ROYAUME DU MAROC
Université Mohammed V
Faculté des sciences
Rabat
جامعة محمد الخامس - كلية العلوم
الرباط

Nº d'ordre 3027

THÈSE DE DOCTORAT

Présentée par:

Zineb IDRISSI

Discipline: Physique
Spécialité: Physique Nucléaire

Titre:

Search for $t\bar{t}$ resonances in highly-boosted lepton+jets final state in proton-proton collisions at $\sqrt{s}$=13 TeV with the ATLAS detector at the Large Hadron Collider

Soutenue le **samedi 14 octobre 2017**

Devant le jury:
Président:

M. Mohamed CHABAB	PES	Faculté des Sciences Semlalia, Université Caddi Ayyad, Marrakech

Examinateurs:

Mme Rajaa CHERKAOUI EL MOURSLI	PES	Faculté des Sciences, Université Mohammed V, Rabat
M. Yahya TAYALATI	PES	Faculté des Sciences, Université Mohammed V, Rabat
Mme Farida FASSI	PH	Faculté des Sciences, Université Mohammed V, Rabat
M. José Fransisco SALT CAIROLS	PES	Consejo Superior de Investigaciones Cientificas, IFIC, Valencia.

To my father, who could not see the woman I have become: I miss you and I miss your warm hugs. If I have gone my further than I ever really imagined, you have been the one in getting me here.

Acknowledgements

The search presented in this thesis was performed in the Laboratory of Nuclear Physics (LPNR) in the Faculty of Science, Rabat under the supervision of Pr. Rajaa Cherkaoui El Moursli and the co-supervison of Pr. Farida Fassi.

Writing this acknowledgment has afforded me the opportunity to think about all the wonderful people who, in different ways, contributed to the success of this research.

Firstly, I would like to express my sincere gratitude to my supervisors Prof. Rajaa Cherkaoui EL Moursli and Prof. Farida Fassi for their constant support, who understood my curiosity and were always there whenever I needed their guidance and motivation. This thesis would never have been possible without their incredible support, inspiration and encouragement.

Next I thank the rest of the members of my examination committee for their time and effort during the defense process.

Thank you Pr. Mohamed Chabab, the president of the examination committee and PES from Caddi Ayyad University, Marrakech.

Thank you Pr. Yahya Tayalati, Reviewer of this thesis and PES from Faculty of Science Rabat.

Thank you Pr. José Fransisco. Salt Cairols, Reviewer of this thesis and Professor de Investigación in Consejo Superior de Investigaciones Cientificas, IFIC, Valencia.

Certain members of the ATLAS collaboration deserve special recognition from my part. First, I would like to thank Zachary Louis Marshall for giving me the opportunity to work with the ATLAS simulation group during my qualification year work, for his support when needed and for his trust by giving me such a nice and interesting subject. I learned a lot from you. My gratitude goes as well to Caterina Doglioni for all that she could teach me during my stay at CERN, her enthusiasm for physics and perseverance have been inspirational, I'm delighted to have you known.

Special thank goes to my mom, who have always been there and constantly encouraging and expecting the best from me, thank you for your unconditional love, strength, sacrifices

and referrals - I wrote this thesis for you.

I would also like to thank my amazing sister and brothers, Leila, Yassine and Abdel Majid. Although our lives were suddenly changed and forced us to contend with the death of our beloved father, my love for you has never been stronger.

And last but certainly not least, I'm beyond grateful to have a person who I could count on for anything, Nacim, words will never be enough -thank you for our life together - the past, present and future... this is for us.

Thank you all.

Abstract

The first part of this thesis is dedicated to the extraction of jet energy scale corrections for the ATLAS experiment at the LHC. The systematic validation of the different Geant4 physics lists are presented. Various extraction methods are described, with a particular interest on corrections extraction using a wider choice of physics lists: QGSP BERT, QGSP BIC, QGSP FTFP BIC, FTFP BERT, QBBC and FTF BIC. These corrections are necessary for all analyses using jets and are already used by the whole collaboration.

The second part presents the searches for new physics beyond the Standard Model(SM). Many models of new physics predict new particles with enhanced coupling to the top quark, which can show up as resonances in the top-antitop invariant mass spectrum. The analysis is performed with the ATLAS detector at the Large Hadron Collider (LHC) at the European Organization for Nuclear Research (CERN) located near Geneva in Switzerland. The data collected in 2015 are produced using proton-proton (pp) collisions at the center-of-mass energy of $\sqrt{s}$ = 13 TeV with an integrated luminosity of 3.2 fb^{-1}. This energy was the highest energy ever reached in a laboratory at this time. Events consistent with top-quark pair production are selected by requiring a single isolated charged lepton, missing transverse momentum, and a high-transverse momentum jet with substructure compatible with that of a hadronic top-quark decay. Jets identified as likely to contain b-hadrons are used to further control the background from other SM processes. The $t\bar{t}$ system is reconstructed using boosted techniques. A comparison between data and estimated background showed no significant excesses. Hence 95% Condfidence Level (C.L) upper limits on the cross section times acceptance have been set for Z' benchmark model.

Key Words: ATLAS, LHC, top quark, jet systematic validation, resonances top-antitop, leptophobic topcolor Z' , Grid Computing, Simulation, GEANT4.

Résumé

La première partie de cette thèse est dédiée à l'extraction des corrections de l'échelle en énergie des jets dans l'expérience ATLAS du LHC. Une validation systématique des différentes listes de physique de Geant4 est présentée. Les diverses méthodes d'extraction sont détaillées, en insistant tout particulièrement sur l'utilisation d'un choix plus large de physics lists : QGSP BERT, QGSP BIC, QGSP FTFP BIC, FTFP BERT, QBBC et FTF BIC. Les corrections ainsi extraites sont primordiales pour les analyses de physique utilisant des jets, et sont d'ores et déjà utilisées par toute la collaboration.

La deuxième partie présente la recherche d'une nouvelle particule lourde au-delà du modèle standard (SM) se désintegrant en paires top-antitop $t\bar{t}$. L'analyse est menée auprès de l'expérience ATLAS au grand collisionneur de Hadrons (LHC) de l'Organisation Européenne pour la Recherche Nucléaire (CERN) situé prés de Genève en Suisse. Les données recueillies en 2015 sont produites à l'aide des collisions proton-proton (pp) avec une énergie inégalée dans le centre de masse de $\sqrt{s} = 13$ TeV et correspondent à une luminosité intégrale de 3,2 fb^{-1}. Cette énergie est la plus élevée jamais atteinte dans un laboratoire à ce jours. Les événements consistant avec la production de paires de quarks tops sont sélectionnés en exigeant un seul lepton chargé isolé, un moment transversal manquant et un jet de moment transversal élevé avec une sous-structure compatible avec celle d'un quark top hadronique. Les jets identifiés comme susceptibles de contenir des hadrons b sont utilisés pour contrôler davantage les bruits de fond des autres processus du Modèle Standard. Le système $t\bar{t}$ est reconstruit en utilisant des techniques boostées, où les produits de désintégration du quark top sont émis dans une même région angulaire du détecteur et pour lequel la reconstruction des quarks top nécessite le déploiement de méthodes de reconstruction innovantes. Une comparaison entre les données et le bruit de fond estimé n'a montré aucun excès d'évènements significatif. Par conséquent, les résultats sont interprétés en termes de limites supèrieures à 95% de niveau de confiance (C.L) posées sur le produit de la section efficace de production par le rapport de branchement en paires de quarks top pour une resonance Z' du modèle leptophobique dans les modèles de technicouleur assistée par topcolor.

Mots clés: ATLAS, LHC, top quark, jet systématique validation, résonances top-antitop, leptophobic topcolor Z', Grid Computing, Simulation, GEANT4.

Résumé détaillé

Le Modèle Standard (MS) de la physique des particules est le cadre théorique qui synthétise les connaissances actuelles sur la physique subatomique. En d'autre terme, il regroupe toutes les interactions fondamentales connues (sauf la gravitation). Jusqu'à présent, aucun résultat expérimental ne contredit fortement les prédictions du MS. Cependant, il ne peut pas prétendre au titre de modèle 'ultime' de la physique des hautes énergies puisqu'il ne nous apprend rien sur la structure quantique de l'interaction gravitationnelle, ne décrit pas l'oscillation des neutrinos, requiert l'ajustement fin des paramètres de renormalisation lors du calcul de la masse du boson de Higgs et une partie de la matière présente dans l'univers lui échappe. Il doit ainsi être envisagé comme un modèle effectif, la limite à basse énergie d'une théorie plus large, dont la nature nous échappe toujours.

De nombreux modèles théoriques se proposant de décrire la physique au delà du MS ont été proposés dans les dernières décennies pour tenter de remédier à une ou plusieurs des lacunes du MS. Certaines de ces théories comme le modèle technicouleur, les modèles de couleurs chiraux et les modèles de Randall-Sundrum avec des dimensions supplémentaires prédisent de nouvelles particules qui se couplent fortement au quark top, en raison de sa masse élevée. Ce couplage implique une désintégration préférentielle en une paires de quark top, $t\bar{t}$.

Pour pouvoir tester ces nouvelles théories, de grand collisionneur de hadrons dimmensioné à hauteur du défi et capables d'atteindre des énergies de l'ordre du TeV sont nécessaires, de même que des détecteurs adaptés pour analyser les événements recueillis. Avec ses 27 kilomètres de circonférence, le LHC (acronyme de Large Hadron Collider) permettra de pousser les tests du modèle standard dans ses derniers retranchements et ouvre la voie à la recherche de nouvelle physique à l'échelle du Téra-électronVolt. Il s'agit d'un accélérateur de particules circulaire, de type synchrotron, désigné pour accélérer des protons et des ions lourds de plomb jusqu'à des énergies de 7 TeV et 5.5 TeV respectivement et produire des collisions proton-proton ou plomb-plomb à hautes fréquences.

Le quark top, ou t-quark, découvert en 1995 auprès des expériences DΦ et CDF du

Tevatron, est la particule la plus lourde du Modèle Standard, sa masse est comparable à celle d'un noyau d'or. et,étant le seul quark qui désintègre avant d'hadroniser, ce quark nous présente des opportunités singulières dans l'étude de nombreuses extensions au modèle standard, au travers de fort couplages à la nouvelle physique.

Le travail de recherche exposé dans cette thèse a été réalisé en utilisant les données collectées par le détecteur ATLAS auprès du LHC. Le LHC est l'accélérateur de particules le plus grand et le plus puissant jamais construit. Il est constitué d'un double anneau de stockage de protons. Environ 10000 physiciens et ingénieurs du monde entier participent aÌĂ cette expérience en développant de nouvelles techniques et approches pour identifier les événements intéressants cachés dans l'environement complexe produit dans des collisions proton-proton. Les premières collisions proton-proton ont eu lieu aÌĂ la fin de l'année 2009 aÌĂ une énergie de 900 GeV dans le système du centre de masse. Le 19 Mars 2010, le LHC a battu un record en augmentant l'énergie des faisceaux à 3.5 TeV, et les premières collisions aÌĂ 7 TeV ont été enregistrées le 30 Mars 2010. L'énergie par faisceau pour l'ensemble de 2011 était de 3.5 TeV, et 4 TeV en 2012. Un an d'arrêt technique en 2013 avant de parvenir à l'énergie de collision de 13 TeV. Quatre expériences se partagent les quatre points de croisement des faisceaux du grand anneau du LHC. Une d'entre elles, ATLAS, est un détecteur généraliste avec un vaste programme de physique.

ATLAS est organisé en un ensemble concentrique de sous détecteurs. Du point d'interaction vers lâĂŹextérieur, les particules produites par les collisions rencontrent successivement: Le détecteur interne de traces émergé dans un champ magnétique de 2 T, offrant une couverture jusqu'à $\eta < 2.5$, le système calorimétrique allant jusqu'à $\eta < 4.9$, le spectromètre à muons dans un champ magnétique toroiÌĹdal et un système de déclenchement composé de trois niveaux. Tous les sous-systèmes ont d'excellentes performances en termes d'efficacité et de reÌĄsolution.

Cette thése comprend deux travaux interconnectés. En premier lieu, elle décrit l'extraction des corrections de l'échelle en énergie des jets dans l'expérience ATLAS du LHC. Une validation systématique des différentes listes de physique de Geant4 est présentée. Les diverses méthodes d'extraction sont détaillées, en insistant tout particulièrement sur l'utilisation d'un choix plus large de physics lists : QGSP BERT, QGSP BIC, QGSP FTFP BIC, FTFP BERT, QBBC et FTF BIC. Les corrections ainsi extraites sont primordiales pour les analyses de physique utilisant des jets, et sont d'ores et déjà utilisées par toute la collaboration.

La deuxième partie présente la recherche d'une nouvelle particule lourde au-delà du modèle standard (SM) se désintegrant en paires top-antitop $t\bar{t}$. L'analyse est menée auprès de l'expérience ATLAS au grand collisionneur de Hadrons (LHC) de l'Organisation Européenne pour la Recherche Nucléaire (CERN) situé prés de Genève en Suisse. Les données

recueillies en 2015 sont produites à l'aide des collisions proton-proton (pp) avec une énergie inégalée dans le centre de masse de $\sqrt{s} = 13$ TeV et correspondent à une luminosité intégrale de 3,2 fb^{-1}. Cette énergie est la plus élevée jamais atteinte dans un laboratoire à ce jours. La recherche de reÌĄsonance $t\bar{t}$ se fait dans le canal lepton plus jets où un boson W venant d'un quark top se desintègre leptoniquement (en un électron ou en un muon et le neutrino correspondant) et l'autre W de façon hadronique (en une paire quark-antiquark). En d'autre terme, les événements consistant sont sélectionnés en exigeant un seul lepton chargé isolé, un moment transversal manquant et un jet de moment transversal élevé avec une sous-structure compatible avec celle d'un quark top hadronique. Les jets identifiés comme susceptibles de contenir des hadrons b sont utilisés pour contrôler davantage les bruits de fond des autres processus du Modèle Standard. Le système $t\bar{t}$ est reconstruit en utilisant des techniques boostées, où les produits de désintégration du quark top sont émis dans une même région angulaire du détecteur et pour lequel la reconstruction des quarks top nécessite le déploiement de méthodes de reconstruction innovantes. Une comparaison entre les données et le bruit de fond estimé n'a montré aucun excès d'évènements significatif. Par conséquent, les résultats sont interprétés en termes de limites supèrieures à 95% de niveau de confiance (C.L) posées sur le produit de la section efficace de production par le rapport de branchement en paires de quarks top pour une resonance Z' du modèle leptophobique dans les modèles de technicouleur assistée par topcolor.

Contents

List of Tables

List of Figures

Introduction

For over twenty centuries people have thought about the smallest physical world components. The ancient Greeks philosophers gave much to the world of physics by developing the basis of fundamental modern principles as the conservation of matter and the atomic theory, that matter is composed of indivisible atoms. Very few new developments occurred in the centuries following the Greek period. However, as the large advancement of scientific progress of the Renaissance entered the field of physics, other great thinkers began to reject the older philosophical approaches and established new ideas based on empirical methods. At the start of the twentieth century, scientists believed that physics was complete and they understood the most fundamental principles of nature. Physicists trusted classical mechanics, most of the important laws of physics had been discovered and the Newton's theory of gravitation was accepted and remained unquestioned until the starting of Einstein's theory of relativity. Scientists realized that their knowledge was far from complete and the field of quantum mechanics beging to completely alter the fundamental precepts of physics. By the mid-1960s, physicists realized that their previous understanding, where all matter is composed of the fundamental protons, neutrons, and electrons, was insufficient to explain the myriad new particles being discovered. The theory that is now called the Standard Model [11] of particles has gradually grown and gained further credence with new evidence from the particle accelerators.

From all the so far discovered elementary particles, the top-quark which has the largest mass with several interesting properties. It belongs to the classification of a *quark* in the Standard Model, of which there are six flavours and interacts through all kinds of forces in the model: the strong, the weak and the electromagnetic interaction. Due to its large mass; m_{top} =172.5 GeV, the top quark may play a special role in nature, as it occupies many interesting position in the Beyond Standard Model (BSM) predictions, proposed in the last years to try to account for one or more of the Standard Model's open questions. Several of these theories anticipate, the existence of new heavy particles that decay predominantly into top quark pairs, specially those related to electroweak Symmetry Breaking (EWSB).

The evidence of the SM and/or of the BSM physics at the TeV scale can be tested through precision measurements at high energy experiments and is at this point that the Large Hadron Collider [12] experiment, located at the European Organization for Nuclear Research (CERN) enters in the scene.

The LHC is the most powerful particle accelerator reconstructed to date. Around 10,000 scientists and engineers from all over the world, are taking part in this experience by developing new techniques and approaches to identify the interesting physics recorded from high energetic proton-proton collisions. The proton (heavy-ion) beams are running in two different beam-pipes, and intersect in four interaction points, where four major detectors that study the high-energy collisions are installed. One of these detectors is ATLAS [13], a multi-purpose particle detector with an extensive initial physics program that includes the test of the SM at new energy regimes and enable the search of signatures of new physics. One of its greatest success so far was the discovery of the Higgs-boson in the year 2012 [14, 15].

This thesis presents the results of the search for new physics that produce resonances in the top-antitop quark pairs in the lepton plus jets final state using a data sample with an integrated luminosity of 3.2 fb^{-1} collected in 2015 by the ATLAS detector from $\sqrt{s} =$ 13 TeV LHC proton-proton collisions. The analysis strategy used is expected to be agnostic of the theoretical model. However the results are quantified by quoting upper limits on the production cross-section times branching ratio for a Z' benchmark model [16]. The search is made in the top and anti-top quark each decay to a bottom quark plus a W boson, one W boson decays leptonically (to an electron or muon plus neutrino) while the W from the second top decays hadronically (to quarks). This leads to a signature with a single isolated high-transverse-momentum lepton, large missing transverse momentum (from the escaping neutrino) and hadronic jets. At least one of the jets must be identified as likely to contain a b-hadron (b-jet). The analysis is optimised to adapt the boosted configuration where the the top-antitop system is highly energetic and the hadronic top decay products overlap in the detector into a single large-radius (large-R) jet. A variety of new physics scenarios give rise to heavy particles that decay preferentially to $t\bar{t}$. The analysis is done in a very generic way, the benchmark signal model used produces a spin 1 colour singlet vector bosons, denoted Z'_{TC2}, according to the topcolour-assisted technicolour model [16]. The parameters of this model are tuned so that the resonance has a width of 1.2% ($\Gamma/m = 1.2\%$). The ATLAS and CMS [17] collaborations set lower limits on the allowed mass for such bosons of $\mathrm{m}(Z'_{TC2}) >$ 1.8 TeV and $\mathrm{m}(Z'_{TC2}) > 2.4$ TeV, respectively. An additional Z'_{TC2} parameter point with a 3% width is tested. A previous search by the ATLAS collaboration sets lower limits on the allowed mass for such bosons of $\mathrm{m}(Z'_{TC2}) > 2.3$ TeV.

This thesis is structured as follows: An overview of the SM of particle physics, its shortcomings and its relation with the relevant aspects of the top quark is given on Chapter 1. The LHC complex at CERN and ATLAS detector are described in Chapter 2, Chapter 3 is devoted to describe the event simulation process in ATLAS and the Monte Carlo generators used to produce the simulated samples used along this thesis. It comprises also the description of the strategy used to validate different Geant4 physics lists. An overview of the reconstruction, identification and calibration strategies for physics objects in the ATLAS is given in Chapter 4. Only the physics objects relevant for this thesis are described: electrons, muons, jets, b-tagging and neutrinos. Chapter 5 details the search strategy, including the selection, the $t\bar{t}$ pair reconstruction methods, the estimation of background processes, the systematic uncertainties, and the limit setting results.

As a final comment, the natural units commonly used by the experimental high energy physics community and which are employed throughout this thesis are such that $\hbar = c = 1$, where c is the speed of light and $\hbar$ is the reduced Planck's constant, sometimes referred to as the Dirac constant. Consequently, the values of a particle's mass (m), momentum (mc) is numerically equal to its energy in its rest frame mc^2 and represented in terms of GeV [1]. The units of length ($\hbar/mc$) and time ($\hbar/mc^2$) are the same and they are the inverse of the units of energy and mass (GeV^{-1}).

$$[mass] = [energy] = [momentum] = [length]^{-1} = [time]^{-1} \tag{1}$$

The Einstein summation notation of summing repeated indices in vectors, matrices and tensors is adopted. [2]

[1] $1GeV = 10^9 eV = 1.602176565(35).10^{-10} J$, the rest mass of proton corresponds roughly to $1GeV$.

[2] In relativistic physics a notational convention for simplifying expressions is included when an index variable appears twice in a single term and is not otherwise defined, it implies summation of that term over all the values of the index.

Chapter 1

Theoretical Foundations

The aim of this thesis is to perform a search of new heavy particles that decay into top-antitop quarks pairs ($t\bar{t}$) using data from proton-proton collisions at the centre-of-mass energy of 13 TeV collected by the ATLAS detector at the Large Hadron Collider. In this chapter we discuss the underlying theory that is related with this studies and we point out its most important aspects.

In section 1.1 a brief overview of the basic concepts of the Standard Model of particle physics is given. Section 1.2 dives into the top quark physics sector: from production of top quark with emphasis to $t\bar{t}$ pair, to the decay channels of the latter. The importance of beyond Standard model theories and the SM known limitations are highlighted in section 1.3. Section 1.4 is devoted to discuss the top-antitop invariant mass spectrum in the beyond standard model theories. The benchmark model used in this search for resonant $t\bar{t}$ production is detailed in Section 1.5.

1.1 The Standard Model

The Standard Model (SM) [11] is a relativistic quantum field theory [18] emerged during the 1960s and 1970s. Since then, it has been tested by several experiments and no significant deviation from its predictions have been observed. It provides an excellent description of the known fundamental particles and their interactions: the electromagnetic, the strong and the weak interaction except the gravitation[1]. All known elementary particles are classified as fermions and bosons. Fermions have half-integer spin and obey the Fermi-Dirac statistic [19], while bosons have integer spin and obey the Bose-Einstein statistic [20]. For each particle, an anti-particle with the same mass, but opposite charge exists. The antiparticle

[1] At high energy scales, the gravitation can be neglected, since its strength is around 43 orders of magnitude smaller than the strong interaction.

can be denoted with an identical symbol used for the particle but with commonly a bar added over it.

Fermions are categorized in three families: three leptons (e, μ, τ), three neutrinos (υ_e, υ_μ, υ_τ) and six quarks(u (up), d (down), c (charm), s (strange), b (bottom), t (top)). The force mediation in the SM is described by the requirement of gauge symmetries, that is, transformations which do not commute in the case of the QCD and weak groups. Gauge invariance leads the existence of a set of gauge boson fields, which transmit the interaction. The SM is a $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ gauge theory. $U(1)_Y$ is the symmetry group of electromagnetic interaction mediated by the photon fields, $SU(2)_L$ of the weak interaction is conveyed by the $W^{\pm}$ and Z fields while $SU(3)_C$ of the strong interaction is transmitted through the gluon fields. C, L and Y correspond to the conserved quantum numbers for each symmetry: color charge, weak isospin and hypercharge, respectively. The last particle to complete the SM list of particles is the Higgs boson which is the consequence of introducing a new doublet of complex scalar fields in the SM theory in order to give mass to the $W^{\pm}$ and Z^0 bosons and fermions. Figure 1.1 summarizes some properties of the SM particles.

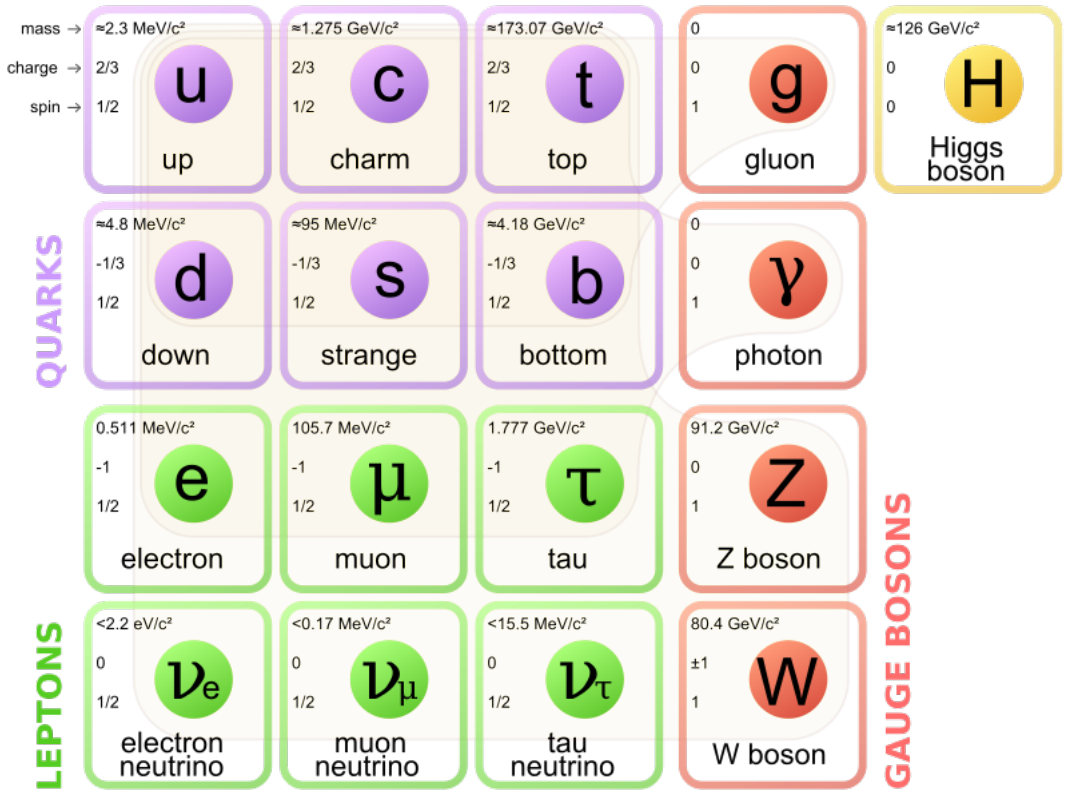


Figure 1.1 – The elementary particles in the Standard Model of particle physics (Figure from [1])

The dynamics of the SM are described by the Lagrangian density of the theory which

can be divided into several pieces [21]

$$\mathcal{L}_{SM} = \mathcal{L}_{Gauge} + \mathcal{L}_{Matter} + \mathcal{L}_{Yukawa} + \mathcal{L}_{Higgs}. \tag{1.1}$$

The matter Lagrangian term includes the kinetic energy of the fermions and their interactions with the gauge fields. The gauge term includes the kinetic energy of gauge fields for strong and electroweak interactions. The interaction of the Higgs field with the fermions is done by the Yukawa interaction terms of the Lagrangian, which provide a dynamical mechanism by which the particles acquire mass. Finally, the Higgs field sector contains the Higgs kinetic energy, its gauge interactions, and the Higgs potential, which causes a non-zero vacuum expectation value.

1.1.1 The Electroweak interactions

The electromagnetic and the weak interaction were unified into electroweak (EW) interaction by a robust theory described by Glashow, Weinberg and Salam in the 60s. The gauge theory that describes both interactions is based on the $SU(2)_L \times U(1)_Y$ symetry group. Fermions are described by their fermionic fields, denoted by ψ_f, that are decomposed into right and a left-handed parts (ψ_R and ψ_L) based on the chirality. therefore:

$$\psi_f = \frac{1-\gamma_5}{2}\psi + \frac{1+\gamma_5}{2}\psi = \psi_L + \psi_R \tag{1.2}$$

The matrices $P_\pm = (1 \pm \gamma_5)/2$ are projectors. They satisfy the relations $P_\pm P_\pm = P_\pm, P_\pm P_\mp = 0, P_+ + P_- = 1$. γ_5 is the fifth gamma matrix of the Dirac matrices (γ^μ) and takes the value $\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3$, while the ψ_L and ψ_R are the left and right-handed projections respectively. The standard EW theory is a chiral theory, ψ_L and ψ_R behave differently under the gauge group. In particular all left-handed fermions transform under both the $SU(2)_L$ and $U(1)_Y$ and form a weak-isospin doublets:

$$\psi_i = \begin{pmatrix} u_i \\ d_i \end{pmatrix} or \begin{pmatrix} \nu_i \\ \ell_i \end{pmatrix} \tag{1.3}$$

while right-handed fermions only transform under $U(1)_Y$ and are represented as weak-isospin singlets.

By imposing these symmetries to the Lagrangian of the free fermionic field that describes the spin-1/2 particles ψ_f (Dirac equation):

$$\mathcal{L} = i\sum_f \bar{\psi}_f \gamma^\mu \partial_\mu \psi_f \tag{1.4}$$

the covariant derivative must take the following form in order to retain the invariance:

$$D_\mu = \partial_\mu + \frac{1}{2} g \tau^i L, R W^i_\mu - \frac{1}{2} i g' Y L, R B_\mu \tag{1.5}$$

where g and g' are the coupling constant associated to SU(2) and U(1), respectively. τ^i are the Pauli matrices, generators associated with the $SU(2)_L$ symmetry group while Y is the hypercharge, generator of $U(1)_Y$. $W^i_\mu, i = 1, 2, 3$ is a triplet vector fiels of SU(2) while B_μ is a single vector fiels introduced for U(1) symmetry group. The conserved quantities for the $SU(2)_L$ and the $U(1)_Y$ are the third component of the weak isospin (I_3) and the hypercharge respectively. They are associated via the Gell-Mann - Nishijima formula $Q = I_3 + Y$. Q is the electric charge. As before, the vector fields that are introduced by each group must also be transformed:

$$B_\mu = B_\mu - \partial_\mu \alpha(x) \tag{1.6}$$

$$W^i_\mu = W^i_\mu - \partial_\mu \beta_i(x) - g \varepsilon_{ijk} \beta_j(x) W^k_\mu \tag{1.7}$$

where ε_{ijk} is the Levi-Civita symbol, and the $\alpha(x)$ and $\beta_j(x)$ are phases. In the end, the Lagrangien of the electroweak interaction ruling the interaction between the gauge fields and fermions, consists of the following terms :

$$\mathcal{L} = \mathcal{L}_{fermion} + \mathcal{L}_{interaction} + \mathcal{L}_{gauge} \tag{1.8}$$

where:

$$\mathcal{L}_{fermion} = i \sum_f \bar{\psi}_f \gamma^\mu \partial_\mu \psi_f = i \bar{L} \gamma^\mu \partial_\mu L + i \bar{R} \gamma^\mu \partial_\mu R, \tag{1.9}$$

$$\mathcal{L}_{interaction} = \bar{L} \gamma^\mu (-\frac{g}{2} \tau^i W^i_\mu - \frac{g'}{2} Y B_\mu) L + \bar{R} \gamma^\mu (-\frac{g'}{2} Y B_\mu) R, \tag{1.10}$$

$$\mathcal{L}_{gauge} = -\frac{1}{4} W^i_{\mu\nu} W^{i,\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}. \tag{1.11}$$

The parameters L and R refer to the left and right-handed fermions, while the $W_{i,\mu\nu}$ and $B_{\mu\nu}$ are the field strength tensors. The first term corresponds to the fermion field with its left-handed and right-handed components, the second term represents the interaction of the fermions with the gauge bosons, and the last term provides the contribution and self-interactions of the vector fields. Rewriting the interaction term, we can separate it into two categories:

$$\mathcal{L}_{charged} = -\frac{g}{2}\bar{L}\gamma^{\mu}(\tau^1 W_{\mu}^1 + \tau^2 W_{\mu}^2)L \tag{1.12}$$

$$\mathcal{L}_{neutral} = -\frac{g}{2}\bar{L}\gamma_{\mu}\tau^3 W_{\mu}^3 L - \frac{g'}{2}\bar{L}\gamma^{\mu} Y B_{\mu} L - \frac{g'}{2}\bar{R}\gamma^{\mu} Y B_{\mu} R. \tag{1.13}$$

$\mathcal{L}_{charged}$ contains only left-handed components, it can be written in the following form:

$$\mathcal{L}_{charged} = -\frac{g}{2}\bar{L}\gamma_{\mu}\begin{pmatrix} 0 & W_{\mu}^1 - iW_{\mu}^2 \\ W_{\mu}^1 + iW_{\mu}^2 & 0 \end{pmatrix} \tag{1.14}$$

where it is possible to extract the definition of the charged gauge boson as the linear combination of the vector fields, given by:

$$W_{\mu}^{+} = \frac{W_{\mu}^1 - iW_{\mu}^2}{\sqrt{2}} \tag{1.15}$$

$$W_{\mu}^{-} = \frac{W_{\mu}^1 + iW_{\mu}^2}{\sqrt{2}} \tag{1.16}$$

The fact that left-handed fermions do not undergo charged current interactions, is incorporated through the vector (V) and axial-vector (A) components that are introduced and is in accordance to the experimental observation of parity violating processes. Equivalently, from the $\mathcal{L}_{neutral}$, which contains both left-handed and right-handed components, we can redefine the fields in the following way:

$$\begin{pmatrix} A_{\mu} \\ Z_{\mu} \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix}\begin{pmatrix} B_{\mu} \\ W_{\mu}^3 \end{pmatrix} \tag{1.17}$$

which gives the neutral gauge bosons as a combination of the introduced vector fields. The parameter $\theta_W$is called the Weinberg weak mixing angle and is defined in the following way:

$$\sin(\theta_W) = \frac{g'}{\sqrt{g^2 + g'^2}} \tag{1.18}$$

$$\cos(\theta_W) = \frac{g}{\sqrt{g^2 + g'^2}} \tag{1.19}$$

where g and g' are the coupling constants.

In conclusion, the introduced vector fields which are required to retain the invariance of the electroweak Lagrangian, give rise to gauge bosons that act as mediators of the elec-

troweak interaction. However, the theory as described so far predicts massless SU(3) gauge fields. This is in contrast to the observations since the $W^{\pm}$ and Z^0 are massive bosons. As a result, an additional mechanism is required to solve this problem.

1.1.2 Electroweak symmetry breaking

The sectors of SM lagrangien $\mathcal{L} = \mathcal{L}_{matter} + \mathcal{L}_{gauge}$ describes a theory of massless, while the electron, muon, tauon, the six quarks, the $W^{\pm}$ and Z are not. The masses have to be generated, but the theory has to remain re-normalizable, meaning that local gauge invariance has to remain conserved. A mechanism has been proposed called the *Brout-Englert-Higgs* with which the mass of the gauge bosons is explained, it consists in introducing an additional complex scalar doublet:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2) \\ \frac{1}{\sqrt{2}}(\phi_3 + i\phi_4) \end{pmatrix} \tag{1.20}$$

The Higgs lagrangian is given by:

$$\mathcal{L}_{Higgs} = (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi), \tag{1.21}$$

where $V(\phi)$ is a scalar potentiel, called Higgs potentiel and is given by:

$$V(\phi) = -\mu^2 |\phi|^2 + \lambda^2 |\phi|^4. \tag{1.22}$$

The covariant derivative in the Lagrangian is also chosen such that it remains invariant under the $SU(2)_L \times U(1)_Y$ transformations. The shape of the potential V depends on the parameters μ^2 and λ. For $\mu^2 < 0$ and $\lambda > 0$, V has its minimum at $\phi^\dagger \phi = -\frac{\mu^2}{2\lambda}$. Spontaneous symmetry breaking is induced if the minimum of V is obtained for non-vanishing ϕ values. Expanding ϕ around a particular minimum, say:

$$\phi_1 = \phi_2 = \phi_4 = 0 \phi_3^2 = -\frac{\mu^2}{\lambda} = u^2, \tag{1.23}$$

we obtain:

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ u + H \end{pmatrix}. \tag{1.24}$$

where H is a Hermitian scalar field which represents the physical Higgs boson. The Lagrangian now becomes:

$$\mathcal{L}_{Higgs} = \frac{1}{8}[g^2(W_1^2 + W_2^2) + (-gW_3 + g'YB_\mu)^2](u+H)^2 + \frac{1}{2}(\partial H)^2 - \mu^2(H)^2 - u\lambda(H)^3 - \frac{1}{4}\lambda(H)^4 \tag{1.25}$$

where using equations 1.17 and 1.14 the first two terms can be written as:

$$\mathcal{L}_{Higgs} = \frac{1}{2}g^2W^+W^-(u+H)^2 + \frac{1}{8}(g^2+g'^2)Z^2(u+H)^2 + 0.A_\mu^2 + ... \tag{1.26}$$

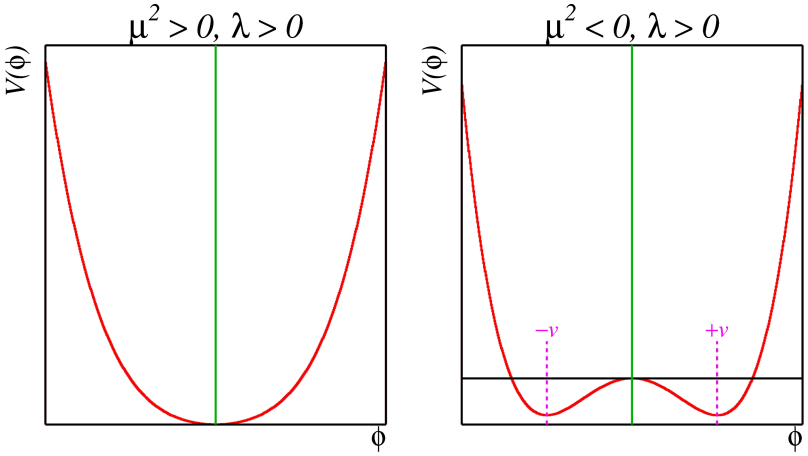


Figure 1.2 – The Higgs potential as a function of $\sqrt{\phi^\dagger\phi}$ for the case of $\lambda^2 > 0$ and $\lambda^2 < 0$.

Under This formulation the mass terms have been included for $W^\pm$ and Z, while the other terms specify the couplings between the Higgs boson and the gauge bosons.

We define the gauge boson mass as

$$m_W = \frac{gu}{2} \tag{1.27}$$

$$m_Z = \frac{u\sqrt{g^2+g'^2}}{2} \tag{1.28}$$

$$m_\gamma = 0, \tag{1.29}$$

Similar to the generation of the gauge boson masses, spontaneous symmetry breaking also generates the masses of the fermions. The coupling between the Higgs field and massless quark and lepton fields are introduced by the Yukawa interactions. Fermions acquire a mass

proportional to u. A Higgs boson has recently been verified experimentally at the LHC. Its mass, $M_{Higgs} = \sqrt{2\lambda u}$ can then be probed experimentally since the λ parameter is entirely arbitrary.

The last missing point corresponds to the discussion of the fermion flavor changes. Weak charged currents are the only interaction in the SM that changes the flavor of the fermions: for example, by exchange of a W boson one quark family is transited to an another quark family, or a υ_e is turned into an e^-. The mass eigenstates of the quarks $(\bar{d}, \bar{s}, \bar{b})$ are not identical to the weak states (d,s,b). The relation is described by a mixing matrix defined by

$$\begin{pmatrix} \bar{d} \\ \bar{s} \\ \bar{b} \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.30}$$

The matrix is a complex, unitary matrix and is called *Cabibbo-Kobayashi-Maskawa (CKM)* matrix. The probability for a quark of flavor i to be transformed to a quark of flavor j, emitting a W boson is proportional to $|V_{ij}|^2$.

1.1.3 The Quantum Chromodynamics

Quantum chromodynamics (QCD) is the gauge theory of strong interactions represented by the non-Abelien $SU(3)_C$ symmetry group. Local gauge invariance in this representation leads to the presence of eight massless gauge bosons, called gluons. The C in $SU(3)_C$ stands for color which plays and important role in QCD. The gluons carry different combinations of colors and anti-colors: red, green and blue and corresponding three anti-colour charges. The Lagrangian is given by:

$$\mathcal{L}_{QCD} = \sum_{flavors} \bar{\psi_q}(i\gamma^\mu D_\mu - m)\psi_q - \frac{1}{4} G^a_{\mu\nu} G^{a,\mu\nu}, \tag{1.31}$$

where the ψ_q are the quark fields, $G^{a,\mu\nu}$ is the gauge invariant field strength tensor, the index a corresponds to the gluon-color index and D_μ is the covariant derivative which is given by:

$$\mathcal{D}_\mu = \partial_\mu + i g_s \lambda_a A^a_\mu, \tag{1.32}$$

where g_s corresponds to the coupling constant and A^a_μ denotes the eight (a=1,...,8) vector gluon fields. The field strength tensor is given by:

$$G^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - g_s f^{ABC} A^B_\mu A^C_\nu \tag{1.33}$$

The gluons which have been referred up to this point, carry naturally the strong force. They do hold color charge which allows them to interact with each other and let the force remain constant with increasing distance. The coupling of the strong interaction is given by the following relation:

$$\alpha_s = \frac{g_s^2}{4\pi} \tag{1.34}$$

The coupling has been found to have a dependence proportional to the momentum scale of the interaction, denoted as Q^2, it is a *running coupling* constant of the QCD as, results from a renormalization process. The value for α_s is given by:

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu_r^2)}{1 + \frac{1}{12\pi}\alpha_s(\mu_r^2)(11N_C - 2N_f)ln\frac{Q^2}{\mu_r^2}}, \tag{1.35}$$

With N_C being the number of color charges ($N_C = 3$), N_f is the number of involved quark flavours and μ_r^2 is a reference mass parameter directly related with the choice of the renormalization scale. The following equation introduces a mass scale Λ_{QCD}, a scale for which the strong coupling constant is infinite:

$$ln(\Lambda_{QCD}^2) = ln(\mu_r^2) - \frac{12\pi}{11N_C - 2N_f\alpha_s(\mu_r^2)}, \tag{1.36}$$

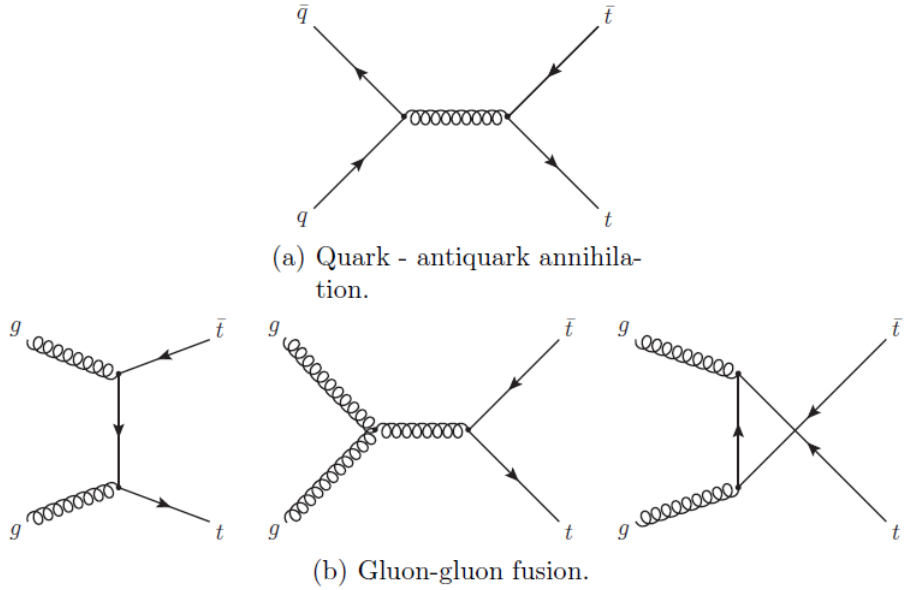
and the eq. 1.34 becomes:

$$\alpha_s(Q^2) = \frac{12\pi}{(11N_C - 2N_f)ln\frac{Q^2}{\Lambda_{QCD}^2}} \tag{1.37}$$

For high enough values of Q^2, the distance of the interaction shrinks, this is called the *asymptotic freedom*, however the strength of strong interaction increases with the Q^2 decreases, which is called as the color confinement and leads to the fact that free quarks and gluons do not exist.

1.2 The top quark

The top quark was discovered in 1995 at the Tevatron by the CDF [22] and DØ [23] Collaborations. Its large mass of about 172.5 GeV/c^2 makes it the heaviest elementary particle observed to date. The Top quark acts differently than other elementary fermions, as it has a lifetime of $\sim 5 \times 10^{-25}s$, it decays before it hadronises, passing directly its momentum and spin information to its decay products. The study of its production, decay modes and properties allow to understand and test the Standard model description of this



(a) Quark - antiquark annihilation.

(b) Gluon-gluon fusion.

Figure 1.3 – Feynman diagrams for top quark pair production processes: (a) quark-antiquark diagram, (b) gluon-gluon scattering diagrams.

quark. Furthermore, the top quark mass is a crucial parameter for a deeper understanding of the electroweek interaction and a powerful probe for new phenomena beyond the standard model.

1.2.1 Top quark production at hadron colliders

Top quarks can be produced via both electroweak and strong interactions. The dominating production mechanism at hadron colliders is through the strong interaction which is also the focus of this thesis. In this case, at the LHC, the top quark is produced in pairs of a top and an anti-top quark $t\bar{t}$ where: ∼80% through *gluon-gluon* fusion $gg \rightarrow t\bar{t}$ and ∼20% through *quark-antiquark* annihilation $q\bar{q} \rightarrow t\bar{t}$. The Feyman diagrams for $t\bar{t}$ pair production processes are shown in Fig 1.3.

At the theoretical scale, the calculation of the production cross section of $t\bar{t}$ from proton-proton interaction is a substantial challenge for QCD calculational techniques. Precise measurements of $\sigma_{t\bar{t}}$ are sensitive to the parton distribution function (PDF) [24] of the proton, the strong coupling constant $\alpha(s)$, the top-quark mass $m_{t\bar{t}}$, and potential enhancements due to physics beyond the Standard Model. Thus it can be written with the following form

$$\sigma_{pp\to t\bar{t}}(s, m_{top}) = \sum_{ij=q,\bar{q},g} \int dx_i dx_j f_i(x_i, \mu_f^2) f_j(x_j, \mu_f^2) . \hat{\sigma}_{ij\to t\bar{t}}(\hat{s}, m_t, \mu_f, \mu_r, \alpha_s) \quad (1.38)$$

Here the sum runs over all quarks and gluons participating to the hard process (3.1.1). pp are the initial state hadrons (proton-proton for LHC), s is the square center-of-mass energy for pp collisions, μ_f and μ_r are the factorization and renormalization scales, respectively. $\hat{s} = x_i x_j s$ is the partonic center-of-mass energy and $f_{i(j)}(x, \mu^2)$ is the parton distribution function (PDF). A PDF describes the probability density of a parton of flavour i to carry a fraction of proton momentum x.

The preliminary resulting calculations of the total cross-section performed by ATLAS [2] are:

$$\sigma_{t\bar{t}} = 803 \pm 8(stat) \pm 27(syst) \pm 45(lumi) \pm 12(beam) pb \quad (1.39)$$

Where the calculation of this inclusive top quark pair production cross-section $\sigma_{t\bar{t}}$ have been made with a data sample of 3.2 fb^{-1} of proton-proton collisions at a centre-of-mass energy of $\sqrt{s} = 13$ TeV collected with the ATLAS detector for a top mass, $m_t = 172.5$ GeV. The theoretical calculation based on next-to-next-to leading order (NNLO) and next-to-next-to-leading logarithmic (NNLL) gives a prediction of $\sigma_{t\bar{t}} = 832^{+40}_{-46} pb$, including PDF, $\alpha(s)$ and QCD scale uncertainties. This value has been calculated using the top++ 2.0 program [25]. The PDF and $\alpha(s)$ uncertainties of ±35 pb were calculated using the PDF4LHC prescription [26] with the MSTW2008 68% CL NNLO [27, 28], CT10 NNLO [29] and NNPDF2.3 5fFFN [30] PDF sets, and added in quadrature to the scale uncertainties to give a final relative uncertainty of $^{+20}_{-29}$pb. The cross-section at $\sqrt{s} = 13$ TeV is predicted to be 3.3 times larger than the cross-section at $\sqrt{s} = 8$ TeV. All results are in agreement with the SM predictions above. Figure 1.4 shows the result for $\sigma_{t\bar{t}}$ together with the most precise ATLAS results at $\sqrt{s} = 13$, 8 and 7 TeV [31, 32]. The data are compared to the NNLO+NNLL predictions as a function of the center-of-mass energy $\sqrt{s}$.

In addition to the top quark pairs production, top quarks can also be produced singly within the standard model via the weak interaction in a process commonly referred to as single top production. There are three mechanisms responsible for the production of single top quarks: an *s-channel* called also a $t\bar{b}$ production $q\bar{q'} \longrightarrow t\bar{b}$ involving the decay of a virtual W-boson, a *t-channel* which proceeds via a virtual W boson exchange, and a *Wt-channel* where the top quark is produced with a real W boson by excitation of a b sea quark. Fig. 1.5 shows the leading-order Feynman diagrams of the single top production. All the three channels are produced via the Wtb vertex.

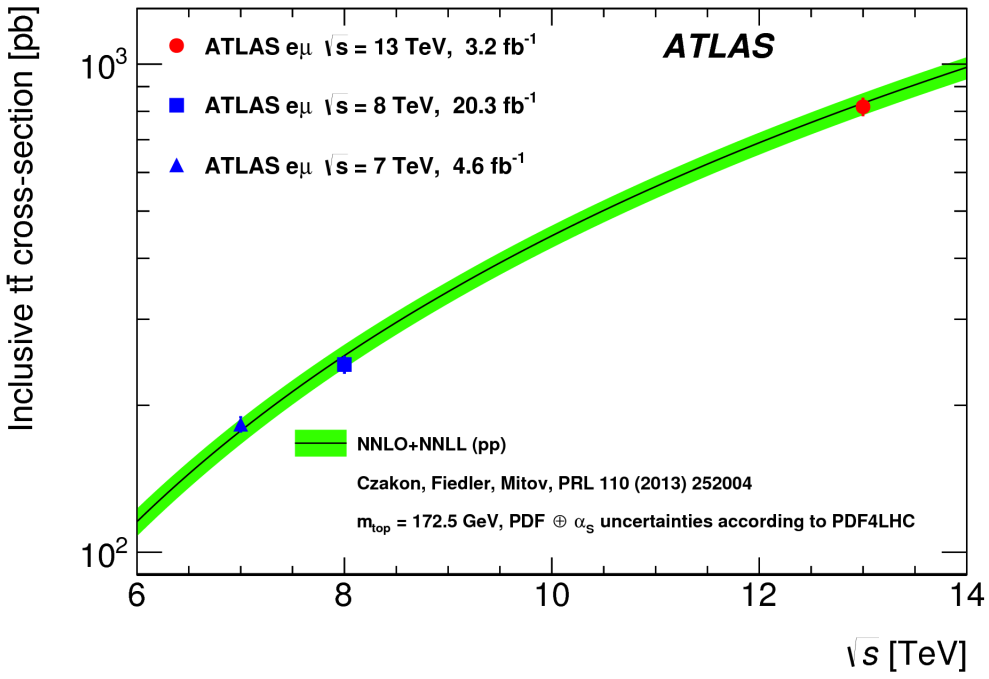
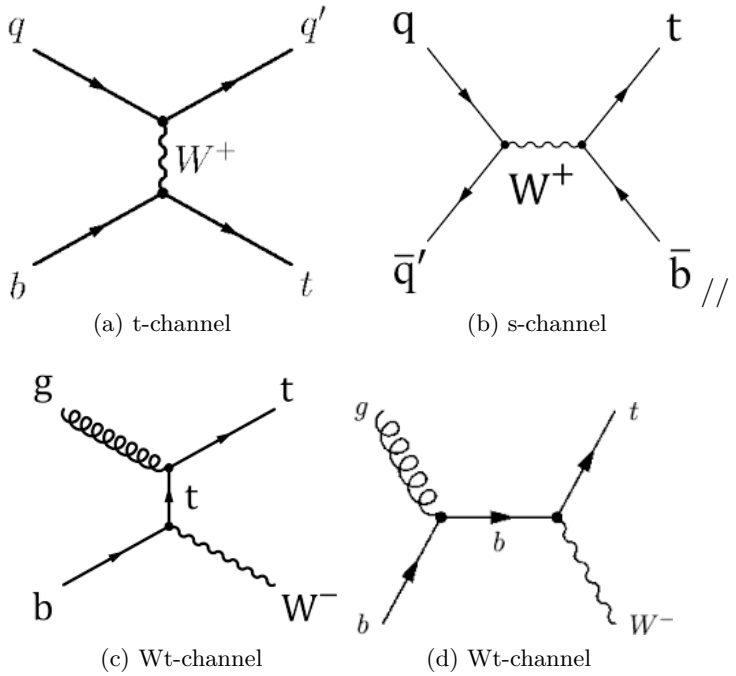


Figure 1.4 – : Cross-section for $t\bar{t}$ pair production in pp collisions as a function of centre-of-mass energy. ATLAS results at $\sqrt{s} = 13$, 8 and 7 TeV are compared to the NNLO+NNLL theoretical predictions. Fig From [2]



(a) t-channel (b) s-channel

(c) Wt-channel (d) Wt-channel

Figure 1.5 – Feynman diagrams at leading order of single top quark production at the LHC.

At the LHC, the predominant mechanism for single top production is the *t-channel*, followed by the *Wt-channel* production, due to the large initial state gluon contribution at the LHC centre-of-mass energy, while the contribution from the *s-channel* is negligible. The production of top and antitop quarks in the t- and s-channel is asymmetric due to the fact that the colliding particles are protons. Both channels require a quark to annihilate with an antiquark (sea quark from proton) and the proton contains two up quarks compared to one down quark. Given the charge of the up quark ($+\frac{2}{3}e$), it is more likely for a W^+ to be produced which decays into a top and $\bar{b}$ quark.

Since single top production proceeds via a weak process, the cross-section depends on the value of the CKM matrix element V_{tb} without assuming the unitarity of CKM matrix and the existence of only three quark families. Therefore, measurements of the single top cross-section provide a test of the unitarity of the CKM matrix. The direct measurements are in good agreement with the theoretical predictions with top quark mass of 172.5 GeV.

1.2.2 Top-quark decays

The top quark decays through the electroweak intercation, predominantly to $t \to Wb$ [2]. The Top quark width is given by [33]:

$$\Gamma_t = \frac{G_F m_t^2}{2\pi\sqrt{2}} |V_{tb}|^2 (1-y)^2 (1+2y)[1 - \frac{2\alpha(s)}{3\pi} f(y)] \tag{1.40}$$

where $y = (\frac{M_W}{m_t})^2$ and $f(y) = \frac{2\pi^2}{3} - 2.5 - 3y + 4.5y^2 - 3y^2 lny$

The average for the top decay width is $\Gamma_t = 1.5$ GeV, assuming a top mass value of $m_t = 172.5$ GeV and $\alpha(s)(M_Z) = 0.118$.

The W boson produced in the decay, in turn decays either leptonically, into a lepton and a neutrino, or hadronically, into quarks, producing jets in the detector. Thus the final states corresponding to a $t\bar{t}$ pair can be classified in three different channels, based on the decays of the two W bosons originating from the decays of the top and the antitop quarks:

- $t\bar{t} \to llννbb$: The dilepton channel, which assumes that both W bosons decay leptoncally, the final state consists of two b quarks, two leptons and two neutrinos.
- $t\bar{t} \to bbjjjj$: The all-hadronic channel, which assumes that both W bosons decay hadronically, the final state consists of six quarks.

[2] In the SM, the only possible top decays are: $t \longrightarrow bW^+$, $t \longrightarrow sW^+$ and $t \longrightarrow dW^+$. Due to the CKM matrix elements, the Ws and Wd final states are highly suppressed relative to the Wb final state

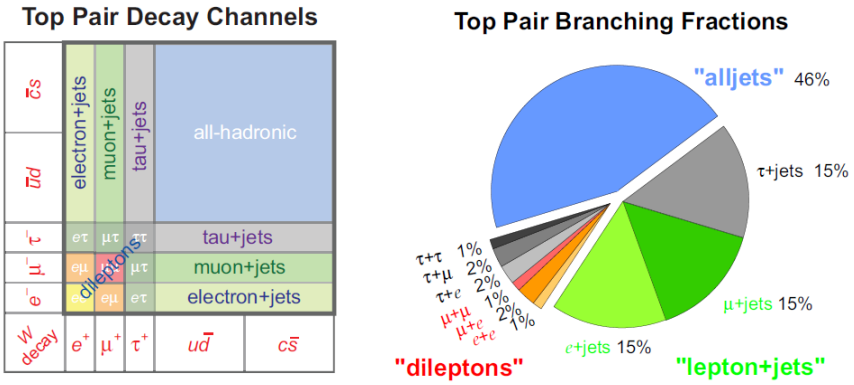


Figure 1.6 – Top pair channel decays and their approximate branching ratios.

This thesis concentrates on the $\ell+jets$ channel [3], where one W boson decays into quarks, while the other decays leptonically and can result in any of the three flavors ($e\nu_e, \mu\nu_\mu or \tau\nu_\tau$). The final state consists of four quarks, one lepton and one neutrino. For this analysis, only the $e+jets$ and $\mu+jets$ final states are considered. The reason for making distinction for the channels with τ-leptons in their final state, is because the τ decays fast, either hadronically ($\sim 65\%$) producing jets, or leptonically via, $\tau \to e\nu_e\nu_\tau$ (17.8%) or $\tau \to \mu\nu_\mu\nu_\tau$ (17.2%). In the first case, it is difficult to distinguish the τ from a jet at the experimental level, while in the second case it is almost impossible to distinguish it from the original W boson decay. The corresponding example of each type of decay and branching ratios are schematically illustrated in Figure 1.6.

The top quark is the only quark whose properties can be extracted from the final particles without the effect from hadronization. The final state decay products carry information about the top quark spin, thus, allows to observe directly the top polarization and measure many observables via the spin correlation of the top quark pair production. The top quark study can be used to test the Standard Model prediction and find new physics more directly.

1.3 Motivation for theories beyond the Standard Model

The SM has proven to be a very successful model, that was able to predict the existence of many particles and describing their properties and processes within a rather simple theoretical framework. Unfortunately, the SM is widely considered an incomplete theory due to its inability to explain the presence of some evident natural phenomena. Although several observations in cosmology revealed that only about 4.9% of the matter in the Universe is made up of the fundamental particles described in the SM, while the remaining of matter

[3] the single lepton channel is chosen , because it provides ample statistics, manageable backgrounds, and a clear signature with a charged lepton in the final state.

can only be detected by its gravitational interactions, dark matter (26.6 %) and dark energy (68.5 %) which is the unknown energy to expand the universe, carry the remaining parts. Indirect evidence for the existence of the **Dark Mater** was attributed to astrophysicist Fritz Zwicky to describe an unknown, invisible, non-luminous and non-absorbing [34] material but has observable gravitional effects, such as the galactic rotation curves [35] and gravitational lensing [36], as well as precision measurements of the power spectrum of the cosmic microwave background (CMB) radiation [37]. Of the known elementary particles, the SM neutrinos have long been considered viable candidates for DM as it was produced in abundance during the *Big Bang* and is the only known particle which doesn't interact via the strong or electromagnetic force. Current constraints on the neutrino relic density, derived from the upper limits on the neutrino masses from direct measurements [34] as well as CMB precision measurements [37], and the fact that neutrinos are relativistic particles, rule them out as DM candidates [38].

One more important puzzle comes from the observed matter-anti-matter asymmetry. The creation of the universe and thermal equilibrium should have resulted in the production of an equal amount of matter and antimatter. However, at the present day, there is a negligible amount of remaining antimatter in the visible universe and There is no sufficient source of baryon asymmetry in the SM to explain why we see only matter in the observable Universe (the only source of anti-matter being from collisions, either via cosmic rays entering the atmosphere or via man-made accelerators). There are also a number of intrinsic questions that cannot be adequately addressed by the SM alone. For example, the discovery of the Higgs boson at the level of 126 GeV mass, completed the model, and solved the question of electroweak symmetry breaking, however, this comes at the expense of the challenge of explaining the mass of the Higgs itself. Since the physical mass of this elementary scalar is not protected against divergent contributions from virtual particle loops and there is no natural bound for the Higgs mass parameter in the Higgs sector Lagrangian and the only way to obtain the physical observed mass, $m_h \approx 126GeV$ is to postulate unnatural cancellations of order the Planck mass, $M_{Pl} \sim 10^{19}GeV$ in the process of renormalization. Notably, the dominant diagram in the higher order divergent contributions to the Higgs mass come from top quark loops and any solution to the hierarchy problem would have to treat top quarks preferentially. From the aesthetic perspective, another unsatisfactory SM feature comes from the large number of free parameters, the arbitrary values which are not predicted in the SM and can only be constrained by the experimental observations. These include masses and mixing angles, with no explanation as to their disparate numerical values. Finally, the SM does not incorporate gravity, therefore integrating a quantum gravity in with the other forces is a necessary step before a model can be considered truly complete.

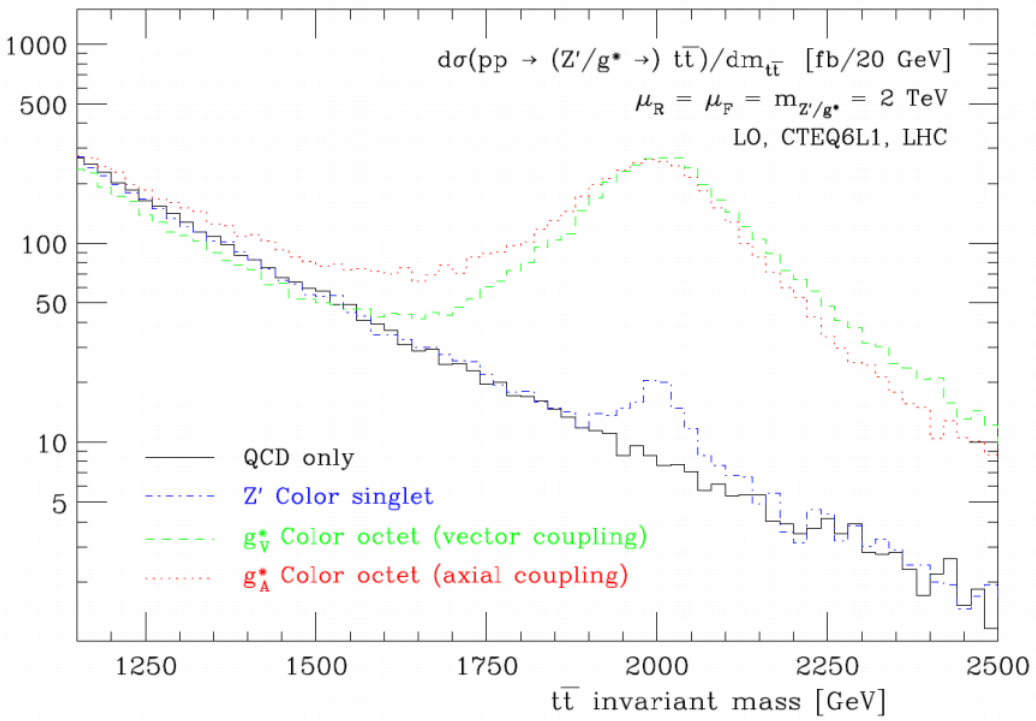


Figure 1.7 – Invariant $t\bar{t}$ spectrum for different resonances at 2 TeV. The disributions show a s-channel Z' color singlet vector boson in blue and color octet vector (axial) bosons coupling in green (red). The y-axis is $d\sigma(pp \to (Z'/g^* \to)t\bar{t})/dM_{t\bar{t}}$ in units of fb/20 GeV. [3]

1.4 The top-antitop invariant mass spectrum in the BSM

The top-antitop invariant mass ($M_{t\bar{t}}$) defined in equation (1.41) forms the discriminant by which deviations from the SM prediction are identified in a variety of models.

$$M_{t\bar{t}} \equiv \sqrt{(p^{top} + p^{antitop})^2} \tag{1.41}$$

In the above expression, p^{top} and $p^{antitop}$ are the top and antitop four-momenta and the $M_{t\bar{t}}$ is the key observable. The top quark mass offers the possibility of studying a wide range of some heavy BSM particles that would primarily decay into a top-antitop quark pair. The Topcolor Z' (discussed in the following section) and KaluzaKlein (KK)-gluon/graviton predicted by certain Randall-Sundrum (RS) scenarios [39] [40] are examples of such new BSM particles that would reveal themselves as resonances in the $M_{t\bar{t}}$ spectrum.

Figure 1.7 and 1.8 show the effects of a color singlet/octet spin-1 particle and a series of KK graviton modes on the $t\bar{t}$ invariant mass spectrum respectively. The mass and width of these resonances are model dependent and somewhat arbitrary at this stage.

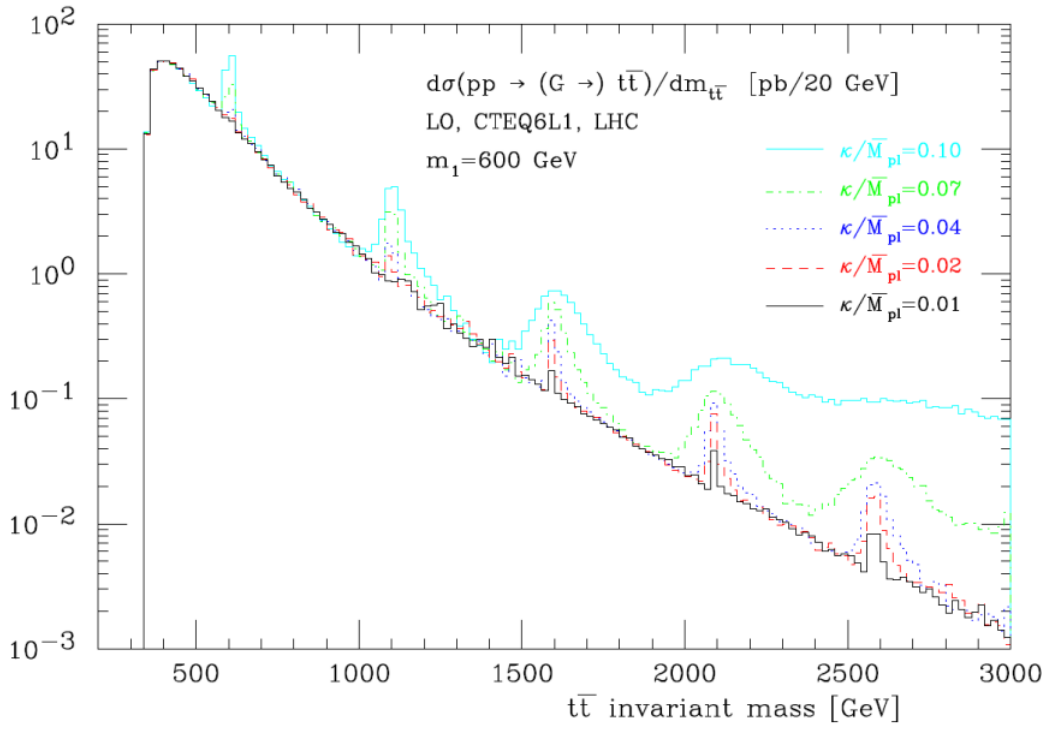


Figure 1.8 – Invariant $t\bar{t}$ spectrum in the presence of a s-channel KK-gravitons. The first mass ($m_1 = 600$ GeV) is arbitrary while the masses of the other KK-gravitons follow from the zeros of the Bessel function J_1(x). The y-axis is $d\sigma(pp \to (G \to)t\bar{t})/dM_{t\bar{t}}$ in units of pb/20 GeV. [3]

1.5 Benchmark model: Spin-1 Colour-Singlet

Top quark is the massive particle ever discovered, significantly heavier than the other quark on its generation, the bottom quark. Motivated by this property, many models assume that the top quark plays a role in the dynamics of electroweak symmetry breaking, which remains a mystery and predict new particles or interactions that couple preferentially to the third generation and in particular to the top quark. One such class of models is based upon topcolor [16]. These non-standard models extend the strong, hypercharge, or weak interactions in such a way that the new groups spontaneously break into their standard-model subgroup at some scale: $SU(3)_h \times SU(3)_l \to SU(3)_C$, $SU(2)_h \times SU(2)_l \to SU(2)_W$, and $U(1)_h \times U(1)_l \to U(1)_Y$, where h represents the third (heavy) generation and l the first two (light) generations. One of the possible topcolor scenarios is topcolor-assisted technicolor [16, 41], recently found to be capable of accounting for top quark problems. It generates a top-antitop pair with high mass from the decay of a massive color-singlet Z-like bosons, denoted Z'_{TC2}. The resonance of this model has a narrow width of 1.2% of the Z' mass ($\Gamma/m = 1.2\%$) which is much smaller than the experimental resolution [4]. There are four variants of the topcolor Z' and the benchmark scenario chosen in this thesis is reffered to as Model IV [41], a leptophobic, topophyllic boson. The dominant part of the interaction Lagrangian for this model is:

$$\begin{aligned}\mathcal{L}_{IV} = (\frac{1}{2}g_1 cot\theta_H)Z'^{\mu}(&\bar{t}_L\gamma_\mu t_L + \bar{b}_L\gamma_\mu b_L + f_1\bar{t}_R\gamma_\mu t_R + f_2\bar{b}_R\gamma_\mu b_R \\ &- \bar{u}_L\gamma_\mu u_L - \bar{d}_L\gamma_\mu d_L - f_1\bar{u}_R\gamma_\mu u_R - f_2\bar{d}_R\gamma_\mu d_R)\end{aligned} \tag{1.42}$$

and the required settings are the following: $f_1 > 0$ (attractive $t\bar{t}$ channel) and/or $f_2 < 0$ (repulsive $b\bar{b}$ channel), where f_1 and f_2 define the couplings to up- and down-quark, respectively. Also, $cot\theta_H >> 1$ to avoid fine-tuning, since it controls the production cross-section and width.

This specific Z' model would therefore be first observed via its decay to top quarks rather than via the dileptonic decay mode, which has fewer background processes and is usually the favoured search channel for similar BSM resonances. All searches performed by the ATLAS, CMS, CDF and $D\emptyset$ collaborations [22, 23] to date, have set lower limits on the allowed mass for such resonance and the strongest mass limit for this boson have been set at $m_{Z'_{TC}} > 2.3$ TeV.

The procedure used in this search is very general, it is not tailored to any particular

[4] The measurement of the experimental resolution depends on the method of reconstructing the $t\bar{t}$ system. However, it is considered to be approximately 6 - 12 % of the resonance mass [42]

model or subsequent decay signature. However, the benchmark model discussed above, is considered to quantify the sensitivity of such search on a number of different resonance mass points in the range 0.4 - 5 TeV. Details on the chosen mass points, the corresponding widths and production cross-sections as well as the simulation of the signal processes are given in Chapter 5.

1.6 Summary

In this chapter the theoretical framework was presented in order to consider a possible Beyond Standard Model as an alternative theory to search for phenomena which cannot be explained within the Standard Model theory. In the context of topcolor assisted technicolor model which can generate a top-antitop pairs. The topcolor model also predicts a Z', which represents a novel class and has predicted cross sections large enough to be experimentally accessible at the Large Hadron Collider (LHC) at CERN. Such Z' resonances couple strongly to the third generation of quarks, and have no significant couplings to the leptons. They are, therefore, leptophobic and topophyllic. The use of this model is solely to aid comparison with results from similar searches and this particular search strives to be model-independent. Several channels are possible to discover this new particle. In this thesis I am focusing my study on lepton+jets channel where a b jets and missing transverse momentum can be found in the final state. This analysis is promising because the SM and the detector backgrounds are low when compared to other channels with no leptons (for the same set of cuts). The entire $\sqrt{s} = 13$ TeV data is considered for this analysis.

Chapter 2

The ATLAS Experiment at the LHC

The access to a whole area of physics at the high energy frontier, the study of basic phenomena that have shaped our universe since the beginning of time and further measurements of the Higgs boson's properties are among the fondumental goals of the experiments located on the Large Hadron Collider (LHC), at CERN. Inside the LHC, two high-energy proton beams circulate in opposite directions and collide up at four particle detectors: ATLAS, described in Sec 2.2, CMS [17], ALICE [43] and LHCb [44]. ATLAS and CMS are multi-purpose detectors, ALICE dedicated to heavy-ion physics and LHCb specializes in B-physics. The ATLAS (A Toroidal LHC ApparatuS) experiment has been conceived to study new phenomena in particle physics at the Terascale, and pursuit a broad and challenging physics program including the search of supersymmetric particles, extra dimensions and dark matter candidates.

This chapter is meant as a quick description of the experimental setup used in this search for the reader who is not familiar with the detector components and its nomenclature. Section 2.1 describes the CERN accelerator complex and the Large Hadron Collider. Section 2.2 briefly introduces the ATLAS detector and its sub-detectors, used to collect the data used for this thesis. Section 2.3 is devoted to describe the measurement of the delivered luminosity. The pile-up effect at the LHC is explained in Section 2.4. In the last sections 2.5, 2.6 and 2.7. An overview of the ATLAS computing and analysis model is given, as the author was directly involved in these activities.

2.1 CERN and the Large Hadron Collider

CERN the European Organization for Nuclear Research, is one of the world's largest physics laboratory in the world, it is located near Geneva, in the border region between France and Switzerland. The LHC is the most recent step in a long succession of particle accelerators approved by the CERN council in 1994, and delivered its first collisions in November 2009, It has taken almost 20 years of design, construction and testing, allowing to proton-proton and heavy-ion collisions to take place at a design maximum center of mass energy of 14 TeV. The operating center-of-mass energies in proton-proton collisions have so far been 13 TeV in 2015 and onwards as the second LHC run (Run-2). The corresponding beam parameters are presented in Tab 2.1.

The LHC is a circular proton-proton (pp) collider consisting of a 27 km ring of strong superconducting magnets and accelerating cavities, where two beams of protons circulate in separate vacuum tubes in opposite directions. The beams are kept on their circular trajectory by 1232 superconducting magnets, and accelerated to their final energy using 16 superconducting radio Frequency (RF) cavities. The system is cooled to -271°C: two degrees above absolute zero, to conduct efficiently electricity without resistance or loss of energy. Protons are accelerated in several stages in the pre-accelerators chain in the LHC, which are both in the CERN accelerator complex, shown in Figure 2.1. Protons from the ionization of hydrogen atoms are first accelerated to 50 MeV in a Linear Collider (LINAC2). Then the beam is injected into the Proton Synchrotron Booster (BOOSTER), which pushes the protons to reach 1.4 GeV, followed by the Proton Synchrotron (PS), which speed up the beam to 25 GeV. Protons are then sent to the Super Proton Synchrotron (SPS), where they are accelerated to 450 GeV. The protons are finally injected into the LHC.

Parameter	2015
Energy (TeV)	6.5
The machine instantaneous luminosity $\mathcal{L}(10^{34}cm^{-2}s^{-1})$	0.5
Number of particles per bunch	1.15 $\times 10^{11}$
Number of bunches	2244
Revolution frequency	11245 Hz
Bunch spacing (ns)	25

Table 2.1 – Run parameters of the LHC in 2015. These values are typical values for the runs with best achieved performance.

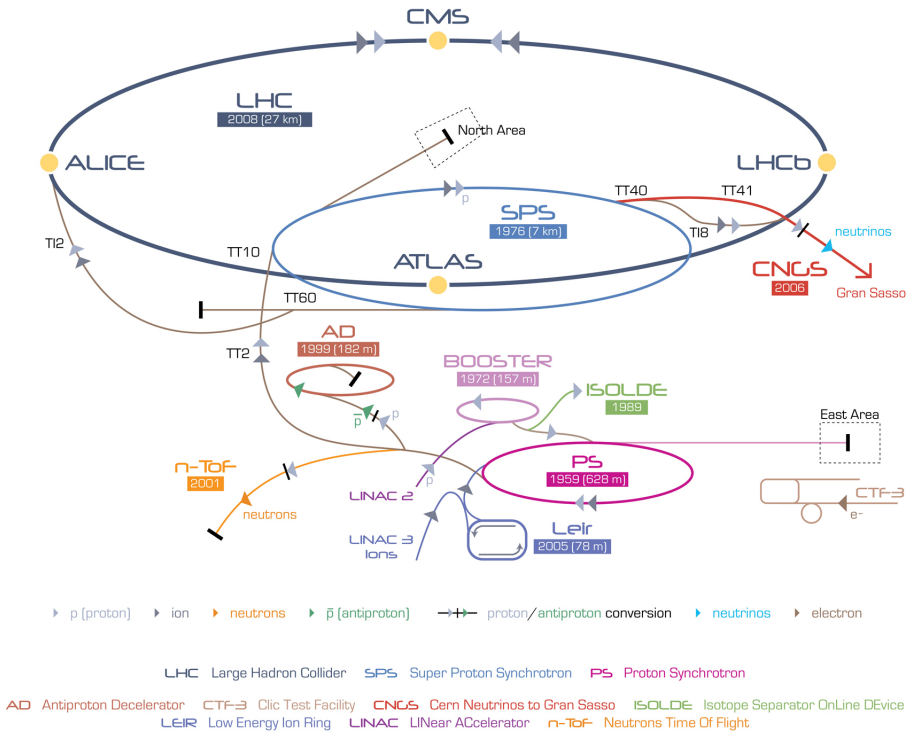


Figure 2.1 – The accelerator complex at CERN

2.2 The ATLAS Detector

ATLAS (A Toroidal LHC ApparatuS) is a general purpose detector located at Interaction Point 1 at the LHC, it has been built to perform well with an expected peak luminosity of $L = 10^{34} cm^{-2} s^{-1}$ (see Sec 2.3), and then, reach a wide physics program that includes precise electroweak interaction measurements, searches for supersymetric signatures and for physics beyond the Standard Model. The Atlas detector has an approximate cylindrical symmetry and organized in a central barrel and two end-caps that close either end. In the barrel, the active detector elements form cylindrical layers around the beam pipe, while in the end-caps, they are organized in disks or wheels. It is 46 m lenght, 25 m height, and has a weight about $7,000$ $tons$. It performs direct measurement of the trajectory and/or energy of different particles, which are reconstructed using the complimentary information from four specialized sub-systems: the magnet system (Sec 2.2.2), the inner detector (Sec 2.2.3), the calorimeter (Sec 2.2.4) and the muon spectrometer (Sec 2.2.5). These are organized in the cylindrical layers around the interaction point as can be seen in the overall layout of the ATLAS detector in Figure 2.2.

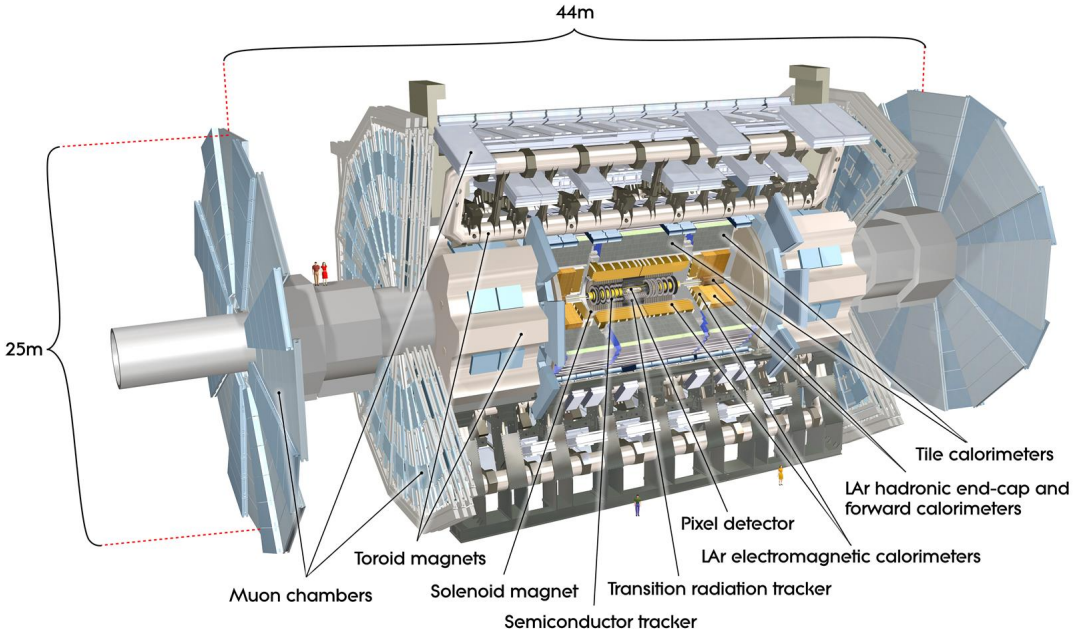


Figure 2.2 – Cut-away view of the ATLAS detector

2.2.1 The ATLAS Coordinate System

To describe directions and positions of objects in the detector and collision products, it is common to use a cylindrical or spherical system of coordinates, besides a standard Cartesian coordinate system. The x-axis is pointing to the center of LHC ring, the y-axis is pointing upwards, and the z axis is directed along the beam line. The transverse plane is defined as the xy-plane as it stands perpendicular to the direction of the colliding beams. The azimuthal angle ϕ, is the angle in the transverse plane measured in the $x-y$ plane with a range of [-π, π], while the polar angle θ is the angle from z-axis. the pp collisions occur along the z-axis, one defines a rapidity as

$$y = \frac{1}{2} ln(\frac{E + p_z}{E - p_z}) \tag{2.1}$$

which is invariant under boosts along the z-axis. For *relativistic* particles, this rapidity is identical to the pseudorapidity, which is simpler for physics analysis and can be defined as

$$\eta = -ln(tan\frac{\theta}{2}) \tag{2.2}$$

The distance ΔR between two objects is an other important Lorentz-invariant quantity and is quantified in the $\eta - \phi$ space as

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \text{ with } \Delta\phi \in [0, \pi] \tag{2.3}$$

The transverse momentum p_T and the kinematics of particles are parametrized as $\vec{p} = (p_T, \eta, \phi)$ with

$$p_x = p_T \cdot \cos\phi \tag{2.4}$$

$$p_y = p_T \cdot \sin\phi \tag{2.5}$$

$$p_z = p_T \cdot \sinh\phi \tag{2.6}$$

$$p - |p| = p_T \cdot \cosh\phi \tag{2.7}$$

The transverse energy E_T and its componnents are defined in the $x - y$ plane as

$$E_T = \frac{E}{\cosh\eta} = E \cdot \sin\theta \tag{2.8}$$

$$E_x = E_T \cdot \cos\phi \tag{2.9}$$

$$E_y = E_T \cdot \sin\phi \tag{2.10}$$

$$\vec{E_T} = \begin{pmatrix} E_x \\ E_y \end{pmatrix} \tag{2.11}$$

2.2.2 The ATLAS Magnet system

The magnet system is an integral part of the ATLAS detector which is a very crucial feature for identification and momentum measurement of the charged particles. It is the largest superconducting magnet in the world with 20 m in diameter and 26 m in length. It has a total weight of 1300 $tons$, is cooled by liquid helium at 4.8 K and stores a magnetic energy of 1600 MJ. The magnet system includes four superconducting magnets: the central solenoid (CS), the barrel toroid (BT) and the two end-caps toroids (ECT). figure 2.3 shows a sketch of the four parts of the ATLAS magnet system and Table 2.2 summarizes its main parameters.

The solenoid magnet: The solenoid magnet is located around the ATLAS Inner Detector and installed along the beam axis, inside of the calorimeter structure. It is a 5.8 m long with an inner and outer diameters of 2.46 m and 2.56 m respectively. It provides an axial field with a strength of 2 T which bends charged particles around the direction of the incoming LHC beams. The solenoid consists of one coil magnet with 1154 turns, operating at a temperature of approximately 4.5 K. It was designed to be lightweight and to keep the

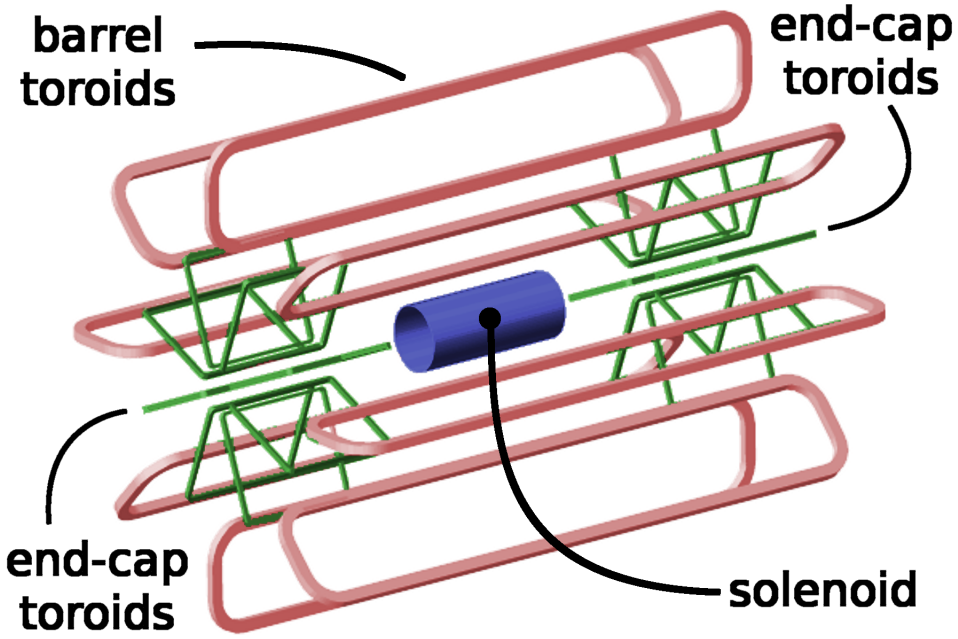


Figure 2.3 – Overview of the magnet system: the solenoid and the toroids magnets are shown.

amount of material, between the calorimeters and the intercation point as low as possible. A minimum radial thickness of 45 mm (corresponding to 0.66 radiation lengths at normal incidence angles) was established in order to achieve a high field and grant a good reliability.

The toroid magnets: The toroid magnet system is made of three individual parts: the barrel toroid, which is situated in the central part of the detector forming a cylindrical volume with an axis that coincides with the beam axis, and the two end-cap toroids, each at one end of the detector. The toroids are made of eight independent coils, arranged with an octagonal symmetry and are located behind the calorimeters and are placed such that a magnetic field with a strength from 3 to 6 T (depending from the pseudorapidity) penetrates all sections of the Muon Spectrometer. The toroid magnet system has an open structure to minimize the contribution of multiple scattering to the momentum resolution; most of the coils is cooled by liquid Helium at 4.5 K.

2.2.3 The ATLAS Inner Detector (ID) System

The ATLAS Inner Detector (shown in Fig 2.4, 2.5), is located close to the proton beam axis and designed to measure the momentum of charged particles and reconstruct their primary and secondary vertices within the region $|\eta| < 2.5$. It is a 6.2 m in length, and has

Property	Unit	Barrel Toroid	End-cap Toroid	Central Solenoid
Size: Inner Diameter	[m]	9.4	1.65	2.46
outer diametre	[m]	20.1	10.7	2.63
axial length	[m]	25.3	5.0	5.30
Number of coils		8	2×8	1
Mass: Conductor	Tons	118	2×20.5	3.8
Cold mass	Tons	370	2×140	5.4
Total assembly	Tons	830	2×239	5.7
Coils: Turns/coil		120	116	1154
Nominal current	KA	20.5	20.5	7.7
Magnet stored energy	GJ	1.08	2× 0.25	0.04
Peak field	T	3.9	4.1	2.6
Conductor: Overall size	mm^2	57× 12	41× 12	30× 4.25
Ratio Al:Cu:NbTi		28:1.3:1	19:1.3:1	15.6:0.9:1
Number of stands		38	40	12
Strand diameter	mm	1.3	1.3	1.22
Critical current @ 5T, 4.2K	KA	58	60	20.4
RRR Al		¿800	¿800	¿400
I/Ic margin @ 4.5K	%	30	30	20
Temperature margin	K	1.9	1.9	2.7
Number of units × lenght	#×m	32×1730	32×800	4×2290
Total lenght	Km	56	2×13	10
Head load: at 4.5K	W	990	330	130
at 60-80 K	kW	7.4	1.7	0.5
Liquid He mass flow	g/s	410	280	7
Cool down time	days	30	4	8

Table 2.2 – The property of ATLAS magnet system.

a radius of 1.2 m. The entire ID is surrounded by a 2 T magnetic field, generated by the superconducting solenoidal magnet. The magnetic field causes the paths of charged particles to bend, therefore allowing the momentum of charged particles to be measured in the ID from the curvature of their tracks. It consists of three independent sub-detector systems: The Pixel detector which is composed of silicon pixels, the SemiConductor Tracker (SCT), which is built up from silicon strip sensors and the Transition Radiation Tracker (TRT), which exploits the transition radiation produced in a gas mixture of Xe, CO_2 and O_2.

Pixel Detector: The Pixel Detector is located closest to the interaction point and covers a pseudo-rapidity range of $|\eta| < 2.5$. The original setup, used in 2012, contains three concentric layers in the barrel and three disks in each end-cap perpendicular to the beam axis. The distance of the barrel layers to the beam line are 50.5 mm, 88.5 mm and 122.5 mm

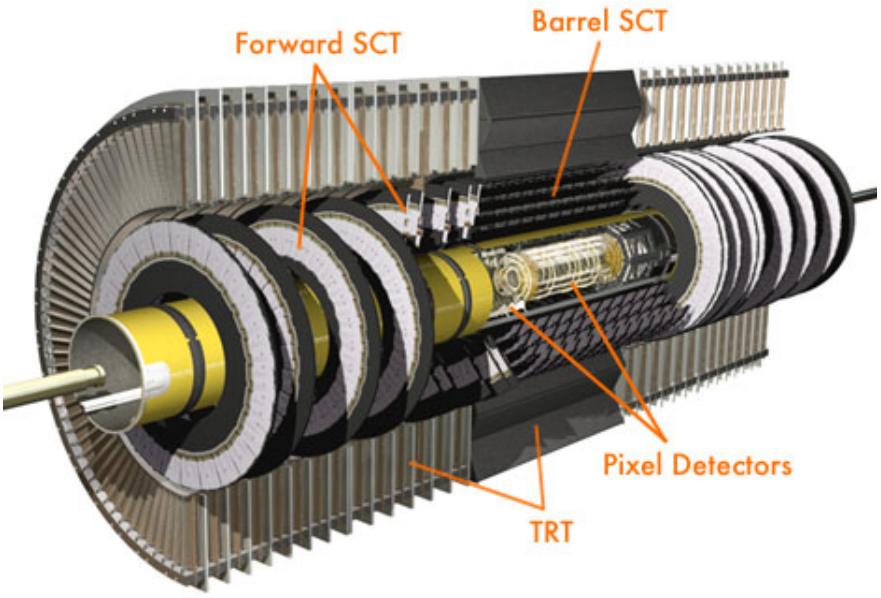


Figure 2.4 – A cut-away view of the ATLAS Inner Detector

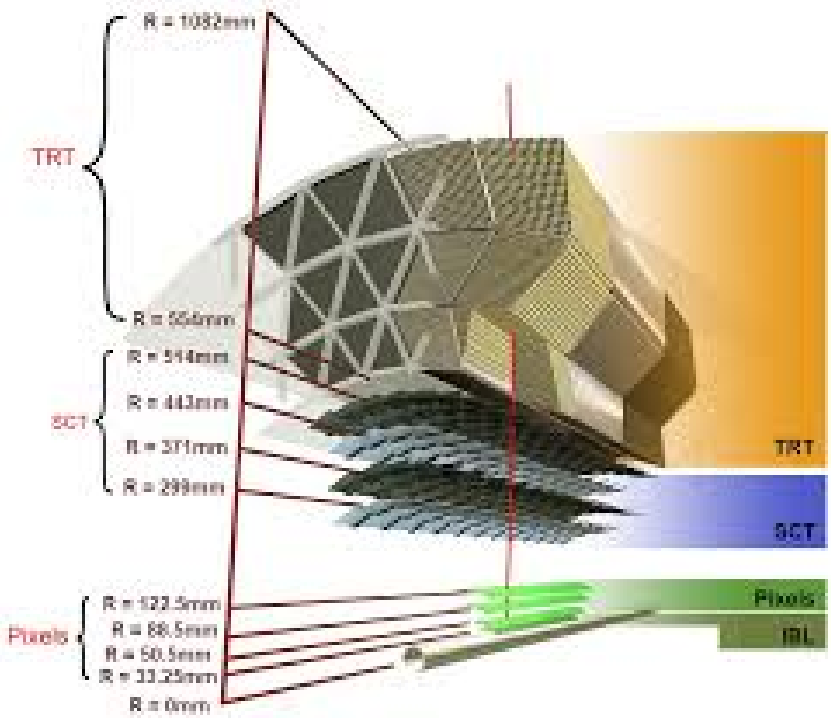


Figure 2.5 – Schematic of the inner detector componnents and their radial distances from the interaction point, including the new insertable B-layer (IBL)

respectively. The layers and disks are equiped with a total of 1744 silicon sensors that are segmented into small rectangles, the pixels. Each pixel measures 40 μm wide in $R-\phi$ and 400 μm long in the z-axis, the silicon has a thickness of 250 μm. The resulting resolution of pixel modules is $\Delta(R\phi) \times \Delta z = 12 \times 66 \mu m^2$.

During the first long shutdown between Run 1 and Run 2, the pixel detector has been complemented with an additional barrel layer, known as the **insertable B-layer, the IBL** [45], installed in-between the pixel detector and a new smaller radius beam pipe. It improves the vertexing capabilities of ATLAS and supplements the previously innermost layer which has been, and will increasingly be, exposed to much radiation. The IBL is composed of 280 modules and consists of 14 staves, instrumented with a mixture of planar and 3D silicon pixel sensors along 332 mm on each side from the center of beam axis. The IBL significantly improves sensitivity to identify b-jets, as well as the measurements of impact parameters of tracks.

Semi-Conductor Tracker (SCT): The Semi-Conductor Tracker (SCT) is situated around the pixel detector. it contains four layers in the barrel region located at 30 cm, 37.3 cm, 44.7 cm and 51.4 cm from the beam-axis and two end-caps each containing nine disks. It covers the range $|\eta| < 2.5$. The barrel region is formed by 2122 modules of silicon micro-strip detectors, while the disks are formed by a total of 1966 modules. The nominal resolution is $\Delta(R\phi) * \Delta z = 17 \times 580 \mu m^2$.

Transition Radiation Tracker (TRT): The Transition Radiation Tracker (TRT) is the last sub-detector of the ID ensemble that a particle traverses. It provides the capability to distinguish electrons from pions through transition radiation emission while traversing this detector. It consists of a large number of polyimid drift tubes or *straws* arranged in a barrel and two endcaps, providing only R-ϕ information, with a spatial resolution of 130 μm per straw. The barrel tubes are 144 cm long, located parallel to the beam axis while the endcaps are 37 cm long tubes arranged radially around the beam axis. The total number of read-out channels is approximately 351 000. Each tube has a diameter of 4 mm where a gold-plated tungsten anode wire of 31 μm diameter is in the middle of each straw held at ground potential and suspended in a gas mixture based on xenon (70%), CO_2 (27%) and O_2 (3%). Once a charged particle traverses the gas it causes ionization and free electrons drift towards the wire. The drift time of the electrons is used to determine the exact incident point of the particle with the tube. A large number of ionizing hits with the gas are typically expected ranging from 30-36 hits per traversing particle.

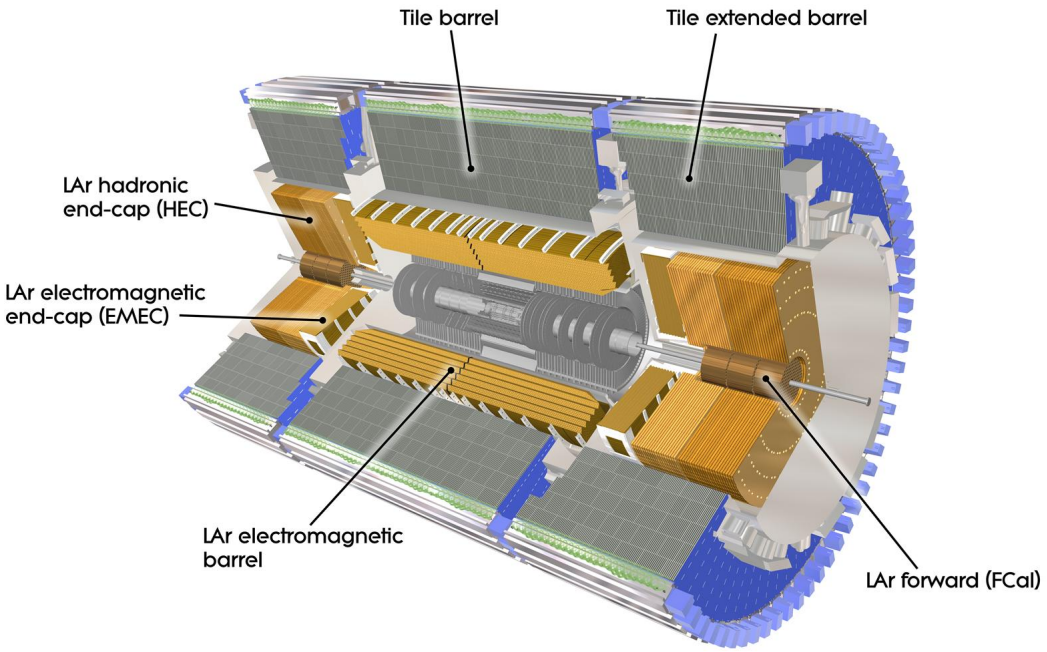


Figure 2.6 – Overview of the ATLAS calorimeter system

2.2.4 The ATLAS Calorimeter System

After The Inner detector, particles enter the calorimeter system, which is designed to measure the energy and direction of charged and neutral objects (electrons, photons, jets and E^T_{miss}) through the absorption of their energy deposits in the calorimeter materials. It is designed to stop most of the particles produced at collisions, except muons and neutrinos. The ATLAS calorimetry systems are realized in ATLAS via a sampling technique: layers of dense absorbers, that stop particles and start the shower alternated with layers of active material, which measure the shower energy. The ATLAS calorimeters cover a range $|\eta| < 4.9$ and an almost full coverage in ϕ and is divided into two part : the electromagnetic calorimeter designed to measure electromagnetic showers, produced by particles that interact primarily through the electromagnetic interaction and the hadronic calorimeter designed for particles that interact mainly through the strong nuclear force. Figure 2.6 shows an overview of the different electromagnetic and hadronic calorimeters of the ATLAS detector.

The ATLAS electromagnetic calorimeter covers a range of $|\eta|$ <3.2 and employs Liquid Argon (LAr) as the active material and lead as absorber. Charged particles traversing the calorimeter create ionisation in the gaps of liquid argon, and the copper electrodes register the signal induced by the ionisation electrons drifting across the gap. The LAr calorimeter

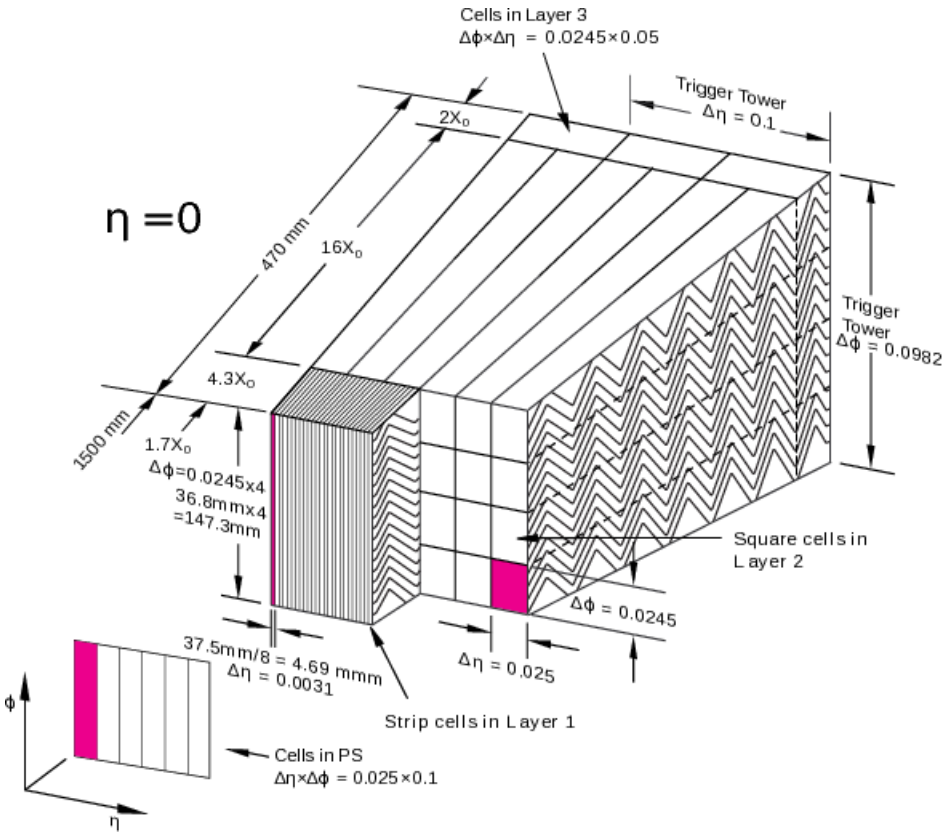


Figure 2.7 – Sketch of LAr barrel module where the different layers and the accordion structure are visible.

can be divided into a barrel part in the region $|\eta| = 1.475$ and two endcaps, consisting of the outer and inner wheels and covering the range $1.375 < |\eta| < 2.5$ and $2.5 < |\eta| < 3.2$ respectively.

The LAr calorimeter barrel has an accordion structure as shown in Fig 2.7, which provides complete coverage in azimuthal angle, ϕ without any cracks and a fast readout of the signal at the rear or at the front of the electrodes. The liquid argon is located in the gaps between the 1024 accordion shaped absorbers, interleaved with readout electrodes. The size of a drift gap on each side of the electrodes in the barrel is 2.1 mm, which correspond approximately to a total drift time of 450 ns for an operating voltage of 2 kV, while the width of gap varies with pseudorapidity in the endcap, so the high voltage needs to vary accordingly to obtain a calorimeter response independent from pseudorapidity. The electromagnetic calorimeters are housed in three cryostats, the liquid argon needs to be kept at a temperature of about 88 K.

The LAr calorimeter barrel has three layers in depth: front, middle and back as shown in Figure 2.8. The front layer is finely segmented along η, being constituted by strips cells of

size $\Delta\eta \times \Delta\phi = 0.0031 \times 0.098$ and has an active depth of almost 4.3 X_0, this geometry provide information on particle identification and allows separation between neutral pions decaying into two photons and jets involving individual photons. The Middle layer, composed by squared towers of size $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$, collects the largest fraction of energy deposited by the shower in its $\approx 16\ X_0$. The back layer only reached by the highest energy particles, and is therefore devised to detect wide clusters with cells having double size in the η coordinate. The front layer is read out at the low-radius side of electrodes, whereas the Middle and back layer are read out at the high radius side of the electrodes. A similar design as for the barrel has been used for the first wheel of the LAr endcaps, which is segmented in three layers with a granularity equal to that of the barrel, while the second wheel (for larger pseudorapidities) has a coarser granularity in η and ϕ, limiting the region devoted to precision physics to $\eta < 2.5$.

In order to improve the precision in the energy reconstruction and correct the energy losses upstream of electrons and photons in the calorimeter, a thin LAr layer (pre-sampler) is used front of the calorimeter at $|\eta| < 1.8$. It is composed of an active layer of LAr with 1.1 cm thickness in the barrel and 0.5 in the endcap. The presampler layer is made of 64 identical azimuthal sectors, each sector is 3.1 m long and 0.28 m wide.

The hadronic calorimeter is a sampling hadronic calorimeter consisting of the tile calorimeter, the liquid-argon hadronic end-cap calorimeter (HEC) and the liquid-argon forward calorimeter (FCal). The Tile calorimeter located in the outward region is subdivided into a central barrel, 5.8 m in length, and two extended barrels, 2.6 m in lenght, covering the range $0 < |\eta| < 0.8$ and $0.8 < |\eta| < 1.7$, respectively. It uses steel as absorber and scintillators as active material, read out by wavelength shifting (WLS) fibers as the active medium. The size of the readout cells is $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ in the region $|\eta| < 2.5$ and $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ for larger $|\eta|$ region. The optical signals read by the WLS fibers are converted into electric signals by photomultipliers (PMTs). It is designed to absorb hadronic showers in the barrel region of the ATLAS experiment.

The end-cap hadronic calorimeter (HEC) is a copper/liquid-argon sampling calorimeter, located outside of the EM calorimeter and covering the region $1.5 < |\eta| < 3.2$. It is structured in two wheels with a flat-plates design. These plates have a thickness of 25 and 50 mm for inner and outer wheel respectively. The active gaps in between the plates all have a thickness of 8.5 mm. The size of the readout cells is $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ for the region $|\eta| < 2.5$ and $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ for larger $|\eta|$ region. The resolution for the hadronic barrel and end-cap is $\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E(GeV)}} \oplus 3\%$.

The forward calorimeter (FCal) is a liquid argon and tungsten/copper detector placed

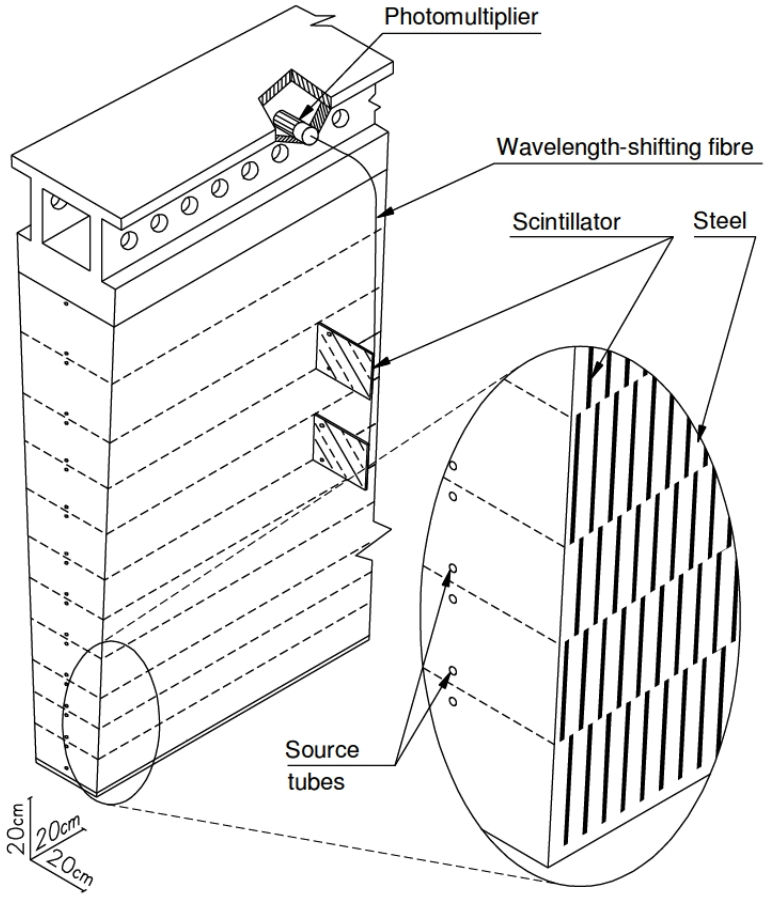


Figure 2.8 – The schematic view of the geometry of a Tile module, showing the placement of the tiles and of the readout components.

4.7 m away from the interaction point and provides energy measurements of electromagnetic and hadronic showers out to $3.1 < |\eta| < 4.9$. The forward calorimeter is constitued by three 45 cm deep modules: one electromagnetic module (FCal1) and two hadronic modules (FCal2 and FCal3) as illustrated in Figure 2.9. In order to optimise the resolution and reduce the heat, the first module is alternating copper chosen as the absorber and LAr and is optimised for electromagnetic shower measurements, while the other two modules (FCal2 and FCal3) are mainly tungsten and LAr used to provide containment and avoid ion build-up and predominantly measure hadronic showers. The resolution for the forward calorimeter is $\frac{\sigma_E}{E} = \frac{100\%}{\sqrt{E(GeV)}} \oplus 10\%$

The electrical signal from the LAr is pre-amplified and then shaped by a bipolar pulse. If the event is retained by the trigger, then five samples of the signal are digitised and saved. The signal is sampled every 25 ns (the LHC bunch crossing period) and for special runs

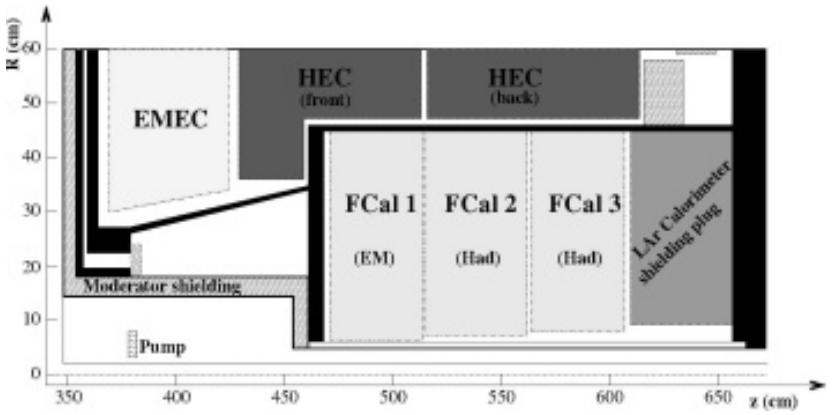


Figure 2.9 – Layout of the calorimeter modules.

more than 5 samples are digitized and recorded. The total energy deposited within a cell in the EM calorimeter, E_{cell}, is then reconstructed from

$$E_{Cell} = F \times \sum_{i}^{N_s} a_i(s_i - p) \tag{2.12}$$

where $N_s = 5$ is the number of samples, p is a constant offset from zero that is determined from dedicated calibration runs, s_i are the samples of the electronic pulse, and a_i are the optimal filter coefficient (OFC) or weights, i = 1,...,n which are calculated by the method with the condition to minimize the noise contribution to the signal. Two sources of noise are foreseen in the calorimeter during operation at LHC: The thermal (electronic) [1] and pileup (or physics) noise. The term F accounts for the cell gain and converts the ionisation current to the total energy deposited by the cell at the EM scale, determined from measurements of test beam data. During the first long shutdown, the number of readout samples was reduced to four in order to reduce the volume of data written to disk in preparation for higher luminosities. A new set of OFCs were therefore derived for the four-sampling readout which can as well take the pile-up noise contribution into account.

2.2.5 The ATLAS Muon System (MS)

The muon spectrometer is the outer part of the ATLAS detector and is designed to mesure precisely the momenta of muons escaping the calorimeters by reconstructing their track and also to trigger on events when a muon is detected. The ATLAS muon spectrometer shown in Figure 2.10, covers the range of $|\eta| < 2.7$ and consists of two kinds of chamber sub-systems: the precision-measurement tracking chambers including the monitored drift tube

[1] The amplitude of the thermal noise depends only upon the characteristics of the detector and the signal processing circuitry.

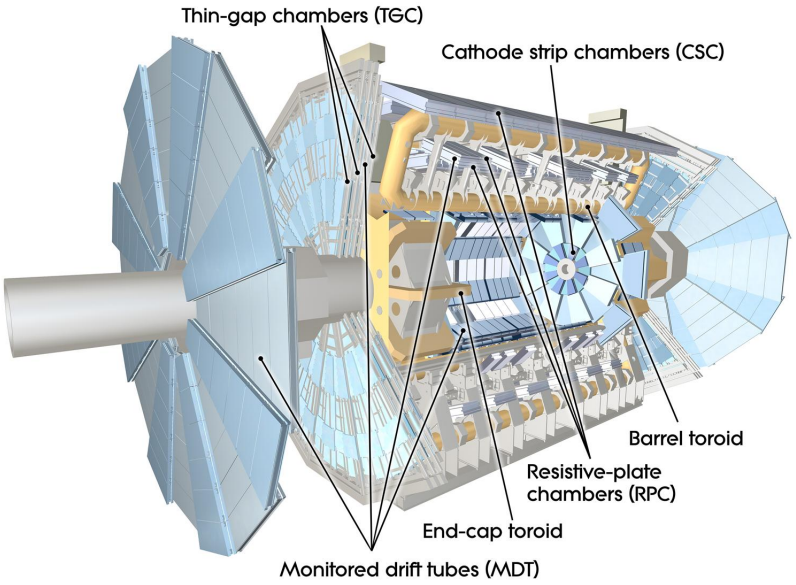


Figure 2.10 – Schematic view of the ATLAS Muon Spectrometer. The Monitored Drift Tube, the Cathode Strip Chambers, the Thin Gap Chambers and the Resistive Plate Chambers are visible.

(MDT) and cathode strip chambers (CSC), and the trigger chambers including the resistive plate chambers (RPC) and the thin gap chambers (TGC). Its structure, shows a cylindrical symmetry around the beam axis. The barrel part is composed by three concentrical layers located at radii 5 m, 7.5 m, 10 m, while the two end-caps are constituted by four disks situated at 7.4 m, 10.8 m, 14 m and 21.5 m from the interaction point.

Tracking chambers: The *monitored drift tub*e (MDTs) ars the main instruments of the Muon Spectrometer for measuring the trajectory of muons. The MDTs are equipped with rectangular chambers containing three to eight layers of drift tubes filled with a gas mixture of 93% CO_2 and 7% Argon at an absolute pressure of 3 bars. In the middle of the tubes a thin anode wire of 50 μm is placed, made of gold plated tungsten and held at a constant voltage of 3 kV.

The *Cathode-Strip Chambers* (CSCs) cover the range $2 < |\eta| < 2.7$. In this region, particle rates exceed 150 Hz/cm^2, the limit for safe operation of the MDTs. The CSCs are constructed from multiwire proportional chambers filled with an Ar/CO_2 gas mixture (The detection principle is the same as in the MDTs). They combine high spatial, time and double track resolution (40 mm per chamber in the bending plane and about 5mm in the transverse plane).

Trigger chambers: The *resistive plate chambers* (RPCs) are a gaseous parallel electrode-plate detector, dedicated for triggering on muons in the barrel region ($|\eta| < 1.05$). The two resistive plates are kept parallel to each other at a gap distance of 2 mm. The gap is filled with a gas mixture of $C_2H_2F_4/Iso - C_4H_{10}/SF6$, a high voltage, about 4.9 kV/mm, is applied between the plates.

The *thin gap chambers* (TGCs) are the muon triggering instruments of the Muon Spectrometer in end-cap region ($1.05 < |\eta| < 2.4$). The TGC detector is similar to CSCs having multi-wire proportional chambers filled with a gas mixture of CO_2/C_5H_{12}, these wires are kept at 2.9 kV and run parallel to the MDTs.

2.2.6 The Trigger System

The ATLAS detector systems process data with a very high rate of events, about 40 millions of bunch crossing per second, corresponding to a rate of 40 MHz at a design luminosity of $10^{34}\ cm^2 s^{-1}$, with each event holds about 1.5 MB disk space of storage. The data acquisition system on ATLAS cannot cope with such high data rates. Furthermore, only a relatively small amount of events are really meaningful and interesting for the physics analyses (In 20 million interactions per second, approximately one $t\bar{t}$ event is produced). As a consequence, a trigger system is set up with the goal to preserve only collision events with interest for the ATLAS physics analyses. In Run 2 the limit of storage rate is reduced to 1 kHz.

The ATLAS Trigger System consists of three levels of event selection: the first level trigger (L1) implemented in hardware and making the initial selection of events, while the second level (L2) and the Event Filter triggers (EF) used *Regions Of Interest* to identify for instance localised high-energy deposits in the calorimeter, using a jet finding algorithm. Each trigger level refined the decision made at the previous level and increased the level of detail available for the decision. A schematic diagram of the ATLAS trigger system (from Run 1) is shown in Figure 2.11.

Level 1: The level L1 trigger is based on the hardware system and implemented in fast custom electronics. It is designed to reduce the bunch crossing rate from 40 MHz to less than 75 kHz according to the decisions made by the central trigger processor (CTS), which treats the information with a maximum overal latency of 2.5 μs. The L1 trigger system implements a pre-selection of the events using a coarser granularity from the muon trigger chambers (via the RPC and TGC chambers), and from the calorimeters (via the trigger towers). The targeted objects are high transverse-momentum muons, electrons, photons, jets, as well as large missing and total transverse energy. Afterwards, if the trigger is passed, the data are sent to the readout drivers (RODs) and buffers (ROBs) for the next trigger level. The L1 trigger has the additional task to define one or more regions-of-interest (RoIs) in η

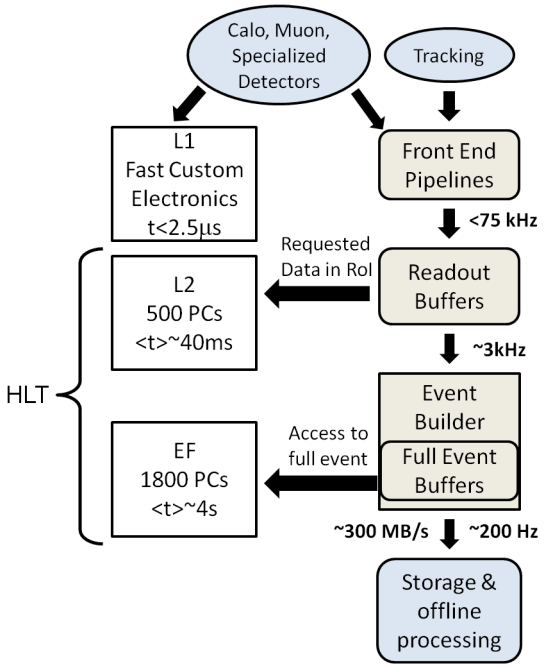


Figure 2.11 – A schematic diagram of the ATLAS trigger system.

and ϕ where the L1 trigger has identified interesting features.

Level 2: The level L2 trigger is based on software system seeded by the RoI from L1. It runs on large computer clusters, reduces the rate to ∼3 kHz, the rejection is about 95% of events passed L1 trigger within an average latency of ∼40 ms/event. The L2 processors run dedicated algorithms on the collected event-fragments in order to reach a decision for the event. If the event is accepted, the data fragments are collected by the event builder and assembled in a full event data structure.

Event Filter: The Event Filter (EF) trigger is the final trigger level. It is also based on the software system, and employs offline physics and event reconstruction algorithms accessing the information of all detectors. At this stage the output rate is reduced to ∼200 Hz, with a latency of ∼4 s/event. The EF allows to perform a global calibration and to check the alignment, this is a fundamental task for the physics quality control, as well as for the detector optimisation and performance. The combination of the two software steps L2 and EF is referred to as High Level Trigger (HLT).

2.3 Luminosity measurement

An important parameter of LHC is the instantaneous luminosity L measured at a specific time, it can be expressed as

$$L = f \times n \frac{N_1 \times N_2}{A} \tag{2.13}$$

where f is the bunch revolution frequency of collisions, n is the number of bunches, N_1 and N_2 are the number of protons in the two bunches, and A is the cross-sectional area of the interaction region. The maximum instantaneous luminosity reached in 2012 was 7.7×10^{33}. If we integrate the previous observable over time we obtain the integrated luminosity usually measured in $femtobarn^{-1}$ or $picobarn^{-1}$, which is used to express the amount of the data collected by the ATLAS detector during some period of time. Processes like the production of the Higgs boson or of massive new particles have very small production cross sections, which means that large amounts of data have to be collected in order to potentially observe them. The integrated Luminosity is mainly related to the total number of events recorded during a time N and the production cross section σ

$$\int L dt = \frac{N_{tot}}{\sigma} \tag{2.14}$$

Figure 2.12 shows the LHC delivered luminosity during stable beams for pp collisions at 13 TeV centre-of-mass energy in 2015. 4.2 fb^{-1} was delivered by LHC of which 3.9 fb^{-1} has been recorded by the ATLAS experiment and 3.2 fb^{-1} was certified to be good quality data (blue).

To measure the ATLAS luminosity, two dedicated detectors are installed in the forward region. The LUCID detectrors (LUminosity measurement using Cerenkov Integrating Detector) and the ALFA detector (Absolute Luminosity For ATLAS). The LUCID consists of two arrays of Cerenkov light detectors made up by aluminium tubes filled with C_4F_{10} gas and located at a distance of 17 m to the interaction point at both side of ATLAS, covering the pseudorapidity range $5.6 < |\eta| < 6.0$. It is conceived to detect charged particles produced in inelastic pp scatterings in the forward direction and so act as an online monitoring of the instantaneous luminosity and the beam conditions. A more precise determination of luminosity is provided with the Absolute Luminosity (ALFA) detector. It consists of scintillating fibers installed in Roman Pots at $\pm$ 240 m from the interaction point and allows to measure elastically scattered protons at very small angles down to 3 μrad and, therefore, provides a direct measurement of the luminosity. In between LUCID and ALFA, at $\pm$140 m the ZDC (Zero degree Calorimeter) detector is located. Its primary purpose is to detect

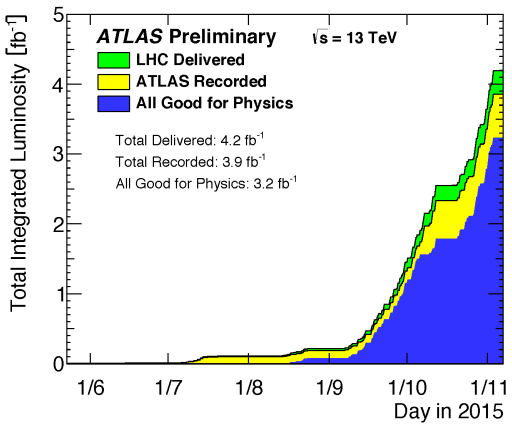


Figure 2.12 – : Integrated luminosity delivered by the LHC (green) and recorded by the ATLAS experiment (yellow) as a function of time during 2015. 3.2 fb^{-1} was certified to be good quality data (blue). Figure from [4]

forward neutrinos in heavy-ion collisions.

2.4 Pile-up

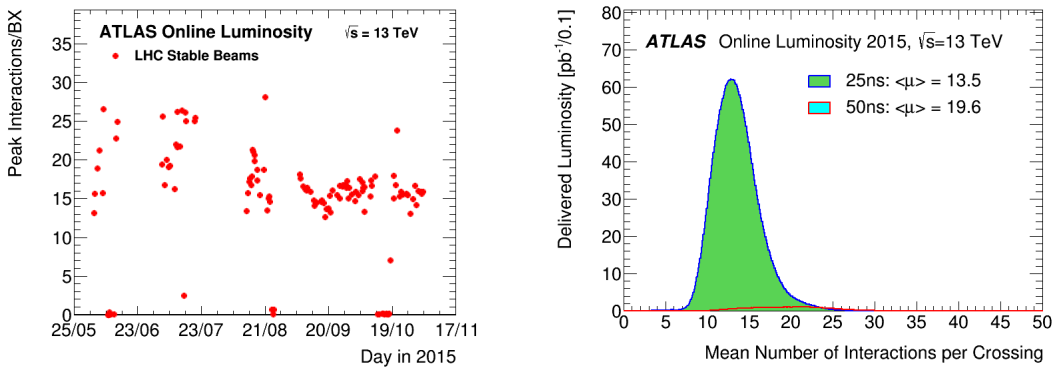
During the high-luminosity running conditions at the LHC, additional inelastic scatterings of proton-proton are generated in the same time window as the hard process of interest. Those additional low energetic pp collisions referred to as pile-up, can influence the measurement of a given event. The pile-up is usually divided into two categories: *in-time pile-up* which results from multiple interactions in the same bunch crossing, and *out-of-time pile-up* which denotes the influence of preceding bunch crossings on the present bunch crossing. They have a strong influence on the design of the LHC detectors and physics analyses. There are also three additional components of the pile-up background which their understanding and modeling is critical for performing analyses in ATLAS.

Cavern background: the gas of thermal neutrons and photons filling the cavern during a typical run of the LHC. These mostly contribute random hits in the muon system.

Beam halo events: the effect of protons from a bunch scraping against collimators far up-stream of the ATLAS detector. The scraping results are mostly muons running approximately parallel to the beam-line.

Beam gas events: collisions between the proton in the LHC beam and residual gas inside the ATLAS beam-pipe. These generally occur off-center in the detector.

Assuming the number of pp interactions per bunch crossing follows a Poisson distribution, we define μ which gives an estimation of the *out-of-time pile-up* and calculated using



(a) (b)

Figure 2.13 – (a) The number of colliding bunches and (b) The distribution of the average number of simultaneous interactions per bunch crossing during the pp runs of 2015. Figure from [4].

the formula 2.15.

$$L =< \mu > \times \frac{N_b \times f_{rev}}{\sigma_{inel}} \tag{2.15}$$

where $< \mu >$ is the mean value of interactions per bunch crossing, f_{rev} is the accelerator revolution frequency, σ_{inel} is the pp inelastic cross-section and N_b is the bunch pairs colliding at the IP. We also use the number of primary vertices N_{PV} variable, sensitive to in-time PU to characterize the amount of pile-up. Figure 2.13 shows the number of colliding bunches and the distribution of the average of simultaneous interactions per bunch crossing during the pp collisions of 2015.

The primary vertices are reconstructed by associating charged-particle tracks to primary interaction points compatible with the beam-spot parameters. Figure 2.14 shows a candidate Z event with 25 reconstructed vertices. This event demonstrates the high pileup environment in 2012 running (β =0.60 m). For this display the track p_T threshold is 0.40 GeV and all tracks are required to have at least 3 Pixel and 6 SCT hits. The vertices shown are reconstructed using tracks with p_T greater than 0.40 GeV, but with tighter requirements on the number of hits on the tracks than in the 2011 reconstruction.

2.5 ATLAS Computing Model

The ATLAS detector produces annually unprecedented amounts of data that need to be stored, processed and made available to thousands of scientists spread around the world. In order to cope with such an amount of data and such a wide community, an advanced

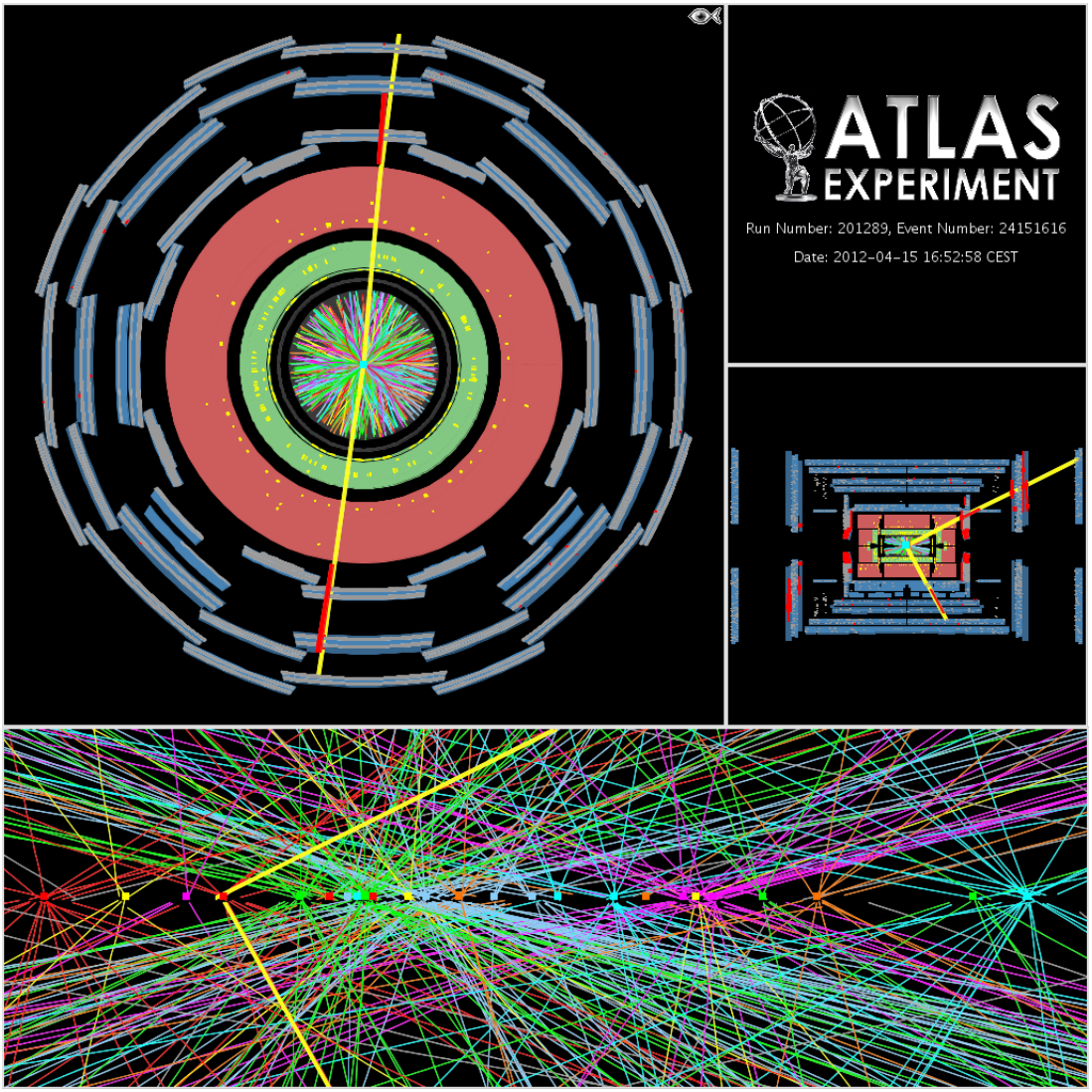


Figure 2.14 – A candidate Z boson event in the dimuon decay. This event was recorded on April 15th 2012 and demonstrates the high pile-up environment in 2012 running. The vertices shown are reconstructed using tracks with p_T greater than 0.4 GeV

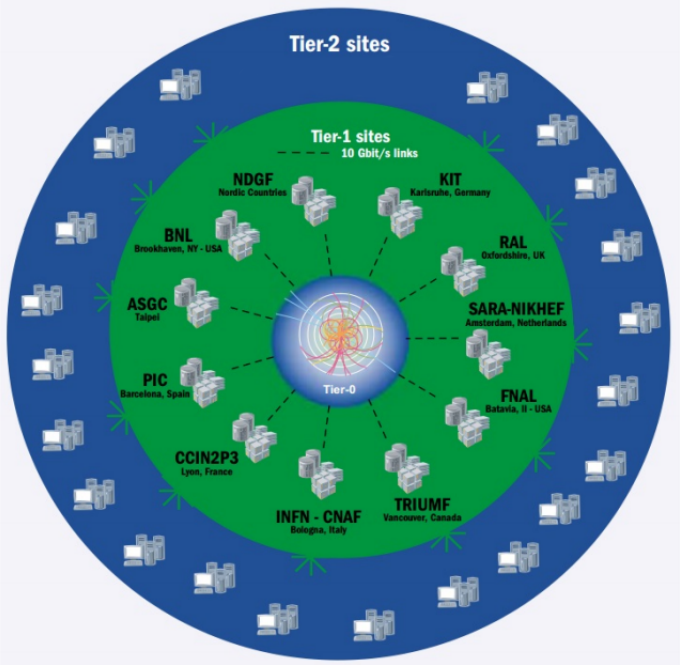


Figure 2.15 – WLCG Tier structure

computing resources to store, distribute and analyse the 30 Petabytes (30 million Gigabytes) for the entire high energy physics community are provided by the *Worldwide LHC Computing Grid* (WLCG) project [46] -a decentralized computing infrastructure arranged in tiers- linking up national and international grid computing infrastructures [2]. The hierarchical **grid** model adopted consists of three tiers, each one is made up of several computer centers and provides a specific set of services. The so called **tier 0** is located mostly at CERN and recently extended with a new data centre hosted by the Wigner Research Centre for Physics in Budapest, Hungary. All of the data from the LHC pass through this central hub, it has a capacity of about 68,000 cores, which is about a third of the grid's total capacity of approximately 235,000 cores. Tier 0 is responsible for the safe keeping of the raw data and performs the first pass at reconstructing the raw data into meaningful information. It is linked with the Tier 1 centers which are typically regional research institutes, reprocesses data when the LHC is not running. In Figure 2.15 you can see the *WLCG* Tier structure.

[2]The *Worldwide LHC Computing Grid Project (WLCG)* was designed in 1999 and approved by the CERN Council on 2001 to operate, build and maintain a distributed computing infrastructure for the simulation, storage and analysis of data coming from the LHC experiments. It federates three large Grid communities, European EGI (the successor of EGEE) [47], US OSG [48] and NDGF [49]. In total it is a global collaboration of more than 170 computing centres in 42 countries. Main purpose of about 200 PB of storage capacity, nearly 350 000 computing cores and mainly 10GB/s network infrastructure of *WLCG* is to provide computational resources to the four main LHC experiments: ATLAS, CMS [17], ALICE [43] and LHCb [44].

Tier 1s consist of 13 computer centres with sufficient storage capacity of LHC data. These sites are mainly responsible for the safe-keeping of a proportional share of raw and reconstructed data, performing large-scale reprocessing, storing the corresponding output and distributing data to the series of Tier 2 computer centres around the world. These Tier 1 sites are linked to each other and to CERN by a dedicated high-width network called *Optical Private Network* (OPN) where the corresponding Optical-fiber links work at 10 gigabits flow per second.

Tier 2s are mostly situated in universities and other scientific institutes, that can store sufficient data and provide adequate computing power for specific analysis tasks. They handle a proportional share of simulated events production and reconstruction. The computing resources are supported by the national funding agencies of the countries where each tier is located and there are around 155 Tier 2 sites around the world.

Individual scientists can access the Grid facilities through local (or **Tier 3s**) computing resources, which can consist of local clusters in a university department or even an individual PC. There is no formal engagement between WLCG and Tier 3 resources.

2.6 ATLAS Event Data Model in Run 2

During the 2013-2014 Shutdown of the LHC, the ATLAS collaboration developed a new event data model for run 2 based on experience gained during Run 1 (Fig 2.16). The model is based on a new ROOT-readable format and called (**xAOD**) produced directly by the reconstruction and replaced the old AOD [3]. A key feature of this model is the separation of the object data from the objects themselves (the *auxiliary store*). Rather than being stored as member variables of the analysis classes, all object data are stored separately, as vectors of simple values. A significant component of the model is the **derivation framework**, used to create the intermediate data products from the xAOD by removing (and adding) information whilst maintaining the structure and EDM used in the original xAOD. Users can access to the full xAOD as is implied by the lower route in Figure 2.16, but the large majority uses centrally produced derived formats (DxAOD). Approximately 100 derived formats are estimated to be used for the full range of activities of ATLAS in Run 2. Collectively these are required to occupy no more disk space than a single copy of the full xAOD. This implies that overall, a given derived format should have around 1% (or less) of the volume of the

[3] **Analysis Object Data (AOD)** is a reduced event representation, derived from ESD data in Run 1. This format was suitable for physical analysis, and contains a summary of the events reconstruction with physics objects such as electrons, muons, jets, etc. However, the AOD files have been too heavy for an analysis in a local machine, and hence were not ROOT-readable. To deal with this, ATLAS converted this complex transient data model to a simpler one which could be written to ROOT directly

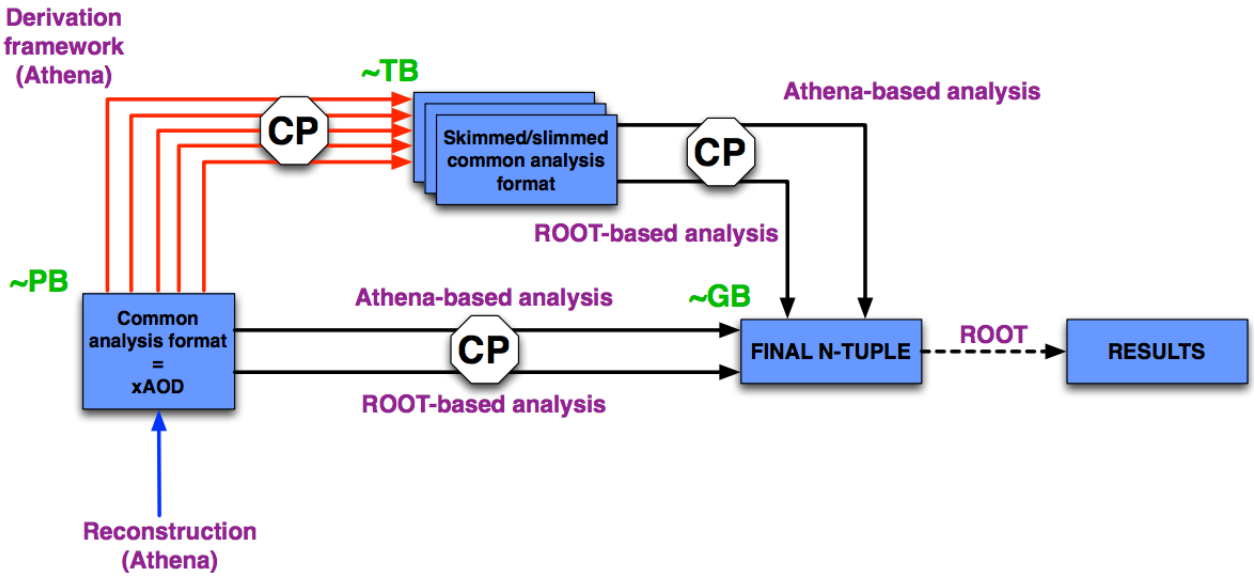


Figure 2.16 – The Event Data Model defined for ATLAS for Run 2

input xAOD. The derivation framework provides a set of services:

- **Skimming** is the removal of whole events, based on some criteria related to the features of the event.
- **Thinning** is the removal of individual objects within an event, based on some criteria related to the features of the object.
- **Slimming** is the removal of variables within a given object type, uniformly across all objects of that type and all events. Unlike the other operations, slimming does not vary depending on any event/object properties: the same variables are removed for every event and object.

The final component of the model is the **analysis framework**, which is used by physicists to read the derived data formats, apply various combined performance tools and produce the final small n**-tuples**, from which plots are produced and upon which statistical analyses are based.

2.7 ATLAS Distributed Computing

The ATLAS Distributed Computing (ADC) is a very complex system used to deliver production tools and support data processing, data management operations and physics analysis

at the ATLAS computing centers located worldwide. In this section we describe the ATLAS workload system, called PanDA (Production ANd Distributed Analysis [50] and the Rucio system, the current generation of ATLAS Distributed Data Computing (DDM). PanDA and DDM systems are critical to perform the analysis presented here. A brief description of both systems is presented in the next sections.

2.7.1 Panda System

PanDA was developed to provide an integrated service architecture with late binding of jobs, tight integration with data management and data flow, minimal operations manpower, advanced error discovery and recovery procedures, and other features. PanDA was initially developed for US based ATLAS production and analysis in late 2005. In 2007 PanDA was adopted by the ATLAS Collaboration as the sole system for distributed processing production across the Collaboration. PanDA throughput has been rising continuously over the years and the current processing rate is 100k-200 jobs/day running at any given time within US ATLAS, and 4 times that number ATLAS-wide.

Architecture and workflow

Jobs are submitted to PanDA via a simple Python/HTTP client interface, so users can define job sets, their associated datasets and the input/output files within them. Then Job specifications are transmitted to the PanDA server with submission information returned to the client. This same interface has been used to implement PanDA front ends for ATLAS production, distributed analysis (Pathena), and other experiments. The PanDA server is the main components which receives work from these front ends into a global job queue, operated by a brokerage module operates to prioritize and assign work on the basis of job type, priority, input data and its locality and available CPU resources. Allocation of job sets to sites is followed by the dispatch of input data to those sites, handled by a data service module interacting with the Rucio Storage Element system (discussed below). Data pre-placement is a strict precondition for job execution: jobs are not released for processing until the data arrives at the processing site. When data dispatch completes, jobs are made available to a job dispatcher.

An independent subsystem manages the delivery of pilot jobs to worker nodes using various scheduling systems. A pilot launched on a worker node contacts the dispatcher and receives an available job suitable to the site. If no appropriate job is available, the pilot may immediately exit or may pause and ask again later, depending on the configuration. An important attribute of this scheme for interactive analysis, where minimal latency from

job submission to launch is important, is that the pilot dispatch mechanism bypasses any latencies in the scheduling system for submitting and launching the pilot itself. The pilot job isolates workload jobs from grid and batch system failure modes (a workload job is assigned only once the pilot successfully launches on a worker node).

JEDI (Job Execution and Definition Interface)

JEDI is an intelligent component in the PanDA server to have capability for task-level workload management. The Main reasons to develop JEDI are to make the PanDA system more task-oriented and to integrate prodDB and PandaDB.

JEDI works with work queues. Tasks are mapped to work queues based on their attributes, such as processing type, working group, production source label and core count. Each task belongs to only one work queue and thus all jobs in the task belong to that particular work queue.

JEDI was put in production in the first week of July 2013 and gradually more and more systems started using its new functionalities. To address and support the new features, extension of the DB implementation took place. Its data segmentation is based on a range partitioning on the JEDI's TASKID column with interval of 100.000 IDs (tasks) on six of the JEDI tables (thus achieving uniform data partitioning). The JEDI data segments are placed on dedicated Oracle data file. Some new JEDI tables were added to the existing PanDA database schema and some of the existing PanDA tables were altered by adding new columns and foreign keys relevant to the JEDI new objects.

2.7.2 Rucio System

The Rucio project [51] is the new generation of ATLAS Distributed Data Management (DDM) system benefiting from recent advances in cloud and *Big Data* computing to address HEP experiments scaling requirements. Rucio allows the ATLAS collaboration to manage very large volumes of data with more than 140 petabytes per year, taken by the detector as well as generated or derived in the ATLAS distributed computing system. In this section we present how data are organized and catalogued by Rucio. We show also an overview of its RESTful architecture and the various software components it uses.

The following concepts define how data is organised, accessed, and catalogued by Rucio.

Rucio uses to manage accounts three pieces (files, datasets and distibuted storage systems) in order to provide an advantage of storing the data physically in separate segments.

Files are the smallest operational units of data and can be grouped into datasets (a named set of files) and datasets can be grouped into containers (a named set of datasets

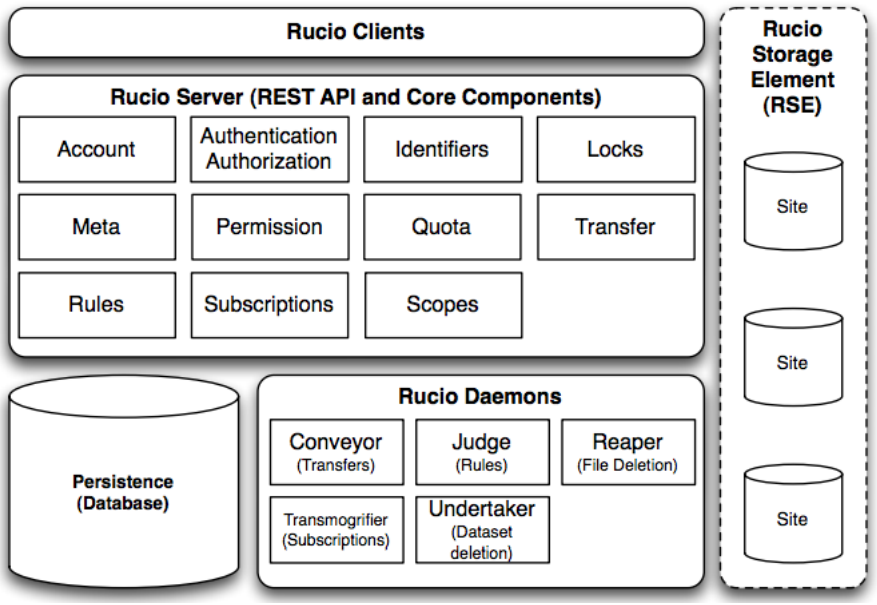


Figure 2.17 – Overview of the Rucio architecture.

or, recursively, containers). All three types follow an identical naming scheme which refer to data, so the term *data identifier set (DIS)* is used to mean any set of file, dataset or container identifiers which is just composed of the scope and the name of a single file, dataset or container.

A Rucio Storage Element (RSE) is a repository for physical files. It is the smallest unit of storage space addressable within Rucio and designed to provide a way to define each storage and all its supported protocols (e.g., file, https, srm) separately and include them in a transparent way for users and developers. It has an unique identifier and a set of properties such as quality of service; storage type (e.g., disk, tape, ...), physical space properties (e.g., used, available, non-pledged); and geographical zone. Rucio Storage Elements can be grouped in many logical ways, e.g., the UK RSEs, the Tier-1 RSEs, or the *good* RSEs. One can reference groups of RSEs by *metadata attributes* or by explicit enumeration of RSEs. Physical files stored on RSEs are identified by their Physical File Name (PFN) [4] and the mapping between the file identifier and the PFN is a deterministic function of the identifier, RSE and protocol.

The Rucio software is based on a distributed architecture and can be decomposed into three horizontal layers and one vertical layer (see in Figure 2.17).

The **Rucio clients** layer offers a command line client for users as well as application

[4]The PFN is a fully qualified path identifying a replica of a file. PFNs may take the form of file names, URIs, or any other identifier meaningful to a RSE

programmer interfaces in Python which can be directly integrated into user programs. All Rucio interactions are transformed by the client into https requests which are sent to the REST [52] interface of the Rucio server. Consequently, external programs can also choose to directly interact with the REST API of the server.

The **Rucio server** is a passive component listening to incoming queries. It connects several Rucio core components together and offers a common, https based, REST interface for external interaction. After a request is received by the REST layer, the authorization component checks the users with their credentials interact first, If they are valid, the permissions of the account to execute the given request are then checked by the permission component. If permitted, the request is passed to the responsible core component for execution. Rucio core components are allowed to communicate with each other, as well as with the Rucio storage element abstraction.

The **Rucio Storage Element (RSE)** abstraction layer is used by the clients, the server as well as the Rucio daemons. It is responsible for all interactions with different grid *middleware* tools and storage systems. It effectively hides the complexity of these tools and combines them into one interface used by Rucio.

The **Rucio daemons** are active components used to orchestrate the collaborative work of the whole system. The following daemons operate requests made by users or by the Rucio core:

- *Conveyor* is in charge of requests transfers.
- *Reaper* and *Undertaker* deal with file replica and datasets deletion with expired lifetime.
- *Transmogrifier* and *Judge* apply subscriptions on newly created/existing data to generate replication rules.

The persistence layer keeps all the logical data as well as for transactional support. Rucio uses SQLAlchemy [53] as an object relational mapper and can therefore support several relational database management systems (RDBMS) like Oracle, MySQL or PostgreSQL. Only the Rucio server or daemons directly communicate with the database.

Chapter 3

ATLAS event simulation chain and Systematic validation of the different Geant4 physics lists

3.1 Monte Carlo event simulation chain

The LHC produces high energy events where the understanding of their final states is an extremely challenging theoretical problem. In simulation, Monte Carlo event generators play the role of the LHC machine and detector simulation programs play the role of the ATLAS detector. During the preparation phase of an experiment, simulation provides the environment to better understand the detector, to develop analysis strategies, to estimate the sensitivity to different physics processes, to develop and validate object reconstruction algorithms, to optimize the trigger menus, and so on. During the running phase of an experiment, simulation is used to compare predictions of theoretical models against the real data. The main purpose of this review is to provide a survey of how all the above components are implemented in the general-purpose event generators and simulation algorithms that are currently available for the simulation of LHC proton-proton collisions and modeling the subsequent interactions of the particles with the detector.

The set of steps required to simulate and reconstruct the different parts of the Standard Model are described in the following (Fig. 3.1).

Event generation: corresponds to the phase of proton-proton (pp) collision events generation. It consists of the production and decay of particles in a given process. Several event generators are available as described in Section 3.1.5.

This first step as built up by the event generators can be described by different aspects (shown in Fig. 3.2): The hard-scatter process that describes the interactions between the two partons (constituents of the proton) that happens at a high energy scale (red circle) and described in Sec. 3.1.1. Additional radiation before and after the hard-scatter occurs has to be taken into account, and modeled correctly as explained in Sec. 3.1.2. As protons are collided in bunches of protons, apart from the interesting hard scatter process, additional interactions between protons are found, this has to be simulated as described in Sec. 3.1.4. As partons cannot be found freely, the building of hadrons is simulated (green ellipse) as outlined in Sec. 3.1.3. Finally, the events generated in this manner are later passed to a Geant4-based simulation [54, 55] software, capable of modeling the interactions of the particles with the detector as described in Sec. 3.1.6.

Digitization: corresponds to the simulation of the detector readout, i.e. the conversion of energy deposited in the detector components to times, currents and voltages for readout electronics. The output format of the simulation is identical to the real detector output format.

Reconstruction: in this step a set of object reconstruction algorithms are applied. These algorithms are applied to both simulation and real data in exactly the same way. This phase is fully described in Chapter 4.

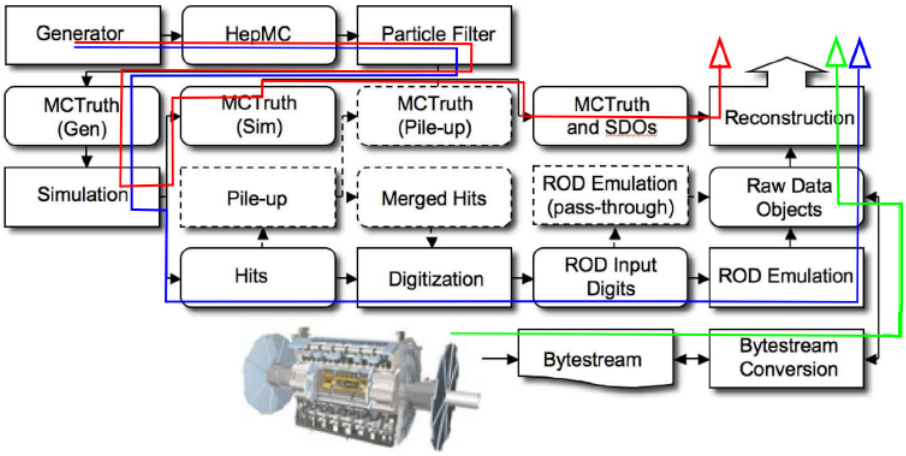


Figure 3.1 – The red path leads to particle level physics objects, the blue path to reconstructed level physics objects, while the green path shows the real data ow to physics objects. SDO stands for Simulated Data Object, ROD for Read Out Driver.

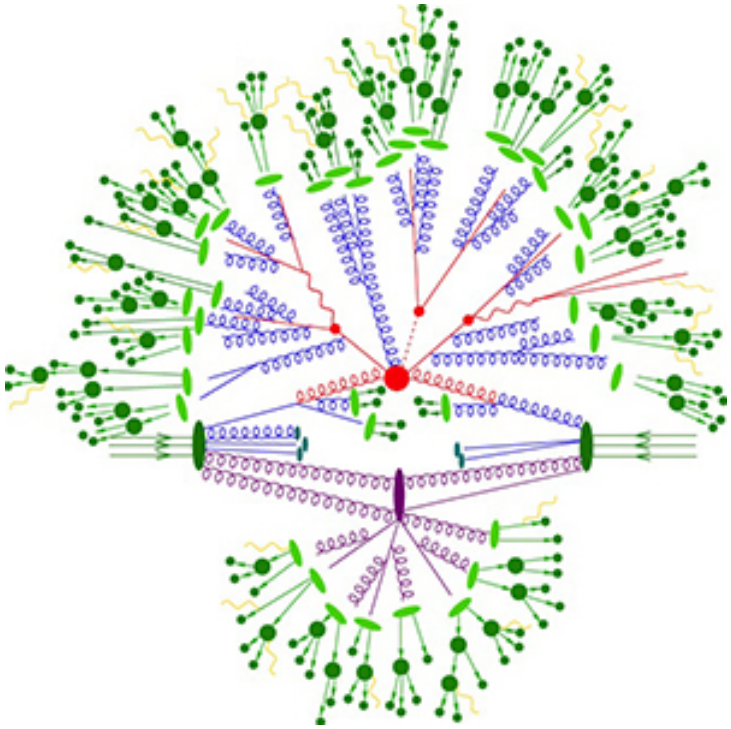


Figure 3.2 – Schematic workflow of a typical Monte Carlo event generation.

3.1.1 Hard Scatter Process

The hard process describes the interaction of partons (quarks and gluons) producing high energy physics event or new physics since it probes distance scales far below the radius of proton. This interaction can be described by matrix elements (ME) 5.4.1, calculable at a fixed-order in perturbation theory (generally leading order (LO) or next to leading order (NLO), in the coupling constant $\alpha(s)$) over a given phase space. This step makes use of the factorization theorem which refers to the set of theorems that establish that the short distance component of the scattering process described by perturbative QCD can be separated from the non-perturbative long distance component, i.e allows to separate parts that describe the final-state hadron formation from the perturbative hard interaction part among the partonic constituents. Factorization is a by-product of a procedure that absorbs singularities into physical quantities (in a similar way as renormalization does). For this reason, a new scale called the factorization scale μ_F^2, is introduced.

$$\sigma(pp \to X) = \sum_{ij} \int dx_1 dx_2 f_{i,p}(x_1, \mu_F^2) f_{j,p)}(x_2, \mu_F^2) \hat{\sigma}_{ij \to X}(x_1 x_2 s, \mu_R^2, \mu_F^2) \tag{3.1}$$

Where $f_{i,p}$ are the parton density functions (PDFs), which depend on the longitudinal

momentum fractions $x_{1,2}$ of the parton 1, 2, and on the factorization scale μ_F^2. The sum here runs over all possible initial state partons that can produce a final state X at a center of mass of $\sqrt{x_1 x_2 s}$. The factorization scale μ_F^2 and the renormalizaton scale μ_R^2 appear in the expressions if they are gathered from truncted expansions in the strong coupling constant. The parton-level cross section denoted by $\hat{\sigma}$ may be calculated at leading order (LO) in the strong coupling α_s, or incorporating next-to-leading order (NLO) or even next-to-NLO (NNLO) corrections.

3.1.2 Parton showers

The previous section describes the generation of a hard interaction of the type $pp \to X$. But as was mentioned before, the hard interaction implies large momentum transfers and therefore the partons participating in it are violently accelerated. Just as accelerated electric charges emit photons, the accelerated colored partons may emit the incoming and outgoing colored participants in the sub-process. These generated strong interacting partons, radiate gluon *bremsstrahlung* and loose energy, until the energy is too small to radiate further. This process is described by parton shower models, which are accurate at leading logarithmic order. These models are applied to simulate the evolution of outgoing partons (final state radiation (FSR)) and the reconstruction of the history of the incoming partons (initial state radiation (ISR)).

In principle, parton showers (PS) represent higher-order corrections to the hard interaction $2 \to n$, corresponding to the production of additional partons in the process $2 \to n+X$. Since radiative corrections at a fixed perturbative order are divergent at low energies (infrared divergence) or small angles (collinear divergence), it is not possible to calculate these corrections exactly, and an approximation scheme (leadinglog) is used where only the main contributions are included. There are three possible processes for QCD emission (splitting): $q \to gq$, $g \to gg$ and $g \to q\bar{q}$. The partonic cross section for the $2 \to n$ process with one extra-emission $2 \to n+1$ can be expressed as the product of the $2 \to n$ partonic cross section and a factor accounting for the splitting probability of one of the partons. Hence, for each splitting process i, the $2 \to n+1$ differential cross section is defined by

$$d\sigma_{2\to n+1} \approx d\sigma_{2\to n} \frac{\alpha_S}{2\pi} \frac{d\theta^2}{\theta^2} dz d\phi P_i(z,\phi) \tag{3.2}$$

where θ and ϕ are the opening angle and azimuthal angle of the splitting, and P_i is the splitting function, which describes the distribution of the fraction z of energy of the original parton, assigned to the new parton. A parton shower is generated for each parton of the hard interaction by applying Eq. 3.2 sequentially. A pseudo-random number is used to

generate values of z, θ and ϕ for each splitting. A major feature of the showering algorithm, not apparent from Eq. 3.2, is the *evolution variable*. The simplest evolution variable is the virtual mass-squared q^2 (also known as the virtuality) of the partons in the shower, which determines the starting and final stage of the evolution of a parton shower. The initial virtuality is required to be smaller than the momentum transfer of the hard process, and the shower is terminated when the virtuality has fallen to low values ($q^2 = Q_0^2 \simeq 1GeV^2$), where Q_0^2 is called hadronization scale.

The virtual contributions to parton emissions are included in the probability of not splitting during evolution from scale q_1^2 to q_2^2. Such probability is referred to as *Sudakov form factor*, defined as

$$\Delta_i(q_1^2, q_2^2) = \exp\left[-\int_{q_2^2}^{q_1^2} \frac{q^2}{dq^2} \frac{\alpha^S}{2\pi \hat{a}\check{A}\acute{c}} \int_{\frac{Q_0^2}{q^2}}^{1-\frac{Q_0^2}{q^2}} dz \int_0^{2\pi} d\phi P_i(z,\phi) \right] \tag{3.3}$$

The evolution of the parton shower is therefore governed by the *Sudakov* form factor. Given the initial scale Q_2, the MC generator solves the equation $\Delta_i(Q^2; q_1^2) = R_1$, where R_1 is a random number uniformly distributed in the interval [0,1], for the virtuality q_1^2 of the first splitting. If the condition $q_1^2 < Q_0^2$ is met, the shower development is terminated and hadronization takes place. Otherwise, the procedure is repeated for each new parton produced by the splitting, taking q_1^2 as initial scale. For each splitting the variables z and ϕ are generated according to the distribution defined by the splitting function.

Final state radiation showers (FSR) : The description above applies the evolution of showers associated with partons produced in the hard interaction, starting at a high energy and a virtuality scale Q_2 set by the hard interaction, and progressively reaching the hadronization scale, Q_0^2. This process is typical of FSR parton showers that are generated from outgoing partons of the hard interaction.

Initial state radiation showers (ISR) : In the case of ISR parton showers, the radiation is emitted by the colliding partons, and there is an important difference in the shower evolution, as the final showering of these partons ceases when they collide to initiate the hard interaction. The partons emitted in the initial state showers may initiate secondary showers that progress in the same way as final state showers.

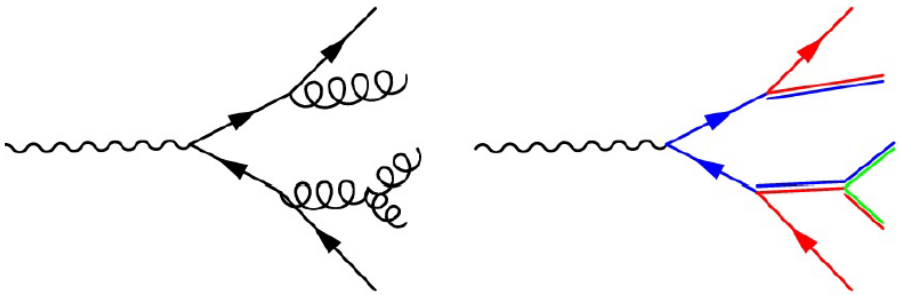


Figure 3.3 – Cluster model: gluons are represented by colour-anticolour lines.

3.1.3 Hadronisation and decay

Once the low virtuality scale is achieved, the phase of hadronization starts, and the dynamics of the parton enters a non-perturbative phase, which leads to the formation of the colour neutral hadrons. The hadron formation happens at energy scales at which $\alpha(s)$ is large and the absence of a well-known theory about a correct treatment of non-perturbative QCD makes the Hadronization one of the most complex steps of the Monte Carlo simulation. Several phenomenological models to simulate the hadronization phase exist. The *Lund string model*, which is implemented in some Monte Carlo generators (like `Pythia` is an advanced model using as starting point the assumption of linear confinement between partons which is induced by the color force and represented by a gluonic string [56]). Based on the experimental observation, at long distances, the string breaks and creates two additional color charges at the endpoints. In the case of a quark-antiquark pair from two adjacent broken strings they combine to form a hadron. Another advanced hadronization model is the *Cluster Model* (Fig. 3.3) where final state gluons are forced to split into light quark-antiquark pairs, and partons are then grouped to form clusters from color-connected pairs. The high mass clusters are decomposed into low mass clusters using string-like mechanism. Cluster splitting continues until all hadrons are built.

The decay products of strings and clusters are mainly unstable resonances, which decay themselves, following the PDG data tables [57].

3.1.4 Underlying events (UE)

In events containing a hard interaction, an additional hadron production mechanism arises from the softer interaction of spectator partons where the perturbative approaches do not hold and phenomenological models are used to model them. Parameters in each model are tuned based on experimental measurements in data. The dominant sub-process of the underlying events (UE) is gluon-gluon scattering, with an inclusive parton-parton cross

section larger than the total proton-proton cross section, indicating that multiple gluon scatterings per proton collision are likely. For this reason the generic soft scattering of partons is referred to as multiple parton interactions (MPI) and is modeled in MC generators as the production of back-to-back jet pairs with little total transverse momentum. The color connection with the beam remnants that are not interacting and the rest of the event is also kept and simulated with phenomenological models.

3.1.5 Monte Carlo event generators

Monte Carlo (MC) event generators are widely used to make theoretical predictions and preparations for the simulation of proton-proton collision event and subsequent passage of outgoing particles through the detectors. MC generators implement a large number of physics processes used to model the physics of the events as they could be occurring in the experiment. This is paramount for the interpretation of results obtained from data, as well as their integration over a phase spaces of huge and variable dimension.

Monte Carlo Generators can be classified as either multi-purpose generators, capable of performing the full simulation chain, or as specialized generators, optimized for an accurate simulation of specific aspects. The characteristics of the main MC generators used in the work described in this dissertation are listed below.

- `Pythia:` [56] is a commonly used event generator in high energy physics, designed to run using different PDF functions, and generate a wide variety of Standard model and beyond the Standard Model processes. It is able to simulate hard interactions in hadron-hadron, hadron-lepton, and lepton-lepton colliders with various customisations, and interfaces to PHOTOS [1] to perform final state radiation of photons. Pythia simulates non-diffractive proton-proton collisions using a $2 \to n$ ($n \leq 3$) matrix element at Leading-order to model the hard subprocess, and uses p_T-ordered parton showers to model additional radiation in the leading-logarithmic approximation. The hadronisation model used is the *Lund string model* [59] and a well developed multiple-interaction model for the underlying event.

- `Sherpa:` [60] is a Monte Carlo event generator for the *Simulation of High-Energy Reactions of PArticles* in lepton-lepton, lepton-photon, photon-photon, lepton-hadron and hadron-hadron collisions. The list of physics processes that can be simulated with Sherpa covers all reactions in the Standard Model and many theories beyond. The entire code has been written in C++ and is based on a modular program. Accordingly, each module encapsulates a different aspect of event generation for high-energy

[1] Photos are a decay packages interfaced to work with Pythia, Herwig [58] and Sherpa

particle reactions.

- `Co-mix:` [61] is a multi-leg tree-level matrix element generator, based on the color dressed Berends-Giele recursive relations [62]. It employs a new algorithm to recursively compute phase-space weights. The module is a useful supplement to older matrix element generators like `AMEGIC++` [63] in the high multiplicity regime. Due to the usage of colour sampling it is particularly suited for an interface with parton shower simulations and can hence be easily employed for the ME-PS merging within `Sherpa`. It is `Sherpa's` default large multiplicity matrix element generator for the Standard Model.

- `Powheg:` *POsitive Weight Hardest Emission Generator* [64] is a multipurpose parton-level Monte Carlo generator. It allows to interface NLO QCD calculations with a parton shower generators. It generates the hardest emission first, with NLO accuracy independently of the parton shower generator used. It uses a MC-independent method and can be interfaced with many parton shower generators (most commonly `HERWIG` [58] and `Pythia`).

- `Powheg-Box:` [65] is a general computer framework for implementing NLO calculations in shower Monte Carlo programs according to the `Powheg` method. It saves the hard event information and can be interfaced with all modern shower Monte Carlo programs that support the *Les Houches Interface for User Generated Processes*.

3.1.6 ATLAS detector simulation

The events generated with an appropriate MC generator are passed through the ATLAS detector simulation GEANT4 [2] [54] [**?**] which reproduces the interaction of these particles with the experimental setup, by propagating all final state particles through the ATLAS detector and converting the energy deposits into electronic signals with the same format as the ATLAS readout system. The interaction of particles with the detector material, taking into account its materials, geometry and readout system, is modeled. The GEANT4 parameters

[2] `GEANT4:` The `GEANT4` toolkit provides a framework for managing the detector simulation where all aspects of the simulation are included: A description of the detector geometry and material composition in combination with an array of packages modeling the response of particles in matter. The packages involve the materials, the fundamental particles of interest, the generation of primary particles of events, the tracking of particles through materials and external electromagnetic fields, the physics processes governing particle interactions, the electromagnetic, hadronic, and optical responses of sensitive detector components, the generation of event data and the decay of long-lived particles and the effects of multiple scattering, the storage of events and tracks, the visualisation of the detector and particle trajectories, and the capture for subsequent analysis of simulation data at different levels of detail, development and refinement. ATLAS maintains its detector description in a separate geometry database, which allows both ease of use in both simulation and reconstruction programs as well as for maintaining records for different detector configurations.

are tuned using test-beam and pp collision data. The accuracy of the detector simulation is based on the information maintained in the geometry database, which contains information about the dimensions, geometry, positions and material properties of each detector volume, while the conditions database contains information about dead channels, temperatures and misalignments in the real detector at a given time. Since conditions and failures vary from run to run, it is important that the detector simulation reproduces as close as possible the real status of ATLAS during a particular data-taking period. To ensure this, simulation samples are reprocessed for each data reprocessings, called also releases.

3.2 Systematic validation of the different Geant4 physics lists

In order to provide a simulation of hadronic shower of high-energy particles on calorimeter based on Geant4 (G4), an extensive physics validation program has been carried out in the past years by ATLAS experiment. In this section we aim to give a short summary of the results as part of this validation activity. The present study is based on two Geant4 (G4) versions G4 9.4 and G4 9.6 used on the early 2015, and with a wider choice of physics lists: QGSP_BERT, QGSP_BIC, QGSP_FTFP_BIC, FTFP_BERT, QBBC and FTF_BIC. Those physics lists are a combinations of physics models chosen in G4 in order to model hadronic and electromagnetic physics processes. An important part of this work is understanding how changing the models considered for the physics lists to set a reasonable uncertainty on the hadronic showers between different physics list, and for the same calorimeter configuration and primary particles.

For a good validity study, various single particle relevant in the hadronic showers have been considered: pions, kaons, protons and neutrons.

The structure of this section is the following: A brief description of used physics lists is given in the section 3.2.1. The calorimeter response for Monte Carlo single particles ran with different physics lists and the variation in response versus physics list as a function of the primary energy, are discussed in section 3.2.2. Section 3.2.3 present The evolution of the physics lists in the two G4 versions 9.4 and 9.6.

3.2.1 Physics lists

Any Geant4 application has to include a physics list, which provides a complete and consistent collection of numerical models that describe the interaction of high energy hadrons in the G4 simulation. Several combination of that models are possible and useful for a

given use-cases. The users are allowed to write their own preferred Physics List, but several pre-defined ones are available in G4 for convenience, and indeed they are used by the large majority of users. The lists of interest for high-energy physics applications are mainly based on one of the two main string models used at high energies ($E > 20$ GeV), the *Quark Gluon String (QGS)* or *Fritiof-like string model (FTF)* showered by two different models, *Binary Cascade (BIC)* and *Bertini-style cascade (BERT)*, which are responsible to track low energy region (below E ∼10 GeV). We will describe here only the main hadronic models used in this work, as present in G4 version 9.4. All values refer to the kinetic energy of primary particle in the interaction.

LEP: *Low-Energy Parameterized model* used for pions, kaons neutrons, and protons interactions with nuclei at energies between 9.5 GeV and 25 GeV. It covers generally the transition region in the described physics list below.

QGSP_BERT: The most recommended physics list used by ATLAS for high Energy physics (HEP) in the particle Monte Carlo production. It is widely used for simulation and has demonstrated the best agreement with data in earlier G4 versions. The list employs the *Bertini nuclear cascade* model (BERT) [66] which handles incident nucleons with kinetic energy below 9.9 GeV. For high energy (> 25 GeV) it uses the *Quark-Gluon String Pre-compound* model (QGSP) while the *low energy parametrised* model (LEP) is used in the intermediate energy region (from 9.5 to 25 GeV). Note that in the overlapping regions between models (9.5-9.9 GeV for *Bertini* and *LEP* for example) a random sampling is performed to decide for each interaction which of the two models to use.

FTFP_BERT: In FTF physics lists, a different string model is used. The FTF model for high energy hadronic interactions is based on the `FRITIOF` description of string excitation and fragmentation. The list includes the BERT model below 5 GeV and the *Fritiof precompound* model above 4 GeV. (The FTFP_BERT was significantly improved in version 9.6 and start to be the most recommended for HEP simulations by the G4 collaboration).

QGSP_BIC: It uses G4 *Binary cascade* model for primary protons and neutrons with energies below 9.9 GeV, thus replacing the use of the LEP model for protons and neutrons. In comparison to the LEP model, *Binary cascade* better describes the rescattering of secondaries particles produced by the *Quark-Gluon model*.

QGSP_FTFP_BERT: The same as QGSP_BERT for both pion and proton showers,

with LEP replaced by FTFP. This transition does not affect the longitudinal shower profile in the energy range of the LHC test beam setups. This physics list is not available before Geant4 version 9.3.

QBBC: It includes combinations of BIC, BERT, QGSP, FTFP and other models and has higher precision than the others for many hadron-ion and ion-ion interactions in a wide energy range. This physics list replace the *Quark-Gluon String* Model by the FRITIOF string fragmentation model for kinetic energy E > 6 GeV; for E < 6 GeV, the *Bertini Cascade* is used for pion interactions, and the *Binary Cascade* for proton interactions; elastic and quasielastic.

FTF_BIC: It uses the *Binary cascade* to rescatter slow products, before the Binary model uses the *Precompound* model for de-excitation. Both FTFP_BERT and FTF_BIC were a promising alternatives for QGSP_BERT, before the choice has been returned in Geant4 9.6 toward FTFP_BERT as a standard physics list.

3.2.2 Calorimeter response

The calorimeter response is defined as the total sum of simulated energy in the calorimeter cells times sampling fraction, divided by the true particle energy. The calorimeter response has been studied for the primary particles with energy ranges from 0.5 to 25 GeV and for $\eta < 0.8$ as a function of the physics lists. The result of this study is illustrated in figures 3.4 and 3.5 for the neutral and charged pions cases. The response of the single neutral pions for the models described in the considered physics lists (see Fig. 3.4) is closing to one when the energy increasing, while the response of the single charged pions gives worse result (Fig. 3.5) as a function of energy and for the tested physics lists.

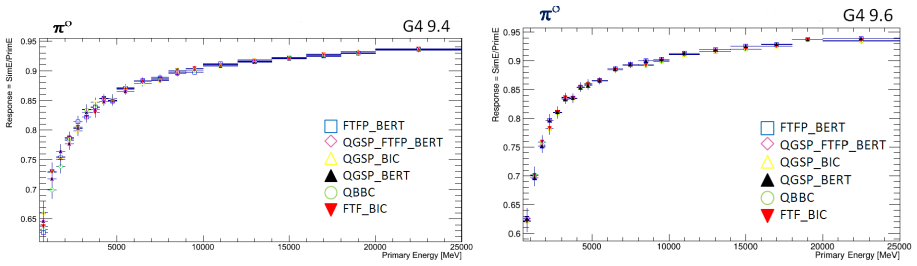


Figure 3.4 – Neutral pion (π^0) response as a function of different physics lists in both G4 version 9.4 and 9.6.

Geant4 provides several *reference* or *standard* physics lists which are routinely validated

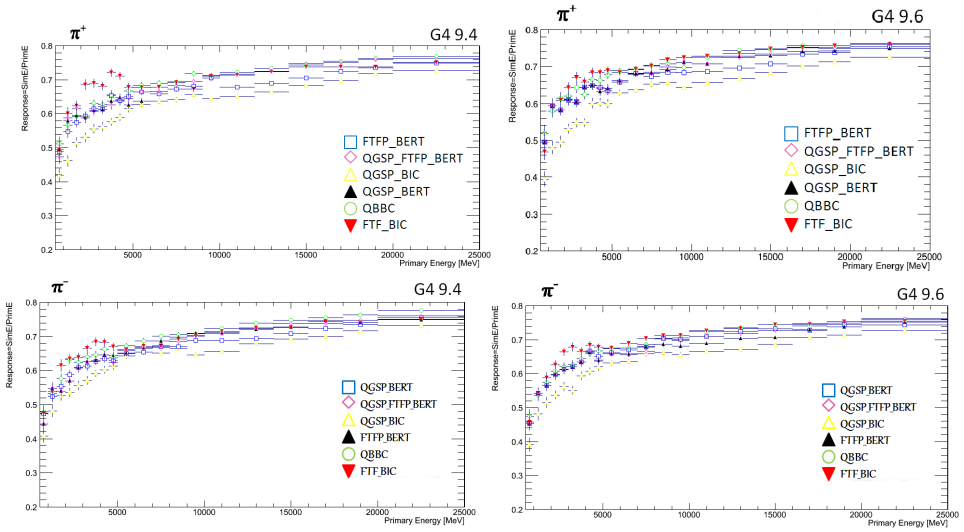


Figure 3.5 – Charged pions (π^+/π^-) response versus physics list for G4 9.4 and G4 9.6 versions

and updated with each release. In G4 version 9.4, the QGSP_BERT is considered as the reference physics list, while in G4 9.6 the standard one is the FTFP_BERT. These physics lists are chosen by the ATLAS experiment as the standard physics lists as well. Figures 3.6, 3.7, 3.8, 3.9 show the calorimeter response versus the physics lists as a function of the energy for various single particles; pions, kaons, neutrons and protons. As illustrated in the figure 3.6 the neutral pions response gives a roughly flat response with a ratio value close to one, and showing similar behavior for all the tested physics list. However the response of the charged pions, neutrons and protons show a significant fluctuation around the ratio close to one for the considered physics list. For single kaons, similar result is shown in figure 3.8 for the set of the physics lists using the same versions of G4 (G9.4 and G9.6).

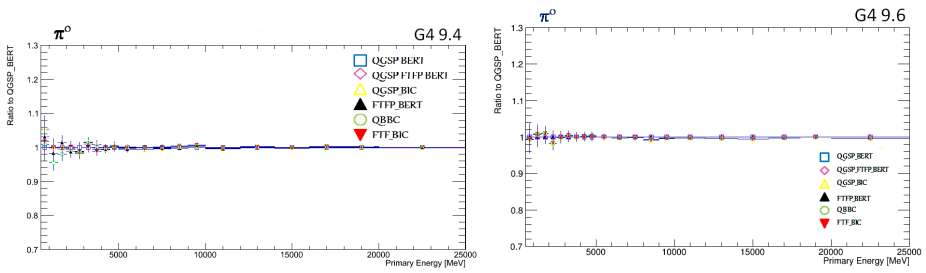


Figure 3.6 – π^0 response variation versus physics list as a function of primary energy.

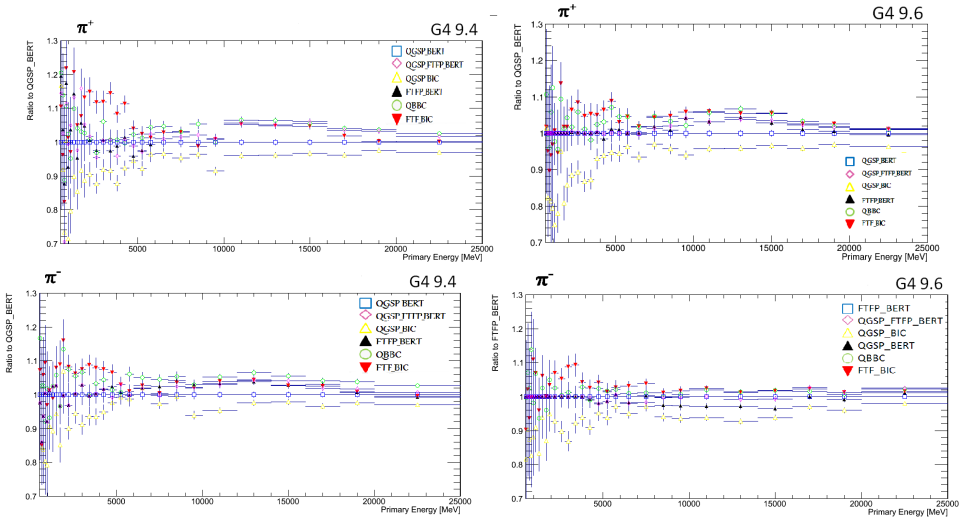


Figure 3.7 – π^+/π^- response variation versus physics list in G4 9.4 and G4 9.6.

3.2.3 Evolution of physics lists in Geant4 9.4 and 9.6 versions

The goal of studying the evolution of physics lists as the calorimeter response to the passage of the particles as a function of the incident energy is to estimate the uncertainties that enter in the Jet Energy Scale (JES) computation. This is necessary because of the JES is a crucial ingredient to several analyses in ATLAS.

For this purpose the evolution of the different physics lists has been evaluated for a set of particles; the charged pions, protons, neutrons and kaons using G4 9.4 and 9.6 releases. The result of this study is illustrated in these figures 3.10, 3.11, 3.15 and 3.16 for different particle energies.

From the plots listed in figures 3.10, 3.11, 3.15 and 3.16 one can see that the response of the charged pions, protons and neutrons is consistent with no significant differences in the evolution of the physics list models, while for kaons the result shows worse agreement between the physics lists, in particular for the K^- and K^+ the response in BIC shows roughly ∼10% shift (see Fig 3.14 and 3.13, the evolution plots of QGSP_BIC and FTF_BIC). Based on these results the physics model used in FTF_BIC for Kaons has been changed from LHEP to BERT for both G4 9.4 and G4 9.6 releases. In addition to the change introduced in QGSP_BIC, which is due to the cross section modification in the model. The complete study is presented in Appendix A.

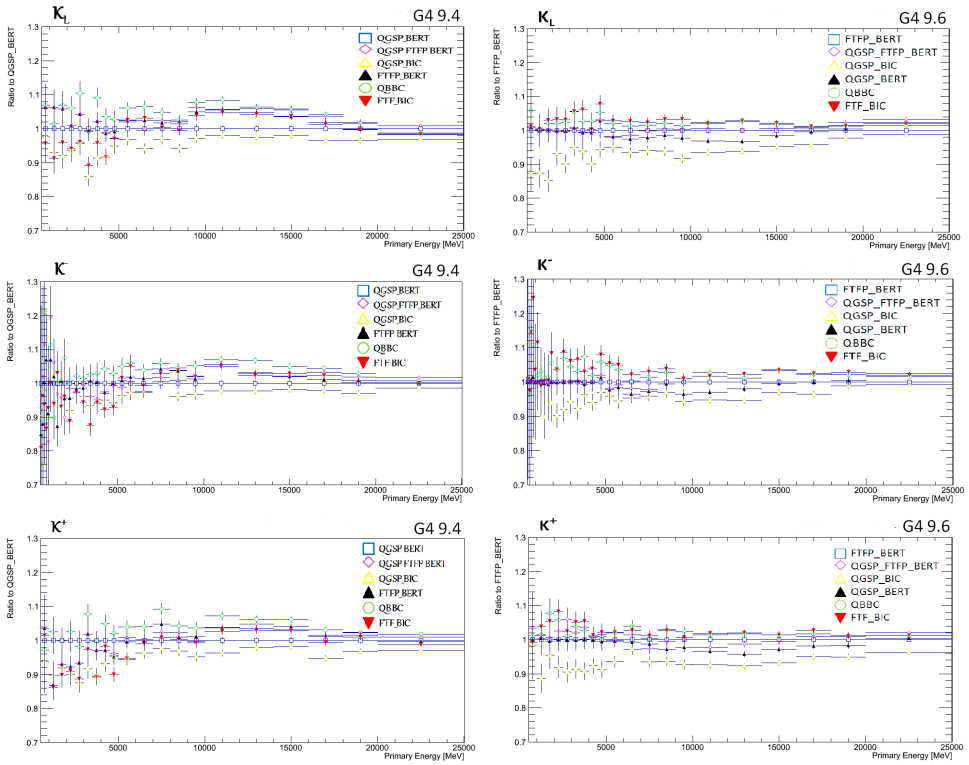


Figure 3.8 – K_L, K^+, K^- response variation versus physics list as a function of primary energy in G4 9.4 and G4 9.6.

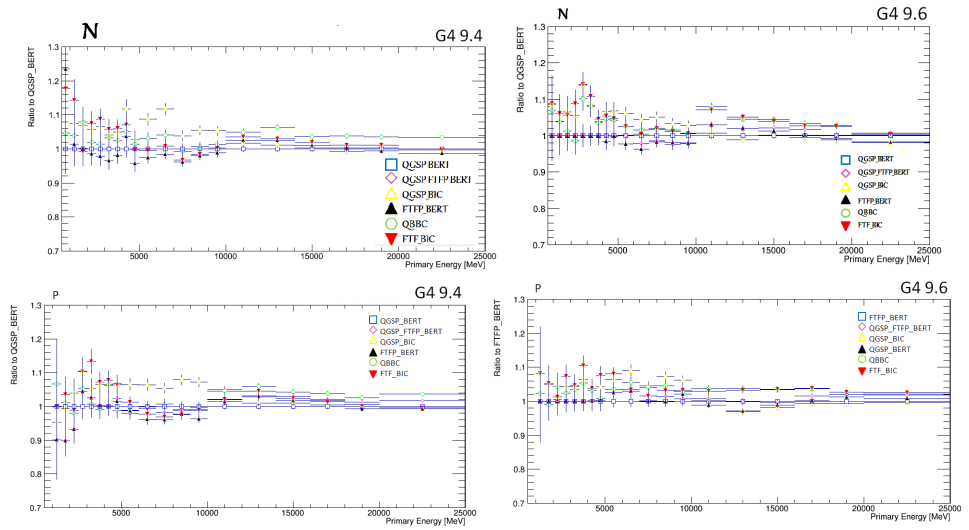


Figure 3.9 – proton, neutron response variation versus physics list as a function of primary energy in G4 9.4 and G4 9.6.

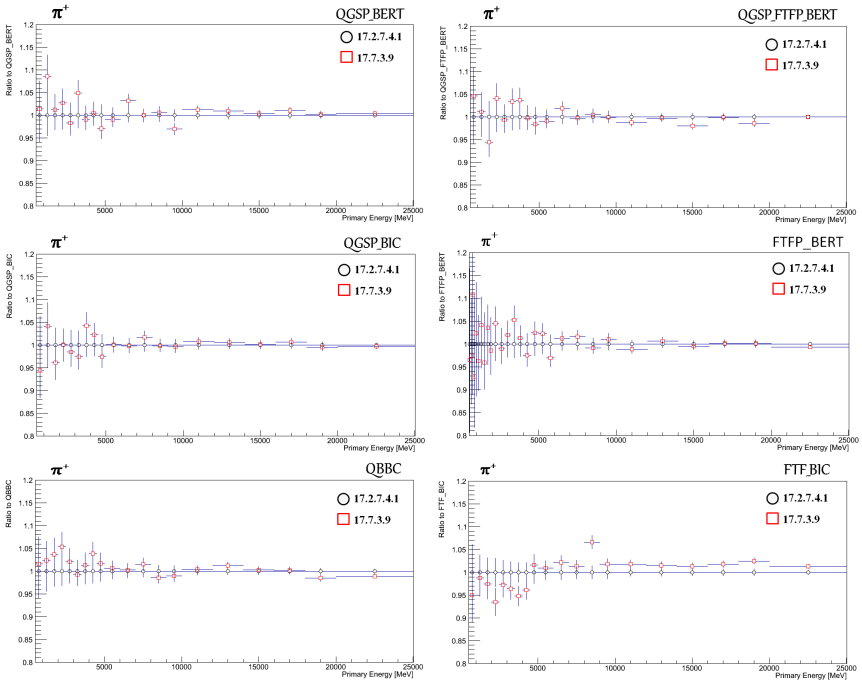


Figure 3.10 – π^+ shows a good agreement between the different physics lists. A minor shift of FTF in FTF_BIC at the high end.

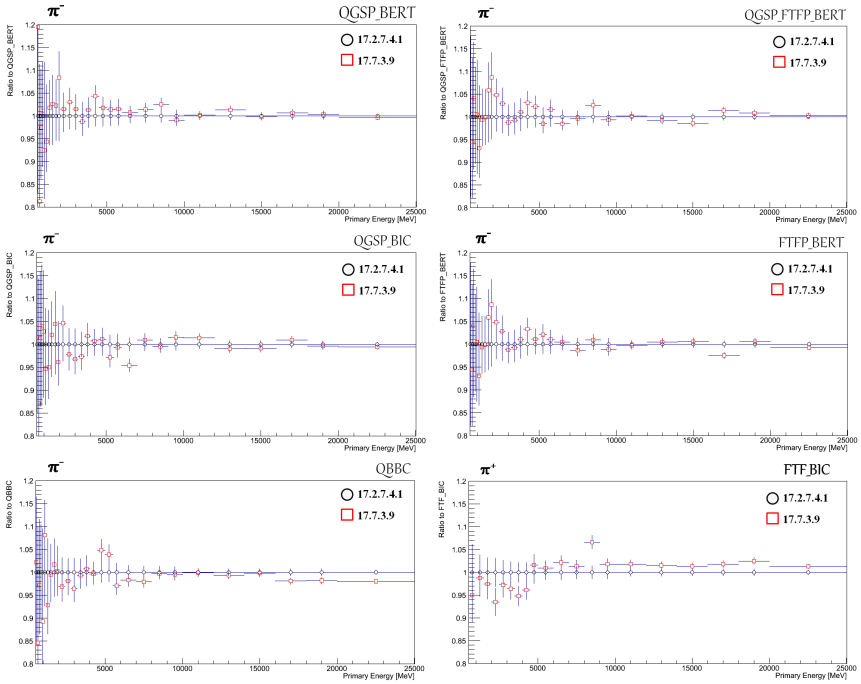


Figure 3.11 – The π^- shows a low shift up in FTF for FTF_BIC, while the other physics lists show a good agreement.

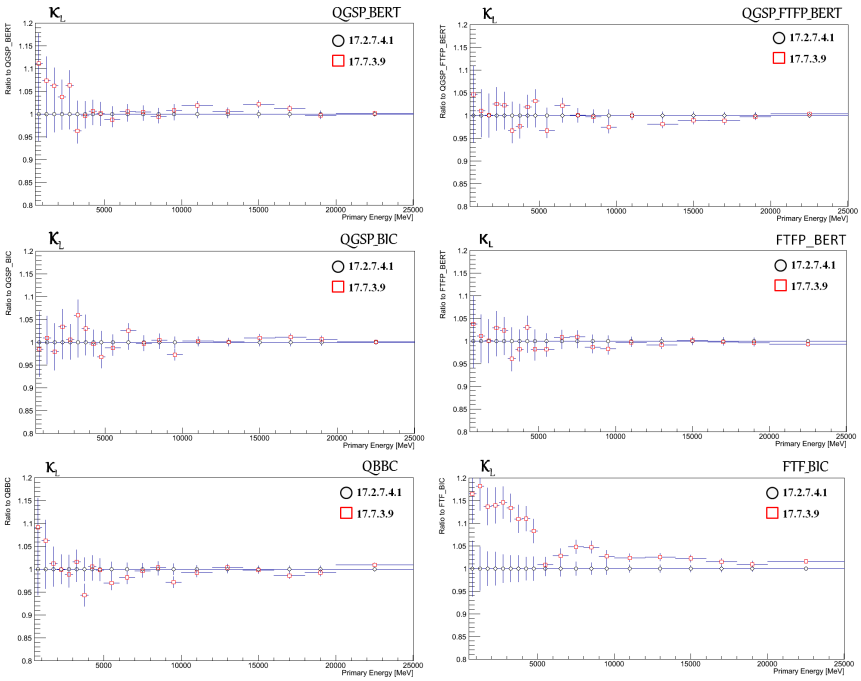


Figure 3.12 – The K_L, shows a shift up in the response in BIC $\sim$ 15% for FTF_BIC but not for QGSP_BIC.

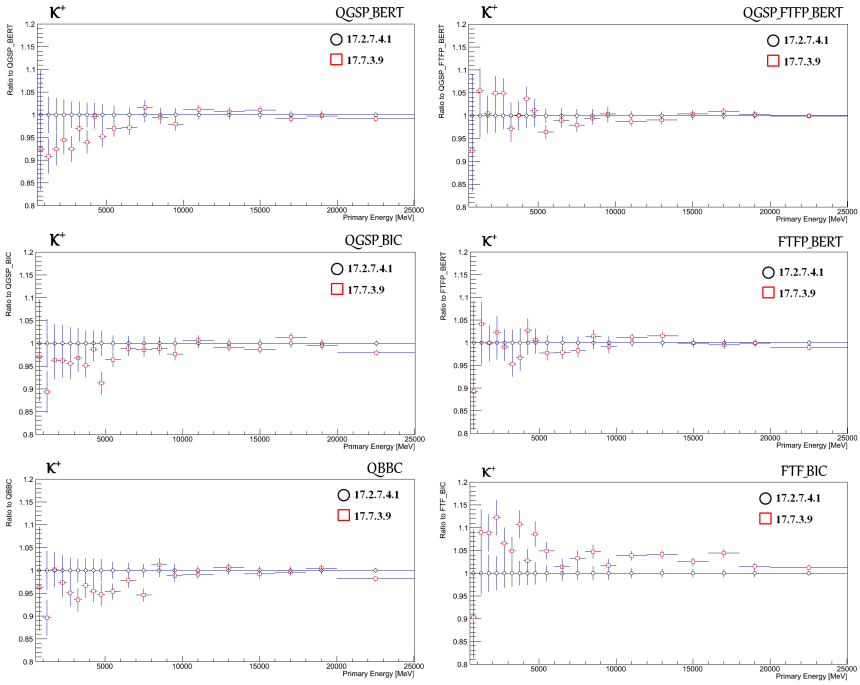


Figure 3.13 – The K^+ shows about 10% shift in BIC response, and it seems also to go in the opposite direction for both G4 versions.

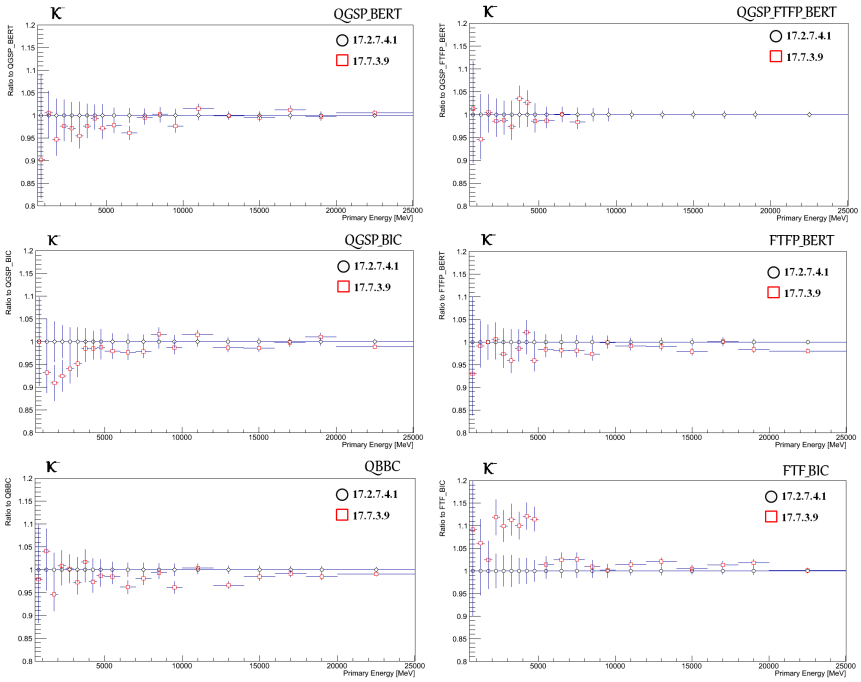


Figure 3.14 – The K^- shows about 10% shift in BIC response, and it seems also to go in the opposite direction for both G4 versions.

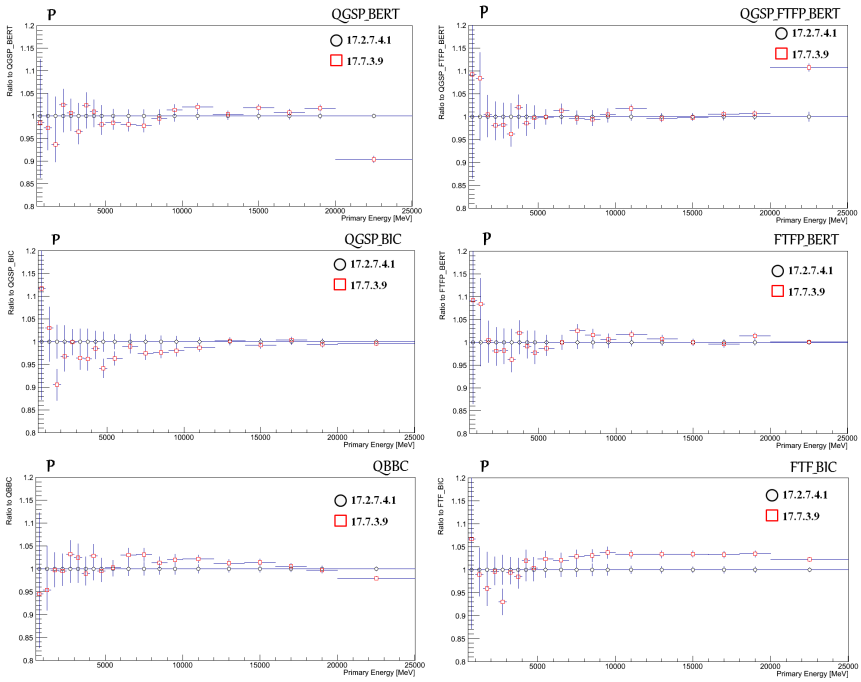


Figure 3.15 – The proton shows a good agreement in the response for the different physics list. A minor shift up ∼3% for FTF in FTF_BIC.

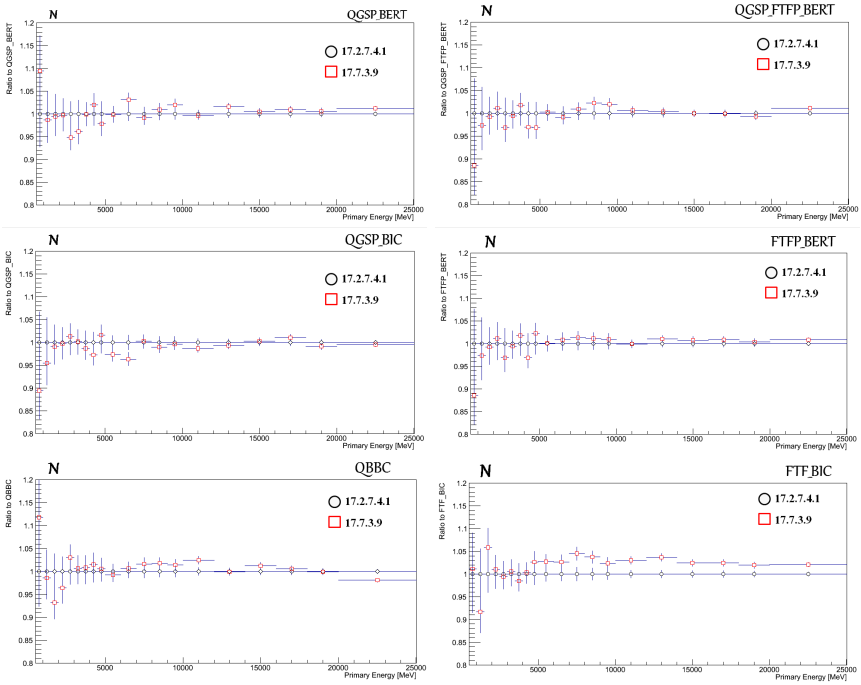


Figure 3.16 – The neutron shows an agreement in response for the different physics lists. Only a low shift up $\sim 3\%$ for FTF in FTF_BIC.

Chapter 4

Object Reconstruction and Particle Identification

As illustrated in the previous chapters, the underlying theory of the lepton+jets final state of the $t\bar{t}$ decay, as well as the experimental setup allowing the identification and the construction of the physical objects is a crucial process. This chapter shows how to establish a connection between the signals grouped and the particles identified and reconstructed by every component of the ATLAS sub-detectors. Separate sections will be thus dedicated to the objects typically occurring in the final state of $t\bar{t}$ decay products such as electrons, muons, jets, b-tagging, and transverse missing momentum, for which the ATLAS combined performance groups carry out detailed studies of these objects, on both real and simulated data, and produce recommendations on the precise definitions to be used in analyses of ATLAS data, including this one.

As the algorithms used to identify the above mentioned objects rely strongly on the primary vertex and tracks, their reconstruction procedures are described as well.

4.1 Tracks and vertices

Tracks of charged particles and primary vertices resulting in a large number of hits are recognized in the ATLAS inner detector. Tracks are used in the electron, muon, and jet reconstruction algorithms. The track reconstruction software has to be robust against pile-up conditions, distinguish the hits from the different charged particles and determine a trajectory that best matches the measurements.

Reconstruction: Tracks are reconstructed using a combination of several tracking

algorithms [67]. The main track reconstruction algorithm is called *inside-out* algorithm which finds the track by extracting space-points (hits) from the pixel and the first SCT layers to form a track seed. These seeds are then extended by adding iteratively hits of the subsequent SCT layers to form a track candidate. This is performed using a simplified *Kalman filtering* [68] and a smoothing approach. Then, the ambiguities in the association of hits and tracks are resolved to reduce the amount of overlapping track candidates sharing one or more hits in the silicon detectors. Fake tracks are rejected by applying quality cuts, in which a certain number of hits in the pixel and SCT detectors are required. Further cuts limit the number of shared hits and the number of *holes*, defined as the passage of the track through a detector element without producing a hit. The track candidates passing these quality requirements are extended into the TRT, where further hits are added.

As the performance of the *inside-out* algorithm depends on a track seed found in the silicon detector, the trajectories of particles, which do not cause a significant number of Pixel or SCT hits, will most be neglected for further processing. This track reconstruction sequence is completed by the so-called *outside-in* algorithm, which starts with hits from the TRT and looks for matching hits inward to the interaction point by adding silicon hits [67] in order to increase the tracking efficiency for e.g. tracks stemming from long-lived particles and recover the remaining trajectory after a large energy loss.

Primary vertex candidates are reconstructed by applying an iterative vertex finding algorithm [69] on tracks that are compatible with originating from the interaction region [70], where all the reconstructed tracks are required to have $p_T > 400$ MeV. In the first step, the algorithm searches for the vertex seeds from the track z-positions along the beam line. The position of this seed and nearby tracks are then used to calculate the vertex position by applying a robust χ^2 fit [71] which makes constraints on tracks that are incompatible with the vertex. Tracks that are displaced by more than 7σ are used as a new vertex seed and the procedure is repeated until no further seed is found in the event. Among several vertices reconstructed in the event, the one with the largest scalar sum of associated track p_T is chosen as the primary vertex for the hard scattering. In top-quark analyses the primary-vertex candidates are required to have more than four reconstructed tracks originating from it, whereas at least 5 tracks associated with the primary vertex are required in the SUSY analysis.

Also secondary and tertiary vertices corresponding to the decay of a heavy flavour hadron contained inside a particle jet are reconstructed using the Inner Detector.

4.2 Electrons

The electron reconstruction is based on the energy deposit (cluster) at the electromagnetic calorimeter (ECAL) and the associated charged tracks in the Inner Detector (ID). The calorimeter signals are used in the first level trigger system (L1), and are combined with tracks to reconstruct electron candidates used for the high level trigger (HLT) decision algorithms. The electron candidates are then further selected against background such as hadrons and electrons originating from photon conversions and heavy flavour hadron decays - using several sets of identification criteria with different levels of background rejection and signal efficiency.

Reconstruction: Electrons follow a curved trajectory in the ID and deposit their energy in the ECAL. Therefore, a track from the ID has to be matched with an energy cluster in the ECAL. The reconstruction starts with finding cluster seeds. A sliding-window algorithm [72] is employed to identify clusters of cells with a size of 3×5 cells in units of 0.025×0.025 and where the transverse energy must be larger than 2.5 GeV. The seed clusters are then matched to tracks reconstructed from hits in the Inner Detector modules using the ATLAS track reconstruction [67]. If a track and a cluster center lie within $\Delta\eta$ <0.05 and $\Delta\phi$ <0.1, the two are considered a match. The larger matching distance in the ϕ direction is to account for *bremsstrahlung*. In the case where several tracks may be matched to the same seed cluster, priority is given to ones showing hits in silicon microstrip detector and the best-match corresponds to tracks representing the smallest distance from the cluster $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$. Finally, to obtain a reconstructed electron, the energy measured in each cell is summed up, and then corrected to account for leakages, as well as for the energy lost in the material before the calorimeter (i.e. in the inner tracker or the solenoid).

Identification: The reconstructed electron candidates are subjected to further refinement criteria with the goal to determine whether the reconstructed electron candidates are signal-like objects or background-like objects such as hadronic jets or converted photons. The baseline ID algorithm used for Run-2 data analyses is the likelihood-based (LH) method [6] which is a multivariate analysis (MVA) technique that simultaneously evaluates several properties of the electron candidates when making a selection decision. The performance of the LH identification algorithm is illustrated in Fig. 4.1.

Three sets of identification criteria are defined, in order to increase background rejection power. These are referred as *loose*, *medium* and *tight*.

- The *Loose* electrons are identified using only information provided by the second layer

of the electromagnetic calorimeter and the hadronic leackage.

- The *Medium* electrons include the loose electron requirements as well as additional requirements on the shower shape variables using the first calorimeter layer and the track quality information provided by the Pixel Detector and SCT.
- The *Tight* electrons definition means a higher background rejection but the efficiency of identification is lower than for the looser selection. It includes the medium electron cuts and additional requirements on the E/p ratio (E is the calorimeter energy and p is the track momentum), it requires also a hit in the vertexing-layer of the pixel detector and information provided by the TRT to reject background caused by hadrons.

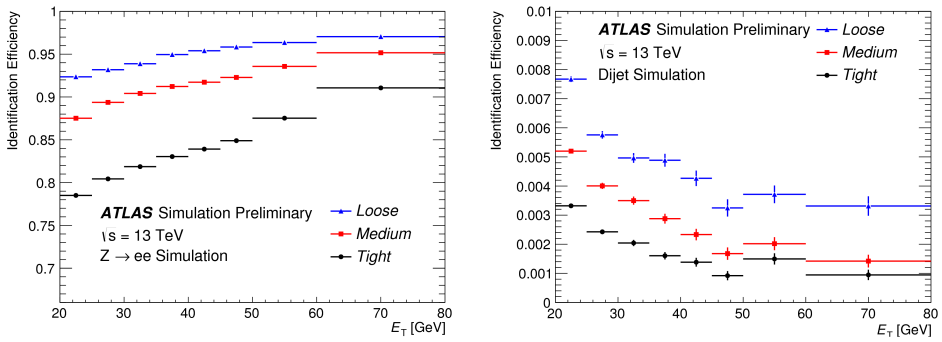


Figure 4.1 – : The efficiency to identify electrons from $Z \longrightarrow ee$ decays (left) and the efficiency to identify hadrons as electrons (background rejection, right) estimated using simulated dijet samples. The efficiencies are obtained using Monte Carlo simulations, and are measured with respect to reconstructed electrons. The candidates are matched to true electron candidates for $Z \longrightarrow ee$ events. Taken from [5, 6].

Calibration: The energy measurement of the electron from clusters in the ECAL needs to be calibrated to fit correctly to the real energy of the particle. The calibration proceeds with the following steps:

- Calibration constants are derived from a MC simulation as an output of a Multivariate Analysis (MVA) optimization comparing the particle's original energy with the energy of EM cluster. The Monte Carlo based calibration has been updated in Run-2 using single particle simulations without pileup with the latest simulation and reconstruction algorithms.
- The differences between the response of the longitudinal layers of the ECAL have to be equalized in data with respect to simulation. They have been derived using data in 2012 and are used in Run-2.
- The derived calibration constants are applied to the reconstructed cluster energies for both MC and data.

- A data-driven corrections optimized to mitigate the non-uniformity and non-stability in detector response is applied to data.
- A simulation-based calibration is applied to data and simulation.
- A data-driven corrections with energy scale factors (applied to data) and correction of the resolution (applied to simulation).
- The response in data is calibrated to the response in simulation using a $Z \longrightarrow ee$ sample. The electron energy in simulation is smeared to fit with the resolution derived from the data.
- At $\sqrt{s}$=13 TeV, the final energy scale for electrons is validated with MC simulation using $Z \longrightarrow ee$.
- The Z resonance is used to set the absolute energy scale of the electrons as the electrons from the Z decays achieve a calibration inaccuracy of less than 0.05%.

4.3 Muons

Reconstruction: Muons are reconstructed firstly by combining information provided by the Muon Spectrometer (MS) and the Inner Detector (ID), while taking into account the small energy deposits in the calorimeters along their trajectory. The information from individual sub-detectors is then combined to form the muon tracks that are used in physics analyses. Muon reconstruction in the MS starts with a search for hit patterns within each layer of the muon chamber to form track segments, then the local track segments from each layer are combined to form a global muon spectrometer track. An MS track must be associated to an ID track, and the latter must have a certain multiplicity of hits in each of the ID sub-detectors. Furthermore, the ID provides an independent measurement of the muon close to the interaction point, and with very high precision spatial and momentum resolution due to the high precision available within the inner detector, so long as the muon candidate can be considered up to $|\eta| < 2.5$.

The hits associated with each track candidate are fitted using a global χ^2 fit. A track candidate is accepted if the χ^2 of the fit satisfies the selection criteria. Hits providing large contributions to the χ^2 are removed and the track fit is repeated. A hit recovery procedure is also performed looking for additional hits consistent with the candidate trajectory. The track candidate is refit if additional hits are found [73].

Depending on the available information from the ID, the MS and calorimeter subsystems one distinguishes the following types of muon candidates, shown also in Fig 4.2

Combined (CB): Muons are reconstructed from tracks measured independently in the ID and MS and then combined into a single track. The majority of muon candidates are reconstructed following an *outside-in* pattern recognition, in which the muons are first reconstructed in the MS and then extrapolated inward and matched to an ID track. An *inside-out* combined reconstruction, in which ID tracks are extrapolated outward and matched to MS tracks, is used as a complementary approach.

Segment-tagged (ST): Tracks reconstructed in ID are extrapolated to the muon spectrometer and associated with at least one MDT or CSC track segment. The parameters matched to such candidates are, however, only those built in the inner detector.

Calorimeter-tagged (CT): muon candidates are described by an ID track points to an energy deposit in the calorimeters which is expected from a minimum-ionising particle. This type of muon has the low purity, but it recovers acceptance in the region where the ATLAS muon spectrometer is only partially instrumented to allow for cabling and services to the calorimeters and inner detector. The identification criteria for CT muons are optimised for that region ($|\eta| < 0.1$) and a momentum range of $15 < p_T < 100$ GeV.

Standalone (SA): Muons trajectory are reconstructed only in the MS. By extrapolating the Muon spectrometer track to the beam pipe, taking into account the energy losses in the calorimeter system and the multiple scattering. This is mostly applicable for the region of $2.5 < |\eta| < 2.7$, where the muon spectrometer exists but the inner detector does not.

The muon candidates used in this analysis are the combined muons, reconstructed according to recommendations of the ATLAS Muon Combined Performance (MCP) group and using the algorithm defined as Chain 3 in [74]. The overall identification and reconstruction efficiency for muons evaluated in 2015 and before is done using both $Z \longrightarrow \mu^+\mu^-$ and $J/\psi \longrightarrow \mu^+\mu^-$ decays and using *tag-and-probe* method. More details on the method are available in [75]. Fig 4.3 shows the reconstruction efficiency results obtained for the Medium identification algorithm for $Z \longrightarrow \mu^+\mu^-$ as a function of the muon pseudo-rapidity. High efficiencies, close to 99%, are observed over the full η range with the exception of the central region ($|\eta| < 0.1$).

Identification: Muon identification is performed by applying a set of quality requirements that suppress background mainly from fake muons coming from pion and kaon decays and guarantee a robust momentum measurement. The measurement of the detector activity around a muon candidate, referred to as *muon isolation*, is therefore a powerful tool for background rejection in the current physics analyses.

Track isolation variables are defined by the sum of the transverse momentum of all tracks (p_T^{trk}) in $\Delta R < 0.2$ or 0.3 around the muon candidate except for muon p_T itself:

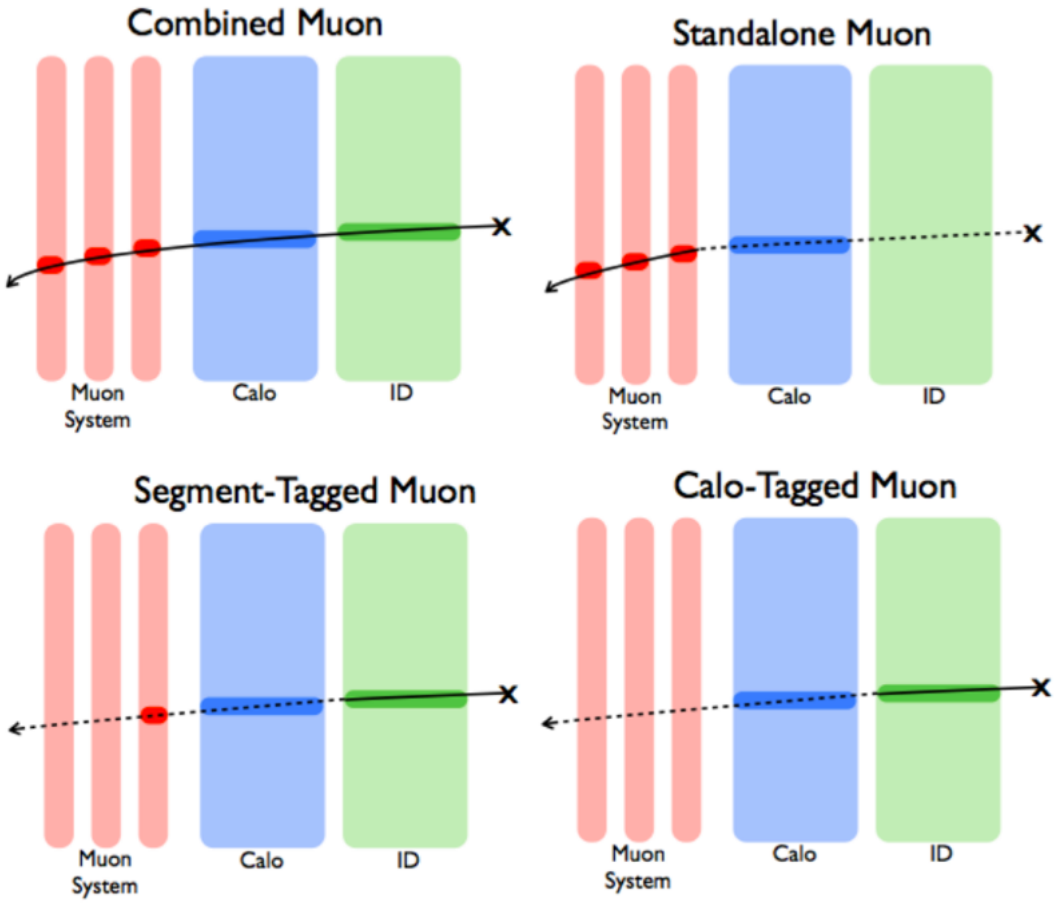


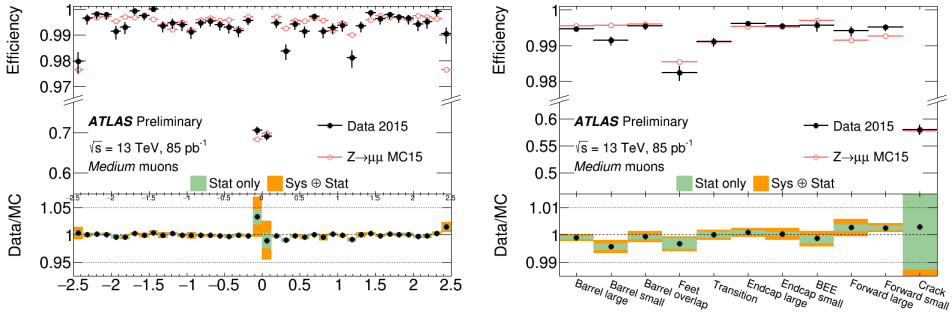
Figure 4.2 – Schematic representation of the different muon reconstructions methods.

$$I_{0.2}^{trk} = \sum_{\Delta R<0.2} p_T^{trk} \tag{4.1}$$

$$I_{0.3}^{trk} = \sum_{\Delta R<0.3} p_T^{trk} \tag{4.2}$$

For muons with $p_T < 25$ GeV, in where the signal regions in this analysis are defined (see Chapter 5), the muon medium [1] isolation requirement as: $I_{0.3} < 0.12\ p_T$ is applied.

[1]The *Medium* identification criteria provides the default selection for muons in ATLAS. This selection minimizes the systematic uncertainties with muon reconstruction and calibration. Only CB and SA tracks are used in this selection. The former are required to have at least three hits on at least two layers of MDT, except for tracks in the $|\eta| < 0.1$ region, where tracks with at least one MDT layer but no more than one MDT hole layer are allowed. The latter are required to have at least three MDT or CSC layers, and are employed only in the $2.5 < |\eta| < 2.7$ region to extend the acceptance outside the ID geometrical coverage. Moreover, A very loose selection on the compatibility between ID and MS momentum measurements is applied to suppress the contamination due to hadrons misidentified as muons. Specifically, the q/p significance is required to be less than seven. In the pseudo-rapidity region $|\eta| < 2.5$, about 0.5 % of the muons classified as Medium originate from the inside-out combined reconstruction strategy.



(a) (b)

Figure 4.3 – Muon reconstruction efficiencies for the Medium identification algorithm measured in $Z \longrightarrow \mu^+\mu^-$ events (a) as a function of the muon pseudorapidity and (b) for various regions of the ATLAS muon spectrometer. The prediction by the detector simulation is depicted as red circles, while black dots indicate the observation in collision data. Green error bands indicate the statistical uncertainty, and orange bands the quadratic sum of statistical and systematic uncertainty. The bottom panel shows the ratio between expected and observed efficiencies. Taken from [7].

The muon's tracks must go past the beam spot: $|d_0/\sigma d_0| < |d_0/\sigma d_0|_{max}$ =3 and the cuts on the longitudinal and transverse impact parameters are applied for the further suppression of backgrounds from in-flight decay of hadrons: $|z_0 sin\theta| < 0.5$ mm (See section 5.3).

4.4 Jets

Jets are objects composed of particles produced by the hadronization of quarks and gluons, which are produced from high energy collisions at the LHC. Jets are defined by the jet algorithm used for their reconstruction and the type of objects chosen as input to this algorithm. Indeed, particles passes through the ATLAS detector produce tracks in the ID and energy deposits inside the calorimeters. These detector signals allow the reconstruction of track jets (reconstructed using track information. see Sec. 4.1) and calorimeter jets (reconstructed using calorimeter information). This section will then focus in explaining the jet reconstruction process for calorimeter jets which consists in three steps: the definition of calorimeter signals, the use of a jet reconstruction algorithm to group the calorimeter signals and finally the jet calibration which corrects the jet energy and momentum for the effects of ATLAS calorimeters non-compensation, dead material, leakage, out of cone and other thresholds effects.

4.4.1 Calorimeter jets

The mechanism of reconstruction starts either from calorimeter towers or topological clusters (*topo-clusters*), considering only towers or topo-clusters with a positive energy. Tower jets consist of radial towers of calorimeter cells with a fixed size of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ in the η-ϕ plane [76]. Tower jets are mostly used at trigger level as the reconstruction of calorimeter towers is considerably faster than that of topological clusters.

Calorimeter jets built from topological clusters are used in most physics analyses in ATLAS. The idea of topological clusters is to follow the shower development in the calorimeter while suppressing efficiently the calorimeter noise. Their formation starts from seed cells with a large signal to noise ratio (above 4). Neighbouring cells with signal to noise ≥ 2 are added iteratively, and finish by including all direct neighbor cells on the outer perimeter. The signal to noise is calculated from the absolute energy deposited in a calorimeter cell divided by the RMS of the energy distribution. The latter is obtained from events triggered at random bunch crossings. In the last step of the cluster formation, a ring of *guard cells*, consisting of all cells neighbouring the proto-cluster, is included. A splitting procedure based on the presence of multiple energy maxima in a cluster is applied to better separate showers initiated by close-by partons. Further details can be found in [77].

The cluster energy is defined to have an energy equal to the sum of energies of all its constituent cells, its mass is set to zero and its angular position is obtained from the energy-weighted averages of the pseudo-rapidities and azimuthal coordinates of the constituent cells. The minimum size of a topological cluster is $\Delta\eta \times \Delta\phi = 0.3 \times 0.3$ for a topological cluster consisting of a single seed cell surrounded by a ring of *guard cells* and assuming a cell size of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ in the central region of the calorimeter.

4.4.2 Jet Algorithms

The main jet finding algorithm used in this study to measure the showers is the so-called **anti-k_T** algorithm [78] which is a sequential recombination algorithm. It allows to reconstruct a four-momentum that would be related to the particle that initiated the shower and lead to the jet formation. In this algorithm, sequential recombination algorithms take topoclusters or towers as input and combine them to form jets according to the distance calculated between all the clusters i and j in the event d_{ij}. A *beamline* distance d_{iB} is also calculated for each particle i:

$$
\begin{aligned}
d_{ij} &= min(p_{T,i}^{2n}, p_{T,j}^{2n}) \frac{\Delta R_{ij}^2}{R^2} \\
d_{iB} &= p_{T,i}^{2n}.R_0^2
\end{aligned}
\tag{4.3}
$$

where p_T refers to the transverse momenta of the object, $\Delta R_{ij}=\sqrt{\Delta\eta^2 + \Delta\phi^2}$ is the distance in the $\eta \times \phi$ plane between the pair of four-momenta and R is a parameter of the algorithm anti-k_T that approximately controls the size of the jet and is chosen to be either R = 0.4, or R = 1.0 for this Analysis.

The topoclusters and towers are defined massless. However, their distribution inside the reconstructed jet leads the jet to have a given mass. In Monte Carlo simulations track jets and calorimeter jets are reconstructed as in real data. In addition, two other types of jets can be defined within ATLAS:

Parton jet which refers to the parton at the matrix-element level which causes the particle shower due to the fragmentation process.

Particle or truth jet are reconstructed from simulated particles produced by the hadronization models of the different Monte Carlo generators, before any detector simulation and with a life time τ that meets the requirement $c_\tau > 10$ mm. This mainly includes stable particles: electrons, photons, pions, kaons, protons and neutrons and their antiparticles. Neutrinos and muons are excluded, since they leave a little or no measurable signal in the calorimeter.

4.4.3 Calibration

The calorimeter provides only a partial measurement of the total energy deposited by hadrons mainly for the following reasons:

- Non instrumented regions or energy losses in inactive regions (dead material) of the detector.
- Leakage of energy deposits outside the calorimeters.
- Incomplete reconstruction of the jet cone, where particles inside the jet are not included in reconstructed jets
- Signal losses due to inefficiencies in calorimeter clustering and jet reconstruction (Thresholds effects).

ATLAS has developed several calibration strategies with different levels of complexity and different sensitivity to systematic effects [79], which made them complementary in the way they contribute to the understanding of the jet energy scale measurement:

EM+JES calibration is a simple Monte Carlo-derived jet calibration scheme used in ATLAS for the first physics analyses. EM+JES applies corrections that restores the reconstructed jet energy to the true jet energy. In this scheme the energy corrections depend on a correction factor which is a function of the reconstructed jet energy at the electomagnetic scale and pseudorapidity. In addition to this energy correction a pile-up as well as a jet origin correction are also applied.

Global Sequential (GS) Calibration is a Monte Carlo-derived jet calibration, which uses longitudinal and transverse properties of the jet structure sequentially to improve the jet energy resolution, while the average jet energy scale remains unchanged. In this scheme jets are reconstructed from topoclusters or towers, then sequential corrections are applied based on the jet-by-jet correction factor which depends on the jet p_T, η and several longitudinal and transverse jet properties.

Global Cell Energy-Density Weighting Calibration (GCW) attempts to compensate for the different calorimeter response to hadrons and electromagnetic energy depositions by weighting each jet constituent cell. The weights, which depend on the cell energy density and the calorimeter layer only, are determined by minimizing the energy fluctuations between the reconstructed and particle jets in Monte Carlo simulation. Jets are found from topoclusters or calorimeter towers at the EM scale, then cells are weighted and a final jet energy scale correction is applied to ensure that good linearity response is achieved.

Local Cluster Weighting (LCW) Calibration [80] classifies topoclusters as hadronic or electromagnetic, based on their properties (such as their energy, depth in the calorimeter, cell energy density, fractional energy deposited in the calorimeter layer and energy measured around it) and then calibrate them individually before applying jet reconstruction. The weights are determined from Monte Carlo simulations of charged and neutral pions.

The final stage of the jet energy calibration consists of various residual corrections, which are applied in the form of calibration factors. To derive the systematic uncertainties on the JES, the jet p_T in data is compared to the one in simulation by exploiting the method called direct p_T balance (DB) between the jet p_T and the p_T of a well measured reference object as follows:

$$[P_T^{jet}/P_T^{ref}]_{data}/[P_T^{jet}/P_T^{ref}]_{MC} \tag{4.4}$$

Firstly, dijets events are exploited as reference to test the direct p_T balance between

a jet and a photon or Z boson decaying to electron or muon pairs. Then events where a system of low p_T jets recoils against a high p_T jet are used to calibrate jets up to the TeV scale.

4.5 b-tagging

In particle colliders, when a bottom quark is produced in an event, it hadronizes into a B-hadron, which has a relatively long lifetime of approximatly 1.5 ps, leading to a few millimiters ($\sim$ 3 mm) travel path in the detector before decaying. it results in a secondary vertex, displaced with respect to the primary interaction vertex (PV). Several sophisticated b-jets (referred to as b-tagging) algorithms have been developped in ATLAS which exploit different properties to determine a probability (b-tag weight) for the jet to originate from the fragmentation of a b-quark. The algorithm used in this physics Analysis is called *MV2c20* [81] which uses a neutral network that combines the output weight of other algorithms, named "IP2D, IP3D", "SV1" and "JetFitter", which are described below and explained in more detail in Ref [82].

- The **IP2D and IP3D: The Impact Parameter based Algorithms** use the transverse and the longitudinal impact parameters of tracks associated with the b-jet candidate. The transverse impact parameter d_0 is defined as the distance of closest approach of the track to the primary vertex point in $r\phi$ projection, while the z coordinate of this point is referred to as the longitudinal impact parameter z_0. IP2D algorithm is using as input only the transverse impact parameter, while IP3D uses both transverse and longitudinal components and their correlation. Their limited resolution is taken into account by using their significances, $d_0/\sigma(d_0)$ and $z_0/\sigma(d_0)$,), instead of the raw values, it gives more weight to tracks measured more precisely.

- The **Secondary Vertex Finding Algorithm (SV)** algorithm attempts to reconstruct an inclusive vertex formed by the decay products of a b-hadron. it combines the remaining two track vertices into a single secondary vertex and exploits this secondary vertex properties, like the invariant mass of all tracks associated to the vertex, the ratio of energy from tracks associated with the secondary vertex to all the jet tracks, the number of 2-track vertices and the direction of the line joining the primary and secondary vertex with respect to the jet axis. These observables are combined using a likelihood ratio technique to assign a b-tagging weight. Details of the likelihood method are given in Section 6 of [83]

- The **Multi-vertex fit (JetFitter)** algorithm exploits the weak b- and c-hadron de-

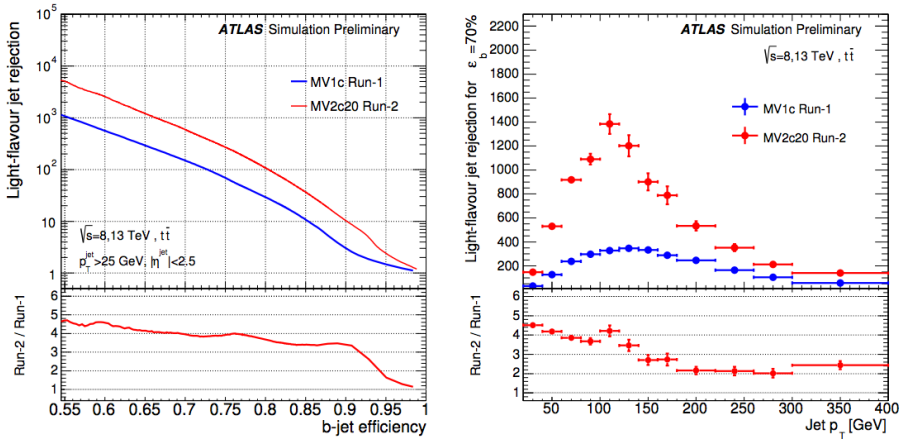
cays topologies inside the jet, as well as information about an approximated flight path of the b-hadron by fitting a common line between PV and the SVs of the b- and c-quark decays using a *Kalman filter* [68]. The advantage is that it can separately identify the b and c vertex, which would otherwise be merged by the procedure employed by SV1 for example.

- The **Multivariate algorithm (MV2)** combines various discriminant observables from the basic above algorithms into a boosted decision tree (BDT) based algorithm [84]. This default algorithm for Run-2, *MV2c20*, is defined as output to a BDT which is trained using b-jets as signal and a mixture of 80% light-flavour jets and c-jets as background (the amount of c-jets in the background is equal to 20% of the amount of light-jets). The kinematic properties (p_T and η) of the jets are included in the training to take advantage of correlations with the other input variables. *MV2c20* constitutes a significant revision of Run-1 main b-tagging algorithm, denoted *MV1*, which combined the outputs of the various b-tagging algorithms using neural network approach. The *MV2c20* algorithm provides better performance and simplifies the software maintenance.

Different methods can be used to measure the efficiency of the b-tagging algorithms in data. Figure 4.4 shows a comparison between the performance of the default Run-2 b-tagging algorithm MV2c20, compared to the default Run-1 b-tagging algorithm, MV1c: light jet rejection vs b-jet efficiency (a) and light jet rejection as a function of jet p_T given a fixed b-jet efficiency of 70% in each bin (b). Comparing the two algorithms, the light-flavour jet rejection is improved by a factor of about 4 compared to run-1 for a 70% b-jet efficiency. This corresponds to a relative 10% improvement in the b-jet efficiency. Improvement at low and medium p_T is mostly due to the addition of the IBL, while the improvement at high p_T is due to algorithm improvements.

4.6 Missing transverse Energy

The missing transverse momentum, or E_T^{miss} is a fundamental observable to spot the large part of the event's energy carried away by the neutrino or possible BSM particles weakly/not interacting with the detector. It can be reconstructed using the difference between the initial state and final state total momentum since the overall momentum of the pp collision at the LHC should sum to zero, as the colliding protons have equal energies and opposite direction. Possible sources of fake E_T^{miss} can also result from insufficient detector coverage, dead or noisy regions and finite detector resolution. E_T^{miss} thus can also serve as an important



(a) (b)

Figure 4.4 – Performance of default Run-2 b-tagging algorithm MV2c20 and the equivalent Run-1 b-tagging algorithm MV1c in simulated $t\bar{t}$ events: light jet rejection vs b-jet efficiency (a) and light jet rejection as a function of jet p_T for a fixed b-jet efficiency of 70% in each bin (b) [8].

measure of the overall event reconstruction performance.

The E_T^{miss} is defined as the event momentum imbalance in the plane perpendicular to the beam axis and calculated as the negative vectorial sum of the transverse energy from every significant energy deposit in the detector, thus creating the vector required to balance the event. The energy deposits in the calorimeters are associated with identified and calibrated high-p_T objects: electrons, photons, hadronically decaying τ-leptons, jets, soft jets and muons. Calorimeter cells and tracks which are not associated to any high-p_T object are also taken into account making up the soft terms. The distinct E_T^{miss} contributions can be expressed as:

$$E_{x(y)}^{miss} = E_{x(y)}^{miss,e} + E_{x(y)}^{miss,\gamma} + E_{x(y)}^{miss,\tau} + E_{x(y)}^{miss,jets} + E_{x(y)}^{miss,SoftTerm} + E_{x(y)}^{miss,\mu} \tag{4.5}$$

and the missing transverse energy is $E_T^{miss} = |\vec{E_T^{miss}}|$.

Each term in Equation 4.5 has his respective calibrations applied, they are in principle affected by the pile-up which spoils the balance in the transverse momenta and worsens the E_T^{miss} resolution. Figure 4.5 shows the E_T^{miss} resolution evaluated in $Z \to \mu\mu$ data and MC events in which no genuine E_T^{miss} is expected. It is plotted as a function of the $\sum E_T$ in the event, calculated using the calorimeter-based soft term, and as a function of the number of primary vertices. The choice of the calorimeter-based soft term allows comparisons with other studies [85].

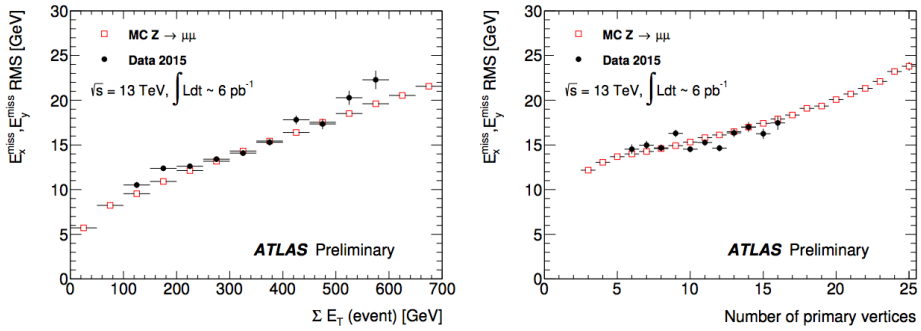


Figure 4.5 – Distributions of Track Soft Term E_x^{miss}, E_y^{miss} resolution as a function of $\sum E_T$ and of the number of primary vertices in $Z \rightarrow \mu\mu$ events. Reasonable agreement is found between data (black circles) and MC simulation (red squares) [9].

Chapter 5

Top-antitop resonances search at $\sqrt{s} = 13$ TeV

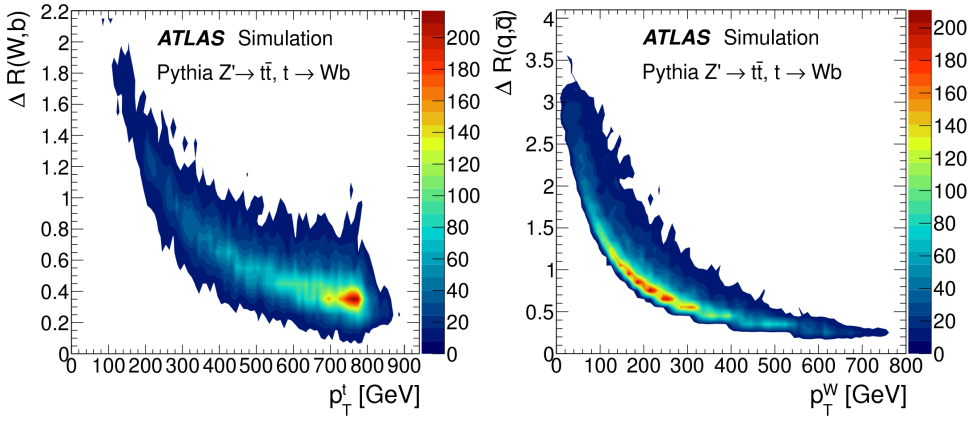
This chapter mainly covers an analysis designed to identify new physics that produce resonances in the top-antitop ($t\bar{t}$) mass spectrum. A variety of new physics scenarios give rise to heavy particles that decay primarily to $t\bar{t}$, thus, the implemented method for this analysis is quite general and can be used to test any model that predicts top-antitop production through non-Standard Model channels, however only the topcolour-assisted-technicolor (TC2) (discussed in Sec. 1.5) benchmark model was used to set limits on the existence of the Z' new particle.

The analysis is organized as follows: A brief outline of the analysis strategy is given in the first Section 5.1; Then details on data and background samples are given in Section 5.2; The event selection is outlined in Section 5.3 and the relevant aspects of the data driven backgrounds methods are given in Section 5.4. Section 5.5 details the reconstruction of the $t\bar{t}$ system. Section 5.6 discusses the diffent systematic uncertainties considered in the analysis; Section 5.7 confronts background expectations with data; Section 5.8 presents the results of the search in terms of expected and observed limits on resonant $t\bar{t}$ production and summarizes the analysis.

5.1 Analysis strategy

The goal of the analysis is to search for new resonance decaying to top-antitop event in the lepton (electon or muon) + jets channel. This final state (introduced in Sec. 1.2.2) is experimentally favored since it allows to effectively suppress the large background contribution from QCD multijet events in the analysis. In this channel, one of the top quarks decays

into a b-quark and a W-boson, which decays into a single lepton and a neutrino. The other top quark generates a final state with two quarks as a result of the W-boson decay and a b-quark. For new heavy particles, whose mass is beyond the TeV scale, the emitted top quark can disintegrate with a major transverse momentum ($p_T > m_t$). This decay topology will be referred to as the *boosted topology*, and the top quark disintegration products are then sent in the same angular region of the detector. This phenomenon is shown in Figure 5.1, where the distances are shown in the (η, ϕ) plan, between the decay products of the hadronic and leptonic top quarks as a function of the transverse momentum of the top quark and W boson. In this case, if no special treatment is given to this final state configuration, these events might be rejected, as background candidates, or $m_{t\bar{t}}$ might be poorly estimated for them.



(a) $t \to Wb$ (b) $W \to q\bar{q}$

Figure 5.1 – Angular separation ΔR of the decay products in the two-body decays (a) $t \to Wb$, and (b) $W \to q\bar{q}$ as a function of the transverse momentum of the top quark and the W boson, respectively. Taken from [10].

However, If the top quarks have low transverse momenta, the decay products are generally well separated in the detector and can be reconstructed individually: The hadronic top-quark decay results in three separate jets, a b-jet and two other jets from the W-boson decay. The lepton from the decay of the other top quark is usually well separated from the corresponding b-jet. This decay topology, will be referred to as the *resolved topology* (typically $p_T < 500$ GeV) in the following. Between the resolved and boosted regimes, there is an intermediate topology, called *resolved-boosted hybrid* approach, where the objects of the final state are partially collimated [1]. Although, for large momentum (above 500 GeV. Fig 5.2), the distance between objects becomes quickly quite lower to the standard reconstruction

[1] For the resolved-boosted hybrid approach only two of the jets from the hadronically decaying top quark, typically those from the decay of the W boson, are merged while the third jet is reconstructed separately.

parameters of the jets.

In this thesis, only the boosted approach is applied on this search for $t\bar{t}$ resonance.

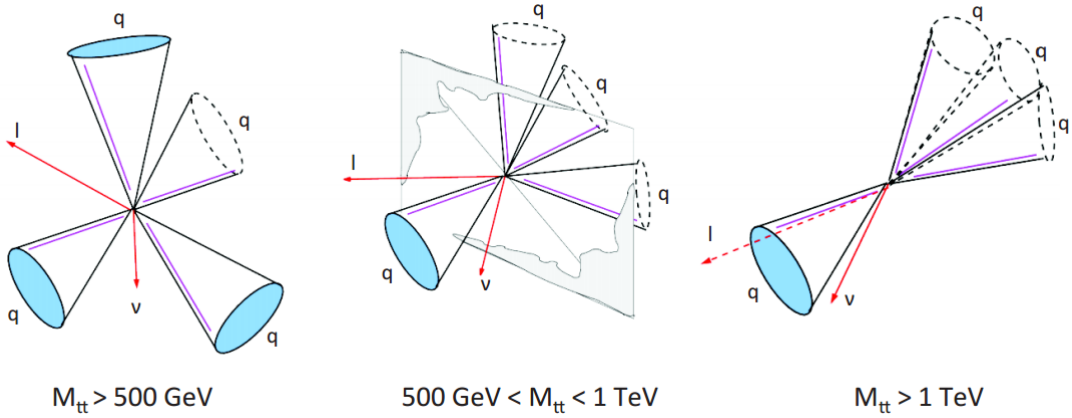


Figure 5.2 – The different decay topologies of the top-antitop system. From left to right, the resolved decay topology, the resolved-boosted-hybrid decay topology and finally boosted decay topology.

The observable used to detect the resonance is the invariant mass of the top-antitop system, ($m_{t\bar{t}}$) which is reconstructed from certain physics objects in the final state of the decay and only the *boosted topology* is applied on the whole spectrum. Using this approach, the neutrino is not detected but expected through the missing transverse energy (E_T^{miss}), the lepton and the b-jet from the leptonic decay of the top might be very close together, which results in loosening the lepton isolation criteria. The b-jet from the leptonic decay of the top quark is reconstructed as a small-R jet and the hadronic top decay should have all three quarks from the top decay in a single high-transverse-momentum large-R jet, which is an Anti-k_T R = 1.0 jet. This procedure splits the analysis in two channels, depending on whether the final state lepton is an electron or muon: *boosted electron-channel* and *boosted muon-channel*. [2]

The selected events in the two categories are then reconstructed based on the assumption that the final state signature stems from the semileptonic decay of a $t\bar{t}$ pair. The aim is to well reconstruct the invariant mass of the $t\bar{t}$ system, $m_{t\bar{t}}$, from the selected objects in order to study the $m_{t\bar{t}}$ spectrum and search for deviations from the SM expectations. Once $m_{t\bar{t}}$ is estimated for both channels, the significance of a data deviation from the

[2]The selection strategy enforce an orthogonality between the boosted and resolved events, since if an event does not pass the selection criteria defined for the boosted analysis event, it passes as input to the resolved one. This way, we maximize the reconstruction efficiency, particularly in the *resolved-boosted-hybrid* analysis, where the establishment of a selection is more difficult since, we do not know how configuration occurs for partial collimated objects. In this analysis only events that pass the boosted topology are used.

background expectation is evaluated by taking into account systematic variations in the background estimation which may partially or completely account for the deviation. If no significant deviation is observed and upper exclusion limits on the production cross-section times branching ratio are derived. The model is excluded with 95% Confidence Level for a certain resonance mass parameter range.

The data which are used to perform this search, was collected by the ATLAS detector at the Large Hadron Collider from proton-proton collisions at a centre-of-mass energy of 13 TeV in the year of 2015. The integrated luminosity of data samples is 3.2 fb^{-1}. It was only recorded for data with the whole relevant sub-detector systems operational and during stable beam conditions.

5.2 Data and Monte Carlo samples

The data used are those in the All Good runlist produced with MC15 in the 13 TeV 2015 running. The Lepton+jets events are collected using single-electron and single-muon triggers with transverse momentum thresholds chosen such that the efficiency is uniform for leptons satisfying offline selections.

The main background processes for this analysis are the *irreducible* Standard Model $t\bar{t}$ production, production of a W boson, Z boson in association with additional jets (W+jets, Z+jets), single-top quark production and diboson production. All of them are evaluated from MC samples. The multi-jets background is directly evaluated from the data.

Simulated Monte Carlo (MC) data samples are used for signal processes, as well as for background processes that produce jets and prompt leptons. The MC samples are employed to develop the event selection, provide SM background estimates, and evaluate signal efficiencies.

The Standard Model $t\bar{t}$ Monte Carlo samples [86] and single-top-quark production in the Wt [87] and s-channel [88] are generated by the `Powheg-Box` v2 [89, 90] generator with the CT10 [91, 92] parton distribution function (PDF) set in the matrix element calculations. Overlap between $t\bar{t}$ and Wt production is treated within the diagram removal (DR) scheme [93]. Electroweak t-channel single top-quark events are generated using the `Powheg-Box` v1 generator [94]. This generator uses the four-flavour scheme for the next-to-leading-order (NLO) matrix elements calculations together with the four-flavour PDF set CT10f4. For this process, the top quarks are decayed using `MadSpin` [95] preserving all spin correlations. For all SM top-quark processes the parton shower, fragmentation, and the underlying event are simulated using `Pythia v6.428` [96] with the CTEQ6L1 [97] PDF set and the corresponding `Perugia` 2012 tune (P2012) [98]. The top mass is set to 172.5 GeV. The `EvtGen` v1.2.0

program [99] is used to model the decays of heavy-flavour hadrons. For the generation of $t\bar{t}$ events, the h_{damp} parameter, which controls the transverse momentum of the first additional emission beyond the Born configuration, is set to the mass of the top quark. The main effect of this is to regulate the high transverse momentum emission against which the $t\bar{t}$ system recoils. The top quark kinematics in all SM $t\bar{t}$ samples were corrected to account for electroweak higher-order effects [100]. This correction is applied after generating the samples, by applying a weight that depends on the flavour and centre-of-mass energy of the initial partons, and the decay angle of the tops in the centre-of-mass frame of the initial partons. The value of the correction factor falls with the invariant mass of the $t\bar{t}$ system from 0.98 at a mass of 400 GeV to 0.87 at a mass of 3.5 TeV.

The W+jets and Z+jets backgrounds are simulated using the `Sherpa` 2.1.1 [60] generator. Matrix elements are calculated for up to two partons at NLO in QCD and four partons at leading order (LO) using the `Co-mix` [61] and `OpenLoops` [101] matrix element generators and merged with the `Sherpa` parton shower [102] using the `ME+PS@NLO` prescription [103]. The CT10 PDF set is used in conjunction with dedicated parton shower tuning developed by the `Sherpa` authors. The W/Z + jets events are normalised to the next-to-next-to-leading-order (NNLO) cross sections [104].

Diboson (WW, WZ, ZZ) production processes with four charged leptons (4ℓ), three charged leptons and one neutrino ($3\ell + \nu$), two charged leptons and two neutrinos ($2\ell + 2\nu$), or one charged lepton and one neutrino plus jets ($\ell\nu q\bar{q}'$) are also simulated using the `Sherpa` 2.1.1 generator. Matrix elements contain all diagrams with four electroweak vertices. They are calculated for zero ($3\ell+\nu, lq\bar{q}'$) or up to one ($4\ell, 2\ell+2\nu$) additional partons at NLO in QCD and up to 3 partons at LO using the `Comix` and `OpenLoops` matrix element generators and merged with the `Sherpa` parton shower using the `ME+PS@NLO` prescription. The same PDF set and parton shower tuning as for the W/Z + jets events are used. The generator cross sections are used for sample normalization in this case.

The signal process $pp \to Z' \to t\bar{t}$ is simulated using `Pythia` v8.165 [105] with the NNPDF23LO [106] PDF set and the A14 [107] tune. This search uses topcolour-assisted-technicolour Z'_{TC2} as a benchmark. The properties of the boson are controlled by three parameters: $\cot\theta_H$, which controls the width and the production cross section, and f_1 and f_2, which are related to the coupling to up-type and down-type quarks, respectively (See also Sec. 1.5). Here $f_1 = 1$ and $f_2 = 0$, which maximises the fraction of Z'_{TC2} that decay to $t\bar{t}$. The parameter $\cot\theta_H$ is tuned for each mass point such that the resonance has a width of 1.2% (or 3%) of its mass. 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 [108, 109]. The interference of these signals with SM electroweak $t\bar{t}$ production

[110] is neglected in this analysis. The MC samples are processed through the full ATLAS detector simulation [111] based on Geant4 [54**?**] or through a faster simulation making use of parameterised showers in the calorimeters [112]. Additional simulated proton-proton collisions generated using `Pythia` v8.165 with the A2M tune [113] and MSTW2008LO PDF set [114] are overlaid to simulate the effects of additional collisions from the same and nearby bunch crossings (pile-up). All simulated events are then processed using the same reconstruction algorithms and analysis chain as would be used for real data.

5.3 Event Selection

The selection criteria imposed on the events are summarised and motivated with regard to the search strategy discussed above.

Only events checked for having fired a number of general quality requirements are accepted for further analysis. These requirements are driven by the need for data of sufficiently high quality and then exclude corrupted events suffering from either a noise bursts in the Luiqid Argon calorimeter or data corruption in the tile calorimeter. Data events are required to be in luminosity blocks contained in the good run list relevant to ATLAS analyses involving top quarks.

Events that are not associated with a collision such as those resulting from background radiation in the cavern or from cosmic rays, are rejected through the requirement that the primary vertex has at least two associated tracks each with the highest $\sum p_{T,track}^2$, where the $p_{T,track}$ is the transverse momentum of tracks associated with the vertex.

The event is then required to have fired a certain initial pre-selection by the triggers, associated potentially with the signal process of the search. Since this analysis is performed on the lepton+jets final state, the natural trigger selection must contain a single lepton candidate (electron or muon, depending on the analysis channel), hadronic jets and large missing energy signature compatible with that expected for a leptonic W-boson decay.

Several trigger configuration are considered and events used in this analysis must pass at least the electron (HLT_e24_lhmedium_L1EM18VH OR HLT_e60_lhmedium OR HLT_e120_lhloose) used to select e+jets events or (HLT_mu20_iloose OR HLT_mu50) used to select μ+jets.

Electron candidates: In the electron channel, events must contain exactly one electron with a minimum transverse energy $E_T(p_T) = 25$ GeV with a pseudo-rapidity of $|\eta| < 2.47$. A tight likelihood-based requirement [115] is used to further suppress the background from multi-jet production. Electrons are required to be isolated using a requirement on the sum of track p_T in a variable-sized cone around the electron direction (excluding the track

associated with the electron).

An electron isolation, *LooseTrackOnly* [116], is implemented by calculating the quantity $I_R = M_I^R/p_T^\ell$, where p_T^ℓ is the lepton p_T and M_I^R is defined as follows (excluding the lepton's track):

$$M_I^R = \sum_{\Delta R(Track,\ell)<R_{cut}} P_T^{Track} \tag{5.1}$$

where R_{cut} is the minimal value between ΔR = 10 GeV/$p_T(e)$, and $R_{max} = 0.2$ for the electron i.e. a cone which increases in size at low p_T up to a maximum of 0.2. To prevent double-counting of electron energy deposits as jets, the closest small-R jet within $\Delta R < 0.2$ of a reconstructed electron is removed. Finally, if the nearest small-R jet surviving this selection is within $\Delta R < 0.4$ of the electron, the electron is rejected, to ensure it is sufficiently separated from nearby jet activity.

Muon candidates: For the muon channel, a single muon is required with $p_T > 25$ GeV and $\eta < 2.5$. Muons are also required to be isolated, using the same track-based variable as for electrons, except that the maximum ΔR in this case is 0.3, with the selection criteria tuned to give similar efficiencies on Z $\rightarrow \mu^+\mu^-$ events. To reduce the background contributions due to muons from heavy flavour decays inside jets, muons are removed if they are separated from the nearest jet by $\Delta R < 0.04 + 10$ GeV/p_T^μ. However, if this jet has fewer than three associated tracks, the muon is kept and the jet is removed instead; this avoids an inefficiency for high-energy muons undergoing significant energy loss in the calorimeter.

Jets: three types of hadronic jets are used in this analysis:

Small-R jets: are reconstructed with the anti-k_T algorithm [78] with radius parameter R = 0.4. They are required to have transverse momentum, p_T, greater than 25 GeV and satisfying the recommended JVT requirements ($JVT > 0.64$ for the jets with $p_T < 50$ GeV and $\eta < 2.4$). These jets are calibrated using a correction that relates the reconstructed jet energy to the true jet energy when reconstructed from stable particles with a lifetime of at least 10 ps (excluding muons and neutrinos) [79]. The analysis uses these small-R jets in the calulation of missing transverse energy, E_T^{miss} as they are calibrated over a wider η range and to much lower p_T.

Large-R jets: The large-R jets are chosen as anti-k_T with radius parameter R = 1.0 jets and calibrated with the local cluster weighting (LCW) procedure [80] which are trimmed with the $R_{sub} = 0.2$ and f = 5% trimming configuration using locally calibrated clusters. In

the LCW calibration procedure, the corrections are obtained from simulations of charged and neutral particles. These jets are further trimmed [117], which has been shown to mitigate the effects of pile-up [118]. In trimming, the constituents of a jet are reclustered into subjets according to the k_t algorithm [119, 120, 121] with radius parameter R_{sub}. Subjets with a transverse momentum smaller than a fraction f_{cut} of the parent jet transverse momentum are then discarded. The surviving subjets are recombined to produce the final trimmed jet. Based on a study of sensitivity to pile-up, the trimming parameters used are $R_{sub} = 0.2$ and $f_{cut} = 0.05$ [122]. The resultant jets are required to have p_T greater than 300 GeV and $\eta < 2.0$. Large-R jets likely to contain the decay products of a hadronically decaying top-quark are identified (top-tagged) using an algorithm [123] based on the invariant mass of the jet and the n-subjettiness ratio τ_{32} [124, 125]. This algorithm has efficiency of approximately 80% for selecting top-quark jets with $p_T > 300$ GeV in simulated SM $t\bar{t}$ events.

Track jets: are built from charged tracks using the anti-k_T algorithm with radius parameter R =0.2. These jets must have p_T greater than 10 GeV and $\eta < 2.5$ and at least two constituent tracks. The charged tracks used to build the jets must themselves have $p_T >$ 0.4 GeV and $|\eta| < 2.5$, and pass quality requirements that test the number of hits used to reconstruct the track and the matching to the hard scatter via the longitudinal impact parameter of the track [126]. Track jets likely to contain a b-hadron are identified (b-tagged) using the MV2c20 algorithm [81]. The b-tagging working point chosen has approximately 70% efficiency for jets containing a b-hadron in simulated SM $t\bar{t}$ events.

Missing transverse momentum (E_T^{miss}): is defined as the magnitude of $\vec{E}_T^{miss}$ and would be characterised, in the signal, as being caused by the neutrino in the semileptonic $t\bar{t}$ decay. The missing transverse energy is calculated from the negative of the global vector sum p_T of all selected physics objects (electrons, muons, jets) as well as specific soft terms accounting for unclassified soft tracks and calorimeter clusters (see Sec 4.6 for more information). In this way, the missing transverse momentum is adjusted to take into account the best calibration of the identified physics objects above [127].

1. lepton selection: Exactly one lepton candidate is required with a minimum transverse momentum of 30 GeV for electron channel or 25 GeV for muon channel. The lepton candidates is required to be matched to the trigger. Events containing a second lepton with a p_T larger than 25 GeV are vetoed.

2. leptonic-top b-jet: Events in the boosted channel are required to contain at least

one anti-k_T R=0.4 small-R jet with $\Delta R(jet, lepton) < 1.5$. If multiple jets satisfy this condition, the one with the highest transverse momentum is chosen and henceforth, referred to as the *selected jet*, j_{sel}. It is identified with the expected b-jet from the leptonic top-quark decay, although no b-tagging requirement is enforced on it.

3. Hadronic-top jet: Another requirement, in the boosted channel, is that the events must contain at least one large-R jet passing the top-tagging requirements. The jet is further required to be well separated from the leptonic-top-quark by requiring it to have a distance in azimuthal angle to the charged-lepton $\Delta\phi$(jet,ℓ) > 2.3 radians and $\Delta\phi$(jet,j_{sel}) > 1.5 . The highest-p_T jet passing all of these requirements is referred to as the hadronic-top jet.

4. b-tagging: The event is required to have at least one b-tagged jet, by b-tagging anti-k_T R=0.2 jets which have tracks as their constituents. The track jets are rejected if they satisfy $\Delta R(lepton, jet) < 0.2$ and they are otherwise considered b-tagged if they satisfy the selection requirement $MV2C20 >$ -0.3098. The b-tagging efficiency is then expected to be 70% (estimated from a mix of jets provided by a SM $t\bar{t}$ MC), while the rejection rates of light- and c-flavors are 173 and 5.7 respectively.

5.4 Data driven backgrounds

Monte Carlo techniques have a great deal of predictive power. However, In some particular cases, it is preferable to estimate background processes from data, since the dependence on knowledge of the detector simulation and the nature of the initial processes is vastly reduced. In the case of QCD multi-jet background a purely data-driven estimation is needed as an alternative way to model the shape of the observable distributions. The shape of the W+jets background is predicted using Monte Carlo simulation while the distributions are normalized using data. These data-driven procedures are briefly introduced in this section.

5.4.1 Multi-jet background

This section describes the estimation of the multi-jet background for the l+jets boosted channel. The multi-jet background in events satisfying the selection criteria consists of events with a jet that is misreconstructed as a lepton or with a non-prompt lepton that satisfies the identification criteria e.g leptons from heavy-flavour hadrons decays, such as B- and D-hadrons. The simulation of the background from sources of non-prompt leptons suffers from large systematic and statistical uncertainties. The behavior is also highly detector-dependent. Therefore, this background must be estimated directly from data and this

is achieved by investigating the phase space regions with leptons of lower reconstruction quality. Such regions are generally more populated by QCD multi-jet events. The event topology and kinematic criteria are chosen to resemble the definition of our signal region, in order to reduce potential systematic uncertainties.

A so-called *matrix method* is used to disentangle the mixture of non-prompt leptons found in the multi-jet background as well as events from processes such as SM $t\bar{t}$ and W/Z+jets production with prompt leptons originating from the decays of W and Z bosons.

Matrix method

The Matrix Method consists on a comparison between two data samples, the selection criteria for which differ only in the quality requirements imposed on the lepton: The first sample is categorized as containing the standard set of requirements for the tight lepton quality criteria described in Sections 4.2 and 4.3. The second data sample is defined in the same way except that the lepton requires a *looser* definition, which is achieved by loosening certain signal selection cuts on the leptons. The first sample is therefore a subsample of the second.

The efficiency ϵ is defined as the probability that a *loose* lepton from prompt sources (W or Z bosons) passes the *tight* selection criteria. The false-identification rate f denotes the equivalent probability that a non-prompt lepton from multi-jets production passes the same *loose* selection.

With f and ϵ derived or validated with data, the QCD background in the signal region is estimated with data events which pass all of the signal selection, except that the loose lepton definition is used, and the events are also allowed to pass prescaled triggers. This sample contains both events from prompt-lepton sources and QCD multi-jet events. The total number of loose leptons, N_L can be defined as

$$N_L = N_{prompt} + N_{QCD} \tag{5.2}$$

Among them, those events with tight leptons should be composed as

$$N_T = \epsilon \times N_{prompt} + f \times N_{QCD} \tag{5.3}$$

Solving these two equations for N_{prompt} and N_{QCD}, makes it possible to estimate the QCD contribution to the signal region as

$$f \times N_{QCD} = \frac{(\epsilon - 1)\, f}{\epsilon - f} N_T + \frac{\epsilon\, f}{\epsilon - f} N_A \tag{5.4}$$

N_T denotes the number of events with a *tight* lepton, and N_A is the number of events with anti-*tight* lepton (i.e. a loose lepton which failed the tight cuts). In addition to predicting the overall QCD yields in our signal region, the method can also be used to estimate the kinematic distributions of the QCD background. A weight can be calculated for each event in the aforementioned sample, from the right hand side of Eqn 5.4, with $(N_T, N_A) = (1, 0)$ if the event passes the tighter lepton selection criteria and $(N_T, N_A) = (0, 1)$ otherwise. So long as any dependency the f and the ϵ have on the variable is sufficiently characterised by the chosen parametrisation, the weighted sample will give the corresponding distribution of the QCD contribution.

Lepton definitions and rates

The definitions of the loose leptons and of the prescaled trigger are given in Table 5.1.

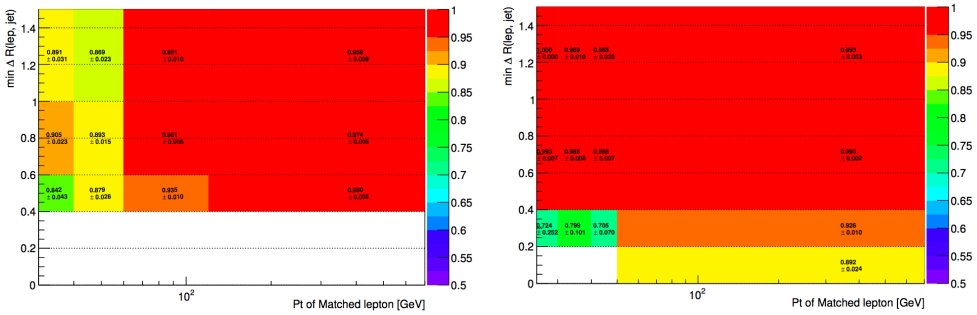
Table 5.1 – Definitions of the loose leptons and of the triggers used for the matrix method QCD estimation.

		electrons	muons
loose	triggers	nominal (Sec. 5.3)	nominal OR `HLT_mu20_L1MU15`
	lepton	`LHMedium` AND no isolation	`Loose` AND no isolation
tight	triggers	nominal	nominal
	lepton	`LHTight` AND `LooseTrackOnly`	`Medium` AND `LooseTrackOnly`

The efficiencies ϵ (one for each lepton flavor are measured from simulated SM samples with the same process mixture as found in the signal region, using the standard selections except the lepton criteria, which are *loose*). The efficiency is calculated as the fraction of reconstructed loose leptons (matched to the true lepton from the W decay to ensure they are not fake leptons) that pass the tight quality requirements. see Figures 4.3.

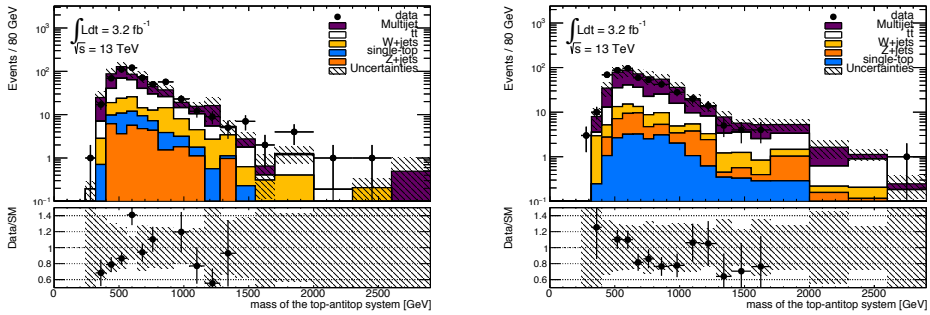
The loose leptons from the QCD multijet process have a smaller probability to pass the tight definition than the prompt lepton. This false-identification f, is measured from data with QCD-enhanced control samples, denoted as $\mathrm{CR}_{\mathrm{QCD}}^{\mathrm{boosted}}$, which follows the usual boosted selection apart for:

- At least one large-R jet with $p_T > 200$ GeV without any cut on τ_{32};
- The large-R jet mass is required to be less than 70 GeV;
- and for the electron channel:
  - The missing energy E_T^{miss} is required to be smaller than 20 GeV;
  - The sum $E_T^{miss} + M_{\mathrm{T}}$ is required to be smaller than 60 GeV;



(a) Efficiency ϵ, electrons (b) Efficiency ϵ, muons

Figure 5.3 – Efficiencies ϵ for loose prompt leptons to be identified as tight, for electrons (left) and muons (right).



(a) $m_{t\bar{t}}$, electrons (b) $m_{t\bar{t}}$, muons

Figure 5.4 – Reconstructed $m_{t\bar{t}}$ in the $\mathrm{CR}_{\mathrm{QCD}}^{\mathrm{boosted}}$ control region, for the electron (left) and muon (right) channels. The grey area indicates the impact of the 50% systematic uncertainty on the prediction of the multi-jet background on top of the statistical uncertainty

.

- and for the muon channel:
  - inverted cut on the impact parameter: $|d_0/\sigma_{d_0}| > 3$ for the muons. It is been checked this invertion does not introduce a bias.

The multijet prediction is then validated in the $\mathrm{CR}_{\mathrm{QCD}}^{\mathrm{boosted}}$ control region. The distributions of the reconstructed $m_{t\bar{t}}$ show a reasonable agreement between the data and the prediction (Figure 5.4) .

5.4.2 W+jets background

The expected background from W+jets is estimated using `Sherpa` MC samples to predict its shape, with data-driven scale factors applied to correct the total normalization. The procedure to derive the scale factors is based on the official approaches of the ATLAS Top

working group with some updates [3].

Overall normalisation: The overall normalization N_W is obtained from a sample of events selected using all cuts that does not include a b-tagging requirement, and requires a large-R jet to fail the top-tagging requirements. The procedure exploits the fact that the resulting production rate of W^++jets is higher compared to W^-+jets production at the LHC pp collider, since there are more up valence quarks than down valence quarks in the protons and that the ratio between the production rates of W^++jets and W^-+jets, r_{MC}, is well calculated from simulation than the rates themselves. W +jets production is hence known as a charge asymmetric process. The contribution to the charge-asymmetry of the backgrounds $t\bar{t}$+V, single-top and diboson has to be taken into account, the other backgrounds $t\bar{t}$, multi-jet and Z+jets are treated as charge-symmetric. The overall normalization is derived from the following formula:

$$N_{Data,W} = N_{Data,W^+} + N_{Data,W^-} = \left(\frac{r_{MC}+1}{r_{MC}-1}\right)(D^+_{\text{asym}} - D^-_{\text{asym}}), \tag{5.5}$$

where $r_{MC} = N_{MC,W^+}/N_{MC,W^-}$ is the ratio of W+jets events with a positive to those with a negative charged lepton using the MC simulation. The $D^{\pm}_{asym}$ is the number of observed events with a positive (negative) lepton, with the prediction for charge-asymmetric events ($t\bar{t}$ +V, single-top, diboson) subtracted using their estimation from MC simulation.

The charge-asymmetry normalization is then

$$CA = N_{Data,W}/N_{MC,W} = \left(\frac{r_{MC}+1}{r_{MC}-1}\right)\frac{(D^+_{\text{asym}} - D^-_{\text{asym}})}{N_{MC,W}} \tag{5.6}$$

The signal region (SR) event yield after the boosted selection is too small for a reliable evaluation of the appropriate scale factors, so a so-called normalization region (NR) is used in order to reduce the statistical uncertainties on the yields and hence on the final SFs. In fact, a looser set of some criteria that yields a normalization region enriched with W+jets events is preferred as it leads to smaller uncertainties. Specifically, the b-tagging requirements are removed, and at least one top-candidate large-R jet ($p_T > 300$ GeV, $|\eta| < 2$, $\Delta\phi(\text{R=1.0 jet}, \ell) > 2.3$ and $\Delta R(\text{R=1.0 jet}, \text{selected jet}) > 1.5$) which is not top-tagged (do not pass one of the mass and τ_{32} requirements) to enhance W+jets fraction. The obtained scale factor for the electron and muon channel are $0.82 \pm 0.18(\text{stat})$ and $(0.73 \pm 0.12(\text{stat}))$ respectively.

[3] The standard W+jets scale factors that are derived centrally by the Top group cannot be used in this analysis because various selection criteria, such as the entire boosted-topology selection or the isolation requirement for electrons, deviate from the default selection criteria used by the ATLAS Top Working Group.

5.5 Event reconstruction

As the objective of the analysis is to study the reconstructed mass of the top-antitop system and search for deviations from the SM expectations, the selected observable must be properly reconstructed. In the boosted channel, one takes the highest transverse momentum large-R jet that passes the top tagging requirements (see Sec. 5.3) as the hadronic top four-momentum estimate. The lepton and the anti-k_t $R = 0.4$ jet with $\Delta R(\ell, \text{jet}) < 1.5$ to the lepton are taken as the lepton and the b-jet from the leptonic top decay. If there is more than one $R = 0.4$ jet satisfying such a constrain, the highest p_T jet is taken.

The neutrino four-momentum is estimated by using the missing transverse energy E_T^{miss} for the estimate of its x and y components, while its z component, p_z, is calculated by constraining the lepton+E_T^{miss} system to have the W boson mass, assuming that the majority of the missing transverse momentum stems from the neutrino and that the neutrino and the charged lepton are the decay products of the W boson. Since m_W is known and the four-momentum of the charged lepton has been fully reconstructed, this quadratic equation can be solved for p_z. If only one real solution exists, the neutrino four-momentum can be reconstructed without ambiguity given that $m_\nu \simeq 0$. In case there are two real solutions, the one with the smallest $|p_z|$ is chosen. When there is no real solution, the E_T^{miss} vector is varied in the transverse plan by the minimal amount in order to get one solution.

The ΔR between the large-R jet and the $R = 0.4$ jet close to the lepton is also required to be greater than 1.5, to avoid overlap between these two objects.

The task of performing the reconstruction is further complicated by the extra radiation emitted. This effect, at simulation level, can be seen in Figures 5.5 and 5.6, on which the generated Z' mass is shown and compared to the mass of the top-antitop system decay products. This effect can reduce the sensitivity of the analysis, particularly at high resonance masses.

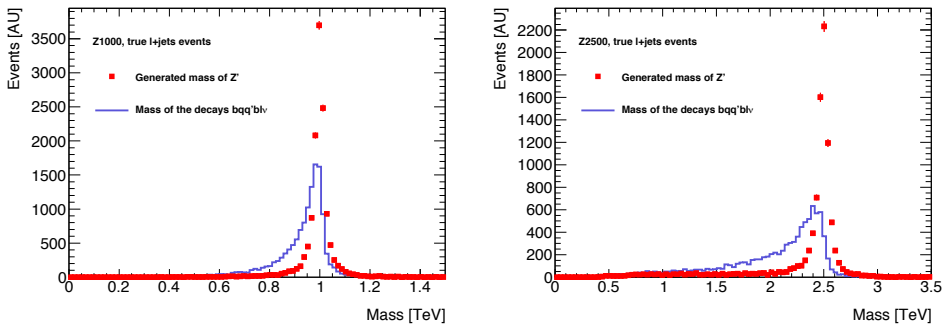


Figure 5.5 – Mass of the Z' signal and of the ttbar decay products for the 1.0 TeV (left) and 2.5 TeV (right) resonances.

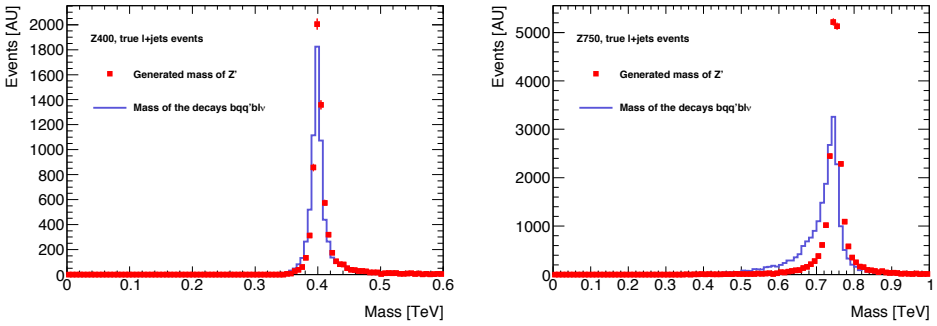


Figure 5.6 – Mass of the Z' signal and of the ttbar decay products for the 400 GeV (left) and 750 GeV (right) resonances.

5.6 Systematic uncertainties

In a search for new physics, such as the one presented here, there are a number of sources of systematic uncertainties associated with the reconstructed physical objects as well as the ones related to the estimations of signal and background modeling. The limited description of the experimental conditions (e.g. luminosity, pile-up) are also considered. Such uncertainties affect both the total number of events in the signal region(s) and hence the normalization of the mass spectrum, the total event yield and the shape of the kinematic distributions. With the exception of the luminosity uncertainty, which affects only the overall normalization.

In the following subsections a summary of all systematic uncertainties considered in this analysis is presented.

5.6.1 Systematic uncertainty of integrated luminosity

For the 2015 dataset used in this analysis, the total uncertainty in the luminosity is estimated to be 5% . This systematic uncertainty is applied as a constant shift to each simulated background (except the data-driven ones, i.e. multi-jets and W+jets).

5.6.2 Systematic uncertainties on the reconstructed objects

All the systematic uncertainties related to the reconstruction, identification and calibration of the different objects used in the search need to be taken into account. These uncertainties may affects only the normalization of the signal and background or affects both the normalization and the shape of the reconstructed $t\bar{t}$ mass distribution $m_{t\bar{t}}^{reco}$. In particular, energy or transverse momentum variations of objects, such as jets, directly affect $m_{t\bar{t}}^{reco}$.

Electron and muon

The lepton reconstruction is connected to many sources of systematic uncertainties. We have considered the trigger, the reconstruction and the identification efficiencies as well as the resolution and the scale on the momemtum. All are evaluated using leptonic Z-boson decays in data and simulation at $\sqrt{s}$= 13TeV.

Small-R jets

The prescription used to estimate the small R-jet energy scale (JES) uncertainty, derived from $\sqrt{s}$= 8 TeV simulation and extrapolated to $\sqrt{s}$ = 13 TeV, is given in Ref. [128]. This tool provides the jet energy scale uncertainties for LC+EM calibrated jets, broken down into 3 or 19 components. To reduce the complexity of inputs to the statistics tool, the simplified set with only 3 orthogonal components is used as uncorrelated nuisance parameters coming from the various in-situ techniques. As recommended, it has been checked that the analysis is not sensitive to this reduction of the number of nuisance parameters.

The jet energy resolution (JER) is dominated by two effects: electronic detector noise and pile-up and measured in dijet events using two different techniques, the dijet balance and the bisector method. The prescription used to estimate the jet energy resolution uncertainty is given in Ref. [129]. It can only be varied *up* (worsening of the resolution). As input to the limit setting code, a *down* shift histogram is created by symmetrizing with respect to the nominal distribution. If bin i has an x% shift in the smeared 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-taggging performances) are broken down into 2 components (Run-I, the uncertainty extracted from 2012 data, and *CrossCalib*, the MC-based cross-calibration uncertainty to extrapolate from 2012 to 2015 data and that takes into account detector, sampling, ... changes).

They are applied using the tool described in [130] with the version `JetUncertainties--00-09-35`. As the top-tagging use a cut on the jet mass, and the signal region is rather pure in $t\bar{t}$ background, the configuration file `Prerec2015_TopTagging_Signal.config` is used.

Large-R and small-R jets

The systematic affecting these jet collections are treated as not correlated. It has been checked (App. C) that other scenarios that assumes some correlations between the jet collections do not significantly change the result of the analysis.

b-tagging of the track-jets

The prescription for implementing the b-tagging uncertainty recommended by the performance group has been followed. The uncertainties are determined from $\sqrt{s} = 8$ TeV data, with some additional uncertainties to account for the presence of the new IBL detector and the extrapolation to $\sqrt{s}$=13 TeV. The eigenvector breakdown approach has been tested to avoid over-profiling. There are 6 (4 and 12) eigenvectors for the b-flavor (c- and light-flavor) quarks, and 2 eigenvectors specifically for the extrapolation of the scale factor in high p_T regimes, which are correlated across the flavors. It has been found (App D) that the eigenvectors 0 of the C and light flavors get constrained by the fitting procedure. To avoid this unwanted constraint, these two eigenvectors are split into six eigenvectors, depending on the track-jet p_T: $p_T < 50$, $50 < p_T < 100$ and $p_T < 100$ GeV. The impact of the limits is within 1% (5%) for the expected (observed) limits as presented in App D.

5.6.3 Systematic uncertainties on the background estimations

Uncertainties affecting only $t\bar{t}$

The uncertainty on the top-antitop cross-section is one of the main sources of systematic uncertainty. It has been estimated ([131]) to be $^{+5.6\%}_{-6.1\%}$ for a top quark mass of 172.5 GeV.

The uncertainty on the electroweak corrections to $t\bar{t}$ production is 10% of their deviation from unity. As the SM $t\bar{t}$ background is the dominant one, there are four sources of systematic uncertainties related to its modeling that have been studied: the modeling of the hard process, the PDF choice and the simulation of the extra QCD radiation ISR and FSR:

- *Generator*: This is evaluated using a shape comparison between `Powheg+Herwig` sample and MC@NLO`+Herwig`, both using `AtlasFastII`. The impact on the shape is symmetrized to get the downward variation.
- *Parton showering*: The uncertainties arise from the shape comparison between `Powheg+Pythia` and `Powheg+Herwig`, both using AtlasFastII. The impact on the shape is symmetrized to get the downward variation.
- *Parton Distrbution Functions (PDF)*: Are calcualted using 30 separate uncertainty eigensets of the `PDF4LHC_nlo_30` [132] prescriptions in an MC@NLO sample.
- *Modelling of extra QCD radiation (ISR/FSR)*: This is evaluated using a shape comparison between a regular `Powheg+Pythia6` and 2 other `Powheg+Pythia6` samples, which have modified shower radiation (P2012 radLo and radHi), factorization and renormalization scale (x 2 and x 0.5) and NLO radiation (hdamp = mtop and 2 x mtop).

Uncertainties affecting only single-top

The cross-section of the single-top backgrounds are assigned an systematic uncertainty of ±5.3% [133]. This corresponds to the theoretical uncertainty on the dominant Wt-channel contribution at approximate NNLO in QCD [134, 135].

5.6.4 Impact of the systematic uncertainties

The impact of the systematic uncertainties in the yields in the electron and muon channels, respectively are presented in Tables 5.2 and 5.3.

Table 5.2 – Systematic impact on the background yields in the boosted electron channel, in percentage values.

Systematics	**Percentage variation**
anti-k_t R=1.0 run 1	11.6
anti-k_t R=1.0 cross calibration	9.8
anti-k_t R=0.4 JER	0.2
anti-k_t R=0.4 JES1	1.0
anti-k_t R=0.4 JES2	0.4
anti-k_t R=0.4 JES3	0.2
b-tagging eff. (E0)	2.1
b-tagging eff. (E1)	0.8
b-tagging eff. (E2)	0.8
b-tagging eff. (E3)	0.8
b-tagging eff. (E4)	0.7
b-tagging c mistag (E0)	1.3
b-tagging c mistag (E1)	0.8
b-tagging c mistag (E2)	0.8
b-tagging c mistag (E3)	0.7
b-tagging extrap.	0.2
b-tagging extrap. (charm)	0.2
b-tagging l mistag (E0)	1.9
b-tagging l mistag (E1)	0.8
b-tagging l mistag (E10)	0.7
b-tagging l mistag (E11)	0.7
b-tagging l mistag (E2)	0.8
b-tagging l mistag (E3)	0.7

Continued on next page

Table 5.2 – *Continued from previous page*

Systematics	**Percentage variation**
b-tagging l mistag (E4)	0.7
b-tagging l mistag (E5)	0.7
b-tagging l mistag (E6)	0.7
b-tagging l mistag (E7)	0.7
b-tagging l mistag (E8)	0.7
b-tagging l mistag (E9)	0.7
electron ID	1.2
electron isolation	0.7
electron reconstruction	0.4
electron res.	0.1
electron scale	0.1
electron trigger	0.5
luminosity	4.5
MET res. para.	0.1
MET res. perp.	0.1
MET scale	0.1
muon ID (stat)	0.0
muon ID (syst)	0.0
muon isolation (stat)	0.0
muon isolation (syst)	0.0
muon res. (ID)	0.0
muon res. (MS)	0.0
muon scale	0.0
muon trigger (stat)	0.0
muon trigger (syst)	0.0
QCD unc.	1.8
single top cross sec.	0.3
tt electroweak corr.	0.4
ttbar gen.	7.9
ttbar ISR/FSR	3.4
ttbar PDF	1.5
ttbar p. shower	8.7
ttbar cross section	5.1

Continued on next page

Table 5.2 – *Continued from previous page*

Systematics	**Percentage variation**
W C/A SF	0.0
Total	21.3

Table 5.3 – Systematic impact on the background yields in the boosted muon channel, in percentage values.

Systematics	**Percentage variation**
anti-k_t R=1.0 run 1	12.0
anti-k_t R=1.0 cross calibration	10.3
anti-k_t R=0.4 JER	0.0
anti-k_t R=0.4 JES1	0.7
anti-k_t R=0.4 JES2	0.4
anti-k_t R=0.4 JES3	0.2
b-tagging eff. (E0)	1.6
b-tagging eff. (E1)	0.2
b-tagging eff. (E2)	0.1
b-tagging eff. (E3)	0.1
b-tagging eff. (E4)	0.1
b-tagging c mistag (E0)	0.7
b-tagging c mistag (E1)	0.2
b-tagging c mistag (E2)	0.1
b-tagging c mistag (E3)	0.1
b-tagging extrap.	0.2
b-tagging extrap. (charm)	0.2
b-tagging l mistag (E0)	0.9
b-tagging l mistag (E1)	0.1
b-tagging l mistag (E10)	0.1
b-tagging l mistag (E11)	0.0
b-tagging l mistag (E2)	0.2
b-tagging l mistag (E3)	0.1
b-tagging l mistag (E4)	0.1
b-tagging l mistag (E5)	0.1
b-tagging l mistag (E6)	0.1
b-tagging l mistag (E7)	0.1

Continued on next page

Table 5.3 – *Continued from previous page*

Systematics	**Percentage variation**
b-tagging l mistag (E8)	0.1
b-tagging l mistag (E9)	0.1
electron ID	0.0
electron isolation	0.0
electron reconstruction	0.0
electron res.	0.0
electron scale	0.0
electron trigger	0.0
luminosity	4.7
MET res. para.	0.0
MET res. perp.	0.1
MET scale	0.1
muon ID (stat)	0.2
muon ID (syst)	0.7
muon isolation (stat)	0.1
muon isolation (syst)	0.3
muon res. (ID)	0.0
muon res. (MS)	0.2
muon scale	0.0
muon trigger (stat)	1.1
muon trigger (syst)	0.5
QCD unc.	0.3
single top cross sec.	0.3
tt electroweak corr.	0.4
ttbar gen.	9.8
ttbar ISR/FSR	3.4
ttbar PDF	1.7
ttbar p. shower	9.1
ttbar cross section	5.3
W C/A SF	0.0
Total	22.4

5.7 Data versus background expectation comparison

5.7.1 Total yield

Having evaluated the W+jets and multijets backgrounds contributions and established the systematic uncertainties, we can begin to study the *signal sample*. Table 5.4 and 5.5 gives the number of expected events for each process as well as a comparison of the total prediction to the observed number of events together with the associated systematic uncertainties. There is a deficit of data compared to expectation for both e+jets and μ+jets selections; however this deficit is consistent with the nominal prediction within the associated systematic uncertainties. The background from SM $t\bar{t}$ production clearly dominates the total SM background. The second largest background contribution in each channel stems from W+jets production, followed by single-top and QCD multi-jet production. The contribution of the latter is larger in the e+jets channels than in μ+jets channels due to the higher rate of *fake* electrons compared to that of *fake* muons.

5.7.2 Selected kinematic variable distributions

The goal of the analysis presented in this thesis is to calculate the $m_{t\bar{t}}$ spectra and use it to set a limit on the benchmark model since a simple comparison of the observed and expected event yields in the signal regions does not provide sufficient sensitivity to BSM processes for which the production cross section times branching ratio is small. A few checks must be made on the relevant kinematic distribution that are not considered directly in the search for BSM resonance. Indeed, after all event selection criteria are applied, the available data are compared to the Monte Carlo simulation in the set of figures below. In this section, only a few representative distributions are shown. The simulation is scaled to the measured data collected luminosity and the agreement is shown in the ratio plots of Figures 5.7 to 5.20. The different level of data/MC agreement between the electron and muon channel ($\frac{data_e}{MC_e} = 0.91$ and $\frac{data_\mu}{MC_\mu} = 0.89$) is compatible with a statistical fluctuation expected to be of the order of 2%. The reconstructed $t\bar{t}$ invariant mass spectra for the electron and muon selections before and after the profiling of the nuisance parameters, are shown in Figure 5.7. Figures 5.8, 5.9 and 5.10 show the shape of E_T^{miss}, the lepton p_T and m_T^w respectively, which are generally well modelled by the simulation. The mass of the leptonically decaying top quark is reconstructed from the lepton, neutrino and the b-jet (closest jet to lepton) in the boosted scenario and it is shown in Figure 5.11. The mass of the hadronically decaying top quark is taken to be the mass of the selected large-R jet and it is shown in 5.12. The distributions of the transverse momentum and the first k_T splitting scale of the

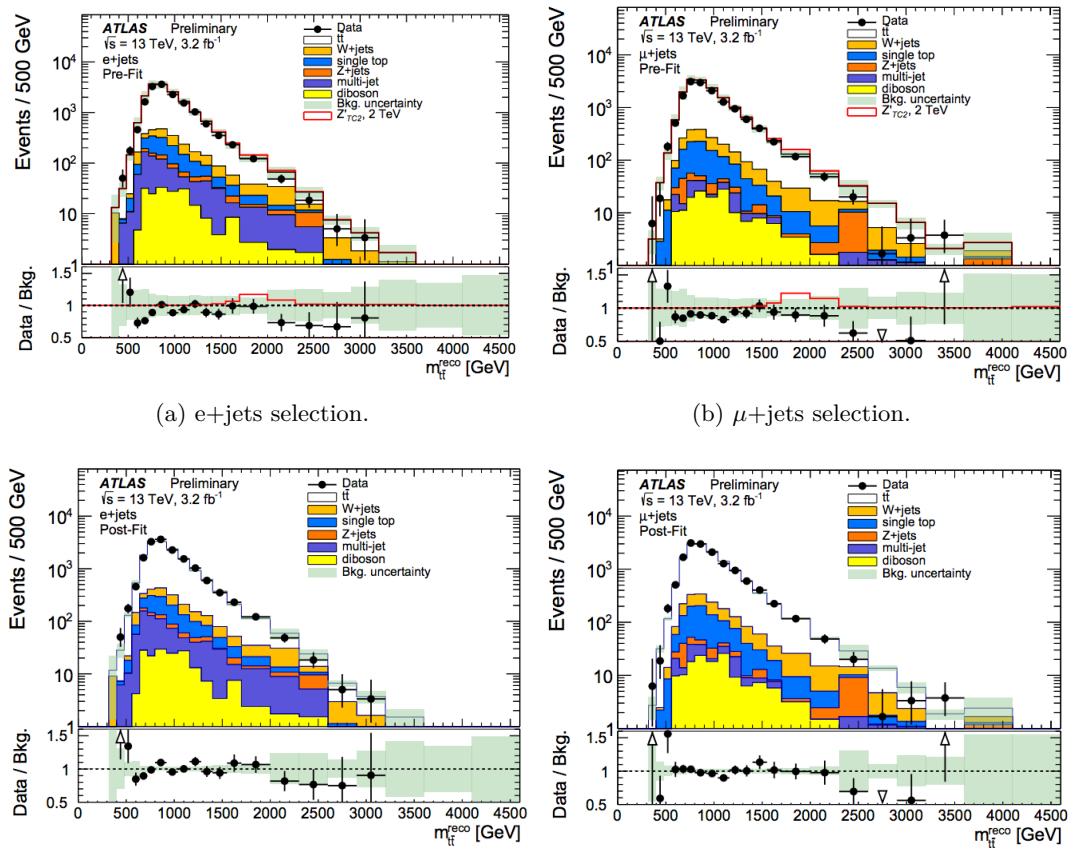
Table 5.4 – Data and expected background event yields for the electron boosted channel. The total systematic uncertainty on the expected background yields is also given.

Type	Yield
$t\bar{t}$	3060.3 ± 703.9
W+jets	201.0 ± 136.0
single top	186.8 ± 37.1
Z+jets	32.8 ± 12.3
multi-jet	128.7 ± 66.5
diboson	45.8 ± 10.9
Total	3655.4 ± 779.3
Data	3352.0

Table 5.5 – Data and expected background event yields for the muon boosted channel. The total systematic uncertainty on the expected background yields is also given.

Type	Yield
$t\bar{t}$	2986.6 ± 716.4
W+jets	198.4 ± 43.6
single top	179.0 ± 36.1
Z+jets	26.4 ± 11.9
multi-jet	18.5 ± 10.6
diboson	36.6 ± 8.2
Total	3445.5 ± 771.0
Data	3074.0

same jet, which are also used in the definition of the signal regions for the boosted topology, are shown in Figures 5.13 and 5.14 respectively. The data-to-prediction agreement is validated furthermore by studying the distributions of several variables of interest such as the τ_{32}^{wta}, ΔR_{min}, the transverse momentum of leading and hardest anti-K_T R = 0.4 jet, η and ϕ of the lepton shown in figures from 5.15 to 5.20., respectively. In this distributions the data generally agree with the expected background with slight shape differences seen in the high-mass and low-mass regions; these deviations are consistent with the nominal predictions within the associated uncertainties. This demonstrates that the variables used in the definition of the signal regions are well-modelled and that the reconstruction of the $t\bar{t}$ system yields reliable results.



(a) e+jets selection.

(b) μ+jets selection.

(c) e+jets selection.

(d) μ+jets selection.

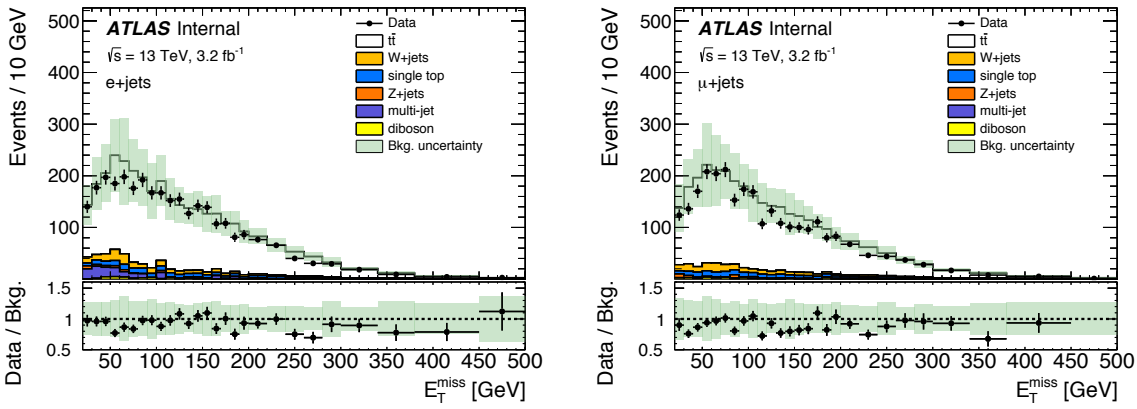
Figure 5.7 – Reconstructed invariant mass of the $t\bar{t}$ system before (a,b) and after (c,d) the profiling of the nuisance parameters for the e+jets and μ+jets selections.

5.8 Results

After the reconstruction of the $t\bar{t}$ mass spectrum, the data and simulation distributions are evaluated in two steps: first the compatibility with the SM-only hypothesis is tested, and then, in the absence of deviations, upper cross section limits for the Z' benchmark model are derived.

5.8.1 Compatibility with the SM-only hypothesis

The search procedure is done systematically with the so-called BumpHunter [136], a hypothesis testing tool that point out the presence of local data excesses or deficits with respect to the expected background. While such deviations are expected in the presence of resonant production of massive particles in Particle Physics, they may also be caused by statistical fluctuations due to the finite number of background events in the search region. BumpHunter, is a test that identify the most significant bump, where the data are most deviant



(a) (b)

Figure 5.8 – The distribution of the E_T^{miss} in the (a) e+jets and (b) μ+jets selections. The SM background components are shown as stacked histograms. The shaded areas indicate the total systematic uncertainties.

from the Null hypothesis. Based on this bump, the test derives a p-value, corresponding to a probability that a given deviation may be caused merely by a statistical fluctuation of the expected background.

In case of a distribution with multiple bins, as the $m_{t\bar{t}}^{reco}$ spectra under consideration, each bin counts as a trial with its own chance of triggering a discovery, and there are many such trials which have to be taken into account. This is referred to as the *look-elsewhere effect* or *trials factor*.

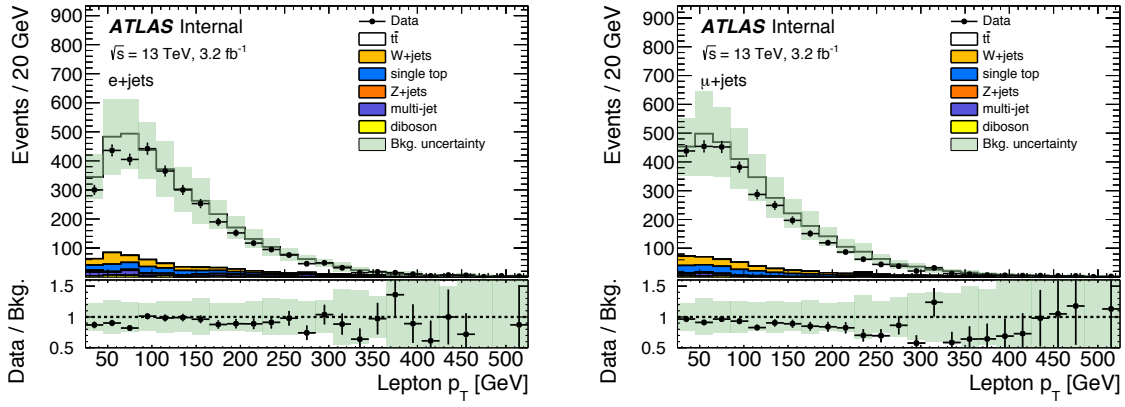
BumpHunter scans the data using a sliding windows of variable size, with a minimum width of two bins, and keeps the window with biggest excess of data compared to the background. For each spectrum, the Poisson probability of the most prominent bump or dip (i.e. the smallest probability) is saved. In each window i, the data count is d_i, the background yield is b_i and the Poisson probability $\mathrm{P}(d_i, b_i)$ is defined as

$$P(d_i, b_i) = \begin{cases} \Gamma(d_i, b_i) = \sum_{n=d_i}^{\infty} \frac{b_i^n}{n!} e^{-b_i} & \text{if } d_i \geq b_i \\ 1 - \Gamma(d_i + 1, b_i) & \text{if } d_i < b_i \end{cases} \tag{5.7}$$

when searching for an excess (the inequality signs are reversed when looking for deficits). Γ is the Gamma function. The 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:

$$P(d_i, b_i) \rightarrow max(P(d_i, b_i + \lambda_i \theta_i) e^{(-\frac{\lambda^2}{2})}) \tag{5.8}$$



(a) (b)

Figure 5.9 – The distribution of the transverse momentum of the lepton in the (a) e+jets and (b) μ+jets selections. The SM background components are shown as stacked histograms. The shaded areas indicate the total systematic uncertainties.

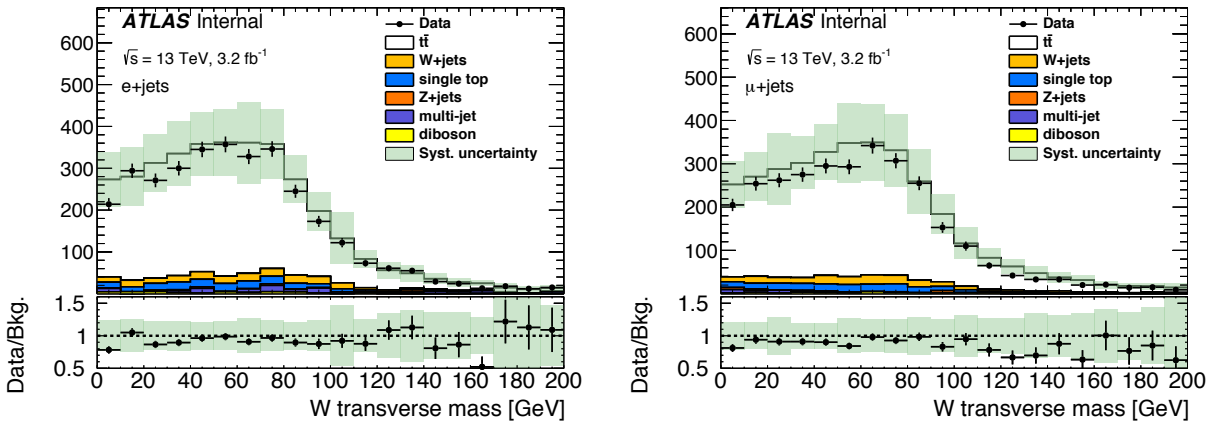


Figure 5.10 – W transverse mass.

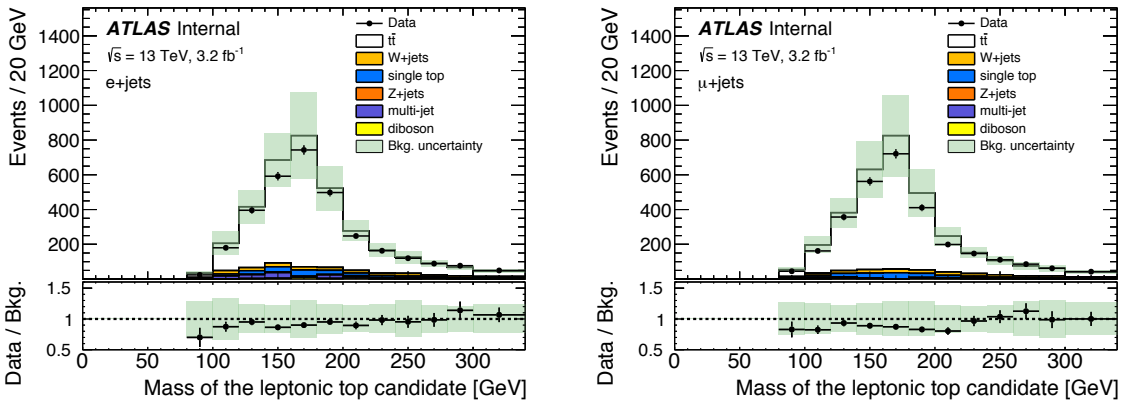
$$P(d_i, b_i) \rightarrow P(d_i, b_i + \lambda_i\theta_i)e^{(-\frac{\lambda^2}{2})} \tag{5.9}$$

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 $\frac{(b_i+\lambda_i\theta_i)^{d_i}}{d_i!}e^{-(b_i+\lambda_i\theta_i)}e^{(-\frac{\lambda^2}{2})}$.

The BumpHunter test statistic t is computed as

$$t = \begin{cases} 0 & \text{if } d_i \leq b_i \\ -\log P_i^{\min} & \text{otherwise} \end{cases} \tag{5.10}$$

The p-value of the most interesting bump is found by comparing the test statistic from data with the test statistics found in at least $N = 10,000$ pseudo experiments, where the pseudodata is generated by Poisson fluctuations of the expected background. The p-value



(a) (b)

Figure 5.11 – The distribution of the reconstructed mass of the leptonic top candidate in the (a) e+jets and (b) μ+jets selections. The SM background components are shown as stacked histograms. The shaded areas indicate the total systematic uncertainties.

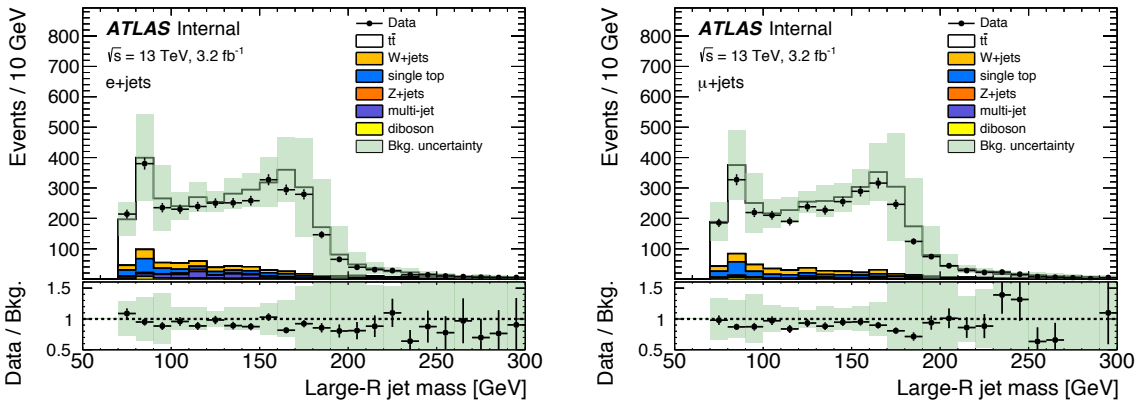
is defined as

$$p - \text{value} = \frac{\int_{t_{obs}}^{\infty} f(t)}{\int_{0}^{\infty} f(t)} \tag{5.11}$$

where $f(t)$ is the distribution of the test statistic 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.

Data has been divided into 2 distinct channels: e+jets and μ+jets boosted, according to the description in Section 5.3. The list of all the excess and deficits with their local p-value can be found in the figures 5.21 and 5.22. In a second step the channels are combined.

The results of the channel combination is based on the fact that if a $t\bar{t}$ resonance exists, bumps will arise in the various spectra at approximately the same mass point. In this case, the most interesting windows found, for each channel, are compared. If they do not overlap, it is called *no overlap*. If they overlap, the combined probability, to observe a larger data count, is taken as the product of the individual probabilities, from which the BumpHUNTER test statistic is computed. The test statistic obtained from data is then compared with pseudo experiments, conducted under the same conditions. In Table 5.6, 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, the expected distribution after



(a) (b)

Figure 5.12 – : The distribution of the mass of the large-R jet in the (a) e+jets and (b) μ+jets selections. The SM background components are shown as stacked histograms. The shaded areas indicate the total systematic uncertainties.

Configuration	Channel	mass range	p-value	σ
excess (syst, prefit)	e	400 - 560	0.896±0.003	0
	μ	400 - 1040	0.836±0.003	0
	comb.	400 - 560	0.284±0.004	0.568 (0.555 - 0.582)
deficit (syst, prefit)	e	3200 - 4100	0.514±0.004	0
	μ	2600 - 3200	0.020±0.001	2.045 (2.017 - 2.075)
	comb.	no overlap		

Table 5.6 – Most significant excess/deficits, using the prefit spectra and the systematic uncertainties.

fitting the systematic is also used to search for bump/dip 5.7.

These results allows us to safely claim there is no visible signal in the data and then limits can be set.

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

Since the current search has not found any evidence to the production of new particle. This leads to exclusion limits on cross section times acceptance and upper limits on masses or scales for New Physics. Two fundamentally different approaches can be used to set limits, the *Frequentist* and the *Bayesian* method. In this section the *frequentist* approach using the *Confidence Level (CL)* method is used to set 95% CL cross section limits on the hypothesis that data is consistent with Standard Model background and signal. TRexFitter [137], RooStats and RooFit were used to perform this limit setting procedure, on which a likelihood for a particular signal assumption is defined as:

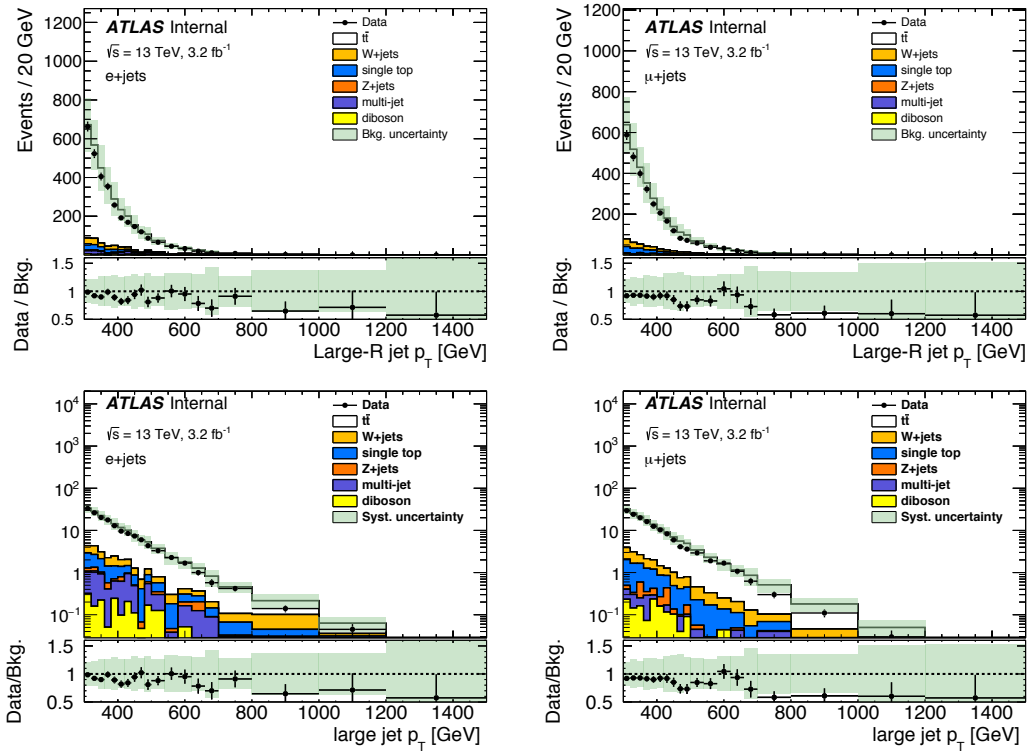


Figure 5.13 – : The distribution of the transverse momentum of the large-R jet in the (left) e+jets and (right) μ+jets selections in linear and log scales. The SM background components are shown as stacked histograms. The shaded areas indicate the total systematic uncertainties.

$$L(\mu,\Theta) = \prod_{i=0}^{\text{channels,bins}} \frac{e^{-\mu a_{Z',i}\sigma_{Z'}+b_i}(\mu a_{Z',i}\sigma_{Z'}+b_i)^{D_i}}{\Gamma(D_i+1)} C(\Theta) \tag{5.12}$$

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:

$$\Lambda(\mu) = \frac{L(\mu,\hat{\hat{\Theta}}(\mu))}{L(\hat{\mu},\hat{\Theta})}, \tag{5.13}$$

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

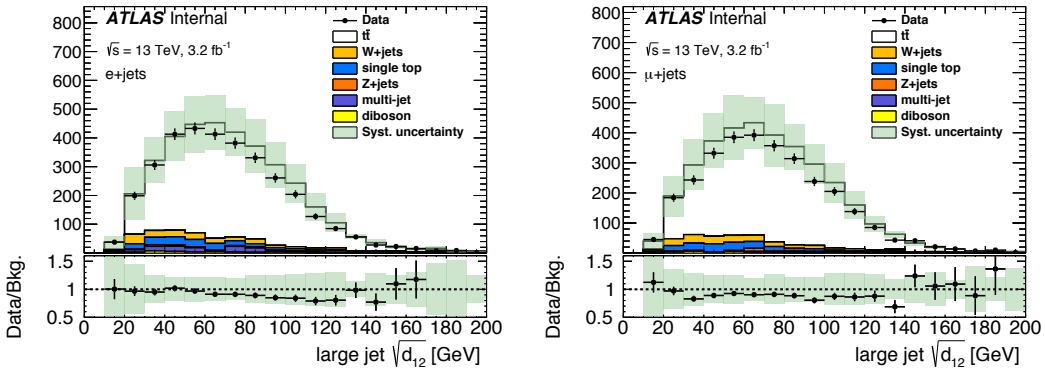


Figure 5.14 – the first k_T splitting scale of of the large-R jet.

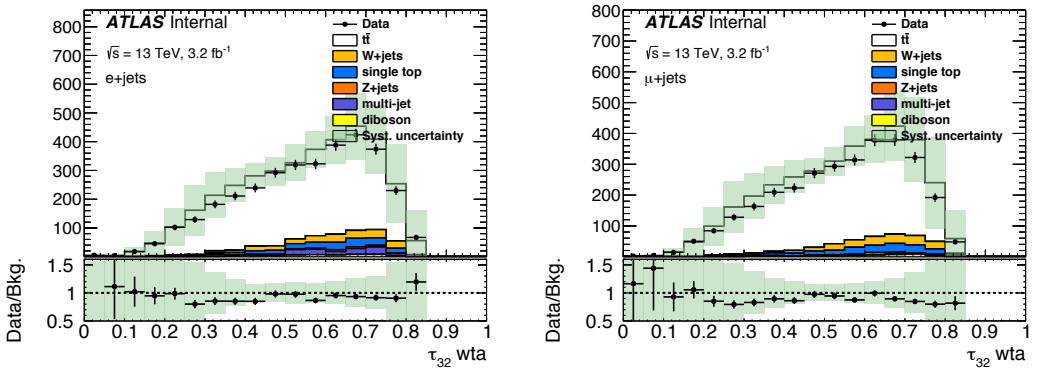


Figure 5.15 – τ_{32}^{wta} of the large-R jet.

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.

The resulting limit for each signal point is shown in Figure 5.23 and an example of the pull of the nuisance parameters on Figure 5.24. 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 due to the large statistic of the signal region. These constrains could already be predicted from a fit on Asimov pseudo-data (i.e 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 largest impacts of the systematic uncertainties (with their pre- and post-fit magnitudes) on the fitted μ is on figure 5.25. The limits on cross-section $\times$ branching ratio can be interpreted as observed (expected) constraints on the Z' mass: 750 <m(Z')<2.1 TeV (2.1 TeV).

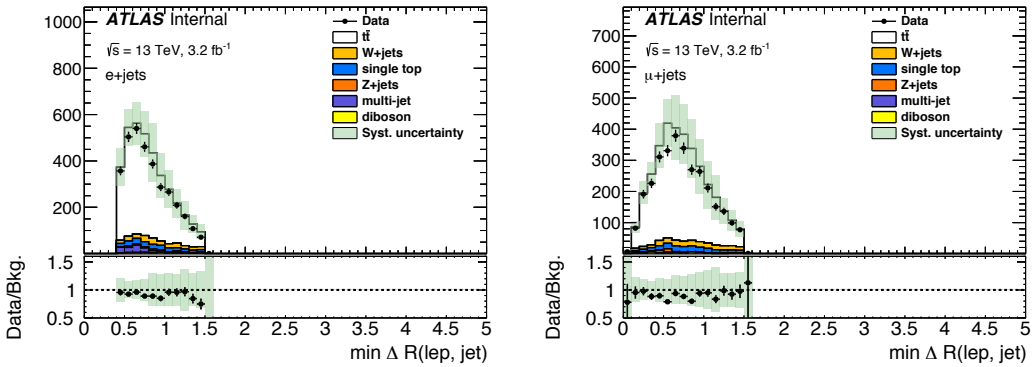


Figure 5.16 – $\Delta R_{min}(\ell, jets)$.

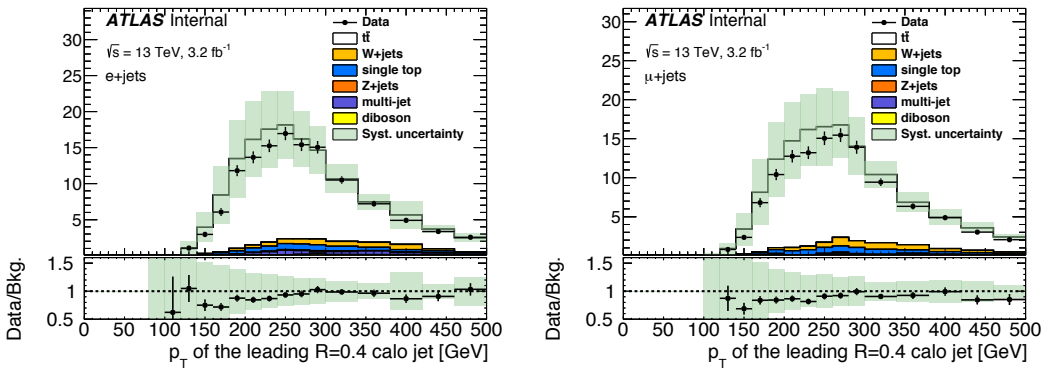


Figure 5.17 – Transverse momentum of the leading anti-k_T $R = 0.4$ jet.

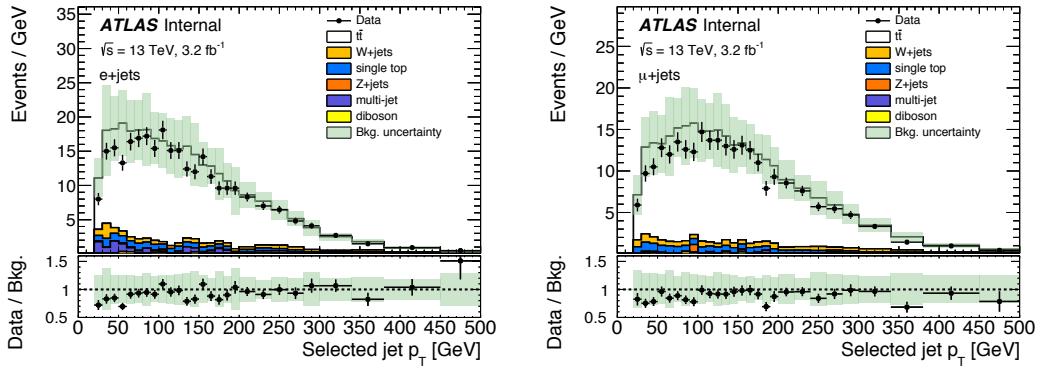


Figure 5.18 – Transverse momentum of the hardest anti-k_T $R = 0.4$ jet with $\Delta R(\ell, \text{jet}) < 1.5$.

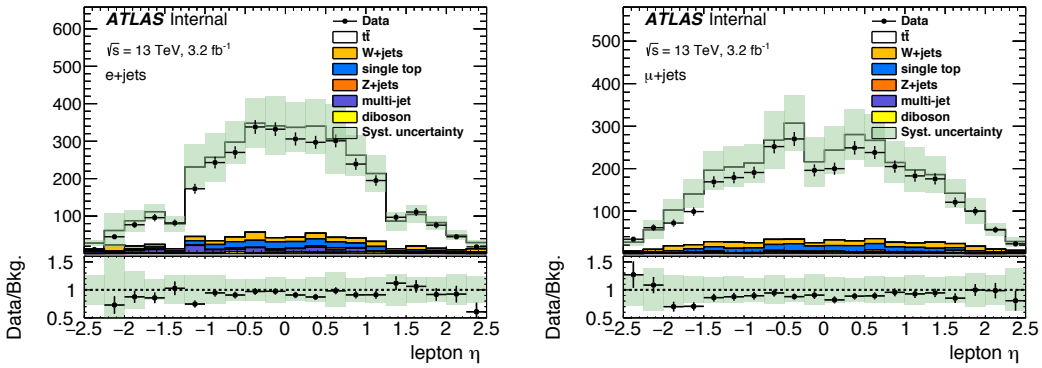


Figure 5.19 – Pseudo-rapidity of the lepton.

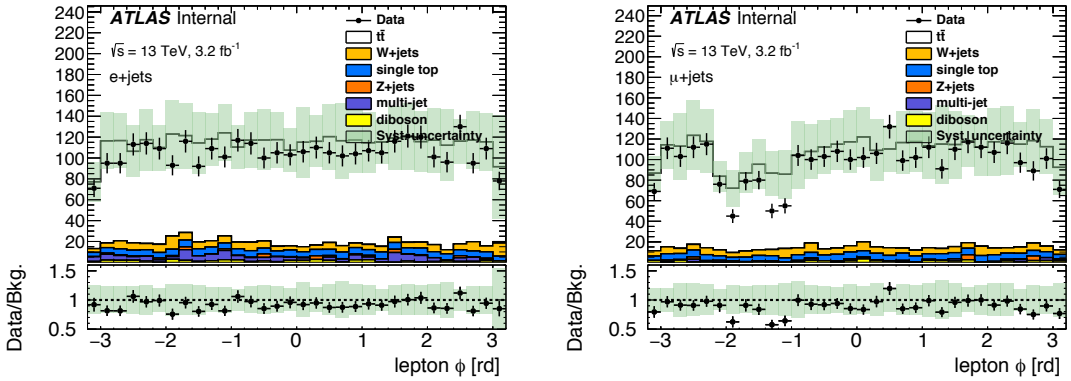
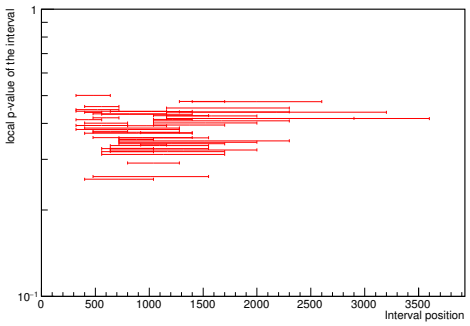


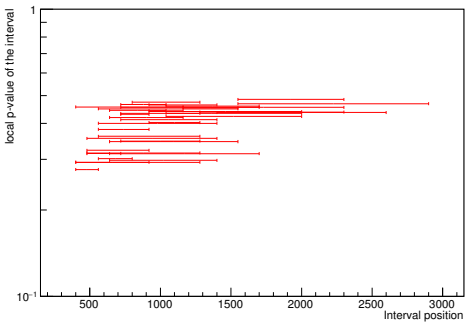
Figure 5.20 – ϕ of the lepton.

Configuration	Channel	mass range	p-value	σ
excess (stat, postfit)	e	400 - 560	0.331±0.004	0.435 (0.422-448)
	μ	320 - 560	0.562±0.004	0
	comb.	400 - 560	0.088±0.002	1.350 (1.333 - 1.368)
deficit (stat, postfit)	e	560 - 720	0.134±0.002	1.106 (1.095 - 1.117)
	μ	2600 - 3200	0.092±0.002	1.328 (1.315 - 1.340)
	comb.	no overlap		

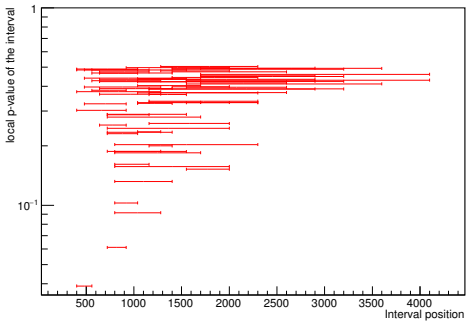
Table 5.7 – Most significant excess/deficits, using the postfit spectra and the systematic uncertainties.



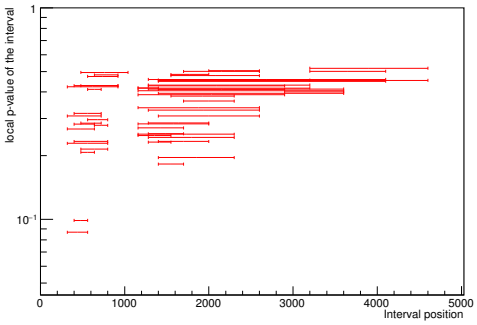
(a) Boosted, e+jets, excess, pre-fit



(b) Boosted, μ+jets, excess, pre-fit

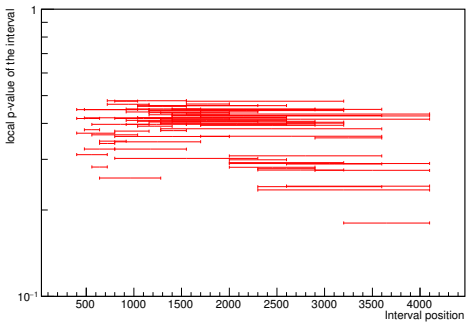


(c) Boosted, e+jets, excess, post-fit

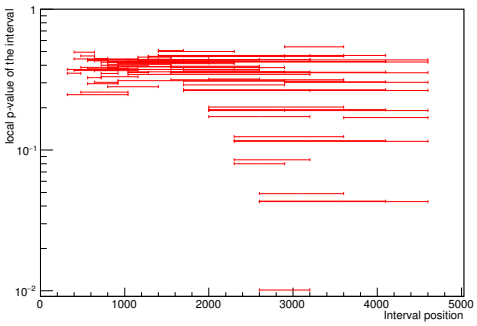


(d) Boosted, μ+jets, excess, post-fit

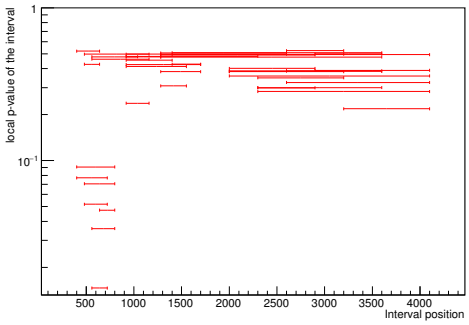
Figure 5.21 – All the considered excesses with their significance for e+jets (left) and μ+jets (right) channels, for m_{tt} spectra before (top) and after (bottom) the fit of the nuisance parameters.



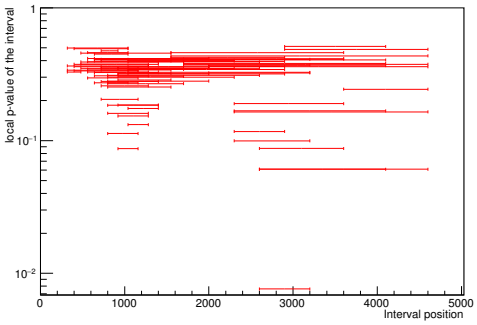
(a) Boosted, e+jets, deficit, pre-fit



(b) Boosted, μ+jets, deficit, pre-fit



(c) Boosted, e+jets, deficit, post-fit



(d) Boosted, μ+jets, deficit, post-fit

Figure 5.22 – All the considered deficits with their significance for e (left) and μ+jets (right) channels, for $m_t\bar{t}$ spectra before (top) and after (bottom) the fit of the nuisance parameters.

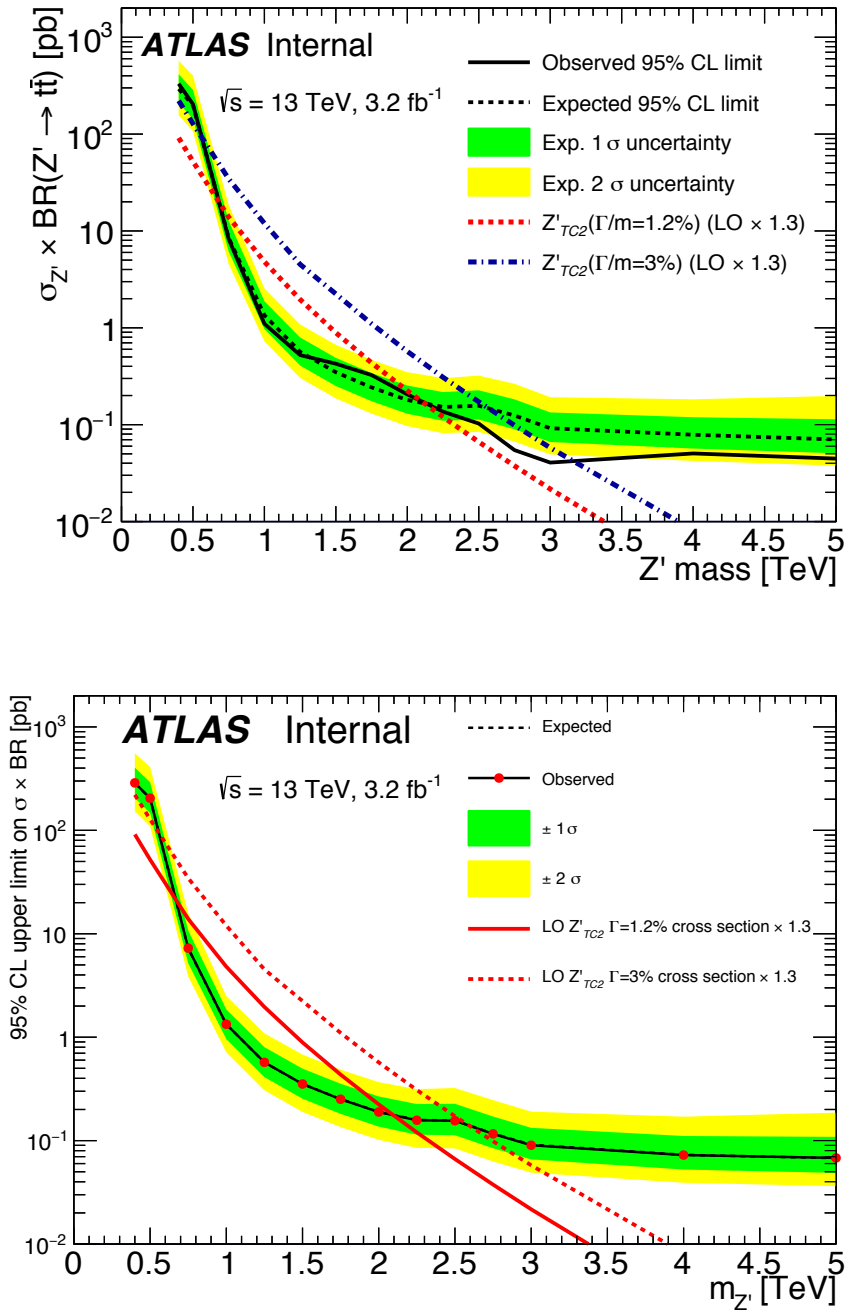


Figure 5.23 – Top: Cross-section 95% CL upper limits on the Z'_{TC2} signal using Atlas data. Bottom: same result but using Asimov pseudo-data, for which the data equal to the expected background. Both expected results (from fits on data or pseudodata) are very similar.

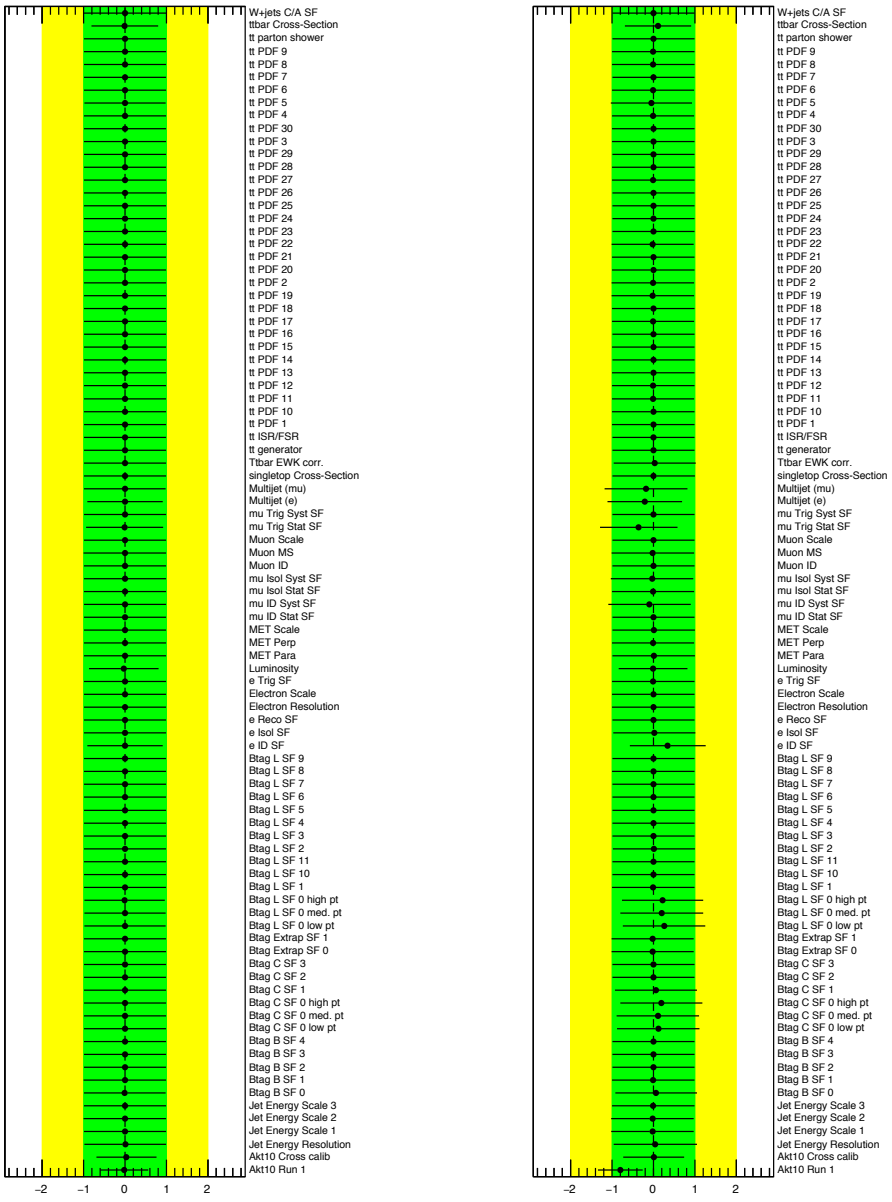


Figure 5.24 – Pull of the nuisance parameters for a fit performed under the hypothesis background+Z'(2 TeV), for the Asimov pseudo-data (left) and Atlas data (right).

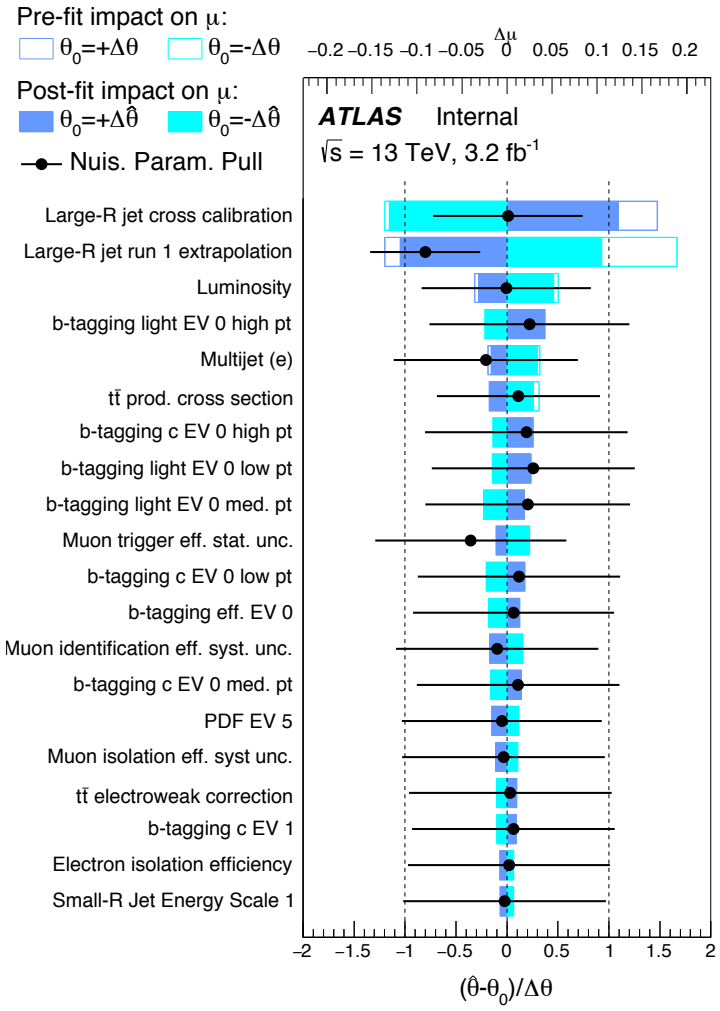


Figure 5.25 – The largest impacts of the nuisance parameters (with their pre- and post-fit magnitudes) on the fitted μ and their pulls for a fit performed under the hypothesis background+Z'(2 TeV) on Atlas data.

Conclusion and Outlook

In this thesis a search for $t\bar{t}$ resonances has been performed. The search is done in final states containing at least one electron or muon in a data sample corresponding to an integrated luminosity of 3.2 fb^{-1} collected with the ATLAS experiment. To find a $t\bar{t}$ resonance a serie of cuts are applied to enhance the $t\bar{t}$ topology. The selection used in the analysis is designed for the particular boosted topology where the selected and reconstructed events are merged into a single, large-R jet with a characteristic substructure. No evidence for a new resonance has been observed and upper limits are set on the cross-section times branching ratio for a narrow ($\leq$ 3% width) Z' boson. Based on these results, the existence of a narrow (width 1.2% of its mass) leptophobic topcolour Z' in the range 0.7 TeV$< m_{Z'} <$ 2.0 TeV is excluded at 95% CL. In addition to the physics analysis of searching for resonances in the top quark sector, the systematic validation of the different Geant4 physics lists are presented. In particular, they are related to the study of the evolution of physics lists as the calorimeter response to the passage of the particles as a function of the incident energy. The target of this work is to estimate the uncertainty that should be taking into account in the Jet Energy Scale determination.

There are several enhancements that may be added to the work established in this thesis. In the short terme, extra data with the partially 33.3 fb^{-1} of 2016 projected integrated luminosity would improve the sensitivity of the search to potential signal processes with smaller cross-sections. Also, the analysis for the resolved topologies can be added. The lepton isolation would also be changed (from LooseTrackOnly to FixedCutTightTrackOnly) to get a better fake lepton rejection, while preserving the signal acceptance. The Graviton benchmark (a spin-2 bulk Randall-Sundrum (RS) Kaluza-Klein graviton G_{kk} [[40] Can be also included. This benchmark has already be considered by a previous Atlas analysis [138], but no constraint have been set on its mass, the limit on its production rate was just above the theoretical cross-section in the 400-500GeV range.

Appendix A

MC-based uncertainties for the different Physics lists in G4 version 9.4 and 9.6

The Monte Carlo-based uncertainties extracted as histograms from the physics lists variations in the two G4 versions studied in Chap 4.

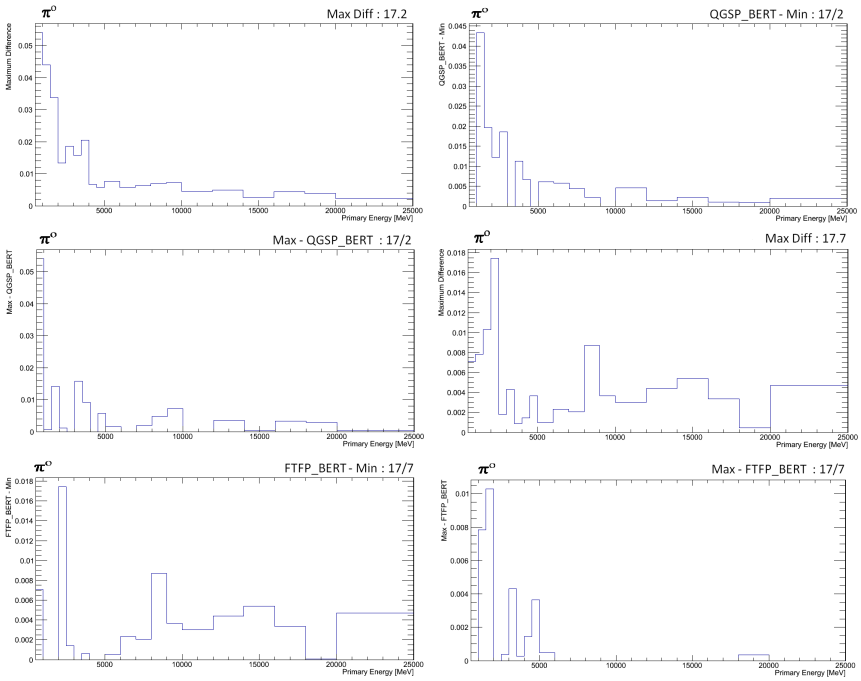


Figure A.1 – π^0: The differences are quite low for both G4 versions

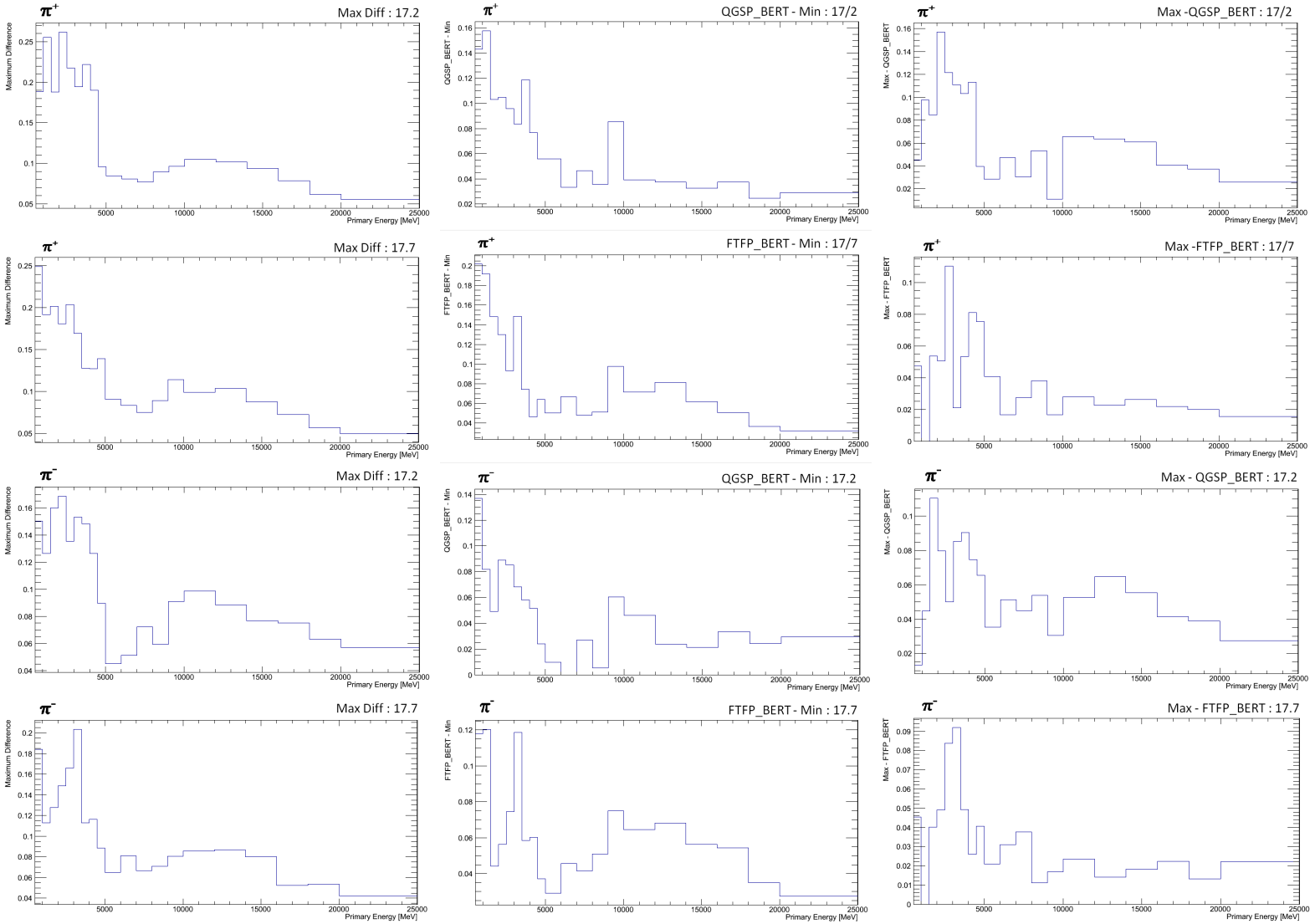


Figure A.2 – π^+, π^-: The differences decrease by increasing energies, They are almost under 10-20% .

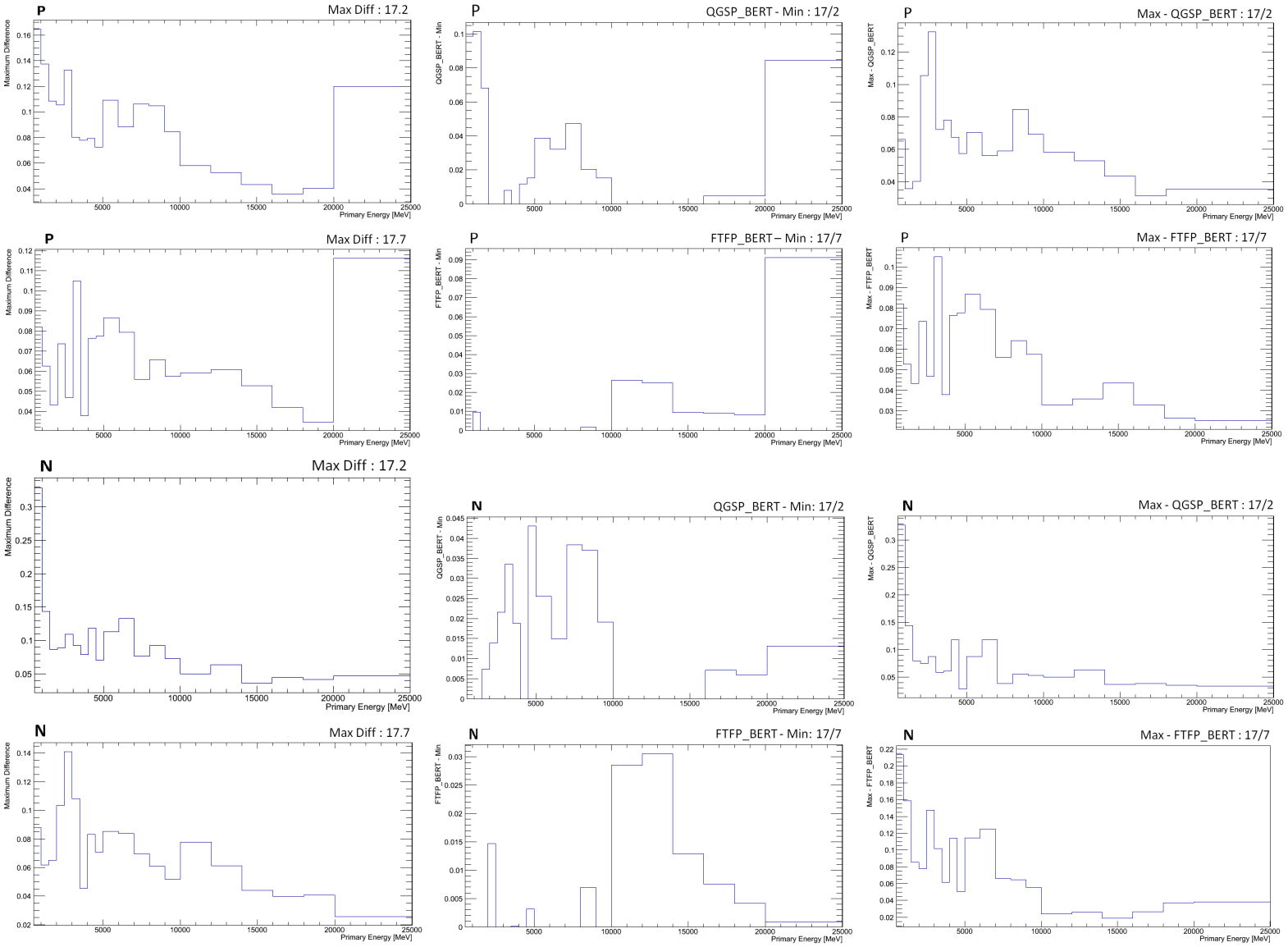
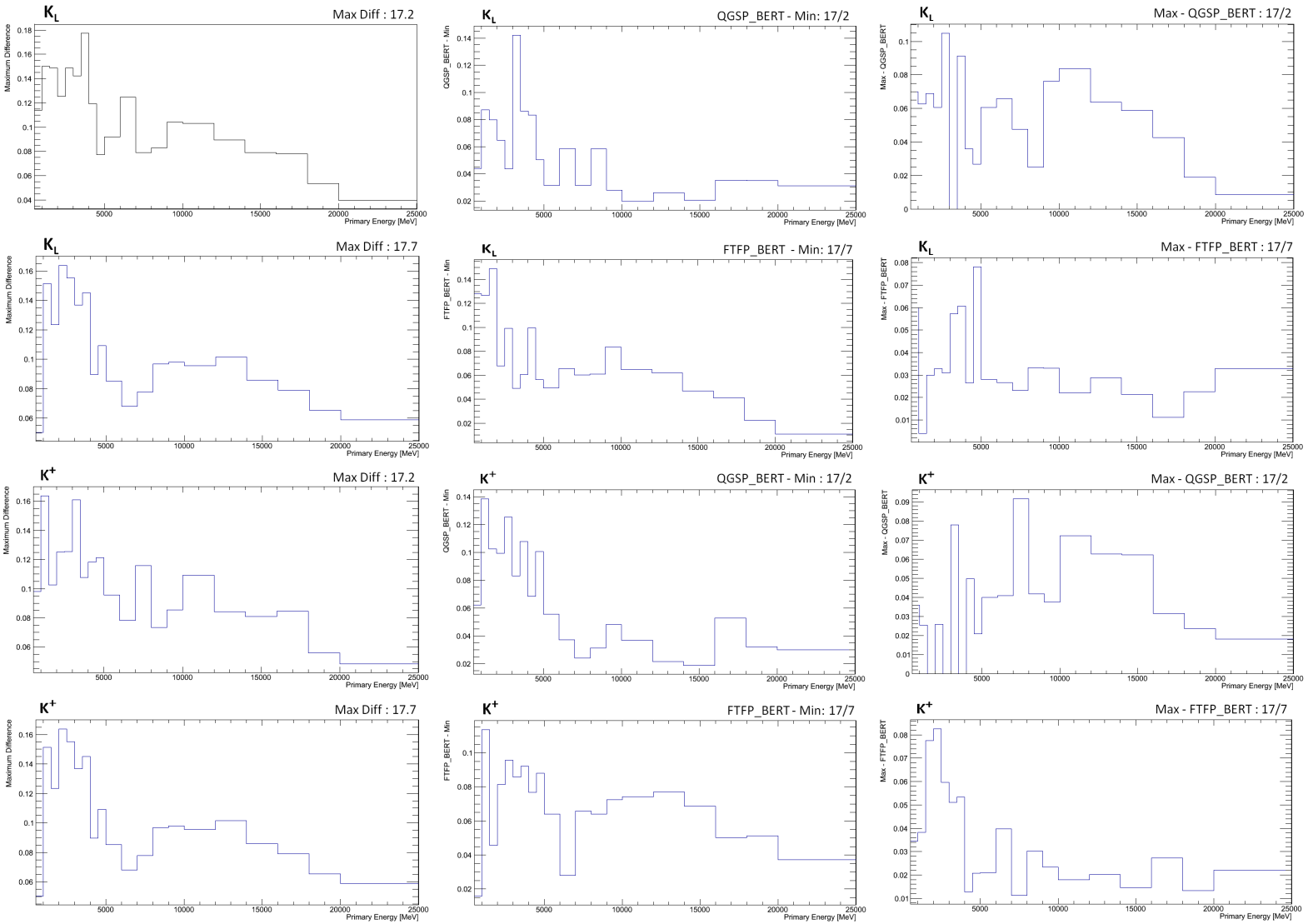


Figure A.3 – P, N: The differences decrease by increasing energies, They are almost under 10-15% .


K_L
Max Diff : 17.2
Maximum Difference
Primary Energy [MeV]
K_L
QGSP_BERT - Min: 17/2
QGSP_BERT - Min
Primary Energy [MeV]
K_L
Max - QGSP_BERT : 17/2
Max - QGSP_BERT
Primary Energy [MeV]
K_L
Max Diff : 17.7
Maximum Difference
Primary Energy [MeV]
K_L
FTFP_BERT - Min: 17/7
FTFP_BERT - Min
Primary Energy [MeV]
K_L
Max - FTFP_BERT : 17/7
Max - FTFP_BERT
Primary Energy [MeV]
K+
Max Diff : 17.2
Maximum Difference
Primary Energy [MeV]
K+
QGSP_BERT - Min: 17/2
QGSP_BERT - Min
Primary Energy [MeV]
K+
Max - QGSP_BERT : 17/2
Max - QGSP_BERT
Primary Energy [MeV]
K+
Max Diff : 17.7
Maximum Difference
Primary Energy [MeV]
K+
FTFP_BERT - Min: 17/7
FTFP_BERT - Min
Primary Energy [MeV]
K+
Max - FTFP_BERT : 17/7
Max - FTFP_BERT
Primary Energy [MeV]

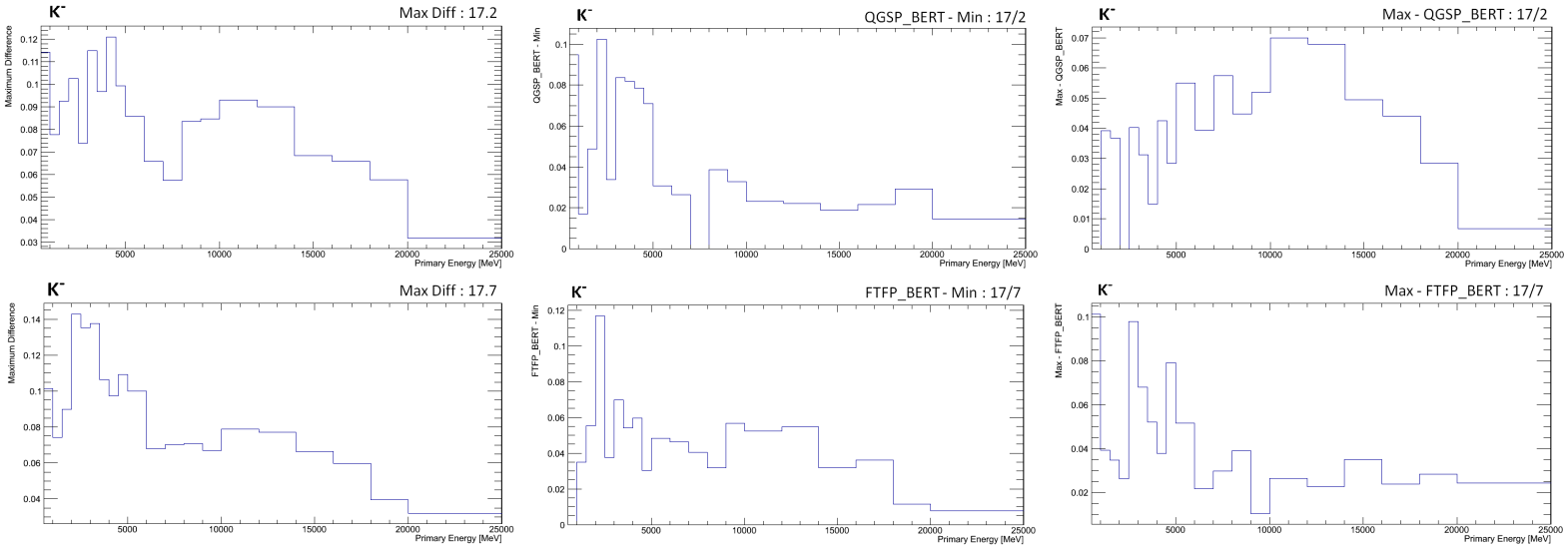


Figure A.4 – K_L, K^+, K^-: The differences generally decrease by increasing energies, They are almost under 10-20% .

Appendix B

List of the datasets used in the analysis

Table B.1 – List of MC15 datasets used for the background and signal estimates.

Sample	Name
$t\bar{t}$	mc15_13TeV.410000.PowhegPythiaEvtGen_P2012_ttbar_hdamp172p5_nonallhad.merge.DAOD_EXOT4.e3698_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.410000.PowhegPythiaEvtGen_P2012_ttbar_hdamp172p5_nonallhad.merge.DAOD_EXOT4.e3698_a766_a810_r6282_p2495
	mc15_13TeV.410001.PowhegPythiaEvtGen_P2012radHi_ttbar_hdamp345_down_nonallhad.merge.DAOD_EXOT4.e3783_s2608_r7326_r6282_p2495
	mc15_13TeV.410002.PowhegPythiaEvtGen_P2012radLo_ttbar_hdamp172_up_nonallhad.merge.DAOD_EXOT4.e3783_s2608_r7326_r6282_p2495
	mc15_13TeV.410003.aMcAtNloHerwigppEvtGen_ttbar_nonallhad.merge.DAOD_EXOT4.e4441_s2726_r7326_r6282_p2495
	mc15_13TeV.410004.PowhegHerwigppEvtGen_UEEE5_ttbar_hdamp172p5_nonallhad.merge.DAOD_EXOT4.e3836_a766_a810_r6282_p2495
	mc15_13TeV.410007.PowhegPythiaEvtGen_P2012_ttbar_hdamp172p5_allhad.merge.DAOD_EXOT4.e4135_s2608_s2183_r6869_r6282_p2452
	mc15_13TeV.410023.Sherpa_CT10_ttbar_SingleLeptonM_MEPS_NLO.merge.DAOD_EXOT4.e3959_a766_a777_r6282_p2459
	mc15_13TeV.410022.Sherpa_CT10_ttbar_SingleLeptonP_MEPS_NLO.merge.DAOD_EXOT4.e3959_a766_a777_r6282_p2459
m_tt sliced $t\bar{t}$	mc15_13TeV.301528.PowhegPythiaEvtGen_P2012_ttbar_hdamp172p5_nonallhad_mtt_1.merge.DAOD_EXOT4.e3954_s2608_r7326_r6282_p2495
	mc15_13TeV.301529.PowhegPythiaEvtGen_P2012_ttbar_hdamp172p5_nonallhad_mtt_2.merge.DAOD_EXOT4.e3954_s2608_r7326_r6282_p2495
	mc15_13TeV.301530.PowhegPythiaEvtGen_P2012_ttbar_hdamp172p5_nonallhad_mtt_3.merge.DAOD_EXOT4.e3954_s2608_r7326_r6282_p2495
	mc15_13TeV.301531.PowhegPythiaEvtGen_P2012_ttbar_hdamp172p5_nonallhad_mtt_4.merge.DAOD_EXOT4.e3954_s2608_r7326_r6282_p2495
	mc15_13TeV.301532.PowhegPythiaEvtGen_P2012_ttbar_hdamp172p5_nonallhad_mtt_5.merge.DAOD_EXOT4.e3954_s2608_r7326_r6282_p2495
Single top	mc15_13TeV.410011.PowhegPythiaEvtGen_P2012_singletop_tchan_lept_top.merge.DAOD_EXOT4.e3824_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.410012.PowhegPythiaEvtGen_P2012_singletop_tchan_lept_antitop.merge.DAOD_EXOT4.e3824_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.410013.PowhegPythiaEvtGen_P2012_Wt_inclusive_top.merge.DAOD_EXOT4.e3753_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.410014.PowhegPythiaEvtGen_P2012_Wt_inclusive_antitop.merge.DAOD_EXOT4.e3753_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.410015.PowhegPythiaEvtGen_P2012_Wt_dilepton_top.merge.DAOD_EXOT4.e3753_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.410016.PowhegPythiaEvtGen_P2012_Wt_dilepton_antitop.merge.DAOD_EXOT4.e3753_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.410025.PowhegPythiaEvtGen_P2012_SingleTopSchan_noAllHad_top.merge.DAOD_EXOT4.e3998_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.410026.PowhegPythiaEvtGen_P2012_SingleTopSchan_noAllHad_antitop.merge.DAOD_EXOT4.e3998_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.361300.Sherpa_CT10_Wenu_Pt0_70_CVetoBVeto.merge.DAOD_EXOT4.e3651_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.361301.Sherpa_CT10_Wenu_Pt0_70_CFilterBVeto.merge.DAOD_EXOT4.e3651_s2586_s2174_r7326_r6282_p2495
	mc15_13TeV.361302.Sherpa_CT10_Wenu_Pt0_70_BFilter.merge.DAOD_EXOT4.e3651_s2586_s2174_r7326_r6282_p2495
	mc15_13TeV.361303.Sherpa_CT10_Wenu_Pt70_140_CVetoBVeto.merge.DAOD_EXOT4.e3651_s2586_s2174_r7267_r6282_p2495
	mc15_13TeV.361304.Sherpa_CT10_Wenu_Pt70_140_CFilterBVeto.merge.DAOD_EXOT4.e3651_s2586_s2174_r7267_r6282_p2495
	mc15_13TeV.361305.Sherpa_CT10_Wenu_Pt70_140_BFilter.merge.DAOD_EXOT4.e3651_s2586_s2174_r7267_r6282_p2495
	mc15_13TeV.361306.Sherpa_CT10_Wenu_Pt140_280_CVetoBVeto.merge.DAOD_EXOT4.e3651_s2586_s2174_r7267_r6282_p2495
	mc15_13TeV.361307.Sherpa_CT10_Wenu_Pt140_280_CFilterBVeto.merge.DAOD_EXOT4.e3651_s2586_s2174_r7267_r6282_p2495
	mc15_13TeV.361308.Sherpa_CT10_Wenu_Pt140_280_BFilter.merge.DAOD_EXOT4.e3651_s2586_s2174_r7267_r6282_p2495
	mc15_13TeV.361309.Sherpa_CT10_Wenu_Pt280_500_CVetoBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361310.Sherpa_CT10_Wenu_Pt280_500_CFilterBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361311.Sherpa_CT10_Wenu_Pt280_500_BFilter.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361312.Sherpa_CT10_Wenu_Pt500_700_CVetoBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361313.Sherpa_CT10_Wenu_Pt500_700_CFilterBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361314.Sherpa_CT10_Wenu_Pt500_700_BFilter.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361315.Sherpa_CT10_Wenu_Pt700_1000_CVetoBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361316.Sherpa_CT10_Wenu_Pt700_1000_CFilterBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361317.Sherpa_CT10_Wenu_Pt700_1000_BFilter.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361318.Sherpa_CT10_Wenu_Pt1000_2000_CVetoBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361319.Sherpa_CT10_Wenu_Pt1000_2000_CFilterBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495

Continued on next page

Table B.1 – *Continued from previous page*

Sample	Name
	mc15_13TeV.361320.Sherpa_CT10_Wenu_Pt1000_2000_BFilter.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361321.Sherpa_CT10_Wenu_Pt2000_E_CMS_CVetoBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361322.Sherpa_CT10_Wenu_Pt2000_E_CMS_CFilterBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361323.Sherpa_CT10_Wenu_Pt2000_E_CMS_BFilter.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361324.Sherpa_CT10_Wmunu_Pt0_70_CVetoBVeto.merge.DAOD_EXOT4.e3651_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361325.Sherpa_CT10_Wmunu_Pt0_70_CFilterBVeto.merge.DAOD_EXOT4.e3651_s2586_s2174_r7267_r6282_p2495
	mc15_13TeV.361326.Sherpa_CT10_Wmunu_Pt0_70_BFilter.merge.DAOD_EXOT4.e3651_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361327.Sherpa_CT10_Wmunu_Pt70_140_CVetoBVeto.merge.DAOD_EXOT4.e3651_s2586_s2174_r7267_r6282_p2495
	mc15_13TeV.361328.Sherpa_CT10_Wmunu_Pt70_140_CFilterBVeto.merge.DAOD_EXOT4.e3651_s2586_s2174_r7267_r6282_p2495
	mc15_13TeV.361329.Sherpa_CT10_Wmunu_Pt70_140_BFilter.merge.DAOD_EXOT4.e3651_s2586_s2174_r7267_r6282_p2495
	mc15_13TeV.361330.Sherpa_CT10_Wmunu_Pt140_280_CVetoBVeto.merge.DAOD_EXOT4.e3651_s2586_s2174_r7267_r6282_p2495
	mc15_13TeV.361331.Sherpa_CT10_Wmunu_Pt140_280_CFilterBVeto.merge.DAOD_EXOT4.e3651_s2586_s2174_r7267_r6282_p2495
	mc15_13TeV.361332.Sherpa_CT10_Wmunu_Pt140_280_BFilter.merge.DAOD_EXOT4.e3651_s2586_s2174_r7267_r6282_p2495
	mc15_13TeV.361333.Sherpa_CT10_Wmunu_Pt280_500_CVetoBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361334.Sherpa_CT10_Wmunu_Pt280_500_CFilterBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361335.Sherpa_CT10_Wmunu_Pt280_500_BFilter.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361336.Sherpa_CT10_Wmunu_Pt500_700_CVetoBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361337.Sherpa_CT10_Wmunu_Pt500_700_CFilterBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361338.Sherpa_CT10_Wmunu_Pt500_700_BFilter.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361339.Sherpa_CT10_Wmunu_Pt700_1000_CVetoBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361340.Sherpa_CT10_Wmunu_Pt700_1000_CFilterBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361341.Sherpa_CT10_Wmunu_Pt700_1000_BFilter.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361342.Sherpa_CT10_Wmunu_Pt1000_2000_CVetoBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361343.Sherpa_CT10_Wmunu_Pt1000_2000_CFilterBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361344.Sherpa_CT10_Wmunu_Pt1000_2000_BFilter.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361345.Sherpa_CT10_Wmunu_Pt2000_E_CMS_CVetoBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361346.Sherpa_CT10_Wmunu_Pt2000_E_CMS_CFilterBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361347.Sherpa_CT10_Wmunu_Pt2000_E_CMS_BFilter.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361348.Sherpa_CT10_Wtaunu_Pt0_70_CVetoBVeto.merge.DAOD_EXOT4.e3733_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361349.Sherpa_CT10_Wtaunu_Pt0_70_CFilterBVeto.merge.DAOD_EXOT4.e3733_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361350.Sherpa_CT10_Wtaunu_Pt0_70_BFilter.merge.DAOD_EXOT4.e3733_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361351.Sherpa_CT10_Wtaunu_Pt70_140_CVetoBVeto.merge.DAOD_EXOT4.e3733_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361352.Sherpa_CT10_Wtaunu_Pt70_140_CFilterBVeto.merge.DAOD_EXOT4.e3733_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361353.Sherpa_CT10_Wtaunu_Pt70_140_BFilter.merge.DAOD_EXOT4.e3733_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361354.Sherpa_CT10_Wtaunu_Pt140_280_CVetoBVeto.merge.DAOD_EXOT4.e3733_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361355.Sherpa_CT10_Wtaunu_Pt140_280_CFilterBVeto.merge.DAOD_EXOT4.e3733_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361356.Sherpa_CT10_Wtaunu_Pt140_280_BFilter.merge.DAOD_EXOT4.e3733_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361357.Sherpa_CT10_Wtaunu_Pt280_500_CVetoBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361358.Sherpa_CT10_Wtaunu_Pt280_500_CFilterBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361359.Sherpa_CT10_Wtaunu_Pt280_500_BFilter.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361360.Sherpa_CT10_Wtaunu_Pt500_700_CVetoBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361361.Sherpa_CT10_Wtaunu_Pt500_700_CFilterBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361362.Sherpa_CT10_Wtaunu_Pt500_700_BFilter.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361363.Sherpa_CT10_Wtaunu_Pt700_1000_CVetoBVeto.merge.DAOD_EXOT4.e3973_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361364.Sherpa_CT10_Wtaunu_Pt700_1000_CFilterBVeto.merge.DAOD_EXOT4.e3973_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361365.Sherpa_CT10_Wtaunu_Pt700_1000_BFilter.merge.DAOD_EXOT4.e3973_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361366.Sherpa_CT10_Wtaunu_Pt1000_2000_CVetoBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361367.Sherpa_CT10_Wtaunu_Pt1000_2000_CFilterBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361368.Sherpa_CT10_Wtaunu_Pt1000_2000_BFilter.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361369.Sherpa_CT10_Wtaunu_Pt2000_E_CMS_CVetoBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361370.Sherpa_CT10_Wtaunu_Pt2000_E_CMS_CFilterBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361371.Sherpa_CT10_Wtaunu_Pt2000_E_CMS_BFilter.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361372.Sherpa_CT10_Zee_Pt0_70_CVetoBVeto.merge.DAOD_EXOT4.e3651_s2586_s2174_r7267_r6282_p2495
	mc15_13TeV.361373.Sherpa_CT10_Zee_Pt0_70_CFilterBVeto.merge.DAOD_EXOT4.e3651_s2586_s2174_r7267_r6282_p2495
	mc15_13TeV.361374.Sherpa_CT10_Zee_Pt0_70_BFilter.merge.DAOD_EXOT4.e3651_s2586_s2174_r7267_r6282_p2495
	mc15_13TeV.361375.Sherpa_CT10_Zee_Pt70_140_CVetoBVeto.merge.DAOD_EXOT4.e3651_s2586_s2174_r7267_r6282_p2495
	mc15_13TeV.361376.Sherpa_CT10_Zee_Pt70_140_CFilterBVeto.merge.DAOD_EXOT4.e3651_s2586_s2174_r7267_r6282_p2495
	mc15_13TeV.361377.Sherpa_CT10_Zee_Pt70_140_BFilter.merge.DAOD_EXOT4.e3651_s2586_s2174_r7267_r6282_p2495
	mc15_13TeV.361378.Sherpa_CT10_Zee_Pt140_280_CVetoBVeto.merge.DAOD_EXOT4.e3651_s2586_s2174_r7267_r6282_p2495
	mc15_13TeV.361379.Sherpa_CT10_Zee_Pt140_280_CFilterBVeto.merge.DAOD_EXOT4.e3651_s2586_s2174_r7267_r6282_p2495
	mc15_13TeV.361380.Sherpa_CT10_Zee_Pt140_280_BFilter.merge.DAOD_EXOT4.e3651_s2586_s2174_r7267_r6282_p2495
	mc15_13TeV.361381.Sherpa_CT10_Zee_Pt280_500_CVetoBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361382.Sherpa_CT10_Zee_Pt280_500_CFilterBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361383.Sherpa_CT10_Zee_Pt280_500_BFilter.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361384.Sherpa_CT10_Zee_Pt500_700_CVetoBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495

Continued on next page

Table B.1 – *Continued from previous page*

Sample	Name
	mc15_13TeV.361385.Sherpa_CT10_Zee_Pt500_700_CFilterBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361386.Sherpa_CT10_Zee_Pt500_700_BFilter.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361387.Sherpa_CT10_Zee_Pt700_1000_CVetoBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361388.Sherpa_CT10_Zee_Pt700_1000_CFilterBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361389.Sherpa_CT10_Zee_Pt700_1000_BFilter.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361390.Sherpa_CT10_Zee_Pt1000_2000_CVetoBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361391.Sherpa_CT10_Zee_Pt1000_2000_CFilterBVeto.merge.DAOD_EXOT4.e4133_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.361392.Sherpa_CT10_Zee_Pt1000_2000_BFilter.merge.DAOD_EXOT4.e4133_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.361393.Sherpa_CT10_Zee_Pt2000_E_CMS_CVetoBVeto.merge.DAOD_EXOT4.e4133_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.361394.Sherpa_CT10_Zee_Pt2000_E_CMS_CFilterBVeto.merge.DAOD_EXOT4.e4133_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.361395.Sherpa_CT10_Zee_Pt2000_E_CMS_BFilter.merge.DAOD_EXOT4.e4133_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.361396.Sherpa_CT10_Zmumu_Pt0_70_CVetoBVeto.merge.DAOD_EXOT4.e3651_s2586_s2174_r7267_r6282_p2495
	mc15_13TeV.361397.Sherpa_CT10_Zmumu_Pt0_70_CFilterBVeto.merge.DAOD_EXOT4.e3651_s2586_s2174_r7326_r6282_p2495
	mc15_13TeV.361398.Sherpa_CT10_Zmumu_Pt0_70_BFilter.merge.DAOD_EXOT4.e3651_s2586_s2174_r7267_r6282_p2495
	mc15_13TeV.361399.Sherpa_CT10_Zmumu_Pt70_140_CVetoBVeto.merge.DAOD_EXOT4.e3651_s2586_s2174_r7267_r6282_p2495
	mc15_13TeV.361400.Sherpa_CT10_Zmumu_Pt70_140_CFilterBVeto.merge.DAOD_EXOT4.e3651_s2586_s2174_r7267_r6282_p2495
	mc15_13TeV.361401.Sherpa_CT10_Zmumu_Pt70_140_BFilter.merge.DAOD_EXOT4.e3651_s2586_s2174_r7267_r6282_p2495
	mc15_13TeV.361402.Sherpa_CT10_Zmumu_Pt140_280_CVetoBVeto.merge.DAOD_EXOT4.e3651_s2586_s2174_r7267_r6282_p2495
	mc15_13TeV.361403.Sherpa_CT10_Zmumu_Pt140_280_CFilterBVeto.merge.DAOD_EXOT4.e3651_s2586_s2174_r7267_r6282_p2495
	mc15_13TeV.361404.Sherpa_CT10_Zmumu_Pt140_280_BFilter.merge.DAOD_EXOT4.e3651_s2586_s2174_r7267_r6282_p2495
	mc15_13TeV.361405.Sherpa_CT10_Zmumu_Pt280_500_CVetoBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361406.Sherpa_CT10_Zmumu_Pt280_500_CFilterBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361407.Sherpa_CT10_Zmumu_Pt280_500_BFilter.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361408.Sherpa_CT10_Zmumu_Pt500_700_CVetoBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361409.Sherpa_CT10_Zmumu_Pt500_700_CFilterBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361410.Sherpa_CT10_Zmumu_Pt500_700_BFilter.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361411.Sherpa_CT10_Zmumu_Pt700_1000_CVetoBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361412.Sherpa_CT10_Zmumu_Pt700_1000_CFilterBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361413.Sherpa_CT10_Zmumu_Pt700_1000_BFilter.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361414.Sherpa_CT10_Zmumu_Pt1000_2000_CVetoBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361415.Sherpa_CT10_Zmumu_Pt1000_2000_CFilterBVeto.merge.DAOD_EXOT4.e4133_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.361416.Sherpa_CT10_Zmumu_Pt1000_2000_BFilter.merge.DAOD_EXOT4.e4133_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.361417.Sherpa_CT10_Zmumu_Pt2000_E_CMS_CVetoBVeto.merge.DAOD_EXOT4.e4133_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.361418.Sherpa_CT10_Zmumu_Pt2000_E_CMS_CFilterBVeto.merge.DAOD_EXOT4.e4133_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.361419.Sherpa_CT10_Zmumu_Pt2000_E_CMS_BFilter.merge.DAOD_EXOT4.e4133_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.361420.Sherpa_CT10_Ztautau_Pt0_70_CVetoBVeto.merge.DAOD_EXOT4.e3733_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361421.Sherpa_CT10_Ztautau_Pt0_70_CFilterBVeto.merge.DAOD_EXOT4.e3733_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361422.Sherpa_CT10_Ztautau_Pt0_70_BFilter.merge.DAOD_EXOT4.e3733_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361423.Sherpa_CT10_Ztautau_Pt70_140_CVetoBVeto.merge.DAOD_EXOT4.e3733_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361424.Sherpa_CT10_Ztautau_Pt70_140_CFilterBVeto.merge.DAOD_EXOT4.e3733_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361425.Sherpa_CT10_Ztautau_Pt70_140_BFilter.merge.DAOD_EXOT4.e3733_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361426.Sherpa_CT10_Ztautau_Pt140_280_CVetoBVeto.merge.DAOD_EXOT4.e3733_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361427.Sherpa_CT10_Ztautau_Pt140_280_CFilterBVeto.merge.DAOD_EXOT4.e3733_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361428.Sherpa_CT10_Ztautau_Pt140_280_BFilter.merge.DAOD_EXOT4.e3733_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361429.Sherpa_CT10_Ztautau_Pt280_500_CVetoBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361430.Sherpa_CT10_Ztautau_Pt280_500_CFilterBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361431.Sherpa_CT10_Ztautau_Pt280_500_BFilter.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361432.Sherpa_CT10_Ztautau_Pt500_700_CVetoBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361433.Sherpa_CT10_Ztautau_Pt500_700_CFilterBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361434.Sherpa_CT10_Ztautau_Pt500_700_BFilter.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361435.Sherpa_CT10_Ztautau_Pt700_1000_CVetoBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361436.Sherpa_CT10_Ztautau_Pt700_1000_CFilterBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361437.Sherpa_CT10_Ztautau_Pt700_1000_BFilter.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361438.Sherpa_CT10_Ztautau_Pt1000_2000_CVetoBVeto.merge.DAOD_EXOT4.e3741_s2608_s2183_r7267_r6282_p2495
	mc15_13TeV.361439.Sherpa_CT10_Ztautau_Pt1000_2000_CFilterBVeto.merge.DAOD_EXOT4.e4133_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.361440.Sherpa_CT10_Ztautau_Pt1000_2000_BFilter.merge.DAOD_EXOT4.e4133_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.361441.Sherpa_CT10_Ztautau_Pt2000_E_CMS_CVetoBVeto.merge.DAOD_EXOT4.e4133_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.361442.Sherpa_CT10_Ztautau_Pt2000_E_CMS_CFilterBVeto.merge.DAOD_EXOT4.e4133_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.361443.Sherpa_CT10_Ztautau_Pt2000_E_CMS_BFilter.merge.DAOD_EXOT4.e4133_s2608_s2183_r7326_r6282_p2495
Dibosons	mc15_13TeV.361081.Sherpa_CT10_WplvWmqq.merge.DAOD_EXOT4.e3836_s2608_s2183_r6869_r6282_p2452
	mc15_13TeV.361082.Sherpa_CT10_WpqqWmlv.merge.DAOD_EXOT4.e3836_s2608_s2183_r6869_r6282_p2452
	mc15_13TeV.361083.Sherpa_CT10_WlvZqq.merge.DAOD_EXOT4.e3836_s2608_s2183_r6869_r6282_p2452
	mc15_13TeV.361084.Sherpa_CT10_WqqZll.merge.DAOD_EXOT4.e3836_s2608_s2183_r6869_r6282_p2452
	mc15_13TeV.361086.Sherpa_CT10_ZqqZll.merge.DAOD_EXOT4.e3926_s2608_s2183_r6869_r6282_p2452
	mc15_13TeV.301322.Pythia8EvtGen_A14NNPDF23LO_zprime400_tt.merge.DAOD_EXOT4.e4061_s2608_s2183_r7326_r6282_p2495

Continued on next page

Table B.1 – *Continued from previous page*

Sample	Name
	mc15_13TeV.301323.Pythia8EvtGen_A14NNPDF23LO_zprime500_tt.merge.DAOD_EXOT4.e4061_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.301324.Pythia8EvtGen_A14NNPDF23LO_zprime750_tt.merge.DAOD_EXOT4.e4061_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.301325.Pythia8EvtGen_A14NNPDF23LO_zprime1000_tt.merge.DAOD_EXOT4.e4061_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.301326.Pythia8EvtGen_A14NNPDF23LO_zprime1250_tt.merge.DAOD_EXOT4.e4061_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.301327.Pythia8EvtGen_A14NNPDF23LO_zprime1500_tt.merge.DAOD_EXOT4.e4061_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.301328.Pythia8EvtGen_A14NNPDF23LO_zprime1750_tt.merge.DAOD_EXOT4.e4061_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.301329.Pythia8EvtGen_A14NNPDF23LO_zprime2000_tt.merge.DAOD_EXOT4.e4061_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.301330.Pythia8EvtGen_A14NNPDF23LO_zprime2250_tt.merge.DAOD_EXOT4.e4061_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.301331.Pythia8EvtGen_A14NNPDF23LO_zprime2500_tt.merge.DAOD_EXOT4.e4061_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.301332.Pythia8EvtGen_A14NNPDF23LO_zprime2750_tt.merge.DAOD_EXOT4.e4061_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.301333.Pythia8EvtGen_A14NNPDF23LO_zprime3000_tt.merge.DAOD_EXOT4.e3723_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.301334.Pythia8EvtGen_A14NNPDF23LO_zprime4000_tt.merge.DAOD_EXOT4.e3723_s2608_s2183_r7326_r6282_p2495
	mc15_13TeV.301335.Pythia8EvtGen_A14NNPDF23LO_zprime5000_tt.merge.DAOD_EXOT4.e3723_s2608_s2183_r7326_r6282_p2495

Appendix C

Testing the correlations for systematic affecting the jets

This analysis assumes there is no correlation between the systematics affecting the large-R jets and the small-R jets. Following the JetEtmiss group recommendations, two others scenarios have been tested to check their impact on the limits:

- scenario 1: correlate `Akt10-Run-I` and `Akt4 eigenvector 2` for JES and JMS;
- scenario 2: correlate `Akt10-Run-I` JES and `Akt4 eigenvector 2` JES on one hand, and `Akt10-Run-I` and `Akt4 eigenvector 2` mass and Tau32 on another hand.

The results, extracted with MC15a samples, are presented in Table C.1. There is an overall agreement between the scenarios, hence the choice of the analysis can not be considered as being agressive.

mass (GeV)	$\sigma^{95CL}_{scenario}/\sigma^{95CL}_{nominal}$			
	scenario 1		scenario 2	
	expected	observed	expected	observed
400	0.97	1.00	0.97	1.00
500	0.98	1.03	0.98	1.04
750	0.95	0.94	1.11	1.05
1000	0.96	0.94	0.99	0.95
1250	0.99	1.07	0.99	1.09
1500	1.03	1.06	1.01	1.07
1750	1.03	1.09	0.99	1.07
2000	1.03	1.05	1.01	1.05
2250	1.03	1.01	1.00	1.00
2500	1.01	1.03	0.99	1.01
2750	0.98	0.98	0.98	1.01
3000	1.01	0.97	1.01	0.97

Table C.1 – Ratio of 95% CL expected and observed limits between the 2 alternative scenarios and the nominal choice of correlation of jets systematic.

Appendix D

Splitting of the b-tagging nuisance parameters

The fit that uses the usual eigenvector breakdown for the b-tagging nuisance parameter leads to a constraint of the eigenvector 0 for the light and C flavors (Fig. D.1). Such constraints are prohibited by the b-tagging group, and that's why the nominal analysis splits these two eigenvectors into six eigenvectors, depending on the track-jet p_T: $p_T < 50$, $50 < p_T < 100$ and $p_T < 100$ GeV. Such splitting allows to not reduce the constraints to a negligible level as seen in figure 5.24. The impact of the limits, presented on table D.1, is within 1% (5%) for the expected (observed) limits.

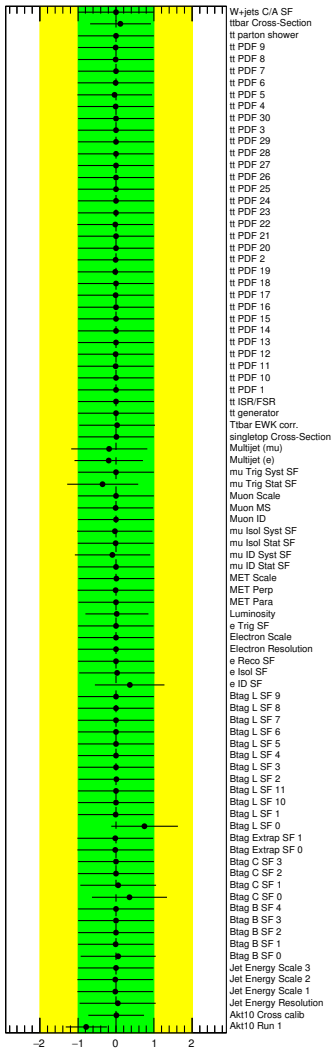


Figure D.1 – Pull of the nuisance parameters for a fit performed under the hypothesis background+Z'(2 TeV) on Atlas data.

mass (GeV)	$\sigma^{95CL}_{standard}/\sigma^{95CL}_{nominal}$	
	expected	observed
	exp	obs
400	1.00	1.00
500	1.00	1.01
750	0.99	1.04
1000	1.00	1.01
1250	1.00	1.00
1500	1.00	0.98
1750	0.99	0.96
2000	0.99	0.95
2250	0.99	0.94
2500	1.00	0.94
2750	1.00	0.97
3000	1.00	0.98
4000	1.02	1.00
5000	1.01	1.05

Table D.1 – Ratio of 95% CL expected and observed limits between the "standard" use of the b-tagging eigenvectors and the one choosen in the analysis.

Bibliography

[1] Wikimedia Commons. Standard model of elementary particles. *File: Standard Model of Elementary Particles. svg.* URL http://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg.

[2] Measurement of the $t\bar{t}$ production cross-section using $e\mu$ events with b-tagged jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. (ATLAS-CONF-2016-005), Mar 2016. URL https://cds.cern.ch/record/2138951.

[3] Rikkert Frederix and Fabio Maltoni. Top pair invariant mass distribution: A Window on new physics. *JHEP*, 01:047, 2009. doi: 10.1088/1126-6708/2009/01/047. URL https://arxiv.org/abs/0712.2355.

[4] ATLAS collaboration. Luminosity public results run ii. URL https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.

[5] The ATLAS collaboration. Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data. 2016. URL http://inspirehep.net/record/1467063.

[6] Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data. Technical Report ATLAS-CONF-2016-024, CERN, Geneva, Jun 2016. URL https://cds.cern.ch/record/2157687.

[7] Muon reconstruction performance in early âĹŽ s=13 TeV data. Technical Report ATL-PHYS-PUB-2015-037, CERN, Geneva, Aug 2015. URL https://cds.cern.ch/record/2047831.

[8] Yulia Rodina. ATLAS b-tagging performance during LHC Run 2 with the new In-

sertable B-layer. Technical Report ATL-PHYS-PROC-2015-179, CERN, Geneva, Dec 2015. URL https://cds.cern.ch/record/2112120.

[9] Performance of missing transverse momentum reconstruction for the ATLAS detector in the first proton-proton collisions at at $\sqrt{s}$= 13 TeV. Technical Report ATL-PHYS-PUB-2015-027, CERN, Geneva, Jul 2015. URL https://cds.cern.ch/record/2037904.

[10] Georges Aad et al. Performance of jet substructure techniques for large-R jets in proton-proton collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector. *JHEP*, 09:076, 2013. doi: 10.1007/JHEP09(2013)076. URL http://arxiv.org/abs/1306.4945.

[11] G Altarelli. The standard model of particle physics, encyclopedia of mathematical physics. URL https://arxiv.org/abs/hep-ph/0510281.

[12] Oliver Sim BrÃijning, Paul Collier, P Lebrun, Stephen Myers, Ranko Ostojic, John Poole, and Paul Proudlock. LHC Design Report. 2004. URL https://cds.cern.ch/record/782076.

[13] ATLAS Collaboration. The atlas experiment at the cern large hadron collider. *Journal of Instrumentation*, 3(08):S08003, 2008. URL http://stacks.iop.org/1748-0221/3/i=08/a=S08003.

[14] Atlas Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett.*, B716:1–29, 2012.

[15] Karl Jakobs and Chris Seez. The Higgs Boson discovery. *Scholarpedia*, 10(9):32413, 2015.

[16] Christopher T. Hill. Topcolor assisted technicolor. *Phys. Lett.*, B345:483–489, 1995. doi: 10.1016/0370-2693(94)01660-5.

[17] CMS Collaboration. Search for resonant $t\bar{t}$ production in proton-proton collisions at $\sqrt{s}$= 8 tev. *Physical Review D*, 93(1):012001, 2016.

[18] Michael Peskin and Dan Schroeder. An introduction to quantum field theory. *Westview Press*, 1995.

[19] Alberto Zannoni. On the quantization of the monoatomic ideal gas. *arXiv preprint cond-mat/9912229*, 1999. URL http://arxiv.org/pdf/cond-mat/9912229v1.pdf.

[20] SN Bose. PlanckâĂŹs law and light quantum hypothesis. *Z. Phys*, 26(1):178, 1924.

[21] Joseph R Incandela, Arnulf Quadt, Wolfgang Wagner, and Daniel Wicke. Status and prospects of top-quark physics. *Progress in Particle and Nuclear Physics*, 63(2): 239–292, 2009. URL https://arxiv.org/abs/0904.2499.

[22] F Abe, H Akimoto, A Akopian, MG Albrow, SR Amendolia, D Amidei, J Antos, C Anway-Wiese, S Aota, G Apollinari, et al. Observation of top quark production in p p collisions with the collider detector at fermilab. *Physical review letters*, 74(14): 2626, 1995. URL https://arxiv.org/abs/0904.2499.

[23] So Abachi, B Abbott, M Abolins, Bannanje Sripath Acharya, I Adam, DL Adams, M Adams, S Ahn, H Aihara, J Alitti, et al. Observation of the top quark. *Physical Review Letters*, 74(14):2632, 1995. URL http://arxiv.org/pdf/hep-ex/9503003.pdf.

[24] Richard D. Ball et al. Parton distributions with LHC data. *Nucl. Phys.*, B867:244–289, 2013. doi: 10.1016/j.nuclphysb.2012.10.003. URL https://arxiv.org/abs/1207.1303.

[25] Michał Czakon and Alexander Mitov. Top++: a program for the calculation of the top-pair cross-section at hadron colliders. *Computer Physics Communications*, 185 (11):2930–2938, 2014. URL http://arxiv.org/abs/1112.5675.

[26] Michiel Botje, Jon Butterworth, Amanda Cooper-Sarkar, Albert De Roeck, Joel Feltesse, Stefano Forte, Alexander Glazov, Joey Huston, Ronan McNulty, Torbjorn Sjostrand, et al. The pdf4lhc working group interim recommendations. *arXiv preprint arXiv:1101.0538*. URL http://arxiv.org/abs/1101.0538.

[27] Michal Czakon and Alexander Mitov. Nnlo corrections to top-pair production at hadron colliders: the all-fermionic scattering channels. *arXiv preprint arXiv:1207.0236*, 2012. URL http://arxiv.org/abs/1207.0236.

[28] Matteo Cacciari, Michał Czakon, Michelangelo Mangano, Alexander Mitov, and Paolo Nason. Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation. *Physics Letters B*, 710(4):612–622, 2012. URL http://arxiv.org/abs/1111.5869.

[29] Jun Gao, Marco Guzzi, Joey Huston, Hung-Liang Lai, Zhao Li, Pavel Nadolsky, Jon Pumplin, Daniel Stump, and C-P Yuan. Ct10 next-to-next-to-leading order global analysis of qcd. *Physical Review D*, 89(3):033009, 2014. URL http://arxiv.org/abs/1302.6246.

[30] Richard D Ball, Valerio Bertone, Stefano Carrazza, Christopher S Deans, Luigi Del Debbio, Stefano Forte, Alberto Guffanti, Nathan P Hartland, José I Latorre, Juan Rojo, et al. Parton distributions with lhc data. *Nuclear Physics B*, 867(2): 244–289, 2013. URL http://arxiv.org/abs/1207.1303.

[31] Georges Aad, Brad Abbott, Jalal Abdallah, S Abdel Khalek, O Abdinov, Rosemarie Aben, Babak Abi, Maris Abolins, OS AbouZeid, Halina Abramowicz, et al. Measurement of the t\ overline {t} production cross-section using e\ mu events with b-tagged jets in pp collisions at\ sqrt {s}= 7 and 8 tev with the atlas detector. *The European Physical Journal C*, 74(10):1–32, 2014.

[32] Georges Aad, B Abbott, J Abdallah, S Abdel Khalek, O Abdinov, R Aben, B Abi, M Abolins, OS Abouzeid, H Abramowicz, et al. Measurement of the top pair production cross section in 8 tev proton-proton collisions using kinematic information in the lepton+ jets final state with atlas. *Physical Review D*, 91(11):112013, 2015.

[33] Joseph R Incandela, Arnulf Quadt, Wolfgang Wagner, and Daniel Wicke. Status and prospects of top-quark physics. *Progress in Particle and Nuclear Physics*, 63(2): 239–292, 2009. URL http://arxiv.org/abs/0904.2499.

[34] K. A. Olive et al. Review of Particle Physics. *Chin. Phys.*, C38:090001, 2014. URL http://inspirehep.net/record/1315584.

[35] Vera C Rubin and W Kent Ford Jr. Rotation of the andromeda nebula from a spectroscopic survey of emission regions. *The Astrophysical Journal*, 159:379, 1970. URL http://articles.adsabs.harvard.edu/full/1970ApJ...159..379R.

[36] Richard Massey, Thomas Kitching, and Johan Richard. The dark matter of gravitational lensing. *Reports on Progress in Physics*, 73(8):086901, 2010. URL http://arxiv.org/abs/1001.1739.

[37] P. A. R. Ade et al. Planck 2015 results. XIII. Cosmological parameters. 2015. URL https://inspirehep.net/record/1343079.

[38] Gianfranco Bertone, Dan Hooper, and Joseph Silk. Particle dark matter: Evidence, candidates and constraints. *Phys. Rept.*, 405:279–390, 2005. doi: 10.1016/j.physrep.2004.08.031.

[39] Ben Lillie, Lisa Randall, and Lian-Tao Wang. The Bulk RS KK-gluon at the LHC. *JHEP*, 09:074, 2007. doi: 10.1088/1126-6708/2007/09/074. URL https://arxiv.org/abs/hep-ph/0701166.

[40] A. Liam Fitzpatrick, Jared Kaplan, Lisa Randall, and Lian-Tao Wang. Searching for the Kaluza-Klein Graviton in Bulk RS Models. *JHEP*, 09:013, 2007. doi: 10.1088/1126-6708/2007/09/013. URL https://arxiv.org/abs/hep-ph/0701150.

[41] Robert M. Harris, Christopher T. Hill, and Stephen J. Parke. Cross-section for topcolor Z-prime(t) decaying to t anti-t: Version 2.6. 1999. URL http://arxiv.org/abs/1112.4928.

[42] Georges Aad et al. A search for $t\bar{t}$ resonances with the ATLAS detector in 2.05 fb^{-1} of proton-proton collisions at $\sqrt{s} = 7$ TeV. *Eur. Phys. J.*, C72:2083, 2012. doi: 10.1140/epjc/s10052-012-2083-1. URL http://arxiv.org/abs/1205.5371.

[43] Kenneth Aamodt, A Abrahantes Quintana, R Achenbach, S Acounis, C Adler, M Aggarwal, F Agnese, G Aglieri Rinella, Z Ahammed, A Ahmad, et al. The alice experiment at the cern lhc. *Journal of Instrumentation*, 3(08):S08002, 2008. URL http://iopscience.iop.org/article/10.1088/1748-0221/3/08/S08002/pdf.

[44] A Augusto Alves Jr, LM Andrade Filho, AF Barbosa, I Bediaga, G Cernicchiaro, G Guerrer, HP Lima Jr, AA Machado, J Magnin, F Marujo, et al. The lhcb detector at the lhc. *Journal of instrumentation*, 3(08):S08005, 2008. URL http://iopscience.iop.org/article/10.1088/1748-0221/3/08/S08005/pdf.

[45] ATLAS collaboration. Track reconstruction performance of the atlas inner detector at $\sqrt{s}$= 13 tev. *CERN, Geneva, Tech. Rep. ATLPHYS-PUB-2015-018*, 2015. URL http://cds.cern.ch/record/2037683/files/ATL-PHYS-PUB-2015-018.pdf.

[46] I. Bird, K. Bos, N. Brook, D. Duellmann, C. Eck, I. Fisk, D. Foster, B. Gibbard, C. Grandi, F. Grey, et al. LHC computing Grid. Technical design report. 2005. URL http://inspirehep.net/record/692916.

[47] P. Andreetto et. al. The glite workload management system. *Journal of Physics: Conference Series*, 119(6):062007, 2008. URL http://stacks.iop.org/1742-6596/119/i=6/a=062007.

[48] Ruth Pordes, Don Petravick, Bill Kramer, Doug Olson, Miron Livny, Alain Roy, Paul Avery, Kent Blackburn, Torre Wenaus, Frank WÃijrthwein, Ian Foster, Rob Gardner, Mike Wilde, Alan Blatecky, John McGee, and Rob Quick. The open science grid. *Journal of Physics: Conference Series*, 78(1):012057, 2007. URL http://stacks.iop.org/1742-6596/78/i=1/a=012057.

[49] M. Ellert et. al. Advanced resource connector middleware for lightweight computational grids. *Future Generation Computer Systems*.

[50] T Maeno. Panda: distributed production and distributed analysis system for atlas. *Journal of Physics: Conference Series*, 119(6):062036, 2008. URL http://stacks.iop.org/1742-6596/119/i=6/a=062036.

[51] URL http://rucio.cern.ch/.

[52] Roy T Fielding and Richard N Taylor. Principled design of the modern web architecture. *ACM Transactions on Internet Technology (TOIT)*, 2(2):115–150, 2002.

[53] Rick Copeland. Essential sqlalchemy. 2008.

[54] G. Aad et al. The ATLAS Simulation Infrastructure. *Eur. Phys. J.*, C70:823–874, 2010. doi: 10.1140/epjc/s10052-010-1429-9. URL http://arxiv.org/abs/1005.4568.

[55] D. Costanzo, A. Dell'Acqua, A Di Simone, M. Gallas, A. Nairz, A. Rimoldi, J. Boudreau, and V. TSulaia. ATLAS detector simulation. 2005. URL http://inspirehep.net/record/1194793.

[56] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 05:026, 2006. URL https://arxiv.org/abs/hep-ph/0603175.

[57] J Beringer et al. Particle data group collaboration. *Phys. Rev. D*, 86:010001, 2012.

[58] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour, and B. R. Webber. HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes). *JHEP*, 01:010, 2001. doi: 10.1088/1126-6708/2001/01/010.

[59] Bo Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand. Parton Fragmentation and String Dynamics. *Phys. Rept.*, 97:31–145, 1983. doi: 10.1016/0370-1573(83)90080-7. URL http://inspirehep.net/record/189583.

[60] T. Gleisberg, Stefan. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter. Event generation with SHERPA 1.1. *JHEP*, 02:007, 2009. doi: 10.1088/1126-6708/2009/02/007. URL http://arxiv.org/abs/0811.4622.

[61] Tanju Gleisberg and Stefan Hoeche. Comix, a new matrix element generator. *JHEP*, 12:039, 2008. doi: 10.1088/1126-6708/2008/12/039. URL https://inspirehep.net/record/793879?ln=fr.

[62] Claude Duhr, Stefan Hoeche, and Fabio Maltoni. Color-dressed recursive relations for multi-parton amplitudes. *JHEP*, 08:062, 2006. doi: 10.1088/1126-6708/2006/08/062. URL http://arxiv.org/abs/hep-ph/0607057.

[63] F. Krauss, R. Kuhn, and G. Soff. AMEGIC++ 1.0: A Matrix element generator in C++. *JHEP*, 02:044, 2002. doi: 10.1088/1126-6708/2002/02/044. URL http://arxiv.org/abs/hep-ph/0109036.

[64] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. *JHEP*, 11:070, 2007. doi: 10.1088/1126-6708/2007/11/070. URL http://arxiv.org/abs/0709.2092.

[65] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. The powheg box user manual: common features. URL http://th-www.if.uj.edu.pl/~erichter/POWHEG-BOX-V2/Docs/manual-BOX.pdf.

[66] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl. Instrum. Meth.*, A506: 250–303, 2003. doi: 10.1016/S0168-9002(03)01368-8. URL http://inspirehep.net/record/593382.

[67] T Cornelissen, M Elsing, S Fleischmann, W Liebig, E Moyse, and A Salzburger. Concepts, Design and Implementation of the ATLAS New Tracking (NEWT). (ATL-SOFT-PUB-2007-007. ATL-COM-SOFT-2007-002), Mar 2007. URL https://cds.cern.ch/record/1020106.

[68] R. K. Bock. Data analysis techniques for high-energy physics experiments. *Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.*, 11:1–434, 2000. URL http://inspirehep.net/record/299776/?ln=fr.

[69] Performance of primary vertex reconstruction in proton-proton collisions at $\sqrt{s}$ =7 TeV in the ATLAS experiment. 2010. URL https://inspirehep.net/record/1204019?ln=fr.

[70] Characterization of Interaction-Point Beam Parameters Using the pp Event-Vertex Distribution Reconstructed in the ATLAS Detector at the LHC. 2010. URL http://inspirehep.net/record/1203962.

[71] Wolfgang Waltenberger. Adaptive vertex reconstruction. 2008.

[72] W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, S. Menke, S. Rajagopalan, D. Rousseau, S. Snyder, and G. Unal. Calorimeter clustering algorithms: Description and performance. 2008.

[73] Georges Aad et al. Muon reconstruction performance of the ATLAS detector in protonâĂŞproton collision data at $\sqrt{s}$ =13 TeV. *Eur. Phys. J.*, C76(5):292, 2016. doi: 10.1140/epjc/s10052-016-4120-y. URL https://inspirehep.net/search?p=find+eprint+1603.05598.

[74] Georges Aad et al. Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC protonâĂŞproton collision data. *Eur. Phys. J.*, C74(11):3130, 2014. doi: 10.1140/epjc/s10052-014-3130-x. URL http://arxiv.org/abs/1407.3935.

[75] Georges Aad et al. Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC protonâĂŞproton collision data. *Eur. Phys. J.*, C74(11):3130, 2014. doi: 10.1140/epjc/s10052-014-3130-x. URL http://arxiv.org/abs/1407.3935.

[76] Georges Aad et al. Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV. *Eur. Phys. J.*, C73(3):2304, 2013. doi: 10.1140/epjc/s10052-013-2304-2. URL http://arxiv.org/abs/1112.6426.

[77] W Lampl, S Laplace, D Lelas, P Loch, H Ma, S Menke, S Rajagopalan, D Rousseau, S Snyder, and G Unal. Calorimeter Clustering Algorithms: Description and Performance. Technical Report ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, CERN, Geneva, Apr 2008. URL https://cds.cern.ch/record/1099735.

[78] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The Anti-k(t) jet clustering algorithm. *JHEP*, 04:063, 2008. doi: 10.1088/1126-6708/2008/04/063. URL http://lanl.arxiv.org/abs/0802.1189.

[79] Georges Aad et al. Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV. *Eur. Phys. J.*, C73(3):2304, 2013. doi: 10.1140/epjc/s10052-013-2304-2. URL http://arxiv.org/abs/1112.6426.

[80] Georges Aad et al. Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1. 2016. URL http://arxiv.org/abs/0912.1342.

[81] Expected performance of the ATLAS *b*-tagging algorithms in Run-2. (ATL-PHYS-PUB-2015-022), Jul 2015. URL https://cds.cern.ch/record/2037697.

[82] Optimisation of the ATLAS *b*-tagging performance for the 2016 LHC Run. Technical Report ATL-PHYS-PUB-2016-012, CERN, Geneva, Jun 2016. URL https://cds.cern.ch/record/2160731.

[83] I Burmeister, I Connelly, A Ferretto, F Filthaut, R Goncalo, R Hawkings, H Esch, F Parodi, D Sidorov, C Schiavi, J Thomas-Wilsker, J Yu, and M Zur Nedden. Measurement of the b-tagging performance in the 2012 ATLAS data. (ATL-COM-PHYS-2013-456), Apr 2013. URL https://cds.cern.ch/record/1542409.

[84] Byron P. Roe, Hai-Jun Yang, Ji Zhu, Yong Liu, Ion Stancu, and Gordon McGregor. Boosted decision trees, an alternative to artificial neural networks. *Nucl. Instrum. Meth.*, A543(2-3):577–584, 2005. doi: 10.1016/j.nima.2004.12.018. URL https://inspirehep.net/record/657972.

[85] Performance of Missing Transverse Momentum Reconstruction in ATLAS studied in Proton-Proton Collisions recorded in 2012 at 8 TeV. Technical Report ATLAS-CONF-2013-082, CERN, Geneva, Aug 2013. URL https://cds.cern.ch/record/1570993.

[86] Stefano Frixione, Paolo Nason, and Giovanni Ridolfi. A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction. *JHEP*, 09:126, 2007. doi: 10.1088/1126-6708/2007/09/126. URL http://arxiv.org/abs/0707.3088.

[87] Emanuele Re. Single-top Wt-channel production matched with parton showers using the POWHEG method. *Eur. Phys. J.*, C71:1547, 2011. doi: 10.1140/epjc/s10052-011-1547-z. URL http://arxiv.org/abs/1009.2450.

[88] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. NLO single-top production matched with shower in POWHEG: s- and t-channel contributions. *JHEP*, 09:111, 2009. doi: 10.1007/JHEP02(2010)011,10.1088/1126-6708/2009/09/111. URL http://arxiv.org/abs/0907.4076. [Erratum: JHEP02,011(2010)].

[89] Paolo Nason. A New method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 11:040, 2004. doi: 10.1088/1126-6708/2004/11/040. URL http://arxiv.org/abs/hep-ph/0409146.

[90] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 06:043, 2010. doi: 10.1007/JHEP06(2010)043. URL http://arxiv.org/abs/1002.2581.

[91] Hung-Liang Lai, Marco Guzzi, Joey Huston, Zhao Li, Pavel M. Nadolsky, Jon Pumplin, and C. P. Yuan. New parton distributions for collider physics. *Phys. Rev.*, D82:074024, 2010. doi: 10.1103/PhysRevD.82.074024. URL http://arxiv.org/abs/1007.2241.

[92] Jun Gao, Marco Guzzi, Joey Huston, Hung-Liang Lai, Zhao Li, Pavel Nadolsky, Jon Pumplin, Daniel Stump, and C. P. Yuan. CT10 next-to-next-to-leading order global analysis of QCD. *Phys. Rev.*, D89(3):033009, 2014. doi: 10.1103/PhysRevD.89.033009. URL http://arxiv.org/abs/1302.6246.

[93] Stefano Frixione, Eric Laenen, Patrick Motylinski, Bryan R. Webber, and Chris D. White. Single-top hadroproduction in association with a W boson. *JHEP*, 07:029, 2008. doi: 10.1088/1126-6708/2008/07/029. URL http://arxiv.org/abs/0805.3067.

[94] Rikkert Frederix, Emanuele Re, and Paolo Torrielli. Single-top t-channel hadroproduction in the four-flavour scheme with POWHEG and aMC@NLO. *JHEP*, 09:130, 2012. doi: 10.1007/JHEP09(2012)130. URL http://arxiv.org/abs/1207.5391.

[95] Pierre Artoisenet, Rikkert Frederix, Olivier Mattelaer, and Robbert Rietkerk. Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations. *JHEP*, 03:015, 2013. doi: 10.1007/JHEP03(2013)015. URL http://arxiv.org/abs/1212.3460.

[96] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 05:026, 2006. doi: 10.1088/1126-6708/2006/05/026. URL http://arxiv.org/abs/hep-ph/0603175.

[97] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, Pavel M. Nadolsky, and W. K. Tung. New generation of parton distributions with uncertainties from global QCD analysis. *JHEP*, 07:012, 2002. doi: 10.1088/1126-6708/2002/07/012. URL http://arxiv.org/abs/hep-ph/0201195.

[98] Peter Zeiler Skands. Tuning Monte Carlo Generators: The Perugia Tunes. *Phys. Rev.*, D82:074018, 2010. doi: 10.1103/PhysRevD.82.074018. URL http://arxiv.org/abs/1005.3457.

[99] Ludovica Sartini, F Simeone, P Pani, N Lo Bue, G Marinaro, A Grubich, A Lobko, G Etiope, A Capone, P Favali, et al. Nuclear instruments and methods in physics research section a: Accelerators, spectrometers, detectors and associated equipment.

Nuclear Instruments and Methods in Physics Research A, 2010. URL http://www.sciencedirect.com/science/article/pii/S0168900201000894.

[100] J. H. KÃijhn, A. Scharf, and P. Uwer. Weak Interactions in Top-Quark Pair Production at Hadron Colliders: An Update. *Phys. Rev.*, D91(1):014020, 2015. doi: 10.1103/PhysRevD.91.014020. URL http://arxiv.org/abs/1305.5773.

[101] Fabio Cascioli, Philipp Maierhofer, and Stefano Pozzorini. Scattering Amplitudes with Open Loops. *Phys. Rev. Lett.*, 108:111601, 2012. doi: 10.1103/PhysRevLett.108.111601. URL http://arxiv.org/abs/1111.5206.

[102] Steffen Schumann and Frank Krauss. A Parton shower algorithm based on Catani-Seymour dipole factorisation. *JHEP*, 03:038, 2008. doi: 10.1088/1126-6708/2008/03/038. URL http://arxiv.org/abs/0709.1027.

[103] Stefan Hoeche, Frank Krauss, Marek Schonherr, and Frank Siegert. QCD matrix elements + parton showers: The NLO case. *JHEP*, 04:027, 2013. doi: 10.1007/JHEP04(2013)027. URL http://arxiv.org/abs/1207.5030.

[104] Ryan Gavin, Ye Li, Frank Petriello, and Seth Quackenbush. W Physics at the LHC with FEWZ 2.1. *Comput. Phys. Commun.*, 184:208–214, 2013. doi: 10.1016/j.cpc.2012.09.005. URL http://arxiv.org/abs/1201.5896.

[105] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. A Brief Introduction to PYTHIA 8.1. *Comput. Phys. Commun.*, 178:852–867, 2008. doi: 10.1016/j.cpc.2008.01.036. URL http://arxiv.org/abs/0710.3820.

[106] Richard D. Ball et al. Parton distributions with LHC data. *Nucl. Phys.*, B867:244–289, 2013. doi: 10.1016/j.nuclphysb.2012.10.003. URL http://arxiv.org/abs/1207.1303.

[107] ATLAS Run 1 Pythia8 tunes. Technical Report ATL-PHYS-PUB-2014-021, CERN, Geneva, Nov 2014. URL https://cds.cern.ch/record/1966419.

[108] Fabrizio Caola, Kirill Melnikov, and Markus Schulze. Complete next-to-leading order QCD description of resonant ZâĂš production and decay into $t\bar{t}$ final states. *Phys.*

Rev., D87(3):034015, 2013. doi: 10.1103/PhysRevD.87.034015. URL http://arxiv.org/abs/1211.6387.

[109] Jun Gao, Chong Sheng Li, Bo Hua Li, C. P. Yuan, and Hua Xing Zhu. 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. doi: 10.1103/PhysRevD.82.014020. URL http://arxiv.org/abs/1004.0876.

[110] Roberto Bonciani, TomÃąs Jezo, Michael Klasen, Florian Lyonnet, and Ingo Schienbein. Electroweak top-quark pair production at the LHC with Z' bosons to NLO QCD in POWHEG. *JHEP*, 02:141, 2016. doi: 10.1007/JHEP02(2016)141. URL http://arxiv.org/abs/1511.08185.

[111] G. Aad et al. The ATLAS Simulation Infrastructure. *Eur. Phys. J.*, C70:823–874, 2010. doi: 10.1140/epjc/s10052-010-1429-9. URL http://arxiv.org/abs/1005.4568.

[112] Coll ATLAS, M Beckingham, M Duehrssen, E Schmidt, M Shapiro, M Venturi, J Virzi, I Vivarelli, M Werner, S Yamamoto, and T Yamanaka. The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim. (ATL-PHYS-PUB-2010-013), Oct 2010. URL https://cds.cern.ch/record/1300517.

[113] Summary of ATLAS Pythia 8 tunes. (ATL-PHYS-PUB-2012-003), Aug 2012. URL https://cds.cern.ch/record/1474107.

[114] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Parton distributions for the LHC. *Eur. Phys. J.*, C63:189–285, 2009. doi: 10.1140/epjc/s10052-009-1072-5. URL http://arxiv.org/abs/0901.0002.

[115] Identification of high transverse momentum top quarks in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. (ATLAS-CONF-2015-036), Aug 2015. URL https://cds.cern.ch/record/2043862.

[116] URL https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/IsolationSelectionTool.

[117] David Krohn, Jesse Thaler, and Lian-Tao Wang. Jet Trimming. *JHEP*, 02:084, 2010. doi: 10.1007/JHEP02(2010)084. URL http://arxiv.org/abs/0912.1342.

[118] Georges Aad et al. Performance of jet substructure techniques for large-R jets in proton-proton collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector. *JHEP*, 09:076, 2013. doi: 10.1007/JHEP09(2013)076. URL http://arxiv.org/abs/1306.4945.

[119] S. Catani et al. New clustering algorithm for multi-jet cross-sections in $e^+e^{\hat{a}}\acute{L}\check{S}$ annihilation. . URL http://www.sciencedirect.com/science/article/pii/037026939190196W.

[120] S. D. Ellis and D. E. Soper. Successive combination jet algorithm for hadron collisions. URL http://arxiv.org/abs/hep-ph/9305266.

[121] S. Catani et al. Longitudinally invariant Kt clustering algorithms for hadron hadron collisions. . URL http://www.sciencedirect.com/science/article/pii/055032139390166M.

[122] Performance of jet substructure techniques in early $\sqrt{s}$ = 13 TeV pp collisions with the ATLAS detector. (ATLAS-CONF-2015-035), Aug 2015. URL https://cds.cern.ch/record/2041462.

[123] Boosted hadronic top identification at ATLAS for early 13 TeV data. Technical Report ATL-PHYS-PUB-2015-053, CERN, Geneva, Dec 2015. URL https://cds.cern.ch/record/2116351.

[124] Jesse Thaler and Ken Van Tilburg. Identifying Boosted Objects with N-subjettiness. *JHEP*, 03:015, 2011. doi: 10.1007/JHEP03(2011)015. URL http://arxiv.org/abs/1011.2268.

[125] Jesse Thaler and Ken Van Tilburg. Maximizing Boosted Top Identification by Minimizing N-subjettiness. *JHEP*, 02:093, 2012. doi: 10.1007/JHEP02(2012)093. URL http://arxiv.org/abs/1108.2701.

[126] Expected Performance of Boosted Higgs ($\to b\bar{b}$) Boson Identification with the ATLAS Detector at $\sqrt{s}$ = 13 TeV. (ATL-PHYS-PUB-2015-035), Aug 2015. URL https://cds.cern.ch/record/2042155.

[127] Performance of missing transverse momentum reconstruction for the ATLAS detector in the first proton-proton collisions at at $\sqrt{s}$= 13 TeV. (ATL-PHYS-PUB-2015-027), Jul 2015. URL https://cds.cern.ch/record/2037904.

[128] . URL https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/JetUncertainties2015Prerec.

[129] . URL https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/JetResolution2015Prerecom.

[130] URL https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JetUncertainties2015PrerecLargeR.

[131] . URL https://twiki.cern.ch/twiki/bin/view/LHCPhysics/TtbarNNLO.

[132] Michiel Botje et al. The PDF4LHC Working Group Interim Recommendations. 2011. URL https://arxiv.org/abs/1101.0538.

[133] . URL https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TopSystematics2015.

[134] Nikolaos Kidonakis. Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-. *Phys. Rev.*, D82:054018, 2010. doi: 10.1103/PhysRevD.82.054018. URL https://arxiv.org/abs/1005.4451.

[135] Nikolaos Kidonakis. NNLL resummation for s-channel single top quark production. *Phys. Rev.*, D81:054028, 2010. doi: 10.1103/PhysRevD.81.054028. URL https://arxiv.org/abs/1001.5034.

[136] Georgios Choudalakis. On hypothesis testing, trials factor, hypertests and the BumpHunter. 2011. URL https://inspirehep.net/record/883244/files/arXiv:1101.0390.pdf.

[137] URL https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/TtHFitter.

[138] Search for heavy particles decaying to pairs of highly-boosted top quarks using lepton-plus-jets events in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. (ATLAS-CONF-2016-014), Mar 2016. URL https://cds.cern.ch/record/2141001.

ROYAUME DU MAROC
Université Mohammed V
Faculté des sciences
Rabat
جامعة محمد الخامس - كلية العلوم
الرباط

Abstract

The first part of this thesis is dedicated to the extraction of jet energy scale corrections for the ATLAS experiment at the LHC. The systematic validation of the different Geant4 physics lists are presented. Various extraction methods are described, with a particular interest on corrections extraction using a wider choice of physics lists: QGSP BERT, QGSP BIC, QGSP FTFP BIC, FTFP BERT, QBBC and FTF BIC. These corrections are necessary for all analyses using jets and are already used by the whole collaboration.

The second part presents the searches for new physics beyond the Standard Model(SM). Many models of new physics predict new particles with enhanced coupling to the top quark, which can show up as resonances in the top-antitop invariant mass spectrum. The analysis is performed with the ATLAS detector at the Large Hadron Collider (LHC) at the European Organization for Nuclear Research (CERN) located near Geneva in Switzerland. The data collected in 2015 are produced using proton-proton (pp) collisions at the center-of-mass energy of $\sqrt{s}$ = 13 TeV with an integrated luminosity of 3.2 fb^{-1}. This energy was the highest energy ever reached in a laboratory at this time. Events consistent with top-quark pair production are selected by requiring a single isolated charged lepton, missing transverse momentum, and a high-transverse momentum jet with substructure compatible with that of a hadronic top-quark decay. Jets identified as likely to contain b-hadrons are used to further control the background from other SM processes. The $t\bar{t}$ system is reconstructed using boosted techniques. A comparison between data and estimated background showed no significant excesses. Hence 95% Condfidence Level (C.L) upper limits on the cross section times acceptance have been set for Z' benchmark model.

Key words: ATLAS, LHC, top quark, jet systematic validation, resonances top- antitop, leptophobic topcolor Z', Grid Computing, Simulation, GEANT4.