ROYAUME DU MAROC
Université Mohammed V - RABAT - جامعة محمد الخامس - الرباط -
Faculté des sciences
كلية العلوم
CENTRED’ETUDESDOCTORALES-SCIENCESETTECHNOLOGIES

N°d’ordre 3736

THESE.

Envuedel’obtentiondu: ***DOCTORAT***

StructuredeRecherche: Équipe de Sciences de la Matière et du Rayonnements
Discipline: Physique
Spécialité: Physique des Hautes Énergies et Analyse des Données

Présentée et soutenue le 24/12/2022 par:

Souad BATLAMOUS

Investigations for new physics using the ATLAS detector: Search of new high-mass resonances in top-antitop and the reinterpretation of a search for dark matter production in association with a Higgs boson

JURY

Driss BENCHEKROUN	PES, Université Hassan II, Faculté des Sciences Casablanca	Président
Aziz BENSALAH	PES, Université Mohamed V, Faculté des Sciences Rabat	Rapporteur/Examinateur
MOHAMED GOUIGHRI	PH, Université Ibn Tofail, Faculté des Sciences Kénitra	Rapporteur/Examinateur
Adil B E L H A J	PH, Université Mohamed V, Faculté des Sciences Rabat	Rapporteur/Examinateur
Yahya TAYALATI	PES, Université Mohamed V, Faculté des Sciences Rabat	Examinateur
RACHID MAZINI	PES,Taiwan Academia Sinica	Co-Directeur de thèse
Rajaâ CHERKAOUI ELMOURSLI	PES, Université Mohamed V, Faculté des Sciences Rabat	Directrice de thèse

AnnéeUniversitaire:2022/2023

Faculté des Sciences, avenue Ibn Battouta, BP.1014 RP,Rabat–Maroc
00212(05)37 771876 00212(05)37774261;http://www.fsr.um5.ac.ma

Dedicace

To my parents who were the first scientists in my life
and to my brothers and sisters.

Remerciements

Although many consider the Ph.D. as a solitary pursuit, that could not be further from reality in my experience. I have gone through a graduate program at an institution I didn't even dream of attending while growing up ; this document and this experience have only been possible because of the support, encouragement, and love of many people throughout my Ph.D, for which I am extremely grateful. I hope to one day thank everyone for their role in helping me reach this point in my career, but I would like to start here by recognizing a few people.

I owe the greatest thanks to my supervisor Mrs. **Rajaa CHERKAOUI EL MOURSLI** Profesor (PES) at university Mohamed V, It is no exaggeration to say that the completion of my D.Phil. would not have been possible without her, particularly the vital encouragement in the first stages of my studies. I valued her keen instinct for realizing worthwhile projects to pursue and the extremely high standards that she expects; the experience of working with her has endowed me with skills that will be useful throughout my life. Realy Without her succinct directions and insightful comments on all aspects of this dissertation, this research wouldn't have been possible.

When I started my Ph.D. in experimental particle physics, it was entirely a new area for me. I am grateful to Mr. **Rachid MAZINI** for helping me find my footing. You have been my primary resource in getting my scientific questions answered, and you have helped me navigate my first ATLAS projects and authorship qualification task. You have been an outstanding supervisor. Your guidance and feedback have been invaluable throughout my graduate studies. Your constant availability, even on short notice, and your flexibility regarding research projects have provided me with the ideal environment in which to carry out my research.

My gratitude goes also to all the jury members, Mr. **Driss BENCHEKROUN** Profesor (PES) at university Hassan II, president of the jury members for accepting to attend the defense of my thesis.
Thanks also go to Mr. **Aziz BENSALAH** Profesor (PES) at university Mohamed V, for going through my thesis and for accepting to be in my examining committee.

I would like to express my sincere gratitude to Mr. **Mohamed GOUIGHRI** Profesor (PES) at university Ibn Tofail, for accepting to review my thesis and providing feedback to improve it.

I am grateful to Mr. **Adil BELHAJ** Profesor (PES) at university Mohamed V, for the evaluation of this thesis.

I gratefully acknowledge Mr. **Yahya TAYALATI** Profesor (PES) at university Mohamed V, for his endless support, dedication to the work, and for always being there when I needed help and for the evaluation of this thesis and their useful comments.

Also, Yahya Tayalati's financial support from the Swedish Research Council for allowing me to travel to CERN has allowed me to visit and participate in various events in CERN. I also acknowledge the financial support by Association Avenir pour la Sciences et le Savoir (AASS) of Oujda, thanks for that. Also many grants from IFIC to attend schools and conferences.

I also wish to thank the members of the ESMaR team for the friendly work environment and support of all my work. Many thanks to the ATLAS Collaboration who allowed me to be part of surely one of the greatest scientific experiments of our times and gave me access to the research facilities. I am also thankful to my colleagues in the group Missing energy transverse for sharing the excitement of learning and developing all those new ideas about MET. I would also like to thank the $t\bar{t}$ resonance group for their immense support and for answering my many questions along the way. Also I would like to thanks my colleages from the High-mass resonances in Photon + MET.

I am especially grateful to my family and friends for their support throughout my PhD studies. I thank you all.

Abstract

The standard model of particle physics (SM) describes all the fundamental particles and their interactions. It is a very successful theory; however, many experimental observations are either not consistent or not explained by the SM. So, it is clear that there has to be a model beyond the SM. Several of such extensions predict new heavy particles that can interact with SM particles. A variety of theories beyond the standard model (BSM) introduce a particular function for the top quark, leading to resonances that are preferentially decaying to $t\bar{t}$ pairs. The invariant mass ($m_{t\bar{t}}$) of the reconstructed $t\bar{t}$ pair is used to test the compatibility of the data with the standard model prediction. The selected events are dominated by top quarks with collimated decay products for the high $m_{t\bar{t}}$ region, and jet substructure techniques have to be used to select $t\bar{t}$ pairs. This is known as boosted approach, while the selected events are dominated by top quarks with well-separated decay products for low $m_{t\bar{t}}$ region, and this is referred to as a resolved approach. In our analyzed mass spectra, no significant excess of events above the expected Standard Model predictions was observed. Hence, upper limits on production cross-section times branching ratio and lower limits on the mass of heavy Z' particles are placed at a 95%. Mass of Z' is expected to be excluded up to 3.7 TeV.

This work also presents the reinterpretation of a research of dark matter production in association with heavy Higgs bosons decaying into a photon and a dark photon, with masses ranging from 400 GeV to 3 TeV, produced either in gluon-gluon Fusion (ggF) or in Vector-Boson-Fusion (VBF) modes. The exclusion limits at 95% CL on the $\sigma \times BR(H \rightarrow \gamma\gamma_d)$ for the combined ggF+VBF productions are found to be in the range from [1.44,24.2]fb and [1.68,28.1]fb for observed and expected respectively.

The technical part of this thesis is dedicated to the performance of algorithms that identify the missing energy transverse soft hadronic activity $E_T^{miss,softterm}$. An estimation of the systematic uncertainties for various aspects of the ATLAS Inner Detector on $E_T^{miss,softterm}$ are also discussed for the $Z \rightarrow \mu\mu$ channel in the 2015 Data at $\sqrt{13}$ TeV.

CERN, LHC, ATLAS, Missing energy transverse, Higgs boson, $Z' \rightarrow t\bar{t}$, $H \rightarrow \gamma\gamma_d$, Resonance

---

Résumé

Le Modèle Standard de la physique des particules (MS) décrit toutes les particules fondamentales et leurs interactions. Cependant, de nombreuses observations expérimentales ne sont pas cohérentes ou ne sont pas expliquées par le SM. Donc, il est clair qu'il doit y avoir un modèle au-delà du SM. Plusieurs extensions du MS prédisent de nouvelles particules lourdes qui peuvent interagir avec les particules MS. Une variété de théories au-delà du modèle standard (BSM) introduisent une fonction particulière pour le quark top, conduisant à des résonances qui se désintègrent préférentiellement en paires $t\bar{t}$. Cette analyse utilise les données des collisions proton-proton recueillies par l'expériences ATLAS auprès du grand collisionneur de hadrons (LHC). Aucune preuve de résonance n'a été trouvée, des limites supérieures sur le produit de la section efficace et le rapport d'enbranchement on the mass of heavy Z' sont extraites à 95% CL. La masse de Z' a été exclue dont des masses jusqu'à 3.7 TeV.

Ce travail présente également une recherche sur la reinterpretation de la matière noire en association avec un boson de Higgs lourds qui se désintégrent en un photon et un photon sombre avec des masses entre 400 GeV et 3 TeV, produites soit en mode fusion gluon-gluon (ggF) ou en mode fusion vecteur-boson (VBF). Les limites d'exclusion à 95 % CL sur le $\sigma \times BR(H \to \gamma\gamma_d)$ pour les productions combinées ggF+VBF se situent entre [1,44, 24,2]fb et [1,68 ,28.1]fb pour observé et attendu respectivement.

La partie technique de cette thèse est consacrée à la réalisation d'un algorithme qui calcule le terme doux d'énergie transverse manquante $E_T^{miss,softterm}$. Des études sur l'estimation des incertitudes systématiques pour différents éléments du détecteur interne ATLAS pour le terme doux de l'énergie transverse manquante sont également discutées pour le canal $Z \to \mu\mu$ en utilisant l'ensemble de données 2015 avec une énergie de faisceau de 13 TeV.

CERN, LHC, ATLAS, Missing energy transverse, Boson de Higgs, $Z' \to t\bar{t}$, $H \to \gamma\gamma_d$, Resonance

---

Contents

List of Tables

List of Figures

Introduction

All matter around us is found to be made from a few basic building blocks of matter called fundamental particles, governed by four fundamental forces. Our best understanding of how these particles and three of the forces are related to each other is encapsulated in the standard model of particle physics.
The Standard Model (SM) of particle physics, is the theory that as of today best describes the experimental results on subatomic particles. It has been enormously successful in predicting a wide range of phenomena. The last missing piece, the Higgs boson, was discovered in 2012 by the ATLAS and CMS experiments at the LHC, represented a great milestone in the history of particle physics. This very important discovery allowed to confirm the mechanism through which all the fundamental particles acquire mass. This successful achievement implies that we are now capable to produce all standard model particles in a controllable environment, allowing us to study those particles in more detail and ultimately discover a deeper nature of our universe that is still unknown to us.

In spite of the nice theoretical features and successful experimental verification of the SM for many years, we have many reasons to think that the SM is not the final theory of Nature. First, there are some phenomenological issues, many coming from cosmology. These consist of the model's its inability to incorporate gravity in its description of the fundamental forces, predict non-zero neutrino masses and account for the existence of Dark Matter in the Universe and the matter-antimatter asymmetry. Numerous extensions known as beyond the Standard Model physics are proposed and explored by Large Hadron Collider experiments where bounds are set upon them. Chapter 1 summarizes some of the basic features of the Standard Model describing our current understanding of particles and their interactions with an overview of its shortcomings. The same chapter also provides a brief overview of theories beyond the Standard Model which introduce the heavy resonant states predominantly decaying to a $t\bar{t}$ pair and dark photon particles in the mono-photon signatures.

It is clear that a robust reconstruction and precise measurement of the missing energy transverse E_T^{miss} is a crucial aspect in the search for the top quark and on the monophoton signature at colliders. An important part of this thesis is devoted to the estimation of the systematic uncertainties for various aspects of the ATLAS Inner Detector on the track based soft term E_T^{miss} (TST). it gives an overview of the sources of systematic uncertainties that will be considered and the way they are propagated to the final result. Such the effects of the variation include track reconstruction efficiency. The evaluation of the fake rate as well as the impact parameter resolution is also shown for completeness, along with systematic uncertainties in the alignment weak mode. In chapter 4 then goes on to discuss the different systematic uncertainties considered in this search and how they propagate to the final result for different object on the ATLAS internal detector.

Variety of final states are predicted in the various BSM scenarios. The top quark is the heaviest known elementary particle and the only quark whose mass is of the order of the electroweak scale. As a consequence of its unique properties, several theoretical extensions of the SM predict the existence of new massive particles, preferably coupled to the top quark and decayed to the top quark ($t\bar{t}$ pair. Important examples of such BSM scenarios are models that predict new types of fundamental interactions described by new gauge symmetries and BSM theories based on the introduction of additional space-time dimensions. This thesis describes the results of searches for resonant $t\bar{t}$ production based on data collected by the ATLAS experiment at the LHC in proton-proton (pp) collisions at center-of-mass energies $\sqrt{s}$ = 13 TeV. The search are performed using the l(=e,μ) + jets channel, in which one of the two W bosons in the final state decays to leptons, while the other decays hadronically. The analyses make use of dedicated techniques to improve the identification of top quark decays produced with very large Lorentz boosts, i.e. with a transverse momentum largely exceeding the top quark mass. Then The topology of the top quark decay is naturally dependent on the transverse momentum, or boost, of the parent particle. Top quarks produced at rest will have well-separated or easily "resolved" decay products. As the level of boost increases, these decay products become increasingly more collimated until they occupy the same region of the detector. Chapter 5 details these techniques for different reconstruction strategies are used for top quarks with small and large momentum, provides also an overview of the simulated signal models and the object definitions and event selections respectivlly. Furthermore some control plots are presented to show the data-background comparison.

On other one interesting scenario on the BSM is to predict a new 'dark force' mediated by a light neutral particle called a dark photon, which opens a window to a complex Dark Sector. Through kinetic mixing a dark photon produced from the decay of a Higgs boson can decay back to SM particles with a sizeable lifetime, giving rise to striking signatures at hadron colliders. This thesis presents a search for a scalar Higgs boson H produced via in both gluon-fusion and vector-boson-fusion channels and decaying to an undetected particle and a photon γ. Chapter 6 describes the search for dark photons in the decays of Higgs bosons produced in association with photons in proton-proton collisions with the ATLAS detector in the RECAST (Request Efficiency Computation for Alternative Signal Theories) framework and its immediate use to test hypotheses of alternative signals. An overview of the reinterpretation at the LHC by RECAST is important to verify the preservation of the analysis. Experimental research for BSM physics could be sensitive to a wide variety of models. However, in most cases only a few models are considered when interpreting the results of experimental research. The number of models that can be evaluated for interpretation only increases over time, and research represents a significant investment of time and resources for experimental physicists. RECAST will preserve the analysis code and the operating environment so that it can be reinterpreted with new physical models.

This thesis presents a physics analysis performed with the ATLAS detector at the Large Hadron Collider with data from the Run 2 (2015-2018) period of proton-proton collision at a center-of-mass energy of 13 TeV, corresponding to an integrated luminosity used for physics analyses of 139 fb^{-1} . In addition, the ATLAS upgrade for the Run 3 (2022-2026) period where the software is migrated to a multi-threading framework is extensively covered along with the new read-out upgrade for the High-Luminosity LHC that will be tested during Run 3. Chapter 2 is devoted to the description of the experimental facilities of the LHC and the ATLAS detector by giving an overview of the operation and details of the sub-detectors, as well as the ATLAS trigger and data acquisition system. Chapter 3 provides a description of the physical object reconstruction from the information recorded for each event in the different ATLAS sub-detectors. A detailed description concerns the hadronic jets, the reconstruction of electrons, photons, muons and tau-leptons, respectively and finally on the missing transverse momentum.

Chapter 1

Theoretical framework

1.1 Introduction

The Standard Model (SM) of particle physics is the theoretical and experimental effort that so far describes the subatomic world, what everything around us is made of, and under what mechanisms all these components interact together. The Chapter generally presents a basic introduction to the building blocks of the SM in section 1.2. and Dark photons and resonant monophoton signatures in Higgs boson decays at the LHC 1.3, and the outstanding questions and limitations of the SM in Section 1.4. Finally, the most important top quark evidences, candidates and detection strategies are presented in Section 1.5.

1.2 Particle content of the Standard Model

The Standard Model of Particle Physics (SM) [1, 2] is a gauge field theory giving a description of all the known subnuclear particles, antiparticles, and three of the four fundamental forces (Electromagnetic, Strong and Weak) that govern their interactions. Only the gravitational force is not included in the theory and its implementation is one of the major goals in the field of particle physics. Many experimental discoveries have been guided by the SM predictions, including the discovery of the top quark [3, 4] and up to the latest one, the observation of the Higgs boson at the LHC in July 2012 [5, 6].
The SM is a renormalisable quantum field theory based on a non-abelian broken symmetry that is characterized by the following product of gauge groups.

$$SU(3)_c \times SU(2)_l \times U(1)_y \tag{1.1}$$

Where SU(3)$_c$ represents the symmetry of the theory describing the strong nuclear force, quantum chromodynamics (QCD). The product group SU(2)$_l$ corresponds to the weak interaction, and $U(1)_Y$ is the electromagnetic interaction described by Quantum Electro-Dynamics (QED). The weak and the electromagnetic interactions are unified in the Electroweak Theory (EW), proposed by Glashow, Weinberg and Salam [1, 2], via the gauge group $SU(2)_l \times U(1)_Y$. According to the SM, all of the

fundamental interactions derive from one general principle, the requirement of local invariance (the invariance of the Lagrangian under a local gauge transformation of a given symmetry group).
In the SM, particles are described as excitations of quantum fields. In the following paragraphs, we introduce the quantum field theory description of fermions and bosons.
The fermion field Particles of spin one-half that obeying to the statistics of Fermi-Dirac. also referred as spinor fields, which are classified in quarks and leptons and subdivided in three of families or generations. Generation of quarks and leptons of different generations differ in their masses and with the same quantum numbers. The lightest and most stable particles make up the first generation and form all stable matter in the universe, whereas the heavier and unstable particles belong to the second and third generations.
Each generation of leptons consists of a charged particle (electron, muon and tau) accompanied by a neutral and massless neutrino. Charged leptons are subject to the weak and electromagnetic interactions while the neutrinos only interact weakly. In the case of quarks, each generation contains a positively charged quark (+2/3) and a negatively charged one (-1/3). Moreover, as in the case of gluons, all quarks carry color charge (red, green, blue). Quarks can interact via strong, weak and electromagnetic interactions. Table 1.2 provides the classification, masses and electric charge of the three families of the SM fermions. Additionally, for each fermion exists an anti-particle, thus doubling the number of fermions. The anti-particles are characterized by having the same masses but opposite quantum numbers such as the charge(s).

Fermion Type	**Field**	T_3	**Y**	**Q**	**Colour**
Lepton	$L_L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}$	$\begin{pmatrix} +1/2 \\ -1/2 \end{pmatrix}$	-1/2	$\begin{pmatrix} 0 \\ -1 \end{pmatrix}$	singlet
	e_R	0	-1	-1	singlet
Quark	$Q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix}$	$\begin{pmatrix} +1/2 \\ -1/2 \end{pmatrix}$	+1/6	$\begin{pmatrix} +2/3 \\ -1/3 \end{pmatrix}$	Triplet
	u_R	0	+2/3	-1/3	Triplet
	d_R	0	+2/3	-1/3	Triplet

Figure 1.1: Summary of fermion field properties for the first generation of fermions. Fermions in the second and third generation have identical properties.

The Boson field particle obeying Bose-Einstein statistics. It responsible for three of the four interactions in nature, have integral spin values and mediate the forces via interactions between each other and with the matter particles. The force carrier the strong that acts on the elementary particle of hadronic matter, i.e. quarks and gluons. The weak interaction is mediated via $W^{\pm}$ (charged weak interactions) and Z (neutral weak interactions) bosons. The photon (γ) mediates the interactions between electrically charged particles (QED interactions). Finally, eight gluons (g) mediate the strong interaction between colored particles (quarks and gluons). Table 1.2 summarizes the classification of bosons in the SM.

Boson	Interaction	Mass	Electric charge	Colour
γ	Electromagnetic	$< 1 \times 10^{-18}$ ev	0	-
$W^{\pm}$	Weak	80.385 ± 0.015 GeV	±1	-
Z	Weak	91.1876 ± 0.0021 GeV	0	-
g	Stong	0	0	8 combinations

Figure 1.2: Overview of the gauge bosons of the SM, given with their properties and associated interactions.

In order to analyze SM lagrangian, one can proceed by dividing the lagrangian into the following terms:

$$\mathcal{L}_{SM} = \mathcal{L}_{gauge} + \mathcal{L}_{Higgs} + \mathcal{L}_{matter} + \mathcal{L}_{Yukawa} + \mathcal{L}_{int}. \tag{1.2}$$

1.2.1 Mathematical framework

1.2.1.1 Quantum Electrodynamics (QED)

Quantum Electrodynamics (QED) is an field theory based on the local gauge symmetry U(1) (Abelian gauge) which describes the interaction between fermion fields with spin-1/2 and their interactions through the electromagnetic force with spin-1. The QED lagrangian described by the relativistic Dirac Lagrangian density:

$$\mathcal{L}_{QED} = \frac{-1}{4} F_{\mu\nu} F^{\mu\nu} + \bar{\Psi}(i\gamma^{\mu} D_{\mu} - m\bar{\Psi})\Psi \tag{1.3}$$

Where Ψ is the Dirac field, m is the mass of the electron or positron, γ^{μ} the Dirac matrices, D_{μ} is the gauge covariant derivative defined as:

$$D_{\mu} = \partial_{\mu} + ieA_{\mu} \tag{1.4}$$

and $F_{\mu\nu}$ the electromagnetic feld tensor:

$$F_{\mu\nu} = \partial_{\mu} A_{\nu} + \partial_{\nu} A_{\mu} \tag{1.5}$$

The property of the QED Lagrangian to be invariant under the local gauge transformations:

$$\begin{aligned} \Psi(x) &\longrightarrow e^{i\alpha(x)}\Psi(x) \\ A_{\mu} &\longrightarrow A_{\mu} + \tfrac{1}{e}\partial_{\mu}\alpha(x) \end{aligned}$$

The expression of the Lagrangian is:

$$\mathcal{L}_{QED} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \bar{\Psi}(\gamma^{\mu}\partial_{\mu} + m)\Psi + ieA_{\mu}\bar{\Psi}\gamma^{\mu}\Psi \tag{1.6}$$

The first term describes the photon propagator, the charged particles propagation and the last term describes the elementary interaction between photons and charged particles.

1.2.2 The Weak interaction (EW)

The electroweak theory describes the weak and the electromagnetic interactions. the symmetry group of this theory is therefore a not Abelian $SU(2)_L \times U(1)_y$ (chiral group). The electromagnetic interaction U(1), mediated by the photon and the weak interaction, mediated by the Z and $W^{\pm}$ bosons for neutral currents and charged currents respectively. The weak interaction field is invariant under $SU(2)_L$ transformations and its gauge group is not-Abelian.
The subscript L on $SU(2)_L$ indicats that the weak interaction couples only left-handed fermions. Left-handed Ψ_L and the right handed Ψ_R fields are defined through the chirality projectors P_L and P_R :

$$\begin{aligned}\Psi_L = P_L\Psi = \tfrac{(1-\gamma_5)}{2}\Psi\\ \Psi_R = P_R\Psi = \tfrac{(1+\gamma_5)}{2}\Psi\end{aligned}$$

Where γ_5 is defined as $\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3$ is a product of the Dirac matrices.
The Electroweak lagrangian density:

$$\mathcal{L}_{EW} = \bar{\Psi_L}(i\gamma^\mu D_\mu)\Psi_L + \bar{\Psi_R}(i\gamma^\mu D_\mu)\Psi_R - \frac{1}{4}W^i_{\mu\nu}W^{\mu\nu i} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \tag{1.7}$$

where the field strength tensors are defined as:

$$W^i_{\mu\nu} = \partial_\nu W^i_\mu - \partial_{mu} W^i_\nu - g\epsilon^{ijk}W^j_\mu W^k_\nu \tag{1.8}$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_{nu} B_\mu \tag{1.9}$$

and the covariant derivate is:

$$D_\mu = \partial_\mu + \frac{ig}{2}I^k W^k_\mu + \frac{ig'}{2}YB_\mu \tag{1.10}$$

Where ϵ^{ijk} is the Levi Civita tensor, I^k the weak isospin operator, of the $SU(2)_L$ and Y the hypercharge associated to the $U(1)_Y$, which are relate to the electric charge Q through the Gell-Mann Nishijima relation:

$$Q = I_3 + Y/2 \tag{1.11}$$

The coupling constants g and g' are related by the weak mixing angle (θ_W):

$$\theta_W = \frac{g'}{g} \tag{1.12}$$

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

$$A_\mu = \cos\theta_W B_\mu + \sin\theta_W W^3_\mu \tag{1.13}$$

$$Z^0_\mu = \sin\theta_W B_\mu + \cos\theta_W W^3_\mu \tag{1.14}$$

$$W^\pm_\mu = \frac{1}{2}(W^1_\mu \mp iW^2_\mu) \tag{1.15}$$

In the model of electroweak interactions the gauge bosons $A, W^{\pm}, Z^0$ are massless itself. To give mass to the bosons of the electroweak theory a process called spontaneous symmetry breaking acts. This happens through a doublet of complex fields called Higgs scalars which generate mass as results of their self-interactions.

1.2.2.1 Quantum Chromodynamics (QCD)

Quantum chromodynamics (QCD) the gauge theory that describes the strong interactions of colored quarks and gluons. is based on the non-Abelian gauge theory SU(3)c color group. The number of colors is therefore $N_c = 3$, and a quark can have any of three colors, usually called red,green and blue. The gauge sector consists of gluons massless gauge bosons that transform in the adjoint representation of the gauge group. There are therefore N_{2c} - $1 = 8$ kinds of gluon. They have to satisfy the commutation relations:

$$[T^{\mu}, T^{\nu}] = i f^{a} bc T^{c} \tag{1.16}$$

where $f^{a}bc$ are the structure constants characterising the SU(3) group.
The Lagrangian density of QCD is the Yang-Mills Lagrangian, coupled to the six fermion fields Ψ_f:

$$\mathcal{L}_{QCD} = \frac{-1}{4} F^{a}_{\mu\nu} F^{\mu\nu}_{a} + \sum_{q=1}^{6} \bar{\Psi}^{i}_{q} (i\gamma^{\mu} D_{\mu} - m) \Psi^{i}_{q} \tag{1.17}$$

Where Ψ^{i}_{q} denote the vector of three components corresponding to the different colors of a given quark type, $\Psi_q = (\Psi_q, \Psi_q, \Psi_q)$, γ^{μ} the Dirac matrices that expresses the vector nature of the strong interaction.
The gauge covariant derivative D_{μ} under the non-Abelian local gauge transformation introduced by :

$$D_{\mu} = \partial_{\mu} - i g_s T_a A^{a}_{\mu} \tag{1.18}$$

Where the g_s is the strong coupling constant and A^{a}_{μ} are the gauge fields.
The tensor field for a gluon $F^{a}_{\mu\nu}$ is givan by :

$$F^{a}_{\mu\nu} = \partial_{\mu} A^{a}_{\nu} + \partial^{a}_{\nu} A^{a}_{\mu} + g f^{abc} A^{b}_{\mu} A^{c}_{\nu} \tag{1.19}$$

where $g f^{abc} A^{b}_{\mu} A^{c}_{\nu}$ term discribes the interactions among the gauge fields, and is resoponsable for the non abelian nature of QCD.
The QCD strong coupling costant is defined, in a leading logarithm approximation, as follows:

$$\alpha_s(Q_2) = \frac{\alpha_s(Q^2_0)}{1 + B\alpha_s(Q^2_0) \ln(\frac{Q^2}{\Lambda^2_{QCD}})} \tag{1.20}$$

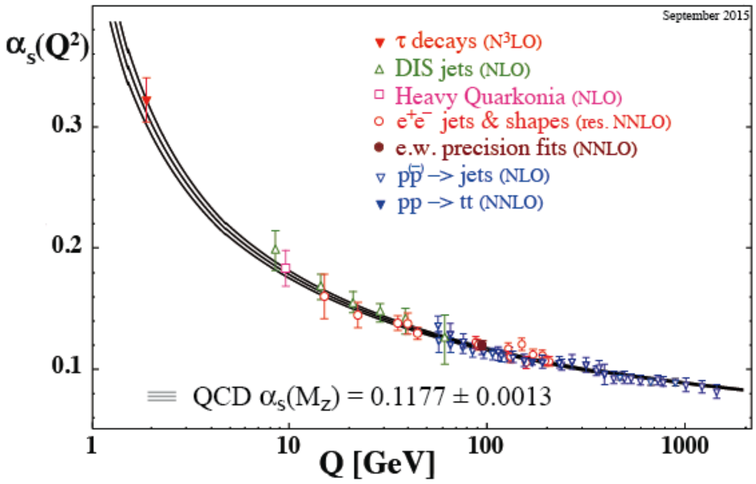


Figure 1.3: Strong running coupling constant distribution α_s with respect to the momentum transfer.

1.2.2.2 Spontaneous breaking of the electroweak symmetry

The requirement of gauge invariance forbids the introduction of mass terms for the fermionic fields and, also, for the gauge fields of the weak interaction. The mechanism through which these fields acquire mass in the SM is so-called spontaneous breaking of the electroweak symmetry (electroweak symmetry breaking, EWSB) is performed via the Brout, Englert and Higgs ("BEH") mechanism.
In the SM, the Higgs field is a doublet in SU(2):

$$\Phi = (\Phi^+, \Phi^0) \tag{1.21}$$

where Φ^+ and Φ^0 are the following complex fields:

$$\Phi^+ = \frac{\Phi_1 + i\Phi_2}{\sqrt{2}}, \tag{1.22}$$

$$\Phi^0 = \frac{\Phi_3 + i\Phi_4}{\sqrt{2}} \tag{1.23}$$

This doublet has hypercharge Y = 1 and isospin $I = \frac{1}{2}$; the first component has positive electric charge, while the second one is electrically neutral.
The Lagrangian of the Higgs field is given by:

$$\mathcal{L}_{Higgs} = \partial_\mu \Phi^+ \partial^\mu \Phi + V(\Phi^+\Phi) \tag{1.24}$$

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

$$V(\Phi^+\Phi) = \mu^2 \Phi^+\Phi + \lambda(\Phi^+\Phi)^2 \tag{1.25}$$

Where λ and μ^2 two real parameters depend respectively to the mass term and the strength of the self-interaction term. The shape of the potential depends on the value of these parameters which has a minimum, which are illustrated in figure 1.4:

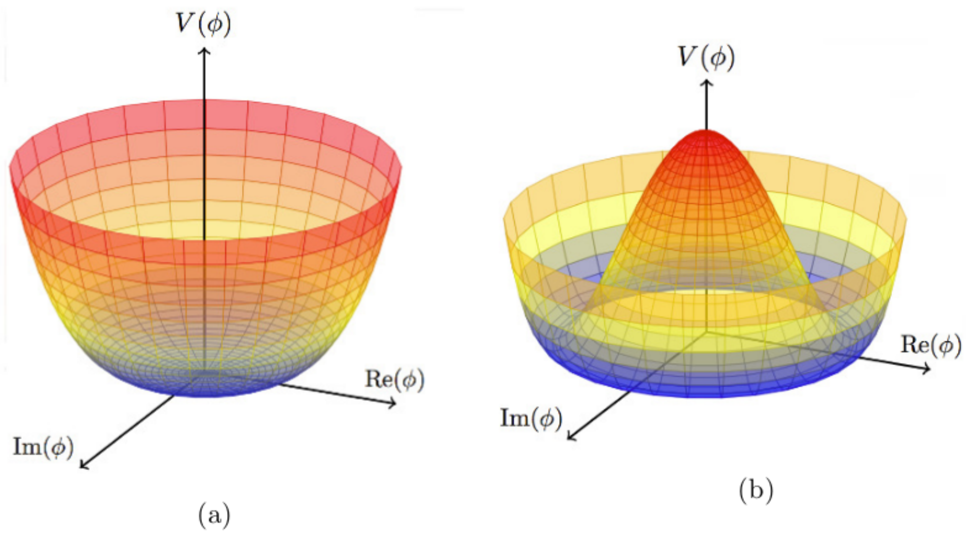


Figure 1.4: Higgs potential in the case (a) $\lambda > 0$ and $\mu^2 > 0$ and (b) $\lambda > 0$ and $\mu^2 < 0$

- if $\lambda < 0$ is unphysical, leading to no stable minima.
- If $\lambda > 0$ and $\mu^2 > 0$ the potential has a single solution to the minimization which corresponds to $|\Phi| = 0$, and gives as vacuum expectation value (VEV) $< \Phi >_0 \equiv 0 < |\Phi| < 0 = 0$.
- if $\lambda > 0$ and $\mu^2 < 0$ the minimum of the potentiel V(Φ) is found in:

$$\Phi^2 = \frac{\mu^2}{2\lambda} \equiv \frac{\nu^2}{2} \quad \text{With} \quad \nu = \sqrt{\frac{-\mu^2}{\lambda}} \tag{1.26}$$

and therefore the field $|\Phi|$ has a non-zero vacuum expectation value $< \Phi >_0 \equiv 0 < |\Phi| < 0 = \frac{\nu}{\sqrt{2}}$. The fundamental vacuum state is no more invariant under the $SU(2)_L \times U(1)_Y$, meaning that these two symmetries are now broken. According to the Goldstone theorem [7], the breaking of a continuous symmetry leads to the appearance of new massless scalar particles as field excitations. These new degrees of freedom are absorbed by the existing gauge bosons, thus giving them mass. Since the photon is the only electroweak boson known to be massless, the choice of the Higgs vacuum has not to break the electromagnetic symmetry, $U(1)_{EM}$, so the component of the Higgs doublet that acquires a VEV is the neutral one:

$$\Phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \nu \end{pmatrix} \tag{1.27}$$

We can then expand the field $|\Phi|$ considering small excitations around the minimum:

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \nu + h \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \nu \end{pmatrix} + \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ h \end{pmatrix} \tag{1.28}$$

Where h being the observable singlet scalar Higgs field, and ν being the vacuum expectation value. Breaking Φ into the two added terms is useful because the origin of $W^{\pm}$ boson and Z boson masses is apparent when plugging the ν term into the EW Lagrangian's Higgs term, and the origin of fermion masses is apparent when plugging the ν term into the EW Lagrangian's Yukawa terms. By substituting equation 1.24 into 1.28 relates the mass of the Higgs boson to the parameters in the potential, some relations between the masses can be obtained, namely:

$$m_H = \sqrt{-2\mu^2} = \sqrt{2\lambda}\nu. \tag{1.29}$$

$$m_w = \frac{1}{2} g\nu \tag{1.30}$$

$$m_Z = \frac{m_w}{\cos\theta_W} \quad \text{and} \quad m_\gamma = 0 \tag{1.31}$$

Where θ is the Weinberg angle and m_H the mass of the Higgs boson.Owing to the experimentally measured mass of the W boson, the vacuum expectation value can be obtained as $\nu = 246$ GeV. The value for the parameter μ and thus for the Higgs boson mass is not predicted by the SM and can only be derived from measurements, which have recently been conducted at the Large Hadron Collider. The newly discovered boson complies with the characteristics expected from a SM Higgs boson, and its mass has been measured to be 125.09 $\pm$ 0.21(stat) $\pm$ 0.11 (syst)[GeV].

1.2.3 Cabibbo-Kobayashi-Maskawa matrix

In the case of fermions, the coupling to the Higgs boson is directly proportional to the fermion mass. In the quark sector, this transformation is encoded in the Cabibbo–Kobayashi–Maskawa (CKM) matrix [8], that describes the mixing of the down-type quarks. In the SM with three generations of fermions, the CKM matrix is parametrized by three angles and one complex phase that provides the only source of CP violation [9] defined as:

$$\begin{pmatrix} d\prime \\ s\prime \\ b\prime \end{pmatrix} = \overbrace{\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}}^{V_{CKM}} \times \begin{pmatrix} d \\ s \\ b \end{pmatrix}$$

where d, s, b denote the mass eigenstates of down-type quarks. While $d\prime, s\prime, b\prime$ are the weak eigenstates. The term $|V_{ij}|$ represent the probability of a quark of flavor i to decay to a quark of flavor j via the a W-boson. The value of the matrix elements are obtained experimentally [10]:

$$V_{CKM} = \begin{pmatrix} 0.97425 & 0.2253 & 0.00413 \\ 0.225 & 0.986 & 0.0412 \\ 0.0084 & 0.0400 & 0.99 \end{pmatrix} \tag{1.32}$$

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

1.3 Dark photons and resonant monophoton signatures in Higgs boson decays at the LHC

Although dark matter (DM) is five times more abundant in the Universe than ordinary baryonic matter [11], its properties are yet unknown. It is plausible that the dark sector, which is weakly coupled to the standard model (SM), possesses rich internal structure and interactions in complete agreement with present astroparticle and cosmological observations. It is also conceivable that a hidden sector could contain an extra long-range force mediator among the dark particles. The most simple example is provided by a new unbroken U(1) gauge group, predicting a dark (or hidden) photon in its spectrum [**essig2013dark**]. Dark-photon scenarios have been extensively considered in the literature in the framework of new physics extensions of the SM gauge group [12, 13]. In cosmology the dark photons may solve the small-scale structure formation problems and, for massless dark photons [14], predict dark discs of galaxies [15]. In astroparticle physics dark photons may induce Sommerfeld enhancement of DM annihilation cross section needed to explain the PAMELA-Fermi-AMS2 positron anomaly [14], may assist light DM annihilations to reach the phenomenologically required magnitude, and make asymmetric DM scenarios phenomenologically viable [16]. Dark/hidden photon scenarios have also been extensively considered in beyond-the-SM frameworks in particle physics [12].
In this section we show that, in the unbroken dark U(1) scenarios, the Higgs-boson two-body decay $H \to \gamma\bar{\gamma}$ to one photon γ and one dark photon $\bar{\gamma}$ can be enhanced despite existing constraints, providing a very distinctive new physics signature of a single photon plus missing energy at the Higgs resonance.

1.3.1 Higgs boson production

Although the theory cannot predict the Higgs boson mass, the Higgs couplings to fermions and bosons are predicted to be proportionate to the associated particle mass. As a result, the Higgs production and decay processes are dominated by channels involving Higgs boson coupling to heavy particles, particularly $W^{\pm}$ and Z bosons and the third generation of fermions. Concerning the other gauge bosons, the Higgs does not couple to photons and gluons at tree level, but only by one-loop diagram, where the main contribution is given by t loops, for the $gg \longrightarrow H$ channel, and by W+W- and t loops for the $yy \longrightarrow H$ channel.
In the SM, Higgs boson production at the LHC mainly occurs through the following processes, listed in order of decreasing cross section at the Run 1 centre-of-mass energies:

Vector boson fusion
The vector boson fusion mechanism (VBF),

$$\mathrm{pp} \longrightarrow \mathrm{q}\bar{q} \longrightarrow \mathrm{qqH},$$

For the Higgs boson at $m_H = 125$ GeV, its contribution to the total cross section is of the order of 8%. However, this channel is very interesting because of its distinct experimental signature: the presence of two spectator jets with high invariant mass in the forward region provides a powerful tool for tagging signal events and distinguishing them from the backgrounds, improving the signal to background ratio despite the low cross section. Up to NNLO, the VBF cross section has been determined.

Associated production The associated production mechanism (VH), described by the process,

$$\mathrm{pp} \longrightarrow \mathrm{q}\bar{q} \longrightarrow \mathrm{VH},$$

is smaller due to the presence of an antiquark that, in a pp machine, does not come from the valence but from the sea. In this case, the selection of an event can be performed by reconstructing the original boson that radiated the Higgs. The VH cross section has been calculated including the QCD corrections up to NNLO, whereas the electroweak corrections are known at NLO.

Associated production with a t$\bar{t}$ pair
The production mechanism of the Higgs boson associated with a t$\bar{t}$ pair,

$$\mathrm{pp} \longrightarrow \mathrm{q}\bar{q} \longrightarrow \mathrm{t}\bar{t}\mathrm{H},$$

is the lowest yielding production mechanism, but the existence of a t$\bar{t}$ pair in the final state offers a clean experimental signature. Furthermore, its cross section grows by a factor of four from 8 to 13 TeV, and this production mode, which provides direct access to top quark coupling, plays an important role in testing the SM. This process has been computed with a precision up to NLO.

Gluon fusion production
The gluon gluon fusion (ggF) mechanism,

$$\mathrm{pp} \longrightarrow \mathrm{gg} \longrightarrow \mathrm{H},$$

In the entire Higgs mass range, is dominant at the LHC. The gluons coupling to the Higgs boson is mediated via a triangular loop of virtual quarks, with the t contribution dominating due to the large top mass. The theoretical cross section includes QCD corrections up to next-to-next-to-leading-order (N3LO) and next-to-next-to-leading-log (NNLL), although electroweak corrections are known at next-to-leading order (NLO).
Recently, a massless dark photon scenario has been foreseen in the framework of a theoretical proposal aimed to naturally solve the flavor hierarchy problem. This model predicts a new Higgs-boson decay channel into a photon (γ) and a massless dark-photon ($\bar{\gamma}$).

1.3.1.1 Dark-photon production via Higgs production at the LHC

In [17] the potential of the Run-1 dataset at the LHC has been studied at parton level by considering the gluon fusion process $gg \to H \to \gamma\bar{\gamma}$ at the center of mass proton energy of 8 TeV. It is particularly informative to look at a comparison of the distribution of the missing transverse mass M_T of the resonating $\gamma\bar{\gamma}$ system in the Higgs signal (where $\Delta\phi$ is the azimuthal distance between the photon and dark-photon three momenta) and of the main irreducible and reducible backgrounds, as shown in Figure 1.5 [17].

$$M_T = \sqrt{2p_T^{\gamma} E_T^{miss}(1 - cos(\Delta\phi))} \qquad (1.33)$$

In that figure, the $\gamma + E_T^{miss}$ transverse invariant mass distribution (in fb/GeV) of the signal for $H \to \gamma\bar{\gamma} = 5\%$ (shown in red) is compared with: i) the γj channel (grey); ii) the $\gamma Z \to \gamma\nu\bar{\nu}$ channel (blue); iii) $jZ \to j\nu\bar{\nu}$ (green), where the jet is misidentified as a photon; and iiii) $W \to e\nu$ (yellow), where the electron (positron) is misidentified as a photon. The resonant structure of the signal emerges clearly from the background.

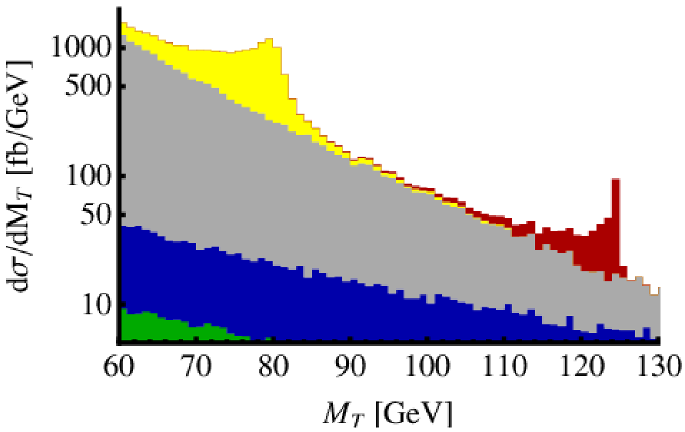


Figure 1.5: Distribution of the $\gamma\bar{\gamma}$ transverse invariant mass assuming $BR(H \to \gamma\bar{\gamma}) = 0.5\%$ (red) compared to the main irreducible and reducible backgrounds, [17].

1.3.1.2 Higgs decay into massless dark photons and dark photon

The new Higgs boson decay channels involving a massless dark photon are $H \to \gamma\bar{\gamma}, \bar{\gamma}\bar{\gamma}, Z\bar{\gamma}$ [18]. In a nondecoupling regime, they can have sizeble rates [17]. The signature of a massless dark photon The Higgs decay into dark photons and photon will then affect the Higgs invisible decay width. Here we consider the single dark photon Higgs signature, as detailed in Figure 1.6, where the effective H$\gamma\bar{\gamma}$ (HZ$\bar{\gamma}$) interaction is described as the low energy approximation of a one-loop diagram mediated by heavy messengers Si. We focus in particular on the $H \to \gamma\bar{\gamma}$ channel giving rise to a resonating gamma-plus-missing-momentum signature, characterized by one monochromatic photon with energy $\frac{m_H}{2}$, and same amount of missing energy. Note that the photon energy is more than twice the one in the resonance corresponding to the rare SM decay $H \to \gamma Z \to \gamma\nu\nu$. While a quantitative prediction for the $H \to \gamma\bar{\gamma}$ decay width will be in general dependent on quite a few model parameters, in a simplified-model analysis [17] one can obtain typical allowed ranges which correspond to branching ratios (BRs) up to about 5%. This is shown in Figure 1.5 [17], where $BR(H \to \gamma\bar{\gamma})$ is plotted versus the dark U(1) fine structure constant $\bar{\alpha}$. Such large $BR(H \to \gamma\bar{\gamma})$ might indeed lead to an observable signal in Higgs production at the LHC.

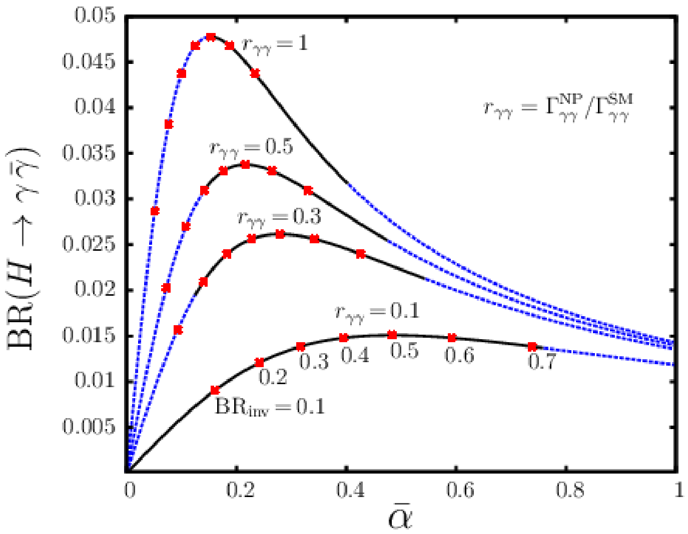


Figure 1.6: $BR(H \to \gamma\bar{\gamma})$ versus the dark U(1) fine structure constant $\bar{\alpha}$. The black solid lines show the allowed BR ranges as a function of the relative new physics contribution to the $H \to \gamma\gamma$ width, and of the $H \to \bar{\gamma}\bar{\gamma}$ contribution to the Higgs invisible BR [17].

1.4 Limitations of the SM

The SM as it stands now is a powerful theory tested by many experimental observations. In addition , the existence of new particles such as the charm and top quarks or the bosons of W, Z, and Higgs was predicted prior to their collider discovery. However some questions that cannot be answered within the SM framework and it is therefore thought that the SM is an effective field theory of a more complete description of the subatomic world. Here are the major questions:

Gravity The SM does not include a quantum field description of gravity, since none of the current attempts for it have been successful. Its inclusion is in particular difficult at high energies (close and above the Planck scale), where infinities due to quantum effects arise.

Neutrino oscillations The SM predicts that neutrinos are massless particles. However, the observation of neutrino oscillations implies non-zero masses. In addition, only left-handed neutrinos are observed, and thus, they cannot acquire their mass via the Yukawa interaction with scalars since that would imply the presence of right-handed neutrinos.

Matter-antimatter asymmetry Particles and antiparticles are created and annihilated in pairs, but the universe today contains mostly matter, while the Big Bang theory predicts equal quantity of matter and antimatter. The presence of a phase in the CKM matrix allows for CP violation, which provides a mechanism for the matter-antimatter asymmetry. However, the amount of CP-violation is too small to account for the asymmetry observed.

Strong CP problem It refers to the fact that, even if QCD does not forbid CP violation in strong interactions, it has not yet been observed in experiments.

Dark Matter and Dark Energy Dark Matter is a type of matter which does not interact via electromagnetic interactions and which existence has been inferred from gravitational effects in astrophysical and cosmological observations. While, according to the SM of cosmology, it should account for a $\approx 25\%$ of the energy content of the universe, the SM does not provide a proper candidate particle. On the other hand, about a 68% of the energy content of the universe is predicted to come from an even more intriguing Dark Energy; responsible of the accelerated expansion of the universe.

Naturalness The SM has 19 arbitrary parameters (including the fermion masses or the strong and electroweak couplings). The naturalness principle states that the dimensionless ratios of the parameters of a theory should be of order unity. This is in clear contradiction with, for example, the masses of the fermions, which range from ≈ 1 MeV for the first generation of fermions to about 173 GeV of the top quark. While this is not a problem of the theory itself, it is often considered as an indication of unknown principles underlying a more complete theory.

Hierarchy problem It arises from the huge gap between two fundamental scales of physics: the EW scale ($\sim$ 10 2 GeV) and the Planck scale ($\sim$ 10 19 GeV). The fact that the Higgs boson mass is well below the Planck scale seems to require new physics between the two scales. If the SM is valid up to the Planck scale instead, the Higgs mass need to be unnaturally fine tuned to cancel out the radiative corrections.

1.5 Top quark in the Standard Model

1.5.1 Top Quark.

The top quark is the heaviest particle of the SM and was discovered 1995 at the Fermilab Tevatron via the strong interaction in pp collisions at $\sqrt{s} = 1.8$ TeV. In the recent years since the LHC has been the world's highest-energy centre for collider physics, vast numbers of top quarks have been produced allowing for the measurement of the properties of what is now the heaviest of all known fundamental particles. it plays a special role in many extensions of the SM. Due to its largest mass with mt $= 173.34 \pm 0.27$ (stat.) ± 0.71 (syst.) GeV of the foundamental particles known, significantly heavier than the other quark in its generation, and Yukawa coupling among the elementary SM fermions. So, it is naturally related to the electroweak symmetry breaking and also offers the unique opportunity to explore the properties of a bare quark, such as its spin and mass. The top quark production is a backgrounds in many searches for physics beyond the SM; for this reason a precise understanding of the top quark signal is crucial to claim new physics. In addition, its very large mass gives rise to large coupling with the Higgs boson.
In this chapter, after an introduction on the top quark theoretical background in the Standard Model, the production and decay modes will be discussed in detail.

1.5.2 Top Quark Production

The Top Quarks can be produced via two main mechanism: Top quark pair production and single top production.

Top quark pair (tt) produced via strong interaction either via quark anti-quark annihilation $q\bar{q} \to t\bar{t}$ or gluon fusion $gg \to t\bar{t}$ at leading order in QCD. Figure 1.7 illustrates the various sub-processes that produce top quark pairs, of which the most dominant process at the LHC is gluon fusion.

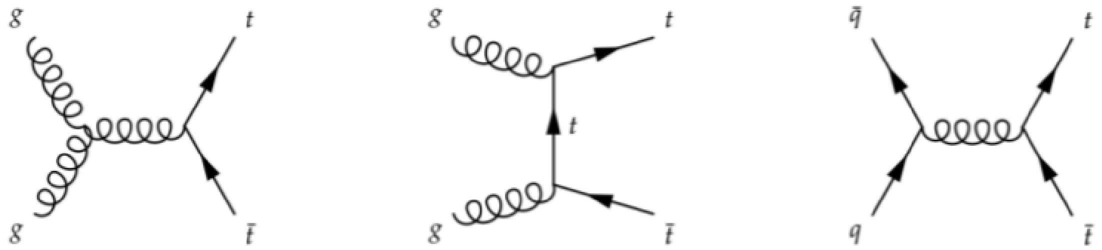


Figure 1.7: Leading order (LO) dominant production modes of top pairs at the LHC: gluon-gluon fusion (left and center) and quark-antiquark annihilation (right).

The single top (t) produced through the electroweak interaction during the hard scattering. specifically depending on the Wtb vertex, as well as the measurement of the Cabibbo–Kobayashi–Maskawa (CKM) matrix element $|V_{tb}|$. Only one top quark is produced thus it is usually referred as the "single-top" production. Figure **??** shows the different diagrams of single quark top production. The single-top production can be categorized into three production modes:

s-channel:the top quark is produced by the exchange of a virtual time-like W boson which produced from two quarks and decaying into.

t-channel: the top quark is produced via the exchange of a virtual space-like W boson. This channel has the largest cross-section and is hence the dominant production mode of single top quarks

Wt-channel: the top quark is produced in association with a real W boson.

1.5.3 Decay Modes and Event Topologies

The large mass of the top quark, it shorts lifetime and it is the only quark that decays before hadronization takes place. Therefore, the top-quark has to be reconstructed from its decay products. The dominant decay mode of the top quark is given by the flavor-changing EW process, $t \longrightarrow bW^+, t \longrightarrow sW^+, t \longrightarrow dW^+$. The probability for each type of the down-type quark to occur as decay product is given by the CKM matrix elements involving top quarks, $|V_tq|^2$ with q=b,s,d. this means that top quarks decay almost uniquely into a W boson and a bottom quark ($t \longrightarrow W^+b, t^- \longrightarrow W^-b^-$), which can be also written in terms of branching ratios BR:

$$BR = \frac{\mathcal{BR}(t \longrightarrow Wb)}{\mathcal{BR}(t \longrightarrow Wq)} = \frac{|V_tq|^2}{|V_tb|^2 + |V_ts|^2 + |V_td|^2} \approx 0.99 \tag{1.34}$$

The final signature of a top quark decay can be characterized based on the decay of the two W bosons: the hadronic decay to a pair of quarks, with $\mathcal{BR}(W \longrightarrow qq)$ is about 64% or to decay leptonically based to a charged lepton and a neutrino, with

$\mathcal{BR}(W \longrightarrow l\nu)$ is about 33%
There are three possible event topologies are classified as dilepton, lepton+jets, and all-jets topologies in figure 1.8. the word lepton always refers to an electron or muon on these three categories and exclude events with one or more tauonic W decays, which are more difficult to reconstruct and provide less mass information than corresponding events with electronic or muonic W decays because of the additional neutrinos from τ decays [19].
Dilepton This channel is characterized by the disintegration of the two bosons W into lepton and neutrino, tt . Its signature is very clear in hadronic collisions, as it contains two leptons and two neutrinos. However, its branching ratio is (10.5%) (including leptons τ), which is quite low, especially for an analysis of the LHC data.
All-hadronic This is the disintegration mode where the two W bosons disintegrate into light quarks tt With a branching ratio of 46.2%, this is the dominant mode of disintegration of a pair tt Its final state is composed of 6 jets which makes it difficult to discriminate against the multijet processes produced abundantly during pp collisions.
Semi-lepton or lepton + jets tt has a branch ratio of 43.5%. which is almost equivalent to the hadronic channel. Its signature includes a lepton, a neutrino and 4 jets, two of which are from quarks b. This signature is a good compromise between a clear signature and a high connection ratio, it is the preferred channel to study the first top quarks produced at the LHC

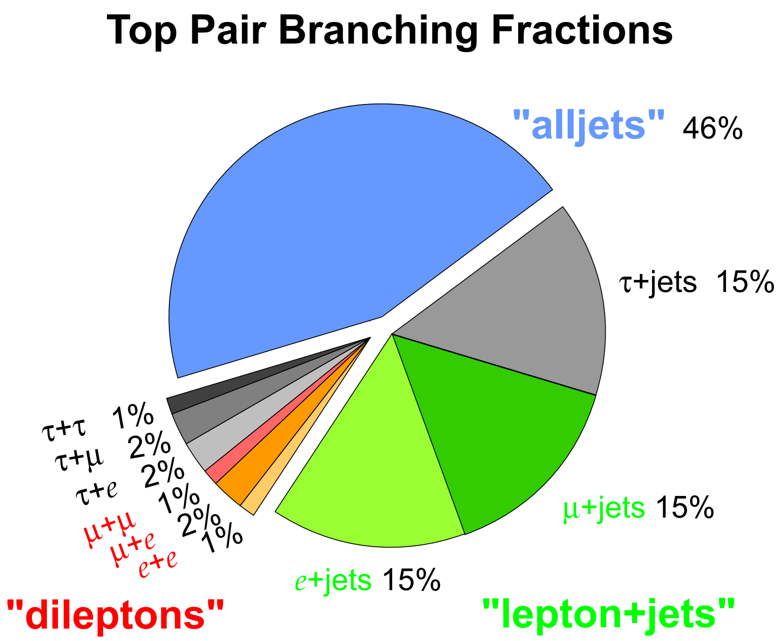


Figure 1.8: Branching ratio for $t\bar{t}$ decay modes.

1.5.4 Top quark mass

The results of different measurements of m_{top} by ATLAS and CMS at the LHC, are summarized in Figure 1.9. With the world average value combined from the four experiments ATLAS and CMS (LHC), CDF and Dϕ(TEVATRON) performed in March 2014.

- The mass of the top quark (m_{top}) is a key parameter of the SM. It is not predicted by the model and its large value is affects many of the observables in top quark production and decay.
- The large m_{top} is also responsible for large contributions to quantum loop corrections to electroweak observables.
- is one of the fundamental ingredients for the calculation of the radiative corrections, which connect electroweak processes. These corrections depend on the masses of the Higgs boson and top quark via the loop quantum corrections. At one loop the gho parameter which relates the W and Z boson masses and the weak angle θ_W
- The most precisely studied property of the top quark is its mass.

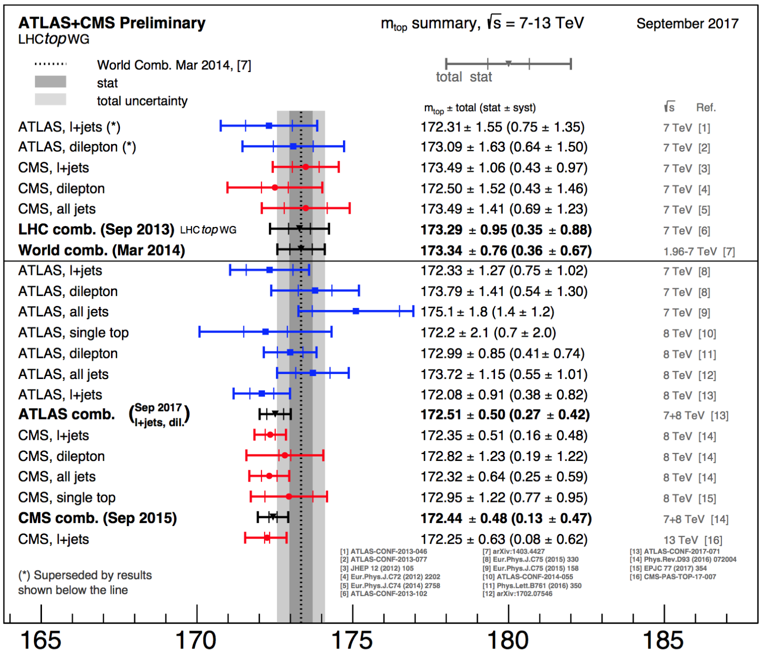


Figure 1.9: Summary plot of recent top quark mass measurements made at the leading LHC (ATLAS and CMS) and Tevatron (CDF and D0) experiments

1.6 Beyond the standard model and the role of the top quark

Following several successful predictions, it is known that the Standard Model can not be a complete and self-consistent theory of particle physics. Indeed, due to various theoretical aspects, the Standard Model is merely an approximation to some more general, yet unknown, theory. The Standard Model does not explain, for instance, why the particles have masses or why the gravity is so weak with respect to the other fundamental forces of Nature. The search for experimental evidence of new phenomena, which can not be taken into account by the Standard Model, is therefore an essential step in establishing the viability of new candidate theories that seek to

address some of the shortcomings of the Standard Model.
Many of these theoretical models predict the existence of one or more new massive particles decaying in large proportion to a pair of top-antitop quarks. In the following sections, two specific models will be briefly outlined.
The reconstructed top-antitop invariant mass is defined as:

$$M_{t\bar{t}} = \sqrt{(P^t + P^{\bar{t}})_\mu (P^t + P^{\bar{t}})^\mu} \tag{1.35}$$

Where p is the top four-momentum and a sum over the components $\mu = 0, ..., 3$ is implicitly assumed. Due to the high mass of the top quark, no real particle predicted by the Standard Model is allowed to decay into the top anti-top pair. Hence, $M_{t\bar{t}}$ is predicted to exhibit a smooth and exponentially-mass distribution. Therefore, a new heavy particle which would decay primarily to a top antitop pair will expose itself as a resonance (a "bump") in the $M_{t\bar{t}}$ spectrum (as expected by the Standard Model), thereby providing a model-independent way of searching for the presence of these new particles. The most striking signal for new top-sector physics can do that through the resonant production of a heavy intermediate state, $[X] \longrightarrow t\bar{t}$. Examples of such heavy resonances include pseudo-scalar Higgs bosons [5], new gauge bosons, Kaluza-Klein gluon and graviton excitations, Technicolor-like dynamic states[25], etc.

1.7 $t\bar{t}$ Resonances

The top quark (t) in the standard model is the most massive known fundamental particle [4, 20]. It has a Yukawa coupling near unity to the Higgs field. This is also closely related to the problem of hierarchy, where the largest corrections to the Higgs mass come from top quark loops. Furthermore, the top quark is of particular interest for the electroweak symmetry breaking mechanism and due to its largest mass and could potentially be connected with new phenomena. Several proposed extensions to the SM predict the existence of heavy particles that decay primarily to top quark pairs.

1.7.1 Benchmark model A: the leptophobic topcolor : $[Z]$ boson

The model of the leptophobic topcolor Z ′ boson is chosen mainly because it has a narrow width, much smaller than the experimental resolution.
The topcolor Z' has been developed to give a natural, dynamic explanation for the large top quark mass as well as to provide a mechanism for the electroweak symmetry breaking EWSB through a strong interaction, known as technicolor-assisted topcolor (TC2) [21]. These interactions have the characteristic energy scale $\bigwedge_{TC} \simeq \bigwedge_{EW} \simeq 1TeV$. The original technicolor manifestation was proposed by Weinberg [22] and Susskind [7] more than 30 years ago. A new gauge interaction encompassing gluons was developed, which provided the longitudinal modes of the W and Z bosons, thus endowing them with mass. TC2 was developed because the original technicolor predicted massless quarks.
The recent evidence for the Higgs boson indicates that EWSB is interpreted by the Higgs mechanism rather than by technicolor. Although it was recently suggested

that the Higgs mechanism could be underpinned by a model similar to technicolor [23]. The use of this Z' boson topcolor as a benchmark model continues irrespective of the validity rather than of the technicolor. It is investigated explicitly pursued because it is a very narrow resonance, meaning that the upper limits determined by the branching ratio of its production cross-section times are to $t\bar{t}(\sigma \times (BR \longrightarrow t\bar{t}))$ are valid for any resonance whose width is narrower than the experimental resolution. There are four main models (Model I, II, III and IV) [35]. In this thesis we use as a benchmark model the Z $\prime$ corresponding to Model IV, a leptophobic, topophyllic boson presented in [27]. This specific Z $\prime$ model would therefore be first observed via its decay to top quarks rather than via the dileptonic decay mode, which has fewer background processes and is usually the favoured search channel for similar BSM resonances. The specific form of the Lagrangian and the partial width for the decays to t$\bar{t}$, u$\bar{u}$ and d$\bar{d}$.

1.7.2 Benchmark modelKaluza-Klein Gluons In The Randall-Sundrum Warped Extra-Dimension Model

The Randall-Sundrum model is theory of extra dimensions. This particular model has a relatively large width 15.3%. This theory of extra dimensions was proposed by Arkani, Dimopoulos and Dvali in 1998. It designed to solve the problem of hierarchy, where the weak scale is generated through an exponential hierarchy from a large scale of the Planck scale. In this particular model, the Standard Model quantum fields and corresponding particles are free to propagate in an extra warped spatial dimension. The excitation of an SM particle into this extra dimension results in massive excited states of excitations of the particle known as Kaluza-Klein (KK). The first excitation of gluon (g_{KK}) in the basic scenario of Randall Sundrum, in particular, is predicted to acquire enough mass to decay into a pair of top-antitop quarks. In the Randall-Sundrum model, the g_{KK} decays primarily to the top quarks (92.5 percent of the time[30]) and would appear to be a broad resonance in the top antitop invariant mass spectrum.The production of g_{KK} takes place in the leading order through the annihilation of light quarks, to which the coupling of g_{KK} is suppressed. The cross section is therefore small, especially since the electroweak precision constraints require that the mass of the g_{KK} be no less than 2 - 3 TeV.

1.8 Summary

In this chapter a theoretical introduction of the Standard Model and the electroweak symmetry breaking is given. The Higgs boson phenomenology at the LHC is then presented, with particular attention to the Higgs to photons and dark photons decay channel. Finally The top quark is integral to searches for phenomena which cannot be explained within the standard model.

Chapter 2

Experimental setup

2.1 Introduction

The Large Hadron Collider (LHC)[24] is the largest and most powerful particle collider ever built. It is located at the European Laboratory for Particle Physics (CERN) on the French-Swiss border close to Geneva (see Figure 2.1), Switzerland. The design was officially approved in 1994 and the first beams were circulated in 2008. Currently in collaboration with over 10,000 scientists and engineers from over 100 countries. CERN currently has 23 Member States, 7 associated Member States, and 6 Observer States, including the Morocco. These states, and many others with scientific contacts or cooperation agreements, finance and operate the LHC and receive access to the collected data. Thus, CERN continues to support its missions of strong international collaboration, scientific discovery, and technological innovation. The accelerator is placed at a depth of 175 m under ground with a circumference of about 27 km and provides two counter-rotating beams of protons or heavy-ions, colliding at four points. The two beams run in opposite directions and intersect at four points along the ring where the detectors ATLAS (A Toroidal LHC Apparatus) [25], CMS (Compact Muon Solenoid)[26], ALICE (A Large Ion Collider Experiment) [**collaboration2008cms**] and LHCb (Large Hadron Collider Beauty) [**clemencic2011lhcb**] are located.
The ATLAS experiment, is a multi-purpose detector, designed to take benefit of the extraordinary energy accessible at the LHC and observe phenomena that contain especially huge particles which were now not observable using formerly lower-energy accelerators. Along with CMS, ATLAS used to be one of the two LHC experiments involved in the discovery of the Higgs boson in July 2012. It was once also designed to search for proof of particle physics beyond the Standard Model (BSM). The analyses in this work use proton-proton collision at a center of mass energy of 13 TeV with ATLAS detector at the LHC during 2015, 2016, 2017 and 2018. This chapter introduces the LHC accelerator in Section 2.1 and the ATLAS detector with its sub-detectors in Section 2.2.

2.2 Large Hadron Collider

2.2.1 The LHC Lead Injection Chain

The LHC [24] is supplied with high-energy protons through a multi-stage, injector chain illustrated in Figure 2.1.is a succession of machines that accelerate particles to increasingly higher energies. Each machine boosts the energy of a beam of particles before injecting the beam into the next machine in the sequence. The source of protons in the LHC is a bottle of hydrogen gas. The Hydrogen gas is used as a source of protons, with their electrons stripped off by an electric field.The protons are initially accelerated in the linear accelerator Linac 2 to energies of 50 MeV. Subsequently, the Proton beam is then split and injected into the four Proton Synchrotron Booster rings where the beams are accelerated to 1.4 GeV, followed by further acceleration in the Proton Synchrotron (PS) to 25 GeV and forms a bunch train with 25 ns spacing. before injection into the LHC, the beam is injected into the Super Proton Synchrotron (SPS) where it reaches an energy of 450 GeV. Once the beams are circulating in the LHC they are further accelerated while keeping the 25 ns spacing. It takes 20 $\sim$ 30 minutes for the protons to reach their maximum energy of 6.5 TeV and both two beams circulate for many hours inside the LHC beam pipes under normal operating conditions. The two beams then collide within four detectors: ALICE, ATLAS, CMS and LHCb, where the total energy at the point of collidation is equal to 13 TeV. The four collision detectors are summarized below:

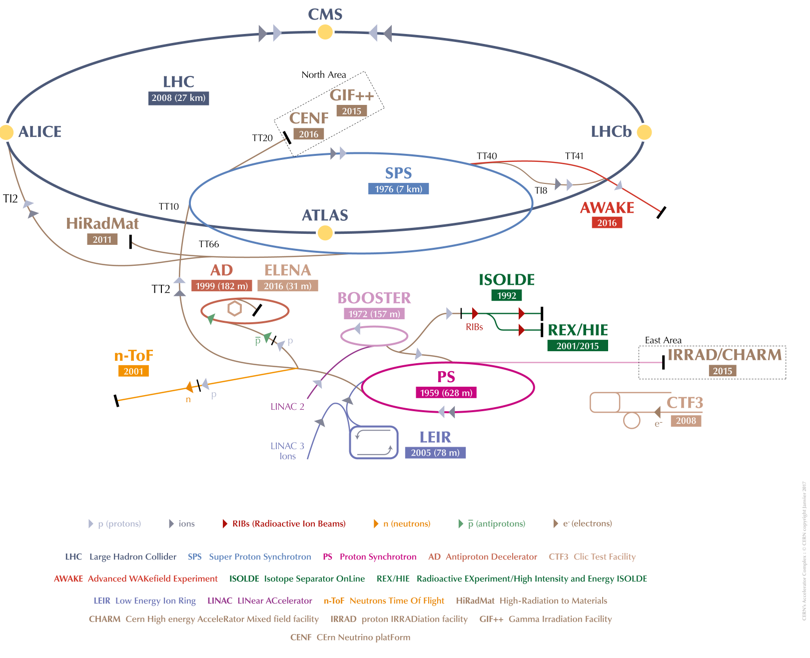


Figure 2.1: Distribution of the LHC complex.

- **ALICE** (**A** **L**arge **I**on **C**ollider Experiment): Optimised for the study of interactions between heavy ions

- **ATLAS** (**A** **T**oroidal **T**HC **A**pparatuS): Optimized for the discovery and measurement of the Higgs boson and the search for new physics, but also pursue a considerable program of Standard Model measurements.
- **CMS** (**C**ompact **M**uon **S**olenoid): The other general-purpose detector, like ATLAS, to study the SM including the Higgs boson and look for clues of new physics.
- **LHCb** (**L**arge **H**adron **C**ollider **b**eauty): Aimed for the particular study of the decays of the b-quarks experiment designed to slight differences between matter and antimatter by studying b hadrons.

2.3 Phenomenology of proton-proton collisions

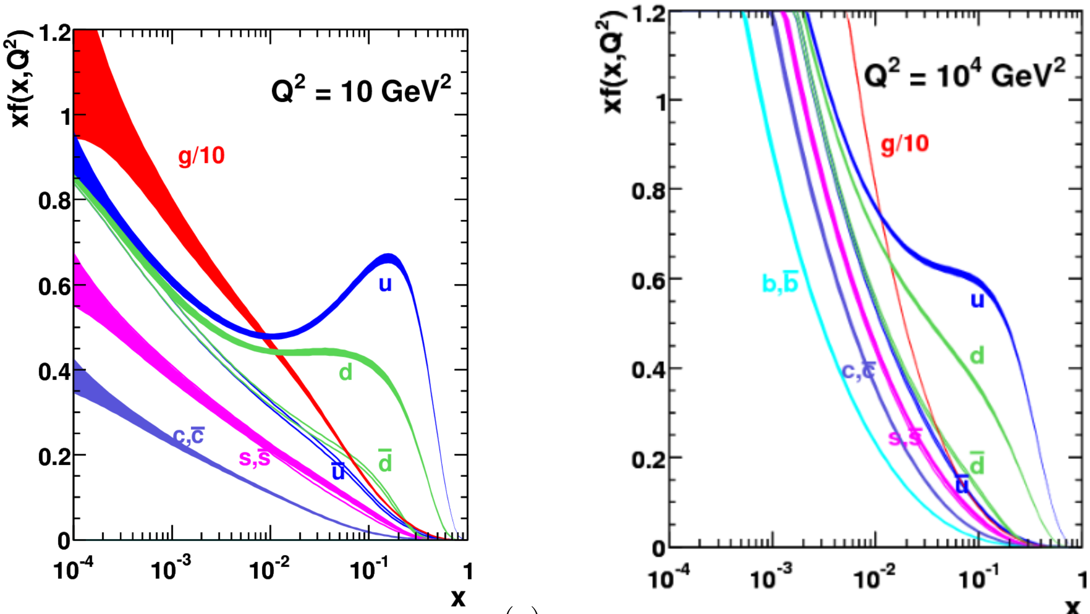
Beams of protons collide at the LHC. One first has to understand the nature of the protons themselves in order to comprehend the outcome of the collisions. A proton is composed of three valence quarks, uud, and the sea quarks coming from gluon splitting into quark and antiquark pairs. At low energies, the momentum of the proton is mainly distributed among the three valence quarks. However, as the energy increases, gluons will also carry some of the momentum. The dynamics of the proton is therefore described by the fraction of its momentum distributed among its constituent partons (quarks and gluons) in the form of parton distribution functions, PDFs. The fundamentals for understanding the hard scattering processes of importance in proton-proton collisions as well as the accompanying interactions are explained in the following subsections, with particular focus on experimental effects such as luminosity and pile-up.

2.3.1 Parton distribution functions

To predict the rates of the various processes it is needed to know the distribution of the partons inside the proton. Parton distribution functions give the probability to find partons in a hadron as a function of the fraction x of the hadron momentum carried by the parton. These parton distribution functions can not be calculated due to the nonperturbative nature of QCD but rather are determined by global fits to data from deep inelastic scattering (DIS), Drell-Yan (DY), and jet measurements. The main experimental inputs come from fixed-target DIS experiments such as NMC, BCDMS, SLAC, HERA or ZEUS. Data from measurements of vector boson production in hadron-hadron colliders (Tevatron and LHC) are also used. The most common used sets of PDFs are provided by the MSTW, the NNPDF and the CTE collaborations. As an example, Figure 2.2 shows parton distribution functions of a proton produced with the MSTW 2008 PDFs at NLO at two different scales, $\mathcal{Q}^2 = 10\ GeV^2$ and $Q^2 = 10^4\ GeV^2$, includind the associated 68% CL uncertainty bands.

2.3.2 The cross section for processes in hadron-hadron collisions

Within the QCD theory, the scattering processes at a hadron collider can be categorized in hard and soft. The hard processes are in general well understood and their



(a)

(b)

Figure 2.2: Parton distribution functions of a proton at two different values of Q^2 using the MSTW 2008 NLO PDF set [27].

properties can be predicted with good precision using perturbation theory. This is not the case for the soft processes in which non-perturbative QCD plays a major role.

One of the physical quantities to connect the theory and experimental measurement is the cross section (σ) for a certain physics process. According to the QCD factorization theorem [36] the cross sections for hard scattering can be factorized into a parton level hard scattering convoluted with the parton distribution functions. The PDF for the parton a inside the hadron A carrying a fraction of the hadron momentum x_a $f_{a/A}(x_a, Q)$ depends on the large momentum scale, Q^2, which characterizes the hard scattering. The total cross section for X production in AB collision (see Figure 2.3), $\sigma_{AB\rightarrow X}$ can thus be written as

$$\sigma_{AB\rightarrow X} = \int dx_a dx_b f_{a/A}(x_a, \mu^2) f_{b/B}(x_b, \mu^2) \hat{\sigma}_{AB\rightarrow X} \tag{2.1}$$

Where Q^2 is chosen to be the factorization scale μ_F , which can be understood as the scale that separates the short- and long-distance physics. The partonic cross-section $\hat{\sigma}_{AB\rightarrow X}$ can be calculated in perturbative QCD giving rise to perturbative ($\mathcal{O}(\alpha_s^n)$) corrections to the leading order cross section

$$\hat{\sigma}_{AB\rightarrow X} = [\hat{\sigma}_{LO} + \alpha(\mu_R)\hat{\sigma}_{NLO} + \alpha(\mu_R^2)\hat{\sigma}_{NNLO} + ...]_{ab\rightarrow X} \tag{2.2}$$

here, μ_R is the renormalization scale for the QCD running coupling. Formally, the cross section calculated to all orders is invariant under changes in μ_F and μ_R . However, To avoid unnaturally large logarithms reappearing in the perturbation series it is sensible to choose μ_F and μ_R values in the order of the typical momentum

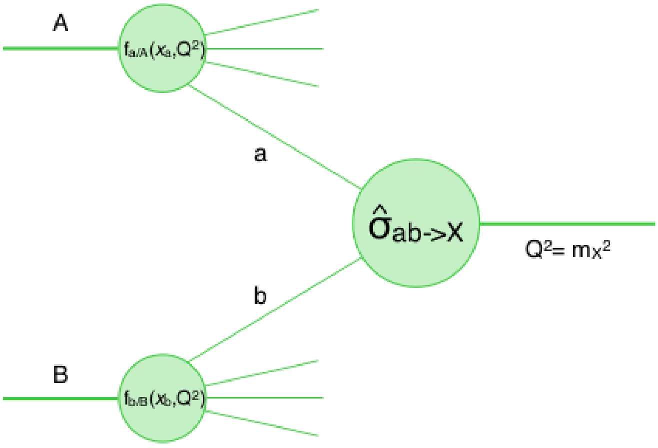


Figure 2.3: Schematic diagram of particle production in a proton-proton collision.

scales of the hard scattering process. A standard choice is $\mu_F = \mu_R = m_X$.

2.3.3 Underlying event

The hard-scattering process is accompanied by a number of softer interactions, collectively referred to as the underlying event (UE). This activity originates from multiple partonic interactions where further partons of the colliding protons undergo interaction, as well as from the proton remnants which are not in a color singlet state. As an illustrative example, Figure 2.4 depicts a proton-proton collision.

The modeling of the underlying event is based on a principle that the $2 \to 2$ parton-parton interaction is the dominant process in the leading oder. These residual particles may lead to multiple parton interactions within one collision which need to be modeled to properly account for its impact on the other high p_T measurements and the detector resolution. A common model to simulate the underlying event, including hard and soft interactions, parton distributions, initial-state and final-state parton showers, multiparton interactions, fragmentation and decay, is implemented in Pythia and assumes that multiple interactions take place in a substantially independent way so that Poisson statistics can be used, and that they can be described by perturbative QCD above a certain scale p_T^{min}, which becomes one of the main parameters of the model. An alternative model is implemented in Herwig.

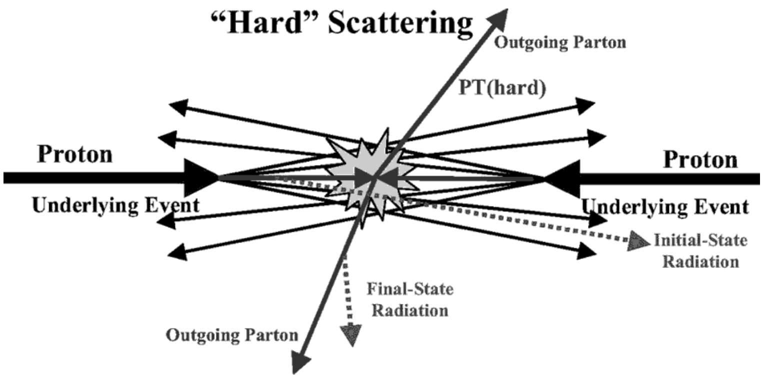


Figure 2.4: Schematic diagram of particle production in a proton-proton collision.

2.3.4 Luminosity and pile-up

The instantaneous luminosity, L, is an important parameter to describe the performance of the proton-proton collider [28], which has units of $cm^{-2}s^{-1}$. In practice, L is relates to the cross-section σ of the inelastic proton-proton process to the corresponding event rate R (number of inelastic proton-proton collisions per second):

$$R = \mathcal{L} \times \sigma \tag{2.3}$$

Since the particle beams come in bunches, it is possible to write the event rate in terms of the average number of inelastic interactions per bunch crossing (μ) and the frequency of bunch crossing:

$$\mathcal{L} = \frac{R}{\sigma} = \frac{\mu f}{\sigma} \tag{2.4}$$

The luminosity $\mathcal{L}$ depends on the beam parameters and can be written as,

$$\mathcal{L} = \frac{N_b^2 n_b f e \nu \gamma_r}{4\pi\epsilon_n \beta^*} F(\theta_c) \tag{2.5}$$

Where N_b is the number of particles per bunch, n_b is the number of bunches per beam, f the revolution frequency, γ the relativistic Lorentz factor, ϵ_n the normalized transverse beam emittance, β^*a function which give the amplitude of oscillations of the protons at the collision point and F(θ_c) a geometric luminosity factor due to the crossing angle at the interaction point (IP, the point where two bunch of protons collide with each other), which can be expressed as:

$$F(\theta_c) = (1 + (\frac{\theta_c \sigma_z}{2\sigma_*})^2)^{1/2} \tag{2.6}$$

Where θ_c is the full crossing angle at the IP, σ_z the root mean square (RMS) of the bunch length and σ_* the transverse RMS of the beam size at the IP. Figure 2.5 shows the instantaneous luminosity peak's delivered to ATLAS during pp collisions at $\sqrt{s} = 13TeV$ in 2018.

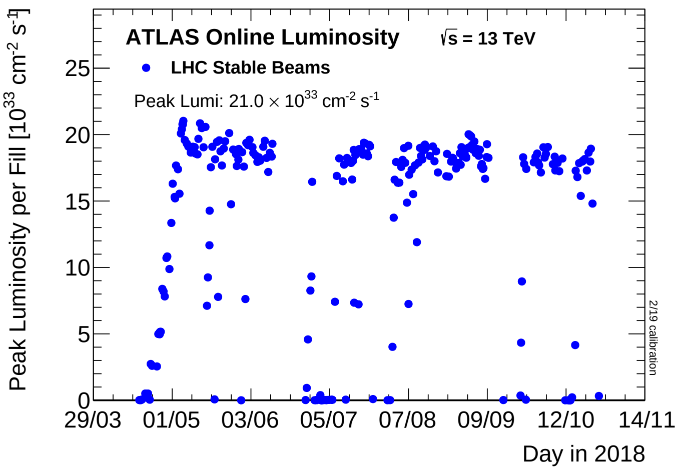


Figure 2.5: Instantaneous luminosity peaks for 2018 data taking.

Figure 2.6 shows the time evolution of integrated luminosity delivered by the LHC (green), recorded by ATLAS (yellow) and the certified runs with good quality data (blue) during pp collisions in Run 2 until the end of 2018.

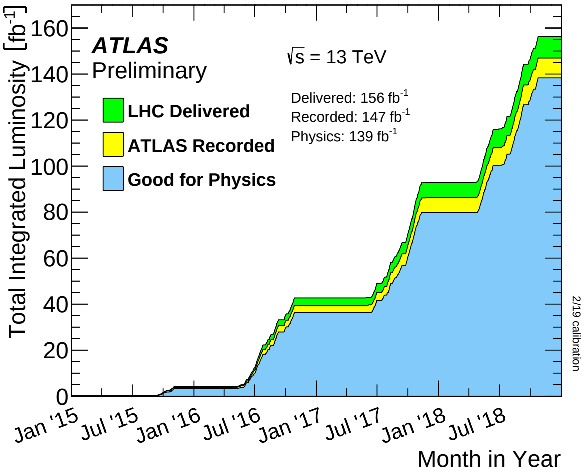


Figure 2.6: Total Integrated Luminosity and Data Quality in 2015-2018.

At the LHC, the average number of pp interactions per bunch crossing are collectively referred as "pile-up" (PU). The collision events from PU are uncorrelated with the hard-scattering process since the reconstructed physics object are affected from additional energy contributions. Multiple pp collisions within the same bunch crossing lead to in-time PU effect. The Out-of-time pile-up is caused by proton–proton interactions taking place in neighboring bunch crossings around the crossing of interest. The amount of PU operation can be calculated from the integrated luminosity over a bunch crossing time as:

$$\mu = \int^{crossin} Ldt \times \sigma_{inelastic} \tag{2.7}$$

Where $\sigma_{inelastic}$ is inelastic interaction cross section, at the center-of-mass energy of $\sqrt{s} = 13$ TeV.
Figure 2.7 shows the integrated luminosity versus the mean number of interactions per crossing averaged over all bunch crossings (denoted as $< \mu >$) for the 2015, 2016, 2017 and 2018 pp collisions data at center-of-mass energy $\sqrt{s} = 13TeV$.

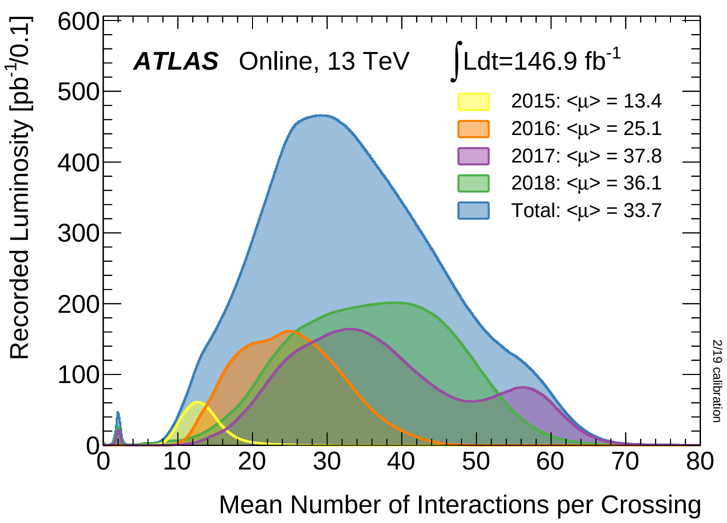


Figure 2.7: Superimposed pileup profiles of various years during $\sqrt{s} = 13TeV$ collisions of Run 2.

2.4 ATLAS detector

The ATLAS detector is a multi-functional instrument for the study of all the processes predicted by the SM, as well as to detect signals of "new physics". It is consists of four major components: the inner detector (ID), the electromagnetic (ECAL)and hadronic calorimeter (HCAL) , the muon spectrometer (MS) and the magnet systems. The ATLAS detector layout is given in Figure 2.8.
The coordinate system used in the ATLAS detector is the Cartesian coordinate system, with its origin at the detector's interaction point (IP).The polar angle θ is defined in the y-z plane and measures the angle from the beam axis, the azimuthal angle Φ is defined in the xy-plane is measured around the beam pipe. The pseudorapidity is related with the polar angle θ as follows :

$$\eta = -\ln(\frac{\theta}{2}) \tag{2.8}$$

The transverse momentum, p_T , and transverse energy, E_T , are defined in the x-y plane and are indicated by the equations:

$$p_T = P\sin(\theta), E_T = E\sin(\theta) \tag{2.9}$$

The distance ΔR in the pseudorapidity-azimuthal angle space is defined as:

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\Phi^2} \tag{2.10}$$

ATLAS offers discriminant detection power and identification between electrons, photons, muons and products from hadronic process leaving large signal called jets. Combined signals from various subdetectors allow an identification of the various particles collectively with an estimation of their energy and direction. The tracks of the charged particles are measured by the Inner Detector (ID). It is installed close to the beam pipe and is immersed in a solenoidal magnetic field. The trajectories of charged particles are bent by the magnetic field which aiding the determination of the sign of their charge and their transverse momentum and its reconstruction. The highly segmented framework also provides a good spatial resolution for reconstruction of the primary and secondary vertices. Surrounding the tracking system, the Electromagnetic and Hadronic Calorimeters, which measure the transverse energy originating from the collisions and are used for electron, photon and hadron identification. Finally, in the outermost part, the muon spectrometer measures the momentum of muons, the only charged particle that can propagate through the calorimeters.

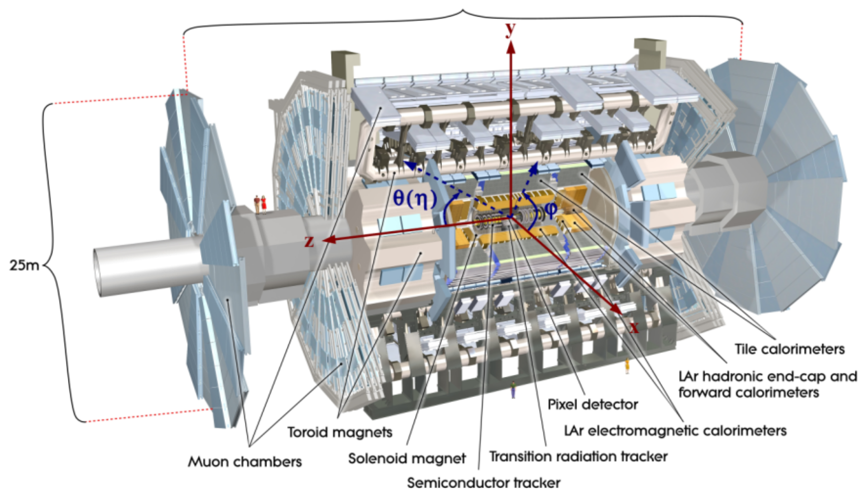


Figure 2.8: Schematic view of the ATLAS detector.

2.4.1 Inner detector

The ATLAS Inner Detector (ID), also known as inner tracking system, consisting of silicon pixels, silicon strips and straw tubes in a 2T magnetic field, the layout of the Inner Detector is shown in Figure 2.9. It is designed to measure the tracks of charged particles coming from the LHC beam-interaction region, with full coverage in ϕ and the acceptance in pseudorapidity is $\mid \eta \mid < 2.5$. The charged particles that traverse the ID ionise the detector material, allowing the reconstruction of the particle track as well as primary and secondary vertices. It also provides excellent momentum resolution in the plane perpendicular to the beam axis, of $\sigma/p_T = 0.05\% \times p_T$ and a transverse impact parameter resolution of $10\mu m$ for high momentum particles in the central η region Full details can be found in[29].

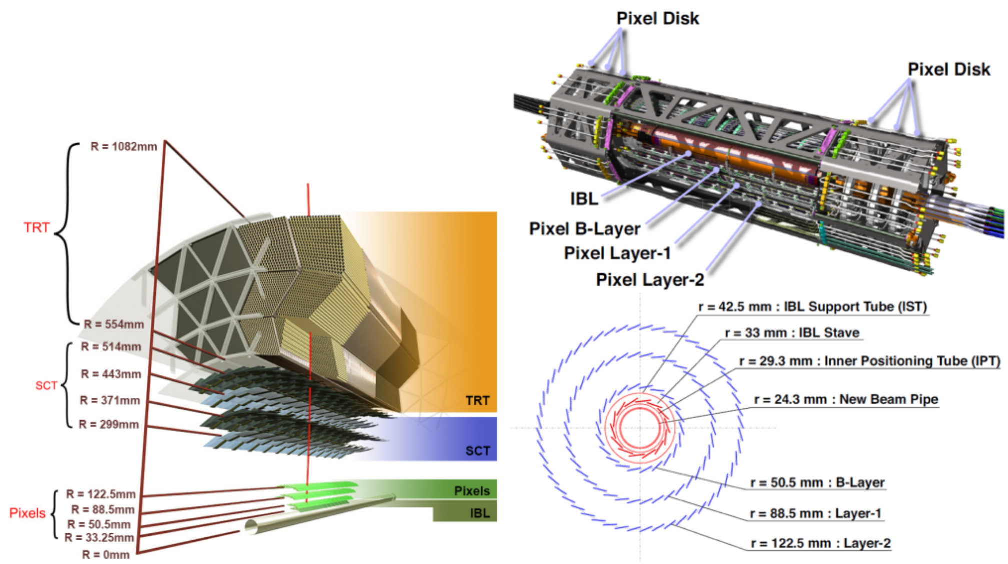


Figure 2.9: Schematic view of the ATLAS 4-Layer Pixel Detector for Run-2.

2.4.1.1 Pixel Detector

The pixel detector system [30], is the inner-most element of the ATLAS tracking system [29], It consists of 4 layers of barrel pixel detector and two end caps of three pixel disks each. The outer 3 barrel layers at radius of 4.3 cm, 10.1 cm and 13.2 cm from the center. The basic unit of the pixel detector is the module. A module is a rectangular active device approximately 6cm by 2 cm with 46080 pixels, each $50\mu m \times 400\mu m$ along the beam. All modules are arranged in 3 concentric cylinders with the axis along the beam (the barrel) at radii of 5 cm, 9 cm and 12 cm from the center, known as B-layer, layer 1, layer 2, plus 3 disks concentric with the beam at each end of barrel. There are 1456 barrel modules and 288 disk modules covering pseudo rapidity $\eta < |2.5|$.
The innermost pixel layer is a newly constructed high-resolution pixel detector, called Insert-able B-Layer (IBL). IBL is a fourth layer of pixel detectors, and has been installed in May 2014 between the existing Pixel Detector and a new smaller radius beam-pipe at a radius of 3.3 cm. to improve the vertex resolution and b-tagging efficiency. Figure 2.9 shows the 4-Layer ATLAS Pixel Detector for LHC Run-2 (left) and the radial placement of concentric pixel barrels, beam pipe and support carbon-fibre cylinders (IPT,IST) (right). The 4-Layer Pixel Detector will be crucial in the reconstruction of primary and secondary vertices which is essential for the detection of long-lived particles, e.g. containing b-quarks, and in searches for new physics at LHC.

2.4.1.2 Semiconductor Tracker

The semiconductor tracker (SCT) [29] is a silicon micro strip detector forming part of the inner tracking system. The SCT consists of 4088 modules of silicon-strip detectors arranged in four concentric barrels and two end-caps of nine disks each, as shown in figure 2.10.Each barrel or disk provides two strip measurements at a stereo angle which are combined to build space-points. The SCT typically provides eight strip measurements (four space-points) for particles originating in the beam-interaction region. The barrel modules are of a uniform design, with strips approximately parallel to the magnetic field and beam axis. Each module consists of four rectangular silicon-strip sensors with strips with a constant pitch of $80\mu m$ pitch in the barrel and of $85\mu m$ on average in the end-caps that are glued back-to-back with a stereo angle of 40 mrad to provide a two dimensional measurement (rφ and z).

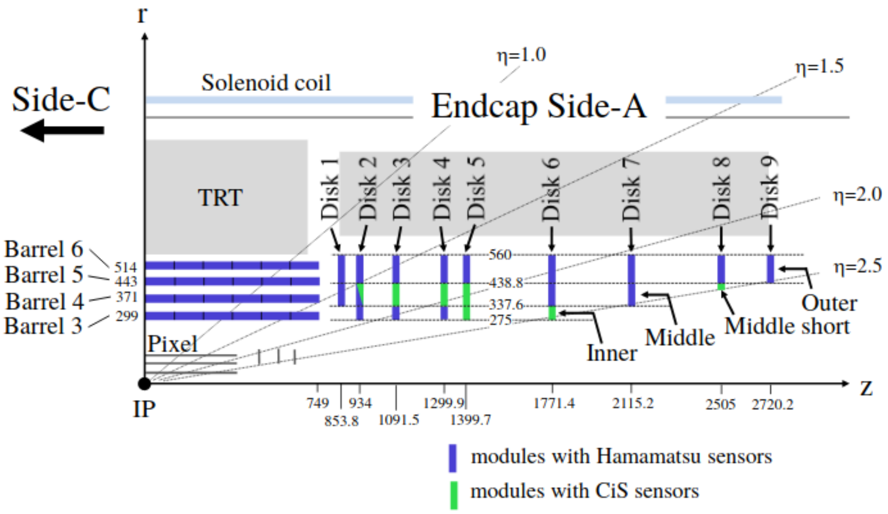


Figure 2.10: A schematic view of one quadrant of the SCT.

2.4.1.3 Transition Radiation Tracker

The transition radiation tracker (TRT) the outermost of the three tracking subsystems of the ATLAS Inner Detector, it is installed at a distance between 55-110 cm from the beam pipe and extends to a radius of 1082 mm from the interaction region. It is based on the use of straw detectors. Each straw has a diameter of 4 mm is filled with gas, and the space between the tubes is filled with a different material where transition radiation photons are produced when a particle traverses it. The resolution of the tubes is 130 μm for charged particle tracks with $|\ \eta\ | < 2.0$ and $p_T > 0.5 GeV$.

2.4.2 Calorimeter

The second layer of the ATLAS detector is occupied by the calorimeter system, shown in Figure 2.11. It is designed to absorb the interacting particles coming from the interaction point, and measuring their energy and position. The incident particle interacts with the calorimeter material through different processes depending of the particle, its kinematics and the material. The calorimeters in ATLAS provide a larger coverage in the range $|\ \eta\ | < 4.9$, compared to the ID, an important feature for an optimal reconstruction of the missing transverse energy. Calorimeters in ATLAS consist of two main components: the electromagnetic calorimeter (ECAL), used to detect particles interacting via the electromagnetic force, such as electrons and photons, and the hadronic calorimeter (HCAL), used to detect particles interacting via the strong force, such as quarks and gluons. In principle, all SM particles, with the exception of muons and neutrinos, are contained within the calorimeters in order to limit punch-through into the muon system and to achieve a better energy measurement from the particles.

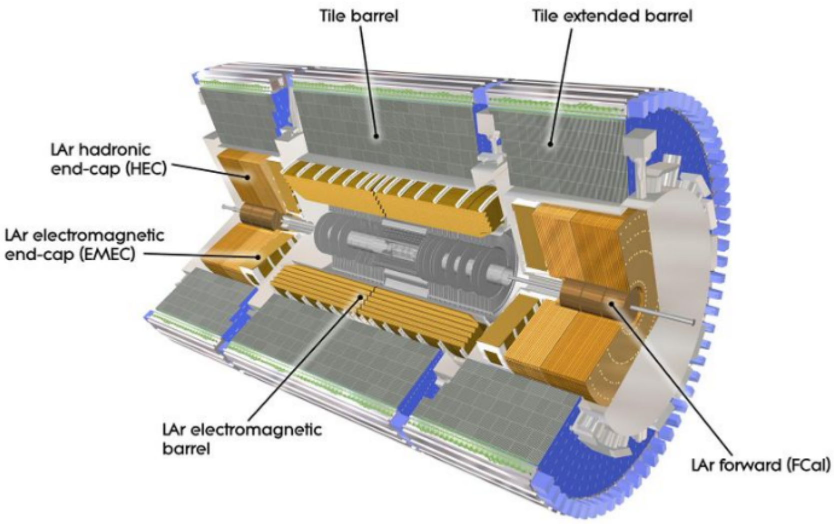


Figure 2.11: Schematic description of the calorimeter system of the ATLAS detector.

2.4.2.1 Electromagnetic Calorimeter

The ATLAS electromagnetic calorimeter [31, 32] is used to detect both photons and electrons. It is consists of layers of lead as absorbed and liquid argon as active material, which are used to initiate particle showers and measure energy deposits, respectively. The LAr measures energy deposits by collecting the charge produced by the particles passing through and ionizing the LAr. The ECAL is divided into a barrel region, covering a pseudo rapidity $\mid \eta \mid< 1.475$ and two end-caps covering $1.375 >\mid \eta \mid< 3.2$. The barrel calorimeter consists of two identical half-barrels, separated by a small gap 4 mm at z= 0. Each end-cap region of the ECAL consists of an inner wheel: an outer wheel covering the region $1.375 >\mid \eta \mid< 2.5$, and an inner wheel covering the region $2.5 >\mid \eta \mid< 3.2$. Figure 2.12 depicts the granularity of these layers in $\Delta\eta \times \Delta\phi$.
The calorimeter is segmented into three longitudinal layers, The first layer called the strips layer, is finely segmented in the η direction in order to distinguish single photons from $\pi^0 \longrightarrow \gamma\gamma$ decays. The cell size in η is for instance 0.003125 in the barrel region. The second layer collects most of the shower energy is deposited, has a granularity of $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$. The third layer of the ECAL is used to capture the tails of the electromagnetic shower,and has a segmentation of $\Delta\eta \times \Delta\phi = 0.05 \times 0.0245$. Such fine segmentation allows precise measurement of electrons and photons with an energy resolution of:

$$\frac{\sigma_E}{E} = \frac{10.1\%}{\sqrt{E}} \oplus 0.17\% \tag{2.11}$$

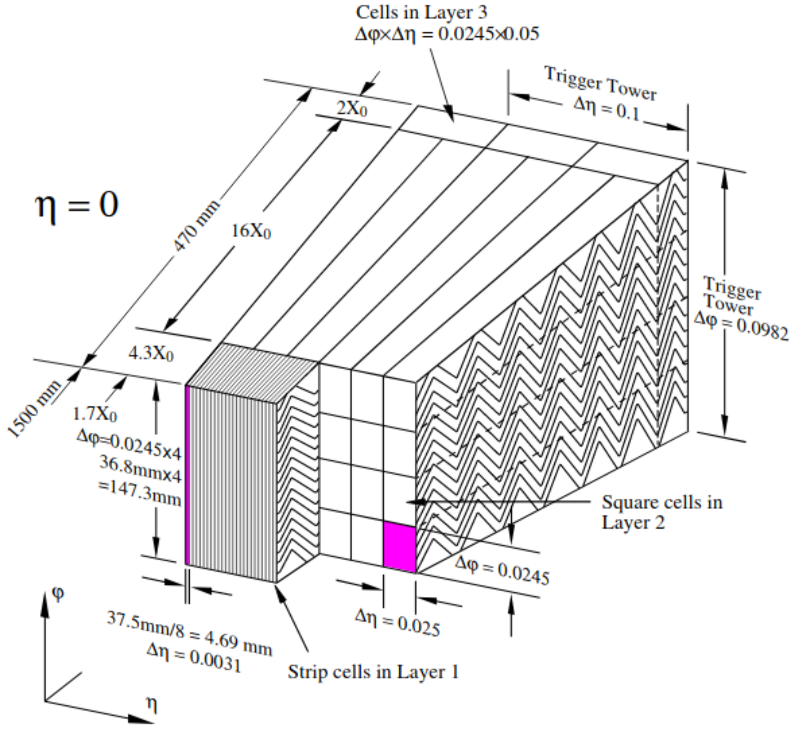


Figure 2.12: Module of the ATLAS electromagnetic calorimeter.

2.4.2.2 Hadronic calorimeters

The hadronic calorimeter is used to detect strongly interacting particles such as hadrons. The ATLAS hadronic calorimeter consists of three main components: the tile calorimeter [33], the LAr hadronic end-cap calorimeter (HEC) [34] and the LAr forward calorimeter (FCal) [35]. The tile calorimeter is a sampling calorimeter using steel as the absorber and scintillating tiles as the active material, located outside the EM calorimeter envelope with a central barrel and two extended barrels, in the region of $|\eta| < 1.0$. The hadronic end-cap calorimeter is a sampling calorimeter uses LAr as its active material and copper as its passive material, and covers $1.5 < |\eta| < 3.2$. Finally, the forward calorimeter uses LAr as its active material, provides full coverage over the range $3.1 < |\eta| < 4.9$.
The energy resolution for Tile and HEC is :

$$\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\% \tag{2.12}$$

Whereas, for FCal it is:

$$\frac{\sigma_E}{E} = \frac{100\%}{\sqrt{E}} \oplus 10\% \tag{2.13}$$

2.4.3 Muon Spectrometer

The Muon Spectrometer (MS), shown in Figure 2.13, is the largest subsystem of the ATLAS detector. It designed to detect tracks over a large region of pseudo-rapidity $|\eta| < 2.7$, is made of a large toroidal magnetic field that allows for transverse momentum (p_T) measurements by measuring the curvature of the muon tracks magnetically deflected in the MS. The two major regions in the MS are the barrel region covering $|\eta| < 1.1$ and the two end-cap regions covering $1.1 < |\eta| < 2.7$. The MS consists of four detectors instrumented with separate trigger and high-precision tracking chambers. Monitoring Drift Tubes (MDT) in the barrel and end-cap sections and Cathode Strip Chambers (CSCs) are used as precision chambers, whereas, Resistive Plate Chambers (RPC) in the Barrel Region and Thin Gap Chambers (TGC) in the End-cap Regions are used as trigger chambers.

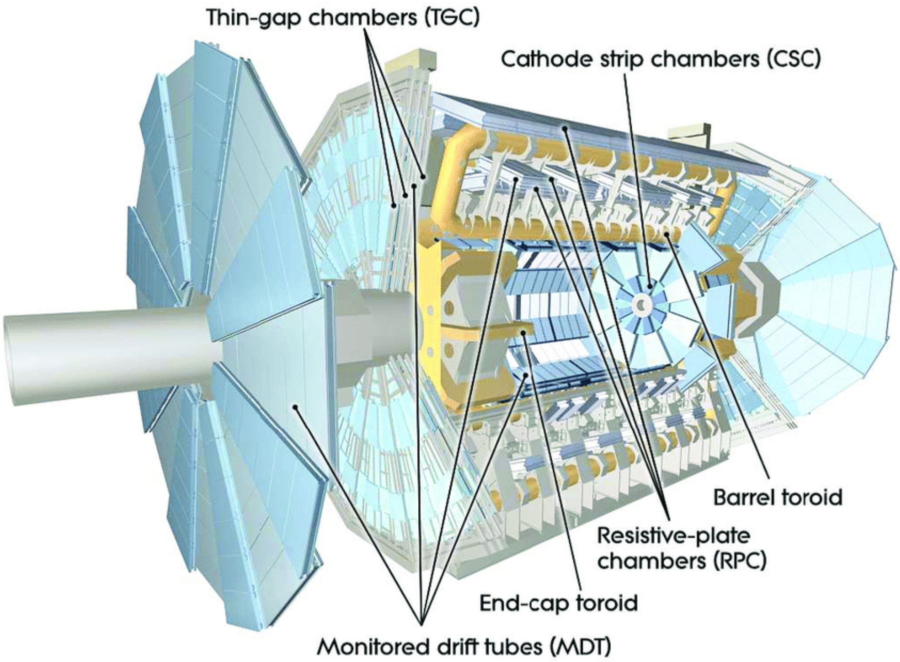


Figure 2.13: The cut-away view of the ATLAS Muon Spectrometer.

- **Monitored Drift Tubes (MDTs)** provides the precision measurement of the track coordinates in the principal bending direction of the magnetic field. In the barrel region ($|\eta| < 1.3$), the MDTs are positioned in three concentric layers around the beam axis, at an approximate radius of 5, 8 and 10 m. To avoid holes in the acceptance, the chambers are partly overlapping. In the end-cap region $1.1 < |\eta| < 2.7$, MDT chambers are assembled onto three wheels, inner, middle and outer layers positioned at z= 7.5, 14 and 22.5 m respectively.

- **Cathode Strip Chambers (CSCs)** are multi-wire proportional chambers with anode wires distributed radially and cathode strips in the $\eta - \phi$ plan One side of the CSC has strips segmented perpendicularly to the wires, allowing for a measurement of the η coordinate, while the other side has its strips segmented parallel to the wires, allowing for a measurement of the ϕ coordinate. Only the cathode strips of the CSCs are read out. The main purpose of the CSCs is to provide precision tracking in the innermost end-cap regions where particle rates are beyond the MDT read out limits they coverthe range $2.0 < |\eta| < 2.7$.

- **Resistive Plate Chambers RPCs** are used for triggering on muon tracks contained within the barrel $|\eta| < 1.05$, of the MS. There are three concentric layers of RPCs around the beam axis: two layers in the middle station are assembled together with the middle MDT chambers, and one layer in the outer station s assembled on the outer MDT chambers. Due to the large lever arm between inner and outer RPCs, the trigger is able to select high momentum muons with thresholds ranging from 9 to 35 GeV. The inner RPCs deliver the low momentum trigger with thresholds from 6 to 9 GeV.

- **Thin Gap Chambers TGPs** provide two functions in the end-cap MS. They cover the endcap regions $1.05 <\mid \eta \mid< 2.7$, with trigger coverage $\mid \eta \mid< 2.7$. TGPs are used for their triggering capabilities (with good time resolution and high rate), and provide a muon track measurement in the ϕ direction. The design muon momentum resolution of the MS is $\frac{\sigma_{p_T}}{p_T[GeV]} = 10\%$ at $p_T = 1GeV$

2.5 Trigger, acquisition and data taking

2.5.1 Trigger

It is impossible to record all collision events in ATLAS at the 40MHz collision rate provided by the LHC. The ATLAS trigger and data acquisition system reduces this input down to a rate of around 1kHz of events, which are recorded and kept for subsequent physics analyses . Figure 2.14 shows a diagram that illustrates the components of the Trigger and Data acquisition system (TDAQ). For Run 2, the ATLAS trigger system is composed on two component : the hardware-based Level-1 (L1) trigger and the software-based higher-level trigger (HLT). Events accepted at the fast initial L1 step are passed on to the HLT, which runs more precise algorithms and makes a final trigger decision.

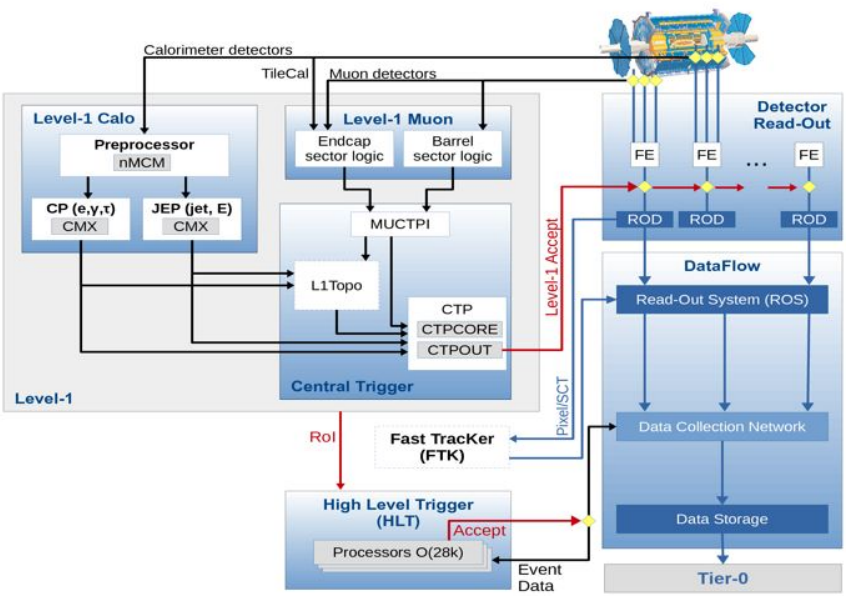


Figure 2.14: The ATLAS Trigger system.

2.5.1.1 Level-1 Trigger

The Level-1 trigger performs a hardware-based online event selection using information of reduced granularity from all the calorimeters and the muon trigger system (RPCs and TGCs) and determine coarse Regions-of-Interest (RoIs) are defined in the detector, and to select events with high-E_T objects including charged leptons, photons, jets, and τ leptons decaying into hadrons, as well as events with large missing transverse momentum (E_T^{miss}). The L1 trigger reduces the event rate to around 100kHz, with a 2.5 μm maximum latency. The Events passing this trigger step are buffered in the Read-Out System (ROS) for further processing in the HLT.

2.5.1.2 High-Level Trigger

After the incoming event rate has been reduced by the L1 trigger system, the software-based HLT system can run more precise and computationally expensive reconstruction algorithms to refine the trigger decision. These reconstruction algorithms generally are run in the RoIs defined at the L1 step. The HLT can make use of the full calorimeter granularity, tracking and muon spectrometer information for the decision and reduces the event rate further down to 2 kHz. Decisions in the HLT are made within 200 ms and events passing this trigger are stored for offline analysis using the data acquisition (DAQ) system. If the events passing the HLT criteria are transferred to local storage at the ATLAS site, and then sent to CERN's computing center for storage and subsequent processing. Aspects of these processes are elaborated on in Appendix A.

2.5.2 Data quality

The ATLAS detector reads out particle collision data fromover 100 million electronic channels at a rate of approximately 100 kHz, with a recordingrate for physics events of approximately 1 kHz. it is essential to monitor the status of the ATLAS hardware and determine the quality of the data being taken in an efficient manner. the trigger data quality is provided in flags reflecting the behavior of the different trigger signatures derived from the online and offline monitoring informations. In the online environment, this information can flag the shifter to take action to prevent taking faulty data. In the offline environment, one can perform more complex checks of the data to determine the quality of the data for various physics groups. A Data Quality (DQ) status flags are determined by the DQ representatives from each sub-detector system, which are used by the combined performance and trigger groups to declare the data as good/flawed/bad for the different physics objects. These flags are set for each luminosity block (approximately 2 minutes long fraction of runs), where the most low-level flags are based on detector control conditions, flagging possible hardware and data-taking problems such as nominal voltages, temperature, humidity, etc. Finally, a "Good Run List" (GRL) are built using all the luminosity blocks for which the sub-detectors useful for a given analysis were operational.
The relative fraction of luminosity associated to data of good quality delivered by the various sub-detector of ATLAS during the 2015 and 2018 data-taking is shown in Figure 2.15. The total fraction of luminosity which is good for physics analysis correspond to 95.6% of the total luminosity delivered by the LHC.

ATLAS pp Run-2: July 2015 – October 2018										
Inner Tracker			Calorimeters		Muon Spectrometer				Magnets	
Pixel	SCT	TRT	LAr	Tile	MDT	RPC	CSC	TGC	Solenoid	Toroid
99.5	99.9	99.7	99.6	99.7	99.8	99.6	100	100	99.8	98.8
Good for physics: 95.6% (139 fb^{-1})										

Luminosity weighted relative detector uptime and good data quality efficiencies (in %) during stable beam in pp collision physics runs with 25 ns bunch-spacing at $\sqrt{s}$=13 TeV for the full Run-2 period (between July 2015 – October 2018), corresponding to a delivered integrated luminosity of 153 fb^{-1} and a recorded integrated luminosity of 146 fb^{-1}. Runs with specialized physics goals are not included. Dedicated luminosity calibration activities during LHC fills used 0.6% of recorded data in 2018 and are included in the inefficiency. Trigger-specific data quality problems (0.4% inefficiency at Level-1) are included in the overall inefficiency. When the stable beam flag is raised, the tracking detectors undergo a so-called "warm start", which includes a ramp of the high-voltage and turning on the pre-amplifiers for the Pixel system. The inefficiency due to this, as well as the DAQ inefficiency, are not included in the table above, but accounted for in the ATLAS data taking efficiency.

Figure 2.15: Performance of the ATLAS detector for the 2015 and 2018 data-taking. Runs with a bunch-spacing of 25 ns at $\sqrt{s} = 13TeV$ for the full Run 2 period between April 2015 and October 2018, corresponding to a delivered integrated luminosity of 139 fb^{-1}.

2.6 Summary

In this chapter we try to describe the full LHC machine and the different technologies for each ATLAS sub-detectors as well as the whole setup needed to better understand the collection of LHC data which are ready to be used for physics analysis, like we use in the coming chapters. The data are produced by the p-p collisions at sqrt(s)=13 TeV, the integrated luminosity recorded by the ATLAS detector during Run2 of the LHC from 2015-2018. After typical data-quality selections, the full Run2 pp data samples correspond to an integrated luminosity of 139 fb^{-1}.

Chapter 3

Physics Object reconstruction

3.1 Introduction

The particles produced in the pp collisions in the center of the ATLAS detector interact with the detector material as discussed in Chapter 2. The information from the various ATLAS subdetectors are recorded of each event in terms of energy deposits in calorimeter cells and hits in the trackers. In order to convert this information into physics observables. An overview for the reconstruction and identification of the different objects used in physics analyses carried out using the ATLAS experiment will be detailed. In particular, Section 3.2 introduces the fundamental objects for the physics observables reconstruction: charged tracks, vertices in the detector; Section 3.3 describes the hadronic jets; the succeeding Sections 3.4 describe the reconstruction of electrons, photons, muons, and tau-leptons, respectively. Finally, Section 3.9 focus on the missing transverse momentum, which plays an important role in the analysis described in Chapters 4.

3.2 Charged tracks and primary vertices

Charged tracks

The trajectories of charged particles relies on the information collected by the subsystems of the ID. These tracks leave a set of hits in these subsystems along their path. These hits are used to reconstruct the tracks and fit their parameters. The track-fitting procedure yields the set of track parameters which minimizes the χ^2. This is done using a Kalman filter, to find the track that has the highest likelihood for the observed hits.
The trajectories of tracks are parameterized in the ATLAS software with the following parameters:

$$M = (d_0, z_0, \phi_0, \theta, q/p) \tag{3.1}$$

Where d_0 and z_0 are respectively the transverse and longitudinal impact parameters, measured with respect to the ATLAS nominal interaction point. The angle ϕ_0 is the azimuthal angle of the track at the perigee (the distance of closest approach to the nominal interaction point), and θ is its polar angle. The ratio q/p is the inverse

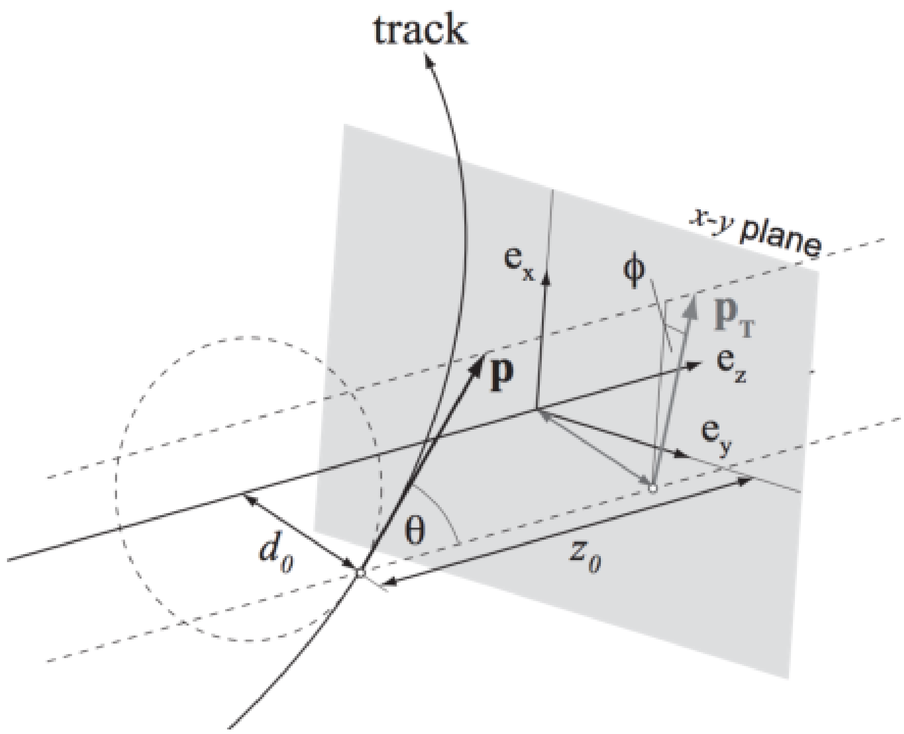


Figure 3.1: Schematic view of the track perigee parameters.

of the particle momentum (p) multiplied by its charge (q). Finally, the transverse momentum, p_T , is computed as $p_T = p \sin\theta$. The track perigee parameters are schematically shown in Figure 3.1

Interaction Vertices

After all ID tracks are identified, a vertex-finding algorithm based on the reconstructed tracks is used in order to find the vertices of each event. The primary vertex (PV) in each event is the vertex with the largest value of $\sum(P_T)^2$, where the scalar sum is taken over all the tracks matched to the vertex. The PV corresponds to the vertex of the hard scatter interaction is very high efficiency. Two sets of quality criteria are studied in this report, based on transverse momentum (p_T), pseudorapidity (η) and the number of hits in the subsystems.

3.3 Hadronic Jets

Due to the color confinement in the strong interaction, the quarks and gluons (called parton) are not observed in isolation and they hadronize. As a consequence of the hadronization, a shower of particles as pions, kaons, protons, neutrons, light leptons from decays, and photons are observed from the hadron decay chain which is limited to a specific solid angle. These hadrons are emitted in a narrow, collimated cone around the direction of the original parton and can be grouped into single objects referred to as jet.

A jet is defined with an algorithm that groups its constituents into jets and computes the momentum of the resulting object. The constituents of the jets can be clusters of topologically connected calorimeter cell signals, "topo-clusters", or tracks in the ID. Various jet collections can be formed targeting different topologies and kinematics. In this work, jets are built using the anti-kt algorithm [36], briefly summarized in Section 3.3.1, and the main differences among them are the radius parameter and the inputs of the algorithm: topo-clusters, ID tracks, or Monte Carlo (MC) generator-level (truth-level) jets. This section presents, the calorimeter jets which are reconstructed from topological calorimeter clusters in Section 3.3.2.

3.3.1 Jet reconstruction: the anti-kt algorithm

The anti-k_T algorithm [36] is a sequential clustering algorithm which is extensively used to reconstruct jets at large hadron colliders, It uses the combination of topoclusters, reconstruction tracks (track jets), or MC events at the particle level (truth jets). The anti-k_t algorithm compute first the all distance between every possible particle pair using equation 3.3.1 and between each particle and beam with equation 3.3.1 and take minimum distance between d_ij and d_iB. If the smaller distance is d_ij, the particle i and j are combine, while if the smallest distance is d_iB, the algorithm calls i a jet and removes it from the list of particle, Then the distance parameters d_{ij} and d_{iB} are recalculated and the procedure repeats iteratively, until there are no particles left in the list.

$$d_{ij} = \min \frac{1}{K_{Ti}^2}, \frac{1}{K_{Tj}^2} \frac{\Delta R_{ij}^2}{R^2} \tag{3.2}$$

$$d_{iB} = \frac{1}{K_{Ti}^2}, \tag{3.3}$$

Where $\Delta R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$ the distance in the $(\eta - \phi)$ plane between particle i and j, and the parameter R defines the radius of cone size for the jets. The analysis will use two kinds of distance parameters R, jets representing quarks and gluons called small-R jets using R = 0.4, while jets representing hadronically decaying massive particles large-R jets, and are reconstructed with R = 1.0.
There are other sequential clustering jet algorithms such as the k_t described in [37] and the Cambridge/Aachen described in [38], however the anti-k_t algorithm used in this analysis has been chosen because it is infared and collinear safe meaning that the configuration of jet does not change when adding a further soft particle, and when substituting one particle with two collinear particles [36]. To perform the jet clustering for these algorithms, the FastJet [39] program is used.

3.3.2 Clustering of calorimeter energy deposits

The calorimeter jets reconstruction is a collection of three dimensional, positive energy topological clusters called "topoclusters", in the ATLAS calorimeters are used for identifying and reconstructing jets. These topoclusters are build from neighbouring calorimeter cells containing a significant energy above a noise threshold. The expected noise takes into account electronic noise from the detector as well as pile-up noise, calculated from simulations. The basic observable controlling this cluster formation is the cell signal significance, given by the following formula:

$$\zeta = \frac{E^{EM}_{noise,Cell}}{\sigma^{EM}_{Cell}} \tag{3.4}$$

Where $E_{noise,Cell}$ is the cell signal and σ^{EM}_{Cell} the average expected noise are measured on the electromagnetic EM energy scale [40].
A three-stage classification is used to build the topo-clusters known as the ATLAS S-N-P with a set of default values 4-2-0 topo-cluster threshold scheme, starting from a calorimeter cell with a highly significant seed signal $\zeta > S$, the corresponding cluster is called "proto-cluster". Then, the cells neighbouring the seed and satisfying $\zeta > N$ are merged into the corresponding cluster. The cluster grows until all direct neighbor cells on the outer perimeter, $\zeta > P$, are included. These topo-clusters can grow and merge together, therefore an additional splitting procedure is performed in order to separate nearby showers. All cells having a local energy maximum above a threshold are designated as seeds and then the clusters are split spatially between them [41].
The final clusters are used as inputs to the anti-kt algorithm described in Section 3.3.1 to provide reconstructed jets. Two categories can be defined depending on the distance parameter used for the anti-kt algorithm: small-R and large-R jets.

3.3.3 Small-R Jets

Small-R jets are reconstructed at the electromagnetic energy scale(EM scale) with the anti-kt algorithm. The radius parameter R = 0.4 in Equation 3.3.1 as implemented in FastJET software package, and which is the frequently used value in ATLAS for the reconstruction of the gluon, light quark, and b-tagged jets. This jets are reconstructed for $p_T > 7GeV$ and $\eta < 4.5$.

3.3.4 Jet calibration

Topoclusters are reconstructed at the EM scale, which correctly estimates the energy in the calorimeter deposited by particles produced in an electromagnetic shower. but does not reflect well the energy for hadronic showers, in particular a large fraction of the energy of the hadronic shower is partly dissipated via nuclear processes, which is invisible to calorimeters and thus the reconstructed jet momentum is not the same as the truth particle jets. This effect is known as calorimeter non compensation. Then these cluster calibrated with the Local Cluster Weighting (LCW) to correctly measure the energy deposited by particles produced in a hadronic shower as well as for further energy losses due to noise thresholds and losses in dead material. The LCW calibration scheme first classifies topoclusters as either electromagnetic or hadronic based on the measured energy density and the longitudinal shower depth. The calibration of the reconstructed jets is typically referred to as the Jet Energy Scale (JES), which undergoes several corrections to reduce the impact of pile-up contamination and recover the energy of the truth particle jets on average. This procedure is shown schematically in Figure 3.2, and the steps are described in the following:

1. **Origin correction**: This correction is applied to change the jet direction in order to make the jet point to the primary event vertex rather than to the

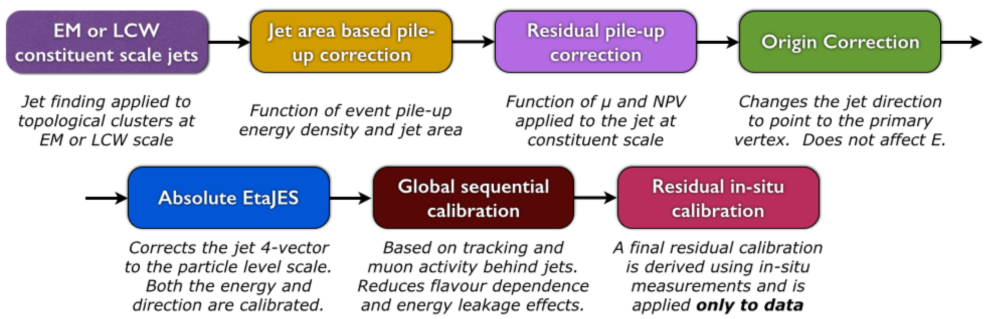
cented of the ATLAS detector. This is obtain by recalculates the four-momenta of the constituent topological clusters using the η and ϕ while keeping the jet energy constant

2. **Pileup correction**: is applied to correct the jet energy and momentum due to the presence of additional pile-up activity that have been absorbed into the jet. It consists of two components: a first correction, dependent on the number of PVs, uses the subtracting the pile-up p_T density integrated over the jet area, A second correction based on the number of PVs and on the number of interactions per bunch crossing is residual correction derived from the MC simulation.

3. **Absolute calibration**: The absolute jet energy scale (JES) restores the energy of the reconstructed jet to that of the corresponding MC particle-level jets. This jet energy scale (JES) factor is calculated from MC as a function of η and p_T.

4. **Global sequential calibration**: improves the JES resolution against fluctuations in the jet particle composition and reduce the flavour dependence of the JES. They are, however, corrects sequentially the dependence of the jet response considering observables related to the fraction of energy deposited in the various layers of the calorimeter and tracks associated with the jet as described in

5. **In-situ calibration**: As the final stage, the data-driven correction is applied to corrects for the differences between data and MC simulation. These are derived from events where the transverse momentum of a jet is balanced against other well-measured objects, such as the Z boson in Z + jets events), γ in γ + jets events and multijet events.



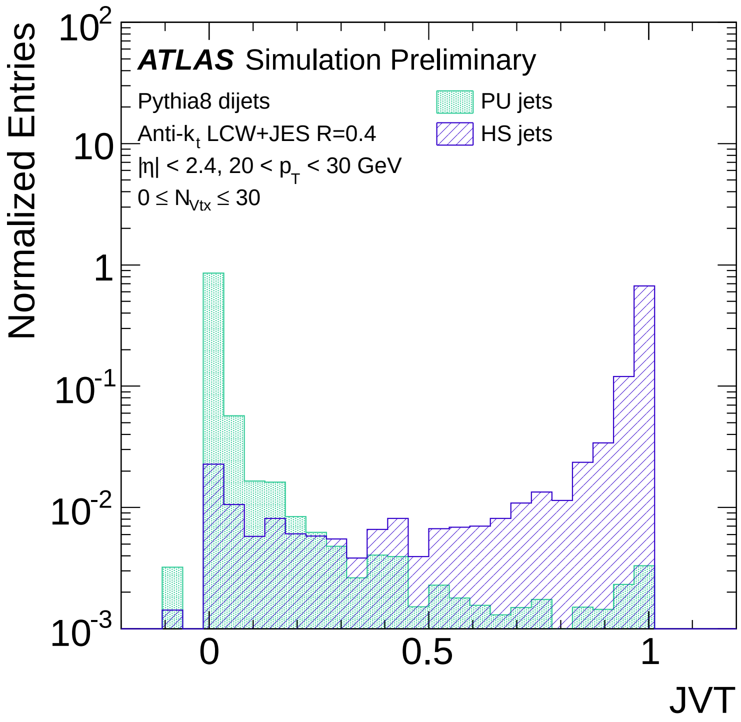
[h]

Figure 3.2: The ATLAS calibration chain used for standard EMTopo jets.

3.3.5 Pile-up suppression via JVT

Multiple collisions of proton bunches in LHC result in Hard-Scatter (HS) interaction and an addition collision called the pile-up (PU) collision, is known as another source of the soft QCD interactions. In addition, pile-up jets may occur from local fluctuations, which are built by a superposition of a random combination of particles from multiple PU vertices. There are two dominant components of PU background, named as "In-time pile-up" and "Out-of-time pile-up" which are discussed in the

chapter 2. Information from the tracks plays a major role in rejecting fake jets originated from pile-up fluctuations, therefore giving information about the pile-up structure and vertex composition of jets within the tracking detector acceptance $|\eta|<2.5$ that can be used for discrimination of pile-up jets.
The Jet Vertex Tagger (JVT) observable is constructed from tracking information, in order to match jets to their primary vertex, and thus determine which came from the vertex pileup. JVT is a discriminating variable calculated from likelihood of two quantities, information about the primary vertex and tracks associated with the jet, which has a good separation power with respect to pile-up jets, and combine them in a 2-dimensional likelihood.
The JVT distribution for hard-scatter and pileup jets with $20 < p_T < 30$ GeV is shown in figure 3.3. For $JVT = -0.1$ is assigned to jets with no associated tracks. For values of the JVT close to zero, the reconstructed jet is likely to come from pile-up interactions, whereas for a JVT value close to one, the jet is more likely to be a hard scatter jet. A recommended JVT score requirement (JVT $>$ 0.59) is applied for jets with $20 < p_T < 60 GeV$ and $|\eta| < 2.4$ is implemented for small-R jets.



[h]

Figure 3.3: Distribution of JVT for pileup and hard-scatter jets with $20 < p_T < 30$ GeV

.

3.3.6 Large-R Jets

Large-R jets are reconstructed by the anti-k_T algorithm (see section 3.3.1) with R = 1.0 using calorimeter clusters calibrated at the hadronic scale using the LCW scheme.

to remove the energy of cluster associated to soft particles originating from PU interactions, a trimming procedure is applied [42]. this procedure reclustering the constituents of the initial jet, using K_T algorithm into subjets with radius parameter R_{sub} and then only the subjets that have a transverse momentum greater than a fraction f_{cut} of the parent jet momentum are used to make final trimmed jet. The values used for the two parameters in ATLAS are: $R_{sub} = 0.2$ and $f_{cut} = 0.05$.

[h]

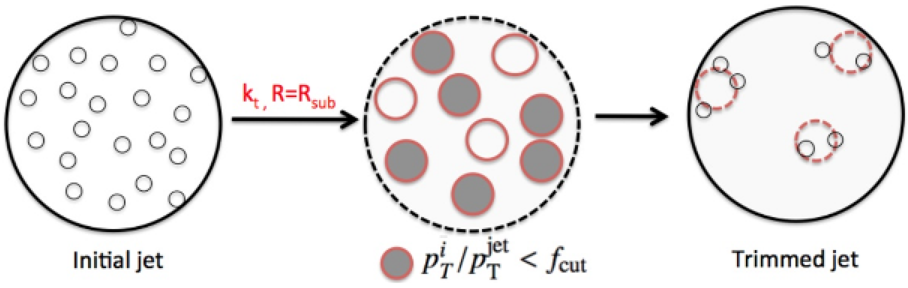


Figure 3.4: Diagram depicting the jet trimming procedure [42]

.

3.3.7 Top tagging

The top tagging algorithm is used to identify hadronically decaying top quarks reconstructed as a single large-R jet with $p_T > 300GeV$ and high efficiency. The top tagging algorithm uses two variables: the calibrated jet mass and subjettiness ratio τ_{32}.

- Jet Mass Calibration:

The calorimeter large-R jet mass is defined as the norm of the four-momentum sum of the jet constituents, which are calibrated calorimeter clusters with energy E_i and momentum p_i, as in Equation 3.3.7:

$$m^{calo} = \sqrt{(\sum_{i \in J} E_i)^2 - (\sum_{i \in J} p_i)^2} \tag{3.5}$$

The reconstructed jet mass is calibrated to the particle level. This correction is calculated similarly to the correction of the jet energy scale.

- N-subjettiness:

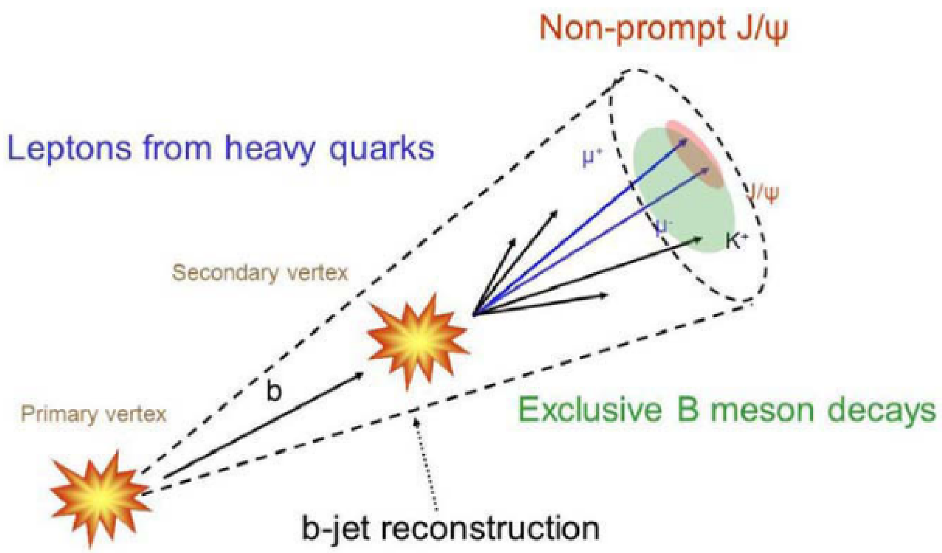
N-subjettiness is an effective discriminating variable for tagging boosted objects it denoted by τ_N. τ_N effectively counts the number of subjets in a given jet. an algorithm computes the N-subjettiness variable τ_N defined by:

$$\tau_N = \frac{1}{N_0} \sum_k p_{T,k}.min(\Delta R_{1,k}, \Delta R_{2,k} .... \Delta R_{N,k}) \tag{3.6}$$

where N_0 is a normalisation factor, $p_{T,k}$ is the transverse momentum of the constituent particle k, $\Delta R_{i,k}$ is the angular distance between a subject i and a constituent particle k. the N-subjettiness ratio $\tau_{21} = \tau_2/\tau_1$ is an effective discriminating variable to identify two-prong objects like boosted W, Z, and Higgs bosons and $\tau_{32} = \tau_3/\tau_2$ is effective for three-prong objects like boosted top quarks [43].

3.3.8 b-tagging algorithm

Jets originating from the hadronization of a b-quark called b-jets. The relatively long lifetime of B-hadrons (about 10^{-12}s), implies that it travels several millimeters before decaying in other particles. In particular, a secondary vertex can be produced where the b hadron decayed and tracks with large d_0. This procedure of identification of such jets is referred to as b-tagging. A schematic view of the topology originating from a jet containing a B-hadron is shown in Figure 3.5



[h]

Figure 3.5: DB-hadrons inside the jets originated from heavy quarks

.

There are three b-tagging algorithms that can be combined through multivariate discriminant. The most relevant are:

- **Impact Parameter based algorithm (IP3D)**: algorithm uses a likelihood ratio discriminant in which input variables related to the impact parameter significance (d_0/σ_{d_0} and $z_0 sin(\theta)/\sigma_{z_0 sin(\theta)}$) of all the tracks in a jet are compared to simulated distributions for b-, c- and light-flavour jet hypotheses.

- **Secondary Vertex Finding (SV1)**: The SV1 depends on the reconstruction of a secondary vertex to identify a b-jet. It uses a likelihood ratio discriminant, that based on the combination of the various variables, such as invariant mass of all tracks associated to the secondary vertex, the ratio between the energy addition of tracks associated to a secondary vertex and tracks from the primary vertex, and the number of two-track vertices.

- **Identification of the Decay Chain (JetFitter)**: This algorithm attempts to reconstruct the topological structure of weak b- and c-hadron decays inside the jet. It tries to find a common line between the PV and the vertices of the b- and c- hadron decays, which corresponds to the b-hadron flight direction [44].

The performance of these b-tagging algorithms are combined in a Boosted Decision Tree (BDT) multivariate analysis. The analysis presented in Chapter 5 performs the b-tagging using charged particle tracks, referred as track jets with radius parameter R fixed to R = 0.2, that improves the performance in boosted regimes.

3.3.9 Overlap removal

Ambiguities within the object identification that arise during reconstruction, i.e. when a reconstructed object matches multiple object hypotheses, are resolved in several steps that give different priority to the thing types. First, electron candidates are removed if they share a track with another electron candidate that features a higher p_T and if they share a track with a muon candidate. Next τ are excluded if they're within $\Delta R = 0.2$ of an electron or muon candidate. For the next step, f an electron and a jet are within a distance of $\Delta R = 0.2$, the electronis selected and the jet is discarded. Muon and electrons are then rejected if they are within $\Delta R = min(0.4, 0.04 + 10GeV/p^{lepton}\ )$ of a small-R jet. Small-R jets are further removed if they are within $\Delta R = 0.2$ of any surviving muon if the jets have fewer than three tracks or if the muon p_T is bigger than half the jet p_T and greater than 70% of the p_T sum of the tracks associated to the jet. Finally, large-R jets are rejected if they are separated from an electron (muon) by less than $\Delta R = 0.1$.

3.4 Objects Physics

3.4.1 Electron

Electrons reconstruction consists of the energy cluster in the electromagnetic calorimeters and matching them to well reconstructed tracks, which originate from a region near to the identified primary vertex. The clusters are formed from the sum of layers of electromagnetic calorimeter into what is referred to as energy towers, with each tower encompassing a region in $\eta - \phi$ space with a granularity of $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$. A sliding window algorithm searches for clusters using a window size of 3×5 towers, where the towers included have energy deposit greater than 2.5 GeV for the purposes of cluster reconstruction. Then the track candidates for the reconstruction of electrons must be situated within the main tracking region of the detector, satisfying $|\eta| < 2.47$. In addition, any electron candidate found in the region of $1.37 < |\eta| < 1.52$, which corresponds to the transition of the EM calorimeter barrel to the end-cap are excluded.

Electron Identification algorithms are applied on the electron candidates in order to ensure high quality electrons for analysis and suppress the mis-identification of other particles (non-prompt electron backgrounds and non-prompt electrons from heavy flavor decays). The baseline identification algorithm is performed using a likelihood (LH) based method. It is a multivariate discriminant that exploiting track features and properties of the clusters. This method provides a good separation between signal electrons (isolated) and fakes electrons. The identification criteria for electron candidates is based on sequential cuts, They comprise the parameters related to the shower shape, as well as track quality (including all information from the TRT) and track cluster matching discriminating variables. Three levels of identification operating points are typically provided for electron ID in order of increasing background rejection: Loose, Medium and Tight. The operating points are defined such that the set of electron candidates with higher background rejection is a subset of the electron candidates with lower background rejection (Tight $\subset$ Medium $\subset$ Loose). Figure 3.6 shows the performance of the LH identification algorithm, for which the efficiencies for electron candidate identification at the three operating points increases with E_T. The efficiencies in MC simulations are compared to the one measured in data with tag-and-probe methods using $Z \to ee$ and $J/\psi \to ee$ decays and the extracted data to MC ratios, as a function of E_T and η, are applied as scale factors at analysis level. For photons, the MC identification efficiencies are cross-checked in data with $Z \to e\gamma$ radiative decays at low momentum.

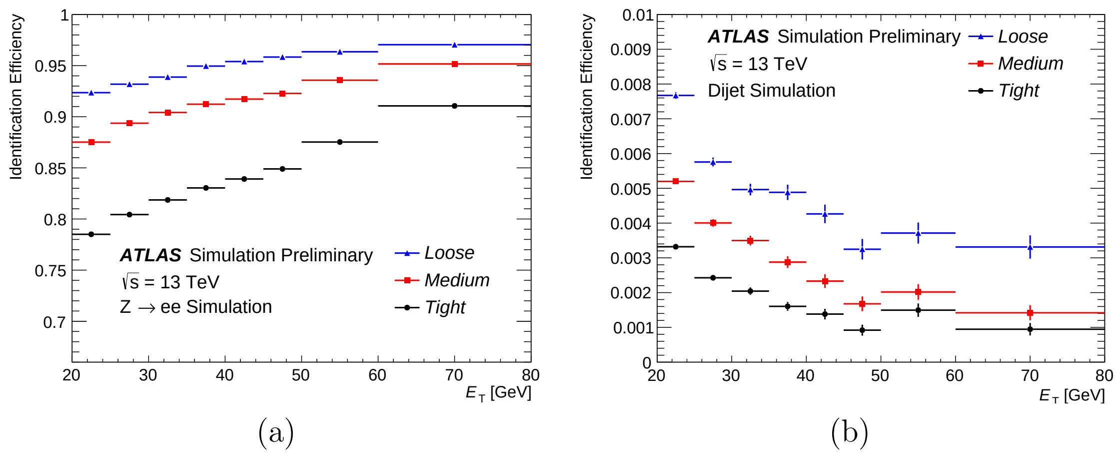


Figure 3.6: Efficiency to identify electrons from Z $\to$ ee decays 3.6 (a) and the efficiency to identify hadrons as electrons estimated using simulated dijet samples 3.6 (b). The efficien- cies are obtained using MC simulations, and are measured with respect to reconstructed electrons

In addition, Signal electrons are required to be isolated in order to suppress the background such as electrons from photon conversion in hadron decays, electrons from heavy flavour hadron decays, and light hadrons. Two discriminant criteria are used: a calorimeter isolation discriminator and a tracking based discriminator.

- Calorimeter isolation energy, known as raw $E_T^{topcone}$ isolation defined as the sum of the transverse energies of topological clusters, within a cone of $\Delta R = 0.2$ centred around the candidate electron cluster, excluding the detector energy.

- A track isolation, $P_T^{varcone}$, is defined by the sum of transverse momenta of tracks satisfying good quality requirements and originating from the primary vertex inside a cone $\Delta R = 0.2$ centred around the electron candidate direction and excluding the electron associated track.

3.4.2 Muons

Muon reconstruction and identification uses information collected by the inner detector system (ID) with that of the muon spectrometer (MS), and also leave very small energy deposits in the calorimeters. In order to combine the muon tracks, four algorithms that are based on the information from the inner detector, muon spectrometer, and calorimeters, are defined, as follows: [45]:

- Combined (CB) muons: are built from a global refit that uses the hits from tracks reconstructed independently in the ID and MS.

- Segment-tagged (ST) muons: If a track from the ID is associated with at least one local track segment in the MDT or CSC chambers of the MS. this track is classified as a segment-tagged muon.

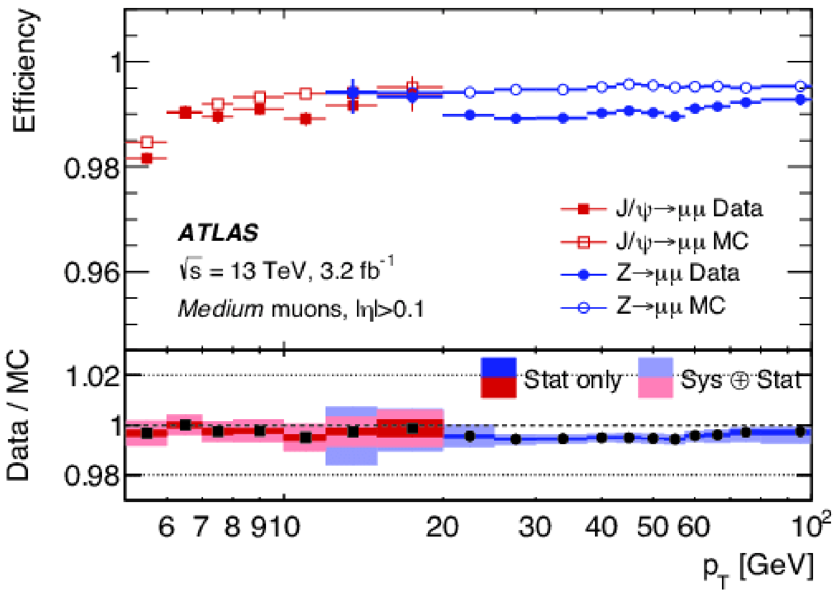
- Calorimeter-tagged (CT) muons: The muon candidate is defined with an ID track matched to a calorimeter energy deposit compatible with a minimum ionising particle.

- Standalone (SA) or extrapolated (ME) muons: muons are reconstructed from MS tracks compatible with originating from the interaction point and with at least two layers of MS.

Muon identification is performed by applying quality requirements that select prompt muons against backgrounds mainly coming from pion and kaon in-flight decays. several variables provide good discriminantbetween prompt muons (signal) and backgrounds:

- q/p significance of the difference of the charge-over-momentum ratios measured by the ID and by the MS
- ρ the difference between the transverse momentum p_T measurements in the ID and MS divided by the p_T of the combined track.
- normalized χ^2 of the combined track fit.

Based on these quantities, four muon identification working point are provided: Loose, Medium, Tight and High-P_T used to address the specific needs of different physics analyses [45]. Figures 3.7 shows the muon reconstruction efficiency as a function of muon p_T in the $Z \to \mu\mu$ and $J/\psi \to \mu\mu$ decay events.



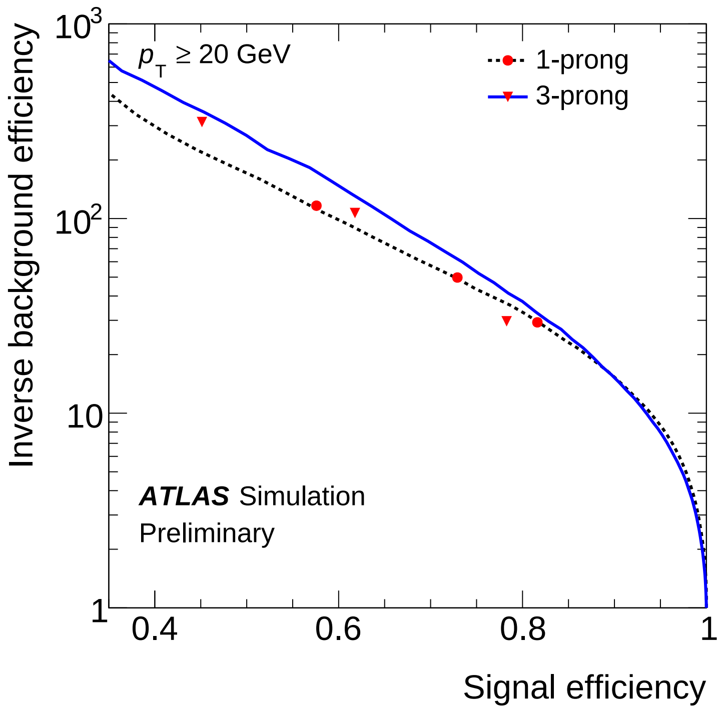
[h]

Figure 3.7: Total reconstruction efficiency for medium as a function of muon p_T in the $Z \rightarrow \mu\mu$ and $J/\psi \rightarrow \mu\mu$ decay events in the region of $0.1 < |\eta| < 2.5$

Additionally, signal muons are required to be isolated from other activities in order to reject backgrounds such as muons from hadron decay, there are two discriminating variables As is done for electrons, a calorimeter-based $E_T^{topcone}$ isolation and a track-based isolation $P_T^{varcone}$ as described in electrons Section.

3.4.3 Tau

The tau lepton is the heaviest mass of $m_\tau = 1776.86 \pm 0.12 MeV$, and due to its short lifetime of $2.9 \times 10^{13} s$ ($c\tau = 87 \mu m$), The tau is the only lepton which decays into both leptons and hadrons, that can decay both leptonically in 35% of the cases and hadronically in the remaining 65%. The summary table for the decay mode is described in Table. 4.4. The majority of the hadronic decay modes are dominated with one or three charged particles, referred to as substructure 1 prongs and 3 prongs, respectively, accompanied by charged pions,$\pi^\pm$, and kaons,$K^\pm$. The 3-prong decays represent very small branching fractions 22% of all hadronic tau decays, and one prong decays represent an additional 72%.
The tau reconstruction algorithm using the anti-kt algorithm, described in Section 3.3.1 with a distance parameter R=0.4 (small-R) applied to clusters of calorimeter cells, also calibrated using a local hadronic calibration ; additionally, tau candidates are required to pass the baseline selections of $p_T > 20 GeV$ and $|\eta| < 2.5$. the primary tau production vertex is assigned by choosing the candidate track vertex with the largest fraction of momentum from tracks within a cone of $\Delta R <$ 0.2 of the tau jet. The tracks must have $p_T >$ 10 GeV and pass requirements on the number of hits, i.e at least two hits in the pixel detectorand seven in the total pixel and SCT layers. 3.8 shows the τ had reconstruction and identification efficiency defined as the fraction of 1-prong (3-prong) hadronic tau decays that are reconstructed as 1-track (3- track) tau candidates passing a given BDT criterion. Three working points provided are Loose, Medium and Tight, corresponding to different identification efficiency values.



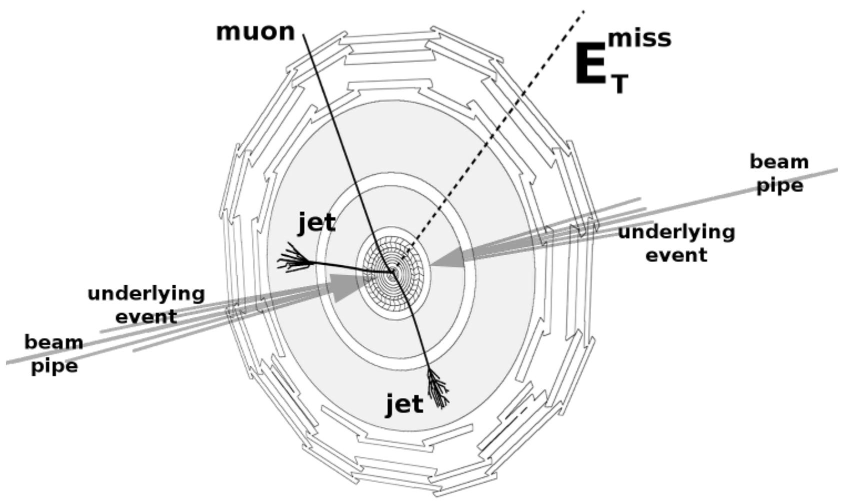
[h]

Figure 3.8: Inverse of the efficiency for mis-tagging QCD jets as a function of the identification efficiency for τ_had_vis candidates. The two lines refer to 1-track and 3-track candidates. The Loose, Medium and Tight working points are shown on these lines. The working points do not stay exactly on the line because they implement variable cuts to achieve a reduced p_T-dependentcy of the efficiency.

The tau identification algorithm is based on the rejection of these backgrounds through boosted decision tree (BDT) methods. The BDT is separately trained for one and three prong decays in order to distinguish tau candidates from jets, a set of discriminating variables based on the shower in the calorimeter, the number of associated tracks and the displaced tau-lepton-decay vertex is used.

3.5 Missing transverse energy

The law of conservation of momentum implies that the total momentum of the proton-proton collisions products in the plane transverse to the beam pipe should be zero. An apparent imbalance in the sum of all the visible transverse momentum, called missing transverse momentum, E_T^{miss}, and indicates the presence of the weakly or non interacting stable particles in the final state. Within the Standard Model, these particles are the neutrino, but also new particles that don't interact with the detector materials such as particles included in BSM models like Supersymmetry ,Dark matter and extra dimensions models. Figure 3.9 shows therelationship between the missing transverse momentum to the physics objects in the transverse plane.



[h]

Figure 3.9: Illustrative diagram of a hypothetical W $\rightarrow \mu\nu$ + jets event, for which the W boson candidate is balanced against jets.

The E_T^{Miss} is obtained from the negative vector sum of transverse momenta of all detected particles in pp colision, i.e:

$$E_T^{miss} = \sum_{i \in visible} P_T^i \tag{3.7}$$

The E_T^{Miss} reconstructed in ATLAS [46, 47] is characterized by two main contributions. The first comes from hard objects containing properly reconstructed and calibrated objects: muons, electrons, photons, tau and jets. The second is a soft term comprising of additional signals that are not associated with any of the hard objects being reconstructed.
The E_T^{Miss} of an event is calculated as the negative vectorial sum of selected and calibrated hard objects, and the soft term:

$$E_T^{miss} = -(\underbrace{\sum_{i \in electron} P_T^i + \sum_{i \in muon} P_T^i + \sum_{i \in \gamma} P_T^i + \sum_{i \in \tau} P_T^i + \sum_{i \in jet} P_T^i}_{hardobject} + \underbrace{\sum_{i \in soft} P_T^i}_{Softobjects}) \tag{3.8}$$

The soft term is reconstructed from the transverse momentum deposited in the detector but not associated with any reconstructed hard object. this can be built either by calorimeter-based methods, defined as Calorimeter Soft Term (CST), or track-based methods, defined as Track Soft Term (TST). The methods are briefly described below.
To avoid energy double counting in the calorimeter cells topoclusters are matched to reconstructed physics objects in the following order muons (μ), electrons (e), photons (γ), hadronically decaying τ-leptons, and jets [47].
From the components $E_{x,y}^{miss}$, the magnitude E_T^{miss} and azimuthal angle ϕ^{miss} can be calculated as:

$$E_T^{miss} = (E_x^{miss}, E_y^{miss}), \tag{3.9}$$

$$E_T^{miss} = (E_x^{miss}, E_y^{miss}), \tag{3.10}$$

$$\phi^{miss} = \tan^{-1}(E_y^{miss}/E_x^{miss}), \tag{3.11}$$

The total transverse energy in the detector $\sum E_T$, is an important quantity to estimate the event activity which is defined as the scalar sum of the transverse momenta of the hard objects and soft term contributions to the E_T^{Miss}:

$$\sum E_T = \sum P_T^e + \sum P_T^\mu + \sum P_T^\gamma + \sum P_T^\tau \sum P_T^{jet} + \sum P_T^{softTerm}, \tag{3.12}$$

A hard object equivalent of this quantity is called H_T defined as the scalar sum of the transverse momenta from all the reconstructed hard objects:

$$\sum H_T = \sum P_T^e + \sum P_T^\mu + \sum P_T^\gamma + \sum P_T^\tau \sum P_T^{jet}, \tag{3.13}$$

3.5.1 Physics Objects Selection

This section describes the selections that are applied to the reconstructed objects physics such as electrons, muons, photons and hadronically decaying tau-leptons and which are considered for the E_T^{miss} calculation. This same selection criteria is considered in Chapter 4. Generally, the object selections for electrons, muons, τ-jets and photons require refinements to achieve an optimal E_T^{miss} reconstruction performance in the context of a given physics analysis, as the criteria used in the $t\bar{t}$ analysis described in Section 5.

3.5.1.1 Hard objects

All the hard objects have its own dedicated calibration, translating detector signals into a totally corrected four-momentum, and a subset of those reconstructed objects are selected using quality and isolation criteria, as introduced in Sections 3.3-3.4.3
Consider the selection of event used in the E_T^{miss} performance studies, photons and hadronically decaying τ-leptons have almost no impact on the E_T^{miss} calculation, but they're documented during this Section to offer a comprehensive description.

- Electron

The electrons candidates in the E_T^{miss} reconstruction are passing the reconstruction quality requirements described in Section 3.4.1 with satisfying the "Loose" isolation and "Medium" likelihood identification criteria. In addition the electrons are required to have a minimum $P_T > 10GeV$ and $|\eta| < 2.47$, or $1.37 < |\eta| < 1.52$, to avoid the transition region between the barrel and end-cap EM calorimeters.

- Photon

Photons are required to pass the tight quality criteria, In addition to the reconstruction quality requirements, photons are normally selected with $P_T > 25GeV$ and $|\eta| < 1.37$ or $1.52 < |\eta| < 2.37$. In the case that to be included in the E_T^{miss} reconstruction.

- Hadronically decaying $\tau - leptons$

Hadronically decaying τ-leptons are reconstructed as described in Section 3.4.3 pass a medium quality criteria, with $|\eta| < 1.37$ and $1.52 < |\eta| < 2.47$.

- Muon

The muons selected for E_T^{miss} reconstruction are reconstructed as described in 3.4.2, and also by requiring the medium reconstruction quality and "Loose" isolation criteria. The muons included in the E_T^{miss} reconstruction are required to have $P_T > 10GeV$ and $|\eta| < 2.5$.

- Jets

Jets are reconstructed as small-R jets with R $= 0.4$ that are used for the E_T^{miss} described in section 3.3.3. After the calibration, they are required to have $p_T > 20GeV$ and $|\eta| < 4.5$. To further reduce the effects of pile-up, jets with $p_T < 60GeV$ and $|\eta| < 2.4$ are required to have JVT > 0.59. For details on jet reconstruction, calibration and pile-up suppression via JVT discriminant, refer to Section 3.3.5.

3.5.1.2 Soft Terms reconstruction

There are mainly two algorithms of soft terms that can be used to reconstruct the E_T^{miss} according to Equation 3.8. The first is called the TST (Track Soft Term) which is based on selected hard scatter vertex. TST is the standard soft term to be used in offline algorithms. An alternative type of the soft term is the calorimeter soft term(CST) which is based on topoclusters. The focus of this Thesis concerns the various corrections for the track-based soft term.

Calorimeter-based soft term (CST)

The calorimeter-based soft term is reconstructed using energy deposits in the calorimeter which are not associated to the high-p_T physics objects used in the E_T^{miss}, grouped into cells belonging to topoclusters and calibrated at the LCW scale. The soft term is referred to as the CST and writen as CST E_T^{miss}.

Track-based soft-term (TST)

The track-based soft term is reconstructed purely from all tracks information that satisfy the track selection criteria described in 3.2 that are not associated with the hard scatter primary vertex which are not used in the reconstruction of any hard physics object [46] which is defined in Chapter 4. This algorithm allows excellent vertex matching for the soft terms E_T^{miss} is largely insensitive to pile-up effects but misses from any soft neutral particles [46].

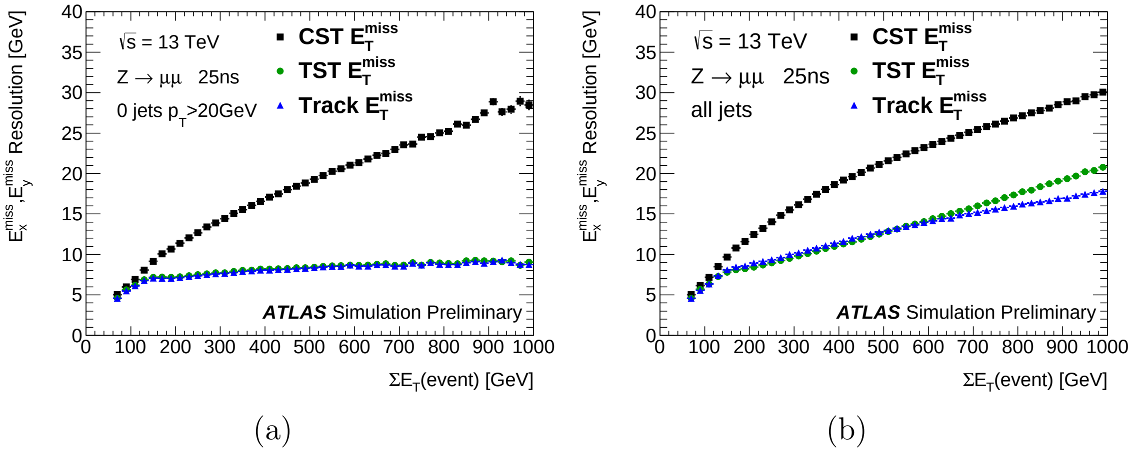


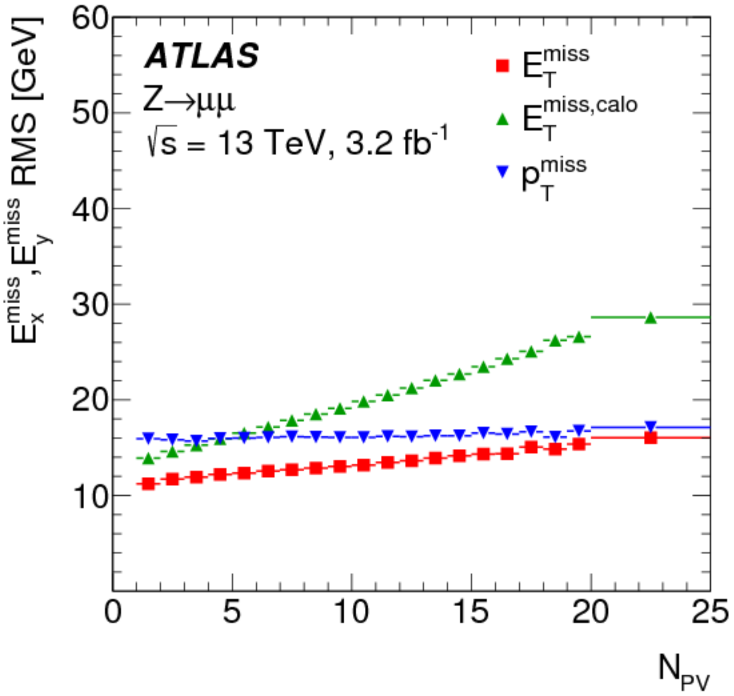
Figure 3.10: Comparison of the performance of E_T^{miss} built from TST and CST, and the p_T^{miss}, as quantified by the resolution, as a function of the CST $\sum E_T$. Powheg+Pythia Z *to* ll simulation is shown [44].

3.5.1.3 E_T^{miss} resolution

The E_x^{miss} and E_y^{miss} are expected to be approximately Gaussian distributed for $z \to ll$ events. Deviations from noise at large $\sum E_T$ values. However, because these distributions have non-Gaussian tails, the root-medium-square (RMS) is used to estimate the resolution. In the definition of the resolution, the RMS is preferred over the width of a Gaussian fit to the distribution to preserve the information on the tail. The resolution of the E_T^{miss} distribution is extracted using the RMS from the combined distribution of E_x^{miss} and E_y^{miss}. For processes with real E_T^{miss}, the true E_x^{miss} and E_y^{miss} are subtracted from their reconstructed quantities in simulation [47]. Figure 3.10 shows the resolution comparison between E_T^{miss} calculated with TST or CST and Track missing transverse momentum as a function of $\sum E_T$ for $z \to \mu\mu$ in the 0-jet and inclusive samples. Figure 3.11 shows the resolution of the E_T^{miss} reconstructed with TST or CST andthe track-only-based variant of the missing transverse momentum p_T^{miss} as a function of the pile-up activity measured in terms of the number of reconstructed vertices N_{PV}.

3.5.1.4 Track missing transverse momentum p_T^{miss}

Track E_T^{miss} and p_T^{miss} are completely reconstructed from ID tracks on their own, reducing pile-up contamination and dependency. It is calculated by taking the negative vector total of the p_T of tracks that satisfy the quality requirements defined in Section 3.2. According to interactions within the ID, the p_T of the electron is determined more accurately with the calorimeter, and therefore the p_T of the electron track is replaced by the calorimeter cluster measurement [48].
Although p_T^{miss} gives a pile-up robust E_T^{miss} calculation, it is insensitive to neutral giving a pile-up of robust particles that don't form tracks within the ID. Also, the $|\eta|$ coverage of the PmissT is restricted to the ID acceptance of $|\eta| < 2.5$, substantially smaller than the calorimeter coverage which extends to $|\eta| = 4.9$



[h]

Figure 3.11: Comparison of the resolution of E_T^{miss}, reconstructed with TST (E_T^{miss}), or with CST $E_T^{miss,calo}$, and the track-only-based variant p_T^{miss}. Resolutions are calculated Z $\rightarrow \mu\mu$ + jets sample.

3.6 Summary

Reconstruction of varied physics objects from the energy signatures within the different ATLAS subdetectors has been defined, along side strategies for their identification and calibration. The search presented during this thesis based on variety of various objects to select and reconstruct events with a semileptonically decaying $t\bar{t}$ pair and also for reconstrcution of E_T^{miss}: High-p_T electrons and muons, hadronic jets with two different R parameters: anti-k_T jets with $\Delta R = 0.4$, and trimmed anti-k_T jets with $\Delta R = 1$ as well as missing transverse energy. in addition, algorithms for the identification of b-jets applied only to anti-k_T $\Delta R = 0.4$ jets, play a crucial role in the selection of $t\bar{t}$ events.

Chapter 4

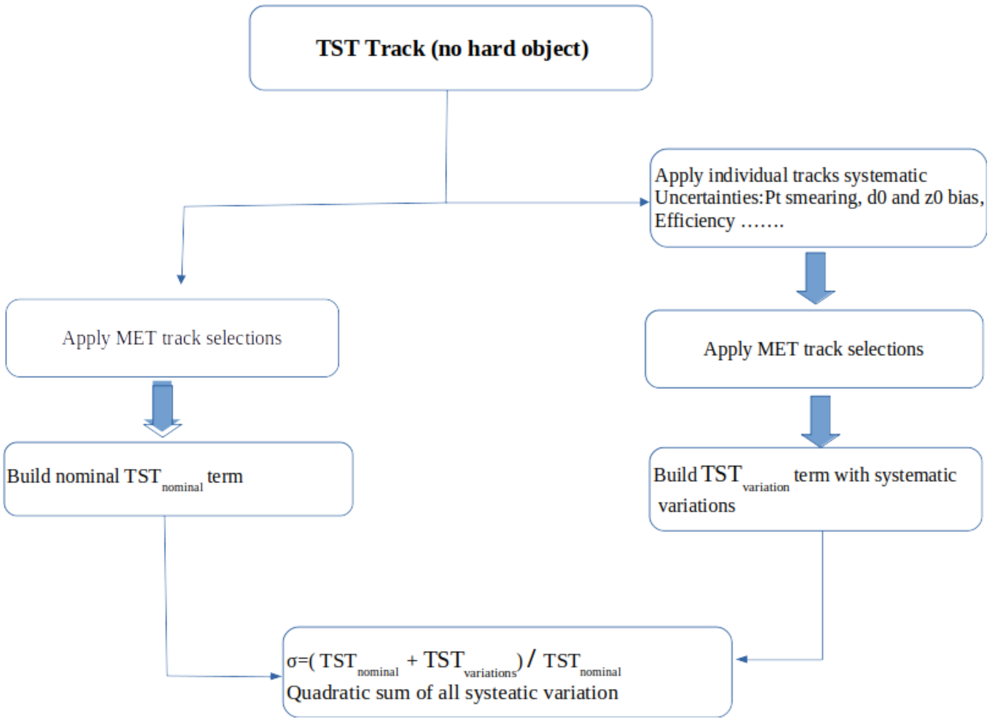
Track based Missing Transverse Momentum Systematics in ATLAS

4.1 Introduction

In a collider event like ATLAS, the missing transverse momentum E_T^{miss} is defined as the momentum imbalance in the x-y plane transverse to the beam axis, as introduced in section 3.9, where the law of momentum conservation is expected. The law implies that the vector momentum imbalance in the transverse plane is obtained from the negative vector sum of the momenta of all particles detected in a proton proton collison should be zero. Such an imbalance may indicative of unseen particles, such as neutrinos or stable particles in the final state, weakly-interacting particles. In aditional the events with large E_T^{miss} is key signatures for new physics such as supersymmetry (SUSY), dark matter particles and extra dimensions. However fake E_T^{miss} can arise from interacting particles which escape the acceptance of the detector, are inaccurately reconstructed, or fail to be reconstructed altogether.
The systematic uncertainty associated with the reconstructed components used to calculate E_T^{miss} is evaluated on the perspective of the knowledge of the reconstructed objects which used to build it. The overall systematic uncertainty of the E_T^{miss} measurement is then calculated by combining the uncertainties of each term corresponding to the reconstructed physical object and the soft term uncertainties discussed in this section.
In this chapter, the estimation of the systematic uncertainties for various aspects of the ATLAS Inner Detector on the track based soft term $E_T^{miss,TST}$ will be presented. It gives an overview of the sources of systematic uncertainties that will be considered and the way they are propagated to the final result. Such the effects of the variation include the track reconstruction efficiency measured is presented for the Loose and Tight track quality selections described in section 4.4.1. The evaluation of the fake rate as well as the impact parameter resolution is also shown for completeness, along with systematic uncertainties in the alignment weak mode.
This chapter is organised as follows, section 4.2 describes the algorithm and definition of the object-based E_T^{miss} soft term uncertainties, next Section 4.3 describes the samples and event selection used in this Chapter, and for the next Section 4.4 presents the performance of the $ET^{miss,TST}$ soft term uncertainties, In particular,

the contribution of the E_T^{miss} is important in Z events using a control sample $Z \to \mu\mu$

4.2 Object-based E_T^{miss} systematics uncertainties algorithm and definition



[h]

Figure 4.1: Simplified organizational chart of the uncertainty on the soft term algorithm.

The greatest contribution to the systematic uncertainty of the E_T^{miss} comes from the soft track term $p_T^{miss,TST}$, as the hard term consists of leptons and jets with well-measured track P_T in the nominal case or the object corrected case, the lepton P_T measurements combined with small systematic uncertainties and jets with a well-defined energy scale and resolution.
To evaluate the uncertainty on the soft term we consider the systematic uncertainty in the reconstruction tracks of charged particles. It is the uncertainty in the description of the material in the inner detector. The procedure consist into creating an algorithm that based first of all to calculate the track soft term E_T^{miss}. A schematic diagram of the algorithm is shown in Figure 4.1. Ideally, the E_T^{miss} soft term ought to be calculated using the track sample for each reconstructed event collected from high-quality tracks emerging from the hard-scatter vertex but not associated with any matched to any reconstructed hard object that presented in section 3.5.1.1. The applied signal-overlap that satisfy the requirements described in the following selection 4.4.1 are removes:

- ID tracks associated with Electron;

- ID tracks associated with muons;
- ID tracks associated with τ ;
- ID tracks ghost-associated with fully or partially contributing jets

The second part of this study shall comprise of applying the "track selection" and the "Systematic variation" to the track soft term E_T^{miss} (TST). Thus there are two kind of tracks, tracks soft term nominal reffered as $TST_{nominal}$ and the tracks soft term variation are reffered as $TST_{variation}$.
To build the nominal term $TST_{nominal}$ the following tracks are vetoed:

- If a track is not associated with a reconstructed object physics, its transverse momentum is added to the calculation of the soft term.
- Applying the E_T^{miss} track selections on the track collection of TST calculated.
- Building the nominal $TST_{nominal}$ term.

Finally, to build the $TST_{variation}$ term with systematic variations is performed according to the following criteria:

- If a track is not associated with a reconstructed object physics, its transverse momentum is added to the calculation of the soft term.
- Applying the track systematic uncertainties on the track collection of TST calculated and giving the modified version of tracks. The remaining tracks required to satisfy the E_T^{miss} track selections described in 4.4.1.
- Thus building $TST_{variation}$ term with systematic variations.

The last part of the algorithm consist into calculate the quadratic sum 4.1 and 4.2 of the statistical using all of the compenents for each of liste uncertainties presented in table 4.1 according to the following equations:

$$\sigma_{TST} = \frac{TST_{Nominal} - TST_{Variation}}{TST_{Nominal}} \tag{4.1}$$

The difference between $TST_{nominal}$ and sigma given by the following equation:

$$\sigma_{TST} = TST_{Nominal} - \frac{TST_{Nominal} - TST_{Variation}}{TST_{Nominal}} \tag{4.2}$$

The implementation of this procedure is based on two software tools:

Tracking Performance Variation Name	Label	LF-Calibration notes
$TRK\_EFF\_LOOSE/TIGHT\_GLOBAL$	eff-GLOBAL-L/T	Not relevant
$TRK\_EFF\_LOOSE/TIGHT\_PhModel$	eff-PhModel-L/T	Not relevant
$TRK\_EFF\_LOOSE/TIGHT\_IBL$	eff-IBL-L/T	Not relevant
$TRK\_EFF\_LOOSE/TIGHT\_PP0$	eff-PP0-L/T	Not relevant
$TRK\_FAKE\_RATE\_LOOSE/TIGHT$	fake-rate	Relevant, Syst. variation
$TRK\_BIAS\_QOVERP\_SAGITTA\_WM$	sagitta-bias WM	Not relevant
$TRK\_RES\_D0/Z0\_DEAD$	*deadmodules*	Not relevant
$TRK\_RES\_D0\_MEAS$	$d_$0-smearing	Relevant
$TRK\_RES\_Z0\_MEAS$	$z_$0-smearing	Relevant
$TRK\_RES\_D0\_MEAS\_UP$	$d_$0 -smearing up	Syst. variation
$TRK\_RES\_Z0\_MEAS\_UP$	$z_$0 -smearing up	Syst. variation
$TRK\_RES\_D0\_MEAS\_DOWN$	$d_$0-smearing dw	Syst. variation
$TRK\_RES\_Z0\_MEAS\_DOWN$	$z_$0-smearing dw	Syst. variation
$TRK\_RES\_D0\_MEAS$	d_0 z_0-smearing	Syst. variation

Table 4.1: List of effects considered for the adjusted-MC calibration of the LF-jet tag rate. For details about the tracking performance variation see [49].

- InDetTrackSystematicsAlgs, an Athena algorithm which takes as input a collection of tracks from the xAOD data-format and produces a new collection of tracks to which the effects of the tracking recommendations are propagated;

- TrackSelectionTool, an Athena algorithm which takes as input the xAOD tracks and evaluate the track selection criteria

Table 4.1 list all the studied tracking performance variations, implemented in the InDetTrackSystematicsAlgs package. For details about the tracking performance variation see [49]

4.3 Physics samples and event selection

4.3.1 Data samples

The analysis discussed in this chapter uses pp collision data collected by the ATLAS detector at the center-of-mass energy 13 TeV in 2015 and 2016. Data can only be analyzed if it has been collected during stable proton beams with nominal magnetic field conditions and if it fulfills the standard ATLAS data quality assessment criteria which include checks that the internal detector, calorimeters and MS are functioning effectively. The resulting integrated luminosity is 36 fb^{-1}.

4.3.2 Monte Carlo samples

Monte Carlo simulated events are used to model Standard Model processes which are normalized using the currently available predictions for their cross-sections. These simulated events were passed through the ATLAS detector simulation [50] based on GEANT 4 [51]. The effects of pile-up were modeled by overlaying minimum bias events, simulated using the soft QCD processes of Pythia 8.1 [52] with the A2 [53] set of tuned parameters and MSTW2008LO [27] parton distribution functions (PDF). For all samples of simulated events, except for those generated using Sherpa [54], the EvtGen v1.2.0 program [55] is used to describe the decays of bottom and charm hadrons.

Events with W or Z bosons decaying into leptons produced in association with jets are simulated using the Sherpa 2.2.1 [54], generator. Matrix elements are calculated for up to 2 partons at NLO and 4 partons at LO using the Comix [56] and OpenLoops [57] matrix element generators and merged with the Sherpa parton shower [58] using the $ME + PS@NLO$ prescription [59]. The W/Z + jets events are normalized to the NNLO cross-sections. The Powheg-Box v2+Pythia8 [60] and MadGraph5_$aMC@NLO$ 2.2.0+Pythia8 [61] Z boson simulation are used to compute the systematic uncertainties on the $ET^{miss,TST}$. In addition to the full Geant4 simulation of the ATLAS detector, dedicated systematic uncertainties are also derived for samples coming from inner detector simulation.

Diboson processes (WW , WZ and ZZ) are simulated using the Powheg-Box v2 generator [60] to the next-to-leading order (NLO) in QCD and after that interfaced with the Pythia 8 generator. Next-to-next- leading order (NNLO) in QCD and NLO electroweak corrections (EW) are added as a function of the invariant mass, $m_Z Z$, of the ZZ process.

For the generation of tt at NLO, the Powheg-Box v2 generator was used. Single top quark events in the $s-$, $t-$ and Wt-channels were generated using the Powheg-Box v1 generator [62, 63]. The top quark mass was set to 172.5 GeV. The overall yield predicted for the tt process is rescaled according to the NNLO cross-section, including the resummation of soft gluon emission at next-to-next-to-leading-logarithm accuracy (NNLL) as available in Top++2.0 [64]. The overall yields predicted for single top quark production in the s-, t-, and Wt channels are rescaled according to their respective NLO cross-sections [65, 66].

4.3.3 Events Selection

Events are selected with single muon or single electron triggers. For 2015 data, electron (muon) triggers had a p_T threshold of 24 (20) GeV. During the 2016 data taking period, the p_T threshold for the electron and muon triggers rose over time to 26 GeV as a result of the increase in the instantaneous luminosity.

Events are required to have two opposite-charge leptons of the same flavor and pass the "medium" identification criteria, see Sections 3.4.1 and 3.4.2 for details, with the more energetic lepton having $p_T > 25$ GeV and the subleading lepton having $p_T > 20$ GeV. If the event has any additional lepton that meets the "loose" identification criteria and having $p_T > 7$ GeV, the event is rejected to reduce the "contamination from the WZ process.

The invariant mass of the lepton pair is required to be in the mass window $|m_Z - m_{\mu\mu}|$ < 15 GeV, where m_Z is equal to the value of the Z boson mass obtained by the Particle Data Group, and reported in Reference [67].Selections listed above would select $Z \to \mu^+\mu^-$ events, but it is convenient to impose another selection based on the number of reconstructed jets; $N_{jets} = 0$ after overlap removal with hard objects and failing the "BadLoose" criteria, which relates to jet cleaning as defined in Reference [68], are removed in data and simulation.

4.3.4 Data and Monte Carlo comparisons

The distribution of the track-based soft term is shown for the $Z \to \mu\mu$ event selections in figure 4.2. The dominant systematic uncertainty bands come from generator differences in the modelling of the underlying event. and this systematic uncertainty on the E_T^{miss} soft term is shown separately to indicate its shape and contribution size. The $Z \to \mu\mu$ event topologies have similar levels of agreement between data and simulation.

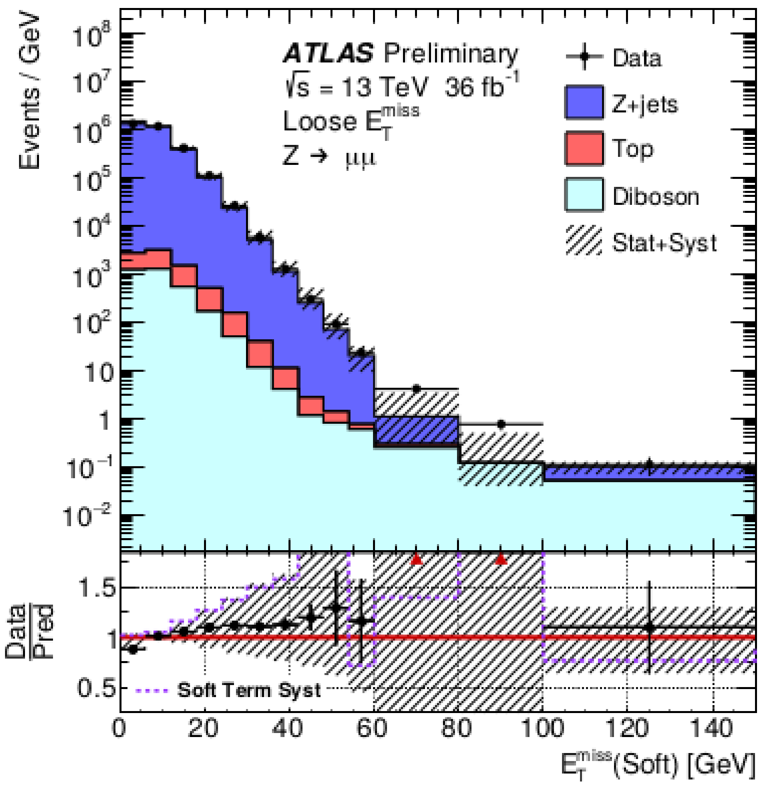


Figure 4.2: E_T^{miss} soft term distributions in data compared to MC predictions including all relevant backgrounds for events satisfying the Z $\to \mu\mu$ + jets.

4.4 Performance of the E_T^{miss} soft term uncertainties

4.4.1 Track Quality Selection

To correctly use tracking information, one must determine whether the track is well reconstructed while associating the track with a vertex in the event. Vertex association allows us to focus on tracks that belong to the hard scatter vertex and thus enhance the performance of E_T^{miss} reconstruction.

Tracks have been reconstructed using the MinBias setup, and they are selected using specific quality criteria for the track. In this thesis, two sets of quality criteria are studied, based on transverse momentum (p_T), pseudorapidity and number of subsystem hits. The tracking algorithm can efficiently reconstruct tracks a higher transverse momentum threshold of 500 MeV was chosen for these studies in order

to keep the rate of fake tracks as low as possible.
Loose

- $p_T > 500MeV$,
- $|\eta| < 2.5$
- Number of Pixel and SCT clusters on track ('silicon hits') ≥ 7
- Number of shared modules ≤ 1
- Number of silicon holes $\leq$2
- Number of pixel holes ≤ 1

Tight Primary (in addition to the Loose selection requirements)

- Number of silicon hits 9 (if $|\eta| \leq 1.65$)
- Number of silicon hits 11 (if $|\eta| \geq 1.65$)
- At least one hit on one of the two innermost pixel layers
- No pixel holes

A shared hit is a hit over one track used. If it has one or more shared hits, a Pixel module is considered shared. A shared module has at least two shared hits 2 in the case of the SCT. Holes are defined as intersections of the reconstructed track trajectory with a sensitive detector element resulting in no hit-essentially a "missing hit" on a track. In order not to penalize late-decaying / secondary particles or tracks that do not reach the detector 's outer silicon layers, holes are only counted between the first and last hit assigned to the track. They are estimated by closely following the track trajectory and comparing the hits-on-track to the intersecting modules. Inactive modules or regions on the silicone sensors such as edge areas are excluded from the definition of the hole. The tracks selected by the above working points may have extensions to the TRT, but there are no explicit criteria for the number of TRT hits applied.
The selection of the Loose track corresponds to the default track requirements applied during the reconstruction [69] of the inside out track and aims to achieve high charged-particle reconstruction efficiency at the cost of a non-negligible fraction of fake tracks. The name Tight Primary reflects the fact that this set of criteria is designed and optimized to select primary tracks and to reject fake tracks at the cost of reduced efficiency in the reconstruction of tracks.
The transverse impact parameter (d_0) is defined to be the shortest distance from the trajectory to the nearby vertex in the transverse plane and the corresponding z value is the longitudinal impact parameter z_0. Each vertex candidate must have at least two tracks that satisfy:

- transverse impact parameter with $|d_0| < 1.5mm$

- longitudinal impact parameters with $|z_0 \sin(\theta)| < 1.5mm$.

Finally, in order to reduce the number of mis-reconstructed tracks, isolated tracks, excluding muon tracks, with $p_T > 200$ GeV and $|\eta| < 1.5$, and with $p_T > 120$ GeV and $|\eta| > 1.5$ are used if:

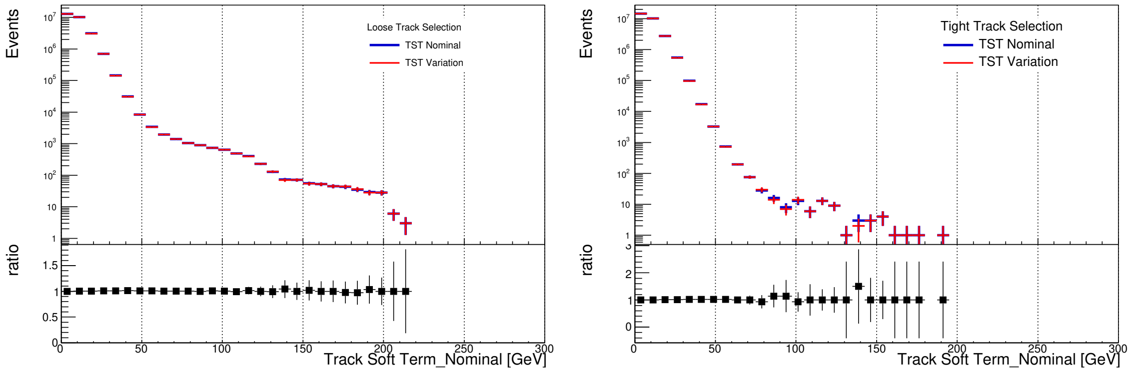
- the relative uncertainty on the charge to track-momentum ratio (q/p) is $\frac{\sigma(q/p)}{(q/p)} < 0.4$
- the energy in calorimeter clusters in a cone of radius 0.1 in the (η, ϕ) plane around the track should reflect the reconstructed track momentum $E_T^0.1/p_T^{tracks} > 0.65$

4.4.2 Track reconstruction efficiency and Systematics

The track reconstruction efficiency, is one of the main ingredient in the charged-particle multiplicities measurements, because its systematic uncertainties represent the main source of uncertainty for this analysis.
The tracking efficiency for primary particles used in this analysis is based on simulation with a correction to take into account material description deficiencies. The systematic uncertainties on the track reconstruction efficiency are based primarily on the knowledge of the material in the detector as well as on the differences in data and simulation in the track selection and the alignment.
The tracking efficiency for primary particles is based on simulation using the primary particle definition and track-to-truth particle association described in Chapter 3 and as defined in Appandex. The efficiency is measured in simulation as a function of p_T and η it is defined as follows:

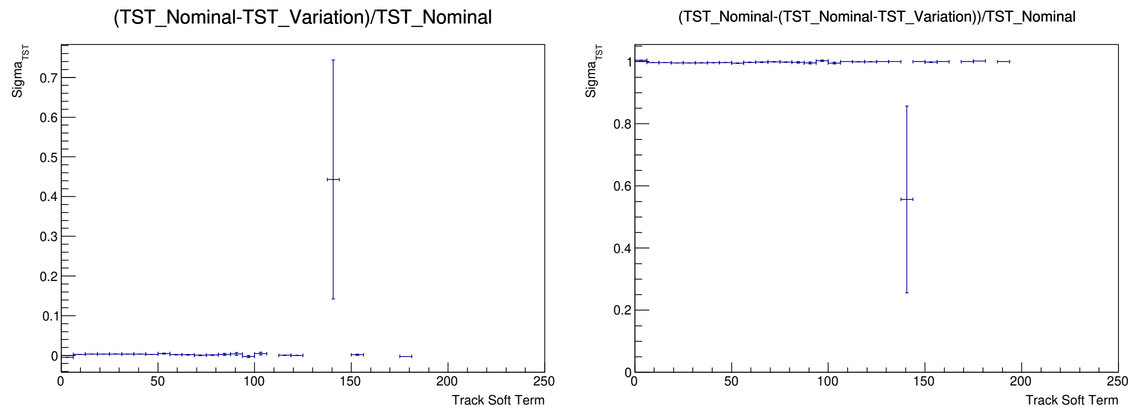
$$\zeta_{track}(p_T, \eta) = \frac{N_{Rec}^{matched}(p_T, \eta)}{N_{gen}(p_T, \eta)} \tag{4.3}$$

Where p_T and η are generated particle properties, $N_{Rec}^{matched}$ is the number of reconstructed tracks ($P_{match} > 0.5$) to a given generated charged-particle and $N_{gen}(p_T, \eta)$ is the number of generated charged-particles in that (p_T, η) bin. Only events which pass the event selection defined in Section 4.4.1 are used. It was tested that the measured efficiency is independent of the number of generated particles in a given event and is therefore independent of the trigger and vertex requirements. This allows to populate phase space regions with a low number of tracks in physics events using a single particle sample and to apply the selection cuts with respect to the generated vertex in these events.
The tracking efficiency is parameterized as a function of the number of selected tracks desribed in 4.4.1 and 3.5.1 with $p_T = 500$ MeV, for $E_T^{miss,TST}$ wich is labeled track soft term for the tight selection as shown in figure 4.4 and for the loose selection 4.5 . The events have an efficiency very near to 100%, is due to particles passing through a larger number of sensitive layers, so in this cases the corrections are not needed.
Figures 4.3 shown the E_T^{miss} track soft term for track efficiency global systematic for loose and tight selection compared to the nominal distribution. The effect of the



(a) (b)

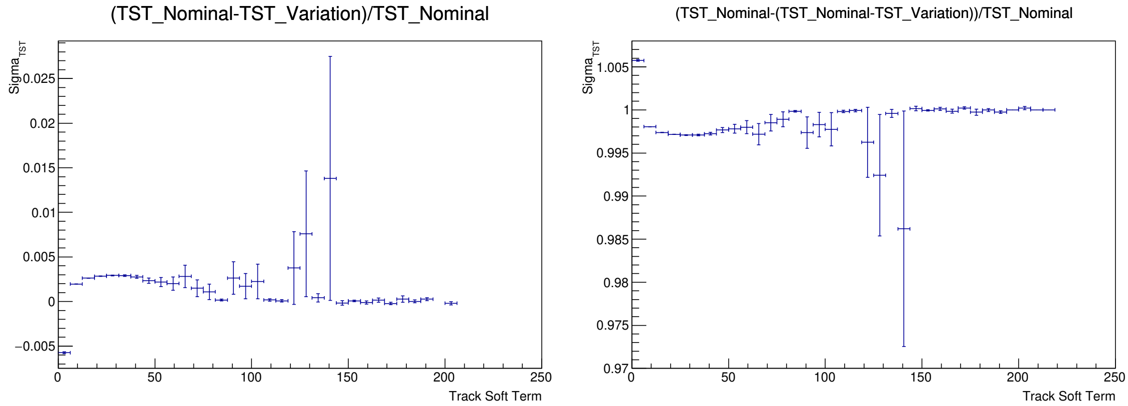
Figure 4.3: Distribution of tracks soft term for Z→ $\mu\mu$ applying the Loose (a) and Tight (b) Primary selections. In (a) and (b), the uncertainty using on the tracks is considered, but are too small to be seen. the uncertainty applied, is $TRK\_EFF\_LOOSE/TIGHT\_GLOBAL$.



(a) (b)

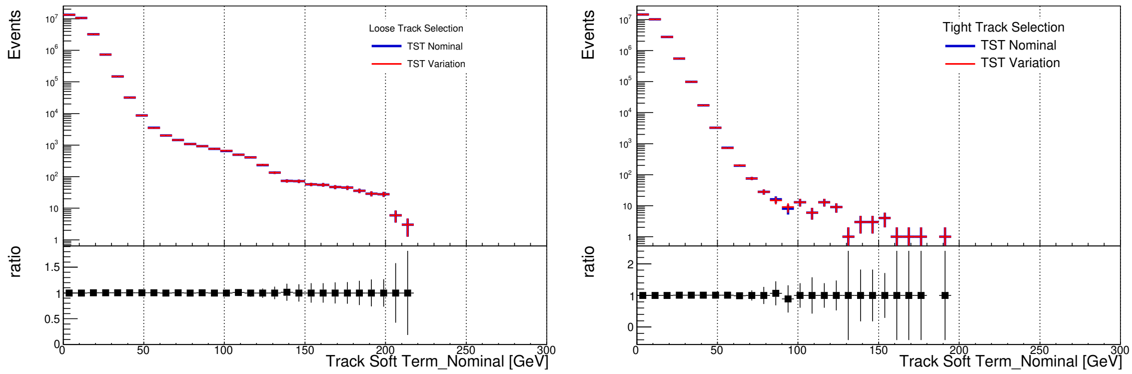
Figure 4.4: Track Reconstruction efficiency for Z→ $\mu\mu$ without jets with $p_T > 500MeV$ with respect to the event selection as a function of the track soft term, the uncertainty used, is $TRK\_EFF\_TIGHT\_GLOBAL$

global variation is negligable for track soft term. The tracking efficiency for physic model, IBL and PP0 materials is shown in Figure 4.7, 4.10 and 4.13 respectively for tight selection and 4.8, 4.11 and 4.14 for loose selection as a function of track soft term [TST]. The lower track reconstruction efficiency in the region TST $<$ 50 GeV is due to the increasing amounts of material that the particles must traverse. and is affected by at most 0.1%. The slight increase in efficiency for TST $>$ 50 GeV very near 100% is due to particles passing through a larger number of sensitive layers. so additional corrections are not needed.
As it can be seen in figures 4.6, 4.9 and 4.12 the setup which shows the E_T^{miss} track soft term distribution for track efficiency Physic model, IBL and PP0 systematic which are listed in table 4.1 for loose and tight selection respectively compared to the nominal distribution. The effect of the global variation is negligable for track soft term.



(a) (b)

Figure 4.5: Track Reconstruction efficiency for Z→ $\mu\mu$ without jets with $p_T > 500MeV$ with respect to the event selection as a function of the track soft term, the uncertainty used, is $TRK\_EFF\_LOOSE\_GLOBAL$



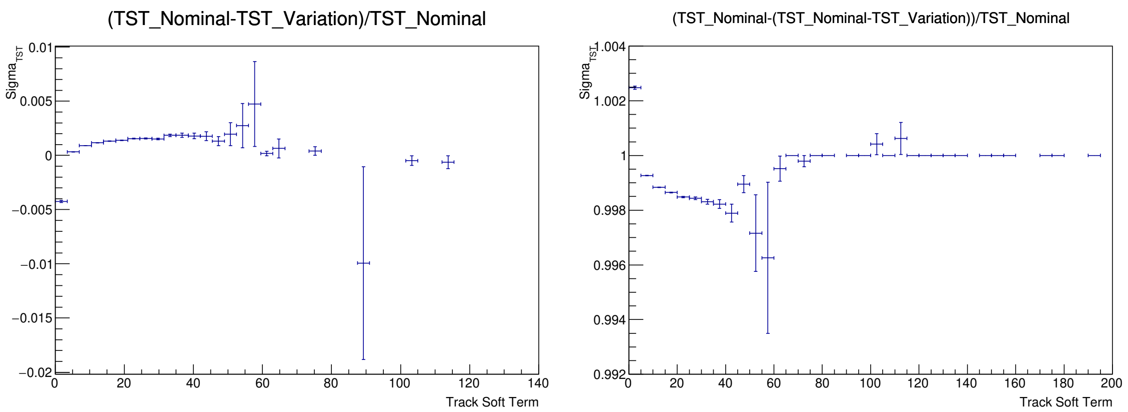
(a) (b)

Figure 4.6: Distribution of tracks soft term for Z→ $\mu\mu$ applying the Loose (a) and Tight (b) Primary selections. In (a) and (b), the uncertainty applied, is $TRK\_EFF\_LOOSE/TIGHT\_PhModel$.

4.4.3 Fake Rate

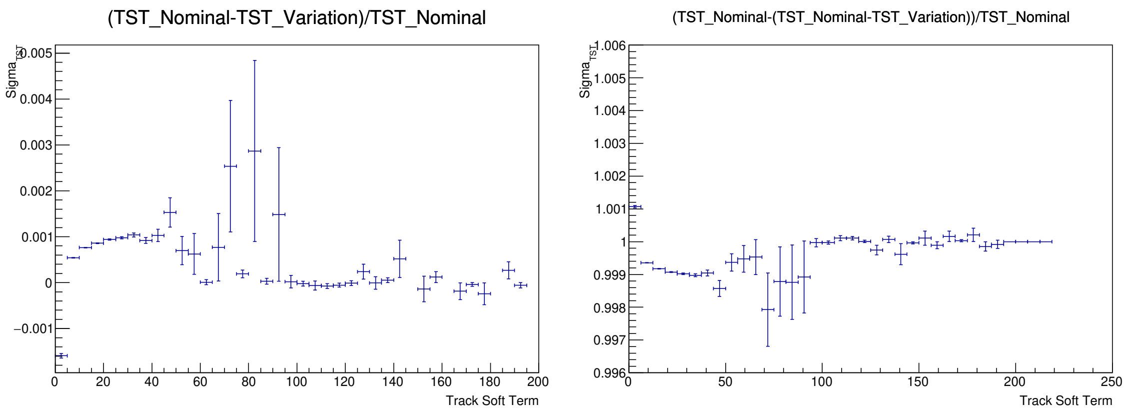
Not all tracks reconstructed in the ID will actually correspond to a charged particle traversing the detector; in particular, combinatorial effects which increase with pile-up can give rise to fake tracks. As a cross-check on the modelling of the tracking fake rate, an estimate of this quantity is made by assuming that the number of non-fake tracks is to first order proportional to the number of pile-up interactions. Any deviation from linearity is therefore assumed to be due to fakes. Figure 4.15 shows the average number of tracks passing the Loose and Tight Primary selections, as a function of μ.
The data in Figure 4.15(a) are fitted with a linear function $f(\mu) = m\mu$ in the region $10 \leq \mu \leq 15$ in which both data and Monte Carlo demonstrate approximately linear behaviour. The relative deviation from the linear fit:

$$\frac{< N_{tracks}(\mu) - f(\mu) >}{< N_{tracks}(\mu)} \tag{4.4}$$



(a) (b)

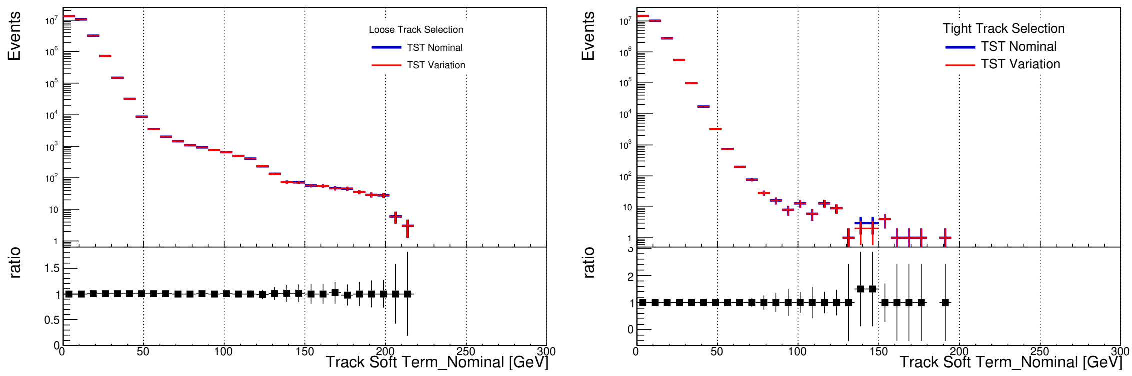
Figure 4.7: Track Reconstruction efficiency for Z→ $\mu\mu$ without jets with $p_T > 500MeV$ with respect to the event selection as a function of the track soft term, the uncertainty used, is $TRK\_EFF\_TIGHT\_PhModel$.



(a) (b)

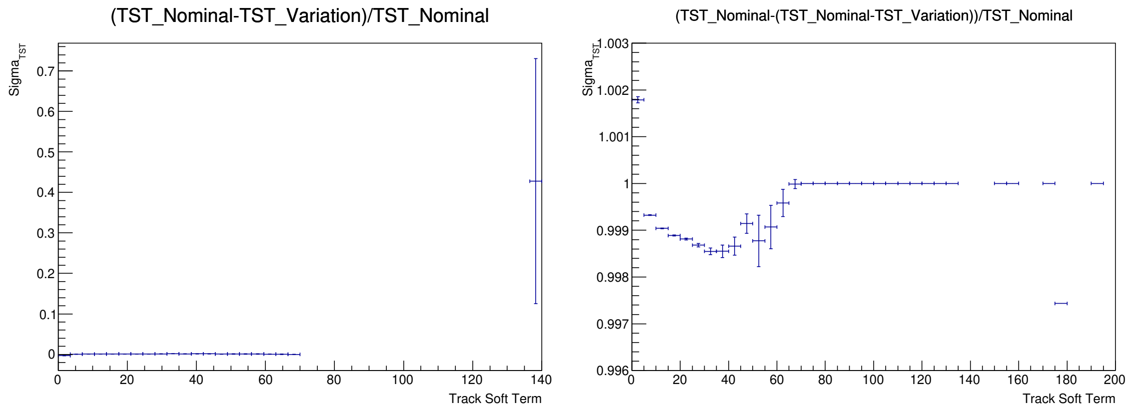
Figure 4.8: Track Reconstruction efficiency for Z→ $\mu\mu$ without jets with $p_T > 500MeV$ with respect to the event selection as a function of the track soft term, the uncertainty used, is $TRK\_EFF\_LOOSE\_PhModel$.

provides an estimate of the fake rate, which is shown in Figure 4.15(b). It is a relative measure, and by construction close to zero in the region of the fit, $\leq \mu \leq 15$. By approximating the fake rate in this manner, the contribution from fake tracks is assumed to be negligible in the fit region; this is a more reliable assumption for the Tight Primary track selection than for the Loose track selection. This method also assumes that the tracking efficiency is independent of μ, and that the fake contribution from sources other than pile-up combinatorics are negligible; these latter two assumptions have been verified to hold in Monte Carlo. The Tight Primary track selection is estimated to have a smaller fake rate, which is also more stable versus μ, than the Loose selection. The recommended systematic uncertainty to be applied on the fake rate is 50%, which is typical of the size of the differences between Data and MC seen in these studies. The TST systematics of the tracks fake are computed following the procedure describing above. The resolution, for $E_T^{miss,TST}$ is presented in Z→ $\mu\mu$ event topology in figure 4.16 for tight and loose selection.



(a) (b)

Figure 4.9: Distribution of tracks soft term for Z→ $\mu\mu$ applying the Loose (a) and Tight (b) Primary selections. In (a) and (b), the uncertainty using on the tracks is considered, but are too small to be seen. the uncertainty applied, is $TRK\_EFF\_LOOSE/TIGHT\_IBL$.



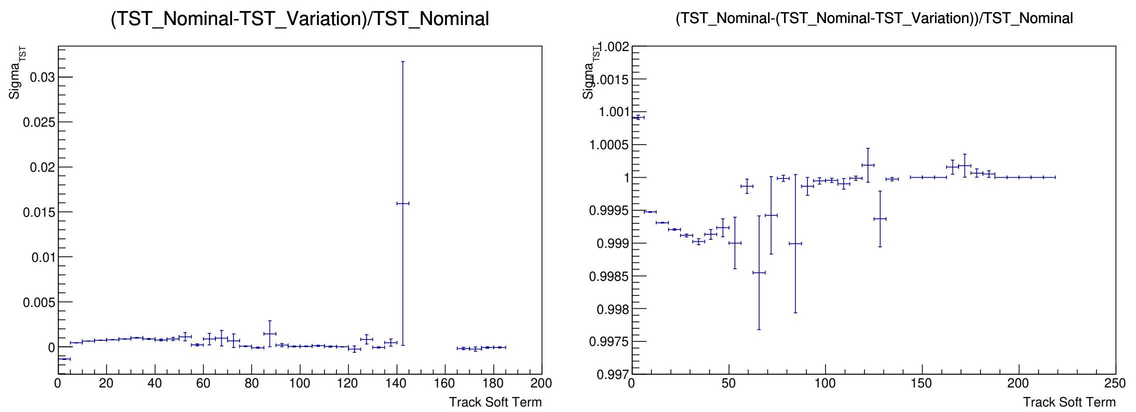
(a) (b)

Figure 4.10: Track Reconstruction efficiency for Z→ $\mu\mu$ without jets with $p_T > 500MeV$ with respect to the event selection as a function of the track soft term, the uncertainty used, is $TRK\_EFF\_TIGHT\_IBL$.

The validation plots obtained from the application of the systematic envelope (Table 4.1) to the $E_T^{miss,TST}$ in Z→ $\mu\mu$ simulation. the majority of the effects were found to be negligible for the tracks within the uncertainties, as it can be seen for example in Figure 4.18 and 4.17 in which the effect of adding 50% more material in the forward region of the Inner Detector was studied for the loose and tight track selection.

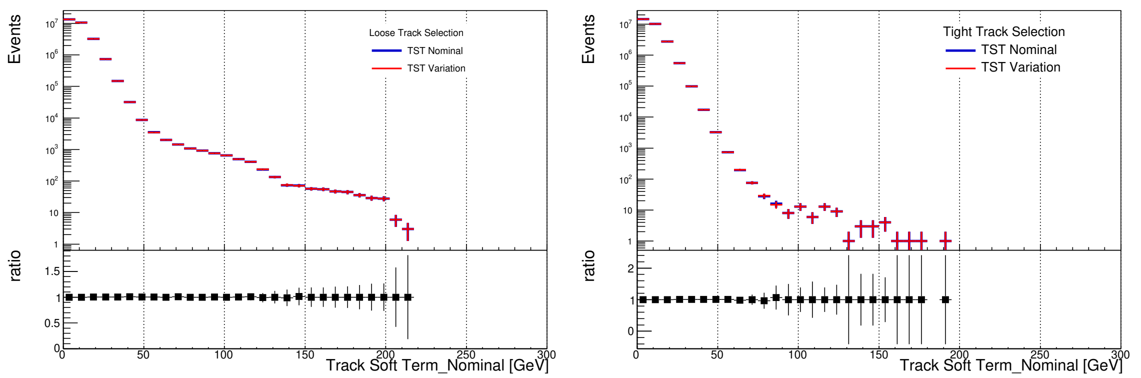
4.4.4 Impact Parameter resolution

The transverse and longitudinal impact parameters, d_0 and z_0 respectively, are important to discriminate tracks originating from primary vertices from tracks originating from secondary vertices. The primary vertices are used as the reference point for the distance of closest approach. The intrinsic resolutions of the track impact parameters with respect to the primary vertex are measured from data and simulation.To extrapolate the systematic uncertainty to higher track P_T on the impact



(a) (b)

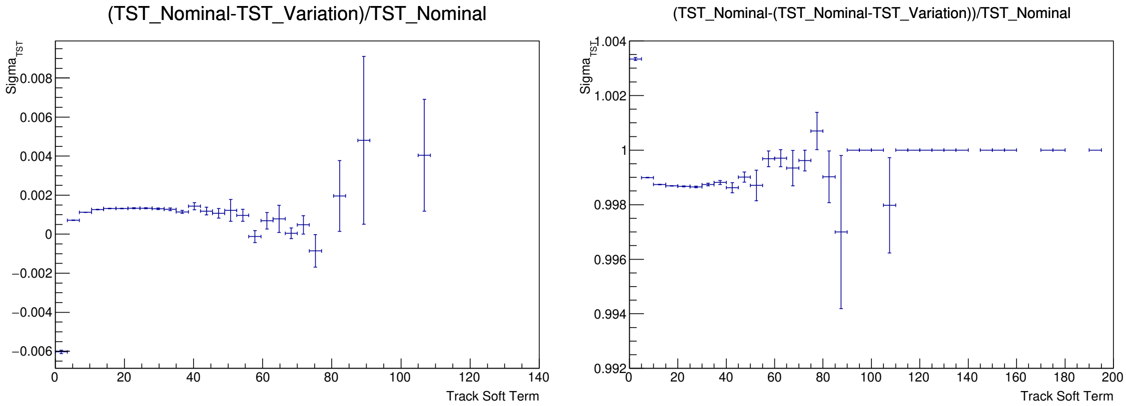
Figure 4.11: Track Reconstruction efficiency for Z→ $\mu\mu$ without jets with $p_T > 500 MeV$ with respect to the event selection as a function of the track soft term, the uncertainty used, is $TRK\_EFF\_LOOSE\_IBL$



(a) (b)

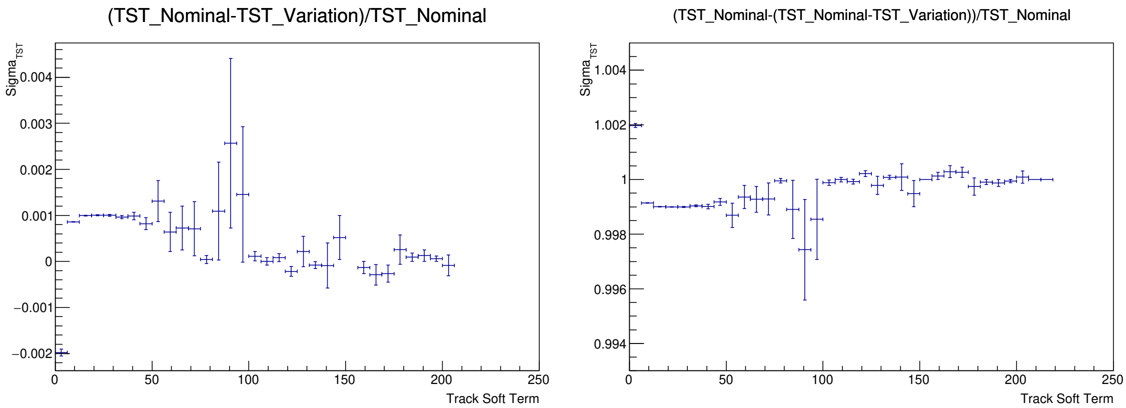
Figure 4.12: Distribution of tracks soft term for Z→ $\mu\mu$ applying the Loose (a) and Tight (b) Primary selections. In (a) and (b), the uncertainty using on the tracks is considered, but are too small to be seen. the uncertainty applied, is $TRK\_EFF\_LOOSE/TIGHT\_PP0$.

parameters, the results using in this study from muon tracks from Z boson decay. Events are required to have a reconstructed primary vertex with at least 10 tracks for the minimum bias data and simulation, while events with one or more additional reconstructed primary vertices with more than 4 tracks are removed to reduce the pile-up effect.
The results of the systematic uncertainties for the impact Parameter for $E_T^{miss,softTerm}$ are also obtained with method described summarized in table 4.1.
The contribution from tracks is found to be less than 0.01%, which is negligible due to the less number of events used in this study However, this method also assumes that impact of z_0 and d_0 uncertainties were investigated and is found to be less than 0.01%, which is negligible, as it can be seen for example in Figures 4.19, 4.20 and 4.21 for the tight track selection. Then No corrections are applied about z_0 and d_0.



(a) (b)

Figure 4.13: Track Reconstruction efficiency for Z$\rightarrow \mu\mu$ without jets with $p_T > 500 MeV$ with respect to the event selection as a function of the track soft term, the uncertainty used, is $TRK\_EFF\_TIGHT\_PP0$.



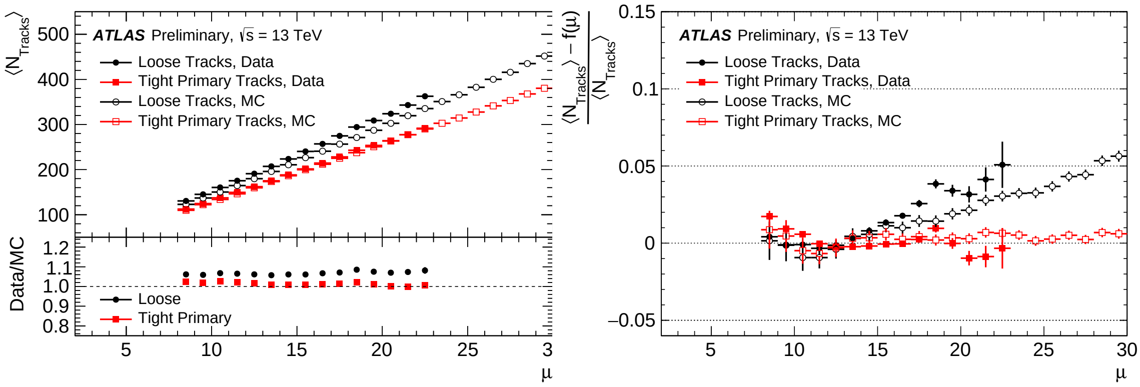
(a) (b)

Figure 4.14: Track Reconstruction efficiency for Z$\rightarrow \mu\mu$ without jets with $p_T > 500 MeV$ with respect to the event selection as a function of the track soft term, the uncertainty used, is $TRK\_EFF\_LOOSE\_PP0$.

4.4.5 Weak Modes

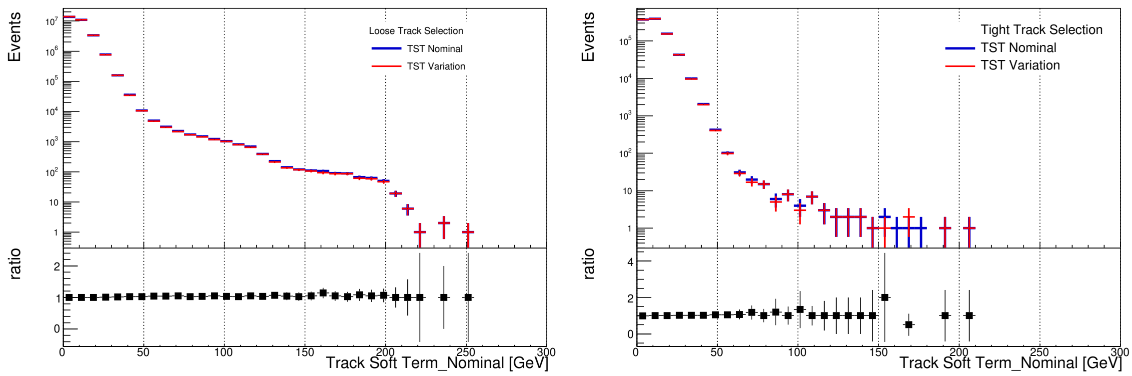
The alignment of the Inner Detector is based on tracking information and minimizes hit-to-track residuals with respect to alignment parameters using a least squares principle [16]. The procedure fundamentally suffers from so-called weak modes of the solution which are detector deformations under which the χ_2 of the track fit remains invariant. The weak mode deformations may, in general, affect reconstructed parameters of the tracks [17].
Since then, only three types of biases on reconstructed track parameters have been considered essential for results in physics: Charge-antisymmetric momentum bias, i.e. bias on the reconstructed Q/p due to sagitta distortion:

- Bias on the reconstructed charge over momuntum Q/p $\longrightarrow$ Q/p + $\delta_s agita$
- Bias on the reconstructed transverse impact parameter: $d_0 \longrightarrow d_0 + \delta_d 0$



(a) (b)

Figure 4.15: Average number of reconstructed tracks as a function of μ for data and minimum bias simulation, applying the Loose and Tight Primary selections (a), and an estimation of the tracking fake rate, derived from the deviation from linearity of a fit to Ntracks as a function of $\mu(b)$.



(a) (b)

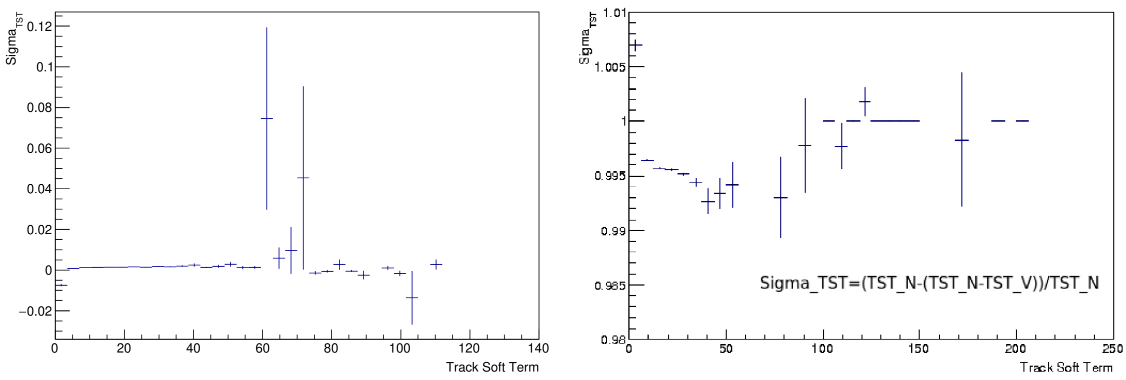
Figure 4.16: Distribution of tracks soft term for Z→ $\mu\mu$ applying the Loose (a) and Tight (b) Primary selections. In (a) and (b), the uncertainty applied, is $TRK\_FAKE\_RATE\_LOOSE/TRK\_FAKE\_RATE\_TIGHT$

- Bias on the reconstructed longitudinal impact parameter: $z_0 \longrightarrow z_0 + \delta_z 0$

The results for the systematic uncertainties of the $E_T^{miss,softTerm}$ obtained with the Weak Modes as it can be seen in Figure 4.23, is measured to be 100% for all expansions in Z_0, d_0, and Q_0/p for tight selection. This decrease show respectively in the ratio on the figure 4.22 that show the distribution of track soft term for Weak Modes systematic for tight selection compared to the nominal distribution. However, the correction can't be treated as as variation.

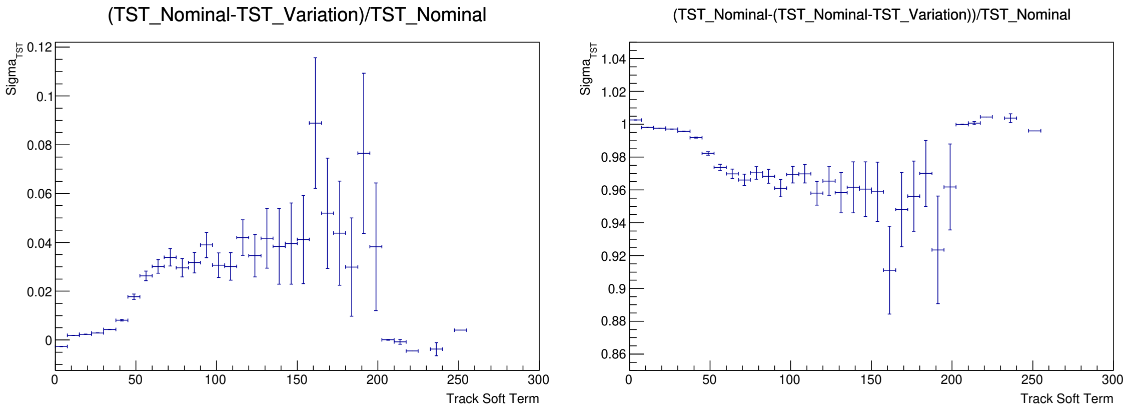
4.5 Summary

Missing transverse energy Soft term $E_T^{miss,softTerm}$ has been measured in $Z \to \mu\mu$ and events selected in the ATLAS detector with 13 TeV collision data corresponding



(a) (b)

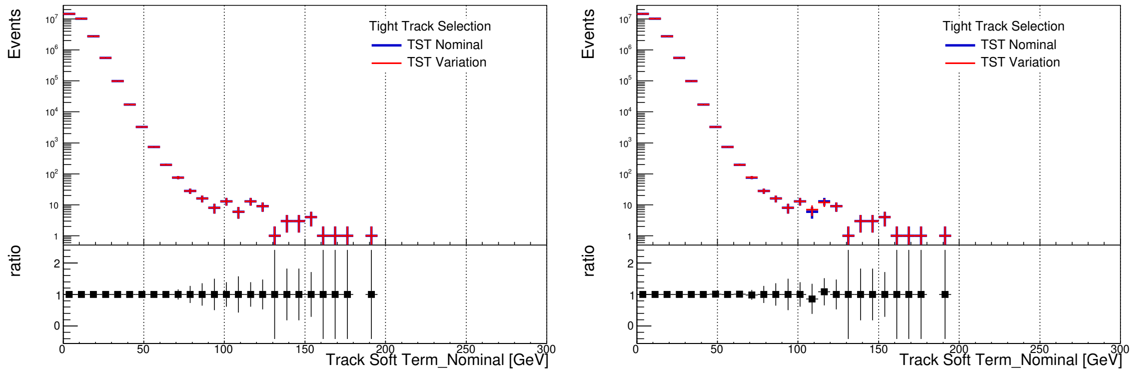
Figure 4.17: Track Reconstruction efficiency for Z→ $\mu\mu$ without jets with $p_T > 500MeV$ with respect to the event selection as a function of the track soft term, the uncertainty used, is $TRK\_FAKE\_RATE\_TIGHT$.



(a) (b)

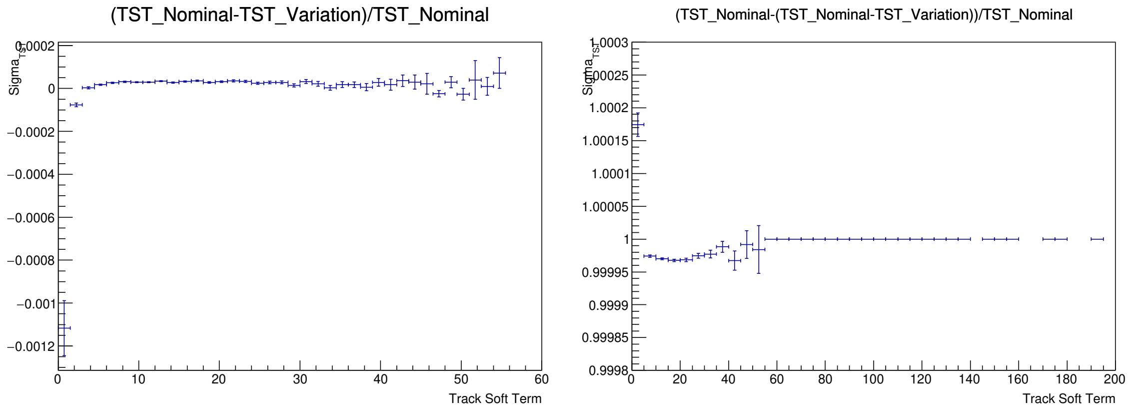
Figure 4.18: Track Reconstruction efficiency for Z→ $\mu\mu$ without jets with $p_T > 500MeV$ with respect to the event selection as a function of the track soft term, the uncertainty used, is $TRK\_FAKE\_RATE\_LOOSE$.

to an integrated luminosity of 36 fb^{-1}. An algorithm for calculate $E_T^{miss,softTerm}$, where the selection of tracks, is provided.
Different systematic uncertainties of the various objects on the inner detector are studied $E_T^{miss,softTerm}$ and compared with the the $E_T^{miss,softTerm}$ nominal. No large variation are observed in the $E_T^{miss,softTerm}$ distribution due to the less number of events of $Z \to \mu\mu$, then the systematic uncertainties and the correction procedures applied to this study are negligible. These results point out the need of a large number of events for more careful optimization for the track selection.



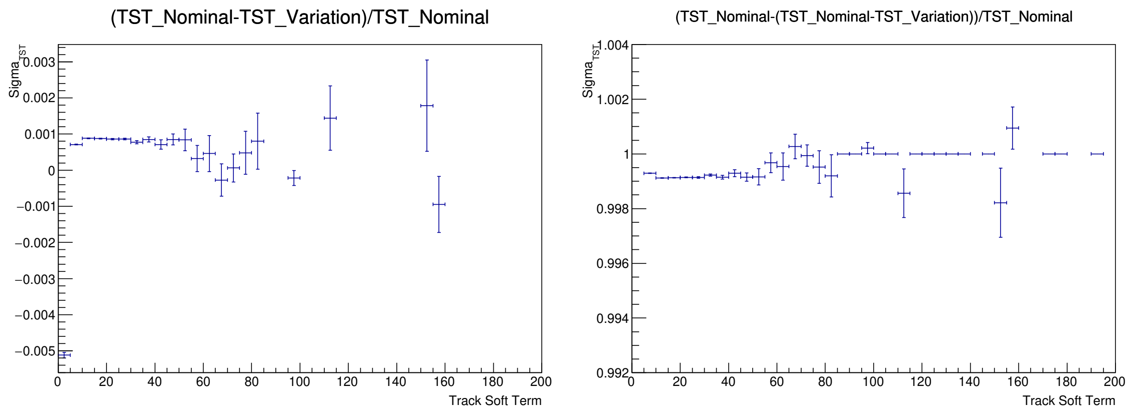
(a) (b)

Figure 4.19: Distribution of tracks soft term for Z→ $\mu\mu$ applying the Tight Primary selections by using the systematic uncertainties for the impact Parameter for d_0 (a)and z_0 (b)



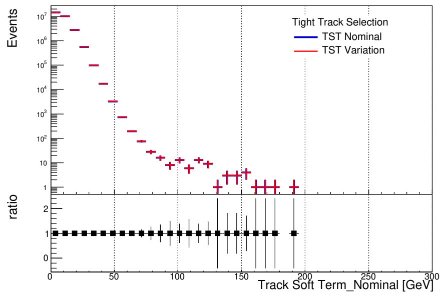
(a) (b)

Figure 4.20: Track Reconstruction efficiency for Z→ $\mu\mu$ without jets with $p_T > 500MeV$ with respect to the event selection as a function of the track soft term, the systematic uncertainty used for the impact Parameter is d_0.

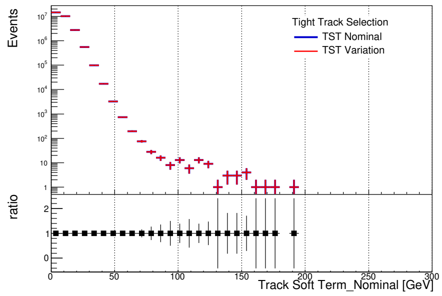


(a) (b)

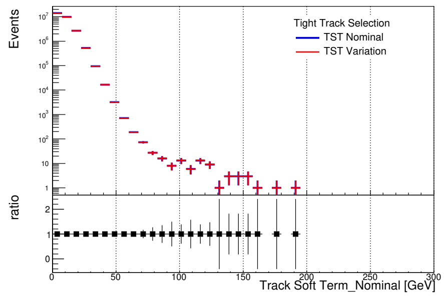
Figure 4.21: Track Reconstruction efficiency for Z→ $\mu\mu$ without jets with $p_T > 500MeV$ with respect to the event selection as a function of the track soft term, the systematic uncertainty used for the impact Parameter is Z_0.



(a)

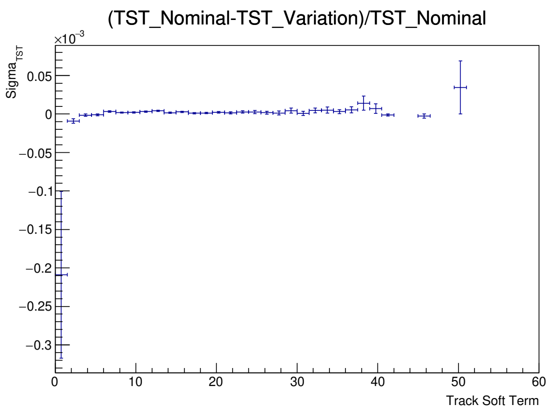


(b)

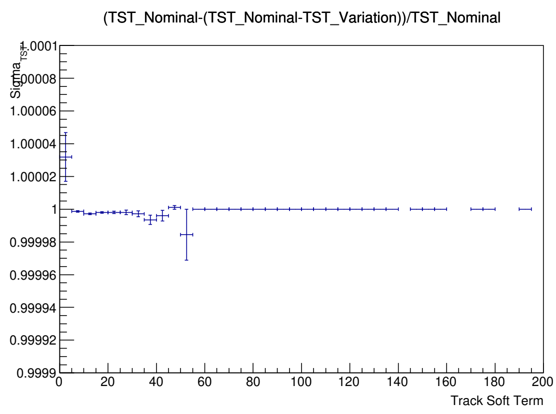


(c)

Figure 4.22: Distribution of tracks soft term for Z→ $\mu\mu$ applying the Tight Primary selections by using minimum bias simulated events of d_0 (a), Z_0 (b), and Q_0/p (c)



(a)



(b)

Figure 4.23: Track Reconstruction efficiency for Z→ $\mu\mu$ without jets with $p_T > 500 MeV$ with respect to the event selection as a function of the track soft term, the uncertainty used depend to d_0 .

Chapter 5

Search for $t\bar{t}$ resonances in the l + jets channel at $\sqrt{s} = 13$ TeV

5.1 Introduction

This chapter presents a search for a resonance decaying into a $t\bar{t}$ pair semileptonic final states, where one of the top-quarks decays leptonically, and the other one decays hadronically. This search is performed with the full Run-2 proton-proton collision data (139 fb^{-1}) collected during 2015-2018. this represented the highest energy reached for hadron-hadron collisions in a particle physics experiment.
The analysis in this thesis focus on searching for hight mass $t\bar{t}$ resonances essentially follows Run 2 analysis, new technique are therefore presented; ehanced b-tagging to obtain hight efficiency in complicated jet environment, a top-tagging for rejection background and new cut to reduce the standard Model $t\bar{t}$ background.

5.2 Analysis Strategy

The search for new resonances decaying to $t\bar{t}$ pairs is performed in the single lepton plus jets (l+jets) final state. This final state is distinguished by the presence of a single isolated (charged) lepton with large transverse momentum derived from the leptonically decaying W-boson, and a certain number of hadronic jets. It is known as the "golden channel" for searches and measurements in $t\bar{t}$ events. The Feynman diagram for this process $Z' \rightarrow t\bar{t} \rightarrow$ qqblν_lb is shown in Figure 5.2.
The analysis searches is for a narrow resonance in the $t\bar{t}$ invariant mass spectrum, $m_{t\bar{t}}$. $m_{t\bar{t}}$ is calculeted from the sum of the four-momentum of the reconstructed physics objects associated to the leptonic and hadronic top-quarks. The top-quarks are reconstructed using different techniques in the resolved and boosted regions. In the boosted channel, a Deep Neural Network (DNN) based top-tagging method was applied to the large radius jets to identify the hadronic tops. While in the resolved region the tops are reconstructed using a chi-squared algorithm using only the small radius jets. To further suppress the non-$t\bar{t}$ backgrounds and increase the sensitivity, b-quarks coming from the top-quark decay are identified using a b-tagging technique. The signal regions are defined based on the number of b-tagged

jets within an event. All the SM background processes, are modeled using Monte Carlo simulations except for the QCD. The QCD multijet background is estimated using a data-driven method. Since no trace of new physics is observed, upper limits on the signal cross-section x branching ratio are set at 95% CL will be set. The lower mass limit on the Z_{TC2} signal with 1.2% relative width is expected to be around 3.7 TeV, which is $\sim$ 700 GeV.

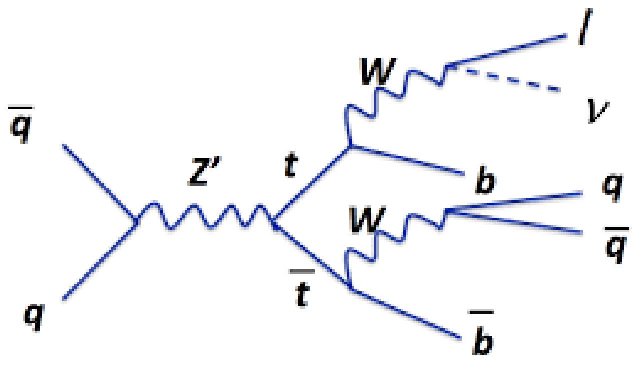


Figure 5.1: Feynman diagram for the signal process of LO lepton-plus-jets top pair production mediated by an interfering photon, Z boson and $Z\prime$ boson.

5.3 Top Decay Topologies

The detector signatures of $t\bar{t}$ events in the final state of l+jets relies on the transverse momenta of the top quarks, as demonstrated in Figure 5.2. For low transverse momentum top-quarks, the decay products are spread out, and the jets are far from each other. In this instance, the ATLAS calorimeter produces six different small-radius jets with R = 0.4. The hadronic top-quark decay results in three separate jets, one of which is a b-jet. The lepton from the decay of the other top quark is usually well separated from the corresponding b-jet. This topology is called the resolved topology. If the top quark has high momentum, then their decay products appear increasingly collimated in the direction of the respective mother particle in the laboratory frame. As a consequence, the hadrons coming from the top-decay become very close to each other, and their responses overlap in the calorimeter forming a large-radius jet with radius parameter (R) of 1. This decay topology, shown in the right pictogram of Figure 5.2, will be referred to as the boosted topology in the following.

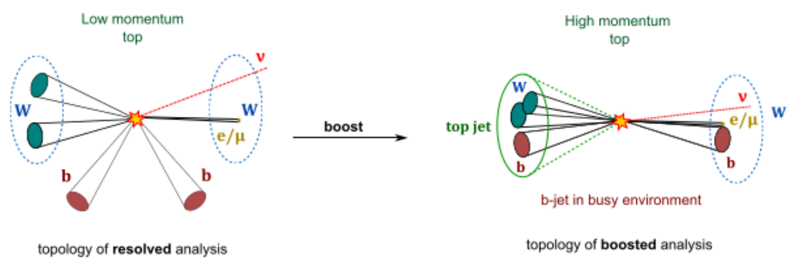


Figure 5.2: Schematic diagram showing the semileptonic $t\bar{t}$ final state in resolved and boosted regime.

5.4 Background processes

In this search, most of the background processes is estimated using Monte Carlo simulation. The main backgrounds for the leptons+jets analysis are irreducible SM $t\bar{t}$ process via s-channel and t-channel contribution. The QCD contributions is estimated using a data-driven technique. The single-top, W + jets, Z+jets, diboson, and associate production of $t\bar{t}$ with a vector boson ($t\bar{t}$V) are pass our event selection and constitute the reducible component of the background.
$t\bar{t}$ MC samples are generated by Powheg along with Pythia 8 using the next-to-leading-order (NLO) with the NNPDF3.0 NLO PDF set in the matrix element calculations. The single-top production is also generated by Powheg and Pythia 8 is used with NNPDF3.0 NLO PDF set for showering and hadronization with A14 tune.
W +jets and Z+jets events are evaluted using MC samples generated using the Sherpa 2.2.1 generator where matrix element calculation includes up to two partons at NLO in QCD, and four partons at the LO using the Comix and OpenLoops matrix element generator. The NNPDF3.0 NLO PDF set is used while generating the events. Generated events are normalized to the NNLO cross-section.
The SM Diboson (WW, WZ, ZZ) processes, with different decay channels (l$\nu q\bar{q}$) and (ll$q\bar{q}$) are also simulated using the Sherpa 2.2.1 generator using the NNPDF 3.0 PDF set with the dedicated parton shower tuning developed by the Sherpa authors. The SM multijet process is simulated using Pythia 8.186, with two jets in the hard process at leading order (LO) with the NNPDF23LO [220] PDF set and the A14 set of tuned parameters for the underlying event.

5.5 Signal processes

The benchmark signal models are generated using Monte Carlo (MC) simulation. The primary benchmark signal model is the Top-color assisted-Technicolor model. The other benchmark model are used sush as spin-1 color-octet Kaluza-Klein gluon g_{KK} and spin-2 color-singlet Kaluza-Klein graviton, G_{KK} are also used as benchmark signal models. The spin-1 color-octet Kaluza-Klein gluon $g_{KK} \to t\bar{t}$ is simulated with Pythia 8.165 at LO with NNPDF2.3 LO PDF set and the A14 set of tuned. The $g_G G \to t\bar{t}$ is simulated using MadGraph5 aMC@NLO NNPDF2.3 LO PDF set, with parton showering performed by Pythia 8.165 with the A14 set of tuned parameters.

5.5.1 Topcolor-assisted-Technicolor model

The signal process Topcolor-assisted-technicolor (TC2) $pp \to Z' \to t\bar{t}$ is simulated using the Pythia 8 generator with default settings for generic Sequential Standard Model (SSM) Z' generation. Samples are generated for a variety of Z' masses, using the NNPDF 2.3 LO PDF set. The widths are all negligible compared to the detector resolution and hence this sample suffices for use to derive limits on Z' production with small widths such as Z'_{TC2} and Z'_{HVT}. The cross-sections at leading-order (LO) at 13 TeV used for Z'_{TC2} are presented in the Table 5.1. Following [3] and the previous searches for $t\bar{t}$ resonances, a multiplicative k-factor of 1.3 is used to bring them at next-to-leading order (NLO). In this analysis, only Z'_{TC2} samples above mass 1.4 TeV are used.

Table 5.1: Leading-order theoretical cross-sections for the Z'_{TC2} signal, before appliying the k-factor of 1.3.

Z'_{TC2} mass	LO cross-section ($\Gamma = 1.2\%$) [pb]
400 GeV	70.3
500 GeV	40.1
750 GeV	10.7
1 TeV	3.70
1.25 TeV	1.51
1.5 TeV	0.684
1.75 TeV	0.334
2. TeV	0.172
2.25 TeV	0.0924
2.5 TeV	0.0511
2.75 TeV	0.0289
3 TeV	0.0167
4 TeV	0.00213
5 TeV	0.000331

5.6 Object definition

The reconstruction and identification of the physics objects used in this search has been considerd in this section 2.5. The selection criteria imposed on these objects are presented and motivated with regard to the search strategy defined above.

Large radius calorimeter jets

Looking at top pair decays with a topology, Large-R jets are built using three-dimensional topological energy clusters calibrated with the local cluster weighting (LCW) procedure, described in 3.3.6 with the radius parameter R = 1.0. The reconstructed jets are trimmed by building the subjets coming from the soft radiations using the radius parameter $R_{trim} = 0.2$, which are required to have $p_T > 300$ GeV and to be within η <2.0. In addition, the jets are required to have mass m > 40 GeV. Large-R jets were furthermore used in this analysis for identify hadronically decaying top-quarks using the DNN-based top-tagging algorithm. Details of the top-tagging method were described in 3.3.7.

Small radius calorimeter jets

In the case of the resolved topology, only small-R jets are a requirement to reconstruct the individual hadronic decay products of the top quarks. Small-R jets are required to have $p_T > 25$ GeV and $\eta < 2.5$, also required to have a jet vertex fraction JVF > 0.5 to reject jets from pile-up vertices within the kinematic range of $p_T < 60 GeV$ and $\eta < 2.4$.

Variable radius (VR) track jets

Track jets are reconstruct using the anti-kt algorithm with with R = 0.4 applied on the inner detector tracks associated with the primary vertex and calorimeter clusters passing the Particle Flow reconstruction algorithm. The input tracks are required to have a transverse momentum greater than 0.5 GeV and $\eta < 2.5$ as discribe in 3.3.1. The overlapping VR track jets in an event are removed. The deep learning based b-tagging technique is used to identify the VR tracks jets to identify the b jets.

Table 5.2: Quality requirements for the three jet containers used in this analysis.

Large radius jet	
Jet radius	R=1
Clustering algorithm	anti-k t on LCTopo
Transverse momentum	$P_T > 300$ GeV
Pseudorapidity coverage	$\eta < 2.5$
Trimming	R_{sub} =0.2, f_{cut}= 0.05
Small radius jets	
Jet radius	R=0.4
Clustering algorithm	anti-k_t on EM topocluster
Transverse momentum	P_T >25 GeV
Pseudorapidity coverage	$\eta < 2.5$
JVT	JVT > 0.59
Small radius jets	
Jet radius	R=0.4
Clustering algorithm	anti-k_t on tracks
Transverse momentum	P_T >10 GeV
Pseudorapidity coverage	$\eta < 2.5$

Electrons

The reconstruction, identification and selection criteria for the electrons is desribed in 3.4.1, The electrons candidates are reconstructed from EM calorimeter clusters with good quality particle tracks from the inner detector within excluding the transition region between the barrel and the endcap calorimeter located at ($|\eta| < 1.37$) excluding the transition region between the barrel and the endcap calorimeter located at $1.52 < |\eta| < 2.47$. Electrons coming from the bad calorimeter clusters and noise bursts are removed. The transverse energy (E_T) of the electrons is required to be greater than 25 GeV. The electrons used in this search are required to be in the tight likelihood-based identification criteria with the additional requirement that the longitudinal impact parameter z_0 must be within -0.5 mm and 0.5 mm, $|z_0 sin0| < 0.5$ mm. The electron candidates are not required to be isolated from the hadronic activity in the event. The efficiency of the tight identification criteria varies from 67% at $E_T = 25$ GeV to 99% at $E_T = 100$ GeV. The electron definition is summarized in Table 5.3.

Table 5.3: Electron object definition.

Selection	Criteria
Pseudorapidity range	$(\|\eta\| < 1.37) \quad \|\| \quad (1.52 < \|\eta\| < 2.47)$
Energy	$\mathrm{E}_T > 25$ GeV
Track to vertex association	$\|d_0^{BL}(\sigma)\| < 5$
	$\|\Delta z_0^{BL} \sin\theta\| < 0.5$mm
Identification	(*MediumLH*) TightLH
Isolation	*No isolation*

Muon

The reconstruction of a muon object in ATLAS is described in section 3.4.2. Combined muons are used in this analysis and are reconstructed using both ID and MS tracks within ($|\eta| < 2.5$). Similar to the electrons, the muon transverse momentum is required to be greater than 25 GeV throughout this thesis. To ensure that the muon candidates are associated with the correct primary vertex, requirements are applied to the impact parameters. The $|d_0^{BL}(\sigma)|$ significance is required to be smaller than 3.0. This requirement has been added in order to further reduce the QCD multi-jet background in the signal regions while preserving the signal efficiency. The muons are required to satisfy the medium identification criteria. The muons are further required to be isolated in order to suppress the background from sources of non-prompt muons. In this case, track-only tight isolation criteria are used for muons. The muon definition is summarized in table 5.4.

Table 5.4: Muon object definition.

Selection	Criteria
Pseudorapidity range	$\|\eta\| < 2.5$
Momentum	$\mathrm{p}_T > 25$ GeV
d_0 significance	$\|d_0^{BL}(\sigma)\| < 3$
$\|\Delta z_0 \sin\theta$	$\|\Delta z_0 \sin\theta\| < 0.5$mm
Identification	(*MediumLH*) Medium
Isolation	*Tight track-only*

Missing transverse energy

Any imbalance in the sum of the transverse momenta is known as missing transverse momentum E_T^{miss}. and indicates the presence of undetectable particles such as neutrinos or new particles which are not detected. The E_T^{miss} is reconstructed as the negative vector sum of the transverse momenta (p_T) of all the detected particles. A summary of MET reconstruction was given in 3.9.

5.7 Overlap removal

Besides a general selection, the overlap removal (OR) are used in this analysis that have been described in Section 3.3.9. In this search, the OR between the jets and the leptons is important. In the boosted topology, the leptons can be very close to a jet. For the muon are reconstructed using the MS and ID tracks, while the jets are reconstructed from calorimeter clusters and ID tracks. Hence if a muon and a jet are close and $\Delta R < 0.4$, it is possible to reconstruct them both correctly. In the electron channel, the problem becomes more complex since both electrons and jets are are built using calorimeter energy clusters and ID tracks. Then in this case double-counting is arising. To discard the double-counting of the calorimeter energy clusters, it is standard to require $\Delta R(el, jet) > 0.4$. But in the resolved topology the lepton responses are usually beyound the jet radius R = 0.4.

5.8 Event selection and reconstruction

The search considered in this studies has two different top-quark decay topologies. Then two different sets of criteria are imposed to seleted. A few requirements are common in both cases. The events must have at least one primary vertex candidate. It is chosen as the vertex with the highest sum of p_T^2 from tracks associated to the vertex candidate. The data events are required to pass the good run list. The events are then cleaned by removing detector noise bursts and corrupted data. All the events are expected to have at least one primary vertex. Furthermore, events are required to have passed the single-electron or a single-muon trigger. The trigger names are listed in table 5.6. Since events in both the electron and muon plus jets channels are required to have $E_T^{miss} > 20GeV$ to suppress the multijet and other non-$t\bar{t}$ events. Hence a second requirement, that also helps to reduce backgrounds, E_T^{miss} + $m_{T,w} > 60$ GeV is imposed. It relies on the transverse mass of the leptonically-decaying W -boson, defined as:

$$m_{T,w} = \sqrt{2.P_T^l.E_x^{miss}.(1 - cos(\Delta\phi_{l\nu}))} \tag{5.1}$$

Where p_T^l is the transverse momentum of the lepton, and $\Delta\phi_{l\nu}$ is the azimuthal angle difference sbetween the lepton and E_T^{miss}. In order to reduce the QCD background coming from the light flavored jets, at least one b-tagged jet is required in the selected events. For "b-tagging" can be used whether calorimeter jets with R = 0.4 or the variable radius track jets.

Classification into Boosted or Resolved selection

After satisfying the later selection criteria, the event is evaluated against the boosted selection and in case of failure against the resolved selection. The specific selection criteria are given in the following sections 5.8.2 and 5.8.3.

Table 5.5: List of the single-lepton high Level triggers (HLT) used in this analysis.

Year	Lepton channel	Trigger
2015	Electron Channel	`HLT_e24_lhmedium_L1EM20VH` OR `HLT_e60_lhmedium` OR `HLT_e120_lhloose`
	Muon Channel	`HLT_mu20_iloose_L1MU15` OR `HLT_mu50`
2016-2018	Electron Channel	`HLT_e26_lhtight_nod0_ivarloose`, `HLT_e60_lhmedium_nod0` and `HLT_e140_lhloose_nod0`
	Muon Channel	`HLT_mu26_ivarmedium` OR `HLT_mu50`

5.8.1 Common reconstruction: Neutrino

Neutrinos can not be detected directly within the ATLAS detector. The longitudinal momentum component P_z^ν, cannot be measured directly, because the longitudinal momentum 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 the sum squared of the four-momenta of the neutrino and the charged lepton from the on-shell decay of a W-boson.

$$P_W = P_l + P_\nu \tag{5.2}$$

Where P_l is the lepton 4-vector and P_ν is the neutrino 4-vector. We get the following equation from the momentum conservation and on-shell W decay approximation:

$$m_W = (P_l + P_\nu)^2 = m_l^2 + P_l.P_\nu \tag{5.3}$$

This quadratic equation can be solved for P_z^ν. if only one real solution exists, the neutrino four-momentum can be reconstructed without ambiguity given that $m_\nu = 0$. In case of no real solution a mis-measurement of E_T^{miss} is assumed and E_T^{miss} is minimally rescaled and rotated until a real solution is found. For a detailed description of sthe procedure see [**187**].

5.8.2 Boosted selection and reconstruction

Boosted $t\bar{t}$ decays are characterized by the decay products of the hadronically decaying top being collimated into a single large jet in the calorimeter. As a consequence, events in the boosted region must have at least one top-tagged large-R jet. The top-tagging is performed with a DNN top-tagger with 80% efficiency. On the leptonic side, the lepton is often near to a high p_T jet produced by b-quark decay. This is done by requiring at least one small-R jet with $\Delta R(lepton, R = 0.4jet) < 2.0$. In case of more than one jet matching this criteria, we select the highest $p_T > 300$ GeV jet and call it the selected jet. The leptonic and hadronic top-quarks are basically back-to-back in the angular planes in the signal events to ensure that, we further require the following criteria $\Delta R(R = 1.0jet, selectedR = 0.4jet) < 1.0$,. This cut will be discussed in the next section.

Boosted top candidates reconstruction

In the boosted topology, it can be simple to reconstruct the $t\bar{t}$ system. The top-tagged large-R jet is used as the hadronic top. The leptonic top is reconstructed using the lepton, the highest-p_T close-by R=0.4 jet and the neutrino calculated from MET.

5.8.3 Resolved selection and reconstruction

In the resolved top-quark decay topology, the event must contain at least four small-R jets (R = 0.4), at least one of which must be b-tagged and the χ^2 algorithm for the reconstruction of the $t\bar{t}$ system, described in section 5.8.3, has to yield a value of $log_{10}(\chi) < 0.9$, to accept the resolved selection.

Resolved top candidates reconstruction

The reconstruction of events in the resolved-topology includes many well separated small-R. The procedure used is to minimise a cost function χ^2 which depends on the assignment of jets to the top decay products. The χ^2 is defined using the 4-momenta of the lepton, neutrino, and all the R=0.4 jets of an event.

$$\chi^2 = [\frac{m_{jj} - m_{Wh}}{\sigma_{Wh}}]^2 + [\frac{m_{jjb} - m_{jj} - m_{th-Wh}}{\sigma_{th-Wh}}]^2 + [\frac{m_{bl\nu} - m_{tl}}{\sigma_{tl}}]^2 + [\frac{(p_{T,jjb} - p_{T,bl\nu}) - (p_{T,th} - p_{T,tl})}{\sigma_{p_{T,th} - p_{T,tl}}}]^2 \quad (5.4)$$

The first term represents the constraint from the dijet mass to from the hadronic W boson, while the second corresponds to term of the hadronic top decay , with the W-boson being subtracted to decouple it from the first term. The combined mass of the lepton, neutrino and a jet is constrained to match to the leptonically decaying top-quark mass. the last term correspond to the constraint of the top quark candidates to have similar transverse momenta. The expected values of the parameters m_{Wh}, m_{th-Wh}, m_{tl} as well as the width values σ_{Wh}, σ_{th-Wh} and $\sigma_{p_{T,th}-p_{T,tl}}$ are estimated from events by matching the reconstructed objects to the truth partons.

5.8.4 Optimisation selection

Besides applying preselection criteria, an optimization of the $\Delta R(Large-Jet, small-jets)$ and $\Delta R(Large - Jet, lepton)$ is performedin order to study the sensitivity of the signal selection against the background. To find the optimal value for these cuts a numerical optimization process was used. It consisted in creating a two-dimensional space of potential cut values and selecting the point with the highest binned Poisson significance σ, as given by:

$$\sigma = \sqrt{2(S_i + B_i) * \log(1 + (\frac{S_i}{B_i})) - S_i} \quad (5.5)$$

where S_i and B_i are respectively the amount of signal and background events in bin i of the reconstructed mass histogram. The optimization was performed in a mass window in 700 - 5000 GeV. Figures 5.3 show the significance study for different cut of angular separation,$\Delta R(R = 1, R = 0.4)$ and $\Delta R(R = 1, l = e/\mu)$ as a function of $t\bar{t}$ mass respectively. There is a high significant for the electron channel with $\Delta R(R = 1, R = 0.4) > 1.0$ and $\Delta R(R = 0.4, l = e/\mu) < 2.0$ (the same for $\mu channel$). Another way to optimize the cuts on variables, we once again calculated the efficiency is defined as the ratio of number of events that pass the different cuts of angular parameters $\Delta R(R = 1, R = 0.4)$ and $\Delta R(R = 0.4, l = e/\mu)$ to the total number of events in signal MC. Figure 5.4 shows the efficiency for different cuts of angular parameters as a function for varying potential resonance mass for Electron Channel. The efficiency increases steadily as decreasing value of angular parameters thus the

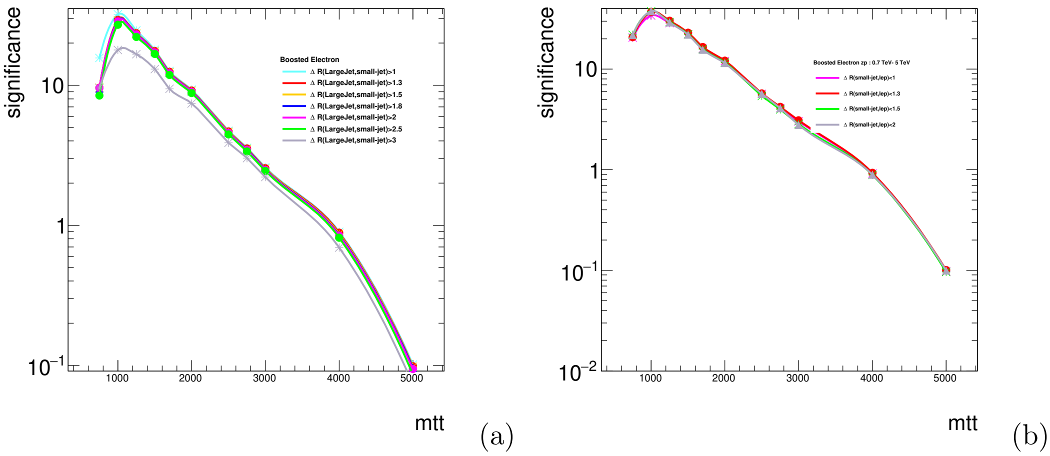


Figure 5.3: Significanse study for diferent cut of (a) $\Delta R(R=1, R=0.4)$ and (b) $\Delta R(R=1, l=e/\mu)$ for Electron channel

optimal values found in lower cut of angular parameters, $\Delta R(R=1, R=0.4) > 1$ and $\Delta R(R=0.4, l=e/\mu) < 2$.

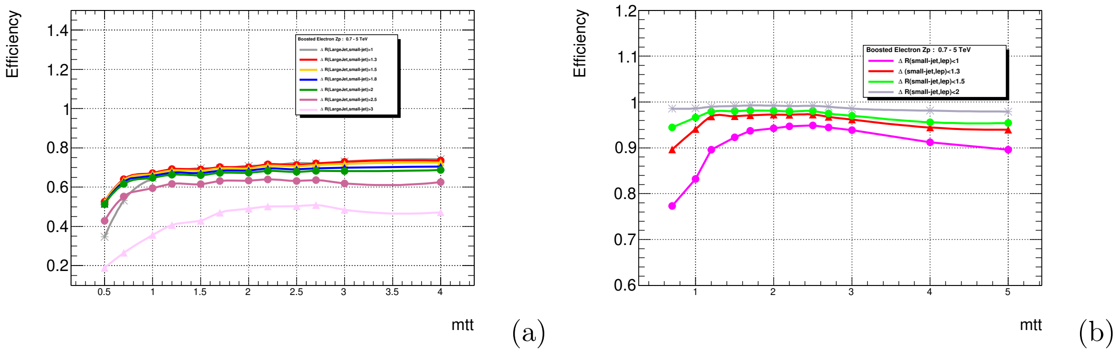


Figure 5.4: Efficiency for different cut of (a) $\Delta R(R=1, R=0.4)$ and (b) $\Delta R(R=0.4, l=e/\mu)$

5.9 Event categorization, acceptance and efficiencies

To enhance the sensitivity of the $t\bar{t}$ resonance search analysis with regard to the limit results, following the event selection described in section 5.8 splitting into four separate b-tag categories is added. The reconstructed hadronic and leptonic tops, as mentioned in section 5.8.2 as well as the b-tagged jet(s), are examined for categorisation. The events are classified into one of four b-tag categories:

- Category 0: there is no b-tagged jet matching the hadronic nor leptonic top candidates.
- Category 1: only the leptonic top candidate has a matching b-tagged jet.
- Category 2: only the hadronic top candidate has a matching b-tagged jet.

- Category 3: both top candidate have a matching b-tagged jet.
- Events in category 0 are removed from the final signal region.

A b-tagged jet is regarded a match if it can be reconstructed within R = 0.4 of the leptonic top candidate's small-R jet or within R = 1.0 of the hadronic top candidate's large-R jet.
The acceptance times efficiency (A x ϵ) for these final signal regions is given in figure 5.5, using the branching ratio for simulated BSM particles decaying into $t\bar{t}$, discussed in section 1.6, in the detector acceptance, because additional decay channels than $t\bar{t}$ are feasible. See [110] for the separated acceptances in the different lepton channels and the b-tag category divided into 1, 2, and 3.

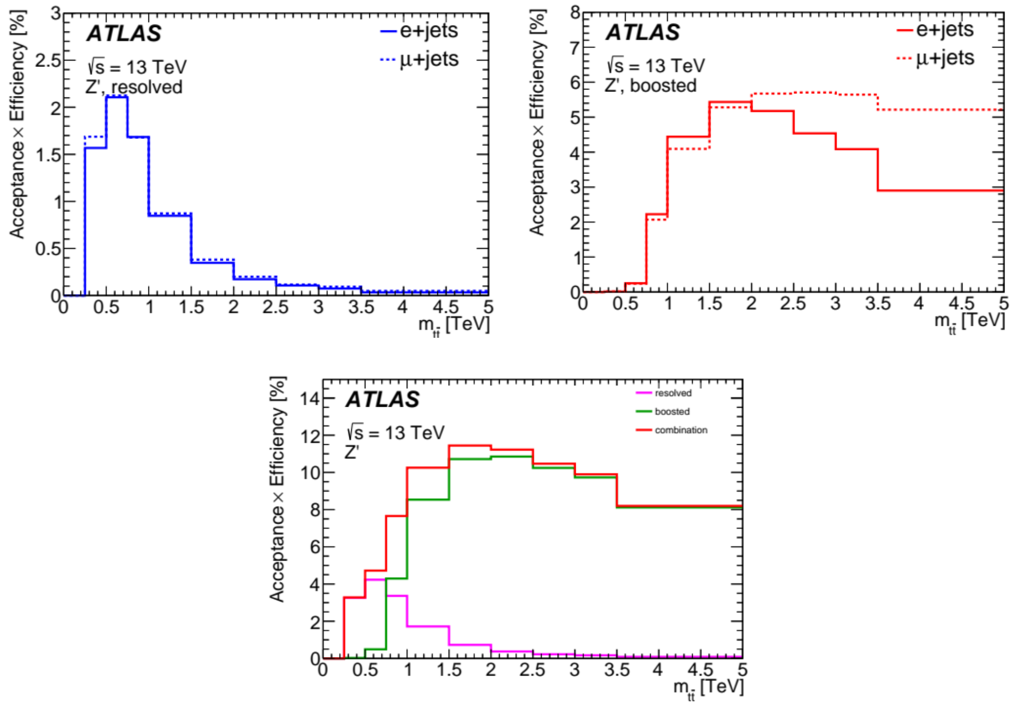


Figure 5.5: Acceptance times efficiency (A x ϵ) including the branching ratio for MC simulated BSM particles decaying into $t\bar{t}$, as a function of the $t\bar{t}$ invariant mass $m_{t\bar{t}}$ for simulated signal events.

5.10 Reconstruction of $t\bar{t}$ the invariant mass

The $t\bar{t}$ invariant mass is built from the sum of 4-vectors of the reconstructed objects identified coming from the top-quark decay products, the leptonic top, t_{lep}, and the hadronic top, t_{had}.

$$m^2_{t\bar{t}} = (t_{lep} + t_{had})^2 \tag{5.6}$$

5.11 Data driven multijet background estimation

The QCD multijet events can pass the signal selection when an addition to two reconstructed jets a fake lepton is selected. A fake lepton is a jet misidentified as an isolated lepton, or a non-prompt lepton that appears to be isolated. The data-driven matrix method is used to estimate this non-negligible source of background by emplying the prompt lepton identification efficiency (real efficiency), ϵ, as well as the non-prompt lepton misidentification probability (fake efficiency), f. These quantities are defined using tight and loose lepton definitions. Tight leptons are the leptons passing the nominal analysis selection of identification and isolation criteria. In the loose definition, the identification criteria are loosened, and the isolation requirement is not applied. The tight and loose lepton definitions are summarized in table 5.6. The number of events with leptons satisfying the loose definition, N_L

Table 5.6: List of the single-lepton high Level triggers (HLT) used in this analysis.

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

is defined as:

$$N_L = N_{prompt} + N_{QCD}, \quad (5.7)$$

where N_{prompt} (N_{QCD}) the number of events containing a prompt and a fake loose lepton. These quantities can be measured directly from data. Therefore, the number of events in the tight lepton sample, N_T can be written as:

$$N_T = \epsilon N_{prompt} + f N_{QCD}, \quad (5.8)$$

Where the real efficiency, ϵ is the probability for a real loose lepton to pass the tight lepton selection defined as **??**, and f is the probability for a fake loose lepton to pass the tight lepton selection denoted as **??**.

$$\epsilon = \frac{N^{tight}_{prompt}}{N^{loose}_{prompt}}, \qquad f = \frac{N^{tight}_{QCD}}{N^{loose}_{QCD}}, \quad (5.9)$$

The value of N_{QCD} is finally determined by combining the above two equations and eliminating the other unknown variable, N_{prompt}. The value of N_{QCD} is therefore given by:

$$f * N_{QCD} = [\frac{\epsilon - 1}{\epsilon - f} f] N_T + [\frac{\epsilon_f}{\epsilon - f}] N_A, \quad (5.10)$$

define the number of anti-tight leptons which pass the loose selection but does not pass the tight selection as $N_A = N_L - N_T$.

Real and fake rate

The fake rate f is estimated by collecting a data sample composed mainly of QCD multijet enriched control region. The Control region is defined by inverting the MET and MET+m_T ,W cut. The fakes rates for electrons vary from 10% to 85%, with the largest value occurring in the region with highest lepton p_T. This behavior is observed because the track-based lepton isolation becomes very loose for high p_T electrons. The fake rates in the muon channel vary from 4% to 74% depending on muon p_T. These rates are used to calculate the number of multijet events in the signal regions.

5.12 Distributions of Selected Kinematic Variables

This study is still ongoing, we are still blinded in the signal region and are not permitted to seeking at the $m_{t\bar{t}}$ distributions. However, we examine the data distributions of other kinematic variables associated to leptons and jets. A comparison between observed data and expected SM background has been carried out for a number of kinematic distributions are shown in this section. These cross-checks must contain variables on which selection constraints are applied, such as transverse momenta, pseudorapidities of leptons, jets (closest to lepton) or E_T^{miss} and large-R jets.

5.12.1 Boosted region

In this section, only a few representative distributions are shown in figures 5.6 for boosted topology. In particular, the distributions for lepton p_T, large-R jet p_T and jet (closest to lepton) p_T for the electron and muon channels, including the total uncertainties of the background estimate. Figure 5.6 shows that the data is consistent with the SM prediction, and there is no visible excess of events.

5.12.2 Resolved

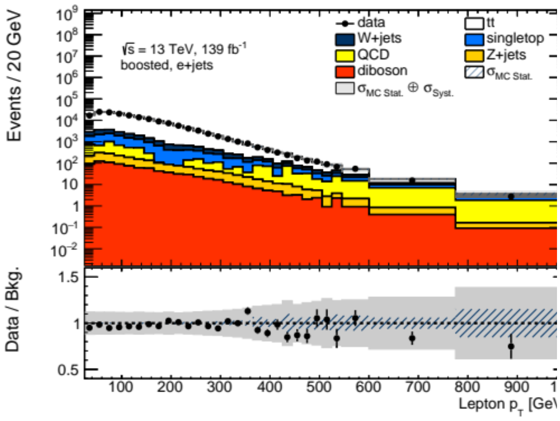
In this section, figures 5.7 show the the p_T distributions of the lepton, the E_T^{miss} and χ^2 in the electron and muon final states. A good agreement between data and expected SM background is observed within the total uncertainty of the background estimate. This demonstrates that the variables used in the definition of the signal regions are well-modelled and that the reconstruction of the $t\bar{t}$ system.

5.13 Statistical Interpretation

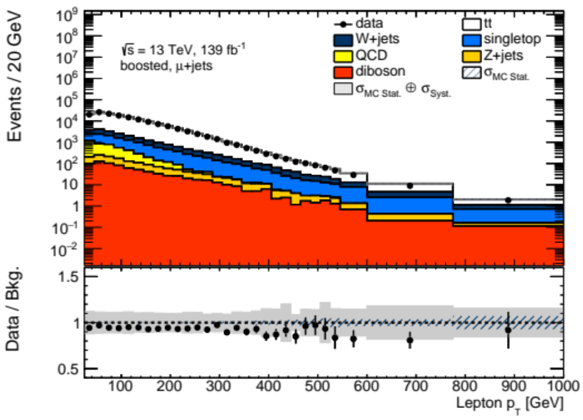
The statistical interpretation consists of two steps: first a search is performed for various signatures of new physics. Second, in the absence of any deviations incompatible with the SM-only hypothesis, upper exclusion limits on the cross-section times branching ratio are set at 95% confidence level (CL), and theoretical cross-sections are overlayed for those models to extract lower limits on the new physics resonance pole mass.

5.13.1 Search Phase

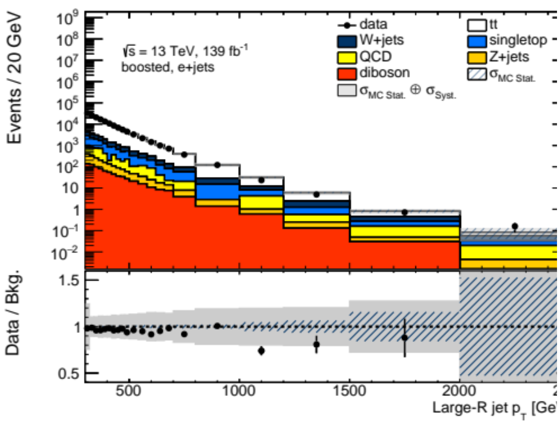
The search phase is performed using BumpHunter, a hypothesis testing tool that agnostically searches for local data excesses or deficits in the data compared to the expected background. The excess/deficit can span any invariant mass range from two bins, to half of the bins in the distribution. The process is performed iteratively comparing data and the expected background in a series of sliding windows of variable size. For each window the Poisson probability is calculated, and the smallest p-value from the scan is recorded. In each window. Systematic uncertainties and trial factors for this procedure are all taken into account. In each window, i, data



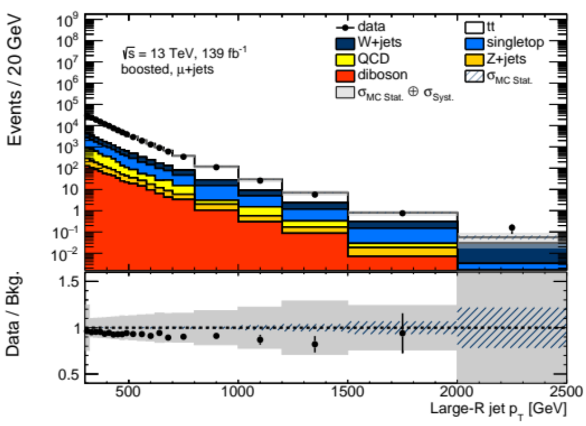
(a) e+jets channel, boosted topology



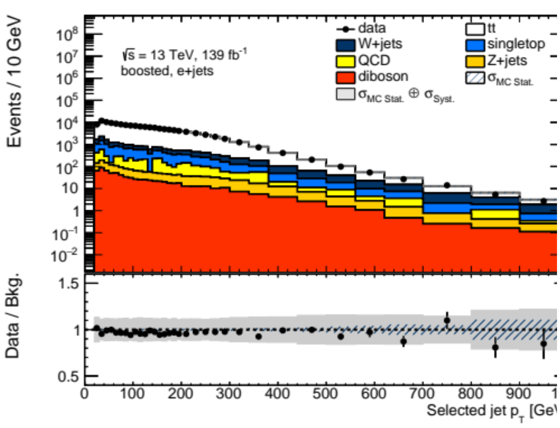
(b) μ+jets channel, boosted topology



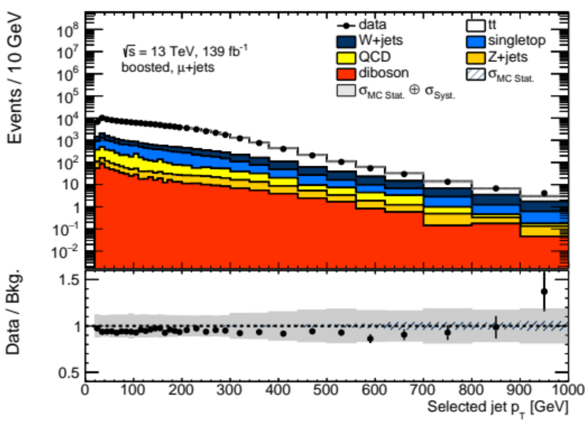
(c) e+jets channel, boosted topology



(d) μ+jets channel, boosted topology



(e) e+jets channel, boosted topology

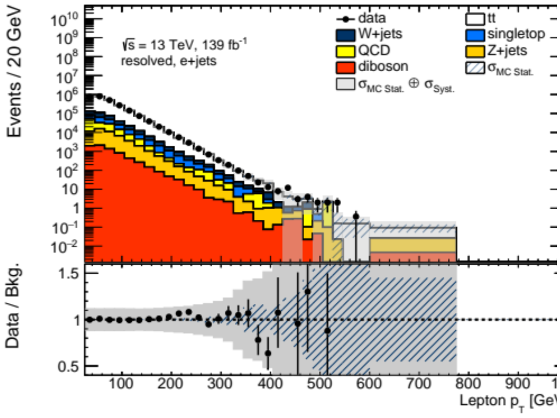


(f) μ+jets channel, boosted topology

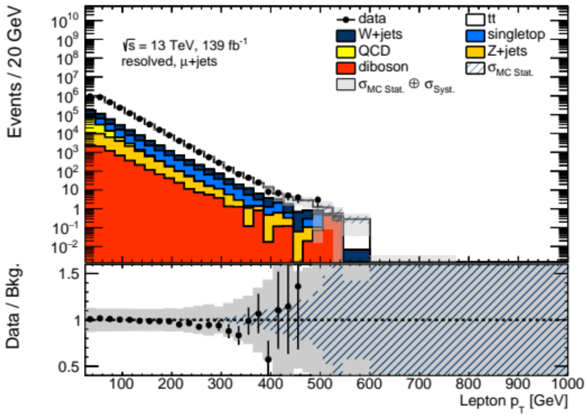
Figure 5.6: Distributions of kinematic variables for events passing the boosted-topology selection P_T, the E_T^{miss} distribution and the distributions for in the (a, c, e) boosted electron and (b,d,f) boosted muon.

count, d_i, and background yield, b_i, the Poisson probability, $P(d_i, b_i)$ is defined as:

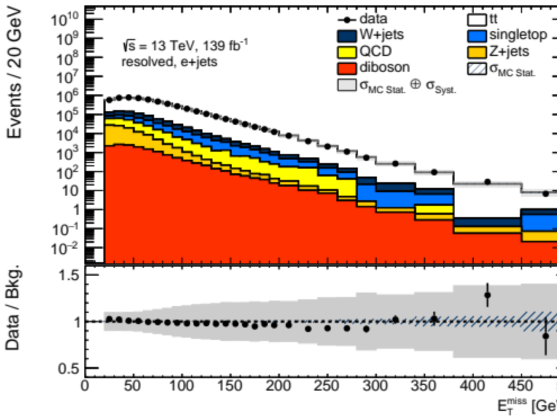
$$P(d_i, b_i) = \begin{cases} \Gamma(d_i, b_i) = \sum_{n=d_i}^{\infty} \frac{b_i^n}{n!} e^{-b_i} & \text{if } d_i \geq b_i \\ 1 - \Gamma(d_i + 1, b_i) & \text{if } d_i < b_i \end{cases} \tag{5.11}$$



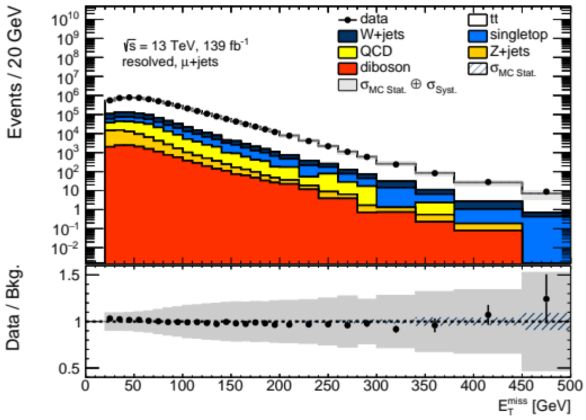
(a) e+jets channel, resolved topology



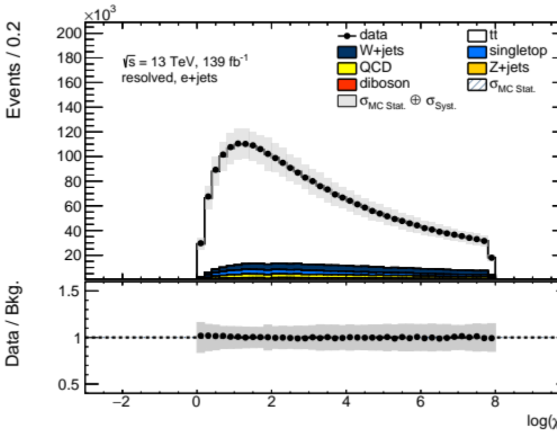
(b) μ+jets channel, resolved topology



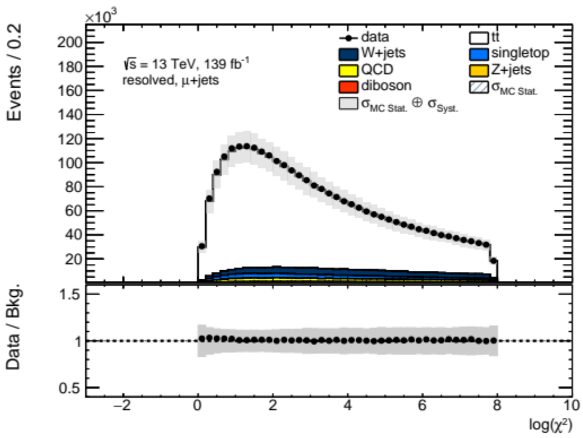
(c) e+jets channel, resolved topology



(d) μ+jets channel, resolved topology



(e) e+jets channel, resolved topology



(f) μ+jets channel, resolved topology

Figure 5.7: Distributions of kinematic variables for events passing the resolved-topology selection P_T, the E_T^{miss} distribution and the distributions for in the (a, c, e) resolved electron and (b, d, f) resolved muon.

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 windows, $P_i^{\min}$ corresponds to the most interesting (discrepant with respect to the background expectation) window. To include the systematic uncertainties, the definition of $P(d_i, b_i)$ is extended to:

$$P(d_i, b_i) \to P(d_i, b_i + \lambda_i \theta_i) e^{(-\frac{\lambda^2}{2})} \tag{5.12}$$

where θ_i is the total combined systematic uncertainty for the considered window, and λ_i is a real number (between -8 and 8) that maximizes the value of $\frac{(b_i+\lambda_i\theta_i)^{d_i}}{d_i!} e^{-(b_i+\lambda_i\theta_i)} e^{(-\frac{\lambda^2}{2})}$. The BumpHunter test statistic t is computed as:

$$t = \begin{cases} 0 & \text{if } d_i \leq b_i \\ -\log P_i^{\min} & \text{otherwise} \end{cases} \tag{5.13}$$

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

$$p - \text{value} = \frac{\int_{t_{obs}}^{\infty} f(t)}{\int_0^{\infty} f(t)} \tag{5.14}$$

where $f(t)$ is the distribution of the test statistic values from the pseudo-data and t_{obs} is the test statistic obtained from data. The p-value is the probability of observing a signal-like excess, at least as significant as the one observed in the data, given the background-only hypothesis is true. 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, i.e. the deviation would be very large. p-values are often also converted into a number of standard deviations to denote significance (σ). In particle physics a 3 σ deviation is considered as "evidence" of new physics, and 5 σ is the required level to claim a discovery. For this search 4 distinct signal regions were studied, with resolved and boosted topologies considered separately for both e+jets and μ+jets channels, according to the description in Section 5.8. The most significant excess/deficit along with their p-value is given for each signal region, and as a second step the channels are combined to provide a p-value for the most significant combined excess/deficit.

5.13.2 Exclusion Limits

In the absence of a significant excess in the data compared to the background expectation, exclusion limits are set on the cross-section times branching ratio at 95% CL using a frequentist approach called the the CLs method. A tool called TRex-Fitter, together with RooStats and RooFit were used to perform the limit setting procedure. The likelihood for a particular signal assumption is defined as:

$$L(\mu, \Theta) = \prod_{i=0}^{\text{channels,bins}} \frac{e^{-\mu a_{Z',i}(\Theta)\sigma_{Z'} + b_i(\Theta)} (\mu a_{Z',i}(\Theta)\sigma_{Z'} + b_i(\Theta))^{D_i}}{\Gamma(D_i + 1)} C(\Theta) \tag{5.15}$$

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

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

$$\Lambda(\mu) = \frac{L(\mu, \hat{\hat{\Theta}}(\mu))}{L(\hat{\mu}, \hat{\Theta})}, \tag{5.16}$$

Where the single circumflex indicates the unconditional maximum likelihood estimate of a parameter, while the double circumflex indicates the maximum likelihood estimate assuming a specific value of μ. Assuming that the test statistic $-2\ln(\Lambda(\mu))$ is distributed according to a χ^2 distribution with one degree of freedom, the variation on the logarithm of the likelihood can be used to set a 95% CL on the upper limit of the signal production cross-section times branching ratio, and to estimate the impact of the systematic uncertainties in the discovery significance. Nuisance parameters (systematic and statistical) which have less than a 0.1% effect on the signal or background normalisation and shape are pruned from the statistical interpretation to improve computing performance with a negligible effect on the result. Although these thresholds may appear rather low, they are justified by the fact the individual bins of the signal regions can contain more than one thousand events. Due to some nuiance parameters suffering from statistical fluctuations which could destabilise the fit, the underlying histograms undergo a smoothing procedure as follows:

- for every bin of a nuisance parameter, define $\delta M = \sqrt{\delta S^2 + \delta N^2}$ (resp. $\delta M = max(\delta S, \delta N)$), the estimate of statistical error in the case where the nominal bin (N) and the varied bin (S) are statistically uncorrelated (resp. correlated).
- for each pair of neighbour bins (i, j), define $X_{i,j} = |\frac{S_j - N_j}{N_j} - \frac{S_i - N_i}{N_i}|$ and $\delta X_{i,j} = \sqrt{\frac{\delta M_j^2}{N_j^2} + \frac{\delta M_i^2}{N_i^2}}$ (error on relative systematic assuming fixed nominal)
- While there are neighbour bins with $X_{i,j} < \delta X_{i,j}$, merge the bins i and j with the largest $\delta X_{i,j}/X_{i,j}$
- As for TRexFitter, use `353QH` on $(S - N)/N$, four times using `TH1::Smooth`

5.13.3 Limit setting

The upper limits derived from this search is optimized for narrow-width Z'_{TC2} with . Expected 95% CL upper limits are calculated for two cases, with all the systematic and statistical uncertainties and with the statistical uncertainty only. Limits are calculated starting from starting from Z_{TC2} 500 GeV to 5 TeV. The limits are calculated for the following points using MC signal simulations:

- 500, 750, 1000, 1250, 1500, 1750, 2000, 2250, 2500, 2750, 3000, 4000, 5000 GeV.

The expected limits on Z'_{TC2} signal cross-section times branching fraction are shown in Figure 5.8. The 95% CL expected limit on Z'_{TC2} (/m = 1.2%) mass is around 3.6 TeV whereas the stat-only limit is around 3.8 TeV.

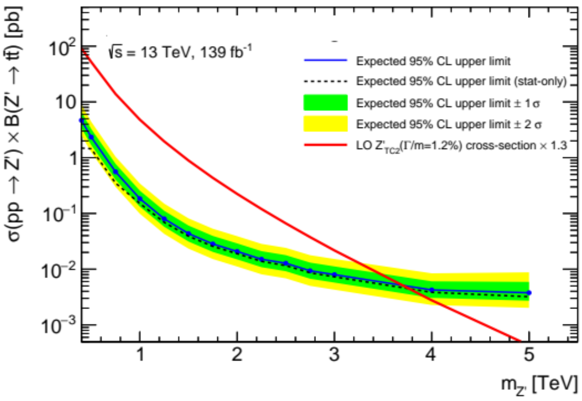


Figure 5.8: The expected cross-section 95% CL upper limits on Z'_{TC2} signal

5.14 Conclusion

A search for $t\bar{t}$ resonances in the single lepton final state is performed, where one top-quark decays leptonically while the other top-quark decays hadronically. In this search, we are seeking for a narrow resonance employing the full Run-2 data (139 fb^1) collected by the ATLAS detector during, the year 2015-2018. The analysis requires various physics objects due to the complicated $t\bar{t}$ final state. This search focuses on both the boosted and the resolved top-decay topologies. The main discriminant is the invariant mass of the $t\bar{t}$ system, where the signal is expected to show up as a bump. As a result, the expected cross-section limits improved by almost a factor of 1.5. The expected mass limit (1.2% width Z'_{TC2}) including only the statistical uncertainties improved from 3 TeV to 3.8 TeV based on the current optimization. This chapter presented the expected sensitivity only on the Z'_{TC2}. model; all the other models will be used in the future.

Chapter 6

Search for high-mass resonances in photon plus missing transverse momentum signatures in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector

6.1 Introduction

This chapter based on the reinterpretation of the ATLAS mono-photon results [**37**] using the RECAST strategy (Request Efficiency Computation for Alternative Signal Theories) [**138**], a framework designed to preserve the same topology of the original mono-photon analysis, reusing the same phase-space and background estimates as well as systematic uncertainties, in the context of searching for dark matter production from heavy Higgs bosons decay to a photon and a massless dark photon. The introduced signal model includes both the gluon-gluon Fusion (ggF) and the VBF production modes of the heavy Higgs bosons as presented in the Feynman diagrams shown in Figure 6.1. To improve the sensitivity of this search, different signal regions (SR) are considered based on the missing transverse momentum E_T^{miss} selections ranging from 200 GeV to 375 GeV, being the primary observable of this search.
The results of this reinterpretation are presented in terms of 95% CL exclusion limits on the $\sigma \times BR(H \to \gamma\gamma_d)$, considering separate ggF and VBF production modes as well as the combined $ggF + VBF$ production for different heavy Higgs boson masses: 400, 600, 800, 1000, 1500, 2000, 2500 and 3000 GeV.

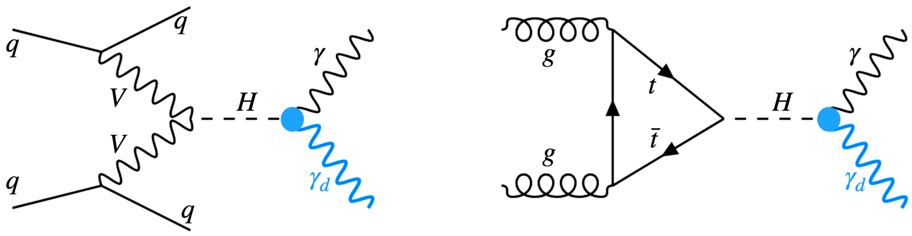


Figure 6.1: Feynman diagrams for H→ $\gamma\gamma_d$ in ggF (right) and VBF (left) production modes.

6.2 Reinterpretation

Experimental searches for BSM physics could be sensitive to a wide variety of models. However, in most cases, only a few models are taken into account when interpreting experimental search results. The number of models that can be evalueted for interpretation only increases with time, and searches are a significant investment of experimental physicists time and resources. To understand the implications of LHC data for the widest feasible set of theories, experimental analyses should be reinterpretable in terms of theories not included in the original analysis publication.

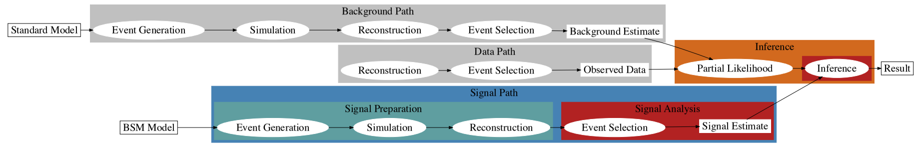
Reinterpretations allow phenomenologists to provide thorough feedback on the original analyses and to better suggest promising alternate destinations for future experimental analysis. Furthermore, Reinterpretability of a data analysis preserves its shelf life and boosts its scientific significance through subsequent work based from it. The preparation of the material to permit reinterpretation invariably facilitates internal reproduction for experimental collaborations, such as when a student responsible for an analysis has transferred to a different position or field. Finally, the availability of such data aids in meeting funders increasingly stringent expectations for the openness and reproducibility of publicly-funded research. A reinterpretation is the process of interpreting an existing experimental result in the context of an alternative BSM physics scenario by recording the data and backgrounds from an analysis and comparing them to a new signal model. Reinterpretations do not give any possibility for making a discovery because the data and backgrounds stay unchanged, but rather allow limitations to be imposed on models that were not considered in the original search.

6.2.1 Recast framework

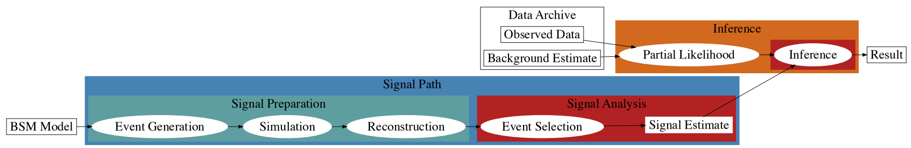
Recast [70] is a software framework that aims to provide the necessary infrastructure to organize systematic reinterpretations of analysis performed in high-energy physics experiments. It is designed to reuse background estimates and observations in the data from searches to test alternative signal hypotheses with only a fraction of the effort required for a new search. The strategies used by Recast to preserve the analysis have also greatly influenced CERN's overall preservation of the analysis [71].

Generally a search analysis tries to identify an effective phase space region for signal model while rejection as much SM background as possible. Defining such a region through the construction of appropriate observables and developing a selection of events constitutes a large part of the analysis work. The search region is examined for reconstructed data from the experiment. In addition, the expected contribution from SM background and BSM physics processes must be determined, these processes must be simulated, reconstructed, and passed through the event selections to estimate their expected contribution to the total event rate. Also, the response to systematic variations in simulation and reconstruction algorithms must be determined. All these inputs then forms the basis of a statistical analysis procedure. This general analysis workflow is shown diagrammatically in Figure 6.2. From the above it is clear that it is unfeasible to construct optimised analyses for the wide range of physics models discussed in the context of the LHC. However, the region Ω of an analysis, optimised for its original model, may be still offer good sensitivity to other models even if it is not optimal. This forms the basis of the reinterpretation approach. The resources required for a reinterpretation far fewer than those necessary for the construction of a new analysis. As the event selection of an analysis is fixed, the background estimates and observed data distributions for the corresponding phase space region Ω do not change in the context of a reinterpretation and only the distributions due to a new signal model must be derived. Once the new signal contribution has been derived, an adapted statistical analysis, which combines the newly derived signal estimate with the archived data measurement and background estimate may then determine whether the model is viable in light of the data observed in the analysis. At a high level of abstraction a reinterpretable analysis provides a *interface* which takes a model and its parameters as inputs and yields inference result. The interface is general enough such that it does not depend on the implementation details of the analysis.

Reinterpretation through the original analysis, compared to e.g. approximate third-party implementations, provides a more accurate reinterpretation result. It allows a faithful selection of events in addition to a more accurate statistical interpretation of the result in the context of a new model. With this in mind, one important point to consider is the possibility of the new signal having an appreciable yield across multiple regions of the analysis beyond the intended signal region in the original analysis. This may affect the background estimates of the analysis and in turn, compromise the validity of the reinterpretation or influence the quoted sensitivity from it. Reinterpretation with the full analysis enables for a complete evaluation of the magnitude of such signal contributions and even allows to incorporate them consistently through a joint statistical analysis, thereby taking them into account properly. However, it is important to note that prior to performing such a statistical interpretation, the extent of the signal contamination should be inspected carefully to evaluate its effect on the final result. With this in mind, there is the possibility of the new signal model having an appreciable yield across multiple regions of the analysis beyond the intended signal region in the original analysis. This may have an impact on the analysis's background estimates, compromising the validity of the reinterpretation or influencing the reported sensitivity from it. Reinterpretation with the full analysis allows for a thorough evaluation of the magnitude of such signal contributions, as well as the ability to consistently include them through a joint statistical analysis, therefore fully accounting for them. It should be noted, however, that before performing such a statistical interpretation, the extent of the signal contamination should be carefully examined to evaluate its impact on the final result.



(A full analysis workflow: Three main data processing paths exist: the estimation of the SM background as well as the BSM signal and the determination of the data counts observed in a phase space region Ω of the ATLAS data. The background path includes simulation of the hard scattering process and subsequent hadronisation (Generation), detector simulation (Simulation), physics object reconstruction (Reconstruction) and event selection to obtain a background estimate. The signal path follows the same set of steps to obtain the signal estimate. The data path begins with the physics object reconstruction using recorded data, followed by the event selection to obtain the observed data. The three are then combined into a statistical model, from which an inference result, such as a limit at a certain confidence level, is derived.)



(A reinterpretation workflow: As the background and data yields are already computed from the original analysis, they can be read from an archive and only the signal path of the workflow must be executed, followed by a statistical analysis.)

Figure 6.2: The computational workflow required for a full analysis compared to that of a reinterpretation.

6.2.2 Analysis preservation in RECAST

Systematic reinterpretation depends on preserving analyses in a manner that permits new collaborators to re-execute them autonomously. This is accomplished in Recast through using explanatory determinations and modern cloud computing technologies. A catalogue of preserved analyses that can be re-executed by the corresponding experimental collaborations for a new signal model may then provide the basis of a streamlined reinterpretation program within the high energy physics community.

To be reusable, an analysis software must be preserved in a portable and deployable format on a wide range of computer infrastructure. This is generally accomplished through the preservation of software in the form of Docker container images. [108,109] Container images contain the whole root filesystem (i.e. all software dependencies) required for the execution of a particular process and strike an appropriate balance between isolation and efficient deployment on various infrastructure. ATLAS delivers appropriate container base images that contain both the ATLAS analytic software release and its dependencies. [110] The analysis release includes the software libraries required to read the ATLAS data files and implement calibrations of the selected physics objects, making it a critical dependency of the user analysis code. The Docker image for the analysis is then constructed on top of this basic image, and includes the code needed for event selection and any additional processing steps.

Basically preserving an analysis software environment is insufficient to allow it to be re-executed. A description is also required to explain the proper use of the command line programs available in these setups. Because reinterpretation is concerned with not just reproducing a given result but also reusing software to derive new results, this description must be parameterised to allow the definition of new input data. As a result, the individual processing steps of the analysis must be job templates with parameters specified in a workflow definition rather than static reproductions of the commands executed to produce the original result.

When given a new input, the templates may be transformed into specific job descriptions. A job description must give three essential pieces of information: a description of the processing step, A job description therefore needs to provide three basic pieces of information: a description of the processing step, often in the form of a command line string; the container image to be used for the process; and a identification of the relevant outputs in order to assemble multiple steps into a workflow. As this scenario, yadage was utilized, with both the job templates and the workflow structure stored in YAML documents, a format selected for its ease of usage and ubiquity. The parameterised job templates specify the instructions needed to execute the user analysis code for various tasks such as physics object calibration, event selection, systematic uncertainty estimate, and statistical inference. By describing individual phases using templates and their input parameters, the workflow orchestrates the separate steps of the analysis. In summary, a reinterpretation using Recast proceeds as follows:

- The original analysis is preserved using Docker containers for the software, job templates for individual processing steps, and a workflow to link the steps together.
- A new signal model is proposed for a reinterpretation, and the central ATLAS

data management system generates the sample required by the analysis.

- The preserved analysis is re-executed on the new signal sample, and a new limit is produced.

6.3 Mono-photon signature

The mono-photon signature consists of events characterized by a large missing transverse momentum recoiling against a high-pT photon, no additional leptons and at most one jet. As opposed to other mono-X signatures, the mono-photon final state is characterized by a low level of background coming from SM processes. Multijet processes, top quark and W + jets and Z + jets processes, which are abundant backgrounds for many searches at the LHC, are strongly reduced by the large E_T^{miss} selection, by the requirement of a highly energetic photon and by the veto on extra jets. Besides the irreducible Z$(\nu\nu)\gamma$ background, an important source of background consists in events with a W boson decaying leptonically, produced in association with a photon, where the lepton is missed.

6.3.1 Event selection and analysis region definition

All the events in the analysis must satisfy beam, detector and data-quality criteria. They are selected by a trigger requiring at least one photon candidate with transverse energy E_T^{γ} above a threshold of 140 GeV and passing 'loose' identification requirements [72]. Events are required to have at least one candidate for primary vertex, defined as the vertex with the highest sum of the squared transverse momenta of its associated tracks and reconstructed from at least two good-quality tracks [73] with $p_t > 0.5$ GeV. Events are removed if they contain a poor-quality photon or jet arising from instrumental problems or non-collision background [68]. The Signal Region (SR) is defined by requiring events to have a selected leading photon that satisfies the criteria defined in Section 6.3.1 and has $E_T^{\gamma} > 150\ GeV$. The 'photon pointing', $|\Delta z_\gamma|$, defined as the separation, measured along the beam line, between the extrapolated origin of the photon and the position of the event's identified primary vertex is required to be smaller than 250 mm. This criterion suppresses the non-collision background in the data-driven method used to estimate the contribution from events in which jets are misidentified as photons. To ensure that the leading photon and E_T^{miss} do not overlap in the transverse plane, $\Delta\phi(\gamma,\ E_T^{miss}) > 0.4$ is required. In order to reduce the number of background events characterised by E_T^{miss} arising from poorly reconstructed physics objects, events are required to have $E_T^{miss} > 200\ GeV$ and E_T^{miss} significance > 8.5, the latter is calculated as $|\vec{E}_{\mathrm{T}}^{\mathrm{miss}}|^2/\,(\sigma_{\mathrm{L}}^2\,(1-\rho_{\mathrm{LT}}^2))$, where σ_{L} is the total standard deviation in the longitudinal direction to the E_T^{miss} and ρ_{LT} is the correlation factor of the longitudinal L and transverse T measurements [**ATLAS-CONF-2018-038**]. To suppress multi-jet background, events with more than one selected jet or with $\Delta\phi(jet,\ E_T^{miss}) < 0.4$ are rejected. Events are required to have no candidate electrons, muons or hadronically decaying τ-leptons passing the requirements described in Section 6.3.1. This lepton veto mainly rejects W/Z events. To improve the sensitivity of the analysis, seven SRs are defined corresponding to different E_T^{miss} ranges:

four inclusive (SRI1–SRI4) and three exclusive (SRE1–SRE3). Table 6.1 summarises the SR definitions and the selection criteria described above.

Table 6.1: Selection criteria for the SRs.

Event cleaning	Detector quality conditions and primary vertex						
Leading γ	$E_T^{\gamma} > 150\ GeV$, $\lvert\eta\rvert < 1.37$ or $1.52 < \lvert\eta\rvert < 2.37$, tight, isolated, $\lvert\Delta z_\gamma\rvert < 250$ mm, $\Delta\phi(\gamma,\ E_T^{miss}) > 0.4$						
E_T^{miss} significance	> 8.5						
Jets	0 or 1 with $p_t > 30\ GeV$, $\lvert\eta\rvert < 4.5$ and $\Delta\phi(jet,\ E_T^{miss}) > 0.4$						
Leptons	veto on e, μ and τ						
	SRI1	SRI2	SRI3	SRI4	SRE1	SRE2	SRE3
E_T^{miss} [GeV]	> 200	> 250	> 300	> 375	200–250	250–300	300–375

6.3.2 Background estimates and results

The SM background in the various SRs arises from a number of processes that generate real photons and from events in which one or more energetic jets or electrons are misidentified as photons. The latter are estimated through the use of control samples including jets or electrons, scaled by misidentification rates determined from data. The background with real photons is dominated by the $Z(\rightarrow \nu\nu)\gamma$ process, secondary contributions come from $W(\rightarrow l\nu)\gamma$ and $Z(\rightarrow ll)\gamma$ production with unidentified electrons, muons or with hadronically decaying τ leptons and from γ+jets events. These contributions are obtained using Monte Carlo (MC) simulations constrained by observed event counts in dedicated control regions (CRs) through the estimation of normalisation factors. Control regions are built by inverting one or more of the selection criteria used to define the SRs, allowing one of the background processes to become dominant but otherwise kinematically similar to the given SR. The contribution of these backgrounds is determined separately for each SR, as defined in Table 6.1, with a maximum-likelihood fit, referred to as the 'background-only fit'. This procedure constrains the normalisation of the dominant backgrounds to the observed event yields in the associated CRs, assuming that no signal is present in the CRs and including the background contribution from the misidentification of electrons and jets as photons estimated with data-driven techniques. The inputs to the fit for each SR include the number of events observed in its associated CRs and the number of events predicted by simulation in the SR and CRs for each background process. The latter are described by Poisson statistics. The systematic uncertainties are included in the fit as nuisance parameters, modelled by Gaussian distributions with widths corresponding to the sizes of the associated uncertainties. Using a simultaneous fit technique allows a straightforward combination of multiple CRs and permits a coherent treatment of the correlation of the systematic uncertainties across the different regions. The product of the various probability distributions forms the likelihood, which the fit maximises by adjusting the background normalisation and the nuisance parameters.

Since the shape of BSM signal models over multiple E_T^{miss} bins is different from the background prediction, a 'simplified shape fit' is applied in the CRs associated to the exclusive SRs plus the inclusive SR corresponding to the highest E_T^{miss} range (SRI4) shown in Table 6.1. This allows normalisation factors for the background contributions in each E_T^{miss} bin to be extracted by exploiting the E_T^{miss} shape information and it is used to set exclusion limits on the considered models when no excess is found in the data.

The background-only simultaneous fit is performed to evaluate the SM background expectations in all regions of the analysis. The event yields in data are consistent with the SM expectations within uncertainties in all signal regions. The observed number of events and the expected SM background contribution resulting after performing the 'simplified shape fit' are reported in Table 6.2 [**MonophotonATLAS:2020**]. The values of the normalisation factors for $W/Z+\gamma$ and $\gamma + jets$ backgrounds are also shown.

Table 6.2: Observed and expected yields from SM backgrounds obtained from the 'simplified shape fit' [**MonophotonATLAS:2020**]. The normalisation factors obtained from the fit are also shown at the bottom of the table. The uncertainty includes both the statistical and systematic uncertainties. The individual uncertainties can be correlated and do not necessarily add in quadrature to equal the total background uncertainty.

	SRE1	SRE2	SRE3	SRI4
Observed events	3023	1164	679	427
Expected SM events	3070 ± 130	1182 ± 75	680 ± 53	448 ± 42
$Z(\rightarrow \nu\nu)\gamma$	1910 ± 110	758 ± 65	468 ± 49	306 ± 40
$W(\rightarrow \ell\nu)\gamma$	394 ± 22	159 ± 15	71.0 ± 8.2	56.7 ± 7.1
$Z(\rightarrow \ell\ell)\gamma$	33.2 ± 2.4	9.32 ± 0.89	4.26 ± 0.48	2.69 ± 0.37
γ + jets	87 ± 35	11.9 ± 4.8	2.7 ± 1.1	3.0 ± 1.2
Fake photons from e	511 ± 48	188 ± 18	100.9 ± 9.5	59.7 ± 5.6
Fake photons from jets	136 ± 28	56 ± 29	33 ± 16	20 ± 11
$k_{Z\gamma}$	0.99 ± 0.08	0.89 ± 0.09	0.90 ± 0.11	0.86 ± 0.12
$k_{W\gamma}$	0.81 ± 0.09	0.84 ± 0.11	0.74 ± 0.11	0.85 ± 0.13
$k_{\gamma+\text{jets}}$	0.82 ± 0.21			

The total uncertainty ranges from 3.5% to 9.5% depending on the SR. The purely statistical uncertainty is dominant, varying from 2.4% to 8.5% and driven by the statistical precision from the CRs adopted to constrain the normalisation of the leading $Z(\rightarrow \nu\nu)\gamma$ background. The relative impact of each source of systematic uncertainty on the total SM background estimates is summarised in Table 6.3 [**MonophotonATLAS:2020**]. A large experimental uncertainty is related to jets misidentified as photons and varies from 1.4% to 4.1%. The impact of the un-

Table 6.3: Summary of the uncertainties (%) in the background estimate for inclusive SRs after the background-only fit and for exclusive SRs after the 'simplified shape fit' [**MonophotonATLAS:2020**]. The individual uncertainties can be correlated and do not necessarily add in quadrature to equal the total background uncertainty.

	SRI1 [%]	SRI2 [%]	SRI3 [%]	SRI4 [%]	SRE1 [%]	SRE2 [%]	SRE3 [%]
Total (statistical+systematic) uncertainty	3.5	4.8	6.2	9.5	4.3	6.3	7.8
Statistical uncertainty	2.4	3.6	5.3	8.5	3.3	5.0	6.7
Fake photons from jets	1.4	2.5	2.8	4.1	1.4	3.6	3.7
Jet energy scale/resol	1.6	2.2	2.5	2.7	2.2	2.2	2.3
Fake photons from electrons	2.1	2.0	2.0	2.0	2.3	2.3	2.1
Electrons reco/id/isolation eff.	1.0	1.2	1.3	1.4	1.0	1.0	1.2
Electron/photon energy scale/resol	0.8	0.6	0.7	0.9	0.9	0.9	0.6
Muon reco/id/isolation eff.	0.7	0.8	0.9	1.0	0.6	0.7	0.9
E_T^{miss} soft term scale/resolution	0.1	0.4	0.7	0.9	0.5	0.2	0.5
Theoretical $W/Z+\gamma$, $\gamma+jets$	0.5	0.5	0.5	0.5	0.5	0.5	0.5
$\langle\mu\rangle$ reweighting in MC simulation	0.3	0.2	0.2	0.2	0.3	0.2	0.2

certainty in the jet energy scale and resolution varies from 1.6% to 2.7%. The uncertainty related to electrons misidentified as photons varies between about 2% and 2.3%. Other experimental systematic uncertainties related to electrons, photons, muons and the E_T^{miss} 'soft term' have a relative impact below 1.5% in all SRs. Theoretical systematic uncertainties in the $W/Z+\gamma$ and $\gamma+jets$ MC estimates have an impact of 0.5% in all SRs.

6.4 Description of the signal model

In the search for high mass resonances from heavy Higgs bosons in $\gamma+E_T^{miss}$ final states carried out in this note, the phase-space defined by the reinterpreted mono-photon analysis leads to dominant contributions from both ggF and VBF production processes as shown in Figure 6.1. For both these processes, dedicated MC generators are used to simulate and model the signal events, at a centre-of-mass energy of 13 TeV following the configurations shown in Table 6.4. The heavy Higgs boson decay to a photon and a massless dark photon $H\to\gamma\gamma_d$ is handled assuming the Hidden-Valley dark photon scenario [**Carloni_2010**] implemented in PYTHIA since 8.150 version.

The signal events produced via ggF are generated using POWHEG[v2 NNLOPS] [**Nason:2004rx**, **Frixione:2007vw**, **Alioli:2010xd**, **Hamilton:2013fea**, **Hamilton:2015nsa**] with the CT[10] set of parton distribution functions (PDFs), calculated at NNLO accuracy in α. The simulation was interfaced with PYTHIA[8.306] using the CTEQ[6L1] PDF set and the AZNLO underlying-event tune to handle the heavy Higgs boson decay, the parton shower and the non-perturbative hadronization effects. The ggF prediction from the MC samples is normalised to the next-to-NNLO cross-section in QCD plus electroweak corrections at NLO [**deFlorian:2016spz**, **Anastasiou:2016cez**, **Anastasiou:2015ema**, **Dulat:2018rbf**, **Harlander:2009mq**, **Harlander:2009bw**,

Harlander:2009my, **Pak:2009dg**, **Actis:2008ug**, **Actis:2008ts**, **Bonetti:2018ukf**]. In addition, a filter is applied at the generator level requiring events to contain one photon with a $p_T > 120\ GeV$ and $|\eta| < 2.5$.

The other signal production process is through VBF. Signal events from this production mode were generated with matrix elements calculated using POWHEG[v2] [**Campbell:2012a**, **Luisoni:2013cuh**] at NLO accuracy in α, using the $PDF4LHC15\_nlo\_30\_pdf$ as set of PDFs [**Butterworth:2015oua**]. In addition, PYTHIA[8.306] [**Skands:2014pea**] was used with the dipole recoil shower variable [**Cabouat_2018**] turned on, to perform the heavy Higgs boson decay as well as the parton shower and the non-perturbative hadronization effects with the CTEQ[6L1] PDF set and the AZNLO underlying-event tune. Moreover, at the generator level, a filter is applied requiring events with E_T^{miss} higher than 75 GeV.

The samples of both ggF and VBF processes were generated with the width set to the SM value of 4 MeV [**HiggsCrossSectionYR4**] and the complex pole scheme [**Goria:2011wa**] turned off. A total of eight MC samples per production process were generated with heavy Higgs boson masses equal to 400, 600, 800, 1000, 1500, 2000, 2500 and 3000 GeV.

The generated events were processed through a simulation [**SOFT-2010-01**] of the ATLAS detector geometry and response using GEANT [**Agostinelli:2002hh**], and through the same reconstruction software as the collected pp collision data. Corrections were applied to the simulated events so that the particle candidates' selection efficiencies, energy scales and energy resolutions match those determined from data control samples. The simulated samples are normalised to the total integrated luminosity assuming their cross-sections, computed to the highest order available in perturbation theory as given in Table 6.5 [**HiggsCrossSectionYR4**]. The detector response was simulated using a fast parameterised simulation of the ATLAS calorimeters [**Atlfast3**] and the full GEANT simulation is used for the other sub-detectors.

Table 6.4: The configurations used for event generation of the signal processes. The matrix element (ME) order refers to the order in the strong coupling constant of the perturbative calculation. The tune refers to the underlying-event tune of the parton shower.

Process	Generator	ME Order	Parton Shower	PDF	Tune
$ggF, H \to \gamma\gamma_d$	POWHEG[v2 NNLOPS]	NNLO	PYTHIA[8.306]	CT[10], CTEQ[6L1]	AZNLO
$VBF, H \to \gamma\gamma_d$	POWHEG[v2]	NLO	PYTHIA[8.306]	PDF4LHC15, CTEQ[6L1]	AZNLO

Theoretical uncertainties: Signals from ggF and VBF are considered for the computation of the theoretical uncertainties in each SR bin of the analysis as shown in Table 6.1. As no dependence on E_T^{miss} has been noticed, the estimates in the most inclusive SRI1 ($E_T^{miss} > 200GeV$) have been considered. Theoretical uncer-

Table 6.5: Cross-section σ values of heavy Higgs boson productions in ggF and VBF processes.

m_H (GeV)	400	600	800	1000	1500	2000	2500	3000
σ_{ggF} (pb)	9.516	2.006	0.449	0.123	8.91×10^{-3}	1.08×10^{-3}	1.78×10^{-4}	3.50×10^{-5}
σ_{VBF} (pb)	0.758	0.327	0.162	0.087	2.28×10^{-2}	7.05×10^{-3}	2.36×10^{-3}	8.25×10^{-4}

Table 6.6: Theoretical uncertainties (%) for ggF and VBF production modes and for different heavy Higgs masses computed in the most inclusive SRI1, with $E_T^{miss} > 200$ GeV.

m_H (GeV)	ggF, $H \to \gamma\gamma_d$		VBF, $H \to \gamma\gamma_d$	
	PDF+scale (%)	PS (%)	PDF+scale (%)	PS (%)
400	(-14.6 +16.9)	2.20	(-1.54 +1.29)	10.0
600	(-14.7 +16.1)	2.20	(-2.57 +2.04)	2.82
800	(-15.0 +16.2)	6.49	(-2.63 +1.82)	2.94
1000	(-15.4 +16.4)	4.54	(-4.21 +3.41)	3.43
1500	(-16.3 +17.2)	3.01	(-5.28 +4.01)	2.51
2000	(-13.3 +13.3)	2.17	(-6.46 +5.05)	1.46
2500	(-22.1 +23.1)	2.36	(-8.15 +7.38)	1.09
3000	(-22.1 +23.1)	3.60	(-9.03 +7.52)	2.28

tainties include uncertainties on the NLO cross-section due to the QCD factorisation and renormalisation scales [**Bothmann:2016nao**] as well as the choice of parton distribution functions. Uncertainties on initial- and final-state radiations due to the choice of parton shower (PS) parameters used with PYTHIA[8.306] are estimated by generating MC samples with the alternative tunes described in Ref. [**ATL-PHYS-PUB-2014-021**]. The uncertainties for both production processes and for each mass point are summarised in Table 6.6. Because of technical reasons, some information were missing in the samples m_H = 400 GeV, 2500 GeV, so the corresponding uncertainties related to samples m_H= 600 GeV, 3000 GeV are considered as a conservative estimate.

6.5 Reinterpretation

The reinterpretation of the Run-2 $\gamma + E_T^{miss}$ search is performed including the heavy Higgs boson signal decaying to a photon and a massless dark photon into the simultaneous likelihood fit described in Section **??**. The signal experimental and theoretical uncertainties are considered in the fit as nuisance parameters. The background estimation has been obtained providing to the fit the same inputs of the $\gamma + E_T^{miss}$ analysis. To illustrate the shape of the different $H \to \gamma\gamma_d$ signals in the SR, their E_T^{miss} distributions have been superimposed to the expected SM background and data after performing the 'simplified shape fit' as shown in Figure 6.3. Signal acceptances for ggF and VBF processes in the SRs used in the 'simplified shape fit' are shown in Table 6.7. Exclusion upper limits at 95% CL on $\sigma \times BR(H \to \gamma\gamma_d)$

Table 6.7: Signal acceptance (%) for ggF and VBF processes in all the exclusive and the last inclusive SRs used in the 'simplified shape fit'.

	SRE1		SRE2		SRE3		SRI4	
m_H (GeV)	ggF (%)	VBF (%)	ggF (%)	VBF (%)	ggF (%)	VBF (%)	ggF (%)	VBF (%)
400	8.15	4.30	0.35	0.49	0.04	0.05	0.00	0.01
600	9.05	4.95	18.9	9.10	7.74	5.44	0.35	0.53
800	3.21	1.96	5.33	3.27	15.4	9.39	15.6	10.5
1000	1.63	1.24	2.50	1.72	5.92	4.01	29.4	21.2
1500	0.50	0.38	0.73	0.69	1.65	1.33	33.3	30.0
2000	0.22	0.21	0.35	0.33	0.67	0.69	32.7	34.3
2500	0.10	0.09	0.16	0.18	0.35	0.41	29.6	38.0
3000	0.04	0.08	0.08	0.11	0.19	0.29	28.9	39.6

are obtained performing the 'simplified shape fit' including both the SRs and their associated CRs by means of the HistFitter package [**Cranmer:2012sba**]. These limits are based on the profile-likelihood-ratio test statistic [**HistFitter**] and $\mathrm{CL_s}$ prescriptions [**0954-3899-28-10-313**], evaluated using the asymptotic approximation [**asymptoticFormulae**]. Results for the heavy Higgs boson signal produced via ggF and VBF separately as well as through a combination in which the two processes have been merged proportionally to their relative theoretical cross-sections [**HiggsCrossSectionYR4**] ggF+VBF are shown in Figure 6.4 for the considered Higgs mass values. The theoretically predicted cross-section of a heavy Higgs boson produced via ggF or VBF is superimposed assuming a $BR(H \to \gamma\gamma_d) = 0.05$. Table 6.8 summarises the observed and the expected exclusion limits at 95% CL on $\sigma \times BR(H \to \gamma\gamma_d)$ for ggF and VBF separately as well as for the combined productions ggF+VBF. The relative impact of the signal theoretical uncertainties on the exclusion limits is below 11.5% for ggF process and below 2.5% for VBF process considering all mass samples, while the signal experimental uncertainties have an overall impact below 7% for both processes.

6.6 Summary

The reinterpretation of the ATLAS full Run-2 $\gamma+E_T^{miss}$ search using the RECAST framework is performed in the context of the production of heavy Higgs boson decaying to a photon and a massless dark photon. Dominant contributions comes from both gluon-gluon Fusion (ggF) and Vector-Boson Fusion (VBF) production modes and different mass hypotheses for the heavy Higgs boson have been considered spanning from 400 GeV to 3 TeV to probe the sensitivity of the $\gamma + E_T^{miss}$ search

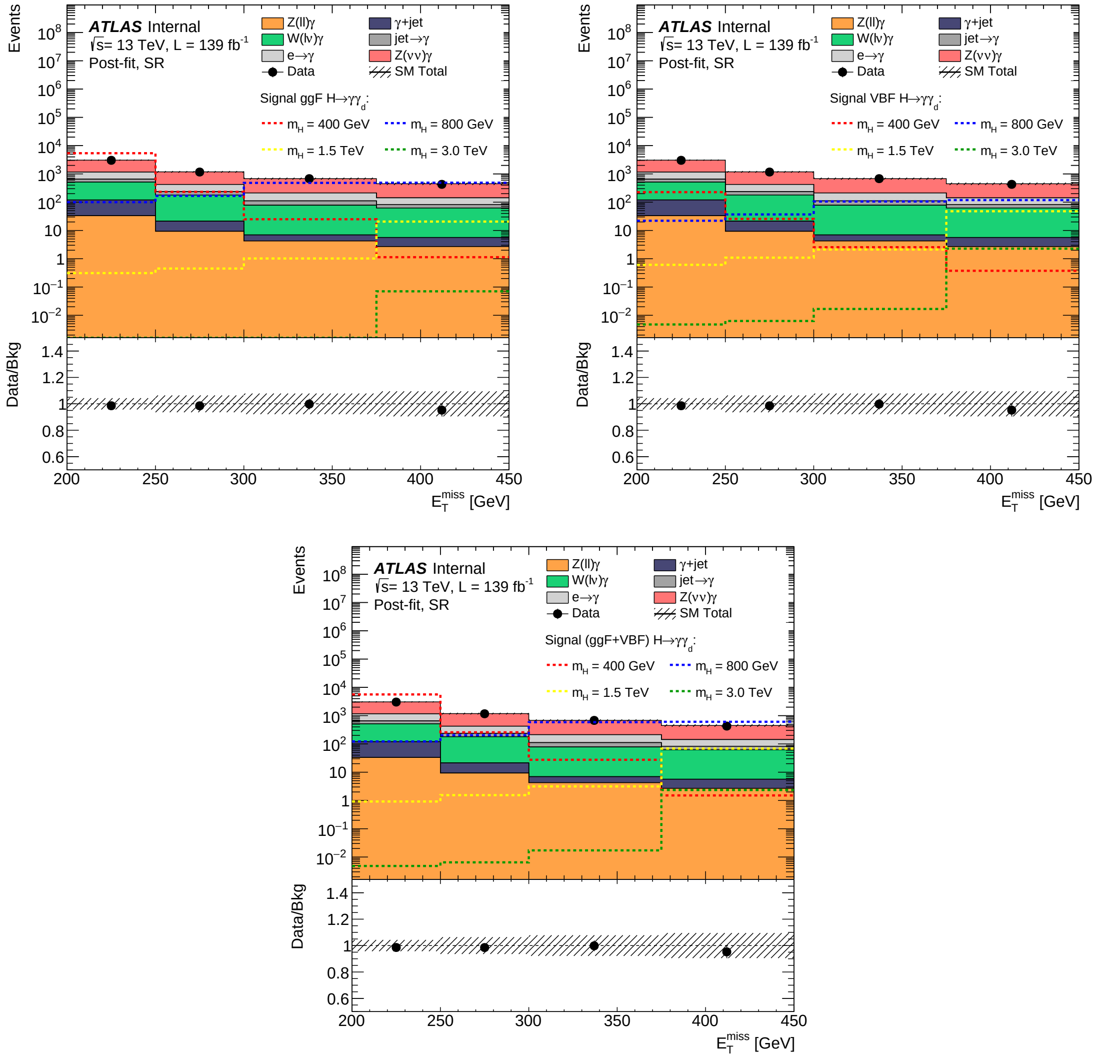


Figure 6.3: Distributions of E_T^{miss} for (a) ggF, (b) VBF, and (c) ggF+VBF production modes of $H \to \gamma\ \gamma_d$ assuming the BR($H \to \gamma\ \gamma_d$) = 0.05. The signals are superimposed to the expected SM background and data after performing the 'simplified shape fit'. The error bars are statistical, and the dashed band includes statistical and systematic uncertainties determined by the fit. The lower panel shows the ratio of data to expected background event yields.

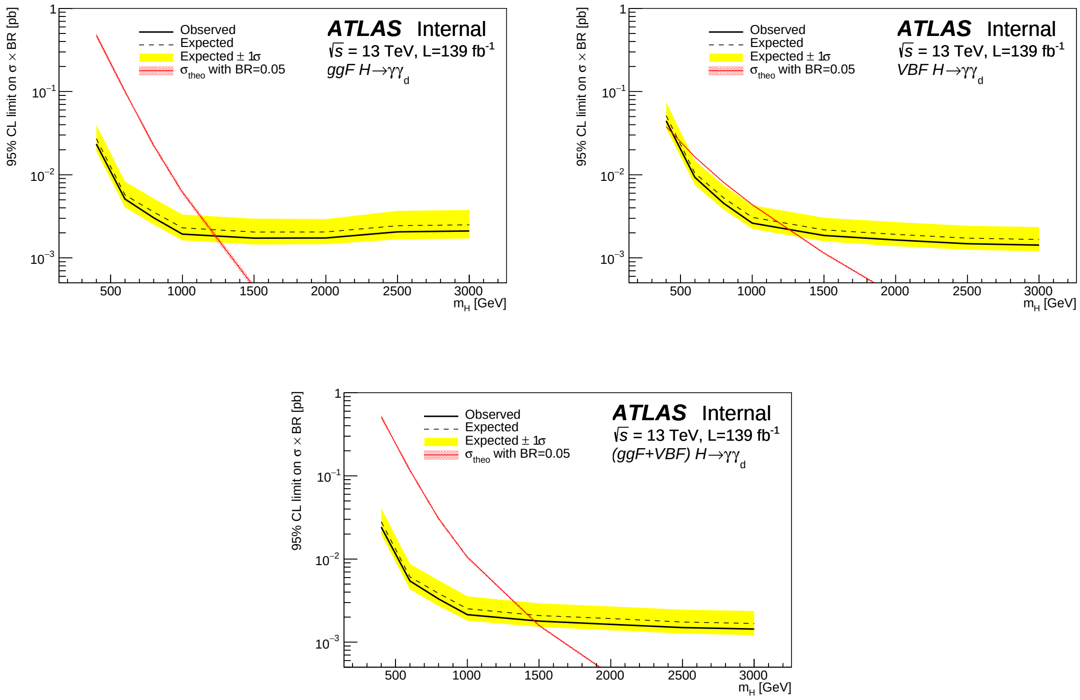


Figure 6.4: Observed and expected exclusion limits at 95% CL on $\sigma\times$ BR($H \rightarrow \gamma\ \gamma_d$) as function of the heavy Higgs mass. The yellow band shows the $\pm1\sigma$ uncertainties. The theoretically predicted cross-section of a heavy Higgs boson produced via (a) ggF, (b) VBF and (c) ggF+VBF is superimposed assuming a $BR(H \rightarrow \gamma\gamma_d) = 0.05$.

Table 6.8: Observed and expected exclusion limits at 95% CL on $\sigma \times BR(H \rightarrow \gamma\gamma_d)$ obtained with the 'simplified shape-fit' for ggF, VBF and combined ggF+VBF channels.

95% CL Limits	$\sigma_{ggF} \times BR$		$\sigma_{VBF} \times BR$		$\sigma_{(ggF+VBF)} \times BR$	
m_H	Obs. (fb)	Exp.$^{+1\sigma}_{-1\sigma}$ (fb)	Obs. (fb)	Exp.$^{+1\sigma}_{-1\sigma}$ (fb)	Obs. (fb)	Exp.$^{+1\sigma}_{-1\sigma}$ (fb)
400 GeV	23.9	$27.3\ ^{+12.7}_{-7.80}$	44.4	$51.4\ ^{+23.1}_{-15.2}$	24.2	$28.1\ ^{+12.7}_{-8.30}$
600 GeV	5.08	$5.72\ ^{+2.60}_{-1.65}$	9.30	$10.43\ ^{+4.35}_{-2.94}$	5.45	$6.10\ ^{+2.66}_{-1.80}$
800 GeV	3.04	$3.56\ ^{+1.63}_{-1.05}$	4.54	$5.28\ ^{+2.17}_{-1.46}$	3.31	$3.86\ ^{+1.69}_{-1.13}$
1 TeV	1.93	$2.30\ ^{+1.02}_{-0.67}$	2.62	$3.08\ ^{+1.21}_{-0.85}$	2.14	$2.52\ ^{+1.04}_{-0.71}$
1.5 TeV	1.73	$2.05\ ^{+0.92}_{-0.60}$	1.86	$2.17\ ^{+0.87}_{-0.59}$	1.79	$2.09\ ^{+0.84}_{-0.57}$
2 TeV	1.74	$2.05\ ^{+0.88}_{-0.59}$	1.64	$1.92\ ^{+0.78}_{-0.52}$	1.64	$1.93\ ^{+0.76}_{-0.54}$
2.5 TeV	2.06	$2.44\ ^{+1.24}_{-0.77}$	1.48	$1.73\ ^{+0.70}_{-0.49}$	1.50	$1.75\ ^{+0.71}_{-0.48}$
3 TeV	2.11	$2.50\ ^{+1.27}_{-0.79}$	1.42	$1.66\ ^{+0.69}_{-0.46}$	1.44	$1.68\ ^{+0.69}_{-0.46}$

towards the $H \rightarrow \gamma\gamma_d$ signal. Exclusion upper limits on the cross-section times the branching ratio for $H \rightarrow \gamma\gamma_d$ are provided for both production modes ggF and VBF separately and for the combined $ggF + VBF$ processes for the considered

Higgs masses. The exclusion limits at 95% confidence level for the $ggF + VBF$ $H \to \gamma\gamma_d$ are found to be in the range from 1.44 (1.68 fb) to 24.2 (28.1 fb) observed (expected) respectively, with the highest limit corresponding to a heavy Higgs boson with a mass of 400 GeV and the lowest limit corresponds to the heaviest considered mass of 3 TeV.

Summary and conclusion

Despite the completeness of the Higgs boson discovery's Standard Model predictions, however certain phenomena in the Standard Model remain questions that fails to explain but are considered beyond the Standard Model physics. The top quark is regarded to be a very interesting probe to search for new physics. One of the features that make this quark unique is its large mass, which translates into a very short lifetime. One result is that its properties are accessible via its decay products. Furthermore, its Yukawa coupling to the Higgs boson is on the order of one, providing access to knowledge on fundamental interactions at the electroweak symmetry-breaking scale and beyond. A further consequence of its large mass is that new particles are expected in many beyond the Standard Model physics preferably to top quarks.

In searches for the t$\bar{t}$ resonnance in the final state, the missing transverse momentum E_T^{miss} is a crucial quantity. It arises in the presence of undetected particles, such as neutrinos in SM processes or hypothetical new particles predicted, for example, by dark matter. In particular a detailed study of different $E_T^{miss,softterm}$ systematic uncertainties for various aspects of the ATLAS Inner Detector has been performed to find the best estimation to be applied to the $E_T^{miss,softterm}$ calculation. The performance of the soft event to the reconstructed E_T^{miss} is determined with Z $\rightarrow$ $\mu\mu$ final states without jets in p-p colision data at the LHC, acquired in 2015 with the ATLAS detector at $\sqrt{s} = 13$ TeV and corresponding to about 36 fb^{-1}. By comparing the behavior between the nominal $E_T^{miss,softterm}$ (without systematic uncertainties) and $E_T^{miss,softterm}$ variation in the different sources of systematic uncertainties listed in table 4.1, we have gained a good understanding of the functioning of the $E_T^{miss,softterm}$ algorithm. No significant deviation between the in the $E_T^{miss,softterm}$ with different variation considered in the search and in the $E_T^{miss,softterm}$ nominal has been observed. These results point out the need of a more careful optimization for the track selection.

This dissertation presented experimental searches for new massive particles decaying to top-antitop t$\bar{t}$ pair events in the semileptonic decay channel. The search has been carried out on the full dataset collected by the ATLAS experiment in Run 2 during 2015 to 2018 in pp collisions at the LHC at a centre-of-mass energy of $\sqrt{s}$ = 13 TeV. The dataset corresponds to an integrated luminosity of 139 fb^{-1} . The search focuses on the distributions of the invariant mass of the reconstructed t$\bar{t}$ pair in the range 0.4 - 5.0 TeV, which are analysed for deviations between observed data and the SM expectation. The search currently is on-going

Two distinct approaches for selecting and reconstructing $t\bar{t}$ events in the l+jets channel in various kinematic regimes are used to improve the sensitivity of the search over the whole $t\bar{t}$ invariant mass range under consideration. The resolved topology selection and reconstruction is focused at events in which the hadronic showers produced by the decay products of the hadronically decaying top quark are well separated in the detector and can be reconstructed as three independent jets with small R parameter. In contrast, boosted-topology selection and reconstruction are optimized for events in which these showers combine into a single, large-R jet with a distinct substructure. All the major SM backgrounds are modeled using Monte Carlo simulation. Since the analysis is still blinded, the dissertation presented the expected sensitivities with the full Run 2 data. Based on the current optimization, the expected limit on Z'_{TC2} mass is 3.7 TeV. With our $t\bar{t}$ resonance search results, we will be able to combine with the other resonance searches in ATLAS using the benchmark Heavy Vector Triplet model.

Following Run-2, the next run of the LHC started in the begining of 2022. The collected data is expected to be about double the amount of current data through the end of Run 3 in 2026. This rise in luminosity will not significantly improve the sensitivity of these searches. As a consequence, we need apply novel techniques to extract even more information from data. We used one discriminant to identify the signals in the search provided in this dissertation. A multivariate technique using likelihood-free inference could help us regain some of the information lost while building a summary test statistics from the discriminant. It could be a potential direction for the Run-3 analyses. Furthermore, the $t\bar{t}$ resonance searches could benefit from the improvement in jet reconstruction in ATLAS.

Summary and conclusion

Despite the completeness of the Higgs boson discovery's Standard Model predictions, however certain phenomena in the Standard Model remain questions that fails to explain but are considered beyond the Standard Model physics. The top quark is regarded to be a very interesting probe to search for new physics. One of the features that make this quark unique is its large mass, which translates into a very short lifetime. One result is that its properties are accessible via its decay products. Furthermore, its Yukawa coupling to the Higgs boson is on the order of one, providing access to knowledge on fundamental interactions at the electroweak symmetry-breaking scale and beyond. A further consequence of its large mass is that new particles are expected in many beyond the Standard Model physics preferably to top quarks.

In searches for the t$\bar{t}$ resonnance in the final state, the missing transverse momentum E_T^{miss} is a crucial quantity. It arises in the presence of undetected particles, such as neutrinos in SM processes or hypothetical new particles predicted, for example, by dark matter. In particular a detailed study of different $E_T^{miss,softterm}$ systematic uncertainties for various aspects of the ATLAS Inner Detector has been performed to find the best estimation to be applied to the $E_T^{miss,softterm}$ calculation. The performance of the soft event to the reconstructed E_T^{miss} is determined with Z $\rightarrow \mu\mu$ final states without jets in p-p colision data at the LHC, acquired in 2015 with the ATLAS detector at $\sqrt{s} = 13$ TeV and corresponding to about 36 fb^{-1}. By comparing the behavior between the nominal $E_T^{miss,softterm}$ (without systematic uncertainties) and $E_T^{miss,softterm}$ variation in the different sources of systematic uncertainties listed in table 4.1, we have gained a good understanding of the functioning of the $E_T^{miss,softterm}$ algorithm. No significant deviation between the in the $E_T^{miss,softterm}$ with different variation considered in the search and in the $E_T^{miss,softterm}$ nominal has been observed. These results point out the need of a more careful optimization for the track selection.

This dissertation presented experimental searches for new massive particles decaying to top-antitop t$\bar{t}$ pair events in the semileptonic decay channel. The search has been carried out on the full dataset collected by the ATLAS experiment in Run 2 during 2015 to 2018 in pp collisions at the LHC at a centre-of-mass energy of $\sqrt{s}$ = 13 TeV. The dataset corresponds to an integrated luminosity of 139 fb^{-1} . The search focuses on the distributions of the invariant mass of the reconstructed t$\bar{t}$ pair in the range 0.4 - 5.0 TeV, which are analysed for deviations between observed data and the SM expectation. The search currently is on-going

Two distinct approaches for selecting and reconstructing $t\bar{t}$ events in the l+jets channel in various kinematic regimes are used to improve the sensitivity of the search over the whole $t\bar{t}$ invariant mass range under consideration. The resolved topology selection and reconstruction is focused at events in which the hadronic showers produced by the decay products of the hadronically decaying top quark are well separated in the detector and can be reconstructed as three independent jets with small R parameter. In contrast, boosted-topology selection and reconstruction are optimized for events in which these showers combine into a single, large-R jet with a distinct substructure. All the major SM backgrounds are modeled using Monte Carlo simulation. Since the analysis is still blinded, the dissertation presented the expected sensitivities with the full Run 2 data. Based on the current optimization, the expected limit on Z'_{TC2} mass is 3.7 TeV. With our $t\bar{t}$ resonance search results, we will be able to combine with the other resonance searches in ATLAS using the benchmark Heavy Vector Triplet model.

Following Run-2, the next run of the LHC started in the begining of 2022. The collected data is expected to be about double the amount of current data through the end of Run 3 in 2026. This rise in luminosity will not significantly improve the sensitivity of these searches. As a consequence, we need apply novel techniques to extract even more information from data. We used one discriminant to identify the signals in the search provided in this dissertation. A multivariate technique using likelihood-free inference could help us regain some of the information lost while building a summary test statistics from the discriminant. It could be a potential direction for the Run-3 analyses. Furthermore, the $t\bar{t}$ resonance searches could benefit from the improvement in jet reconstruction in ATLAS.