

N° d'ordre : 3702

THESE

En vue de l'obtention du : **DOCTORAT**

Structure de Recherche : Laboratoire de Matière Condensée et Sciences
interdisciplinaires : Unité de Recherche Labellisée CNSR

Discipline : Physique

Spécialité : Physique des particules

Présentée et soutenue le 29/10/2022 par :

Badr-eddine NGAIR

**Search for new physics in Top-Antitop Quarks Pair-Production with the ATLAS
detector at the LHC and missing transverse energy performance**

Devant le Jury

Mohamed CHABAB	PES, Université Cadi Ayyad, Faculté des Sciences Semlalia, Marrakech	Président
Mohammed LOULIDI	PES, Université Mohammed V, Faculté des Sciences, Rabat	Rapporteur/Examineur
Abdesslam ARHRIB	PES, Université Abdelmalek Essaadi, Faculté des Sciences et Techniques, Tanger	Rapporteur/Examineur
Hamid EZ-ZAHRAOUI	PES, Université Mohammed V, Faculté des Sciences, Rabat	Rapporteur/Examineur
Abdeslam HOUMMADA	PES, Université Hassan II, Académie Hassan II des Sciences et Techniques, Casablanca	Examineur
Farida FASSI	PH, Université Mohammed V, Faculté des Sciences, Rabat	Directeur

Année Universitaire : 2022/2023

Dedication

I dedicate my thesis work to my family and my friends. A special feeling of gratitude to my loving parents, Lahcen and Fatima whose words of encouragement and push for tenacity ring in my ears. I didicated this to my sister Wissal, my brothers Reda and Marouan for your constant love and support.

Acknowledgment

Foremost, I am sincerely grateful to **Mr. Hamid EZ-ZAHRAOUY**, Professor at the faculty of sciences of Rabat and director of "Laboratory of Condensed Matter and Interdisciplinary Sciences: Unit of Research Labeled CNSRT". The office's door and the cell phone of him were always open whenever I ran into trouble or had a question.

I want to express my most sincere and deep gratitude to my supervisor **Mme. Farida FASSI**, professor at faculty of sciences Rabat for giving me the unique opportunity to work in the most exciting international environments for scientific research, as the CERN is. It is needless to say that without her this thesis wouldn't be what it is, so I will focus on what is not so obvious. Her brilliance and kindness really inspired me. Her guidance, constant support and patience during my PhD meant a lot to me, and this work would not have been possible without all that.

My sincere gratitude goes to **Mr. Mohamed CHABAB**, professor at the faculty of sciences Semlalia Marrakesh, who accepted the presidency of the jury. I appreciate his comments and discussions about this thesis work.

My Acknowledgment goes also to **Mr. Mohammed LOULIDI**, professor at faculty of sciences Rabat for accepting to be reporter and examiner of my thesis report and providing feedback to improve it.

I am especially grateful to **Mr. Abdeslam ARHRIB**, professor at faculty of sciences and technology Tangier, who is reporter and examiner of my thesis work, for his useful comments.

My sincere gratitude also goes to **Mr. Hamid EZ-ZAHRAOUY**, professor at faculty of sciences Rabat for being reporter and examiner of my thesis work. I really appreciate his detailed comments and interesting discussions.

I would like to thank **Mr. Abdeslam HOUMMADA**, professor at the Académie Hassan II of Sciences and Technology, who is examiner of this thesis. I am extremely thankful and indebted for his comments about my work.

Résumé

La pièce maîtresse de cette thèse consiste en la recherche de nouvelles résonances se désintégrant en paire de quarks $t\bar{t}$ dans le canal semi-leptonique, ainsi de mettre des limites supérieures sur la section efficace de production de particules lourdes prédites par plusieurs modèles de références. La masse invariante du système $t\bar{t}$ est l'observable principal de cette recherche, qui est utilisée pour tester la compatibilité des données réelles par rapport aux prédictions du MS. Pour la région à grand mass, les quarks top sont produits avec une grande impulsion et leurs produits de désintégrations sont collimatés dans un seul jet possédant un grand rayon, ce qui nécessite l'utilisation des techniques de sous-structure de jet pour identifier les paires $t\bar{t}$ (boosté), tandis que pour de basses valeurs de mass, les événements sélectionnés sont dominés par des quarks top où les produits de désintégration sont bien séparés (résolu). Des techniques améliorées sont spécifiquement adaptées à l'estimation du bruit de fond multi-jets, aussi la sélection et la reconstruction des événements sont développées et utilisées. Des résultats préliminaires sur les limites supérieures du rapport d'embranchement fois la section efficace de production sont établis à un niveau de confiance de 95% pour le boson Z' .

Mots clés : Quark Top, section efficace, QCD-multi-jet, semi-leptonic, Modèle Standard

Abstract

The main part of this dissertation is the search for a heavy resonance decaying into a $t\bar{t}$ pair in the lepton plus jets final state, also establishes upper bounds on the production cross section of new heavy particles in a number of benchmark scenarios. The invariant mass of the $t\bar{t}$ system is the principal observable on this search, which is introduced to evaluate the degree of compatibility of the data with respect to the SM-only hypothesis. For the high mass region, the chosen events are dominated by top-quarks with collimated decay products, then, jet substructure techniques should be used to pick out the $t\bar{t}$ pairs (boosted scenario), while for low mass, the chosen events are dominated by top-quarks with well separated decay products (resolved scenario). Improved techniques specifically tailored to the QCD multi-jet background estimation, event selection and reconstruction are developed and used. No evidence for a resonant Z' boson is found and, as a result, constraints are set on theoretical models. Preliminary results on the upper limits on the production cross-section times branching ratio are established at a 95% credibility level for Z' boson.

Key Words: Quark Top, cross-section, QCD-multi-jet, semi-leptonic, Standard Model

Résumé détaillé

Le modèle standard (MS) de la physique des particules s'est imposé comme une théorie ayant de solides fondements expérimentaux ce qui lui permet de fournir actuellement une meilleure compréhension de la phénoménologie des particules connues. Cependant, le MS n'est pas la théorie ultime de la physique car plusieurs observations et interrogations théoriques suggèrent l'existence d'une nouvelle physique, peut être à l'échelle d'énergie de quelques TeV. De nombreuses extensions du MS prédisent la présence des particules lourdes qui se désintègrent en paire de quarks top-antitop ($t\bar{t}$). Cette thèse porte sur la recherche de nouvelles résonances dans le secteur de quark top, qui est souvent considéré comme une fenêtre vers une nouvelle physique grâce à ses propriétés uniques.

Les collisions de particules à hautes énergies sont un moyen permettant de mettre à l'épreuve les théories qui tentent d'aller au-delà du Modèle Standard. Le Large Hadron Collider et le détecteur ATLAS sont des instruments qui offrent la possibilité de réaliser des collisions très énergétiques et de détecter les particules produites, ouvrant ainsi potentiellement la voie vers la découverte de la nouvelle physique. Les travaux présentés dans ce manuscrit exploitent les données issues des collisions proton-proton à 13 TeV observés par le détecteur ATLAS auprès du collisionneur LHC pendant la période 2015 et 2018, ce qui correspond à une luminosité de 139 fb^{-1} .

Le quark top est la particule la plus lourde du modèle standard, par conséquent, il a un très grand couplage avec le boson de Higgs et à d'autres résonances dans de nombreux modèles exotiques. Beaucoup de ces modèles exotiques proposent également un mécanisme alternatif de rupture de symétrie électrofaible, qui incorpore de nouvelles particules avec un couplage plus important au quark top qu'aux quarks plus légers. La pièce maîtresse de cette

thèse consiste en la recherche de nouvelles résonances se désintégrant en paire de quarks $t\bar{t}$ dans le canal semi-leptonique, ainsi de mettre des limites supérieures sur la section efficace de production de particules lourdes prédites par plusieurs modèles de références. La masse invariante du système $t\bar{t}$ ($m_{t\bar{t}}$) est l'observable principal de cette recherche, qui est utilisée pour tester la compatibilité des données réelles par rapport aux prédictions du MS. Pour la région à grand $m_{t\bar{t}}$, les quarks top sont produits avec une grande impulsion et leurs produits de désintégrations sont collimatés dans un seul jet possédant un grand rayon (large-R jet), ce qui nécessite l'utilisation des techniques de sous-structure de jet pour identifier les paires $t\bar{t}$ (scénario boosté), tandis que pour de basses valeurs de $m_{t\bar{t}}$, les événements sélectionnés sont dominés par des quarks top où les produits de désintégration sont bien séparés (scénario résolu). Des techniques améliorées sont spécifiquement adaptées à l'estimation du bruit de fond multi-jets, aussi la sélection et la reconstruction des événements sont développées et utilisées. Aucune preuve de résonance n'est trouvée, par conséquent, des contraintes sont établies sur les modèles théoriques. Des résultats préliminaires sur les limites supérieures du rapport d'embranchement fois la section efficace de production sont établis à un niveau de confiance de 95% pour le boson Z'_{TC2} .

L' E_T^{miss} est reconstruit à partir des signaux de toutes les particules détectées dans l'état final, et en tant que tel fortement dépend de l'échelle d'énergie et de la résolution des « objets durs » reconstruits et calibrés : électrons, photons, muons, leptons τ à décomposition hadronique et jets. Flux de quantité de mouvement reconstruit non attribué à tous les objets durs sont reconstruits comme le "terme souple" d' E_T^{miss} . E_T^{miss} parfaitement reconstruit doit être égal à la composante transverse de la somme des impulsions des particules invisibles dans l'événement. Cependant, comme E_T^{miss} est une qualité compliquée au niveau de l'événement construite à l'aide de nombreux signaux différents dans tous les sous-systèmes de détection, elle est susceptible de nombreuses imperfections. La deuxième partie de ce manuscrit est consacrée à des études approfondies de suppression du chevauchement électron(photon)-jet et à l'obtention de l'énergie transversale manquante optimale. Cette performance est d'une importance capitale pour de nombreuses analyses de physique du LHC, en particulier pour l'analyse principale de cette thèse en raison de la présence de neutrino dans l'état final.

List of publications

- ATLAS collaboration. "Search for new phenomena with top quark pairs in final states with one lepton, jets, and missing transverse momentum in p p collisions at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS detector", J. High Energ. Phys. 2021, 174 (2021) [1].
- ATLAS collaboration. "Performance of the missing transverse momentum triggers for the ATLAS detector during Run-2 data taking". J. High Energ. Phys. 2020, 80 (2020) [2].
- ATLAS collaboration. "Differential cross-section measurements for the electroweak production of dijets in association with a Z boson in proton–proton collisions at ATLAS". Eur. Phys. J. C 81, 163 (2021). [3]
- ATLAS Collaboration. "Configuration and performance of the ATLAS b-jet triggers in Run 2". Eur. Phys. J. C 81, 1087 (2021). [4]
- ATLAS Collaboration. "Measurements of the inclusive and differential production cross sections of a top-quark–antiquark pair in association with a Z boson at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS detector". Eur. Phys. J. C 81, 737 (2021). [5]
- Badr-eddine Ngair and Farida Fassi, "Search of new resonances decaying into top quark pairs in the lepton+jet final state in proton-proton collisions at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS detector", arXiv:2206.12978v1 [hep-ex] 26 Jun 2022 [7]
- ATLAS collaboration. "Search for $t\bar{t}$ resonances in the lepton+jet channel using 139 fb^{-1} p p collisions at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS experiment" ANA-EXOT-2018-49-INT1. (under review, to be published).
 - Scopus author ID: [57211206478](#).
 - ORCID: [0000-0002-5807-8535](#).
 - INSPIRE ID: [1781470](#).

List of contributions

- ATLAS Qualification Task with Jet/Etmiss subgroup: Tuning of electron(photon)-jet overlap removal and performance of the.
 - * Update the METPerfAnalysis package to reoptimize the overlap removal for electrons/photons over jets in the E_T^{miss} calculation.
- Estimation of QCD multi-jet background using the Matrix Method for $t\bar{t}$ resonances search in the $\ell + jets$ final state analysis:
 - * Development and integration of the Matrix Method to the offline reconstruction package: “TopNtupleAnalysis”.
 - * Production of the loose and tight samples from real data and MC simulation.
 - * Update and implementation of a C++ code for the derivation of the fake and real efficiencies.
 - * Prediction of the overall yields and the kinematic distributions of the QCD background in several control regions.
 - * Implementation of a ΔR correction factor to the real and fake parameterisations to improve the QCD multi-jet estimation.
 - * Estimation of a conservative uncertainty associated to the QCD background.
 - * Update and implementation of a C++ code for smoothing the nominal distribution of the QCD multi-jet background.
- Implementation of the DL1r b-tagging algorithm with the 77% efficient working point in "HQTtResonancesTools" package used for the identification of b-jet.
- Study of the performance of the χ^2 algorithm used for the selection and reconstruction of events in the resolved regime.

-
- Develop a Kinematic Fit code using python to optimize the χ^2 parameters values.
 - Development and integration of python code studying the acceptance times efficiency ($A \times \epsilon$) for signal benchmark scenarios.
 - Reconstruction of the $t\bar{t}$ invariant mass ($m_{t\bar{t}}^{reco}$) used to search for deficit or excess with respect to the $t\bar{t}$ SM prediction.
 - Contribution to LOI for ASFAP strategy “ Searches for top-antitop quark resonances at ATLAS detector.”

List of abbreviations

Name	Symbol
Standard Model	SM
Beyond Standard Model	BSM
Spontaneous Symmetry Breaking	SSb
Muon Spectrometer	MS
Inner Detector	ID
A Toroidal LHC Apparatus	ATLAS
Large Hadron Collider	LHC
Compact Muon Solenoid	CMS
Hadronic Calorimetre	HadCal
Electromagnetic Calorimetre	EMCal
Matrix Method	MM
Electromagnetic Topo Cluster	EMTopo
Particle Flow	PFlow
Quantum Chromo Dynamics	QCD
Dark Matter	DM
Missing Transverse Momentum	MET (E_T^{miss})
Overlap Removal	OR
Top anti-Top quark pair	$t\bar{t}$
Mass of the top anti-top system	$m_{t\bar{t}}$

List of Tables

1.1	Fundamental forces and its corresponding mediators with charge types described by the Standard Model [31].	5
1.2	The SM fermion quantum numbers (weak hypercharge (Y), electric charge (Q) and its third component (T_3)) and generations.	8
1.3	Production cross sections of the single top in pp collisions at $\sqrt{s} = 13$ TeV [10].	18
4.1	Contributions to E_T^{miss} and $\sum E_T$ from hard objects such as photons (γ), electrons (e), muons (μ), τ – leptons (τ_{had}), jets, and tracks from the soft term [28].	60
4.2	PDF sets, Generators, and MC tunes employed in the OR (photon/electron, jet) analysis for the E_T^{miss}	63
4.3	E_T^{miss} performance for $Z \rightarrow ee$ events using EMTopo jets collection. The jet is treated as a real jet if $\Delta p_T(e, jet)$ and $Ratio_E(e, jet)$ are larger than a tuned threshold. The mean and RMS are relatively unaffected.	70
4.4	E_T^{miss} performance for $Z \rightarrow ee$ events using PFlow jets collection. The jet is treated as a real jet if $\Delta p_T(e, jet)$ and $Ratio_E(e, jet)$ are larger than a tuned threshold. The mean and RMS are relatively unaffected.	71
4.5	E_T^{miss} performance for $Z \rightarrow ee$ simulated events using EMTopo jets collection.	71
4.6	E_T^{miss} performance for $Z \rightarrow ee$ simulated events using PFlow jets collection.	72
4.7	E_T^{miss} performance for $t\bar{t} + \gamma$ events using EMTopo jets collection. The jet is treated as a real jet if $\Delta p_T(\gamma, jet)$ and $Ratio_E(\gamma, jet)$ are larger than a tuned threshold. The mean and RMS are relatively unaffected.	75
4.8	E_T^{miss} performance for $t\bar{t} + \gamma$ events using PFlow jets collection. The jet is treated as a real jet if $\Delta p_T(\gamma, jet)$ and $Ratio_E(\gamma, jet)$ are larger than a tuned threshold. The mean and RMS are relatively unaffected.	76

4.9	E_T^{miss} performance for $t\bar{t} + \gamma$ simulated events using EMTopo jets collection. . . .	77
4.10	E_T^{miss} performance for $t\bar{t} + \gamma$ simulated events using PFlow jets collection. . . .	77
6.1	List of the single-lepton high Level triggers (HLT) used in this analysis. . . .	106
6.2	Definitions of the leptons used for the matrix method QCD estimation. . . .	111
6.3	Impact of uncertainties on the expected number of total background events and 4 TeV (pole mass) Z'_{TC2} signal events in the four signal regions.	132
6.4	Observed data and background event yields after the boosted and resolved selections in the $e + jets$ and $\mu + jets$ channels. The total uncertainty on the expected background yields is derived by summing all systematic uncertainties and the MC statistical uncertainty in quadrature.	136

List of Figures

1.1	The SM fundamental particles. Information from [8].	6
1.2	Interaction vertices of QCD.	11
1.3	The ground state of the Higgs-Potential for $\lambda > 0$, $\mu^2 > 0$ (a) and $\mu^2 < 0$ (b). Pictures taken from Reference [9].	12
1.4	m_t measurements done with the CMS and ATLAS experiments at LHC, compared with LHC and Tevatron+LHC results. Systematic and statistics uncertainties are shown [10].	17
1.5	Feynman diagrams at the LO for the top quark pairs production. Bottom: $q\bar{q}$ annihilation process. Top: gluon-gluon fusion processes [11].	18
1.6	$t\bar{t}$ production cross-section at LHC and Tevatron, in term of the centre of mass energy relative to the NNLO QCD calculation. These calculations assuming $m_t = 172.5$ GeV [10].	19
1.7	Feynman diagram at LO for the single top production, s-channel [11]. . . .	20
1.8	Feynman diagram at LO for the single top production, t-channel [11]. . . .	20
1.9	Feynman diagram at LO for the single top production, Wt-channel [11]. . .	21
1.10	$t \rightarrow Wb$ Feynman diagram, illustrating the coupling related to $ V_{tb} $ [11]. . .	21
1.11	From right to left: di-leptonic, semi-leptonic and all hadronic top pair decay Feynman diagram [11].	21
1.12	Top quark pairs branching ratios [11].	22
1.13	Summary of several Standard Model total production cross section measurements, corrected for branching fractions, compared to the corresponding theoretical expectations [12].	23
1.14	Feynman diagrams at Leading-order for the studied signal processes in this search. The Z' (a) and Kaluza–Klein gluons (g_{KK}) have spin 1 (b), while the Kaluza–Klein graviton (G_{KK}) has spin 2 (c).	24

2.1	The total integrated luminosity that ATLAS received and captured between 2015 and 2018 while collecting stable beam pp collision data at $\sqrt{s} = 13 \text{ TeV}$ [13]. . . .	28
2.2	The distribution of the mean number of interactions per bunch crossing, μ , for the entire Run2 pp collision dataset at $\sqrt{s} = 13 \text{ TeV}$, is luminosity-weighted [13]. . . .	28
2.3	The CERN accelerator complex [14].	30
2.4	Projected LHC performance through 2037, showing preliminary dates for long shutdowns of the LHC and projected luminosities. Image source [15].	31
2.5	The ATLAS detector with labels on each of its sub-detectors. A few human beings are present in the image as a scale reference. Figure from the ATLAS public web pages [16].	32
2.6	Diagram of the ATLAS coordinate system. The positive x – axis points towards the centre of the LHC, and the y – axis points upwards.	34
2.7	The three toroids positioned between the outer layers of the muon system and the solenoid encircling the ID in the ATLAS magnetic system [17].	35
2.8	A sketch of the ATLAS inner detector, labelling the individual components. Credit CERN [16].	36
2.9	Components of the ATLAS Inner Detector shown from the side. Credit CERN [16].	37
2.10	Sketch of the ATLAS calorimeters. Credit CERN [16].	38
2.11	a picture of the ECAL module at barrel $\eta = 0$ [18].	39
2.12	The ATLAS Muon Spectrometer and toroid magnets. Credit CERN [16]. . . .	40
2.13	The components important for triggering in the ATLAS TDAQ system in Run 2. In 2015, L1Topo and FTK were put into service [19].	41
2.14	In a fill taken in 2018, HLT physics trigger group rates are shown in term of time [20].	43
3.1	An illustration of the ATLAS particle detection process [21].	45
3.2	Illustration of the trajectory of an electron (in solid red) through the tracking detectors and the ecal. The dashed line shows the trajectory of a photon emitted by the electron in the tracking system. Taken from Ref [22]	46
3.3	Example of a supercluster presenting a seed electron cluster and a satellite photon cluster [23].	47

3.4	Total electron reconstruction efficiency as a function of true electron. E_T calculated from single electron simulation sample. (b) Reconstructed efficiency relative to reconstructed clusters, given by the number of reconstructed electron candidates N_{reco} , divided by the number of EM-cluster candidates $N_{cluster}$, as a function of electron transverse energy, E_T calculated from $Z \rightarrow ee$ events. Figure taken from [22]	47
3.5	The efficiency of electron identification as a function of E_T using different working points. Taken from Ref [24]	48
3.6	Muon reconstruction techniques based on signals from several ATLAS detector regions.	50
3.7	Muon identification techniques based on signals from several ATLAS detector regions [25].	51
3.8	Using the anti-kt technique, a parton-level event with several arbitrary soft contributions was grouped with a $R=1$ value. Nice circular distribution of jet with highest p_T is achieved.	52
3.9	JES calibration procedure for the jet energy scale [26].	53
3.10	Shape of hard-scatter (HS) and JVT for pileup jets. While HS jets collect for high JVT values, pile-up jets only manage to do so for low JVT levels. Jets without an accompanying track are impacted by $JVF = -1$ [27].	54
3.11	An example of a schematic diagram that compares secondary vertex generation from long-lived particle decay to other jets started from the main vertex.	55
4.1	Distributions of E_T^{miss} , E_x^{miss} and E_y^{miss} for sample with inclusive jets of $Z \rightarrow \mu\mu$ extracted from data with respect to the MC simulations including all relevant backgrounds. The last bin of each distribution includes the overflow, and the first bin contains the underflow in 1 (b) and 1 (c). The respective ratios between data and expectation are shown in the lower panel with the shaded areas indicating the total uncertainty for MC simulations. Figures from Ref [28].	62

4.2	Matching approach employed in the E_T^{miss} map to remove duplicates. Each circle refers to a cluster. The solid circle in purple, green, and red means the cluster is associated with the object. The cluster is then removed if it is matched with an object, otherwise the cluster is recognized as a jet in blue. The lines connect clusters and objects through a ΔR matching [29].	65
4.3	Overlapping electrons/photons and jets. (1) Real jet close to real electron/photon. (2) Jets from pile-up or electrons/photons faking jets in the calorimeter. (3) Real jet and fake electron/photon.	66
4.4	Comparison of $\Delta p_T(e, jet)$ (GeV) and $Ratio_E(e, jet)$ performance for the jets matching the truth (blue) and not matching the truth (red) using $Z \rightarrow ee$ simulated events.	68
4.5	Comparison of $\Delta p_T(\gamma, jet)$ (GeV) and $Ratio_E(\gamma, jet)$ performance for the jets matching the truth (blue) and not matching the truth (red) using $t\bar{t} + \gamma$ simulated events.	68
4.6	$\Delta p_T(e, jet)$ (GeV) as a function of $Ratio_E(e, jet)$ for the jets matching the truth (right-hand) and not matching the truth (left-hand) for events from $Z \rightarrow ee$ process.	69
4.7	Comparison of the E_T^{miss} (TST MET means track-based algorithm is used for the soft term) performance between old (magenta) and new (black) jet-electron overlap removal (OR) procedure in $Z \rightarrow ee$ simulation. The E_T^{miss} tail has been reduced using EMTopo jets collection.	72
4.8	Comparison of the E_T^{miss} (TST MET means track-based algorithm is used for the soft term) performance between old (magenta) and new (black) jet-electron overlap removal procedure in $Z \rightarrow ee$ simulation. The E_T^{miss} tail has been reduced using PFlow jets collection.	72
4.9	$\Delta p_T(\gamma, jet)$ (GeV) as a function of $Ratio_E(\gamma, jet)$ for the jets matching the truth (left-hand) and not matching the truth (right-hand) for events from $t\bar{t} + \gamma$ process.	73
4.10	Comparison of the $ Reconstructed - Truth E_T^{miss}$ performance between old (magenta) and new (black) jet-photon overlap removal procedure in $t\bar{t} + \gamma$ simulation.	77
4.11	Comparison of the $ Reconstructed - Truth E_T^{miss}$ performance between old (magenta) and new (black) jet-photon overlap removal procedure in $t\bar{t} + \gamma$ simulation.	78
4.12	$P_{ }$ and $P_{\perp}$ projections of p_T^{soft} along p_T^{hard} for $Z \rightarrow \mu\mu$ for a final state (a) without any jets and (b) with one jet. Figures from Ref [28].	80

4.13	(a) mean value of $\langle P_{\parallel} \rangle$ (b) variance $RMS_{\parallel}^2$ for $Z \rightarrow \mu\mu$ events with $N_{jet} = 0$, for data and two different MC simulations, displayed in term of p_T^{hard} . The variance $RMS_{\perp}^2$ is displayed in (c) for similar event samples. Figures from Ref [28].	82
4.14	(a) Shape of E_T^{miss} , (b) $\sum E_T$ and (c) the soft term displayed for both data and MC events satisfying the $Z \rightarrow \mu\mu$ selection. The data/MC is shown in the lower panel of each figure. The red band shows the associated systematic and MC statistical uncertainties. Figures from Ref [28].	85
4.15	(a) Distributions of the total E_T^{miss} in the inclusive case, (b) with $N_{jet} = 0$ case, as well as (c) the soft term $E_T^{miss,soft}$ reconstructed in $N_{jet} = 0$. The data/MC is shown in the lower panel of each figure. The red band shows the associated systematic and MC statistical uncertainties. Figures from Ref [28].	87
5.1	The number of observed events in a specific bin, n_{obs} assuming H_0 and H_1 according to the PDFs in Equations 5.3 and 5.4 (left). Test statistic t distribution estimated from the number of observed events where the test statistic is employed here is the ratio of the log likelihood (right).	90
5.2	(left) Demonstration of how to calculate the p-value given the observed test statistic value, t_{obs} . The p-value is the integral of the right-hand tail of the test statistic distribution. (right) The relationship between the p-value and the significance Z is shown using a standard normal distribution.	92
5.3	Distributions of the test statistic q under the s + b and b hypotheses in an example where sensitivity to the signal model is very small (left). Schematic diagram showing the test statistic distribution under b and s + b hypotheses (right). The observed test statistic value is indicated by the dotted line. The s+b p-value is shown by the red filled area, whereas the area filled with skyblue shows $(1-p_b)$	96
6.1	Feynman diagram for the Z'_{TC2} signal process, in which one of the top quarks decays hadronically and the other leptonically.	100
6.2	Angular separation ΔR of the decay products in the two-body decays (left-hand) $t \rightarrow W b$, and (right-hand) $W \rightarrow q\bar{q}$ as a function of the transverse momentum of the top quark and the W boson, respectively [30].	101
6.3	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	101

6.4	Acceptance times efficiency of the Z'_{TC2} signals.	109
6.5	Acceptance times efficiency of the Z'_{TC2} signals.	110
6.6	Acceptance times efficiency of the Z'_{TC2} signal as a function of the true $t\bar{t}$ mass for the combined boosted and resolved topology selection (red), the boosted (green) and resolved (magenta) selections.	110
6.7	Real efficiency in the muon channel for $\Delta R_{\min}(\mu, jet) > 0.4$ (left-hand) and for $\Delta R_{\min}(\mu, jet) < 0.4$ (right-hand).	113
6.8	Real efficiency in the electron channel for $\Delta R_{\min}(e, jet) > 0.4$ (left-hand) and for $\Delta R_{\min}(e, jet) < 0.4$ (right-hand).	113
6.9	Fake rate in the muon channel for $\Delta R_{\min}(\mu, jet) > 0.4$ (left-hand) and for $\Delta R_{\min}(\mu, jet) < 0.4$ (right-hand).	114
6.10	Fake rate in the electron channel for $\Delta R_{\min}(e, jet) > 0.4$ (left-hand) and for $\Delta R_{\min}(e, jet) < 0.4$ (right-hand).	115
6.11	Fake rates f (right-hand) and real rates ϵ (left-hand) for the loose μ to satisfy the tight selection, as a function of its p_T and min $\Delta R(\mu, jets)$, measured in $CR_{\text{QCD}}^{4j,1b}$	117
6.12	Fake rates f (right-hand) and real rates ϵ (left-hand) for loose e to satisfy the tight selection, as a function of its p_T and min $\Delta R(e, jets)$, measured in $CR_{\text{QCD}}^{4j,1b}$	118
6.13	Distribution of ΔR (left-hand), ΔR_{corr} (right-hand) in the $CR_{\text{QCD}}^{4j,1b}$, (top) $VR1_{\text{QCD}}^{4j,1b}$ (middle) and $VR2_{\text{QCD}}^{4j,1b}$ (bottom) for the muon channel. In addition to the statistical uncertainty, the gray band represents a flat systematic error band of 50% linked to the QCD multi-jet.	119
6.14	Distribution of ΔR (left-hand), ΔR_{corr} (right-hand) in the $CR_{\text{QCD}}^{4j,1b}$, (top) $VR1_{\text{QCD}}^{4j,1b}$ (middle) and $VR2_{\text{QCD}}^{4j,1b}$ (bottom) for the electron channel. In addition to the statistical uncertainty, the gray band represents a flat systematic error band of 50% linked to the QCD multi-jet.	120
6.15	E_T^{miss} (left-hand), $m_{t\bar{t}}$ (right-hand) for the muon channel. In addition to the statistical uncertainty, the gray band represents a flat systematic error band of 50% linked to the QCD multi-jet.	121
6.16	E_T^{miss} (left-hand), $m_{t\bar{t}}$ (right-hand) for the electron channel. In addition to the statistical uncertainty, the gray band represents a flat systematic error band of 50% linked to the QCD multi-jet.	122
6.17	Pre(post)-fit comparison between data and background in the 0bCRs.	124

6.18	Generated Z'_{TC2} and invariant mass of the top quarks decay products at the partonic level for the 1.0 TeV (left-hand) and 2.5 TeV (right-hand) resonances.	125
6.19	Efficiencies of matching the different partons from top-quark decays as function of the pole mass of the Z'_{TC2} samples, after the full selection of the resolved (left-hand) and boosted (right-hand) selections, before the selection on the b-tagging category.	128
6.20	Efficiency to select the correct combination form reconstructible eventss for the resolved (left-hand) and boosted (right-hand) selections as function of the pole mass of the Z'_{TC2} samples, for the categories 1 (top), 2 (middle) and 3 (bottom).	129
6.21	Reconstructed $t\bar{t}$ invariant mass distributions for the Z'_{TC2} signal model considered in the search. The events satisfy the criteria of the resolved-topology selection before (empty) or after (filled) the veto on events that also pass the boosted-topology criteria. All distributions are normalised to unit area.	130
6.22	Resonstracted $t\bar{t}$ versus true $t\bar{t}$ invariant mass for boosted electron (top-right), resolved electron (top-left), boosted muon (bottom-right) and resolved muon (bottom-left) for SM $t\bar{t}$ production. Both reconstructible and non-reconstructible events are included.	131
6.23	The p_T of the lepton (top), (middle) the E_T^{miss} , and (bottom) Mass of the hadronic top distributions in the resolved (a, c, e) electron and (b, d, f) muon channel. The ratio of data and background is mostly within the systematic band.	137
6.24	The mass of the leptonic top (top), (middle) mass of the hadronic W boson, and (bottom) jets multiplicity distributions in the resolved (a, c, e) electron and (b, d, f) muon channel. The ratio of data and background is mostly within the systematic band.	138
6.25	The p_T of the lepton (top), (middle) the E_T^{miss} , and (bottom) Large-R jet p_T distributions in the boosted (a, c, e) electron and (b, d, f) boosted muon channel. The ratio of data and background is mostly within the systematic band.	139
6.26	The p_T of the selected jet (top), (middle) mass of the leptonic top, and (bottom) jets multiplicity distributions in the boosted (a, c, e) electron and (b, d, f) muon channel. The ratio of data and background is mostly within the systematic band.	140

6.27	Background composition in the 16 SRs. The pie chart shows relative contributions per process in the different regions.	141
6.28	The $m_{t\bar{t}}^{reco}$ distributions for the boosted electron (left-hand) and the boosted muon (right-hand) after the profiling of the nuisance parameters.	144
6.29	The $m_{t\bar{t}}^{reco}$ distributions for the resolved electron (left-hand) and the resolved muon (right-hand) after the profiling of the nuisance parameters.	145
6.30	95% CL Exclusion Limits on the Z_{TC2} signal cross-section times branching ratio, as a function of $m_{Z_{TC2}'}.$ All signal regions (electron and muon channels, for both the resolved and boosted selection) are combined. All statistical and systematic uncertainties are included. The theoretical cross-section is overlaid as a red line, and used to extract the mass exclusion limit.	147

Contents

Dedication	i
Acknowledgement	ii
Résumé	iii
Abstract	iv
Résumé détaillé	vi
List of abbreviations	x
List of Tables	xi
List of Figures	xiii
Introduction	1
1 Theoretical framework	4
1.1 The Standard Model of particle physics	4
1.2 The gauge sector: Local $SU_C(3) \otimes SU_L(2) \otimes U_Y(1)$ gauge invariance	6
1.2.1 Gauge Principle: Quantum electrodynamics (QED)	7
1.2.2 Electroweak interactions	8
1.2.3 Strong interactions	10
1.2.4 Spontaneous Symmetry Breaking: Higgs mechanism	11
1.3 Top Quark	16
1.4 Top quark pairs Production	18
1.5 Top quark decay	19

1.6	Beyond the standard model	23
1.6.1	Spin-1 color singlet Z'_{TC2}	24
1.6.2	Spin-2 color singlet G_{KK}	24
1.6.3	Spin-1 colour octet g_{KK}	25
2	The Large Hadron Collider and the ATLAS Detector	26
2.1	The Large Hadron Collider	26
2.2	Luminosity	26
2.3	The LHC acceleration complex	29
2.3.1	LHC Program	29
2.3.2	The Worldwide LHC Computing Grid (WLCG)	30
2.4	The ATLAS detector	31
2.4.1	ATLAS coordinate system	32
2.4.2	The magnetic systems	33
2.4.3	The ATLAS Inner Detector (ID)	34
2.4.3.1	The Pixel detector	34
2.4.3.2	The Semiconductor Tracker (SCT)	35
2.4.3.3	The Transition Radiation Tracker (TRT)	36
2.4.4	The ATLAS Calorimeters	36
2.4.4.1	Electromagnetic Calorimeter	37
2.4.4.2	Hadronic Calorimeter	38
2.4.5	The muon spectrometer (MS)	39
2.4.6	Trigger and Data Acquisition	40
2.4.6.1	Level-1 Trigger	41
2.4.6.2	Level-2 Trigger	42
2.4.6.3	Event Filter	42
3	Physics Objects	44
3.1	Electrons	44
3.1.1	Electron identification and reconstruction	44
3.1.2	Electron isolation	45
3.2	Muon	49
3.2.1	Muon identification and reconstruction	49
3.2.2	Muon isolation	50
3.3	Jets	51

3.3.1	Jet reconstruction	51
3.3.2	Jet calibration	52
3.3.3	Jet pile-up rejection	53
3.3.4	Jet flavour tagging	54
3.3.5	Jet trimming	54
3.3.6	Top-tagging	55
3.4	Missing Transverse Energy	56
4	Overlap Removal and performance of E_T reconstruction	57
4.1	Missing transverse momentum measurement in ATLAS	57
4.2	E_T basics	58
4.2.1	E_T terms definition	59
4.2.2	E_T^{miss} reconstruction	59
4.2.3	Electron and photon	60
4.2.4	Muon	61
4.2.5	τ – <i>lepton</i>	61
4.2.6	Soft term	61
4.3	Jet selection	62
4.3.1	EMTopo Jet	62
4.3.2	PFlow Jet	63
4.4	Monte Carlo Samples	63
4.5	Overlap Removal	64
4.6	Jets and $e(\gamma)$ overlap removal	65
4.7	Systematic uncertainties	78
4.7.1	Methodology	78
4.7.1.1	Observables	79
4.7.1.2	Procedures	80
4.7.1.3	Systematic uncertainties in E_T^{miss} response and resolution	81
4.8	Performance of the missing transverse energy E_T^{miss}	83
5	Statistical Frameworks	88
5.1	Evaluation of Agreement between Observed Data and Expectation	89
5.2	Trials factor	92
5.3	Parameter Estimation	93
5.4	Confidence Interval Definition	93

5.5	The Likelihood ratio test-statistic	95
5.6	The CL_s value and exclusion limits	95
6	Search for $t\bar{t}$ resonances in ATLAS	98
6.1	Analysis Strategy	99
6.2	Object definitions	102
6.2.1	Leptons	102
6.2.2	Jets	103
6.2.3	The missing transverse momentum	103
6.2.4	Overlap Removal	103
6.3	Data and Monte Carlo samples	104
6.4	Event Selection	106
6.4.1	Preselection	106
6.4.2	Boosted Selection	107
6.4.3	Resolved Selection	108
6.4.4	b-tag categories	108
6.4.5	Signal acceptance times efficiency	108
6.5	Data driven backgrounds	111
6.5.1	Multi-jet background modelling	111
6.5.1.1	The Matrix Method	111
6.5.1.2	Estimation of fake and real rates	112
6.5.1.3	ΔR correction factor to the real and fake rates	114
6.5.1.4	Validation of QCD multi-jet modeling	115
6.5.2	W+jets background normalisation	123
6.6	Event Reconstruction	125
6.6.1	Neutrino Reconstruction	125
6.6.2	Boosted channel topology	126
6.6.3	Resolved channel topology	126
6.6.4	Performance of the reconstruction	127
6.6.5	$t\bar{t}$ invariant mass spectrum	129
6.7	Systematic Uncertainties	131
6.7.1	Experimental uncertainties	131
6.7.2	Theory uncertainties	133
6.8	Comparison Expectation to Data	135

6.9 Compatibility of data with the SM-only hypothesis	142
6.10 Upper production cross section limits on $t\bar{t}$ resonances	146
7 Summary and outlook	148
Bibliography	149

Introduction

The Standard Model of particle physics (SM), developed in the 1960's and finalized in the 1970's, is a very successful theory whose predictions have been verified many times, notably at large-scale particle accelerator experiments. The observation of a new scalar particle by the ATLAS and CMS collaborations at the Large Hadron Collider (LHC) at CERN in 2012 marked yet another milestone in the history of particle physics. The origin of mass of the elementary particles is one of the most fundamental questions in particle physics. It is answered by the electroweak symmetry breaking (EWSB), for which the W and Z bosons become massive and from which a new particle, the Higgs boson, arises and it gives mass to the fermions through its Yukawa interaction with them. The properties of this particle have been found to be consistent with those of the long sought-after Higgs boson in subsequent and increasingly precise measurements. In spite of the SM's considerable success, it is believed that there should be a more universal theory underlying the SM which sheds light to unanswered questions such as the existence of Dark Energy and Dark Matter, the matter-antimatter asymmetry of the Universe or the neutrino mass.

The top quark, discovered in 1995 at the Tevatron, is the heaviest known elementary particle, and its unusual features need special study. The top quark's significance stems from its enormous mass, which correlates to an extremely short lifetime, causing it to disintegrate quickly through EW interactions before it has opportunity to hadronize. As a result, it allows researchers to investigate its characteristics using decay products. Another consequence of its enormous mass is that it is the only quark with a Yukawa coupling to the Higgs boson on the order of one; hence, a thorough understanding of its characteristics can provide light on basic interactions at the EWSB and beyond. Various searches for novel physics have also focused on the top quark. It would have a strong coupling to new heavy particles predicted by different extensions of SM theories known as beyond SM (BSM). Evidence of such a particle would be a $t\bar{t}$ resonance that may therefore be sought through direct searches for new particles in high-energy particle collisions -a local deficit or excess in the invariant mass

spectrum of $t\bar{t}$ pairs produced in proton-proton collisions with respect to the SM expectation around the mass of the new particle.

In this thesis, a direct search for a $t\bar{t}$ resonance using $t\bar{t}$ pairs produced at proton-proton collisions at a center-of-mass energy of 13 TeV recorded by ATLAS detector at the LHC during Run 2 collected over 2015-2018, corresponding to 139 fb^{-1} of integrated luminosity is presented. The search presented in this thesis relies on two different reconstruction strategies to optimise its sensitivity over a large mass regime between 0.4 and 5.0 TeV, resolved topology, where the decay of the low-energetic top quarks is well separated in the detector, the other targeted boosted topology, where the highly energetic top quarks decay is reconstructed as a single object. In particular, an improved reconstruction strategy for semi-leptonically decaying boosted top quarks has been developed specifically for this search. In addition, a Deep Neural Network (DNN) based top-tagging method was applied to the large radius jets to identify the hadronic tops. Whereas the tops are reconstructed using a chi-squared algorithm (χ^2) in the resolved region using only the small radius jets. A deep-learning-based b-tagging method is used here to identify the b-jets coming from top decay b-tagging helps to suppress non- $t\bar{t}$ backgrounds further. Special emphasis is paid to the estimation of the QCD multi-jet background using a data-driven techniques as well as the W+jets background normalization. No evidence for resonant is found and, as a consequence, the findings of the search are explained in terms of a heavy boson, Z'_{TC2} , mass limit using several benchmark BSM models. A similar search was carried out with partial Run-2 ATLAS data (36.1 fb^{-1}), and no trace of a new signal was found [32].

The second part of this manuscript is devoted to the determination of the electron-jet and photon-jet overlap removal studies and the achievement of the optimal missing transvers energy performance. The missing transverse momentum (E_T^{miss}) is a crucial variable serving as an experimental representative of the transverse momentum carried by particles not interact with the detector. The measurement of E_T^{miss} is a real challenge because it involves different signals of all detected particles. Hence, the energy deposits in the detector should be precisely associated to the jets or particles that produced them. When hard objects are created near to each other, they might share or overlap a quantity of energy deposits which need to be correctly treated. A dedicated study for electron-jet and photon-jet overlap removals have been developed as well as the study of the E_T^{miss} performance.

The manuscript is structured as follows: Chapter 1 gives more detailed insights into the theory describing interactions among elementary particles and the expectations related to the observation of the new particle searched. Chapter 2 describes the experimental setup,

the ATLAS detector at the LHC used to carry out the search for a new particle. Chapter 3 details the algorithms employed for reconstructing and identifying the different experimental signatures based on the information recorded by the detector. Chapter 4 details the electron (photon)-jet overlap removal techniques and discusses the E_T^{miss} performance study. The statistical techniques employed to compare the SM $t\bar{t}$ mass spectrum to the data $t\bar{t}$ mass spectrum and therefore impose constraints on the cross-sections of new particles are highlighted in Chapter 5. A detailed account of the search strategy, including the definition of the signal regions, the estimation of background processes, systematic uncertainties, results and their statistical interpretation is presented in Chapter 6 and a final conclusion is presented in Chapter 7.

Chapter 1

Theoretical framework

The Standard Model (SM) of fundamental particles is a relativistic quantum field theory (QFT) [33] first developed in 1960's by Glashow, Weinberg and Salam to explain the unified electroweak interaction. The SM is the foremost successful theoretical framework, approved by numerous experiments to describe our best perception of the fundamental bricks of matter, and their interactions, with the exception of gravitational interactions.

This chapter presents a summary of the SM's essential concepts related to the analysis displayed in this thesis. The elementary particles categorization within the SM is briefly described. Additionally, the gauge formalism of the SM theory is presented with the mechanism of Brout-Englert-Higgs [34] through which the SM's core building blocks gain their masses.

1.1 The Standard Model of particle physics

According to the SM, the ordinary matter (as it were 4% of our Universe) is made up of structureless particles named fermions by Paul Dirac, and classified by several intrinsic characteristics such as: quantum numbers and mass. There are twelve fermions plus their conjugates (anti-particles with the same mass, but opposite electromagnetic charge, color and opposite component of the weak isospin) characterised with half-integer spin, and governed by Fermi-Dirac statistics. Fermions interact through weak, electromagnetic and strong forces by the exchange of intermediate gauge bosons of spin-one obey Bose-Einstein statistics. Including gravity to the SM as a fourth force represents one of the relevant challenge waiting for answer. Some details about these fundamental interactions are summarized in Table 1.1.

The fermions [35] are categorized into two categories (quarks and leptons) based on the way they interact or by the charges they hold (Figure 1.1). Firstly, the six leptons are classified in three generations of charged leptons together with their corresponding neutral

Interaction	Gauge boson	Mass (GeV)	Charge	coupling strength	range [fm]
Electromagnetic	γ	0	Electrical	$\alpha_{em} = \frac{e^2}{4\pi\hbar c} \approx \frac{1}{137}$	∞
weak	$W^\pm$	80.4	Weak isospin	$Gm^2 = 1.01 \cdot 10^{-5}$	$< 2 \cdot 10^{-3}$
	Z	91.2			
Strong	8 gluon	0	colours r, g, b, $\bar{r}$, $\bar{g}$, $\bar{b}$	$\alpha_S = \frac{g^2}{4\pi\hbar c} \approx \frac{1}{137}$	< 1
—	Higgs	125	—	—	—

Table 1.1: Fundamental forces and its corresponding mediators with charge types described by the Standard Model [31].

leptons called neutrinos which are handled as massless despite the perception of the neutrino oscillations¹ [36]: (electron (e), electron neutrino ν_e), (muon (μ), muon neutrino ν_μ) and (tau (τ), tau neutrino ν_τ). Higher generations individuals have bigger mass than the earlier generations. The μ , e and τ carry electrical charge equal to -1² ($q = -1$), so they interact electromagnetically and weakly whereas the associated neutrinos don't ($q=0$), therefore, their motion is affected by the weak force. Each lepton has a corresponding anti-particle with the same mass and sign-reversed quantum numbers (charges and parity).

Furthermore, fundamental fermions incorporate six quarks known as: down (d), up (u), charm (c), strange (s), top (t) and bottom (b), classified in three generations which separates them in several doublets of weak isospin: (u,d), (c,s) and (b,t). Quarks have a charge for color, subsequently, they interact through the strong interaction. Beside this, quarks also hold weak isospin and electric charge, so they interact with the rest of fermions both electromagnetically and through the weak interaction. There exists an anti-quark to each quark which carries anti-colour charge. In nature, Quarks are all time found confined by shaping colourless hadrons (do not exist as free particles) [37]. There are two sorts of hadrons: mesons formed by one anti-quark and one quark while the baryons are formed by three quarks. Visible matter is made exclusively by the up and down quarks forming the neutrons and protons (hadrons). The vast bulk of the Universe's stable matter is made up of fermions corresponding to the first-generation, while the last two generations are primarily relevant to high energy phenomena including astrophysical processes involving cosmic radiation and collider physics. Because of the mass hierarchy, particles of second and third generations have bigger mass with short time of live and generally decay to the first generation via the

¹Neutrino flavour oscillations, which indicate evidence of non-zero neutrino masses, are predicted by the SM.

²We use natural units where $\hbar = c = 1$. Multiples of the fundamental electric charge e are used to express the electric charges.

weak force.

The bosons: The SM fundamental forces are mediated by gauge bosons (integer spin). The electromagnetic force intermediate is the photon. The weak interaction is mediated by three massive bosons: Z^0 and $W^\pm$. The strong force is carried by massless spin-1 particles called gluons (g) which are eight in number. Figure 1.1 provides a global view of the basic SM particles properties.

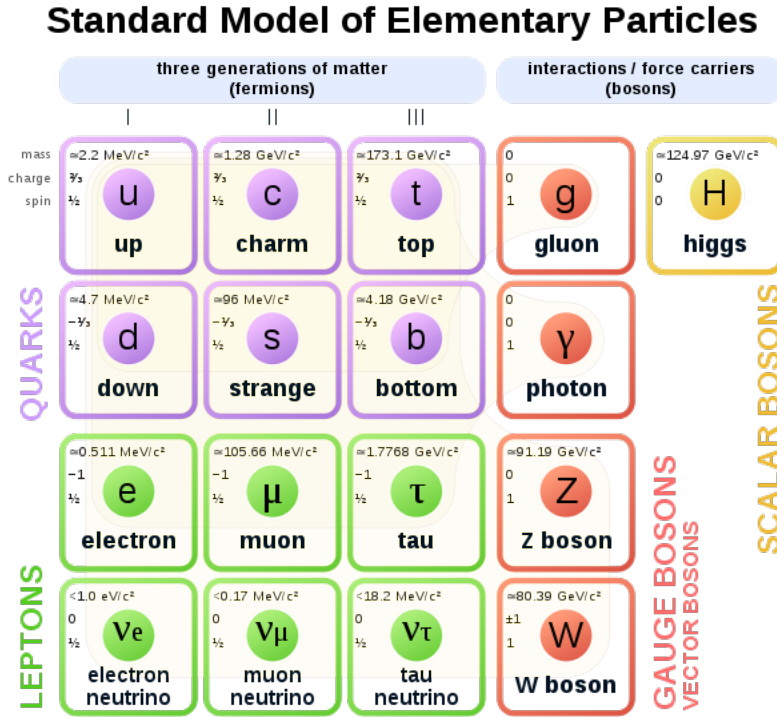


Figure 1.1: The SM fundamental particles. Information from [8].

1.2 The gauge sector: Local $SU_C(3) \otimes SU_L(2) \otimes U_Y(1)$ gauge invariance

The fundamental forces in the SM theoretical framework are built on the gauge symmetry group $SU_C(3) \otimes SU_L(2) \otimes U_Y(1)$. The quantum chromodynamics theory (QCD) [38] is the one that describes the strong interaction, based on $SU_C(3)$ colour gauge symmetry (3×3 unitary matrices with unit determinant and the color charge is denoted by C) while the gauge group $SU_L(2) \otimes U_Y(1)$ represents the electro-weak interaction, Y represents the weak hypercharge, while L stands for the left-handed fields.

1.2.1 Gauge Principle: Quantum electrodynamics (QED)

The SM is based on quantum field theory which underlies the local gauge symmetry defined as the transformation of the field in every point with a phase factor. A unitary transformation of the field $\phi(x)$ with a phase factor $e^{i\theta(x)}$ is given by:

$$\phi(x) \rightarrow \phi'(x) = e^{i\theta(x)} \cdot \phi(x) \quad (1.1)$$

Where the new field $\phi'(x)$ describes the same physical processes. The expression of a group of fields transformed is:

$$\phi'_i(x) = U_i^k \phi_k(x) = (e^{i\frac{g}{2}\theta^\alpha(x)T_\alpha})_i^k \phi_k(x) \quad (1.2)$$

$U_i^k = (e^{i\frac{g}{2}\theta^\alpha(x)T_\alpha})_i^k$ is a unitary transformation, g the coupling constant, T_α are the generators of group $U(N)$, and $\theta(x)$ correspond to the transformation angles. In the case if the determinant of U is equal to 1, this group is noted $SU(N)$ and called special unitary group with N^2-1 generators T_α . For $N=2$, the generators of $SU(2)$ are the Pauli-matrices:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (1.3)$$

The Dirac equation for free particles corresponds to a Lagrangian with an expression:

$$\mathcal{L} = \bar{\Phi}(i\gamma^\mu \partial_\mu - m)\Phi \quad (1.4)$$

m represents the particle mass and γ^μ refers to the Dirac matrices. The transformation cited in Equation 1.2, breaks the symmetry of $\mathcal{L}$. To restore the invariance assuming a local transformation, the derivation has to be modified to covariantly transform under local transformation:

$$\partial_\mu \rightarrow D_\mu = \partial_\mu - ieA_\mu \quad (1.5)$$

A_μ represents a vector field obey transformation like:

$$A_\mu \rightarrow A'_\mu = A_\mu + \frac{1}{e}\partial_\mu\theta_\alpha \quad (1.6)$$

Furthermore it can be shown, that for this gauge field of massless free particles, the local gauge invariance is restored. This guides to a physical explanation of the vector field A_μ for

$U(1)_{em}$: field of the photon coupling to charged particles. Together with the field strength tensor:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad (1.7)$$

The Lagrangian of QED may be expressed as follows:

$$\mathcal{L} = \bar{\Phi}(i\gamma^\mu \partial_\mu - m)\Phi + e\bar{\Phi}A_\mu\Phi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \quad (1.8)$$

Where $\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$ correspond to the Kinetic energy of the gauge particle.

1.2.2 Electroweak interactions

In the SM, the electromagnetic and weak interactions are both represented by the $SU_L(2) \otimes U_Y(1)$ gauge symmetry group [39]. The $SU_L(2)$ is a left-handed group coupling to the left-handed fermions f_L which is defined in the weak isospin charge space, T (its generators are $T_i = \frac{\sigma_i}{2}$ where $i=1,2,3$ and σ_i are the Pauli matrices). $U_Y(1)$ is a weak hypercharge group defined in the hypercharge Y space linked to the weak isospin's third component (T_3) and the electromagnetic charge and as:

$$Y = 2(Q - T_3) \quad (1.9)$$

Both leptons and quarks are divided in three groups composed by right-handed and left-handed components with some of the quantum numbers are listed in Table 1.2:

	generation I	generation II	generation III	Q[e]	T_3	Y
leptons	$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L$	$\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L$	$\begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L$	0	+1/2	-1
	e_R	μ_R	τ_R	-1	-1/2	-1
				-1	0	-2
Quarks	$\begin{pmatrix} u \\ d' \end{pmatrix}_L$	$\begin{pmatrix} c \\ s' \end{pmatrix}_L$	$\begin{pmatrix} t \\ b' \end{pmatrix}_L$	+2/3	+1/2	+1/3
	u_R	c_R	t_R	-1/3	-1/2	+1/3
	d_R	s_R	b_R	+2/3	0	+4/3
				-1/3	0	-2/3

Table 1.2: The SM fermion quantum numbers (weak hypercharge (Y), electric charge (Q) and its third component (T_3)) and generations.

Local gauge invariance is provided with introduction of triplet vector field W_μ^i ($i = 1, 2, 3$) associated to the $SU_L(2)$ components and single vector field B_μ for $U(1)_Y$. For the quark and lepton, the electroweak Lagrangian density is defined as follows:

$$\mathcal{L}_{EW} = \bar{f}\gamma^\mu i D_\mu f - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}W_{\mu\nu}^i W_i^{\mu\nu} + \mathcal{L}_{GF} + \mathcal{L}_{FP} \quad (1.10)$$

Where f correspond to the fermion field. To be gauge-invariant, in the Lagrangian kinetic term, the covariant derivative D_μ must be used in place of the ordinary derivative ∂_μ , where:

$$D_\mu = \partial_\mu - i\frac{g}{2}\sigma_i W_\mu^i - i\frac{g'}{2}Y B_\mu \quad (1.11)$$

The field tensors also, $W_{\mu\nu}$ and $B_{\mu\nu}$, provide the gauge boson kinetic self-interaction terms:

$$W_{\mu\nu}^i = \partial_\mu W_\nu^i - \partial_\nu W_\mu^i - g\epsilon_{ijk}W_\mu^j W_\nu^k \quad (1.12)$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \quad (1.13)$$

The physical fields representing the real gauge bosons of the electromagnetic and weak interactions are obtained:

- The charged bosons $W^{\mu\pm}$ of the weak force are identified as a superposition of W_1^μ and W_2^μ vector fields:

$$W^{\mu\pm} = \frac{1}{\sqrt{2}}(W_1^\mu \pm iW_2^\mu) \rightarrow W^\pm \quad (1.14)$$

- The neutral boson Z^0 of the weak force:

$$Z^\mu = -B^\mu \sin\theta_W + W_3^\mu \cos\theta_W \rightarrow Z^0 \quad (1.15)$$

- The photon field for the QED:

$$A_\mu = B^\mu \cos\theta_W + W_3^\mu \sin\theta_W \rightarrow \gamma \quad (1.16)$$

The coupling constants are linked to θ_W called the mixing angle, generally referred to as the Weinberg angle:

$$\sin\theta_W = \frac{g'}{\sqrt{g^2 + g'^2}} \quad (1.17)$$

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

They are also connected in terms of electric charge, as follows:

$$g' \cos\theta_W = g \sin\theta_W = e. \quad (1.19)$$

Finally, the gauge fixing Faddeed-Popov terms required to quantize the gauge theory are presented by the $\mathcal{L}_{GF}$ and $\mathcal{L}_{FP}$ terms (details is outside the scope of this thesis).

1.2.3 Strong interactions

Quantum chromodynamics (QCD) [38] is the name for the theory of the strong force at short range. It explains the interaction between gluons and quarks leading to their confinement to compose hadronic states. The associated spinor fields to the quark ψ_i (i=1,2,3- three colour indices) transform under $SU(3)_C$ colour-symmetry group (non abelian). $SU(3)_C$ has eight ($3^2 - 1$) generators, so eight gauge fields (massless gluon) required to provide local gauge invariance. The full Lagrangian density of QCD can be formulated as:

$$\mathcal{L}_{QCD} = \bar{\psi} \gamma^\mu D_\mu \psi - \frac{1}{4} G_{\mu\nu}^\alpha G_\alpha^{\mu\nu} \quad (1.20)$$

With the covariant derivative:

$$D_\mu = \partial_\mu - i \frac{\lambda^a}{2} g_s G_\mu^a \quad (1.21)$$

The gluon field's strength tensor is:

$$G_{\mu\nu}^\alpha = \partial_\mu G_\nu^\alpha - \partial_\nu G_\mu^\alpha - g_s f_{ijk} G_\mu^b G_\nu^c \quad (1.22)$$

While λ^α ($\alpha = 1, \dots, 8$) represent the generators of the basic representation of $SU(3)_C$, G_μ^α are the gauge boson fields related to the gluons. In Equation 1.22, the third term is coming from the fact that the $SU(3)_C$ matrices not commuting where f^{abc} are the structure constants³.

The strong force Lagrangian density may be divided into three major parts:

$$\mathcal{L}_{QCD} = \mathcal{L}_{quark} + \mathcal{L}_{int} + \mathcal{L}_{gluon} \quad (1.23)$$

³The matrices generators for $SU_C(3)$ λ^α , are with trace=0 and meet: $[\lambda^\alpha, \lambda^b] = 2i f^{abc} \lambda_c$ with f^{abc} the $SU_C(3)$ structure constants which are real and totally antisymmetric

$\mathcal{L}_{quark}$ describes the motion of free quarks with a given flavour and colour. The spin-1 gluon field interacting with these quarks, like the photon in QED, restores local gauge invariance, involving the gluon-quark interaction term:

$$\mathcal{L}_{int} = g_s \bar{q}_i \frac{(\lambda^a)_k^i}{2} \gamma^\mu q^k G_\mu^\alpha \quad (1.24)$$

$\mathcal{L}_{gluon}$ represents the gluon fields propagation. The gauge bosons are massless, according to local gauge invariance. The gluon Lagrangian is more troublesome and results in triple and four gluon vertices (Figure 1.2), in contrast to its QED equivalent $B_{\mu\nu}$. Despite the fact that gluons have no mass, the strong interaction has a limited range.

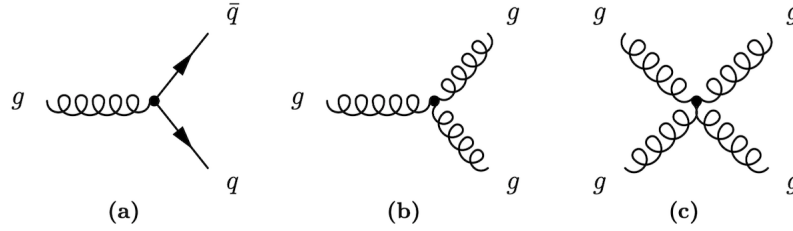


Figure 1.2: Interaction vertices of QCD.

While the strong coupling strength is too strong to use perturbation theory over long distances (effective theories or phenomenological models are needed). The asymptotically free quarks and gluons are described by perturbation theory for the vanishing coupling strength on short scales.

1.2.4 Spontaneous Symmetry Breaking: Higgs mechanism

Although the predictive power of the Standard Model, there are several experimental evidence that point out a weakness of this theoretical framework: only massless fields are considered by the QCD and electroweak lagrangians while measurements reveal for instance that Z and $W^\pm$ have masses and only the gluons and the photon (g and γ) are actually massless. Furthermore, SM fails on explaining neutrino-oscillations which have already been observed, and does not offer a dark matter candidate required to describe observed phenomena in astrophysics.

Only massless fields are present in the electroweak and QCD Lagrangians. This is true for both bosons and fermions. Adding mass terms like: $\frac{M^2}{2} W_\mu W^\mu$ should destroy local gauge invariance. The Brout-Englert-Higgs Mechanism is a method for dealing with this feature in the SM. A new term must be introduced into the Lagrangian in order to preserve the

symmetry of the SM Lagrangian with a non-symmetric vacuum. A degenerate collection of minimum energy states that convert nontrivially will be present in the new Lagrangian. The symmetry is considered as being spontaneously broken (SSB) [34] because a specific one degenerate state was arbitrarily chosen as the ground state.

In a practical manner, the solution to the mass problem is to include to the model a complex scalar field Φ that retains a non-zero value in the ground state and thereby breaks electroweak symmetry, also a certain potential V is introduced.

$$\Phi(x) = \begin{pmatrix} \Phi^+ \\ \Phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \Phi_1 + i\Phi_2 \\ \Phi_3 + i\Phi_4 \end{pmatrix} \quad (1.25)$$

$$V(\Phi) = \mu^2 \Phi^+ \Phi + \lambda (\Phi^+ \Phi)^2, \lambda > 0, \quad (1.26)$$

The Lagrangian density is unchanged under $SU(2)$, where $\Phi(x) \rightarrow \Phi'(x) = e^{i\theta(x)} \Phi(x)$.

$$\mathcal{L}_{SSB} = (D_\mu \Phi)^+ (D^\mu \Phi) - V(\Phi) \quad (1.27)$$

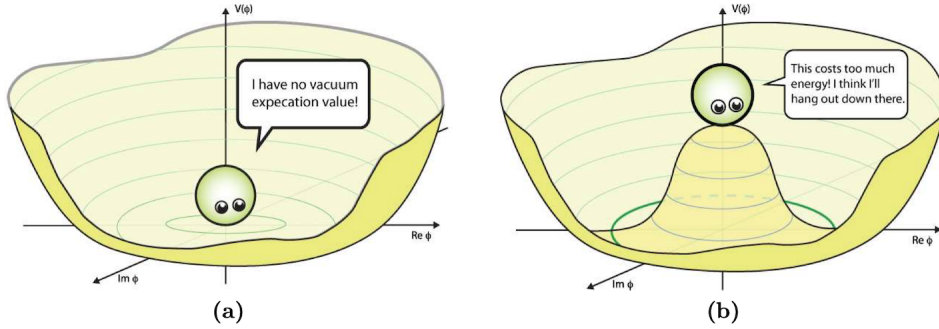


Figure 1.3: The ground state of the Higgs-Potential for $\lambda > 0$, $\mu^2 > 0$ (a) and $\mu^2 < 0$ (b). Pictures taken from Reference [9].

The potential V minimized for $|\Phi| = 0$ if $\mu^2 > 0$ (trivial minimum), if $\mu^2 < 0$, the ground state of V is currently degenerate (Figure 1.3) and defined as follows:

$$\Phi^+ \Phi = \frac{1}{2} (\Phi_1^2 + \Phi_2^2 + \Phi_3^2 + \Phi_4^2) = -\frac{\mu^2}{2\lambda} \equiv \frac{v^2}{2}, V(\Phi_0) = -\lambda v^4. \quad (1.28)$$

Therefore, according to the Brout-Englert-Higgs Mechanism, choosing a particular vacuum, such as $(\Phi_1 = \Phi_2 = \Phi_4 = 0 \text{ and } \Phi_3 = v)$, the symmetry is said spontaneously broken for

minimum of the potential V at vacuum expectation value (VEV) of Φ :

$$\langle \Phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \quad (1.29)$$

Utilizing the unitary gauge, the form of the Lagrangian is moreover studied by including the vacuum expectation value (Equation 1.29) to it which reduce the covariant derivative term of $SU_L(2) \otimes U_Y(1)$ to:

$$\begin{aligned} D_\mu \langle \Phi \rangle &= (\partial_\mu - i \frac{g}{2} \sigma_i W_\mu^i - i \frac{g'}{2} B_\mu) \langle \Phi \rangle \\ &= \begin{pmatrix} -i \frac{g}{2\sqrt{2}} (W_\mu^1 - i W_\mu^2) v \\ -\frac{i}{2\sqrt{2}} (g' B_\mu - g W_\mu^3) v \end{pmatrix} \end{aligned} \quad (1.30)$$

$(D_\mu \langle \Phi \rangle)^\dagger (D^\mu \langle \Phi \rangle)$ brings out several terms quadratic in the fields:

$$\begin{aligned} (D_\mu \langle \Phi \rangle)^\dagger (D^\mu \langle \Phi \rangle) &= \frac{1}{8} v^2 g^2 ((W_\mu^1)^2 + (W_\mu^2)^2) + \frac{1}{8} v^2 (g' B_\mu - g W_\mu^3)^2 \\ &= (\frac{1}{2} v g)^2 W_\mu^+ W^{-,\mu} + \frac{1}{4} v^2 (g^2 + g'^2) Z_\mu Z^\mu \end{aligned} \quad (1.31)$$

The field Φ is expanded by performing perturbation about Its VEV, where $H(x)$ is the scalar Higgs boson field.

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \quad (1.32)$$

As consequence, the acquired masses by means of the Higgs mechanism for the electroweak and Higgs bosons are all directly proportional to v , where as expected, the photons does not acquire mass:

$$M_\gamma = 0 \quad (1.33)$$

$W^\pm$ acquires a mass term :

$$M_{W^\pm} = \frac{1}{2} g v \quad (1.34)$$

The Z boson mass can be written as :

$$M_Z = \frac{1}{2}v\sqrt{g^2 + g'^2} \quad (1.35)$$

The Higgs Boson Mass:

$$M_H = \sqrt{2\lambda}.v \quad (1.36)$$

The vacuum expectation value can be estimated using the $\frac{G_F}{\sqrt{2}} = \frac{g^2}{8M_W^2}$ relation obtained from the effective V-A theory:

$$v = \frac{1}{\sqrt{G_F\sqrt{2}}} \simeq (246\text{GeV}) \quad (1.37)$$

The Fermi constant is denoted by G_F . The mass of the Z and $W^\pm$ boson might be calculated at first order using the the estimation of the Weinberg angles, and the evaluated value for v , the acknowledged value of the fine structure constant. We measure the W boson mass, M_W , using data corresponding to 8.8 fb^{-1} of integrated luminosity collected in proton-antiproton collisions at a 1.96 TeV center-of-mass energy with the CDF II detector at the Fermilab Tevatron collider.

$$M_W = \sqrt{\left(\frac{\pi\alpha}{G_F\sqrt{2}}\right)} \frac{1}{\sin\theta_W} \simeq 80.385 \pm 0.016 \text{ GeV}. \quad (1.38)$$

The mass of the Z boson, M_Z has been determined by combining the data from the four LEP experiments ALEPH, DELPHI, L3 and OPAL.

$$M_Z = \frac{M_W}{\sin\theta_W} \simeq 91.18 \pm 0.007 \text{ GeV}. \quad (1.39)$$

By including Yukawa couplings to the Higgs doublet, it is possible to produce Fermion masses. The Yukawa Lagrangian is a gauge invariant fermion-scalar coupling:

$$\mathcal{L}_{YK} = \lambda_e \bar{l}_L \Phi e_R + \lambda_\mu \bar{q}_L \bar{\Phi} u_R + \lambda_d \bar{q}_L \Phi d_R + h.c \quad (1.40)$$

where:

$$l_L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}; q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix}; \bar{\Phi} = i\sigma_2 \Phi^* = \begin{pmatrix} \Phi_0^* \\ -\Phi^- \end{pmatrix} \quad (1.41)$$

Following the same previous procedure, expanding the Lagrangian around the vacuum and rotating the weak eigenstates to the mass eigenstates, lead to a new set of mass terms:

$$\mathcal{L}_{YK} = -(\lambda_e \frac{v}{\sqrt{2}}) \bar{e}'_L e'_R - (\lambda_\mu \frac{v}{\sqrt{2}}) \bar{\mu}'_L \mu'_R - (\lambda_d \frac{v}{\sqrt{2}}) \bar{d}'_L d'_R + \dots \quad (1.42)$$

Associated to the massive fermions,

$$m_e = \lambda_e \frac{v}{\sqrt{2}}; m_u = \lambda_u \frac{v}{\sqrt{2}}; m_d = \lambda_d \frac{v}{\sqrt{2}}; \dots \quad (1.43)$$

λ_f being the Yukawa coupling constant of each fermion f with the Higgs boson field. We can rewrite $\mathcal{L}_{SSB}$ and $\mathcal{L}_{YK}$ in function of physical scalar fields, they may be viewed as the Higgs field's free and interaction terms.

$$\mathcal{L}_{SSB} + \mathcal{L}_{YK} \rightarrow \mathcal{L}_H^{free} + \mathcal{L}_H^{int} + \dots \quad (1.44)$$

Where the Higgs free Lagrangian can be written as:

$$\mathcal{L}_H^{free} = \frac{1}{2} \partial_\mu H \partial^\mu H - \frac{1}{2} M_H^2 H^2 \quad (1.45)$$

And the interaction Lagrangian related to the gauge coupling and to the particle masses (including the Higgs mass [40]) is written as:

$$\begin{aligned} \mathcal{L}_H^{int} = & -\frac{gM_H^2}{4M_W} H^3 - \frac{g^2 M_H^2}{32M_W^2} H^4 - \frac{gm_f}{2M_W} \bar{f} H f + M_W^2 W_\mu^+ W^{\mu-} (1 + \frac{g}{M_W} H + \frac{g^2}{4M_W^2} H^2) \\ & + \frac{M_Z^2 Z_\mu Z^\mu}{2} (1 + \frac{g}{M_Z \cos \theta_W} H + \frac{g^2}{4M_Z^2 \cos^2 \theta_W} H^2) \end{aligned} \quad (1.46)$$

Equation 1.46 shows the Higgs self-couplings: the couplings to the weak interaction gauge bosons also to fermions.

The Brout-Englert-Higgs mechanism gives a clear explanation on how particles gain masses via their interaction with the Higgs field, but the exact value of their masses is undetermined by the SM predictions. On 4 July 2012, in the LHC, a new state compatible with the Higgs boson properties have been discovered [41] with a mass of $M_H = 125.03$ GeV [40]. Considering this mass value, the Higgs boson is taken as the second heaviest elementary particle so far discovered, after the top quark, which the measured mass is $m_t \approx$

173 GeV [31].

1.3 Top Quark

After the discovery of the bottom quark (b) in 1977, the presence of other quark carrying a charge of $+2/3$ was predicted to close the third Isospin doublet analogous to the first two quark generations. It took about 23 years to observe the heaviest known elementary particle partner of the b quark at the Fermilab Tevatron by CDF [42] and D0 experiment [43]. The top-quark has the greatest interaction with the Higgs boson. To better comprehend the electroweak sector in the Standard Model (SM) or beyond Standard Model (BSM) ideas, top-quark physics experiments are essential. Due to its large mass ($m_t = 173.34 \pm 0.27$ (stat) ± 0.71 (syst) GeV), according to [44], the top quark has a huge expected decay width $\Gamma = 1.32$ GeV, which results in a very brief mean life-time $\approx 10^{-25}$ s. The CMS and ATLAS direct m_t measurements obtained at the LHC are compared to LHC and Tevatron+LHC m_t combinations in Figure 1.4.

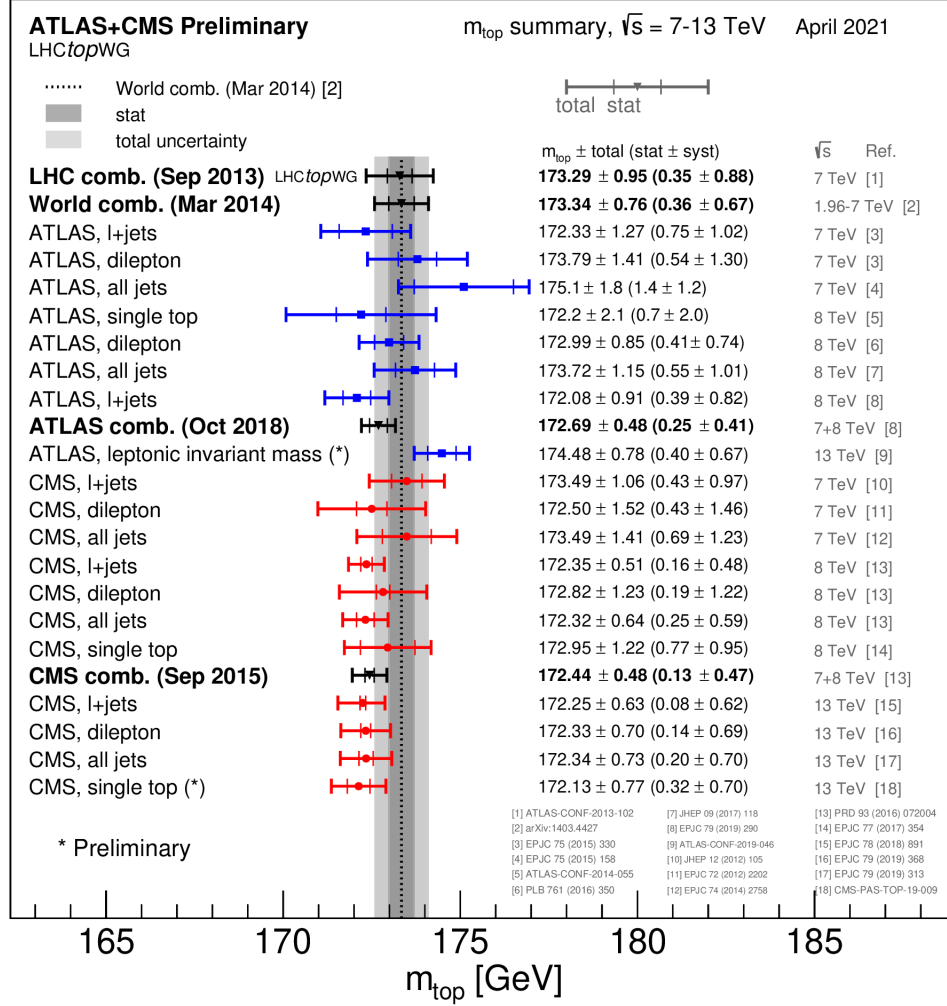


Figure 1.4: m_t measurements done with the CMS and ATLAS experiments at LHC, compared with LHC and Tevatron+LHC results. Systematic and statistics uncertainties are shown [10].

1.4 Top quark pairs Production

The top quark pairs production at LHC is mainly occurs through strong interactions via the gluon-gluon fusion. Figure 1.5 presents the leading order (LO) Feynman diagrams for production of top quark pairs: $q\bar{q}$ annihilation and gluon-gluon fusion. The gluon gluon fusion is the dominating process which happens in 87% of the cases at the LHC because of the high $\sqrt{s}$ value also due to the fact that both initial-state beams are made of protons (Up until high x, gluons predominate in the proton's parton distribution function). As displayed in Figure 1.6, the predicted theoretical cross-section, estimated up to the next-to-next-to-leading-order (NNLO), for top-quark pair production at the LHC at $\sqrt{s} = 13$ TeV is $\sigma_{t\bar{t}} = 832^{+46}_{-46}$ pb [10].

In addition, as demonstrated in the Feynman diagrams in Figures 1.7,1.8,1.9, it is feasible to manufacture single top quarks by a variety of processes. Several electroweak processes can produce a single top quark, however electroweak cross section creation is less common than strong cross section generation. Table 1.3 describes the cross sections of the single top in pp collisions at $\sqrt{s} = 13$ TeV .

Process	Total cross section pb	Total cross section pb (%)	AntiTop cross section pb (%)
t channel	$216.99^{+9.04}_{-7.71}$	136.02 (62.68%)	80.95 (37.31%)
s channel	$10.32^{+0.40}_{-0.36}$	6.35 (61.53%)	3.97 (38.47%)
Wt channel	71.7 ± 3.85	35.85 (50.00%)	35.85 (50.00%)

Table 1.3: Production cross sections of the single top in pp collisions at $\sqrt{s} = 13$ TeV [10].

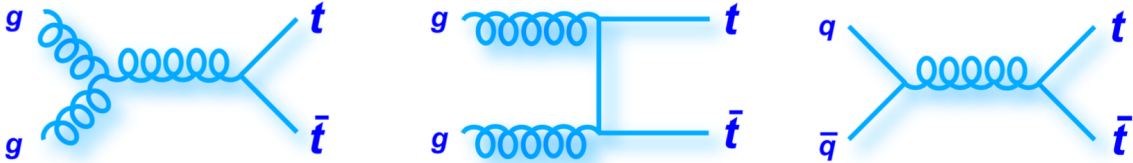


Figure 1.5: Feynman diagrams at the LO for the top quark pairs production. Bottom: $q\bar{q}$ annihilation process. Top: gluon-gluon fusion processes [11].

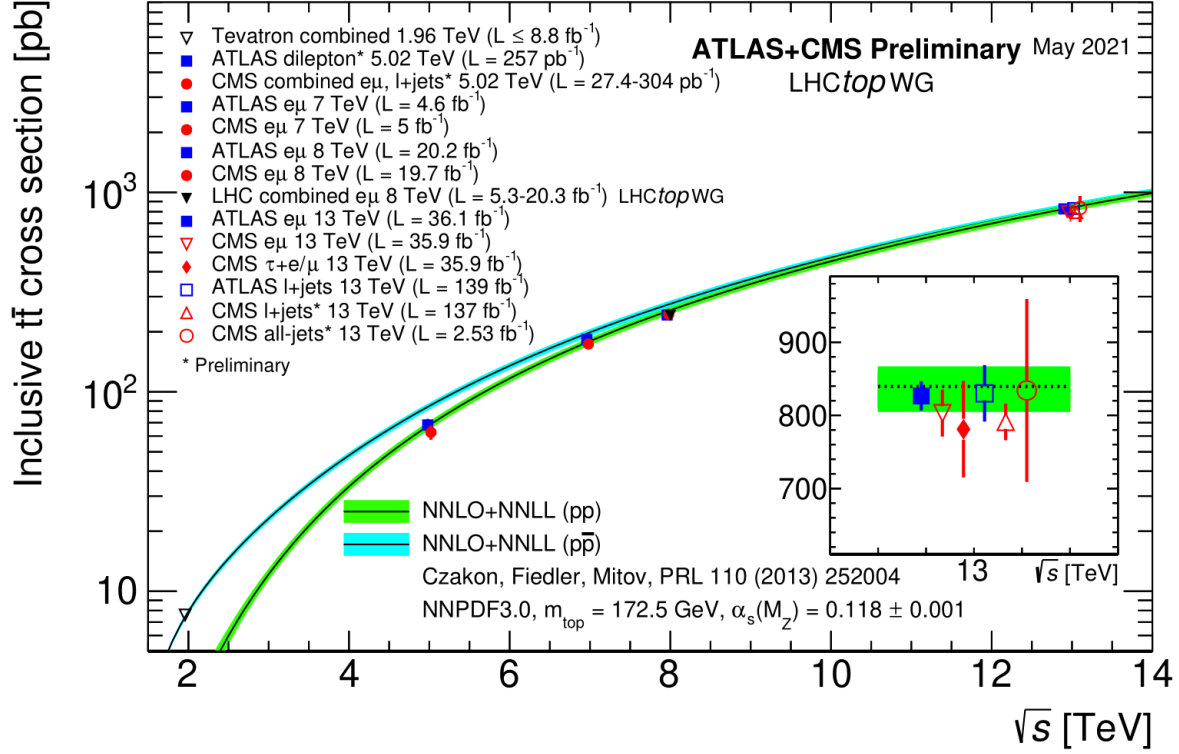


Figure 1.6: $t\bar{t}$ production cross-section at LHC and Tevatron, in term of the centre of mass energy relative to the NNLO QCD calculation. These calculations assuming $m_t = 172.5$ GeV [10].

1.5 Top quark decay

Due to its enormous mass, the top quark decays with probability close to one through electroweak force in the channels, $t \rightarrow Wb$, $t \rightarrow Wc$ and $t \rightarrow Ws$. The top quark is extremely short-lived $\tau_s \simeq 0.5 \times 10^{-24}s$ which is even small compared to the hadronization time scale $\simeq 10^{-23}s$ [8] and shorter compared to the spin decorrelation time $\simeq 10^{-21}s$, then, detectors have the unique opportunity to only measure top quark's decay. The top quark mainly decays into a W boson and a bottom quark ($t \rightarrow Wb$) due to the fact that the Cabbibo-Kobayashi-Maskawa (CKM) matrix element $|V_{tb}|$ is close to one. The top quark decay amplitude $\Gamma(t \rightarrow Wb)$ may be estimated via the formula in Equation 1.47.

$$\Gamma(t \rightarrow Wb) = \frac{G_F m_t^3}{8\pi\sqrt{2}} |V_{tb}|^2 \approx 1.74 GeV \quad (1.47)$$

Figure 1.10 displays the Feynman diagram for the $t \rightarrow Wb$ process.

The leading pair-production process's ultimate states may be split into three groups. The decay characteristics of the top quark or anti-top quark are separated into di-leptonic, hadronic, and semi-leptonic decay modes, which correspond to the final states of the W boson (Figures 1.11). Figure 1.12 summarizes the respective decay probabilities.

- All hadronic: both of the W bosons decay into quarks, $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow q\bar{q}'bq\bar{q}'\bar{b}$, (45.7%)
- Semi leptonic: one of the W bosons decays into quarks, and the other into $\ell + \nu_\ell$, $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow q\bar{q}'bl^-\bar{\nu}\bar{b} + l^+\nu bq\bar{q}'\bar{b}$, (43.8%)
- Di-leptonic: both of the W bosons decay into $\ell + \nu_\ell$, $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow l^+\nu bl^-\bar{\nu}\bar{b}$, (10.5%)

Assuming universality for lepton interactions, the quarks in the final state develop into jets of hadrons and $\ell = \mu, e, \tau$ leptons in all of the scenarios described.

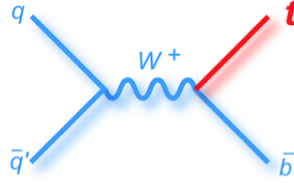


Figure 1.7: Feynman diagram at LO for the single top production, s-channel [11].

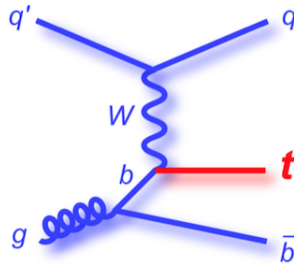


Figure 1.8: Feynman diagram at LO for the single top production, t-channel [11].

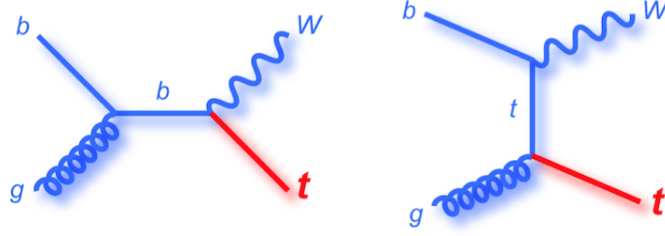


Figure 1.9: Feynman diagram at LO for the single top production, Wt-channel [11].

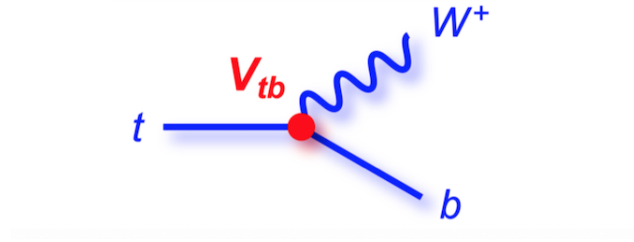


Figure 1.10: $t \rightarrow Wb$ Feynman diagram, illustrating the coupling related to $|V_{tb}|$ [11].

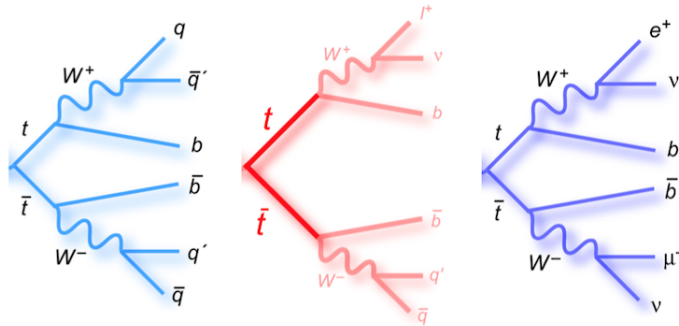


Figure 1.11: From right to left: di-leptonic, semi-leptonic and all hadronic top pair decay Feynman diagram [11].

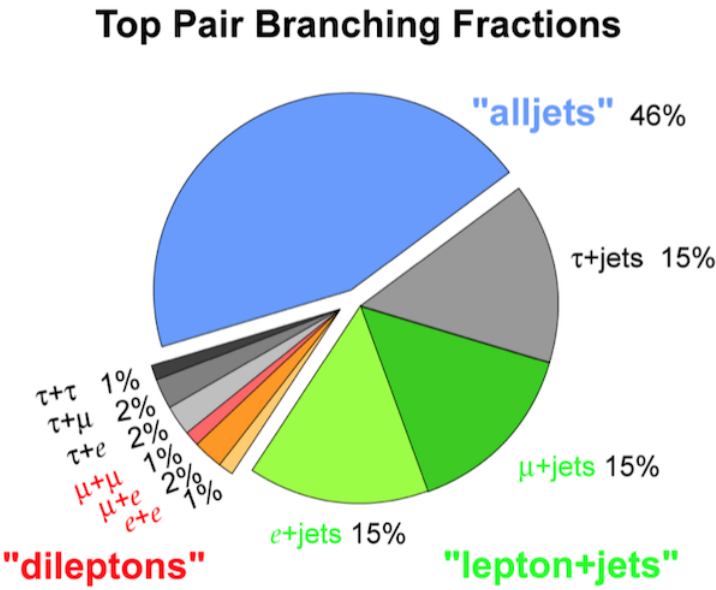


Figure 1.12: Top quark pairs branching ratios [11].

1.6 Beyond the standard model

Despite its incredible success and compatibility with majority of experimental results as displayed in Figure 1.13, the SM is regarded as merely an effective theory. Several open questions stay without answer regarding for instance it does not include gravity and does not explain why there are three generations of quarks and leptons. The top quark is crucial for examining physics BSM since it is the heaviest particle in the SM and has a very large Yukawa-coupling.

This section presents the identification of heavy particles that produce resonances in the top-antitop $t\bar{t}$ mass spectrum. Particular set of benchmark models are discussed : a spin-1 gauge boson Z' (1.6.1), a spin-1 Kaluza-Klein gluon (1.6.3) and a spin-2 Kaluza-Klein graviton (1.6.2).

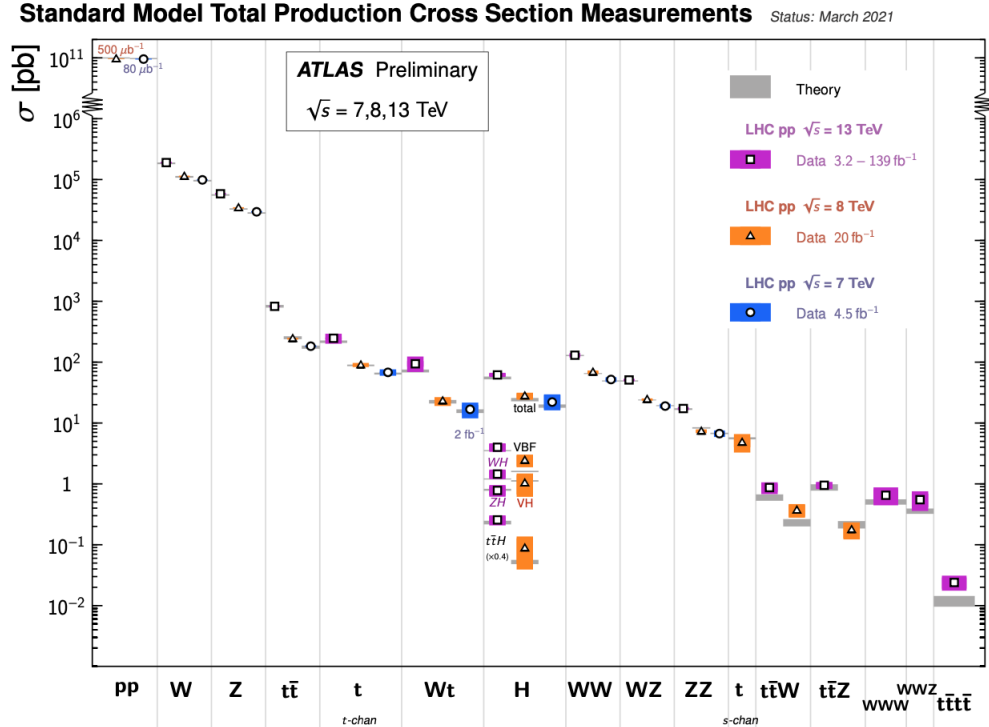


Figure 1.13: Summary of several Standard Model total production cross section measurements, corrected for branching fractions, compared to the corresponding theoretical expectations [12].

1.6.1 Spin-1 color singlet Z'_{TC2}

Numerous BSM modules predict that the LHC will produce additional neutral heavy particle, denoted Z' vector boson. One of searches uses topcolour-assisted technicolour Z'_{TC2} [45] as a benchmark. This is a leptophobic boson, couple preferentially to the first- and third-generation quarks. The primary production mode for Z'_{TC2} is $q\bar{q}$ annihilation as illustrated in Figure 1.14(a). The properties of Z'_{TC2} are managed by the topcolour filtering parameter, $\cot\theta_H$, controlling the width and the production cross-section, also f_1 and f_2 maximizing the fraction of Z'_{TC2} bosons that decay into $t\bar{t}$ if they take the values: $f_1 = 1$ and $f_2 = 0$. The parameters of this model are scanned in order to achieve a resonance with a width of its mass of 1.2%. At a 95% confidence level, previous searches by the CMS [46] and ATLAS [47] collaborations put lower bounds of $m(Z'_{TC2}) > 1.8$ TeV and $m(Z'_{TC2}) > 2.5$ TeV, respectively, on the permissible mass for these bosons.

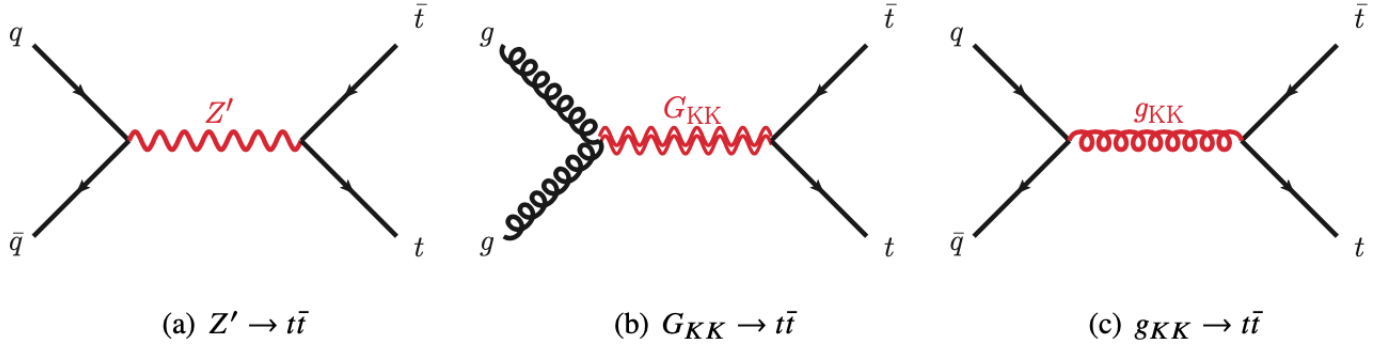


Figure 1.14: Feynman diagrams at Leading-order for the studied signal processes in this search. The Z' (a) and Kaluza–Klein gluons (g_{KK}) have spin 1 (b), while the Kaluza–Klein graviton (G_{KK}) has spin 2 (c).

1.6.2 Spin-2 color singlet G_{KK}

In theories that predict extra dimensions of space, spin-2 color-singlet bosons are created, resulting in graviton Kaluza–Klein excitations. One of these models is the Randall–Sundrum (RS) framework with extra dimensions where the SM fields are in the warped bulk and the fermions are localised properly to explain the flavour structure of the SM [48]. This kind of graviton (G_{KK}) is commonly referred to as a ‘Bulk’ RS graviton and is characterised by a dimensionless coupling constant $k/\bar{M}_{Pl} \approx 1$, where k is the curvature of the warped extra dimension and $\bar{M}_{Pl} = M_{Pl}\sqrt{8\pi}$ is the reduced Planck mass. As shown in Figure 1.14(b),

the principal mechanism of synthesis is gluon-gluon fusion. For masses between 400 and 600 GeV, the branching ratio of G_{KK} decaying into a top and antitop quark pair grows from 18% to 50%, plateauing at 68% for masses larger than 1 TeV.

1.6.3 Spin-1 colour octet g_{KK}

Theories that predict extra dimensions of space, spin-1 color-octet bosons are produced, leading to Kaluza–Klein gluon excitations. Randall–Sundrum (RS) benchmark considers heavy Kaluza–Klein gluons, g_{KK} , production with a single warped extra dimension in space [49], where widths varying from 10% to 40% of the g_{KK} mass. The primary production mode is $q\bar{q}$ annihilation as shown in Figure 1.14(c). Previous searches using $\sqrt{s} = 7$ TeV ATLAS data exclude a g_{kk} with a mass less than 2.2 TeV [47]. Using slightly different benchmark models the CMS Collaboration sets exclusion limits for g_{KK} masses of less than 3.3 TeV at 95% CL [46].

Chapter 2

The Large Hadron Collider and the ATLAS Detector

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [50],[51] is the highest and the largest energy particle collider ever built to date in the world. Placed on the border of France and Switzerland near Geneva, it first started the operations in 2008. It's built in 27-kilometre circumference of superconducting magnets long tunnel with a centre of mass collision energies of up to 14 TeV and $10^{34}cm^{-2}s^{-1}$ nominal luminosity. LHC collides ions and protons bundled in two beams, moving in opposite directions at unprecedented energy scale (several TeV). LHC includes four detectors corresponding to four experiments to study the produced collisions. A total of 1200 dipole niobium-titanium magnets are employed to maintaining the protons in a circular orbit, with extra 392 quadrupole, 168 octupole, and 688 sextupole magnets needed to properly focus the beam. The magnets perform at 1.9 K and can create a magnetic field of up to 8.3 T. The protons in the LHC circulate in bunches of around 1.15×10^{11} protons, with a maximum of 2808 bunches inside one beam at a time. In the laboratory frame, the bunch spacing is 25 ns.

2.2 Luminosity

The luminosity of a collider measures its capability. The number of events that the LHC creates per second is given as:

$$N_{event/s} = \mathcal{L}\sigma_{event} \tag{2.1}$$

Integrating the luminosity yields the total number of events for a studied process:

$$N_{event} = \sigma_{event} \int \mathcal{L} dt \quad (2.2)$$

Where $\mathcal{L}$ is called the instantaneous machine luminosity that measures the total number of events for a given process per cross section and time, the cross section of the studied event is represented by σ_{event} . The dimension of Luminosity is $[\mathcal{L}] = cm^{-2}s^{-1}$. For the Gaussian beam distribution, the machine luminosity definition depends on the collider parameters [52]:

$$\mathcal{L} = \frac{N_b^2 n_b f}{4\pi\sigma_x\sigma_y} \quad (2.3)$$

Where n_b the number of bunches per beam, N_b is the number of particles per bunch, f is the frequency of bunch crossings, finally, σ_x and σ_y ($19 \mu m$) are the Gaussian beam widths in the transverse plane compared to the bunch direction. In Equation 2.3, the beams are assumed to be round and have equal parameters. The Number of bunches in one beam might reach 2808 and the number of protons per bunch can exceed 1.1×10^{11} . To measure the quantity of data generated in a specific period of time, the luminosity integral over time $\int \mathcal{L}$ is used. It is measured in the inverse cross section units. The Run-II dataset collected by ATLAS and CMS from 2015 to 2018 was close to $140 fb^{-1}$ in value. The remarkable instantaneous luminosity of the LHC, which peaked during the 2017 and 2018 data at over $2.10^{34} cm^{-2}s^{-1}$, allowed for the vast size of the dataset. Figure 2.1 shows the total integrated luminosities collected by the ATLAS detector during 2015-2018. The large luminosity achieved by LHC indicates that each single bunch crossing causes several proton-proton interactions, characterised by distribution of mean number of interactions per bunch crossing $\langle \mu \rangle$, illustrated in Figure 2.2. This implicates more challenges in the reconstruction of the pp collisions coming from the in-time and out-of-time pile-up.

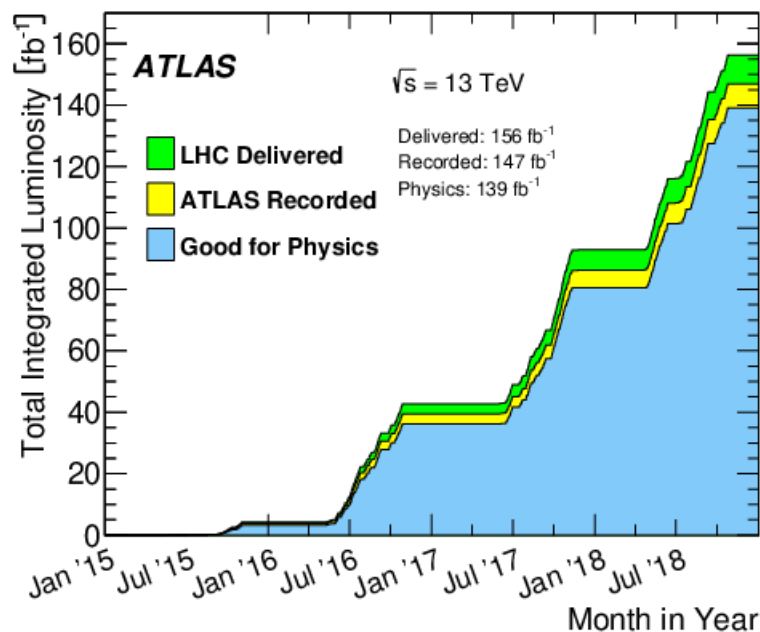


Figure 2.1: The total integrated luminosity that ATLAS received and captured between 2015 and 2018 while collecting stable beam pp collision data at $\sqrt{s} = 13 \text{ TeV}$ [13].

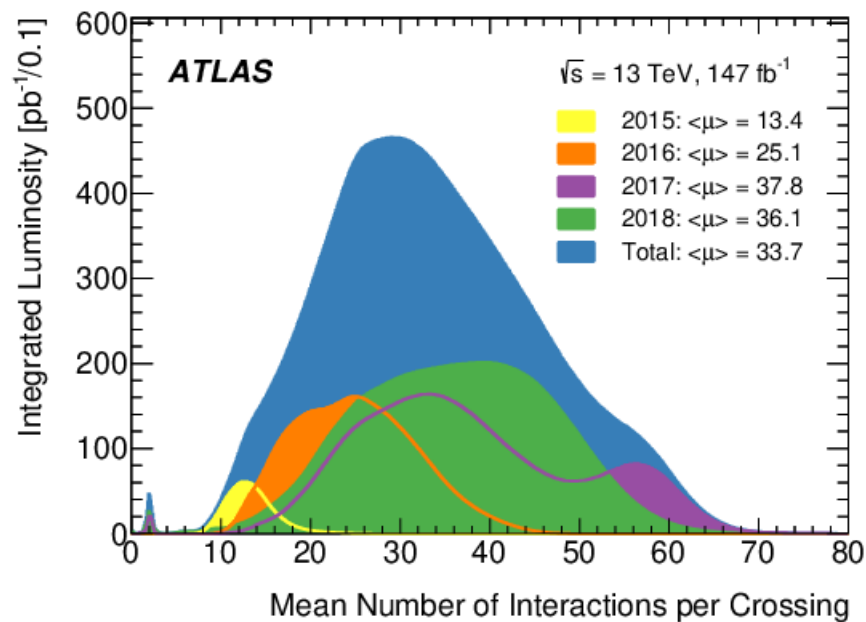


Figure 2.2: The distribution of the mean number of interactions per bunch crossing, μ , for the entire Run2 pp collision dataset at $\sqrt{s} = 13 \text{ TeV}$, is luminosity-weighted [13].

2.3 The LHC acceleration complex

The location of the series of linear detectors at the LHC and an overview of the CERN accelerator complex for proton beams are schematically shown in Figure 2.3. Proton beams are created by ionising Hydrogen atoms gas and accelerated in several steps with the smaller accelerating complex. Before getting into the LHC, first, proton beams are directed into the linear accelerator (LINAC 2) and accelerated to achieve an energy of 50 MeV. Then protons are injected into the Proton Synchrotron BOOSTER (PSB) circular accelerator where they are accelerated to 1.4 GeV. The acceleration of protons continues in a bigger circular accelerator: the Proton Synchrotron (PS), which allows protons to reach 26 GeV, and the next step sent protons to the Super Proton Synchrotron (SPS), which has a radius of 7 km and a beam energy of 450 GeV. After the SPS, protons are finally injected into two beam pipes of the LHC where they circulate in opposite directions with acceleration up to 6.5 TeV per beam containing $\approx 10^{11}$ particles each, with a minimum time spacing between two consecutive bunches of 25 ns. Protons are not the only particles accelerated in the LHC, the lead ion collisions start from a source of vaporized lead and enter Linac 3 before being collected and accelerated in the Low Energy Ion Ring (LEIR). Then, They follow the same path as protons to achieve their maximum energy. The two bunches in the LHC are injected into collision inside the four major detectors (ATLAS, CMS, LHCb and ALICE).

2.3.1 LHC Program

The LHC began operations in 2009 with collisions at a center of mass energy of $\sqrt{s} = 7$ TeV, followed by collisions at $\sqrt{s} = 8$ TeV in 2012. Run-1 refers to the data collected between 2010 and 2012. During a Long Shutdown (LS1) between 2013 and 2015, the LHC machine and detectors were fixed and updated. In 2015, operations for Run-2 were resumed at $\sqrt{s} = 13$ TeV, and the machine's generated luminosity was gradually increased until the end of 2018, when another shutdown, LS2, began for maintenance and improvements. A sketch of the LHC tentative plan is shown in Figure 2.4. A third LS will precede the High Luminosity configuration after Run-3, which is expected for 2021-2024. (HL-LHC). The machine will be upgraded significantly during LS3 to boost its brightness by at least a factor of five over its nominal design value. To ensure sufficient statistics for completing precise measurements, a large increase in brightness is required.

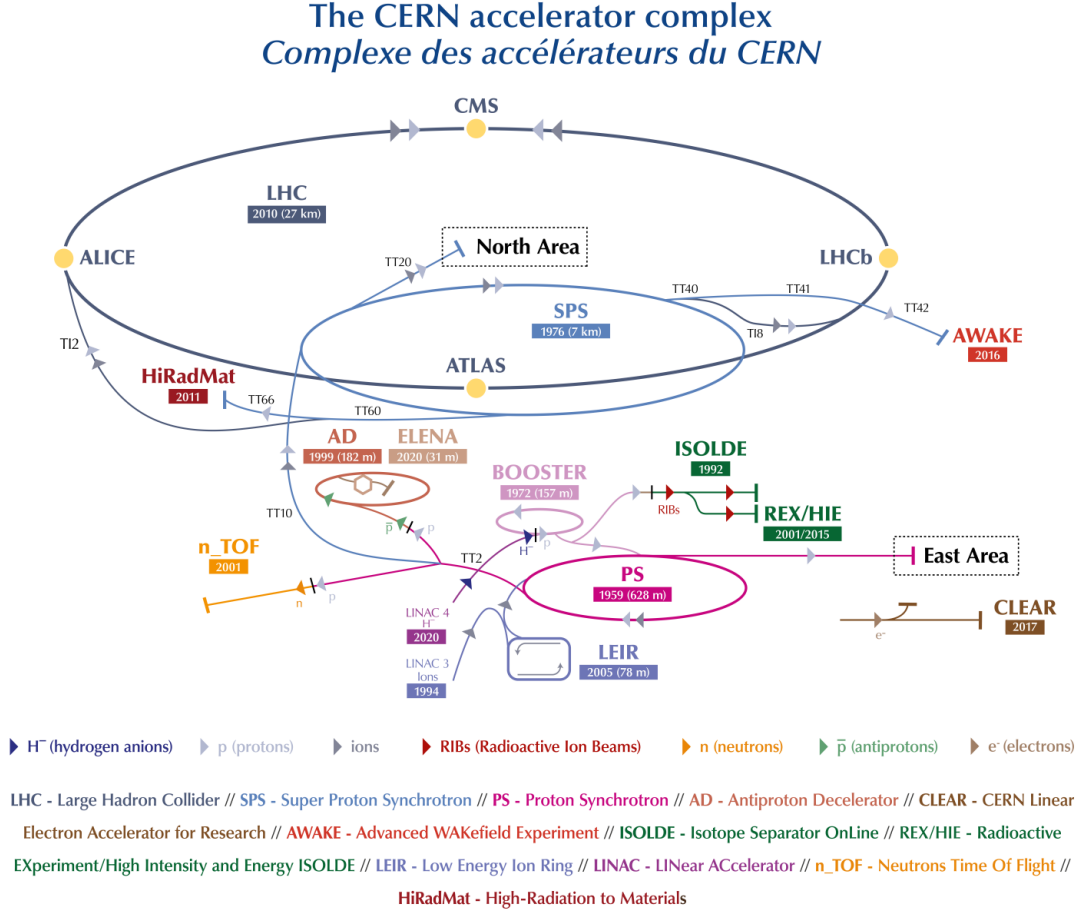


Figure 2.3: The CERN accelerator complex [14].

2.3.2 The Worldwide LHC Computing Grid (WLCG)

The Worldwide LHC Grid Computing Grid is an international partnership that brings together more than 170 computing kernels from more than 40 countries, bringing together national and international grid infrastructures. The computing pattern was developed with the purpose of developing and maintaining computer technology capable of handling the massive amounts of data generated at the LHC. Grid technologies are used in this computer approach. As a result, the WLCG provides global computing resources for saving, distributing, and interpreting the LHC's approximately 30 Pb (petabytes) of data. The WLCG must not only be able to handle massive amounts of data, but it must also be able to allow access to these data to a sizable user base (≈ 6000) all over the world, allowing them to do their analyses virtually and simultaneously. The LHCComputing Grid (LCG) in Europe,

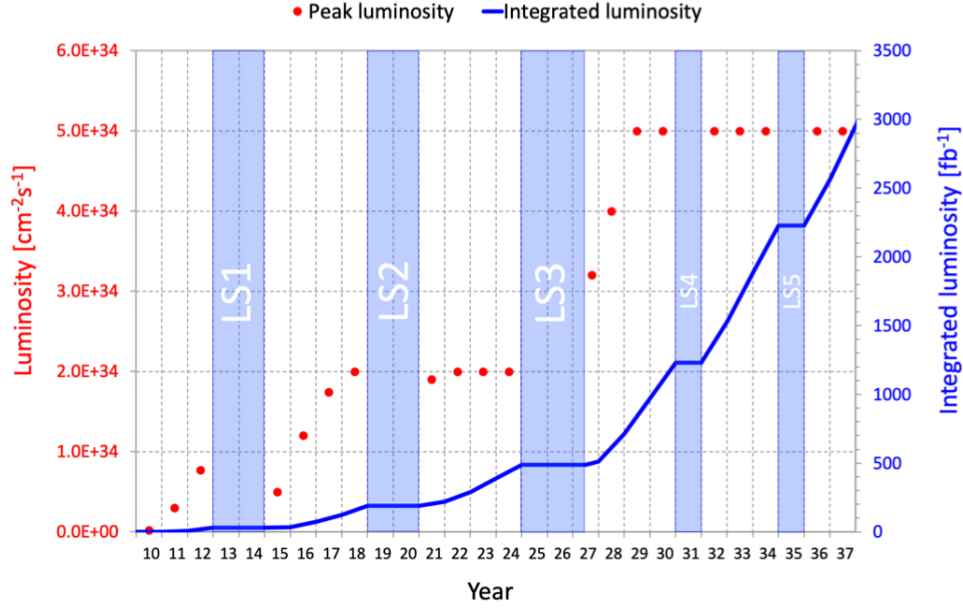


Figure 2.4: Projected LHC performance through 2037, showing preliminary dates for long shutdowns of the LHC and projected luminosities. Image source [15].

the NorduGrid/ARC in Europe but only for the Nordic nations, and the Open Science Grid (OSG) in the United States make up the WLCG.

Tiers are used to organize the LHC’s distributed computing model. The Tier0 center is the first and is located at CERN. The LHC data is collected and analyzed in this center as a primary backup. The data is then distributed to Tier 1 centers following this first phase. There are 11 Tier 1 locations in the world, each of which defines a cloud. Tier 1 centers convert data and send it to Tier 2 centers (140 over the world). Tier2 centers are made up of one or more computing facilities that can stock data and allow enough processing capacity for a given analysis. The analysts from the closest Tier3 computer resources, who employ the local resources of their search centers, gain access to these facilities. In this method, jobs submitted by the user to undertake any study could run where the data is kept, avoiding as much data transmission as possible.

2.4 The ATLAS detector

The ATLAS experiment [53] is a multi-purpose particle detector studying the signatures left by particles emerging from the pp collisions in its components. It is a forward-backward symmetric with cylindrical geometry (4π coverage) around the interaction point. It consists

in successive layers with diameter 25 m, length of 44 m, and a weight of 7000 tons, as displayed in Figure 2.5. The ATLAS detector was designed with two features in mind. The tracking system is ionized by charged particles, and both charged and neutral particles interact with calorimeters, resulting in particle cascades that produce more particles.

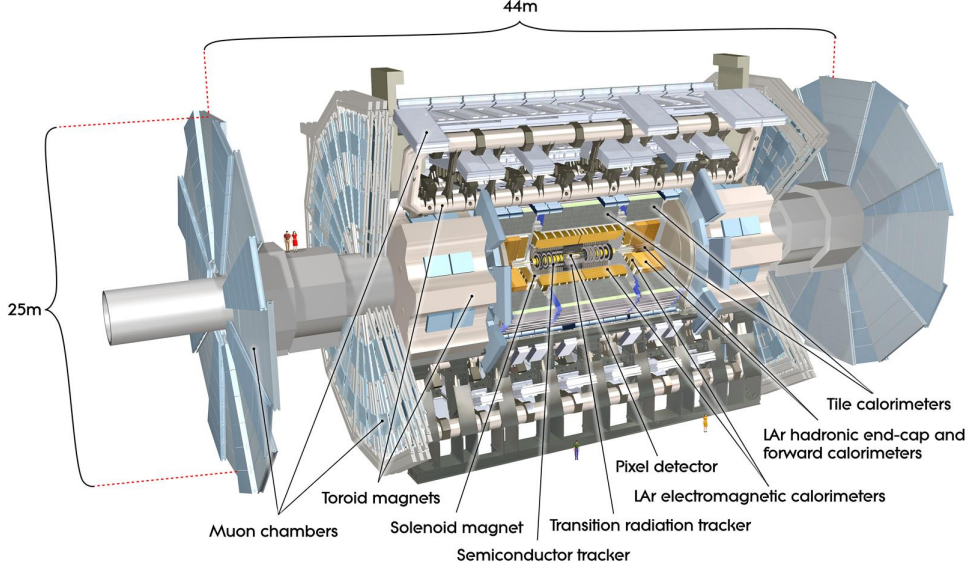


Figure 2.5: The ATLAS detector with labels on each of its sub-detectors. A few human beings are present in the image as a scale reference. Figure from the ATLAS public web pages [16].

The ID is described in sub-section 2.4.3. The ATLAS calorimeters (see sub-section 2.4.4) are made from the hadronic calorimeter and the electromagnetic calorimeter are mandatory for energy measurements of photons, electrons, and jets. The final layer in the detector is the muon spectrometer, needed to precisely reconstruct muon momenta, is described in sub-section 2.4.5

2.4.1 ATLAS coordinate system

The ATLAS detector coordinate system employs a right-handed Cartesian system with axes (X, Y, Z), as shown in Figure 2.6, such that the origin coincides with the nominal interaction point. The z -axis is taken to be parallel to the beam pipe. The x - y plane is defined orthogonally to the beam, where the positive x -axis points towards the centre of the LHC ring and the positive y -axis upwards to ensure that coordinate system is right-handed. Therefore, it is important to mention that momentum conservation is required

in the transverse plane. Given that the detector is cylindrical, it seems sense to convert to spherical coordinates (r, θ, ϕ) , where the azimuthal angle ϕ in the transverse plane with respect to the $x - axis$ direction covers the range of $-\pi$ to π around the beam axis. The polar angle θ indicates the orientation in relation to the $z - axis$. It is often useful to employ pseudo-rapidity η instead of θ for relativistic particles ($m \ll E \approx p$), due to $\Delta\eta$ being Lorentz-invariant under boosts along the $z - axis$. Pseudo-rapidity is defined as follows:

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right) \quad (2.4)$$

The longitudinal momentum component of colliding quarks is unknown after the collision, but it is known that total transverse momentum is equal to zero before the collision. The transverse momentum of a particle of four-momentum, $p_\mu = (E/c, \vec{p})$, is defined as :

$$p_T = p \sin(\theta) = \sqrt{(p_x)^2 + (p_y)^2}, p_z = p \cos(\theta) \quad (2.5)$$

Due to the conservation of transverse momentum, it is anticipated that the total transverse momenta of all particles created following a proton-proton collision in the transverse plane will equal zero. A particle's transverse energy is defined as:

$$E_T = E \sin(\theta) \quad (2.6)$$

Distances between particles or other reconstructed objects in the $\eta - \phi$ are parametrised via distance metric:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \quad (2.7)$$

which, when subjected to longitudinal Lorentz boosts, is likewise a quantity invariant.

2.4.2 The magnetic systems

The ATLAS magnet system (Figure 2.7) consists of a combination of one central superconducting solenoid and three outer toroids (one barrel and two end-caps). The central solenoid (CS) surrounds the ID is designed to produce an axial magnetic field of 2T parallel to the beam axis, minimising the radiative thickness and energy loss in order to improve the momentum and the energy measurements.

Two end-caps toroids are installed on the detector's end-caps area, creating a 1 T magnetic field to deflect muons in the $1.4 < |\eta| < 2.7$ region. The central barrel toroid provides

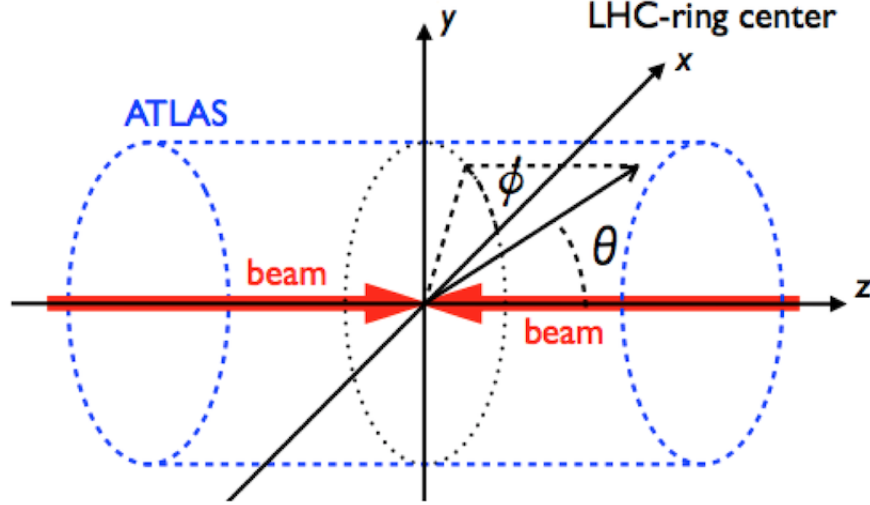


Figure 2.6: Diagram of the ATLAS coordinate system. The positive x -axis points towards the centre of the LHC, and the y -axis points upwards.

a magnetic field of around 0.5 T after the calorimeter system to deflect muons in the area $|\eta| < 1$ and so measure their momentum.

2.4.3 The ATLAS Inner Detector (ID)

The ATLAS inner detector [54] (ID) extended over a length of 6.2 m with a diameter of 2.1 m (Figures 2.8, 2.9), consists of three sub-detector located inside the solenoid magnets: the pixel detector, the Semiconductor Tracker (SCT), and the Transition Radiation Tracker (TRT). The entire ID is submerged in a 2 T magnetic field generated by a superconducting solenoid that deflects trajectories of charged particles (tracks) in the region near to the interaction point, within a range of $|\eta| < 2.5$ which allows the measurement of their electric charge and momentum with high precision.

2.4.3.1 The Pixel detector

The pixel detector, is the innermost subsystem of the ID [55], surrounding the beam pipe which highly suffers from a massive radiation coming from the roughly 1000 particles produced from the collision vertex every 25 ns. The pixel detector is made of oxygen-doped silicon and has three barrel layers with roughly 67 million squares pixels and two end-caps of three disks with 13 million pixels each, for a total of 92 million readout channels. The minimum pixel sensors size is $50 \times 400 \mu\text{m}^2$ with intrinsic spatial measurement accuracy of

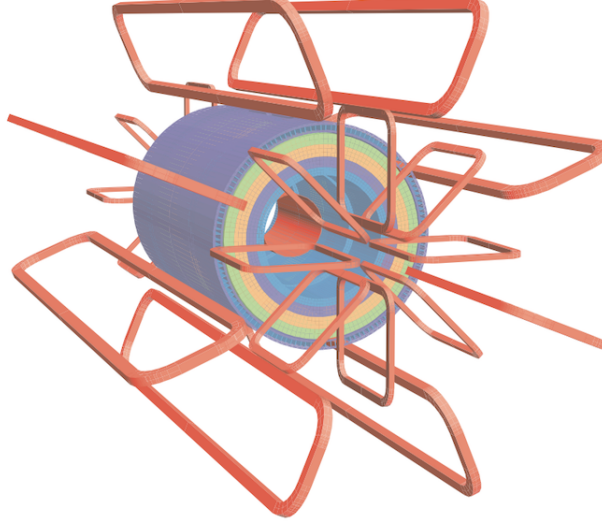


Figure 2.7: The three toroids positioned between the outer layers of the muon system and the solenoid encircling the ID in the ATLAS magnetic system [17].

$10 \mu m$ in the transverse plane ($R - \phi$) and $115 \mu m$ in z for the barrel, while for the end-caps, a spatial resolution of $10 \mu m$ in the $R - \phi$ plane and $115 \mu m$ in the longitudinal direction (z) are achieved. The transit of charged particles across the silicon creates electron-hole pairs drifting back to back under a potential difference in the range between 150 to 600 V recorded as hits.

2.4.3.2 The Semiconductor Tracker (SCT)

The Semiconductor Tracker (SCT), the middle layer of the ID is a silicon micro-strip detector technology [56]. It consists of a barrel layers made of four cylinders and two end-caps each composed of nine disks. It is basically dedicated to allow high track precision reconstruction and contribute to the reconstruction of impact parameter, momentum, and define the vertices localization in the intermediate radial range. The full SCT has a silicon-covered surface area of approximately $61 m^2$ where the detection principle are similar to the pixel detector but has less granularity. The SCT contains only 6.3 million readout channels which lead to an intrinsic spatial resolution of $17 \mu m$ in the $R - \phi$ plane and $580 \mu m$ in the z direction.

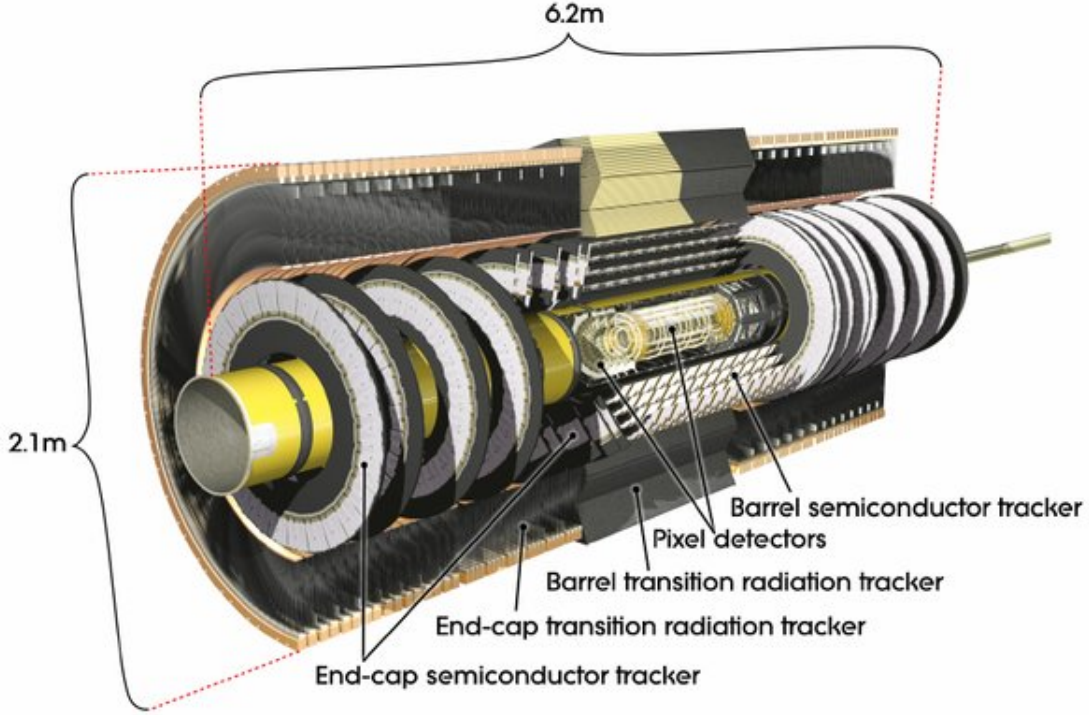


Figure 2.8: A sketch of the ATLAS inner detector, labelling the individual components. Credit CERN [16].

2.4.3.3 The Transition Radiation Tracker (TRT)

The Transition Radiation Tracker (TRT) [57], the outermost layer and the largest of the three tracking subsystem of the ID. It consists of about 350.000 drift tubes (called strawtubes) filled with a xenon-based gas, spread out in a barrel and two end-caps on each side of this barrel. Strawtubes in the barrel region of TRT are 144 cm in length with a diameter of 4 mm, oriented in the same direction of the beam, while 37cm in length for the end caps radially aligned to the collision axis. A charged particle traversing the TRT ionise the gas inside the straws which produce on average 35-40 one-dimensional $R - \phi$ hits with an intrinsic resolution of $130 \mu m$. The generated electrical signals are then recombined to improve the determination of a particle's trajectory and momentum in the transverse plane. Moreover, TRT allows a good discrimination between electrons and pions since the amount of transition radiation changes in function of the particle mass.

2.4.4 The ATLAS Calorimeters

The ATLAS calorimeters system comprises electromagnetic (ECAL) [58] and hadronic (HCAL) [59] subsystems (Figure 2.10). The purpose of calorimeter detectors is the energy

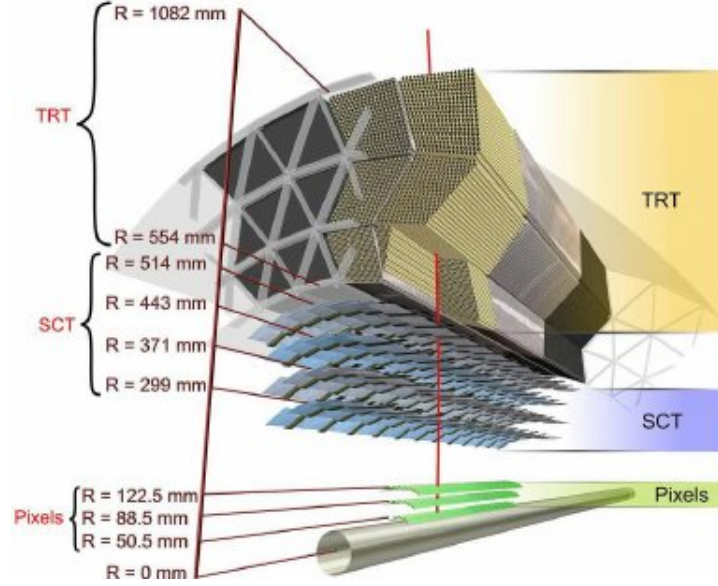


Figure 2.9: Components of the ATLAS Inner Detector shown from the side. Credit CERN [16].

measurement including both charged and neutral particles with the highest possible precision. The measurement is based on the formation of a shower of particles due to particles absorption which ionize the liquid active medium (typically scintillators).

2.4.4.1 Electromagnetic Calorimeter

The first calorimeter, just outside the ID is called the electromagnetic calorimeter (ECAL) dedicated to identify electromagnetic showers initiated by electrons or photons and the measurement of their energy. The ECAL use liquid argon (LAr) as ionization medium combined with the lead plates absorbers arranged in an accordion geometry ensuring a full coverage symmetry in the ϕ coordinate. The ECAL is divided into three main part: a barrel calorimeter covering the pseudo-rapidity region $|\eta| < 1.475$ and two end-cap calorimeters covering $1.375 < |\eta| < 3.2$. LAr was chosen as the active medium [60] because of its linear response and radiation hardness. When particles traverse the lead absorbers, they create photons via bremsstrahlung process; these photons in turn convert into e^+e^- pairs when ($E_\gamma \geq 1\text{MeV} = 2m_0^e \approx 1\text{MeV}$), and these high energetic electrons will interact with the liquid-argon and ionises its atoms, producing an electromagnetic shower (more photons). The result being a shower of electrons and photons which continues until all their energy is absorbed by the calorimeter. Based on the amount of charge deposited on the electrodes in its path, we can measure the energy of the original electron or photon that entered the

electromagnetic calorimeter

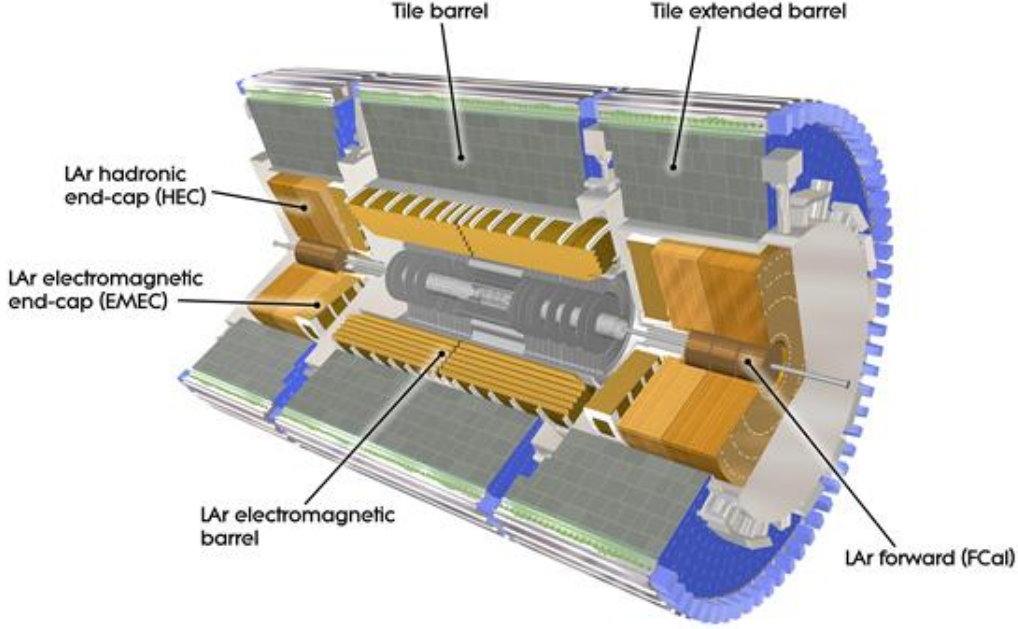


Figure 2.10: Sketch of the ATLAS calorimeters. Credit CERN [16].

2.4.4.2 Hadronic Calorimeter

The Hadronic Calorimeter (HCAL) [61] located behind the corresponding parts of the (ECAL). It consists of three parts: tile calorimeter covering the region $|\eta| < 1.7$, hadronic end-cap calorimeter in the range $1.5 < |\eta| < 3.24$ and forward calorimeter covering $3.1 < |\eta| < 4.9$. Steel serves as the absorber, and plastic scintillators serve as the active ingredient in the TileCal. while the end-caps use LAr as the active medium due to the increased radiation close to the beam pipe interleaved with copper plates used as an absorber. HCAL is dedicated to measure the energy depositions of hadronic jets caused by hadrons colliding with the nuclei of the materials they are traversing via inelastic scatterings at energies of the order of ≈ 100 GeV. New particles are created as a result of these nuclear interactions, resulting in a hadronic shower that is significantly larger and emerge trough large distances than an electromagnetic shower. The absorber chosen by HCAL seeks to reduce the number of interaction lengths indicated in the TileCal ($\lambda_0 \approx 9$).

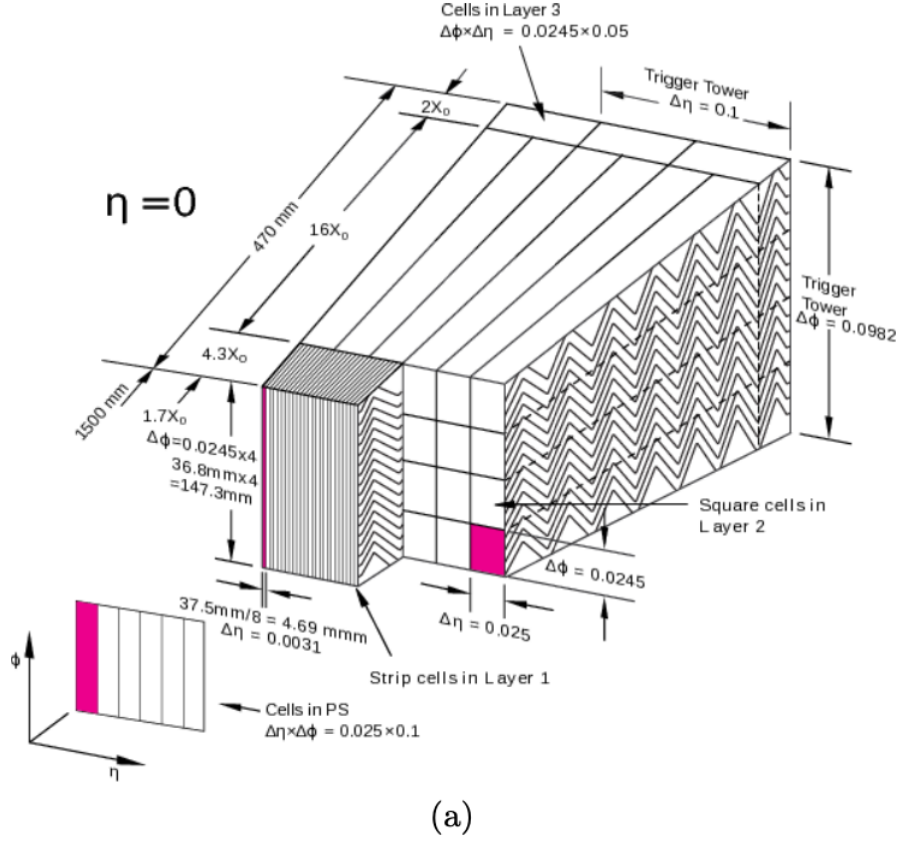


Figure 2.11: a picture of the ECAL module at barrel $\eta = 0$ [18].

2.4.5 The muon spectrometer (MS)

The muon spectrometer (MS) (Figure 2.12) is the outermost part of the ATLAS detector which is designed to the detection and the momentum measurement of muons based on their tracks curvature. Muons pass the ATLAS detector more than any other particle with very low interaction probability since they are not sensitive to the strong interactions and are relatively heavy particles [62]. After passing through the calorimeters, muon tracks are deflected along the beam direction using a toroidal magnetic field. The MS integrates four detector subsystems: resistive plate chambers (RPC), cathode strip chambers (CSC), and narrow gap chambers (TGC), as well as barrel and end-cap toroid magnets.

The MDT chambers used most for the precise measurement of particles momenta. Each MDT chamber is made of multi-layers of aluminium drift tubes with a diameter of 3 cm. In total there are 1088 chambers, each containing three to eight MDTs, extended along tree concentric layers perpendicular to the beam axis. The drift spatial resolution of each

individual MDTs is ($80 \mu\text{m}$) with a typical drift time of 700 ns. The CSCs are composed of multi-wire proportional chambers in orthogonal directions, covering the forward region $2 < |\eta| < 2.7$, designed to provide high precision tracking close to the beam pipe. When muons translate through the weirs, they ionise the gas and lead to production of electron-ion pairs: ions drift to the cathode while electrons deflect to the anode, creating an electric signal. Due to the perpendicular orientation of the anode and cathode strips, two spatial coordinates with a spatial granularity of $40 \mu\text{m}$ in the radial direction (R) and 5 mm in the longitudinal direction (ϕ) may be recorded. The timing resolution is approximately 7 nanoseconds. RPCs can provide high spatial resolution bi-dimensional space coordinates (2 to 6 mm in R-plane and 3 to 7 mm in the ϕ –plane) while a time resolution of just 1 ns. TGCs are a multi-wire proportional chambers with a 2.8 mm gap used by the muon trigger system in the end-cap region to ensure a high granularity. The wires allow a measurement of the muon track coordinate η and ϕ with time resolution smaller than 25 ns for 99% of tracks, which lead these detectors to act as trigger chambers.

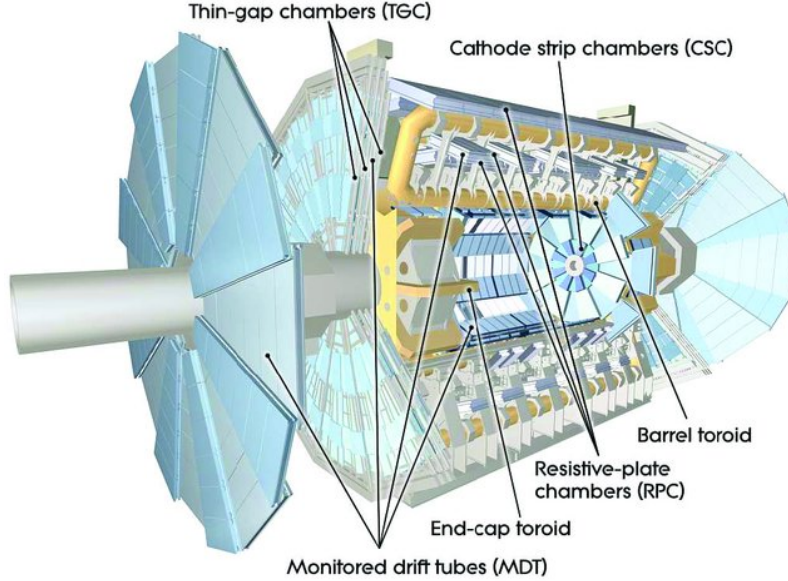


Figure 2.12: The ATLAS Muon Spectrometer and toroid magnets. Credit CERN [16].

2.4.6 Trigger and Data Acquisition

To decrease the massive amount of data collected by the LHC (up to 40 MHz) to approximately 1 kHz where only events with promising properties that make them intriguing for physics analyses are stored, ATLAS implemented a specialised multi-level Trigger System

that consists of three layers: Level-1, Level-2 and event filter

The Level-1 (L1) [63] trigger use informations from the calorimeters and muon spectrometer, while all ATLAS sub-detectors, including the inner detector, are used in the Level-2 (L2) trigger. L2 together with Event Filter (EF), compose the High-Level Trigger (HLT) [64]. A block diagram of the trigger and data acquisition systems is shown in Figure 2.13

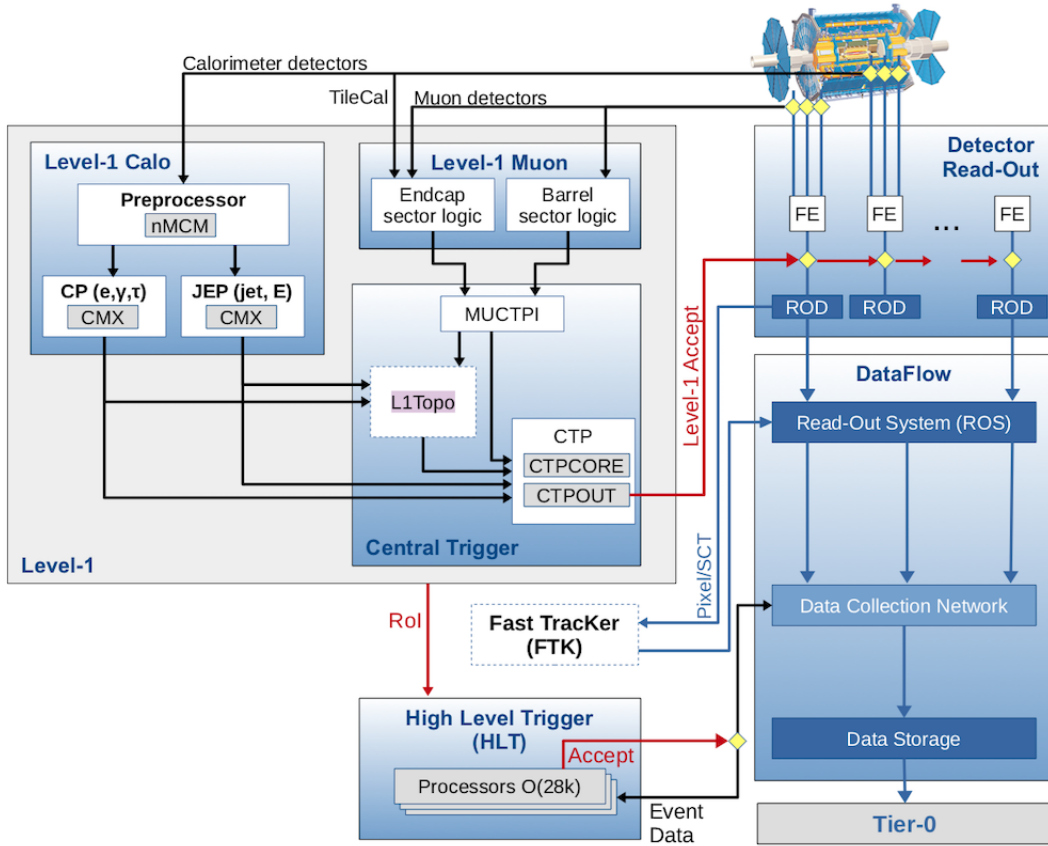


Figure 2.13: The components important for triggering in the ATLAS TDAQ system in Run 2. In 2015, L1Topo and FTK were put into service [19].

2.4.6.1 Level-1 Trigger

The Level-1 (L1) trigger is a hardware-based event that marks the initial trigger on proton-proton collisions. Calorimeter Level-1 (L1Calo) and muon Level-1 (L1Muon) are the two types of Level-1 triggers.

L1Calo uses information recieved from calorimeter sub-detectors (ECAL and HCAL) to identifies physical objects with high transvers energy, like eletrons, photons, τ leptons and

jets. L1Muon looks for physical signatures originating from high- p_T muons in the MS, or large missing momentum to select events with interesting signature among the dominant background process. The L1 trigger, for example, had a latency of roughly $2.5 \mu s$ during Run-1 data collection and lowered the events rate to 70 kHz. The L1 trigger identifies physical regions-of-interest (RoIs) in the ATLAS detector that will be utilized as a target in the filtering phase that follows.

2.4.6.2 Level-2 Trigger

To perform basic event reconstruction, the Level-2 (L2) trigger start form information was decreased in the regions-of-interest (RoIs) identified by the L1 trigger. Within a latency of 200 milliseconds, the L2 reduces the quantity of data from the L1 up to 70 kHz to around 6.5 kHz.

2.4.6.3 Event Filter

The event filter (EF) represents the last step of the trigger process. It manages the fully reconstructed event using a large computer farm, to look for interesting objects from the full granularity of the detector. EF reduces the events rates to 1kHz which are subsequently transferred to CERN's central data-recording facility. The HLT trigger rate of the primary physics streams in term of time for data collected in 2018 with a peak luminosity of $L = 2.0 \times 10^{34} cm^{-2} s^{-1}$ and a peak mean number of interactions per crossing of $\langle \mu \rangle = 56$ is shown in Figure 2.14. The HLT trigger rates diminish exponentially as the brightness increases throughout a fill. The spikes are caused by detector noise, while the dips are seeded by dead time.

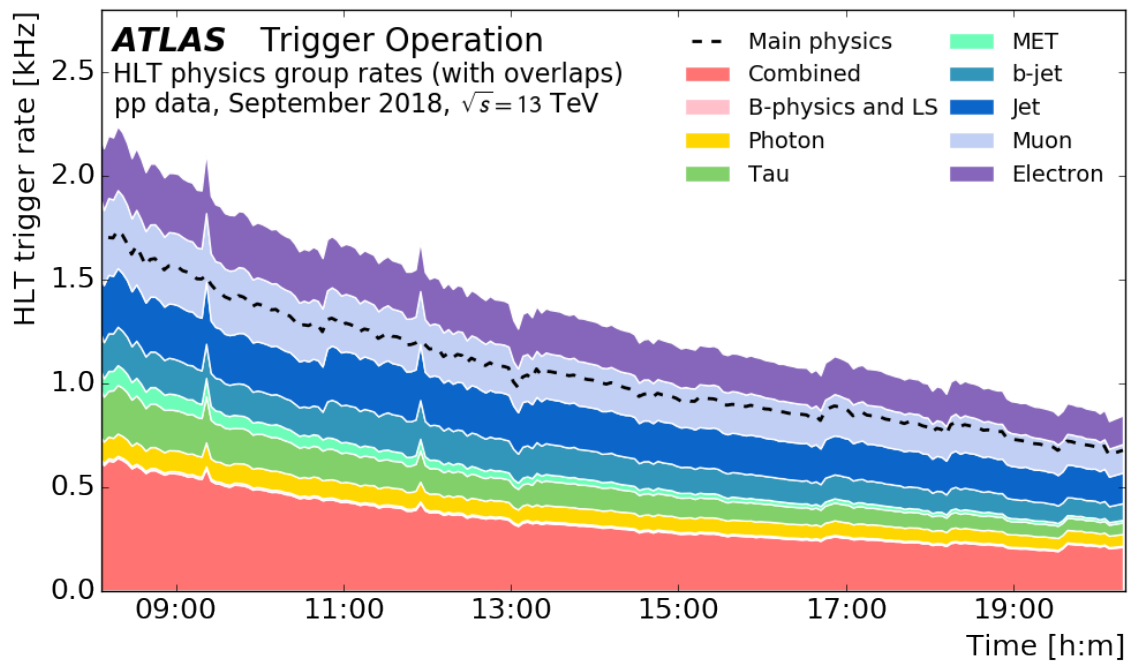


Figure 2.14: In a fill taken in 2018, HLT physics trigger group rates are shown in term of time [20].

Chapter 3

Physics Objects

In ATLAS, the electronic signals produced in proton-proton collisions might be converted into physical entities (muons, photons, electrons, taus) or physical processes like jets. The objects reconstruction is done using dedicated software algorithms, translating the detector responses to energy cluster, track and four-momenta which is calibrated to correct the detector effects. Figure 3.1 provides a graphic representation of how ATLAS finds particles.

In this chapter, we give a general view of the algorithms used for particle reconstruction in the ATLAS detector. We will focus only on the objects crucial for the search of $t\bar{t}$ -resonances such as: leptons (electrons and muons), hadronic jets created by either gluons or quarks while the E_T^{miss} is described in chapter 4

3.1 Electrons

3.1.1 Electron identification and reconstruction

Electrons are identified by tracks in the tracking layers of the ATLAS detector, also their energy deposits in the ECAL as shown in the schematic Figure 3.2. In case with at least one track is associated with its cluster, the electrons in the core zone of the Inner Detector ($|\eta| < 2.5$) are properly reconstructed. The signal from each ECAL cell is processed in intervals of 25 ns. The space of ECAL is distributed into a grid of $N_\eta \times N_\phi = 200 \times 256$ elements of size $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ rad towers which combine the calorimeter layers. A sliding window algorithm [65] then scans the whole ECAL clusters candidates and records only cluster seeds containing energy higher than 2.5 GeV as displayed in Figure 3.3. The efficiency of this clustering algorithm depends on E_T and $|\eta|$. It ranges from 65% at $E_T = 4.5$ GeV, to 96% at $E_T = 7$ GeV, to more than 99% above $E_T = 15$ GeV as shown in Figure 3.4. Once the seed clusters are picked, they are tentatively matched to the well-reconstructed ID tracks

utilizing differences in η and ϕ separating the track's impact point on the ECAL and the electron cluster barycentre to identify the electron candidates.

Electron identification (ID) algorithms are employed to separate particles like pions and converted photons from electrons. A likelihood-based (LH) identification method [22] is used to discriminate the prompt electrons. It is a multivariate analysis approach employed to evaluate many characteristics of the electron candidates from background-like non-prompt electrons. The LH method has three "working points" (in order of increasing background rejection): Loose, Medium, and Tight, corresponding to a 93%, 88%, and 80% efficiency to reconstruct a $E_T = 40$ GeV electron, respectively. Figure 3.5 shows the identification efficiency of these working locations as a function of the electrons' transverse energy (E_T).

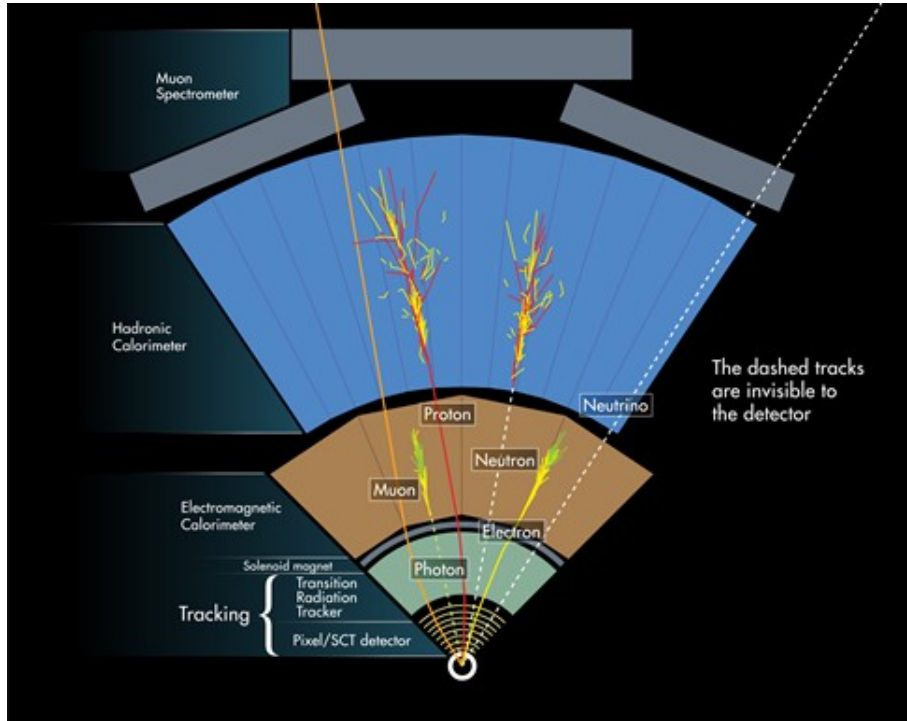


Figure 3.1: An illustration of the ATLAS particle detection process [21].

3.1.2 Electron isolation

The electron isolation criteria linked to the p_T of near tracks to those of the electron candidate and the energy deposited by other physical objects close in the ECAL. They are further employed to differentiate prompt electrons from non-prompt electrons resulting from hadrons mis-reconstruction and photon conversions. The first Isolation variable is calculated as the energy deposited in a cone of radius $\Delta R = 0.2$ around the associated cluster to the

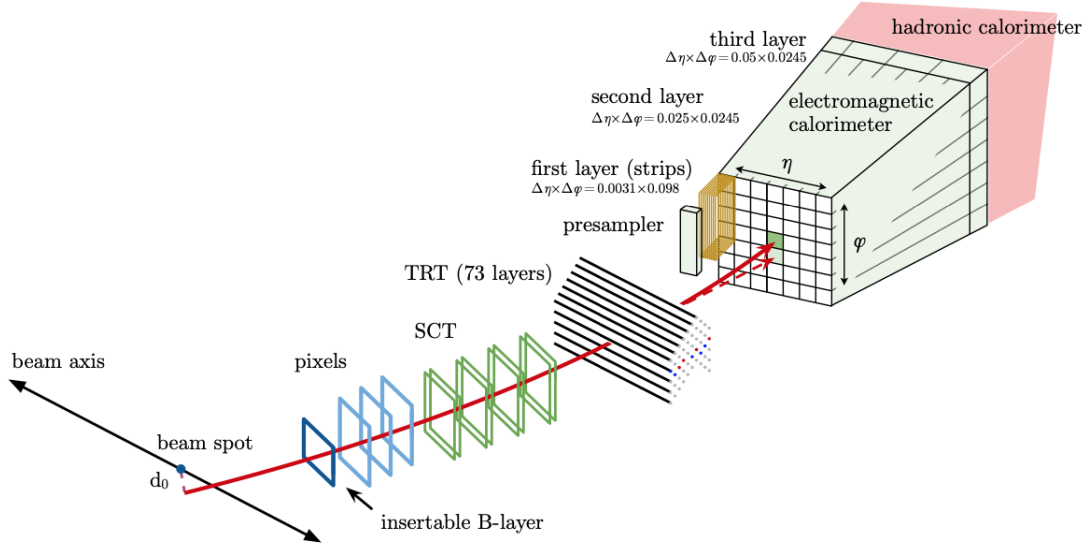


Figure 3.2: Illustration of the trajectory of an electron (in solid red) through the tracking detectors and the ecal. The dashed line shows the trajectory of a photon emitted by the electron in the tracking system. Taken from Ref [22]

electron candidate in the ECAL [22], $E_T^{\Delta cone 0.2}$, and the second parameter is measured as the sum of the p_T of tracks surrounding the electron track, $p_T^{\Delta cone 0.2}$, after subtraction of tracks associated to the electron. Based on isolation selection and efficiency requirements, various isolation operating points (OP) are defined. Isolation OPs are defined in three categories: track-only, calo-only and combined track and calo.

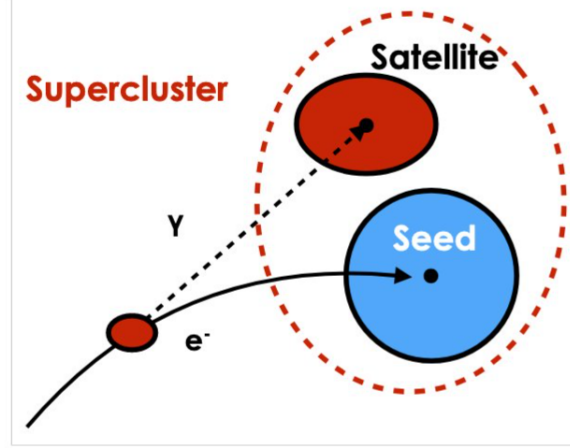
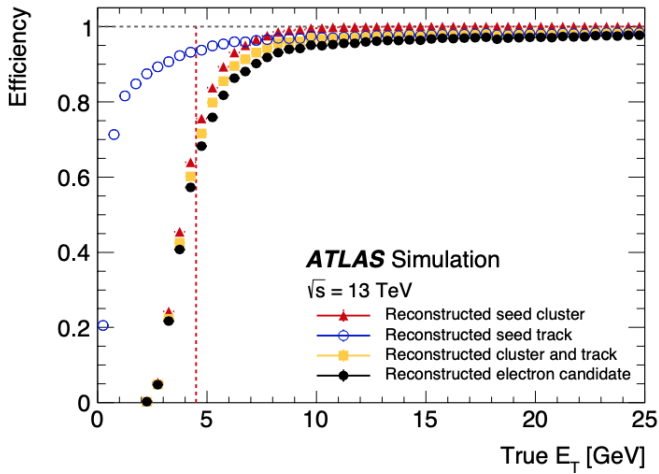
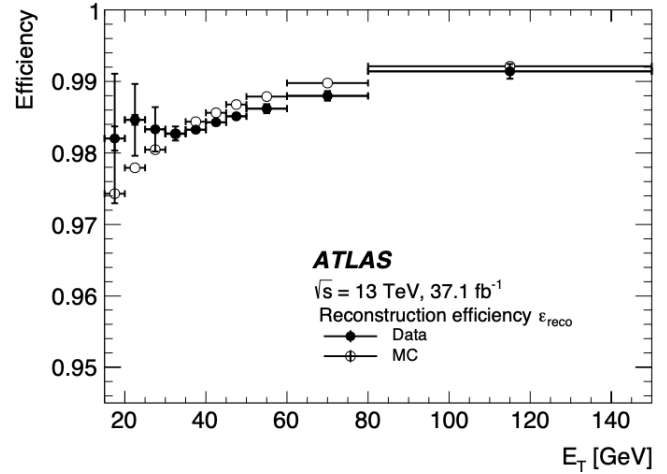


Figure 3.3: Example of a supercluster presenting a seed electron cluster and a satellite photon cluster [23].



(a)



(b)

Figure 3.4: Total electron reconstruction efficiency as a function of true electron. E_T calculated from single electron simulation sample. (b) Reconstructed efficiency relative to reconstructed clusters, given by the number of reconstructed electron candidates N_{reco} , divided by the number of EM-cluster candidates $N_{cluster}$, as a function of electron transverse energy, E_T calculated from $Z \rightarrow ee$ events. Figure taken from [22]

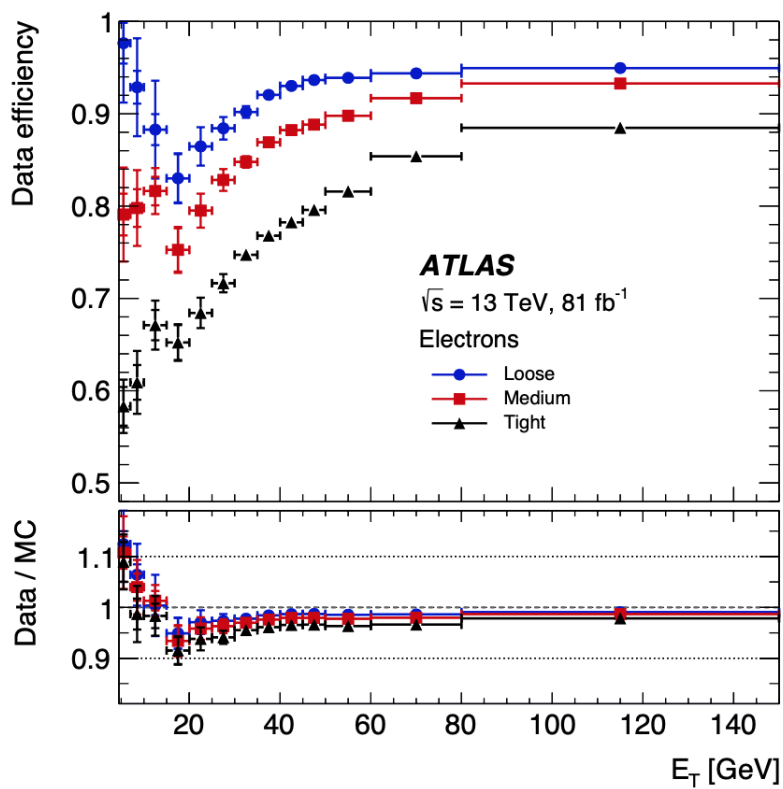


Figure 3.5: The efficiency of electron identification as a function of E_T using different working points. Taken from Ref [24]

3.2 Muon

3.2.1 Muon identification and reconstruction

Muon reconstruction relies on combining information from both tracks in the Muon Spectrometer (MS) and in the inner detector to form the muon tracks used in physics analyses. Hits deposit in the MS are firstly linked into tracks, then associated to charged tracks in the ID. Numerous algorithms are available to locate and reconstruct muons, as a result, there are four different types of muon candidates that can exist as illustrated in Figure 3.6:

- CB: “combined muons” when the association algorithm successfully finds combination of ID and MS tracks. This category offer the most accurate muon candidates with lowest momentum resolution.
- ST: “segment-tagged” candidates are built by projecting a segment from the MS to the ID tracks. ST muons are used at low transverse momentum and when muons traverse only one layer of the MS chambers.
- CT: “calorimeter-tagged” muons have track in the ID matched to an energy deposit in the calorimeter, if it is compatible with a minimum-ionising particle. The lowest purity muons are CT muons, however, they regain acceptance outside of the MS for $|\eta| < 0.1$. but they recover acceptance in the $|\eta| < 0.1$ zone not within the MS.
- ME: “muons extrapolated” are built based on the MS track with a loose requirement on the track to interaction point association.

Muon identification is performed to distinguish between background candidates and prompt muons, which are mainly produced by hadrons decay (like kaons and pions). Muon identification is performed to make difference between prompt muons and other background candidate mostly coming from hadrons decay (like kaons and pions). The applicable selection and discriminating factors, such as the muon p_T , the energy deposit, charge, and the fitting-method for track reconstruction, provide the basis for identification. According to the signal effectiveness and background rejection, muons are divided into four identification operating points: High- p_T , Medium, Loose and Tight. For the first three categories, the identification efficiencies vary from 92% to 98% in the range $|\eta| < 2.5$ and $20 \text{ GeV} < p_T < 100 \text{ GeV}$. Muon identifications criteria are optimized using $Z \rightarrow \mu\mu$ and J/Ψ decay [25] as shown in Figure 3.7.

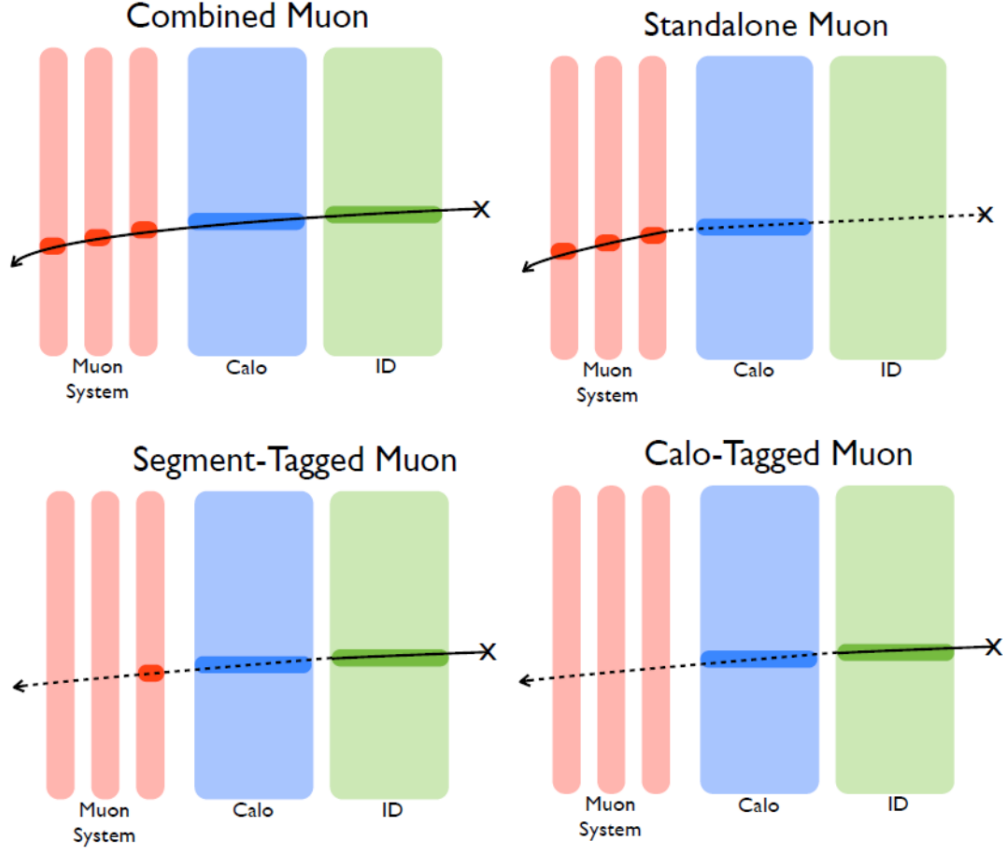


Figure 3.6: Muon reconstruction techniques based on signals from several ATLAS detector regions.

3.2.2 Muon isolation

Muon isolation considers comparable method as that employed to separate the prompt electrons, where two additional parameters are employed to evaluate the muon isolation [66]: $E_T^{topocone20}$ identified as the sum of the transverse energy of the topo-clusters around a cone of size $\Delta R = 0.2$ surrounding the muon after removing the muon energy deposits and correcting the pileup effects. The second variable is $p_T^{varcone30}$ defined as the sum of the transverse momenta of the tracks with $p_T > 1$ GeV in a cone of size $\Delta R = \min(10 \text{ GeV}/p_T^\mu, 0.3)$ around the muon p_T , excluding the muon track. Based on the isolation selection, several isolation operating points are divided into tree groups: track-only isolation, combined track and calorimeter based isolation. In this thesis trackonly isolation operating point TightTrackOnly_VarRad, is required for prompt muon isolation.

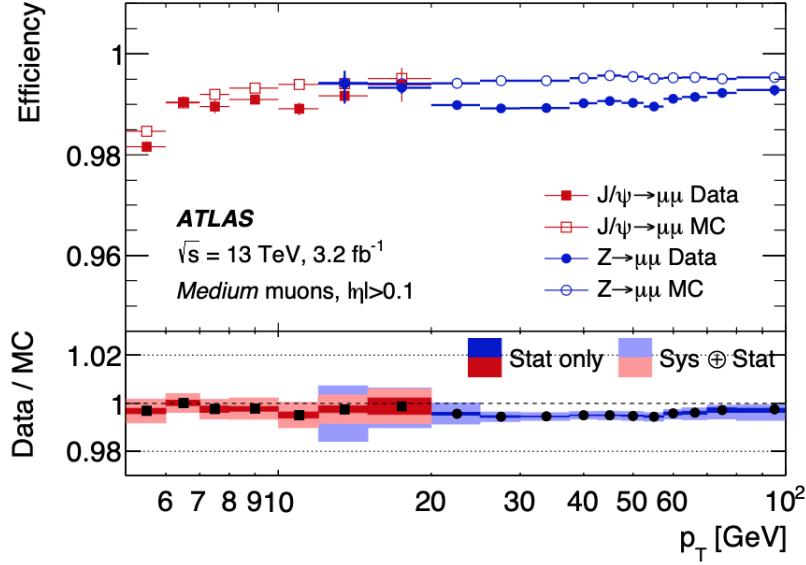


Figure 3.7: Muon identification techniques based on signals from several ATLAS detector regions [25].

3.3 Jets

3.3.1 Jet reconstruction

The QCD confinement forces quarks and gluons (partons) to create a shower of hadrons and produce so-called jets. Using a special anti- k_t method [67], jets are collimated sprays of energetic hadrons (Figure 3.8) from the fraction energies deposit in the calorimeters or tracks of charged particles in the tracking systems. This algorithm iterates over a set of topoclusters objects to evaluate the separation distance d_{ij} for all pairs of objects detected i, j and the distance measure d_i^B of each object i to the beam:

$$d_{ij} = \min(p_{T,i}^{-2}, p_{T,j}^{-2}) \frac{\Delta R_{ij}^2}{R^2} \quad (3.1)$$

$$d_i^B = p_{T,i}^{-2} \quad (3.2)$$

where $p_{T,j}$ and $p_{T,i}$ represent the p_T of the jet candidates j , and i , respectively. Also, $\Delta R_{ij} = \sqrt{(\Delta\eta_{ij})^2 + (\Delta\phi_{ij})^2}$ represents the angular separation between the jet candidates (i and j) for a jet with cone radius R . Algorithm starts by identifying the lowest distance between jet candidates d_{ij}^{min} and between the jet compared to the beam axis d_i^B . If $d_{ij}^{min} <$

$d_i^{B,min}$, we combine the two jets to form a pseudo-jet by adding the four-momenta of jet i and j . Then, the new jet candidate is incremented back to the jet candidates list.

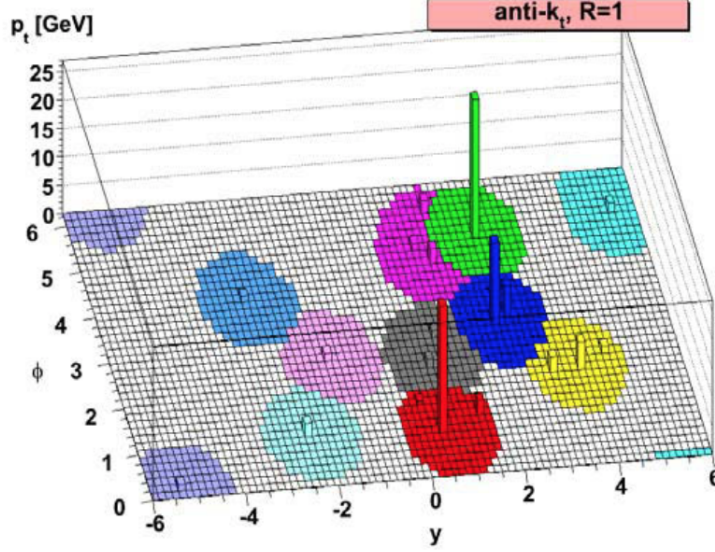


Figure 3.8: Using the anti-kt technique, a parton-level event with several arbitrary soft contributions was grouped with a $R=1$ value. Nice circular distribution of jet with highest p_T is achieved.

3.3.2 Jet calibration

The uncertainty caused by the jet energy calibration might affect with large fraction the total measurement uncertainty, therefore, good calibration of jets is of high importance. Several calibration strategies are developed to correct the jet energy to account for effects like pile-up, non-compensating calorimeter response, discrepancy between data and MC etc. The calibrations bellow are applied subsequently, as displayed in Figure 3.9

- Origin correction: corrects the jet direction so that the analysis can focus on the principal vertex that was chosen.
- Pile-up correction: corrects for the energy offset produced by pile-up using an area-based subtraction procedure and residual corrections. The latter is a function of the number of primary vertices, N_{PV} , and the average number of interactions per bunch crossing $\langle \mu \rangle$, and captures residual in-time and out-of-time effects. Derived from Monte Carlo simulated events.

- Energy and η calibration: calibrates the jet energy by applying p_T and η -dependent corrections derived from Monte Carlo simulated events. Biases in η due to different calorimeter technologies in different regions are corrected for, too.
- Global sequential corrections: reduce the dependence of the jet energy on the longitudinal and transverse structure of the jet and additionally correct for jets that are only partially contained in the calorimeter. Derived from data events.
- Residual in situ corrections: calibrating remaining (residual) issues, derived in situ, i.e. from data events

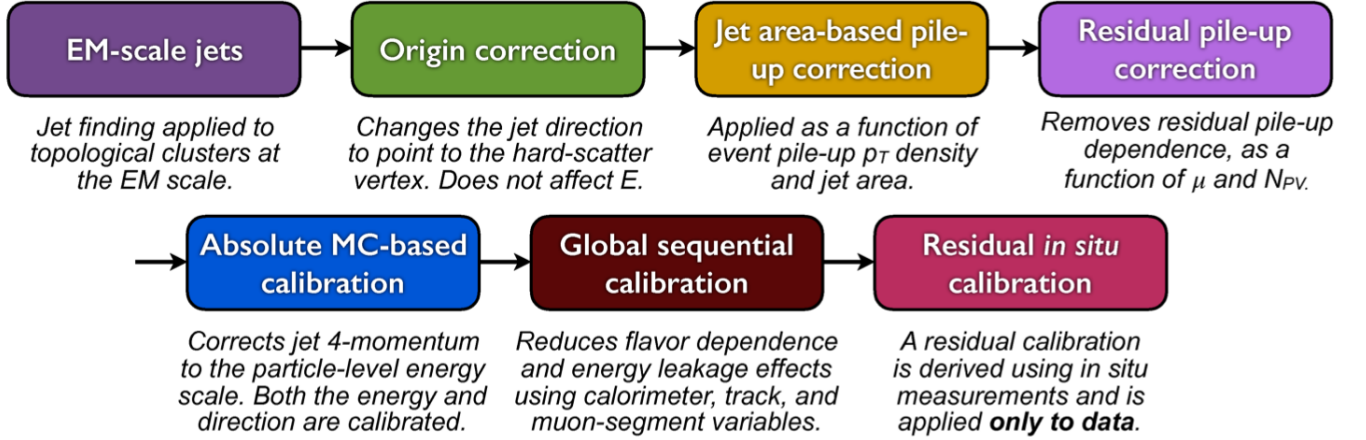


Figure 3.9: JES calibration procedure for the jet energy scale [26].

3.3.3 Jet pile-up rejection

The suppression of pile-up jets originating from pile-up interactions [27] with respect to the jets contribution from the hard-scatter interactions, is essential for proper measurement of the process of high interest. Since the hadronic calorimeter does not give any indices about the origin of clusters, the pile-up jet rejection is based on the characteristics of tracks associated to jets. Pileup jets are suppressed by the jet-vertex-tagger (JVT) discriminant which is a tagging algorithm that identifies jets originating from the primary vertex (PV). It use track information derived from a multivariate study utilizing the R_{P_T} and corrJVF variables [68]. Jets must satisfy a JVT cut greater than a fixed values which according to three working points: Tight, Loose and Medium. Figure 3.10 displays the JVT distributions for jets created from primary vertex and those from pile-up.

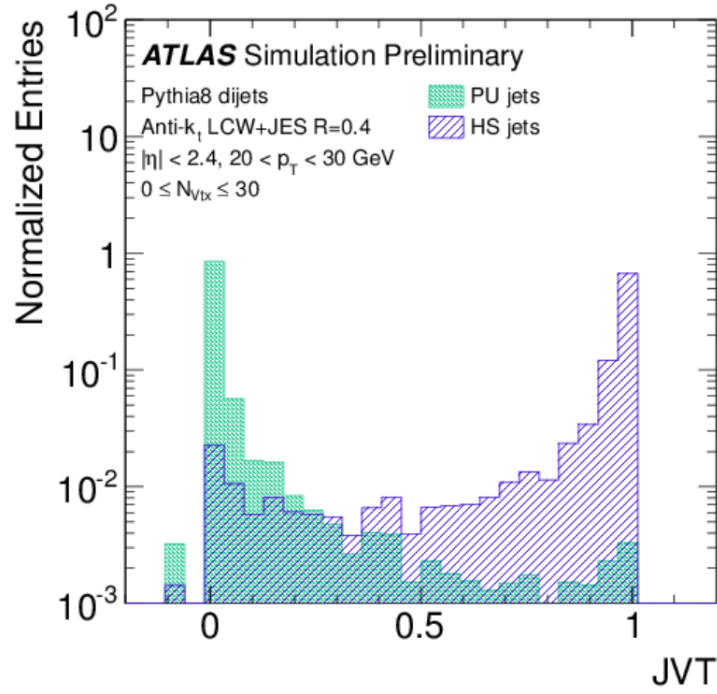


Figure 3.10: Shape of hard-scatter (HS) and JVT for pileup jets. While HS jets collect for high JVT values, pile-up jets only manage to do so for low JVT levels. Jets without an accompanying track are impacted by $JVF = -1$ [27].

3.3.4 Jet flavour tagging

It is decisive for many analyses to identify jets originating from the hadronisation process of b-quarks. This is known as b-tagging aiming to decrease the b-jets from the jets produced by the hadronization of light-jets and a c-quark. The b-jets have properties including a lengthy lifespan of order 1.5 ps and may cross a distance of $\approx 1 \text{ mm}$ before decaying. This extensive travel that happens in the ID may be recognized as a rebuilt vertex, known as the secondary vertex (PV2), which is shifted in relation to the PV's actual localisation as displayed in Figure 3.11

The 77% efficiency as implemented in the DL1r tagging method was employed in this thesis for the search for $t\bar{t}$ resonances. The b-tagging algorithm has multiple efficiency levels available.

3.3.5 Jet trimming

Jet trimming is a grooming method considered for large- R jets to decrease the effect of contained energy from pile-up contamination and underlying events. The constituents of the

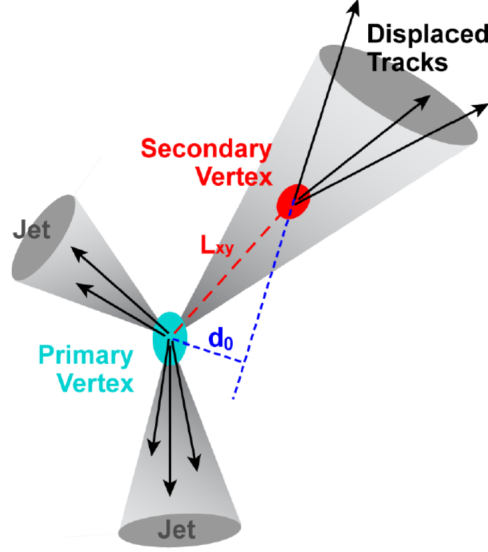


Figure 3.11: An example of a schematic diagram that compares secondary vertex generation from long-lived particle decay to other jets started from the main vertex.

jet are reclustered into subjets with the k_t algorithm¹ and the radius parameter R_{sub} is the first step. In a second step all subjets within a jet with p_T of less than a fraction f_{cut} of the ungroomed jet are discarded [69]. The trimming parameters used for the presented analysis are $R_{sub} = 0.2$ and $f_{cut}=5\%$.

3.3.6 Top-tagging

The identification of heavy particle decays, in particular top decays, discriminates signal from background and can be achieved via jet substructure techniques. A so called Smooth Top Tagger [70], can be used, which applies cuts on one or two substructure observables only. As the latter is in general depending on the jet momentum an important algorithm criterion is its stability concerning the signal efficiency for a wide range of jet p_T . A flexible procedure has been investigated to find the optimal two variable combination with smoothed jet p_T dependent cuts at fixed signal efficiency working points. The following substructure variables are used for the Smooth Top Tagger:

- k_t splitting scales: k_t splitting scales [71] are determined by re-clustering a large R-jet using the k_t algorithm, which tends to cluster high p_T ingredients with large distances

¹The k_t algorithm is a sequential clustering algorithm for jet reconstruction, same as anti- k_t

last. They are identified as the distance $\sqrt{d_{ij}}$ between the subjets i and j in the k_t history starting with its last step.

$$\sqrt{d_{ij}} = \min(p_{T,i}, p_{T,j}) \Delta R_{ij} \quad (3.3)$$

The splitting scale $\sqrt{d_{23}}$ is defined from the second to last clustering step, expected to cover subjets originating from the W boson decay with $\sqrt{d_{23}} \approx (\frac{m_W}{2})(\sqrt{d_{12}} \approx \frac{m_t}{2})$

- N-subjetiness: The N-subjetiness variables τ_N [72] are an indicator of the subjet multiplicity of a large-R jet. To calculate τ_N the k_t algorithm is applied to a jet, merging the constituents until N subjets are composed. τ_N is determined as the sum of all p_T weighted distances between the constituent k and the closest subjet axis and can be interpreted as a measure of how well the substructure of the jet is described with N subjets. The N-subjetiness ratio $\tau_{32} = \frac{\tau_3}{\tau_2}$ is employed to discriminate top events against soft gluons, light quark jets or W bosons.

3.4 Missing Transverse Energy

Particles like neutrinos and hypothesized dark matter interact with other SM particles and the ATLAS detector in a subtle way. Depositions do not have a signature that may be stated with them, such as jets and leptons. Instead, the energy they transport away from an interaction is used to determine their presence in an interaction.

Momentum should be conserved in interactions even in collider contexts. Because of PDFs, it is impossible to identify separately interacting parton velocities in the z (or beam) direction, although the momentum transverse to this direction is known to be tiny. The momentum should be conserved in the detector's transverse plane as well. Any equilibrium in transverse momentum shows that which is transported away by a particle that did not interact with the detector once all physics objects in an event have been determined. The negative vector sum of all physical objects and evaluating tracks in the event are employed to measure the E_T^{miss} . More detail will be provided in Chapter 4.

Chapter 4

Overlap Removal and performance of E_T reconstruction

The missing transverse momentum (E_T^{miss}) [73] is a critical observable performing as an experimental representative for the p_T associated to particles not interact with any of the detector components created via proton–proton (pp) collisions. E_T^{miss} is reconstructed from the calorimeter, tracks built in the muon spectrometer and the inner detector with the ATLAS detector. A precise measurement of the E_T^{miss} in terms of linearity, resolution and the absence of tails is a key for many physics studies in ATLAS, especially the search of new particles introduced by physics beyond the Standard Model such as: dark matter, SUSY and extra dimensions.

4.1 Missing transverse momentum measurement in ATLAS

In the direction perpendicular to the beam axis, perfect momentum conservation indicates that the sum of the p_T of all particles in the final state should be effectively zero. Every momentum inequality seen (Known as E_T^{miss}) may be a sign that invisible particles to the detector systems were created in the collision, like neutrinos or possibly exotic particles BSM that pass through the ATLAS detector without being noticed.

The measurement of E_T^{miss} is a real challenge since it involves different signals of all detected particles in various ATLAS detector subsystems, therefore, many sources of noise and imperfections can affect its calculation like: Object momentum miscalibration, mis-measurement, and particles going through dead zones of the detector. Another source of noise is the contributions from additional p-p collisions, called “pileup”, concurrent with the

hard-scatter process. Pile-up impact must be suppressed or at least minimized since its momentum contributions might enter into the E_T^{miss} calculation, contaminating the true E_T^{miss} . All the previous effects create distorted distributions compared to the real E_T^{miss} signatures and contribute to arise the fake E_T^{miss} .

Using either energy deposits in the calorimeters, tracks reconstructed in the ID, or a combination of both, the E_T^{miss} is rebuilt as the negative vector sum of the transverse momenta p_T of all identified particles. The E_T^{miss} reconstruction is characterised by two contributions: the reconstructed hard and soft terms. The E_T^{miss} "hard term" comprise fully reconstructed and calibrated muons, electrons, τ -leptons, photons, and jets in each event. The E_T^{miss} "soft term" refers to signals made up of reconstructed charged-particle tracks linked to the primary vertex but not to the hard physics objects.

4.2 E_T basics

The missing transverse momentum vector $\vec{E}_T^{miss}$ components (E_x^{miss} , E_y^{miss}) is reconstructed via the magnitude across the x and y axes of the various contributions:

$$E_{x(y)}^{miss} = E_{x(y)}^{miss,\mu} + E_{x(y)}^{miss,e} + E_{x(y)}^{miss,\gamma} + E_{x(y)}^{miss,\tau} + E_{x(y)}^{miss,jet} + E_{x(y)}^{miss,soft} \quad (4.1)$$

Where each component $E_{x(y)}^{miss}$ is determined as the negative vectorial sum of $p_{x(y)}$ of energy deposits or trajectories corresponding to muons (μ), electrons (e), photons (γ), hadronically decaying τ -leptons, and jets.

$$E_{x(y)}^{miss} = - \sum_{i \in \text{hard objects}} p_{x(y),i} - \sum_{j \in \text{soft signal}} p_{x(y),j} \quad (4.2)$$

With the magnitude of $\vec{E}_T^{miss}$ is given as:

$$E_T^{miss} = |\vec{E}_T^{miss}| = \sqrt{(E_x^{miss})^2 + (E_y^{miss})^2} \quad (4.3)$$

And its transverse plane orientation is described in function of the azimuthal angle ϕ^{miss} :

$$E_T^{miss} = \tan^{-1}(E_y^{miss} / E_x^{miss}) \quad (4.4)$$

The resulting transverse momenta in the detector, labelled as $\sum E_T$ evaluates the total event activity. It is a crucial variable for well understand the resolution of the E_T^{miss} , particularly when pileup contributions clearly increase, calculated as the sum of the p_T of hard objects

and soft-term signals.

4.2.1 E_T terms definition

Each physical objects (particle and jet) in a given event has its own selection cuts which provide a direct effect on the E_T^{miss} and $\sum E_T$ calculation. Hence, the reconstructed objects have its dedicated calibration and must satisfy isolation criteria in order to reject fake or otherwise problematic signatures.

In this approach, the total E_T^{miss} (Equation: 4.5) is calculated as the vectorial sum of transverse momentum terms $E_T^{miss,p}$, with $p \in (\mu, e, \tau_{had}, \gamma, jet)$ built using the $p_T = (p_x, p_y)$ of recognized physics objects including jets respecting the priority-ordered reconstruction. The soft term $E_T^{miss,soft}$ is built of soft event signals associated to tracks linked to the hard-scatter vertex except hard objects. The pile-up contributions to neutral particle signals from the calorimeter are not considered in the soft term calculation. The applied selections on the physical objects are summarised in Table 4.1.

$$E_T^{miss} = - \underbrace{\sum_{\text{selec muons}} p_T^\mu - \sum_{\text{selec electrons}} p_T^e - \sum_{\text{selec photons}} p_T^\gamma - \sum_{\text{selec } \tau} p_T^\tau - \sum_{\text{selec jets}} p_T^{jets}}_{\text{hard term}} - \underbrace{\sum_{\text{unused tracks}} p_T^{track}}_{\text{soft term}} \quad (4.5)$$

The transverse momenta of hard objects taken into account in the E_T^{miss} reconstruction scalar sums, as well as the soft term, are computed as:

$$\sum E_T = \underbrace{\sum_{\text{selec muons}} p_T^\mu + \sum_{\text{selec electrons}} p_T^e + \sum_{\text{selec photons}} p_T^\gamma + \sum_{\text{selec } \tau} p_T^\tau + \sum_{\text{selec jets}} p_T^{jets}}_{\text{hard term}} + \underbrace{\sum_{\text{unused tracks}} p_T^{track}}_{\text{soft term}} \quad (4.6)$$

4.2.2 E_T^{miss} reconstruction

In the calculation of $E_{x(y)}^{miss}$ the contributing objects to the hard event part must be built from exclusive detector signals. Generally, the order utilized for the E_T^{miss} reconstruction starts from the hard-object contribution: electrons, then photons, hadronically decaying τ -leptons, muons and lastly jets. The cuts flow applied to built particles and jets in order to achieve better E_T^{miss} reconstruction performance in a particular study is discussed in this

CHAPTER 4. OVERLAP REMOVAL AND PERFORMANCE OF E_T RECONSTRUCTION

Type	Selections	Variable	Comments
e	$ \eta < 1.37$ or $1.52 < \eta < 2.47$ $p_T > 10 \text{ GeV}$	$E_T^{miss,e}$ $\sum E_T^e$	all $e^\pm$ satisfying medium reconstruction quality and kinematic criteria
γ	$ \eta < 1.37$ or $1.52 < \eta < 2.47$ $p_T > 25 \text{ GeV}$	$E_T^{miss,\gamma}$ $\sum E_T^\gamma$	all γ satisfying tight quality and kinematic criteria in reconstruction, without signal overlap.
τ_{had}	$ \eta < 1.37$ or $1.52 < \eta < 2.47$ $p_T > 20 \text{ GeV}$	$E_T^{miss,\tau_{had}}$ $\sum E_T^{\tau_{had}}$	all τ_{had} satisfying medium reconstruction quality and kinematic criteria, and without signal overlap.
μ	$ \eta < 2.7$ $p_T > 10 \text{ GeV}$	$E_T^{miss,\mu}$ $\sum E_T^\mu$	all μ satisfying medium quality and kinematic criteria in reconstruction.
jet	$ \eta < 4.5$ $p_T > 60 \text{ GeV}$	$E_T^{miss,jet}$ $\sum E_T^{jet}$	all jets satisfying reconstruction quality (jet cleaning) and kinematic criteria, and without signal overlap.
ID track	$p_T < 400 \text{ MeV}$ $ d_0 < 1.5 \text{ mm}$ $ z_0 \sin(\theta) < 1.5 \text{ mm}$ $\Delta R(track, e - / \gamma \text{ cluster}) > 0.05$ $\Delta R(track, \tau_{had}) > 0.05$	$E_T^{miss,soft}$ $\sum E_T^{soft}$	all ID tracks originating from the hard-scatter matched with a jet vertex satisfying reconstruction quality and kinematic criteria, and not and not associated with any particle or ghost-matched with a jet

Table 4.1: Contributions to E_T^{miss} and $\sum E_T$ from hard objects such as photons (γ), electrons (e), muons (μ), τ – leptons (τ_{had}), jets, and tracks from the soft term [28].

sub-section. The criteria used in this study are only one example of a set of criteria.

4.2.3 Electron and photon

Electrons are built using their shower shapes in the electromagnetic calorimeter (ECAL) which are connected to ID tracks [74]. Electrons to enter in the E_T^{miss} calculation must be identified in the range of $|\eta| < 2.47$, avoiding the area between the end-cap and the barrel of an EM calorimeter, $1.37 < |\eta| < 1.52$. The EM scale [75] is used to calibrate them, also more simulation-based and in-situ corrections [76]. Electrons must satisfy the “Medium” likelihood and “Loose” isolation [74] with $p_T > 10 \text{ GeV}$ to be included in the E_T^{miss} reconstruction. The reconstruction and identification of photons are based on the electromagnetic cascades created in the calorimeters [77]. Selected photons are calibrated using “Tight” criteria discussed in [78]. Moreover, the transverse momentum of photons must satisfy $p_T > 25 \text{ GeV}$ and be included in the region $1.52 < |\eta| < 2.37$ or $|\eta| < 1.37$ to be considered in the reconstruction of the E_T^{miss} . In the same way, photons and electrons created within $1.37 < |\eta| < 1.52$ can participate to E_T^{miss} as a jet.

4.2.4 Muon

Muons candidates are built for η satisfying $|\eta| < 2.5$ using an association of MS and ID tracks. Muons seen outside of the ID area coverage, are defined in the region with $2.5 < |\eta| < 2.7$ from a track fit to MS track segments alone. Further, selected muons for E_T^{miss} calculation by demanding the "Medium" reconstruction quality as cited in [66] with $p_T > 10$ GeV, and a matching to the hard-scatter vertex for muons within $|\eta| < 2.5$.

4.2.5 τ – lepton

Hadronically decaying τ – leptons are built from narrow jets with small matching track number [79]. τ – leptons should satisfy the "Medium" criteria discussed here [80], also pass the kinematic cut: $p_T > 20$ GeV also being in the range $1.52 < |\eta| < 2.47$ or $|\eta| < 1.37$. Each τ – lepton not verifying these τ – identification criteria could participate to E_T^{miss} if satisfying the discussed jet identification selection.

4.2.6 Soft term

Regarding the soft term, there are two different ways to calculate the energy associated to it and they are described below.

- The most inclusive approach is performed starting from the calorimeter topo-clusters (Calorimeter Soft Term) not matched to reconstructed hard objects. The performance of this approach is affected by a large dependence on pile-up.
- For each event, high-quality ID tracks originating from the hard-scatter vertex but not matched to any muon, electron, τ -lepton or jet are considered (Track Soft Term). The Track Soft Term is mainly unaffected by pile-up effects since only reconstructed tracks matched with the hard-scatter vertex are employed, but it excludes contributions from any soft neutral particles.

The high momentum tracks originating from hard vertex, contribute to $E_T^{miss,soft}$ for each reconstructed event if they are not matched with any muon, electron, τ -lepton or jet already entered into E_T^{miss} reconstructed. The required signal-overlap resolution takes out:

- ID tracks with $\Delta R(track, electron/photon\ cluster) < 0.05$;
- ID tracks with $\Delta R(track, \tau - lepton) < 0.2$;
- ID tracks matched with muons;

- ID tracks ghost-associated with fully or partially contributing jets.

The track-based $E_T^{miss,soft}$ is generally insensitive to pile-up impacts and does not consider contributions from any soft neutral particles, despite being created by the hard-scatter interaction, for the soft term because we only consider reconstructed tracks matched with the hard-scatter vertex.

The performance of E_T^{miss} calculation in Monte Carlo simulation and data is displayed in Figure 4.1. The quality of the modelling of E_x^{miss} , E_y^{miss} and E_T^{miss} is evaluated for an inclusive sample of $Z \rightarrow \mu\mu$ events by comparing the distributions of these observables in data recorded by the ATLAS experiment during 2015 and in Monte Carlo.

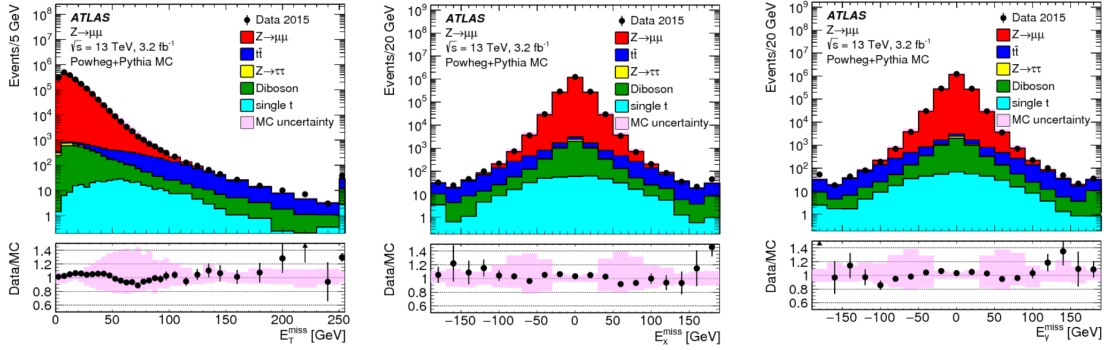


Figure 4.1: Distributions of E_T^{miss} , E_x^{miss} and E_y^{miss} for sample with inclusive jets of $Z \rightarrow \mu\mu$ extracted from data with respect to the MC simulations including all relevant backgrounds. The last bin of each distribution includes the overflow, and the first bin contains the underflow in 1 (b) and 1 (c). The respective ratios between data and expectation are shown in the lower panel with the shaded areas indicating the total uncertainty for MC simulations. Figures from Ref [28].

4.3 Jet selection

4.3.1 EMTopo Jet

Jets are built using three dimensional topological clusters so-called (topo-clusters) [65] from the amount of energy deposits in the calorimeter using the *anti* - k_t algorithm [67] with a distance parameter $R = 0.4$. These EMTopo jets need to satisfy $p_T > 20 \text{ GeV}$ after the full calibration, then, to reduce the pileup jets contribution and select hard-scatter jets, the jet vertex tagging (JVT) algorithm is used to provide a cut on the JVT¹ discriminant

¹The JVT score ranges from 0 (pileup-like) to 1 (hard-scatter-like).

requirement: $JVT > 0.59$ for jets with $p_T < 60\text{GeV}$ and $|\eta| < 2.4$. Moreover, the forward-jet-vertex tagger (fJVT) approach is used to reject additional jets originating from pileup with $|\eta| > 2.5$ and $20 < p_T < 50\text{GeV}$

4.3.2 PFlow Jet

In order to enhance the E_T^{miss} calculation of each event, PFlow jets are reconstructed from a combination of high quality signals from the the Calorimeter and the Inner Detector instead of considering only track based corrections in the jet calibration.

The PFlow algorithm described in detail in [81], avoid energy deposits in calorimeter originating from pile-up physics object and consider the momentum calculation from tracks every time the tracker granularity is improved compared to the calorimeter resolution using *anti* - k_t algorithm with a distance parameter $R = 0.4$. JVT cut is applied on PFlow jets by requiring this parameter to be larger than 0.2 for jet with $p_T < 60\text{GeV}$ and $|\eta| < 2.4$ to ensure including only jets matched to hard-scatter vertex.

4.4 Monte Carlo Samples

All samples of simulated events are passed through the ATLAS detector simulation bades on Geant4 [82] and are reconstructed with the standard ATLAS reconstruction softwar. The effects of pileup from multiple interactions in the same and nearby bunch crossings are modelled by overlaying minimum-bias events, simulated using the soft QCD of pythia8.186 [83] with the A2 [84] set of tuned parameters (tune). The EvtGen v1.2.0 program [85] is used to describe the decays of bottom and charm hadrons. Simulation of Z bosons decaying to a pair of opposite sign leptons (e) is produced by Sherpa2.2.1 [86], and the simulation is corrected to the NNLO order cross-section [87]. The parton showering and hadronization are performed with Sherpa2.2.1 [88]. A summary of all the generators used for the simulation of the signal and background processes is shown in Table 4.2.

Sample	$Z \rightarrow ee$	$t\bar{t} + \gamma$
Generator	Sherpa 2.2.1	MadGraph Pythia 8 EvtGen
ME PDF set	NNPDF3.0NNLO	A14NNPDF23
PS and Hadronization	Sherpa 2.2.1	MadGraph Pythia 8 EvtGen
UE Model TUNE	Default	p3663

Table 4.2: PDF sets, Generators, and MC tunes employed in the OR (photon/electron, jet) analysis for the E_T^{miss} .

$Z \rightarrow ee$ event selection criteria

The $Z \rightarrow ee$ event topology is particularly convenient to study the fake E_T^{miss} effects. The leptonic decay of the Z boson is not only extremely easy to trigger but also abundantly produced in pp collision and well modeled in the current ATLAS simulation. The selection criteria for $Z \rightarrow ee$ event are as follows:

- The events must contain two opposite-sign electrons that are the same flavor, satisfying "Medium" identification, isolated with "Loose" working point, and matched to the hard-scatter primary vertex (PV). The leading electron in the pair is asked to have $p_T > 30$ GeV and the sub-leading one $p_T > 20$ GeV.
- No additional electrons with $p_T > 7$ GeV and satisfying "Loose" PID are identified.
- The chosen di-electrons invariant mass must fall between a range of 25 GeV: $|m_{ee} - m_Z| < 25$ GeV

$t\bar{t} + \gamma$ event selection criteria

The $t\bar{t} + \gamma$ process in the semi-leptonic final states identified by the presence of a photon candidate and the decay products of a pair of top quarks, where momentum imbalance is expected due to the presence of a neutrino (ν) resulting from the leptonically decay of W boson, which allow this topology to be a good candidate to study the gamma-jet overlap removal.

- select events with exactly 1 electron or exactly 1 muon with $p_T > 30$ GeV
- at least 4 jets, without any b-tagging requirements.
- require at least 1 photon with $p_T > 25$ GeV.

4.5 Overlap Removal

Physical objects considered in the calculation of E_T^{miss} and $\sum E_T$ must be reconstructed from exclusive signals deposited in the different detector sub-systems to avoid double counting of the same signal. Hence, the energy deposits in the detector should be precisely associated to the jets or particles that produced them. When hard objects are created near to each other, they might share or overlap a quantity of energy deposits which need to be correctly treated. Next section provides a treatment of $e-jet$ and $\gamma-jet$ overlap removal for two type of jet collections: EMTopo and PFlow jets. Figure 4.2 provides the general plan of overlap

handling in the measurement of E_T^{miss} . Once the physical objects are designed, any jet that is near to the high-priority object is rejected from the event based on the ΔR matching criteria (The approach tests for overlap based on a minimal distance ΔR between pairs of objects for a collection of baseline items).

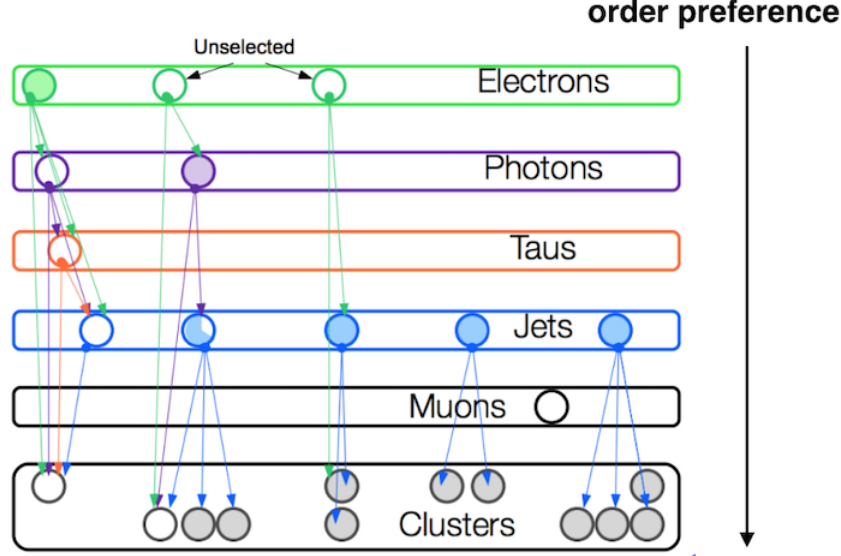


Figure 4.2: Matching approach employed in the E_T^{miss} map to remove duplicates. Each circle refers to a cluster. The solid circle in purple, green, and red means the cluster is associated with the object. The cluster is then removed if it is matched with an object, otherwise the cluster is recognized as a jet in blue. The lines connect clusters and objects through a ΔR matching [29].

4.6 Jets and $e(\gamma)$ overlap removal

It is common for hadronic jets to decay in ways like $\pi^0 \rightarrow \gamma\gamma$, depositing a portion of their energy in the electromagnetic calorimeter. As discussed earlier, the informations we get from both hadronic and electromagnetic calorimeters are employed to evaluate the jet four-momentum with the best accurate possible. All electrons passing the selections cuts described in Section 4.2.3 are computed in the E_T^{miss} reconstruction as the highest priority signature in the calorimeter. The lower-priority reconstructed particles (γ , τ -lepton, and jets added in order of priority) are totally removed whether they share their calorimeter signal (e.g. topoclusters are ΔR matched to the hard objects) with order already considered in the E_T^{miss} reconstruction. Conventionally, jets are excluded whenever they overlap with accepted higher-priority particles. Hence, the jet contribution to E_T^{miss} is dependent on

the signal ambiguity resolution. Therefore, fake missing transverse momentum can appear in scenarios where electrons or photons are close to jets ($\Delta R(e(\gamma), jet) < 0.4$). Some of overlapping electrons/photons and jets configurations are shown in Figure 4.3. As it is illustrated, jets from electrons/photons that are faking jets in the calorimeter, in addition to the fake electron/photon that contaminates the real jets, lead to the jets mis-calibration and events double counting in the calorimeter.

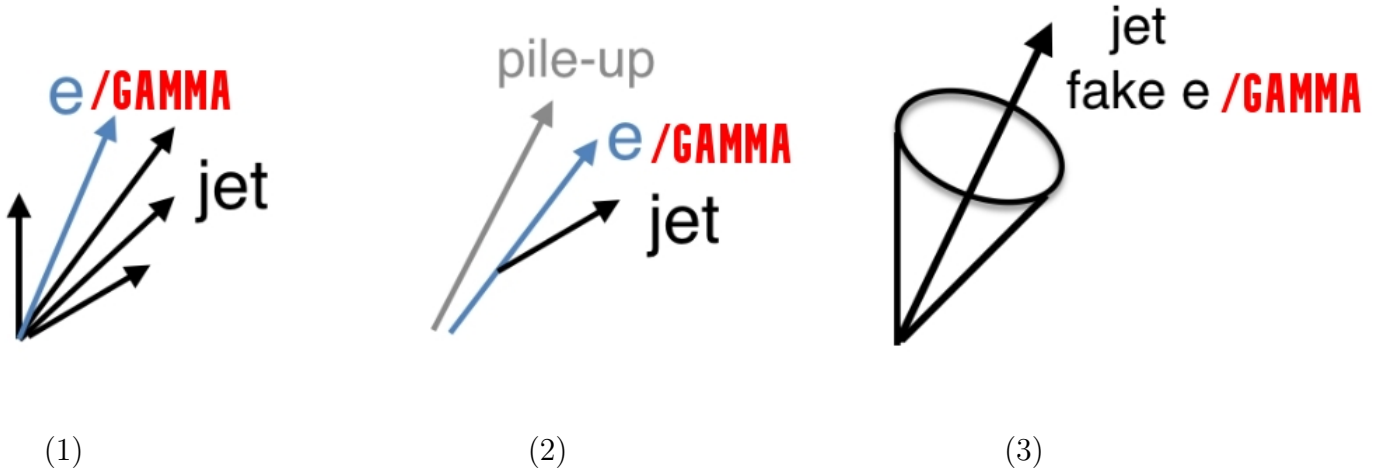


Figure 4.3: Overlapping electrons/photons and jets. (1) Real jet close to real electron/photon. (2) Jets from pile-up or electrons/photons faking jets in the calorimeter. (3) Real jet and fake electron/photon.

As in E_T^{miss} definition [73], in order to estimate the real jet energy signature that was handled as a real jet and assigned to the E_T^{miss} jet term, a so-called variable, $Ratio_E(jet, e(\gamma))$ was used, and it is calculated as the difference between the jet energy, and the electron (photon) energy to the jet energy calibrated at the electromagnetic scale:

$$Ratio_E(jet, e(\gamma)) = \frac{E_{jet}^{EM} - E_{e(\gamma)}^{EM}}{E_{jet}^{EM}} \quad (4.7)$$

This $Ratio_E(jet, e(\gamma))$ is used to substrate the fraction of the electron (photon) energy in a given jet, and decide for the jet's energy quantity that has to be computed in the E_T^{miss} reconstruction. Figures 4.4 and 4.5 (right-hand) show the distributions of the $Ratio_E(jet, e(\gamma))$ when the reconstructed jet is associated with the truth particle jet using $Z \rightarrow ee$ and $t\bar{t} + \gamma$ simulation data processes. Both figures also illustrate the ratio distribution when the matching criterion to the truth-particle jets is not considered. The matching criteria is based

on requiring an angular cut distance of $\Delta R(jet_{reco}, jet_{truth}) < 0.3$, where ΔR is the pseudorapidity and azimuthal angle separation between the two jets added in quadrature, i.e. $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$. The angular criterion $\Delta R < 0.3$ is chosen to be smaller than the jet radius parameter used for ATLAS analyses ($R = 0.4$ or larger) but much larger than the jet angular resolution. As it is expected the matching to the truth particles jets improves the selection efficiency. The behavior of the ratio spectrum, when $Ratio_E(jet, e(\gamma)) \geq 0.5$ might be explained by the presence of a higher p_T jet near an electron (photon). In this case, the jet is retained in E_T^{miss} reconstruction with its p_T scaled by the $Ratio_E(jet, e(\gamma))$ variable. While, in the case where $Ratio_E(jet, e(\gamma)) < 0.5$ the jet was associated to the soft term, and only the tracks associated with the jet is included in the E_T^{miss} computation. However, this approach leads to underestimating results; when both jet and electron (photon or hadronically decaying τ -lepton) are real and overlapping, since the soft term in this study is track-based only and does not account for neutral particles. Neutral energy is lost and manifests as spurious E_T^{miss} . To reduce the number of real jets that are wrongly matched to the soft term and hence improve the E_T^{miss} calculation, the $\Delta p_T(jet, e(\gamma))$ (Equation 4.8) discriminant variable is defined as the p_T difference at the electromagnetic scale of the jet and electron (photon). Figures 4.4 and 4.5 (left-hand) show Δp_T distributions for the truth-matched jets (blue) and not matching truth-particles jets (red) in $Z \rightarrow ee$ and $t\bar{t} + \gamma$ simulated events, respectively. A better performance is obtained when matching truth-particle jet is considered.

$$\Delta p_T(jet, e(\gamma)) = p_T^{EM, jet} - p_T^{EM, e(\gamma)} \quad (4.8)$$

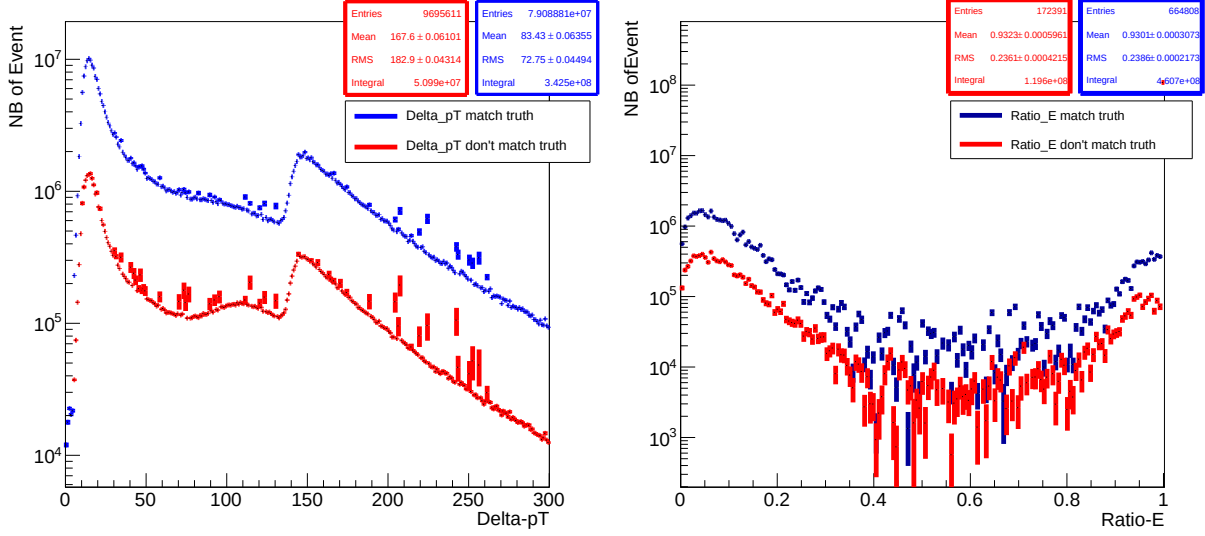


Figure 4.4: Comparison of $\Delta p_T(e, jet)$ (GeV) and $Ratio_E(e, jet)$ performance for the jets matching the truth (blue) and not matching the truth (red) using $Z \rightarrow ee$ simulated events.

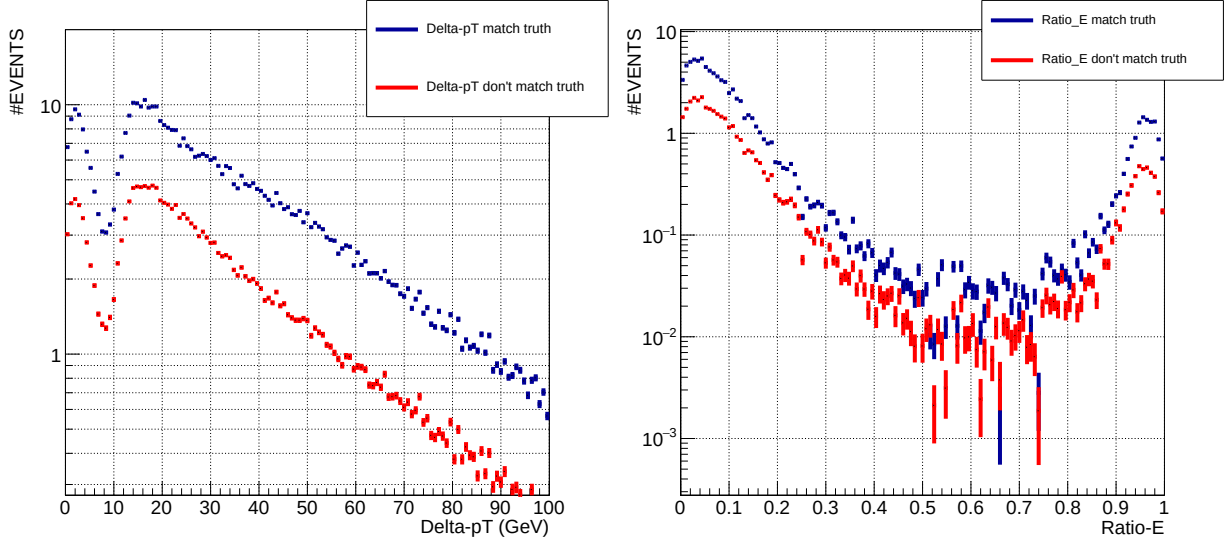


Figure 4.5: Comparison of $\Delta p_T(\gamma, jet)$ (GeV) and $Ratio_E(\gamma, jet)$ performance for the jets matching the truth (blue) and not matching the truth (red) using $t\bar{t} + \gamma$ simulated events.

$e - jet$ Overlap Removal

To avoid labeling the same detector signature as more than one object, an overlap removal

(OR) procedure is applied. The overlap between electrons and jets generally happen when they deposit their energy in the same clusters which might be considered twice as duplicated energy. This situation leads to fake E_T^{miss} and results in large E_T^{miss} tail. Therefore, an optimization of the electron-jet OR is performed using $Z \rightarrow ee$ simulated events. To reduce the fake E_T^{miss} tail, a scan of the $Ratio_E(e, jet)$ and $\Delta p_T(e, jet)$ variables have been studied. The correlation between both variables has been evaluated in Figure 4.6. They show a low dependency which motivate their simultaneous employment to handle the electron-jet OR. Tables 4.3 and 4.4 summarize the tuning values that have been considered for $Ratio_E(e, jet)$ and $\Delta p_T(e, jet)$ for EMTopo and PFlow jets collections, respectively. The E_T^{miss} mean and the Root Mean Square (RMS) resulting values are listed as well. To improve the E_T^{miss} tail performance, the integral of the E_T^{miss} over the GeV ranges of [100,1000] and [200,1000] have been estimated.

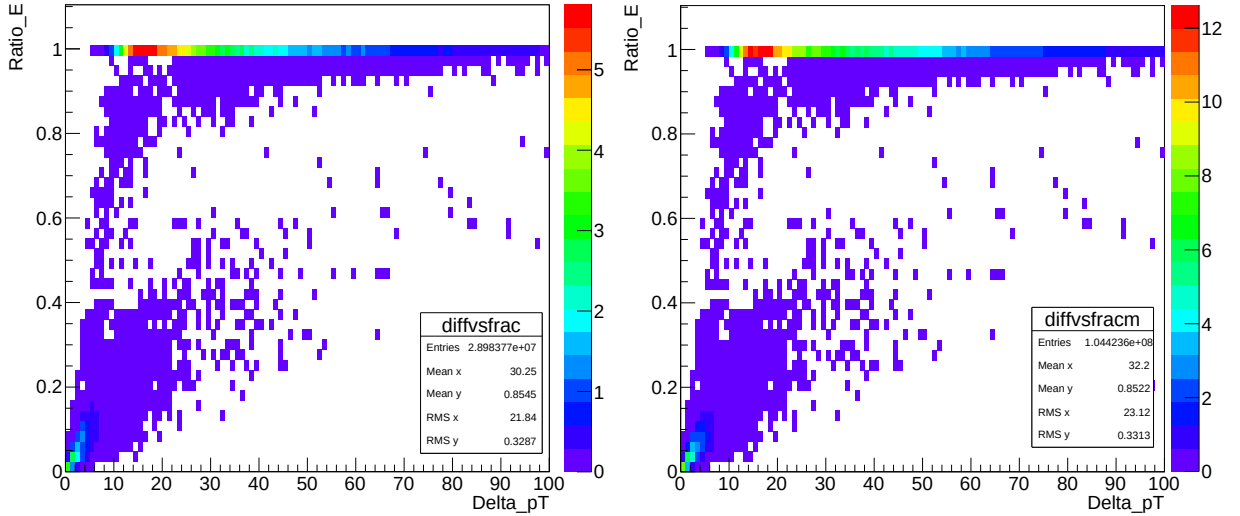


Figure 4.6: $\Delta p_T(e, jet)$ (GeV) as a function of $Ratio_E(e, jet)$ for the jets matching the truth (right-hand) and not matching the truth (left-hand) for events from $Z \rightarrow ee$ process.

Table 4.5 presents the combination of the tuned values for $Ratio_E(e, jet)$ and $\Delta p_T(e, jet)$ using EMTopo jets collection, including E_T^{miss} integral over the GeV range of [150,1000]. This optimization reduces only 1% of the E_T^{miss} tail. Figure 4.7 shows the resulting E_T^{miss} performance. The same optimization was done for PFlow jets collection as can be seen in Table 4.6. As a result, the improvement of the E_T^{miss} performance is about 6% (Figure 4.8).

CHAPTER 4. OVERLAP REMOVAL AND PERFORMANCE OF E_T RECONSTRUCTION

$Ratio_E > 0.5$	$\Delta p_T > 10 \text{ GeV}$	$\Delta p_T > 20 \text{ GeV}$	$\Delta p_T > 30 \text{ GeV}$
$\langle E_T^{miss} \rangle (\text{GeV})$	28.03	28.03	28.03
$E_T^{miss} \text{ RMS}$	21.08	21.08	21.08
$\int_{100}^{1000} E_T^{miss} (\text{GeV})$	$6.60 \cdot 10^{-04}$	$6.60 \cdot 10^{-04}$	$6.60 \cdot 10^{-04}$
$\int_{200}^{1000} E_T^{miss} (\text{GeV})$	$5.62 \cdot 10^{-06}$	$5.62 \cdot 10^{-06}$	$5.62 \cdot 10^{-06}$
$Ratio_E > 0.2$	$\Delta p_T > 10 \text{ GeV}$	$\Delta p_T > 20 \text{ GeV}$	$\Delta p_T > 30 \text{ GeV}$
$\langle E_T^{miss} \rangle (\text{GeV})$	28.41	28.13	28.02
$E_T^{miss} \text{ RMS}$	21.15	21.11	21.07
$\int_{100}^{1000} E_T^{miss} (\text{GeV})$	$6.80 \cdot 10^{-04}$	$6.90 \cdot 10^{-04}$	$6.60 \cdot 10^{-04}$
$\int_{200}^{1000} E_T^{miss} (\text{GeV})$	$5.55 \cdot 10^{-06}$	$5.48 \cdot 10^{-06}$	$5.47 \cdot 10^{-06}$
$Ratio_E > 0$	$\Delta p_T > 10 \text{ GeV}$	$\Delta p_T > 20 \text{ GeV}$	$\Delta p_T > 30 \text{ GeV}$
$\langle E_T^{miss} \rangle (\text{GeV})$	28.90	28.18	28.03
$E_T^{miss} \text{ RMS}$	21.46	21.18	21.09
$\int_{100}^{1000} E_T^{miss} (\text{GeV})$	$7.10 \cdot 10^{-04}$	$6.50 \cdot 10^{-04}$	$6.60 \cdot 10^{-04}$
$\int_{200}^{1000} E_T^{miss} (\text{GeV})$	$5.62 \cdot 10^{-06}$	$5.41 \cdot 10^{-06}$	$5.45 \cdot 10^{-06}$

Table 4.3: E_T^{miss} performance for $Z \rightarrow ee$ events using EMTopo jets collection. The jet is treated as a real jet if $\Delta p_T(e, jet)$ and $Ratio_E(e, jet)$ are larger than a tuned threshold. The mean and RMS are relatively unaffected.

CHAPTER 4. OVERLAP REMOVAL AND PERFORMANCE OF E_T RECONSTRUCTION

$(Ratio_E > 0.5)$	$\Delta p_T > 10 \text{ GeV}$	$\Delta p_T > 20 \text{ GeV}$	$\Delta p_T > 30 \text{ GeV}$
$\langle E_T^{miss} \rangle (\text{GeV})$	27.12	27.12	27.12
$E_T^{miss} \text{ RMS}$	20.59	20.59	20.59
$\int_{100}^{1000} E_T^{miss} (\text{GeV})$	$6.00 \cdot 10^{-04}$	$6.00 \cdot 10^{-04}$	$6.00 \cdot 10^{-04}$
$\int_{200}^{1000} E_T^{miss} (\text{GeV})$	$7.99 \cdot 10^{-06}$	$7.99 \cdot 10^{-06}$	$7.99 \cdot 10^{-06}$
$(Ratio_E > 0.2)$	$\Delta p_T > 10 \text{ GeV}$	$\Delta p_T > 20 \text{ GeV}$	$\Delta p_T > 30 \text{ GeV}$
$\langle E_T^{miss} \rangle (\text{GeV})$	27.18	27.06	27.05
$E_T^{miss} \text{ RMS}$	21.46	20.39	20.38
$\int_{100}^{1000} E_T^{miss} (\text{GeV})$	$6.80 \cdot 10^{-04}$	$6.00 \cdot 10^{-04}$	$5.00 \cdot 10^{-04}$
$\int_{200}^{1000} E_T^{miss} (\text{GeV})$	$7.75 \cdot 10^{-06}$	$7.71 \cdot 10^{-06}$	$7.71 \cdot 10^{-06}$
$Ratio_E > 0$	$\Delta p_T > 10 \text{ GeV}$	$\Delta p_T > 20 \text{ GeV}$	$\Delta p_T > 30 \text{ GeV}$
$\langle E_T^{miss} \rangle (\text{GeV})$	27.18	27.06	27.05
$E_T^{miss} \text{ RMS}$	20.42	20.39	20.38
$\int_{100}^{1000} E_T^{miss} (\text{GeV})$	$6.80 \cdot 10^{-04}$	$5.00 \cdot 10^{-04}$	$5.00 \cdot 10^{-04}$
$\int_{200}^{1000} E_T^{miss} (\text{GeV})$	$7.82 \cdot 10^{-06}$	$7.70 \cdot 10^{-06}$	$7.72 \cdot 10^{-06}$

Table 4.4: E_T^{miss} performance for $Z \rightarrow ee$ events using PFlow jets collection. The jet is treated as a real jet if $\Delta p_T(e, jet)$ and $Ratio_E(e, jet)$ are larger than a tuned threshold. The mean and RMS are relatively unaffected.

	$Ratio_E > 0 \text{ and } \Delta p_T > 20 \text{ GeV}$	$Ratio_E > 0.5 \text{ and } \Delta p_T > 0 \text{ GeV}$
$\int_{150}^{1000} E_T^{miss} (\text{GeV})$	$1.39 \cdot 10^{-05} \pm 6.31 \cdot 10^{-07}$	$1.40 \cdot 10^{-05} \pm 6.14 \cdot 10^{-07}$

Table 4.5: E_T^{miss} performance for $Z \rightarrow ee$ simulated events using EMTopo jets collection.

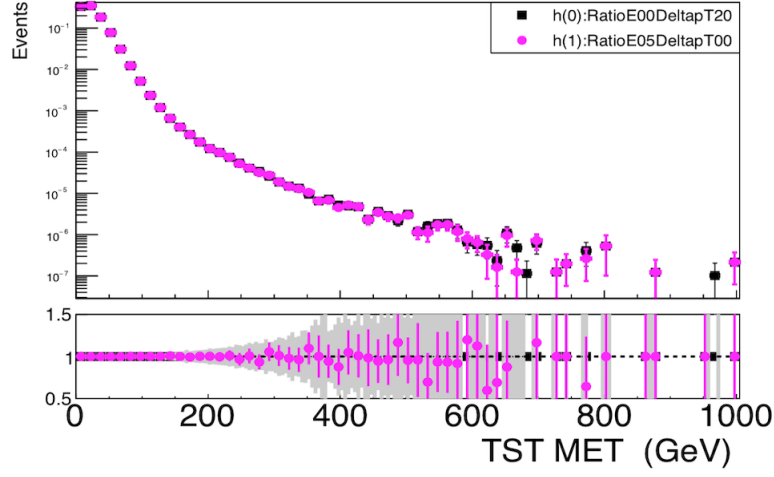


Figure 4.7: Comparison of the E_T^{miss} (TST MET means track-based algorithm is used for the soft term) performance between old (magenta) and new (black) jet-electron overlap removal (OR) procedure in $Z \rightarrow ee$ simulation. The E_T^{miss} tail has been reduced using EMTopo jets collection.

	$Ratio_E > 0 \text{ and } \Delta p_T > 20 \text{ GeV}$	$Ratio_E > 0.5 \text{ and } \Delta p_T > 0 \text{ GeV}$
$\int_{150}^{1000} E_T^{miss} \text{ (GeV)}$	$1.83 \cdot 10^{-05} \pm 8.53 \cdot 10^{-07}$	$1.96 \cdot 10^{-05} \pm 8.70 \cdot 10^{-07}$

Table 4.6: E_T^{miss} performance for $Z \rightarrow ee$ simulated events using PFlow jets collection.

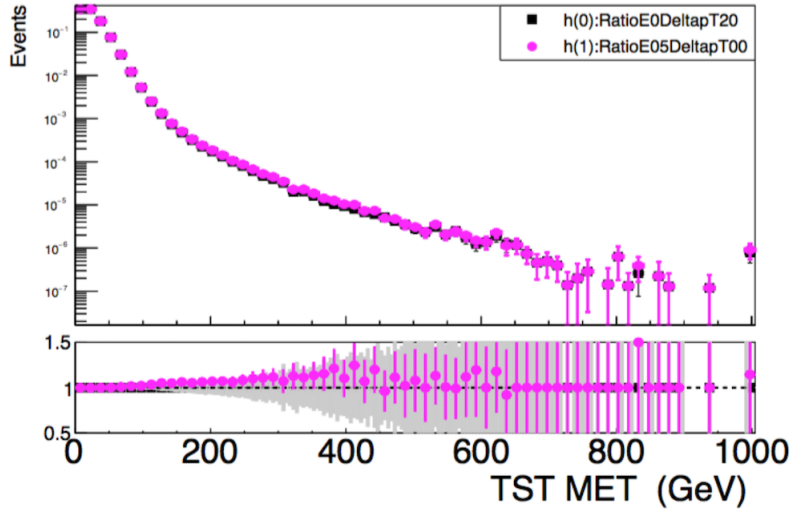


Figure 4.8: Comparison of the E_T^{miss} (TST MET means track-based algorithm is used for the soft term) performance between old (magenta) and new (black) jet-electron overlap removal procedure in $Z \rightarrow ee$ simulation. The E_T^{miss} tail has been reduced using PFlow jets collection.

γ - jet Overlap Removal

As discussed for the electrons and jets, an overlap removal procedure is carried out to handle cases where photons and jets deposit their energy in the same clusters. Such scenarios give rise to a large fake E_T^{miss} since double counting of the same energy can happen. To optimize the treatment of assigning energy deposits to the photon or the jet that generated them, a tuning of photon-jet OR is performed using simulated $t\bar{t} + \gamma$ events where a significant real E_T^{miss} is expected, arising from the presence of neutrino. To reduce the fake E_T^{miss} tail defined as: $|Reconstructed - Truth| E_T^{miss}$, a scan of the $Ratio_E(\gamma, jet)$ and $\Delta p_T(\gamma, jet)$ variables have been done. Figure 4.9 shows that these discriminating parameters are uncorrelated. Hence, they are relevant to study the photon-jet OR. Tables 4.7 and 4.8 recapitulate the scanning values that have been considered for $Ratio_E(\gamma, jet)$ and $\Delta p_T(\gamma, jet)$ for EMTopo and PFlow jets collections, respectively. The E_T^{miss} mean and the RMS resulting values are listed as well. To improve the fake E_T^{miss} tail performance, the integral of the $|Reconstructed - Truth| E_T^{miss}$ over the GeV ranges [50,500] and [100,500] have been estimated.

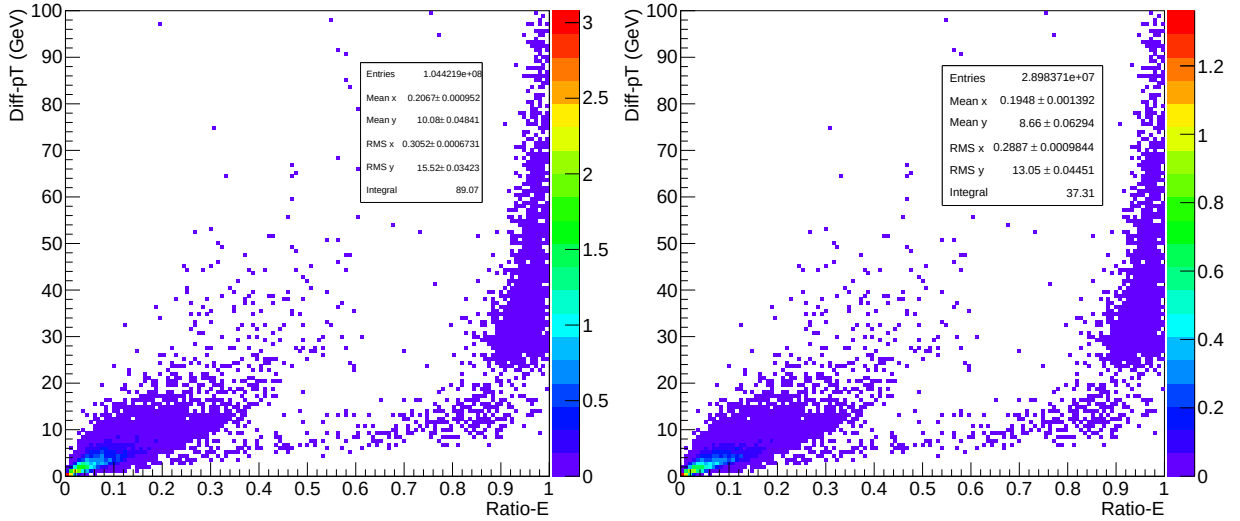


Figure 4.9: $\Delta p_T(\gamma, jet)$ (GeV) as a function of $Ratio_E(\gamma, jet)$ for the jets matching the truth (left-hand) and not matching the truth (right-hand) for events from $t\bar{t} + \gamma$ process.

Table 4.9 illustrates the combination of the tuned values for $Ratio_E(\gamma, jet)$ and $\Delta p_T(\gamma, jet)$ using EMTopo jets collection, including the $|Reconstructed - Truth| E_T^{miss}$ integral over the

GeV range of $[150, 500]$. This optimization decreases the fake E_T^{miss} tail of about 19%. Figure 4.10 presents the resulting fake E_T^{miss} performance. Similar optimization was done for PFlow jets collection as can be seen in Table 4.10. In consequence, the enhancement of the $|Reconstructed - Truth|$ E_T^{miss} performance is about 11% (Figure 4.11).

CHAPTER 4. OVERLAP REMOVAL AND PERFORMANCE OF E_T RECONSTRUCTION

	$Ratio_E > 0$	$\Delta p_T > 0 \text{ GeV}$	$\Delta p_T > 10 \text{ GeV}$	$\Delta p_T > 20 \text{ GeV}$	$\Delta p_T > 30 \text{ GeV}$
$\langle E_T^{miss} \rangle (\text{GeV})$		68.31	68.10	67.99	67.99
$E_T^{miss} \text{ RMS}$		42.13	41.96	41.87	41.84
$\int_{100}^{500} reco - truth E_T^{miss} (\text{GeV})$		$2.31 \cdot 10^{-3}$	$2.30 \cdot 10^{-3}$	$2.27 \cdot 10^{-3}$	$2.25 \cdot 10^{-3}$
$\int_{50}^{500} reco - truth E_T^{miss} (\text{GeV})$		$1.26 \cdot 10^{-1}$	$1.25 \cdot 10^{-1}$	$1.23 \cdot 10^{-1}$	$1.23 \cdot 10^{-1}$
	$Ratio_E > 0.1$	$\Delta p_T > 0 \text{ GeV}$	$\Delta p_T > 10 \text{ GeV}$	$\Delta p_T > 20 \text{ GeV}$	$\Delta p_T > 30 \text{ GeV}$
$\langle E_T^{miss} \rangle (\text{GeV})$		68.16	68.06	67.99	67.99
$E_T^{miss} \text{ RMS}$		41.93	41.91	41.87	41.84
$\int_{100}^{500} reco - truth E_T^{miss} (\text{GeV})$		$2.27 \cdot 10^{-3}$	$2.28 \cdot 10^{-3}$	$2.27 \cdot 10^{-3}$	$2.25 \cdot 10^{-3}$
$\int_{50}^{500} reco - truth E_T^{miss} (\text{GeV})$		$1.25 \cdot 10^{-1}$	$1.25 \cdot 10^{-1}$	$1.23 \cdot 10^{-1}$	$1.23 \cdot 10^{-1}$
	$Ratio_E > 0.2$	$\Delta p_T > 0 \text{ GeV}$	$\Delta p_T > 10 \text{ GeV}$	$\Delta p_T > 20 \text{ GeV}$	$\Delta p_T > 30 \text{ GeV}$
$\langle E_T^{miss} \rangle (\text{GeV})$		68.03	68.02	67.99	67.99
$E_T^{miss} \text{ RMS}$		41.86	41.88	41.85	41.84
$\int_{100}^{500} reco - truth E_T^{miss} (\text{GeV})$		$2.28 \cdot 10^{-3}$	$2.28 \cdot 10^{-3}$	$2.26 \cdot 10^{-3}$	$2.25 \cdot 10^{-3}$
$\int_{50}^{500} reco - truth E_T^{miss} (\text{GeV})$		$1.24 \cdot 10^{-1}$	$1.24 \cdot 10^{-1}$	$1.23 \cdot 10^{-1}$	$1.23 \cdot 10^{-1}$
	$Ratio_E > 0.3$	$\Delta p_T > 0 \text{ GeV}$	$\Delta p_T > 10 \text{ GeV}$	$\Delta p_T > 20 \text{ GeV}$	$\Delta p_T > 30 \text{ GeV}$
$\langle E_T^{miss} \rangle (\text{GeV})$		67.98	67.98	67.98	67.98
$E_T^{miss} \text{ RMS}$		41.86	41.86	41.86	41.85
$\int_{100}^{500} reco - truth E_T^{miss} (\text{GeV})$		$2.25 \cdot 10^{-3}$	$2.25 \cdot 10^{-3}$	$2.25 \cdot 10^{-3}$	$2.25 \cdot 10^{-3}$
$\int_{50}^{500} reco - truth E_T^{miss} (\text{GeV})$		$1.24 \cdot 10^{-1}$	$1.24 \cdot 10^{-1}$	$1.24 \cdot 10^{-1}$	$1.24 \cdot 10^{-1}$
	$Ratio_E > 0.4$	$\Delta p_T > 0 \text{ GeV}$	$\Delta p_T > 10 \text{ GeV}$	$\Delta p_T > 20 \text{ GeV}$	$\Delta p_T > 30 \text{ GeV}$
$\langle E_T^{miss} \rangle (\text{GeV})$		67.98	67.98	67.98	67.98
$E_T^{miss} \text{ RMS}$		41.86	41.86	41.86	41.85
$\int_{100}^{500} reco - truth E_T^{miss} (\text{GeV})$		$2.25 \cdot 10^{-3}$	$2.25 \cdot 10^{-3}$	$2.25 \cdot 10^{-3}$	$2.25 \cdot 10^{-3}$
$\int_{50}^{500} reco - truth E_T^{miss} (\text{GeV})$		$1.24 \cdot 10^{-1}$	$1.24 \cdot 10^{-1}$	$1.24 \cdot 10^{-1}$	$1.24 \cdot 10^{-1}$
	$Ratio_E > 0.5$	$\Delta p_T > 0 \text{ GeV}$	$\Delta p_T > 10 \text{ GeV}$	$\Delta p_T > 20 \text{ GeV}$	$\Delta p_T > 30 \text{ GeV}$
$\langle E_T^{miss} \rangle (\text{GeV})$		67.99	67.99	67.99	67.99
$E_T^{miss} \text{ RMS}$		41.84	41.84	41.84	41.84
$\int_{100}^{500} reco - truth E_T^{miss} (\text{GeV})$		$2.28 \cdot 10^{-3}$	$2.28 \cdot 10^{-3}$	$2.28 \cdot 10^{-3}$	$2.28 \cdot 10^{-3}$
$\int_{50}^{500} reco - truth E_T^{miss} (\text{GeV})$		$1.26 \cdot 10^{-1}$	$1.26 \cdot 10^{-1}$	$1.26 \cdot 10^{-1}$	$1.26 \cdot 10^{-1}$

Table 4.7: E_T^{miss} performance for $t\bar{t} + \gamma$ events using EMTopo jets collection. The jet is treated as a real jet if $\Delta p_T(\gamma, jet)$ and $Ratio_E(\gamma, jet)$ are larger than a tuned threshold. The mean and RMS are relatively unaffected.

CHAPTER 4. OVERLAP REMOVAL AND PERFORMANCE OF E_T RECONSTRUCTION

	$Ratio_E > 0$	$\Delta p_T > 0 \text{ GeV}$	$\Delta p_T > 10 \text{ GeV}$	$\Delta p_T > 20 \text{ GeV}$	$\Delta p_T > 30 \text{ GeV}$
$\langle E_T^{miss} \rangle (\text{GeV})$	73.13	72.75	72.70	72.70	
$E_T^{miss} \text{ RMS}$	52.37	52.36	52.36	52.37	
$\int_{100}^{500} reco - truth E_T^{miss} (\text{GeV})$	$3.70 \cdot 10^{-3}$	$3.70 \cdot 10^{-3}$	$3.60 \cdot 10^{-3}$	$3.70 \cdot 10^{-3}$	
$\int_{50}^{500} reco - truth E_T^{miss} (\text{GeV})$	$1.3 \cdot 10^{-1}$	$1.28 \cdot 10^{-1}$	$1.28 \cdot 10^{-1}$	$1.29 \cdot 10^{-1}$	
	$Ratio_E > 0.1$	$\Delta p_T > 0 \text{ GeV}$	$\Delta p_T > 10 \text{ GeV}$	$\Delta p_T > 20 \text{ GeV}$	$\Delta p_T > 30 \text{ GeV}$
$\langle E_T^{miss} \rangle (\text{GeV})$	72.81	72.72	72.71	72.71	
$E_T^{miss} \text{ RMS}$	52.28	52.33	52.36	52.37	
$\int_{100}^{500} reco - truth E_T^{miss} (\text{GeV})$	$3.70 \cdot 10^{-3}$	$3.70 \cdot 10^{-3}$	$3.60 \cdot 10^{-3}$	$3.70 \cdot 10^{-3}$	
$\int_{50}^{500} reco - truth E_T^{miss} (\text{GeV})$	$1.29 \cdot 10^{-1}$	$1.28 \cdot 10^{-1}$	$1.28 \cdot 10^{-1}$	$1.29 \cdot 10^{-1}$	
	$Ratio_E > 0.2$	$\Delta p_T > 0 \text{ GeV}$	$\Delta p_T > 10 \text{ GeV}$	$\Delta p_T > 20 \text{ GeV}$	$\Delta p_T > 30 \text{ GeV}$
$\langle E_T^{miss} \rangle (\text{GeV})$	72.73	72.72	72.70	72.70	
$E_T^{miss} \text{ RMS}$	52.34	52.35	52.36	52.37	
$\int_{100}^{500} reco - truth E_T^{miss} (\text{GeV})$	$3.70 \cdot 10^{-3}$	$3.70 \cdot 10^{-3}$	$3.70 \cdot 10^{-3}$	$3.70 \cdot 10^{-3}$	
$\int_{50}^{500} reco - truth E_T^{miss} (\text{GeV})$	$1.29 \cdot 10^{-1}$	$1.29 \cdot 10^{-1}$	$1.29 \cdot 10^{-1}$	$1.29 \cdot 10^{-1}$	
	$Ratio_E > 0.3$	$\Delta p_T > 0 \text{ GeV}$	$\Delta p_T > 10 \text{ GeV}$	$\Delta p_T > 20 \text{ GeV}$	$\Delta p_T > 30 \text{ GeV}$
$\langle E_T^{miss} \rangle (\text{GeV})$	72.72	72.72	72.71	72.71	
$E_T^{miss} \text{ RMS}$	52.38	52.38	52.38	52.39	
$\int_{100}^{500} reco - truth E_T^{miss} (\text{GeV})$	$3.70 \cdot 10^{-3}$	$3.70 \cdot 10^{-3}$	$3.70 \cdot 10^{-3}$	$3.70 \cdot 10^{-3}$	
$\int_{50}^{500} reco - truth E_T^{miss} (\text{GeV})$	$1.29 \cdot 10^{-1}$	$1.29 \cdot 10^{-1}$	$1.29 \cdot 10^{-1}$	$1.30 \cdot 10^{-1}$	
	$Ratio_E > 0.4$	$\Delta p_T > 0 \text{ GeV}$	$\Delta p_T > 10 \text{ GeV}$	$\Delta p_T > 20 \text{ GeV}$	$\Delta p_T > 30 \text{ GeV}$
$\langle E_T^{miss} \rangle (\text{GeV})$	72.73	72.73	72.73	72.73	
$E_T^{miss} \text{ RMS}$	52.41	52.41	52.41	52.41	
$\int_{100}^{500} reco - truth E_T^{miss} (\text{GeV})$	$3.80 \cdot 10^{-3}$	$3.80 \cdot 10^{-3}$	$3.80 \cdot 10^{-3}$	$3.80 \cdot 10^{-3}$	
$\int_{50}^{500} reco - truth E_T^{miss} (\text{GeV})$	$1.31 \cdot 10^{-1}$	$1.31 \cdot 10^{-1}$	$1.31 \cdot 10^{-1}$	$1.31 \cdot 10^{-1}$	
	$Ratio_E > 0.5$	$\Delta p_T > 0 \text{ GeV}$	$\Delta p_T > 10 \text{ GeV}$	$\Delta p_T > 20 \text{ GeV}$	$\Delta p_T > 30 \text{ GeV}$
$\langle E_T^{miss} \rangle (\text{GeV})$	72.75	72.75	72.75	72.76	
$E_T^{miss} \text{ RMS}$	52.47	52.47	52.47	52.47	
$\int_{100}^{500} reco - truth E_T^{miss} (\text{GeV})$	$3.70 \cdot 10^{-3}$	$3.70 \cdot 10^{-3}$	$3.70 \cdot 10^{-3}$	$3.70 \cdot 10^{-3}$	
$\int_{50}^{500} reco - truth E_T^{miss} (\text{GeV})$	$1.32 \cdot 10^{-1}$	$1.32 \cdot 10^{-1}$	$1.32 \cdot 10^{-1}$	$1.32 \cdot 10^{-1}$	

Table 4.8: E_T^{miss} performance for $t\bar{t} + \gamma$ events using PFLow jets collection. The jet is treated as a real jet if $\Delta p_T(\gamma, jet)$ and $Ratio_E(\gamma, jet)$ are larger than a tuned threshold. The mean and RMS are relatively unaffected.

CHAPTER 4. OVERLAP REMOVAL AND PERFORMANCE OF E_T RECONSTRUCTION

	$Ratio_E > 0.5$ and $\Delta p_T > 0$ GeV	$Ratio_E > 0$ and $\Delta p_T > 0$ GeV
$\int_{150}^{500} reco - truth E_T^{miss}$ (GeV)	$2.20 \cdot 10^{-4} \pm 29.35 \cdot 10^{-05}$	$2.70 \cdot 10^{-4} \pm 29.79 \cdot 10^{-05}$

Table 4.9: E_T^{miss} performance for $t\bar{t} + \gamma$ simulated events using EMTopo jets collection.

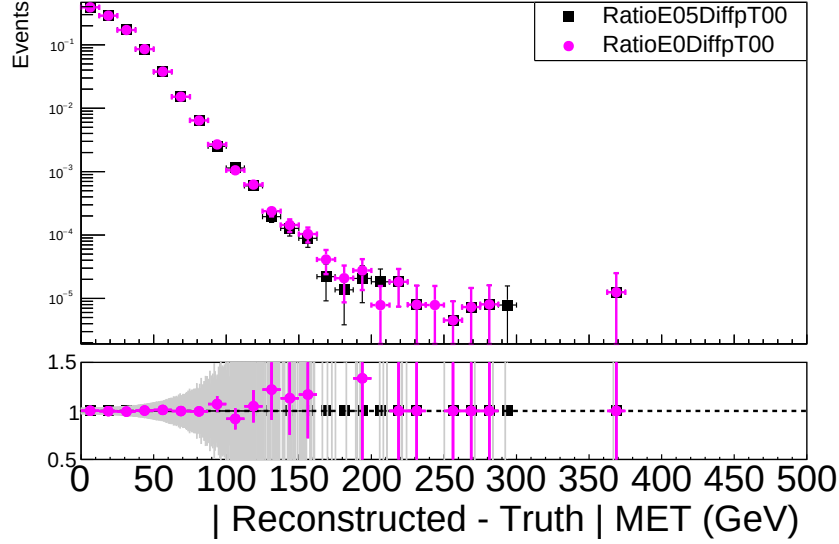


Figure 4.10: Comparison of the $|Reconstructed - Truth| E_T^{miss}$ performance between old (magenta) and new (black) jet-photon overlap removal procedure in $t\bar{t} + \gamma$ simulation.

	$Ratio_E > 0.1$ and $\Delta p_T > 10$ GeV	$Ratio_E > 0.5$ and $\Delta p_T > 0$ GeV
$\int_{150}^{500} reco - truth E_T^{miss}$ (GeV)	$6.20 \cdot 10^{-4} \pm 4.53 \cdot 10^{-05}$	$6.90 \cdot 10^{-4} \pm 4.69 \cdot 10^{-05}$

Table 4.10: E_T^{miss} performance for $t\bar{t} + \gamma$ simulated events using PFlow jets collection.

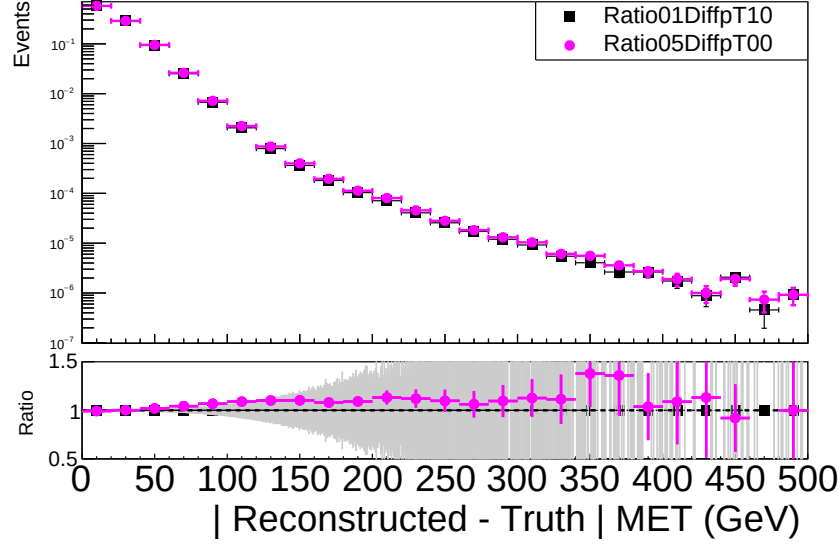


Figure 4.11: Comparison of the $|Reconstructed - Truth| E_T^{miss}$ performance between old (magenta) and new (black) jet-photon overlap removal procedure in $t\bar{t} + \gamma$ simulation.

4.7 Systematic uncertainties

The E_T^{miss} systematic uncertainties are computed for both the response (E_T^{miss} scale) and for the resolution. For the hard component, systematic uncertainties on the electrons, muons, jets, τ -leptons, and photons are propagated to the E_T^{miss} as part of E_T^{miss} measurement, When systematic uncertainty correlations for the same physical item are commonly viewed. Nevertheless, systematic uncertainties associated to non-identical objects are assumed to be uncorrelated. Evaluation of the E_T^{miss} resolution and scale uncertainties due to the soft component $E_T^{miss,soft}$ is discussed in the next sub-section.

4.7.1 Methodology

The systematic uncertainties associated to the E_T^{miss} measurement is estimated based on data to simulation comparisons of several variables distributions evaluating the contribution of $E_T^{miss,soft}$ to the total E_T^{miss} .

4.7.1.1 Observables

The resulting transverse momentum vector of all physical objects p_T originating from a hard scattering process in proton-proton collision (p_T^{HS}) is given by:

$$p_T^{HS} = \underbrace{\sum p_T^e + \sum p_T^\gamma + \sum p_T^\tau + \sum p_T^\mu + \sum p_T^{jets}}_{p_T^{obs}(observable)} + \underbrace{\sum p_T^\nu}_{p_T^{inv}(not\ observable)} \quad (4.9)$$

p_T^ν refers to as the transverse momenta of non interacting particles, which are added to compose p_T^{inv} . The sum of the transverse momenta of visible particles in the event form p_T^{obs} . In the transverse plan, p_T^{HS} should ideally sum to 0 ($p_T^{HS} = |p_T^{HS}| = 0$), according to the momentum conservation law. Due to the hard object momentum miscalibration, mismeasurement, also particles passing through detector zones that aren't instrumented, only an approximation (p_T^{hard}) for the observable particle contribution p_T^{obs} might be evaluated. p_T^{hard} is calculated from high p_T physical objects as described in section 4.2.1.

$$p_T^{hard} = \underbrace{\sum_{electrons} p_T^e + \sum_{photons} p_T^\gamma + \sum_{\tau\ leptons} p_T^\tau + \sum_{muons} p_T^\mu + \sum_{jets} p_T^{jets}} \quad (4.10)$$

p_T^{hard} is expected to be positive and different than p_T^{obs} , including $p_T^{soft} = -E_T^{miss,soft}$ ($E_T^{miss,soft}$ is introduced in Equation 4.5), to p_T^{hard} that ameliorates the estimation of the transverse momentum carried out by the invisible particles. For event topology with zero true E_T^{miss} (e.g. $Z \rightarrow \mu\mu$), the soft term momentum p_T^{soft} is expected to be balanced against the hard term momentum p_T^{hard} ($p_T^{hard} = p_T^{soft}$). This equality expectation is distorted due to the detector inefficiencies, However, for events with $N_{jets} = 0$, p_T^{soft} points in the opposite direction of p_T^{hard} , i.e. in the hadronic recoil direction. The divergence from this prediction is evaluated as a function of p_T^{soft} 's perpendicular ($P_\perp$) and parallel ($P_\parallel$) projections along p_T^{hard} . Figure 4.12 displays these projections quantities for events with exactly one jet and without jets.

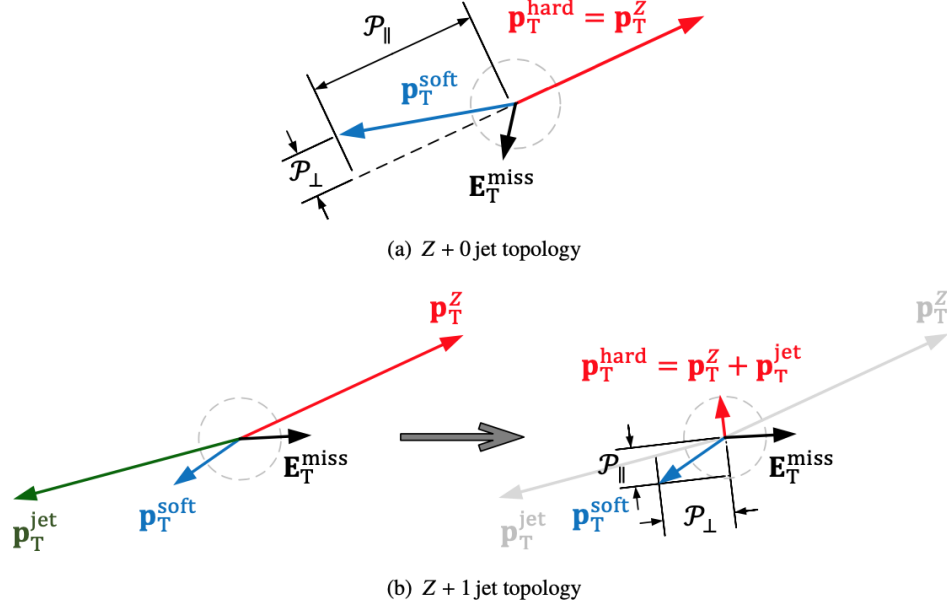


Figure 4.12: $P_{\parallel}$ and $P_{\perp}$ projections of p_T^{soft} along p_T^{hard} for $Z \rightarrow \mu\mu$ for a final state (a) without any jets and (b) with one jet. Figures from Ref [28].

The mean $\langle P_{\parallel} \rangle$ for an identified bin k of phase space identified by p_T^{hard} evaluates the $E_T^{miss,soft}$ response, with $\langle P_{\parallel} \rangle = \langle p_T^{hard} \rangle_K$ demonstrating an accurate response in this bin. The E_T^{miss} resolution contribution from the soft term is evaluated in terms of two components, the fluctuations of the angular deviation and the fluctuations in response around the p_T^{hard} axis, labeled as $RMS_{\parallel}^2$ and $RMS_{\perp}^2$, respectively. These fluctuations are defined in function of variances, as follows:

$$RMS_{\parallel}^2 = \langle (P_{\parallel})^2 \rangle - \langle P_{\parallel} \rangle^2 \quad (4.11)$$

and

$$RMS_{\perp}^2 = \langle (P_{\perp})^2 \rangle \quad \text{with} \quad \langle P_{\perp} \rangle = 0 \quad (4.12)$$

4.7.1.2 Procedures

The evaluation of the systematic uncertainties generated by the $E_T^{miss,soft}$ into the E_T^{miss} calculation is based on data to MC comparisons of $\langle P_{\parallel} \rangle(p_T^{hard})$ for the response, and $RMS_{\parallel}^2(p_T^{hard})$ and $RMS_{\perp}^2(p_T^{hard})$ for the resolution. Various MC samples are employed, each with a different detector simulation or event generator. The exclusive $Z \rightarrow \mu\mu$ final state with

$N_{jet} = 0$ is the foundation for identifying the systematic uncertainty matched with data and all MC simulations for the largest effect of $E_T^{miss,soft}$ on E_T^{miss} . In this scenario, the hard term is the only component from the reconstructed Z boson ($p_T^{hard} = p_T^Z$) as displayed in Figure 4.12(a). Comparing the resulting $P_\perp$ and $P_\parallel$ spectra obtained from MC simulations and data, in bins of p_T^{hard} , one can extract the $E_T^{miss,soft}$ components uncertainties. The smearing of the width and the response both produce longitudinal resolution and scale offsets for the parallel ($P_\parallel$) projection of p_T^{soft} onto p_T^{hard} . Only smearing of the width is considered when assessing transverse resolution offsets in the perpendicular ($P_\perp$) projection.

As a result, the fitted offsets, which were carried out for several MC samples, result in systematic uncertainty with regard to a particular MC modeling setup. In actuality, Gaussian smearing is applied in simulation to the longitudinal and transverse components of $E_T^{miss,soft}$ linked to the direction of p_T^{hard} to account for the resolution offsets. The longitudinal component of $E_T^{miss,soft}$ is scaled up and down to provide an uncertainty band in order to account for discrepancies in response between data and simulation.

4.7.1.3 Systematic uncertainties in E_T^{miss} response and resolution

The E_T^{miss} scale systematic uncertainty is shown in Figure 4.13. The average longitudinal projection $\langle P_\parallel \rangle$ as function of the p_T^{hard} is illustrated in Figure 4.13(a) for the usual Powheg+Pythia8-based simulations also the alternative MC simulation MadGraph, comparing to data. Within the uncertainties related to data, all MC simulation outcomes are predicted to have $\langle P_\parallel \rangle$ MC. The ratio $[\langle P_\parallel \rangle_{MC} / \langle P_\parallel \rangle_{data}]$ remains within the systematic uncertainty band over the full p_T^{hard} range, as can be seen from the lower panel of Figure 4.13(a).

The E_T^{miss} resolution systematic uncertainty that are determined based on the variances of the perpendicular ($RMS_\perp^2$) and the parallel ($RMS_\parallel^2$) projections of E_T^{miss} onto p_T^{hard} are shown in Figures 4.13(b) and 4.13(c), respectively. The p_T^{hard} dependence of $RMS_\parallel^2$ measured for the exclusive $Z \rightarrow \mu\mu$ event in data and two MC simulations is illustrated as well as the variances $(RMS_\parallel^2)_{MC}$ calculated for both set of simulations. The results fit well within the systematic uncertainties of $(RMS_\parallel^2)_{data}$ with the data as can be seen in the bottom of the figure, where the ratio $[(RMS_\parallel^2)_{MC} / (RMS_\parallel^2)_{data}]$ is shown in term of p_T^{hard} .

Figure 4.13(c) summarizes the results of the determination of the variances ($RMS_\perp^2$) of the perpendicular projection as a function of p_T^{hard} , along with the resulting p_T^{hard} dependence of the ratio $[(RMS_\perp^2)_{MC} / (RMS_\perp^2)_{data}]$. All discrepancies between MC simulations are covered by the systematic uncertainties of the data.

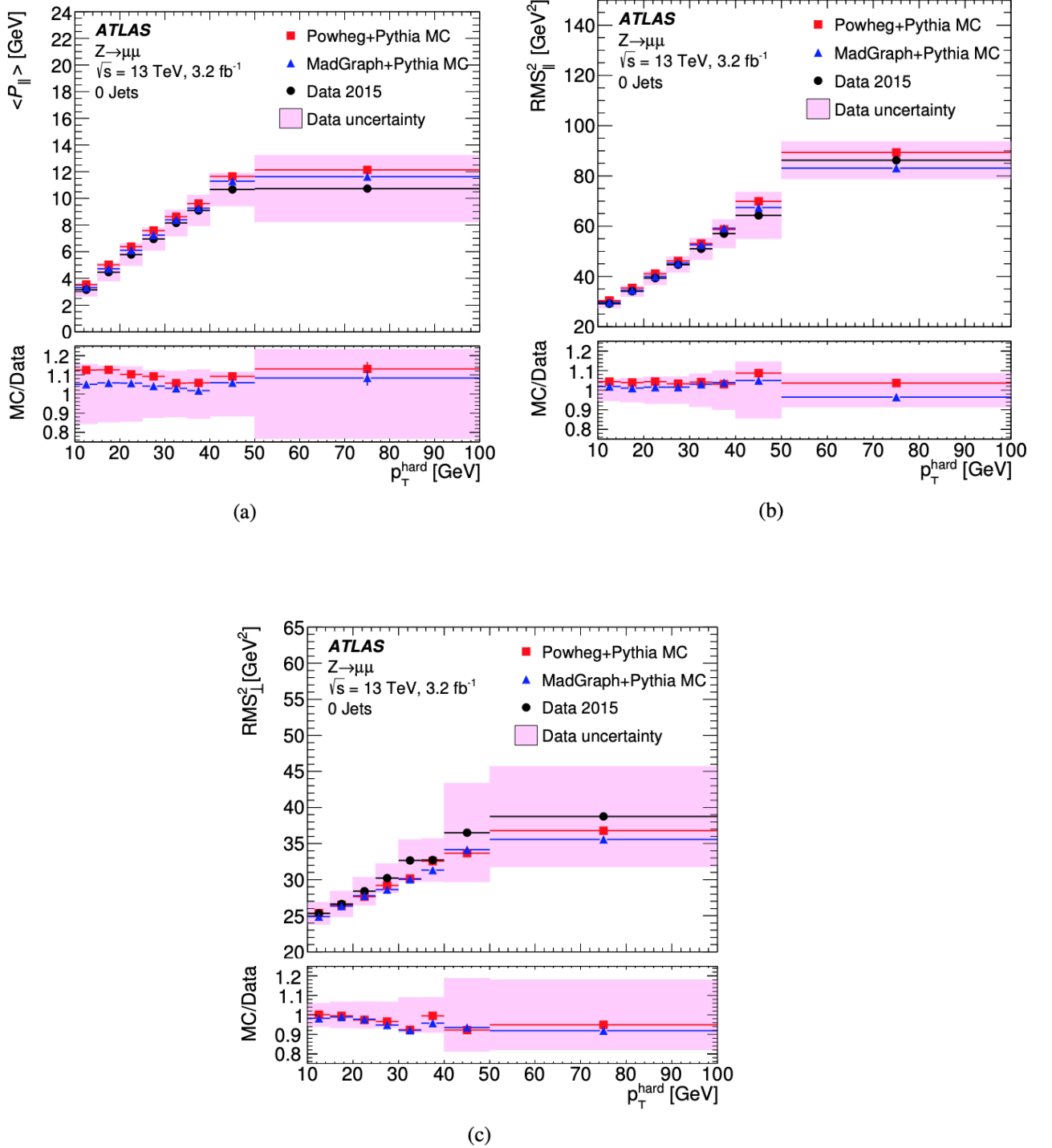


Figure 4.13: (a) mean value of $\langle P_{\parallel} \rangle$ (b) variance $RMS_{\parallel}^2$ for $Z \rightarrow \mu\mu$ events with $N_{jet} = 0$, for data and two different MC simulations, displayed in term of p_T^{hard} . The variance $RMS_{\perp}^2$ is displayed in (c) for similar event samples. Figures from Ref [28].

4.8 Performance of the missing transverse energy E_T^{miss}

After the proton-proton scattering, most SM particles produced leave traces inside the detector, however, SM neutrinos and some theorised exotic particles may escape the detector without producing a signal. Since the momenta of the colliding protons in the transverse direction is essentially zero (in comparison to their boost along the beam direction), the presence of these particles can be inferred by a resultant momentum imbalance in the transverse plane. Ideally the E_T^{miss} should be calculated using the final state particles from the hard pp interaction only, summing the momentum contributions from both charged and neutral particles over the full 4π solid angle. In practice, however, the detector phase space, reconstruction efficiencies, and the presence of additional “pileup” interactions from other pp scattering vertices in the same bunch crossing event, all affect the measurement. Both performance characteristics alter in terms of total event activity and are influenced by pile-up since the E_T^{miss} response is resolution dependent.

As a result, the performance of the E_T^{miss} reconstruction is evaluated by comparing a set of reconstructed E_T^{miss} -linked variables in both MC simulations and data for the same final-state selection, using the same object and event choices. Both response and resolution systematic uncertainties in the E_T^{miss} are obtained from these comparisons and are employed to evaluate the degree of agreement of the data with a given physical model.

The quality of the detector simulation is established individually for all reconstructed particles, jets, and ID tracks, and the total E_T^{miss} uncertainty for each event may therefore be propagated. The distributions of the observables, as well as their average behavior, are compared to relevant scales characterizing the overall kinematic activity, whether it is a hard-scatter event or a pile-up event.

For a total comparison with the data, the samples simulating all the background processes should be taken into account, with the correct weighting procedure for the relative cross sections and luminosity scaling required. The full comparison is mandatory, not only for determining the breadth of our understanding of the event at the truth level, but also in physics searches where cuts on the E_T^{miss} may be applied and disagreements could worsen the systematic uncertainties for the analysis. Moreover, variations in the distributions with different event generators can indicate regions where one particular generator describes the event better than another.

W and Z boson production is relatively abundant at the LHC, and their leptonic decays

provide clear experimental signatures with well understood topologies that make them ideal events for measuring the performance of our detector and reconstruction algorithms. The absence of real E_T^{miss} in $Z \rightarrow \ell\ell$ events allows us to measure the intrinsic resolution and performance of the p_T^{hard} that results mainly from detector response and object reconstruction efficiencies, whereas in $W \rightarrow \ell\nu$ events the neutrino provides a good metric against which to measure the p_T^{hard} scale and direction. Specific exclusive samples are selected from these by adding criteria on the number of rebuilt jets. Zero jet ($N_{jet} = 0$) samples without any jet with $p_T > 20 \text{ GeV}$ (completely calibrated) as well as $|\eta| < 4.9$ are very valuable for examining the soft term's performance. For analyzing the contribution of jets to E_T^{miss} , samples with events selected on the basis of a non-zero number of reconstructed jets with $p_T > 20 \text{ GeV}$ are useful. While the p_T response of jets is properly calibrated and offers a better assessment of the total event p_T -flow, pile-up affects the p_T resolution of jets, which can have a negative impact on E_T^{miss} reconstruction performance. The plots in figure 4.14 show the quality of the MC modelling of E_T^{miss} and $\sum E_T$ that has been evaluated in an inclusive sample $Z \rightarrow \mu\mu$ by comparing the distributions of these observables to data. The agreement between data and MC simulations is within 20% for the bulk of the E_T^{miss} distribution shown in Figure 4.14(a), while clear incompatibility not included within the overall (statistical and systematic) uncertainties of the distributions for high E_T^{miss} . Figure 4.14(b) illustrates the $\sum E_T$ distributions, where the disagreement in the Data-MC modelling significantly larger than the overall uncertainties for $200 \text{ GeV} < E_T^{miss} < 1.2 \text{ TeV}$. This indicates the degree of mis modelling of the final state mostly in function of hard-object composition in MC simulations. Data versus Monte Carlo simulation comparisons for the $E_T^{miss,soft}$ term in Figure 4.14(c). Agreement between data and MC simulations is better, with differences up to 10% for $E_T^{miss,soft} < 30 \text{ GeV}$. However, clear disagreement for larger $E_T^{miss,soft}$ have been seen included in the uncertainties band.

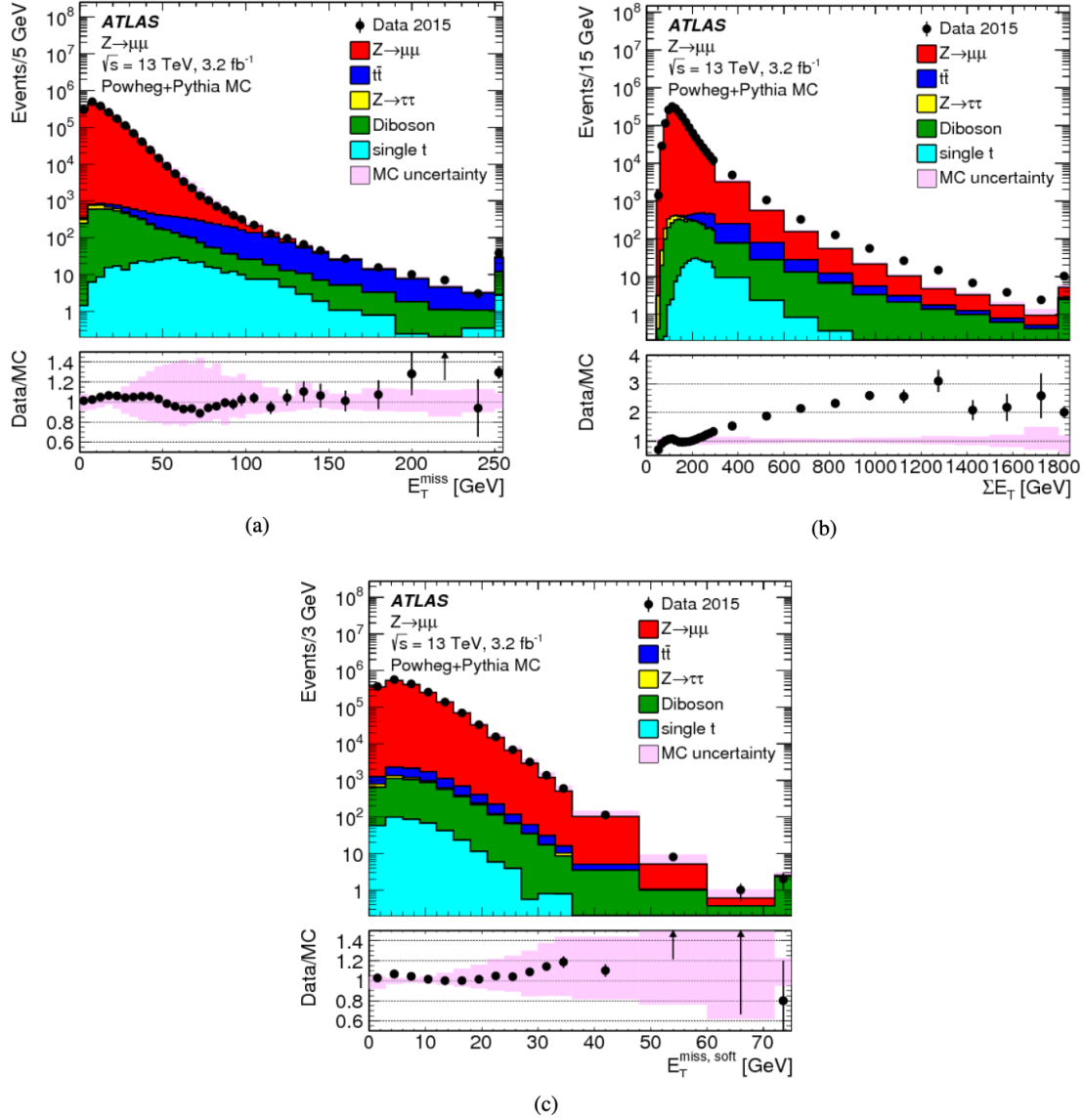


Figure 4.14: (a) Shape of E_T^{miss} , (b) ΣE_T and (c) the soft term displayed for both data and MC events satisfying the $Z \rightarrow \mu\mu$ selection. The data/MC is shown in the lower panel of each figure. The red band shows the associated systematic and MC statistical uncertainties. Figures from Ref [28].

The total reconstructed E_T^{miss} and the soft term in $W \rightarrow e\nu$ events from data, where a genuine E_T^{miss} is expected, are compared with the expected distributions from the MC samples. Figure 4.15(a) shows the description of the data by the MC for inclusive $W \rightarrow e\nu$ events, where the level of agreement is about 20 %, within the associated uncertainties of the measurement. In exclusive $W \rightarrow e\nu$ events shown in Figure 4.15(b), data and relevant

background MC distributions are compared and they are well below 20%, but show a trend to larger discrepancies for decreasing $E_T^{miss} < 40$ GeV. Some background contribution from QCD processes in these events is also expected, particularly in the low E_T^{miss} region. The extraction of this contribution is exceedingly inefficient and only feasible with high statistical uncertainty because these processes are not properly modelled. There are relatively rare events with just one jet that is unintentionally reconstructed as an electron, even in very large MC samples of multi-jet final states. Therefore, the QCD contamination is not included in the background set, giving rise to the trend in the data/background ratio in this region. The $E_T^{miss,soft}$ term distributions are compared between data and MC simulations in Figure 4.15(c). A good comparison is obtained within the uncertainties for $E_T^{miss,soft} > 10$ GeV. The discrepancy seen in the MC distribution for decreasing $E_T^{miss,soft} < 10$ GeV are due to the fact that the multi-jet contribution is missed.

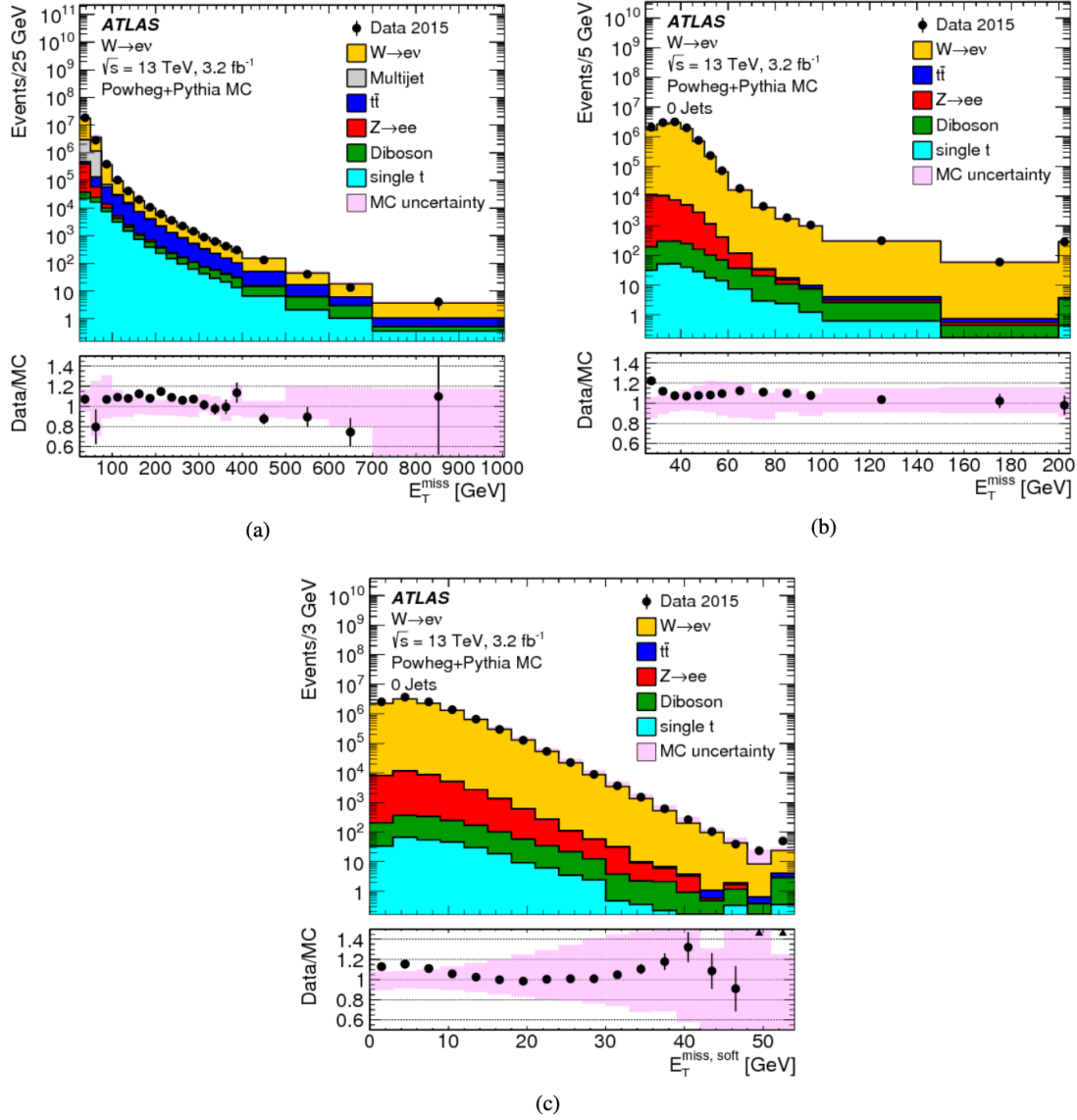


Figure 4.15: (a) Distributions of the total E_T^{miss} in the inclusive case, (b) with $N_{jet} = 0$ case, as well as (c) the soft term $E_T^{miss,soft}$ reconstructed in $N_{jet} = 0$. The data/MC is shown in the lower panel of each figure. The red band shows the associated systematic and MC statistical uncertainties. Figures from Ref [28].

Chapter 5

Statistical Frameworks

In high-energy physics, a clearly specified statistical framework is essential for claiming a new discovery or determining the model's exclusion limits. The approach particular to this investigation to calculate the statistic findings will be discussed in the next chapter. The search for new physics will be based on a couple of possibilities: background-only (SM), and signal-plus-background (BSM+SM)¹. The goal is to evaluate the consistency of the observed data with background-only expectation, where the presence of large deviation may indicate the existence of a signal. This is accomplished by focusing on the observables that are most sensitive to the new physics. Particles which can couple to top quarks through resonant production of a $t\bar{t}$ pair would modify a peak structure in the $t\bar{t}$ invariant mass distribution. At the LHC, it would be worthwhile searching for a bump structure generating a local deficiency or excess regarding the prediction surrounding the mass of the exotic particle. As a result, the number of $t\bar{t}$ events that occurring within a given local mass range provides an observable helping to look for new resonances. Predictions concerning such observable, which is a random variable, may be represented as probability density functions (PDFs) under the background-only and signal+background hypotheses. A Poisson distribution is the most nearly suited PDF in a $t\bar{t}$ resonance search. The parameter of the Poisson distribution, ν , represents the average number of events in a specific bin and Equation 5.1 quantifies the probability of identifying n $t\bar{t}$ events in a given $m_{t\bar{t}}$ bin.

$$P(n; \nu) = \frac{e^{-\nu} \nu^n}{n!} \quad (5.1)$$

Because the mass of the sought-after particles is undetermined, the full $t\bar{t}$ mass spectrum up to 6 TeV is investigated. The joint PDF representing the chance of discovering $\mathbf{n} = \{n_i\}$

¹Several signal models can have an impact on the same observable and can all be investigated in the same search.

events across all mass bins is determined in Equation 5.2, where ν_i represents the average number of events in the i^{th} bin.

$$P(n; \nu) = \prod_{i=1}^{N_{bins}} \frac{e^{-\nu_i} \nu_i^{n_i}}{n_i!} \quad (5.2)$$

By identifying the joint PDF in Equation 5.2, with n_i set to the observed events in the i^{th} bin and ν_i set to the average of n_i corresponding to that model, the data may be compared to a specific signal assumption. The outcome, referred to as the likelihood² of the model, L , evaluate how likely it was to detect the data n_i assuming the validity of the model. Data might look compatible with a false model or inconsistent with a real model due to statistical fluctuations, regardless of whether the data is explained by the background+signal hypothesis or the background-only hypothesis. The data n_i should converge to the mean of the Poisson distribution that describe them correctly as more events are recorded, however there is an inherent statistical uncertainty that necessitates the use of rigorous statistical techniques to get reasonable conclusions about how good a model to simulate the real data when comparing several hypotheses and deciding the one validated using data.

The methods discussed in the following chapters use a frequentist method to statistics. If the correct kind of PDF is used to represent the real data, the non identified parameters in the PDF are assumed to take the real values that may be determined in principle from data if $N_{events} \rightarrow \infty$. On the other hand, finite data may be utilized to construct claims concerning the truthful value with an uncertainty level connected to N_{events} . This section introduces the statistical methods employed in the subsequent chapters, which will be developed while looking for a new resonance BSM.

5.1 Evaluation of Agreement between Observed Data and Expectation

A statistical standard to begin with is a hypothesis test of the data collected for a search to assess how the real data is compatible with the SM expectation. The SM hypothesis is referred to as the null hypothesis, H_0 , while the competing Beyond SM hypothesis is named the alternative hypothesis, H_1 . A PDF of the random vector or random variable being calculated corresponding to every hypothesis. For instance, If the measurement is the total

²The likelihood is not a true probability. It refers to how likely a model is given observed data rather than how likely a collection of data is given a specific model.

number of events n_i across all N bins, labeled $\mathbf{n}$, therefore H_0 and H_1 may be written as:

$$H_0 : f(\mathbf{n}; H_0) = \prod_{i=1}^{N_{bins}} \frac{e^{-B_i} B_i^{n_i}}{n_i!} \quad (5.3)$$

$$H_1 : f(\mathbf{n}; H_1) = \prod_{i=1}^{N_{bins}} \frac{e^{-(B_i+S_i)} (B_i + S_i)^{n_i}}{n_i!} \quad (5.4)$$

where B_i is the average of predicted events in a given bin assuming the background-only hypothesis, and $S_i + B_i$ is the average of predicted events in the i^{th} bin assuming the S+B hypothesis.

For the observed data, a test statistic t is evaluated as an estimate of the degree of consistency between the data and H_0 . It may be easy to describe the entire dataset with a single scalar number. The sensitivity of a good test statistic to this type of variation produced by the presence of a signal is optimized. While searching for a $t\bar{t}$ resonance search with the events are split in several $m_{t\bar{t}}$ bins, this can be a local excess of events compared to the SM prediction where the most clear excess corresponds to larger test statistic values. The test statistic is considered as a random variable since it is related to the observed data and has a different PDF under H_0 and H_1 . Figure 5.1 shows a critical region, w , determined such that if $t \in w$ then H_0 is removed.

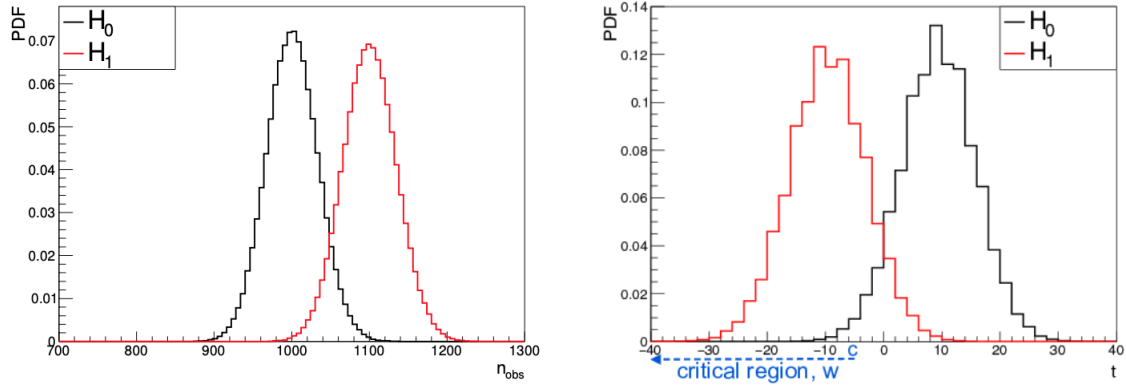


Figure 5.1: The number of observed events in a specific bin, n_{obs} assuming H_0 and H_1 according to the PDFs in Equations 5.3 and 5.4 (left). Test statistic t distribution estimated from the number of observed events where the test statistic is employed here is the ratio of the log likelihood (right).

In this approach, a hypothesis test may remove hypotheses, but it can never show them to be real. H_0 is simply not excluded if $t \notin w$. The value of w affects how confident one

can be in rejecting the hypothesis. The significance degree of the hypothesis test is the probability α that t will be seen in the critical zone under H_0 . This is the probability that H_0 will be removed when true, and is referred to as a type-I error. A type-II error is defined as the probability β that H_0 is not rejected while H_1 is true. The power of a test to reject H_0 when it is not true is designed to be $1 - \beta$, depending to the truth alternative hypothesis H_1 . Figure 5.1 shows an example of selected critical region where α and β can be written as follows:

$$\alpha = P(t \in w; H_0) = \int_{-\infty}^c f(t; H_0) dt \quad (5.5)$$

$$\beta = P(t \notin w; H_1) = \int_c^{\infty} f(t; H_1) dt \quad (5.6)$$

The chance of finding data equal or less uniform with H_0 than the real data was, as measured by the test statistic (t_{obs}), is referred to as the p-value defined as follows:

$$p - value = \int_{t_{obs}}^{\infty} f(t; H_0) dt \quad (5.7)$$

The lower the p-value, the wider is the disagreement among the data with respect to the null hypothesis. The critical zone can be defined as a cut on the p-value of the data. If $t_{obs} \in w$ imply the p-value to be $\leq \alpha$, small enough that the data is supposed to be enough inconsistent with H_0 which lead its rejection. Figure 5.2 (left) presents an example test statistic distribution and the shaded area indicates the observation's p-value. In particle physics, the p-values are usually represented in terms of significance, Z , defined as:

$$Z = \Phi^{-1}(1 - p) \quad (5.8)$$

where Φ^{-1} is the inverse cumulative function of a unit Gaussian pdf with mean zero and unit standard deviation. The Z values are derived from one-sided tail probabilities of a unit Gaussian distribution. Figure 5.2 (right) presents the link between Z and the p-value. The following expression can be used to express p-values in terms of Z values:

$$p - value = \frac{1 - erf(Z/\sqrt{2})}{2} \quad (5.9)$$

where erf denotes the error function. It is defined as:

$$erf z = \frac{2}{\pi} \int_0^z e^{-t^2} dt \quad (5.10)$$

It is widely acknowledged in the particle physics field that a significance of at least $Z = 5$ (often referred to as 5σ) is required to reject the background hypothesis (H_0) in order to claim discovery of a new particle which correspond to $p\text{-value} = 2.87 \cdot 10^{-7}$. For the purpose of rejecting a specific signal hypothesis (H_1), a $p\text{-value} = 0.05$ (corresponds to 95% confidence level) threshold is utilized; the associated Z value is 1.64.

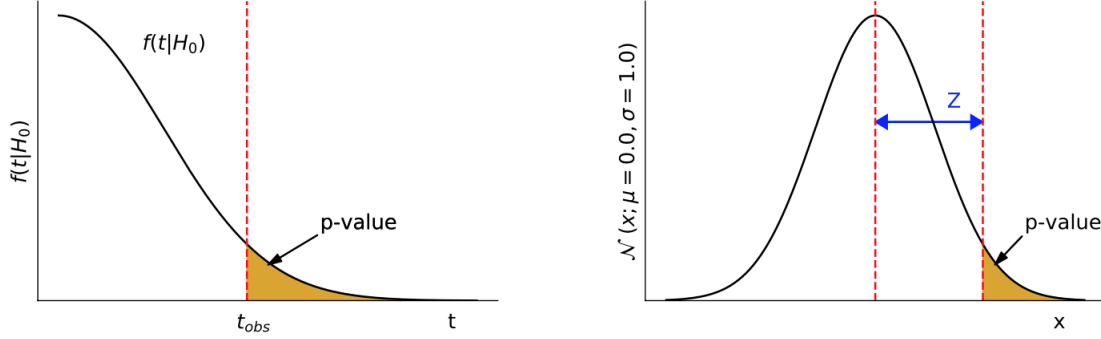


Figure 5.2: (left) Demonstration of how to calculate the p-value given the observed test statistic value, t_{obs} . The p-value is the integral of the right-hand tail of the test statistic distribution. (right) The relationship between the p-value and the significance Z is shown using a standard normal distribution.

5.2 Trials factor

Resonance searches try to identify data discrepancies in a collection of neighboring bins when assessing the consistency between data and H_0 . Equation 5.4 accounts for data variations among all bins; in case where the test statistic was built following this PDF, the sensitivity of a resonance search would be reduced. The signal models don't predict the location of the excess, and after the data has been exposed, selecting a set of bins showing the biggest excess of events corrupts the hypothesis test. It is possible that selecting the major statistical fluctuation out of N independent groups of bins to generate the test statistic will yield a $p\text{-value} < \alpha$, leading to reject H_0 when true. This is generally occur for larger N . Due to the range of test statistics that may be employed in the experiment, a hypothesis test that employs a test statistic reliant on the results of the experiment no longer has type-I error *alpha* [89]. The chances that each of the test statistics returning a $p\text{-value} < \alpha$ is higher

compared to the chance that a particular one among them does the same:

$$P(\text{at least one } p\text{-value} \leq \alpha) = 1 - \prod_{i=1}^{N_{bins}} P(p\text{-value} > \alpha) = 1 - (1 - \alpha)^N \quad (5.11)$$

To avoid this problem, a hypertest is determined relied on the results of N independent hypothesis tests applied to the data by sampling all bins. Its test statistic is:

$$t = -\log_e(\min_i \{p\text{-value}_i\}) \quad (5.12)$$

the negative log of the smallest p-value among the N tests. This hypertest evaluates each bin groups separately and yields the appropriate type-I error α as shown in Equation 5.5.

5.3 Parameter Estimation

A fundamental statistical problem in the analysis of collision data is the evaluation of an unidentified parameters based on a predetermined PDF. The signal strength, μ is frequently the parameter of interest in a new physics search. In a signal+background model, the average number of predicted events in a given bin ν_i may be parameterized as $\nu_i = B_i + \mu S_i$, with $\mu = 1$ related to the nominal S+B hypothesis $\mu = 0$ proportional to the background-only hypothesis, and $\mu \in (0, 1)$ related to a signal+background model with a decrease of the signal cross-section. If H_0 is not rejected by a hypothesis test with significance α , the next step is to investigate several H_1 hypotheses:

$$f(n; H_1) = \prod_{i=1}^{N_{bins}} \frac{e^{-(B_i + S_i)} (B_i + S_i)^{n_i}}{n!} \quad (5.13)$$

when the signal strength μ is sufficient to prevent H_0 from being removed. If the nominal background and signal rates S_i and B_i are recognized, the parameter μ may be deduced from the observed data by evaluating the parameter $\nu = \mu S + B$. Because H_0 is not removed, μ is usually calculated in function of an upper limit on its value. If H_0 is rejected by a hypothesis test with significance α , then H_1 are investigated to calculate the cross-section of the observed signal.

5.4 Confidence Interval Definition

Confidence interval gives an indication of the degree of the statistical uncertainty of an unknown parameter ν . An experiment's observed data is used to create a confidence

interval. It includes a set of possible parameter values and has the property of containing the true value ν_t in a particular fraction of similar incorrelated experiments as the number of experiments $N \rightarrow \infty$. A parameter estimator is a function of the observed data. For instance, the parameter's Maximum Likelihood Estimator $\hat{\nu}$, estimated from observed data n_{obs} when the PDF is a Poisson distribution of mean ν , is simply $\hat{\nu} = n_{obs}$. The estimator's PDF is therefore just dependent on the parameter's value: $f(\hat{\nu}; \nu)$, and may be employed to identify a value a (b) for which the estimator is bigger (less) than a (b) with probability α (β):

$$\alpha = P(\hat{\nu} \leq a(\nu)) = \int_{a(\nu)}^{\infty} f(\hat{\nu}; \nu) d\hat{\nu} \quad (5.14)$$

$$\beta = P(\hat{\nu} \geq b(\nu)) = \int_{-\infty}^{b(\nu)} f(\hat{\nu}; \nu) d\hat{\nu} \quad (5.15)$$

The probability that the estimator $\hat{\nu}$ lies in the interval $[b, a]$ can be written as follows:

$$P(\hat{\nu} \in [b(\nu), a(\nu)]) = 1 - \alpha - \beta \quad (5.16)$$

This can be changed in order to learn more about ν . From Equation 5.14 and 5.15, it follows that:

$$\alpha = P(a^{-1}(\hat{\nu}) \leq \nu) \quad (5.17)$$

$$\beta = P(b^{-1}(\hat{\nu}) \geq \nu) \quad (5.18)$$

resulting an interval containing ν with probability $1 - \alpha - \beta$:

$$P(\nu \in [a^{-1}(\hat{\nu}), b^{-1}(\hat{\nu})]) = 1 - \alpha - \beta \quad (5.19)$$

This is true for all conceivable values of ν , but because the estimator is based on observed data, it is a ν_t estimator, hence the built interval applies to ν_t . It is thus allowed to construct a range that includes the unknown parameter with a predetermined probability using observed data and an identified PDF of the estimator. The lower threshold of the interval is set to zero to set an upper limit on a parameter, and the resulting upper bound is bigger than the parameter with the given probability.

5.5 The Likelihood ratio test-statistic

The likelihood ratio is the ratio of the probability densities for a given experimental result for two hypotheses. In searches for new particles, before the process of identifying a confidence interval for the signal strength is discussed, the ratio of the log-likelihood should be described. The approach requires testing data against S+B hypotheses with augmenting signal strength μ , via the log-likelihood ratio test statistic. For this, the roles of the hypotheses are inverted, with S+B becoming H_0 and background-only becoming H_1 . For rejecting the null hypothesis, the likelihood ratio is the strongest statistic test to use. As a function of likelihood parameters α , the likelihood ratio may be represented as:

$$\lambda(\alpha) = \frac{L(\alpha; x)}{L(\hat{\alpha}; x)} \quad (5.20)$$

where x denotes the data and $\hat{\alpha}$ denote the MLEs. When the likelihood approaches a Gaussian in the large sample limit, Wilk's theorem [90] states that $-2\ln\lambda(\theta)$ follows a chi-squared distribution (χ^2) with m degrees of freedom, where m is the number of parameters in the model. In this limit, the quantiles $\chi^2_{1-\alpha}$ of the χ^2 distribution can be utilized to calculate $100(1-\alpha)\%$ confidence intervals.

When the likelihood depends on several parameters but only one parameter μ and its uncertainty are of relevance, a profile likelihood ratio can be used as:

$$\lambda(\alpha) = \frac{L(\mu; \hat{\hat{\theta}})}{L(\hat{\mu}; \hat{\theta})} \quad (5.21)$$

where $\hat{\hat{\theta}}$ is the conditional MLE of θ that maximizes L for a given value of μ is in the numerator. The denominator is the unconditional maximum likelihood function, i.e., $\hat{\mu}$ and $\hat{\theta}$ that maximize the likelihood, L . In the limit of high statistics, the profile likelihood ratio similarly follows a χ^2 distribution. For some of the research in this dissertation, the likelihood ratio is employed, and the profile likelihood ratio is always used in the hypothesis test for significance and limit calculation.

5.6 The CL_s value and exclusion limits

The CL_s approach is commonly employed to stipulated a confidence interval to a signal cross section in a S+B model. In order to test the signal+background hypothesis H_{s+b} against the background-only hypothesis H_b in each traditional hypothesis testing for calcu-

lating the upper limit. The p-value, p_{s+b} , is defined as the probability of finding a value q that is equally or less compatible with H_{s+b} in comparison to the data, q_{obs} . We use p_{s+b} to perform the test in the standard " CL_{s+b} " technique. At a confidence level $(1 - \alpha)$, if we find $p_{s+b} < \alpha$, H_{s+b} is ruled out.

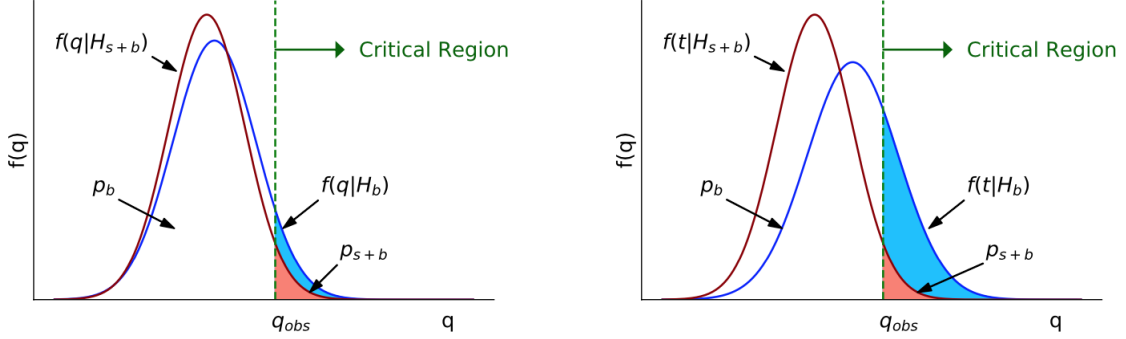


Figure 5.3: Distributions of the test statistic q under the $s + b$ and b hypotheses in an example where sensitivity to the signal model is very small (left). Schematic diagram showing the test statistic distribution under b and $s + b$ hypotheses (right). The observed test statistic value is indicated by the dotted line. The $s+b$ p-value is shown by the red filled area, whereas the area filled with skyblue shows $(1-p_b)$.

The difficulty with using only the $s+b$ p-value, p_{s+b} , is that it's possible to rule out a signal hypothesis with a probability close to α even if the signal model's sensitivity is very low. The q distribution under the b and $s+b$ hypotheses practically overlap in the schematic Figure 5.3 (left), illustrating this scenario. However, based on the p_{s+b} , the $s+b$ hypothesis might be excluded. In general, H_{s+b} will be rejected if the expected number of signal events (s) is relatively low in comparison to the background events (b) and the observed data displays downward variation. Intuitively, in the limit cases where $s \ll b$, $H_{s+b} \approx H_b$, therefore the exclusion probability must go to zero. But in CL_{s+b} method it approaches α .100%.

The test statistic distribution under the b hypothesis, H_b , can be built and calculate the appropriate p-value p_b . A schematic diagram of the distributions and p-values are shown in Figure 5.3 (right). The " CL_s " method [91] is used in most recent particle physics experiments to defend against this problem. Rather than explicitly using p_{s+b} , we use " CL_s " which is defined as:

$$CL_s = \frac{p_{s+b}}{1 - p_b} \quad (5.22)$$

The p_{s+b} value is effectively penalized by dividing it by the test's power, $1 - p_b$. If the H_{s+b} and H_b test statistic distributions are well separated, then $(1 - p_b) \gg p_{s+b}$ and the exclusion based on CL_s will be quite comparable to the exclusion based on p_{s+b} . However, if the two distributions almost overlap on each other, then $(1 - p_b) \approx p_{s+b}$. As a result, CL_s becomes quite the same, and the ability to exclude the s+b hypothesis decreases. The signal strength (μ) exclusion limit is determined with a 95 percent confidence level by determining the μ for which $CL_s < \alpha$, where $\alpha = 0.05$.

Chapter 6

Search for $t\bar{t}$ resonances in ATLAS

The ATLAS detector at the LHC is used in this chapter to simulate the hunt for new particles decaying into top-quark pairs. Various theories BSM proposes the presence of similar heavy resonances (Section 1.6 in Chapter 1), which have an infinitesimal production cross section with respect to the SM $t\bar{t}$ pair production. The presented analysis seeks for inconsistencies with the SM prediction in $t\bar{t}$ production that can point out the attendance of a new particle decaying into a $t\bar{t}$ pair, which would be visible as a local deficit or excess in the $t\bar{t}$ invariant mass ($m_{t\bar{t}}$) spectrum. Consequently, the discriminant observable used in the search is the reconstructed $t\bar{t}$ mass spectrum. The top pair system is made up of six decay products: one muon or electron, four jets, and a neutrino (ν). Nonetheless, top quarks with a significant p_T relative to their mass may be generated at LHC energy.

The search is carried out by creating and assessing the $m_{t\bar{t}}$ spectrum's compatibility with data in terms of expectation. The $m_{t\bar{t}}$ distribution in data is created by choosing events that are compatible with $t\bar{t}$ decay, with the majority of nominated events being $t\bar{t}$ signature with substantial input from $t\bar{t}$ backgrounds. By performing a similar criteria to identify SM $t\bar{t}$ events and contribution of non $t\bar{t}$ SM background fulfilling the signal region (SR) requirements, the anticipated background $m_{t\bar{t}}$ spectrum is reconstructed. When simulation fails to adequately explain acceptable SM processes, data driven methodologies are commonly employed to assess their contribution to the $m_{t\bar{t}}$ spectrum.

The $t\bar{t}$ decay channel investigated in this search conforms to the semi-leptonic ($\ell + jets$) topology as discussed before. It has a smaller branching ratio than the entire hadronic topology. Nonetheless, the existence of a lepton in the final state provides an experimental advantage by reducing the QCD multi-jet background contamination in the study. The dileptonic topology, on the opposite side, has a distinct signature since two leptons and two neutrinos must be detected in the final state, but the branching percentage is lower

and the complete reconstruction of the $t\bar{t}$ causes many uncertainties. The existence of an extra neutrino in the dileptonic channel discourages the E_T^{miss} and the p_T of neutrino from matching. Only the final state topologies " $\mu + \text{jets}$ " and " $e + \text{jets}$ " are taken into account. They come from a decay chain in which one of the top-W quark's bosons decays leptonically into an electron and an electron-neutrino, or hadronically into a quark-antiquark pair, while the second W boson decays hadronically into a quark-antiquark pair. Given the difficulty of detecting taus, the " $\tau + \text{jets}$ " channel is not investigated. When τ decays leptonically, it may be mis-reconstructed as muon or electron. When taus decay hadronically, they can form narrow jets with restricted track multiplicity, which can be mis-reconstructed as jets or electrons. As a result, a subset of " $\tau + \text{jets}$ " occurrences may be chosen for study.

The chapter is ordered as follows: A quick outline of the analysis strategy is presented in the first Section 6.1; Then the selection criteria imposed on the physics objects used in this search is discussed in Section 6.2; details on data and background samples are given in Section 6.3; the event selection is outlined in Section 6.4. The relevant aspects of the data driven backgrounds methods are given in Section 6.5. Section 6.6 details the reconstruction of the $t\bar{t}$ system. Section 6.7 discusses the different systematic uncertainties considered in the analysis. Section 6.8 confronts background expectations with data; Section 6.9 discusses the compatibility of data with SM only hypothesis and Section 6.10 discuss the results of the search in terms of expected limit on resonant $t\bar{t}$ production.

6.1 Analysis Strategy

The search for new resonances decaying to $t\bar{t}$ pairs is conducted in the single-lepton plus jets ($\ell + \text{jets}$) final state, using 139 fb^{-1} of p-p collected data at $\sqrt{s} = 13 \text{ TeV}$ at the LHC. This final state (introduced above) is experimentally favored since it allows to effectively suppress the large contribution from QCD multi-jet background in the analysis. According to this topology, one of the selected top quarks decays into a W-boson and a b-quark, while the W boson decays subsequently into a single lepton with its associated neutrino. Two quarks coming from the W-boson decay and a b-quark are produced by the second top quark in the final state. A Feynman diagram of this process is shown in Figure 6.1. Nonetheless, top quarks with a significant p_T relative to their mass may be generated at high energy at the LHC. They are referred to as boosted tops. For resonance masses greater than around 1 TeV, boosted tops are anticipated to be formed [92]. Normally, since they are produced close to each other, the decay products of the boosted top cannot be resolved separately. This search's event selection has been altered to contain occurrences where it is

likely that two of the hadronic decay products were reassembled as a unique jet. The three top decay products are rebuilt as a single entity in an extensively enhanced topology fit. [93] describes an ATLAS search for extremely boosted $t\bar{t}$ resonances.

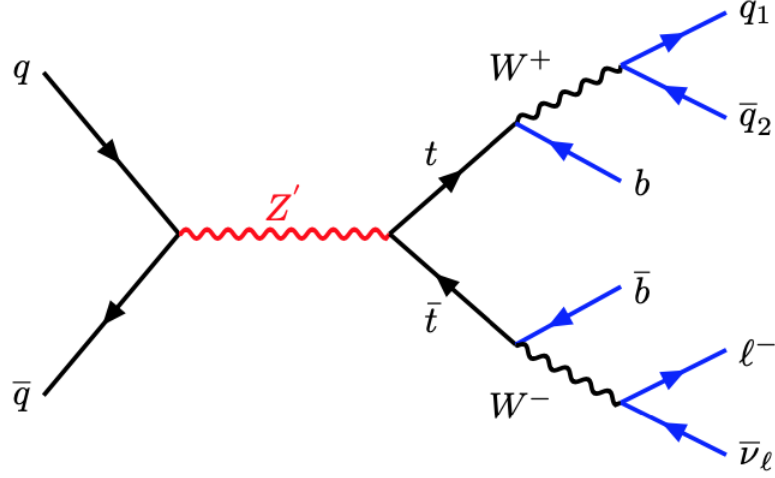


Figure 6.1: Feynman diagram for the Z'_{TC2} signal process, in which one of the top quarks decays hadronically and the other leptonically.

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 is called the boosted topology, and the top quark disintegration products are then sent in the same angular region of the detector. This phenomenon is presented in Figure 6.2, where the distances are shown in the (η, ϕ) plan, between the decay products of the hadronic and leptonic top quarks in term of the p_T of the top quark and W boson.

In this case, if no particular handling is made for this kind of events signature, they can be dismissed as background or the $m_{t\bar{t}}$ can be inadequately estimated for them. However, If the top quarks have low transverse momenta, the decay products are generally well isolated in the detector and can be reconstructed individually: the hadronic top-quark decay results in three separate jets, two jets originating from the W-boson decay and a b-jet. The lepton coming from the decay of the other top quark is usually well separated from the corresponding b-jet. This decay topology is called 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 for the resolved-boosted hybrid approach only two of the jets from the hadronically

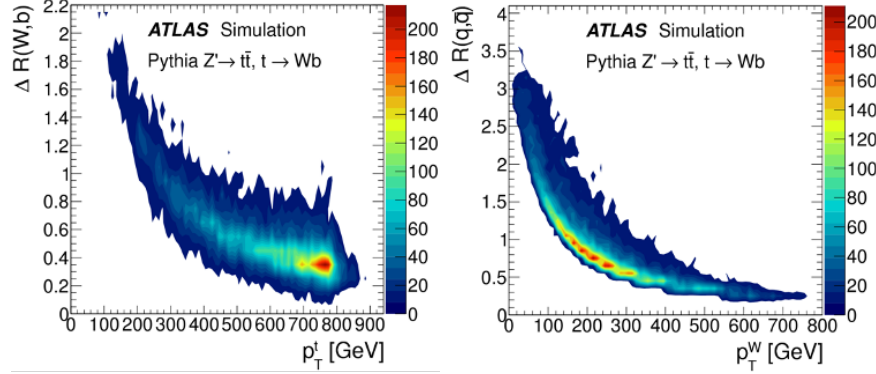


Figure 6.2: Angular separation ΔR of the decay products in the two-body decays (left-hand) $t \rightarrow W b$, and (right-hand) $W \rightarrow q\bar{q}$ as a function of the transverse momentum of the top quark and the W boson, respectively [30].

decaying top quark, typically those from the decay of the W boson, are merged while the third jet is reconstructed separately. Although, for large momentum (above 500 GeV. Figure 6.3), the distance between objects becomes quickly quite lower to the standard reconstruction parameters of the jets.

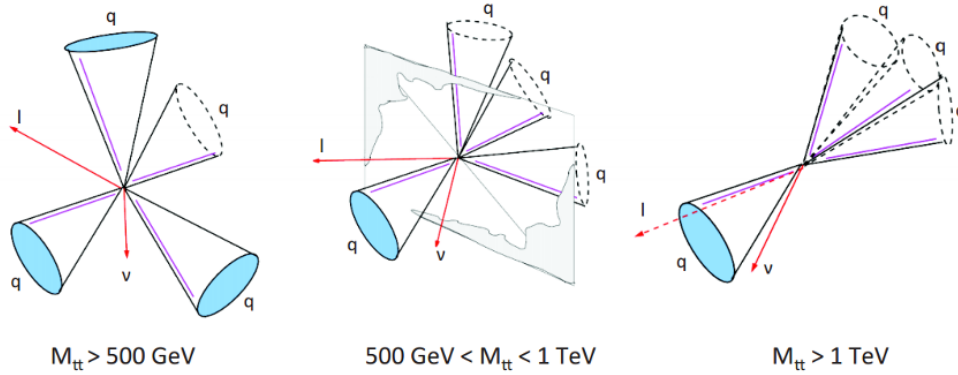


Figure 6.3: The different decay topologies of the top-antitop system. From left to right, the resolved decay topology, the resolved-boosted-hybrid decay topology and finally boosted decay topology

The selection criteria and reconstruction techniques are designed to various decay scenarios to secure a high selection efficiency and $m_{t\bar{t}}$ calculation. The analysis can take benefit of the collision at high energy and persist sensitive to an extensive variety of resonance masses since the nominated top quark pair events are split into four decay mods depending on the flavour of the lepton and the top quark decay channel. In addition, the selection is optimized to accurately built the decay products of a hadronic boosted top as a single Large-R jet in

order to reduce the non- $t\bar{t}$ backgrounds. Afterwards, a statistic test is performed to validate the consistency of the $t\bar{t}$ mass spectrum with the SM only hypothesis.

The search yielded two results. First, a p-value is attributed to the real data by mean of a test statistic sensitive to a top anti-top resonance but independent of any signal benchmark, i.e. sensitive to the yield of events in a random $m_{t\bar{t}}$ bin, as described in Chapter 5. Then, the degree of agreement between data and prediction is measured. The p-value used to determine whether or not there is evidence of a new finding particle not predicted by the SM. Second, assuming there is no significant disagreement, upper bounds for the cross-sections times branching ratio of signals in three benchmark scenarios for a given range of resonance widths and masses are determined at 95% confidence level (CL). The upper limit of a confidence interval (0, upper limit) that include the true cross-section value 95% of the time is known as the 95% CL upper limit.

The approach followed in this analysis is general, the search for signals with no presupposition on the signal shape. Furthermore, expected limit is set on one benchmark model: topcolor assisted technicolor Z'_{TC2} (described in Chapter 1). Other benchmark scenarios, such as Randall-Sundrum (RS) warped extra-dimensions, were traditionally explored including the bulk Kaluza-Klein (KK) graviton and gluon decaying into a $t\bar{t}$ pair (Simulations are not attainable for the time scale of this thesis). The free parameters in such models permit for setting restrictions on several widths, such as small or broad resonances, where recording the final signal is problematic [94].

6.2 Object definitions

This analysis makes use of leptons (electron, muon), missing transverse energy (E_T^{miss}) from neutrinos and jets (including b-tagged jets). Candidate object for both data and simulation is asked to satisfy specific selection and pass overlap removal if needed. In Chapter 3, more information on how the physics objects are rebuilt using detector signatures is provided. The supplemental cuts applied on each candidate item are highlighted in this section. Only objects satisfying the quality cuts are used in the event selection and further stages of the analysis.

6.2.1 Leptons

The reconstruction of electrons and muons used in this search have been described in Chapter 3, in Section 3.1 and 3.2 respectively. Lepton candidates are required to have a transverse momentum greater than 25 GeV, and $|\eta|$ within the Inner detector acceptance ($|\eta|$

< 2.5 for muons and $|\eta| \in [0, 1.37) \cup (1.52, 2.47]$ for electrons). The TightLH [95] identification quality is imposed on the electron, while Medium [96] quality is used for the muon to suppress as much as possible the background from non-prompt leptons. In order to make sure the lepton candidates are associated to the correct primary vertex, cuts are applied on the impact parameters, d_0^{BL} and z_0^{BL} . For electron, the d_0^{BL} is required to satisfy $|d_0^{BL}(\sigma)| < 5$ and the longitudinal impact parameter with respect to the beam-line is required to be $|z_0^{BL} \sin \theta| < 0.5$ mm. The muon candidates are further required to satisfy $|z_0^{BL} \sin \theta| < 0.5$ mm and $|d_0^{BL}(\sigma)| < 3$. Both leptons must be isolated from hadronic activities. In this study, TightTrackOnly_VarRad isolation working point [97] is used for both electrons and muons.

6.2.2 Jets

Large-R jets with $R = 1$ are used in this analysis to collimate the decay products of the top that decay hadronically in the boosted regime using the DNN based top tagging algorithm at the 80% working point [98]. They are reconstructed using local-calibrated topoclusters (LCTopo) and trimmed using the radius parameter $R_{trim} = 0.2$ and $f_{cut} = 0.05$. Large-R jets are asked to have p_T greater than 350 GeV, $|\eta| < 2.0$ and mass > 40 GeV.

Small-R jets with $R = 0.4$ are built using the calorimeter clusters passing the Particle Flow reconstruction algorithm. They should have $|\eta| < 2.5$ and $p_T > 25$ GeV. For jets with $|\eta| > 2.4$ and $p_T < 60$ GeV, a threshold on the jet vertex tagger (JVT) (described in Chapter 3), $JVT > 0.50$, is applied to suppress pile-up.

The b-tagging assignment are applied to the small-R jets using the Deep Learning algorithm DL1r [99] operating at 77% working point.

6.2.3 The missing transverse momentum

The missing transverse momentum, E_T^{miss} , is specified as the negative vector sum of the p_T of all selected physics objects in the following order: electrons, muons, photons and jets. In top of this, a set of 'soft terms' for tracks that don't fit the high- p_T objects picked. More details can be found in Chapter 4.

6.2.4 Overlap Removal

Overlap between reconstructed jets and leptons occurs in both the boosted and resolved event topologies, therefore, an overlap removal (OR) procedure is used to avoid duplicates of the same physical object and to isolate electrons and muons from jets.

To handle the overlap between jet and electron, electron-in-jet subtraction method is used to increase the signal acceptance. For every jet in a specific event, if an electron is present

within a $\Delta R(e, \text{jet}) < 0.4$, the electron 4-vector is subtracted from the jet 4-vector and the resulting jet is defined as electron-subtracted jet. A p_T dependent cut is defined to check if an electron-subtracted jet has high enough p_T to still be considered as a jet. If $\frac{p_T^{el-sub}}{p_T^{jet} + b} < C$, where b is a bias term and C is the threshold term¹, then it is assumed that the jet is really part of the electron so the jet is removed and the electron is kept, otherwise the ΔR between the electron and the modified jet is recalculated, and the following steps are done:

- **If $\Delta R(e, \text{e-subtracted jet}) > 0.2$:** we keep both the electron-subtracted jet and the electron. We use the electron-subtracted jet as a proxy of the actual jet.
- **If $\Delta R(e, \text{e-subtracted jet}) < 0.2$:** we assume that the electron is coming from a B-hadron decay. Then the electron is removed, its p_T added back to the electron-subtracted jet, and the original jet is kept in the event.

The muon-jet OR is done in two steps. First, jets close to the muon with $\Delta R(\mu, \text{jet}) < 0.2$ or with less than three associated tracks are removed. Then, a distance parameter depending on the muon's p_T is used to remove muons for which:

$$\Delta R(\mu, \text{jet}) < 0.04 + \frac{10 \text{ GeV}}{p_T^\mu} \quad (6.1)$$

As the muon's p_T increases, the $\Delta R(\mu, \text{jet})$ threshold value becomes very small, allowing to keep jet against muon since it is likely fake.

6.3 Data and Monte Carlo samples

This search considers the full Run-2 dataset collected at $\sqrt{s} = 13$ TeV proton-proton collisions during 2015 (3.2 fb^{-1}), 2016 (33.0 fb^{-1}), 2017 (44.3 fb^{-1}), and 2018 (58.5 fb^{-1}), combined for a total integrated luminosity of 139 fb^{-1} . To maintain steady beam conditions, the data was received from a Good Runs List (GRL) that was prepared centrally also it is recorded while all relevant ATLAS sub-system systems should be operational to ensure that the events are perfectly reconstructed. Single-lepton triggers were used to gather lepton+jet event candidates. The major background mechanisms for this analysis are the irreducible Standard Model (SM) $t\bar{t}$ production, production of a W boson, Z boson together with additional jets (W+jets, Z+jets), single-top quark production and diboson production. All of them are evaluated using simulation. The QCD multi-jet background is directly evaluated from the data. Simulated MC samples are employed to evaluate backgrounds that

¹The bias (b) and threshold (C) terms are configurable in AnalysisTop. We are using $b = 20 \text{ GeV}$ and $C = 0.37$, which are found to be optimal for semi-leptonic and di-lepton channel.

generate jets and prompt leptons and signal processes. The event selection, SM background estimations, and signal efficiencies are all developed, provided, and evaluated using the MC samples. More information about the MC samples used in this analysis is provided below.

Standard Model $t\bar{t}$ and single top

The SM $t\bar{t}$ samples were generated by POWHEG [100]+PYTHIA 8 [83], using the NNPDF 3.0 NLO parton distribution function (PDF) [101] for the next-to-leading order (NLO) matrix element (ME) and the NNPDF 2.3 LO PDF [102] for the showering and hadronization together with the A14 tune. For masses of the $t\bar{t}$ system larger than 1.1 TeV, dedicated samples with a filtering on $m_{t\bar{t}}$ are used. They are normalized to a next-to-next-to-leading order (NNLO) cross-section ($\sigma_{\text{BR}} = 451.59 \text{ pb}^{-1}$) computed with top++ 2.0 [103] that includes resummation of next-to-next-to-leading logarithmic (NNLL) soft gluon terms. The shape of the $t\bar{t}$ spectrum is also reweighted to NNLO using an ATLAS tool [104]. These corrections are available either as function of top p_T (up to 3TeV), as well as a function of $t\bar{t}$ p_T and $t\bar{t}$ mass (up to top $p_T < 800 \text{ GeV}$). The single top production is generated by POWHEG+PYTHIA 8 using the NNPDF 3.0 NLO PDF for the NLO ME and the NNPDF 2.3 LO PDF for the showering and hadronization together with the A14 tune.

W+jets, Z+jets and diboson

SHERPA 2.2.1 [105] was used to generate the Z+jets and W+jets events, with the NNPDF 3.0 NNLO PDF set [101] and the generated events are normalized to the NNLO cross-section. To guarantee sufficient statistical precision over the whole phase space, filters on the vector boson p_T and heavy flavor content have been utilized. A charge-asymmetry normalization is used as a W+jets correction derived from data, as explained in Section 6.5.2. Diboson (WW, WZ, ZZ) process is generated using the Sherpa 2.2.1 generator. The NNPDF3.0 NLO PDF set is used while generating the events and normalized to the NNLO cross-section.

Signal modeling

The signal process $pp \rightarrow Z'_{TC2} \rightarrow t\bar{t}$ is simulated using the PYTHIA 8 generator with the default settings for generic Sequential Standard Model (SSM) Z'_{SSM} generation. Samples are generated for a variety of Z'_{TC2} masses from 0.4 to 5 TeV, using the NNPDF 2.3 LO PDF [102] set. The Z'_{TC2} is generated with width of 3% of its mass which is narrower compared to the detector resolution and branching fraction to $t\bar{t}$ of 33%. Following [45], a multiplicative

k-factor of 1.3 is used to bring the cross- sections used for the Z'_{TC2} to the next-to-leading order (NLO).

6.4 Event Selection

Only events satisfying a number of general quality cuts are accepted for further analysis. This set of constraints are applied to simulated events and data to reduce the non-top-antitop final states. To achieve this, a serie of fundamental cuts are applied to all events to pick out $t\bar{t}$ like events, pursued by a further strict selection dividing the chosen events into four categories by the decay mode (resolved or boosted) and lepton flavor (muon or electron).

6.4.1 Preselection

An initial pre-selection is applied on each event to be considered as signal candidate. These requirements are common for both regimes: boosted or resolved and they are as follows:

Trigger

The selected events must pass either a single electron or a muon trigger. The exact trigger names are listed in Table 6.1

Year	lepton channel	Trigger
2015	electron muon	e24_lhmedium_L1EM20VH, e60_lhmedium, e120 lhloose mu20_iloose_L1MU15, mu50
2016-2018	electron muon	e26_lhtight_nod0_ivarloose, e60_lhmedium_nod0, e140_lhloose_nod0 mu26_ivarmedium, mu50

Table 6.1: List of the single-lepton high Level triggers (HLT) used in this analysis.

The primary vertex (PV)

At least one primary vertex with two tracks must have been triggered by each event. The vertex with the highest $\sum p_{T,track}^2$, where $p_{T,track}$ is the transverse momentum of tracks associated with the vertex, is picked as the primary vertex.

Event quality

To ensure no detector noise creates disagreements between data and prediction, the events must pass the GoodCalo criteria and any event having at least one jet flagged as LooseBad is rejected.

Lepton selection

All the events must have exactly one reconstructed lepton, either an electron or a muon with $p_T > 30$ GeV. The lepton must be matched to the trigger. Events including a second lepton with $p_T > 25$ GeV are vetoed.

Leptonic-W selection

Due to the required leptonic decay of one of the W-bosons, we expect to see E_T^{miss} coming from the neutrino. Then, only events with $E_T^{miss} > 20$ GeV are considered to avoid the QCD multi-jet background and other non- $t\bar{t}$ contributions. In addition to the E_T^{miss} cut, a combination of E_T^{miss} and W boson transverse mass selection, $E_T^{miss} + m_T^W > 60$ GeV is applied, where:

$$m_T^W = \sqrt{2 \cdot p_{T,\ell} \cdot E_T^{miss} (1 - \cos(\Delta\phi))} \quad (6.2)$$

using the transverse momentum of the lepton $p_{T,\ell}$, and the angular distance in ϕ between lepton and the missing transverse momentum.

b-tagging

The b-tagging requirement is necessary to distinct signal events from background events with only light flavor jets. Then, event should comprise at minimum one b-tagged jet. For the b-tagging purpose, DL1r algorithm with 77% efficiency working point is used.

At this point the resolved and boosted channel selections diverge. Fulfilling the later selection criteria, event is evaluated toward the boosted selection and in case of failure it is asked to pass the resolved selection.

6.4.2 Boosted Selection

When the hadronic top quark (t_{had}) decay products are build as one large-R jet, events have a boosted topology. This topology is seen commonly in top quark decays with high transverse momentum. Boosted events should include one large-R jet or more with $p_T > 300$ GeV that is top-tagged. Events in this topology should include one small-R jet or more within $\Delta R < 1.5$ from the lepton. If many jets satisfy this cut, we select the one with the highest p_T and labeled as the selected jet, j_{sel} , which is marked as the b-jet from the top that decay leptonically. One b-tagged small-R jet or more is demanded in each event, however we don't enforce any b-tagging on the j_{sel} . Selecting j_{sel} as b-jet, without necessity to be b-tagged, provides an enhanced $m_{t\bar{t}}$ resolution with respect to employing b-tagging information to select it. If the top quarks are sufficiently boosted, their decay products will also be well separated in the detector. This is reflected by the requirement of:

- $\Delta\Phi(large - R, \ell) > 2.3$
- $\Delta R(large - R, j_{sel}) > 1.5$.

6.4.3 Resolved Selection

The selection in this topology is almost the same than for the boosted decay mode, apart the requirement of at least four well separated in (η, ϕ) small-R jets instead of at least one large-R jet. The $t\bar{t}$ system is built using a χ^2 minimization algorithm, discussed in detail in section 6.6.3, to chose the appropriate four jet combination in each event and assign them to the top quarks that decay hadronically or leptonically. The combination providing the lowest χ^2 value is chosen. We kept only the events for which $\log_{10}(\chi^2) < 0.9$ to remove backgrounds that satisfy the resolved selection cut.

6.4.4 b-tag categories

To enhance the sensitivity of the $t\bar{t}$ resonance search analysis with respect to the limit results, events are split into four different categories in function of the b-tagged calorimeter jets:

- **Category 0:** $\nexists$ b-tagged jet associated to either of the top quark candidates.
- **Category 1:** Only the leptonic top candidate has an associated b-tagged jet.
- **Category 2:** Only the hadronic top candidate has an associated b-tagged jet.
- **Category 3:** Both the leotonic and hadronic top candidates have b-tagged jets associated to them.

6.4.5 Signal acceptance times efficiency

The acceptance² times efficiency³ ($A \times \epsilon$) as a function of truth $t\bar{t}$ invariant mass ($m_{t\bar{t}}^{truth}$) is shown for the Z'_{TC2} signal in Figure 6.4 where $\mu/e + \text{jets}$ final states are included separately in the plots for the boosted and resolved regimes. Note that all the different mass points used for the signals are combined here, as if they had all the same cross section. For the boosted regime, clear difference between the $e + \text{jets}$ and the $\mu + \text{jets}$ channels are observed for $m_{t\bar{t}}^{truth} > 2$ TeV with the $A \times \epsilon$ in the latter case being up to a factor of two higher than in the electron case. This is a consequence of the differences in the overlap removal

²The fraction of the signal events, that is visible to and recorded by the detector systems.

³The fraction of signal events, that pass the selection.

procedures for the two lepton flavor: in the case of the e-jet overlap removal, the electron is rejected if $\Delta R(e, jet) < 0.2$, whereas the muon-jet overlap removal has a looser $\Delta R(\mu, jet)$ requirements, as mentioned in the Section 6.2, which allows it to keep higher $A \times \epsilon$ at higher $m_{t\bar{t}}^{truth}$ values. $A \times \epsilon$ for the resolved regime is relevant at low mass and looks similar for both $\mu + jets$ and e+jets channels with a resonance at 0.75 TeV.

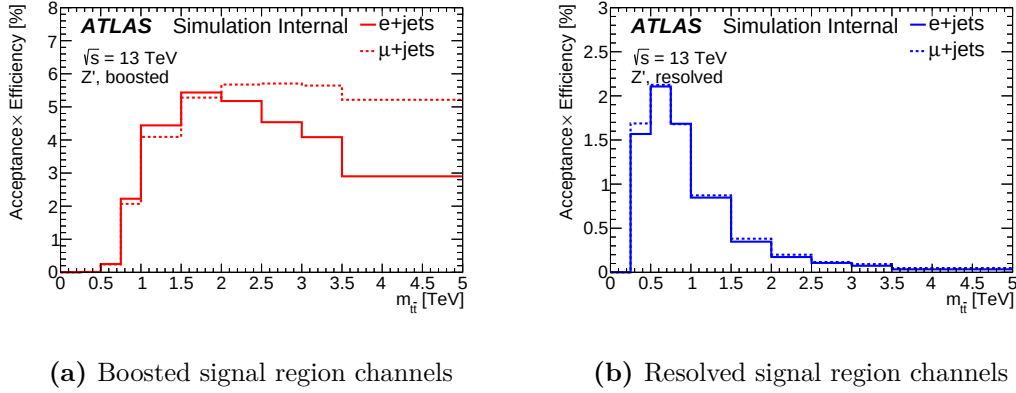
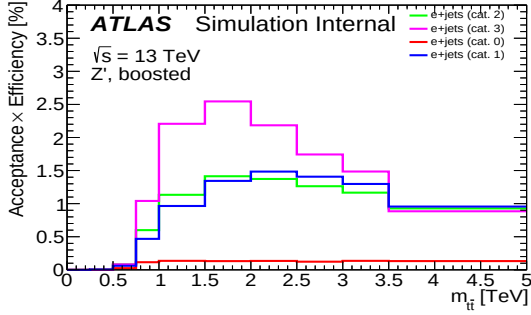


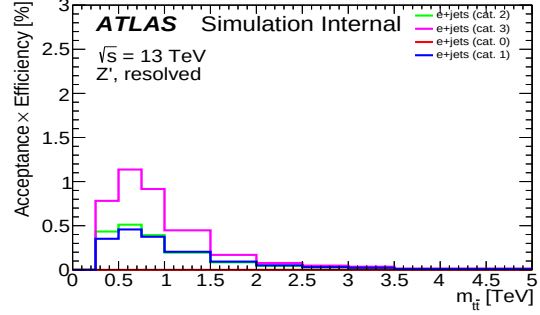
Figure 6.4: Acceptance times efficiency of the Z'_{TC2} signals for the boosted (left-hand) and the resolved topology (right-hand) selections.

Further, $A \times \epsilon$ plots are split based on the b-tagged calorimeter jets into four category in Figure 6.5. For resolved topology, category 0 component is expected to be empty since the χ^2 algorithm ask events to have at least one b-tagged calorimeter jet to be considered in the analysis, also we remark that category 3 where the leptonic top candidate and the hadronic top candidate are associated to b-tagged jet shows high $A \times \epsilon$ values either for electron or muon channels. For the boosted regime category 3 presents high $A \times \epsilon$ values and category 0 has clearly worse performances for the reconstruction either for electron or muon channel.

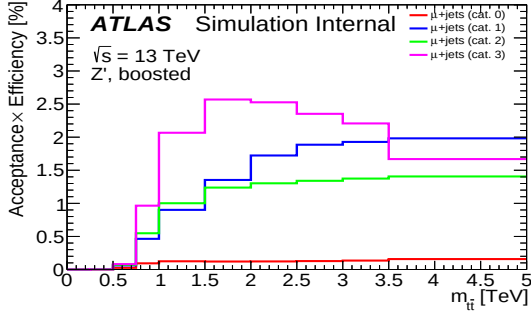
The overall $A \times \epsilon$ presented in Figure 6.6 is mostly dominated by the resolved-topology at the low mass region below 0.75 TeV. In higher mass region, most signal events instead are selected by the boosted approach.



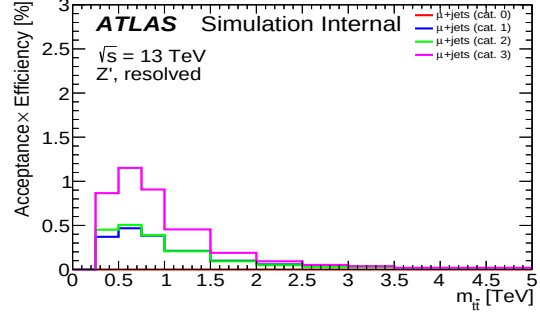
(a) Boosted electron signal region channel



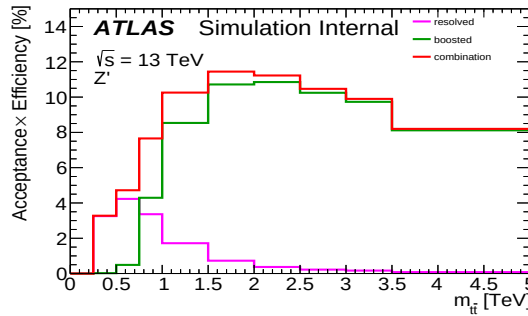
(b) Resolved electron signal region channel



(c) Boosted muon signal region channel



(d) Resolved muon signal region channel

Figure 6.5: Acceptance times efficiency of the Z'_{TC2} signals split by b-tagging categories.

Figure 6.6: Acceptance times efficiency of the Z'_{TC2} signal as a function of the true $t\bar{t}$ mass for the combined boosted and resolved topology selection (red), the boosted (green) and resolved (magenta) selections.

6.5 Data driven backgrounds

6.5.1 Multi-jet background modelling

6.5.1.1 The Matrix Method

BSM physics may be unveiled by studying events with a large number of outgoing jets, created at the LHC with energies above the TeV scale (energetic multi-jet events). Such events are dominated by QCD processes, where the calculations rely on some sort of approximation. Therefore, it is important to develop a robust approach for modelling such events that could prove the existence of BSM signals. Hence, QCD multi-jet background contribution is evaluated directly from data using a data-driven method, a so-called “matrix method”. The normalization and shape of this background can be extracted from data through the reweighting of individual data events. Two separate selections on charged leptons are applied leading to two different datasets for both muons and electrons. First a data sample is selected requiring the lepton to satisfy loose isolation criteria; this is thereafter referred to as the loose sample. Then a sub-sample is selected requiring the lepton to satisfy tight isolation criteria. The leptons selections used in the matrix method are summarized in Table 6.2.

Lepton isolation	Electron	Muon
Loose	MediumLH and No isolation	Loose and No isolation
Tight	TightLH and TightTrackOnly	Medium and TightTrackOnly_VarRad

Table 6.2: Definitions of the leptons used for the matrix method QCD estimation.

Given those two samples, the number of events fulfilling the loose and tight lepton identification criteria can be summarized in this system of equations:

$$N_L = N_{\text{prompt}} + N_{\text{QCD}} \quad (6.3)$$

$$N_T = \epsilon N_{\text{prompt}} + f N_{\text{QCD}} \quad (6.4)$$

Where $N_L(N_T)$ is the number of data events satisfying the loose (tight) selections, N_{prompt} is the number of signal-like leptons coming from the W decay and N_{QCD} is the number of misidentified fake leptons. The quantity ϵ , also referred to as real efficiency,

denotes the fraction of loose signal-like leptons subsequently passing the tight cut; f , also referred to as fake rate, denotes the fraction of loose QCD leptons passing the tight cut. Real and fake efficiencies can be defined as:

$$\epsilon = \frac{N_{\text{prompt-tight}}}{N_{\text{prompt-loose}}} \quad (6.5)$$

$$f = \frac{N_{\text{QCD-tight}}}{N_{\text{QCD-loose}}} \quad (6.6)$$

Solving Equations 6.3 and 6.4 for N_{prompt} and N_{QCD} , it is possible to evaluate the QCD multi-jet contribution to the signal region as:

$$w_{\text{QCD}}(N_{\text{T}}, N_{\text{A}}) = f \times N_{\text{QCD}} = \frac{(\epsilon - 1)f}{\epsilon - f} N_{\text{T}} + \frac{\epsilon f}{\epsilon - f} N_{\text{A}} \quad (6.7)$$

Where N_{T} is the number of events containing a tight lepton, and N_{A} is the number of events containing anti-tight lepton (i.e. a loose lepton which unsuccessful the tight cuts). This result however cannot be used directly as the parameters ϵ and f can depend on event kinematics and have to be modelled accordingly. To take those dependencies into account, an unbinned approach is employed in which a weight is constructed. With this technique, all the events are used in the background estimation, as the event weight is calculated differently for a loose event and for a tight event. In the expression of the weight $w_{\text{QCD}}(N_{\text{T}}, N_{\text{A}})$, N_{T} (N_{A}) is equal to 1 if the event satisfies both the loose and tight criteria and 0 if the event passes only the loose criteria. The QCD expectation in the final sample is therefore the sum of weights calculated across the loose sample.

6.5.1.2 Estimation of fake and real rates

The **real rate** ϵ (signal efficiency) can be measured from the SM $t\bar{t}$ simulation samples with the same process mixture as found in the signal region, using the standard selections except the lepton criteria, which are “loose”. The real efficiency is equal to the fraction of reconstructed loose leptons (matched to the true lepton from the W decay to provide an enriched real lepton sample) that pass the tight cuts. The best parametrisations of the real efficiency ϵ are to be in term of the lepton isolation (topoetcone20) and the lepton transverse momentum p_{T} . Figures 6.7 and 6.8 show the real efficiency for the muon and the electron channel respectively split based on the angular separation (ΔR) between a lepton and its closest small-R jet.

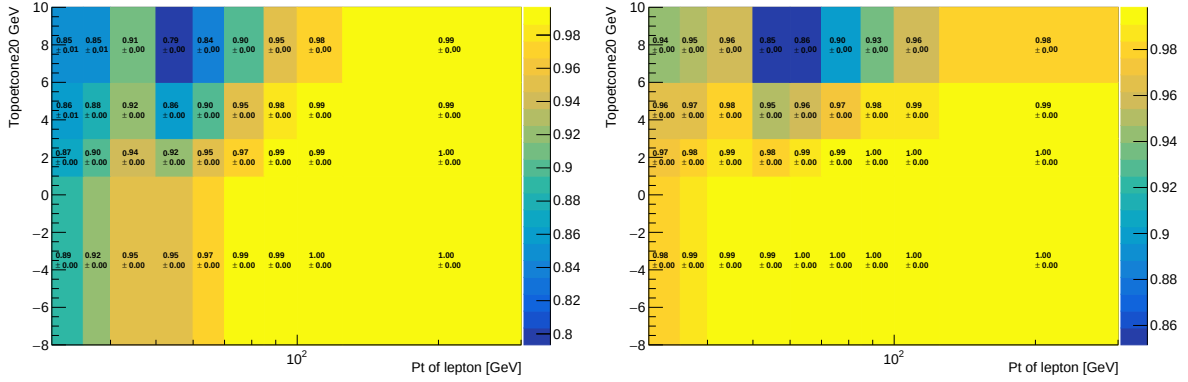


Figure 6.7: Real efficiency in the muon channel for $\Delta R_{\min}(\mu, jet) > 0.4$ (left-hand) and for $\Delta R_{\min}(\mu, jet) < 0.4$ (right-hand).

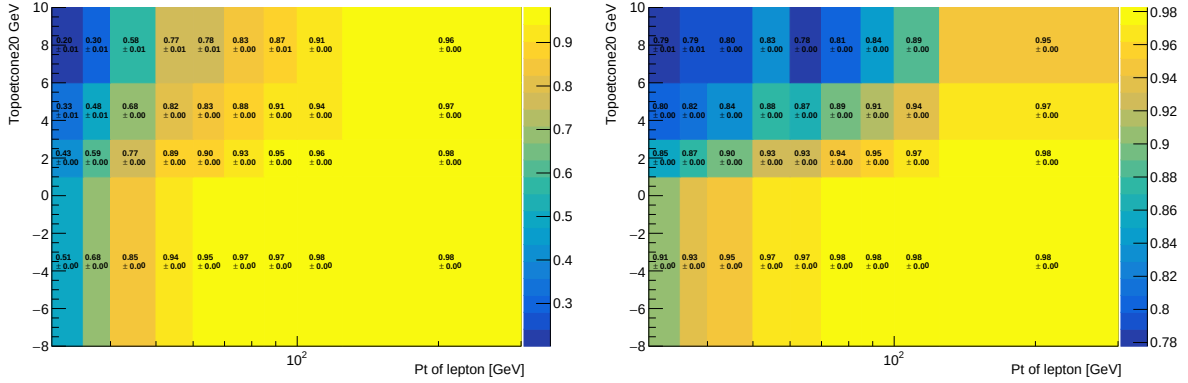


Figure 6.8: Real efficiency in the electron channel for $\Delta R_{\min}(e, jet) > 0.4$ (left-hand) and for $\Delta R_{\min}(e, jet) < 0.4$ (right-hand).

The **fake rate**, f , can be interpreted as the probability for a QCD-like lepton passing the loose selection cut to subsequently pass the tight selection cut. It is measured from data in the QCD-enriched control region ($CR_{\text{QCD}}^{4j,1b}$) after subtracting the prompt lepton contamination in both tight and loose lepton samples as shown in Figures 6.9 and 6.10 for the muon and electron channels respectively. The $CR_{\text{QCD}}^{4j,1b}$ is built by inverting some cuts of the signal region considered to reject the QCD multi-jet events as follows:

- **muon channel:**
 - exactly one loose muon

- no tight electron
 - $E_T^{miss} < 20 \text{ GeV}, E_T^{miss} + m_T^W < 60 \text{ GeV}$
 - ≥ 4 jets
 - ≥ 1 b-tagged jet
- **electron channel:**
 - exactly one loose electron
 - no tight muon
 - $E_T^{miss} < 20 \text{ GeV}, E_T^{miss} + m_T^W < 60 \text{ GeV}$
 - ≥ 4 jets
 - ≥ 1 b-tagged jet

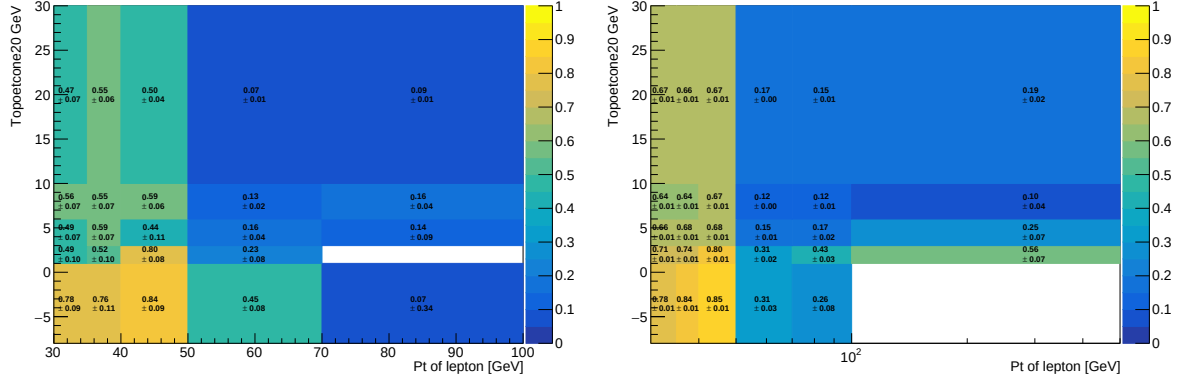


Figure 6.9: Fake rate in the muon channel for $\Delta R_{\min}(\mu, jet) > 0.4$ (left-hand) and for $\Delta R_{\min} \mu, jet) < 0.4$ (right-hand).

6.5.1.3 ΔR correction factor to the real and fake rates

The parameterisation is further improved for the muon and the electron channels by applying a ΔR correction factor to the real and fake rates. The correction is implemented as illustrated in Equations 6.8 and 6.9:

$$\epsilon_{\text{corr}}(\Delta R, P_T, \text{topoetcone20}) = \epsilon(P_T, \text{topoetcone20}) \times \frac{\epsilon(P_T, \Delta R)}{\epsilon(P_T)} \quad (6.8)$$

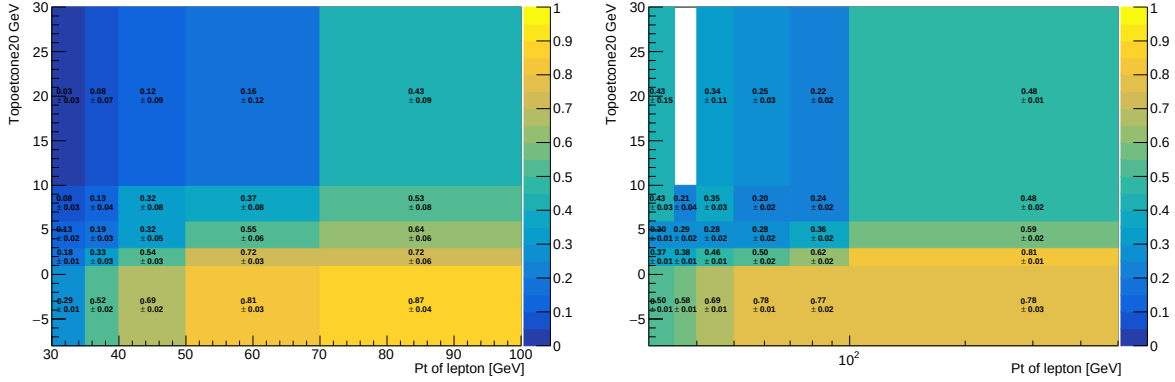


Figure 6.10: Fake rate in the electron channel for $\Delta R_{\min}(e, jet) > 0.4$ (left-hand) and for $\Delta R_{\min}(e, jet) < 0.4$ (right-hand).

$$f_{\text{corr}}(\Delta R, P_T, \text{topoetcone20}) = f(P_T, \text{topoetcone20}) \times \frac{f(P_T, \Delta R)}{f(P_T)} \quad (6.9)$$

This correction allows a better description of the rates when $\Delta R_{\min}(lep, jet) < 0.4$, which is particularly relevant for the boosted topologies. It has been verified that this extra correction improves the QCD multi-jet description in a validation region for muons, where a loose angular separation requirement for the overlap removal between muons and jets is used. The parametrisation used in the corrections is shown in Figure 6.11 where the real and fake rates show a strong dependency on the angular separations between the closest jet and muon ($\Delta R_{\min}(mu, jet)$). Similar dependence on ΔR in the electron channel is also studied and shown in Figure 6.12. The impact of the ΔR correction for the electron and muon channels is illustrated in Figures 6.13 and 6.14 respectively.

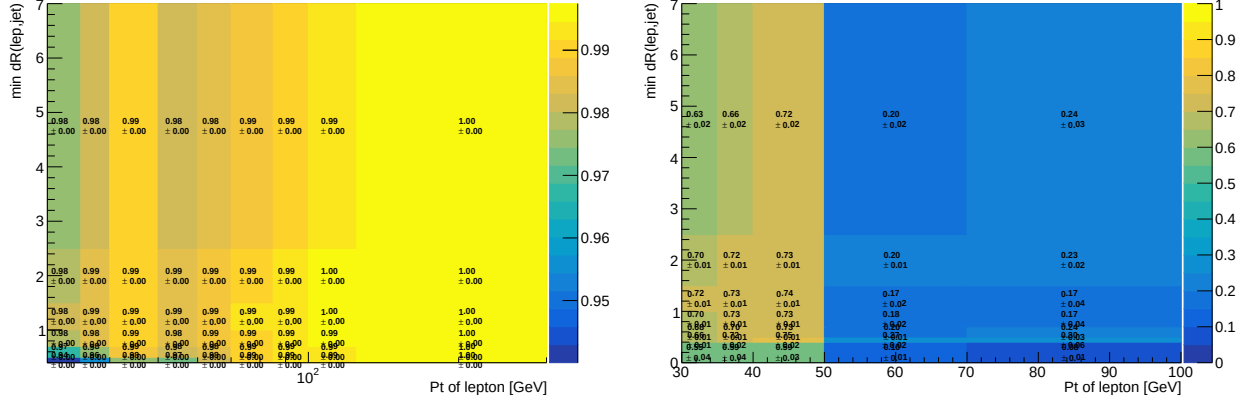
6.5.1.4 Validation of QCD multi-jet modeling

To study the performance of the QCD multi-jet prediction, the distributions of several observables have been validated in the $CR_{\text{QCD}}^{4j,1b}$ and two validation regions, $VR1_{\text{QCD}}^{4j,1b}$ and $VR2_{\text{QCD}}^{4j,1b}$ built using the same selection considered in $CR_{\text{QCD}}^{4j,1b}$, except the cuts on E_T^{miss} and m_T^W :

- $VR1_{\text{QCD}}^{4j,1b}$: $E_T^{\text{miss}} < 20 \text{ GeV}$ and $E_T^{\text{miss}} + m_T^W > 60 \text{ GeV}$
- $VR2_{\text{QCD}}^{4j,1b}$: $E_T^{\text{miss}} > 20 \text{ GeV}$ and $E_T^{\text{miss}} + m_T^W < 60 \text{ GeV}$

Good agreement between data and the prediction is achieved in the control and both validation regions for the muon and electron channels as can be seen in Figures 6.15 and 6.16.

This closure test indicates that the parameterization of real and fake rates is sufficient to provide a good shape prediction. The W+jets background normalization explained in the next sub-section requires a QCD multi-jet estimation in a control region without b-tagged jets (0bCR). Hence dedicated rates have been derived so that they can be applied for the estimation of non-prompt lepton background used in the calculation of the normalization scale factors of the W+jets background.


 (a) $\epsilon(p_T, \min \Delta R)$, muons

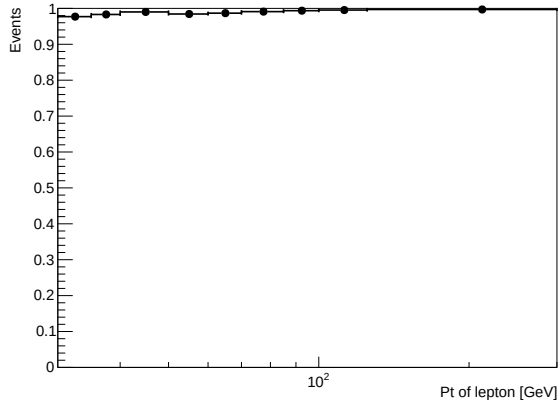
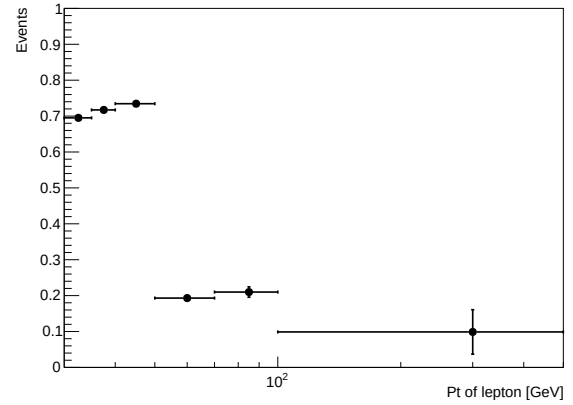
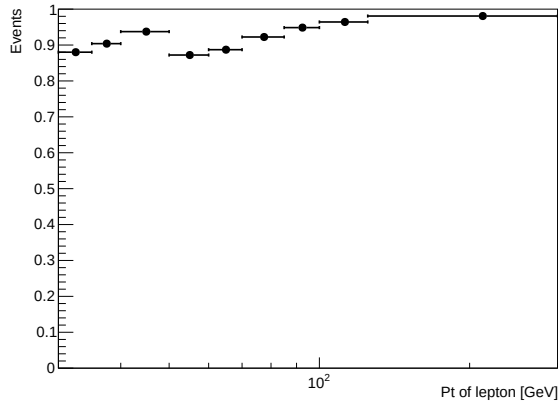
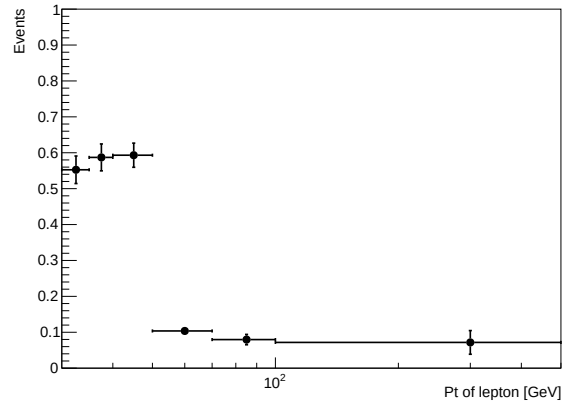
 (b) $f(p_T, \Delta R)$, muons

 (c) $\epsilon(p_T)$, for $\Delta R(\mu, \text{jet}) > 0.4$

 (d) $f(p_T)$, for $\Delta R(\mu, \text{jet}) > 0.4$

 (e) $\epsilon(p_T)$, for $\Delta R(\mu, \text{jet}) < 0.4$

 (f) $\epsilon(p_T)$, for $\Delta R(\mu, \text{jet}) < 0.4$

Figure 6.11: Fake rates f (right-hand) and real rates ϵ (left-hand) for the loose μ to satisfy the tight selection, as a function of its p_T and $\min \Delta R(\mu, \text{jets})$, measured in $CR_{\text{QCD}}^{4j, 1b}$.

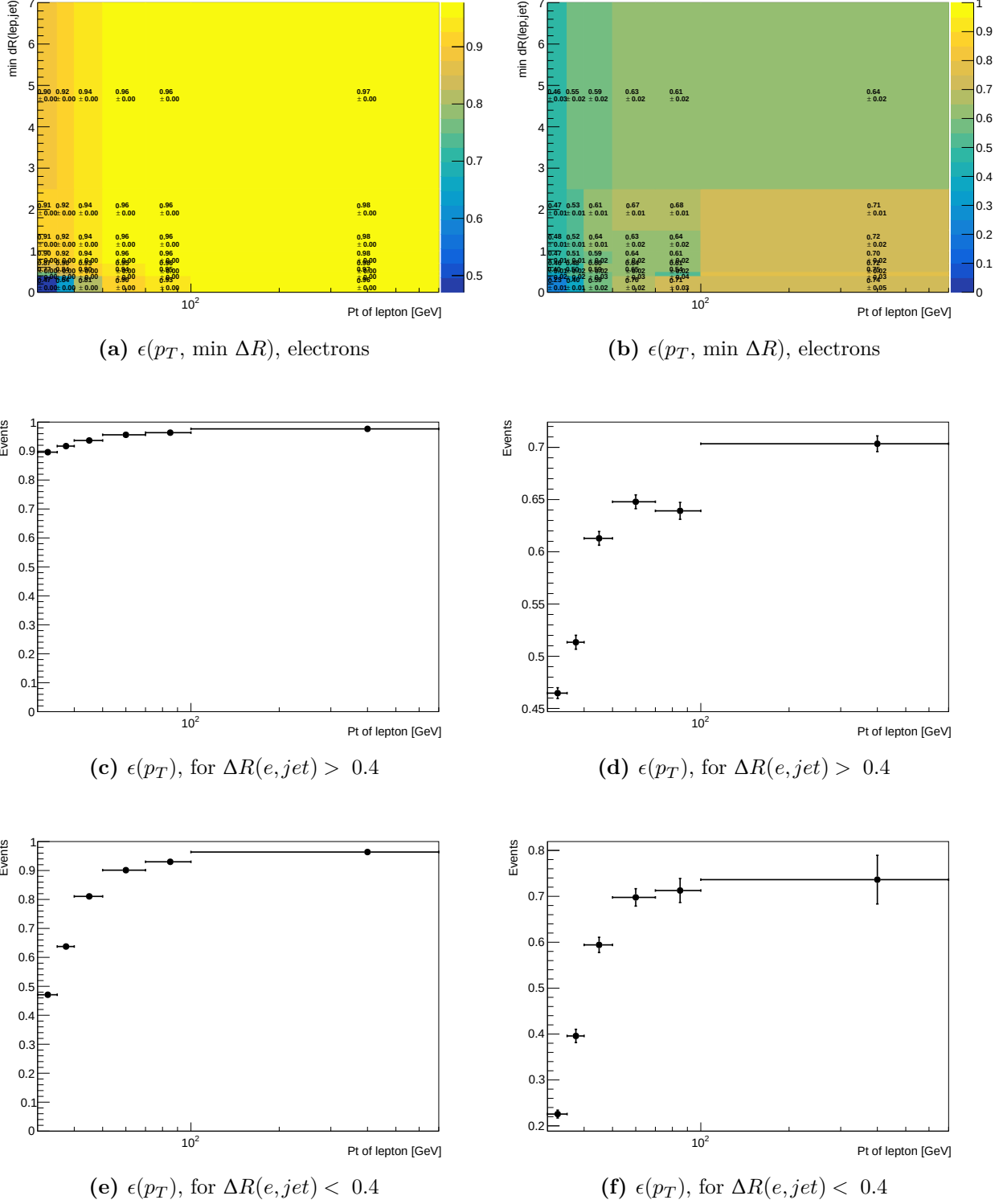


Figure 12: Fake rates f (right-hand) and real rates ϵ (left-hand) for loose e to satisfy the tight selection, as a function of its p_T and $\min \Delta R(e, jets)$, measured in $CR_{\text{QCD}}^{4j,1b}$.

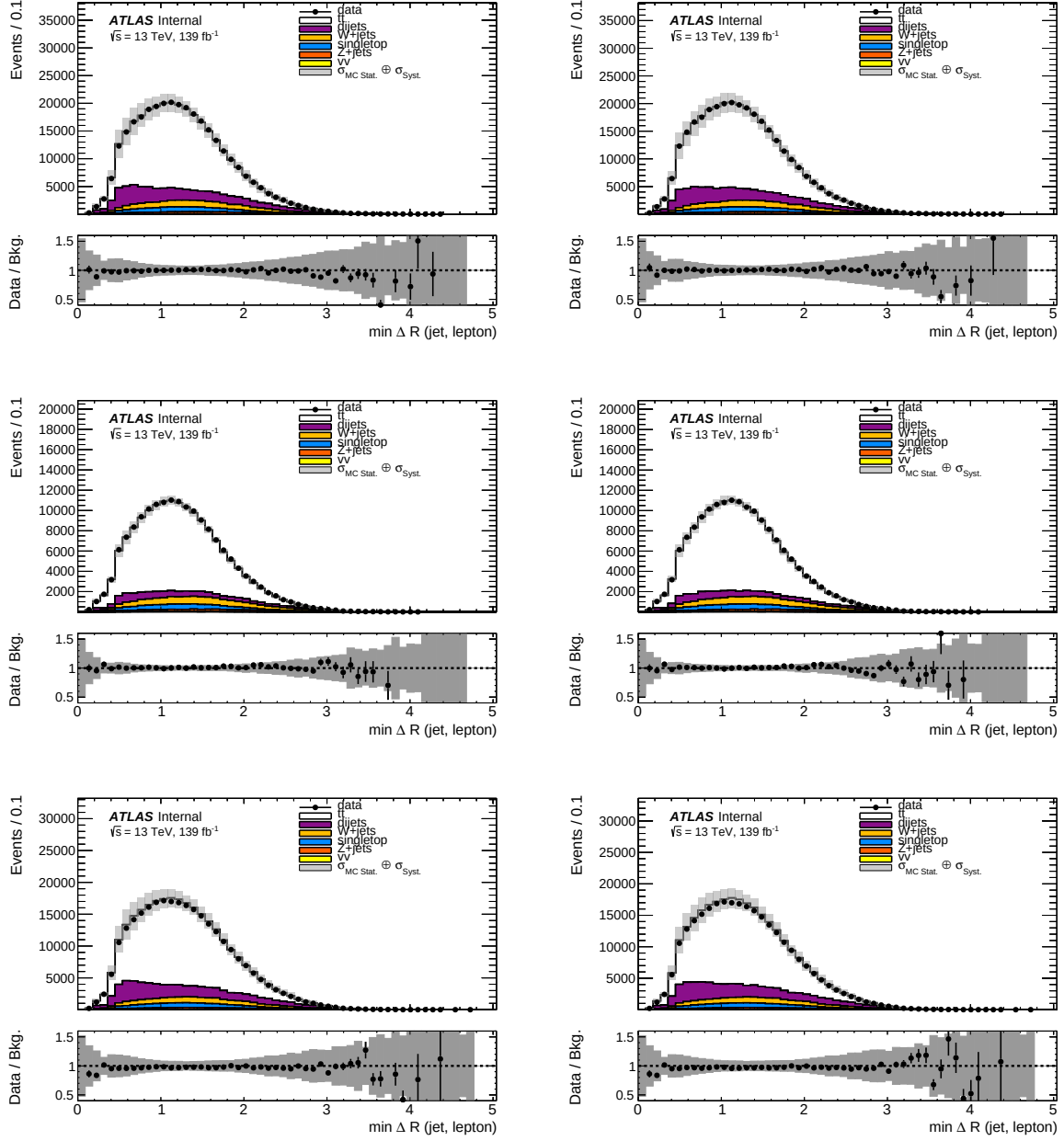


Figure 6.13: Distribution of ΔR (left-hand), ΔR_{corr} (right-hand) in the $CR_{\text{QCD}}^{4j,1b}$ (top) $VR1_{\text{QCD}}^{4j,1b}$ (middle) and $VR2_{\text{QCD}}^{4j,1b}$ (bottom) for the muon channel. In addition to the statistical uncertainty, the gray band represents a flat systematic error band of 50% linked to the QCD multi-jet.

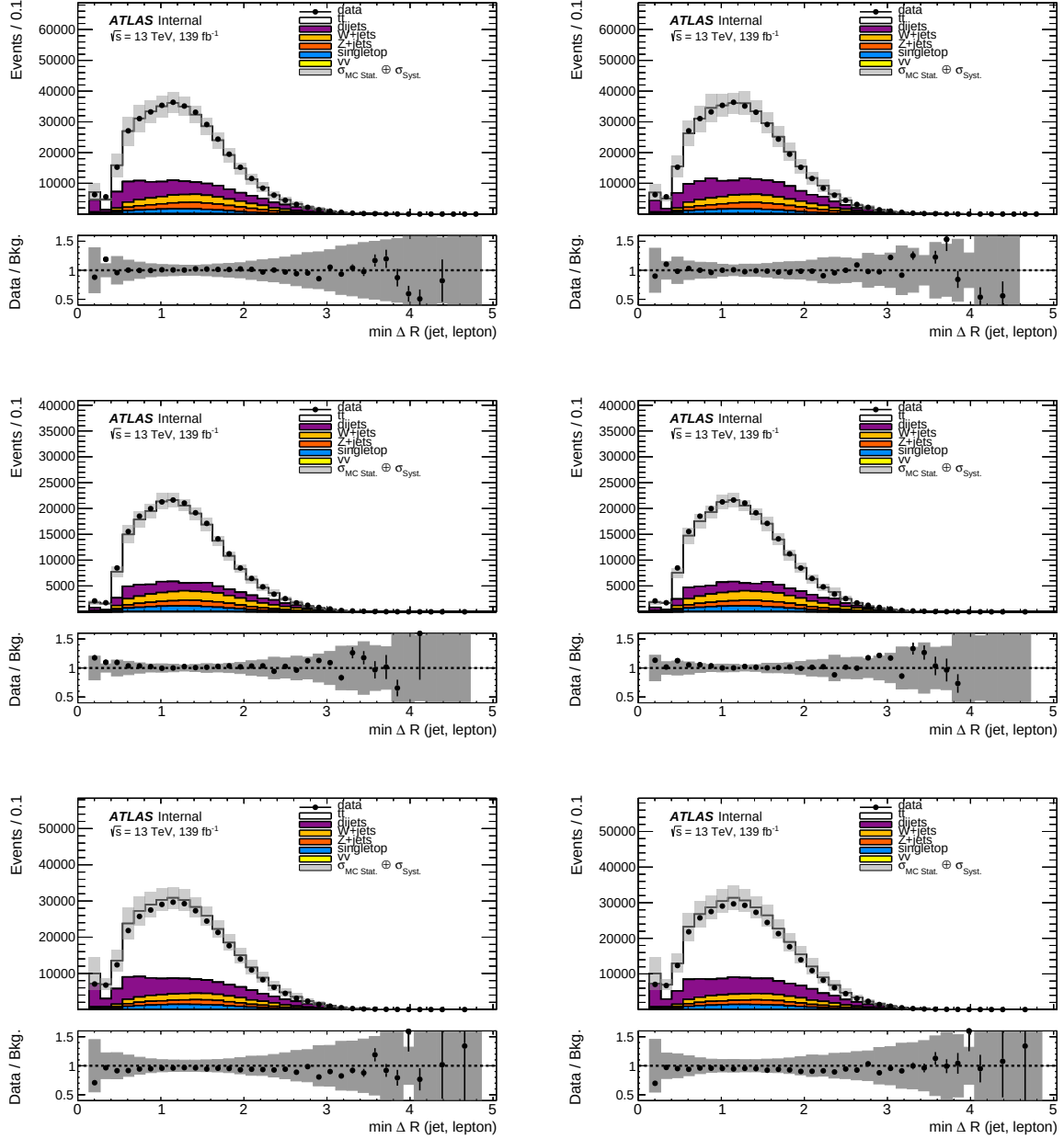


Figure 6.14: Distribution of ΔR (left-hand), ΔR_{corr} (right-hand) in the $CR_{\text{QCD}}^{4j,1b}$, (top) $VR1_{\text{QCD}}^{4j,1b}$ (middle) and $VR2_{\text{QCD}}^{4j,1b}$ (bottom) for the electron channel. In addition to the statistical uncertainty, the gray band represents a flat systematic error band of 50% linked to the QCD multi-jet.

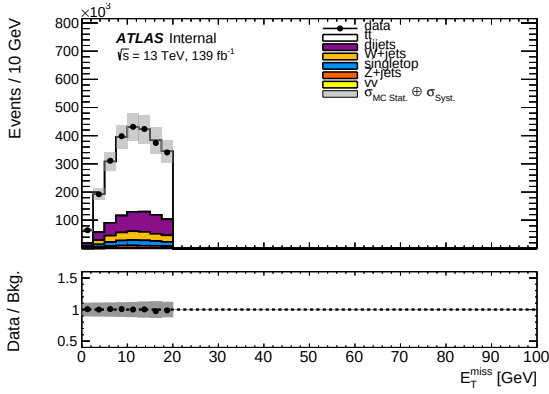
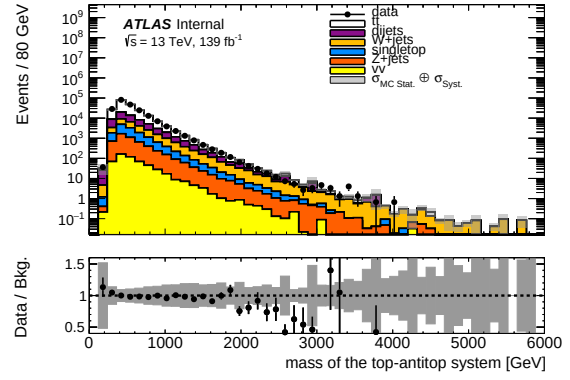
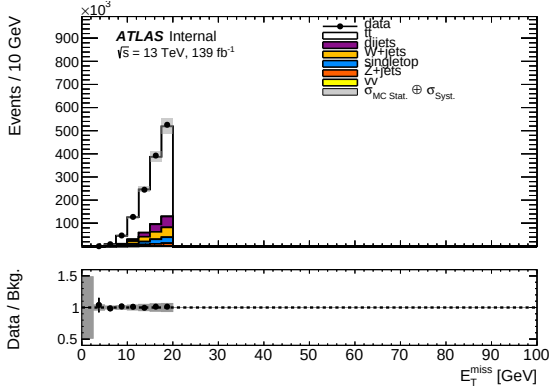
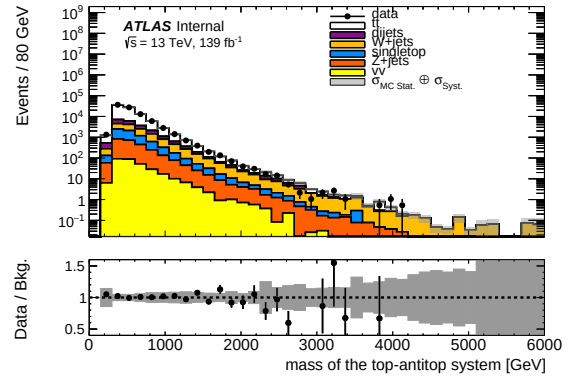
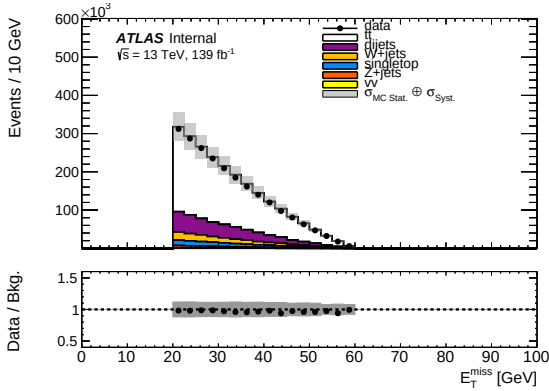
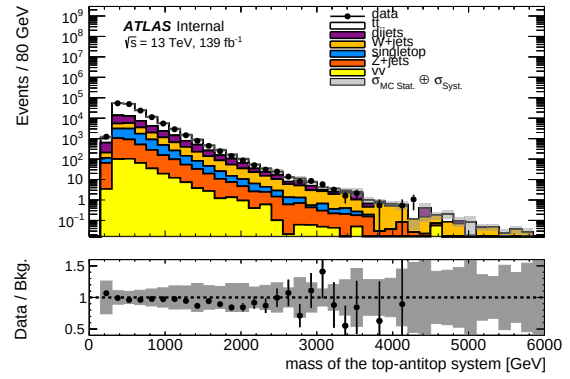

 (a) Lepton p_T - $CR_{\text{QCD}}^{4j,1b}$ for μ channel

 (b) $m_{t\bar{t}}$ - $CR_{\text{QCD}}^{4j,1b}$ for μ channel

 (c) Lepton p_T - $VR1_{\text{QCD}}^{4j,1b}$ for μ channel

 (d) $m_{t\bar{t}}$ - $VR1_{\text{QCD}}^{4j,1b}$ for μ channel

 (e) Lepton p_T - $VR2_{\text{QCD}}^{4j,1b}$ for μ channel

 (f) $m_{t\bar{t}}$ - $VR2_{\text{QCD}}^{4j,1b}$ for μ channel

Figure 6.15: E_T^{miss} (left-hand), $m_{t\bar{t}}$ (right-hand) for the muon channel. In addition to the statistical uncertainty, the gray band represents a flat systematic error band of 50% linked to the QCD multi-jet.

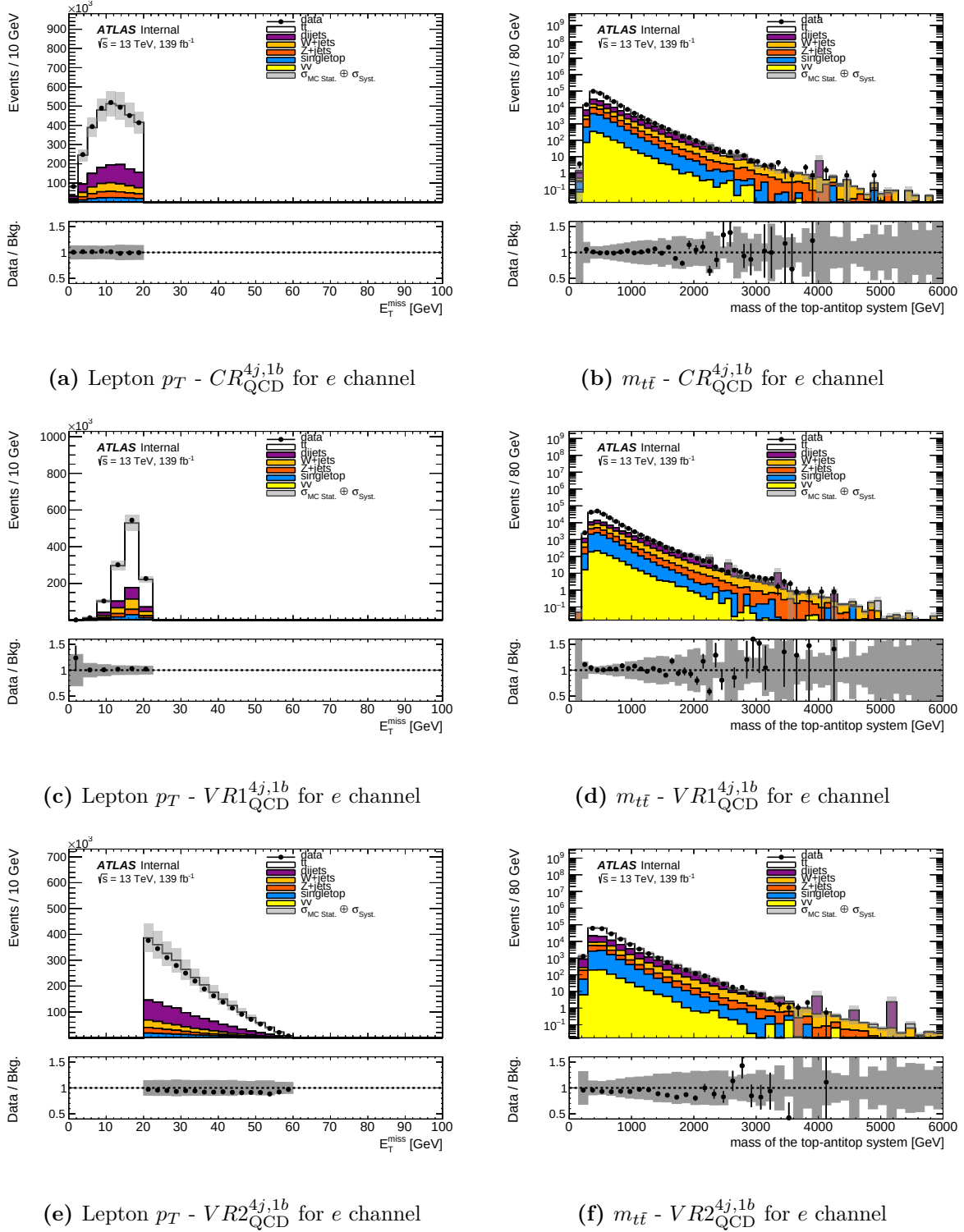


Figure 6.16: E_T^{miss} (left-hand), $m_{t\bar{t}}$ (right-hand) for the electron channel. In addition to the statistical uncertainty, the gray band represents a flat systematic error band of 50% linked to the QCD multi-jet.

6.5.2 W +jets background normalisation

The expected background from W +jets is estimated using Sherpa MC samples, with data-driven scale factors which correct the normalization. The scale-factors are computed independently for each signal regions as the fraction of heavy flavor may be impacted by the selections. Only the overall normalization correction is considered in the analysis.

The W boson is theoretically well understood, having been studied in great detail at previous electron-positron colliders such as LEP and proton-anti-proton colliders such as the Tevatron. At the LHC, the W boson, which comes in both positive and negative charge, is predominantly produced via Drell-Yan annihilation processes, in which a quark and anti-quark annihilate, creating a W boson. In this scenario the W +jets cross-section has a large theoretical uncertainty, furthermore the process has a significant contribution on the final yields. In order to cope with this, the MC simulation is renormalized by usage of a data driven method that offers a more precise prediction. The approach consists on measuring the asymmetry for the generation of the bosons W^+ and W^- [106], where the ratio is theoretically known. In both simulation and data, the normalisation factor is identified by comparing the W -boson charge asymmetry. The measured ratio of cross-section, $\sigma_{W^+}/\sigma_{W^-}$ is around 1.46. This number is also theoretically well modelled. In this analysis, the overall normalization f_W is obtained from the events of a signal region i , by exploiting the fact that the W charge-asymmetry in W +jet production is predicted with better precision than the overall normalization. The advantage of this method is that flavor symmetric backgrounds like $t\bar{t}$, QCD multi-jet and the Z +jets, the charge asymmetry is zero. The charge asymmetry is defined as the difference between the number of events with W^+ and W^- :

$$A^W = N^+ - N^- , \quad (6.10)$$

where $N^\pm$ represent the number of events with negative or positive charged lepton. The charge asymmetry is calculated for data ($N^{\pm \text{ data}}$) by subtracting the MC contributions from all the charge-asymmetric processes ($N^{\pm \text{ MC, asym}}$) like $t\bar{t} + V$, singletop (s -, t -channel) and diboson. So, we define corrected number of data events with positive and negative charged leptons as following

$$N^{\pm \text{ data, corr}} = N^{\pm \text{ data}} - N^{\pm \text{ MC, asym}} \quad (6.11)$$

The charge-asymmetry in corrected data is compared with that of the W +jets pro-

cess from MC simulation to calculate the charge-asymmetry normalization. The charge-asymmetry normalization in the signal region i is used as the W +jets normalization and can be written as:

$$f_i^W = \frac{N_i^{W+jets, \text{ data}}}{N_i^{W+jets, \text{ MC}}} \approx \frac{A_i^{W, \text{ data}}}{A_i^{W, \text{ MC}}} = \frac{N_i^{+, \text{ data, corr}} - N_i^{-, \text{ data, corr}}}{N_i^{+, \text{ MC, W+jets}} - N_i^{-, \text{ MC, W+jets}}} \quad (6.12)$$

The Charge-asymmetry normalization factors thus obtained for the various signal regions are: 1.30 ± 0.05 for the resolved channel and 1.38 ± 0.07 for the boosted channel.

W+jets modelling study

In order to validate the modelling of the W +jets process, dedicated control regions (CR), dominated by the W +jets process, are defined. The CRs are defined as requiring exactly zero b -tagged PFlow jets in the event (0bCRs), keeping all the other cuts the same. There are four CRs for muon and electron channels in the resolved and boosted final states.

Figure 6.17 (left-hand) shows the pre-fit event yield distribution comparing data and background in the 0bCRs. There is a disagreement between the data and the background especially in the muon channel. A background-only fit is performed with four normalization factors for the W +jets background. None of the systematic uncertainties are included in the fit. The post-fit distribution are shown in Figure 6.17 (right-hand) where data matches closely to the total background.

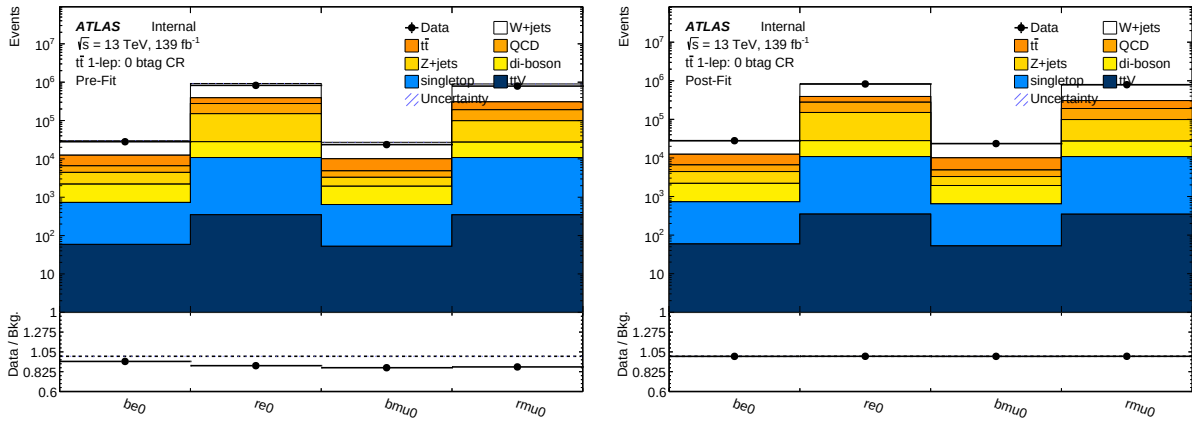


Figure 6.17: Pre(post)-fit comparison between data and background in the 0bCRs.

6.6 Event Reconstruction

The $t\bar{t}$ events satisfying the aforementioned criteria are considered for calculating the $t\bar{t}$ invariant mass, $m_{t\bar{t}}$, which is built from the physics objects included in the event that are matched with top quark decays. The invariant mass is calculated by adding the four-momenta of each item associated with a $t\bar{t}$ decay product to obtain the four-momentum of the $t\bar{t}$ pair. As a result of final state radiation, this method yields in an offset among the reconstructed mass of the decay products and the mass of the produced $t\bar{t}$ pair. This effect, at simulation level, can be seen in Figure 6.18, on which the generated Z'_{TC2} mass before FSR is displayed and compared to the mass of the $t\bar{t}$ system decay products taken at partonic-truth level after the FSR. This limitations may diminish the sensitivity of the analysis, notably for the strongly radiating highly boosted top-quarks from a heavy resonance.

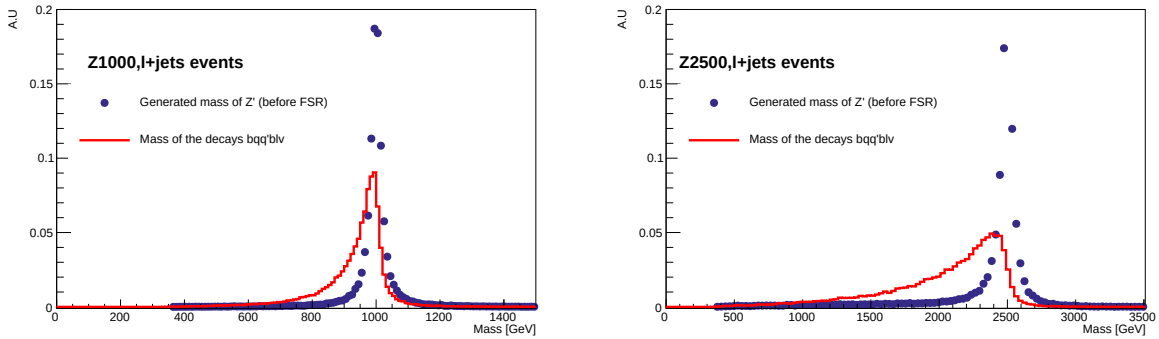


Figure 6.18: Generated Z'_{TC2} and invariant mass of the top quarks decay products at the partonic level for the 1.0 TeV (left-hand) and 2.5 TeV (right-hand) resonances.

6.6.1 Neutrino Reconstruction

In both regimes, the first reconstruction step is the determination of the four-momentum of the neutrino. So, we assume the transverse momentum of a neutrino, p_T^ν , is same as the missing transverse energy, E_T^{miss} . The longitudinal momentum component, p_z^ν , cannot be measured directly since the initial p_z of the centre-of-mass frame of the hard interaction with respect to the laboratory frame is unknown. To determine the neutrino vector z-component, we use momentum conservation

$$p_W = p_\ell + p_\nu \quad (6.13)$$

where p_ℓ is the lepton 4-vector and p_ν is the neutrino 4-vector. We also assume the mass of a neutrino to be zero, $m_\nu = 0$. Which implies $p_\nu^2 = 0$. We further assume that the lepton and the neutrino originating from a W boson decay and impose the W mass constraint on the lepton-neutrino system. We get the following equation from the momentum conservation and on-shell W decay approximation.

$$m_W^2 = (p_\ell + p_\nu)^2 = m_\ell^2 + 2p_\ell \cdot p_\nu \quad (6.14)$$

Solving this equation we get the neutrino momentum z-component, $p_{z,\nu}$ in terms of the other known quantities. We find

$$p_{z,\nu} = \frac{4c_1 p_{z,\ell} \pm \sqrt{(4c_1 p_{z,\ell})^2 - 4.4(E_\ell^2 - p_{z,\ell}^2) \cdot 4(p_{T,\nu}^2 E_\ell^2 - c_1^2)}}{2.4(E_\ell^2 - p_{z,\ell}^2)} \quad (6.15)$$

where

$$c_1 = m_W^2 - m_\ell^2 + 2(p_{x,\ell} p_{x,\nu} + p_{y,\ell} p_{y,\nu}). \quad (6.16)$$

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. If two real solutions are found, the one with the smallest $|p_z|$ is used in the boosted-topology case. In the resolved-topology reconstruction, the choice is made by the χ^2 algorithm.

6.6.2 Boosted channel topology

In the boosted topology, large- R jet with the highest transverse momentum and top tagged is marked as the hadronic top quark, while the selected jet, lepton (muon or electron) and neutrino are taken to form the leptonic top quark. Finally, $m_{t\bar{t}}$ is calculated by adding together the four-momenta of these four objects.

6.6.3 Resolved channel topology

The reconstruction of events in the resolved-topology is more complicated since there are several manner to select and assign three small- R jets to the hadronic top quark (t_{had}) and one small- R jet have to be linked to the b-jet from the leptonic top quark (t_{lep}). In order to do this, a χ^2 value is associated to every probable mixture of four jets in the event. The χ^2 value evaluates the compatibility of a specific mixture of jets is with resolved $t\bar{t}$ decay using the constraints from W boson, the expected top quark masses and the 4-vectors of the

reconstructed physics objects.

$$\chi^2 = \underbrace{\left[\frac{m_{jj} - m_W}{\sigma_W} \right]^2}_{\text{hadronic W}} + \underbrace{\left[\frac{m_{jjb} - m_{jj} - m_{th-W}}{\sigma_{th-W}} \right]^2}_{\text{hadronic top}} + \underbrace{\left[\frac{m_{jl\nu} - m_{tl}}{\sigma_{tl}} \right]^2}_{\text{leptonic top}} + \underbrace{\left[\frac{(p_{T,jjb} - p_{T,jl\nu}) - (p_{T,th} - p_{T,tl})}{\sigma_{diffpT}} \right]^2}_{\text{top quark transverse momentum}} \quad (6.17)$$

The first term compares the mass of the two jets associated to the W boson, m_{jj} with the expected mass m_W . The second term compares the mass of the reconstructed top that decay hadronically (m_{jjb}) to expectation, but since m_{jj} and m_{jjb} are heavily correlated, m_{jj} is subtracted from m_{jjb} and this is compared to the expected mass of t_{had-W} , m_{th-W} . The third term compares the mass of the reconstructed leptonically decaying top quark t_{lep} , $m_{jl\nu}$, to the expected mass m_{tl} , and the last term compares the p_T difference between the two reconstructed tops, $p_{T,jjb} - p_{T,jl\nu}$, to the expected p_T difference, $p_{T,th} - p_{T,tl}$. The parameter values are determined from reconstructed MC events in which the semi-leptonic $t\bar{t}$ decay products have been matched to the reconstructed physics objects using truth level information. This is done on a mix of Z'_{TC2} samples of masses from 0.5 to 2 TeV, as we want to optimize the algorithm for a search in this range. The values of the parameters are: $m_W = 80.36$ GeV, $m_{th-W} = 86.12$ GeV, $m_{tl} = 165.95$ GeV, $\sigma_W = 12.50$ GeV, $\sigma_{th-W} = 16.64$ GeV, $\sigma_{tl} = 26.03$ GeV, $p_{T,th} - p_{T,tl} = -3.86$ GeV and $\sigma_{diffpT} = 37.42$ GeV where the masses and Δp_T parameter values represent the means of the Gaussians fitted to the relevant distributions, and the σ values are the standard deviations. In this analysis, all possible jet permutations are tried and only the permutation with the lowest χ^2 value is used.

6.6.4 Performance of the reconstruction

The effectiveness of matching the different partons from the top quarks decay to the reconstructed objects after the full selection of the resolved and the boosted channel is shown in Figure 6.19. The matching is evaluated in term of the fraction of the reconstructible Monte Carlo events, i.e. semileptonic $t\bar{t}$ events where all the decays products of the $t\bar{t}$ quarks are detected in the acceptance of the detector therefore they are matched to the reconstructed physics objects. This efficiency of matching is measured as a function of the pole mass of the Z'_{TC2} signal form 0.4 to 5 TeV. The events for which at least one of the partons from the decay of the top-quark escapes the detection are not considered to evaluate the

reconstruction performances. The b-partons (b_H and b_L)⁴ must be associated to truth jets also those truth jets are demanded to be associated to reconstructed jets. At each step, the matching is performed using a cone of radius $\Delta R = 0.4$. In order to consider the hadronic W boson (W_H) as reconstructible, its both decay products have to be linked to a reconstructed jet using a cone of radius $\Delta R = 0.4$. Similarly, the hadronic top is considered reconstructible if b_H and W_H are reconstructible for the resolved case, or if it is matched to a large-R jet using a cone of radius 1 for the boosted case. The total fraction of "reconstructible" events varies from 60 to 70% for the resolved selection, and between 80 and 95% for the boosted selection.

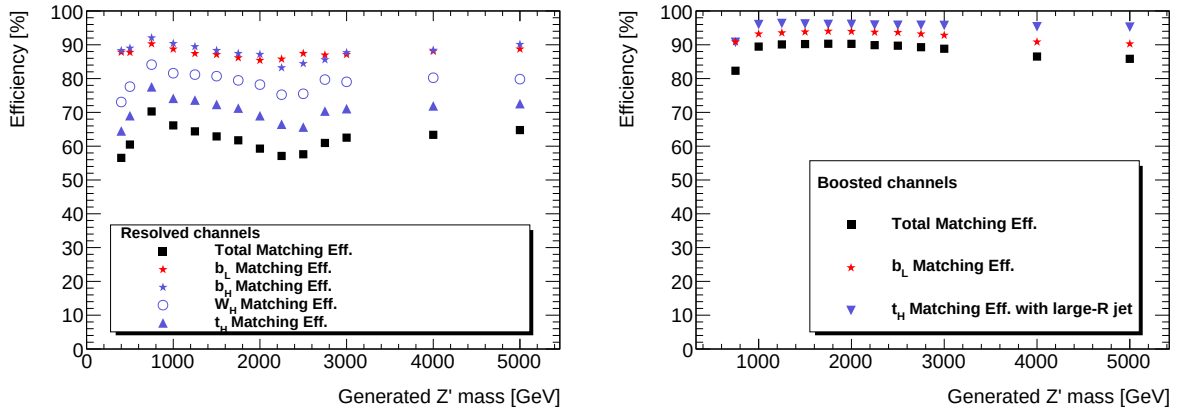


Figure 6.19: Efficiencies of matching the different partons from top-quark decays as function of the pole mass of the Z'_{TC2} samples, after the full selection of the resolved (left-hand) and boosted (right-hand) selections, before the selection on the b-tagging category.

The performance of the reconstruction algorithms are evaluated among the reconstructible events as the efficiency to select the correct jet combination which minimize the χ^2 algorithm for the resolved regime, and when the partons are associated exactly to the reconstructed jets instead of the ΔR matching for the boosted regime (Figure 6.20). For the resolved reconstruction, a swap is allowed between the partons of the hadronic top as it leads to the same result from the analysis point of view. The performance of the reconstruction is split based on four b-tagging categories. As expected the category 3 has the best performance with b-tagged jets assigned to the leptonic and the hadronic b-parton. The category 1, with a b-tagged jet assigned only to leptonic b-jet, has the a good performance to find the correct jet for b_L , while category 2 (b-tagged jet assigned only to hadronic b-jet) has better performance for the b_H side.

⁴L and H differentiate the leptonic top from the hadronic one.

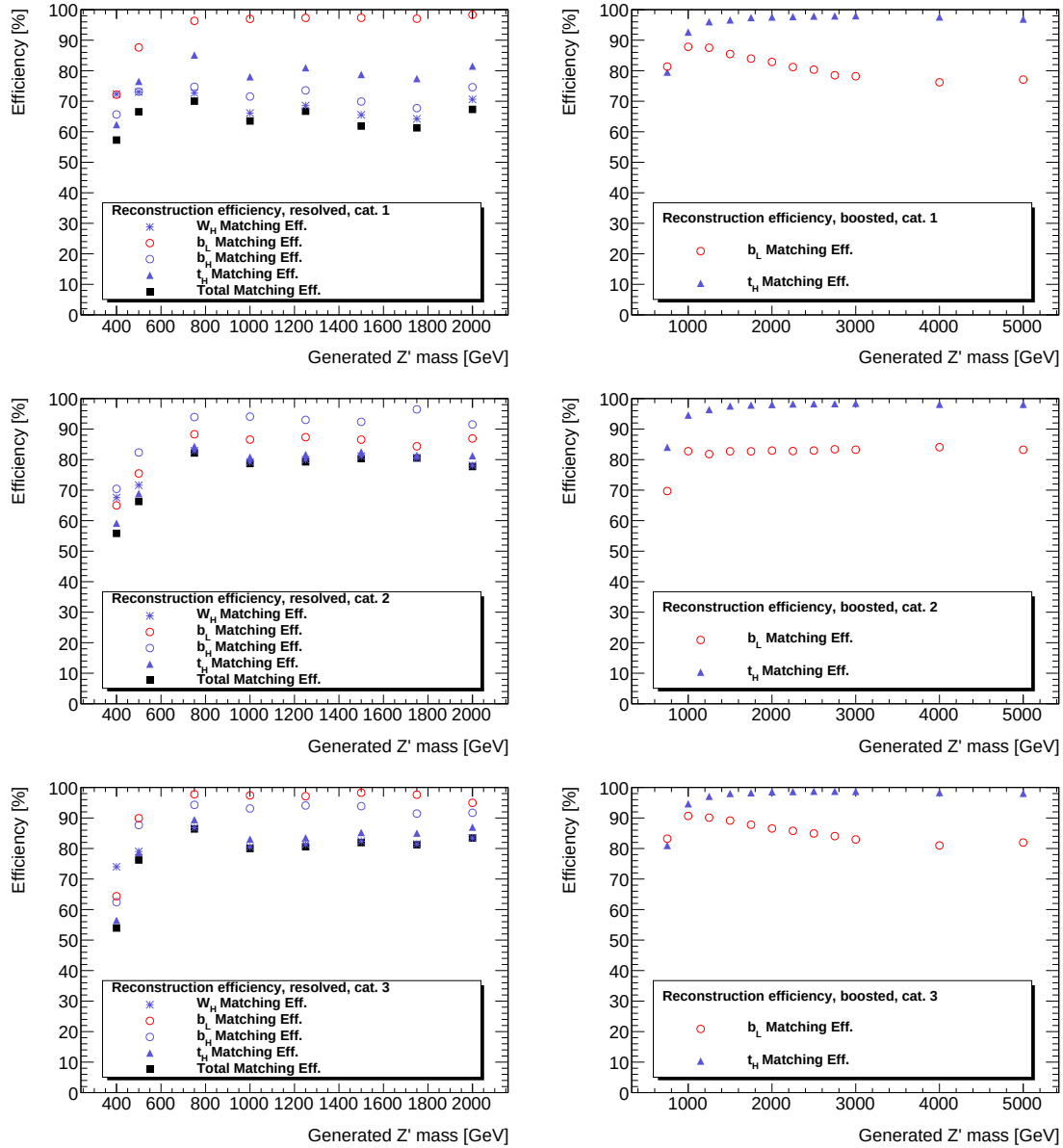


Figure 6.20: Efficiency to select the correct combination form reconstructible eventss for the resolved (left-hand) and boosted (right-hand) selections as function of the pole mass of the Z'_{TC2} samples, for the categories 1 (top), 2 (middle) and 3 (bottom).

6.6.5 $t\bar{t}$ invariant mass spectrum

The reconstructed $t\bar{t}$ invariant mass ($m_{t\bar{t}}^{reco}$) distributions of Z'_{TC2} signal samples for all events fulfilling the resolved and boosted topology selections are shown in Figure 6.21. The mass distributions peak at the resonance mass of its corresponding sample. Most dis-

tributions exhibit tails towards lower masses which are especially significant for broader resonances and larger resonance masses. They are caused mainly by the final-state radiation effect (FSR) from any of the objects in the $t\bar{t}$ system that is not included in the reconstruction. This effect is particularly important at high resonance masses where the top quarks and their decay products are more energetic and therefore the probability for the emission of hard gluon radiation is higher. The true $t\bar{t}$ invariant mass after FSR, is defined as the combined invariant mass of the six $t\bar{t}$ decay products at generator level. The $m_{t\bar{t}}^{reco}$ is displayed in term of true $t\bar{t}$ invariant mass in Figure 6.22. High correlation between these two variables is seen for both resolved and boosted regimes.

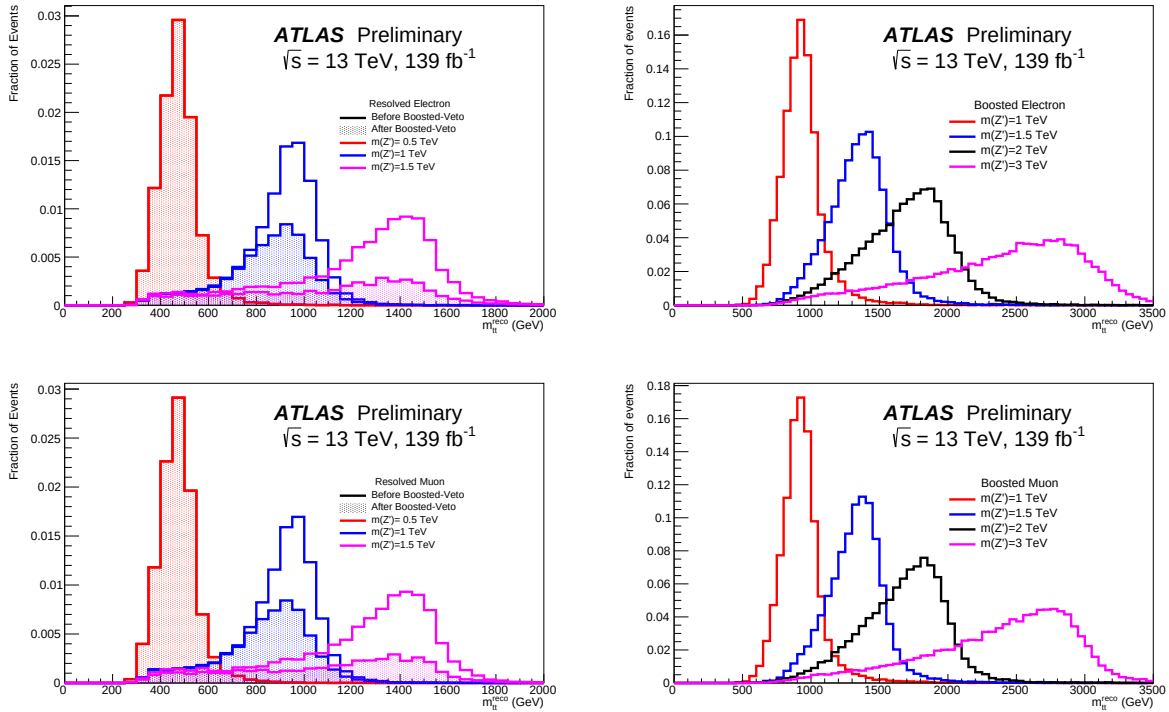


Figure 6.21: Reconstructed $t\bar{t}$ invariant mass distributions for the Z'_{TC2} signal model considered in the search. The events satisfy the criteria of the resolved-topology selection before (empty) or after (filled) the veto on events that also pass the boosted-topology criteria. All distributions are normalised to unit area.

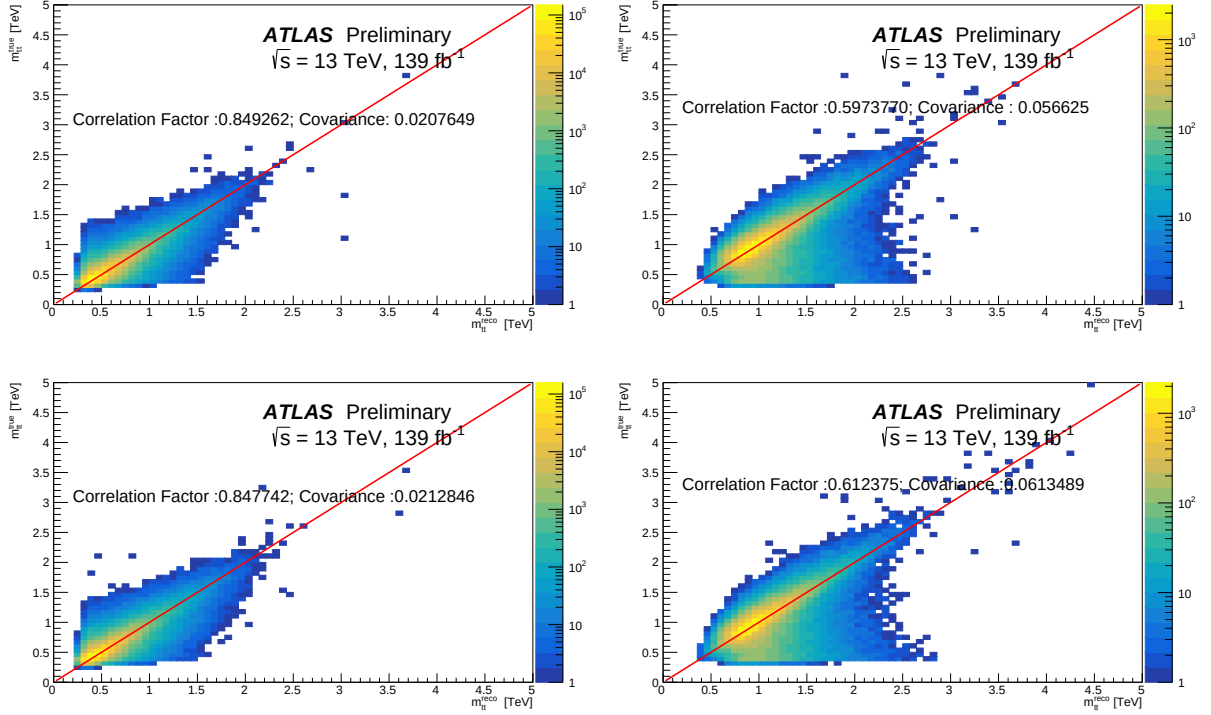


Figure 6.22: Resonstracted $t\bar{t}$ versus true $t\bar{t}$ invariant mass for boosted electron (top-right), resolved electron (top-left), boosted muon (bottom-right) and resolved muon (bottom-left) for SM $t\bar{t}$ production. Both reconstructible and non-reconstructible events are included.

6.7 Systematic Uncertainties

Several sources of systematic uncertainties may alter the predicted normalization and distribution of the signal and the background $m_{t\bar{t}}^{\text{reco}}$ distribution. The impact of systematic variations on the event yield must be quantified in order to be taken into account like a nuisance parameter while setting the limit process. This section describes the details of the main sources of systematic uncertainties considered to validate the agreement of the observed data with the prediction. Table 6.3 summarizes the effect of the most considerable systematic uncertainties on the total number of events in the resolved and boosted channels.

6.7.1 Experimental uncertainties

Various sources of experimental uncertainties are taken into account. As a result of these uncertainties, both the the normalisation and the shape of signal and background discriminant distributions are impacted as evaluated in dedicated studies, described below.

Source	Boosted el (%)		Boosted mu (%)		Resolved el (%)		Resolved mu (%)	
	Bkg	4 TeV Z'	Bkg	4 TeV Z'	Bkg	4 TeV Z'	Bkg	4 TeV Z'
Anti- k_t R = 0.4 (Comb)	5.19	0.64	4.87	0.32	31.41	6.03	32.34	4.77
Anti- k_t R = 1.0 (Comb)	4.40	0.12	4.54	0.48	0.09	0.61	0.09	1.06
E_T^{miss}	1.63	0.02	1.01	0.06	2.73	0.23	2.06	0.32
EGamma	3.02	2.47	0.01	0.00	2.08	0.97	0.00	0.00
Muon	0.00	0.00	13.18	1.30	0.02	0.00	2.96	0.77
b-tagging	4.68	1.03	4.66	2.43	5.80	0.54	5.75	0.77
top-tagging	6.67	8.08	5.75	11.02	0.13	3.31	0.12	3.90
$t\bar{t}$ modelling, HS	2.91	-	2.57	-	< 1.0	-	< 1.0	-
$t\bar{t}$ modelling, had	1.89	-	2.48	-	2.62	-	2.18	-
$t\bar{t}$ modelling, hdamp	1.29	-	12.56	-	1.42	-	< 1.0	-
$t\bar{t}$ μ_R Scale	8.98	-	9.18	-	10.42	-	10.50	-
$t\bar{t}$ μ_F Scale	3.57	-	3.69	-	1.99	-	2.10	-
$t\bar{t}$ ISR (α_S)	<1.0	-	<1.0	-	<1.0	-	< 1.0	-
$t\bar{t}$ FSR (α_S)	3.47	-	3.24	-	2.13	-	2.04	-
$t\bar{t}$ pdf	< 1.0	-	< 1.0	-	< 1.0	-	< 1.0	-
$t\bar{t}$ NNLO	< 1.0	-	< 1.0	-	< 1.0	-	< 1.0	-

Table 6.3: Impact of uncertainties on the expected number of total background events and 4 TeV (pole mass) Z'_{TC2} signal events in the four signal regions.

For the analysis presented in this chapter, principal sources of uncertainty are given by the resolution and calibration of the jet energy mass and scale, as well as uncertainties on the b-tagging misidentification and identification efficiencies. Additional details concerning the considered sources of uncertainties are specified in the following:

Luminosity and pile-up

The total uncertainty on the integrated luminosities are 2.0-2.4% for each year of 2015-2018, and 1.7% on the full data sample [107]. It is implemented as a flat uncertainty to each simulated sub-background and signal. The MC samples are reweighted in order to get the same luminosity profile than the data. The uncertainty associated to this re-weighting is extracted from the official tool [108].

The lepton

The lepton reconstruction is connected to many sources of systematic uncertainties. The sources of uncertainty considered in this analysis were the efficiencies related to: the trigger, the reconstruction and the identification, as well as the energy/momentum scale and resolution uncertainties. The modeling of these sources is studied by employing Z boson that decay leptonically in both simulation and data at $\sqrt{s} = 13$ TeV.

Small- R jets

Using the uncertainties for calibrated jets divided into more than 100 baseline nuisance

parameters, the effect of the jet energy scale (JES) uncertainty for small-R jets is evaluated. To reduce the complexity of inputs to the statistics tool, the set with only 30 components is used as uncorrelated nuisance parameters. The jet energy resolution (JER) may only be diversified “up” (worsening of the resolution). As input to the limit setting code, a “down” shift histogram is created by symmetrizing compared to the nominal shape. Bin i in the symmetrized “down” histogram has a $-x\%$ shift compared to the nominal bin value if bin i in the smeared histogram has a $x\%$ shift.

Large- R jets

In a similar method to the small-R jets, large-R jets are assigned systematic uncertainties that impact their energy scale (20 nuisance parameters) and resolution, but also their mass resolution (1 nuisance parameter) and scale (6 nuisance parameters). The systematic uncertainties affecting the small and large- R jet collections are treated as uncorrelated.

b -tagging

The prescription for implementing the b -tagging uncertainty recommended by the flavor tagging performance group has been followed. The eigenvector breakdown approach has been tested to avoid over-profiling. There are 4, 3, and 4 eigenvectors respectively for the b -, c -, and light-flavor quarks, and 2 eigenvectors particularly for the extrapolation of the scale factor into high p_T regimes, which are correlated over all jets flavors.

The reconstruction of the missing transverse energy

The E_T^{miss} is altered consistently when one of its inputs (small-R jets or leptons) is modified by a systematic uncertainty. Moreover 3 systematic uncertainties associated with the soft track term are used [28], as recommended by the JetEtMiss group.

6.7.2 Theory uncertainties

By comparing each theory model’s assumptions to alternative simulated samples or by using event-based variation weights that are built into the nominal generator, it is possible to estimate how each theory model’s assumptions will affect the expected cross sections, acceptance, and kinematic distributions of background and signal events. We assess the three main causes of uncertainty.

The SM $t\bar{t}$ background

The uncertainty affecting the top-antitop cross-section is one of the relevant sources of systematic uncertainty. It has been estimated to be $^{+5.6\%}_{-6.1\%}$. Approximately an uncertainty of 10% is attributed to the correction for EW loops. The correction for electroweak loops has an uncertainty of 10%. As the $t\bar{t}$ process is the dominant background in this search, the

following uncertainties related to its modeling are considered:

- Parton shower (PS) and matrix element (ME) uncertainties are investigated by comparison to alternative samples where the nominal ME or parton shower generator was replaced by an alternative one. ME uncertainties target the calculation of the ME itself as well as the matching of the ME to the parton shower. The PS uncertainties are also investigated based on a comparison to several set of tuned parameters for the same shower generator.
- The effect of uncertainties from a specific PDF set is estimated by means of event reweighting according to each of the 30 eigensets of the PDF4LHC_nlo_30 in an aMC@NLO sample that are used and the relative effect in that sample is taken as a systematic uncertainty.
- The impact of QCD scale uncertainties is estimated by varying both the renormalisation scale μ_r and the factorisation scale μ_f independently by factors of 0.5 and 2 from their default values $\mu_r = \mu_f = m_{top}$. The envelope of the resulting cross-section values is taken as the scale uncertainty. For hadronic processes, variations in the showering parameters are applied in order to account for increased or decreased initial state radiation (ISR) final state radiation (FSR).
- The QCD scale and PDF sources of uncertainty are treated as a single combined “PDF+ α_s ” uncertainty because the conventions concerning the choice of α_s differ between PDF sets. The procedure for deriving this uncertainty is based on the PDF4LHC [109] recommendations, but uses NNLO PDF sets.

Single-top

The cross-section uncertainty for the single top background is estimated to be $\pm 5.3\%$. In addition, the diagram removal method is compared with the diagram subtraction method in the Wt channel and the difference is added as an uncertainty in the single top background. The effect of the uncertainty on the modeling of ISR/FSR has been found to be negligible and has not been used in the fit.

Multi-jets processes

The QCD multi-jet background is evaluated directly from data using a purely data-driven techniques as described in 6.5.1. The specification of the control zone that was utilized for the computation of the fake rate f determines the number of predicted QCD multi-jet background occurrences. The number of expected QCD multi-jet events relies on the

reconstruction of its CR employed for the derivation of the fake efficiency f . A conservative uncertainty of 50% is large enough to cover the data-simulation disagreement in the control and validation regions.

6.8 Comparison Expectation to Data

Even though the major goal of this analysis is the accurate estimation of the $m_{t\bar{t}}$ spectra and eventually employ it to fix a limit on the benchmark models, a study should be performed on the kinematics variables of the events in data, to check either the results are compatible with the SM as a first order. This is predicted since the benchmark models reduce to the SM as an effective theory, therefore any variation in the results may be only incompatible with the SM with a small level. This section presents the comparison of the available data with the MC simulation in the signal region for several kinematic distributions that are not considered directly in the search for BSM resonances. Such observables are expected to show a good consistency between data and the corresponding background prediction within the uncertainties as can be seen in Figures 6.23, 6.24, 6.25 and 6.26 for the resolved and boosted regimes. This proves that these variables employed in the definition of the signal region are well-modelled and that the reconstruction of the $t\bar{t}$ system provides truthful results. The b -tag categories 1, 2 and 3 are included together in these plots. The simulation is scaled to the measured data collected luminosity and the agreement is shown in the lower panel of each plot where the sky-blue band represents the systematic and statistical uncertainties on the background. Table 6.4 presents the observed and expected event yields after the resolved and boosted topology selections for both $e + jets$ and $\mu + jets$ channels. The total uncertainty on the background estimate is derived by summing all systematic uncertainties and the MC statistical uncertainty in quadrature.

In addition, the relative contribution coming from each background process in different SRs is summarized in the pie chart presented in Figure 6.27. The QCD contribution is estimated using the method described in Section 6.5.1. Category-3 is the most pure region and is mostly dominated by $t\bar{t}$ events. Category-0 has the largest contribution from non- $t\bar{t}$ processes out of the four b -tag categories.

Type	Boosted e	Boosted μ	Resolved e	Resolved μ
$t\bar{t}$	146469.0 ± 11643.2	135022.6 ± 10379.6	2227361.8 ± 220826.3	2280214.7 ± 231945.8
W +jets	9071.9 ± 855.9	9064.1 ± 816.7	159342.9 ± 43371.3	168989.5 ± 46553.3
Single top	9344.6 ± 773.7	8372.9 ± 697.0	102555.3 ± 15704.3	103277.7 ± 17319.0
QCD	3596.8 ± 299.4	3284.5 ± 121.6	89525.0 ± 1092.4	90938.7 ± 777.0
Z +jets	1276.5 ± 138.2	811.5 ± 78.3	45519.0 ± 13323.1	27731.4 ± 8203.2
Diboson	876.8 ± 67.6	764.9 ± 67.4	7274.5 ± 1523.3	7130.6 ± 1402.0
Total	170635.6 ± 13065.0	157320.4 ± 11699.2	2631578.3 ± 294887.5	2678282.6 ± 307191.2
Data	166661.0 ± 408.2	151149.0 ± 388.8	2633761.0 ± 1622.9	2703119.0 ± 1644.1

Table 6.4: Observed data and background event yields after the boosted and resolved selections in the $e + jets$ and $\mu + jets$ channels. The total uncertainty on the expected background yields is derived by summing all systematic uncertainties and the MC statistical uncertainty in quadrature.

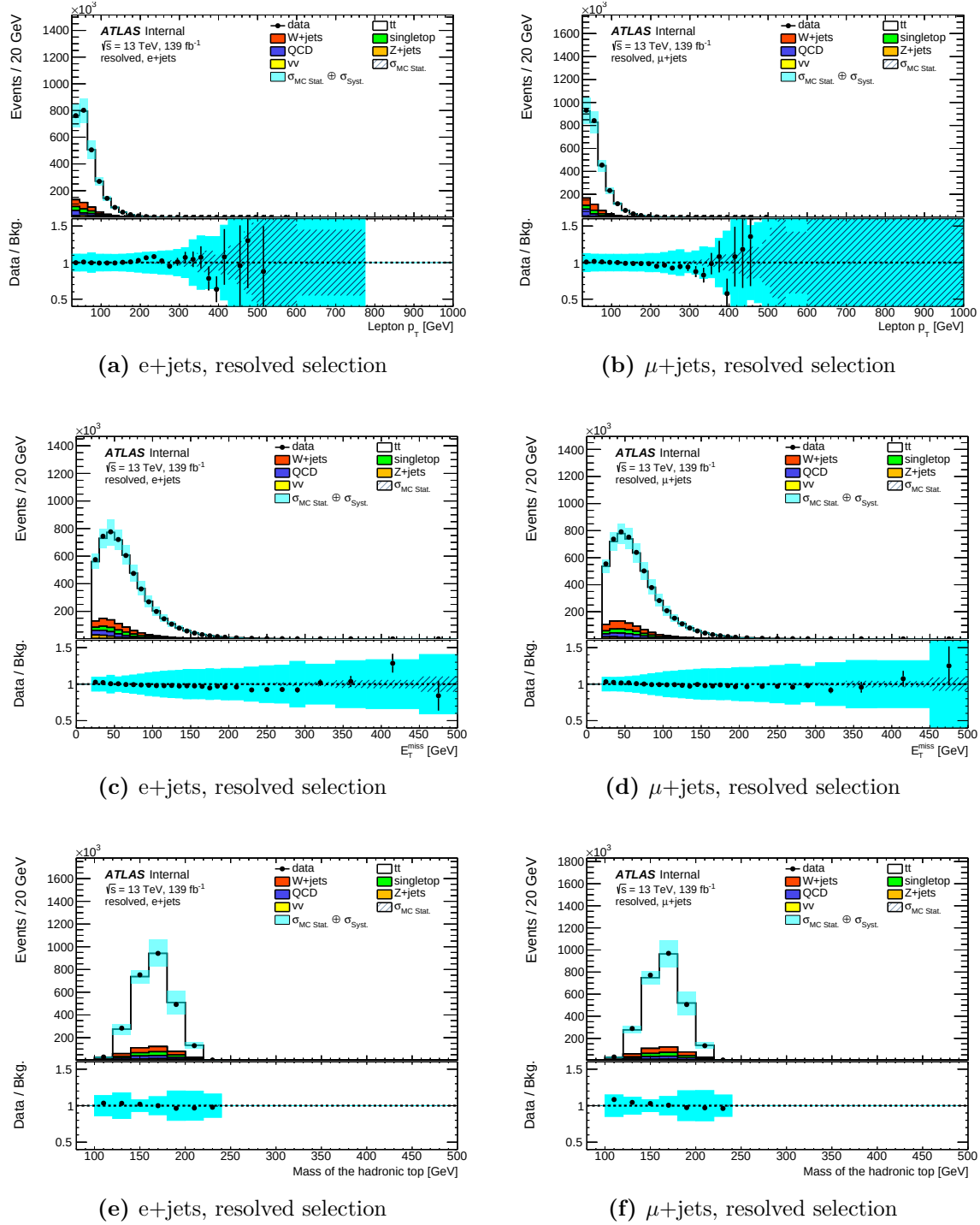


Figure 6.23: The p_T of the lepton (top), (middle) the E_T^{miss} , and (bottom) Mass of the hadronic top distributions in the resolved (a, c, e) electron and (b, d, f) muon channel. The ratio of data and background is mostly within the systematic band.

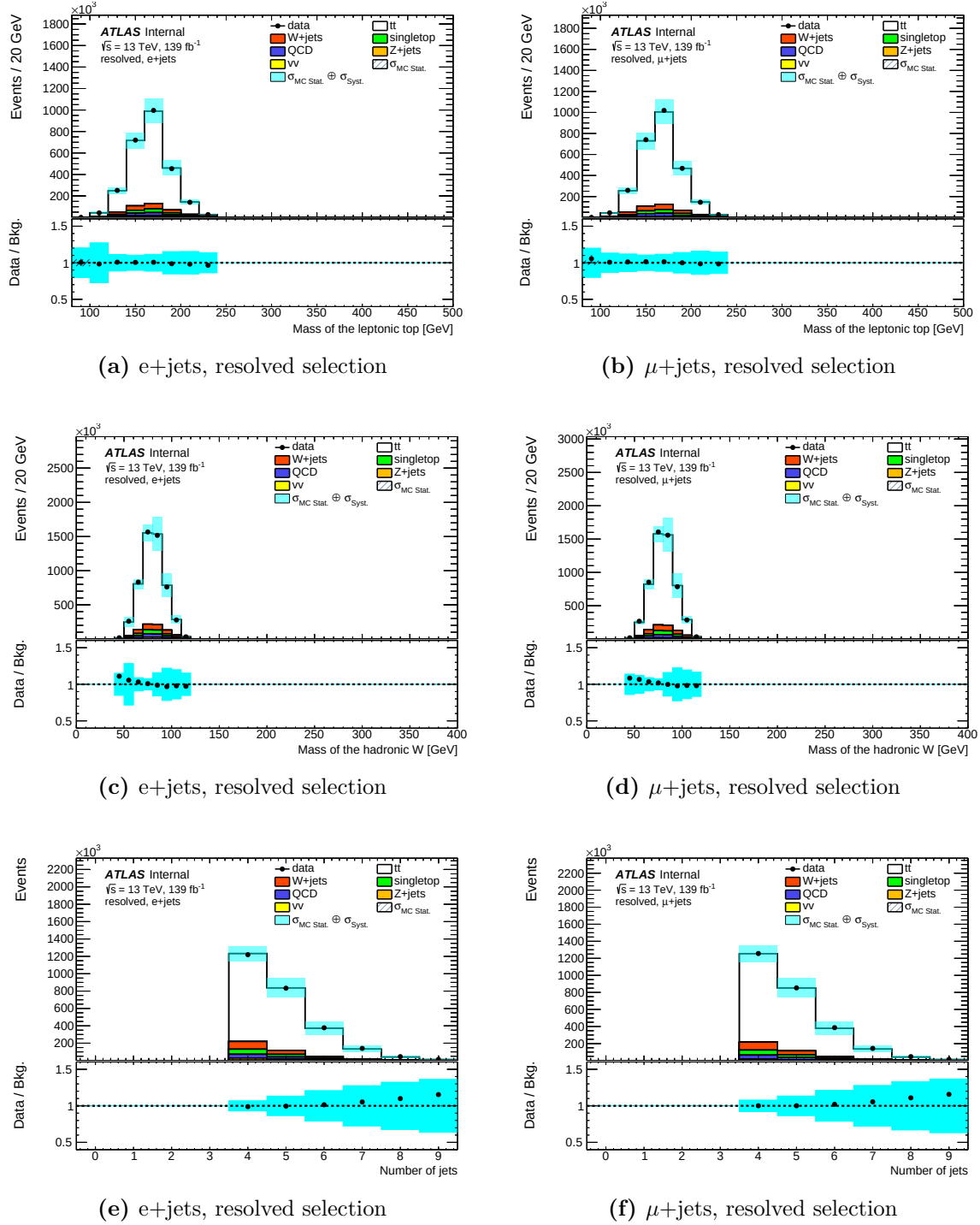


Figure 6.24: The mass of the leptonic top (top), (middle) mass of the hadronic W boson, and (bottom) jets multiplicity distributions in the resolved (a, c, e) electron and (b, d, f) muon channel. The ratio of data and background is mostly within the systematic band.

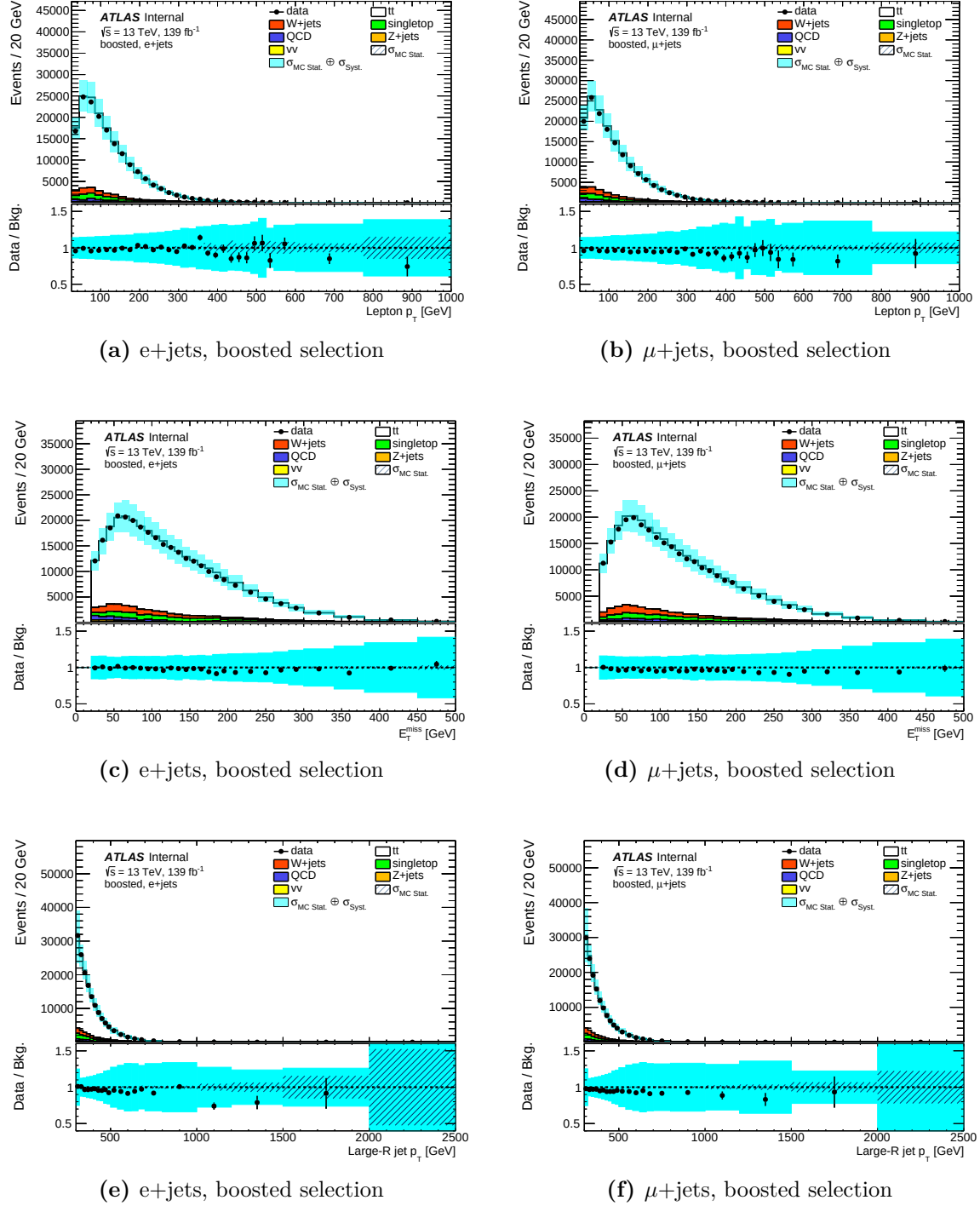


Figure 6.25: The p_T of the lepton (top), (middle) the E_T^{miss} , and (bottom) Large-R jet p_T distributions in the boosted (a, c, e) electron and (b, d, f) boosted muon channel. The ratio of data and background is mostly within the systematic band.

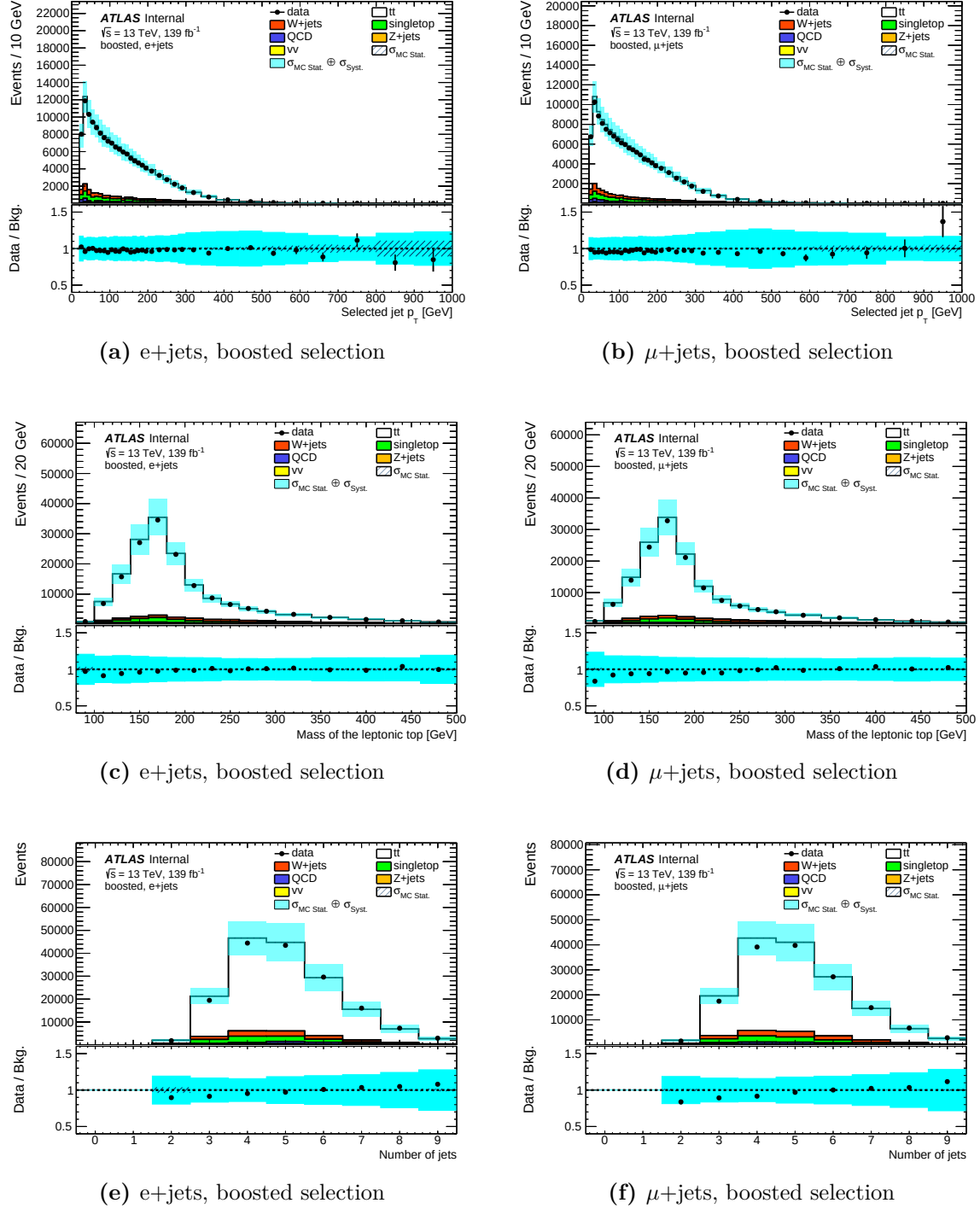


Figure 6.26: The p_T of the selected jet (top), (middle) mass of the leptonic top, and (bottom) jets multiplicity distributions in the boosted (a, c, e) electron and (b, d, f) muon channel. The ratio of data and background is mostly within the systematic band.

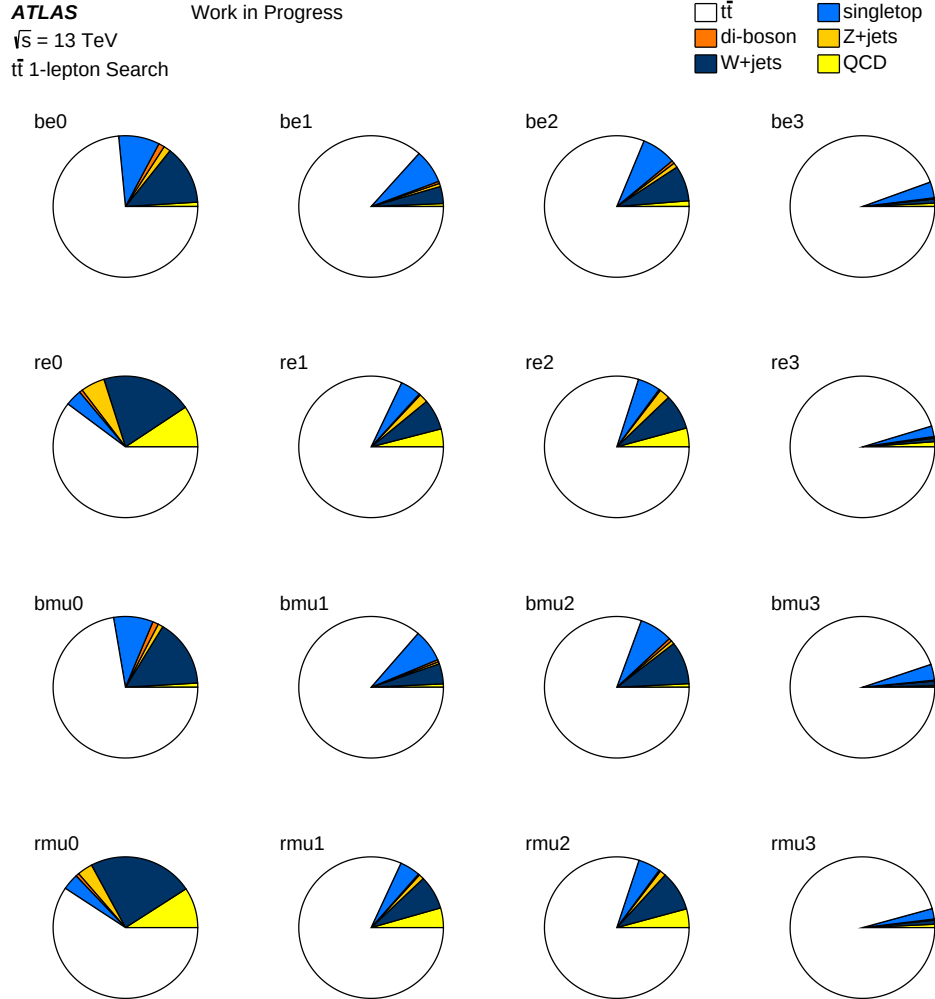


Figure 6.27: Background composition in the 16 SRs. The pie chart shows relative contributions per process in the different regions.

6.9 Compatibility of data with the SM-only hypothesis

A statistical test called the BumpHunter [89] [110] tool calculates the level of agreement between some data and a given hypothesis. These tests evaluate a data set's consistency using a null hypothesis, in this case the background-only hypothesis, designated H_0 . Various alternative signal hypotheses may be checked when a discrepancy with H_0 has been established to see if there is a clear sign of a resonance like signal present in the data.

The whole mass spectrum in resolved (0.3 TeV to 3 TeV) and boosted (0.4 TeV to 6 TeV) topologies is traversed by sliding windows of configurable size, with a minimum width of two bins, to compare the data and SM prediction using the BumpHunter algorithm. For each spectrum, the Poisson probability of the most prominent bump or dip (i.e. the smallest probability) is kept. In each window i , the number of background events is b_i , the data count is d_i , and the Poisson probability $P(d_i, b_i)$ is defined as:

$$P(d_i, b_i) = \begin{cases} \Gamma(d_i, b_i) = \sum_{n=d_i}^{\infty} \frac{b_i^n}{n!} e^{-b_i} & \text{if } d_i \geq b_i \\ 1 - \Gamma(d_i + 1, b_i) & \text{if } d_i < b_i \end{cases} \quad (6.18)$$

The equation above is defined when searching for an excess, with the inequality signs reversed when looking for deficits. Γ is the Gamma function. The smallest $P(d_i, b_i)$ from all the tested bins, $P_i^{\min}$ refers to the most interesting (discrepant with respect to the background expectation) window. To take into account the systematic uncertainties, the definition of $P(d_i, b_i)$ is extended to:

$$P(d_i, b_i) \rightarrow P(d_i, b_i + \lambda_i \theta_i) e^{(-\frac{\lambda_i^2}{2})} \quad (6.19)$$

where λ_i is a real integer in the range from -8 to 8 that optimizes the value of $\frac{(b_i + \lambda_i \theta_i)^{d_i}}{d_i!} e^{-(b_i + \lambda_i \theta_i)} e^{(-\frac{\lambda_i^2}{2})}$ and θ_i is the total cumulative systematic uncertainty for the studied window.

The BumpHunter test statistic t is computed as:

$$t = \begin{cases} 0 & \text{if } d_i \leq b_i \\ -\log P_i^{\min} & \text{otherwise} \end{cases} \quad (6.20)$$

The test statistic from the data is compared to the test statistic seen in at least 10,000 pseudo-data produced by Poisson fluctuations of the anticipated background, in order to

determine the p -value of the most inconsistent bump. The p -value is defined as:

$$p - \text{value} = \frac{\int_{t_{obs}}^{\infty} f(t)}{\int_0^{\infty} f(t)} \quad (6.21)$$

where t_{obs} refers to the test statistic derived from data and $f(t)$ represents the distribution of test statistic values from the pseudo-data. A p -value of 0 would mean that none of the pseudo-experiments produced a deviation from the background expectation that was as large as the deviation observed in the data. For reference, $m_{t\bar{t}}$ spectra after the profiling of the nuisance parameters (Post-Fit distributions) are displayed in Figures 6.29 and 6.28 for 12 distinct channels, divided by lepton flavor (e, μ), topology (resolved, boosted) and the three b-tagging categories (1,2,3). No hints for the existence of a visible signal have been seen in the observed data and hence limits may be set.

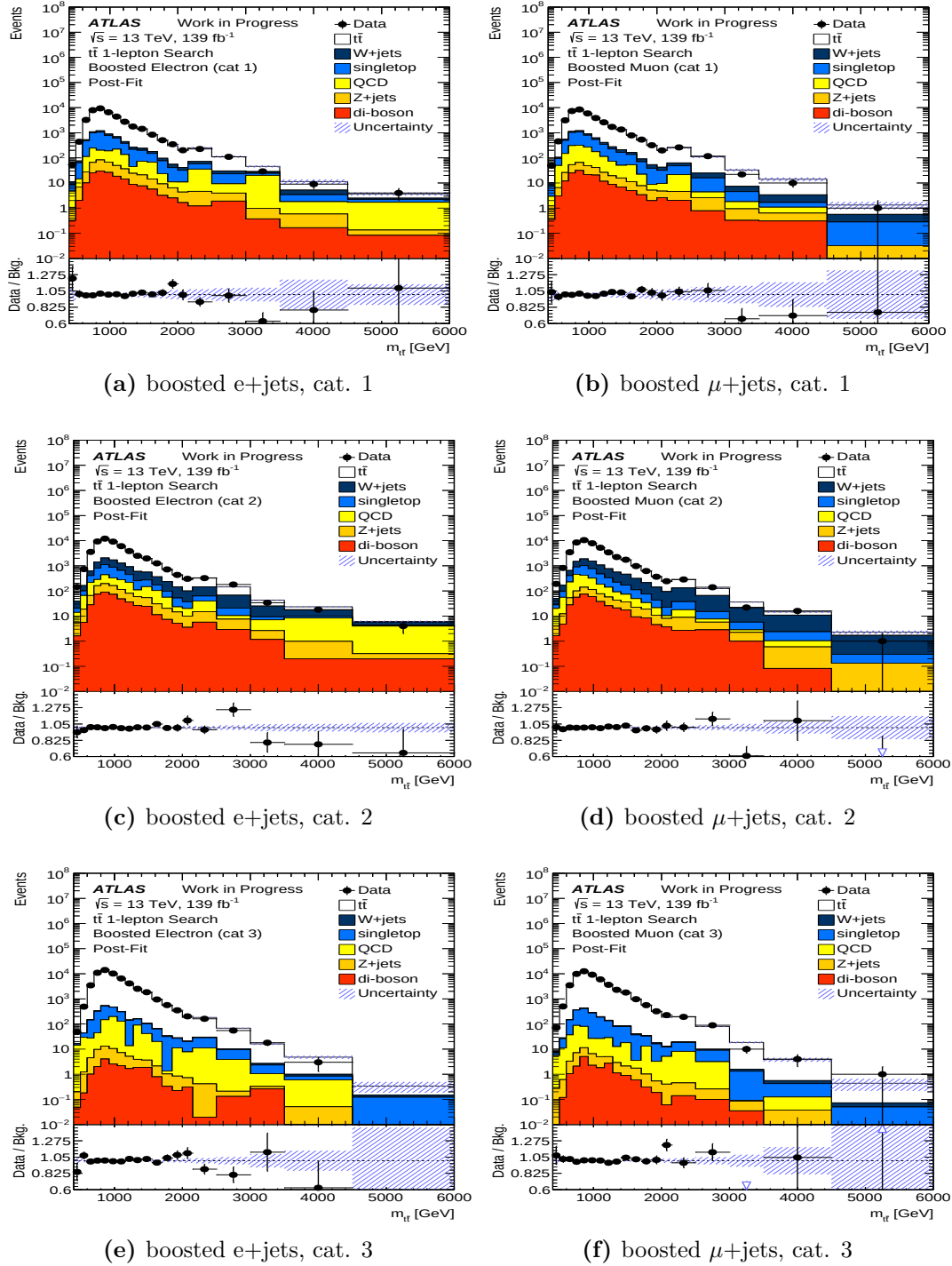


Figure 6.28: The $m_{t\bar{t}}^{reco}$ distributions for the boosted electron (left-hand) and the boosted muon (right-hand) after the profiling of the nuisance parameters.

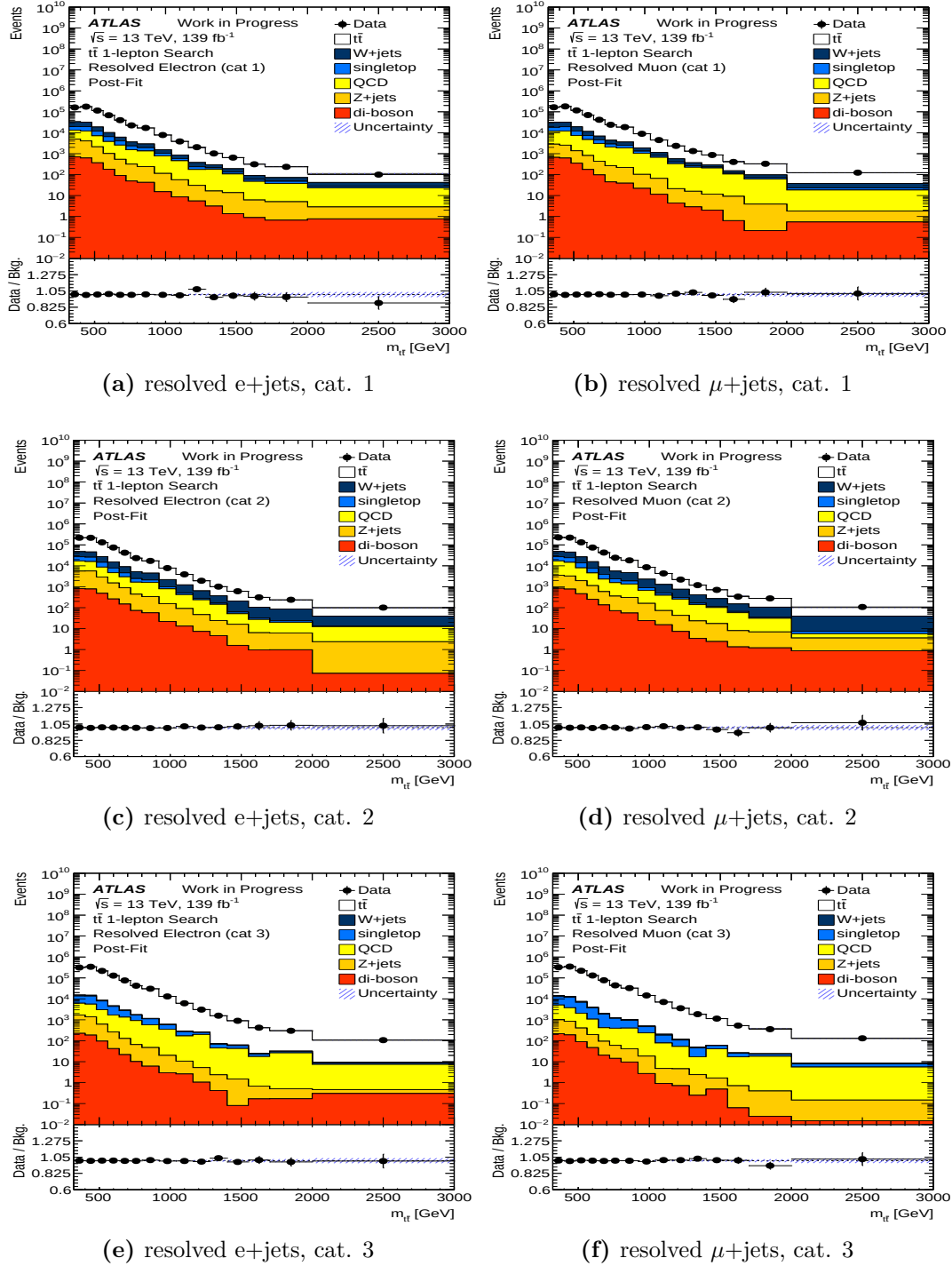


Figure 6.29: The $m_{t\bar{t}}^{reco}$ distributions for the resolved electron (left-hand) and the resolved muon (right-hand) after the profiling of the nuisance parameters.

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

A frequentist procedure using the CL_s approach have been employed to set an upper limits on the cross-section times branching ratio of Z'_{TC2} benchmark scenario at 95% Confidence Level (CL). A tool called TRexFitter [111], together with RooStats and RooFit were used to perform the limit setting process, on which a likelihood for a specific signal model is determined as follows:

$$L(\mu, \Theta) = \prod_{i=0}^{\text{channels, bins}} \frac{e^{-\mu a_{Z',i}(\Theta) \sigma_{Z'} + b_i(\Theta)} (\mu a_{Z',i}(\Theta) \sigma_{Z'} + b_i(\Theta))^{D_i}}{\Gamma(D_i + 1)} C(\Theta) \quad (6.22)$$

where for each bin i , b_i is the number of expected background events, D_i is the number of observed data events, $a_{Z'}$ is the acceptance of the signal, $\sigma_{Z'}$ is the cross section of the Z'_{TC2} signal and μ is the signal strength, which is the parameter of interest. The function $C(\Theta)$ denotes the set of constraints applied on the nuisance parameters Θ , including the systematic uncertainties in the background and signal expectation, as well as the uncertainty associated with luminosity measurement. The profile likelihood ratio test statistic, or Λ is the foundation for the hypothesis testing:

$$\Lambda(\mu) = \frac{L(\mu, \hat{\hat{\Theta}}(\mu))}{L(\hat{\mu}, \hat{\Theta})}, \quad (6.23)$$

where the single circumflex denotes the parameter's maximum likelihood estimate in its unconditional form, and the double circumflex denotes the parameter's maximum likelihood estimate for a particular value of μ . Presuming that the test statistic $-2\ln(\Lambda(\mu))$ is following the χ^2 distribution with one degree of freedom, the variation on the logarithm of the likelihood may be employed to fix an upper limit of the signal production cross-section times branching ratio a 95% CL and to evaluate the impact of the systematic uncertainties in the discovery significance.

The expected limit on the cross-section times branching ratio is calculated for Z'_{TC2} (1.2% width) and interpolated between each mass point from 0.4 to 5 TeV. Preliminary results presented in Figure 6.30 shows that the existence of a Z'_{TC2} boson for masses less than 3.7 TeV is excluded at 95% CL.

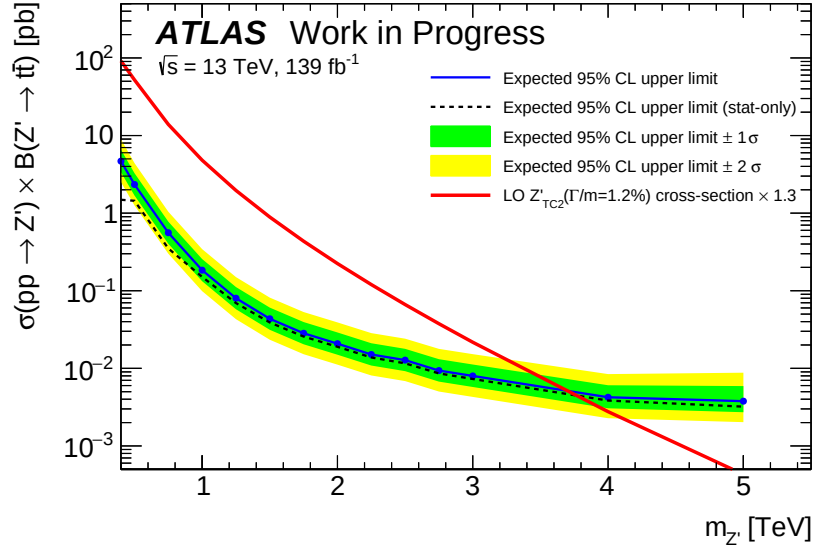


Figure 6.30: 95% CL Exclusion Limits on the Z'_{TC2} signal cross-section times branching ratio, as a function of $m_{Z'}$. All signal regions (electron and muon channels, for both the resolved and boosted selection) are combined. All statistical and systematic uncertainties are included. The theoretical cross-section is overlaid as a red line, and used to extract the mass exclusion limit.

Chapter 7

Summary and outlook

The research work presented in this thesis addresses two main topics of challenges in experimental high energy physics: detector performance studies and analyses of physics data from the experiment. On one hand, the performance and validation of the missing transverse momentum, on the other hand, a search for new massive resonances decaying to a top-antitop ($t\bar{t}$) pair in the $\ell + jets$ channel. The missing transverse momentum (E_T^{miss}) is a variable of great interest, since enormous interesting physical processes produce non-interacting particles that could only be measured using the E_T^{miss} . E_T^{miss} is defined as the negative sum of the transverse momentum of all visible particles. The performance and features of E_T^{miss} are evaluated for selected event samples with and without genuine E_T^{miss} . Within the systematic uncertainty, there is often high agreement between the data from the detector and the matching MC simulations in the covered phase space. Additional performance measures considered in these studies include the optimization of the electron-jet and photon-jet overlap removal the technique and the reduction of tails in the E_T^{miss} distribution. A great improvement of the E_T^{miss} performance is obtained.

A heavy particle would be created at uncertain rates in LHC proton-proton collisions and decay subsequently to a top-antitop quark pair, according to several extensions beyond the SM. These theories of physics include a wide range of topics, from an expanded gauge group to additional dimensions, and they go along divergent routes to more basic models of nature. A search for $t\bar{t}$ resonances in the final states containing at least one electron or muon has been presented in this thesis, exploiting the full Run 2 dataset of proton-proton collisions at a center-of-mass energy of 13 TeV recorded by ATLAS at the LHC, amounting to about 139 fb^{-1} . The reconstruction of the $t\bar{t}$ invariant mass required the development of dedicated identification and reconstruction techniques. The resolved-topology selection and reconstruction is tailored towards events in which the hadronic showers initiated by the

decay products of the hadronically decaying top quark are well separated in the detector and can be reconstructed as three individual jets with small- R parameter using χ^2 method. The boosted-topology selection and reconstruction, in contrast, are optimised for events in which these showers are merged into a single, large- R jet with a characteristic substructure and Deep Neural Network based top-tagger. A special focus in this thesis is the estimation of the QCD multi-jet background, which is a small but relevant background due to its increasing importance with the data-taking conditions at the LHC characterised by a high pile-up activity. Excesses in the reconstructed $M_{t\bar{t}}$ distribution were searched for using the BumpHunter test; no statistically significant deviations were found with respect to the SM expectations. Upper limits on the cross-section times branching ratio were set on the production of a new particle decaying to $t\bar{t}$ in the context of the Topcolor model (Z' leptophobic boson, Z'_{TC2}). Within the measurement no new particles could be discovered but new exclusion limits for the mass of Z'_{TC2} with 1.2% relative width was excluded below 3 TeV using partial Run-2 ATLAS data (36.1 fb^{-1}). The analysis improvements, including the background-rejection, b-tagging and top-tagger algorithms considered to improve the previous limits for the full Run-2 datasets by a factor of 4. The mass limit on the Z' model would improve by almost 1 TeV. But such analysis is still ongoing and under the review. Moreover, for future rounds of this research, the existing production cross section limitations determined for other standard benchmark models will be modified.

The sensitivity of the analysis will be increased by a better detector and reconstruction performance. However, higher pile-up conditions, e.g. in Run 3 or especially at the high-luminosity LHC, can have an opposite effect on the sensitivity. First, more data collected at the LHC will allow for setting stronger limits on various BSM scenarios. In this context, the usage of a more sophisticated techniques for analysis selection, object reconstruction and backgrounds reduction can enhance the sensitivity of the analysis in the future.

Bibliography

- [1] ATLAS collaboration, “Search for new phenomena with top quark pairs in final states with one lepton, jets, and missing transverse momentum in pp collisions at $s = 13$ TeV with the ATLAS detector,” *Journal of High Energy Physics*, 04 2021.
- [2] ATLAS collaboration, “Performance of the missing transverse momentum triggers for the ATLAS detector during Run-2 data taking,” *JHEP*, vol. 2008, p. 080. 53 p, May 2020. [Online]. Available: <https://cds.cern.ch/record/2718582>
- [3] ATLAS collaboration, “Differential cross-section measurements for the electroweak production of dijets in association with a Z boson in proton–proton collisions at ATLAS,” *Eur. Phys. J. C*, vol. 81, p. 163. 42 p, Jun 2020. [Online]. Available: <https://cds.cern.ch/record/2722304>
- [4] ATLAS collaboration, “Configuration and performance of the ATLAS b -jet triggers in Run 2,” *Eur. Phys. J. C*, vol. 81, p. 1087. 61 p, Jun 2021. [Online]. Available: <https://cds.cern.ch/record/2771727>
- [5] ATLAS collaboration, “Measurements of the inclusive and differential production cross sections of a top-quark–antiquark pair in association with a Z boson at $\sqrt{s} = 13$ TeV with the ATLAS detector,” *Eur. Phys. J. C*, vol. 81, no. 8, p. 737, 2021.
- [6] B.-e. Ngair and F. Fassi, “Search of new resonances decaying into top quark pairs in the lepton+jet final state in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector,” 6 2022.
- [7] B.-e. Ngair and F. Fassi, “Search of new resonances decaying into top quark pairs in the lepton+jet final state in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector,” 2022. [Online]. Available: <https://arxiv.org/abs/2206.12978>

- [8] M. Tanabashi *et al.*, “Review of Particle Physics, 2018-2019,” *Phys. Rev. D*, vol. 98, no. 3, p. 030001. 1898 p, 2018. [Online]. Available: <https://cds.cern.ch/record/2636832>
- [9] “Quantum diaries,.” [Online]. Available: <http://www.quantumdiaries.org/2011/11/21/>
- [10] “Top Working Group summary plots,” <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCTopWGSummaryPlots>.
- [11] “Top figures,” <https://www-d0.fnal.gov/Run2Physics/top/>.
- [12] ATLAS Collaboration, “Standard Model Summary Plots March 2021,” CERN, Geneva, Tech. Rep., Mar 2021. [Online]. Available: <https://cds.cern.ch/record/2758261>
- [13] ATLAS Collaboration, “Atlas luminosity public results.” [Online]. Available: <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2>.
- [14] CERN, “Cern accelerator complex.” [Online]. Available: <http://cds.cern.ch/record/1621583/files/CERN\T1\textquoterightsaccelerator-complex2013.jpg?version=1>.
- [15] “Schedules and luminosity forecasts.” <https://lhcb-commissioning.web.cern.ch/lhcb-commissioning/schedule/HL-LHC-plots.htm>.
- [16] ATLAS Collaboration, “Atlas multimedia,.” [Online]. Available: <http://www.atlas.ch/photos/index.html>,
- [17] ATLAS Collaboration, “The atlas experiment at the cern large hadron collider,” *JINST*, vol. 3, p. S08003, 2008.
- [18] J. Pequeno, “Computer generated image of the ATLAS inner detector,” Mar 2008. [Online]. Available: <https://cds.cern.ch/record/1095926>
- [19] ATLAS Collaboration, “Performance of the ATLAS Trigger System in 2015.” 2016. [Online]. Available: <https://cds.cern.ch/record/2235584>
- [20] ATLAS Collaboration, “Trigger operation public results, (2018).” [Online]. Available: twiki.cern.ch/twiki/bin/view/AtlasPublic/TriggerOperationPublicResults
- [21] J. Pequeno, “Event Cross Section in a computer generated image of the ATLAS detector.” Mar 2008. [Online]. Available: <https://cds.cern.ch/record/1096081>

-
- [22] ATLAS Collaboration, “Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at $\sqrt{s} = 13$ TeV,” *Eur. Phys. J. C*, vol. 79, no. 8, p. 639, 2019.
- [23] ATLAS Collaboration, “Electron and photon reconstruction and performance in ATLAS using a dynamical, topological cell clustering-based approach,” Tech. Rep., Dec 2017. [Online]. Available: <https://cds.cern.ch/record/2298955>
- [24] ATLAS Collaboration, “ATLAS data quality operations and performance for 2015–2018 data-taking,” *JINST*, vol. 15, no. 04, p. P04003, 2020.
- [25] ATLAS Collaboration, “Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13$ TeV,” *The European Physical Journal C*, vol. 76, no. 5, May 2016. [Online]. Available: <http://dx.doi.org/10.1140/epjc/s10052-016-4120-y>
- [26] ATLAS Collaboration, “Monte Carlo Calibration and Combination of In-situ Measurements of Jet Energy Scale, Jet Energy Resolution and Jet Mass in ATLAS,” CERN, Geneva, Tech. Rep., Aug 2015. [Online]. Available: <https://cds.cern.ch/record/2044941>
- [27] ATLAS Collaboration, “Tagging and suppression of pileup jets with the ATLAS detector,” CERN, Geneva, Tech. Rep., May 2014. [Online]. Available: <https://cds.cern.ch/record/1700870>
- [28] ATLAS Collaboration, “Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at $\sqrt{s} = 13$ TeV,” *The European Physical Journal C*, vol. 78, Nov 2018. [Online]. Available: <http://dx.doi.org/10.1140/epjc/s10052-018-6288-9>
- [29] T. Khoo, “Missing transverse momentum (talk).” [Online]. Available: https://indico.cern.ch/event/565211/contributions/2293331/attachments/1347838/2033199/SWTut_070616.pdf
- [30] G. A. et al, “Performance of jet substructure techniques for large-R jets in proton-proton collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector,” vol. 2013, no. 9, sep 2013. [Online]. Available: <https://doi.org/10.1007%2Fjhep09%282013%29076>

-
- [31] M. Tanabashi *et al.*, “Review of Particle Physics,” *Phys. Rev. D*, vol. 98, no. 3, p. 030001, 2018.
- [32] ATLAS collaboration, “Search for heavy particles decaying into top-quark pairs using lepton-plus-jets events in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector,” *Eur. Phys. J. C*, 2018.
- [33] R. P. Feynman, “Mathematical formulation of the quantum theory of electromagnetic interaction,” *Phys. Rev.*, vol. 80, pp. 440–457, Nov 1950. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRev.80.440>
- [34] P. W. Higgs, “Spontaneous symmetry breakdown without massless bosons,” *Phys. Rev.*, vol. 145, pp. 1156–1163, May 1966. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRev.145.1156>
- [35] K. A. Meissner and H. Nicolai, “Standard model fermions and infinite-dimensional r symmetries,” *Phys. Rev. Lett.*, vol. 121, p. 091601, Aug 2018. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.121.091601>
- [36] Ahmad and et al, “Direct evidence for neutrino flavor transformation from neutral-current interactions in the sudbury neutrino observatory,” *Phys. Rev. Lett.*, vol. 89, p. 011301, Jun 2002. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.89.011301>
- [37] K. G. Wilson, “Confinement of Quarks,” *Phys. Rev. D*, vol. 10, pp. 2445–2459, 1974.
- [38] A. Grozin., “Quantum chromodynamics,,” *arXiv:1205.1815 [hep-ph]*.
- [39] A. Salam and J. C. Ward, “Weak and electromagnetic interactions,” *Nuovo Cim.*, vol. 11, pp. 568–577, 1959.
- [40] ATLAS Collaboration, “Measurement of the Higgs boson mass from the $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow 4\ell$ channels with the ATLAS detector using 25 fb^{-1} of pp collision data,” *Phys. Rev. D*, vol. 90, no. 5, p. 052004, 2014.
- [41] ATLAS Collaboration, “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC,” *Physics Letters B*, vol. 716, no. 1, Sep 2012. [Online]. Available: <http://dx.doi.org/10.1016/j.physletb.2012.08.020>

-
- [42] F. Abe *et al.*, “Observation of top quark production in $\bar{p}p$ collisions,” *Phys. Rev. Lett.*, vol. 74, pp. 2626–2631, 1995.
- [43] S. Abachi, B. Abbott, M. Abolins, B. S. Acharya, I. Adam, D. L. Adams, M. Adams, S. Ahn, H. Aihara, J. Alitti, and et al., “Observation of the top quark,” *Physical Review Letters*, vol. 74, no. 14, p. 2632–2637, Apr 1995. [Online]. Available: <http://dx.doi.org/10.1103/PhysRevLett.74.2632>
- [44] J. Gao, C. S. Li, and H. X. Zhu, “Top-quark decay at next-to-next-to-leading order in qcd,” *Phys. Rev. Lett.*, vol. 110, p. 042001, Jan 2013. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.110.042001>
- [45] R. M. Harris and S. Jain, “Cross sections for leptophobic topcolor Z' decaying to top–antitop,” *The European Physical Journal C*, vol. 72, no. 7, Jul 2012. [Online]. Available: <http://dx.doi.org/10.1140/epjc/s10052-012-2072-4>
- [46] CMC Collaboration, “Search for $t\bar{t}$ resonances in highly boosted lepton+jets and fully hadronic final states in proton-proton collisions at $\sqrt{s} = 13$ tev,” *Journal of High Energy Physics*, vol. 2017, no. 7, Jul 2017. [Online]. Available: [http://dx.doi.org/10.1007/JHEP07\(2017\)001](http://dx.doi.org/10.1007/JHEP07(2017)001)
- [47] ATLAS Collaboration, “A search for $t\bar{t}$ resonances using lepton-plus-jets events in proton-proton collisions at $\sqrt{s} = 8$ tev with the atlas detector,” *Journal of High Energy Physics*, vol. 2015, no. 8, Aug 2015. [Online]. Available: [http://dx.doi.org/10.1007/JHEP08\(2015\)148](http://dx.doi.org/10.1007/JHEP08(2015)148)
- [48] L. Fitzpatrick, J. Kaplan, L. Randall, and L.-T. Wang, “Searching for the kaluza-klein graviton in bulk rs models,” *Journal of High Energy Physics*, vol. 2007, no. 09, p. 013–013, Sep 2007. [Online]. Available: <http://dx.doi.org/10.1088/1126-6708/2007/09/013>
- [49] B. Lillie, L. Randall, and L.-T. Wang, “The bulk rs kk-gluon at the lhc,” *Journal of High Energy Physics*, vol. 2007, no. 09, p. 074–074, Sep 2007. [Online]. Available: <http://dx.doi.org/10.1088/1126-6708/2007/09/074>
- [50] M. Benedikt, P. Collier, V. Mertens, J. Poole, and K. Schindl, “LHC Design Report,” 2004. [Online]. Available: <https://cds.cern.ch/record/823808>

- [51] O. S. Bruning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, “LHC Design Report,” 2004. [Online]. Available: <https://cds.cern.ch/record/782076>
- [52] W. Herr and B. Muratori, “Concept of luminosity,” 2006. [Online]. Available: <https://cds.cern.ch/record/941318>
- [53] ATLAS collaboration, “The ATLAS Experiment at the CERN Large Hadron Collider,” *JINST*, vol. 3, p. S08003. 437 p, 2008, also published by CERN Geneva in 2010. [Online]. Available: <https://cds.cern.ch/record/1129811>
- [54] ATLAS Collaboration, “ATLAS pixel detector electronics and sensors,” *JINST*, vol. 3, p. P07007, 2008. [Online]. Available: <https://cds.cern.ch/record/1119279>
- [55] H. Pernegger, “The Pixel Detector of the ATLAS experiment for LHC Run-2,” *JINST*, vol. 10, no. 06, p. C06012, 2015.
- [56] MoEDAL Collaboration, “Operation and performance of the ATLAS semiconductor tracker. Operation and performance of the ATLAS semiconductor tracker,” *JINST*, vol. 9, no. CERN-PH-EP-2014-049. CERN-PH-EP-2014-049, p. P08009. 80 p, Apr 2014. [Online]. Available: <https://cds.cern.ch/record/1698966>
- [57] E. Abat *et al.*, “The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: Design and performance,” *JINST*, vol. 3, p. P02013, 2008.
- [58] ATLAS Collaboration, “Readiness of the ATLAS Liquid Argon Calorimeter for LHC Collisions,” *Eur. Phys. J. C*, vol. 70, no. arXiv:0912.2642. CERN-PH-EP-2010-041, May 2010. [Online]. Available: <https://cds.cern.ch/record/1228823>
- [59] ATLAS Collaboration, “Readiness of the ATLAS Tile Calorimeter for LHC collisions,” *Eur. Phys. J. C*, vol. 70, no. arXiv:1007.5423. CERN-PH-EP-2010-024, Jul 2010. [Online]. Available: <https://cds.cern.ch/record/1282535>
- [60] ATLAS Collaboration, *ATLAS liquid-argon calorimeter: Technical Design Report*, ser. Technical design report. ATLAS. Geneva: CERN, 1996. [Online]. Available: <https://cds.cern.ch/record/331061>
- [61] ATLAS Collaboration, “ATLAS tile calorimeter: Technical Design Report,” 1996. [Online]. Available: <https://cds.cern.ch/record/331062>

-
- [62] ATLAS Collaboration, “ATLAS muon spectrometer: Technical Design Report,” 1997. [Online]. Available: <https://cds.cern.ch/record/331068>
- [63] ATLAS Collaboration, “ATLAS level-1 trigger: Technical Design Report,” 1998. [Online]. Available: <https://cds.cern.ch/record/381429>
- [64] ATLAS Collaboration, “ATLAS high-level trigger, data-acquisition and controls: Technical Design Report,” 2003. [Online]. Available: <http://cds.cern.ch/record/616089>
- [65] ATLAS Collaboration, “Calorimeter Clustering Algorithms: Description and Performance,” CERN, Geneva, Tech. Rep., Apr 2008. [Online]. Available: <https://cds.cern.ch/record/1099735>
- [66] ATLAS Collaboration, “Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}=13$ TeV.” *Eur. Phys. J. C*, vol. 76, p. 292. 45 p, Mar 2016. [Online]. Available: <https://cds.cern.ch/record/2139897>
- [67] M. Cacciari, G. P. Salam, and G. Soyez, “The anti-ktjet clustering algorithm,” *Journal of High Energy Physics*, vol. 2008, no. 04, p. 063–063, Apr 2008. [Online]. Available: <http://dx.doi.org/10.1088/1126-6708/2008/04/063>
- [68] ATLAS Collaboration, “Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector,” *Eur. Phys. J. C*, vol. 76, no. 11, p. 581, 2016.
- [69] ATLAS Collaboration, “Impact of Alternative Inputs and Grooming Methods on Large-R Jet Reconstruction in ATLAS,” 12 2017.
- [70] ATLAS Collaboration, “Boosted hadronic top identification at ATLAS for early 13 TeV data,” CERN, Geneva, Tech. Rep., Dec 2015. [Online]. Available: <https://cds.cern.ch/record/2116351>
- [71] J. M. Butterworth, B. E. Cox, and J. R. Forshaw, “WW scattering at the cern lh,” *Phys. Rev. D*, vol. 65, p. 096014, May 2002. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevD.65.096014>
- [72] J. Thaler and K. Van Tilburg, “Maximizing boosted top identification by minimizing n-subjettiness,” *Journal of High Energy Physics*, vol. 2012, no. 2, Feb 2012. [Online]. Available: [http://dx.doi.org/10.1007/JHEP02\(2012\)093](http://dx.doi.org/10.1007/JHEP02(2012)093)

-
- [73] ATLAS Collaboration, “ E_T^{miss} performance in the ATLAS detector using 2015-2016 LHC p-p collisions,” CERN, Geneva, Tech. Rep., Jun 2018. [Online]. Available: <https://cds.cern.ch/record/2625233>
- [74] ATLAS Collaboration, “Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data,” *Eur. Phys. J. C*, vol. C74, no. CERN-PH-EP-2014-153. CERN-PH-EP-2014-153, p. 3071. 48 p, Jul 2014. [Online]. Available: <https://cds.cern.ch/record/1744017>
- [75] ATLAS Collaboration, “Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data,” no. ATLAS-CONF-2016-024, Jun 2016. [Online]. Available: <https://cds.cern.ch/record/2157687>
- [76] ATLAS Collaboration, “Electron and photon energy calibration with the ATLAS detector using data collected in 2015 at $\sqrt{s} = 13$ TeV,” CERN, Geneva, Tech. Rep. ATL-PHYS-PUB-2016-015, Aug 2016. [Online]. Available: <https://cds.cern.ch/record/2203514>
- [77] ATLAS Collaboration, “Measurement of the photon identification efficiencies with the atlas detector using lhc run-1 data,” *The European Physical Journal C*, vol. 76, no. 12, Dec 2016. [Online]. Available: <http://dx.doi.org/10.1140/epjc/s10052-016-4507-9>
- [78] ATLAS Collaboration, “Electron and photon energy calibration with the atlas detector using lhc run 1 data,” *The European Physical Journal C*, vol. 74, no. 10, Oct 2014. [Online]. Available: <http://dx.doi.org/10.1140/epjc/s10052-014-3071-4>
- [79] ATLAS Collaboration, “Reconstruction of hadronic decay products of tau leptons with the atlas experiment,” *The European Physical Journal C*, vol. 76, no. 5, May 2016. [Online]. Available: <http://dx.doi.org/10.1140/epjc/s10052-016-4110-0>
- [80] ATLAS Collaboration, “Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment in pp collisions at $\sqrt{s}=8$ TeV,” *Eur. Phys. J. C*, vol. 75, no. 7, p. 303, 2015.
- [81] ATLAS Collaboration, “Jet reconstruction and performance using particle flow with the atlas detector,” *The European Physical Journal C*, vol. 77, no. 7, Jul 2017. [Online]. Available: <http://dx.doi.org/10.1140/epjc/s10052-017-5031-2>

-
- [82] ATLAS Collaboration, “The atlas simulation infrastructure,” *The European Physical Journal C*, vol. 70, no. 3, p. 823–874, Sep 2010. [Online]. Available: <http://dx.doi.org/10.1140/epjc/s10052-010-1429-9>
- [83] T. Sjöstrand, S. Mrenna, and P. Skands, “A brief introduction to pythia 8.1,” *Computer Physics Communications*, vol. 178, no. 11, p. 852–867, Jun 2008. [Online]. Available: <http://dx.doi.org/10.1016/j.cpc.2008.01.036>
- [84] ATLAS Collaboration, “Summary of ATLAS Pythia 8 tunes,” CERN, Geneva, Tech. Rep. ATL-PHYS-PUB-2012-003, Aug 2012. [Online]. Available: <https://cds.cern.ch/record/1474107>
- [85] D. J. Lange, “The EvtGen particle decay simulation package,” *Nucl. Instrum. Meth. A*, vol. 462, pp. 152–155, 2001.
- [86] T. Gleisberg, S. Höche, F. Krauss, M. Schönherr, S. Schumann, F. Siegert, and J. Winter, “Event generation with sherpa 1.1,” *Journal of High Energy Physics*, vol. 2009, no. 02, p. 007–007, Feb 2009. [Online]. Available: <http://dx.doi.org/10.1088/1126-6708/2009/02/007>
- [87] S. Catani, L. Cieri, G. Ferrera, D. de Florian, and M. Grazzini, “Vector boson production at hadron colliders: A fully exclusive qcd calculation at next-to-next-to-leading order,” *Physical Review Letters*, vol. 103, no. 8, Aug 2009. [Online]. Available: <http://dx.doi.org/10.1103/PhysRevLett.103.082001>
- [88] S. Schumann and F. Krauss, “A parton shower algorithm based on catani-seymour dipole factorisation,” *Journal of High Energy Physics*, vol. 2008, no. 03, p. 038–038, Mar 2008. [Online]. Available: <http://dx.doi.org/10.1088/1126-6708/2008/03/038>
- [89] G. Choudalakis, “On hypothesis testing, trials factor, hypertexts and the bumhunter,” 2011.
- [90] S. S. Wilks, “The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses,” *Annals Math. Statist.*, vol. 9, no. 1, pp. 60–62, 1938.
- [91] A. L. Read, “Presentation of search results: The CL(s) technique,” *J. Phys. G*, vol. 28, pp. 2693–2704, 2002.

-
- [92] ATLAS Collaboration, “Prospects for top anti-top resonance searches using early ATLAS data.” CERN, Geneva, Tech. Rep., Jul 2010. [Online]. Available: <https://cds.cern.ch/record/1278454>
- [93] ATLAS Collaboration, “A search for $t\bar{t}$ resonances in lepton+jets events with highly boosted top quarks collected in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector.” *JHEP*, vol. 09, p. 041. 45 p, Jul 2012. [Online]. Available: <https://cds.cern.ch/record/1461225>
- [94] ATLAS Collaboration, “A search for $t\bar{t}$ resonances in the lepton plus jets final state with ATLAS using 14 fb $^{-1}$ of pp collisions at $\sqrt{s} = 8$ TeV,” CERN, Geneva, Tech. Rep., May 2013. [Online]. Available: <https://cds.cern.ch/record/1547568>
- [95] <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/EGammaIdentificationRun2>.
- [96] ATLAS Collaboration, “Muon reconstruction performance in early $\sqrt{s} = 13$ TeV data,” CERN, Geneva, Tech. Rep., Aug 2015. [Online]. Available: <https://cds.cern.ch/record/2047831>
- [97] <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/RecommendedIsolationWPs>.
- [98] ATLAS Collaboration, “Performance of Top Quark and W Boson Tagging in Run 2 with ATLAS,” CERN, Geneva, Tech. Rep., Aug 2017. [Online]. Available: <https://cds.cern.ch/record/2281054>
- [99] ATLAS Collaboration, “Optimisation and performance studies of the ATLAS b -tagging algorithms for the 2017-18 LHC run,” CERN, Geneva, Tech. Rep., Jul 2017. [Online]. Available: <https://cds.cern.ch/record/2273281>
- [100] C. Oleari, “The powheg box,” *Nuclear Physics B - Proceedings Supplements*, vol. 205-206, p. 36–41, Aug 2010. [Online]. Available: <http://dx.doi.org/10.1016/j.nuclphysbps.2010.08.016>
- [101] R. D. Ball, V. Bertone, S. Carrazza, C. S. Deans, L. Del Debbio, S. Forte, A. Guffanti, N. P. Hartland, J. I. Latorre, and et al., “Parton distributions for the LHC run II,” *Journal of High Energy Physics*, vol. 2015, no. 4, Apr 2015. [Online]. Available: [http://dx.doi.org/10.1007/JHEP04\(2015\)040](http://dx.doi.org/10.1007/JHEP04(2015)040)

-
- [102] R. D. Ball, V. Bertone, S. Carrazza, C. S. Deans, L. Del Debbio, S. Forte, A. Guffanti, N. P. Hartland, J. I. Latorre, J. Rojo, and M. Ubiali, “Parton distributions with LHC data,” *Nuclear Physics B*, vol. 867, no. 2, p. 244–289, Feb 2013. [Online]. Available: <http://dx.doi.org/10.1016/j.nuclphysb.2012.10.003>
- [103] M. Beneke, P. Falgari, S. Klein, and C. Schwinn, “Hadronic top-quark pair production with NNLL threshold resummation,” *Nuclear Physics B*, vol. 855, no. 3, pp. 695–741, feb 2012. [Online]. Available: <https://doi.org/10.1016%2Fj.nuclphysb.2011.10.021>
- [104] TTbarNNLOReweighter, <https://gitlab.cern.ch/pinamont/TTbarNNLOReweighter>.
- [105] E. Bothmann, G. Singh Chahal, S. Höche, J. Krause, F. Krauss, S. Kuttimalai, S. Liebschner, D. Napoletano, M. Schönherr, H. Schulz, and et al., “Event generation with sherpa 2.2,” *SciPost Physics*, vol. 7, no. 3, Sep 2019. [Online]. Available: <http://dx.doi.org/10.21468/SciPostPhys.7.3.034>
- [106] C.-H. Kom and W. J. Stirling, “Charge asymmetry in $W + \text{jets}$ production at the LHC,” *Eur. Phys. J. C*, 2010.
- [107] ATLAS Collaboration, “Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC,” CERN, Geneva, Tech. Rep., Jun 2019. [Online]. Available: <https://cds.cern.ch/record/2677054>
- [108] W. Buttinger, “Using Event Weights to account for differences in Instantaneous Luminosity and Trigger Prescale in Monte Carlo and Data,” CERN, Geneva, Tech. Rep., May 2015. [Online]. Available: <https://cds.cern.ch/record/2014726>
- [109] M. Botje, J. Butterworth, A. Cooper-Sarkar, A. De Roeck, J. Feltesse, S. Forte, A. Glazov, J. Huston, R. McNulty, T. Sjostrand, and R. Thorne, “The pdf4lhc working group interim recommendations.”
- [110] <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/BumpHunter>.
- [111] <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TtHFFitter>.

Résumé

La pièce maîtresse de cette thèse consiste en la recherche de nouvelles résonances se désintégrant en paire de quarks $t\bar{t}$ dans le canal semi-leptonique, ainsi de mettre des limites supérieures sur la section efficace de production de particules lourdes prédites par plusieurs modèles de références. La masse invariante du système $t\bar{t}$ est l'observable principal de cette recherche, qui est utilisée pour tester la compatibilité des données réelles par rapport aux prédictions du MS. Pour la région à grand mass, les quarks top sont produits avec une grande impulsion et leurs produits de désintégrations sont collimatés dans un seul jet possédant un grand rayon, ce qui nécessite l'utilisation des techniques de sous-structure de jet pour identifier les paires $t\bar{t}$ (boosté), tandis que pour de basses valeurs de mass, les événements sélectionnés sont dominés par des quarks top où les produits de désintégration sont bien séparés (résolu). Des techniques améliorées sont spécifiquement adaptées à l'estimation du bruit de fond multi-jets, aussi la sélection et la reconstruction des événements sont développées et utilisées. Des résultats préliminaires sur les limites supérieures du rapport d'embranchement fois la section efficace de production sont établis à un niveau de confiance de 95% pour le boson Z' .

Mots- clés : *Quark Top, MET, QCD-multi-jet, semi-leptonic, Modèle Standard*

Abstract

The main part of this dissertation is the search for a heavy resonance decaying into a $t\bar{t}$ pair in the lepton plus jets final state, also establishes upper bounds on the production cross section of new heavy particles in a number of benchmark scenarios. The invariant mass of the $t\bar{t}$ system is the principal observable on this search, which is introduced to evaluate the degree of compatibility of the data with respect to the SM-only hypothesis. For the high mass region, the chosen events are dominated by top-quarks with collimated decay products, then, jet substructure techniques should be used to pick out the $t\bar{t}$ pairs (boosted scenario), while for low mass, the chosen events are dominated by top-quarks with well separated decay products (resolved scenario). Improved techniques specifically tailored the QCD multi-jet background estimation also events selection and reconstruction are developed and used. No evidence for resonant is found and, as a result, constraints are set on theoretical models. Preliminary results on the Upper limits on the production cross-section times branching ratio are established at a 95% credibility level for Z' boson.

Key Words : *Top Quark, QCD Multi-jet, BSM, Missing Et, Z' boson, Standard Model.*