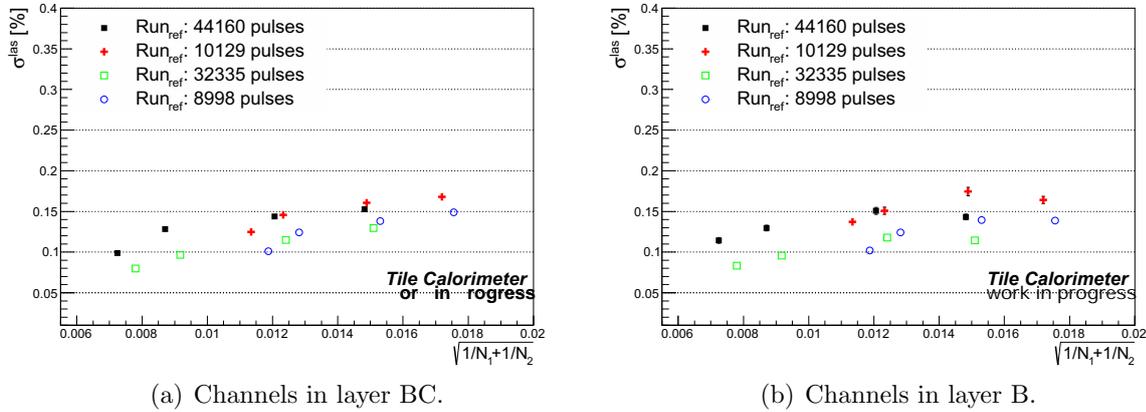


Figure 3.8: $\sigma^{las,stat}$ as function of $\sqrt{\frac{1}{N_1} + \frac{1}{N_2}}$, using the fit results of the laser gain variation distributions. The dots of the same mark are correlated, having the same reference run. The error bars correspond to the uncertainty provided by the χ^2 fit.

3.3.4 Compatibility between C_{las} and C_{Cesium}

A way to check whether the dispersion of the true response can explain the previous result, is to compare the gain variation seen by the laser and Cesium systems. Indeed the effect of the variation of the true response cancel out in the ratio (computed for each channel):

$$\rho_i = \frac{C_{laser}^i}{C_{Cesium}^i}. \quad (3.9)$$

The compatibility between the calibration systems cannot be perfect, since several effects can play an important role in this comparison, for instance the resolution of each calibration systems, any effects detected only by the Cesium system, or the optical effects due to the light splitting in the distribution box of the laser system.

The ρ distributions are fitted by a Gaussian function from which the parameters $\mu^{las/cs}$ and $\sigma^{las/cs}$ are extracted (Figure 3.9). The fit is performed in a ± 2 RMS window excluding channels with incompatible deviations from laser and Cesium systems. But the χ^2/ndf of the fit indicates that they differ from a pure Gaussian function. Consequently, from now on the ρ distributions are considered after splitting the channels in layers (Figures 3.10 and 3.11 for layers in the long- and extended barrel partitions respectively). The values of χ^2/ndf are then closer to the unity, demonstrating the shape of the ρ distributions are almost Gaussian (except for layer A in EB where the cells A12-A14 drift more than the others).

The $\sigma^{las/cs}$ values shows little dependence with respect to the number of laser pulses as showed in Figure 3.12. However the systematic associated to the Cesium calibration⁹ should be subtracted as:

$$\sigma^{las/cs} = \sigma^{las} \oplus \sigma^{Cesium}. \quad (3.10)$$

Using the pair of runs with the lowest statistics, a value of 0.25% is used to roughly estimate $\sigma^{las/CS}$, leading to the laser gain variation of $O(0.15 \%)$, consistent with the one obtained in section 3.3.3.

3.3.5 Source of systematic uncertainty of the C_{laser}

One of the main sources of systematics uncertainties that affect C_{laser} has been identified as originating from the assumption that the PMTs of the distant cells are not affected by the radiation. Any deviation from this assumption would lead to a systematic bias, that can be estimated by evaluating how much the average calibration of these cells differ from 1, i.e. by considering the $|\mu^{las/cs} - 1| = \langle |\Delta_{f(i)}| \rangle$ distributions for the distant cells. Using the fit results of the ρ distributions in the cells belonging to layer D for long-barrel and to the corner cells (B13-15 and D5-6) for extended barrel, the $|\mu^{las/cs} - 1|$ distributions are presented in Figure 3.13. An estimation of the systematic uncertainty is extracted from the respective mean values: 0.09% for long- and 0.17 % for extended barrels, consistent with the results of the previous sections.

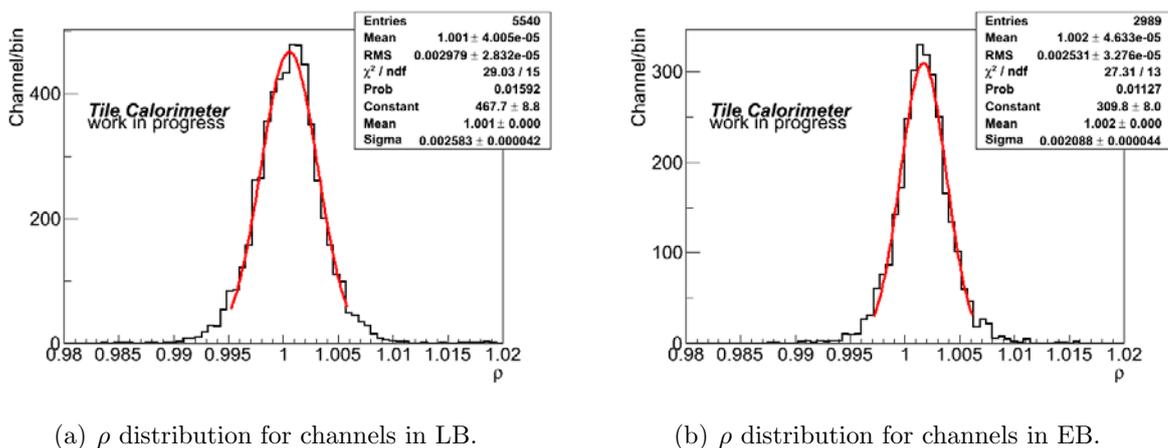
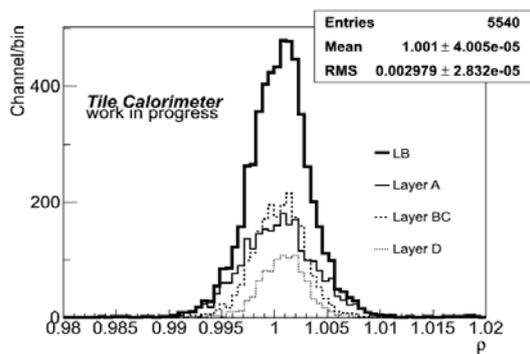
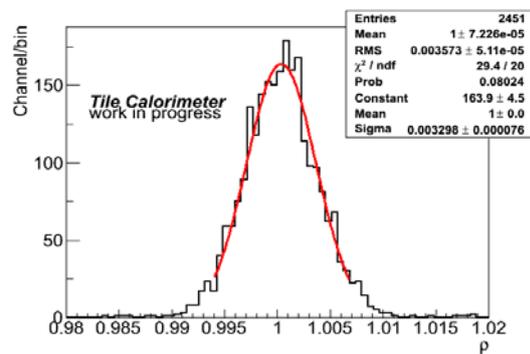


Figure 3.9: ρ distributions for the channels in TileCal fitted with a Gaussian function.

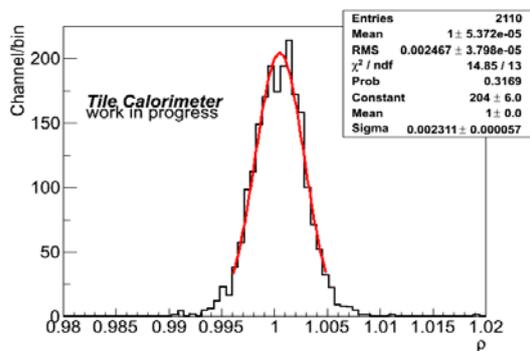
⁹A $\sigma^{Cesium} \sim 0.2 \%$ value is taken from a private communication with Cesium calibration experts.



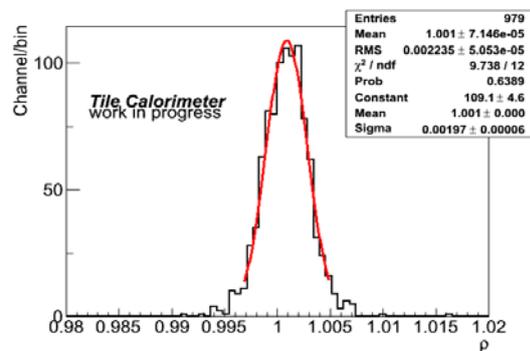
(a) ρ distribution in layer components.



(b) Fit result for the ρ distribution in layer A.

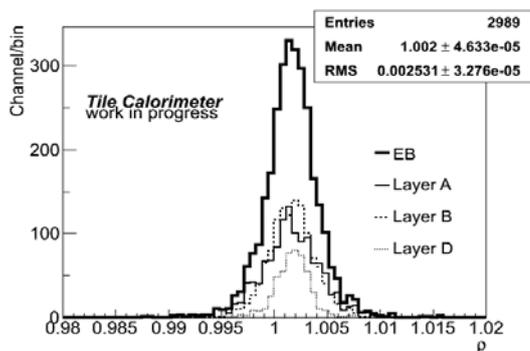


(c) Fit result for the ρ distribution in layer BC.

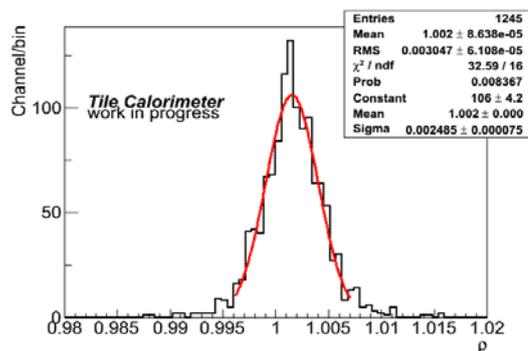


(d) Fit result for the ρ distribution in layer D.

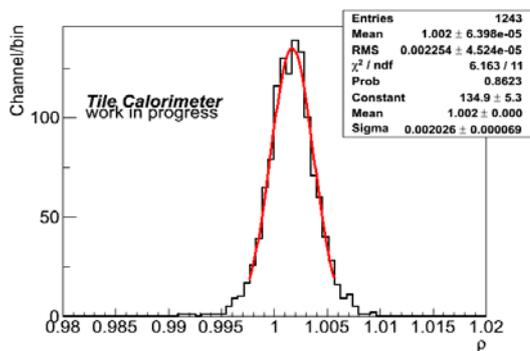
Figure 3.10: ρ distribution separated in layer components for channels in LB. The laser in-time runs used to compute the gain variation are: 208125 and 208629.



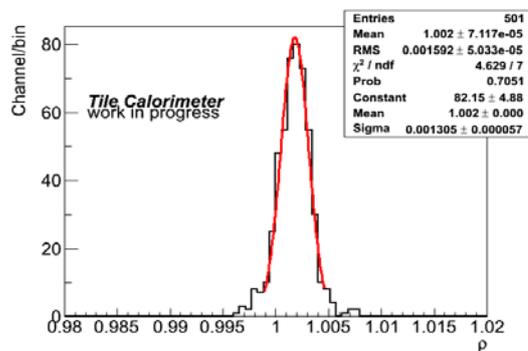
(a) ρ distribution in layer components.



(b) Fit result for the ρ distribution in layer A.



(c) Fit result for the ρ distribution in layer B.



(d) Fit result for the ρ distribution in layer D.

Figure 3.11: ρ distribution separated in layer components for channels in EB. The laser in-time runs used to compute the gain variation are: 208125 and 208629.

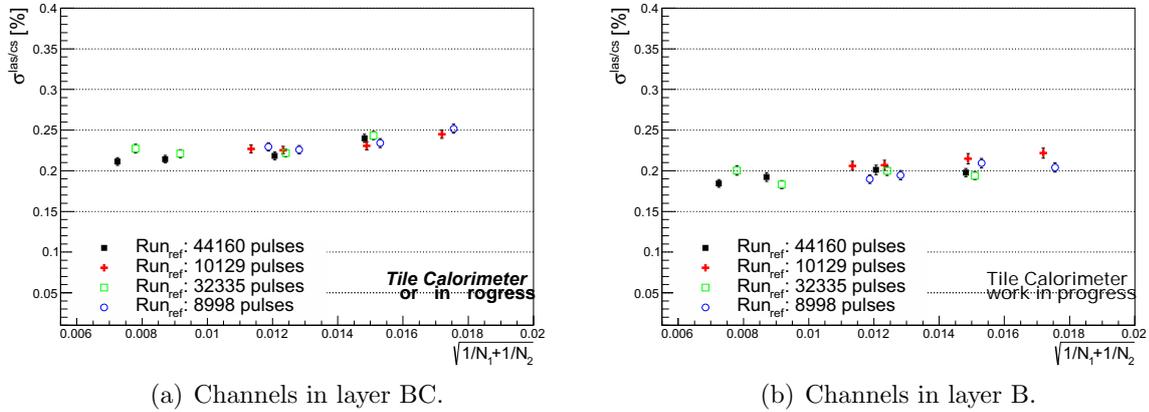


Figure 3.12: $\sigma^{las/cs}$ as function of the $\sqrt{\frac{1}{N_1} + \frac{1}{N_2}}$, using the fit results of the ρ distributions.

3.4 Statistical uncertainty studies using pseudo-data

The results of the previous section have been checked with an emulation of the effects that can alter the laser and Cesium calibration. Laser pseudo-data are generated to emulate the configuration used to estimate the statistical precision of a laser-in-time calibration. The response R_i^{MC} belonging to different emulated laser runs are shifted to reproduce a typical gain variation (following the results from section 3.3.3). Then, the effects due to the fiber correction and Cesium calibration are emulated and introduced progressively to check the impact on the estimation of the statistical precision.

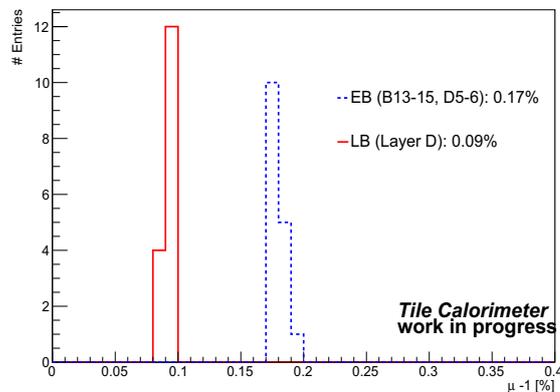


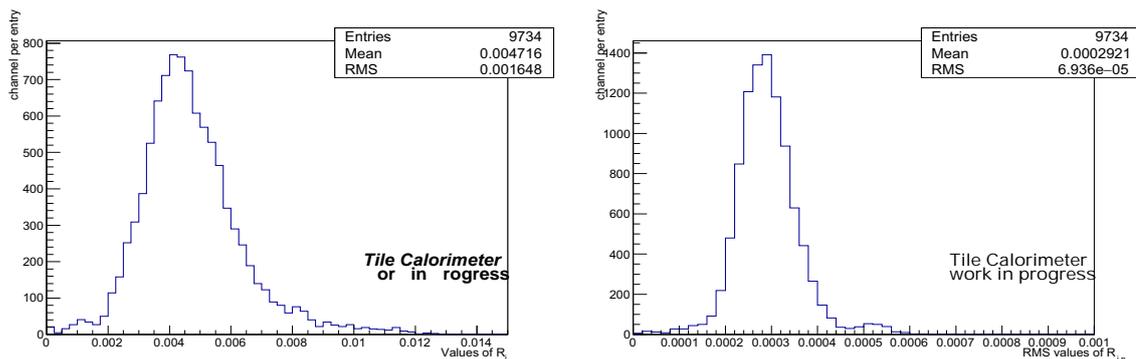
Figure 3.13: $|\mu^{las/cs} - 1|$ distribution used to estimate the systematic uncertainty.

3.4.1 Generation of the $R_{i,p}^{MC}$ distribution

As mentioned in section 3.3.2, the corrected gain variation measured using the laser system is computed using the mean of the $R_{i,p}$ distribution over certain number of pulses (R_i). An emulated gain variation distribution (Δ_{las}^{MC}) of 0.2 % is used to reproduce a typical value.

The R_i^{MC} value of a channel i can be randomly generated using a Gaussian function with parameters μ_{las}^{MC} and σ_{las}^{MC} , which are extracted from a typical laser in-time run (Figure 3.14). Then for the other emulated laser run needed to compute $\Delta_{i,las}^{MC}$, R_i^{MC} is generated following the same method but increasing μ_{las}^{MC} by 0.2% in the mean value but keeping the same σ_{las}^{MC} . This step is reproduced for all the 2200 channels generated.¹⁰

Finally, the $R_{i,p}^{MC}$ distributions are generated with different numbers of laser events. Several sets of simulated run pairs are generated with different numbers of laser events, covering the observed range in data: 5k, 10k, 20k, 30k, 35k, 40k and 45k pulses. For instance, a $R_{i,p}^{MC}$ distribution generated with 45000 laser events is shown in Figure 3.15.



(a) R_i , i.e. the mean values of the $R_{i,p}$ distribution.

(b) RMS values of the $R_{i,p}$ distributions.

Figure 3.14: Distributions of the mean values (a) and RMS's (b) of the $R_{i,p}$ distributions for all the channels in run 208125. The μ_{las}^{MC} parameter is extracted from the mean value of the R_i distribution (4.7×10^{-3}), while the σ_{las}^{MC} parameter is extracted from the mean value of the RMS's of $R_{i,p}$ distribution (2.9×10^{-4}).

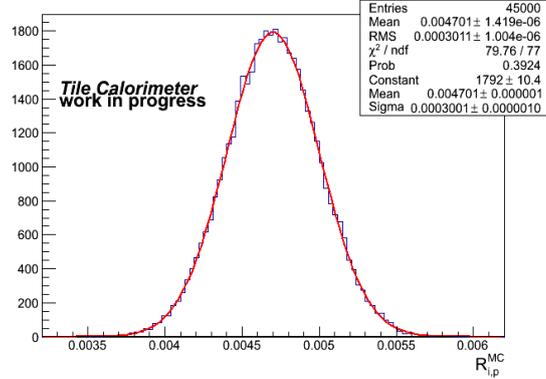
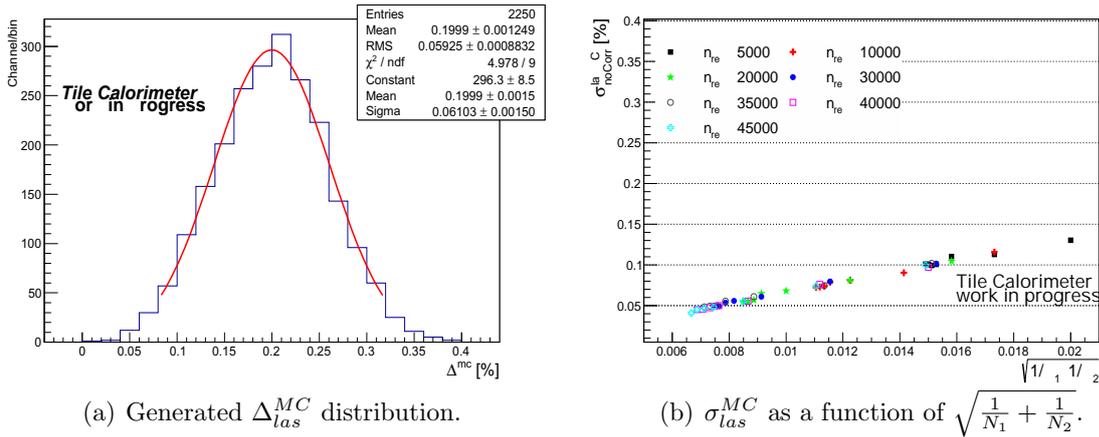


Figure 3.15: Normalized response distribution $R_{i,p}^{MC}$ generated for a channel with 45000 laser pulses.

3.4.2 The statistical dependence of the laser precision

With the pairs of the $R_{i,p}^{MC}$ distributions for the different number of laser events in the range selected, the Δ_{las}^{MC} distributions are computed. For instance in Figure 3.16 (a), the Δ_{las}^{MC} distribution is obtained from a pair of simulated runs with 20k and 35k laser events. Once the Δ_{las}^{MC} distributions are obtained using all possible combination between the laser runs, it is possible to reproduce the statistical precision studies using the σ_{las}^{MC} values extracted from a fit with a Gaussian function



(a) Generated Δ_{las}^{MC} distribution.

(b) σ_{las}^{MC} as a function of $\sqrt{\frac{1}{N_1} + \frac{1}{N_2}}$.

Figure 3.16: Statistical studies considering only the laser calibration without the emulation of the fiber corrections (MC laser data).

¹⁰Number of channels in the BC layer.

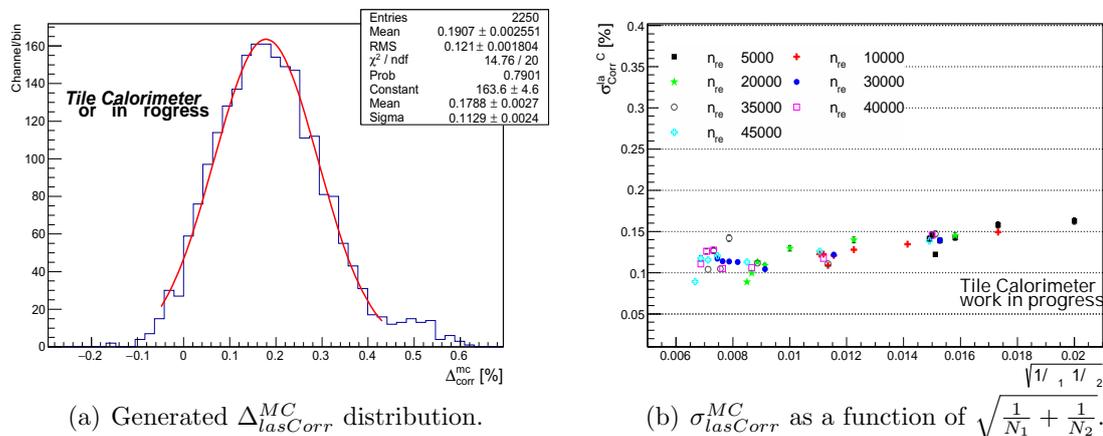


Figure 3.17: Statistical studies including the fiber corrections to the gain variation (MC laser data).

(Figure 3.16 (a)). The statistical dependence of the emulated laser deviation is shown in Figure 3.16 (b). The slope is due to the statistical precision which is improved when N_1 and N_2 increases. Notice that for a large number of laser pulses for both runs, the $\sigma_{\text{las}}^{\text{MC}}$ values are 50 % smaller than for the laser data (Figure 3.8).

Emulation of the light splitting correction

As second step, the emulation includes an estimation of the uncertainty on the fiber corrections to determine its impact on the calibration precision. Typically, the uncertainty on this correction is of order of 0.1% over one month [77]. The emulation of the gain variation due to the fibers is obtained with a random gaussian generator, using as parameters $\mu = 0.0$ and $\sigma = 0.001$ to introduce a noise of the same magnitude than the error on the fiber corrections. All the channels associated to a same fiber get their response varied by the same random number. The statistical studies using pseudo-data (Figure 3.17 (b)) shows an expected smaller slope when the corrections are added. The impact of the light splitting corrections on the gain variation is not negligible.

3.4.3 Emulation of the Cesium scans

For the last case, the emulation of the Cesium calibration allows to obtain the ratio of the emulated calibration constants ($\rho^{\text{MC}} = C_{\text{las}}^{\text{MC}} / C_{\text{Cesium}}^{\text{MC}}$). A Gaussian function is used to emulate the measurements of PMT response by the Cesium system. The $\mu_{\text{Cesium}}^{\text{MC}}$ parameter is set to obtain the same gain variation chosen for the emulation of the laser runs (0.2 %). In addition, the $\sigma_{\text{Cesium}}^{\text{MC}} = 0.2$

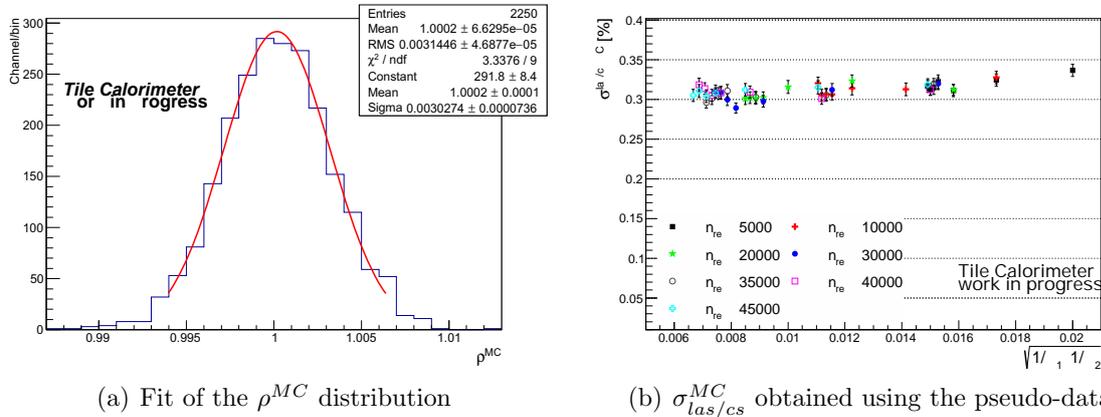


Figure 3.18: Statistical studies including fiber corrections and the uncertainty of the Cesium calibration (pseudo-data).

% is chosen for both generated Cesium scans to match the precision of the calibration constants in the BC layer (see section 3.3.4). Then, a random gaussian generator is used to emulate the Cesium scans considering only one entry per channel.¹¹

The Figure 3.18 (a) shows the ρ^{MC} distribution including the light splitting corrections in the laser emulations. The $\sigma^{MC}_{las/cs}$ values extracted from these distribution are of the same order of magnitude than for the data. The Cesium uncertainty brings a flat contribution (Figure 3.18 (b)) and it has a large impact compared to the statistical uncertainty of the laser calibration. This rough emulation explains the small improvement on the $\sigma^{las/cs}$ seen in the data when the number of laser events is increased.

Outlook

The TileCal is an important detector in particular for a search for $t\bar{t}$ resonances, and its good calibration is crucial to be able to spot a small excess over the large $t\bar{t}$ background.

In this chapter, it has been demonstrated that a calibration using the laser in-time runs could work as good as the usual laser calibration, that would allow to get more frequent calibrations, with the particular interesting feature that the calibrations would be done with the same experimental condition than the collision data.

¹¹Since the Cesium system does not depend on the laser pulses.

CHAPTER 3. THE TILE CALORIMETER AND CALIBRATION BY LASER IN-TIME RUNS

There is an improvement of the laser light splitting box for the second version of the laser system for Run 2. Therefore, we expect to avoid the optical corrections to the gain variation and the associated uncertainty in the coming laser analysis. Also, another step for this analysis could be the comparison between the laser standalone runs and the collision runs (This was not possible with 2012 data as different attenuation filters were used for laser in-time and the standard laser calibration runs).

Chapter 4

pp collisions: simulation and physic object identification

Contents

4.1	Simulation of the proton-proton collisions	81
4.1.1	Event simulation	81
4.2	Track and vertex reconstructions	84
4.3	Electrons	85
4.3.1	Electron reconstruction	85
4.3.2	Electron identification	86
4.4	Muon	87
4.4.1	Muon reconstruction	87
4.4.2	Muon identification	88
4.5	Lepton isolation and correction	89
4.5.1	Lepton isolation	89
4.6	Jets of hadrons	91
4.6.1	Jet reconstruction algorithms	92
4.6.2	Jet calibration	93
4.6.3	Track-based pile-up jets suppression	94
4.6.4	Large-R jets and Top-tagging	96
4.6.5	b-tagging algorithm	97
4.7	Overlap removal	99
4.8	Missing transverse energy	100

CHAPTER 4. PP COLLISIONS: SIMULATION AND PHYSIC OBJECT IDENTIFICATION

4.9 Outlook 101

The proton-proton collisions at LHC will be compared to the theoretical predictions. In order to be able to perform accurate comparisons, the theoretical prediction can be fully simulated, from the interaction of the partons inside the protons, until the description of the full detection chain of the ATLAS detector. The first part of this Chapter provides an overview of the fundamental concepts for the production of Monte Carlo simulations, which are used for the results of the physics analysis exposed in Chapter 5. In a second part, the “off-line event reconstruction” that is applied both on data and simulation will be also described. The information from the detectors has to be processed with algorithms to reconstruct and to identify the physics objects used to develop the analysis: tracks, primary and secondary vertices, jets, electron, muon and “neutrino”.

4.1 Simulation of the proton-proton collisions

Events from Monte Carlo simulations are extensively used, for example for developing algorithms for the physics object reconstruction, for estimating the sensitivity of physics analyses or for estimating the background composition of the data. An event simulation takes into account the full chain from the production and decay of particles according to a given process, as well as the hadronisation and the interaction with the detector.

The event simulation (Figure 4.1) begins with the proton-proton collisions, where the interaction between the partons (quarks and gluons) from each proton is known as the hard interaction [82]. The scattering of the partons with large momentum exchange; or the production of massive particles are the outcomes from the hard interactions. The radiation emitted by the incoming partons before the collision is known as initial state radiation (ISR), while the radiation emitted by the partons created in hard interactions is known as final state radiation (FSR). The outgoing colored particles from the hard interactions evolve in parton showers, where a non-perturbative interaction converts the showers into outgoing hadrons. Additional soft activity are produced in addition to the hard scatter processes, know as underlying event (composed mainly of multi-parton interactions and beam remnants) and pile-up (PU) events.

4.1.1 Event simulation

Hard scatter

An event simulation typically begins with a simple subprocess produced from high energetic collision of the partons. The hard process (HS) with a large momentum exchange can be described by the perturbative QCD theory, since the partons behave with asymptotic freedom. The cross section for a pp scattering is then given by:

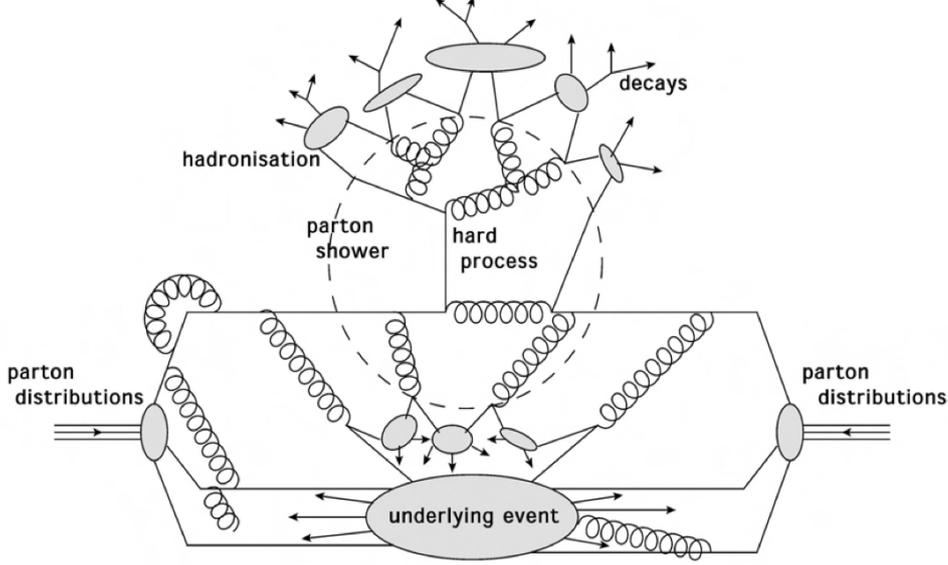


Figure 4.1: General structure of the proton-proton collisions. The partons inside the protons are initially held together by the strong force mediators. In the collision, those partons can interact producing large momentum transfers (hard processes), producing interesting events for the various physics analysis.

$$\sigma_{pp \rightarrow X} = \sum_{a,b} \int dx_1 dx_2 \int f_a^{h_1}(x_a, \mu_F) f_b^{h_2}(x_b, \mu_F) d\hat{\sigma}_{q_a q_b \rightarrow X}(\mu_F, \mu_R) \quad (4.1)$$

where $f_k^h(x_k, \mu_F)$ are the Parton Distribution Functions (PDF) associated to the incoming partons k belonging to the proton h . The PDFs and the factorisation scale μ_F parametrize the probability for a parton k to have a fraction x_k of the proton momentum. $\hat{\sigma}_{q_a q_b \rightarrow X}(\mu_F, \mu_R)$ is the parton-level cross section for the production of X from the initial partons a and b , which depends on μ_F and the renormalisation scale μ_R and computed from the associated matrix elements.

Parton showers

The colored partons (from the HS, ISR and FSR can radiate) can radiate gluons, and a gluon can produce $q\bar{q}$ pairs. Both phenomena lead to a parton showers. These showers represent higher-order corrections to the hard subprocess, but it is not feasible to calculate exactly these complex corrections. Instead, the most dominant contributions are considered for the HS simulation.

Hadronisation

The hadronisation refers to the phenomenological models used by an event generator to describe the transition from the partonic final state to an hadronic final state. The perturbation theory cannot be used to describe the hadronisation since it happens at large α_s scales. The string and cluster models are used to describe the hadronisation process [84]. The first one represents the strong force by a string between partons and the second one is based on the colored structure of the showers. Then, the results of the simulation of the hadron decays in lighter particles can be used as input for detector simulations.

Monte Carlo simulation softwares

The generators can be devoted to describe a part or all the steps of generation of the collision simulation. In the incoming Chapters, combinations of the following generators are used:

Powheg [87]: Hard scatter Monte Carlo generator with fixed next-to-leading order (NLO) QCD computations, used in this analysis for the SM $t\bar{t}$ and single-top productions.

MC@NLO [90]: a Monte Carlo event generator with NLO in QCD calculations of the matrix elements. In this analysis it is used as an alternative to Powheg in order to get an estimate of the uncertainty that could affect the modelling of the hard scatter..

Pythia 6 [85] and 8 [86]: standard tool for the generation of high-energy collisions, including a coherent set of physics models for the evolution from a few-body hard process to a complex multi-hadronic final states. The proton-proton configuration use a matrix element at leading order (LO) to model the hard-scatter process. There are libraries of containing the matrix elements at leading order (LO) hard processes, models for the initial- and final-state parton showers, the multiple parton-parton interactions and the particle decays.

Herwig [88]: general-purpose Monte Carlo event generator at LO matrix element, which includes the simulation of hard scattering processes (lepton-lepton, lepton-hadron and hadron-hadron) and soft hadron-hadron collisions. It is supplemented with an angular-ordered parton shower simulation and the cluster model is implemented for the hadronisation process.

Sherpa [89]: Monte Carlo event generator for the Simulation of High-Energy Reactions in Particle collisions. The list of physics processes that can be simulated with Sherpa covers all reactions

in the Standard Model and many theories beyond. It is used for simulations of the hard scatter processes and parton shower.

Full/Fast ATLAS simulations [91]: are softwares developed for simulation of the particle interaction with the ATLAS detector and the digitation of the energy deposited in the sensitive region of the detector. The full simulation uses a software package for GEometry ANd Tracking (GEANT4) [92] using Monte Carlo methods to accurately simulate the passage of particles through matter. All aspects of the simulation process have been included in the toolkit, such as the geometry and materials involved of the detector system, the tracking of particles through materials and electromagnetic fields and the response of sensitive detector components. The fast simulation uses a mix of GEANT4 and a simplified detector geometry and physics description, which significantly speeds up the simulation production compared to the full simulation.

The events generated in the MC simulation samples are “reconstructed” (i.e. the physics objects are built from the detector information) in the same way than the data events recorded by ATLAS. The following sections are dedicated to the offline reconstruction of the events.

4.2 Track and vertex reconstructions

Tracks are reconstructed from the information of the silicon detectors, and then extrapolated to include measurements in the TRT [93, 94]. Then the vertices associated to hard scatter interactions, known as the primary vertices (PV), are reconstructed from the track collection. The tracks are parametrised by p_T , η , ϕ and the transverse and longitudinal impact parameters (d_0 and z_0 respectively) computed with respect to the beam spot position. The quality selection required for the tracks used in a PV reconstruction is [95]:

- $p_T > 500$ MeV
- $|\eta| < 2.5$
- Number of hits in the silicon detector:
 - ≥ 9 if $|\eta| \leq 1.65$
 - ≥ 11 if $|\eta| > 1.65$
- IBL hits + B-layer hits ≥ 1
- SCT holes ≤ 1

where a hole is defined as a silicon sensor crossed by a track but without any associated hit [96]. A vertex seed is obtained from the global maximum in the distribution of z coordinates of the

tracks, computed at the closest point to the beam spot centre. The vertex position is determined with χ^2 fit, using the seed position and the tracks around it. The tracks incompatible with this vertex are used to seed another vertex. Finally, the PV with at least two tracks and the highest value of the sum of squared p_T of the tracks is chosen as the one associated to the hardest scatter, and the track's d_0 and z_0 are recomputed with its respect.

In addition, the presence of nearby interactions to the PV increases in high PU environments, increasing the probability to reconstruct a fake PV. Therefore, a selection for tracks associated to the PV is implemented: $|d_0| < 1.5$ mm and $|\Delta z_0 \times \sin \theta| < 1.5$ mm are required, where Δz_0 is the longitudinal impact parameter difference between the track and the PV; and θ is the polar angle of the track.

Long-lived particles can be produced from hadronization processes (like B-hadrons), where the vertices associated are called “secondary vertices”. It has a lower track multiplicity than the PV. The reconstruction of secondary vertices is important for the flavor tagging of jets of hadrons. In particular, the identification of b-jets is very important for the discrimination of $t\bar{t}$ events. A B-hadron can be formed by the bottom-quark has a relatively long lifetime of about 1×10^{-12} s and with a decay length ($c\tau$) 3 mm before decaying (section 4.6.5).

4.3 Electrons

4.3.1 Electron reconstruction

The electrons in the central region of ATLAS ($|\eta| < 2.5$) are successfully reconstructed if at least one track is matched to its cluster (otherwise the cluster is classified as an unconverted photon candidate). The $\eta \times \phi$ space of the EM calorimeter is divided into a grid of $N_\eta \times N_\phi = 200 \times 256$ elements of size $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ called towers (granularity of the middle layer of the EM calorimeter). Electron clusters with $E_T > 2.5$ GeV are searched using a sliding-window algorithm, with a size of 3×5 in tower units. The cluster reconstruction efficiency for electrons from the Z boson decay is about about 99% for electrons with $E_T > 20$ GeV [97]. Then, the reconstructed tracks with $p_T > 0.5$ GeV are loosely matched to the electron cluster using deviations in η and ϕ between the impact point of the track on the EM calorimeter and the electron cluster barycentre. The matching criteria takes into account the curving direction of the track and the number of precision hits in the silicon detector. After a successful track-cluster matching, the energy of the electron cluster is re-formed using a size of 3×7 tower units for the central region.

4.3.2 Electron identification

Algorithms for electron identification (ID) are established to distinguish the electrons from other particles [97]. The background electrons are composed mostly by non-isolated electrons, hadrons misidentified as electrons and electrons from photon conversions. The ID algorithms use calorimeter shower shapes, as information from the TRT, track-cluster matching related quantities, track properties, and variables measuring bremsstrahlung effects. A likelihood (LH) ratio is the multivariate analysis technique chosen for electron identification, using signal and background PDFs which depends on the discriminating variables. Then, an overall discriminant, d_L is computed for a given electron:

$$d_L = \frac{L_s}{L_s + L_b}; \quad L_{s/b}(\vec{x}) = \prod_{i=1}^n P_{s/b,i}(x_i) \quad (4.2)$$

where \vec{x} is a vector of discriminating variables (x_i); $P_{s,i}(x_i)$ and $P_{b,i}(x_i)$ are the signal and background probability density function extracted from experimental measurements. Three levels of identification operating points are provided for electron ID, in order of decreasing signal efficiency

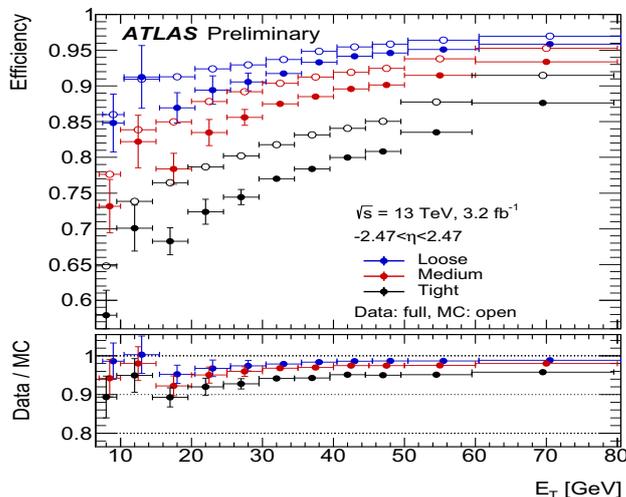


Figure 4.2: Combined electron reconstruction and identification efficiencies using 2015 p-p collision data at $\sqrt{s} = 13$ TeV. The data collected by the ATLAS detector, corresponds to an integrated luminosity of 3.2 fb^{-1} . Tag-and-probe method is used selecting $Z \rightarrow ee$ events to measure the efficiency for the three electron quality working points, as a function of the E_T and integrated over the full pseudorapidity range. The uncertainties are obtained with pseudo-experiments, treating the statistical uncertainties from the different (E_T, η) bins as uncorrelated [97].

and increasing background rejection: LH `loose`, LH `medium` and LH `tight`; based on different selections of the discriminating variables. The signal efficiencies for electron candidates of $E_T > 25$ GeV are in the range from 70% to 90%, increasing with E_T (Figure 4.2). The LH technique is the baseline ID algorithm in Run 2, since it allows a better background rejection for a given signal efficiency than the “cut-based” algorithms used in Run 1, for which the identification criteria were implemented with sequential cuts allowing a hierarchical increasing background rejection power for the identification of electrons.

In addition, the electron measurements are performed by requiring compatibility between the tracks associated to the electron candidates and the PV to reduce the background from conversions and secondary particles: $|d_0/\sigma_{d_0}| < 5$, where σ_{d_0} is the error on the transverse impact parameter. To ensure the compatibility of the electron with the PV, the $|\Delta z_0 \sin\theta| < 0.5$ mm is required.

4.4 Muon

4.4.1 Muon reconstruction

The reconstruction of the muons is performed independently in the inner detector and MS, and then the results are combined to form the muon tracks used in physics analyses [98]. The MS part is built by forming segments in each layer from the hit patters, and then by fitting together these segments.

Four types of muons are defined depending on the sub-detectors used in the reconstruction:

Combined (CB) muon: the track reconstruction is performed independently in the inner detector and MS, and a combined track is obtained with a global fit using the hits from both sub-detectors.

Segment-tagged (ST) muons: if a track in the inner detector is associated to at least one track segment in the MDT or CSC chambers.

Calorimeter-tagged (CT) muons: when a track in the inner detector can be matched to an energy deposit in the calorimeter compatible with a minimum-ionizing particle. This type recovers acceptance in a region where the MS is partially instrumented to due to the cabling services.

Extrapolated (ME) muons: the muon trajectory is reconstructed based only on the MS track and a loose requirement for muons coming from the interaction point. The acceptance of ME muons covers the region $2.5 < |\eta| < 2.7$, which is not covered by the tracking.

4.4.2 Muon identification

Several variables are used for the muons identification in order to discriminate between prompt muons and background muon candidate mainly originating from hadron decays (like pions and kaons) [98]: the charge over momentum ratio significance ($\frac{q/p}{\sigma_{q/p}}$) of the muon candidate measured in the inner detector or the MS, the absolute value of the p_T difference between the measurements in the inner detector and the MS divided over the p_T of the combined track (ρ') and a normalised χ^2 value from the fit of the combined tracks.

Robust momentum measurements are archived by requiring a certain number of hits in the inner detector and MS: at least one hit on Pixel, at least five hits on SCT and that at least 10% of the TRT hits assigned to the tracks included in the final fit.

Three muon identification selections are defined:

Loose muons: selection designed to maximise the reconstruction efficiency while providing good-quality muon tracks. All types of muon tracks are used where the ME muons are used only in the $2.5 < |\eta| < 2.7$ region. In the region $|\eta| < 2.5$, about 97.5% of the loose muons are CB muons, 1.5% are CT muons and the remaining 1% are reconstructed as ST muons.

Medium muons: aiming to minimise the systematic uncertainties associated with the reconstruction and calibration of the muon candidates. Only CB and ME muons are used, where the CB tracks are required to have > 3 hits in at least two MDT layers¹ and ME tracks are required to have at least three MDT/CSC layers. A cut in the q/p significance ($|\frac{q/p}{\sigma_{q/p}}| < 7$) is required to suppress the contamination due to hadrons misidentified as muons.

Tight muons: allow to maximise the purity of muons but loosing some efficiency. Only CB muons with at least two hits in stations of the MS are considered, satisfying the Medium selection criteria. A normalised χ^2 of the combined track is required together with a two-dimensional cut

¹Except in the $|\eta| < 0.1$ region where tracks with at least one MDT layer but no more than one MDT hole layer are allowed.

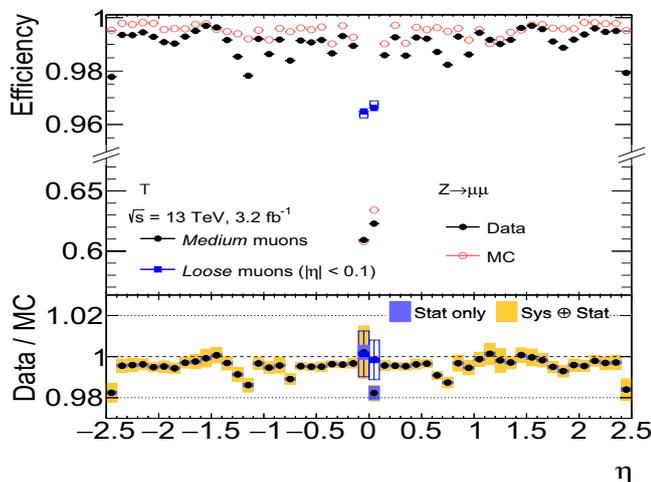


Figure 4.3: For $Z \rightarrow \mu\mu$ events, the muon reconstruction efficiency is shown as a function of η for muons with $p_T > 10$ GeV for the Medium (black dots) and Loose (red boxes) muon selections. The error bars on the efficiencies indicate the statistical uncertainty. Panels at the bottom show the ratio of the measured to predicted efficiencies, with statistical and systematic uncertainties.

in ρ' and q/p significance variables as a function of the muon p_T .

The medium and loose muon qualities are used in the analysis presented in Chapter 5 and 6. Their reconstruction and identification efficiencies for these two muon selections are shown in Figure 4.3, for which are obtained using $Z \rightarrow \mu\mu$ events.

4.5 Lepton isolation and correction

4.5.1 Lepton isolation

Prompt electrons or muons are those not originated from hadronic decay products. The sources of non-prompt or fake electrons are mostly coming from semi-leptonic decays of b- and c-quarks, photon conversions and pion decays. The sources of non-prompt muons are mostly semi-leptonic decay of heavy flavors and jets with large electromagnetic energy. The isolation variables allow to reject more fake leptons.

The calorimetric isolation is the addition of the transverse energy of the calorimeter cells inside a cone centred around the lepton direction (known as raw topocone isolation) not counting the

electron cluster. Such isolation would not be appropriate in boosted-top topologies² since a radius smaller than the calorimeter “cell size” cannot be used. Therefore, track isolation is preferred for the analysis presented in Chapter 5.

A track isolation, $p_T\text{varcone}$, is defined by the sum of transverse momenta of tracks (satisfying good quality requirements) inside a cone centred around the lepton direction. For this isolation the cone radius ΔR gets smaller with the transverse momentum of the lepton:

$$\Delta R = \min\left(\frac{k_T}{p_T}, R\right) \quad (4.3)$$

where k_T is a constant fixed to 10 GeV and R is the maximal radius of the cone used, which is 0.2 (0.3) for the electron (muon). The cut on $p_T\text{varcone}$ can be fixed at 6% of the lepton p_T (`FixCutTightTrackOnly` isolation) or can be set in order to get an efficiency of 99 % independent of the lepton p_T (`LooseTrackOnly` isolation). The analysis in Chapter 5 uses the `LooseTrackOnly` working point to select isolated leptons (Figure 4.4).

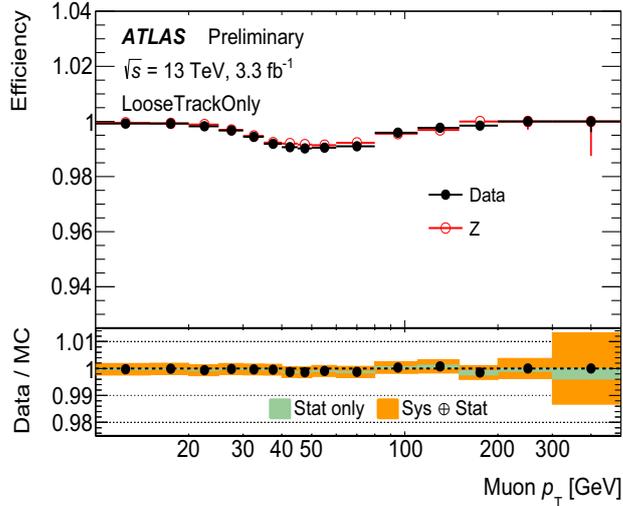


Figure 4.4: Efficiency of the muon isolation for the LooseTrackOnly working point, as a function of the muon p_T measured in $Z \rightarrow \mu\mu$ events. Such working point aims a flat efficiency of about 99%. Black dots indicate the efficiency measured using 2015 data while red circles indicate the prediction from MC samples. The statistical uncertainties (green) and combination of statistical and systematic uncertainties (orange) are shown in the bottom panel [98].

²In these topologies, the lepton can be very close to a b-jet.

Lepton corrections

Scale factors (SF) are derived to take into account the differences in the lepton reconstruction and selection efficiencies between data and simulation, and they are applied to the simulation samples as an event weight. Electron (muon) SF's are estimated using tag-and-probe methods in $Z \rightarrow ee(\mu\mu)$ and $J/\Psi \rightarrow ee(\mu\mu)$ samples, for trigger, reconstruction, identification and isolation efficiencies. The momentum of the lepton are calibrated using the well known mass of the Z and J/Ψ as reference.

In the simulation, the momentum of the leptons are smeared in order to reach the same momentum resolution than the data.

4.6 Jets of hadrons

Jets are collections of nearby clusters in the calorimeter (“calorimeter jets”) or tracks (“track jets”) of charged particles in the tracking systems. In chapter 5, the calorimeter jets will be used for the event selection and the invariant mass reconstruction of the $t\bar{t}$ system while the track jets, having only information about the charged particles inside them, are used for calorimeter jet calibration uncertainty and to identify the jets containing B-hadrons.

The particles interacting with the calorimeter produce energy deposits in various calorimeter cells. These deposits are clustered in topoclusters using a dynamical combination of cells around the calorimeter seeds passing a certain signal-to-noise ratio threshold [99]. The formation of topoclusters is a sequence of seed-and-collect algorithm, defining a cell signal significance ζ_{cell}^{EM} as:

$$\zeta_{cell}^{EM} = \frac{E_{cell}^{EM}}{\sigma_{EM}^{noise,cell}}, \quad (4.4)$$

where E_{cell}^{EM} is the energy measured in the cell and σ_{cell}^{noise} is the RMS of the respective noise distribution, both measured at the electromagnetic energy scale. The calorimeter seeds are the cells with $\zeta_{cell}^{EM} > 4$. Subsequently, a looser signal-to-noise ratio threshold ($\zeta_{cell}^{EM} > 2$) is demanded in the adjacent cells to the seed, to be included in the topocluster. Finally, all the calorimeter cells neighbouring the formed topocluster are added.

4.6.1 Jet reconstruction algorithms

Different sequential clustering algorithms can be used to reconstruct a four-momentum related to the particle which initiates the shower leading to the jet formation [100, 101]. The combination of topoclusters or tracks into jets is performed by evaluating the distance d_{ij} between objects i and j , and the distance d_i between object i and the beam:

$$d_{ij} = \min(p_{T,i}^{2k}, p_{T,j}^{2k}) \frac{\Delta R_{ij}^2}{R^2}, \quad (4.5)$$

$$d_i = p_{T,i}^{2k}, \quad (4.6)$$

where $\Delta R_{ij}^2 = \sqrt{\Delta\eta_{ij}^2 + \Delta\phi_{ij}^2}$ is the distance in the $\eta \times \phi$ plane, k and R are parameters of the algorithm. The value of k differentiates three kind of reconstruction techniques:

- $k = 1$: k_t algorithm,
- $k = 0$: Cambridge/Aachen algorithm,
- $k = -1$: anti- k_t algorithm [100].

The four-momentum of a pair of elements i and j with the smallest d_{ij} are added to form a new element when $d_{ij} < d_i$ and $d_{ij} < d_j$. Otherwise, the algorithm looks for another pair of elements to merge. The anti- k_t algorithm (Figure 4.5), which is used to reconstruct the jets in the analysis presented, provides an iteratively combination of the momenta of pairs of clusters. The reconstruction is not affected by the presence of soft particles (infrared safety).

For a jet reconstruction, the radius parameter R provides a relative measure of the distances between the constituent elements inside the jets. The analysis will use two kinds of calorimeter jets, using $R = 0.4$ for “small- R jets” and $R = 1.0$ for “large- R jets”. The large- R jets are interesting objects to study as their substructure can be used to reconstruct boosted top-quark decays. In addition, truth jets are similarly reconstructed using MC simulation of stable particles ($c\tau > 10$ mm) in the final state of the hard-scatter interaction. It is often useful to associate the tracks or simulated truth particles to the calorimeter jets, using the ghost association algorithm [102].

Jets can be initially reconstructed at the electromagnetic (EM) scale, meaning their energy measurement assumes the calorimeter response and the energy loss in the dead material are those of an electromagnetic shower. However, these assumption are no more valid when the shower is produced by hadrons and this correction of this effect is of primary importance for the jet calibration.

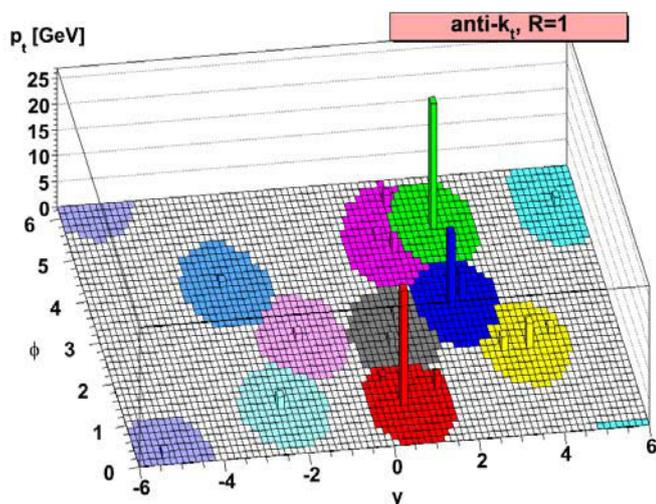


Figure 4.5: A parton-level event with many random soft contributions clustered with the anti-kt algorithm using a $R=1$. One can notice the nice circular shape of the highest p_T jet.

The jet reconstruction can also be performed on a Local Cluster Weighting (LCW) calibrated topocluster. A LCW scheme classifies the topoclusters as electromagnetic or hadronic, and assigns it an energy that depends on this classification. The cluster shape characterizes the topology of the energy deposits of electromagnetic or hadronic showers, defining observables derived from the cluster energy [104] as presented in Figure 4.6.

4.6.2 Jet calibration

Several calibration schemes (Figure 4.6) are developed to correct the jet energy measurement while reducing PU contributions, with different levels of complexity and different sensitivities to systematic effects [103].

First, the jet direction is corrected to point to the primary vertex instead of the geometrical detector centre. The correction improves the angular resolution of the jets without changing its energy. Then, the pile-up effects are removed using an area-based subtraction process. As a third step, Monte Carlo-based calibration called Absolute EtaJES, corrects the non-linear correlation between the reconstructed energy in the calorimeter and the energy of the particles generating the jets. It is followed by the “Global sequential calibration” that reduces the fluctuations in the energy measurements of the jets, improving its resolution. Finally, a residual data calibration is derived from *in-situ* measurements, where techniques are used to adjust the jet calibration in data to the one in MC by the comparison with well-calibrated objects.

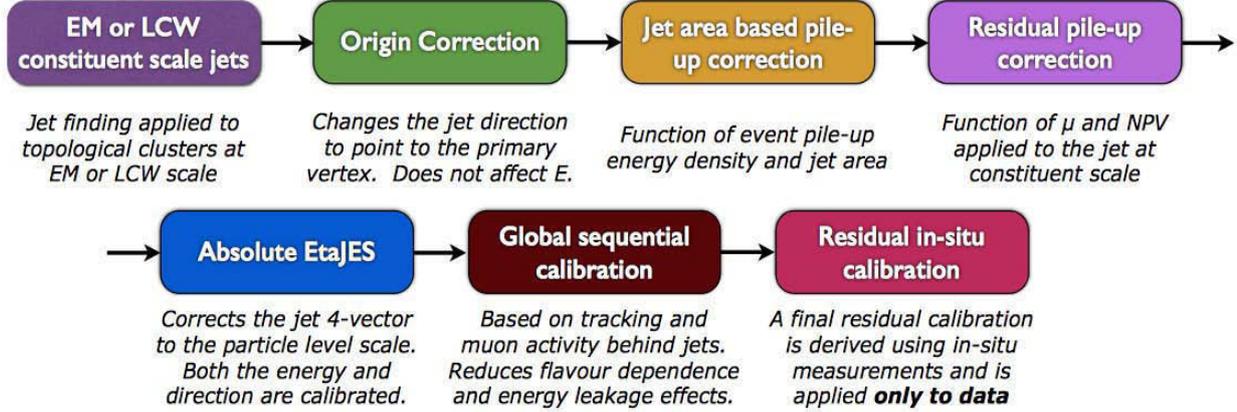


Figure 4.6: Stages used in the calibration of calorimeter jets.

In the steps that have been presented, the jet energy scale is measured and characterised by the jet energy response, defined as the ratio between the reconstructed jet energy and the energy of the corresponding truth particle in the simulation:

$$R(E, \eta) = \left\langle \frac{E^{reco}}{E^{truth}} \right\rangle \quad (4.7)$$

The calibration constants are derived as a function of the truth jet energy and pseudorapidity, and a numerical inversion is performed to express them as function of redo-level variables. In addition, further corrections are implemented to correct jets which are partially contained in the calorimeter [105, 106].

4.6.3 Track-based pile-up jets suppression

To distinguish the jets coming from the hard-scatter interactions from the jets coming from PU, several algorithm have been developed [112]:

Jet Vertex Fraction (JVF): quantity defined as the scalar sum of the transverse momentum of the tracks associated to a jet and originating from the hard-scatter vertex, divided by the scalar sum of all the jet tracks:

$$JVF = \frac{\sum_k p_T^{trk_k}(PV_0)}{\sum_l p_T^{trk_l}(PV_0) + \sum_{n \geq 1} \sum_j p_T^{trk_j}(PV_n)} \quad (4.8)$$

where PV_0 represents the hard-scattering vertex and PV_k corresponds to vertices from pile-up. The JVF quantity is bound between 0 and 1, measuring the fraction of transverse momentum associated to jets coming from the hard-scatter vertex. In average, the JVF decreases with the number of reconstructed primary vertices in an event; therefore a correction is implemented using the average scalar sum of the transverse momentum of the pile-up tracks for a jet $\langle p_T^{PU} \rangle$:

$$corrJVF = \frac{\sum_k p_T^{trk_k}(PV_0)}{\sum_l p_T^{trk_l}(PV_0) + \frac{\sum_{n \geq 1} \sum_j p_T^{trk_j}(PV_n)}{(k \cdot n_{trk}^{PU})}} \quad (4.9)$$

where $\langle p_T^{PU} \rangle = k \cdot n_{trk}^{PU}$, k is a scaling factor roughly taken from the slope of the average scalar sum of the pile-up tracks and n_{trk}^{PU} is the number of pile-up interactions per event.

Charged p_T fraction (R_{p_T}): is another important variable for the suppression of the PU jets, defined as the scalar p_T sum of the tracks originating from the hard-scatter vertex divided by the p_T of the fully calibrated jet:

$$R_{p_T} = \frac{\sum_k p_T^{trk_k}(PV_0)}{p_T^{jet}} \quad (4.10)$$

which is interpreted as a charged p_T fraction for jets from hard-scattering. The mean value and spread of R_{p_T} are larger for jets originated from hard-scatter vertex than for the pile-up jets.

Jet Vertex Tagger (JVT): the R_{p_T} and corrJVF variables are combined to construct another discriminant called JVT. The k-nearest neighbour algorithm [113] is used to build a 2-D likelihood, where a relative probability for a jet to be signal-type is computed for each point of the corrJVF - R_{p_T} plane, as the ratio of the number of HS jets divided by the number of HS plus PU jets found in a local neighborhood around the point used (Figure 4.7). A recommended JVT requirement (JVT > 0.59) for jets with $p_T < 60$ GeV and $|\eta| < 2.4$ is implemented for small-R jets used in Chapter 5.

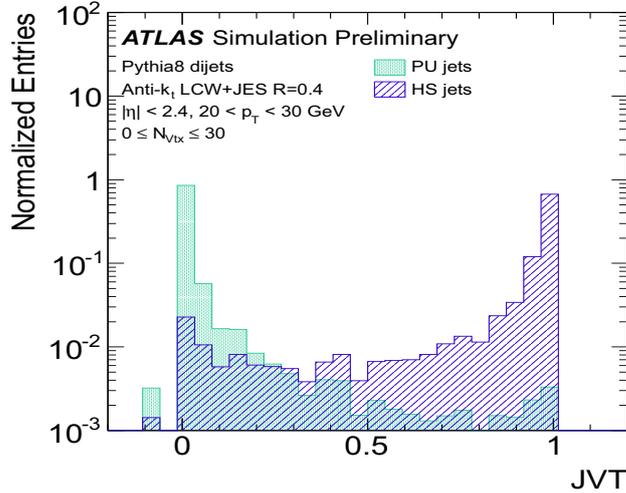


Figure 4.7: Distribution of JVT for pileup (PU) and hard-scatter (HS) jets. PU jets tends to accumulate at low JVT values while HS jets accumulate at high JVT values. A value $JVF = -1$ is assigned to jets with no associated track [112].

4.6.4 Large-R jets and Top-tagging

The large-R jets [109] are reconstructed using calorimeter clusters calibrated at the hadronic scale with the LCW method using the anti- k_t algorithm with $R=1$. A jet trimming [108] is used to suppress the clusters associated to soft particles which are likely to come from PU: the jet constituents are reclustered using the k_T algorithm into sub-jets, with a characteristic radius $R_{\text{sub}} = 0.2$, and the sub-jet is dropped from the large-R jet if its p_T is smaller than a fraction $f_{\text{cut}} = 5\%$ of the total jet p_T .

The top-tagging identification techniques [110, 111] are important for analyses searching for heavy particles decaying into $t\bar{t}$. The analysis presented in Chapter 5 will use a simple top-tagging algorithm [111] that has a strong and reliable performance to identify top-quarks with $p_T > 300$ GeV with a high efficiency. Two substructure-related variables are used in this top-tagging algorithm: the calibrated jet mass and the N-subjettiness ratio (τ_{32}).

The calorimeter-based large-R jet mass (m^{calo}) is defined as the norm of the four-momentum sum of the jet constituents, which are calibrated calorimeter clusters with energy E_i and momentum

\vec{p}_i :

$$m^{calo} = \sqrt{\left(\sum_{i \in J} E_i\right)^2 - \left(\sum_{i \in J} \vec{p}_i\right)^2}. \quad (4.11)$$

The jet mass is then calibrated to the particle-level jet mass with a procedure analogous to the step ‘‘Absolute EtaJES’’ step of the jet energy scale [104].

The τ_{32} variable allows the discrimination between jets containing a three-prong structure over jets containing a two-prong structure. After having reconstructed the large-R jet constituent into N sub-jets, an algorithm computes the N -subjettiness variable τ_N defined as:

$$\tau_N = \frac{1}{N_0} \sum_k p_{T,k} \cdot \min(\Delta R_{1,k}, \Delta R_{2,k}, \dots, \Delta R_{N,k}), \quad (4.12)$$

where N_0 is a normalisation factor, $p_{T,k}$ is the transverse momentum of the constituent particle (topocluster) k , and $\Delta R_{i,k}$ is the angular distance between a subject i and a constituent particle k . τ_N quantifies how well a jet can be described as containing N sub-jets, where $\tau_N \rightarrow 0$ means that the constituents are aligned with the candidate sub-jet directions, so the large-R jet is likely to be composed of N sub-jets. The ratio $\tau_{21} = \tau_2/\tau_1$ allows a discrimination of W-bosons (containing a two-prong structure) from pile-up jets and radiation jets. Similarly, the ratio $\tau_{32} = \tau_3/\tau_2$ allows a discrimination of top-quarks, which is a three-prong structure.

4.6.5 b-tagging algorithm

It is crucial for analyses using top-quarks to identify the jets coming from b-quarks (known as ‘‘b-jets’’). The b-tagging algorithms aim to distinguish the b-jets from the jets originated from the hadronisation of a c-quark (called ‘‘c-jets’’), and jets coming from gluons or u, d, s quarks (jointly called ‘‘light-jets’’). The b-tagging algorithms can be used to identify b-jets as they contain long-lived hadrons, that can produce a secondary vertex and tracks with large d_0 (Figure 4.8).

The ATLAS collaboration has implemented three b-tagging algorithms, which are later combined in a multivariate discriminant [114]:

Impact parameter based algorithms

The IP3D algorithm uses the signed impact parameters significance of the jet tracks to build a likelihood function using simulation of the signal (b-jets) and backgrounds (c-jets and light-jets).

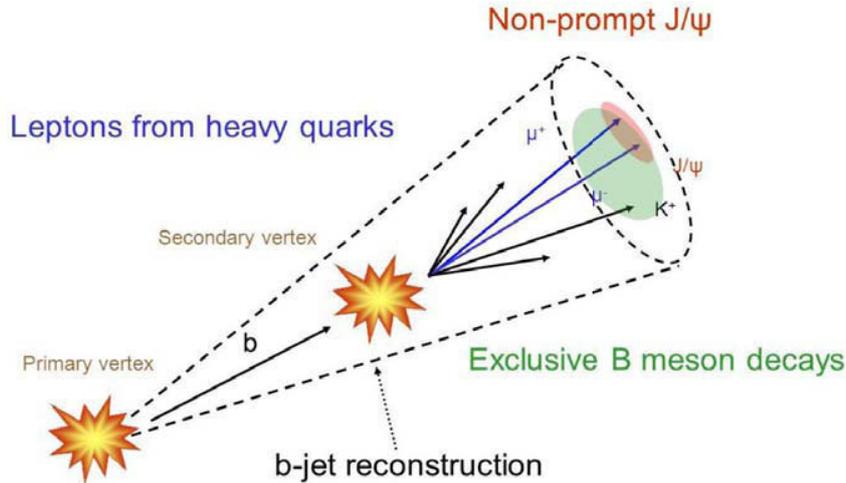


Figure 4.8: B-hadrons inside the jets originated from heavy quarks

The significance of the longitudinal impact parameter, z_0/σ_{z_0} , and the significance of the transverse impact parameter, d_0/σ_{d_0} is obtained of each track associated to a given jet, where σ_{d_0} and σ_{z_0} corresponds to the respective uncertainties. The significances for the impact parameters are signed based on whether the track intercepts the jet cone axis before the PV or not. The likelihood function for signed significances is asymmetric for b-jets, while it is symmetric for light-jets.

Inclusive secondary vertex reconstruction algorithm

The SV1 algorithm reconstruct a secondary vertex to identify a b-jet. It uses a likelihood ratio based on variables related to the secondary vertex position and multiplicity, the invariant mass of all the tracks associated to a secondary vertex; and the ratio between the energy addition of tracks associated to a secondary vertex and tracks from the primary vertex.

Decay chain multi-vertex reconstruction algorithm

The JetFitter reconstructs the topological structure of weak b- and c-hadron decays (PV \rightarrow b-quark \rightarrow hadron decay chain). The algorithm tries to find a common line between the primary vertex and the vertices of the hadron candidates.

Boosted decision tree

Finally, the b-jet discrimination is performed with a boosted decision tree (BDT) algorithm combining the input variables obtained from the three previous algorithms. The MV2c10 and MV2c20 b-tagging algorithms are implemented in the analysis, where the BDT training is performed assigning b-jets as signal; and a mixture of c- and light jets as background, with compositions of 10% and 20% of c-jets for MV2c10 and MV2c20 respectively. The analysis presented in Chapter 5 performs the b-tagging using track jets with $R = 0.2$, which improves the performance in boosted regimes (dense environments).

The performance of the b-tagging algorithm is defined as the fraction of true b-jets which are tagged by the algorithm. The probabilities of mistakenly b-tagged a jet originating from a c-quark or a light-flavor is referred as a mistag rate. The rejection rate of non-b-jets can be also determined, and it is defined as the number of true light-jets which are rejected when one passes the b-tagging requirement. Several operating points are established to be used with a single cut on the MV2 output distribution (Table 4.1). Each of them provides a specific b-tagging efficiency which is obtained using $t\bar{t}$ MC sample.

The b-tagging, c-tagging and mistag efficiencies can be different for simulation and data. Therefore, the simulations are corrected with scale factors [115].

Cut Value	b-jet Efficiency [%]	c-jet Rejection	τ -jet Rejection	Light-jet Rejection
0.4496	60	21	93	1900
-0.0436	70	8.1	26	440
-0.4434	77	4.5	10	140
-0.7887	85	2.6	3.8	28

Table 4.1: Operating points for the MV2c20 b-tagging algorithm and their corresponding efficiency or rejections. The small statistical uncertainties are not shown.

4.7 Overlap removal

A given physics object can sometimes appears in several collection of reconstructed objects (for example an electron can be seen as a jet cluster). A selection devoted to remove the geometric overlap between jet, electron and muons is implemented with objects with $p_T > 25$ GeV. If any selected small-R jets has an angular distance smaller than 0.2 with an electron ($\Delta R < 0.2$), the jet is rejected. Furthermore, if an electron is found to be close to a jet such that with an angular distance between 0.2 and 0.4 ($0.2 < \Delta R < 0.4$), then the electron is rejected. Then, if the angular distance between a muon and jet satisfies $\Delta R < 0.04 + 10\text{GeV}/p_T^\mu$, the muon is rejected if the jet

has at least 3 tracks originating from the primary vertex. The jets with less than 3 tracks which overlap a muon are rejected.

4.8 Missing transverse energy

The non-interacting particles such as neutrinos cannot be detected in ATLAS. They are inferred from the measurement of the missing transverse energy.

The transverse momentum component of the initial state of the collision is null, so any imbalance in the final state is the signature that a particle escaped the detector. The longitudinal momentum is unknown since the colliding partons have only a fraction of the proton's energy. The E_T^{miss} is defined as the transverse momentum that would be needed to get no transverse momentum in the final state. Ideally, the E_T^{miss} should be null in the case where non-interacting particles are not produced in an event. But there are many sources of fake E_T^{miss} , like the finite detector resolution, the limited detector coverage or the presence of non instrumented regions in the detector.

The reconstructed track-based soft term (TST) E_T^{miss} is separated in two components, a “hard term” considering the reconstructed and calibrated physics objects (e, jets, μ); and a “soft term” considering the reconstructed tracks associated to the hard-scatter vertex which are not associated to the hard term objects:

$$E_{x(y)}^{miss} = E_{x(y)}^{miss,e} + E_{x(y)}^{miss,jets} + E_{x(y)}^{miss,\mu} + E_{x(y)}^{miss,soft}, \quad (4.13)$$

where each term correspond to the negative vectorial sum of the momentum of the respective calibrated objects. The soft term can be reconstructed with the TST algorithm, aimed to reduce the impact of pile-up interactions [116]. With the $E_{x(y)}^{miss}$ components, the magnitude of the E_T^{miss} and the azimuthal angle ϕ_{miss} are calculated:

$$E_T^{miss} = \sqrt{(E_x^{miss})^2 + (E_y^{miss})^2}. \quad (4.14)$$

The reconstruction of the $Z \rightarrow ll$ can be used for the estimation of the performance of the E_T^{miss} calculation, since no missing transverse momentum is expected. The reconstruction of $W \rightarrow l\nu$ can be also used to test the performance of the E_T^{miss} estimation. An estimator of the E_T^{miss} performance is its resolution, which is parametrised with the sum of transverse energy in the detector:

$$\sigma(E_T^{miss}) = k\sqrt{\sum E_T}. \quad (4.15)$$

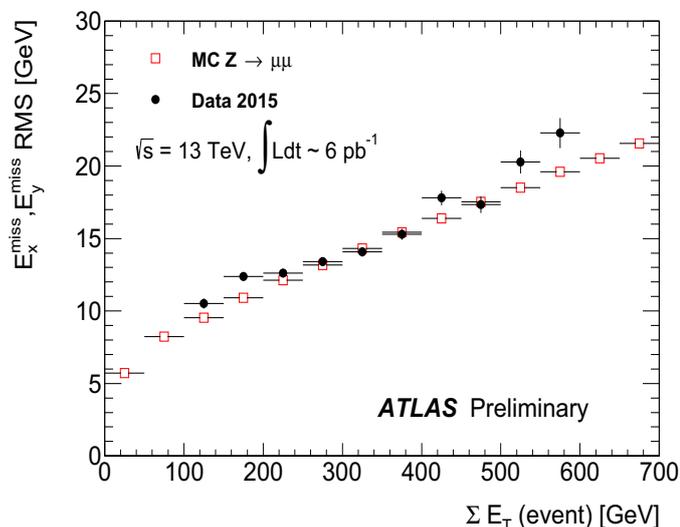


Figure 4.9: Distributions of E_x^{Miss} , E_y^{Miss} resolution as a function of $\sum E_T$. The value of the resolution is estimated from the RMS of the combined of E_x^{Miss} and E_y^{Miss} in each bins of $\sum E_T$. The data (black circles) and MC simulation (red squares) are overlaid. As expected, the resolution of the missing transverse energy increases with $\sum E_T$ [116].

where k is a fitted parameter and $\sum E_T$ is defined as the scalar sum of the transverse momentum of the objects from the hard-scatter interactions and soft term contributions (Figure 4.9).

4.9 Outlook

A display of an event candidate for boosted top quark pair production that use all the elements presented in this Chapter is shown in figure 4.10, where the red line corresponds to the muon, the dashed line shows the direction of the missing transverse momentum, the green and yellow bars indicate energy deposits in the liquid argon and tile calorimeters. Four small-radius ($R=0.4$) jets are identified, where three of these jets are re-clustered into the leading large-radius ($R=1.0$) jet (not shown explicitly). One of these three jets is identified as having originated from b-quarks.

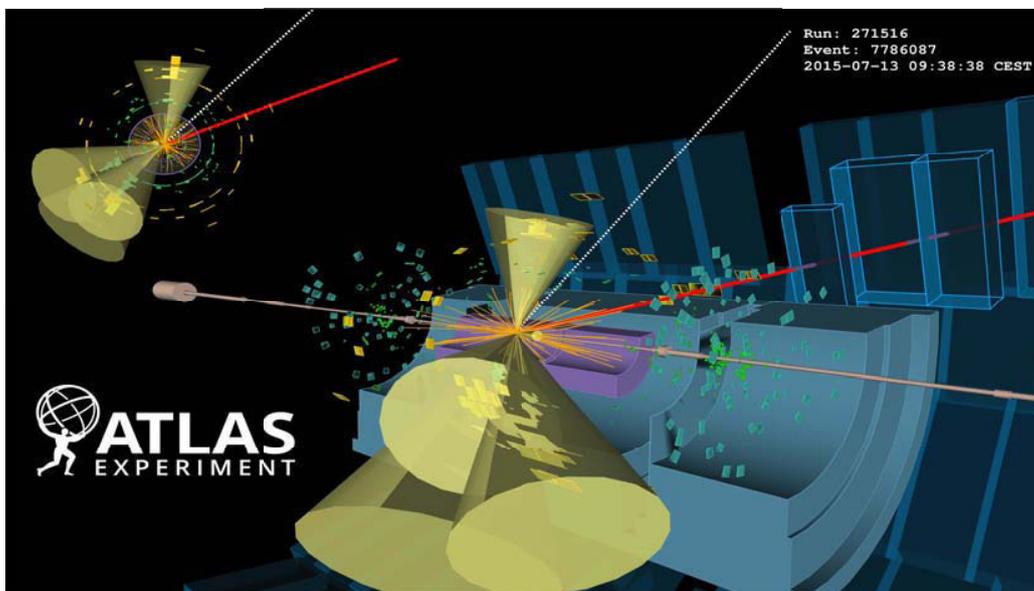


Figure 4.10: Display of event candidate for boosted top quark pair from LHC pp collisions in 2015.

Chapter 5

Search for $t\bar{t}$ resonances in boosted regimes

Contents

5.1	Strategy	104
5.2	Event selection	105
5.2.1	$t\bar{t}$ topologies	105
5.2.2	Event pre-selection	106
5.2.3	Boosted selection	107
5.3	Invariant $t\bar{t}$ mass reconstruction	108
5.4	Signal simulation	110
5.5	Background processes	110
5.6	Data driven backgrounds	112
5.6.1	Non-prompt leptons	112
5.6.2	W +jets background normalisation	117
5.7	Systematic Uncertainties	117
5.8	Control plots	119
5.9	Compatibility with SM only hypothesis	119
5.10	Upper production cross section limits on $t\bar{t}$ resonances	126

The Standard Model has many successful predictions, but its theoretical and experimental limitations motivate the development of many alternative Exotic models, as described in chapter 1. As the top-quark has a strong coupling with the Standard Model Higgs boson it is likely it has some connections with the naturelness issue. Therefore, it is appealing to search for “new physics” in the top-quark physics sector.

The author has participated in the effort for the integration of the $t\bar{t}$ analysis tools into a common software framework for analyses related to top-quark physics for ATLAS (named “AnalysisTop”). The object definitions and working points have been set in the “HQTtResonancesTools” package. Another package developed for this results is named “TopPartons”, which contains the tools to read the truth information from the signal simulations and it has been used by other members of the analysis team. This package has been used to measure the real rates and to perform the reconstruction studies in Chapter 6. In addition, the author has participated in the effort to estimate the QCD multi-jet background for this analysis. The tools have been integrated to the offline reconstruction package: “TopNtupleAnalysis”.

5.1 Strategy

The exposed analysis aims to search for $t\bar{t}$ resonances in the “lepton+jets” channel, using 3.19 fb^{-1} of p-p collision data at $\sqrt{s} = 13 \text{ TeV}$, which represents approximately six times less data

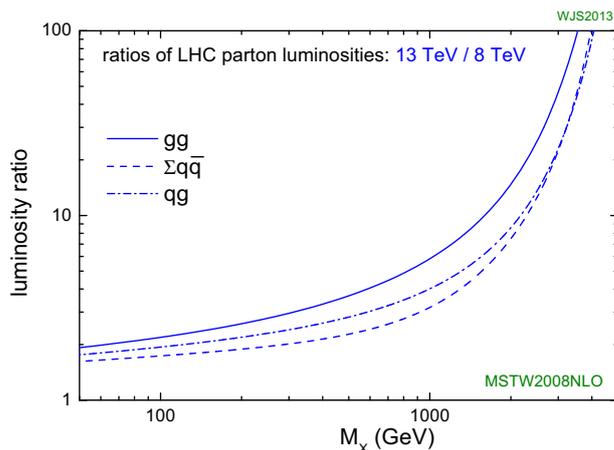


Figure 5.1: Parton luminosity ratio (13 TeV over 8 TeV) as function of the of the centre-of-mass energy, for the production of gluon-gluon, quark-antiquark and quark-gluon interactions at LHC [117].

than the one used to search for $t\bar{t}$ resonances at $\sqrt{s} = 8$ TeV (Run 1) [40]. The physics potential dependence with \sqrt{s} can be estimated from the parton luminosity calculations. The ratio of the parton luminosity obtained at $\sqrt{s} = 13$ TeV and 8 TeV, is presented in Figure 5.1 as function of the beam energy. For high parton energies (> 1 TeV), the parton luminosity ratio increases significantly, compensating the smaller amount of data collected at early Run 2. Instead, the sensitivity for the production of new physics at low parton energies is not significantly improved with only 3.19 fb^{-1} of collision data. Therefore for early Run 2, this analysis focuses on the regime of high mass resonances that would produce boosted top-quarks.

The analysis selection is optimised to properly reconstruct the decay products of an hadronic boosted-top within a single large- R jet while reducing the non- $t\bar{t}$ backgrounds. Then, a statistic test is performed to check the compatibility of the $t\bar{t}$ mass spectrum with the Standard Model only hypothesis. If the data is compatible with the SM-only hypothesis, an exclusive limit is set at 95% of confidence level (CL) for a spin 1 Z' benchmark model.

The procedure used in the current analysis is quite generic, the search for signals being without assumption on the signal shape. Then, limits are set over one benchmark particle: topcolor assisted technicolor Z' (TC2) [29, 30, 118]. Another benchmark models were historically used, like Randall-Sundrum (RS) warped extra-dimensions, which includes bulk Kaluza-Klein (KK) gluon and graviton decaying into a $t\bar{t}$ pair (signal simulations are not available for the time-scale of this manuscript). The free parameters in such models allow to set limits on different widths, being narrow or broad resonances, where the last one are more difficult to be detected [?, 119, 120].

5.2 Event selection

5.2.1 $t\bar{t}$ topologies

As mentioned in section 1.3, the semileptonic decay of the $t\bar{t}$ system is the most sensitive channel for the search for $t\bar{t}$. The invariant mass ($m_{t\bar{t}}$) is the main observable used to detect heavy resonances, which is computed with the sum of the four-momentum of the reconstructed physics objects associated to the leptonic and hadronic top-quarks. The event topology depends on the available objects in the final state which are going to be used for the $m_{t\bar{t}}$ reconstruction. The low $m_{t\bar{t}}$ region is dominated by the resolved scenario, where the decay products of the top-quarks are well separated in the detector (discussed in section 6.1). But for top-quarks with p_T larger than twice its own mass, its decay products start to become collimated, merging the jets in a single region on the calorimeter. This configuration is known as boosted scenario, where sophisticated reconstruction techniques are implemented to identify the substructure elements of the hadronic top decay.

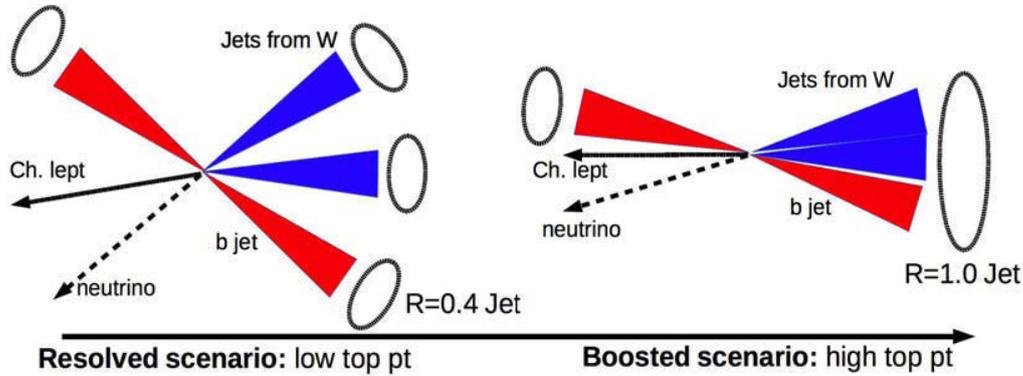


Figure 5.2: Kinematic topology of $t\bar{t}$ events. Resolved and Boosted scenarios require different reconstruction techniques to reconstruct the invariant mass of the $t\bar{t}$ system.

5.2.2 Event pre-selection

An initial pre-selection can be defined for single lepton events containing the physics objects described in previous chapter 4:

Primary vertex

The events must have at least one primary vertex candidate. It is chosen as the vertex with the highest sum of p_T^2 from tracks associated to the vertex candidate.

Triggers

The selected events must pass either the single electron triggers (for the e +jets channel):

- `HLT_e24_lhmedium_L1EM20VH`¹ OR `HLT_e60_lhmedium` OR `HLT_e120_lhloose`

or the single muon triggers (for the μ +jets channel):

- `HLT_mu20_loose` OR `HLT_mu50`

where the trigger name follows the convention defined in section 5.2.2.

¹`HLT_e24_lhmedium_L1EM20VH` was not available for the simulation, being replaced by `HLT_e24_lhmedium_L1EM18VH`. However, scale factors take into account this difference.

Lepton selection

The quality of the electron and muon candidates is selected to reject mis-reconstructed and non-prompt leptons. `LH Tight` quality is demanded for the electron while `Medium` quality is used for the muon (see sections 4.3.2 and 4.4.2). In addition, compatibility between the lepton tracks with the primary vertex are required demanding $|d_0/\sigma_{d_0}| < 5(3)$ and $|\Delta z_0 \sin \theta| < 0.5$ mm for the electron (muon).

Exactly one electron or muon within the detector acceptance is demanded in the analysis selection. The p_T of the lepton should be larger than the lowest trigger threshold, leading to an offline p_T threshold of 30 GeV for the electron and 25 GeV for the muon. The events having a second lepton with p_T larger than 25 GeV are vetoed to reject $Z \rightarrow ll$ and dileptonic $t\bar{t}$ processes. In addition, an electron in the crack region of the calorimeters is excluded. The `LooseTrackOnly` working point is chosen for the lepton isolation with a average selection efficiency of about 99% which is almost constant in the considered lepton p_T range (section 4.5.1). In addition, the lepton is also required to be matched to the lepton trigger candidate to ensure the validity of the lepton trigger SF (section 4.5.1).

Leptonic-W selection

It is important to suppress the backgrounds without neutrinos in the final state, like events from the QCD multi-jets background. Therefore, the selection is tightened to select topologies containing a W-boson decaying leptonically using the E_T^{miss} ($E_T^{miss} > 20$ GeV) and the transverse mass (m_T^W) between the lepton and the E_T^{miss} ($E_T^{miss} + m_T^W > 60$ GeV), computed as:

$$m_T^W = \sqrt{2 \cdot p_{T,l} \cdot E_T^{miss} \left(1 - \cos(\Delta\phi)\right)}, \quad (5.1)$$

where $p_{T,l}$ is the transverse momentum of the lepton, and $\Delta\phi$ is the azimuthal angle difference between the lepton and E_T^{miss} (see section 4.8).

5.2.3 Boosted selection

The pre-selected events are then required to pass the following “boosted” selection [121]:

Leptonic top b-jet Events are required to have at least one small-R jet with a transverse momentum $p_T > 25$ GeV and $|\eta| < 2.5$. In addition, a topological cut of distance $\Delta R < 1.5$ between the lepton and its closest jet, known as the *selected jets*. The *selected jet* with the largest

p_T is associated to the b -jet from the leptonic top-quark, but no b -tagging requirement is enforced on them.

Hadronic-top selection At least one large- R jet (section 4.6.2) is demanded with a transverse momentum $p_T > 300$ GeV and $|\eta| < 2$, which is top-tagged. The top-tagger defined in section 4.6.4 is based on the calibrated mass m^{calib} and the τ_{32} ratio. A dynamical optimisation was performed by testing the possible lower thresholds on m^{calib} and the upper thresholds on τ_{32} , as function of the p_T of the large- R jet. For the selected working point of 80%, the threshold values (m^{calib}, τ_{32}) varies from (> 70 GeV, < 0.85) at $p_T = 200$ GeV, until (> 135 GeV, < 0.7) for p_T larger than 1600 GeV [111].

In addition, two topological cuts are implemented in the boosted selection: requiring a minimum angular separation between the large- R jet and the selected lepton $\Delta\phi > 2.3$, and a minimum angular separation between the large- R jet and the *selected jet* $\Delta R > 1.5$.

b-tagging

Events are required to have at least one b -tagged track-jet, considering only the track-jet far from the lepton ($\Delta R > 0.2$). The b -tagging is done using the MV2C20 algorithm with the 70% efficient working point (see section 4.6.5).

5.3 Invariant $t\bar{t}$ mass reconstruction

Following event selection, the observable $m_{t\bar{t}}^{reco}$ is obtained from the reconstructed physics objects in the final state (Figure 5.3). The four-momentum of the leading (in p_T) *top-tagged jet* is used as the hadronic top-quark candidate. The leptonic top-quark candidate is constructed by summing the four-momenta of the charged lepton, the neutrino candidate and the leading *selected jet*. The neutrino candidate is estimated from the E_T^{miss} and by assuming an on-shell W-boson decay.

Neutrino four-momentum

For semi-leptonic $t\bar{t}$ events, the neutrino is the only particle that cannot be detected. This means that the missing transverse energy can be considered as a measurement of the transverse component of the neutrino's momentum ($p_{T,\nu}$). The longitudinal component of the neutrino's momentum ($p_{z,\nu}$) is estimated using the mass of the W-boson as a constrain ($m_W^2 = E_W^2 - |p_W|^2$).

The W-boson on-shell assumption should not have a significant effect since the width of the W-boson is 2.085 ± 0.042 GeV is much smaller than the resolution of the E_T^{miss} that goes as

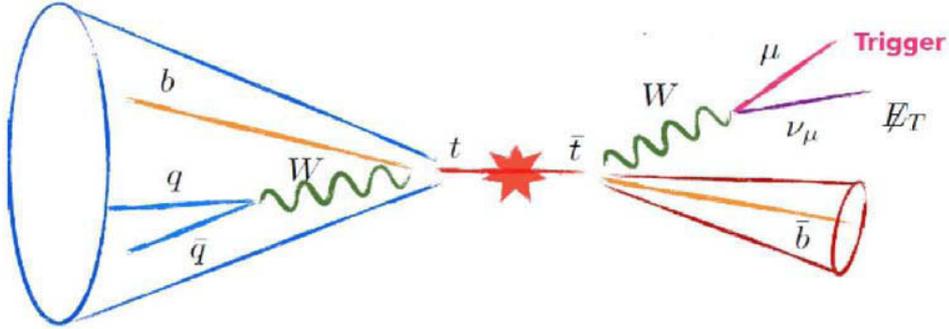


Figure 5.3: Scheme of the invariant mass reconstruction of $t\bar{t}$ signature in boosted scenarios.

$\sigma_{E_T^{miss}} = O(\sqrt{\sum E_T})$ GeV (section 4.8). The W mass constrain on the lepton + neutrino system leads to a quadratic equation for p_z^ν :

$$p_{z,\nu}^2 - 2 \frac{\mu \cdot p_{z,l}}{E_l^2 p_{z,l}^2} \cdot p_{z,\nu} + \frac{E_l^2 \cdot p_{T,\nu}^2 - \mu^2}{E_l^2 p_{z,l}^2} = 0 \quad (5.2)$$

$$\mu = \frac{1}{2} M_W^2 + p_{T,l} \cdot p_{T,\nu} \cdot \cos(\Delta\phi) \quad (5.3)$$

where μ is the discriminant of the quadratic equation; E_l , $p_{T,l}$ and $p_{z,l}$ are the energy, and the transverse and longitudinal momentum components of the charged lepton respectively; M_W is the known mass of the W-boson ($M_W = 84.4$ GeV is extracted from the PDG [39]) and $\Delta\phi$ is the azimuthal angle between the charged lepton and the missing transverse energy.

In the case of the quadratic equation has real roots, the solution with the smallest $|p_{z,\nu}|$ is used. However, the detector resolution can smear the measurement of the missing transverse energy, leading to complex solutions to eq. 5.2. In that case, several alternatives can be used to extract the longitudinal component of the neutrino. In the current analysis, the transverse component of the neutrino momentum is rotated and scaled by the smallest amount until a real solution to eq. 5.2 is obtained.

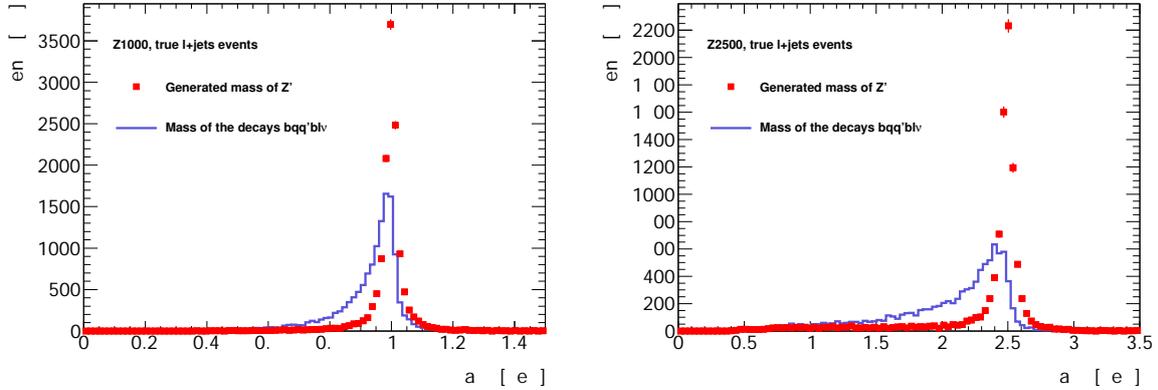


Figure 5.4: Mass of the Z' signal and of the $t\bar{t}$ decay products for the 1 TeV (left) and 2.5 TeV (right) resonances.

5.4 Signal simulation

This search uses topcolor-assisted technicolor (TC2) Z' as a benchmark. The signal process $pp \rightarrow Z' \rightarrow t\bar{t}$ is modelled using the SSM model [122] in Pythia 8 generator using the NNPDF2.3 PDF set at leading-order (which has a 3% width) and then is reinterpreted in the TC2 framework. The properties of the TC2 model used are parametrised with three variables: $\cot\theta_H$, drives the coupling to the quarks, and f_1 and f_2 , alter this coupling to right-handed up-type and down-type quarks, respectively. The cross-section are compared for 2 widths (1.2% or 3%) by changing $\cot\theta_H$. To account for higher-order contributions to the cross section, the leading-order calculation is multiplied by a factor of 1.3 based on calculations performed at NLO in QCD [123]. The interference of these signals with SM electroweak $t\bar{t}$ production is neglected in this analysis.

The generated Z' mass is compared to the reconstructed mass of the $t\bar{t}$ system at truth-level (Figure 5.4). The latest one can be significantly different from a Breit-Wigner shape (typical for a resonance), because the high mass resonances are more likely to be produced off-shell, and also because the highly boosted top-quarks from a heavy resonance tends to radiate strongly. The task of performing the reconstruction is further complicated by the extra radiation emitted. These effects can reduce the sensitivity of the analysis, particularly for very high resonance masses.

5.5 Background processes

There are certain physics processes which give a similar final state configuration than the signal, known as background processes. The source of background can be also due to experimental effects,

like the misidentification of jets as leptons for example. The selection procedure aims to reject significantly more background than signal, but it is necessary to estimate the residual background in the data after the selection either using MC simulation or data-driven technique.

The irreducible SM $t\bar{t}$ and the single-top processes are generated with Powheg + Pythia 6 with the matrix element calculations at next-to-leading order (NLO) in QCD with the CT10 PDF set. In addition, the cross-section is computed at NNLO for the SM $t\bar{t}$ simulation [124, 125] and NLO for single-top simulation [126]. The $t\bar{t}$ events are weighted to account for the interferences of electroweak process diagrams (Figure 5.5) for the quark- and gluon-induced processes [127, 128]. The impact of this correction depends of the kinematic, being a few percents at low $m_{t\bar{t}}$ and about $O(10\%)$ at 2.5 TeV.

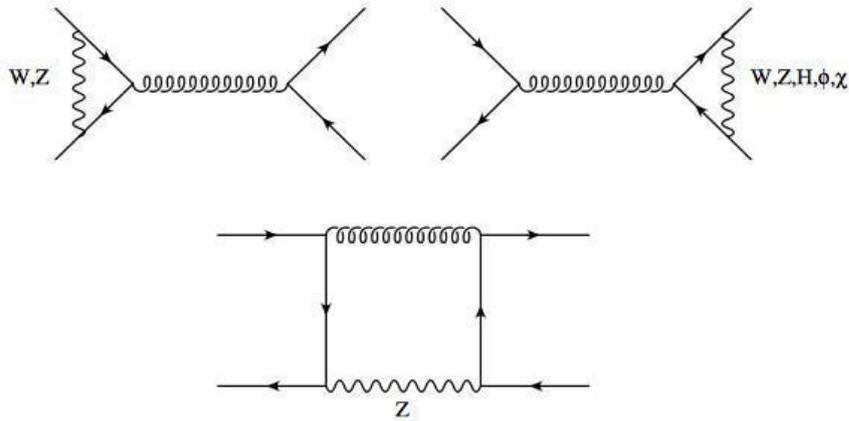


Figure 5.5: Example of Feynman diagrams used for the NLO electroweak corrections to the quark-induced process of the $t\bar{t}$ production.

The production of W+jets, Z+jets and di-boson are evaluated using Monte Carlo samples generated by Sherpa 2.1.1, the matrix elements calculated at NLO in QCD with using the CT10 PDF set. The W/Z + jets events are normalised to the next-to-next-to-leading-order (NNLO) cross sections [129]. For the specific cases of W and Z+jets production, filters on vector boson p_T and heavy flavor content have been used to guaranty enough statistics in the full phase space. The normalisation of the W+jets background is corrected by a data-driven method (see section 5.6.2).

Diboson (WW, WZ, ZZ) productions are also simulated using the Sherpa 2.1.1 generator, including multi-lepton processes considering all diagrams with four electroweak vertices. The predicted cross-sections for this sample normalization is evaluated at NLO precision in QCD [130].

The multi-jets background is evaluated with a data driven method (as it is presented in the next section).

5.6 Data driven backgrounds

5.6.1 Non-prompt leptons

The matrix method

The selection of events is based on the identification of a “prompt” isolated lepton coming from the decay of W-bosons (“real” leptons). Acceptance, quality and isolation requirements are applied to select these leptons, but non-prompt leptons and non-leptonic particles may also satisfy such selection criteria although rarely (“fake” leptons). In that case, given the huge QCD multi-jet background cross-section in the proton-proton collisions, the contribution of such background in the signal region is not negligible. Such small phase space of the QCD multi-jet is difficult to model using MC simulations. Therefore, to achieve the estimation of this background, a data-driven method so-called “matrix method” is implemented.

This method requires a lepton selection looser than the one demanded in the analysis (“loose lepton”) to built a sample with a large fraction of QCD multi-jet events. The leptons passing the selection demanded in the analysis are known as “tight leptons”, while the lepton passing the looser selection but failing the tight one are called “anti-tight leptons”. Therefore, the loose lepton sample is composed by tight and anti-tight leptons. The total number of events in loose lepton sample, N_L can be defined as:

$$N_L = N_{prompt} + N_{QCD}, \quad (5.4)$$

where the N_{prompt} and N_{QCD} are the number of events containing a prompt and a fake loose lepton. On the other hand, the tight lepton sample is also composed of prompt and QCD components. An efficiency ϵ is defined as the probability for a real loose lepton to pass the tight lepton selection. The false-identification rate f denotes the probability for a fake loose lepton to pass the tight lepton selection. Therefore, the number of events in the tight lepton sample, N_T can be written as:

$$N_T = \epsilon \times N_{prompt} + f \times N_{QCD}. \quad (5.5)$$

Solving eq. 5.4 and 5.5 for N_{prompt} and N_{QCD} , it is possible to estimate the real and fake lepton contributions in the tight lepton sample:

$$f \times N_{QCD} = \frac{\epsilon f}{\epsilon - f} N_A - \frac{(1 - \epsilon) f}{\epsilon - f} N_T, \quad (5.6)$$

where N_A is the number of events with anti-tight leptons. It will be presented later how ϵ is derived in region dominated by real leptons and f in a region dominated by non-prompt and fake leptons.

The QCD multi-jet background is then estimated with data events passing the analysis selection with a looser lepton quality, where the events with tight and anti-tight lepton are weighted according to Eq. 5.6. The parametrisation used should not introduce a large miss-modelling or trends to the shapes of the observables considered in the analysis. This is difficult to achieve since an ideal parametrisation, using all the meaningful observables, can not be used because the datasets are statistically limited.

Loose lepton definitions

The lepton quality for the loose selection is LHMedium electrons and Loose muons, without implementing the lepton isolation (Table 5.1). In addition, the prescaled trigger HLT_mu20_L1MU15 is also used to increase the statistics of the anti-tight sample at low muon p_T . Its prescale factor is about 10 during 2015 data-taking to regulated its high rates, due to the absence of isolation.

Selection	objects	electrons	muons
loose	triggers lepton	nominal (section 5.2.2) LHMedium AND no isolation	nominal OR HLT_mu20_L1MU15 Loose AND no isolation
tight	triggers lepton	nominal LHTight AND LooseTrackOnly	nominal Medium AND LooseTrackOnly

Table 5.1: Definitions of the loose and tight lepton selection for the matrix method QCD estimation.

Efficiency measurement

Real efficiencies: a typical calculation of the ϵ_{real} involves the reconstruction of Z-boson using a tag-and-probe method using a selection which would be composed in electron or muon pairs. Then, the ϵ_{real} term could be calculated selecting a lepton satisfying the loose selection criteria and verifying how often another lepton satisfying the tight selection criteria can be found in the Z-boson mass window used. But for the current analysis, the real rates are extracted from the SM $t\bar{t}$ simulation (with the lepton scale factors implemented) to get rates from a topology closer to the

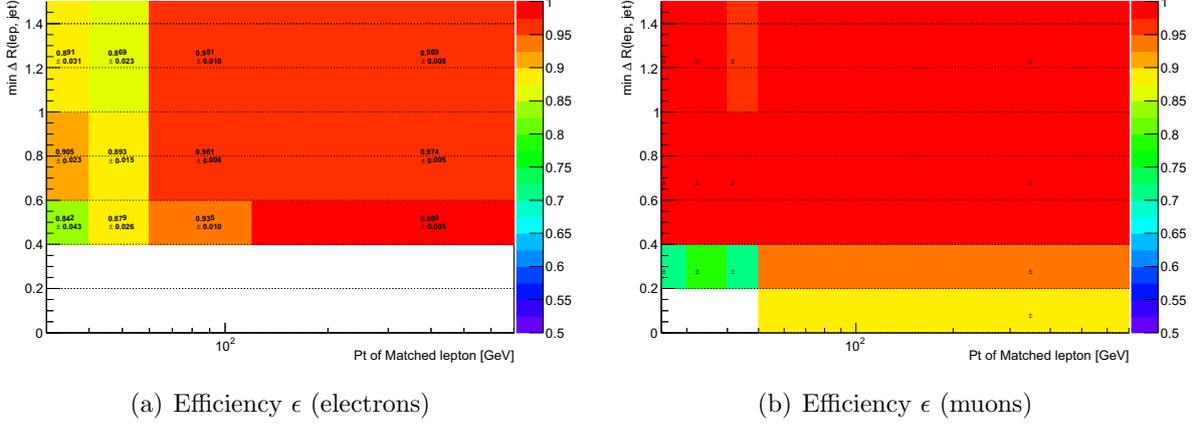


Figure 5.6: (a, b) Efficiencies ϵ for loose prompt leptons to be identified as tight, for electrons (left) and muons (right). For the electron channel, the overlap removal rejects events if the minimum $\Delta R(\text{lepton}, \text{jet}) < 0.4$.

one demanded on the analysis and small statistical uncertainty. The events are selected according to the standard analysis selection with a truth-matching requirement ($\Delta R(l^{reco}, l^{truth}) < 0.4$), to ensure an enriched real lepton sample. The efficiencies ϵ are measured for each lepton flavor using both types of lepton selection criteria: “tight” and “loose” lepton. Finally, the $\frac{N_T}{N_L}$ ratio is defined as the real rates (see Figures 5.6). The ϵ_{real} shows a strong dependency on the lepton p_T and on the angular separations between a jet and a lepton (Figure 5.6).

Fake efficiencies: a control region is defined to enhance the QCD multi-jet contribution in data, denominated CR_{fake} . The QCD multi-jet contribution in CR_{fake} is obtained from the data in this control region after a “background” subtraction to remove the prompt lepton contamination in both tight and loose lepton samples (Figures 5.7) and, the false-identification efficiency $f = \frac{N_T^{QCD}}{N_L^{QCD}}$ is computed.

The CR_{fake} is defined by reverting the cuts of the standard analysis selection devoted to reject the QCD multi-jet events. Such selection depends on the lepton flavor since the sources of fake leptons are different. The requirements $E_T^{miss} > 20$ GeV and $m_T^W + E_T^{miss} > 60$ GeV in the standard selection strongly reject the QCD multi-jet contribution (see section 5.2.2). Therefore, reverting these cuts can be used to build the CR_{fake} for the electron channel. An alternative selection can be used to find a CR_{fake} for the muon channel, by reverting the direction of the $|d_0/\sigma_{d_0}|$ cut. In addition to these change, a looser large-R jet selection is required to reduce the statistical uncertainty of the fake rates and the “top-tagging” is modified to reject the $t\bar{t}$ events.

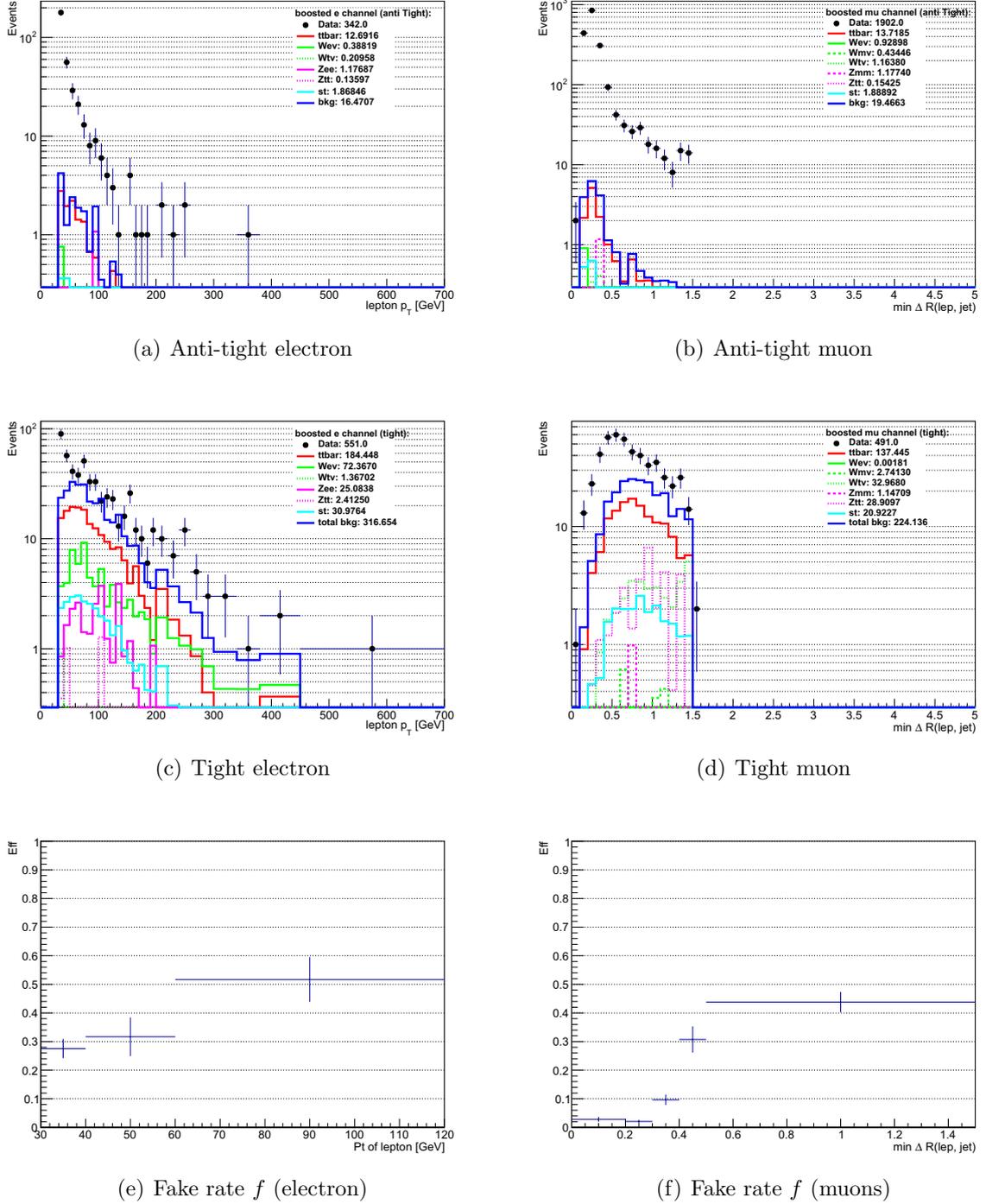


Figure 5.7: Data and MC predictions in CR_{fake} which is dominated by multi-jet events. Anti-tight (a, b) and tight (c, d) lepton samples (to judge their purity) and (e, f) fake rates f for the fake loose leptons to be identified as tight, as a function of the lepton p_T and minimal ΔR with a jet, for electrons (left) and muons (right).

The following changes are implemented to the analysis selection to build the CR_{fake} :

- At least one large-R jet with $p_T > 200$ GeV without any cut on τ_{32} and $m \leq 70$ GeV;
- and for the electron channel:
 - $E_T^{miss} < 20$ GeV;
 - $E_T^{miss} + m_T^W < 60$ GeV;
- and for the muon channel:
 - Inverting the direction of the cut on the impact parameter ²: $|d_0/\sigma_{d_0}| > 3$.

The fake rates show a dependency with the p_T of the lepton for the electron channel (Figure 5.7(e)) and the minimum $\Delta R(\text{jet}, \text{lepton})$ for the muon channel (Figure 5.7(f)). Despite all the efforts provided to increase the statistics, it can be seen fake rates are affected by relative large statistical uncertainties.

The QCD multi-jet background is first estimated in the CR_{fake} , since it is expected to fill the gap between the data and the electroweak backgrounds. A good agreement between the data and backgrounds is expected as the fake rates have been measured in these regions. The distributions of the reconstructed $m_{t\bar{t}}$ (Figure 5.8) shows a reasonable agreement between the data and the prediction, covered by a conservative 50% flat error and the statistical uncertainties.

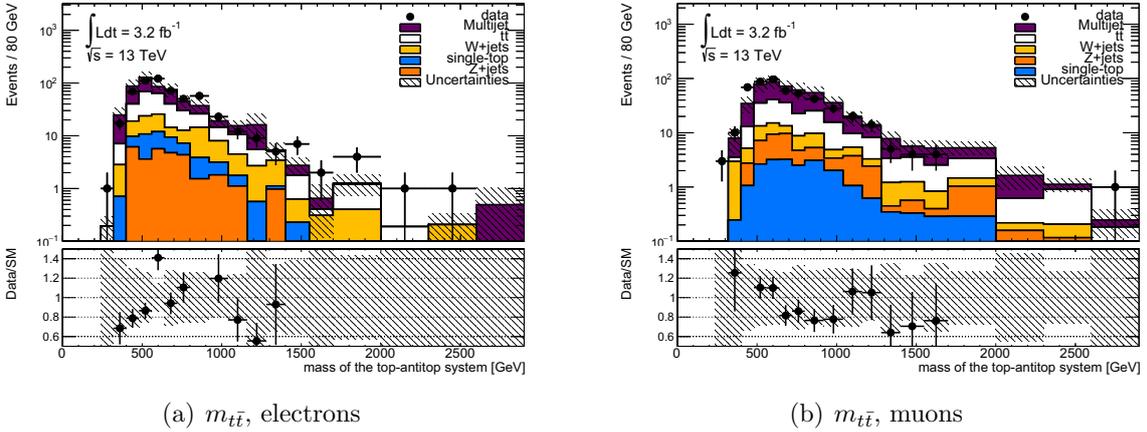


Figure 5.8: Reconstructed $m_{t\bar{t}}$ in the $CR_{fake}^{\text{boosted}}$ control region, for the electron (left) and muon (right) channels. The grey area indicates the impact of the 50% systematic uncertainty on the multi-jet background prediction on top of the statistical uncertainty.

²It is been checked that the reversion of this cut does not introduce a strong bias.

5.6.2 W+jets background normalisation

Because of the large theoretical uncertainty on the cross-section of W+jets, and its significant contribution on the final yields (see section 5.8). The MC simulation is renormalised using a data-driven technique that provides a more accurate prediction. It is based on the measurement of the asymmetry for the production of the bosons W^+ and W^- in the LHC proton-proton collisions [137], where the rate is well known theoretically. The normalisation factor is determined by the comparison of the W-boson charge asymmetry, in data and simulation. An overall normalisation is obtained in a W + jets dominated control sample, where the events are selected using the boosted selection (section 5.2.3), but removing the b-tagging requirement and inverting the top-tagging requirements in the large-R jet, to decrease the statistical uncertainty of the normalisation factor.

The total number of W+jets events measured in data ($N_{Data,W^+} + N_{Data,W^-}$) is given by:

$$N_{Data,W^+} + N_{Data,W^-} = \left(\frac{r_{MC} + 1}{r_{MC} - 1} \right) (D_{\text{asym}}^+ - D_{\text{asym}}^-), \quad (5.7)$$

where r_{MC} is defined as the ratio of the number of W+jets events in MC with a positive and negative charge lepton ($r_{MC} = N_{MC,W^+}/N_{MC,W^-}$) and D_{asym}^\pm is the number of observed events with a positive (negative) lepton. While the contribution to the charge-asymmetry by the backgrounds $t\bar{t} + V$, single-top and diboson has to be subtracted from the D_{asym}^\pm (right side of eq. 5.7), the other backgrounds $t\bar{t}$, multi-jet and Z+jets are charge-symmetric, and cancel themselves out. Scale factors (CA) applied on top of the W + jets normalisation are calculated as the ratio of $N_W^+ + N_W^-$ evaluated from data and predicted from MC simulation:

$$CA = N_{Data,W}/N_{MC,W} = \left(\frac{r_{MC} + 1}{r_{MC} - 1} \right) \frac{(D_{\text{asym}}^+ - D_{\text{asym}}^-)}{N_{MC,W}} \quad (5.8)$$

The value obtained for CA and its respective statistical errors in e+jets events is 0.82 ± 0.18 , while for μ +jets events is 0.73 ± 0.12 [121]. Any bias induced by the selection was found to be negligible compared to the statistical uncertainty of such scale factors [119].

5.7 Systematic Uncertainties

Many sources of systematic uncertainties can alter the predicted shape of $m_{t\bar{t}}^{\text{reco}}$ distribution for the signal and the backgrounds. This section describes de main sources of systematic uncertainties that are taken into account to access the compatibility of the data with the SM prediction.

Systematic uncertainty of integrated luminosity The integrated luminosity in data was measured with Van der Meer scans. For the 2015 data set, the total uncertainty in the luminosity is 5% [131], which is applied as a constant shift to each simulated background except for W+jets.

Systematic uncertainties on the physics objects:

Lepton reconstruction and identification The modelling of the electron and muon trigger efficiencies, identification efficiencies, energy scales and resolutions are studied using leptonic Z-boson decays in data and simulation at $\sqrt{s} = 13$ TeV. The modelling of the isolation requirements depends on the hadronic activity near the lepton, and the measurements are extrapolated to the $t\bar{t}$ environment giving uncertainties for both electrons and muons [132, 98].

Small- R jets An estimation of the impact of the jet energy scale uncertainty for small- R jets is performed using the uncertainties for calibrated jets broken down into 3 components built from a combination of the various systematic uncertainties affecting the jet calibration. The jet energy resolution uncertainty is extrapolated to $\sqrt{s} = 13$ TeV using $\sqrt{s} = 8$ TeV data [133]. The energy resolution can only be varied “up” (worsening of the resolution). The “down” shift distribution is created by symmetrizing the “up” variation with respect to the nominal distribution. If bin i has an $x\%$ shift in the “up” histogram, then bin i in the symmetrized “down” histogram has a $-x\%$ shift with respect to the nominal bin value.

Large- R jets The scale uncertainties on the large- R jets (which is propagated to the top-tagging performances) are broken down into 2 components (“Run 1”, the uncertainty extracted from 2012 data, and “CrossCalib”, the MC-based cross-calibration uncertainty to extrapolate from 2012 to 2015 data). The large- R jet energy, mass scales and τ_{32} scale are varied in simulation according to the uncertainties derived from $\sqrt{s} = 8$ TeV simulation and in-situ calibration [134].

The systematic associated to large- R and small- R jet collections are treated as non-correlated, it has been checked that other scenarios assuming some correlations between the components of the jet collections do not significantly change the result of the analysis.

b -tagging of the track-jets The b -tagging uncertainties are determined from $\sqrt{s} = 8$ TeV data extrapolated to $\sqrt{s} = 13$ TeV including additional uncertainties to account the presence of the IBL detector. They has been obtained with the eigenvector breakdown for the b -flavor (c - and light-flavor) quarks, and 2 additional components specifically for the extrapolation of the scale factor in high p_T regimes, which are correlated across the flavors.

Systematic uncertainties on the background modelling:

The uncertainty on the $t\bar{t}$ cross-section is one of the main sources of systematic uncertainty, $^{+5.6\%}_{-6.1\%}$, which is implemented as a normalisation variation of the SM $t\bar{t}$ Monte Carlo simulation [124, 125, 135]. The uncertainty on the modelling of the hard scatter and parton showering are extracted from a shape comparison between Powheg + Herwig and MC@NLO + Herwig, both using ATLASFASTII. The impact on the shape is symmetrized to get the downward variation. Uncertainties for the hadronisation modelling is extracted from a shape comparison between a regular Powheg + Pythia 6, taking into account the shower radiation. In addition, the correction for electroweak loops has an uncertainty of 10%.

The normalisation SF for the W+jets background derived with a data-driven technique (section 5.6.2) has a dominant statistical uncertainty. A conservative approach has been adopted to set an uncertainty to QCD multi-jets estimate, which are dominated by the measurements on the fake rates. Another control region CR_{fake2} is defined identical to CR_{fake} but without requirements on the largeR jets. A new fake rate parametrization can be set, where the yields predicted in the signal region are changing by 50%, which are set as the systematic uncertainty on this background. The single-top cross-section the production modes are assigned an systematic uncertainty of $\pm 5.3\%$ [136].

5.8 Control plots

The available data are compared to the MC predictions in Table 5.8 and Figures 5.9 to 5.15. After the boosted event selection criteria are applied (see section 5.2.3), 3352 events remains in the e+jets selection and 3074 events in the μ +jets selection. The comparisons are shown for lepton kinematics in Figure 5.9 and E_T^{miss} in Figure 5.10. The kinematics of the closest jet to the lepton in Figure 5.11. Then the kinematics of the hadronic top jet in Figures 5.13 and 5.14. In general, the agreement is satisfactory between data and expectation. Finally the reconstructed top-quark $t\bar{t}$ mass is shown in Figure 5.15, where a Z_{TC2} signal with mass of 2 TeV is overlaid. A test of the compatibility between the data and the SM prediction can now be performed.

5.9 Compatibility with SM only hypothesis

The BumpHunter [138] algorithm is a statistical test which evaluates the compatibility of some dataset with respect to a specific hypothesis. Such tests, often called “hypothesis tests” or “goodness of fit”,³ test the consistency of a dataset with a null hypothesis (the background-only hypoth-

³like Pearson’s χ^2 and the Kolmogorov-Smirnov test.

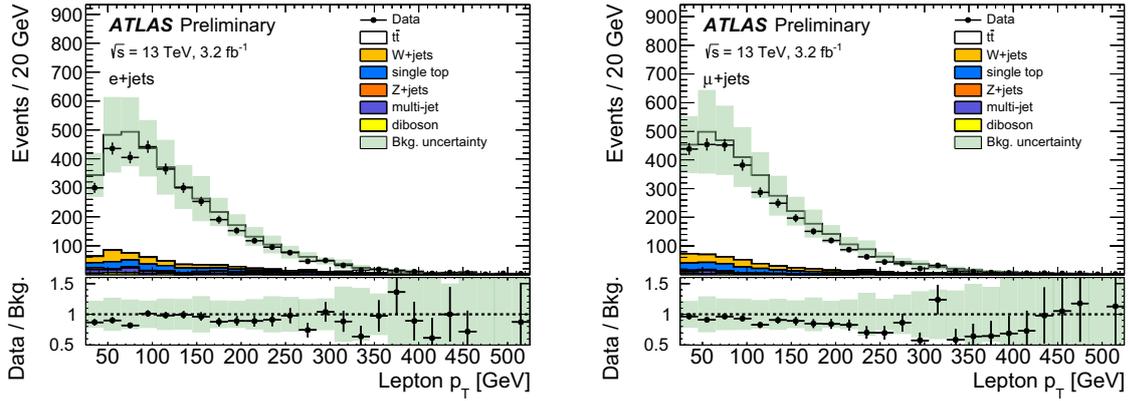


Figure 5.9: The distribution of the transverse momentum of the lepton in the (a) $e+jets$ and (b) $\mu+jets$ selections.

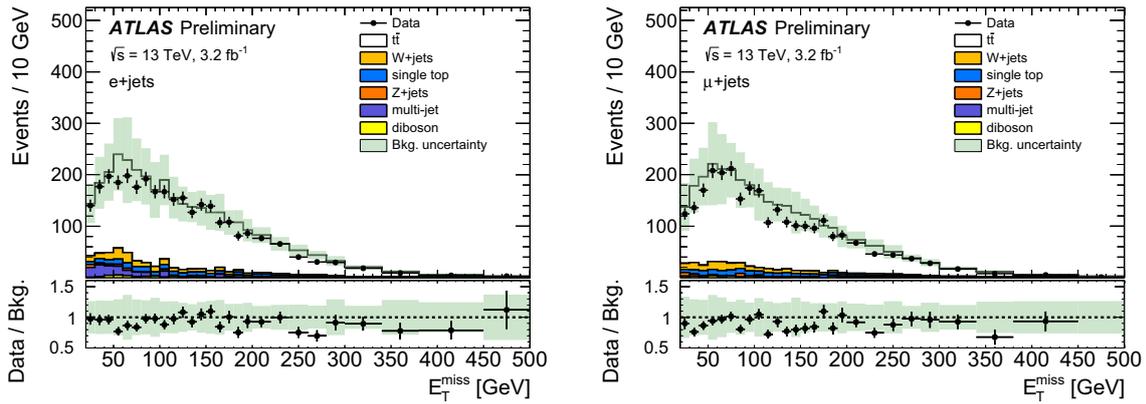


Figure 5.10: The distribution of the E_T^{miss} in the (a) $e+jets$ and (b) $\mu+jets$ selections.

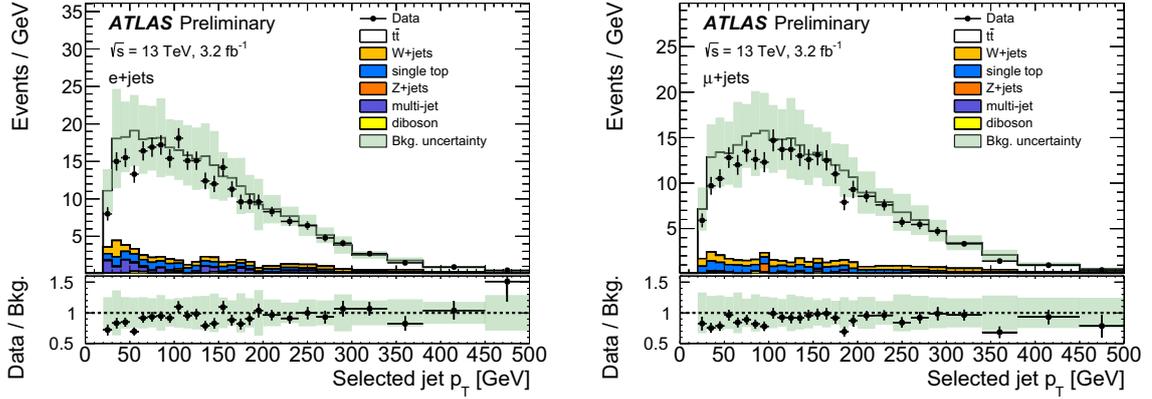


Figure 5.11: The distribution of the transverse momentum of the hardest small-R jet with $\Delta R(l, \text{jet}) < 1.5$ in the (a) e+jets and (b) μ +jets selections. The SM background components are shown as stacked histograms.

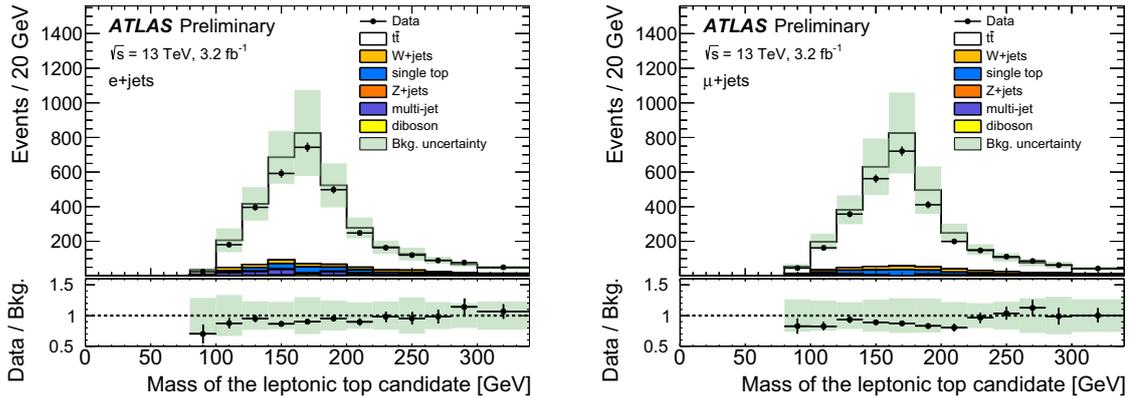


Figure 5.12: Reconstructed mass of the leptonic top candidate.

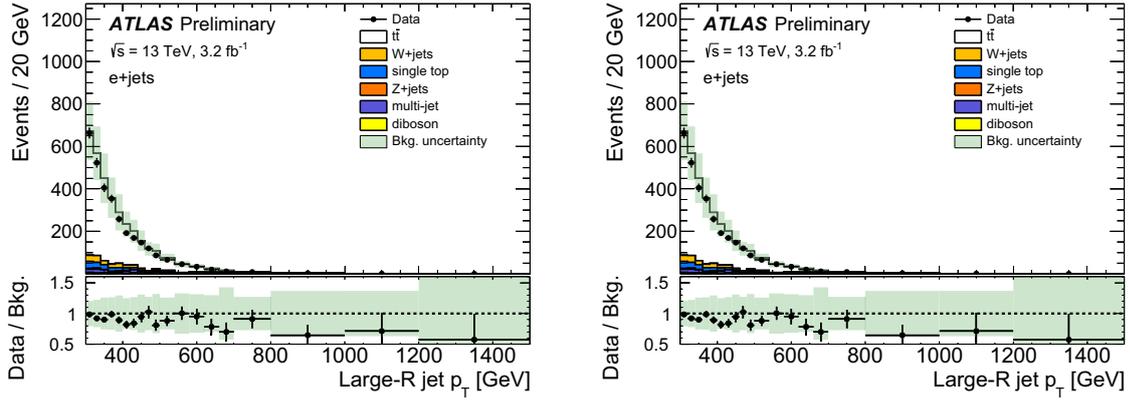


Figure 5.13: Transverse momentum of the large- R jet in linear scale.

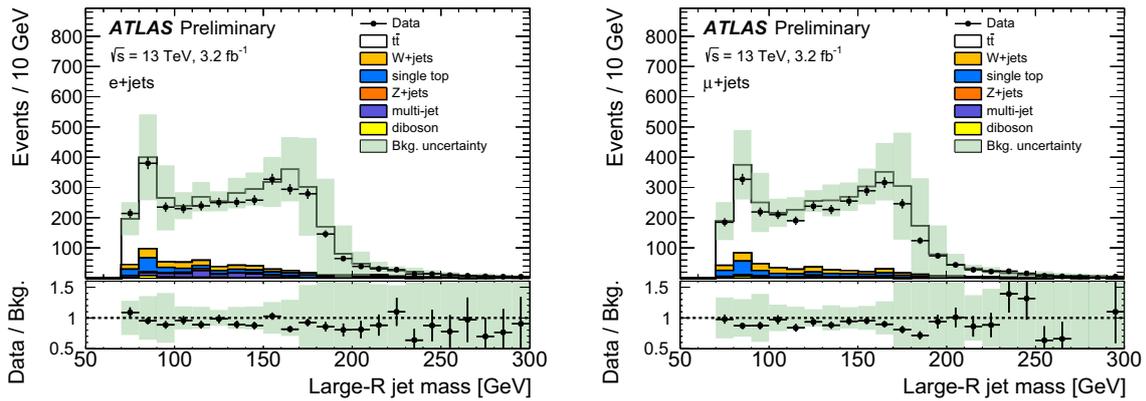


Figure 5.14: Mass of the large- R jet.

Type	Yield (e + jets)	Yield (μ + jets)
$t\bar{t}$	3000 ± 700	3000 ± 700
W +jets	200 ± 140	200 ± 40
single top	190 ± 40	180 ± 40
Z +jets	33 ± 12	26 ± 12
multi-jet	130 ± 70	19 ± 11
diboson	46 ± 11	37 ± 8
Total	3700 ± 800	3400 ± 800
Data	3352	3074

Table 5.2: Data and expected background event yields after the e+jets and μ +jets selections with their respective total systematic uncertainty on the expected background yields.

esis in this case) denoted H_0 . Once an inconsistency with H_0 is established, several alternative signal hypotheses can be tested to characterise a discovery.

When the data is compared to H_0 , their difference is quantified by a single number called “statistical test”. High values of the statistical test are associated to high inconsistency between data and H_0 . The statistical significance associated to the observation of new phenomena is usually expressed using a p-value approach, which evaluates the probability to get the result observed in the data given H_0 .

The mass of the $t\bar{t}$ resonances is a free parameter of BSM models, so a signal could appear anywhere in the $m_{t\bar{t}}^{reco}$ spectrum. Then, the test of H_0 has to be done for all the possible ranges of the $m_{t\bar{t}}^{reco}$. However, the probability to observe a disagreement because of a fluctuation can be high if the statistical test is performed in all the possible range. This is known as the “Look Elsewhere Effect” (LEE). The objective of BumpHunter is to search the location of the most significant deviation of the data with respect to the null hypothesis, and assign it a p-value free from the LEE.

The BumpHunter algorithm is the following: the data and expected background discrimination are compared in sliding windows of variable size, with a minimum width of two bins. The p-value of the most prominent bump (or dip when a deficit is saved). In each window i , the data count is d_i , the background yield is b_i and the p-value (i.e the probability to observe at least data under H_0) $P(d_i, b_i)$ is defined when searching for an excess as:

$$P(d_i, b_i) = \begin{cases} \Gamma(d_i, b_i) = \sum_{n=d_i}^{\infty} \frac{b_i^n}{n!} e^{-b_i} & \text{if } d_i \geq b_i \\ 1 - \Gamma(d_i + 1, b_i) & \text{if } d_i < b_i \end{cases} \quad (5.9)$$

where Γ is the Gamma function and the inequality signs are reversed to look for deficits. The

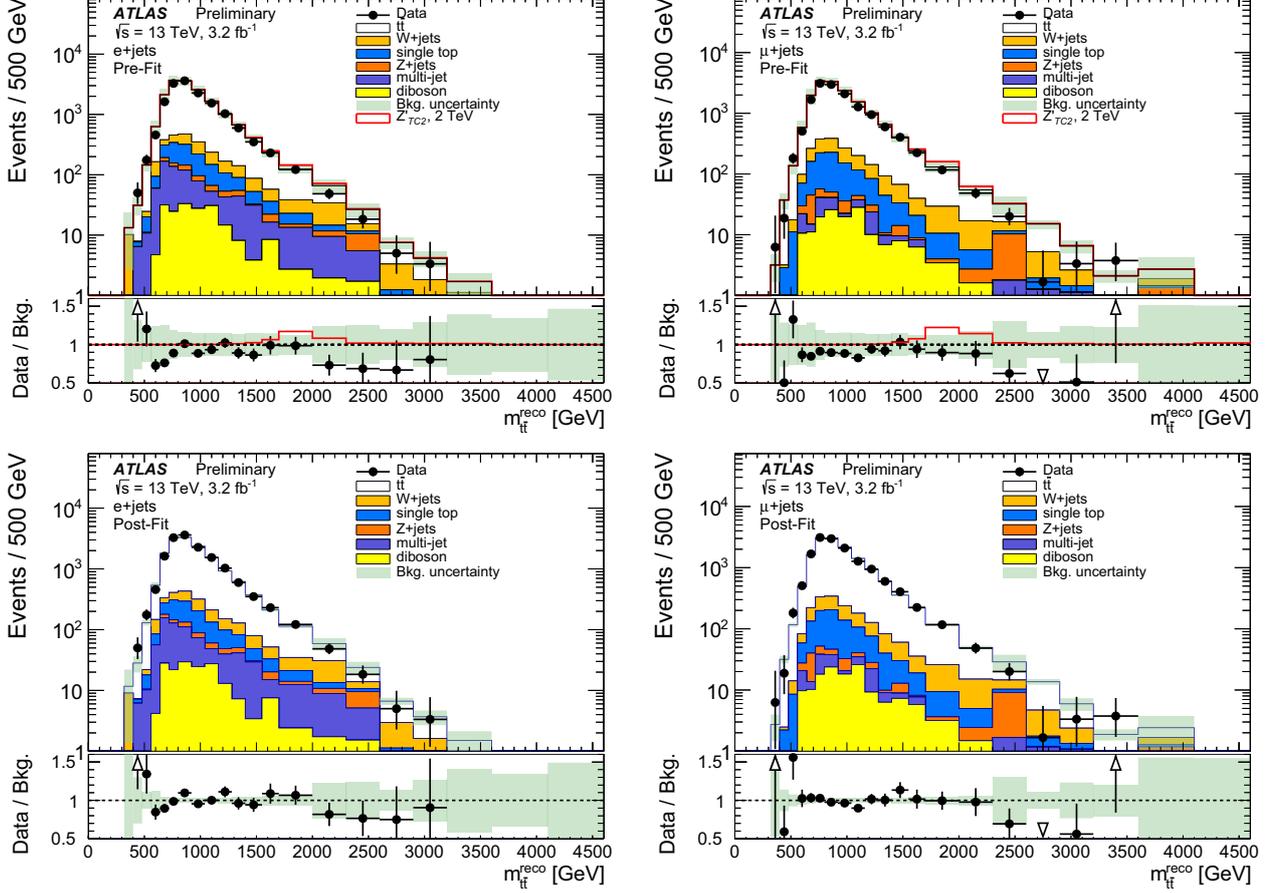


Figure 5.15: $m_{t\bar{t}}$ distributions before (top) and after (bottom) the profiling of the nuisance parameters for the electron (left) and muon (right) channels.

smallest $P(d_i, b_i)$ from all the windows, P_i^{\min} corresponds to the most interesting (discrepant) window. To include the systematic uncertainties, the definition of $P(d_i, b_i)$ is redefined as:

$$P(d_i, b_i) \rightarrow P(d_i, b_i + \lambda_i \theta_i) e^{(\frac{\lambda_i^2}{2})} \quad (5.10)$$

where θ_i is the total systematic uncertainties for the considered window, and λ_i is a real number (between -8 and 8) that maximizes the value of $\frac{(b_i + \lambda_i \theta_i)^{d_i}}{d_i!} e^{-(b_i + \lambda_i \theta_i)}$.

The p -value free of LEE (named global p -value) of the most interesting bump is found by comparing the statistical test from data with the statistical test found in at least $N = 10,000$ pseudo experiments, where the pseudodata is generated by Poisson fluctuations of the expected back-

ground considering the systematic variation. The BumpHunter test statistic t is computed as when searching for an excess:

$$t = \begin{cases} 0 & \text{if } d_i \leq b_i \\ -\log P_i^{\min} & \text{otherwise} \end{cases} \quad (5.11)$$

and the global p -value is defined as:

$$p\text{-value} = \frac{\int_{t_{obs}}^{\infty} f(t)}{\int_0^{\infty} f(t)} \quad (5.12)$$

where $f(t)$ is the distribution of the statistical test values from the pseudodata and t_{obs} is the test statistic obtained from data. A p -value of 0 means that no deviation was observed in the pseudo-experiments that is bigger than the one obtained in data, i.e. the deviation is very large (and more pseudo-experiments are needed).

If a $t\bar{t}$ resonance exists, its presence should be detected in the various spectra at approximately the same mass point. The BumpHunter uses this property by requiring an overlap of discrepancy across the channels that allows to strongly reduce the LEE and hence improve the sensitivity.

Configuration	Channel	mass range	p-value
excess (syst, prefit)	e	400 - 560	0.896±0.003
	μ	400 - 1040	0.836±0.003
	comb.	400 - 560	0.284±0.004
deficit (syst, prefit)	e	3200 - 4100	0.514±0.004
	μ	2600 - 3200	0.020±0.001
	comb.	no overlap	

Table 5.3: Most significant excess and deficit with their associated global p -values, using the prefit spectra and the associated systematic uncertainties. The compatibility with the SM-only hypothesis is tested for e+jets and μ +jets channels, and for their combination. The significance is not enough for the p -value of 0.020 for the muon channel (2.045σ).

In Table 5.3, the p -values and mass ranges of the most interesting deviations are listed, as well as the corresponding significance in sigmas, when taking the systematic uncertainties into account. In order to avoid fake bumps and dips created by a systematical effect and no benefit of the bring by the profiling, the expected distribution after fitting the systematic is also used to search for bump/dip (Table 5.4). These results allows us to safely claim there is no visible signal in the data and then limits can be set.

Configuration	Channel	mass range	p-value
excess (syst, postfit)	e	400 - 560	0.331±0.004
	μ	320 - 560	0.562±0.004
	comb.	400 - 560	0.088±0.002
deficit (syst, postfit)	e	560 - 720	0.134±0.002
	μ	2600 - 3200	0.092±0.002
	comb.	no overlap	

Table 5.4: Most significant excess and deficit with their associated global p-values, using the postfit spectra and the associated systematic uncertainties. The compatibility with the SM-only hypothesis is tested for e+jets and μ +jets channels, and for their combination.

5.10 Upper production cross section limits on $t\bar{t}$ resonances

A frequentist approach using the CLs method is used to set 95% Confidence Level (CL) limits on cross-section of the Z' signal. The limit setting procedure uses a likelihood for a particular signal assumption that is defined as [139]:

$$L(\mu, \Theta) = \prod_{i=0}^{\text{channels, bins}} \frac{e^{-\mu a_{Z',i} \sigma_{Z'} + b_i} (\mu a_{Z',i} \sigma_{Z'} + b_i)^{D_i}}{\Gamma(D_i + 1)} C(\Theta) \quad (5.13)$$

where D is the expected data yield, b is the expected background yield, $\sigma_{Z'}$ is the cross section of the Z' signal, $a_{Z'}$ is the acceptance of the signal and μ is the signal strength, which is the parameter of interest. The function C indicates the set of constraints applied on the nuisance parameters Θ , such as the systematic uncertainties in the background and signal, and the luminosity measurement uncertainty.

The hypothesis testing is based on the profile likelihood ratio test statistic Λ , defined as follows:

$$\Lambda(\mu) = \frac{L(\mu, \hat{\hat{\Theta}}(\mu))}{L(\hat{\mu}, \hat{\Theta})}, \quad (5.14)$$

where the single circumflex indicates the unconditional maximum likelihood estimate of a parameter, while the double circumflex indicates the maximum likelihood estimate assuming a specific value of μ . Assuming that the test statistic $-2 \ln(\Lambda(\mu))$ is distributed according to a χ^2 distribution with one degree of freedom, the variation on the logarithm of the likelihood can be used to set a 95% CL on the upper limit of the signal production cross-section times branching ratio, and to estimate the impact of the systematic uncertainties in the discovery significance.

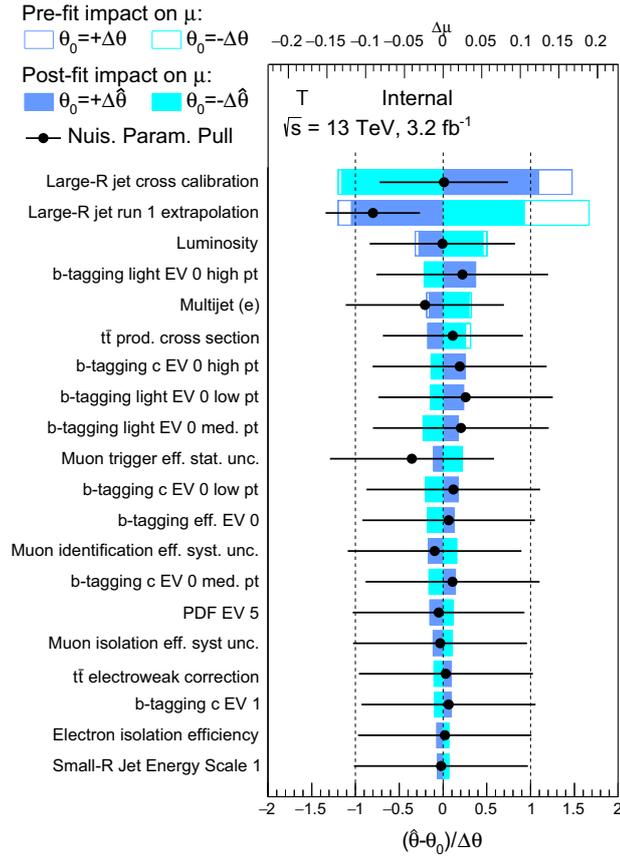


Figure 5.16: The largest impacts of the nuisance parameters (with their pre- and post-fit magnitudes) on the fitted signal strength μ and their pulls for a fit performed under the hypothesis background plus a Z' signal with a mass of 2 TeV. The impact of an uncertainty on the best-fit value is obtained by fixing the nuisance parameter θ to the one-sigma range limits and repeating the fit. The most significant uncertainty is the JES of the large-R jets, impacting the signal strength in 10%.

The largest impacts of the systematic uncertainties (with their pre- and post-fit magnitudes) on the fitted μ is on figure 5.16. Some large nuisance parameters (SM $t\bar{t}$ cross-section and the 2 systematics related to the large-R jets that have large impact on the $m_{t\bar{t}}$ shape) are logically constrained thanks to the large statistic of the signal region. These constraints could already be predicted from a fit on Asimov pseudo-data (ie pseudodata equal to the expected background). It is worth to notice that the expected limits extracted from fit on data and pseudo-data are very similar, confirming the nuisance parameter have not been strongly pulled.

The limits on cross-section \times branching ratio can be interpreted as observed (expected) constraints on the Z' mass: $0.75\text{TeV} < m(Z') < 2.0\text{ TeV}$. For reference, $m_{t\bar{t}}$ distributions after the profiling of the nuisance parameters are presented in Figure 5.15. The resulting limit for each signal point is shown in Figure 5.17.

Outlook

A search for heavy particles decaying into $t\bar{t}$ quarks has been performed in the boosted-top scenario regime. The search uses the data corresponding to an integrated luminosity of 3.2 fb^{-1} of proton-proton collisions at a centre-of-mass energy of 13 TeV. No evidence of process beyond Standard Model predictions have been observed in the invariant mass spectra of the $t\bar{t}$ system. As described in the analysis strategy, upper limits are set on the possible cross-section \times branching ratio on a narrow Z' benchmark model (smaller than 3%). Based on these results, the existence of a narrow leptophobic topcolor Z' (width 1.2% of its mass) in a range $0.75\text{ TeV} < m_{Z'} < 2.0\text{ TeV}$ is excluded at 95% CL.

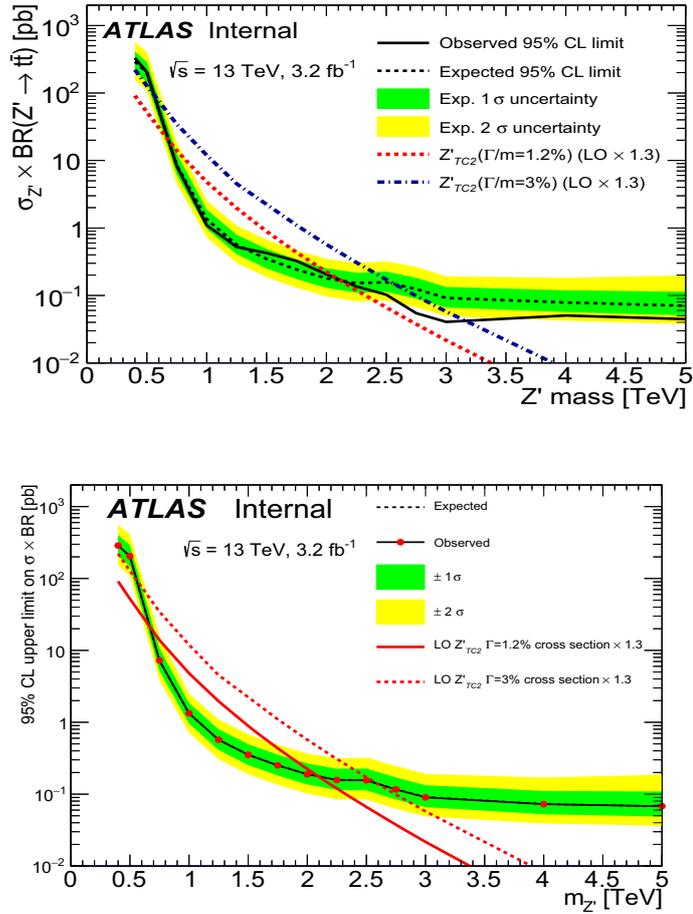


Figure 5.17: Top: Cross-section 95% CL upper limits on the Z'_{TC2} signal using Atlas data. Bottom: same result but using Asimov pseudo-data, for which the data equal to the expected background. Both expected results (from fits on data or pseudodata) are very similar.

Chapter 6

Improving the sensitivity at low $m_{t\bar{t}}$

Contents

6.1	Resolved selection	132
6.1.1	Low mass topology	132
6.1.2	Boosted and resolved combination	133
6.2	Estimation of the QCD multi-jet background	133
6.2.1	Requirements for the QCD multi-jet estimation	134
6.2.2	Changes with respect the previous QCD multi-jet estimation	134
6.2.3	Estimation of the real and fake rates	135
6.2.4	Performance of the QCD multi-jet background estimation	140
6.3	Reconstruction of the invariant mass of the $t\bar{t}$ system	147
6.3.1	χ^2 minimisation algorithm	147
6.4	Improving the $t\bar{t}$ invariant mass resolution	149
6.4.1	Impact of the neutrino reconstruction	149
6.4.2	Jet rescaling method	150
6.4.3	Impact of the rescaling	155

The search for $t\bar{t}$ resonances presented in the previous chapter has a low sensitivity in the low invariant mass region of the $t\bar{t}$ system, which is dominated by events with the resolved topology. As demonstrated by previous results [40, 119], a combination between the boosted and resolved channels would lead to an improvement of the sensitivity. The search for $t\bar{t}$ resonance for the resolved topology using 2015 datasets is under development at the time where this manuscript is written, but a status of the estimation of the QCD multi-jet background will be discussed in the first part of this Chapter.

The author have participated in the integration of the code for the selection and reconstruction of the $t\bar{t}$ events in the low mass region. The rates and parametrisation for the estimation of the QCD multi-jet background for the resolved analysis are also included in the reconstruction package (TopNtupleAnalysis). There are technical issues to solved first before doing the combination between boosted and resolved analyses using the Run 2 dataset.

The detector effects on the resolution of the reconstructed physics objects can hide the searched signal, and consequently reduce the sensitivity to new physics. In the second part of this chapter, the $m_{t\bar{t}}$ reconstruction method for the resolved topology is exposed, followed by an in situ calibration that could be implemented to balance detector resolution effects.

The studies using 8 TeV signals where developed in the Run 1 setup. The author have tested the jet rescaling method which improves the $t\bar{t}$ mass resolution, and other members of the team are currently working to test this method in the current packages used, and testing the impact on the sensitivity in the upper cross-section limits.

6.1 Resolved selection

6.1.1 Low mass topology

In the low $m_{t\bar{t}}$ range, the decay products of the top-quark can be individually recovered in the detector (Figure 6.1) leading to final states with 4 jets, 1 lepton and E_T^{miss} . Therefore, in addition to the pre-selection defined in section 5.2.2, the selection requires at least 4 small-R jets reconstructed with anti- k_T algorithm with $R = 0.4$, using EM+JES calibration (section 4.6.2). These jets are required to have a transverse momentum $p_T > 25$ GeV and $|\eta| < 2.5$.

The selected events must also contain at least 1 b-tagged track jet. The track jets are reconstructed with anti- k_T algorithm with $R = 0.2$, having at least two constituent tracks. A $p_T > 10$ GeV and

$|\eta| < 2.5$ is demanded for such jets.

A χ^2 algorithm is used to choose the physics objects which will be used to calculate the invariant mass of the $t\bar{t}$ system (described in section 6.3.1). A low χ^2 value reflects a good compatibility with a resolved $t\bar{t}$ topology. Then, only the events with a combination of physics objects that satisfy $\log_{10}(\chi^2) < 0.9$ are kept.

6.1.2 Boosted and resolved combination

The strategy for the combination of the boosted and the resolved scenarios is based on the orthogonalisation between both selections to avoid a statistical correlation for the statistical interpretation of the results. If an event does not satisfy the boosted selection, then it is asked to pass the resolved one [40]. This procedure splits the analysis in four channels, depending on whether the lepton in the final state is an electron or muon; and on whether the event topology is boosted or resolved. After an event is selected and tagged in one of these four categories, $m_{t\bar{t}}$ is calculated from the physics objects available in the events.

6.2 Estimation of the QCD multi-jet background

The matrix method (MM) described in the section 5.6.1, is reconsidered for a possible improvement on the estimate of the QCD multi-jets background, since the fake rates obtained had a large statistical uncertainty. The looser selection of the resolved topology allows to develop a better

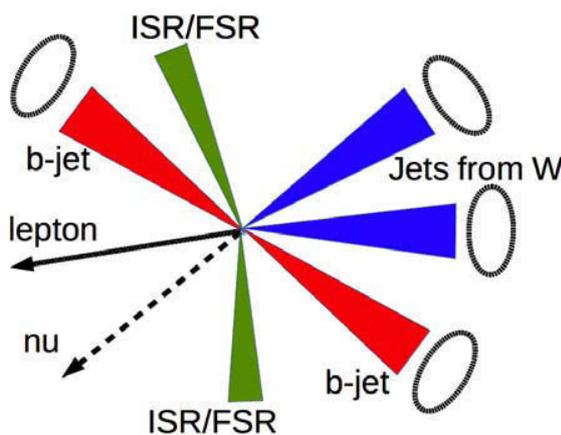


Figure 6.1: Scheme of a typical $t\bar{t}$ event in the resolved scenario. The χ^2 minimisation algorithm (section 6.3.1) is used to reject the jets from ISR/FSR for the reconstruction of $t\bar{t}$ invariant mass.

parametrisation thanks to the higher statistics, that could be used in the boosted channel.

6.2.1 Requirements for the QCD multi-jet estimation

The QCD multi-jet background needs to be estimated in the following regions for different purposes in the analysis:

Region enriched with W+jets events: a good QCD multi-jet estimation is needed to extract the W+jets normalisation using the method described in section 5.6.2. After the pre-selection, the events of this region are required to have exactly one loose lepton and at least two jets. In the future, the estimation of W+jets normalisation factors can be improved with an iterative estimation of the QCD multi-jet background using the corrected W+jets production in the subtraction procedure for the fake rate measurement.

Region enriched with $t\bar{t}$ events: which can be the boosted analysis (section 5.2.3) or the resolved analysis (section 6.1). Both regions are optimized to reject QCD multi-jet background, but some residual remanent should be estimated.

6.2.2 Changes with respect the previous QCD multi-jet estimation

Some changes have been performed with respect to the object definitions described in section 5.6.1 for this iteration of the QCD multi-jet estimation:

The isolation working point: it has been changed from `LooseTrackOnly` to `FixCutLooseTrackOnly` which is looser at high p_T and tighter at low p_T . This allows to increase the fake lepton rejection rate, which is important to secure the stability of the MM weights (see denominator in eq. 5.6). The impact on the signal acceptance is negligible.

The prescaled trigger: it has been dropped since the new QCD multi-jet enriched control region (see 6.2.3) has enough events to get a small statistical uncertainty.

Electron quality for the loose selection: it has been changed from `LHMedium` to `LHTight` electrons, without implementing the lepton isolation (summarised in Table 6.1). In this way, the loose and tight lepton selections are separated only by the isolation. This allows a stronger dependence of the fake rates on the p_T of the lepton and the isolation variables.

Selection	electrons	muons
loose	LHTight AND no isolation	Loose AND no isolation
tight	LHTight AND FixCutLooseTrackOnly	Medium AND FixCutLooseTrackOnly

Table 6.1: Summary of the tight and loose lepton selections.

The b-tagging algorithm: the MV2C20 algorithm has technical issues in the release used to produce this study. Therefore, it is replaced by MV2C10, which has a different training configuration but the working point used has the same 70% b-tagging efficiency than the previous algorithm.

In addition, the fake rates are obtained for two categories, depending on whether there is a b-tagged track jet (known as the “b-tagged” category) or not (known as the “non-b-tagged” category). The fraction of fake leptons produced by heavy flavor decays is expected to be larger in the b-tagged category. Therefore, an improvement on the QCD multi-jet description for both regions is expected thanks to this splitting, since the fake rate better takes into account the flavor dependence.

6.2.3 Estimation of the real and fake rates

It is crucial for the modelling of the QCD multi-jet shapes to find the observables for which ϵ and f have the largest dependence. A multi-dimensional parametrisation is created to take into account the correlation between the observables, allowing to ease the extraction of the shapes for the estimation of the QCD multi-jet background. The binning used for the parametrisation needs to be as thinner as possible to also improve the shape modelling. But high statistical fluctuation of the fake rates can introduce shape miss-modellings when they are applied in the signal regions. Moreover, it is important to have the lepton p_T in the parametrisation as it drives the “choose” of trigger and the cut on the isolation.

Real rate parametrization

The ϵ is extracted using truth-matched events in the W+jets dominated control region, using the SM $t\bar{t}$ simulations with the implementation of the lepton scale factors (lepton SF) derived from data.

Electron channel: the isolation requirement `pTvarcone` discriminates between loose and tight leptons, therefore the rates are expected have a strong dependence with lepton p_T and any isolation variable. The calorimetric isolation variable `topoetcone20` it has been chosen since it has been

seen to have a very strong dependence. A 2-dimensional parametrisation is implemented with those variables (Figure 6.2(a)).

Muon channel: the real rates are showing the strongest dependency with the p_T of the lepton and the minimum angular separation between a jet and a lepton, ΔR_{min} (Figure 6.2(b)). As expected, the variable R isolation is tighter at low lepton p_T and at small angular difference with respect to a jet.

The real rates, estimated with the truth matching of the SM $t\bar{t}$ simulation, has been checked to reproduce a similar dependence than computing the tight over loose ratio using the W+jets samples in a high m_T^W region, which is expected to be dominated by a real lepton contribution.

Fake rate parametrization

The CR_{fake} is defined for the measurement of f , where the loose leptons have a large probability to pass the tight definition. After the subtraction of the real lepton contribution using the MC, the ratio of the tight over loose events in each bin of the parametrisation, the fake rates are computed.

The CR_{fake} for electron channel is defined as:

- Exactly one loose electron
- A tight muon veto
- $E_T^{miss} < 20$ GeV and $E_T^{miss} + m_T^W < 60$ GeV
- At least 2 jets

The f are also expected to have a strong dependence with the isolation. A 2-dimensional parametrisation based in lepton p_T and `topoetcone20` is also implemented for the fake rates (Figures 6.3).

The CR_{fake} for muon channel is defined as:

- Exactly one loose muon, with $|d_0/\sigma_{d_0}| > 5^1$
- A tight electron veto
- $E_T^{miss} > 20$ GeV and $E_T^{miss} + m_T^W > 60$ GeV

¹It is a two-step procedure. Exactly one loose muon is demanded without a cut on $|d_0/\sigma_{d_0}|$. Then, this muon must satisfy $|d_0/\sigma_{d_0}| > 5$. This is important to ensure that E_T^{miss} is not strongly biased by a “real” muon that would have $|d_0/\sigma_{d_0}| < 3$.

- At least 2 jets

The observables for which the fake rate shows the stronger dependence are: m_T^W , E_T^{miss} and ΔR_{min} . The rates are expressed as function of m_T^W and E_T^{miss} for three regions of ΔR_{min} (Figures 6.4). In the bins in which the fake rates have large statistical fluctuation, f is extracted from the previous adjacent bins instead.

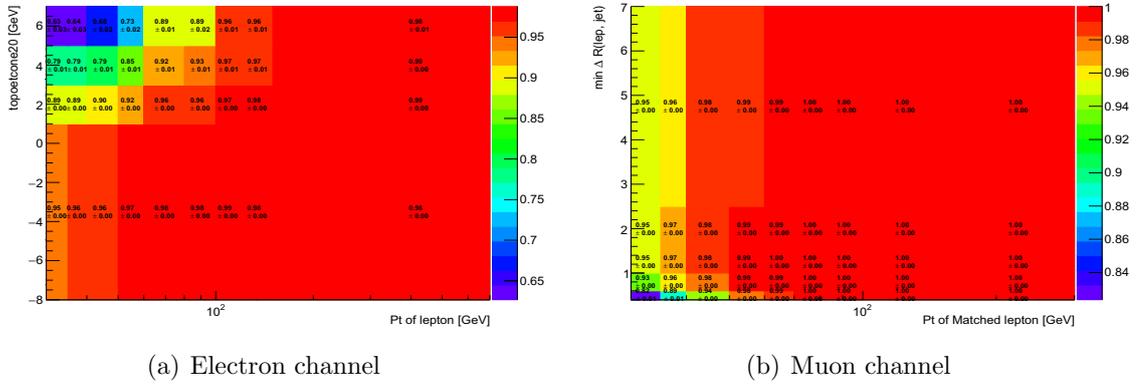


Figure 6.2: Efficiencies ϵ for loose prompt leptons to be identified as tight, for electrons (top) and muons (bottom) from simulations of 2015.

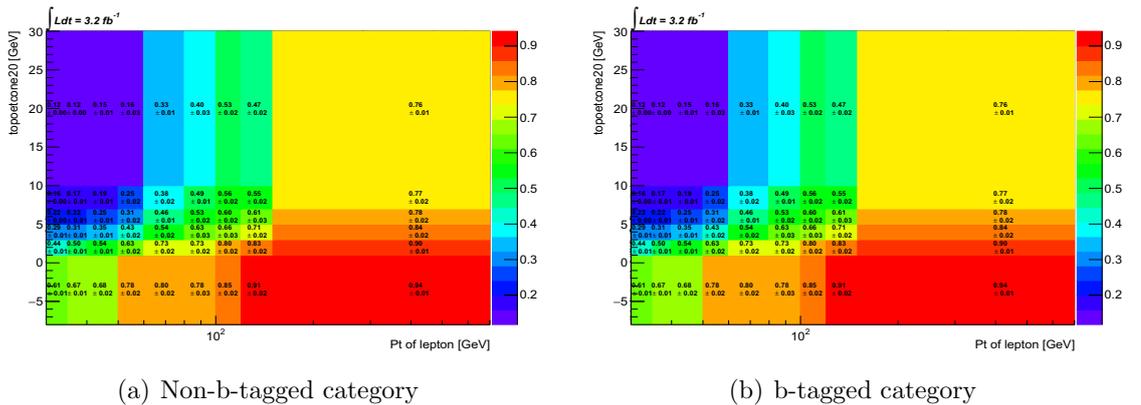


Figure 6.3: Fake rates f for the loose electrons to be identified as tight, as a function of the lepton p_T and $topoetcone$, for event with (left) and without (right) at least one b-tagged jet, as measured in CR_{fake} for 2015 data and simulations.

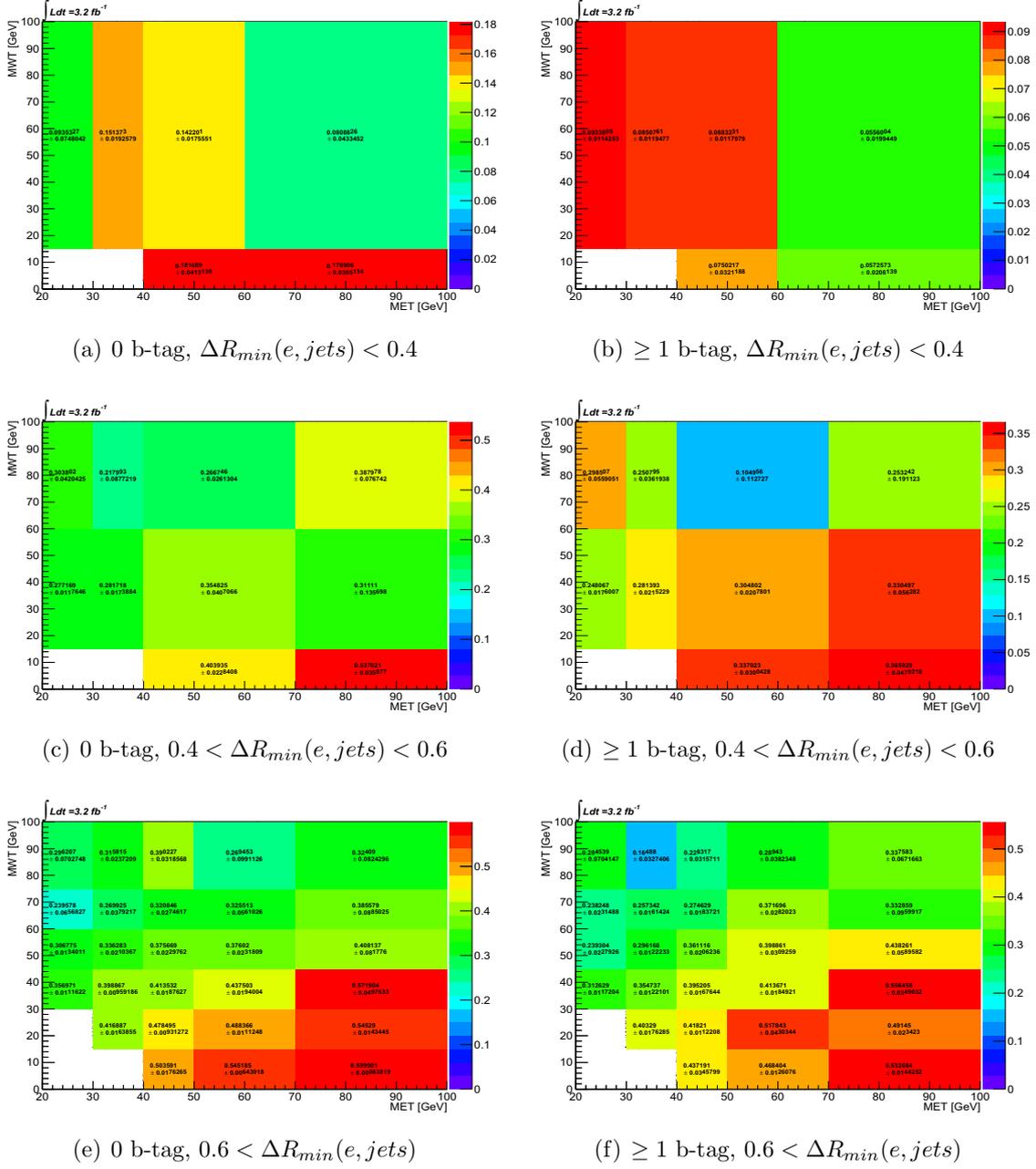


Figure 6.4: Fake rates f for the loose muons to be identified as tight, as a function of the E_T^{miss} , m_T^W and ΔR_{min} , for event with (left) and without (right) at least one b-tagged jet, as measured in CR_{fake} for 2015 data.

Closure test

The real and fake rates are then validated in the CR_{fake} , where the QCD multi-jet background is expected to fill the gap between the data and electroweak backgrounds. The estimation of the QCD multi-jet background seems to fit well in the data-MC gap (Figure 6.6 and 6.7).

Systematic uncertainties

The method used to obtain preliminary systematic uncertainties associated to the QCD multi-jet estimate is based on the measurement of this background in orthogonal regions to the signal region and the CR_{fake} . Such intermediary regions are called validation regions. Using the real and fake rates, the QCD contribution is obtained in the validation regions, and a flat systematic uncertainty is chosen to cover the disagreements between data and predictions. The figure 6.5 illustrates the different control regions used for electron channel. As the CR_{fake} is obtained reverting the direction of the cuts involving the E_T^{miss} and m_T^W :

- $E_T^{miss} < 20$ GeV and $E_T^{miss} + m_T^W < 60$ GeV,

the intermediary regions between the CR_{fake} and the signal region are:

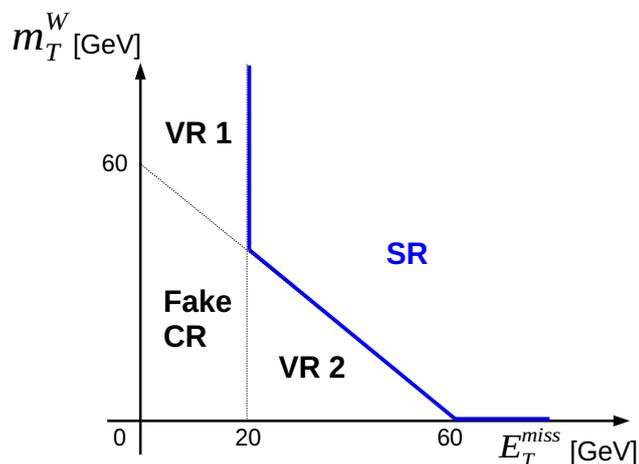


Figure 6.5: Scheme of the phase space parametrisation for the electron channel. A conservative 50% uncertainty covers the disagreements of the QCD measured in both validation regions VR 1 and 2.

²Reverting the cuts with respect to the signal region selection (see section 5.2.2)

- VR 1: $E_T^{miss} < 20$ GeV and $E_T^{miss} + m_T^W > 60$ GeV
- VR 2: $E_T^{miss} > 20$ GeV and $E_T^{miss} + m_T^W < 60$ GeV.

A similar prescription is used to estimate the systematic uncertainty in the muon channel. As the CR_{fake} is constructed by reverting the cut on the d_0 significance ($|d_0/\sigma_{d_0}| > 5$), and the signal region is achieved when $|d_0/\sigma_{d_0}| < 3$, the validation region proposed for the muon channel is:

- VR: $3 < |d_0/\sigma_{d_0}| < 5$

The results show that a flat systematic error band of 50% is enough to cover the disagreements in the electron and muon channels.

6.2.4 Performance of the QCD multi-jet background estimation

Using the real and fake rates obtained, the QCD multi-jet background can be estimated in the selection used for the analysis (6.2.1):

W+jets enriched control region

In the region dominated by W+jets events, the QCD multi-jet background is expected to be accumulated at low m_T^W and low lepton p_T . The figures 6.8 and 6.9 show a reasonable agreement. The miss-modelling that can be attributed to the QCD estimation being within the QCD multi-jets systematic.

$t\bar{t}$ enriched control region

Finally, the QCD multi-jet background is measured in a $t\bar{t}$ dominated region, the signal region (Figures 6.10 and 6.11). It is constructed by requiring at least 4 jets and at least 1 b-tagged track-jet. The real and fake rates derived in section 6.2.3 are used to estimate the QCD multi-jet background, which has a small contribution in this region. Then, the invariant mass of $t\bar{t}$ system can be reconstructed implementing the χ^2 minimisation algorithm but without requiring $\log_{10}(\chi^2) < 0.9$ to keep enough QCD events to test their impact on the shapes.

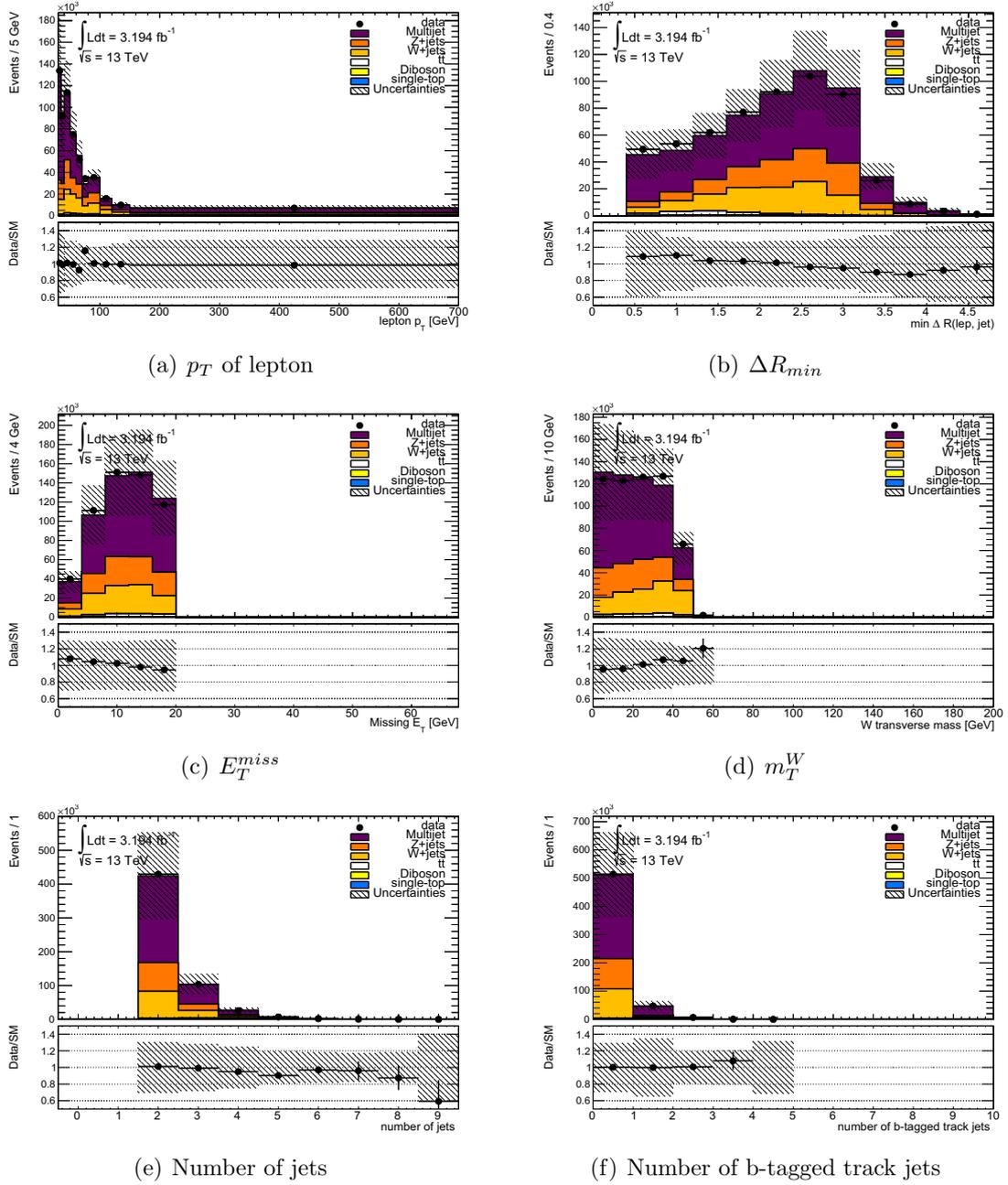


Figure 6.6: Closure test for electron channel.

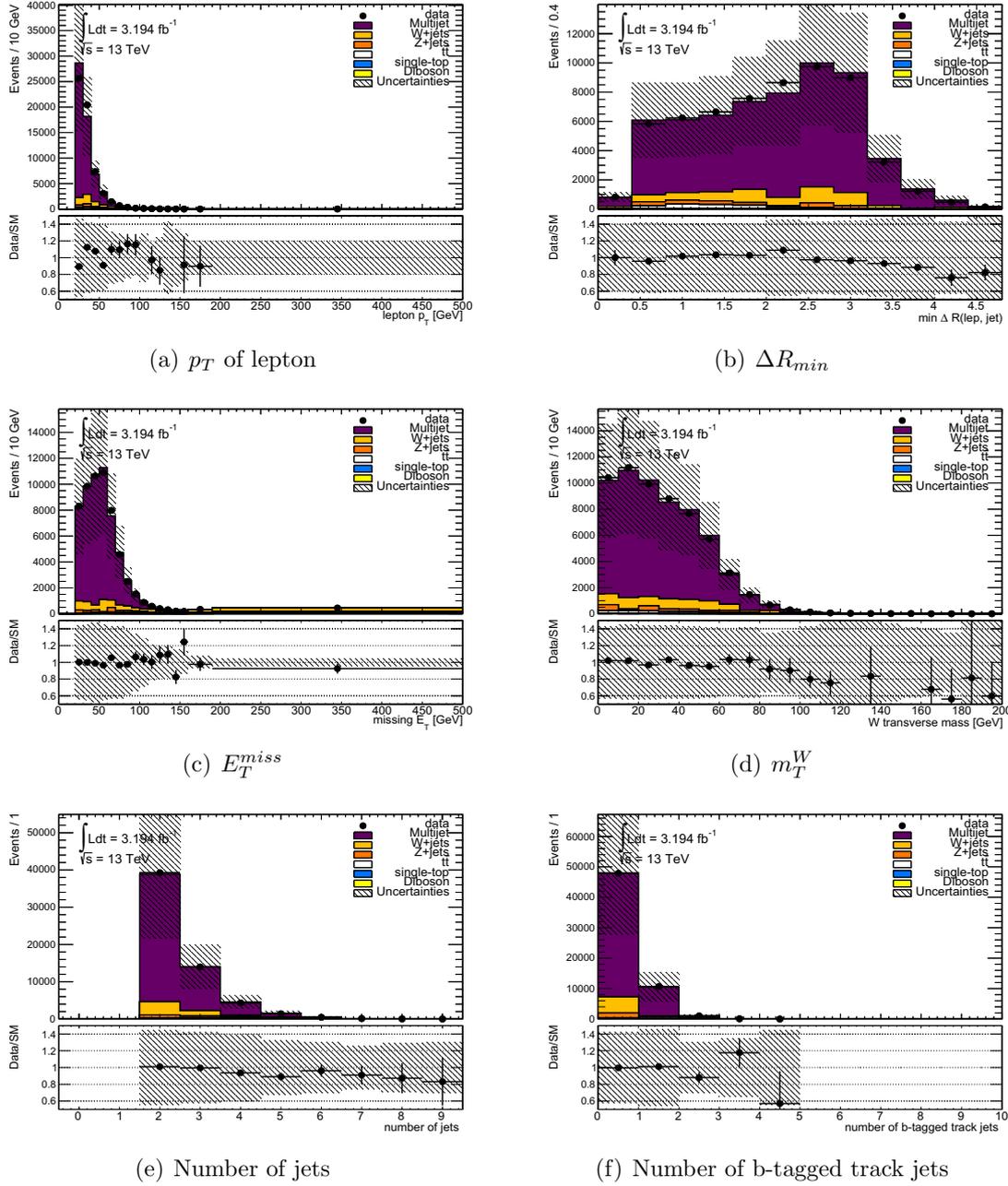


Figure 6.7: Closure test for muon channel.

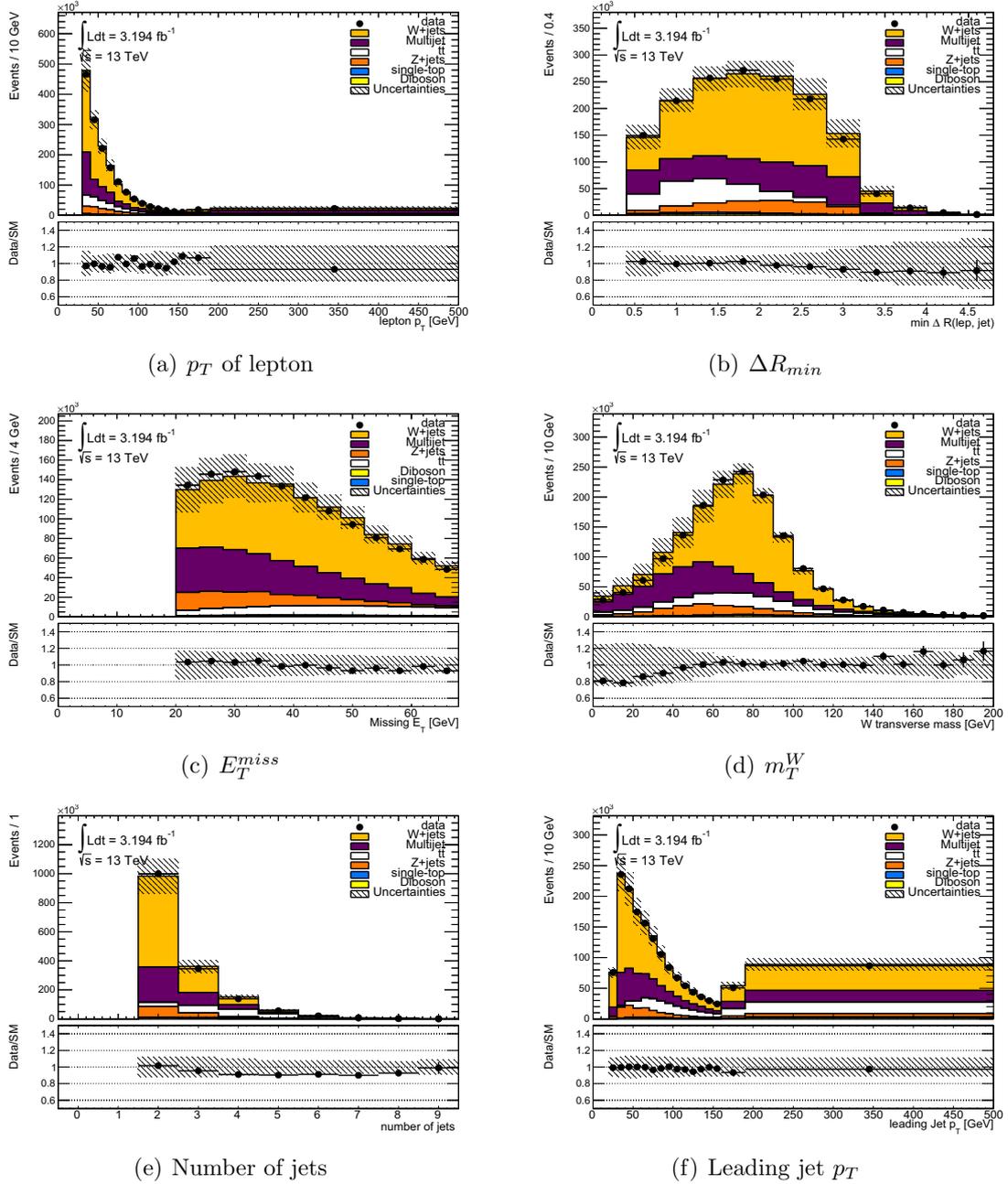


Figure 6.8: Estimation of the QCD multi-jet background in the W+jets enriched CR (electron channel).

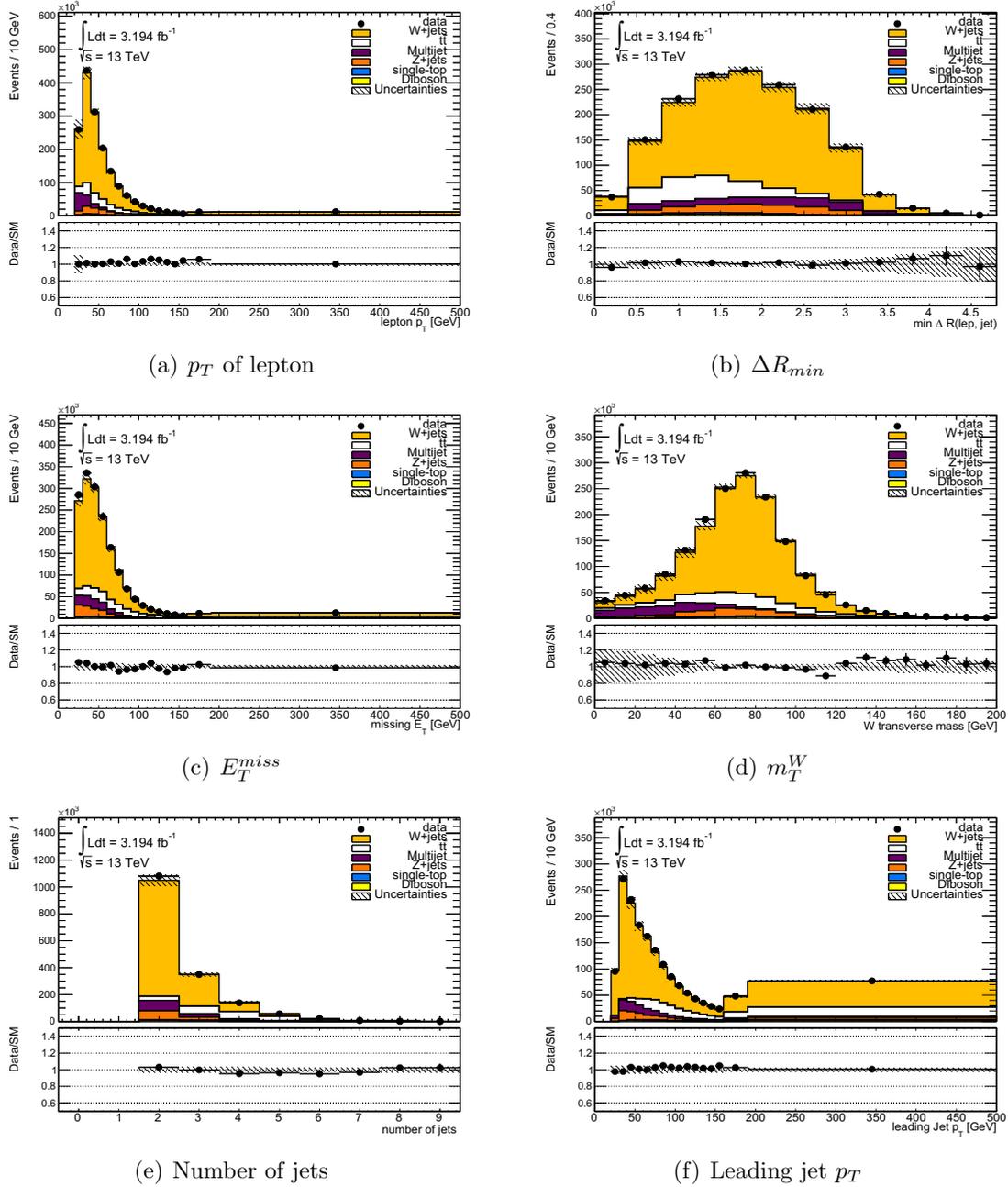


Figure 6.9: Estimation of the QCD multi-jet background in the W+jets enriched CR (muon channel).

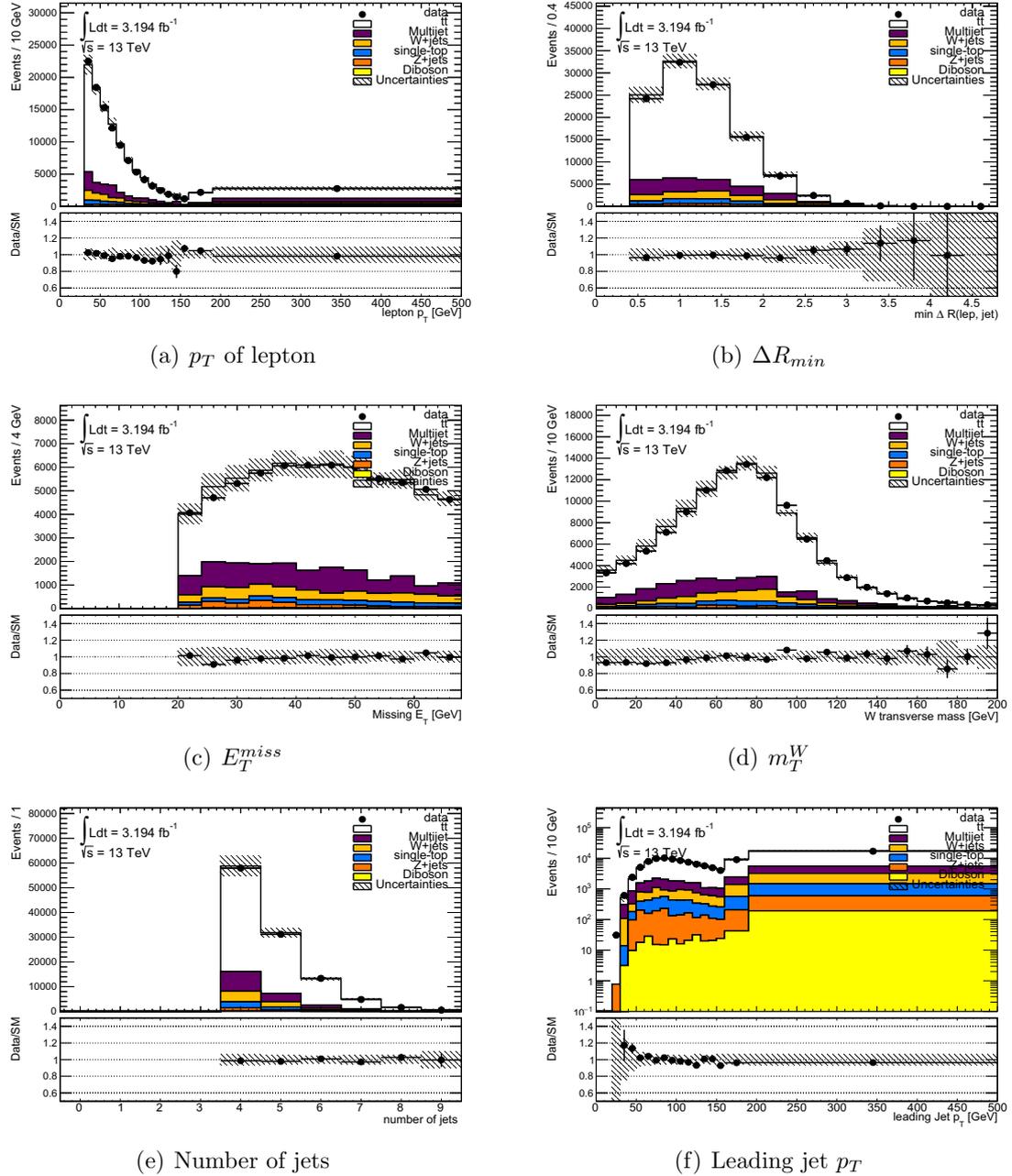


Figure 6.10: Estimation of the QCD multi-jet background in a $t\bar{t}$ enriched CR (electron channel). Events with at least 4 jets and at least 1 b-tagged track jet are demanded.

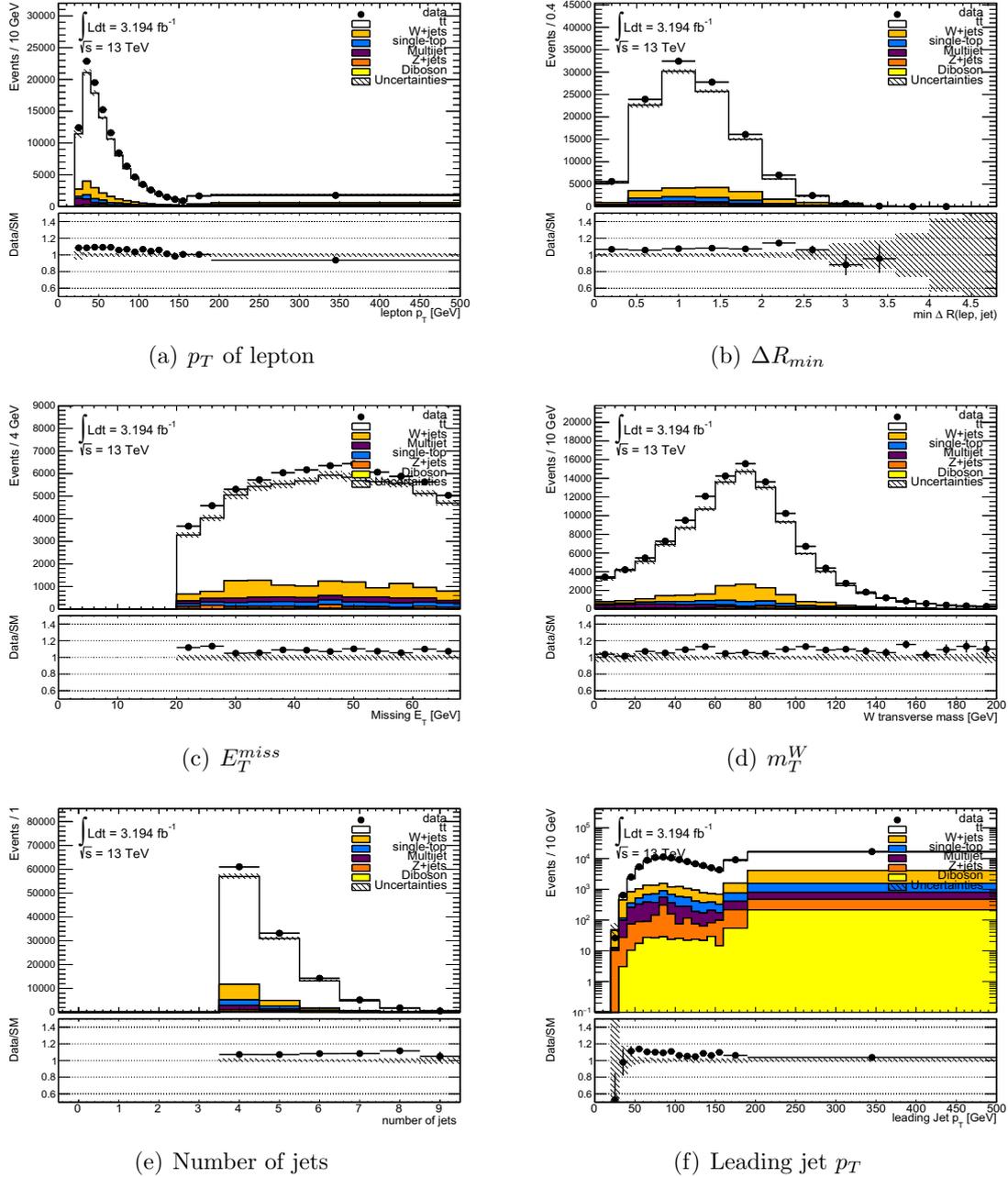


Figure 6.11: Estimation of the QCD multi-jet background in a $t\bar{t}$ enriched CR (muon channel). Events with at least 4 jets and at least 1 b-tagged track jet are demanded.

6.3 Reconstruction of the invariant mass of the $t\bar{t}$ system

The $t\bar{t}$ invariant mass ($m_{t\bar{t}}^{reco}$) is built from the sum of the four-momentum of the reconstructed objects identified as coming from the top-quark decay products, including the calculation of the neutrino four-momenta (section 5.3). A χ^2 algorithm selects the physics objects to be used in the reconstruction of $m_{t\bar{t}}^{reco}$.

6.3.1 χ^2 minimisation algorithm

The χ^2 minimisation algorithm is constructed using constraints from the expected top-quark and W-boson masses. All the possible permutations of the reconstructed physics objects which are available in an event are used to compute a χ^2 value. Then, the configuration with the lowest χ^2 is chosen:

$$\chi^2 = \left[\frac{m_{jj} - m_W}{\sigma_W} \right]^2 + \left[\frac{m_{jjb} - m_{jj} - m_{t_h - W}}{\sigma_{t_H - W}} \right]^2 + \left[\frac{m_{jl\nu} - m_{t_l}}{\sigma_{t_l}} \right]^2 + \left[\frac{(p_{T,jjb} - p_{T,jl\nu}) - (p_{T,t_h} - p_{T,t_l})}{\sigma_{p_T \text{ diff}}} \right]^2 \quad (6.1)$$

The first constraint comes from the hadronically decaying W-boson (W_h), where the invariant mass of two candidate jets (m_{jj}) is compared to the expected mass of the reconstructed W-boson distribution. The expected mass is modelled by a Gaussian function (m_W, σ_W), where the parameters are extracted from the invariant mass distribution of jets pairs matched to the decay products of the W-boson. Similarly, the second term corresponds to the hadronic-top t_h . Since m_{jj} and m_{jjb} are strongly correlated, the W_h is subtracted to consider an uncorrelated constrains term. The third constrain is associated to the leptonic top-quark decay, where in addition to testing the jet combinations, both solutions to the neutrino four-momentum are tried. The last term constrains the difference between the two top quarks transverse momentum, which helps for the discrimination between the jets coming from the top-quarks and the ISR/FSR jets.

The values of the parameters are determined from MC events using the “reconstructible events”. They are defined in MC simulations as events for which it is possible to match the semi-leptonic $t\bar{t}$ decays products to the reconstructed physics objects. The lepton at the partonic-level is matched to the selected reconstructed lepton using a cone of radius $\Delta R = 0.2$. The quarks from the top-quark decay are required to be matched to truth jets and then the truth jets are required to be matched to reconstructed jets, using in both steps a cone with $\Delta R=0.4$. The transverse component of the neutrino’s momentum and the E_T^{miss} have to point in the same direction: $\Delta\phi(\nu, E_T^{miss}) < 1$ radian. The matching efficiency between the different partons and the reconstructed objects in the simulation after the full resolved selection is about 50 % at low mass, and it reaches to 70% in the high mass region [40]. All the Z' samples of masses from 0.5 to 2 TeV are used in the computation

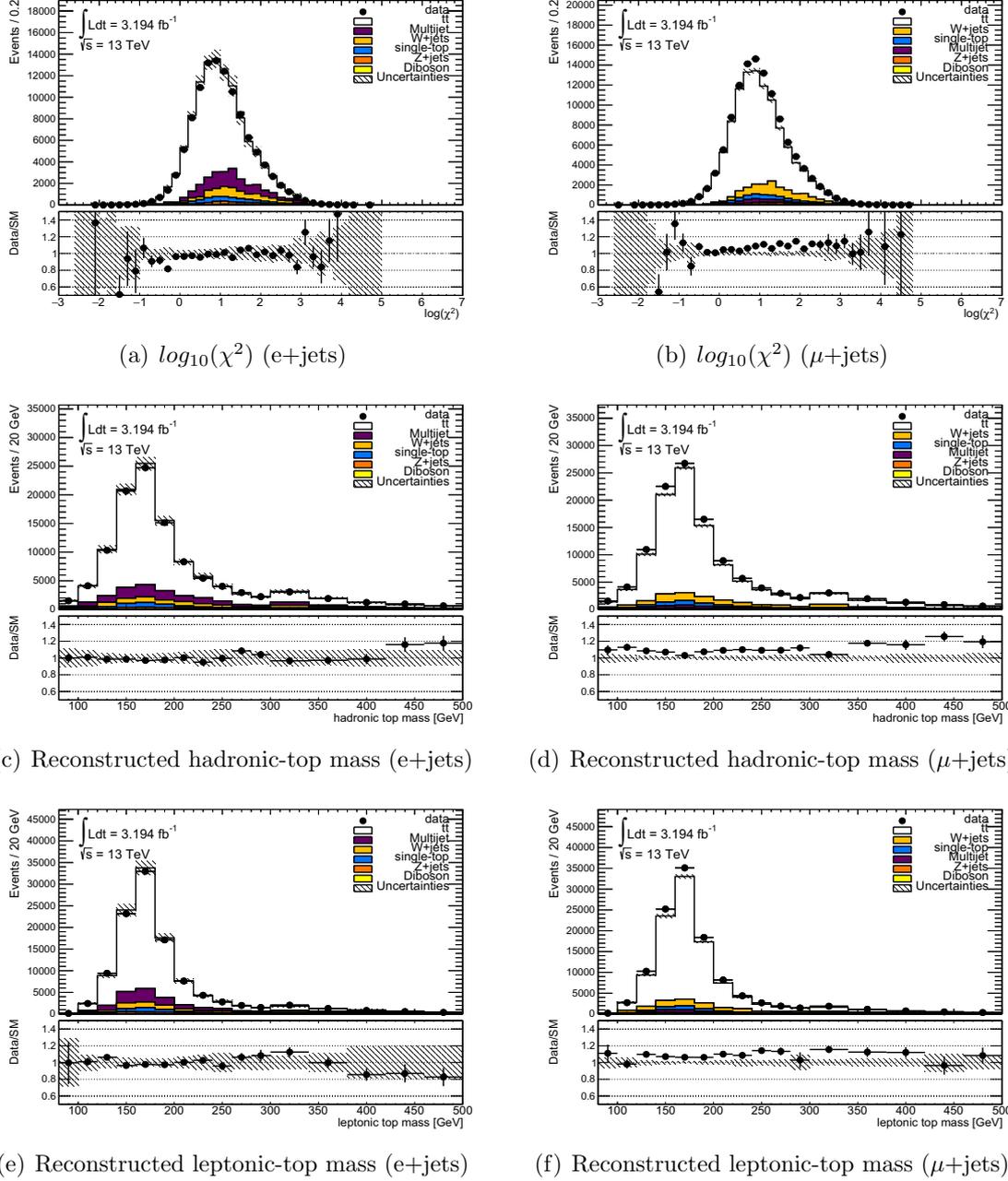


Figure 6.12: $t\bar{t}$ enriched CR: output of the χ^2 algorithm and reconstructed leptonic- and hadronic-top masses using the χ^2 algorithm (no- χ^2 cut is implemented).

of the χ^2 parameters to optimize the search in the “low” mass range. Then, a fit with a Gaussian function is used to extract the values of these parameters.

The inputs for the χ^2 parametrisation has been derived using the Z' signal samples (see section 5.4) are: $m_W = 80.51$ GeV, $\sigma_W = 12.07$ GeV, $m_{t_h-W} = 85.17$ GeV, $\sigma_{t_h-W} = 16.05$ GeV, $m_{t_l} = 167.36$ GeV, $\sigma_{t_l} = 25.41$ GeV, $p_{T,t_h} - p_{T,t_l} = 0.23$ GeV and $\sigma_{p_T \text{ diff}} = 18.85$ GeV. The reconstructed hadronic- and leptonic-top mass are reconstructed using the selection from the χ^2 (Figures 6.12). The $m_{t\bar{t}}$ distribution is not showed since it is still blinded for the time scale of this manuscript.

6.4 Improving the $t\bar{t}$ invariant mass resolution

The resolution on $m_{t\bar{t}}$ is expected to have a large impact on the sensitivity of this analysis. Two studies have been performed: firstly to assess the impact of the neutrino reconstruction on the $m_{t\bar{t}}$ resolution, secondly to develop a method which aims to improve the resolution.

These studies will use the average detector resolution ($\Delta m_{t\bar{t}}$) that can be computed using the shape of the reconstructed $t\bar{t}$ invariant mass $m_{t\bar{t}}^{reco}$ after the subtraction of the shape from the $t\bar{t}$ invariant mass at truth level $m_{t\bar{t}}^{truth}$:

$$\Delta m_{t\bar{t}} = \frac{m_{t\bar{t}}^{reco,MA} - m_{t\bar{t}}^{truth}}{m_{t\bar{t}}^{truth}} \quad (6.2)$$

where $m_{t\bar{t}}^{truth}$ is obtained after FSR (see section 5.4) and $m_{t\bar{t}}^{reco,MA}$ is computed using the reconstructible events for which $m_{t\bar{t}}^{reco}$ is computed from reconstructed objects matched to the partons from the top-quark decay. Then, a fit with a gaussian function is used to extract the detector resolution. The configuration of the Run-1 analysis is used for these studies [40]. The signal simulation samples for a Z' and scalar resonance (2HDM³) were available for Run 1 setup at different mass points.

6.4.1 Impact of the neutrino reconstruction

As mentioned in section 5.3, the neutrino reconstruction procedure is based on the mass constraint of the leptonic decaying W-boson. The E_T^{miss} module and direction are used to infer the longitudinal component of the neutrino's momentum, but a large fluctuation of the E_T^{miss} measurement can lead to non-real solutions for the quadratic equation of the longitudinal neutrino's momentum component. An estimation on the impact of this effect on the resolution of the $t\bar{t}$ invariant mass is exposed.

The study is based on the comparison of the detector resolution obtained in two scenarios:

³Without taking into account interference effects with the SM backgrounds.

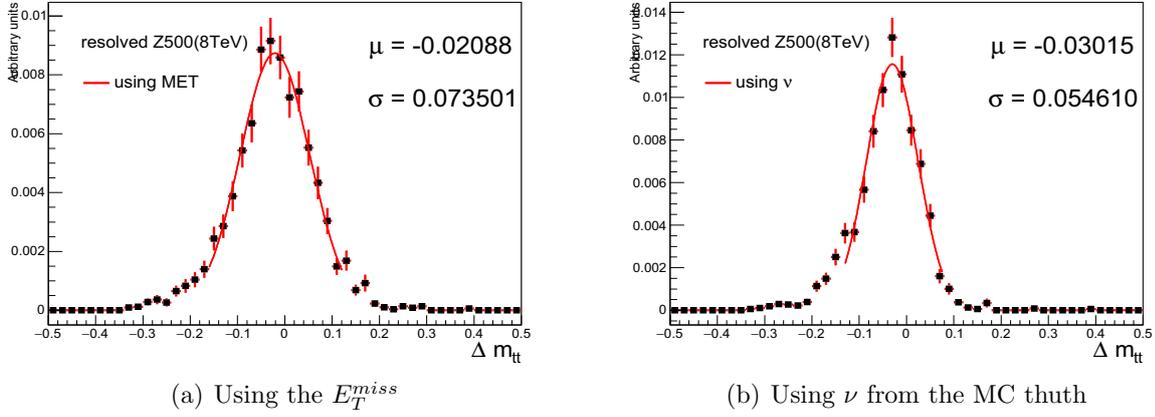


Figure 6.13: Fits for the extraction of the $t\bar{t}$ invariant mass resolution using a Z' sample of 500 GeV. The mass resolution gets improved when using the true neutrino four-momenta to compute $m_{t\bar{t}}^{reco,MA}$ after FSR.

- When $m_{t\bar{t}}^{reco,MA}$ is obtained using the neutrino reconstruction procedure described in section 5.3, deriving the longitudinal momentum component which depends on the E_T^{miss} (Figure 6.13 (a)).
- When $m_{t\bar{t}}^{reco,MA}$ is obtained using the truth neutrino four-momentum from the simulation samples (Figure 6.13 (b)).

The figures 6.14 show the evolution of the detector resolution as function of the Z' mass. In particular, the figures 6.14 (c, d) shows the degradation of the resolution due to the E_T^{miss} :

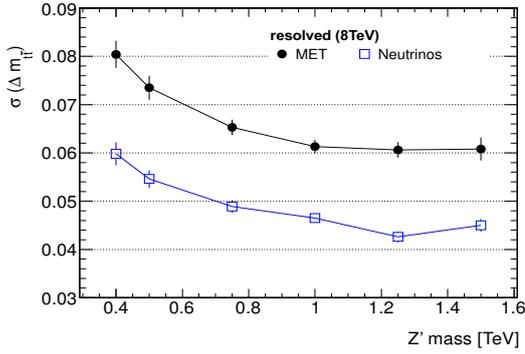
$$\Delta\sigma = \frac{\sigma(E_T^{miss}) - \sigma(\nu_{truth})}{\sigma(\nu_{truth})} \quad (6.3)$$

where $\sigma(E_T^{miss})$ and $\sigma(\nu_{truth})$ are obtained from the spread of the detector resolution for the corresponding case. It can be deduced the missing information due to the escaping neutrino is responsible of a degradation of the detector resolution of $\sim 35\%$ for the resolved regime and of $\sim 20\%$ in the boosted regime.

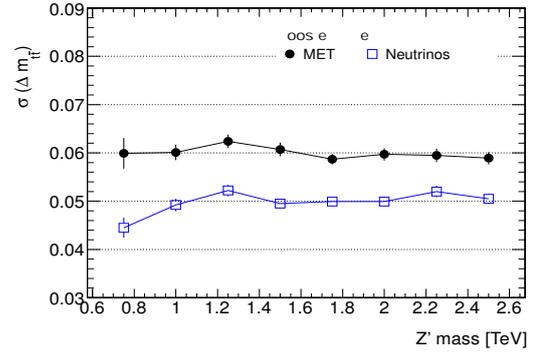
6.4.2 Jet rescaling method

Procedure

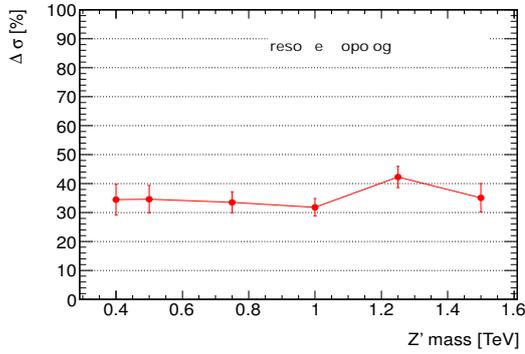
As mentioned, the detector resolution effects degrade the sensitivity for a possible signal detection, but the known masses of the top-quark and W-boson can be used to improve the resolution on



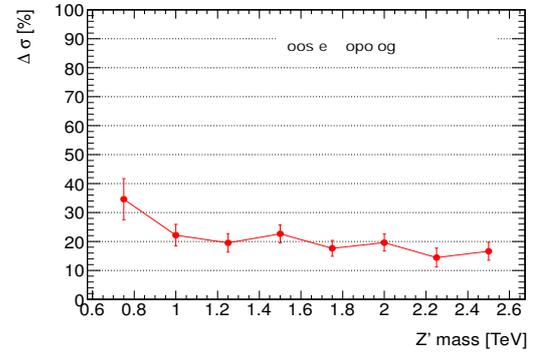
(a) $m_{t\bar{t}}$ resolution for resolved channel.



(b) $m_{t\bar{t}}$ resolution for boosted channel.



(c) Impact of neutrino in the detector resolution (resolved).



(d) Impact of neutrino in the detector resolution (boosted).

Figure 6.14: $m_{t\bar{t}}^{reco}$ resolution to check the impact due to the neutrino reconstruction procedure. The blue points represent the detector resolution when $m_{t\bar{t}}^{reco,MA}$ is computed using the truth neutrino information instead of inferring neutrino momentum from E_T^{miss} (black points).

the $t\bar{t}$ invariant mass. Constraints are implemented in several stages of the mass reconstruction, producing constants which “calibrate” the jets in the rescaling procedure.

This study uses the samples at $\sqrt{s} = 8$ TeV for a scalar resonance signal, which are generated for different mass points. In addition, the reconstructible events for which the reconstructed physics objects (electron, muon, jets, ...) are matched to the truth information at the partonic-level are used.

Hadronic W mass scaling

The first step consists in reconstructing the invariant mass of the hadronic W-boson (W_h) using the four-momentum of the jets \mathbf{P}_1 and \mathbf{P}_2 matched to the W’s decays ($|\mathbf{P}_1 + \mathbf{P}_2|^2 = (m_{W_h}^{noRes})^2$). The two jet four-momenta are rescaled by the factor α obtained from the known mass of the W-boson ($M_W = 80.4$ GeV) such as:

$$M_W^2 = |\alpha\mathbf{P}_1 + \alpha\mathbf{P}_2|^2 \quad (6.4)$$

Figure 6.15 (a) shows the impact of the jet rescaling on the W_h mass. As expected, the distribution is transformed from a broad peak to a very narrow one. Then, the mass of the hadronic top-quark (t_h) and the $t\bar{t}$ system are reconstructed to observe the impact of the jet rescaling, which improve the resolution on both (Figure 6.15 (b) and (c)).

Top-quark mass scaling

The second step is the rescaling of the b-jets from the top-quark mass constrain. The value of the top-quark mass used for the constrain is $M_{top} = 172.5$ GeV. One needs to find the rescaling factor β that, applied to the four-momentum of the jet associated to the b-quark in the hadronic top decay (\mathbf{P}_b), would gives:

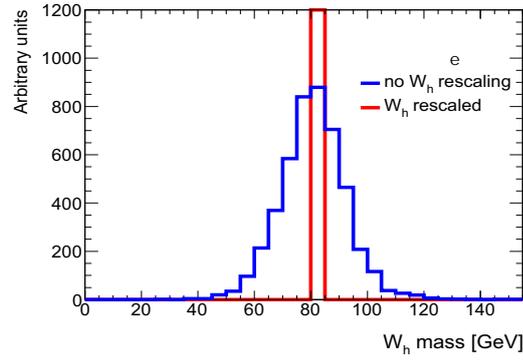
$$M_{top}^2 = |\mathbf{P}_w + \beta\mathbf{P}_b|^2 \quad (6.5)$$

To obtain the scaling factor, solutions of the quadratic equation ($a\beta^2 + b\beta + c$) needs to be found:

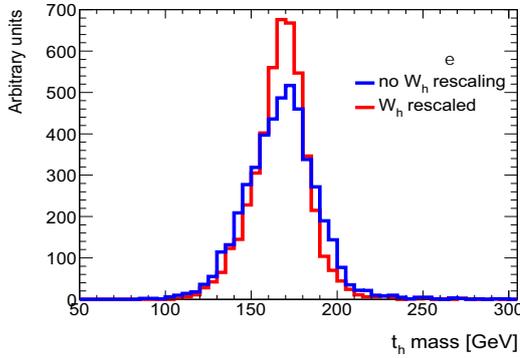
$$a = m_b^2 \quad (6.6)$$

$$b = m_{jjb}^2 - (m_{jj}^{res})^2 - m_b^2 \quad (6.7)$$

$$c = m_{jj}^2 - M_{top}^2 \quad (6.8)$$



(a) Rescaling the jets from hadronic W-boson mass.



(b) Impact on hadronic top-quark mass.

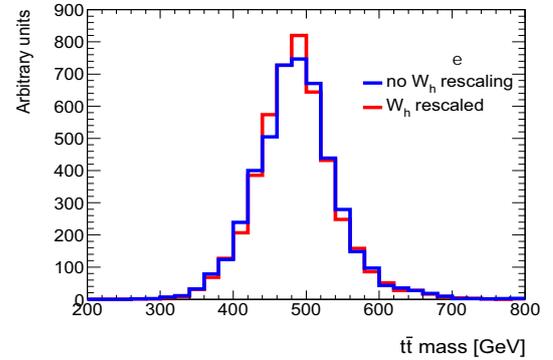
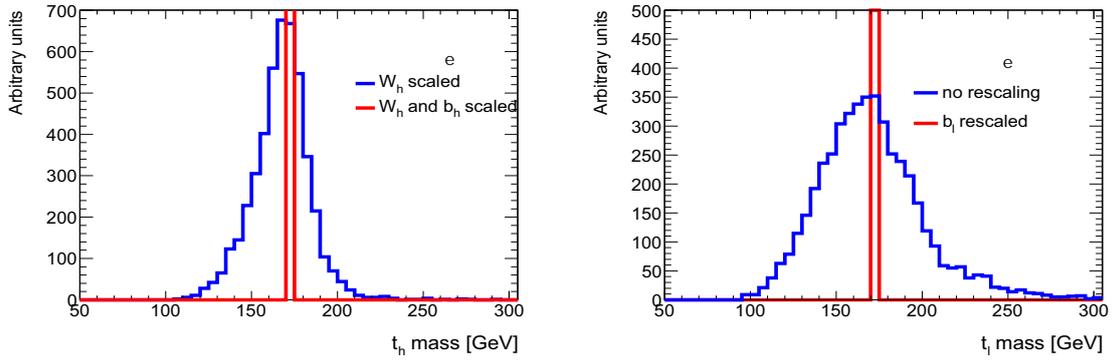
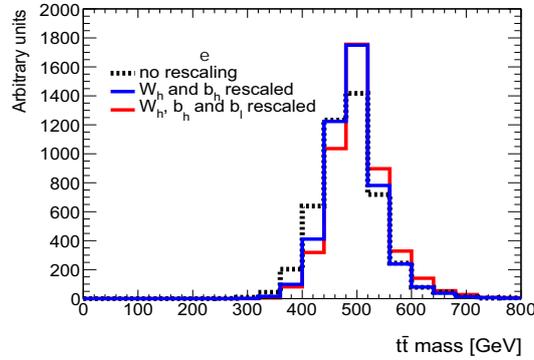

 (c) Impact on the invariant mass of the $t\bar{t}$ system.

Figure 6.15: Invariant mass distribution of the hadronic W-boson before and after the rescaling. Propagation of the jet rescaling to the top-quark mass and invariant mass of the $t\bar{t}$ system. Scalar benchmark simulation at $\sqrt{s} = 8$ TeV with a mass of 500 GeV, using only matchable events.



(a) Rescaling the jets from hadronic-top mass.

(b) Rescaling the jets from leptonic-top mass.



(c) Impact on the invariant mass of the $t\bar{t}$ system.

Figure 6.16: Rescaling of the top-quarks and their impact to the reconstructed $t\bar{t}$ mass spectra before and after the different rescaling. Scalar benchmark simulation at $\sqrt{s} = 8$ TeV with a mass of 500 GeV, using only matchable events.

where m_{jj}^{res} is used after the W_h scaling, m_b is the invariant mass of the jet associated to the b-quark, and m_{jjb} is the invariant mass of the hadronic top-quark without scaling. Then, the jet associated to the b-quark is rescaled by a factor β which is computed for each event. Figure 6.16 (a) shows an illustration of the b_h rescaling. Finally, the rescaling of the b-jet from the leptonic top-quark (b_l) is obtained in the same way (Figure 6.16 (b)).

6.4.3 Impact of the rescaling

The impact of the jet rescaling can be measured by computing the detector resolution as mentioned in eq. 6.2. It is extracted using a fit with a Gaussian function, for the available scalar samples at different mass points. Figure 6.17 (a) shows the detector resolution at each step of the rescaling procedure. The rescaling of the jets from the hadronic-top improves the detector resolution, but the rescaling of b_l degrades the resolution at high mass. Indeed in the boosted regime the neutrino is expected to be close to the lepton. Then, the resolution on the E_T^{miss} direction has a large impact on the W_l momentum reconstruction (through the neutrino p_z reconstruction), which is a key point for the rescaling of b_l . In figure 6.17 (b), the improvement of the resolution with respect to the no-rescaling case is quantified. The detector resolution is improved by at least $\sim 20\%$ from the rescaling of the jets from the hadronic-top.

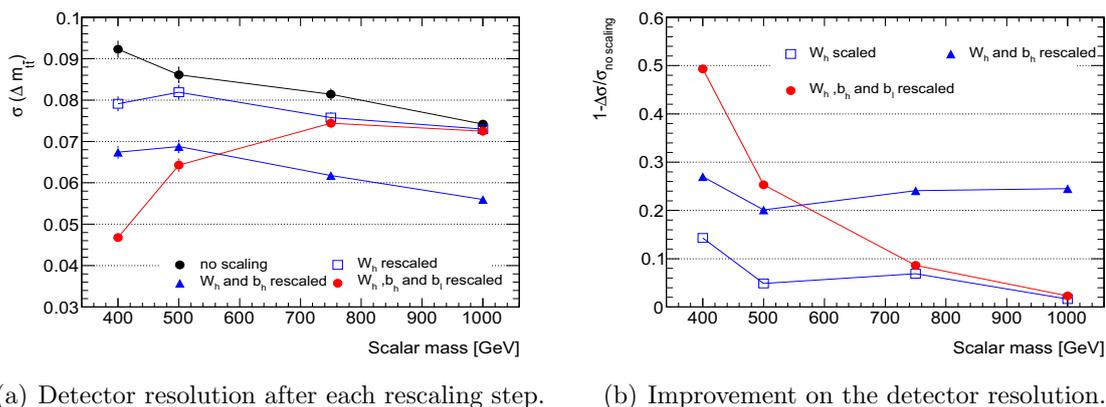


Figure 6.17: Impact of the jet rescaling method in the resolution of $t\bar{t}$ invariant mass for different scalar mass point. Improvement of the resolution for low signal masses, but degradation start to appears in the high signal mass region.

Outlook

A status of the resolved analysis has been presented in this chapter. The QCD multi-jet is estimated for the region dominated by W +jets events and the region dominated by $t\bar{t}$ events. The systematic uncertainties has not been included yet, but based in the historical systematic uncertainty of approximately 20 % for the resolved channel, the disagreement in the $t\bar{t}$ enriched control region will be covered by this uncertainty band. It has been also proved that this new parametrisation provides a better shape description of the QCD multi-jet background in the boosted channel.

For future iterations of the resolved channel, the jet rescaling method can be implemented to improve the sensitivity at low mass region. The detector resolution is improved by almost $\sim 20\%$ after the rescaling of the jets from the hadronic-top, for all the invariant $t\bar{t}$ mass range. In addition, the rescaling of all the jets improves significantly the very low mass region, close to the “two times the top mass” threshold (~ 345 GeV). It could be also interesting to check in future iterations of the resolved analysis, the impact of the full rescaling on the limit setting for this very low mass region.

Conclusion

Two main subjects have been developed in the current thesis manuscript, which expose the study of the possible use of laser in-time calibration in the future and the analysis to search for $t\bar{t}$ resonances using the datasets recorded in the 2015 collisions of the LHC.

The laser system is currently used in dedicated calibration runs to derive the laser calibration constants. But the laser in-time runs can lead to a more frequent calibration during the physics runs. Using laser in-time runs with at least 5000 events is enough since the systematic uncertainties are still dominant over the statistical one. Therefore, the analysis developed does not find any reasons to not calibrate the channels of TileCal using these laser in-time runs. In addition, there is an upgrade on the light splitting box for the second version of the laser system (implemented for Run 2), which improves the homogeneity and stability of the light after the splitting. This could change the calibration method in the future since the stability assumption in the most distant cells for the fiber correction could not longer be valid (due to the increase of the energy of the proton beams).

The second part of the manuscript is about for the search for $t\bar{t}$ resonances using 2015 datasets. As expected, the production cross-section limits set on the Z'_{TC2} boson of width 1.2% has been improved, since its existence is excluded at 95% CL in a range $0.75 \text{ TeV} < m_{Z'_{TC2}} < 2.0 \text{ TeV}$. In previous production cross-section limits set in Run 1, the Z'_{TC2} of width 1.2% is excluded for masses lower than 1.8 TeV (see Chapter 1). A combination with the resolved analysis could improve the sensibility in the low mass region, but is not expected to be very significant considering only 3.2 fb^{-1} of collision data.

The matrix method has been used to estimate the contribution of the QCD multi-jet background specifically for this analysis. The contribution of this background is small in the boosted analysis (see Chapter 5), where the fake rates are having large statistical uncertainties. Instead, the contribution of the QCD multi-jet background is larger in the resolved analysis and the statistical uncertainties of the fake rates are small enough to extract the proper shapes of the main observables (see Chapter 6). The estimation performed looks very promising in the electron channel but a disagreement persists in the muon channel. This issue is currently under investigation.

The luminosity peak has been increased for the 2016 data-taking, so the resolved analysis could be considered to improve the previous limits using the 2016 datasets with approximately 30 fb^{-1} of collision data at $\sqrt{s} = 13 \text{ TeV}$, but such analysis is still on development. In addition, the current production cross section limits obtained for other historical benchmark models will be updated for the coming iterations of this analysis.

As showed in Chapter 6, the resolution of the invariant mass of $t\bar{t}$ system can be affected by the reconstruction techniques for example. This could lead to a smearing of the BSM signal peak which induces to loose sensitivity for its detection. The jet rescaling method looks promising to improve the resolution on the invariant mass in the low mass region, while the uncertainty on the neutrino direction introduces a degradation on the resolution for high masses.

Bibliography

- [1] S. Weinberg. **A Model of Leptons**. Phys. Rev. Lett. 19(1967) 1264-1266
- [2] S. L. Glashow. **Partial Symmetries of Weak Interactions**. Nucl. Phys. 22(1961) 579-588
- [3] J. Goldstone, A. Salam, S. Weinberg. **Broken Symmetries**. Phys. Rev. Lett. 127(1962) 965-970
- [4] G Altarelli. **The Standard model of particle physics**. To appear in the “Encyclopedia of Mathematical Physics”, Elsevier. arXiv:0510281v1 [hep-ph].
- [5] M. Baak, M. Goebel, J. Haller, A. Hoecker, D. Ludwig, et al. **Updated Status of the Global Electroweak Fit and Constraints on New Physics**. Eur.Phys.J. C72 (2012) 2003. arXiv:1107.0975.
- [6] S.M. Bilenky, C. Giunti, W. Grimus. **Phenomenology of Neutrino Oscillations**. Prog.Part.Nucl.Phys.43:1-86,1999 arXiv:9812360 [hep-ph].
- [7] R. Barate et al. **Search for the standard model Higgs boson at LEP**. Phys. Rev. Lett., B565:61 75, 2003.
- [8] The ATLAS Collaboration. **Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC**. Phys. Rev. Lett. B716 (2012) 1-29 arXiv:1207.7214 [hep-ex].
- [9] The CMS Collaboration. **Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC**. Phys. Lett. B 716 (2012) 30 arXiv:1207.7235 [hep-ex].
- [10] The ATLAS and CMS Collaborations. **Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments**. Phys. Rev. Lett. 114, 191803 (2015) arXiv:1503.07589 [hep-ex].
- [11] The ATLAS Collaboration. **Evidence for the spin-0 nature of the Higgs boson using ATLAS data**. Phys. Lett. B 726 (2013) 120, arXiv:1307.1432 [hep-ex].
- [12] D.H. Perkins. **Introduction To High Energy Physics**. Cambridge University Press, 2000.

BIBLIOGRAPHY

- [13] F. Halzen and A.D. Martin. **Quarks and leptons: an introductory course in modern particle physics.** Wiley, 1984.
- [14] E. Eichten and K. Lane. **Dynamical breaking of weak interaction symmetries.** Physics Letters B, 90(1-2):125-130, 1980.
- [15] F. Englert, R. Brout (1964). **Broken Symmetry and the Mass of Gauge Vector Mesons.** Phys. Rev. Lett. 13 (9): 321-323.
- [16] P.W. Higgs (1964). **Broken Symmetries and the Masses of Gauge Bosons.** Phys. Rev. Lett. 13 (16): 508-509.
- [17] G. Dissertori, I.G. Knowles, and M. Schmelling. **Quantum chromodynamics: high energy experiments and theory.** International series of monographs on physics. Clarendon Press, Oxford, 2003.
- [18] G. P. Salam. **Elements of QCD for hadron colliders.** CERN Yellow Report CERN-2010-002, pp. 45-100 arXiv:1011.5131v2 [hep-ph].
- [19] M. Kobayashi and T. Maskawa. **CP Violation in the Renormalizable Theory of Weak Interaction.** Prog. Theor. Phys. 49 (1973) 652.
- [20] The UA1 Collaboration, G. Arnison et al. **Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540$ GeV.** Phys. Lett. B 122 (1983) 103.
- [21] The CDF Collaboration. **Observation of top quark production in $\bar{p}p$ collisions.** Phys. Rev. Lett. 74 (1995) 2626.
- [22] The DØ Collaboration. **Observation of the top quark.** Phys. Rev. Lett. 74 (1995) 2632.
- [23] M. Baak and R. Kogler. **The global electroweak Standard Model fit after the Higgs discovery.** Rencontres de Moriond, La Thuile, Valle d'Aosta, Italy, 2-16 March, 2013 arXiv:1306.0571 [hep-ph].
- [24] The ATLAS Collaboration. **SM summary plots for ICHEP 2016.** ATL-COM-PHYS-2016-1026
- [25] G. F. Giudice. **Beyond the standard model.** Lectures delivered at the 1995 European School of High-Energy Physics in Dubna, Russia. arXiv:9605390v1 [hep-ph].
- [26] G. Bertone, D. Hooper, and J. Silk. **Particle dark matter: Evidence, candidates and constraints.** Phys. Rept., 405:279-390, 2005.
- [27] The ATLAS Collaboration.
<https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/EXOTICS/>

BIBLIOGRAPHY

- [28] C. T. Hill. **Topcolor: top quark condensation in a gauge extension of the standard model.** Phys. Lett, B266(3-4):419-424, 1991.
- [29] C. T. Hill. **Topcolor assisted technicolor.** Phys. Lett., B345:483-489, 1995.
- [30] R. M. Harris, C. T. Hill and S. J. Parke. **Cross-section for topcolor Z-prime(t) decaying to t anti-t: Version 2.6 (1999).** Fermilab-FN-687 arXiv:9911288 [hep-ph].
- [31] R. Foadi, M. Frandsen, F. Sannino. **125 GeV Higgs from a not so light Technicolor Scalar.** CP3-Origins-2012-028. arXiv:1211.1083 [hep-ph].
- [32] L. Randall and R. Sundrum. **A Large mass hierarchy from a small extra dimension.** Phys. Rev. Lett. 83:3370-3373, 1999. arXiv:9905221 [hep-ph].
- [33] L. Randall and R. Sundrum. **An Alternative to Compactification.** Phys. Rev. Lett. 83:4690-4693, 1999. arXiv:9906064 [hep-th].
- [34] M. Gremm. **Four-dimensional gravity on a thick domain wall.** Phys.Lett. B478 (2000) 434-438 arXiv:9912060 [hep-th].
- [35] R. Guerrero, A. Melfo, N. Pantoja, and O. Rodriguez. **Close to the edge: Hierarchy in a double braneworld (2006).** Phys. Rev. Lett. D74:084025. arXiv:0601161 [hep-th].
- [36] B. Lillie, L. Randall, and L. T. Wang. **The Bulk RS KK-gluon at the LHC.** JHEP, 0709:074, 2007.
- [37] G. C. Branco, P. M. Ferreira, L. Lavoura, M. N. Rebelo, Marc Sher, and Joao P. Silva. **Theory and phenomenology of two-Higgs-doublet models.** 10.1016/j.physrep.2012.02.002 arXiv:1106.0034 [hep-ph].
- [38] The ATLAS Collaboration.
<https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/TOP/>
- [39] K.A. Olive et al. **Particle Data Group.** Chin. Phys. C, 38, 090001 (2014) and 2015 update.
- [40] The ATLAS Collaboration. **A search for $t\bar{t}$ resonances using lepton-plus-jets events in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector.** arXiv: 1505.07018 [hep-ex].
- [41] The CDF collaboration. **Search for $t\bar{t}$ resonances in the lepton plus jets final state in p-p collisions at $\sqrt{s} = 1.96$ TeV.** Phys. Lett. B668 (2008) 98-104, arXiv:1108.4755 [hep-ex].
- [42] The CMS Collaboration. **Search for $t\bar{t}$ resonances in boosted semileptonic final states in pp collisions at $\sqrt{s} = 8$ TeV.** CMS PAS B2G-15-002
- [43] The DØ Collaboration. **Search for new physics in top events with the DØ detector.** Submitted for the SUSY07 conference proceeding arXiv:0709.2625 [hep-ex].

BIBLIOGRAPHY

- [44] http://www.nobelprize.org/nobel_prizes/physics/laureates/2013/press.htm
- [45] L. Evans and P. Bryant. **LHC Machine**. JINST, 3 :S08001, 2008.
- [46] The ALICE Collaboration. **The ALICE Experiment at the CERN LHC**. JINST, 3(S08002), 2008.
- [47] The ATLAS Collaboration. **The ATLAS Experiment at the CERN Large Hadron Collider**. JINST, 3(S08003), 2008.
- [48] The CMS Collaboration. **The CMS Experiment at the CERN LHC**. JINST, 3(S08004), 2008.
- [49] The LHCb Collaboration. **The LHCb Experiment at the LHC**. JINST, 3(S08005), 2008.
- [50] R. Alemany and al. **LHC Modes**. LHC-OP-ES-0005.
- [51] R. Schmidt and al. **Protection of the cern large hadron collider**. New. J. Phys., 8(290), 2006.
- [52] The ATLAS Collaboration.
<https://twiki.cern.ch/twiki/bin/view/AtlasPublic/ApprovedPlotsATLASDetector>
- [53] The ATLAS Collaboration.
<https://atlas.web.cern.ch/Atlas/GROUPS/DATAPREPARATION/PublicPlots/D>
- [54] The ATLAS Collaboration. **Performance of the ATLAS Detector in Run 2**.
<https://cds.cern.ch/record/2194314/files/ATL-COM-PHYS-2016-811.pdf>
- [55] The ATLAS Collaboration. **ATLAS magnet system : Technical Design Report**. ATLAS-TDR-6.
- [56] The ATLAS Collaboration. **ATLAS Inner Detector Technical Design Report I**. CERN/LHCC/97-16.
- [57] The ATLAS Collaboration. **Alignment of the ATLAS Inner Detector and its Performance in 2012**. ATLAS-CONF-2014-047.
- [58] The ATLAS Collaboration. **ATLAS Insertable B-Layer Technical Design Report**. ATLAS-TDR-19-ADD-1.
- [59] The ATLAS Collaboration. **ATLAS pixel detector : Technical Design Report**. ATLAS-TDR-11.
- [60] The ATLAS Collaboration. **Particle Identification Performance of the ATLAS Transition Radiation Tracker**. ATLAS-CONF-2011-128.
- [61] The ATLAS Collaboration. **ATLAS Calorimeter Performance Technical Design Report**. CERN/LHCC/96-40.

BIBLIOGRAPHY

- [62] The ATLAS Collaboration. **ATLAS Liquid Argon Calorimeter Technical Design Report**. CERN/LHCC/96-41.
- [63] The ATLAS Collaboration. **Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data**. 2014. Eur. Phys. J. C (2014) 74:3071 arXiv:1407.5063 [hep-ex].
- [64] The ATLAS Collaboration. **ATLAS Tile Calorimeter Technical Design Report**. CERN/LHCC/96-42.
- [65] G. Paara. **Integrator based readout in Tile Calorimeter of the ATLAS experiment**. ATL- TILECAL-PROC-2011-010.
- [66] A. Artamonov, D. Bailey, G. Belanger, M. Cadabeschi, T. Chen, and al. **The ATLAS forward calorimeters** 13th ICATPP Conference on Astroparticle, Particle, Space Physics and Detectors for Physics Applications, Como, Italy, 3 - 7 Oct 2011.
- [67] The ATLAS Collaboration. **ATLAS Muon Spectrometer Technical Design Report**. CERN/LHCC/97-22.
- [68] The ATLAS Collaboration. **ATLAS TDAQ system: Phase-1 Upgrade Technical Design Report**. ATLAS-TDR-023.
- [69] The ATLAS Collaboration. **2015 start-up trigger menu and initial performance assessment of the ATLAS trigger using Run-2 data**. ATL-DAQ-PUB-2016-001.
- [70] R. Brun and F. Rademakers. **ROOT - An Object Oriented Data Analysis Framework**. Proceedings AIHENP'96 Workshop, Lausanne, Sep. 1996.
- [71] The ATLAS Collaboration. **Performance of the ATLAS Detector in Run 2**. <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/RunStatsPublicResults2010#2015>
- [72] The ATLAS Tile Calorimeter system. **The Laser calibration of the Atlas Tile Calorimeter during the LHC run 1**. Instrumentation and Detectors (physics.ins-det). arXiv:1608.02791v1 [physics.ins-det]
- [73] The ATLAS Tile Calorimeter system. **Calibration of the ATLAS Tile Calorimeter: Laser events in the empty bunch crossing**. ATL-COM-TILECAL-2014-068.
- [74] D. Calvet, J. Carvalho, R. Febbraro, B. Galhardo, C. Santoni, F. Vazeille, S. Viret. **Commissioning of the ATLAS tile calorimeter LASER calibration system**. ATL-TILECAL-INT-2011-004.
- [75] O. Lundberg. **Calibration and monitoring of the ATLAS Tile calorimeter**. ATL-TILECAL-SLIDE-2016-261.
- [76] P. Adragna, and al. **Testbeam Studies of Production Modules of the ATLAS Tile Calorimeter**. ATL-COM-TILECAL-2009-004.

BIBLIOGRAPHY

- [77] E. Dubreuil, D. Boumediene, D. Pallin. **Calibration of the ATLAS Tile Calorimeter channels using the LASER system.** ATL-TILECAL-INT-2012-006.
- [78] S. Fracchia. **Irradiation effect on the response of the scintillators in the ATLAS Tile Calorimeter.** ATL-TILECAL-PROC-2014-013.
- [79] D. Boumediene, D. Pallin. **Calibration of the Tile Hadronic Calorimeter of ATLAS at LHC.** ATL-TILECAL-PROC-2014-001.
- [80] P. Gris. **Upgrade of the Laser Calibration System of the Atlas Hadron Calorimeter.** ATL-TILECAL-2015-088.
- [81] D. Calvet, R. Marchionini, S. Viret, P. Lafarguette, D. Boumediene. **The TileCal Laser calibration system control and monitoring software** ATL-TILECAL-INT-2012-002
- [82] A. Buckley and al. **General-purpose event generators for LHC physics.** Cavendish-HEP-10/21 arXiv:1101.2599 [hep-ph].
- [83] G. Rodrigo, S. Catani, D. de Florian, W. Vogelsang. **Collinear splitting, parton evolution and the strange-quark asymmetry of the nucleon in NNLO QCD.** 10.1016/j.nuclphysbps.2004.09.048. arXiv:0406338 [hep-ph].
- [84] B.R. Webber. **Fragmentation and Hadronization.** Int.J.Mod.Phys.A15S1:577-606,2000. arXiv:9912292 [hep-ph].
- [85] T. Sjostrand, S. Mrenna, P. Skands. **PYTHIA 6.4 Physics and Manual.** FERMILAB-PUB-06-052-CD-T arXiv:0603175 [hep-ph].
- [86] T. Sjostrand, S. Mrenna, P. Skands. **A Brief Introduction to PYTHIA 8.1.** CERN-LCGAPP-2007-04. arXiv:0710.3820 [hep-ph].
- [87] S. Frixione, P. Nason, and C. Oleari. **Matching NLO QCD computations with Parton Shower simulations: the POWHEG method.** Bicocca-FT-07-9 arXiv: 0709.2092 [hep-ph].
- [88] J. Bellm et al. **Herwig 7.0 / Herwig++ 3.0 Release Note** Eur.Phys.J. C76 (2016) no.4, 196.
- [89] E. Gerwick, S. Hoeche, S. Marzani, S. Schumann. **Soft evolution of multi-jet final states.** SLAC-PUB-16143 arXiv:1411.7325 [hep-ph].
- [90] S. Frixione, F. Stoeckli, P. Torrielli, B. Webber, C. White. **The MC@NLO 4.0 Event Generator.** Cavendish-HEP-10/12. arXiv:1010.0819 [hep-ph].
- [91] The ATLAS Colaboration. **The ATLAS Simulation infrastructure.** CERN-PH-EP-2010-044. arXiv:1005.4568 [physics.ins-det].

BIBLIOGRAPHY

- [92] The Geant4 Collaboration (S. Agostinelli et al). **GEANT4: A Simulation Toolkit**. Published in Nuclear Instruments and Methods in Physics Research, NIM A 506 (2003), 250-303.
- [93] The ATLAS Collaboration. **Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)** ATL-SOFT-PUB-2007-007.
- [94] The ATLAS Collaboration. **Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment**. ATLAS-CONF-2012-042.
- [95] The ATLAS Collaboration. **Vertex Reconstruction Performance of the ATLAS Detector at $\sqrt{s} = 13$ TeV**. ATL-PHYS-PUB-2015-026.
- [96] The ATLAS Collaboration. **Performance of primary vertex reconstruction in proton-proton collisions at $\sqrt{s} = 7$ TeV in the ATLAS experiment**. ATLAS-CONF-2010-069.
- [97] The ATLAS Collaboration. **Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data**. ATLAS-CONF-2016-024.
- [98] The ATLAS Collaboration. **Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s} = 13$ TeV**. Eur.Phys.J.C70:823-874,2010. arXiv:1603.05598 [hep-ex].
- [99] The ATLAS Collaboration. **Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1** CERN-PH-EP-2015-304. arXiv:1603.02934 [hep-ex].
- [100] M. Cacciari, G. Salam, G. Soyez. **The anti-k_t jet clustering algorithm**. JHEP 0804:063,2008. arXiv:0802.1189 [hep-ph].
- [101] S. Ellis and D. Soper. **Successive combination jet algorithm for hadron collisions**. Phys. Rev. D48:3160-3166,1993. arXiv:9305266 [hep-ph].
- [102] M. Cacciari and G. P. Salam. **Pileup subtraction using jet areas**. Phys.Lett.B659:119-126,2008. arXiv:0707.1378 [hep-ph].
- [103] The ATLAS Collaboration. **Monte Carlo Calibration and Combination of In-situ Measurements of Jet Energy Scale, Jet Energy Resolution and Jet Mass in ATLAS**. ATL-COM-PHYS-2015-071.
- [104] The ATLAS Collaboration. **Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV**. Eur. Phys. J. C, 73 3 (2013) 2304. arXiv:1112.6426 [hep-ex].
- [105] The ATLAS Collaboration. **Inputs to Jet Reconstruction and Calibration with the ATLAS Detector Using Proton-Proton Collisions at $\sqrt{s} = 900$ GeV**. Technical Report ATLAS-CONF- 2010-016.

BIBLIOGRAPHY

- [106] The ATLAS Collaboration. **Properties of Jets and Inputs to Jet Reconstruction and Calibration with the ATLAS Detector Using Proton-Proton Collisions at $\sqrt{s}=13$ TeV.** ATL-PHYS-PUB-2015-036.
- [107] The ATLAS Collaboration. **Performance of jet substructure techniques for large-R jets in proton-proton collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector.** JHEP09 (2013) 076. arXiv:1306.4945 [hep-ex].
- [108] D. Krohn, J. Thaler, and L. Wang. **Jet Trimming.** JHEP 1002:084,2010. arXiv:0912.1342v2 [hep-ph].
- [109] The ATLAS Collaboration. **Measurement of the jet mass scale and resolution uncertainty for large radius jets at $\sqrt{s}=8$ TeV using the ATLAS detector.** ATLAS-CONF-2016-008.
- [110] J. Thaler and K. Van Tilburg. **Identifying Boosted Objects with N-subjettiness** JHEP 1103:015,2011. arXiv:1011.2268 [hep-ph].
- [111] The ATLAS Collaboration. **Boosted hadronic top identification at ATLAS for early 13 TeV data.** ATL-PHYS-PUB-2015-053.
- [112] The ATLAS Collaboration. **Tagging and suppression of pileup jets with the ATLAS detector.** ATLAS-CONF-2014-018.
- [113] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, H. Voss. **TMVA 4. Toolkit for Multivariate Data Analysis with ROOT.** PoS ACAT:040,2007. arXiv:physics/0703039 [physics.data-an].
- [114] The ATLAS Collaboration. **Expected performance of the ATLAS b-tagging algorithms in Run-2.** ATL-PHYS-PUB-2015-022.
- [115] The ATLAS Collaboration. **Performance of b-Jet Identification in the ATLAS Experiment.** JINST 11 P04008 (2016). arXiv:1512.01094 [hep-ex].
- [116] The ATLAS Collaboration. **Performance of missing transverse momentum reconstruction for the ATLAS detector in the first proton-proton collisions at $\sqrt{s} = 13$ TeV.** ATL-PHYS-PUB-2015-027
- [117] W.J. Stirling, private communication.
<http://www.hep.ph.ic.ac.uk/wstirlin/plots/plots.html>
- [118] R. M. Harris and S. Jain. **Cross Sections for Leptophobic Topcolor Z' Decaying to Top-Antitop.** Eur. Phys. J. C72 (2012) 2072, arXiv: 1112.4928 [hep-ph].
- [119] The ATLAS collaboration. **Searches for new physics using the $t\bar{t}$ invariant mass distribution in p-p collisions at $\sqrt{s} = 8$ TeV.** Phys. Rev. Lett. 111 (2013) 211804. arXiv:1309.2030 [hep-ex].

BIBLIOGRAPHY

- [120] The ATLAS collaboration. **A search for $t\bar{t}$ resonances in the lepton plus jets final state using 20 fb^{-1} of p-p collisions at $\sqrt{s} = 8\text{ TeV}$.** tech. rep. ATL-COM-PHYS-2014-003.
- [121] The ATLAS Collaboration. **Search for heavy particles decaying to pairs of highly-boosted top quarks using lepton-plus-jets events in proton-proton collisions at $\sqrt{s} = 13\text{ TeV}$ with the ATLAS detector.** ATLAS-CONF-2016-014.
- [122] T. Jezo et al. **NLO+NLL limits on W' and Z' gauge boson masses in general extensions of the Standard Model.** LPSC-14-220. arXiv:1410.4692v1.
- [123] J. Gao et al. **Next-to-leading order QCD corrections to the heavy resonance production and decay into top quark pair at the LHC.** Phys.Rev.D82:014020,2010. arXiv:1004.0876 [hep-ph].
- [124] M. Cacciari et al. **Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation.** Phys. Lett. B 710 (2012) 612, arXiv:1111.5869 [hep-ph].
- [125] M. Beneke et al. **Hadronic top-quark pair production with NNLL threshold resummation.** Nucl. Phys. B 855 (2012) 695, arXiv:1109.1536 [hep-ph].
- [126] P. Kant et al. **HATHOR for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions.** Comput.Phys.Commun. 191 (2015) 74-89. arXiv:1406.4403 [hep-ph].
- [127] J.H.Kuhn, A.Scharf and P. Uwer. **Electroweak corrections to top-quark pair production in quark-antiquark annihilation.** CERN-PH-TH/2005-145. arXiv:0508092 [hep-ph].
- [128] J.H.Kuhn, A.Scharf and P. Uwer. **Weak Interactions in Top-Quark Pair Production at Hadron Colliders: An Update.** Phys.Rev. D91 (2015) 1, 014020. arXiv:1305.5773 [hep-ph].
- [129] R. Gavin et al., **W Physics at the LHC with FEWZ 2.1.** Comput. Phys. Commun. 184 (2013) 208, arXiv:1201.5896 [hep-ph].
- [130] The ATLAS Collaboration. **Multi-boson simulation for 13 TeV ATLAS analyses.** ATL-PHYS-PUB-2016-002.
- [131] The ATLAS Collaboration. **Improved luminosity determination in pp collisions at $\sqrt{s} = 7\text{ TeV}$ using the ATLAS detector at the LHC.** Eur. Phys. J. C 73 (2013) 2518. arXiv:1302.4393 [hep-ex].
- [132] The ATLAS Collaboration. **Electron identification measurements in ATLAS using $\sqrt{s} = 13\text{ TeV}$ data with 50ns bunch spacing.** ATL-PHYS-PUB-2015-043, 2015.

BIBLIOGRAPHY

- [133] The ATLAS Collaboration. **Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s}=13$ TeV.** ATL-PHYS-PUB-2015-015.
- [134] The ATLAS Collaboration. **Identification of high transverse momentum top quarks in pp collisions at $\sqrt{s}=8$ TeV with the ATLAS detector (2016).** JHEP 06 (2016) 093. arXiv: 1603.03127 [hep-ex].
- [135] M. Czakon, P. Fiedler and A. Mitov. **The total top quark pair production cross-section at hadron colliders through $\mathcal{O}(\alpha_s^4)$.** Phys. Rev. Lett. 110 (2013) 252004, arXiv: 1303.6254 [hep-ph].
- [136] N. Kidonakis. **Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production.** Phys. Rev. D 83 (2011) 091503, arXiv: 1103.2792 [hep-ph].
- [137] C.H. Kom and W.J. Stirling. **Charge asymmetry in W+jets production at the LHC.** Eur.Phys.J.C69:67-73,2010. arXiv:1004.3404 [hep-ph].
- [138] G. Choudalakis. **On hypothesis testing, trials factor, hypertests and the BUM-PHUNTER.** Prepared for PHYSTAT2011. arXiv:1101.0390 [physics.data-an].
- [139] A. L. Read. **Presentation of search results: The CL(s) technique.** J. Phys. G 28 (2002) 2693.