Search for Dark Matter Produced in Association with Heavy Standard Model Particles at $\sqrt{s} = 13$ TeV with the ATLAS Detector at the LHC

Dissertation
zur Erlangung des Doktorgrades
an der Fakultät für Mathematik, Informatik und Naturwissenschaften
Fachbereich Physik
der Universität Hamburg

vorgelegt von

Xuanhong Lou

Hamburg
2020

Gutachter/innen der Dissertation:	Dr. Krisztian Peters Prof. Dr. Johannes Haller
Zusammensetzung der Prüfungskommission:	Prof. Dr. Dieter Horns Prof. Dr. Johannes Haller Prof. Dr. Gregor Kasieczka Dr. Krisztian Peters Dr. Kai Schmidt-Hoberg
Vorsitzende/r der Prüfungskommission:	Prof. Dr. Dieter Horns
Datum der Disputation:	27.05.2020
Vorsitzender Fach-Promotionsausschusses PHYSIK:	Prof. Dr. Günter Sigl
Leiterdes Fachbereichs PHYSIK:	Prof. Dr. Wolfgang Hansen
Dekander Fakultät MIN:	Prof. Dr. Heinrich Graener

Abstract

Dark Matter composes a significant part of the Universe, while its physical nature remains unknown. This thesis presents two searches for Dark Matter produced in association with heavy Standard Model particles using pp collision data at a center-of-mass energy of $\sqrt{s} = 13$ TeV recorded by the ATLAS detector at the Large Hadron Collider. Moreover, jet mass scale calibrations for variable-radius calorimeter jets are performed to improve the reconstruction performance of heavy particles in boosted event topologies. A data sample corresponding to an integrated luminosity of 36.1 fb^{-1} is analyzed in the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search, which selects processes with hadronic decays of W and Z bosons in association with large missing transverse energy. The $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search is performed using 126.7 fb^{-1} of collision data and targets events containing fully-hadronically decaying top quark pairs and medium missing transverse energy. No significant excess over the Standard Model prediction is observed in both analyses. The results of the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search are interpreted in terms of constraints on the parameter space of spin-1 vector mediator simplified model and mediator masses of up to 650 GeV are excluded for Dark Matter masses of up to 250 GeV at 95% confidence level with a dark sector coupling of 1.0 and a coupling to Standard Model particles of 0.25. The results of the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search are interpreted in the framework of spin-0 mediator simplified models with unitary couplings. For scalar and pseudo-scalar mediators, masses below 190 GeV and 240 GeV are excluded assuming a Dark Matter mass of 1 GeV, respectively. The implications of these results are discussed and compared to results from current direct detection experiments.

Zusammenfassung

Das Universum besteht zum großen Teil aus dunkler Materie, deren physikalische Grundlagen unbekannt sind. Diese Arbeit präsentiert zwei Suchen nach dunkler Materie in Proton-Proton-Kollisionen bei einer Schwerpunktsenergie von $\sqrt{s} = 13$ TeV mit dem ATLAS-Experiment am Large Hadron Collider. Zudem wird die Massenkalibration von Kalorimeterjets mit variablem Radius präsentiert. Sie ermöglicht eine verbesserte Rekonstruktion schwerer Teilchen in kollimierten Topologien. Die $E_{\mathrm{T}}^{\mathrm{miss}} + V$-Suche selektiert Ereignisse mit hadronisch zerfallenden W- und Z-Bosonen sowie großem fehlenden Transversalimpuls aus einem Datensatz von 36.1 fb^{-1}. Für die $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$-Suche werden aus einem größeren Datensatz von 126.7 fb^{-1} Ereignisse mit vollhadronisch zerfallenden Top-Quark-Paaren und mittlerem fehlenden Transversalimpuls ausgewählt. In keiner der Analysen wird eine signifikante Abweichung von der Standardmodellvorhersage beobachtet. Die Resultate der $E_{\mathrm{T}}^{\mathrm{miss}} + V$-Suche werden in einem vereinfachten Spin-1-Vektor-Mediator-Modell interpretiert, wobei Mediatormassen kleiner als 650 GeV bei dunkle Materie mit Massen kleiner als 250 GeV mit 95-prozentiger Sicherheit ausgeschlossen werden können. Hierbei wird angenommen, dass die Kopplungsstärke zum dunklen Sektor 1 und zu Standardmodellteilchen 0.25 beträgt. Die $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$-Suche wird im Rahmen vereinfachter Spin-0-Mediator-Modelle mit unitären Kopplungen interpretiert. Für dunkle Materie mit einer Masse von 1 GeV können skalare und pseudoskalare Mediatoren leichter als 190 GeV bzw. 240 GeV mit 95-prozentiger Sicherheit ausgeschlossen werden. Abschließend werden die Konsequenzen dieser Ergebnisse erörtert und mit aktuellen Resultaten direkter Suchen nach dunkler Materie verglichen.

Contents

Chapter 1

Introduction

The Standard Model provides remarkably accurate descriptions of nearly all observed physical phenomena, except for gravitation, and can be considered as one of the most successful theories of the last century. With the discovery of the Higgs boson by the ATLAS and the CMS experiments at the Large Hadron Collider in 2012, the last piece of the Standard Model fell into place, leaving no hints of new physics. Nevertheless, a number of open questions are still unaddressed, including the hierarchy problem, the unification with General Relativity, the matter-antimatter asymmetry as well as the nature of Dark Energy and Dark Matter.

The existence of Dark Matter is well-established by a wide range of cosmological and astrophysical observations, while the abundance of Dark Matter is expected to be more than five times larger compared to the visible components in the Universe. Over the past decades, various theories beyond the Standard Model consisting of hypothetical Dark Matter particles have been proposed, with one of the best motivated candidate referred to as the weakly interacting massive particles (WIMPs). Under such scenario, the pair production of Dark Matter can be probed at hadron colliders and, moreover, can be described by the models involving mediator particles which couple to both the Standard Model particles and the dark sector.

The Large Hadron Collider, the largest particle accelerator ever built in human history, allows for precise measurements of proton-proton collision events using data collected by the ATLAS detector. In addition, the increased center-of-mass energy and collision rate in Run 2 data-taking period of the Large Hadron Collider starting from 2015 grant countless new opportunities to extend the boundaries of knowledge of Dark Matter. The simplified models of Dark Matter, which provide minimal extensions of the Standard Model sector by assuming the existence of a single mediator, are proposed to serve as the benchmarks in many analyses. As the detection of Dark Matter particles at hadron colliders is extremely challenging due to their weak coupling to the luminous matter, an indirect measurement of Dark Matter production is performed based on the momentum conservation in the transverse plane. The missing transverse energy, $E_{\mathrm{T}}^{\mathrm{miss}}$, is defined as the imbalance of the total transverse momentum to evaluate the kinematic properties of all the invisible particles in the final state. A large variety of Dark Matter searches have been carried out in ATLAS targeting different $E_{\mathrm{T}}^{\mathrm{miss}} + X$ signatures, where X may represent a jet, a photon, a vector boson or heavy-flavor quarks.

Among these distinctive signatures, the sensitivity of searches for Dark Matter in 0 lepton final state strongly depend on the accuracy of jet measurements. The jet mass scale calibration, which corrects, on average, the invariant mass of jets to that at truth-level is of great importance for any searches exploring such topologies. In the scope of this work, jet mass scale calibration functions are derived for variable-radius calorimeter jets, which are reconstructed using algorithms specially designed for optimal performance of physical object reconstruction in highly boosted event topologies. Significant improvements can be observed in response closure and mass resolution after the calibration over the full kinematic phase space.

Two searches for Dark Matter using pp collision data recorded by the ATLAS detector at the center-of-mass energy of $\sqrt{s} = 13$ TeV are presented in this thesis, which are performed in a way that both the spin-1 and the spin-0 mediator simplified models can be probed. The $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search is designed to select events in which Dark Matter particles are produced through the s-channel exchange of a spin-1 vector mediator and in association with a Standard Model vector boson. The signals are characterized by at least one large-R jet or two small-R jets with kinematic properties compatible with a hadronically decaying W/Z and large missing transverse energy. The $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search is optimized in search of Dark Matter produced through the s-channel exchange of a spin-0 scalar or pseudo-scalar mediator and in association with a pair of top quarks. The targeted event signatures include multiple small-R jets, among which two jets originate from b-hadron decays, with kinematic patterns compatible with two hadronically decaying tops and medium missing transverse energy.

This dissertation is organized as follows.

In Chapter 2, a brief theoretical overview of the Standard Model of particle physics is given. The observations pointing to the existence of Dark Matter, the possible Dark Matter candidates and searches for Dark Matter production at the Large Hadron Collider are also included.

Chapter 3 discusses the structure of the Large Hadron Collider and the ATLAS detector which are used to produce and collect the pp collision data, followed by a detailed explanation of the reconstruction of physical objects in the ATLAS experiment.

Chapter 4 details the reconstruction and calibration of variable-radius jets, especially the studies done to improve the closure performance of their mass response.

Chapter 5 is dedicated to the search for Dark Matter produced in association with a hadronically decaying vector boson. With a similar structure, Chapter 6 documents the search for Dark Matter produced in association with a pair of top quarks and medium missing transverse energy in 0 lepton final state.

All these works are summarized in Chapter 7, alongside an outlook for the future Dark Matter searches at the Large Hadron Collider.

1.1 Author's contributions

As one of the most complex scientific projects in the field of high energy physics, the ATLAS collaboration involves more than 3000 researchers from all across the world and a large number of distinctive groups focusing on different aspects of the experiment, including the design, construction, maintenance and upgrade of the ATLAS detector, the reconstructed object performance as well as the physical analyses. Therefore, all results presented in this work rely on the joint effort of many people, while the author's contributions are summarized in this section.

Regarding to the derivation of jet mass scale calibration for variable-radius calorimeter jets, the author contributed to the development of the most recent framework used to evaluate the calibration functions, validated it with respect to the previous version of code and derived the calibration factors for variable-radius calorimeter jets with three different mass definitions, which can be of assistance to any Run 2 analyses exploiting the highly boosted event topologies.

The $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search is mainly performed by the author in cooperation with another Ph.D. student, Paul Philipp Gadow at the Max Planck Institute for Physics (Werner Heisenberg Institute), Munich. In addition, the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger calibration documented in Section 5.2.3.1 is carried out by Stanislav Suchek at the Kirchhoff Institute for Physics, Heidelberg University. The results of this analysis were published in *Journal of High Energy Physics* in 2018 [1]. The author designed and optimized the analysis regions, maintained the frameworks used for event selection and statistical analysis, evaluated the theoretical systematic uncertainties for signals, derived the final exclusion limits and contributed to all the studies and their documentations towards the publication.

The $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search is mainly performed by the author, except for the multijet estimation documented in Section 6.6.4, which is carried out by Jonas Neundorf as part of his master dissertation at Universität Hamburg. Since this analysis is done for the first time in ATLAS, the author designed and optimized the analysis regions at both truth- and reconstruction-level from scratch, developed and maintained the frameworks used for event selection and statistical analysis, configured the Monte Carlo generation of signal events, evaluated the theoretical systematic uncertainties for backgrounds and signals, derived the final exclusion limits and contributed to the documentation of all the above mentioned studies. These results are currently in the internal review process in ATLAS to be published in a peer reviewed journal.

Chapter 2

Theoretical overview

Since the very ancient times, philosophers and scientists had been pondering the nature of the Universe. What is the world made of? What are the fundamental units which assemble all the material existence? Countless physicists tried to answer these questions, leading to the development of the Standard Model of particle physics. However, the Standard Model is only capable to describe 5% of the energy content of the Universe. Astrophysical measurements indicate that a dominant part of the mass of the Universe is composed of Dark Matter, yet its connection to the Standard Model sector remains unknown.

This chapter gives a theoretical overview of the Standard Model and the properties of Dark Matter. Section 2.1 summarizes the elementary particles and the fundamental interactions in the context of the Standard Model, as well as the open questions and limitations of this framework. Section 2.2 describes the observation of Dark Matter, the possible Dark Matter candidates, the main approaches being exploited in search of Dark Matter and the simplified models of Dark Matter. A brief discussion of the motivation of the two analyses presented in this thesis is attached at the end of this chapter.

2.1 The Standard Model

The Standard Model (SM) of particle physics [2, 3, 4, 5] is by far the most successful theoretical model which illustrates the nature of fundamental particles and the interaction between them in the language of quantum field theory (QFT). The validity of the SM is supported by many experimental results including the observation of the W boson [6], Z boson [7], gluon [8], top quark [9, 10], charm quark [11, 12] and Higgs boson [13, 14]. As described in the SM, all fundamental particles can be categorized into two groups with respect to their spin quantum number:

- **bosons**: particles with integer spin, following Bose-Einstein statistics [15];

- **fermions**: particles with half-integer spin, following Fermi-Dirac statistics [16, 17].

The fundamental particles can interact with each other via electromagnetic (EM), weak, strong and gravitational forces. In the SM only the first three interactions are described while the integration of gravitational interaction is still left open in the current frame of QFT. Nevertheless, for elementary particles the gravitational force is ~ 39 orders of magnitude weaker than the electromagnetic force, which makes the gravitational interaction negligible when performing the SM calculations.

In most cases, the bosons serve as the mediators (or the force-carriers) of EM, weak and strong interactions, while the fermions constitute the fundamental matter.

2.1.1 The Standard Model particles

Based on the spin number, bosons can be divided into gauge bosons[1] (spin = 1) and scalar bosons (spin = 0). Gauge bosons include the photon (γ), eight different types of gluons (g), the neutral weak boson (Z) and the charged weak bosons ($W^{\pm}$). For scalar bosons, only one particle falls into this category: the Higgs boson. As illustrated in Table 2.1.1, the EM and the strong interactions are mediated via virtual photons and gluons, respectively, while the weak interaction is carried by the exchange of W and Z bosons.

	Name	**Charge**	**Mass [GeV]**	**Interaction**
gauge boson	photon (γ)	0	$< 1 \times 10^{-24}$	electromagnetic
	gluon (g)	0	0[2]	strong
	Z boson (Z)	0	91.19	weak
	W boson ($W^{\pm}$)	± 1	80.38	
scalar boson	Higgs boson (H)	0	125.18	

Table 2.1.1: List of bosons in the Standard Model and their basic properties. The numbers are taken from the Particle Data Group 2018 [18].

Fermions include six leptons, six quarks as well as their corresponding antiparticles and can be further categorized into three generations based on the differences in their masses. Among the six leptons, three of them are charged particles, including the electron (e^-), the muon (μ^-) and the tau (τ^-), while the other three are neutral, including the electron neutrino (ν_e), the muon neutrino (ν_μ) and the tau neutrino (ν_τ). For each of the charged leptons there exists a corresponding antilepton which shares the same mass but has opposite charge and for each neutrino there is an antineutrino with the opposite lepton number and chirality. The quark sector includes three up-type quarks (the up, the charm and the top quark) with an electric charge of $+\frac{2}{3}e$ and three down-type quarks (the down, the strange and the bottom quark) with a charge of $-\frac{1}{3}e$. Similar

[1] Also known as vector bosons. The name 'vector boson' is inspired by the fact that for massive bosons with spin 1, three eigenvalues of spin are expected, which comes from the degree of freedom of the rotation group and equals to that of a vector in three-dimensional space. However, for massless particles, no rest frame is available and this forbids the longitudinal polarization, which reduces the degree of freedom of the rotation group by 1. Hence, it is not appropriate to refer massless bosons, i.e. the photons and the gluons, as 'vector bosons', although they all come from the excitation of vector fields.

[2] Theoretical value.

to the charged leptons, each quark possesses an antiquark with an electric charge of equal magnitude but opposite sign. In the notation, the antiparticles is marked by the same symbol of the original particle with a '$^{-}$' on top, e.g. b and $\bar{b}$.

As shown in Table 2.1.2, each generation of fermions contains one charged lepton, one neutrino, one up-type quark and one down-type quark. With the rising of generations, the mass of charged leptons and quarks becomes higher[3] and the particle tends to be more unstable. Additionally, the quark masses shown in Table 2.1.2 are the estimation of the current quark masses (or 'the bare quark masses'), which are calculated by subtracting the constituent quark covering from the constituent quark masses. For light quarks, e.g. up (u), down (d) and strange (s), the major contribution of constituent quark mass comes from the quantum chromodynamics binding energy (QCBE) of gluons surrounding the quark, which leads to large deviation between the constituent mass and the current mass. For heavy quarks, e.g. charm (c), top (t) and bottom (b), the two definitions of mass show almost no differences.

Generation	**Name**	**Charge**	**Mass [GeV]**
1st generation	electron (e)	-1	5.11×10^{-4}
	electron neutrino (ν_e)	0	—
	up quark (u)	$+\frac{2}{3}$	2.2×10^{-3}
	down quark (d)	$-\frac{1}{3}$	4.7×10^{-3}
2nd generation	muon (μ)	-1	0.10566
	muon neutrino (ν_μ)	0	—
	charm quark (c)	$+\frac{2}{3}$	1.275
	strange quark (s)	$-\frac{1}{3}$	0.095
3rd generation	tau (τ)	-1	1.77686
	tau neutrino (ν_τ)	0	—
	top quark (t)	$+\frac{2}{3}$	173.0
	bottom quark (b)	$-\frac{1}{3}$	4.18

Table 2.1.2: List of fermions in the Standard Model and their basic properties. The numbers are taken from the Particle Data Group 2018 [18].

By exchanging the boson mediators, leptons interact via electromagnetic and weak forces, while for quarks the strong force is also involved. As a result, in addition to the electric and weak charge, quarks can carry another type of charge which corresponds to the strong interaction, namely 'the color charge'. There are three arbitrarily defined color charges, 'red' (r), 'green' (g) and 'blue' (b), and three anticolors, 'antired ($\bar{r}$)', 'antigreen ($\bar{g}$)' and 'antiblue ($\bar{b}$)'. Each quark carries a color and each antiquark carries an anticolor. Due to the color confinement, quarks and antiquarks cannot be observed in free state under normal circumstances[4] and can only be observed in bound states called hadrons.

[3]For neutrinos, the flavor eigenstates (ν_e, ν_μ and ν_τ) are not identical to the mass eigenstates, which makes it impossible to define a mass property for neutrinos in any of the three generations.

[4]The confinement is only valid under the Hagedorn temperature ($\sim 2 \times 10^{12}$ K). At higher energy, the

Hadrons can be classified into two categories:

- **mesons**: particles formed by a quark-antiquark pair, where the quark carries a color and the antiquark carries the corresponding anticolor, e.g. $r\bar{r}$;
- **baryons**: particles formed by three quarks or three antiquarks, where each quark (antiquark) carries a different color (anticolor), e.g. rgb.

In either case, hadrons will be color-neutral.

2.1.2 Gauge symmetries and fundamental interactions

The Standard Model is expressed in the mathematical framework of QFT and every fundamental particle in the SM can be written as a quantum field, including:

- the **vector fields**, A_μ, which describe the gauge bosons and transform like a four-vector under the Lorentz transformation;
- the **Higgs field**, ϕ, which describes the Higgs boson and remains invariant under the Lorentz transformation, also know as the scalar field;
- the **fermion fields**, ψ, which describes the fermions and transform like a spinor under the Lorentz transformation.

Following the Lagrangian formalism and the notion of symmetries, the SM is based upon the gauge symmetry of the group:

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y, \tag{2.1.1}$$

where:

- $SU(3)$ is the three-dimensional special non-abelian unitary group which describes the strong interaction and the **color symmetry**, while C represents the conserved current associated to this symmetry based on the Noether's theorem [19], the color charge;
- $SU(2)$ is the two-dimensional special non-abelian unitary group which describes the weak interaction and the **isospin symmetry**, while the notion L specifies that this symmetry applies to only left-handed fermion fields;
- $U(1)$ is the one-dimensional abelian group which describes the electromagnetic interaction and the **hypercharge symmetry**, while Y represents the conserved current associated to this symmetry, the weak hypercharge $Y = 2(Q - T_3)$[5].

The $U(1)$, $SU(2)$ and $SU(3)$ symmetry groups give rise to the three fundamental interactions in the Universe and determine the variety of gauge bosons in the Standard Model.

boundings between quarks and antiquarks break and the formation of quark-gluon plasma starts to take place.

[5] Q represents the electric charge and T_3 represents the z component of the weak isospin.

2.1.2.1 Electromagnetic interaction

The electromagnetic interaction in the SM is described by Quantum Electrodynamics (QED), where the corresponding Lagrangian density $\mathcal{L}_{\text{QED}}$ (or Lagrangian for short) is invariant under the $U(1)$ transformation. Consider a free fermion with mass m, its Lagrangian can be written as:

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi, \tag{2.1.2}$$

where ψ represents the fermion field spinor, $\bar{\psi}$ is the Dirac adjoint of ψ, defined as $\psi^\dagger\gamma^0$, and γ^μ represents the Dirac matrices. Note that this Lagrangian is not invariant under the local transformation of $U(1)$ group:

$$\psi(x) \rightarrow e^{ie\theta(x)}\psi(x), \tag{2.1.3}$$
$$\bar{\psi}(x) \rightarrow e^{-ie\theta(x)}\bar{\psi}(x), \tag{2.1.4}$$
$$\mathcal{L} \rightarrow \mathcal{L} - e\bar{\psi}(x)\gamma^\mu\partial_\mu\theta(x)\psi(x). \tag{2.1.5}$$

However, by replacing the partial derivative ∂_μ with the covariant derivative D_μ,

$$D_\mu \equiv \partial_\mu - ieA_\mu, \tag{2.1.6}$$

where e is the electric charge and A^μ stands for a vector field which fulfills

$$A^\mu(x) \rightarrow A^\mu(x) + \partial^\mu\theta(x), \tag{2.1.7}$$

the Lagrangian $\mathcal{L}$ preserves its invariance under the gauge transformation:

$$\bar{\psi}D_\mu\psi \rightarrow \bar{\psi}D_\mu\psi. \tag{2.1.8}$$

Adding the kinematic term of A^μ,

$$\mathcal{L}_A = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + m_A A^\mu A_\mu, \tag{2.1.9}$$

where the field strength tensor is defined as $F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu$, the QED Lagrangian can be written as:

$$\mathcal{L}_{\text{QED}} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi + e\bar{\psi}\gamma^\mu A_\mu\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} + m_A A^\mu A_\mu. \tag{2.1.10}$$

This Lagrangian is invariant under the local $U(1)$ transformation as long as $m_A = 0$, which implies the fact that the mediator of the electromagnetic field A^μ, the photon, is massless:

$$\mathcal{L}_{\text{QED}} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi + e\bar{\psi}\gamma^\mu A_\mu\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}. \tag{2.1.11}$$

2.1.2.2 Strong interaction

The strong interaction in the SM is described by Quantum Chromodynamics (QCD), where the corresponding Lagrangian $\mathcal{L}_{\text{QCD}}$ is invariant under the $SU(3)$ transformation. Similar to the construction of the QED Lagrangian, consider the fermion fields of six quarks, the Lagrangian can be written as:

$$\mathcal{L} = \sum_f \bar{\psi}^f (i\gamma^\mu \partial_\mu - m^f)\psi^f, \tag{2.1.12}$$

where $\psi^f = (\psi^f_r, \psi^f_g, \psi^f_b)$ is a three-component vector of the Dirac spinors with the three colors. Again, the partial derivative ∂_μ is replaced by the covariant derivative D_μ,

$$D_\mu \equiv \partial_\mu - ig_s \frac{\lambda_a}{2} G^a_\mu, \tag{2.1.13}$$

where g_s is the gauge coupling constant, λ_a are the eight Gell-Mann matrices and G^a_μ are the eight gluon fields ($a = 1...8$). Adding the kinematic term of G^a_μ,

$$\mathcal{L}_G = -\frac{1}{4} G^a_{\mu\nu} G^{\mu\nu}_a, \tag{2.1.14}$$

where the field strength tensor is defined as $G^a_{\mu\nu} \equiv \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + g_s f^a_{bc} G^b_\mu G^c_\nu$, the QCD Lagrangian can be written as:

$$\mathcal{L}_{\text{QCD}} = \sum_f \left(\bar{\psi}^f (i\gamma^\mu \partial_\mu - m^f)\psi^f + g_s \bar{\psi}^f \gamma^\mu \frac{\lambda_a}{2} G^a_\mu \psi^f \right) - \frac{1}{4} G^a_{\mu\nu} G^{\mu\nu}_a. \tag{2.1.15}$$

2.1.2.3 Weak interaction and Electroweak Unification

The weak interaction in the SM is described together with the EM interaction by the Glashow-Weinberg-Salam (GWS) theory [3, 4, 20], also known as the Electroweak Unification, where the corresponding Lagrangian $\mathcal{L}_{EW}$ is invariant under the $SU(2) \otimes U(1)$ transformation. The first part of this symmetry group, $SU(2)$, introduces three W bosons of the weak isospin symmetry (T), and the second part $U(1)$ introduces the B boson of the weak hypercharge symmetry (Y).

In the Electroweak (EW) theory, the left- and the right-handed fermion fields are in the form of:

$$\psi_{R,L} = \frac{1}{2}(1 \pm \gamma^5)\psi, \tag{2.1.16}$$

where $\gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3$ and $\frac{1}{2}(1 \pm \gamma^5)$ are the chirality operators. Furthermore, the left-handed fermions form the doublets ($T = \frac{1}{2}$) and only participate in the charged-current interactions, while the right-handed fermions form the singlets ($T = 0$) and only participate in the neutral-current interactions.

Similar to the construction of the QED and the QCD Lagrangian, considering the fermion fields of leptons and quarks, the Lagrangian can be written as:

$$\mathcal{L} = \sum_f \bar{\psi}^f i\gamma^\mu \partial_\mu \psi^f, \tag{2.1.17}$$

where the mass term is removed because a simple mass term will mix the left- and the right-handed fields and spoil the gauge symmetry. The partial derivative ∂_μ is replaced by the covariant derivative D_μ,

$$D_\mu \equiv \partial_\mu - ig\frac{\sigma_i}{2}W_\mu^i - ig'\frac{Y}{2}B_\mu, \tag{2.1.18}$$

where g and g' are the gauge coupling constants of the weak and the EM interaction, σ_i are the three Pauli matrices, W_μ^i are the three W boson fields ($i = 1...3$), Y is the weak hypercharge and B_μ is the B boson field. The physical state of the charged weak bosons can be obtained by the mixing of W_μ^1 and W_μ^2:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2), \tag{2.1.19}$$

and the neutral bosons (including the neutral weak boson Z_μ and the photon A_μ) can be obtained by the mixing of W_μ^3 and B_μ:

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix}, \tag{2.1.20}$$

where θ_W is the weak mixing angle, fulfilling:

$$\cos\theta_W = \frac{g'}{\sqrt{g^2 + g'^2}}, \tag{2.1.21}$$

$$\sin\theta_W = \frac{g}{\sqrt{g^2 + g'^2}}. \tag{2.1.22}$$

Adding the kinematic terms of W_μ^i and B_μ,

$$\mathcal{L}_{W,B} = -\frac{1}{4}W_{\mu\nu}^i W_i^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}, \tag{2.1.23}$$

where the field strength tensors are defined as $W_{\mu\nu}^i \equiv \partial_\mu W_\nu^i - \partial_\nu W_\mu^i + g\epsilon_{jk}^i G_\mu^j G_\nu^k$ and $B_{\mu\nu} \equiv \partial_\mu B_\nu - \partial_\nu B_\mu$, the EW Lagrangian can be written as:

$$\mathcal{L}_{\mathrm{EW}} = \sum_f \left(\bar{\psi}^f i\gamma^\mu \partial_\mu \psi^f + g\bar{\psi}^f \gamma^\mu \frac{\sigma_i}{2} W_\mu^i \psi^f + g'\bar{\psi}^f \gamma^\mu \frac{Y}{2} B_\mu \psi^f \right) - \frac{1}{4}W_{\mu\nu}^i W_i^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}. \tag{2.1.24}$$

2.1.2.4 The Higgs Mechanism

To preserve the EW symmetry, bosons and fermions have to be massless. However this is not in agreement with the experimental observations in which the masses of the W and Z boson, the charged leptons and the quarks are measured in high precision. Therefore, the EW symmetry has to be broken to a certain extent, allowing for the existence of massive gauge fields. This is achieved by the mechanism of Spontaneous Symmetry Breaking (SSB), also known as the Higgs Mechanism, where the symmetry group $SU(2) \otimes U(1)$ breaks down to $U(1)$. With SSB, although the Lagrangian in general remains invariant under the transformation, the vacuum ground state is dependent on the choice of gauge. Consider a complex $SU(2)$ doublet ϕ:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \tag{2.1.25}$$

where ϕ^+ is a field with positive electric charge and ϕ^0 is a electric neutral field. The Lagrangian can be written as:

$$\mathcal{L} = (D_\mu\phi)^\dagger(D_\mu\phi) - V(\phi), \tag{2.1.26}$$

where D_μ is the covariant derivative of EW and $V(\phi)$ is in the form of:

$$V(\phi) = \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2. \tag{2.1.27}$$

This field, ϕ, is called the Higgs field and $V(\phi)$ is the Higgs potential[6]. Among the two floating parameters, $\lambda > 0$ has to be fulfilled since the case $\lambda < 0$ will lead to non-existent ground state, while μ^2 can be either positive or negative. If $\mu^2 > 0$, $V(\phi)$ reaches its minimum only at $\phi^+ = \phi^0 = 0$, leading to a QED-like Lagrangian containing a scalar field with mass μ. On the other hand, if $\mu^2 < 0$, the Higgs potential will have an infinite number of degenerate ground states with minimum energy at $\phi^\dagger\phi = -\frac{\mu^2}{2\lambda} \equiv \frac{\nu^2}{2}$. For simplicity, the minimum is chosen as:

$$\phi = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ \nu \end{pmatrix}, \tag{2.1.28}$$

where ν represents the vacuum expectation value (VEV) of ϕ. Expanding the Higgs field around the minimum,

$$\phi' = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ \nu + h \end{pmatrix}, \tag{2.1.29}$$

[6] The renormalizability of $V(\phi)$ forbids the appearance of $\phi^\dagger\phi$ in higher orders.

the Lagrangian of the Higgs field becomes:

$$\mathcal{L}_{\text{Higgs}} = (D_\mu\phi)^\dagger(D_\mu\phi) - \frac{1}{2}(-2\mu^2)h^2 - \lambda\nu h^3 - \frac{1}{4}\lambda h^4, \tag{2.1.30}$$

where the mass of the Higgs field can be extracted from the second term:

$$m_H = \sqrt{-2\mu^2} = \nu\sqrt{2\lambda}. \tag{2.1.31}$$

Expanding the first term with the EW covariant derivative,

$$D_\mu\phi = (\partial_\mu - ig\frac{\sigma_i}{2}W^i_\mu - ig'\frac{Y}{2}B_\mu)\phi, \tag{2.1.32}$$

the mass term of the electroweak bosons should have the form of:

$$\begin{aligned}
\left|(-ig\tfrac{\sigma_i}{2}W^i_\mu - ig'\tfrac{Y}{2}B_\mu)\phi\right|^2 &= \frac{1}{8}\left|\begin{pmatrix} gW^3_\mu + g'B_\mu & g(W^1_\mu - iW^2_\mu) \\ g(W^1_\mu + iW^2_\mu) & -gW^3_\mu + g'B_\mu \end{pmatrix}\begin{pmatrix} 0 \\ \nu \end{pmatrix}\right|^2 \qquad (2.1.33)\\
&= \frac{1}{2}\nu^2 g^2\left[(W^1_\mu)^2 + (W^2_\mu)^2\right] + \frac{1}{8}\nu^2(gW^3_\mu - g'B_\mu)^2 + 0(g'W^3_\mu + gB_\mu)^2\\
&= (\frac{\nu g}{2})^2 W^+_\mu W^-_\mu + \frac{1}{2}(\frac{\nu\sqrt{g^2+g'^2}}{2})^2 Z^2_\mu + \frac{1}{2}\cdot 0\cdot A^2_\mu,
\end{aligned}$$

where the masses of the gauge bosons can be extracted as:

$$m_{W^+} = m_{W^-} = \frac{\nu g}{2}, \tag{2.1.34}$$

$$m_Z = \frac{\nu\sqrt{g^2+g'^2}}{2}, \tag{2.1.35}$$

$$m_\gamma = 0. \tag{2.1.36}$$

The masses of fermions are generated via the Yukawa coupling between the fermion fields and the Higgs field:

$$\mathcal{L}_{\text{Yukawa}} = \sum_f -g^f_Y(\bar{\psi}^f_L\phi\psi^f_R + \bar{\psi}^f_R\bar{\phi}\psi^f_L), \tag{2.1.37}$$

where g^f_Y represent the Yukawa coupling constants for fermions f. The fermion masses therefore are given by:

$$m_f = \frac{\nu g^f_Y}{\sqrt{2}}. \tag{2.1.38}$$

2.1.3 Open questions of the Standard Model

Despite being a huge success in the field of particle physics, the Standard Model is still far from perfection and there are several phenomena left unaddressed by the Standard Model:

- **Gravity**. The Standard Model is currently incompatible with General Relativity, the most successful theory of gravitational interaction. On the other hand, attempts to describe the gravity in the language of QFT all failed on renormalization, leading to prediction of infinite values to certain observables which is apparently unphysical.
- **Neutrino masses**. As predicted in the Standard Model, neutrinos are massless particles, while various experiments [21, 22] reached the opposite conclusion with the observation of neutrino oscillations. The flavor eigenstates of neutrinos are mixed by the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix, and thus forming the physical eigenstates with non-zero masses. To date, the absolute scale of neutrino masses and the mass hierarchy of the three neutrinos are still unknown.
- **Dark Matter**. According to cosmological observations, only around 5% of the energy content of the Universe is the 'luminous' matter described in the Standard Model. About 27% of the contribution comes from Dark Matter (DM). The description 'dark' indicates that they cannot be detected by most experimental techniques since they interact with neither the EM force (thus emit photons) nor the strong force and the word 'matter' corresponds to the distinctly possible particle nature of these energy components. The only DM candidate in the SM, the neutrino, has its energy fraction in the Universe constrained by the measurements of the Planck Satellite $\Omega_\nu h^2 < 0.0025$ [23] and cannot cover the required energy density of DM. The most promising DM candidates are the weakly-interacting massive particles (WIMPs) and many beyond the Standard Model (BSM) theories predict the existence of WIMPs.
- **Dark Energy**. Aside from the luminous matter and DM, Dark Energy (DE) forms the remaining 68% of the Universe. Very little is known about the physical nature of DE and it is believed to interact only gravitationally. One of the commonly accepted theory of DE describes it as an intrinsic property of spacetime and a constant energy density filling the entire Universe. This allows DE to account for the accelerating expansion of the Universe.
- **Matter-antimatter asymmetry**. The Standard Model predicts that matter and antimatter should have been created in comparable amounts in the early phase of the Universe, but the stability of the modern Universe shows a dominance of matter over antimatter. Although any asymmetrical production of matter-antimatter leads to a violation of the CP symmetry, the CP violation provided in the SM is far not sufficient to explain the observed asymmetry.
- **Hierarchy problem**. The masses of particles are introduced by their coupling to the Higgs field in the Standard Model, which causes the spontaneous symmetry breaking of the $SU(2)$ group. This indicates that the electroweak symmetry breaks at the scale of $\mathcal{O}(100 \text{ GeV})$ (around the Higgs mass $m_H = 125$ GeV). Furthermore, the unification scale of the EW and strong forces is at $\mathcal{O}(10^{16} \text{ GeV})$ and the unification with the gravity is (at least) at $\mathcal{O}(10^{19} \text{ GeV})$, the Planck scale. The hierarchy of these different scales raises suspicion of the naturalness of the possible fine tuning

on the quantum correction to the Higgs mass. Consider a loop level correction to the Higgs mass, it should be proportional to the square of the scale Λ where the SM breaks down, which is at $\mathcal{O}(10^{38}$ GeV). This implies either severe fine tuning or correction from some new physics which cancels out the loop, thus allows the Higgs mass to stay at 125 GeV. Naturally, the Standard Model does not provide any solution to the problem, since the Higgs mass is not a calculable parameter in the Standard Model.

2.2 Dark Matter

The Standard Model particles and the interactions between them build up the everyday life of human society, but in the scale of galaxies, they are just icebergs above water while huge mysteries are hidden beneath. In the early 1930s, the first observations suggesting the existence of certain 'unknown matter' in the Universe [24, 25] was made and the hypothesis of such matter had arisen even earlier. However, not until the 1990s are the physicists able to determine that the portion of ordinary matter is only $\sim$ 5% among the mass-energy of the whole Universe. While $\sim$ 68% of the contribution is coming from the so-called 'Dark Energy' which corresponds to the acceleration rate of the expansion of the Universe, the remaining 27% belongs to 'Dark Matter' which is assumed to be mostly composed of some undiscovered (meta)stable particles. These particles are likely to only couple weakly to the SM fields and thus makes it extremely challenging to measure their properties. So far, the observations of DM are all based on its gravitational interaction with ordinary matter and the particle nature of DM as well as the DM density throughout the Universe remains unclear.

The unsolved puzzle of Dark Matter has always been one of the most important topics in the field of particle physics. Various BSM models have been proposed by the theorists to bridge the SM and DM sector, while three types of experimental approaches are used chasing for the presence of DM, including direct detection, indirect detection and collider search. The simplified models of Dark Matter, which provide (almost) minimal extensions of the SM sector while maintain the generality at high energy[7] by introducing a DM particle and a mediator that couples to it, balance the predictiveness and the complexity of parameter space in a reasonable way, hence are broadly applied by the experimentalists in search of Dark Matter.

This dissertation, in particular, is focusing on searching for DM produced at the Large Hadron Collider (LHC) via proton-proton collisions in association with Standard Model particles, often referred to as $E_{\mathrm{T}}^{\mathrm{miss}} + X$ searches. Two searches with distinguishing final states are covered in the scope of this work, namely the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search and the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search.

2.2.1 The existence of Dark Matter

The existence of DM is supported by a huge variety of astrophysical observations, while the most paradigmatic ones include studies on the galaxy rotation curves [26, 27], the collision of galaxy clusters [28] and the spectrum of the cosmic microwave background [29, 30, 31].

[7] With respect to the Effective Field Theories, which focus on building effective operators involving both the SM and DM fields but meanwhile sacrifice the validity at different energy scales.

2.2.1.1 Galaxy rotation curves

In the 1960s, Vera Rubin and Kent Ford noticed during their measurements of the velocity curve of edge-on spiral galaxies that the luminous objects in the galaxy move much faster than expected, if one only takes the gravitational interaction into consideration [26]. Based on the classic theory of gravity and Newton's law of motion, the rotation velocity $v(r)$ of any object inside a galaxy should have:

$$\frac{GM(r)}{r^2} = \frac{v(r)^2}{r}, \tag{2.2.1}$$

where r represents the distance between the object and the center of galaxy and $M(r)$ is the total mass of the galaxy within radius r,

$$M(r_0) = 4\pi \int_0^{r_0} r^2 \rho(r) \mathrm{d}r. \tag{2.2.2}$$

The galaxy rotation curve thus can be defined as the profile of rotation velocity against r and measurements show that $v(r)$ becomes approximately constant when far from the center, as illustrated in Figure 2.2.1. To form such distribution, $M(r)$ and $\rho(r)$ need to fulfill:

$$M(r) \propto r, \tag{2.2.3}$$
$$\rho(r) \propto r^{-2}. \tag{2.2.4}$$

More specifically,

$$\rho(r) = \frac{v(r)^2}{4\pi G r^2}(1 + 2\frac{\mathrm{d}\log v(r)}{\mathrm{d}\log r}), \tag{2.2.5}$$

with $v(r) \to$ const. and $\frac{\mathrm{d}\log v(r)}{\mathrm{d}\log r} \to 0$ at large distance.

This is not in agreement with the hypothesis that the galaxy is only composed of luminous matter, e.g. stars and hot gases, and suggests the existence of spherical halos around the galaxy, which consists of invisible matter, e.g. Dark Matter.

2.2.1.2 Bullet clusters

The most decisive evidence of DM is provided by the weak gravitational lensing effect around two colliding galaxy clusters, also known as the bullet cluster. As suggested by General Relativity, mass curves the spacetime and thus distort the path of light. If an extremely massive object, e.g. a black hole, is located between the background light source and the observer, the light will be deflected when passing by the foreground object which serves as a gravitational lens. This leads to visible distortions of light source such as partial rings (arcs), or even Einstein rings [32] when the source-lens-observer

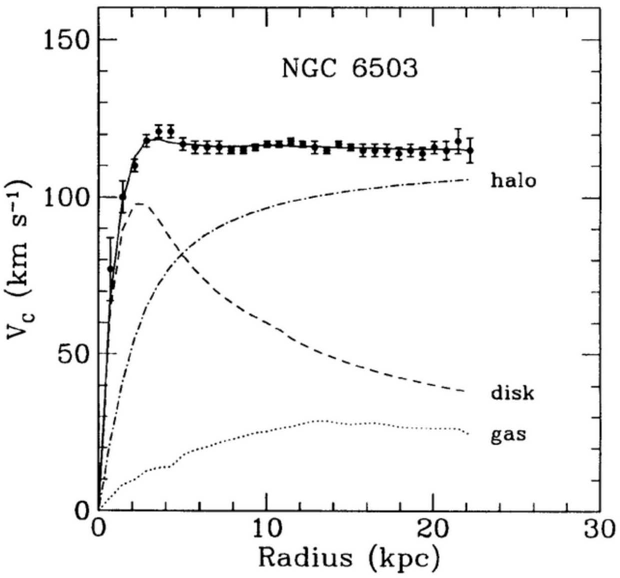


Figure 2.2.1: Rotation curve of the galaxy NGC6503 with a three-parameter dark halo fit (solid curve) to the observed data [27]. The dashed curve shows the rotation curve for visible matter. The dotted curve shows the rotation curve for gas component. The dash-dot curve shows the rotation curve for the dark halo. The parameters of the fit include the mass-to-light ratio of the disk (M/L), the halo core radius (r_c) and the halo asymptotic circular velocity (V_h).

alignment is perfect and the mass distribution of the lens is axially symmetric. However, if the mass of the object in between is not large enough, the bending of spacetime will be too weak to cause detectable distortions on the shape of any individual light source, but instead leads to a tangentially stretch on the background image around the lensing object. This effect, known as the weak gravitational lensing, can only be measured via statistical estimation of the ellipticity of the background galaxies, and allows for the reconstruction of mass distribution of the foreground galaxy.

Figure 2.2.2 shows the reconstructed cluster surface mass density κ of bullet cluster 1E 0657-558 [28]. An $\sim 8\sigma$ deviation is observed for both clusters between the center of mass density κ and the center of mass of the respective plasma cloud determined by X-ray data. This rules out the possibility of Modified Newtonian Dynamics (MOND), which modifies the law of gravity to account for the flat galaxy rotation curves without introducing the theory of Dark Matter, and serves as a direct proof of the existence of DM.

2.2.1.3 Cosmic microwave background

The cosmic microwave background (CMB) is a special kind of electromagnetic radiation emitted at the early stage of the Universe, roughly 379,000 years after the Big Bang. With the continuous expansion of the Universe, protons and electrons cooled down until the temperature fell below their binding energy. Neutral atoms (mostly hydrogen) emerged and consequently photons started to decouple from electrons due to the rapidly dropping density of free electrons. These photons formed a relic radiation, namely the cosmic microwave background radiation, which had propagated freely through the space ever since their last scattering billions of years ago. Up until now, the tempera-

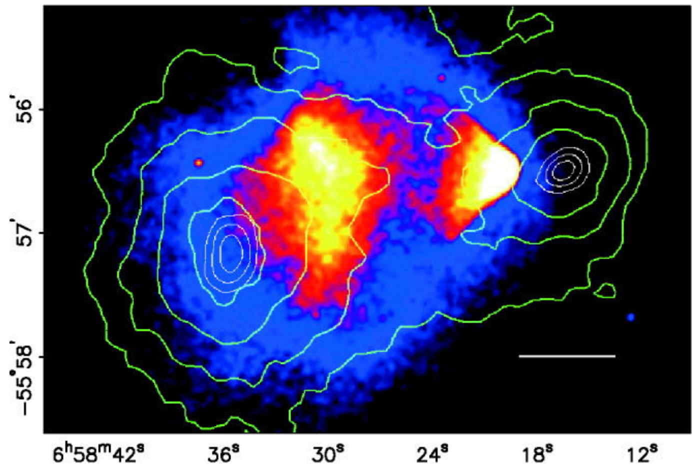


Figure 2.2.2: Illustration of bullet cluster 1E 0657-558 measured by the Chandra X-ray Observatory [28]. The green contours mark the different levels of reconstructed cluster surface mass density κ based on weak gravitational lensing effect. The white contours show the errors on the positions of the κ peaks and correspond to 68.3%, 95.5%, and 99.7% confidence levels. The colored area indicates the mass density of X-ray plasma clouds from the two colliding clusters. Significant deviations between the κ peaks and the mass peaks of plasma clouds can be observed.

ture of the Universe dropped from $\sim$ 3000 K (at the time of decoupling) to $\sim$ 2.7 K, while measurements of the cosmic microwave background show that its exact spectrum corresponds to a blackbody radiation at 2.7255 ± 0.0006 K [29].

The existence of the CMB was first predicted by Georg Gamow, Ralph Alpher and Robert Herman in 1948 [33, 34] and later confirmed by Arno Penzias and Robert Woodrow Wilson at Bell Labs using Echo satellites in 1964 [35]. According to multiple astrophysical experiments, the CMB temperature is proven to be anisotropic with fluctuations at the level of 10^{-5} K. This anisotropy, aside from the instrumental noise of the detector, is the result of many factors combined:

- The **dipole anisotropy**. Due to the relative velocity between the CMB rest frame and the Earth frame, a redshift is expected at the direction backwards to the motion of the solar system while a blueshift is expected at the direction of motion. As shown in Figure 2.2.3, a smooth variation between relatively hot and relatively cold areas can be observed.
- The **primary anisotropy**, which refers to the directional dependence caused by effects occurring before and when the last scattering happened. These effects mainly include gravitational perturbations, that photons scattered from high density regions lost energy to climb up the potential well, intrinsic perturbations, that photons scattered from these regions also needed to be hotter to decouple with matter, and Doppler perturbations, that the baryon-electron-photon plasma with non-zero velocity caused Doppler shift on the photon energy.
- The **secondary anisotropy**, which refers to the directional dependence caused by effects occurring after the last scattering. These effects mainly include CMB lensing, that photons could be deflected when passing through the galaxies, and thermal Sunyaev-Zel'dovich effect, that photons could be inverse Compton scattered when passing through hot gases.

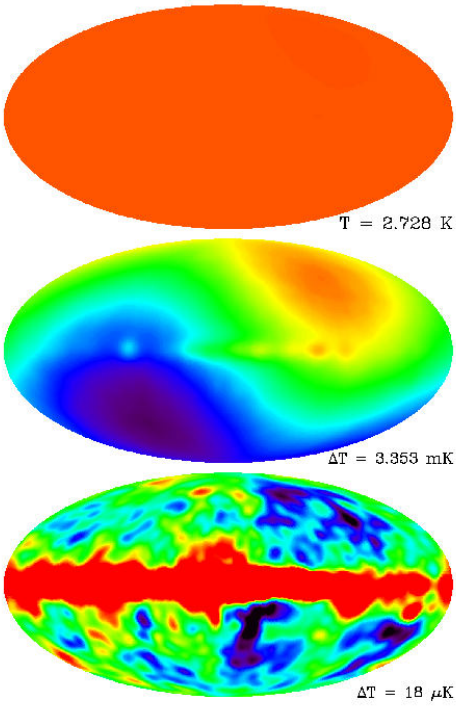


Figure 2.2.3: Four-year DMR sky maps [36] measured at the 53 GHz band and smoothed to an effective angular resolution of 10 degrees. The top map is plotted on a scale from 0 - 4 K, showing the near-uniformity of the CMB temperature. The middle map is plotted on a scale intended to enhance the contrast due to the dipole anisotropy. The bottom map is the CMB distribution after subtracting the dipole component.

The CMB spectrum can be described using the spherical harmonics expansions:

$$\frac{\Delta T}{T} = \sum_{l,m} a_{lm} Y_{lm}(\theta, \phi), \tag{2.2.6}$$

$$C_l^{TT} = \frac{1}{2l+1} \sum_{m=-l}^{m=+l} |a_{lm}|^2, \tag{2.2.7}$$

where C_l^{TT} is defined as the observed angular power spectrum for mode l. The $l = 0$ mode corresponds to the CMB temperature, 2.7255 ± 0.0006 K, while the $l = 1$ mode corresponds to the dipole anisotropy which is measured to be 3.3645 ± 0.0020 K [29]. Figure 2.2.4 shows the full CMB sky map and Figure 2.2.5 shows the power spectrum computed with the Planck 2015 data. Additionally, a Lambda Cold Dark Matter (ΛCDM) model based fit is applied to the observed data, allowing to determine the ΛCDM parameters in high precision. Knowledge about the curvature of the Universe, the baryon density and the Dark Matter density can be extracted from the position of the peaks in the CMB spectrum [31]:

$$\Omega_b h^2 = 0.02226 \pm 0.00023, \tag{2.2.8}$$
$$\Omega_{\mathrm{DM}} h^2 = 0.1186 \pm 0.0020. \tag{2.2.9}$$

According to the ΛCDM model, the energy density for Dark Energy can be inferred:

$$\Omega_{\mathrm{DE}} = 0.692 \pm 0.012, \tag{2.2.10}$$

which leads to the present picture of mass-energy composition of the Universe: 4.9% ordinary matter, 26.8% Dark Matter and 68.3% Dark Energy.

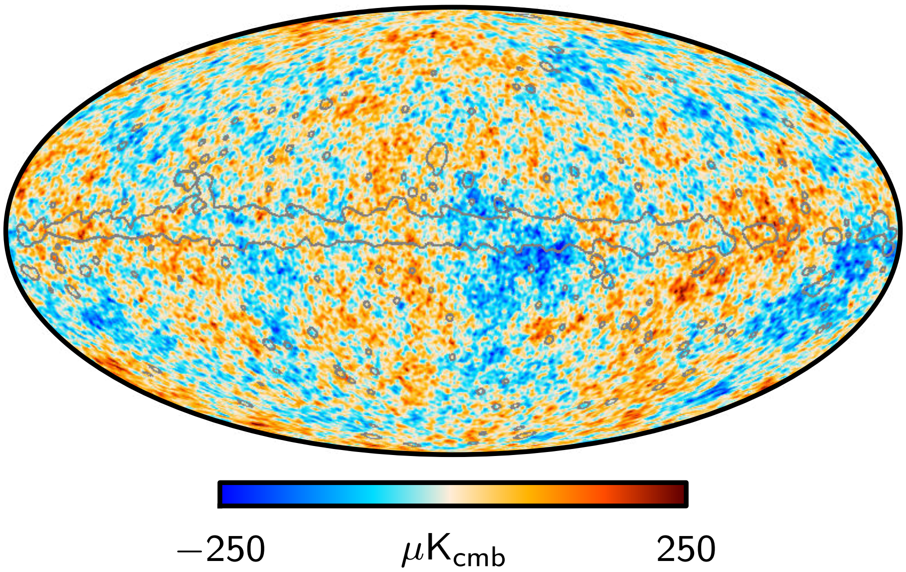


Figure 2.2.4: Commander CMB temperature map derived from the Planck 2015, nine-year WMAP and 408 MHz Haslam observations [30] et al.

2.2.2 Dark Matter candidates

Although the existence of Dark Matter is evident, very little is known about its physical nature. Based on the relativity of the DM particles, the DM hypotheses can be categorized into three types which give distinctive predictions [37]:

- **Cold Dark Matter**. Under this scenario, the Dark Matter particles had already been non-relativistic when the decoupling with ordinary matter took place. This happened at an early phase of cosmic evolution, known as the 'freeze-out' process. Measurements of the CMB radiation and the structure formation of galaxy clusters provide sound support for this hypothesis. Excessive dwarf galaxies and too few empty regions are predicted compared to observation [38, 39], while the supernova feedback may provide a viable solution to the problem [40].
- **Warm Dark Matter**. It is also possible that when the decoupling between ordinary matter and DM happened the DM particles were still relativistic, but later cooled down in unison with the expansion of the Universe and became non-relativistic. This hypothesis can give fitting prediction to the galaxy density [41] and the most promising Warm Dark Matter candidates are sterile neutrinos.

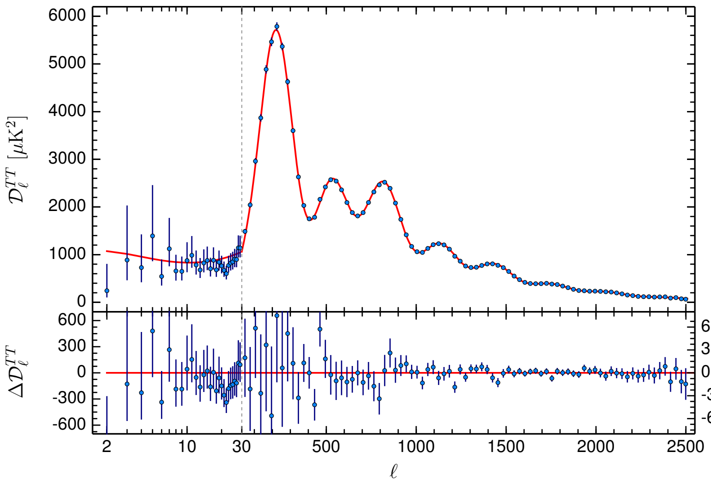


Figure 2.2.5: Planck 2015 CMB spectrum with the base ΛCDM fit to data (red line) [30]. $D_l^{TT} = l(l+1)C_l^{TT}/(2\pi)$ is shown for better visualization. The horizontal scale changes from logarithmic to linear at the 'hybridization' scale, $l = 29$.

- **Hot Dark Matter**. When the DM particles are light enough ($<$ 1 eV), even the cosmic expansion could not cool them down to the non-relativistic state. Neutrinos, which possess almost zero rest mass, are the most commonly known example of Hot Dark Matter. However, simulations have shown that neutrinos alone are not able to account for the inferred DM mass density in the galaxies [42] and the upper limit on the energy fraction of neutrinos is much lower than the expected total energy density of DM in the Universe [23].

At present, the scenario of Cold Dark Matter (CDM) is the most recognized DM hypothesis due to its simplicity and predictiveness. A wide variety of CDM models have arisen and the relevant DM candidates have been proposed during the past decades, with discussions below.

- **Massive Compact Halo Objects** (MACHOs). MACHOs are condensed and non-luminous astronomical bodies composed of baryonic matter, such as black holes, neutron stars and planets unassociated to any planetary systems. Since MACHOs emit (almost) no radiation, they can only be observed via gravitational lensing as explained in Section 2.2.1.2. However, the theory of Big Bang nucleosynthesis and measurements of the cosmic microwave background together put constraints on the abundance of baryonic matter in the Universe, which rule out the possibility for MACHOs to be the major Dark Matter component.

- **Axions**. The theory of axions was first proposed by Roberto Peccei and Helen Quinn in 1977 [43] as a solution to the strong CP problem, that no CP violation is observed in the strong interaction, by introducing a new $U(1)$ symmetry which is spontaneously broken and a complex scalar field, namely the axion field. Axions are predicted to have extremely weak coupling to photons and negligible masses, $m_a \lesssim 1$ eV, while remaining non-relativistic due to Bose-Einstein condensation [44]. These properties make axion one of the leading candidates of CDM. Searches for axions have been performed including the 'light shining through a wall' experiments, e.g. ALPS [45], and the axion haloscope searches, e.g. ADMX [46]. Bounds on axion mass, 10^{-6} eV $< m_a < 10^{-2}$ eV, have been set by the cosmological and astrophysical measurements [47].

- **Weakly Interacting Massive Particles** (WIMPs). WIMPs arise naturally in many BSM extensions, e.g. the lightest supersymmetric particle in Supersymmetry [48], the Kaluza-Klein photon in Universal Extra Dimension [49] and the lightest T-odd particle in Little Higgs model [50]. Compared to axions and axion-like particles, WIMPs, denoted by χ, are assumed to be massive with m_χ ranging from $\mathcal{O}(1\text{ GeV})$ to $\mathcal{O}(10\text{ TeV})$. The upper bound on m_χ at 100 TeV is derived from perturbative unitarity [51]. Additionally, the thermal production of WIMPs with weak scale cross section naturally leads to the observed Dark Matter relic abundance $\Omega_\chi h^2 \sim 0.1$, often referred to as the 'WIMP miracle', which further motivates the experimental searches.

This dissertation is dedicated to searches for Dark Matter under the WIMP scenario. For simplicity, the discussion of DM candidates will be restricted to WIMPs from here on.

2.2.3 Search for Dark Matter via proton-proton collisions

Due to its weak coupling, Dark Matter has only been observed via gravitational interaction at cosmological scale, which does not reveal much of the kinematic characteristics of Dark Matter particles. So far, the only precisely measured property of DM is the relic abundance Ω_χ and this naturally leads to degeneracies when testing theories with rather complex parametrization. In order to have further understandings of interactions between the SM and DM, three experimental approaches are used to shed light on the Dark sector, as illustrated in Figure 2.2.6:

- **Direct detection searches**, aiming to probe the elastic scatterings of DM particles against ordinary matter. The Earth is expected to travel through DM particles from the halos in the Milky Way during its rotation around the Sun. This motivates the direct detection searches which measure the $\mathcal{O}(10\text{ keV})$ scale recoils of nuclei due to the possible scattering with DM particles with masses from $\mathcal{O}(\text{GeV})$ to $\mathcal{O}(\text{TeV})$. The detector used in such experiments is often composed of cryogenic crystals such as germanium, silicon (CDMS [52]) and $CaWO_4$ (CRESST [53]), crystal scintillators such as NaI(TI) (DAMA/LIBRA [54]) or liquid noble gases such as xenon (LUX [55], XENON1T [56]) and argon (DarkSide [57]). Moreover, the interaction between DM particles and nuclei can be either spin-dependent or spin-independent based on the coupling structure. The spin-independent interaction only couples to the nucleus mass, while the spin-dependent interaction also couples to the spin of the nucleus. The combination of direct detection searches covers the full mass range of the WIMP scenario and is ceaselessly pushing the exclusion limits to the 'neutrino floor', a state where the background of neutrino-nucleus scattering becomes significant.

- **Indirect detection searches**, aiming to detect the production of the SM particles via DM self-annihilation. In regions with high DM density, e.g. the center of galaxies, the annihilation of a pair of DM particles is more probable to happen, possibly resulting in γ-rays, neutrino fluxes or charged particle flows. Some of the most well known indirect detection experiments include space detectors like Fermi Large Area Telescope [58], PAMELA [59], ground-based detectors like H.E.S.S. [60] and IceCube [61].

- **Collider searches**, aiming to measure the DM production via the collision of higher energy particles, e.g. protons, electrons and positrons. Generally, three types of analyses are used in search of DM particles: measurements of electroweak observables with high precision, where any deviation from the SM prediction might indicate the loop level contribution from BSM particles, searches for narrow resonances, where the DM mediators produced at the collider decay back into a pair of SM particles, and searches for events with imbalanced total momentum, where the DM pair produced in association with SM particles escapes the detector without leaving a trace. One of the most famous particle colliders in the world is the Large Hadron Collider (LHC) at Conseil Européen pour la Recherche Nucléaire (CERN) near Geneva, Switzerland. Seven experiments in total take place at the LHC, while three of them, including ATLAS, CMS and LHCb, are probing DM with the collision data [62, 63, 64].

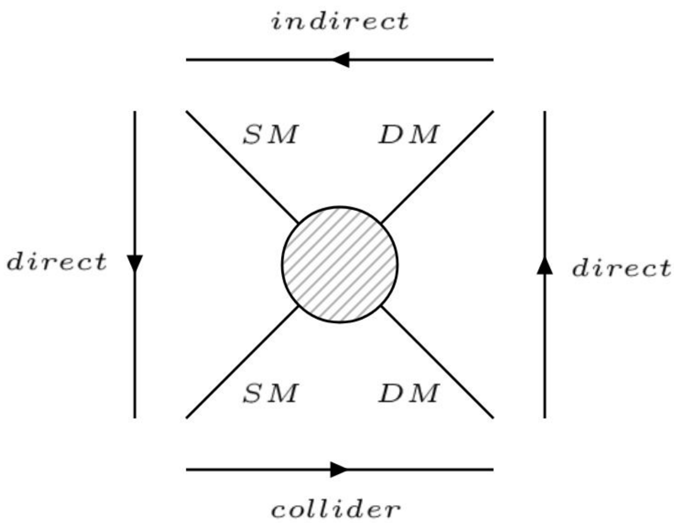


Figure 2.2.6: A simple illustration of the physical processes probed by the three types of DM searches. From top to bottom (bottom to top) is the scattering between the SM and DM particles (direct detection search). From right to left is the SM production via DM self-annihilation (indirect detection search). From left to right is the DM pair production at high energy colliders (collider search).

This dissertation is mainly composed of two searches for DM in events with imbalanced total momentum using data collected by the ATLAS detector at the LHC. For DM particles, their weak coupling to the SM fields makes it almost impossible for detectors to catch their traces, just like for neutrinos. However, in hadron colliders, while the longitudinal momentum of any collision event remains unknown, the net momentum in the transverse direction should always be zero. As a result, the presence of DM can be inferred from any apparent imbalance in the transverse plane with respect to the interaction point of the collision. Such imbalance represents the total transverse momentum of invisible particles and can be computed as the negative vector sum of the momenta of all visible particles in the final state, the 'missing transverse momentum', or $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$, of which the magnitude is often referred to as the 'missing transverse energy', $E_{\mathrm{T}}^{\mathrm{miss}}$. Note that this only happens when the DM mediator recoils against other SM particles, otherwise it will result in a pair of back-to-back DM particles and a negligible $E_{\mathrm{T}}^{\mathrm{miss}}$. The search for DM production at colliders can thus be transformed into searches for final states with recoiled SM particles and large $E_{\mathrm{T}}^{\mathrm{miss}}$, also known as the $E_{\mathrm{T}}^{\mathrm{miss}} + X$

final states.

Compared to the resonance searches of DM mediators which look for a narrow peak in the invariant mass distribution of two SM particles, the $E_{\mathrm{T}}^{\mathrm{miss}} + X$ searches focus on small discrepancies on the shape of certain variables, e.g. $E_{\mathrm{T}}^{\mathrm{miss}}$ and the significance of $E_{\mathrm{T}}^{\mathrm{miss}}$ (see Section 3.3.4.1), between the SM contribution and the observed data. The experimental feasibility of the $E_{\mathrm{T}}^{\mathrm{miss}} + X$ searches and the multiplicity on the choice of SM particle X demonstrate the huge potential of such analyses by allowing to constrain a broad range of DM models in the corresponding final states. Various definitions of X have been explored by the ATLAS and the CMS experiments, including a jet [65, 66], a photon [67, 68], a leptonically decaying Z boson [69, 70], a hadronically decaying W or Z boson [1, 66], a heavy-flavor quark [62, 63], a heavy-flavor quark pair [71, 63] and a Higgs boson [72, 73].

2.2.4 The simplified models

A crucial step in the search for Dark Matter is choosing a fitting description of signals which allows for the comparison between data and prediction, and also between results from different experiments. Two approaches are commonly used to build such theoretical models, namely the effective field theories (EFT) and the complete theories.

The EFT framework introduces a set of non-renormalizable operators which parametrize the coupling between the Standard Model and the Dark Matter sector in terms of one effective scale Λ and one Dark Matter mass m_χ. By treating the DM particle as the only new accessible degree of freedom compared to the SM, the mediators are integrated out and this allows the EFTs to give predictions in a more model-independent way. Furthermore, for a given choice of the spin of the DM particle, one can enumerate every possible operator which couples to the SM fields and place limits on each of them. The major disadvantage of the EFTs lies within its intrinsic dependence on the energy scale. When the energy exceeds the effective scale Λ, the perturbativity of the theory will be severely challenged and the high dimensional correction to the scattering amplitudes will be almost comparable to the lower order contributions. Such ultraviolet (UV) incompleteness harms the prediction power of the EFTs especially at collider searches, where the energy exchanges between partons is often at the TeV scale, while the expect exclusion on Λ may only be at $\mathcal{O}(100\ \mathrm{GeV})$ [74].

The complete theory, on the other hand, provides a complete extension of the Standard Model and tends to be UV complete up to some very large scale, e.g. $\mathcal{O}(10^{16}\ \mathrm{GeV})$. One of the most famous complete theories containing DM candidates is the Minimal Supersymmetric Standard Model (MSSM) which predicts the existence of neutralinos. These theories are generally able to give precise predictions in the fields of cosmology, astrophysics and particle physics, while at the same time grant compatibility to the results measured in different experiments. However, unlike the EFTs, the complete theories necessarily involve many free parameters and therefore pose great difficulties on drawing any general or less model-dependent conclusions. The structure of the complete theories are so rich that it is nearly impossible to extract the underlying physics from a finite number of measurements, leading to the so-called 'inverse problem' [75].

Inspired by the simplicity of the EFTs and the predictiveness of the complete theories, the simplified models of Dark Matter are constructed with the inclusion of the minimal complete extensions to the SM sector. This comes with the benefit that while a simplified model is able to correctly describe the kinematic properties of DM production up to a rather high energy scale, the complexity of its parameter space still remains at an acceptable level. Compared to an EFT which integrates out all but the Dark Matter particle, a simplified model can be viewed as the approximation of a complete theory which integrates out all but the lightest Dark Matter sector, typically composed of one (meta)stable DM candidate and one mediator. Moreover, the simplified model Lagrangians should be renormalizable and also consistent with Lorentz invariance and gauge symmetries.

Based on the spin of the mediator, the simplified models used in the searches for Dark Matter presented in this dissertation can be divided into two categories:

- **spin-1**: vector and axial-vector mediator simplified models;
- **spin-0**: scalar and pseudo-scalar mediator simplified models.

2.2.4.1 Vector and axial-vector mediator simplified models

Consider the interaction between the SM quarks and DM particle in the form of a Dirac fermion through the exchange of a spin-1 mediator, denoted by Z', in s-channel. The Lagrangian of such a model can be written as following [76]:

$$\mathcal{L}_V \supset \frac{1}{2}m_{Z'}^2 Z'_\mu Z'^\mu - m_\chi \bar{\chi}\chi - g_\chi Z'_\mu \bar{\chi}\gamma^\mu \chi - g_q \sum_{q=u,d,s,c,b,t} Z'_\mu \bar{q}\gamma^\mu q, \tag{2.2.11}$$

$$\mathcal{L}_A \supset \frac{1}{2}m_{Z'}^2 Z'_\mu Z'^\mu - m_\chi \bar{\chi}\chi - g_\chi Z'_\mu \bar{\chi}\gamma^\mu \gamma_5 \chi - g_q \sum_{q=u,d,s,c,b,t} Z'_\mu \bar{q}\gamma^\mu \gamma_5 q, \tag{2.2.12}$$

where $\mathcal{L}_V$ and $\mathcal{L}_A$ stand for the vector and the axial-vector models, respectively, $m_{Z'}$ and m_χ are the masses of the mediator and the DM particle, g_χ is the coupling strength to the Dark sector and g_q is the coupling to the SM sector. The universality of g_q between different generations guarantees that the above Lagrangians are minimal flavor violating (MFV), meaning that all higher dimensional operators constructed from the SM and DM fields are invariant under the flavor and the CP transformations [77]. In particular, all flavor- and CP-violating transitions should be determined by the Cabibbo-Kobayashi-Maskawa (CKM) matrix.

The minimal decay width of the mediator can be defined as the summed partial widths of all decays to DM and quarks which are kinematically accessible. For vector mediators, the partial widths are in the form of:

$$\Gamma_V^{\chi\bar{\chi}} = \frac{g_\chi^2 m_{Z'}}{12\pi}(1-\frac{4m_\chi^2}{m_{Z'}^2})^{\frac{1}{2}}(1+\frac{2m_\chi^2}{m_{Z'}^2}), \tag{2.2.13}$$

$$\Gamma_V^{q\bar{q}} = \frac{g_q^2 m_{Z'}}{12\pi}(1-\frac{4m_q^2}{m_{Z'}^2})^{\frac{1}{2}}(1+\frac{2m_q^2}{m_{Z'}^2}). \tag{2.2.14}$$

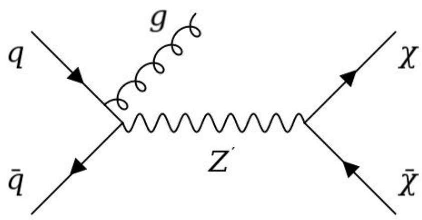
And for axial-vector mediators:

$$\Gamma_A^{\chi\bar{\chi}} = \frac{g_\chi^2 m_{Z'}}{12\pi}(1 - \frac{4m_\chi^2}{m_{Z'}^2})^{\frac{3}{2}}, \tag{2.2.15}$$

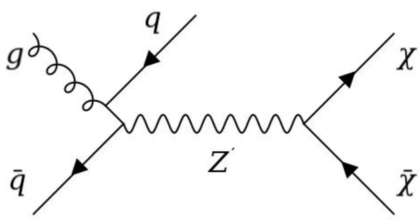
$$\Gamma_A^{q\bar{q}} = \frac{g_q^2 m_{Z'}}{12\pi}(1 - \frac{4m_q^2}{m_{Z'}^2})^{\frac{3}{2}}. \tag{2.2.16}$$

It is obvious that for $m_{Z'} < 2m_{\chi,q}$, the corresponding partial width simply vanishes and only off-shell decays are allowed.

As an example, Figure 2.2.7 shows two of the possible Feynman diagrams for the DM production in association with a quark or a gluon under the description of a vector or axial-vector mediator simplified model.



(a)



(b)

Figure 2.2.7: Feynman diagrams of DM production in association with a quark or a gluon under the description of a vector or axial-vector mediator simplified model.

2.2.4.2 Scalar and pseudo-scalar mediator simplified models

Consider the interaction between the SM quarks and DM particle in the form of a Dirac fermion through the exchange of a spin-0 mediator, denoted by ϕ, in s-channel. The Lagrangian of such a model can be written as following [76]:

$$\mathcal{L}_S \supset -\frac{1}{2}m_\phi^2\phi^2 - m_\chi\bar{\chi}\chi - g_\chi\phi\bar{\chi}\chi - g_q\frac{\phi}{\sqrt{2}}\sum_{q=u,d,s,c,b,t} g_Y^q\bar{q}q, \tag{2.2.17}$$

$$\mathcal{L}_P \supset -\frac{1}{2}m_\phi^2\phi^2 - m_\chi\bar{\chi}\chi - ig_\chi\phi\bar{\chi}\gamma_5\chi - ig_q\frac{\phi}{\sqrt{2}}\sum_{q=u,d,s,c,b,t} g_Y^q\bar{q}\gamma_5 q, \tag{2.2.18}$$

where $\mathcal{L}_S$ and $\mathcal{L}_P$ stand for the scalar and the pseudo-scalar models, respectively, m_ϕ and m_χ are the masses of the mediator and the DM particle, g_χ is the coupling strength to the Dark sector, g_q is the coupling to the SM sector and g_Y^q are the SM quark Yukawa couplings. Similarly, the above Lagrangians are compatible with the MFV criterion.

Additionally, the scalar (pseudo-scalar) mediator can decay into a pair of gluons, e.g. via t-quark triangle loop, thus another term needs to be considered for minimal width computation. For scalar mediators, the partial widths are in the form of:

$$\Gamma_S^{\chi\bar{\chi}} = \frac{g_\chi^2 m_\phi}{8\pi}(1 - \frac{4m_\chi^2}{m_\phi^2})^{\frac{3}{2}}, \tag{2.2.19}$$

$$\Gamma_S^{q\bar{q}} = \frac{3g_q^2 g_Y^{q2} m_\phi}{16\pi}(1 - \frac{4m_q^2}{m_\phi^2})^{\frac{3}{2}}, \tag{2.2.20}$$

$$\Gamma_S^{gg} = \frac{\alpha_s^2 g_q^2 m_\phi^3}{32\pi^3 v^2}|f_S(\frac{4m_t^2}{m_\phi^2})|^2, \tag{2.2.21}$$

where v is the vacuum expectation value of the Higgs field and $f_S(\tau)$ is defined as:

$$f_S(\tau) = \tau\left[1 + (1-\tau)\arctan^2(\frac{1}{\sqrt{\tau - 1}})\right]. \tag{2.2.22}$$

And for pseudo-scalar mediators:

$$\Gamma_P^{\chi\bar{\chi}} = \frac{g_\chi^2 m_\phi}{8\pi}(1 - \frac{4m_\chi^2}{m_\phi^2})^{\frac{1}{2}}, \tag{2.2.23}$$

$$\Gamma_P^{q\bar{q}} = \frac{3g_q^2 g_Y^{q2} m_\phi}{16\pi}(1 - \frac{4m_q^2}{m_\phi^2})^{\frac{1}{2}}, \tag{2.2.24}$$

$$\Gamma_P^{gg} = \frac{\alpha_s^2 g_q^2 m_\phi^3}{32\pi^3 v^2}|f_P(\frac{4m_t^2}{m_\phi^2})|^2, \tag{2.2.25}$$

where $f_P(\tau)$ is defined as:

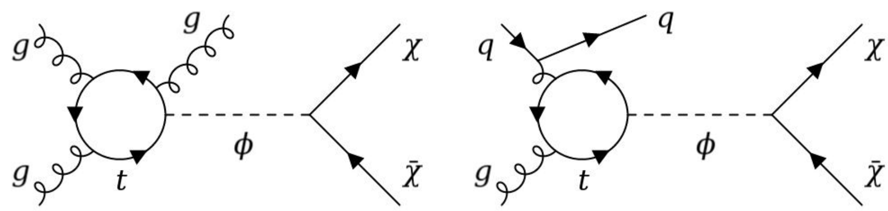
$$f_P(\tau) = \tau\arctan^2(\frac{1}{\sqrt{\tau - 1}}). \tag{2.2.26}$$

As an example, Figure 2.2.8 shows two of the possible Feynman diagrams for the DM production in association with a quark or a gluon under the description of a scalar or pseudo-scalar mediator simplified model.

2.2.4.3 Motivation of the $E_{\mathrm{T}}^{\mathrm{miss}}$ + V and $E_{\mathrm{T}}^{\mathrm{miss}}$ + $t\bar{t}$ search

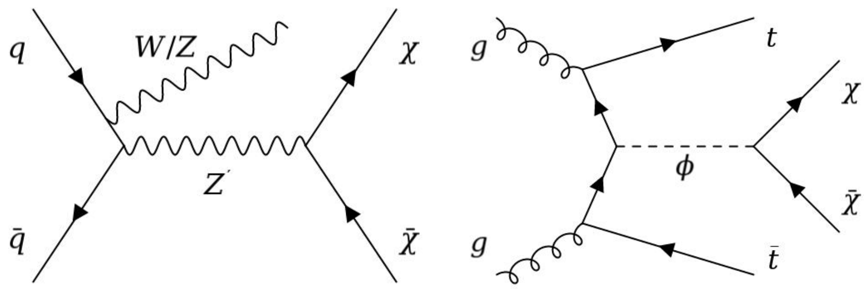
To test various simplified models of Dark Matter regardless of the spin of the mediator, two searches performed will be discussed in this thesis: the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search and the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search, with the corresponding Feynman diagrams shown in Figure 2.2.9.

The vector or axial-vector mediator simplified models can be examined in processes where the DM particles are produced in association with a SM vector boson originating from the initial state radiation. This is often referred to as $E_{\mathrm{T}}^{\mathrm{miss}} + V$ final state and three different analyses can be carried out based on the decay modes of the vector boson: hadronic ($W/Z \to qq$), semi-leptonic ($W \to l\nu$) and dileptonic ($Z \to ll$). Compared



(a) (b)

Figure 2.2.8: Feynman diagrams of DM production in association with a quark or a gluon under the description of a scalar or pseudo-scalar mediator simplified model.



(a) (b)

Figure 2.2.9: (a) Feynman diagram of DM production in association with a W or Z boson under the description of vector or axial-vector mediator simplified model. (b) Feynman diagram of DM production in association with a pair of top quarks under the description of scalar or pseudo-scalar mediator simplified model.

to the leptonic decay modes, the hadronic channel benefits from its large cross section, which allows for finer tuning of event selections while keeping enough event yields in the signal regions. Additionally, in $W/Z \to qq$ final states, the $E_{\mathrm{T}}^{\mathrm{miss}}$ is expected to purely come from the transverse momentum of the escaping DM particles, which is not the case for $W \to l\nu$. This helps to avoid extra layers of complexity in the design of an analysis strategy. Despite the fact that much lower background contamination is expected in the leptonic channels, the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) DM search is still able to provide the best sensitivity among the three decay modes, second only to the $E_{\mathrm{T}}^{\mathrm{miss}}$+jet and the $E_{\mathrm{T}}^{\mathrm{miss}} + \gamma$ searches.

For models with scalar or pseudo-scalar mediators, the MFV criterion implies that the interaction between the mediator and the SM particle is proportional to the mass of the SM particle via Yukawa-type couplings, which leads to sizeable cross sections of the production of DM in association with heavy-flavor quarks, e.g. $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$. Similar to $E_{\mathrm{T}}^{\mathrm{miss}} + V$, three types of analyses can be considered dependent on the decay modes of the top quark pair: fully-hadronic ($t\bar{t} \to bqqbqq$), semi-leptonic ($t\bar{t} \to bqqbl\nu$) and dileptonic ($t\bar{t} \to bl\nu bl\nu$), while the fully-hadronic channel has the largest branching ratio of around 46%. Searches in final state events characterized by fully-hadronically decaying $t\bar{t}$ and large missing transverse energy have been performed in ATLAS targeting the supersymmetric (SUSY) partner of the top quark, as documented in Ref. [78]. However, these searches are not optimal for DM signals due to the different kinematic properties predicted by the SUSY theory and the simplified models. The $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) DM search is thus designed to exploit regions where rather low $E_{\mathrm{T}}^{\mathrm{miss}}$ is expected, which are not fully covered by the previous analyses.

Chapter 3

The ATLAS experiment at the Large Hadron Collider

An overview of the Large Hadron Collider and the ATLAS experiment is presented in this chapter. Section 3.1 introduces the Large Hadron Collider and Section 3.2 details the structure of the ATLAS detector with its sub-systems. The reconstruction of various physical objects in the ATLAS experiment is discussed in Section 3.3.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [79, 80, 81] at CERN, based near Geneva, Switzerland, has been serving as the world's largest and most powerful particle accelerator since the beginning of its operation in 2008. With a circumference of 26.7 km, the LHC stores protons and heavy ions accelerated to unprecedentedly high energies and collides them in the tunnels located up to 175 meters beneath ground and built for its predecessor, the Large Electron-Positron collider (LEP) [82, 83] which had been operational until 2000.

The collisions of particles happen at four different locations along the LHC ring, where four of the major LHC experiments take place: A Toroidal LHC ApparatuS (ATLAS), a Compact Muon Solenoid (CMS), A Large Ion Collider Experiment (ALICE) and a Large Hadron Collider beauty (LHCb). The ATLAS and the CMS experiments are both multi-purpose high-luminosity experiments, while ALICE and LHCb are designed to cover more specific physical topics with lower luminosities. In general, the ALICE experiment focuses on Pb-Pb collisions as well as the quark-gluon plasma and the LHCb experiment targets interactions involving b-physics. Three other experiments, namely a TOTal Elastic and diffractive cross section Measurement (TOTEM), a Monopole and Exotics Detector at the LHC (MoEDAL) and a Large Hadron Collider forward (LHCf), also use the LHC data but for much more specialized research purposes.

The LHC accelerates and collides both protons and ions. For proton beams, before the particles reach the designed final energy of up to 7 TeV, they have already been accelerated in a succession of machines, including the linear accelerator LINAC 2, the Proton Synchrotron Booster (BOOSTER), the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS). The first accelerator in the system, LINAC 2, increases the energy of the protons from ionized hydrogen atoms to 50 MeV and injects them into the next accelerator, the Proton Synchrotron Booster, where the protons are brought to the energy of 1.4 GeV. The third accelerator is the Proton Synchrotron, where the protons are accelerated to 25 GeV and start to be squeezed into a condensed beam structure, and the last accelerator before the LHC is the Super Proton Synchrotron, the second largest circular accelerator in the LHC complex with a circumference of 7 km, where the beams are further accelerated and injected into the LHC ring with an energy of 450 GeV in two opposite injection points, close to where the ALICE and the LHCb experiment take place. A more detailed illustration of the LHC complex can be found in Figure 3.1.1.

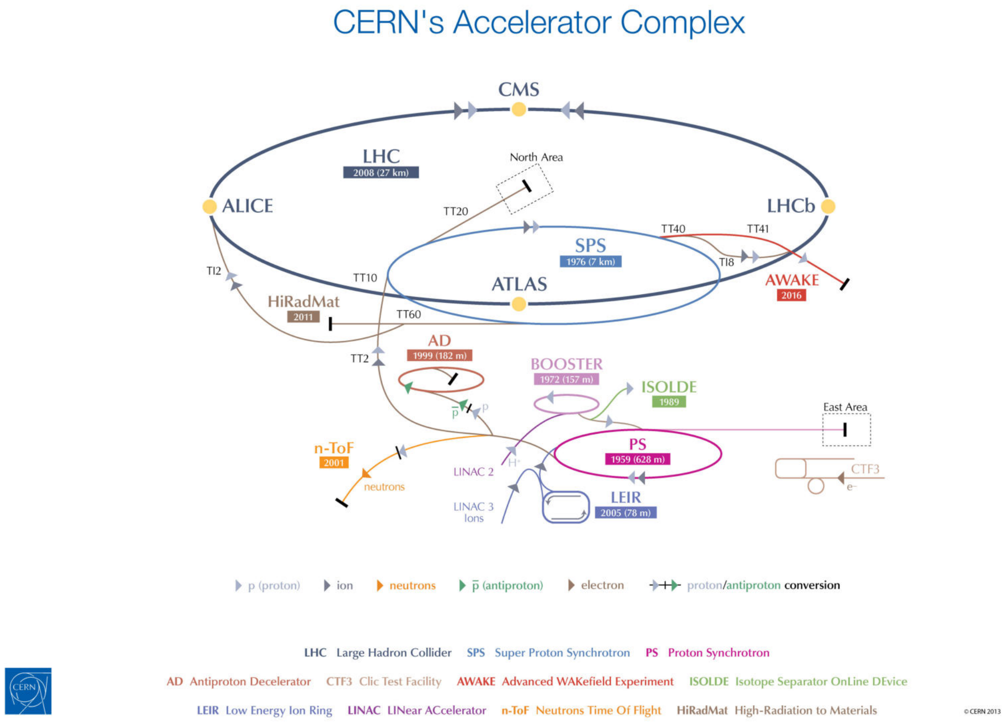


Figure 3.1.1: Illustration of the LHC accelerator complex. The beam acceleration starts at LINAC 2 and then through the Proton Synchrotron Booster (BOOSTER), the Proton Synchrotron (PS), the Super Proton Synchrotron (SPS) and finally the LHC [84]. The beams cross at four interaction points located along the LHC ring where the four major experiments take place, while the other three smaller experiments are not shown in the plot.

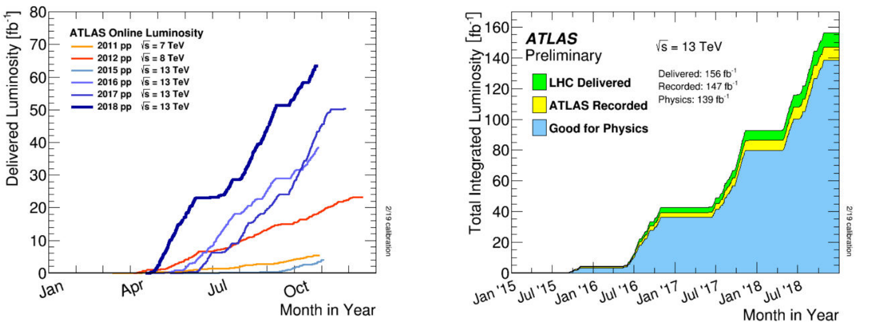
The proton beams consist of separated units called bunches, each formed by $\sim 1.15 \times 10^{11}$ protons and with a spacing of 25 ns (approximately 7.5 m) in between. The acceleration of a single bunch of protons roughly takes 17 seconds from rest to the energy of 450 GeV, while it takes $\sim$ 20 minutes to accelerate all bunches into the LHC ring. The size and the crossing frequency of bunches together determine the instantaneous luminosity ($\mathcal{L}$) of collisions given the assumption of Gaussian beams:

$$\mathcal{L} = \frac{N_b^2 n_b f_r \gamma_r}{4\pi\epsilon_n\beta^\star} F(\theta), \tag{3.1.1}$$

where $N_b \sim 1.15 \times 10^{11}$ is the number of protons per bunch, n_b is the number of bunches per beam, f_r is the revolution frequency, γ_r is the relativistic Lorentz factor, $4\pi\epsilon_n\beta^\star$ represents the overlap of the two crossing beams and $F(\theta)$ denotes a geometric reduction factor due to the crossing angle θ between beams at the interaction point, defined as [85]:

$$F(\theta) = \frac{1}{\sqrt{1 + (\frac{\theta\sigma_z}{2\sigma_{x,y}})^2}}, \tag{3.1.2}$$

where σ_z ($\sigma_{x,y}$) is the longitudinal (transversal) Gaussian width of the colliding beams.



(a) (b)

Figure 3.1.2: (a) Integrated luminosity versus time delivered to ATLAS during stable beams for high energy pp collisions in 2011 - 2018 and (b) integrated luminosity versus time delivered to ATLAS (green), recorded by ATLAS (yellow), and certified to be good quality data (blue) during stable beams for pp collisions at 13 TeV center-of-mass energy from 2015 to 2018 [86].

Figure 3.1.2 (a) shows the integrated luminosity, defined as $\int \mathcal{L}dt$, delivered to ATLAS in 2011 - 2018. Then the expected number of collision events is given by:

$$N_{\text{events}} = \sigma \int \mathcal{L}dt, \tag{3.1.3}$$

with σ representing the total cross section. The integrated luminosity delivered to ATLAS (green), recorded by ATLAS (yellow) and certified to be good quality data (blue) during pp collision at 13 TeV center-of-mass energy from 2015 to 2018, often referred to as the LHC Run 2 data-taking period, are shown in Figure 3.1.2 (b).

Due to the high luminosity of beam collision, multiple pp interactions may happen within a single bunch crossing, resulting in that every event recorded by the detector may contain final products from other collisions, which is called pile-up. This leads to additional energy deposits in the detector and can be divided into in-time pile-up and out-of-time pile-up, depending on whether the additional interactions take place in the same bunch crossing or not, respectively. In principle, increasing N_b causes higher rate of in-time pile-up, while increasing n_b results in less bunch separation and thus higher rate of out-of-time pile-up. The distribution of mean number of interactions per crossing, $\langle \mu \rangle$, recorded by ATLAS is shown in Figure 3.1.3.

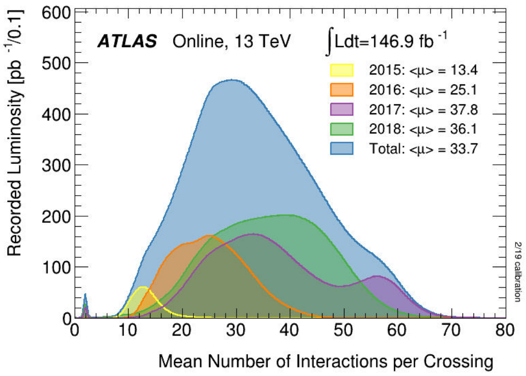


Figure 3.1.3: Luminosity weighted distribution of the mean number of interactions per crossing recorded by ATLAS for pp collisions at 13 TeV center-of-mass energy from 2015 to 2018 [86].

3.2 The ATLAS detector

The ATLAS detector is a multi-purpose detector located at one of the four interaction points of the LHC and allows for exploration of a wide range of physical topics, including precise measurements of the SM processes as well as searches for new physics beyond the SM, e.g. DM production at high energies. With a cylindrical shape of 44 m in length and 25 m in diameter, the ATLAS detector covers almost 4π radians of solid angle around the central interaction point and is considered to be the largest particle detector ever built in the history with a weight of approximately 7000 tons.

The ATLAS detector consists of several layers of sub-detector systems which are designed to identify and measure the kinematic properties of collision products. There are four major components of the detector. As illustrated in Figure 3.2.1, the innermost layer belongs to the Inner Detector (ID), which is composed of three sub-detectors (see Section 3.2.2) and is used to measure the charge and momentum of charged particles. Enclosing the ID is a thin solenoid which provides a magnetic field of 2 T and causes the bending of charged particle trajectories, allowing to measure the momentum of particles by the curvature of their tracks. The Calorimeter system, composed of a electromagnetic calorimeter and a hadronic calorimeter (see Section 3.2.3), absorbs the energy of charged and neutral particles traversing the ID and measures their energy deposits. The outermost part of the detector, the Muon Spectrometer (MS), is designed to determine the position and energy of muons with the help from a 4 T magnetic field, provided by a long barrel toroid and two end-cap toroids surrounding the spectrometer.

Figure 3.2.2 is a schematic illustration of the ATLAS detector in the transverse plane where different types of particles interact with the detector in different layers. Neutrinos and other invisible particles (e.g. DM particles) are expected to have very weak

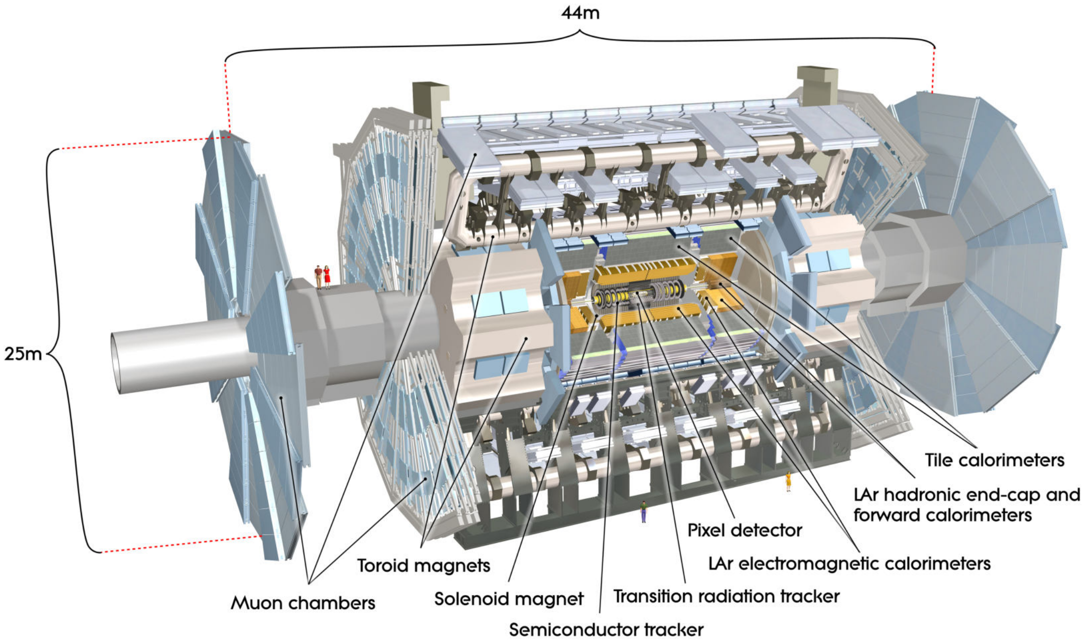


Figure 3.2.1: An illustration of the whole ATLAS detector [87] which consists of four major components, namely the Inner Detector (tracking system), the Calorimeter (calorimeter system), the Muon Spectrometer (muon system) and the magnet system.

interaction with the detector material and thus escape the detection, resulting in non-zero $E_{\mathrm{T}}^{\mathrm{miss}}$ as discussed in Section 2.2.3.

Due to the limitations on the capacity of data storage, ATLAS uses a trigger and data acquisition system (TDAQ) to reduce the event rate from 40 MHz to 1 kHz and filter the most interesting inelastic collision events for physical analyses, which will be discussed in Section 3.2.5.

For convenience, a right-handed coordinate system with its origin at the interaction point is applied in ATLAS. The z-axis is defined along the beam direction, with the x-axis pointing towards the center of the LHC ring and the y-axis pointing upwards. Spherical coordinates (θ, ϕ) are used, where the azimuthal angle ϕ is measured on the x-y plane and the polar angle θ is measured from the z-axis. Since θ is not Lorentz invariant, a pseudorapidity η is introduced as $\eta \equiv -\ln\tan(\theta/2)$ so that the difference in η between two objects, $\Delta\eta$, remains invariant under the Lorentz transformation. The angular distance in this new coordinate system can be then defined as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

3.2.1 Magnet system

The ATLAS magnet system is designed to provide strong magnetic fields deflecting the trajectories of charged particles and allow for precise measurements of particle momenta. Four superconducting magnets are constructed to form the magnet system: a central solenoid magnet which provides a magnetic field of 2 T and bends the charged particles inside the Inner Detector, a long barrel toroid magnet and two end-cap toroid magnets which provide a magnetic field of 4 T and bend the muon tracks inside the Muon Spectrometer.

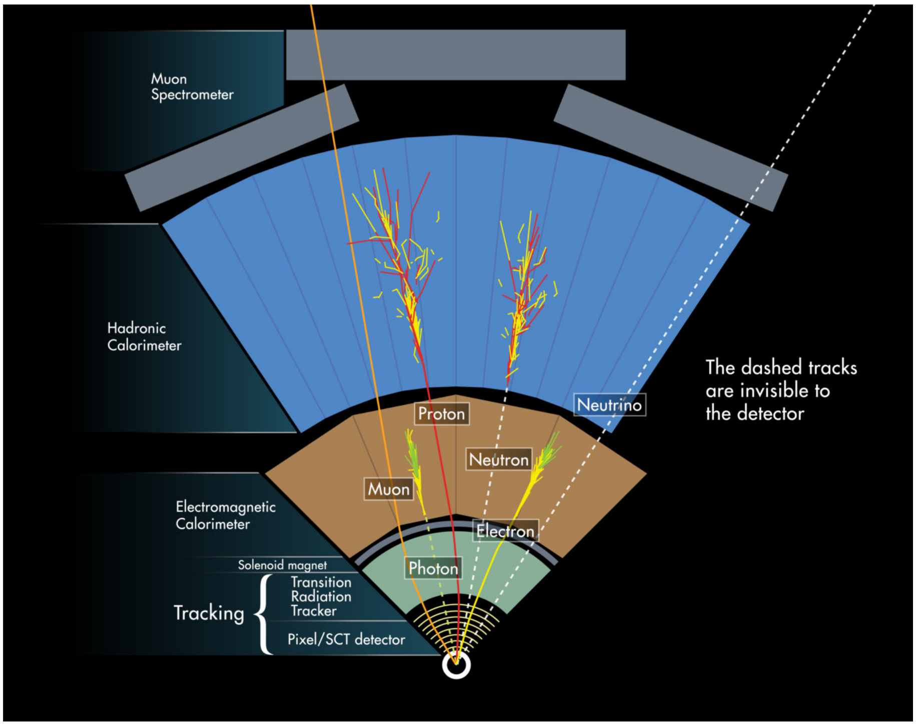


Figure 3.2.2: An illustration of the ATLAS detector in the transverse plane where different types of particles interact with the detector material in different layers [88]. The toroids are not plotted.

The central solenoid, with a length of 5.3 m, a diameter of 2.4 m and 4.5 cm in thickness, is located between the Inner Detector and the Electromagnetic Calorimeter and is composed of 9 km of superconducting wire. The barrel toroid has a length of 25.3 m and an outer diameter of 20.1 m. The two end-cap toroids are inserted in the barrel at each end, with an diameter of 10.7 m and an axial length of 5.0 m. The three toroid magnets are all cooled down to 4.5 K using liquid helium, while the central solenoid is designed to be indirectly cooled down by the cold mass of the liquid argon Electromagnetic Calorimeter surrounding it, thus minimizing its thickness.

An illustration of the layout of the ATLAS magnet system is in Figure 3.2.3.

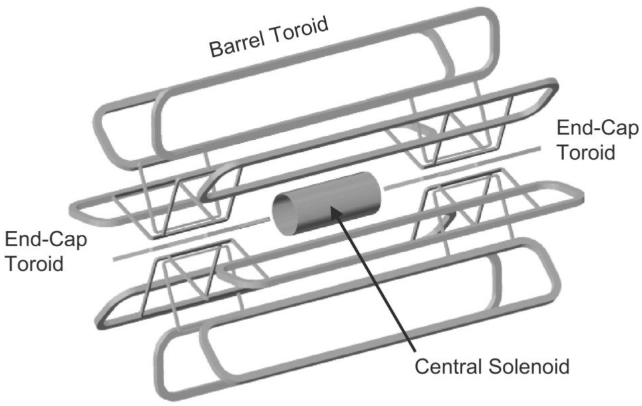


Figure 3.2.3: An illustration of the layout of the ATLAS magnet system [89].

3.2.2 Tracking system

The ATLAS Inner Detector (ID) is designed to identify the charged particles with high accuracy and measure their momenta with high resolution within the pseudorapidity range of $|\eta| < 2.5$. Charged particle tracks with a transverse momentum above 0.5 GeV are reconstructed and then used to rebuild the primary and secondary vertices, which can be taken as input for pile-up suppression as well as b- and τ-hadron identification. Figure 3.2.4 (a) shows the general layout of the ID and Figure 3.2.4 (b) shows its three major sub-detector layers: the Pixel Detector (PD), the Semi-Conductor Tracker (SCT) and the outermost Transition Radiation Tracker (TRT). Note that there is another sub-system in the innermost layer of the ID, namely the Insertable B-Layer (IBL). It is included as part of the ATLAS Run 2 upgrade and is designed to improve the track and vertex reconstruction performance at higher luminosities and under higher pile-up conditions.

As discussed in Section 3.2.1, all sub-detector systems are fully contained within a homogeneous magnetic field of 2 T produced by the central solenoid.

3.2.2.1 Pixel Detector

The Pixel Detector is composed of three cylindrical layers in the central region and three disks in each end-cap region. Since the PD is the closest sub-detector system to the beam pipe, extremely high particle fluxes are expected and a high granularity of pixel modules is required. Semiconducting sensors made of silicon are used to build the modules, each with a cell size of 50×400 μm^2 and a thickness of 250 μm. This provides an intrinsic spatial resolution of 10 μm per layer in the R-ϕ plane transverse to the beam and 115 μm along the z-axis (R-direction) for the barrel (end-cap) layers. The electrical signals from the pixels are then read out via 80.4 million channels.

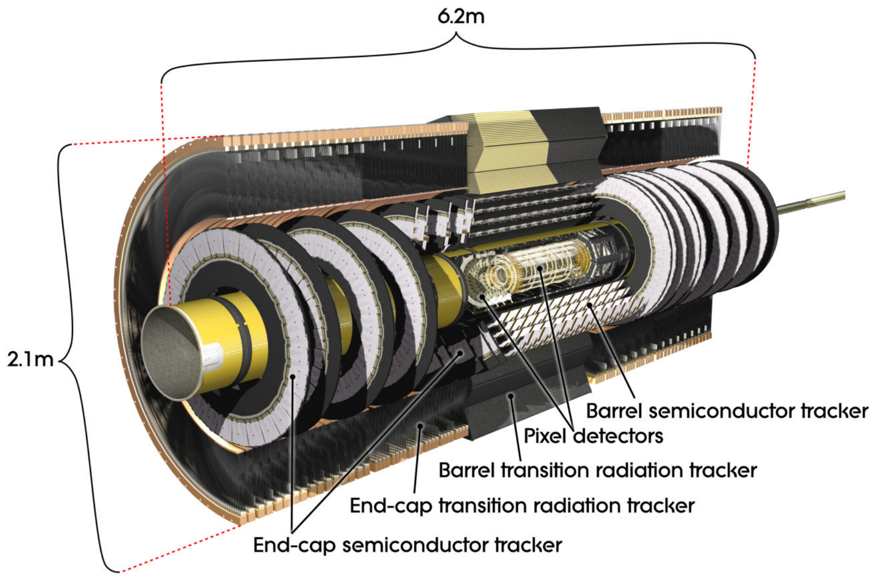
A fourth Insertable B-Layer was installed between a new Beryllium beam pipe and the inner pixel layer in the barrel region during the LHC shutdown between Run 1 and Run 2 and aims to recover the loss of sensitivity due to radiation damage with its additional 8 million pixels. Each pixel in the IBL has a cell size of 50×250 μm^2, providing a spatial resolution of 8 μm in the R-ϕ plane and 40 μm along the z-axis. The IBL upgrade allows for better reconstruction of tracks and vertices and leads to improved b-jet identification [92].

3.2.2.2 Semi-Conductor Tracker

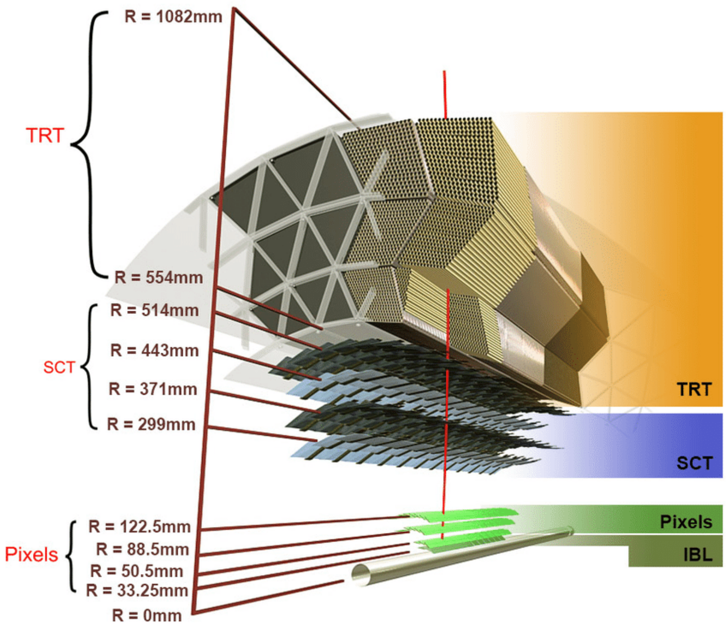
The Semi-Conductor Tracker lies within the middle layer of the ID and consists of four layers of silicon strips in the barrel region and nine disks in each end-cap region. All strips have a length of 12 cm with a constant pitch of 80 μm and are mounted back-to-back in pairs with a 40 mrad stereo angle between two nearby layers, forming the so-called double layer structure. The spatial resolution of the SCT is 17 μm in the R-ϕ plane and 580 μm along the z-axis (R-direction) for the barrel (end-cap) layers. 6.3 million read-out channels in total are connected to the strip sensors and the SCT barrel covers the pseudorapidity range of $|\eta| < 1.5$, which is then extended to $|\eta| = 2.5$ by the end-cap disks.

3.2.2.3 Transition Radiation Tracker

The Transition Radiation Tracker is in the outermost layer of the ID, which consists of 50000 polyimide straw tubes (73 layers) in the barrel and 250000 straws in both end-caps



(a)



(b)

Figure 3.2.4: Illustrations of (a) the general layout of the ATLAS Inner Detector and (b) the central region barrel of the ATLAS Inner Detector showing the Pixel Detector (PD), the Semi-Conductor Tracker (SCT), and the Transition Radiation Tracker (TRT) [90]. The new Insertable *B*-Layer (IBL) is also shown in (b), as part of the ATLAS Run 2 upgrade which aims to mitigate the impact of radiation damage to the Pixel Detector [91].

(160 planes). All straw tubes have a diameter of 4 mm with a 0.03 mm diameter gold-plated tungsten wire in the center and are filled with a gas mixture of Xe + CO_2 + O_2 with the proportion of 70%, 27% and 3%, respectively. Xenon is very sensitive to the transition radiation emitted when particles traversing material with different dielectric constants, while CO_2 + O_2 helps to increase the drift velocity of electrons. A negative

electric potential of 1.5 kV are kept between the tungsten wire and the wall of each tube. When charged particles pass through the straw tube, the gas mixture will be ionized, forming electrons and positive ions drifting in the electric field in two opposite directions. Signals proportional to the energy deposit of the traversing particles thus can be detected and send to the 351000 read-out channels connected to the straws. Since the transition radiation is suppressed for heavy particles, e.g. pions, electrons passing through the straw tubes will release a significantly larger amount of radiation, which serves as the most crucial evidence for electron identification.

The TRT covers the pseudorapidity range of $|\eta| < 2.0$ and records 36 hits on average for a charged particle with $p_{\mathrm{T}} > 0.5$ GeV. The only exception is the barrel-end-cap transition region ($0.8 < |\eta| < 1.1$), where on average 22 hits per track are expected. Compared to the PD and the SCT, the TRT provides a much coarser spatial resolution of 130 µm in the R-ϕ plane and no resolution in the direction parallel to the straws, e.g. z-direction for the barrels and R-direction for the end-caps.

3.2.3 Calorimeter system

The ATLAS calorimeter system is designed to stop particles traversing the Inner Detector and measure their positions and energies within a pseudorapidity range of $|\eta| < 4.9$. It consists of sampling calorimeters that contain alternating layers of dense passive material with high atomic number, also known as absorber material, as well as active material serving as medium. The particles can interact with the passive material either electromagnetically or hadronically, resulting in cascades of secondary particles, often referred to as 'particle showers', which induce signals in the active material in the forms of ionization or scintillation and proportional to the total energy deposit. Such design is mainly based on the consideration that building calorimeters purely with active material, e.g. plastic scintillators, yields to unreasonably high cost and bulky detector volume, since much more layers would be needed to fully absorb the energies of scattering particles without the dense passive material. However, the usage of absorbers may also lead to energy losses during the shower process, defined as:

$$f_{\mathrm{samp}} = \frac{E_{\mathrm{active}}}{E_{\mathrm{active}} + E_{\mathrm{passive}}}, \tag{3.2.1}$$

which needs to be later corrected by various calibration techniques.

In general, two types of showers take place in the calorimeters, depending on the type of incident particles and their interactions with the detector material. Electrons and photons lose their energies via bremsstrahlung and electron-positron pair production respectively in the absorber material. This is called electromagnetic shower and can be evaluated in terms of the radiation length X_0, which is defined as $7/9$ of the mean free path for photons and the average distance after which the energy of traversing particles are reduced to $1/e$ for electrons. On the other hand, hadrons typically lose their energies through inelastic collisions with the dense material, causing hadronic showers which can be evaluated in terms of the hadronic interaction length λ. The radiation length X_0 and the hadronic interaction length λ are both crucial to calorimeter design, as the calorimeters should be thick enough to provide good containment of various particles and prevent them from entering the Muon Spectrometer, which leads to underestimations of the total energy. In practice, the ATLAS calorimeter system can be divided into the Electromagnetic Calorimeter (ECal) and the Hadronic Calorimeter (HCal), optimized for measurements of the electromagnetic and the hadronic showers, respectively. Figure 3.2.5 shows the general layout of the ATLAS calorimeters.

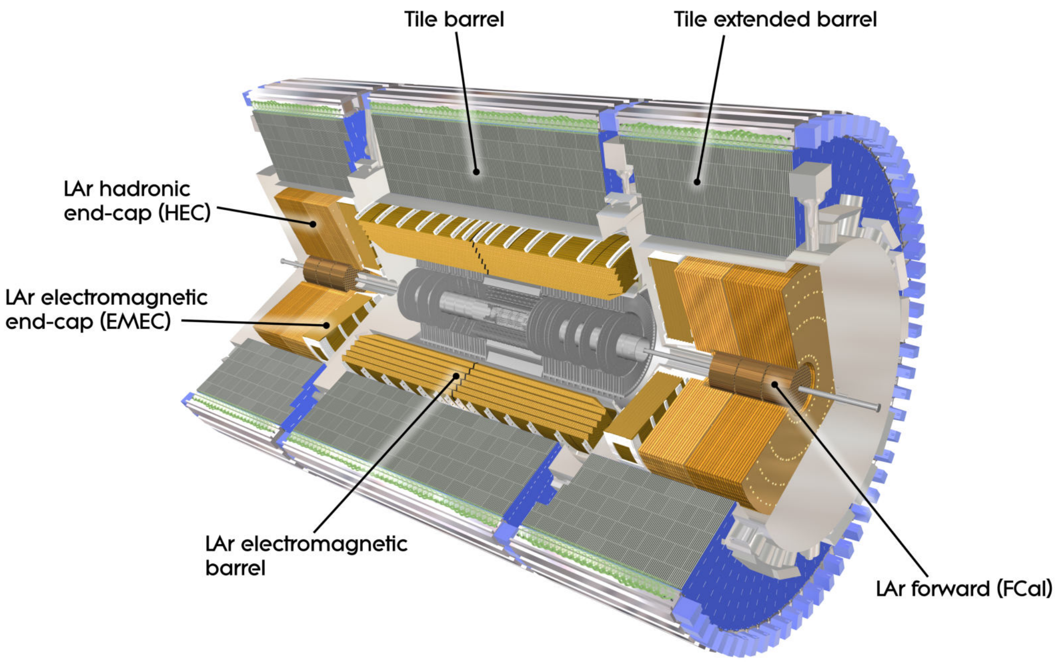


Figure 3.2.5: The general layout of the ATLAS calorimeter system [93].

The fractional resolution of calorimeters can be measured as a function of the total energy, in the form of:

$$\frac{\sigma_E}{E} = \frac{N}{E} \oplus \frac{S}{\sqrt{E}} \oplus C, \tag{3.2.2}$$

where N represents the noise of the measurements, S stands for the stochastic uncertainty, C is a constant term which reflects the nonuniformities in the detector and $\oplus$ denotes that the terms are added in quadrature.

3.2.3.1 Electromagnetic Calorimeter

The Electromagnetic Calorimeter aims to capture electromagnetically interacting particles, e.g. photons and electrons, within the pseudorapidity range of $|\eta| < 3.2$, composed of a barrel component which covers $|\eta| < 1.475$ and two end-cap wheels which cover $1.375 < |\eta| < 3.2$. It is arranged in alternating layers of lead as the absorber and liquid argon (LAr) as the active medium, while the copper read-out electrodes are attached between the absorber plates. Lead has a short radiation length of $X_0 = 0.56$ cm, while LAr is radiation-hard and produces stable ionization proportional to the input energy. An accordion geometry is chosen for the absorbers and the electrodes of the barrel and end-cap calorimeters, providing a full ϕ coverage without any cracks.

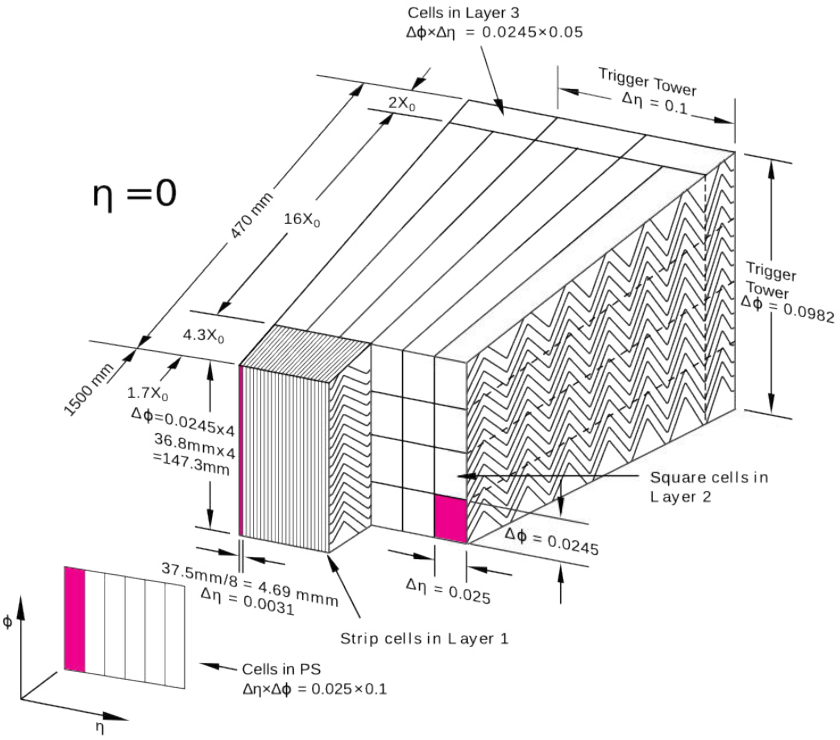


Figure 3.2.6: Sketch of a barrel module in the Electromagnetic Calorimeter at $\eta = 0$ [87]. The granularities of the cells in three different layers in η and ϕ are also shown.

The barrel calorimeter has a minimal depth of 22 X_0 at $\eta = 0$ and a maximum depth of 33 X_0 at higher $|\eta|$. Figure 3.2.6 is a sketch of a barrel module located at $\eta = 0$ and shows the granularities of the cells in three different layers. The first layer contains extremely thin strip cells with a depth of 4.3 X_0, allowing for identification of photons from $\pi_0 \to \gamma\gamma$ decay. The second layer consists of cells with a size of $(0.025, 0.0245)$ in (η, ϕ) space and a large depth of 16 X_0. Most energy of the photons and electron-positron pair cascades are absorbed at this stage. The third layer is designed to measure the remnants of radiation, thus has the coarsest granularity. Similarly, the end-cap wheels are constructed in layers with different granularities depending on $|\eta|$. The maximum granularity is arranged at $1.5 < |\eta| < 2.5$, while for regions with $1.375 < |\eta| < 1.5$ and $2.5 < |\eta| < 3.2$ both the number of layers and the granularity of cells are reduced.

The fractional resolution of the ECal, after noise subtraction, is measured to be

$$\frac{\sigma_E}{E} = \frac{10.1\%}{\sqrt{E}} \oplus 0.2\% \tag{3.2.3}$$

in the barrel region [94], and

$$\frac{\sigma_E}{E} = \frac{12.1\%}{\sqrt{E}} \oplus 0.4\% \tag{3.2.4}$$

in the end-cap region [95].

3.2.3.2 Hadronic Calorimeter

The Hadronic Calorimeter surrounds the Electromagnetic Calorimeter and aims to measure the energies of hadronically interacting particles within the pseudorapidity range of $|\eta| < 4.9$, composed of three sub-systems: the tile calorimeter (TileCal) in the central region which covers $|\eta| < 1.7$, the liquid argon end-cap (HEC) which covers $1.5 < |\eta| < 3.2$ and the liquid argon forward calorimeter (FCal) which covers $3.1 < |\eta| < 4.9$, as shown in Figure 3.2.5.

The TileCal can be further divided into three parts, a central component which covers $|\eta| < 1.0$ and two extended barrels which cover $0.9 < |\eta| < 1.7$. Each part is divided into 64 wedge-shaped modules in ϕ. Similar to the ECal, the TileCal contains sampling calorimeters which are made of alternating layers of steel as the absorber and scintillating tiles as the active medium. The photons from scintillation are then read out by wavelength shifting fibers, which deliver the signals to photomultiplier tubes (PMTs). Figure 3.2.7 is a sketch of a barrel module and shows how the electronics are attached to the layers of steel and scintillators. At $\eta = 0$ the barrel has a minimal depth of 7.4 λ and a maximum depth of 9.7 λ is reached at the outer edge of the tile-instrumented region. The measured energy resolution of the TileCal, after noise subtraction, can be parametrized as:

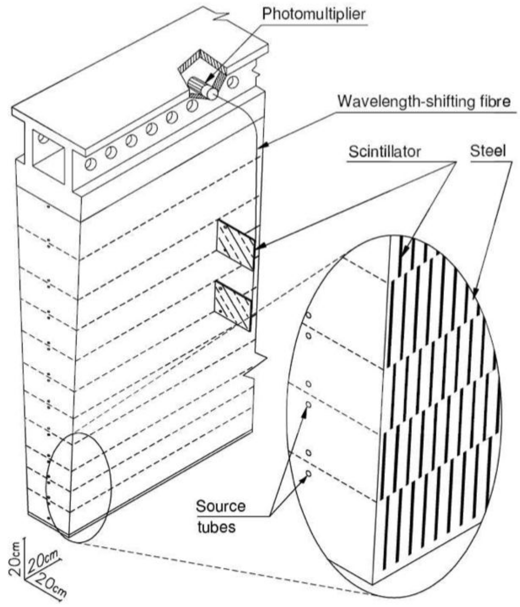


Figure 3.2.7: Sketch of a barrel module in the Tile Calorimeter [87], showing the major components of the optical read-out, including the tiles, the fibers and the PMTs.

$$\frac{\sigma_E}{E} = \frac{52.0\%}{\sqrt{E}} \oplus 3.0\%. \tag{3.2.5}$$

The HEC uses copper as the absorber and liquid argon as the active material. Each of the end-caps consists of a front and a rear cylindrical wheel, both composed of 32 identical wedge-shaped modules and providing a radiation depth of 12 λ when combined. At $|\eta| < 2.5$, the read-out cells have the size of (0.1, 0.1) in (η, ϕ) space and the granularity decreases at higher pseudorapidity. The measured energy resolution of the HEC is:

$$\frac{\sigma_E}{E} = \frac{70.6\%}{\sqrt{E}} \oplus 5.8\%. \tag{3.2.6}$$

The FCal consists of three layers in each end-cap disk. In the first layer, copper is used as the absorber and liquid argon is used as the active material, while in the second and the third layers tungsten is used as the absorber. A combined depth of 10 λ is given and the energy resolution of the FCal is measured as:

$$\frac{\sigma_E}{E} = \frac{94.2\%}{\sqrt{E}} \oplus 7.5\%. \tag{3.2.7}$$

3.2.4 Muon system

The ATLAS Muon Spectrometer is the outermost layer of the ATLAS detector and is designed to capture muons and provide precise measurements of their momenta within the range of $|\eta| < 2.7$. Surrounded by the barrel toroid and two end-cap toroids (see Section 3.2.1), muons traversing the MS are deflected in the strong magnetic field and thus can be accurately identified in the trigger chambers and measured in the precision chambers. The MS aims to track muons with transverse momenta between 3 GeV and up to a few TeV and consists of three barrel layers at $R = 5$ m, 7.5 m and 10 m and three wheels at $|z| = 7.4$ m, 14 m and 21.5 m in each end-cap. At $|\eta| \approx 0$, a gap in the coverage of the spectrometer is left open for cables connected to the solenoid magnet, the calorimeters and the ID. Two types of trigger chambers and two type of precision chambers are used to construct the spectrometer, namely the Resistive Plate Chambers (RPCs), the Thin Gap Chambers (TGCs), the Monitored Drift Tubes (MDTs) and the Cathode Strip Chambers (CSCs), as illustrated in Figure 3.2.8.

The RPCs provide fast tracking information at the central region of $|\eta| < 1.05$, composed of three concentric cylindrical layers. The first and second RPC layers (inner RPC) are assembled together with the second MS layer and permit low p_{T} muon trigger in the range of 6 - 9 GeV, which is then extended to 35 GeV when combined with the outer RPC assembled on the third MS layer. Each RPC unit consists of two parallel electrode plates separated by 2 mm of distance and filled with a gas mixture of $C_2H_2F_4$ + Iso-C_4H_{10} + SF_6 with the proportion of 94.7%, 5% and 0.3%, respectively. An electric potential of 9.8 kV is posed between the two plates and allows for a good timing resolution of 2 ns.

The TGCs provide fast tracking information at the forward region up to $|\eta| < 2.4$ and is also used to determine the azimuthal coordinate ϕ of muons. Each TGC unit is a multi-wire proportional chamber consisting of two parallel cathode plates filled with a highly quenching gas mixture of CO_2 and n-C_5H_{12} and wires placed in between serving as anodes. A potential of 2.9 kV is applied across the wires, allowing for a timing resolution of ~ 4 ns.

The MDTs provide precise measurements of muon tracks over the entire coverage of $|\eta| < 2.7$, except for the innermost wheel in the end-cap regions at $2.0 < |\eta| < 2.7$ where the CSCs are installed. The MDT module consists of multiple layers of aluminum tubes with a diameter of 3 cm and filled with a gas mixture of Ar + CO_2 with the proportion of 93% and 7%, respectively. A tungsten-rhenium wire is positioned at the center of each tube with an electric potential of 3 kV. This allows for measurements of the ionization

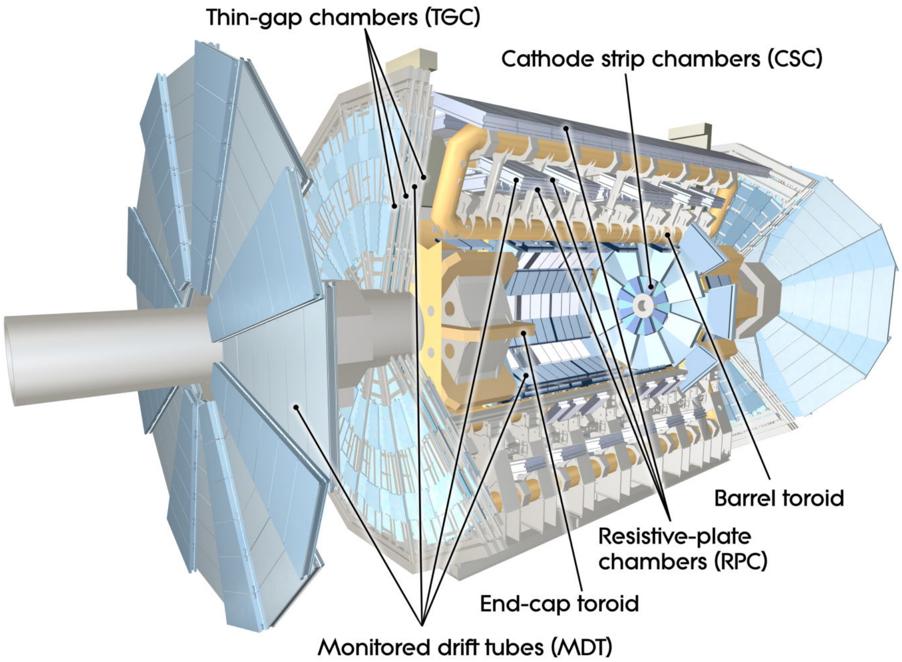


Figure 3.2.8: The general layout of the ATLAS Muon Spectrometer [96].

caused by the incident muons and an average spatial resolution of 80 μm is achieved per tube, which can be improved to 35 μm after the combination of multiple tube layers in each chamber.

The CSCs provide precise measurements of muon tracks in the very forward region of $2.0 < |\eta| < 2.7$ and are only installed on the innermost wheel in the end-cap regions as a replacement of the MDTs. Compared to the MDTs, CSCs have higher rate capacity and better timing resolution, thus preferable for the high pseudorapidity region where the particle flux is expected to be much larger. The CSCs consist of multi-wire proportional chambers filled with a gas mixture of Ar + CO_2 with the proportion of 80% and 20%, respectively and a potential of 1.9 kV between the wires and the strip cathodes, leading to a spatial resolution of 60 μm in the CSC plane and 5 mm in the non-bending direction where the cathode segmentation is much coarser.

The combined muon momentum resolution provided by the ATLAS Muon Spectrometer is p_{T} dependent. At $p_{\mathrm{T}} = 100$ GeV, the fractional resolution $\sigma_{p_{\mathrm{T}}}/p_{\mathrm{T}}$ is around 3%. At $p_{\mathrm{T}} = 1$ TeV, the resolution is at the level of 10%.

3.2.5 Trigger and data acquisition system

In LHC Run 2, the proton bunches are separated by a spacing of 25 ns, which indicates a collision frequency of 40 MHz. Consequently, around 1.7 billion proton-proton collisions take place in the ATLAS detector per second, while each individual collision event, containing information collected in all sub-detector systems as well as the reconstructed collision products, takes about 1.5 MB of disk space and 15 s of CPU time. This will cause huge pressure on computing, e.g. data storage consumption of more than 60 PB per second, if events are recorded without any selections.

The ATLAS trigger and data acquisition system is designed to run in parallel to the data-taking process and determine whether a collision event should be kept for physical analyses. As shown in Figure 3.2.9, a two-level trigger system is applied, consisting of the hardware-based level-1 (L1) triggers and the software-based high-level triggers (HLTs). The L1 triggers help to reduce the event rate from 40 MHz to 100 kHz, while the HLTs further reduce the number to 1 kHz, denoting an acceptable data flow of 1.5 TB per second.

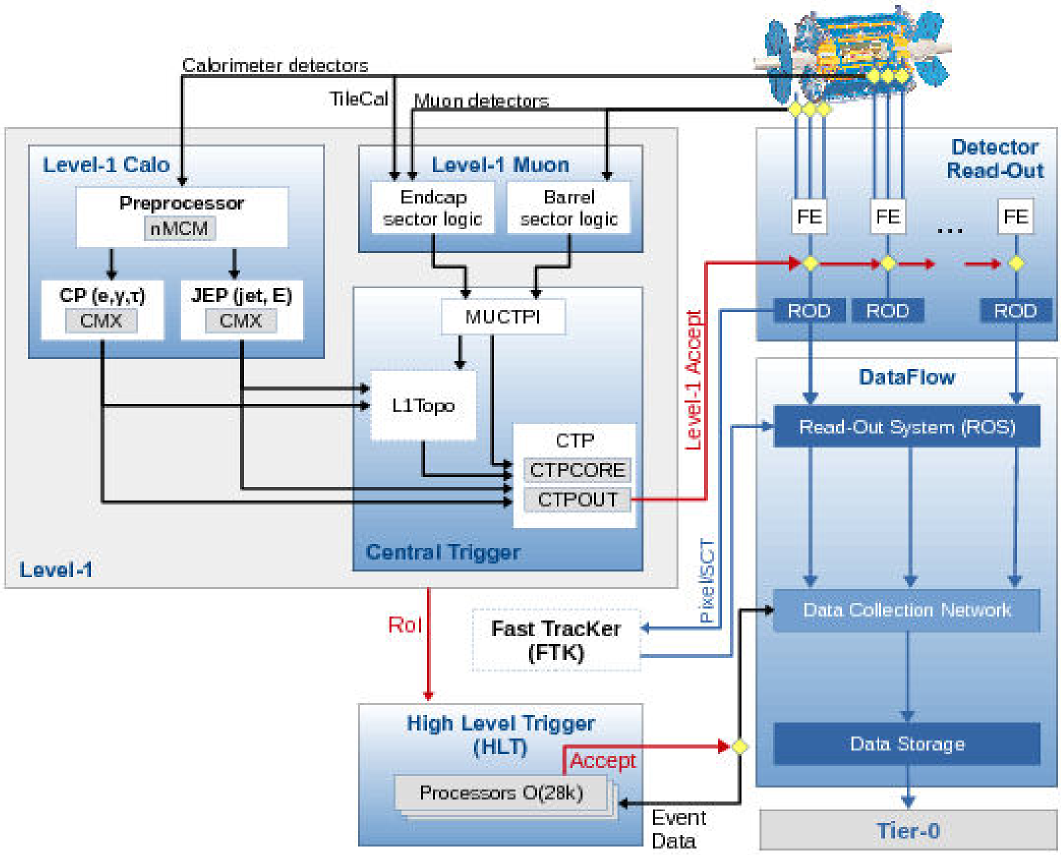
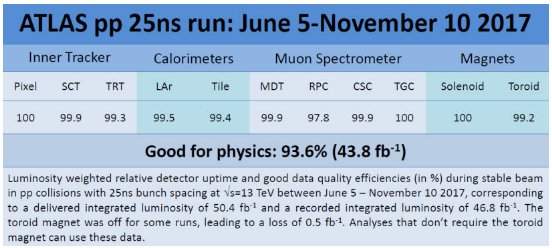


Figure 3.2.9: The ATLAS trigger and data acquisition system (TDAQ) in Run 2 [97]. The hardware-based level-1 (L1) triggers, including L1 Calo, L1 Muon and L1 Topo, and the software-based high-level triggers (HLTs) together reduce the event rate from 40 MHz to 1 kHz. The L1 Topo system has been commissioned in 2017, while the FTK is currently being integrated into the ATLAS trigger system, aiming to improve the performance of particle identification at high luminosity.

Immediately after a collision is read out by the electronics, the signals are delivered to the L1 triggers where the initial event selections are performed. The L1 trigger system is composed of four sub-systems, namely the L1 Calorimeter triggers (L1 Calo), the L1 Muon triggers (L1 Muon), the L1 Topological triggers (L1 Topo) and the Central Trigger Processor (CTP). L1 Calo is designed to select events containing all objects except for muons based on information collected in the calorimeters. L1 Muon, on the other hand, identifies muons using data collected by the RPC and the TGC in the Muon Spectrometer. Results from L1 Calo and L1 Muon are then transferred to L1 Topo where the topological association between trigger objects rebuilt in L1 Calo and L1 Muon are evaluated. Lastly, the CTP utilizes the inputs from L1 Calo, L1 Muon and L1 Topo and makes the final decision based on a certain combination of predefined trigger item selections, referred to as the 'trigger menu'. Events will be accepted by the L1 triggers and passed to the HLTs as long as they fulfill at least one set of requirements, e.g. L1_MU10 which asks for at least one muon with p_{T} above 10 GeV, on the trigger menu. The L1 trigger system is optimized to provide ultra fast decisions, more specifically, within 2.5 μs from readout to acceptance or rejection.

The outputs from L1 are processed in the HLTs, which can be further divided into two phases, the level-2 (L2) trigger system and the event filter. L2 analyzes the event topology using information from the Regions of Interest, which are defined as regions in the detector where possible trigger objects are recognized by the L1 triggers. The event rate is then reduced to below 3.5 kHz. The event filter is applied after the offline reconstruction using the standard ATLAS analysis applications and makes use of the full granularity of the ATLAS detector, especially the precise tracking information from the Inner Detector which has not been considered in the preceding trigger chain. The final output of the HLTs is at a rate of around 1 kHz.

Additionally, a quality check is run over the full data and events recorded during a period when possible detector problems, e.g. abnormal voltages and temperatures, dead cells, take place are marked as 'bad'. Such flags are set for every luminosity block, which corresponds to approximately 2 minutes long fraction of data-taking, and are documented in the so-called 'Good Run List' (GRL), allowing for the removal of events recorded by malfunctional sub-detectors. Figure 3.2.10 shows the relative fraction of luminosity associated to 'good' data delivered by the various sub-detector systems.



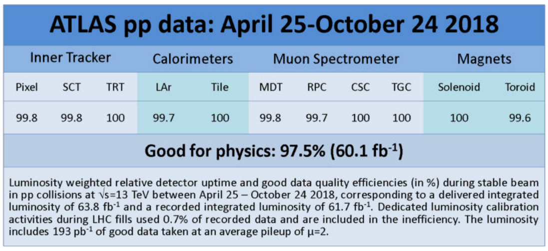
ATLAS pp 25ns run: June 5-November 10 2017

Inner Tracker			Calorimeters		Muon Spectrometer				Magnets	
Pixel	SCT	TRT	LAr	Tile	MDT	RPC	CSC	TGC	Solenoid	Toroid
100	99.9	99.3	99.5	99.4	99.9	97.8	99.9	100	100	99.2

Good for physics: 93.6% (43.8 fb^{-1})

Luminosity weighted relative detector uptime and good data quality efficiencies (in %) during stable beam in pp collisions with 25ns bunch spacing at √s=13 TeV between June 5 – November 10 2017, corresponding to a delivered integrated luminosity of 50.4 fb^{-1} and a recorded integrated luminosity of 46.8 fb^{-1}. The toroid magnet was off for some runs, leading to a loss of 0.5 fb^{-1}. Analyses that don't require the toroid magnet can use these data.

(a)



ATLAS pp data: April 25-October 24 2018

Inner Tracker			Calorimeters		Muon Spectrometer				Magnets	
Pixel	SCT	TRT	LAr	Tile	MDT	RPC	CSC	TGC	Solenoid	Toroid
99.8	99.8	100	99.7	100	99.8	99.7	100	100	100	99.6

Good for physics: 97.5% (60.1 fb^{-1})

Luminosity weighted relative detector uptime and good data quality efficiencies (in %) during stable beam in pp collisions at √s=13 TeV between April 25 – October 24 2018, corresponding to a delivered integrated luminosity of 63.8 fb^{-1} and a recorded integrated luminosity of 61.7 fb^{-1}. Dedicated luminosity calibration activities during LHC fills used 0.7% of recorded data and are included in the inefficiency. The luminosity includes 193 pb^{-1} of good data taken at an average pileup of μ=2.

(b)

Figure 3.2.10: Luminosity weighted relative fraction of good quality data delivery by the various components of the ATLAS detector sub-systems during stable beams for pp collisions between (a) 5 June and 10 November 2017, (b) 25 April and 24 October 2018 [98].

3.3 Physical object reconstruction in the ATLAS experiment

The pp collisions are recorded by the ATLAS detector in terms of energy deposits in the calorimeters and hits in the trackers. However, these information cannot be directly used by the offline event selections and need to be rebuilt as compound physical objects such as photons, electrons, muons, jets (especially b-jets) and $E_{\mathrm{T}}^{\mathrm{miss}}$ using different reconstruction algorithms. Due to various detector effects, the identification of objects will not be 100% efficient and deviations between the reconstructed and original kinematic properties are always expected. The former leads to corresponding systematic uncertainties in the statistical analyses while the latter can be reduced to some extent using proper calibration methods. The standard recommendations of physical objects identification, reconstruction and calibration within the ATLAS collaboration are provided by the Combined Performance (CP) groups, in order to keep the consistency in object definitions across different physical analysis groups.

The reconstructed objects can be categorized into two types: the basic objects, including tracks, vertices and topological clusters, and the composite objects, e.g. jets. The basic objects serve as the building blocks for the reconstruction of all composite objects, with brief introduction below:

- **Tracks**. Charged particles are deflected in the magnetic field and their trajectories can be rebuilt by connecting the hits in different layers of the Inner Detector. For muons, the hits recorded by the Muon Spectrometer are also considered. A sequence of tracking algorithms [99, 100] are used for track reconstruction. Starting from a seed containing three hits in the silicon detectors, the inside-out algorithm extends the track into the TRT by adding hits moving away from the interaction point using a combinatorial Kalman filter. Primary charged particles produced in the pp collision (or as the decay products of extremely short-living intermediate particles) are expected to be reconstructed using this algorithm. Then, an outside-in algorithm is used to rebuild tracks starting from the TRT segments and extrapolate them inwards by adding silicon hits not picked in the previous step. This is also known as back-tracking, which is designed to reconstruct secondary particles, the decay products of primary particles. All tracks are required to have $p_{\mathrm{T}} > 400$ MeV and $|\eta| < 2.5$ based on the coverage of the ID and can be parametrized by a set of five parameters, $(d_0, z_0, \phi, \theta, q/p_{\mathrm{T}})$, where d_0 and z_0 are the transverse and the longitudinal impact parameters, ϕ and θ are the azimuthal and the polar angle of the tracks and q/p_{T} represents the ratio between the charge and the transverse momentum.

- **Vertices**. Vertices denote locations where the particle collisions or decays take place. The reconstruction of vertices takes tracks as input and mainly consists of two steps: vertex finding and vertex fitting. During vertex finding, well-reconstructed tracks are matched to possible vertex candidates constrained by the beam spot position. The precise z-coordinate of each vertex is later determined through an iterative χ^2 fit, where the displaced tracks are used to build additional vertices. All vertices are required to contain at least two tracks, while the vertex with the largest sum of squared transverse momenta of all tracks associated is referred to as the primary vertex (PV).

- **Topological clusters**. The energy deposits of incident particles can be grouped into clusters of topologically connected calorimeter cells, where each topological cluster (topo-cluster) might contain the full or partial energy response of a single particle, or a combined response of multiple particles. The clustering starts from seed cells called 'proto-clusters'. Subsequently, cells adjacent to the proto-clusters with their signal significance exceeding certain thresholds are added into the proto-clusters. This procedure, namely the 'collect' step, will be iterated until no neighboring cells with signal significance passing the threshold are left unmerged. However, the proto-clusters at this stage might be too large to provide good spatial resolution and are split according to the three-dimensional distribution of local signal maxima fulfilling $E_{\mathrm{cell}}^{\mathrm{EM}} > 500$ MeV, forming the final topo-clusters. The total energy deposited in each topo-cluster is then calibrated to different scales with respect to the type of physical objects.

More complex factors are involved in the reconstruction of composite objects. Section 3.3.1 describes the reconstruction of electrons as well as the corresponding identification and isolation criteria. The reconstruction, identification and isolation of muons are discussed in Section 3.3.2. Section 3.3.3 focuses on various jet definitions, including track jets, small-radius (small-R) jets and large-radius (large-R) jets, and the flavor tagging algorithms used to identify jets originating from b-hadrons. Section 3.3.4 is dedicated to the computation of missing transverse momentum $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$, which plays an important role in the analyses documented in Chapter 5 and Chapter 6, and the evaluation of $E_{\mathrm{T}}^{\mathrm{miss}}$ significance based on the object resolution.

3.3.1 Electrons

Electrons and photons deposit significant amount of energies when passing through the calorimeter, as explained in Section 3.2.3, and can be distinguished using track information from the ID. Figure 3.3.1 illustrates the hypothetical path of a traversing electron in the ATLAS detector. Typically, an electron leaves 12 silicon hits in the tracker and exhausts most of its energy in the ECal. Only $\sim 2\%$ of the total energy are deposited in the HCal.

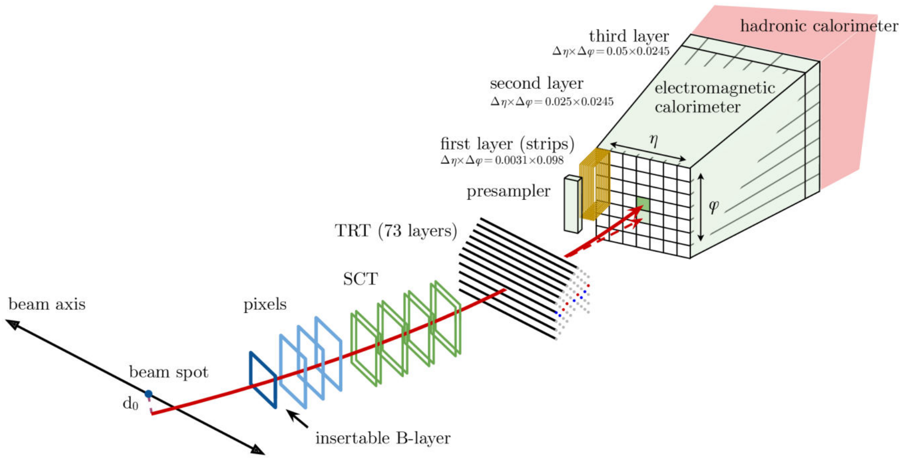


Figure 3.3.1: Illustration of the hypothetical path of an electron traversing the ATLAS detector [101]. The solid red trajectory represents the path of the electron, which first traverses the tracking system and then enters the ECal. The dashed red trajectory represents a secondary photon produced by the interaction between the incident electron and the detector material.

The reconstruction of electron is limited within the angular coverage of the ID, $|\eta| < 2.47$, and mainly consists of four steps:

- **Cluster reconstruction**. The energy deposit of electrons in the calorimeters are clustered using a sliding window algorithm [102]. Different from the formation of the topological clusters, the reconstruction of electrons starts from fixed-size seeds composed of 3×5 cells each with a size of (0.025, 0.025) in (η, ϕ) space, which corresponds to the granularity of the second layer of the ECal. The total transverse energy E_{T} of the electron is calculated as the sum of energy deposits in the seed across all three layers and is maximized by adjusting the position of the window. All seed clusters are required to have $E_{\mathrm{T}} > 2.5$ GeV in order to optimize the reconstruction efficiency.

- **Track association**. Once the clusters are determined, attempts to associate the reconstructed ID tracks with the seed clusters are made. If the distance between the track impact point and the cluster fulfills $|\Delta\eta| < 0.05$ and $\Delta\phi < 0.1$, the matching is considered to be successful. In the case that more than one tracks are matched to a same cluster, the track with the smallest ΔR (defined in Section 3.2) is chosen.
- **Track refit**. The track associations are then refined by taking the energy losses due to bremsstrahlung into account, using an optimized Gaussian Sum Filter. In order to suppress the background contamination from secondary particles, the matched tracks are required to be originating from the PV with $d_0/\sigma_{d_0} < 5$ and $|z_0 \sin\theta| < 0.5$ mm. Additionally, clusters associated to secondary vertices are marked as converted photons, which are defined as photons converting into electron-positron pairs when interacting with the tracker material.
- **Candidate reconstruction**. Lastly, the energy of electrons are recomputed with enlarged seed clusters with a size of 3×7 in the barrel region and 5×5 in the end-caps. This allows for corrections on electron energies in the different regions in the calorimeter, especially where the EM shower is not fully contained. The four-momenta of the electron candidates are then determined using information from both the calibrated total energy in the enlarged clusters (for the energy component) and the best track matched to the original cluster (for the η and ϕ components).

The electron reconstruction efficiency has a dependence on the transverse energy E_{T}, as shown in Figure 3.3.2.

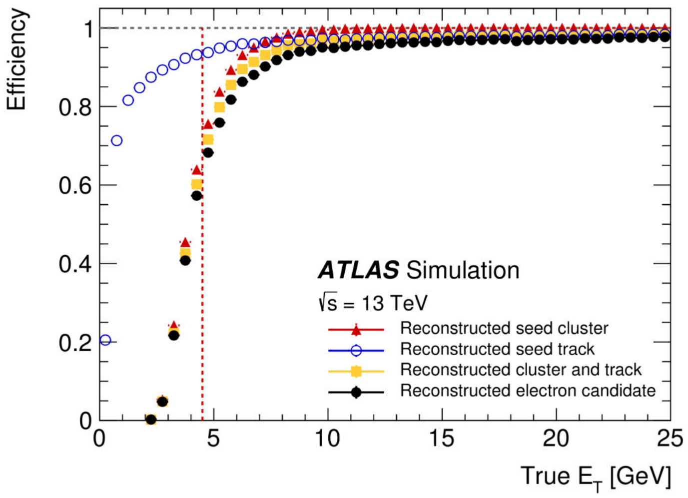


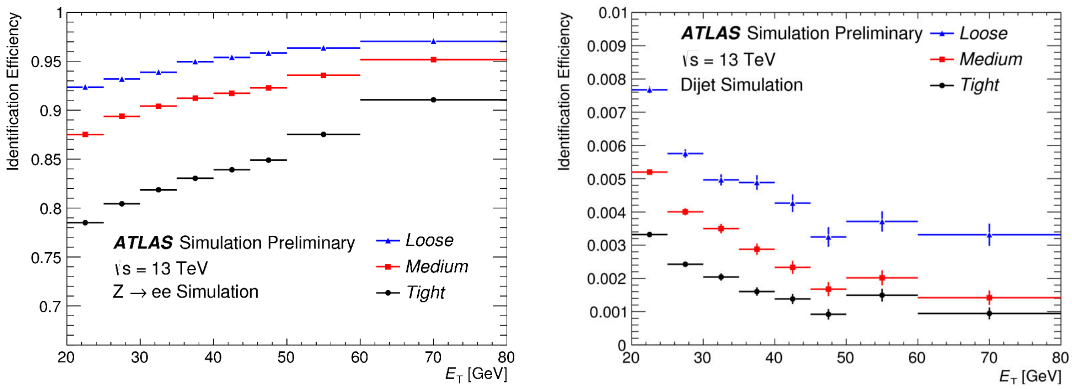
Figure 3.3.2: Reconstruction efficiency for simulated electrons in a single electron sample as a function of the truth-level transverse energy E_{T} for each step of the electron candidate formation [101]: cluster reconstruction (red triangles), track association (blue empty circles), track refit (yellow squares) and candidate reconstruction (black full circles).

Identification algorithms are developed to discriminate signal electrons from background particles, e.g. charged pions and photons. A multivariate likelihood-based (LH) method is applied taking various cluster and track variables as input, including calorimeter shower shapes, track conditions and track-cluster matching. Three levels of identification working points are provided, namely Loose, Medium and Tight, in order of increasing background rejection. These working points are defined such that the electrons selected by the Medium criterion are subsets of Loose electrons and the electrons selected by the Tight criterion are sub-sets of Medium electrons, while each working point uses the same set of variables to define the LH discriminant but with a different selection value:

- **Loose**: cuts on the shower shape variables of the first and the second layers of the ECal and the hadronic leakage variables;
- **Medium**: cuts on the track quality in the IBL and the Pixel Detector, the TRT high-threshold fraction, the track-cluster matching and the transverse impact parameter d_0, in addition to a tighter version of the Loose selections;
- **Tight**: cuts on the track quality in the TRT, the ratio between the cluster energy and the track momentum and the photon conversion suppression variables, in addition to a tighter version of the Medium selections.

Aside from the standard Loose, Medium and Tight criteria, a variation of the Loose working point, LooseAndBLayer, is defined by adding the requirement of at least one hit in the IBL on top of the Loose selections, allowing for more flexibility in analysis selection design.

A detailed list of variables used for electron identification can be found in Ref. [103]. Figure 3.3.3 (a) shows the electron identification efficiency evaluated with $Z \to ee$ process and (b) shows the rate of misidentification where hadrons are identified as electrons evaluated with dijet production for the Loose, Medium and Tight working points.



(a) (b)

Figure 3.3.3: (a) Electron identification efficiency evaluated with simulated $Z \to ee$ process and (b) the rate of misidentification where hadrons are identified as electrons evaluated with simulated dijet production for Loose (blue triangles), Medium (red squares) and Tight (black circles) as functions of the candidate electron E_T [103].

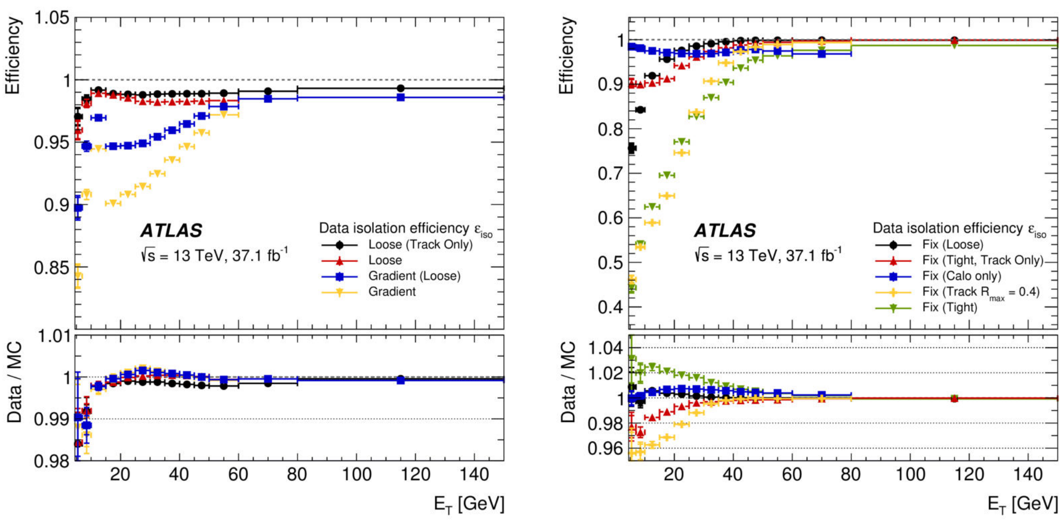
Furthermore, isolation requirements need to be fulfilled in order to reduce the contribution from background electrons originating from hadronic showers, e.g. electrons in jets. Motivated by the fact that the signal electrons are expected to be isolated while the electrons from weak decays are mixed with other shower products, two types of discriminating variables are used to quantify the energy of particles around the reconstructed electrons, namely the track isolation $p_{\mathrm{T}}^{\mathrm{varcone0.2}}$ ($p_{\mathrm{T}}^{\mathrm{varcone0.3}}$) and the calorimeter isolation energy $E_{\mathrm{T}}^{\mathrm{cone0.2}}$.

The track isolation $p_{\mathrm{T}}^{\mathrm{varcone0.2}}$ is defined as the sum of the transverse momenta of all the tracks in a cone with E_{T} dependent size $R = \min\{0.2, 10\ \mathrm{GeV}/E_{\mathrm{T}}\}$ around the candidate electron (excluding the electron track) and originating from the same PV. Similarly $p_{\mathrm{T}}^{\mathrm{varcone0.3}}$ is defined with a cone size of $R = \min\{0.3, 10\ \mathrm{GeV}/E_{\mathrm{T}}\}$. The calorimeter isolation energy $E_{\mathrm{T}}^{\mathrm{cone0.2}}$ is defined as the sum of the transverse energies of all the calorimeter cells in a cone with fixed size $R = 0.2$ around the candidate electron (excluding the electron cluster) and calibrated at the electromagnetic scale.

Combinations of different cuts on $p_{\mathrm{T}}^{\mathrm{varcone0.2}}$ ($p_{\mathrm{T}}^{\mathrm{varcone0.3}}$) and $E_{\mathrm{T}}^{\mathrm{cone0.2}}$ lead to a variety of isolation working points. For simplicity, only the ones used in the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search and the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search will be discussed:

- **LooseTrackOnly**: cuts on $p_{\mathrm{T}}^{\mathrm{varcone0.2}}$ targeting a constant isolation efficiency of 99%, uniform in E_{T} and η of the electron;
- **FixedCutLoose**: $p_{\mathrm{T}}^{\mathrm{varcone0.3}}/p_{\mathrm{T}} < 0.15$, $E_{\mathrm{T}}^{\mathrm{cone0.2}}/p_{\mathrm{T}} < 0.20$;
- **FixedCutTight**: $p_{\mathrm{T}}^{\mathrm{varcone0.3}}/p_{\mathrm{T}} < 0.06$, $E_{\mathrm{T}}^{\mathrm{cone0.2}}/p_{\mathrm{T}} < 0.06$.

Figure 3.3.4 shows the electron isolation efficiency evaluated with $Z \to ee$ process for different working points, including LooseTrackOnly, FixedCutLoose and FixedCutTight, as functions of E_{T}.



(a) (b)

Figure 3.3.4: Electron isolation efficiency evaluated with simulated $Z \to ee$ process for different working points, including LooseTrackOnly (black circles, left), FixedCutLoose (black circles, right) and FixedCutTight (green inverted triangles, right), as functions of the candidate electron E_{T} [101].

3.3.2 Muons

Muons can be reconstructed using track information from the ID and the MS, or even the energy deposits in the calorimeter. Typically, a muon in the central region leaves three pixel hits, eight SCT hits, and 30 TRT hits in the tracker. Four types of muons are defined depending on which sub-detector information and which combination algorithms are used for reconstruction, including combined (CB) muons, segment-tagged (ST) muons, calorimeter-tagged (CT) muons and extrapolated (ME) muons.

- **Combined** muons. Tracks reconstructed independently from the ID and the MS are combined with a global fit. The outside-in method is used to first rebuild the trajectory of muons in the MS, which are then extrapolated towards the interaction point and matched to the ID tracks. During this procedure, the reconstruction efficiency can be optimized by adding or removing MS hits from the fitted tracks. An alternative approach of the inside-out reconstruction, where the ID tracks are extrapolated and fitted to the MS tracks, serves as the complement of the outside-in method. In general, CB muon candidates have the highest muon purity compared to other muon definitions.
- **Segment-tagged** muons. When muons traverse only one layer in the MS, track reconstruction purely based on MS information will be challenging since very few hits are expected in the spectrometer. ST muons are then reconstructed by extrapolating the ID tracks to the MS where at least one local track segment in the MDTs or the CSCs can be associated. This muon definition helps to recover the acceptance of muons with low p_{T} or passing through regions with limited MS coverage.
- **Calorimeter-tagged** muons. An ID track will be recognized as a CT muon if it can be matched to an energy deposit compatible with a minimum ionizing particle after extrapolation to the calorimeters. As mentioned in Section 3.2.4, a gap on the MS barrel at $|\eta| \approx 0$, which allows for cabling and services of the ID, the calorimeters and the solenoid magnet, leads to an non-instrumented region at $|\eta| < 0.1$. The identification criteria for CT muons are thus optimized to enhance the acceptance in this region, with a momentum range of $15 < p_{\mathrm{T}} < 100$ GeV.
- **Extrapolated** muons. The reconstruction of ME muons only takes the MS tracks compatible with the interaction point (by extrapolation to the beam line) as input. Hits in at least two independent MS layers in the central region or at least three layers in the forward region are required for precise track measurements. The energy losses of muons traversing the calorimeters are taken into consideration as well. ME muons are mainly used to extend the muon acceptance to the region of $2.5 < |\eta| < 2.7$, which already exceeds the ID coverage.

Identification algorithms are developed to discriminate signal muons from background muons, e.g. muons from the decays of pions and kaons. For CB muons, the major discriminants of muon identification include:

- q/p significance: defined as the absolute value of the difference between the ratio of the charge and the momentum of the muon measured in the ID and the MS ($|q/p_{\mathrm{ID}} - q/p_{\mathrm{MS}}|$) divided by the sum in quadrature of the corresponding uncertainties ($\sqrt{\sigma_{\mathrm{ID}}^2 + \sigma_{\mathrm{MS}}^2}$), providing good separation between real and fake muon candidates;

- ρ': defined as the absolute value of the difference between the transverse momentum measured in the ID and the MS ($|p_{\mathrm{T}}^{\mathrm{ID}} - p_{\mathrm{T}}^{\mathrm{MS}}|$) divided by the CB muon p_{T}, providing good separation between signal and background muons;

- normalized χ^2 of the combined fit.

Four muon identification working points, Loose, Medium, Tight and High-p_{T}, are provided to satisfy the needs of different physical analyses. The Loose, Medium and Tight working points are defined in the order of increasing background rejection and are inclusive categories such that muons selected by the tighter criteria are also included in the looser categories. The High-p_{T} working point, on the other hand, is specifically designed to maximize the momentum resolution for muons with $p_{\mathrm{T}} > 100$ GeV at the cost of a reduced reconstruction efficiency. Among them, two working points are used in the scope of this dissertation, namely Loose and Medium.

- **Loose**: maximizes the reconstruction efficiency while providing good quality muon tracks. All muon types are included. CB muons are required to have q/p significance < 7 and at least three hits in at least two MDT layers, except for tracks within $|\eta| < 0.1$, where hits in only one MDT layer are allowed. ME muons are required to have hits in at least three MDT/CSC layers and $2.5 < |\eta| < 2.7$, while ST and CT muons are required to have $|\eta| < 0.1$. In the region of $|\eta| < 2.5$, about 97.5% of the Loose muons are CB muons, while ST and CT muons amount to 1% and 1.5%, respectively.

- **Medium**: minimizes the systematic uncertainties introduced by muon reconstruction and calibration. Only CB and ME muons are included in this category, with same selections as the Loose criterion. The Medium identification working point serves as the default selection criterion for muons in the ATLAS experiment.

Figure 3.3.5 shows the muon identification efficiency evaluated with $Z \to \mu\mu$ process for the Loose, Medium and Tight working points. Clear drop in efficiency can be seen at $|\eta| \approx 0$ for Medium and Tight, while for the Loose working point this is mostly compensated by the inclusion of ST and CT muons.

Similar to electrons, isolation requirements need to be fulfilled in order to reduce the contribution from background muons originating from hadronic showers, e.g. muons in jets. The track isolation $p_{\mathrm{T}}^{\mathrm{varcone0.3}}$ and the calorimeter isolation energy $E_{\mathrm{T}}^{\mathrm{cone0.2}}$, as defined in Section 3.3.1, are also used to select isolated muons. Combinations of different cuts on $p_{\mathrm{T}}^{\mathrm{varcone0.3}}$ and $E_{\mathrm{T}}^{\mathrm{cone0.2}}$ lead to a variety of isolation working points. For simplicity, only the ones used in the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search and the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search will be discussed:

- **LooseTrackOnly**: cuts on $p_{\mathrm{T}}^{\mathrm{varcone0.3}}$ targeting a constant isolation efficiency of 99%, uniform in E_{T} and η of the muon;

- **FixedCutTightTrackOnly**: $p_{\mathrm{T}}^{\mathrm{varcone0.3}}/p_{\mathrm{T}} < 0.06$;

- **FixedCutLoose**: $p_{\mathrm{T}}^{\mathrm{varcone0.3}}/p_{\mathrm{T}} < 0.15$, $E_{\mathrm{T}}^{\mathrm{cone0.2}}/p_{\mathrm{T}} < 0.30$;

- **FixedCutTight**: $p_{\mathrm{T}}^{\mathrm{varcone0.3}}/p_{\mathrm{T}} < 0.04$, $E_{\mathrm{T}}^{\mathrm{cone0.2}}/p_{\mathrm{T}} < 0.15$.

Figure 3.3.6 shows the muon isolation efficiency evaluated with $Z \to \mu\mu$ process for LooseTrackOnly and FixedCutLoose as functions of p_{T}.

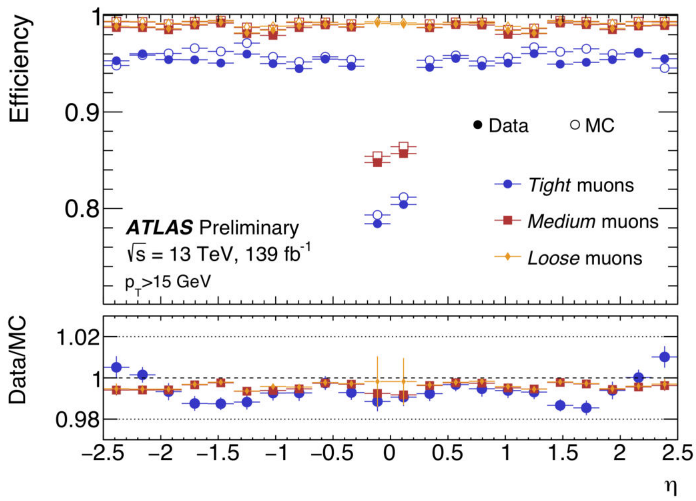
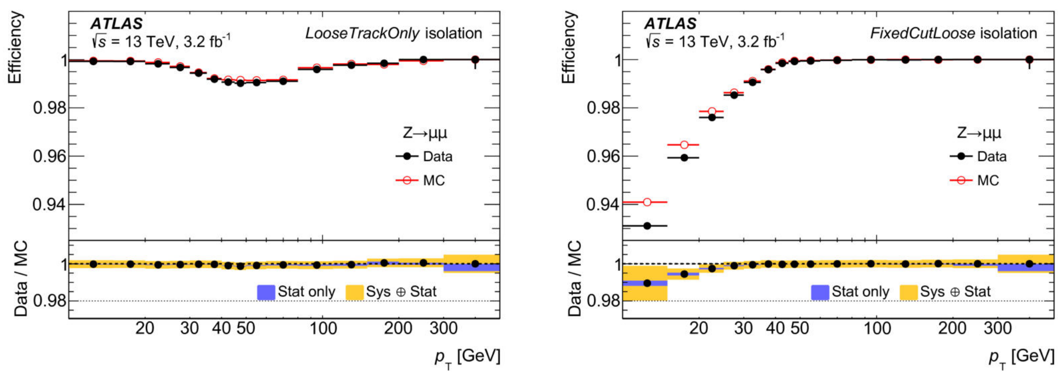


Figure 3.3.5: Muon identification efficiency evaluated with simulated $Z \to \mu\mu$ process for Loose (orange diamonds), Medium (red squares) and Tight (blue circles) as functions of the candidate muon η [104]. The empty markers denote predictions given by the Monte Carlo simulation and the full markers denote the data observation.



(a) (b)

Figure 3.3.6: Muon isolation efficiency evaluated with simulated $Z \to \mu\mu$ process for (a) LooseTrackOnly and (b) FixedCutLoose as functions of the candidate muon p_{T} [105]. The empty markers denote predictions given by the Monte Carlo simulation and the full markers denote the data observation.

3.3.3 Jets

Unlike electrons and muons, quarks and gluons cannot be directly observed in the detector in isolation due to the color confinement. Instead, they form sprays of color neutral particles via hadronization processes and are then reconstructed as jets in the detector. In practice, jets can be built from topological clusters in the calorimeters, tracks in the ID or truth hadrons in Monte Carlo (MC) simulation, while a large variety of jet definitions are established based on the choices of jet clustering algorithms.

Calorimeter jets, or simply jets, are reconstructed with energy deposits in the calorimeter, while track jets take the charged particle tracks recorded by the ID as input. In general, calorimeter jets show a better energy resolution due to the inclusion of neutral particle energy deposits and track jets have better spatial resolution thanks to the high granularity of the Inner Detector. Alternatively, jets can also be rebuilt from truth stable particles in simulated samples after the exclusion of non-interacting particles e.g. neutrinos and muons, thus can be referred to as particle-level jets or truth jets. Truth jets are widely used in truth-level studies especially in the early phase of the analyses, with one of the examples shown in Section 6.2.3.1.

Various jet definitions based on different clustering algorithms are designed to regroup the input objects in a hadron spray, e.g. tracks and topo-clusters, into a four-momentum which corresponds to the kinematic properties of the initial hard-scattering parton and are optimized for different event topologies. In ATLAS this step is achieved using the FastJet [106] package and the most commonly used jet clustering algorithms are the sequential recombination algorithms, which rely on the distance between object i and object j (d_{ij}) and the distance between object i and the beam (d_{iB}) to iteratively group the input objects into jets. The two distances are defined as:

$$d_{ij} = \min\{p_{\mathrm{T}i}^{2n}, p_{\mathrm{T}j}^{2n}\}\Delta R_{ij}^2, \tag{3.3.1}$$

$$d_{iB} = p_{\mathrm{T}i}^{2n} R^2, \tag{3.3.2}$$

where $\Delta R_{ij} = \sqrt{\Delta\eta_{ij}^2 + \Delta\phi_{ij}^2}$ represents the directional distance between object i and j and R is a predefined value, also known as the radius of the jets. Distances between all input objects are evaluated at the beginning of the clustering process and the algorithm identifies the smallest of d_{ij} and d_{iB}. If it is a d_{ij}, the two objects with the smallest distance, denoted as pseudojet i and pseudojet j, will be replaced by their vector sum (pseudojet k), while all the distance calculations involving pseudojet i or j will be updated with the kinematic properties of pseudojet k. If it is a d_{iB}, the pseudojet i will be removed from the pool of available inputs and added to the list of final jet candidates. This step is repeated until no pseudojet is left in the pool, marking the end of jet clustering.

The parameter n determines the type of jet clustering algorithms. The case $n = 1$ corresponds to the k_t algorithm [107, 108], where pseudojets with low p_T are combined in priority. The case $n = 0$ corresponds to the Cambridge/Aachen algorithm [109, 110], where only the geometrical distance between objects are taken into consideration. The case $n = -1$ corresponds to the anti-k_t algorithm [111], where pseudojets with high p_T are combined in priority.

In theory, a good jet algorithm should be both collinear- and infrared-safe. The collinear-safety refers to the stability of the algorithm against the number of objects within a hadron shower. The boundaries of jets should not be affected if an initial object is replaced by two collinear objects carrying the same amount of energy in total. The infrared-safety refers to the stability of the algorithm against the soft radiation from the initial partons. This requires the algorithm to be minimal sensitive to the soft energy deposits between overlapping jets.

The anti-k_t algorithm fulfills both requirements and in addition to this, is able to provide jet candidates with better reconstruction efficiency, pile-up suppression and exact circular shape on the η-ϕ cylinder compared to other algorithms. As a result, the anti-k_t algorithm is recognized as the standard jet algorithm within the ATLAS collaboration and jets used in this thesis are all anti-k_t jets reconstructed with different radius R. Section 3.3.3.1 is dedicated to the calibration and flavor tagging of small-radius (small-R) jets, which are reconstructed from topo-clusters with $R = 0.4$. Section 3.3.3.2 describes the grooming and calibration of large-radius (large-R) jets, reconstructed from topo-clusters with a radius of $R = 1.0$. In Section 3.3.3.3, jets reconstructed from charged particle tracks with radius $R = 0.2$ are discussed.

3.3.3.1 Small-radius calorimeter jets

Small-radius calorimeter jets, or simply small-R jets, are reconstructed from topological clusters within the coverage of $|\eta| < 4.5$ using the anti-k_t algorithm with $R = 0.4$. The input topo-clusters are rebuilt from calorimeter cells with deposited energy beyond a certain threshold, which aims to reject noises originating from electronics and pile-up interactions, and are then calibrated at the electromagnetic (EM) scale [102], which provides corrections on energy measurements of electromagnetic showers. Lastly, a $p_T > 7$ GeV cut is applied to all jets used in physical analyses.

The energy scale of the reconstructed small-R jets can be restored to that of the truth jets reconstructed from truth-level particles by a series of calibrations derived from data and MC simulation. As illustrated in Figure 3.3.7, the sequential calibration scheme for small-R jets can be divided into six steps, namely origin correction, jet area-based pile-up correction, residual pile-up correction, MC-based calibration, global sequential calibration (GSC) and residual in-situ calibration.

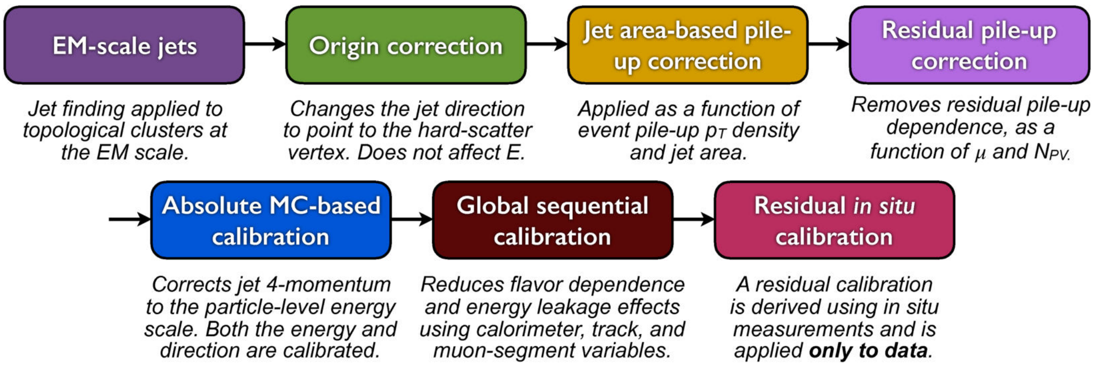
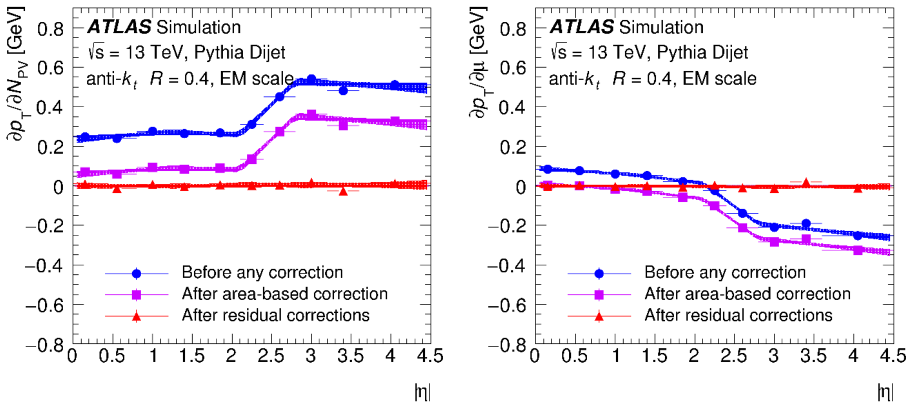


Figure 3.3.7: Overview of the sequential calibration scheme for small-R jets [112]. All corrections are applied to the four-momentum of the jet as scale factors, except for the origin correction which recalculates the jet direction in order to point to the hard-scatter PV while keeping the jet energy unchanged.

- **Origin correction**. The origin correction recalculates the jet direction in order to point to the hard-scatter PV instead of the geometric center of the detector, while keeping the jet energy unchanged. This helps to improve the η resolution of the jets.
- **Jet area-based pile-up correction**. The area-based pile-up correction subtracts the per-event pile-up contribution to jet p_T estimated based on the median p_T density ρ, which is defined as the averaged ratio between p_T and jet area, $\langle p_T/A \rangle$. Jets used for the calculation of ρ are reconstructed from positive energy topo-clusters within the coverage of $|\eta| < 2.0$ using the k_t algorithm with $R = 0.4$. The choice of the k_t algorithm is mainly motivated by its sensitivity to soft radiations, hence

it is able to provide a better estimation of the pile-up constituents compared to the standard anti-k_t jets.

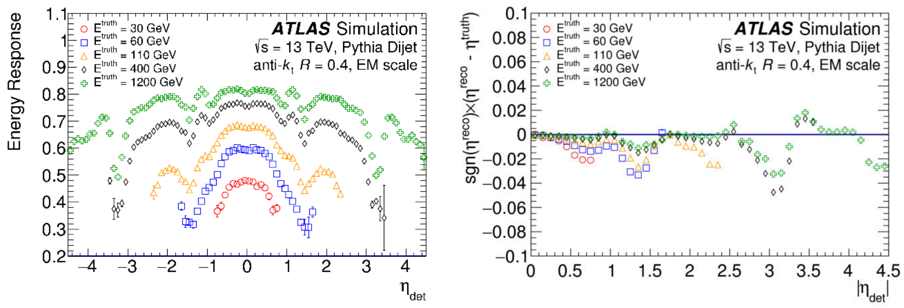
- **Residual pile-up correction**. Since the area-based pile-up correction is derived from k_t jets rebuilt in the central region, naturally it cannot fully describe the pile-up condition at high pseudorapidity and thus the residual pile-up dependence in jet p_T needs to be accounted for. The residual pile-up correction is computed with respect to the number of primary vertices (N_{PV}) and the number of interactions per crossing (μ), which are sensitive to in-time pile-up and out-of-time pile-up, respectively. Figure 3.3.8 shows the dependence of small-R jets on (a) in-time and (b) out-of-time pile-up as functions of $|\eta|$ before and after the pile-up corrections.



(a) (b)

Figure 3.3.8: Dependence of small-R jets on (a) in-time and (b) out-of-time pile-up as functions of $|\eta|$ with $p_T^{truth} = 25$ GeV before the pile-up corrections (blue circle), after the jet area-based correction (violet square) and after the residual correction (red triangle) [112].

- **MC-based calibration**. The absolute MC-based calibration aims to correct the four-momentum of the reconstructed jet to the particle-level energy scale and account for mismodelings due to various detector effects, e.g. energy deposits in inactive regions, shower products outside of the jet radius, energy losses in hadronic showers which cannot be covered by the EM scale correction and biases in η reconstruction in the transition region of the calorimeter. It consists of two steps, jet energy scale (JES) calibration and η calibration, both factorized as functions of E^{reco} and η_{det}, the pseudorapidity measured towards the center of the detector instead of the primary vertex, which allows for the derivation of a geometrically independent calibration map. The JES calibration factor is taken as the inverse of the average energy response $\langle R_{energy}\rangle$, defined as the mean of a Gaussian fit to the ratio between the reconstructed and the truth energy, E^{reco}/E^{truth}. In practice, $\langle R_{energy}\rangle$ is firstly evaluated in bins of E^{truth} and η_{det} and then parametrized in terms of E^{reco} and η_{det} using a numerical inversion procedure [113]. On the other hand, correction on η is defined as the difference between η^{reco} and η^{truth}. Similar to the JES calibration, the numerical inversion method is used to derive the corrections in bins of E^{reco} from E^{truth}. Figure 3.3.9 shows (a) the average energy response and (b) the signed difference between reconstructed and truth η as functions of η_{det} for small-R jets with different p_T^{truth}. Jets calibrated with the full JES and η calibration are considered to be at the EM+JES scale.



(a) (b)

Figure 3.3.9: (a) Average energy response and (b) signed difference between reconstructed and truth η as functions of η_{det} for small-R jets with $p_{\text{T}}^{\text{truth}} = 30, 60, 110, 400, 1200$ GeV before the MC-based calibration [112].

- **Global Sequential Calibration**. The global sequential calibration is a series of corrections designed to remove the residual JES dependencies on average particle composition and shower shape, which vary depending on the initiating particles, especially between the quark- and the gluon-initiated jets. A quark-initiated jet tends to contain hadrons with higher p_{T} which traverse further into the calorimeter, while a gluon-initiated jet typically contains more soft constituents which results in a lower calorimeter response and a wider transverse profile. To account for the energy scale dependence on the topology of calorimeter clusters and associated tracks, five observables are used in the GSC procedure, where for each observable a correction on jet four-momentum is derived using the numerical inversion method and as a function of $p_{\text{T}}^{\text{reco}}$ and $|\eta_{\text{det}}|$.
- **Residual in-situ calibration**. As the last step of the calibration chain, the residual in-situ calibration is applied to cover the differences between data and MC simulation and account for the possible MC mismodelings. The in-situ calibration is derived by balancing the p_{T} of the target jet against other well-measured reference object(s), e.g. a photon, a Z boson or a collection of well-calibrated jets. The vector boson balancing [114] uses a well-calibrated photon or a Z boson decaying into a pair of electrons or muons to correct the p_{T} response of central jets with $|\eta| < 0.8$ and p_{T} up to about 950 GeV. The multijet balance [115] further extends the $|\eta|$ coverage to 1.2 and the p_{T} coverage to around 1.7 TeV by measuring the responses of high-p_{T} jets recoiling against a collection of low-p_{T} jets whose energy scales have been calibrated in the previous step. Finally, the η-intercalibration [115] equalizes the forward JES ($0.8 < |\eta| < 4.5$) to the scale of central jets ($|\eta| < 0.8$) in the dijet system and with p_{T} up to 1.2 TeV. The calibration factors derived in each step are then combined into a final in-situ calibration map covering the full kinematic region, as shown in Figure 3.3.10.

The full set of systematic uncertainties related to the jet energy scale calibration contains more than 90 components. However, most of these uncertainties are subdominant in physical analyses and thus the reduced sets of JES uncertainties are introduced to minimize the complexity of the statistical models. For example, a strongly reduced set consists of four terms, including one term for the non-closure uncertainty

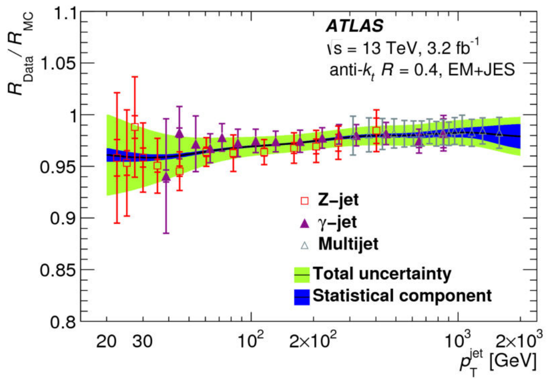


Figure 3.3.10: Ratio of the EM+JES jet response between data and MC simulation as functions of jet p_{T} for Z+jet, γ+jet, and multijet in-situ calibrations [112]. The final correction (black) and its statistical (dark blue) and total (light green) uncertainty are also shown.

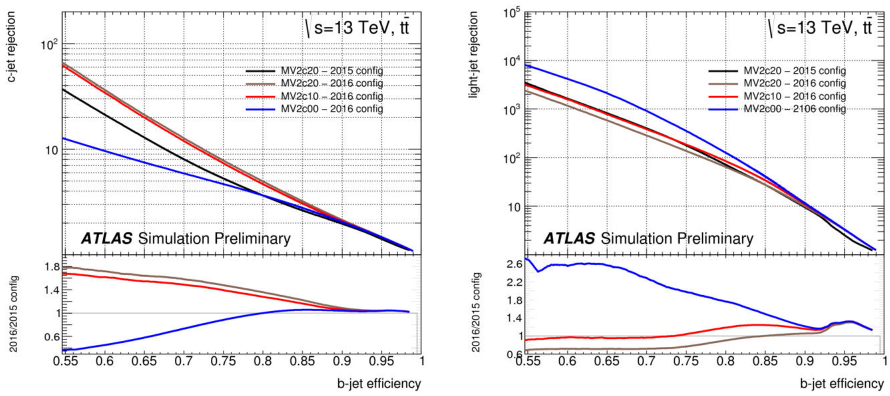
on η-intercalibration and three terms representing the combined uncertainty in low-p_{T} regime, medium-p_{T} regime and high-p_{T} regime, is considered in the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search, as described in Section 5.7.1.

The identification of jets coming from the hadronization of bottom quarks is often referred to as b-tagging. Compare to c- and τ-hadron, the most distinguishing characteristic of b-hadron is its relative long life time ($\sim$ 1.5 ps), which corresponds to a proper decay length of 450 µm. Various b-tagging algorithms are developed to identify this signature, including impact parameter based algorithm (IP3D), secondary vertex finder algorithm (SV1) and topological multi-vertex algorithm (JetFitter) [116], with brief introduction below:

- **The IP3D algorithm** uses both the transverse and the longitudinal impact parameter significance, d_0/σ_{d_0} and $z_0 \sin\theta/\sigma_{z_0 \sin\theta}$, to build a two-dimensional log-likelihood ratio (LLR) discriminant. The computation of the LLR discriminant is based on the template probability density functions (PDF) extracted from MC simulation for b-jet, c-jet and light-flavor jet hypotheses and is performed on a per-track basis.

- **The SV1 algorithm** is optimized to rebuild a single displaced secondary vertex which corresponds to the decay of heavy-flavor hadrons using tracks which fulfills a set of track quality requirements. Additionally, all two-track vertices compatible with the decays of long-lived particles, converted photons or interactions with detector material are removed from the candidates. The remaining tracks are used to form the secondary vertex which is significantly displaced from the PV using a Kalman based χ^2 fit [117]. The LLR discriminant is then constructed using various vertex variables, e.g. the invariant mass of all associated tracks, the decay length significance and the number of two-track vertices.

- **The JetFitter algorithm** exploits the topological structure of weak b- and c-hadron decays inside the jet and tries to reconstruct the full b-hadron decay chain. A Kalman filter [118] is used to find a common line where the primary vertex and the vertices of the b- and c-hadron decays lie, which can be used to determine the flight direction of the heavy-flavor jets.



(a) (b)

Figure 3.3.11: (a) c-jet and (b) light-flavor jet rejection versus b-tagging efficiency for MV2c10 (red) and MV2c20 (gray for 2016 configuration) [119].

The standard ATLAS b-tagging algorithm, MV2 [119], combines the discrimination variables provided by the above mentioned algorithms, also known as the basic algorithms, using a boosted decision tree (BDT) implemented in the Toolkit for Multivariate Data Analysis (TMVA) package [120]. The MV2 variants are trained with different background compositions, while two of them, namely MV2c10 score and MV2c20 score, are used in the analyses discussed in this thesis. The MV2c10 score is trained on b-jet signals with backgrounds consisting of 7% c-jets and 93% light-flavor jets, while the MV2c20 score is trained with backgrounds consisting of 20% c-jets and 80% light-flavor jets. Figure 3.3.11 shows the c-jet and the light-flavor jet rejection versus b-tagging efficiency for MV2c10 and MV2c20. Figure 3.3.12 shows the MV2c10 score for b-jets, c-jets and light-flavor jets evaluated with simulated $t\bar{t}$ events.

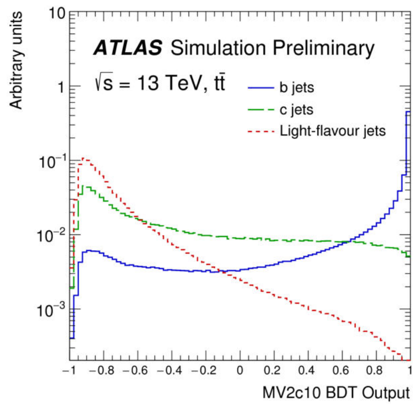
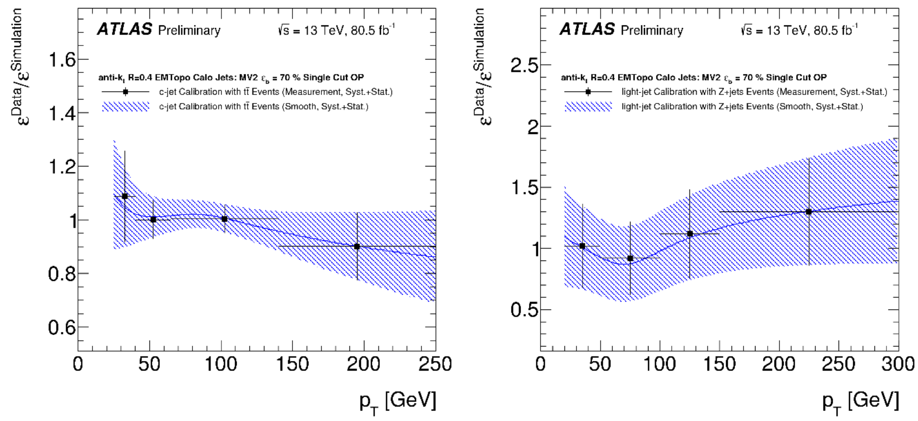


Figure 3.3.12: MV2c10 score for b-jet (solid blue), c-jets (dashed green) and light-flavor jets (dotted red) evaluated with simulated $t\bar{t}$ events [119].

Several fixed-cut working points are used in the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search and the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search, which are optimized for EM scale calorimeter jets clustered using the anti-k_{t} algorithm with $R = 0.4$, following the recommendations from the ATLAS Flavor Tagging group. The b-tagging efficiency as well as the mistag rates of c-jets and light-flavor jets are compared between data and MC. Corrections on simulated events are derived in terms of scale factors, which account for mismodelings of the input variables used by MV2c10 and MV2c20, as illustrated in Figure 3.3.13.



(a) (b)

Figure 3.3.13: (a) Corrections on c-jet mistag rate of MV2 algorithm for small-R jets with a 70% fixed-cut working point evaluated with simulated $t\bar{t}$ events using a likelihood fit [121]. (b) Corrections on light-flavor jet mistag rate of MV2 algorithm for small-R jets with a 70% fixed-cut working point evaluated with simulated Z+jets events using a negative-tag method [122]. The plots are taken from Ref. [123].

3.3.3.2 Large-radius calorimeter jets

Large-radius calorimeter jets, or simply large-R jets, are reconstructed from topological clusters using the anti-k_{t} algorithm with $R = 1.0$. The input topo-clusters are rebuilt from calorimeter cells with deposited energy beyond the noise threshold and are then calibrated using the local cell weighting (LCW) scheme [124] in order to correct for response differences between particles originating from electromagnetic shower and hadronic shower. Similar to the origin correction applied to the small-R jets, the four-momenta of these topo-clusters are adjusted to point towards the hard-scatter PV, instead of the geometric center of the detector, while keeping the energy constant.

In events containing heavy intermediate particle with a significant Lorentz boost, the decay products of the boosted particle can be highly collimated and may be reconstructed as a single large-R jet with substructure information. Compared to jets reconstructed with $R = 0.4$, large-R jets are more sensitive to pile-up contamination, soft radiations and multiple parton interactions due to the choice of the radius parameter. These contributions are typically much softer than the hard-scattering partons as well as their final state radiations and therefore jet grooming is introduced to remove the soft components inside the large-R jets. Subsequently, the MC-based calibration, consisting of jet energy scale (JES) calibration, η calibration and jet mass scale (JMS) calibration, is applied to the groomed large-R jets, which corrects the response of the reconstructed jets to particle-level on average and accounts for mismodelings due to various detector effects. Figure 3.3.14 provides an overview of the reconstruction, grooming and calibration for large-R jets used in ATLAS analyses.

- **Jet grooming**. The grooming procedure aims to remove contributions to a given large-R jet that are irrelevant or detrimental to resolving the hard decay products

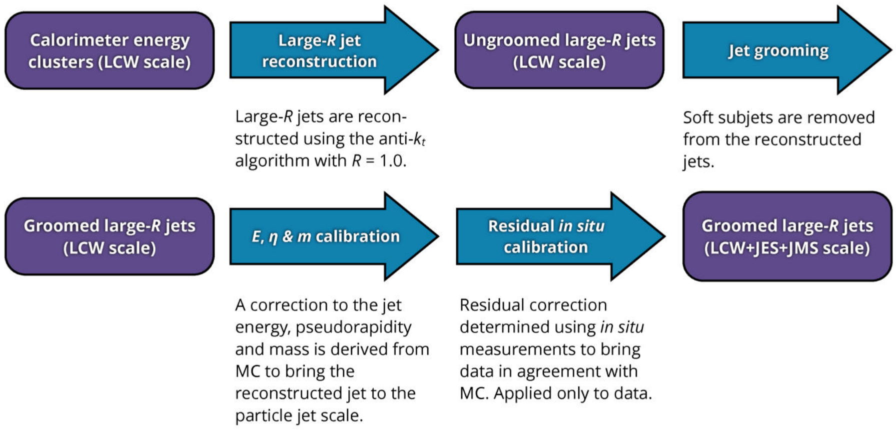
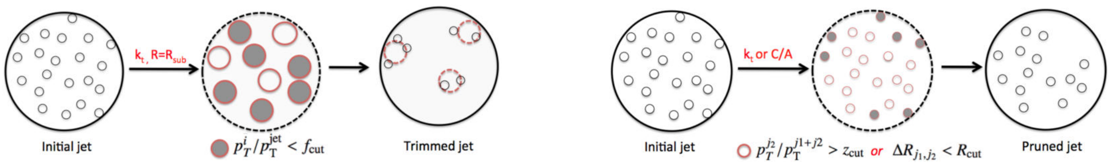


Figure 3.3.14: Overview of the reconstruction, grooming and calibration for large-R jets [125]. The calorimeter energy clusters from which jets are reconstructed have already been adjusted to point towards the hard-scatter PV.

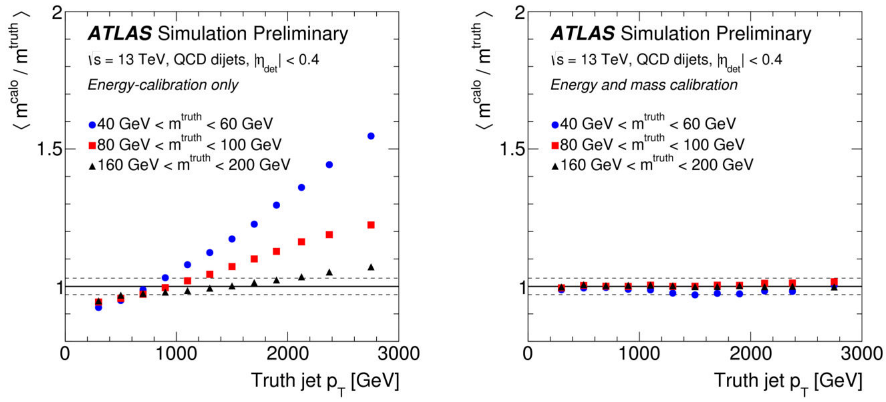
from a boosted object while retaining the corresponding jet substructure, thereby improving the mass resolution and pile-up mitigation. Several jet grooming algorithms have been developed, as illustrated in Figure 3.3.15. Among them, the trimming algorithm [126] serves as the standard ATLAS jet grooming algorithm. It starts from building 'subjets' from the constituents of the input large-R jet (the 'parent jet') using the k_t algorithm with $R_{\text{sub}} = 0.2$, which clusters the low energy components in priority. Any subjet with $p_T^{\text{subjet}}/p_T^{\text{parent jet}} < f_{\text{cut}}$ will be then removed, where $f_{\text{cut}} = 0.05$ denotes the threshold parameter of the algorithm, and the four-momentum of the trimmed large-R jet is recalculated as the vector sum of the remaining constituents.



(a) (b)

Figure 3.3.15: Illustration of two jet grooming algorithms [127]. (a) Diagram depicting the trimming procedure, the standard ATLAS jet grooming algorithm for large-R jet which removes reclustered subjets with p_T below a certain threshold. (b) Diagram depicting the pruning procedure, an alternative method to build groomed large-R jets by introducing additional requirements on ΔR_{ij}, p_{Ti} and p_{Tj} during the jet clustering process in order to veto wide-angle and soft pseudojets.

- **MC-based calibration**. The MC-based calibration corrects, on average, the kinematic properties of the trimmed jets to particle-level by applying calibration factors obtained from MC simulation. Three types of calibration are performed in sequential order during this procedure, i.e. jet energy scale (JES) calibration, η calibration and jet mass scale (JMS) calibration. The JES and η calibration are derived following the same procedure of small-R jets and factorized as functions of E^{reco} and η_{det}, while the JMS calibration is applied in bins of $(E^{\text{reco}}, |\eta_{\text{det}}|, m^{\text{reco}})$. Figure 3.3.16 shows the average jet mass response before and after the JMS calibration. More details about the MC-based large-R jet calibration can be found in Chapter 4. Jets calibrated with the full JES, η and JMS calibration are considered to be at the LCW+JES+JMS scale.



(a) (b)

Figure 3.3.16: Average calorimeter-based mass response as functions of $p_{\mathrm{T}}^{\mathrm{truth}}$ for large-R jets in bins of m^{truth} (a) before and (b) after the JMS calibration [128].

- **Residual in-situ calibration**. Similar to small-R jets, the residual in-situ calibration is applied to account for the differences between data and MC simulation as the final step of large-R jet calibration. The calibration is performed in two separate steps. First, the JES response is measured with the same methods used for small-R in-situ calibration, as described in Section 3.3.3.1. Then, the JMS response is derived using the R_{trk} method [127], which relies on the tracker to provide an independent measurement of the jet mass scale and its associated uncertainty, and the forward folding [129], which fits the mass peaks of the W boson and the top quark in $t\bar{t}$ production and measures the relative energy and mass scales as well as the corresponding resolutions between data and MC.

However, the in-situ JES and JMS calibrations were not fully available at the time when the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search and the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search were performed, thus will not be included as part of the large-R jet calibration in the scope of this work.

Multiple mass definitions of large-R jet can be used in physical analyses, each with strength in specific kinematic regimes. The standard large-R jet mass used in ATLAS, the calorimeter-based mass m^{calo}, is defined as:

$$m^{\mathrm{calo}} = \sqrt{\left(\sum_{i \in J} E_i\right)^2 - \left(\sum_{i \in J} \vec{p}_i\right)^2}, \tag{3.3.3}$$

where J denotes the large-R jet, E_i and $\vec{p}_i$ represent the energy and the momentum of topo-cluster constituent i. For particle with a sufficiently high Lorentz boost, the angular spread in the decay products, which will be suppressed by a factor of $1/p_{\mathrm{T}}$, may be comparable to the calorimeter granularity. This limits the performance of m^{calo} at highly boosted event topologies ($p_{\mathrm{T}} > 1$ TeV), where the track-assisted mass, m^{TA}, is introduced for compensation:

$$m^{\mathrm{TA}} = \frac{p_{\mathrm{T}}^{\mathrm{calo}}}{p_{\mathrm{T}}^{\mathrm{track}}} \cdot m^{\mathrm{track}}, \tag{3.3.4}$$

where $p_{\mathrm{T}}^{\mathrm{calo}}$ is the transverse momentum of the large-R jet, $p_{\mathrm{T}}^{\mathrm{track}}$ is the vector sum of the tracks associated to the large-R jet and m^{track} is evaluated as the invariant mass of these associated tracks, with the mass of each track set to pion mass m_{π}. The ratio between $p_{\mathrm{T}}^{\mathrm{calo}}$ and $p_{\mathrm{T}}^{\mathrm{track}}$ corrects for the neutral particle contribution, which cannot be measured by the tracker, improving the resolution of the track-only mass m^{track}.

Since the calorimeter-based mass is not used in the computation of the track-assisted mass, possibility of taking advantage of the independent mass measurements by the calorimeter and the ID arises. The combined mass, m^{comb}, is then defined as a linear combination of m^{calo} and m^{TA}:

$$m^{\mathrm{comb}} = \left(\frac{\sigma_{\mathrm{calo}}^{-2}}{\sigma_{\mathrm{calo}}^{-2} + \sigma_{\mathrm{TA}}^{-2}} \right) \cdot m^{\mathrm{calo}} + \left(\frac{\sigma_{\mathrm{TA}}^{-2}}{\sigma_{\mathrm{calo}}^{-2} + \sigma_{\mathrm{TA}}^{-2}} \right) \cdot m^{\mathrm{TA}}, \tag{3.3.5}$$

where σ_{calo} and σ_{TA} are the expected resolution function extracted after the MC-based JMS calibration applied to the calorimeter-based mass and the track-assisted mass, respectively. No additional JMS calibration is applied to m^{comb}. By construction, it outperforms both m^{calo} and m^{TA} in terms of mass resolution over the full p_{T} range, as shown in Figure 3.3.17. In both analyses presented in this thesis, the combined mass definition is applied, ensuring optimal mass performance for large-R jets.

The differences between data and MC simulation for large-R jets are accounted by the scale uncertainties. These uncertainties are derived using the R_{trk} and the forward folding method mentioned above and are then categorized into four uncorrelated groups. In the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search, where the kinematic properties of large-R jets are used as discriminants, the full set of large-R jet uncertainties consisting of 15 terms is considered in the statistical model. A reduced set of systematic uncertainties is applied for simplicity in the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search, where only the large-R jet multiplicity and mass are used for event categorization, as discussed in Section 6.7.1.

3.3.3.3 Track jets

Track jets [130] are reconstructed from charged particle tracks recorded by the ID within the coverage of $|\eta| < 2.5$ using the anti-k_{t} algorithm with $R = 0.2$. All tracks used for clustering need to have $p_{\mathrm{T}} > 400$ MeV and consist of at least one pixel hit and at least six SCT hits. In order to reject tracks originating from pile-up vertices, requirements on impact parameters are introduced to ensure the tight matching between tracks and the primary vertex. This includes a cut on the transverse impact parameter, $d_0 < 1.5$ mm,

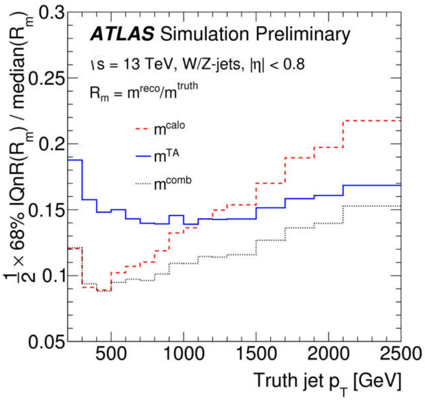


Figure 3.3.17: Jet mass resolution with respect to p_T for the calorimeter-based mass (solid blue), the track-assisted mass (dashed red), and the combined mass (dotted black) evaluated with W/Z+jets events [128].

and a cut on the longitudinal impact parameter, $|z_0 \sin\theta| < 1.5$ mm, where θ stands for the polar angle. No energy scale calibration is applied to the track jets.

In events with boosted topology, where the merging of multiple small-R jets start to take place and large-R jets are used for the reconstruction of W, Z or top quark's decay products, track jets are used to assist the identification of b-hadrons within the large-R jets. Compared to small-R jets with $R = 0.4$, $R = 0.2$ track jets allow for a more sophisticated description of jet substructure information and therefore improve the resolution and efficiency of the directional measurements of b-hadrons. However, in environments with dense hadronic activities, matching between track jets and calorimeter objects can lead to a certain degree of ambiguity if only the geometric distance between objects are considered. As a result, the 'ghost-association' technique [127, 131, 132] is developed to uniquely match track jets to large-R jets by introducing the 'ghosts', defined as the four-vector of track jets with the momentum set to an infinitesimal value, in the clustering process. The infrared-safety of the anti-k_t algorithm ensures that the reconstruction of large-R jets is not altered by the inclusion of ghost particles, while the reclustered jets are expected to have the same boundaries and kinematic properties compared to the original ones, only with the addition of the ghost particles retained as constituents. Track jets can be then 'ghost-associated' to a large-R jet if their corresponding ghosts are contained within the catchment area of this large-R jet.

Similar to small-R jets, the MV2c10 algorithm is used for the b-tagging of track jets. In the scope of this work, a 70% fixed-cut working point is chosen, which is optimized for track jets clustered using the anti-k_t algorithm with $R = 0.2$, following the recommendations from the ATLAS Flavor Tagging group. This corresponds to a b-jet purity of 93.5% and a c-jet (light-flavor jet) rejection of 7.1 (119.7). The b-tagging efficiency as well as the mistag rates of c-jets and light-flavor jets are compared between data and MC. Corrections on simulated events are derived in terms of scale factors, which account for mismodelings of the input variables used by MV2c10, while the relative uncertainties are treated as uncorrelated terms with respect to the ones for small-R jets.

3.3.4 Missing transverse momentum

According to momentum conservation, the transverse momenta of all collision products should sum to zero since the initial partons are expected to carry negligible momenta in the transverse direction. The missing transverse momentum, $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$, refers to the total transverse momentum of invisible final state particles, e.g. neutrinos and Dark Matter particles, and can be evaluated as the negative vector sum of p_{T} of all visible particles, while its magnitude is often referred to as the missing transverse energy, $E_{\mathrm{T}}^{\mathrm{miss}}$.

$$\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}} = -\sum_{i\in\text{visible}} \boldsymbol{p}_{\mathrm{T}}^{i}. \tag{3.3.6}$$

The reconstructed $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ consists of two major components. The first component, the hard term, is defined as the momenta of all reconstructed and well-calibrated objects (hard objects), including electrons, photons, muons, hadronic taus and small-R jets, while the second component, the soft term, represents the additional correction accounting for detector signals not associated with any hard objects. These unassociated signals can be either topo-clusters in the calorimeter, from which the calorimeter-based soft term (CST) is built, or charged particle tracks in the ID, from which the track-based soft term (TST) is built.

$$\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}} = -\underbrace{\left(\sum_{i\in e}\boldsymbol{p}_{\mathrm{T}}^{i} + \sum_{i\in\gamma}\boldsymbol{p}_{\mathrm{T}}^{i} + \sum_{i\in\mu}\boldsymbol{p}_{\mathrm{T}}^{i} + \sum_{i\in\tau}\boldsymbol{p}_{\mathrm{T}}^{i} + \sum_{i\in\text{jets}}\boldsymbol{p}_{\mathrm{T}}^{i}\right)}_{\text{hard term}} - \sum_{i\in\text{soft term}}\boldsymbol{p}_{\mathrm{T}}^{i}. \tag{3.3.7}$$

Additionally, a track-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ can be reconstructed from ID tracks satisfying certain track quality criteria [133]. Despite the fact that track-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ is insensitive to the contribution of neutral particles, it can be used to reduce pile-up contamination and non-collision background, as discussed in Section 5.5.2.

Figure 3.3.18 shows the distribution of the x component of CST $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$, TST $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ and track-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ evaluated with $Z \to \mu\mu$ process, where zero $E_{\mathrm{T}}^{\mathrm{miss}}$ is expected at truth-level. Compared to CST $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ and track-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$, TST $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ presents a good compromise between the good angular resolution and pile-up resilience from the track-based approach and the precise energy measurement from the calorimeter-based approach, therefore serves as the standard $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ definition in ATLAS Run 2 analyses.

A variety of $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction working points are defined to match the requirement of analyses focusing on different kinematic topologies. Among them, two working points, namely Loose and Tight, are used in the analyses documented in this thesis:

- **Loose**. Jets with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 4.5$ are used. For better pile-up suppression, jets with $p_{\mathrm{T}} < 60$ GeV and $|\eta| < 2.4$ are required to pass the Jet Vertex Tagger (JVT) [134] selection with JVT score > 0.59.

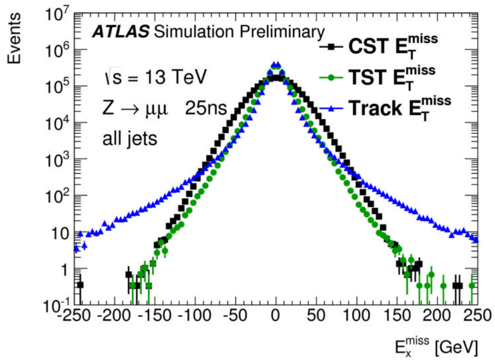


Figure 3.3.18: Distribution of the x component of CST $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ (black squares), TST $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ (green circles) and track-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ (blue triangles) evaluated with $Z \to \mu\mu$ process [133].

- **Tight**. In addition to the Loose criterion, forward jets ($2.4 < |\eta| < 4.5$) must fulfill $p_{\mathrm{T}} > 30$ GeV. For better pile-up suppression, requirements on the forward Jet Vertex Tagger (fJVT) [135] can be combined with the Tight working point.

Figure 3.3.19 shows the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ resolution as functions of the number of primary vertices (N_{PV}) evaluated with simulated semi-leptonically decaying $t\bar{t}$ events. Three $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ definitions are shown, differing only in the forward jet selection.

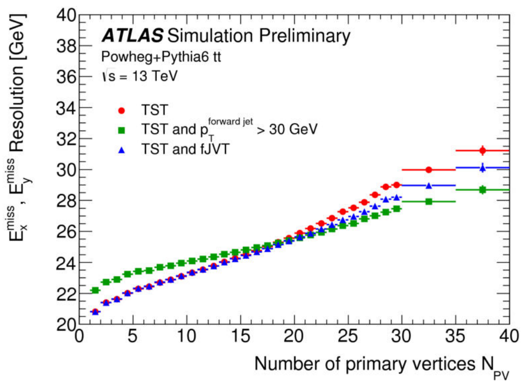


Figure 3.3.19: $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ resolution as functions of the number of primary vertices (N_{PV}) evaluated with simulated semi-leptonically decaying $t\bar{t}$ events [136]. Three $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ definitions are shown, differing only in the forward jet selection. Red circles denote $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ reconstructed with the Loose working point, where the input forward jets are required to have $p_{\mathrm{T}} > 20$ GeV. Green squares denote $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ reconstructed with the Tight working point, where the p_{T} selection for forward jets is tightened to 30 GeV. Blue triangles denote $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ reconstructed with the Loose working point with additional fJVT requirements.

3.3.4.1 Object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance

At reconstruction-level, a non-zero $E_{\mathrm{T}}^{\mathrm{miss}}$ may arise from two sources:

- **Weakly interacting particles**. These can be neutrinos in the context of the Standard Model or some theoretically predicted new particles, e.g. the neutralinos in the

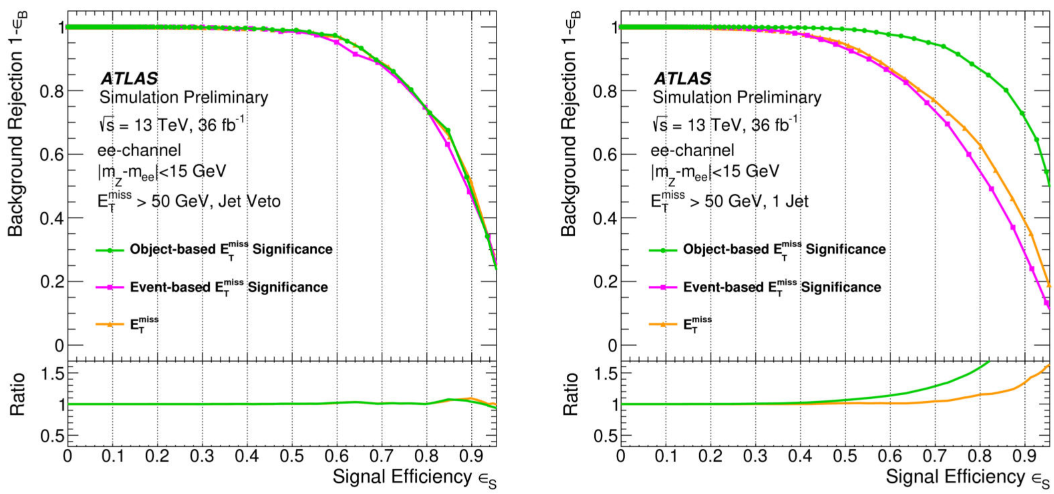
SUSY theory or the DM particles in the simplified models. In this case, the reconstructed missing transverse energy is referred to as real $E_{\mathrm{T}}^{\mathrm{miss}}$ and serves as the key variable in many searches for new physics.

- **Detector effects**. The reconstruction of $E_{\mathrm{T}}^{\mathrm{miss}}$ involves all detector sub-systems, thus it is susceptible to mismeasurements, miscalibration, limits on detector acceptance including the dead regions and the signal remnants from additional interactions, i.e. the pile-up contaminations. In this case, the reconstructed missing transverse energy can be referred to as fake $E_{\mathrm{T}}^{\mathrm{miss}}$.

In order to distinguish between events with real $E_{\mathrm{T}}^{\mathrm{miss}}$ and fake $E_{\mathrm{T}}^{\mathrm{miss}}$, an object-based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance [137] is constructed to quantify the degree to which the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ is consistent with the hypothesis of $E_{\mathrm{T,\,truth}}^{\mathrm{miss}} = 0$, given the information of all particles entering the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation, e.g. resolution on p_{T} scale and direction of hard objects, energy resolution of the soft term, etc. The object-based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance can be written in the form of:

$$\text{object-based } E_{\mathrm{T}}^{\mathrm{miss}} \text{ significance} = \frac{E_{\mathrm{T}}^{\mathrm{miss}}}{\sqrt{\sigma_{\mathrm{L}}^2(1-\rho_{\mathrm{LT}}^2)}}, \tag{3.3.8}$$

where σ_{L} represents the total variance in the longitudinal direction to $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ and ρ_{LT} represents the correlations between the variances in the longitudinal and transverse direction. In events containing multiple jets, object-based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance can provide very good discrimination between signal and background, as shown in Figure 3.3.20.



(a) (b)

Figure 3.3.20: Background rejection versus signal efficiency evaluated with simulated $Z \to ee$ and $Z \to ee\nu\nu$ samples [137]. The performance is shown for $E_{\mathrm{T}}^{\mathrm{miss}}$ (green circles), event-based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance (pink squares) and object-based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance (orange triangles) in events with (a) jet veto or (b) exact one jet, pre-selected with $|m_Z - m_{ee}| < 15$ GeV and $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV.

Chapter 4

Jet mass scale calibration for variable-radius calorimeter jets

With the conclusion of the four-year-long LHC Run 2 at the end of October 2018, around 147 fb^{-1} of proton-proton collision data is ready to be used in all ATLAS analyses, allowing for probing new physics at extremely boosted event topologies which was previously obstructed by statistical limitation. New techniques have been developed in order to improve the performance of physical object reconstruction, including an alternative jet clustering algorithm designed for high p_{T} decaying particle, the variable-radius jet algorithm [138].

This chapter is dedicated to the reconstruction and calibration of variable-radius calorimeter jets. Section 4.1 details the clustering and grooming algorithms optimized for variable-radius jets, while the standard workflow of deriving the jet mass scale calibration functions are described in Section 4.2. Section 4.3 documents the Monte Carlo samples considered in the calibration procedure as well as the selections applied to the simulated events. The closure performance for variable-radius jets with different mass definitions after the calibration are provided in Section 4.4. A brief conclusion of these studies is drawn in Section 4.5.

4.1 Variable-radius calorimeter jets

The variable-radius (VR) algorithm is motivated by the fact that the average angular separation between the decay products of a boosted particle decreases inversely with respect to its transverse momentum, which can be described by an empirical formula:

$$\Delta R \approx 2m/p_{\mathrm{T}}. \tag{4.1.1}$$

At low p_T, the decay products appear well-separated in the detector and can be reconstructed as multiple isolated small-R jets, often referred to as the 'resolved' topology. However, at high p_T, these particles tend to become collimated in the direction of the decaying particle due to the Lorentz boost, preventing them from being reconstructed as individual small-R jets. Hence, the definition of large-R jets, which allows to capture the hadronically decaying boosted object in a single cone with large radius parameter, is proposed and widely applied in analyses focusing on boosted signatures.

Due to the large fixed jet radius, large-R jets are typically more sensitive to pile-up and initial state radiation (ISR), which cannot be fully resolved by the jet grooming procedure. Also, for particles with very high transverse momentum, their decay products are expected to be highly concentrated, even to the extent that they can be contained within one small-R jet. Under such conditions, the choice of large R actually comes to be a disadvantage, as more underlying interactions might be introduced while the size of the hadronic shower is likely to be overestimated.

The variable-radius jets are derived as an alternative to the standard fixed-R jet algorithms for the reconstruction and identification of highly boosted objects, e.g. the W boson, Z boson, top quark and Higgs boson [139]. The most distinctive feature of VR jets is that their size shrink in proportion to $1/p_T$, leading to a more robust pile-up and ISR rejection. Similar to jets with fixed radius, VR jets can be built from either topo-clusters or charged particle tracks. The former refers to VR calorimeter jets, which are especially optimized for boosted top quark decays, while the latter refers to VR track jets, often used to provide high resolution b-hadron identification within the shower of high p_T objects. Section 4.1.1 documents the general clustering algorithm of VR jets, a modification of the sequential recombination algorithms. In Section 4.1.2, the choice of the jet grooming technique for VR calorimeter jets will be discussed. Details on VR track jets will not be covered in the scope of this chapter and more information can be found in Ref. [140].

4.1.1 Jet clustering

As discussed in Section 3.3.3, the sequential recombination algorithms for fixed-R jet clustering relies on two distance parameters, d_{ij} and d_{iB}. In VR algorithm, the predefined jet radius R is replaced by the effective radius, $R_{\text{eff}} \equiv \rho / p_{\text{T}i}$:

$$d_{ij} = \min\{p_{\text{T}i}^{2n}, p_{\text{T}j}^{2n}\}\Delta R_{ij}^2, \tag{4.1.2}$$

$$d_{iB} = p_{\text{T}i}^{2n} R_{\text{eff}}^2 = p_{\text{T}i}^{2n-2}\rho^2, \tag{4.1.3}$$

The new parameter ρ determines how fast the effective jet radius decreases with the rising of the transverse momentum of the jet. Additionally, $n = -1$, which corresponds to the anti-k_t algorithm, is required for VR jet clustering, simply due to the fact that the definition of 'effective radius' indicates a circular shape of the reconstructed jet. This can only be achieved by the anti-k_t algorithm, in which the combination of pseudojets begins with the hardest four-momenta.

The idea of constructing p_T inversely proportional R_{eff} originates from the reconstruction of hadronic showers initiated by two partons from the decay of a heavy resonance. If the resonance decays to two jets, these jets are expected to have a uniform size in spherical coordinates (θ, ϕ), regardless of their orientation with respect to the beam axis. However, at the LHC the partonic center-of-mass frame is not fixed, as the momentum exchange in z-axis is unknown, and pseudorapidity η is introduced to maintain the boost invariance. For anti-k_t algorithms performed with a fixed radius parameter, the geometric distance between pseudojets are defined as $\Delta R_{ij} = \sqrt{\Delta\eta_{ij}^2 + \Delta\phi_{ij}^2}$, ensuring circular jet boundaries in (η, ϕ) coordinates. This naturally leads to the unphysical result that forward jets reconstructed using fixed-R algorithms always have smaller angular size compared to jets in the central region.

The VR algorithm, on the other hand, is able to provide jet candidates uniform in (θ, ϕ) coordinates, as the distance between pseudojet i and the beam axis, d_{iB}, is suppressed by a factor of $p_{\mathrm{T}i}^{-2}$. In order to prevent the jets from becoming too large at low p_{T} (high $|\eta|$) and from shrinking below the detector resolution at high p_{T} (low $|\eta|$), the upper and lower cut-offs on R_{eff} are set, denotes by R_{max} and R_{min}:

$$R_{\mathrm{eff}} = \begin{cases} R_{\mathrm{max}} & \text{, if} \quad \rho/p_{\mathrm{T}i} > R_{\mathrm{max}}; \\ \rho/p_{\mathrm{T}i} & \text{, if} \quad R_{\mathrm{min}} \leq \rho/p_{\mathrm{T}i} \leq R_{\mathrm{max}}; \\ R_{\mathrm{min}} & \text{, if} \quad \rho/p_{\mathrm{T}i} < R_{\mathrm{min}}. \end{cases} \tag{4.1.4}$$

In theory, the choice of the VR parameters $(\rho, R_{\mathrm{max}}, R_{\mathrm{min}})$ should depend on the physical purpose of the reconstructed jets. Considering the clustering of two pseudojet i and j, the combination will only happen if

$$d_{ij} < \min\{d_{iB}, d_{jB}\}, \tag{4.1.5}$$

as explain in Section 3.3.3. Assuming $p_{\mathrm{T}i} > p_{\mathrm{T}j}$, $\min\{d_{iB}, d_{jB}\} = d_{iB}$ is expected in the anti-k_t VR algorithm, therefore the two pseudojets are combined when

$$\frac{d_{ij}}{d_{iB}} = \frac{\Delta R_{ij}^2}{R_{\mathrm{eff}}^2} < 1, \tag{4.1.6}$$

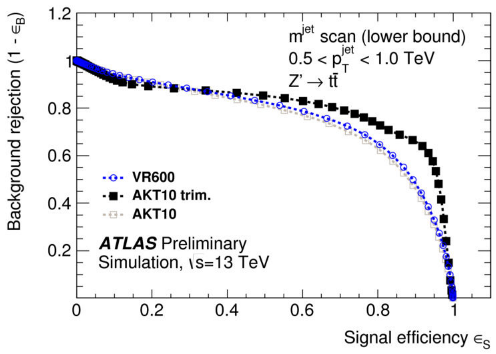
$$\Delta R_{ij} < R_{\mathrm{eff}}. \tag{4.1.7}$$

The maximum separation between the decay products is given by Equation 4.1.1, therefore the value of ρ can be estimated by:

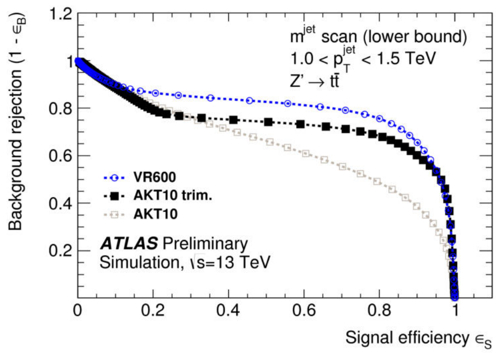
$$2 \cdot m/p_{\mathrm{T}} \approx \rho/p_{\mathrm{T}}. \tag{4.1.8}$$

For hadronic top quark decays, $\rho \approx 2 \cdot m_{\mathrm{top}} = 350$ GeV, while additional shower and hadronization effects may increase the size of top jets. Reconstruction-level studies have been done to determine the optimal choice of the VR parameters, as documented in Ref. [141]. A ρ-value of 600 GeV is chosen as it yields the best overall performance. The maximum jet size is set to $R_{\mathrm{max}} = 1.0$ so that the VR jets and the large-R jets perform identically in low p_{T} regime ($p_{\mathrm{T}} < \rho/R_{\mathrm{max}} = 600$ GeV). The minimum jet size is set to $R_{\mathrm{min}} = 0.2$ to ensure that the effective radius of VR jets is well above the maximum size of the topo-clusters, which will be reached when $p_{\mathrm{T}} > \rho/R_{\mathrm{min}} = 3$ TeV.

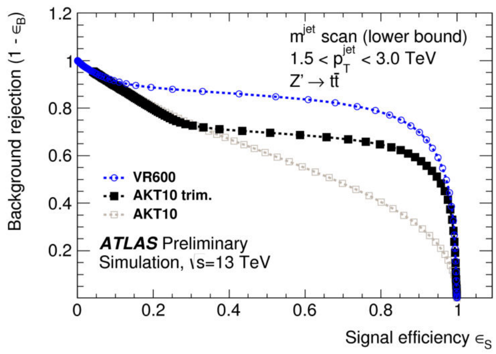
Figure 4.1.1 compares the background rejection versus signal efficiency for ungroomed VR jets, trimmed large-R jets and ungroomed large-R jets in different p_{T} regimes. The curves are obtained by varying the value of a hypothetical lower bound on the jet mass, evaluated with simulated Z' signal and dijet background. For events with boosted topology, VR jets already provide better discrimination power before jet grooming.



(a)



(b)



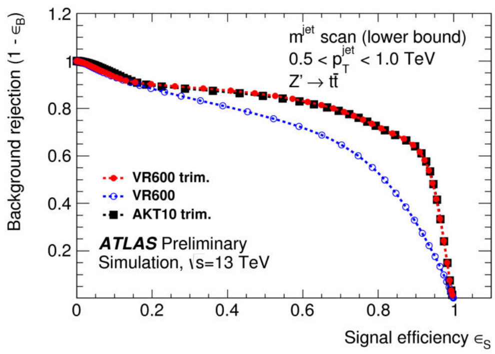
(c)

Figure 4.1.1: Background rejection versus signal efficiency evaluated with simulated Z' signal and dijet background [141]. The performances are shown for ungroomed VR jets (blue circles), trimmed large-R jets (black full squares) and ungroomed large-R jets (gray empty squares) in events with (a) 0.5 TeV $< p_{\mathrm{T}}^{\mathrm{jet}} <$ 1.0 TeV, (b) 1.0 TeV $< p_{\mathrm{T}}^{\mathrm{jet}} <$ 1.5 TeV and (c) 1.5 TeV $< p_{\mathrm{T}}^{\mathrm{jet}} <$ 3.0 TeV. The curves are obtained by varying the value of a hypothetical lower bound on the jet mass.

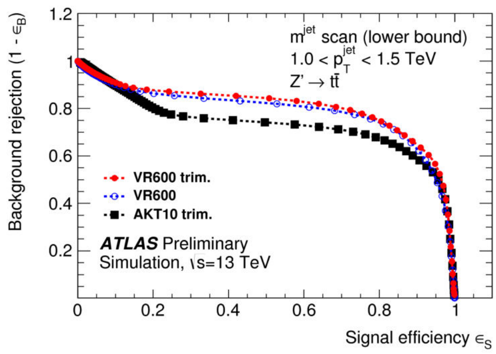
4.1.2 Jet grooming

As shown in Figure 4.1.1, ungroomed VR jets outperform trimmed large-R jets in the identification of jets coming from highly boosted top quarks with $p_{\mathrm{T}} > 1$ TeV. In the resolved regime, however, the performance of ungroomed VR jets is by definition identical to that of ungroomed large-R jets and accordingly affected by contaminations from pile-up and initial state radiations. In order to selectively remove the soft components inside the VR cone and meanwhile provide results comparable to trimmed large-R jets at $p_T < 600$ GeV, a grooming procedure has to be applied during the VR reconstruction, using trimming algorithm with the same set of parameters, $R_{\text{sub}} = 0.2$ and $f_{\text{cut}} = 0.05$, as used for the standard large-R jets. Given the fixed radius of the k_t subjets reclustered for jet trimming, VR jets with large transverse momentum will be mostly insensitive to the grooming procedure. Under extreme circumstances, the k_t reclustering may result in one single subjet composed of all constituents inside the VR jet while no correction is done to the jet substructure after trimming.

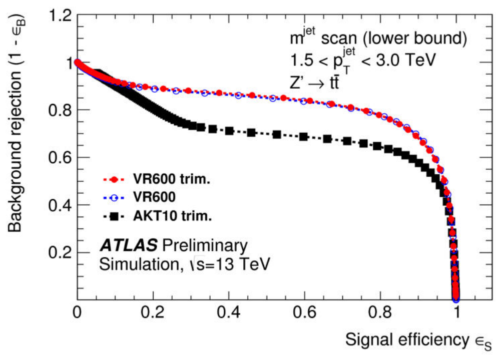
Figure 4.1.2 compares the background rejection versus signal efficiency for trimmed VR jets, ungroomed VR jets and trimmed large-R jets in different p_T regimes. The curves are obtained by varying the value of a hypothetical lower bound on the jet mass, evaluated with simulated Z' signal and dijet background. After the trimming procedure, VR jets are able to provide comparable results with respect to trimmed large-R jets in the resolved regime (0.5 TeV $< p_T <$ 1.0 TeV) while maintaining its discrimination against QCD background in the boosted regime (1.0 TeV $< p_T <$ 3.0 TeV).



(a)



(b)



(c)

Figure 4.1.2: Background rejection versus signal efficiency evaluated with simulated Z' signal and dijet background [141]. The performances are shown for trimmed VR jets (red full circles), ungroomed VR jets (blue empty circles) and trimmed large-R jets (black squares) in events with (a) 0.5 TeV $< p_T^{\text{jet}} <$ 1.0 TeV, (b) 1.0 TeV $< p_T^{\text{jet}} <$ 1.5 TeV and (c) 1.5 TeV $< p_T^{\text{jet}} <$ 3.0 TeV. The curves are obtained by varying the value of a hypothetical lower bound on the jet mass.

4.2 Jet mass scale calibration workflow

The MC-based calibration is applied to the trimmed jets in order to correct, on average, the reconstructed kinematic properties to truth-level by applying calibration factors derived from simulated multijet samples. It consists of three components, namely jet energy scale (JES) calibration, η calibration and jet mass scale (JMS) calibration, which are performed in a sequential order.

The JES and η calibration for large-R and VR jets are derived following the same procedure as used by the calibration of small-R jets, as discussed in Ref. [113]. No explicit pile-up correction is applied, as the residual pile-up dependence is expected to be negligible after the jet grooming process. As the last step of the MC-based calibration, the JMS calibration is derived using isolated large-R or VR jets exclusively matched to isolated particle-level truth jets and is crucial for the identification of hadronically decaying boosted particles.

For each matched pair of reconstruction-level ('reco') and truth jets, the individual jet mass response is defined as:

$$R_{\text{reco}} = m^{\text{reco}} / m^{\text{truth}}. \tag{4.2.1}$$

To extract the JMS calibration factor from Monte Carlo simulation, the input jets are binned in a phase space of reconstruction- or truth-level variables, e.g. $(E^{\text{truth}}, |\eta^{\text{truth}}|, m^{\text{truth}})$, and in each bin the average jet mass response $\langle R_{\text{mass}} \rangle$ is evaluated as the mean of a Gaussian fit to the R_{mass} distribution. Naturally, in order to be able to apply the derived calibration in data, the calibration factor c should depend only on reconstruction-level quantities:

$$m_{\text{calib}} = c(E^{\text{calib}}, |\eta^{\text{calib}}|, m^{\text{reco}}) \cdot m^{\text{reco}}. \tag{4.2.2}$$

Similarly the calibrated mass response can be defined:

$$R_{\text{calib}} = m^{\text{calib}} / m^{\text{truth}} = c(E^{\text{calib}}, |\eta^{\text{calib}}|, m^{\text{reco}}) \cdot R_{\text{reco}}. \tag{4.2.3}$$

A good 'closure' of the calibration is then reached when $R_{\text{calib}} = 1$ in full kinematic phase space.

The numerical inversion technique is applied to obtain the calibration factor c when $\langle R_{\text{reco}} \rangle$ depends on m^{reco} itself:

$$\begin{aligned} c(E^{\text{calib}}, |\eta^{\text{calib}}|, m^{\text{reco}}) &\equiv c(E^{\text{calib}}, |\eta^{\text{calib}}|, \langle R_{\text{reco}} \rangle \cdot m^{\text{truth}}) \\ &= \frac{1}{\langle R_{\text{reco}} \rangle}, \end{aligned} \tag{4.2.4}$$

where $\langle R_{\text{reco}} \rangle \cdot m^{\text{truth}}$ is the average reconstructed mass for jets with m^{truth}, or the 'numerically inverted' mass of m^{truth}. Equation 4.2.5 defines that the calibration factor for jets with m^{reco}, which corresponds to a numerically inverted m^{truth}, is the inverse of the average response at m^{truth}.

In practice, the derivation of the JMS calibration is performed in four steps, namely 'fit raw response', 'smooth raw response', 'fit numerically inverted response' and 'smooth numerically inverted response'.

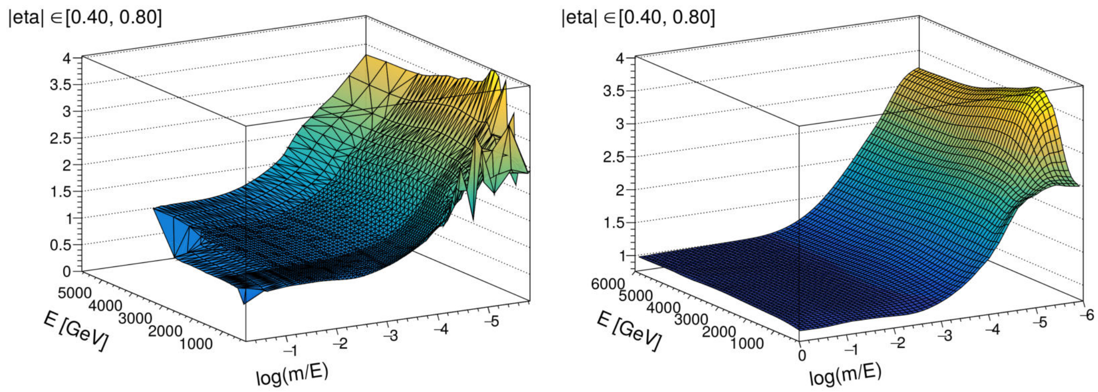
4.2.1 Fit raw response

In this step, selected jets are binned in truth-level kinematic properties to obtain the average jet mass response. Different binning systems are tested for $\langle R_{\text{mass}} \rangle$ calculation, while the $(E, \log(m/E), |\eta_{\text{det}}|)$ binning, as suggested by many studies, is applied to reach the optimal closure performance for the JMS calibration.

The uncalibrated mass response R_{reco} for each jet is computed, while $\langle R_{\text{reco}} \rangle$ is evaluated as the Gaussian mean of the R_{reco} distribution. To avoid biases introduced by the non-Gaussian tail in the mass response, three sequential fits are performed iteratively in truth bins $(E^{\text{truth}}, \log(m^{\text{truth}}/E^{\text{truth}}), |\eta^{\text{truth}}|)$. The range of each fit is defined as $[\mu - \sigma, \mu + \sigma]$, where μ and σ are the mean value and the standard deviation obtained from the last Gaussian fit.

4.2.2 Smooth raw response

Following the Gaussian fits, a two-dimensional Gaussian kernel smoothing is applied to the fitted $\langle R_{\text{reco}} \rangle$ in the $(E^{\text{truth}}, \log(m^{\text{truth}}/E^{\text{truth}}))$ phase space to reduce the fluctuations caused by various non-linear effects. Figure 4.2.1 shows the $\langle R_{\text{reco}} \rangle$ distribution for jets with $|\eta^{\text{truth}}| \in [0.4, 0.8]$ before and after the smoothing procedure.



(a) (b)

Figure 4.2.1: Average jet mass response $\langle R_{\text{reco}} \rangle$ as functions of E^{truth} and $\log(m^{\text{truth}}/E^{\text{truth}})$ for uncalibrated VR jets with $|\eta^{\text{truth}}| \in [0.4, 0.8]$ (a) before and (b) after the two dimensional Gaussian kernel smoothing.

The combination of smoothed $\langle R_{\text{reco}} \rangle$ of different $|\eta^{\text{truth}}|$ bins provides the uncalibrated mass response as a function of E^{truth}, $\log(m^{\text{truth}}/E^{\text{truth}})$ and $|\eta^{\text{truth}}|$. The calibration function can be then expressed in terms of truth variables:

$$c(E^{\text{truth}}, |\eta^{\text{truth}}|, m^{\text{truth}}) = \frac{1}{\langle R_{\text{reco}}\rangle(E^{\text{truth}}, \log(m^{\text{truth}}/E^{\text{truth}}), |\eta^{\text{truth}}|)}. \tag{4.2.5}$$

4.2.3 Fit numerically inverted response

For each jet, the numerically inverted mass can be calculated as

$$m^{\text{ni}} = \langle R_{\text{reco}}\rangle \cdot m^{\text{truth}}, \tag{4.2.6}$$

while the average response at m^{truth} can be obtained from the smoothed $\langle R_{\text{reco}}\rangle$ map obtained in the previous step.

Similar to the 'fit raw response' step, $\langle R_{\text{reco}}\rangle$ is evaluated using three iterative Gaussian fit on the R_{reco} distribution in numerically inverted bins $(E^{\text{calib}}, \log(m^{\text{ni}}/E^{\text{calib}}), |\eta^{\text{calib}}|)$.

4.2.4 Smooth numerically inverted response

The two-dimensional Gaussian kernel smoothing is applied to the fitted $\langle R_{\text{reco}}\rangle$ with respect to E^{calib} and $\log(m^{\text{ni}}/E^{\text{calib}})$. The combination of smoothed $\langle R_{\text{reco}}\rangle$ of different $|\eta^{\text{calib}}|$ bins provides the uncalibrated mass response as a function of E^{calib}, $\log(m^{\text{ni}}/E^{\text{calib}})$ and $|\eta^{\text{calib}}|$. The calibration function can be then expressed in terms of reconstruction-level variables:

$$\begin{aligned} c(E^{\text{calib}}, |\eta^{\text{calib}}|, m^{\text{reco}}) &\equiv c(E^{\text{calib}}, |\eta^{\text{calib}}|, m^{\text{ni}}) \\ &= \frac{1}{\langle R_{\text{reco}}\rangle(E^{\text{calib}}, \log(m^{\text{ni}}/E^{\text{calib}}), |\eta^{\text{calib}}|)}. \end{aligned} \tag{4.2.7}$$

Figure 4.2.2 shows the jet mass response distribution of VR jets in truth bins before and after the calibration. Closure of $\langle R_{\text{reco}}\rangle$ can be observed.

4.3 Monte Carlo samples and event selection

Samples of Monte Carlo simulated events are used for studying the detector response of mass reconstruction for VR and large-R jets. The calibration map is derived with dijet samples simulated with the Pythia 8 [142] generator using NNPDF23LO PDF set [143] and A14 tune [144]. Moreover, the MC simulation takes into account the data condition, trigger menu and μ profile in 2017 data-taking period, and therefore the dijet events are reweighted so that the distribution of the mean number of interactions per crossing $\langle\mu\rangle$ in MC matches to that observed in data.

Two set of alternative MC samples are used for calibration closure validation. An independent set of dijet events are simulated with Pythia 8 using NNPDF23LO PDF set and A14 tune, but reweighted to match the pile-up condition in 2015 + 2016 data. To evaluate the JMS closure on exotic signals, $W' \rightarrow WZ \rightarrow qqqq$ samples are produced

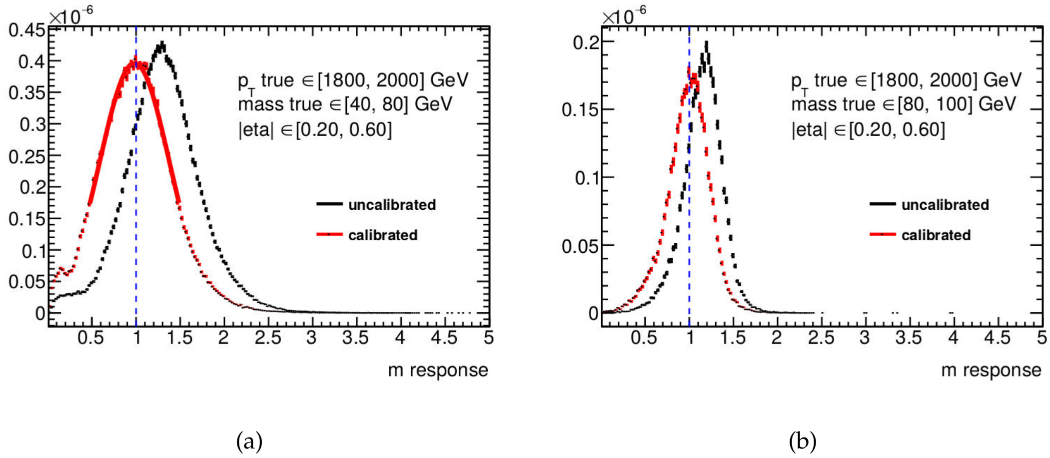


Figure 4.2.2: Jet mass response for VR jets with $p_{\mathrm{T}}^{\mathrm{truth}} \in [1800, 2000]$ GeV, $|\eta| \in [0.2, 0.6]$ and (a) $m^{\mathrm{truth}} \in [40, 80]$ GeV, (b) $m^{\mathrm{truth}} \in [80, 100]$ GeV before (black) and after (red) the calibration.

with Pythia 8 using NNPDF23LO PDF set and A14 tune, reweighted to match the pile-up condition in 2017 data.

To precisely measure the individual mass response, matching between isolated truth jets and isolated reconstructed jets is required. A shown in Table 4.3.1, for VR jets, $\Delta R < 0.6 \cdot R_{\mathrm{eff}}^{\mathrm{truth}}$ has to be fulfilled to declare a truth-reco matching, while the pair having the minimum $\Delta|p_{\mathrm{T}}|$ are selected as the final input. The isolation is defined such that no other truth (reco) jet with $p_{\mathrm{T}} > 100$ GeV can be found with $\Delta R = 2.5 \cdot R_{\mathrm{eff}}^{\mathrm{truth,\,max}}$ ($1.5 \cdot R_{\mathrm{eff}}^{\mathrm{reco,\,max}}$). Obviously, at $p_{\mathrm{T}} < 600$ GeV, the matching and isolation requirements are identical for large-R and VR jets.

selection criterion	**truth-reco matching**	**truth-truth isolation**	**reco-reco isolation**
large-R jet	$\Delta R < 0.6$ minimize $\Delta|p_{\mathrm{T}}|$ when matched to multiple jets	$\Delta R > 2.5$ for truth jets with $p_{\mathrm{T}} > 100$ GeV	$\Delta R > 1.5$ for reconstructed jets with $p_{\mathrm{T}} > 100$ GeV
VR jet	$\Delta R < 0.6 \cdot R_{\mathrm{eff}}^{\mathrm{truth}}$ minimize $\Delta|p_{\mathrm{T}}|$ when matched to multiple jets	$\Delta R > 2.5 \cdot R_{\mathrm{eff}}^{\mathrm{truth,\,max}}$ for truth jets with $p_{\mathrm{T}} > 100$ GeV	$\Delta R > 1.5 \cdot R_{\mathrm{eff}}^{\mathrm{reco,\,max}}$ for reconstructed jets with $p_{\mathrm{T}} > 100$ GeV

Table 4.3.1: Matching and isolation requirements for large-R and VR jets used to derived the JMS calibration function.

Additional requirements on reconstruction quality are applied to veto unphysical results in MC simulation. Reconstructed jets with uncalibrated calorimeter-based mass exceeding its energy are abandoned, while for jets used in track-assisted mass calibration, at least one track needs to be associated to the large-R or the VR jet to ensure a positive track-only mass m^{track}. At event-level, the average p_{T} of the leading and the sub-leading reco jets is compared to the p_{T} of the leading truth jet, fulfilling $(p_{\mathrm{T}}^{\mathrm{reco1}} + p_{\mathrm{T}}^{\mathrm{reco2}})/(2 \cdot$

$p_{\mathrm{T}}^{\mathrm{truth1}}) < 1.4$.

4.4 Closure performance for variable-radius calorimeter jets

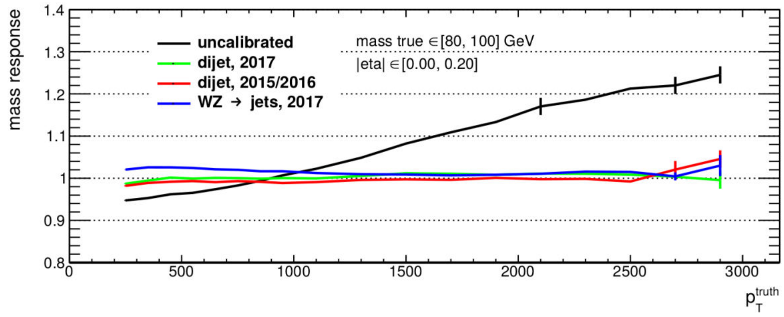
To validate the JMS calibration functions derived for VR jets, the calibrated average mass response $\langle R_{\mathrm{calib}} \rangle$ is compared to one in each bin. The 68.0% coverage interval of the interquartile range[1] of the jet mass response, instead of the standard deviation from a Gaussian fit, is used to evaluate the expected mass resolution, as the jet mass response does not always follow a Gaussian distribution. In Section 4.4.1 and Section 4.4.2, the closure performance of the calorimeter-based mass and the track-assisted mass are shown, respectively, while the closure plots of the combined mass are given in Section 4.4.3.

4.4.1 Calorimeter-based mass

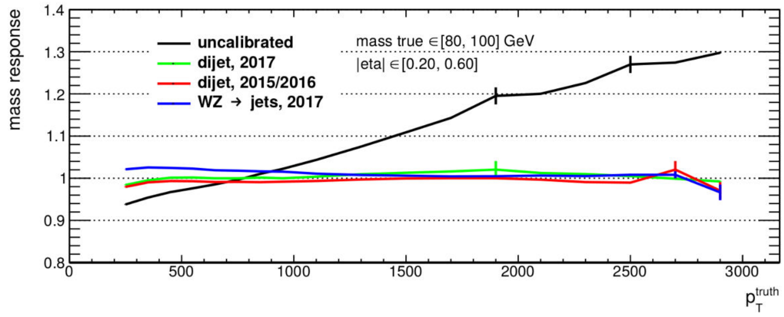
As stated in Section 4.3, Pythia dijet samples reweighted to match the pile-up condition in 2017 are used to derive the JMS calibration, while the corresponding calibrated calorimeter-based mass response is shown in Figure 4.4.1. Additionally, the same calibration function is applied to the alternative MC samples, namely Pythia dijet samples reweighted to match the pile-up condition in 2015 and 2016 and $W' \to WZ \to qqqq$ samples reweighted to match the pile-up condition in 2017. For all samples, the closure responses of m^{calo} get significantly improved after the calibration. Residual non-closure can be seen at high p_{T} regions and $|\eta| \in [0.6, 1.0]$, which is mainly due to the non-linear detector response introduced at the transition region of the TileCal at $|\eta| \in [0.8, 1.1]$.

Figure 4.4.2 presents the calibrated calorimeter-based mass resolution for the above mentioned samples. As expected, the resolutions of the calibrated m^{calo} are no worse than those before the JMS calibration.

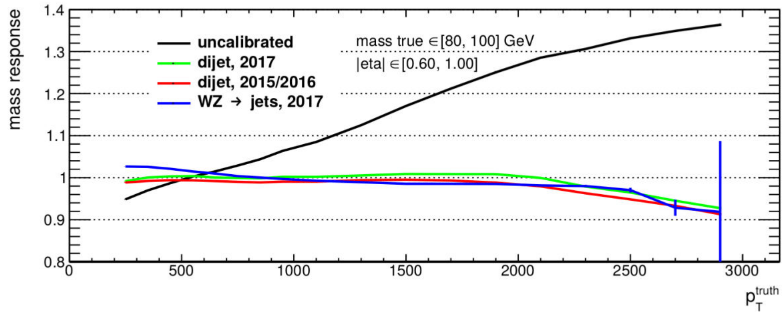
[1] This is defined as $q_{84\%} - q_{16\%}$, where $q_{84\%}$ and $q_{16\%}$ are the 84^{th} and the 16^{th} percentiles of a given distribution.



(a)

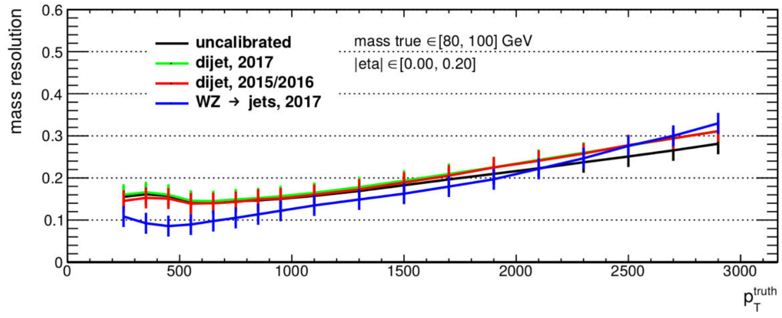


(b)

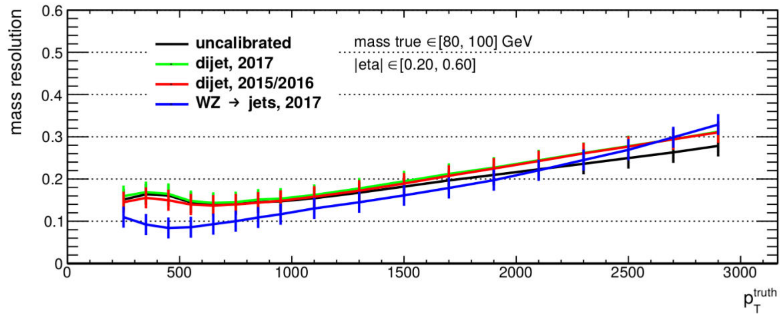


(c)

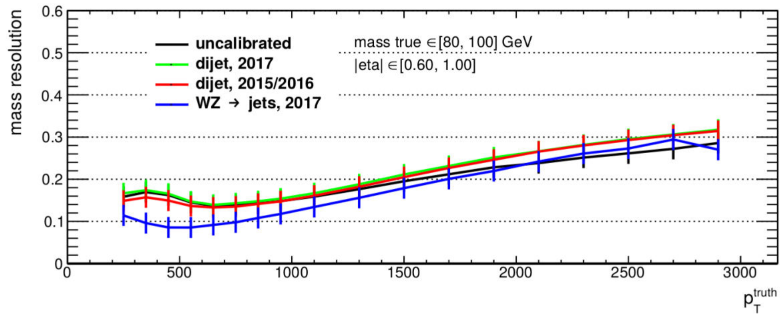
Figure 4.4.1: Average calibrated calorimeter-based mass response for VR jets with $m_{\text{truth}} \in [80, 100]$ GeV and (a) $|\eta| \in [0.0, 0.2]$, (b) $|\eta| \in [0.2, 0.6]$, (c) $|\eta| \in [0.6, 1.0]$ as functions of $p_{\text{T}}^{\text{truth}}$. The calibration is applied to Pythia dijet samples reweighted to match the pile-up condition in 2017 (green), dijet samples reweighted to match the pile-up condition in 2015 and 2016 (red) and $W' \to WZ \to qqqq$ samples reweighted to match the pile-up condition in 2017 (blue). The uncalibrated response for Pythia dijet samples reweighted to match the pile-up condition in 2017 (black) is also drawn. The error bars shown in the plots correspond to the statistical uncertainty of the fit.



(a)



(b)

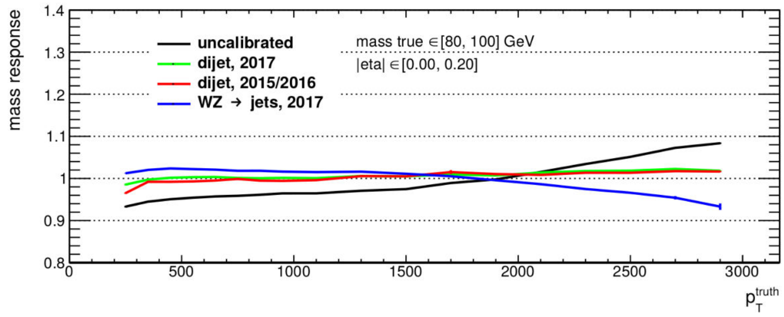


(c)

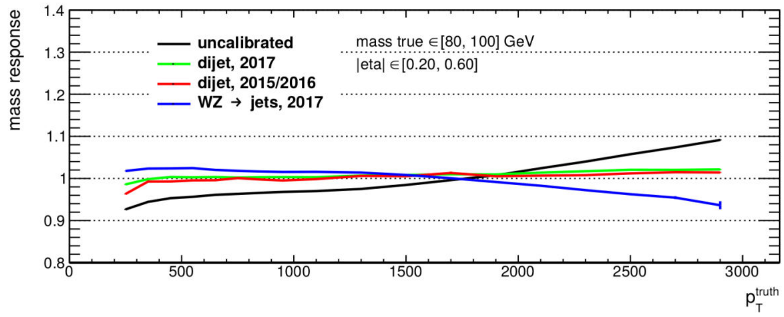
Figure 4.4.2: Average calibrated calorimeter-based mass resolution for VR jets with $m_{\text{truth}} \in [80, 100]$ GeV and (a) $|\eta| \in [0.0, 0.2]$, (b) $|\eta| \in [0.2, 0.6]$, (c) $|\eta| \in [0.6, 1.0]$ as functions of $p_{\text{T}}^{\text{truth}}$. The calibration is applied to Pythia dijet samples reweighted to match the pile-up condition in 2017 (green), dijet samples reweighted to match the pile-up condition in 2015 and 2016 (red) and $W' \to WZ \to qqqq$ samples reweighted to match the pile-up condition in 2017 (blue). The uncalibrated resolution for Pythia dijet samples reweighted to match the pile-up condition in 2017 (black) is also drawn. The error bars shown in the plots correspond to the statistical uncertainty of the fit.

4.4.2 Track-assisted mass

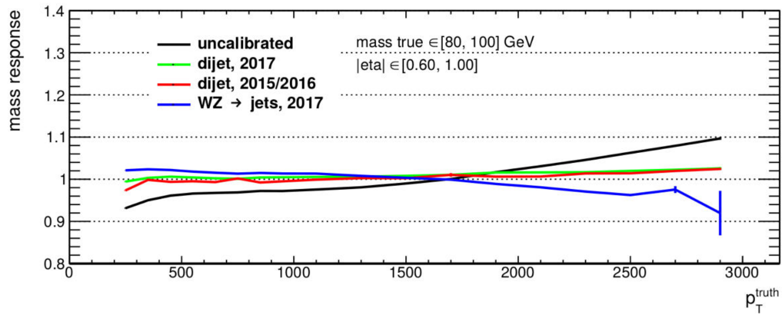
Following the similar procedure of m^{calo} calibration validation, the calibrated track-assisted mass response for different MC samples are shown in Figure 4.4.3. Good closure performances of m^{TA} are reached for the two Pythia dijet samples, while slight residual p_{T} dependence can be observed in the closure response of the $WZ \to qqqq$ samples. As this non-closure can be fully covered by the jet mass resolution (JMR) uncertainties, no additional corrections are derived for $WZ \to qqqq$.



(a)



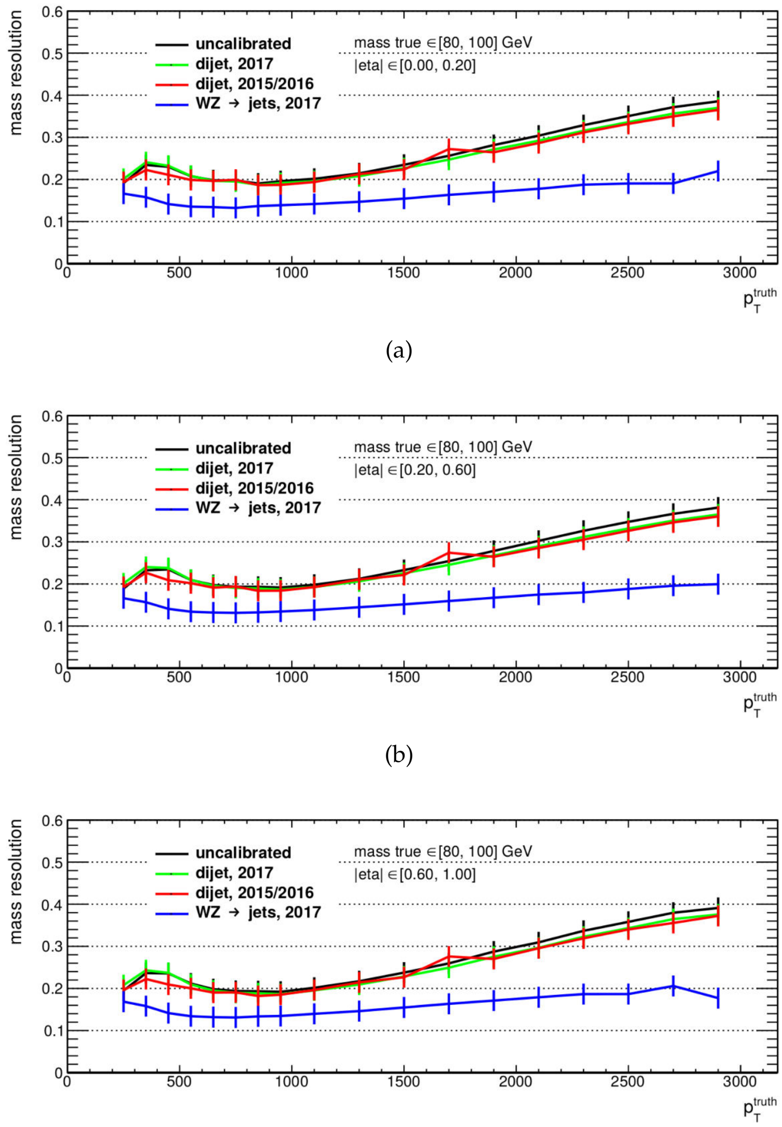
(b)



(c)

Figure 4.4.3: Average calibrated track-assisted mass response for VR jets with $m_{\text{truth}} \in [80, 100]$ GeV and (a) $|\eta| \in [0.0, 0.2]$, (b) $|\eta| \in [0.2, 0.6]$, (c) $|\eta| \in [0.6, 1.0]$ as functions of $p_{\text{T}}^{\text{truth}}$. The calibration is applied to Pythia dijet samples reweighted to match the pile-up condition in 2017 (green), dijet samples reweighted to match the pile-up condition in 2015 and 2016 (red) and $W' \to WZ \to qqqq$ samples reweighted to match the pile-up condition in 2017 (blue). The uncalibrated response for Pythia dijet samples reweighted to match the pile-up condition in 2017 (black) is also drawn. The error bars shown in the plots correspond to the statistical uncertainty of the fit.

Figure 4.4.4 presents the calibrated track-assisted mass resolution for these samples. For each process, the uncalibrated and the calibrated resolutions of m^{TA} are maintained roughly at the same level. Note that the mass resolution of $WZ \to qqqq$ is already much better than those of the dijet backgrounds before the JMS calibration, which is likely due to the cleaner decay patterns expected in the signals events.



(a)

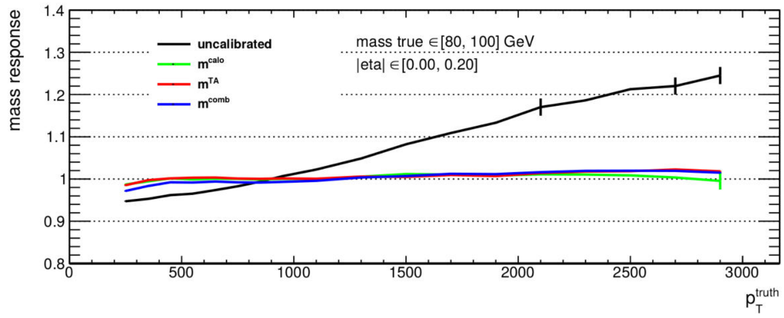
(b)

(c)

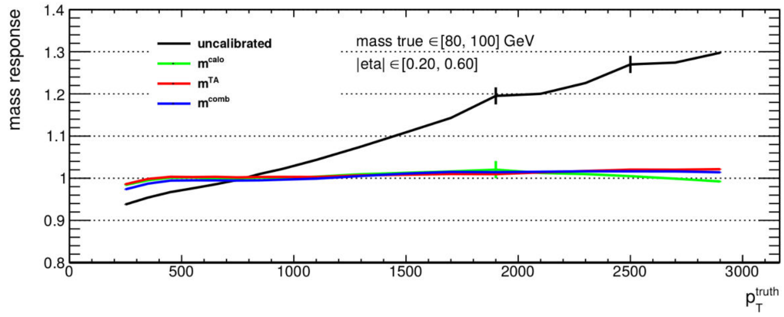
Figure 4.4.4: Average calibrated track-assisted mass resolution for VR jets with $m_{\mathrm{truth}} \in [80, 100]$ GeV and (a) $|\eta| \in [0.0, 0.2]$, (b) $|\eta| \in [0.2, 0.6]$, (c) $|\eta| \in [0.6, 1.0]$ as functions of $p_{\mathrm{T}}^{\mathrm{truth}}$. The calibration is applied to PYTHIA dijet samples reweighted to match the pile-up condition in 2017 (green), dijet samples reweighted to match the pile-up condition in 2015 and 2016 (red) and $W' \to WZ \to qqqq$ samples reweighted to match the pile-up condition in 2017 (blue). The uncalibrated resolution for PYTHIA dijet samples reweighted to match the pile-up condition in 2017 (black) is also drawn. The error bars shown in the plots correspond to the statistical uncertainty of the fit.

4.4.3 Combined mass

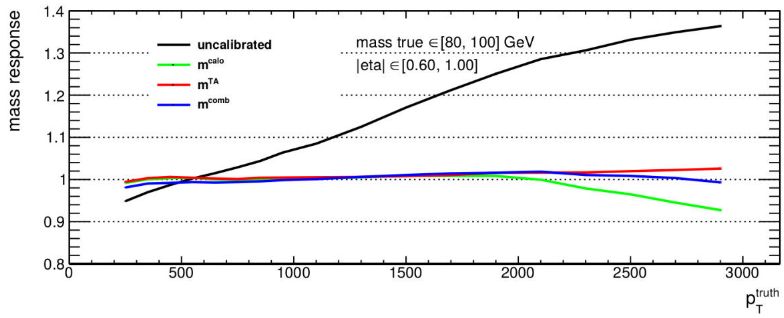
Lastly, the combined mass response is checked with respect to the closure of m^{calo} and m^{TA}. As shown in Figure 4.4.5, m^{comb} is able to provide comparable, if not better, mass response closure over the full p_{T} range. Since m^{comb} is constructed as a linear combination of m^{calo} and m^{TA}, no specific calibration functions are derived for m^{comb} and therefore the closure comparison between alternative MC samples is not to the purpose of calibration validation.



(a)



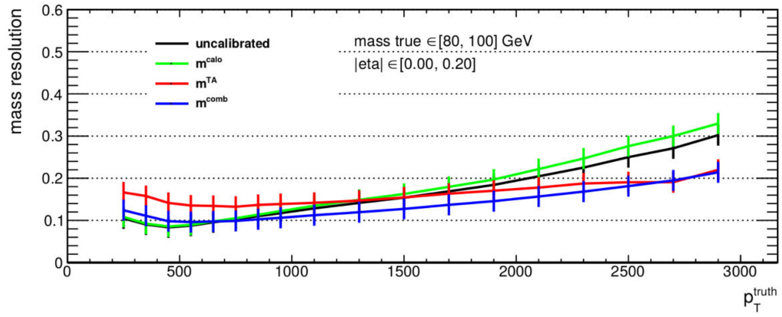
(b)



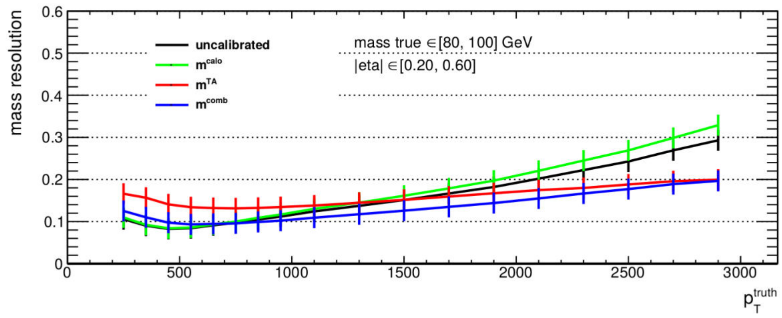
(c)

Figure 4.4.5: Average mass response of the calibrated calorimeter-based mass (green), the calibrated track-assisted mass (red) and the combined mass (blue) for VR jets with $m_{\text{truth}} \in [80, 100]$ GeV and (a) $|\eta| \in [0.0, 0.2]$, (b) $|\eta| \in [0.2, 0.6]$, (c) $|\eta| \in [0.6, 1.0]$ as functions of $p_{\text{T}}^{\text{truth}}$. The uncalibrated calorimeter-based mass response for PYTHIA dijet samples reweighted to match the pile-up condition in 2017 (black) is also drawn. The error bars shown in the plots correspond to the statistical uncertainty of the fit.

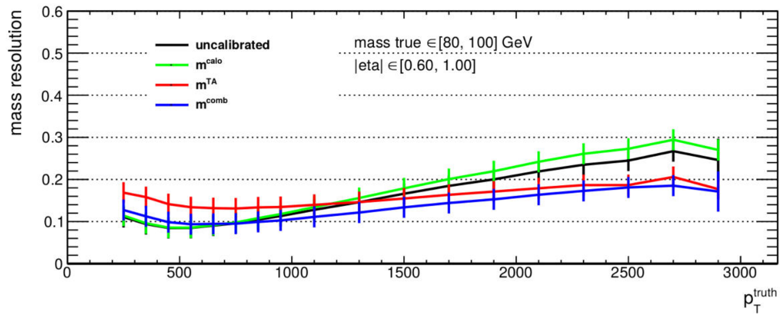
Figure 4.4.6 compares the resolution of the calibrated calorimeter-based mass, the calibrated track-assisted mass and the combined mass. As expected, m^{comb} outperforms both m^{calo} and m^{TA} in terms of mass resolution over the full p_{T} range.



(a)



(b)



(c)

Figure 4.4.6: Average mass resolution of the calibrated calorimeter-based mass (green), the calibrated track-assisted mass (red) and the combined mass (blue) for VR jets with $m_{\text{truth}} \in [80, 100]$ GeV and (a) $|\eta| \in [0.0, 0.2]$, (b) $|\eta| \in [0.2, 0.6]$, (c) $|\eta| \in [0.6, 1.0]$ as functions of $p_{\text{T}}^{\text{truth}}$. The uncalibrated calorimeter-based mass resolution for Pythia dijet samples reweighted to match the pile-up condition in 2017 (black) is also drawn. The error bars shown in the plots correspond to the statistical uncertainty of the fit.

4.5 Summary

The JMS calibration functions are derived for VR jets using MC simulated dijet samples and validated with alternative background and signal processes. Good closure performance can be observed for all mass definitions, while m^{comb} outperforms m^{calo} and m^{TA} over the full p_{T} range.

Figure 4.5.1 illustrates the value of the calibration factors with a fixed m^{calo} as functions of E^{calib} ($p_{\text{T}}^{\text{calib}}$) and $|\eta^{\text{calib}}|$. A sudden increase in $c(E^{\text{calib}}, |\eta^{\text{calib}}|, m^{\text{calo}})$ at $|\eta| \in [1.0, 1.1]$ can be seen in all plots. This corresponds to the transition region between the central barrel and the extended barrels of the TileCal at $|\eta| \in [0.8, 1.1]$.

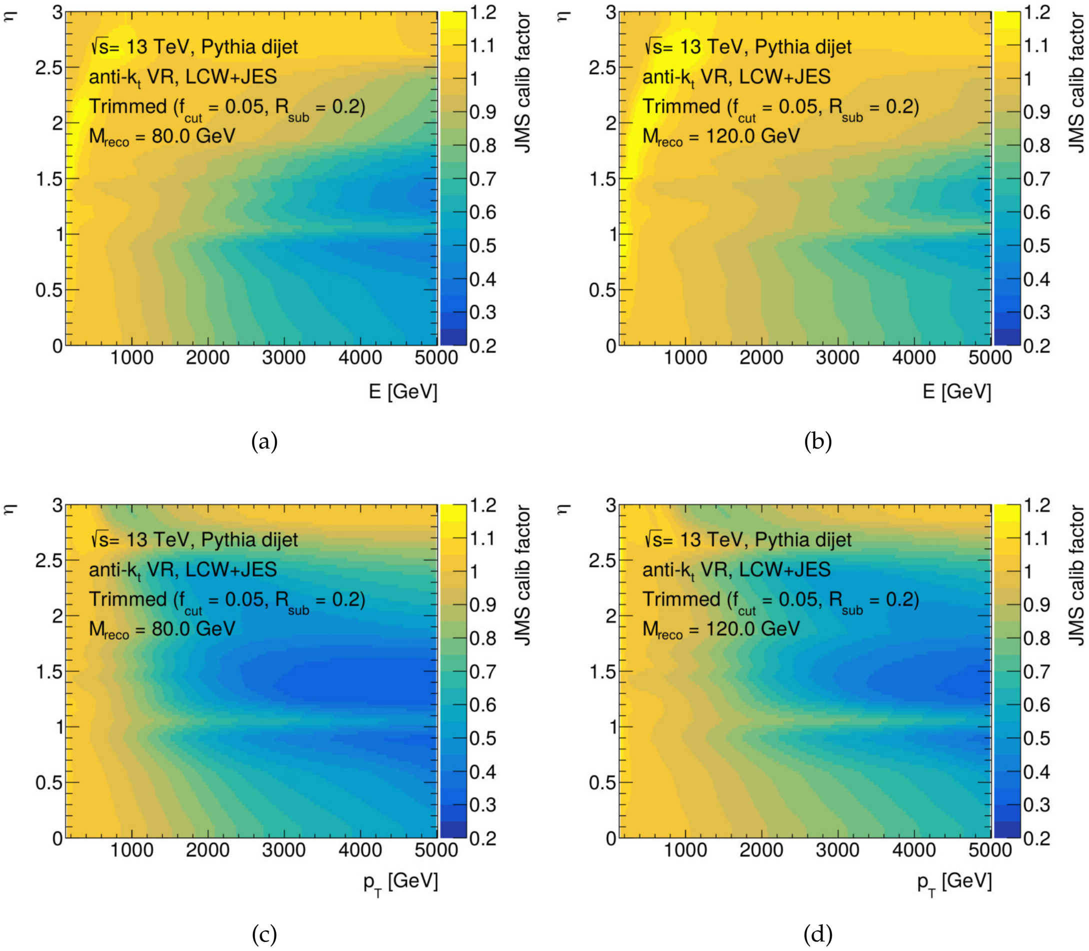


Figure 4.5.1: Calibration maps of the calorimeter-based mass for VR jets with $m^{\text{calo}} = 80$ GeV (left) and $m^{\text{calo}} = 120$ GeV (right) as functions of $|\eta^{\text{calib}}|$ and E^{calib} (top) or $p_{\text{T}}^{\text{calib}}$ (bottom).

Figure 4.5.2 illustrates the value of the calibration factors with a fixed m^{TA} as functions of E^{calib} ($p_{\text{T}}^{\text{calib}}$) and $|\eta^{\text{calib}}|$. The spikes at $|\eta| \in [1.0, 1.1]$ observed in the calibration maps of m^{calo} are not present, while the fluctuations at $|\eta^{\text{calib}}| \in [1.2, 1.5]$ are mostly attributed to the transition between the barrel and the end-caps of the SCT.

To conclude, the JMS calibration is performed using the numerical inversion procedure for VR calorimeter jets and agreement between calibrated and truth-level jet mass is reached on average after the correction applied. The derived calibration functions are ready to be used by any ATLAS analyses using full Run 2 data, allowing for a more accurate description of high p_{T} objects and better pile-up mitigation in extremely boosted event topologies.

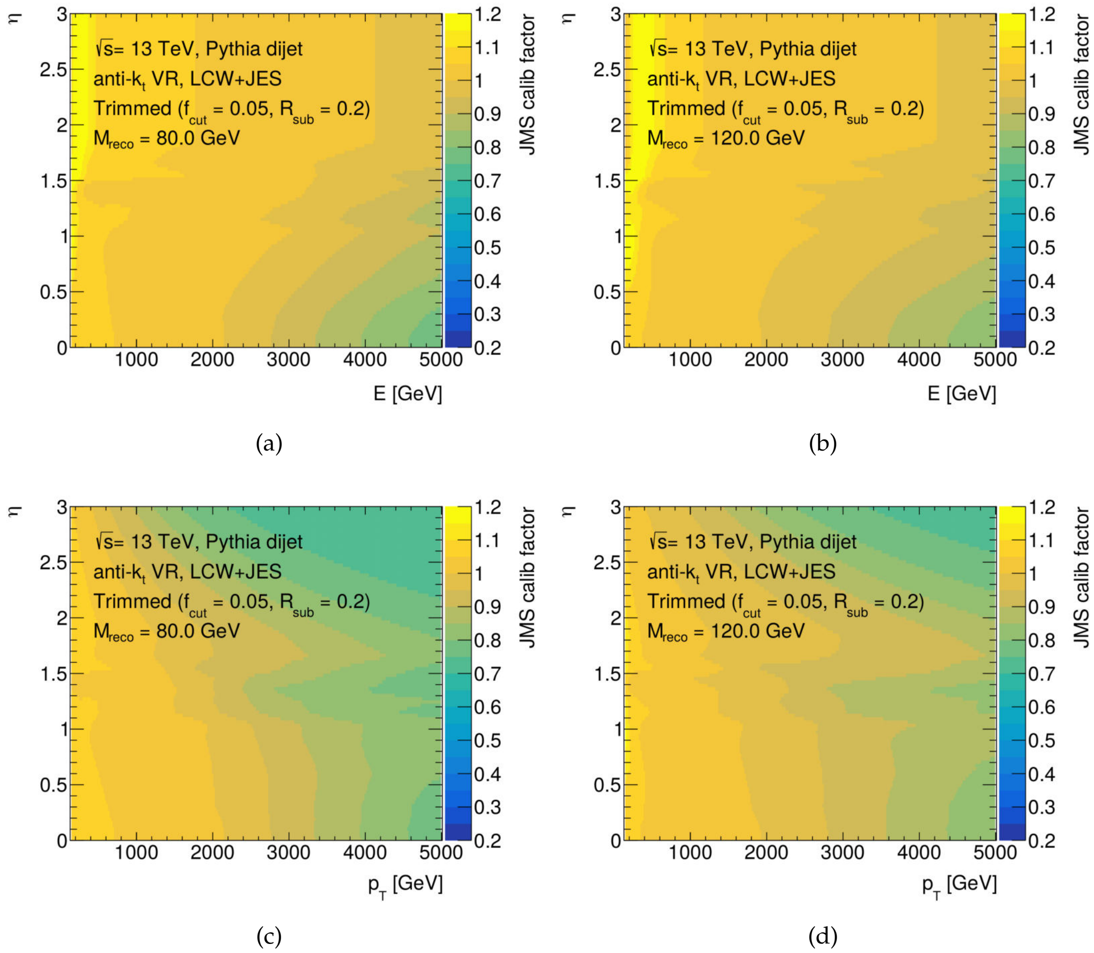
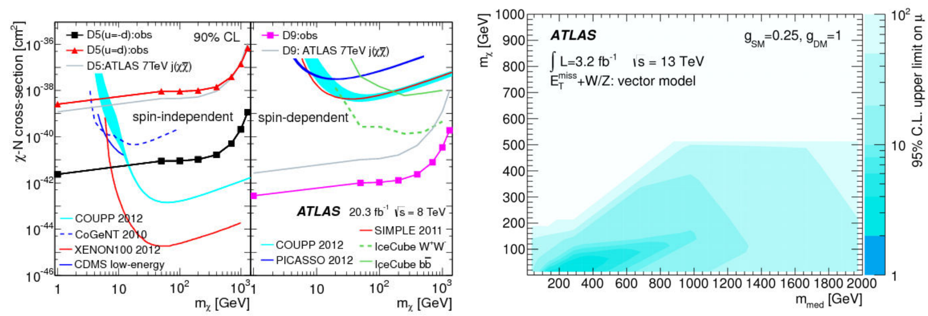


Figure 4.5.2: Calibration maps of the track-assisted mass for VR jets with $m^{\mathrm{TA}} = 80$ GeV (left) and $m^{\mathrm{TA}} = 120$ GeV (right) as functions of $|\eta^{\mathrm{calib}}|$ and E^{calib} (top) or $p_{\mathrm{T}}^{\mathrm{calib}}$ (bottom).

Chapter 5

Search for Dark Matter produced in association with a hadronically decaying vector boson

Searches for Dark Matter in $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) final state have been performed by the ATLAS collaboration at the center-of-mass energy of $\sqrt{s} = 8$ TeV using 20.3 fb^{-1} of pp collision data [145] and at $\sqrt{s} = 13$ TeV using 3.2 fb^{-1} of pp collision data [146], while the latter set limits on the parameters of simplified DM models which include an s-channel vector mediator, as shown in Figure 5.0.1.



(a) (b)

Figure 5.0.1: (a) Observed limit on the DM-nucleon scattering cross section of the spin-independent effective field theory as a function of the DM mass m_{χ} [145], obtained using 20.3 fb^{-1} of pp collision data at $\sqrt{s} = 8$ TeV. (b) Observed limit on the signal strength, μ, for the vector mediator simplified model in the plane of the DM particle mass, m_{χ}, and the mediator mass, m_{med} [146], obtained using 3.2 fb^{-1} of pp collision data at $\sqrt{s} = 13$ TeV.

With the accumulation of $\sqrt{s} = 13$ TeV data and improvements in the detector performance in LHC Run 2, the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search with enhanced sensitivity is performed using data recorded during 2015 and 2016, with an integrated luminosity of 36.1 fb^{-1}.

Section 5.1 focuses on the vector mediator simplified models of Dark Matter, while the data and Monte Carlo samples as well as the triggers applied in the search are described in Section 5.2. Section 5.3 and Section 5.4 summarize the analysis strategy and the definitions of the physical objects being considered in the analysis, respectively. The event selection of the signal regions is given in Section 5.5. Control regions are defined to constrain the dominant background compositions, as presented in Section 5.6. Section 5.7 lists all the systematic uncertainties taken into account in the statistical model. Finally, the results of the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search are shown in Section 5.8.

5.1 Signal models

The vector mediator simplified model of Dark Matter considered in the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search consists of six parameters:

- g_χ, coupling of the DM particles to the mediator;
- g_q, coupling of the SM quarks to the mediator;
- g_l, coupling of the SM leptons to the mediator;
- Γ, decay width of the mediator;
- m_{med}, mass of the mediator;
- m_χ, mass of the DM particle.

Among these parameters, the couplings to the SM and DM particles are chosen following the recommendations of the dark matter forum [147]:

$$g_\chi = 1,\ g_q = 0.25,\ g_l = 0. \tag{5.1.1}$$

The choice of g_χ and g_q is made taking into account the results from the $E_{\mathrm{T}}^{\mathrm{miss}}$+ jet search, the most sensitive channel in all $E_{\mathrm{T}}^{\mathrm{miss}} + X$ final states, as well as the dijet constraints on the processes where the mediator decays back into a pair of SM particles from the resonance searches. The lepton coupling g_l is set to zero to resolve the possible overlap between the dilepton searches.

On the other hand, the minimal total decay width Γ is assumed to allow only for mediator decaying into DM or quark, hence its value can be determined by the choice of the couplings g_χ and g_q. For $g_\chi = 1$ and $g_q = 0.25$, $\Gamma_{\mathrm{min}}/m_{\mathrm{med}} \lesssim 0.06$ is expected, fulfilling the narrow width approximation.

A scan over the two-dimensional m_χ-m_{med} plane is done to exploit the kinematic properties of the regions of interest in the parameter space. As mentioned in Section 2.2.4.1, for a given mediator mass m_{med}, the mass of DM m_χ can be divided into three categories:

- **On-shell**. When $m_{\text{med}} \gg 2m_\chi$, the hardness of the ISR is mainly determined by m_{med} and the kinematic properties of the radiated vector boson do not strongly depend on the value of m_χ. This makes rescaling the results according to models with a same m_{med} and different m_χ applicable, thus no fine scan along m_χ is needed. The on-shell region is where the LHC $E_{\text{T}}^{\text{miss}} + X$ searches have best sensitivity.
- **Threshold**. When $m_{\text{med}} \approx 2m_\chi$, the production of DM will be resonantly enhanced with a much higher dependence on the two mass parameters. A finer scan on m_{med} and m_χ is required for stringent limit setting in this region.
- **Off-shell**. When $m_{\text{med}} \ll 2m_\chi$, only off-shell DM production is allowed, alongside with a significant cross section suppression on hard ISR. The LHC $E_{\text{T}}^{\text{miss}} + X$ searches are expected to have minimal sensitivity to such signals.

As a result, 28 mass points of the vector mediator simplified model signal are generated and tested in the $E_{\text{T}}^{\text{miss}} + V$ (hadronic) search, most of which fall into the on-shell region. An interpolation procedure is then applied to obtain the limits on the signal strength at other mass points, with more details given in Section 5.8.4.

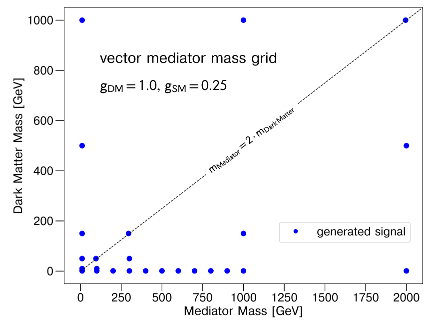


Figure 5.1.1: Grid of generated signal model samples with different configurations for mediator mass m_{med} and DM mass m_χ.

5.2 Data and Monte Carlo simulation

Hard-scatter pp collision data measured by the ATLAS detector are used for various physical analyses, while the signal and background processes are modeled by simulating the detector response to particles produced with Monte Carlo (MC) event generators. Section 5.2.1 provides a brief overview of the data samples used in the $E_{\text{T}}^{\text{miss}} + V$ (hadronic) search. The corresponding signal and background simulations are documented in Section 5.2.2. L1 triggers and HLTs are applied to stabilize the event rate with the increasing instantaneous luminosity during LHC Run 2 data-taking and to reject the overwhelming QCD background which harms the sensitivity of the analysis. Section 5.2.3 is dedicated to the different types of triggers included in the event selection and details on the $E_{\text{T}}^{\text{miss}}$ trigger efficiency calibration are given in Section 5.2.3.1.

5.2.1 Data samples

The $E_{\text{T}}^{\text{miss}} + V$ (hadronic) search uses pp collision data collected by the ATLAS detector during 2015 and 2016 under the premise of stable beam conditions.

All analyzed events are required to meet the standard ATLAS data quality assessment criteria (the 'Good Run List'), ensuring all sub-detector systems operating with full functionality, as described in Section 3.2.5. This leads to a 93% (92%) data-taking efficiency and an integrated luminosity of 3.2 (32.9) fb^{-1} in 2015 (2016), resulting in a total luminosity of 36.1 fb^{-1}.

The average number of pp interactions per bunch crossing $\langle\mu\rangle$ was around 13 in 2015 and 25 in 2016, as shown in Figure 3.1.3.

5.2.2 Monte Carlo samples

Monte Carlo simulated event samples are used to evaluate the signal and background contributions as well as estimate the systematic uncertainties of the statistical analysis. The standard ATLAS MC production consists of two steps: event generation and detector simulation.

The event generation can be divided into two stages. First, the hard-scattering process between two incident partons are calculated with a given perturbative accuracy. As QCD does not provide knowledge of the parton constituents inside the proton, parton distribution functions (PDFs) obtained from fits to experimental observations are used to model the interaction at this stage, often referred to as 'parton-level'. Next, shower and hadronization processes are simulated with different sets of tunable parameters, while the decaying partons start to form color neutral bound states. MC simulations at this stage are referred to as 'particle-level' or 'truth-level' and can be used to calibrate the kinematic properties of reconstructed objects, as detailed in Section 4.2. Different MC generators are applied for event generation, depending on the physical process to be simulated.

The generated truth-level events are then put into detector simulation, which is designed to simulate the interactions between the final particles and the detector materials. This is achieved with the help of the Geant4 toolkit [148].

In this search, the simplified model DM signals are generated in a grid of mediator mass m_{med} and the DM particle mass m_χ, following the recommendations of the dark matter forum [147], with 28 mass points in total. MadGraph5 [149] is used to provide the leading-order (LO) matrix element calculation for the hard-scattering process using NNPDF23LO PDF set [143] and Pythia 8 [142] is used for the simulation of parton shower and hadronization using A14 tune [144].

The major background processes include the production of W and Z bosons in addition to partons (W+jets and Z+jets), top quark production ($t\bar{t}$ and single top), as well as diboson production. The W+jets and Z+jets events are simulated with the Sherpa 2.2.1 [150] generator using NNPDF30NNLO PDF set [151] and normalized to the next-to-next-to-leading order (NNLO) cross section. For $t\bar{t}$ and single top processes, the Powheg [152] generator is used and interfaced to Pythia 6 [153] for the modeling of parton shower, hadronization and underlying events using Perugia2012 tune [154]. Diboson production, including WW, WZ and ZZ, is simulated with Sherpa 2.1 [155] with CT10 PDF set [156].

A summary of the MC generators, PDF and tune sets used for the production of simulated signal and background samples can be found in Table 5.2.1.

Process	Generator	PDF / tune
Vector mediator models	MadGraph5 + Pythia 8 (LO)	NNPDF23LO + A14
V + jets		
W + jets	Sherpa 2.2.1 (NNLO)	NNPDF30NNLO
Z + jets	Sherpa 2.2.1 (NNLO)	NNPDF30NNLO
$t\bar{t}$	Powheg + Pythia 6 (NLO)	Perugia2012
Single top		
s-channel	Powheg + Pythia 6 (NLO)	Perugia2012
t-channel	Powheg + Pythia 6 (NLO)	Perugia2012
Wt	Powheg + Pythia 6 (NLO)	Perugia2012
Diboson		
$Wl\nu Wqq$	Sherpa 2.1 (NLO)	CT10
$Wl\nu Zqq$	Sherpa 2.1 (NLO)	CT10
$WqqZll$	Sherpa 2.1 (NLO)	CT10
$WqqZ\nu\nu$	Sherpa 2.1 (NLO)	CT10
$ZqqZll$	Sherpa 2.1 (NLO)	CT10
$ZqqZ\nu\nu$	Sherpa 2.1 (NLO)	CT10

Table 5.2.1: List of the MC generators, PDF and tune sets used for the production of simulated signal and background samples in the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search.

For all MC simulated samples used in the analysis, the amount of pile-up contribution is adjusted by a reweighting procedure to match the actual condition in the corresponding data-taking period.

5.2.3 Triggers

Two types of triggers are applied to data and MC simulated events depending on the lepton (e/μ) multiplicity. $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers are used for both signal region selection in 0 lepton final state and control region selection in 1 muon final state, while single-lepton triggers, including single-electron and single-muon triggers, are used in 2 lepton final state.

Unlike the offline $E_{\mathrm{T}}^{\mathrm{miss}}$ computation which takes into account all well-calibrated objects, the decision of $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger is made purely based on calorimeter information. As muons are expected to deposit negligible amount of energy inside the calorimeter, their contribution to trigger-level $E_{\mathrm{T}}^{\mathrm{miss}}$ can be considered as equivalent to that of invisible particles. This motivates the usage of $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger in 0 lepton and 1 muon final states, selecting signals as well as semi-leptonic W+jets and $t\bar{t}$ backgrounds with non-zero trigger-level $E_{\mathrm{T}}^{\mathrm{miss}}$.

The $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers are constructed with two levels, L1 and HLT (as described in Section 3.2.5). Trigger towers (with a coarse granularity in $\Delta\eta \times \Delta\phi$ of 0.1×0.1 in the barrel region and 0.4×0.4 in the end-caps) calibrated at the EM scale are used for L1 $E_{\mathrm{T}}^{\mathrm{miss}}$ computation, e.g. L1XE50 which filters events with L1 $E_{\mathrm{T}}^{\mathrm{miss}}$ above 50 GeV. For

HLTs, different methods are used for $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation in 2015 and 2016. During 2015 data-taking, the 'xe' algorithm which uses calorimeter clusters calibrated at the EM scale with full granularity was employed, while the 'mht' algorithm which defines $E_{\mathrm{T}}^{\mathrm{miss}}$ as the negative vector sum of jets reconstructed from topo-clusters using the anti-k_t algorithm with $R = 0.4$ was used in 2016 instead.

Similarly, the single-lepton triggers operate with a two-level structure. At L1 the transverse energy (E_{T}) of electrons and muons are measured using trigger towers in the ECal and trigger chambers in the spectrometer (RPCs and TGCs), respectively. Thresholds of E_{T} are set depending on the data-taking period, e.g. L1EM20VH which filters events with L1 electron E_{T} above 20 GeV and L1MU15 which filters events with muon E_{T} above 15 GeV. Tighter cuts on lepton transverse energy are required in HLTs, ranging from 24 GeV to 300 GeV for electrons and from 20 GeV to 50 GeV for muons. Additional selections on lepton identification, namely lhloose, lhmedium and lhtight, and isolation criteria, namely iloose and ivarmedium, may be applied as well, including requirement on the transverse impact parameter nod0.

A summary of all triggers used in the analysis can be found in Table 5.2.2.

Period	0 lepton	1 lepton	2 lepton
2015	HLT_xe70	HLT_xe70	HLT_e24_lhmedium_L1EM18VH (MC) **OR** HLT_e24_lhmedium_L1EM20VH (data) **OR** HLT_e60_lhmedium **OR** HLT_e120_lhloose **OR** HLT_mu20_iloose_L1MU15 **OR** HLT_mu50
2016 (A)	HLT_xe90_mht_L1XE50	HLT_xe90_mht_L1XE50	HLT_e24_lhtight_nod0_ivarloose **OR** HLT_e60_lhmedium_nod0 **OR** HLT_e60_medium **OR** HLT_e300_etcut **OR** HLT_e140_lhloose_nod0 **OR** HLT_mu24_iloose_L1MU15 (MC) **OR** HLT_mu24_iloose (data) **OR** HLT_mu40
2016 (B - D3)	HLT_xe90_mht_L1XE50	HLT_xe90_mht_L1XE50	HLT_e24_lhtight_nod0_ivarloose **OR** HLT_mu24_ivarmedium **OR** HLT_mu50
2016 (D4 - E3)	HLT_xe100_mht_L1XE50 **OR** HLT_xe110_mht_L1XE50	HLT_xe100_mht_L1XE50 **OR** HLT_xe110_mht_L1XE50	HLT_e26_lhtight_nod0_ivarloose **OR** HLT_mu24_ivarmedium
2016 (F - L)	HLT_xe110_mht_L1XE50	HLT_xe110_mht_L1XE50	HLT_e26_lhtight_nod0_ivarloose **OR** HLT_e60_lhmedium_nod0 **OR** HLT_e60_medium **OR** HLT_e300_etcut **OR** HLT_e140_lhloose_nod0 **OR** HLT_mu26_ivarmedium **OR** HLT_mu50

Table 5.2.2: List of triggers applied in the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search using data recorded during 2015 and 2016. Triggers marked with 'MC' ('data') are applied to MC (data) only.

5.2.3.1 $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger calibration

As different algorithms are used for trigger-level and reconstruction-level ('reco') $E_{\mathrm{T}}^{\mathrm{miss}}$ computation, the efficiency of $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger can be defined as:

$$\text{efficiency} = \frac{N_{\text{events}} \text{ passing reco } E_{\mathrm{T}}^{\mathrm{miss}} \text{ cut } \textbf{AND} \text{ trigger requirement}}{N_{\text{events}} \text{ passing reco } E_{\mathrm{T}}^{\mathrm{miss}} \text{ cut}}. \tag{5.2.1}$$

With the increasing instantaneous luminosity in LHC Run 2, the threshold of $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers gradually rises in order to maintain a stable event rate, which leads to reduced trigger efficiency for events with rather low reconstruction-level $E_{\mathrm{T}}^{\mathrm{miss}}$. Typically, the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers are only fully efficient with $E_{\mathrm{T}}^{\mathrm{miss}} > 250$ GeV, as shown in Figure 5.2.1.

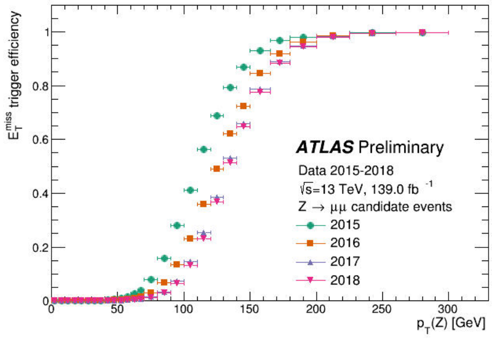


Figure 5.2.1: Combined L1 and HLT efficiency of the lowest unprescaled $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers for the years 2015 to 2018 as functions of Z boson p_{T} in events passing a $Z \rightarrow \nu\nu$ selection [157]. Muons are treated as invisible objects by the triggers concerned.

The $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search tries to exploit the DM signatures over the full $E_{\mathrm{T}}^{\mathrm{miss}}$ range of above 150 GeV. In parts of the phase space included in the analysis where the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers are not fully efficient, corrections are derived to account for the trigger inefficiencies and the possible MC mismodelings with respect to the data. Given the fact that the calculation of trigger-level $E_{\mathrm{T}}^{\mathrm{miss}}$ includes no muon information, one can assume that the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger selections are uncorrelated with the kinematic properties of muons. Thus, the calibration functions for $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers can be evaluated with events passing single-muon trigger requirements. To mimic the event signature in 0 lepton final state at the trigger-level, the muon contributions are subtracted from the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$, denoted by $E_{\mathrm{T,\,no}\,\mu}^{\mathrm{miss}}$, while all other event selections are consistent with the resolved signal regions (see Section 5.5).

The efficiencies of $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers are parametrized as functions of $E_{\mathrm{T,\,no}\,\mu}^{\mathrm{miss}}$ in the range of 120 GeV $< E_{\mathrm{T,\,no}\,\mu}^{\mathrm{miss}} <$ 300 GeV in events with exactly one muon:

$$f\left(E_{\mathrm{T,\,no}\,\mu}^{\mathrm{miss}}\right) = 0.5 \cdot \left[1 + Erf\left(\frac{E_{\mathrm{T,\,no}\,\mu}^{\mathrm{miss}} - p_0}{\sqrt{2}p_1}\right)\right], \tag{5.2.2}$$

where p_0 and p_1 are the two floating parameters of the fit. Corrections on the residual difference between data and MC are then defined as the ratio of their efficiency functions, derived inclusively in b-tag multiplicity. However, the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers are sensitive to the presence of the b-tagged jets in the event, due to the different calorimeter responses to light- and heavy-flavor jets. This contributes as the major source of systematic uncertainty of the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger calibration and can be estimated by taking the difference between calibrations obtained from events with any b-jet multiplicity and with at least one b-jet. Figure 5.2.2 compares the scale factors derived inclusively and in events with ≥ 1 b-jet in all data-taking periods from 2015 to 2016. The total calibration uncertainty is evaluated as the sum in quadrature of b-tag composition uncertainty and the 1σ confidence interval of the fit.

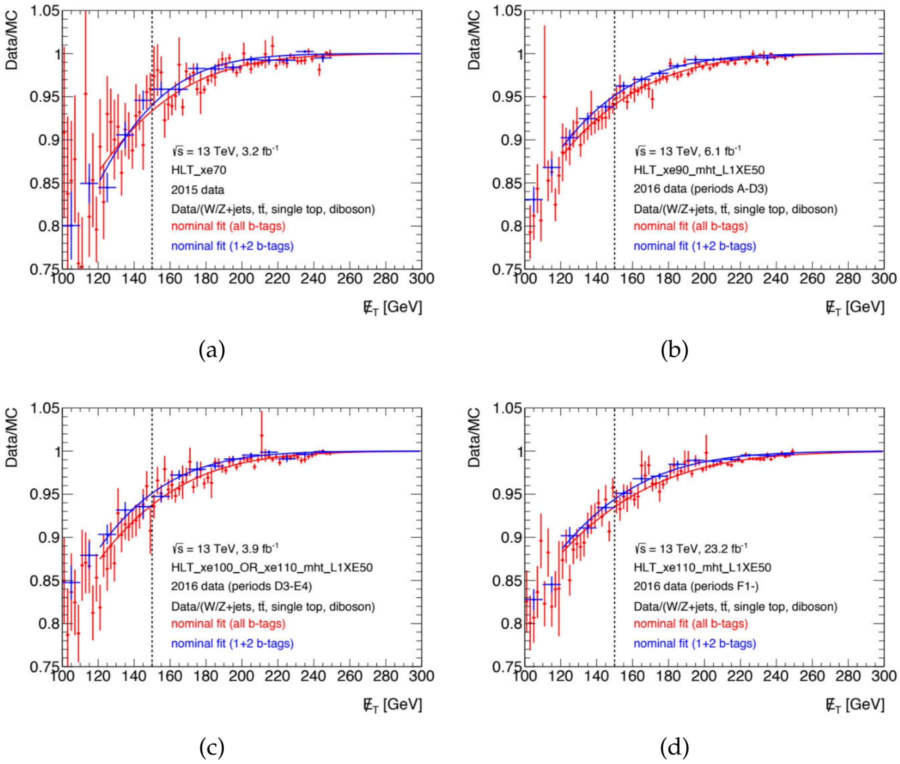


Figure 5.2.2: Comparisons between scale factors derived inclusively in b-tag multiplicity (red) and in events with ≥ 1 b-jet (blue), as functions of $E_{\mathrm{T,\,no}\,\mu}^{\mathrm{miss}}$ for (a) HLT_xe70, (b) HLT_xe90_mht_L1XE50, (c) HLT_xe100_mht_L1XE50 **OR** HLT_xe110_mht_L1XE50 and (d) HLT_xe110_mht_L1XE50.

5.3 Analysis strategy

A general overview of the analysis strategy of the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search is given in this section, including event categorization, main discriminants for signal event selection and treatment of the dominant SM backgrounds. A brief summary of all signal regions (SRs), control regions (CRs) and validation regions (VRs) used in the analysis is presented in Section 5.3.3.

5.3.1 Signal region strategy

The final state of the analyzed signal scenarios contains DM particles recoiling against a vector boson, while the missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ attributed to DM is expected to be of the same order as the vector boson p_{T}. In events with large $E_{\mathrm{T}}^{\mathrm{miss}}$, the two-pronged hadronic decay of the vector boson can be reconstructed within one single large-R jet, which is often referred to as the boosted topology, as discussed in Section 3.3.3.2. In addition to this, events with rather low $E_{\mathrm{T}}^{\mathrm{miss}}$ are also considered in this search, where the separation between the decay products is sufficiently large to be recognized as two separate small-R jets, namely the resolved topology.

Two sets of distinctive event selections are considered in this analysis, optimized for the boosted and the resolved topology, respectively. Events in the boosted regime are required to have $E_{\mathrm{T}}^{\mathrm{miss}} > 250$ GeV, while for the resolved regime the $E_{\mathrm{T}}^{\mathrm{miss}}$ cut is relaxed to 150 GeV. Since the kinematic region of $E_{\mathrm{T}}^{\mathrm{miss}} > 250$ GeV is shared by both regimes, a priority-boosted strategy is applied to resolve the possible overlap. For events satisfying both the boosted and the resolved criteria, priority is always given to the boosted as the majority of the sensitivity of this search comes from the boosted regime.

Event categorization based on b-jets multiplicity is introduced to recover sensitivities to the signal processes involving heavy-flavor components, especially DM production in association with $Z \to bb$ decay. This also helps to suppress the V+jets and $t\bar{t}$ background in the signal regions. Three b-tag categories, 0 b-tag, 1 b-tag and 2 b-tag, are defined for each kinematic regime. Track jets ghost-associated to large-R jets are used for b-tagging in the boosted regime, while small-R jets are used in the resolved regime.

The main discriminant of the analysis is the combined mass of the leading[1] large-R jet or the invariant mass of the leading two small-R jets. All signal region events are required to pass a mass window selection around the mass of W/Z. An additional discriminant which models the substructure of the large-R jet is applied in the boosted selection, with more details documented in Section 5.5.3.

Based on the above mentioned requirements, in total 16 regions are defined in 0 lepton final state, including six high purity SRs (two kinematic regimes with three b-tag categories each), two low purity SRs (relaxed substructure requirement for events with 0 or 1 b-tag multiplicity in the boosted regime) and eight W/Z mass sideband regions, each corresponding to one of the high/low purity SRs. The mass sidebands consist of events passing all signal region selections except for the mass requirement and serve as validation regions of the analysis. All these regions are included in the combined profile likelihood fit, as discussed in Section 5.8.1.

5.3.2 Control region strategy

The common experimental signature searched in this analysis consists of jets from vector boson decay, rather significant $E_{\mathrm{T}}^{\mathrm{miss}}$ and no isolated leptons. The dominant background sources are:

- $Z \to \nu\nu$ plus additional jets from QCD gluon splitting or the 'scattering' diagrams;
- $W \to l\nu$ plus additional jets from QCD gluon splitting, where the lepton is either lost or misidentified as a jet;
- $t\bar{t}$ production containing $W \to l\nu$ decay, where the lepton is either lost or misidentified as a jet;
- WZ/ZZ diboson production, where a vector boson decays hadronically and the Z boson decays to $\nu\nu$, the irreducible background of the analysis.

In addition, subdominant background components include:

- single top production containing $W \to l\nu$ decay, where the lepton is either lost or misidentified as a jet;
- WW diboson production containing $W \to l\nu$ decay, where the lepton is either lost or misidentified as a jet.

[1] All jets are ranked in decreasing p_{T}, thus the jet with the highest p_{T} is often referred to as the 'leading' jet.

The background contributions are modeled with MC simulations and constrained by the usage of CRs, specifically designed to be enriched in certain background processes. The 1 muon CRs are used to constrain the W+jets and $t\bar{t}$ backgrounds, while the 2 lepton CRs are used to constrain the $Z \to \nu\nu$ background using $Z \to ll$ events, since the hard-scattering processes for these two backgrounds are very similar. The WZ/ZZ background, however, is irreducible in this analysis and cannot be distinguished from the DM signal. Fortunately the rather small diboson cross section prevents it from becoming a major component in the SRs. The $WZ \to qq\nu\nu$ process can be constrained in the 1 muon CRs through $WZ \to l\nu qq$, while $ZZ \to qq\nu\nu$ can be constrained in the 2 lepton CRs through $ZZ \to qqll$.

The QCD multijet production, on the contrary, is hard to be modeled with pure MC due to the large amount of computing resources required for the event simulation in this case. Multijet events with fake $E_{\mathrm{T}}^{\mathrm{miss}}$ coming from poorly measured jets may contaminate the SRs, while anti-QCD cuts (Section 5.5.2) are specially designed to suppress this process. The remaining multijet contribution is estimated using a data-driven 'template' method, as discussed in Section 5.6.4.

Both the 1 muon and the 2 lepton CRs are divided into 12 disjoint regions, defined by the event topology being either boosted or resolved, passing or failing the mass selection and the three b-tag multiplicities. Note that as the large-R jet substructure requirements are dropped in the CRs for higher event yields, the number of CRs in either 1 muon or 2 lepton final state is less than that in 0 lepton final state.

5.3.3 Summary of analysis regions

Table 5.3.1 provides a summary of all signal regions, control regions and validation regions defined in the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search.

Kinematic regime		**Boosted**					**Resolved**		
b-tag multiplicity		0		1		2	0	1	2
substructure requirement		⧄	⊠	⧄	⊠				
0 lepton	mass window	●	◐	●	◐	●	●	●	●
	mass sideband	◇	◇	◇	◇	◇	◇	◇	◇
1 lepton	mass window	○		○		○	○	○	○
	mass sideband	○		○		○	○	○	○
2 lepton	mass window	○		○		○	○	○	○
	mass sideband	○		○		○	○	○	○

Table 5.3.1: Summary of the high purity signal regions '●', low purity signal regions '◐', control regions '○' and validation regions '◇' defined in the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search.

5.4 Object definition

A large variety of physical object definitions are applied in the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search, including different kinematic requirements, identification and isolation criteria, etc. A brief introduction is given as follows.

Two types of electron definitions are considered:

- **Loose**. Loose electrons are used in the signal region selection and are required to have $p_{\mathrm{T}} > 7$ GeV and $|\eta| < 2.47$. The LooseLH identification criterion and the LooseTrackOnly isolation criterion have to be fulfilled, as defined in Section 3.3.1.
- **Medium**. Medium electrons are used in the 2 lepton CR selection and are required to have $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.47$. Same identification and isolation requirements as for the loose definition are applied.

Three types of muon definitions are considered:

- **Loose**. Loose muons are used in the signal region selection and are required to have $p_{\mathrm{T}} > 7$ GeV and $|\eta| < 2.7$. The Loose identification criterion and the LooseTrackOnly isolation criterion have to be fulfilled, as defined in Section 3.3.2.
- **Medium**. Medium muons are used in the 2 lepton CR selection and are required to have $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$. Same identification and isolation requirements as for the loose definition are applied.
- **Tight**. Tight muons are used in the 1 muon CR selection and are required to have $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$. The Medium identification criterion and the FixedCut-TightTrackOnly isolation criterion have to be fulfilled, as defined in Section 3.3.2.

The small-R jets used in the analysis are reconstructed from topo-clusters using the anti-k_{t} algorithm with $R = 0.4$ and calibrated at the EM+JES scale. Furthermore they can be divided into two categories based on their pseudorapidity:

- **Central jets**, with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$;
- **Forward jets**, with $p_{\mathrm{T}} > 30$ GeV and $2.5 \leq |\eta| < 4.5$, used in the anti-QCD selections only (Section 5.5.2).

For better pile-up suppression, low p_{T} central jets (i.e. $p_{\mathrm{T}} < 60$ GeV) are required to pass the Jet Vertex Tagger (JVT) [134] selection with JVT score > 0.59. Correspondingly, a JVT efficiency scale factor is applied to the MC event weight for each event passing the selection, which corrects the residual difference in JVT score evaluation between data and MC simulation.

The MV2c10 discriminant (Section 3.3.3.1) is used to identify small-R jets originating from b-hadrons. A fixed-cut working point providing 70% b-tagging efficiency is applied in the resolved regime selection.

The large-R jets used in the analysis are reconstructed from topo-clusters using the anti-k_{t} algorithm with $R = 1.0$ and calibrated at the LCW+JES+JMS scale. The trimming procedure is applied to the large-R jets by reclustering the initial jet constituents using the k_{t} algorithm with $R_{\mathrm{sub}} = 0.2$ and then removing any subjets with p_{T} less than $f_{\mathrm{cut}} = 0.05$ times the p_{T} of the parent jet. Lastly, the trimmed large-R jets are required to have $p_{\mathrm{T}} > 200$ GeV, $m > 30$ GeV and $|\eta| < 2.0$, in order to select high p_{T} central jets with a good overlap between the ID and the calorimeter in a phase space where the p_{T} and the mass calibrations are available.

Track jets reconstructed from charged particle tracks using the anti-k_{t} algorithm with $R = 0.2$ are used for b-hadron identification in events with boosted topology, while the 'ghost-association' technique is applied to uniquely match track jets to the selected large-R jets. All track jets are required to have $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.5$, within the full coverage of the ID. The track-based MV2c10 discriminant (Section 3.3.3.3) is used to identify track jets originating from b-hadrons, giving a fixed 70% b-tagging efficiency in the boosted regime selection.

The missing transverse momentum is reconstructed using well-calibrated physical objects and the track-based soft term (TST), while the Loose working point is applied considering the rather low $\langle\mu\rangle$ in 2015 and 2016 data-taking. In the 1 muon and the 2 lepton CRs, modified versions of $E_{\mathrm{T}}^{\mathrm{miss}}$, namely $E_{\mathrm{T,\,no}\,\mu}^{\mathrm{miss}}$ and $E_{\mathrm{T,\,no\,lepton}}^{\mathrm{miss}}$, are computed by subtracting the lepton contribution from the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector. This is achieved by removing all components associated to the target lepton, including the calorimeter clusters and ID tracks, from the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction input, which ensures the invisibility of the leptons.

The track-based missing transverse momentum is reconstructed from ID tracks satisfying certain track quality criteria [133] and is used to reject events with $E_{\mathrm{T}}^{\mathrm{miss}}$ originating from mismeasured QCD processes. Similarly, modified versions of track-based $E_{\mathrm{T}}^{\mathrm{miss}}$, namely track-based $E_{\mathrm{T,\,no}\,\mu}^{\mathrm{miss}}$ and track-based $E_{\mathrm{T,\,no\,lepton}}^{\mathrm{miss}}$, are applied in the 1 muon and the 2 lepton CR selections, respectively.

A concise summary of the various object definitions used in this analysis is presented in Table 5.4.1.

Object	**Kinematics**	**Type, Quality**
Electrons	loose: $p_{\mathrm{T}} > 7$ GeV, $|\eta| < 2.47$	LooseLH, LooseTrackOnly
	medium: $p_{\mathrm{T}} > 25$ GeV, $|\eta| < 2.47$	LooseLH, LooseTrackOnly
Muons	loose: $p_{\mathrm{T}} > 7$ GeV, $|\eta| < 2.7$	Loose, LooseTrackOnly
	medium: $p_{\mathrm{T}} > 25$ GeV, $|\eta| < 2.5$	Loose, LooseTrackOnly
	tight: $p_{\mathrm{T}} > 25$ GeV, $|\eta| < 2.5$	Medium, FixedCutTightTrackOnly
Small-R jets	central: $p_{\mathrm{T}} > 20$ GeV, $|\eta| < 2.5$	anti-k_t $R = 0.4$, EMTopo
	forward: $p_{\mathrm{T}} > 30$ GeV, $|\eta| = [2.5, 4.5)$	b-tag: MV2c10, fixed-cut 70%
Large-R jets	$p_{\mathrm{T}} > 200$ GeV, $m > 30$ GeV, $|\eta| < 2.0$	anti-k_t $R = 1.0$, LCTopo trimmed, $R_{\mathrm{sub}} = 0.2$, $f_{\mathrm{cut}} = 0.05$
Track jets	$p_{\mathrm{T}} > 10$ GeV, $|\eta| < 2.5$	anti-k_t $R = 0.2$ b-tag: MV2c10, fixed-cut 70%
$E_{\mathrm{T}}^{\mathrm{miss}}$	$E_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV	track-based soft term Loose working point
Track-based $E_{\mathrm{T}}^{\mathrm{miss}}$	track-based $E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV	ID tracks $p_{\mathrm{T}} > 500$ MeV

Table 5.4.1: Summary of the reconstructed objects used in the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search.

5.4.1 Resolving overlapping objects

Since the reconstruction algorithms of different physical objects are mostly independent, potential ambiguities of one single detector response being assigned to multiple final state objects may arise. Hence, an overlap removal procedure is applied to resolve such ambiguities of object definition and avoid any double-counting of detector signals. This is done in the following order:

- **Electron-muon overlap removal**. Electron candidates are removed if they share the same ID tracks with a CB muon. In the case of CT muons, the muon candidates are removed instead.

- **Electron-jet overlap removal**. This is performed in two steps. First, small-R jets within $\Delta R = 0.2$ of any well-identified electrons are removed. Then, electrons within $\min\{0.4, 0.04 + 10\text{ GeV}/p_{\mathrm{T}}^{\mathrm{electron}}\}$ of any surviving small-R jets are removed to avoid double-counting of energy.

- **Muon-jet overlap removal**: Similarly, small-R jets within $\Delta R = 0.2$ of any well-identified muons are removed if the jet has fewer than three associated tracks or if the muon p_{T} is greater than half the jet p_{T} and greater than 70% of the summed p_{T} of all tracks associated to the jet. Then, muons within $\min\{0.4, 0.04 + 10\text{ GeV}/p_{\mathrm{T}}^{\mathrm{muon}}\}$ of any surviving small-R jets are removed to avoid double-counting of energy. This avoids the inefficiency of high p_{T} muons undergoing significant energy loss in the calorimeter.

Note that the distance metric used to define the overlapping objects is $\Delta R = \sqrt{\Delta y^2 + \Delta\phi^2}$, where the rapidity y is used instead of the pseudorapidity η, as recommended in Ref. [158].

5.5 Event selection

As stated in Section 5.3, the signal events of the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search are characterized by jets with mass (and substructure) compatible with a hadronically decaying W/Z, rather significant $E_{\mathrm{T}}^{\mathrm{miss}}$ and no isolated leptons. The SR selections are designed and optimized for such final states, while event categorization based on kinematic topologies and b-jet multiplicities are introduced to further enhance the sensitivity to DM signals. Section 5.5.1 discusses the baseline selection applied to all regions in the analysis. Section 5.5.2 details the anti-QCD requirements applied for multijet background suppression. The full list of the SR selections is presented in Section 5.5.3.

5.5.1 Baseline selection

All events are required to satisfy the following basic quality criteria:

- **GRL and event cleaning**: data events failing the standard ATLAS data quality assessment are vetoed.

- **Trigger**: events in 0 lepton or 1 muon final state are required to pass the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger selection, while the 2 lepton events are required to pass the single-lepton trigger selection, see Section 5.2.3.

- **Vertex selection**: at least one reconstructed vertex with at least two associated tracks is required.

- **Jet cleaning**: events containing jets flagged by the BadLoose jet cleaning requirement [159] are vetoed to ensure a good measurement of $E_{\mathrm{T}}^{\mathrm{miss}}$.

5.5.2 Anti-QCD requirements

To select signal events in 0 lepton final state, events with loose electrons or muons are vetoed and $E_{\mathrm{T}}^{\mathrm{miss}}$ greater than 150 and 250 GeV is required for the resolved and the boosted selection, respectively. However, multijet events may enter the SRs due to mismeasurement or miscalibration of jet p_{T}, causing unbalanced momentum in the transverse plane. Given its large cross section at the LHC, QCD multijet production can be considered as one of the most dominant background processes in hadronic final states. This adds layers of complexity to the analysis, as no precise multijet simulation is available due to computational limitations. In order to reduce the contribution of QCD multijet to a negligible amount, a set of requirements aiming to reject events with $E_{\mathrm{T}}^{\mathrm{miss}}$ originating from mismeasured jets, often referred to as the 'anti-QCD' requirements, is applied after the baseline selection, with descriptions below:

- Track-based $E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV. Events with large $E_{\mathrm{T}}^{\mathrm{miss}}$ originating from calorimeter mismeasurements can be suppressed using the track-based missing transverse momentum reconstructed with ID tracks. The track-based $E_{\mathrm{T}}^{\mathrm{miss}}$ selection also helps to significantly reduce the non-collision background introduced by beam interactions with the gas or the pipe wall, for which no charged particle tracks are expected.
- $\Delta\phi_{\mathrm{min}}(E_{\mathrm{T}}^{\mathrm{miss}}$, small-$R$ jets) $> 20°$. The minimal azimuthal angle between $E_{\mathrm{T}}^{\mathrm{miss}}$ and small-R jets can be used to veto multijet background where one jet is significantly mismeasured, resulting in large fake $E_{\mathrm{T}}^{\mathrm{miss}}$. Under such conditions, the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector will be highly collinear with the mismeasured jet, while for signals the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ and small-R jets are expected to be well-isolated. Additionally, in events with no small-R jets present, the default value of π is assigned to this variable.
- $\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}$, track-based $E_{\mathrm{T}}^{\mathrm{miss}}) < 90°$. Assuming all jets in the event are well-measured, $E_{\mathrm{T}}^{\mathrm{miss}}$ and track-based $E_{\mathrm{T}}^{\mathrm{miss}}$ should be aligned as they both represent the energy flow of the undetected final state particles. However, in the case of a multijet event with a mismeasured jet, the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector tends to point closer to the direction of that jet, while the track-based $E_{\mathrm{T}}^{\mathrm{miss}}$ remains mostly unaffected, thus no directional preference can be observed.
- $\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{V_{\mathrm{hadronic}}}) > 120°$: Given the signal topology where a pair of DM particles recoil against the vector boson, clear angular separation between the reconstructed $p_{\mathrm{T}}^{V_{\mathrm{hadronic}}}$ (calculated as either the leading large-R jet p_{T} or the p_{T} sum of the leading two small-R jets) could be expected. This cut effectively rejects multijet events and moreover, is able to reduce backgrounds with real $E_{\mathrm{T}}^{\mathrm{miss}}$ originating from $W \to l\nu$ decay, one of the major sources of background contamination in the SRs.

The exact selection values are chosen by examining the distribution of these observables and removing regions in which poor MC modeling of data is observed.

5.5.3 Signal region selection

Following the baseline selection and the anti-QCD requirements, events are categorized into either the boosted or the resolved regime, while the SR selections are optimized for each kinematic topology.

In events containing highly boosted W/Z boson, the decay products may be reconstructed as one large-R jet. Therefore, at least one large-R jet is required in the boosted regime selection, while its mass can be used to distinguish between jets originating from light particles, e.g. gluon and light-flavor quark, and heavy particles, e.g. top quark, W and Z. Figure 5.5.1 shows the expected mass distributions of the leading large-R jet for vector mediator models in the boosted regime. Obvious peaks at around W/Z mass can be seen for both signals, while more continuous spectra are expected for the reducible backgrounds.

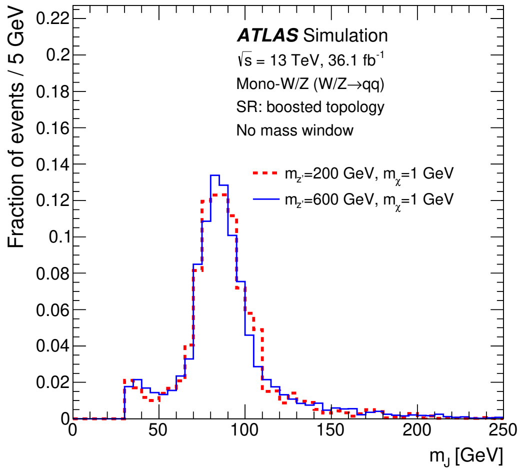


Figure 5.5.1: Expected invariant mass distributions of the leading large-R jet, normalized to unit area, for the vector mediator simplified model with $m_{\text{med}} = 200$ GeV (dashed red) and 600 GeV (solid blue) in the boosted regime [1].

The primary tool used to identify large-R jets from hadronically decaying W and Z bosons is the boosted boson tagger [160, 161], which exploits the characteristics of the jet constituents using the unique radiation pattern of boson jets. A large variety of physical observables have been studied in Run 1 [162] and Run 2 [160] in order to determine the optimal set of input variables which provides the best tagging performance.

A two-variable cut-based W/Z tagger is considered in this analysis, which takes the combined mass of the large-R jet m_J and a substructure variable $D_2^{(\beta=1)}$ [163] as input. The definition of $D_2^{(\beta=1)}$ is given by:

$$D_2^{(\beta)} \equiv \frac{e_3^{(\beta)}}{(e_2^{(\beta)})^3}, \tag{5.5.1}$$

$$e_2^{(\beta)} = \frac{1}{(p_{\mathrm{T}}^J)^2} \sum_{i<j\in J} p_{\mathrm{T}}^i p_{\mathrm{T}}^j (R_{ij})^\beta, \tag{5.5.2}$$

$$e_3^{(\beta)} = \frac{1}{(p_{\mathrm{T}}^J)^3} \sum_{i<j<k\in J} p_{\mathrm{T}}^i p_{\mathrm{T}}^j p_{\mathrm{T}}^k (R_{ij}R_{jk}R_{ik})^\beta. \tag{5.5.3}$$

$e_2^{(\beta)}$ and $e_3^{(\beta)}$ are the two- and three-point jet energy correlation functions, hence $D_2^{(\beta=1)}$ is often referred to as the energy correlation ratio. $D_2^{(\beta=1)}$ separates between the two-pronged decaying W/Z jets and the one-pronged QCD jets, serving as a more sophisticated version of the subject multiplicity. Small values of $D_2^{(\beta=1)}$ are expected for boson jets, while the QCD backgrounds generally have large $D_2^{(\beta=1)}$. A p_{T} dependent upper cut on the $D_2^{(\beta=1)}$ variable is implemented in the tagger, providing fixed boson tagging efficiency over the full kinematic range, as illustrated in Figure 5.5.2. In the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search, a logical 'OR' of the W and Z tagger is used for signal event selection with the 50% fixed-cut working point.

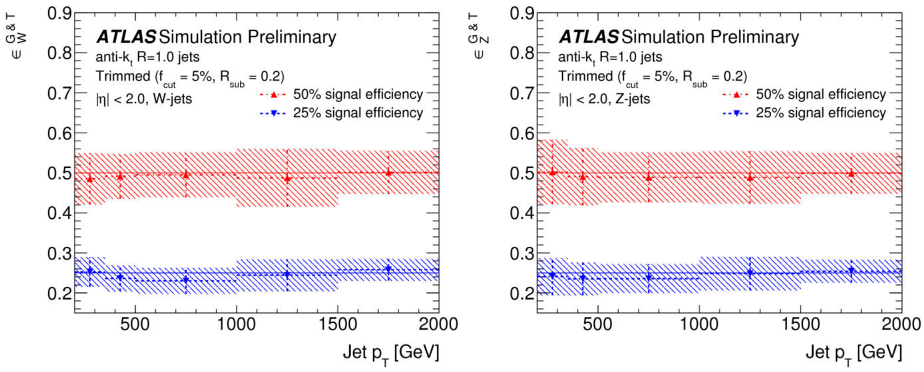


Figure 5.5.2: Signal efficiency versus p_{T} for large-R jets originating from (a) W bosons and (b) Z bosons with 25% (blue) and 50% (red) fixed-cut working point [160]. The combined systematic and statistical uncertainties are shown in shades.

As mentioned in Section 5.3.1, track jets matched to the leading large-R jet are used to define the b-tag categories in the boosted regime. Moreover, events with non-associated b-tagged track jets are vetoed to suppress the V+jets and $t\bar{t}$ background containing heavy-flavor jets. Different selection criteria on the invariant mass and $D_2^{(\beta=1)}$ variable are applied to events with different b-jet multiplicities:

- **0 *b*-tag category**. All events need to pass the p_{T} dependent m_J selection of the W/Z tagger. In addition to this, events fulfilling the $D_2^{(\beta=1)}$ requirement are categorized into the high purity SR, while those failed are recycled in the low purity SR.

- **1 *b*-tag category**. Same selection as 0 b-tag category applies, including the tagger mass cut and the $D_2^{(\beta=1)}$ requirement which divides events into the high purity and the low purity SRs.

- **2 *b*-tag category**. Signal events in this category come predominantly from $Z \to bb$ decays, hence the W/Z tagging is replaced by a mass window requirement of 70 GeV $< m_J <$ 100 GeV, specially optimized for $E_{\mathrm{T}}^{\mathrm{miss}} + Z$ final state. No low purity SR is defined in this category.

The resolved regime selection is designed for events containing well-separated small-R jets as the decay products of W/Z boson. At least two central jets are required under this scenario, with the leading jet $p_{\mathrm{T}} >$ 45 GeV. Events with $\Delta\phi(j_1, j_2) \geq 140°$ are vetoed in order to further suppress the multijet contribution, in which jets recoil against each other, forming the 'back-to-back' topology. To improve the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency modeling in MC simulation, the scalar sum of the transverse momenta of the leading two (three) jets is required to be above 120 (150) GeV for events with two (at least three) central jets.

The invariant mass of the leading two jets, denoted by m_{jj}, provides very good discrimination power between signal and background processes. Figure 5.5.3 shows the expected m_{jj} distributions for vector mediator models in the resolved regime. Obvious peaks at around W/Z mass can be seen for both signals, while more continuous spectra are expected for the reducible backgrounds.

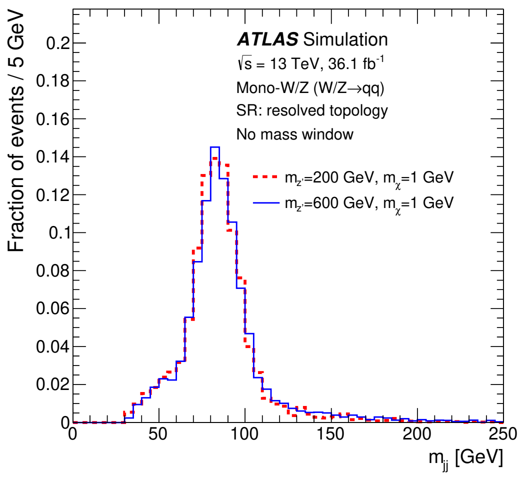


Figure 5.5.3: Expected invariant mass distributions of the leading two central jets, normalized to unit area, for the vector mediator simplified models with $m_{\text{med}} = 200$ GeV (dashed red) and 600 GeV (solid blue) in the resolved regime [1].

As mentioned in Section 5.3.1, central jets are used to defined the three b-tag categories in the resolved regime, with the following selection criteria applied to each category:

- **0 *b*-tag category**. 65 GeV $< m_{jj} <$ 105 GeV, $\Delta R(j_1, j_2) < 1.4$.
- **1 *b*-tag category**. 65 GeV $< m_{jj} <$ 105 GeV, $\Delta R(j_1, j_2) < 1.25$.
- **2 *b*-tag category**. 65 GeV $< m_{jj} <$ 100 GeV, $\Delta R(j_1, j_2) < 1.25$.

The full list of the signal region event selections used in the $E_{\text{T}}^{\text{miss}} + V$ (hadronic) search can be found in Table 5.5.1, while columns from left to right correspond to the selection criteria of the low purity boosted SRs, the high purity boosted SRs and the high purity resolved SRs.

5.6 Background estimation

The major background processes of the $E_{\text{T}}^{\text{miss}} + V$ (hadronic) search are Z+jets, W+jets and $t\bar{t}$, accounting for more than 90% of the total background in the SRs. In order to further constrain these components, control regions (CRs) are designed to estimate various background contributions by extracting their normalization using high purity data sample in regions enriched in the corresponding processes. Lepton multiplicity is used to distinguish the CRs: the 1 muon CRs are defined for W+jet and $t\bar{t}$, whereas the 2 lepton CRs are used to constrain the Z+jets background.

<table>
<tr><th colspan="2">Boosted regime</th><th>Resolved regime</th></tr>
<tr><td colspan="3">baseline selection</td></tr>
<tr><td colspan="3">0 loose electron/muon</td></tr>
<tr><td colspan="3">anti-QCD requirements</td></tr>
<tr><td colspan="2">$E_{\mathrm{T}}^{\mathrm{miss}} > 250$ GeV</td><td>$E_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV</td></tr>
<tr><td colspan="2">≥ 1 large-R jet</td><td>≥ 2 central jets</td></tr>
<tr><td colspan="2">0 non-associated b-tagged track jet</td><td>$p_{\mathrm{T}}^{j_1} > 45$ GeV</td></tr>
<tr><td colspan="2" rowspan="2">pass W/Z tagger m_J cut for 0, 1 b-tag
75 GeV $< m_J <$ 100 GeV for 2 b-tag</td><td>$\Delta\phi(j_1, j_2) < 140°$</td></tr>
<tr><td>$\sum_{i=1}^{2(3)} p_{\mathrm{T}}^{j_i} > 120(150)$ GeV</td></tr>
<tr><td rowspan="2">fail W/Z tagger $D_2^{(\beta=1)}$
cut AND 0, 1 b-tag</td><td rowspan="2">pass W/Z tagger $D_2^{(\beta=1)}$
cut OR 2 b-tag</td><td>$\Delta R(j_1, j_2) < 1.4(1.25)$ for 0, 1 (2) b-tag</td></tr>
<tr><td>65 GeV $< m_{jj} <$ 105 GeV for 0, 1 (2) b-tag</td></tr>
</table>

Table 5.5.1: List of the signal region event selections used in the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search.

The baseline selection is applied to the CRs, ensuring the quality of analyzed data and MC simulated events. Similar to the selection of signal events, all the CRs are split into two kinematic regimes, boosted and resolved, based on the value of $E_{\mathrm{T,\,no}\,\mu}^{\mathrm{miss}}$ and $E_{\mathrm{T,\,no\,lepton}}^{\mathrm{miss}}$. As discussed in Section 5.4, $E_{\mathrm{T,\,no}\,\mu}^{\mathrm{miss}}$ ($E_{\mathrm{T,\,no\,lepton}}^{\mathrm{miss}}$) is constructed by subtracting the muon (lepton) components from $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation, which mimics the kinematic behavior of $E_{\mathrm{T}}^{\mathrm{miss}}$ in the leptonic final states. Events with $E_{\mathrm{T,\,no}\,\mu}^{\mathrm{miss}}$ ($E_{\mathrm{T,\,no\,lepton}}^{\mathrm{miss}}$) above 250 GeV are preferentially used for the boosted regime selections, while events with 150 GeV $< E_{\mathrm{T,\,no}\,\mu}^{\mathrm{miss}}$ ($E_{\mathrm{T,\,no\,lepton}}^{\mathrm{miss}}$) $<$ 250 GeV are considered in the resolved regime. Three b-tag categories are defined based on the number of associated b-tagged track jets and b-tagged central jets in the boosted and the resolved regimes, respectively. This helps to distinguish the W+jets background and $t\bar{t}$ in the 1 muon CRs, where W+jets is mostly included in 0 b-tag and $t\bar{t}$ dominates events with ≥ 1 b-tag. Moreover, categorization based on b-jet multiplicity allows for better understanding of the Z+jets process in the 2 lepton CRs containing heavy-flavor decays, e.g. $Z \to bb$, $Z \to bc$ and $Z \to bl$, which may not be perfectly described by MC simulation. The full event selections of the 1 muon CRs and the 2 lepton CRs are presented in Section 5.6.1 and Section 5.6.2.

In addition to the CRs, validations regions (VRs) are defined to check the validity of the control-to-signal region extrapolation and to constrain the major backgrounds in 0 lepton final state. The selection criteria for VRs are kept orthogonal to the SR selections, featuring a low signal expectation. This is achieved by selecting events outside of the W/Z mass window, with more details to be found in Section 5.6.3.

The CRs and VRs are used in the combined profile likelihood fit (Section 5.8.1) to extract the normalization factor of the major backgrounds, i.e. Z+jets, W+jets and $t\bar{t}$. The subdominant backgrounds, on the other hand, are set to their predicted value based on MC simulation, with corresponding theory uncertainties assigned. The only background which cannot be estimated by means of MC is the multijet production due to its large hadronic cross section and extremely small selection efficiency, hence an unreasonably large amount of simulated samples are needed so that the multijet prediction are not dominated by statistical fluctuations. As a result, a dedicated data-driven approach, namely the ‘template’ method, is applied to evaluate the multijet contamination in the SRs by using the QCD enriched regions, as documented in Section 5.6.4.

5.6.1 1 muon control regions

The 1 muon control regions are defined to constrain W+jets and $t\bar{t}$ using events containing exactly one tight muon and no additional loose muons. As discussed in Section 5.3.2, most of the W+jets and $t\bar{t}$ events in the SRs include a lost or misidentified lepton coming from W boson decay, which itself is either produced via ISR or originates from the decay of a top quark. By studying the kinematic properties in the 1 muon CRs, where the leptons are well-identified and well-measured, one can gain more insights into the physical nature of these processes and moreover, partially cancel the theoretical and experimental uncertainties introduced due to various simulation issues.

To avoid any bias on modeling in different kinematic phase space, the event topologies in the SRs and the 1 muon CRs are kept as similar as possible, with the only extrapolation applied to the lepton. Despite the fact that the multijet contamination is negligible in the 1 muon CRs, the anti-QCD requirements are included as part of the selection criterion, while all $E_{\mathrm{T}}^{\mathrm{miss}}$ related variables are recalculated with $E_{\mathrm{T,\,no}\,\mu}^{\mathrm{miss}}$, allowing for the usage of the same $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers as the SRs:

- Track-based $E_{\mathrm{T,\,no}\,\mu}^{\mathrm{miss}} > 30$ GeV;
- $\Delta\phi_{\mathrm{min}}(E_{\mathrm{T,\,no}\,\mu}^{\mathrm{miss}}$, small-$R$ jets$) > 20°$;
- $\Delta\phi(E_{\mathrm{T,\,no}\,\mu}^{\mathrm{miss}}$, track-based $E_{\mathrm{T,\,no}\,\mu}^{\mathrm{miss}}) < 90°$;
- $\Delta\phi(E_{\mathrm{T,\,no}\,\mu}^{\mathrm{miss}}, p_{\mathrm{T}}^{V_{\mathrm{hadronic}}}) > 120°$.

The remaining event selections are identical to those in the SRs, except for the W/Z tagger $D_2^{(\beta=1)}$ cut in the boosted regime and the $\Delta R(j_1, j_2)$ cut in the resolved regime, which are dropped in order to increase the event yields. Table 5.6.1 presents the full list of the 1 muon control region event selections used in the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search.

<table>
<tr><th>Boosted regime</th><th>Resolved regime</th></tr>
<tr><td colspan="2">baseline selection</td></tr>
<tr><td colspan="2">1 tight muon, 0 loose muon</td></tr>
<tr><td colspan="2">anti-QCD requirements</td></tr>
<tr><td>$E_{\mathrm{T,\,no}\,\mu}^{\mathrm{miss}} > 250$ GeV</td><td>$E_{\mathrm{T,\,no}\,\mu}^{\mathrm{miss}} > 150$ GeV</td></tr>
<tr><td>≥ 1 large-R jet</td><td>≥ 2 central jets</td></tr>
<tr><td>0 non-associated b-tagged track jet</td><td>$p_{\mathrm{T}}^{j_1} > 45$ GeV</td></tr>
<tr><td rowspan="3">pass W/Z tagger m_J cut for 0, 1 b-tag
75 GeV $< m_J <$ 100 GeV for 2 b-tag</td><td>$\Delta\phi(j_1, j_2) < 140°$</td></tr>
<tr><td>$\sum_{i=1}^{2(3)} p_{\mathrm{T}}^{j_i} > 120(150)$ GeV</td></tr>
<tr><td>65 GeV $< m_{jj} <$ 105(100) GeV for 0, 1 (2) b-tag</td></tr>
</table>

Table 5.6.1: List of the 1 muon control region event selections used in the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search.

5.6.2 2 lepton control regions

The 2 lepton control regions are defined to constrain $Z(\to \nu\nu)$+jets using $Z(\to ll)$+jets events containing exactly two loose electrons or muons, at least one of which fulfills the medium criterion. Given that the hard-scattering processes of these two final states are expected to be identical, one can extract the $Z(\to \nu\nu)$+jets normalization from $Z(\to$

ll)+jets measurements. Furthermore, as the transverse momentum of Z does not depend on its decay mode, p_T^{ll} serves as a good emulation of the missing transverse energy of the Z+jets events in the SRs, which is approximately equivalent to $p_T^{\nu\nu}$.

The event topologies in the SRs and the 2 lepton CRs are mostly identical, with the only extrapolation applied to the dilepton system, i.e. the Z boson. Similar to the 1 muon CRs, the anti-QCD requirements are applied to the 2 lepton CRs, while all E_T^{miss} related variables are recalculated with $E_{T,\text{ no lepton}}^{miss}$, resembling the corresponding kinematic properties in the SRs:

- Track-based $E_{T,\text{ no lepton}}^{miss} > 30$ GeV;
- $\Delta\phi_{min}(E_{T,\text{ no lepton}}^{miss}$, small-$R$ jets) $> 20°$;
- $\Delta\phi(E_{T,\text{ no lepton}}^{miss}$, track-based $E_{T,\text{ no lepton}}^{miss}) < 90°$;
- $\Delta\phi(E_{T,\text{ no lepton}}^{miss}, p_T^{V_{\text{hadronic}}}) > 120°$.

On top of the anti-QCD requirements, events failing 66 GeV $< m_{ll} <$ 116 GeV are vetoed to effectively reduce background processes with non-resonant lepton pair, e.g. $t\bar{t}$, single top and WW. The remaining event selections are identical to those in the SRs, except for the W/Z tagger $D_2^{(\beta=1)}$ cut in the boosted regime and the $\Delta R(j_1, j_2)$ cut in the resolved regime, which are dropped in order to increase the event yields. Table 5.6.2 presents the full list of the 2 lepton control region event selections used in the $E_T^{miss} + V$ (hadronic) search.

<table>
<tr><th>Boosted regime</th><th>Resolved regime</th></tr>
<tr><td colspan="2">baseline selection</td></tr>
<tr><td colspan="2">2 loose electrons/muons, among which ≥ 1 medium electron/muon</td></tr>
<tr><td colspan="2">anti-QCD requirements</td></tr>
<tr><td colspan="2">66 GeV $< m_{ll} <$ 116 GeV</td></tr>
<tr><td>$E_{T,\text{ no lepton}}^{miss} > 250$ GeV</td><td>$E_{T,\text{ no lepton}}^{miss} > 150$ GeV</td></tr>
<tr><td>≥ 1 large-R jet</td><td>≥ 2 central jets</td></tr>
<tr><td>0 non-associated b-tagged track jet</td><td>$p_T^{j_1} > 45$ GeV</td></tr>
<tr><td rowspan="3">pass W/Z tagger m_J cut for 0, 1 b-tag
75 GeV $< m_J <$ 100 GeV for 2 b-tag</td><td>$\Delta\phi(j_1, j_2) < 140°$</td></tr>
<tr><td>$\sum_{i=1}^{2(3)} p_T^{j_i} > 120(150)$ GeV</td></tr>
<tr><td>65 GeV $< m_{jj} <$ 105(100) GeV for 0, 1 (2) b-tag</td></tr>
</table>

Table 5.6.2: List of the 2 lepton control region event selections used in the $E_T^{miss} + V$ (hadronic) search.

5.6.3 Validation regions

The extrapolation of backgrounds from control to signal regions are validated and additionally constrained using validation regions defined in 0 lepton final state. These VRs, also known as the mass sideband regions, consist of events passing all signal selections but the W/Z mass cut and are expected to have a similar background composition to that in the SRs. Note that only the upper sideband above the mass window and below 250 GeV is taken into consideration, as it is very difficult to simultaneously model both the high and the low mass region with rich event yields.

Following the 0 lepton VRs, mass sidebands are introduced for 1 muon and 2 lepton final states as well, providing knowledge about the mass spectra of the major backgrounds. Moreover, the inclusion of mass sideband regions ensures the complementarity between the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search and other analyses featuring $E_{\mathrm{T}}^{\mathrm{miss}}$+jet(s) signatures, which might possibly show up in both the SRs and the sideband regions.

Table 5.6.3 presents the full list of the 0 lepton validation region event selections used in the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search. The event selections of the sideband regions in 1 muon and 2 lepton final states are in similar forms.

<table>
<tr><th colspan="2">Boosted regime</th><th>Resolved regime</th></tr>
<tr><td colspan="3">baseline selection</td></tr>
<tr><td colspan="3">0 loose electron/muon</td></tr>
<tr><td colspan="3">anti-QCD requirements</td></tr>
<tr><td colspan="2">$E_{\mathrm{T}}^{\mathrm{miss}} > 250$ GeV</td><td>$E_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV</td></tr>
<tr><td colspan="2">≥ 1 large-R jet</td><td>≥ 2 central jets</td></tr>
<tr><td colspan="2">0 non-associated b-tagged track jet</td><td>$p_{\mathrm{T}}^{j_1} > 45$ GeV</td></tr>
<tr><td colspan="2">W/Z tagger mass window $< m_J <$ 250 GeV for 0, 1 b-tag</td><td>$\Delta\phi(j_1, j_2) < 140°$</td></tr>
<tr><td colspan="2">100 GeV $< m_J <$ 250 GeV for 2 b-tag</td><td>$\sum_{i=1}^{2(3)} p_{\mathrm{T}}^{j_i} > 120(150)$ GeV</td></tr>
<tr><td rowspan="2">fail W/Z tagger $D_2^{(\beta=1)}$ cut AND 0, 1 b-tag</td><td rowspan="2">pass W/Z tagger $D_2^{(\beta=1)}$ cut OR 2 b-tag</td><td>$\Delta R(j_1, j_2) < 1.4(1.25)$ for 0, 1 (2) b-tag</td></tr>
<tr><td>105(100) GeV $< m_{jj} <$ 250 GeV for 0, 1 (2) b-tag</td></tr>
</table>

Table 5.6.3: List of the 0 lepton validation region event selections used in the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search.

5.6.4 Multijet estimation

After the anti-QCD requirements, the contribution of multijet process is expected to be subdominant in the SRs compared to other backgrounds, while in the 1 muon and the 2 lepton CRs it can be assumed as negligible. Hence, the multijet estimation is only performed in 0 lepton final state using a data-driven template method, which consists of two steps:

- **Template generation**. In this step, the multijet templates are constructed using the QCD enriched regions, which are designed to have kinematically similar event selections to the SRs but dominated by multijet events. This is achieved by inverting the dominant anti-QCD cut $\Delta\phi_{\mathrm{min}}(E_{\mathrm{T}}^{\mathrm{miss}}$, small-$R$ jets) $> 20°$. The expected multijet contribution can be then obtained by subtracting the prediction of all non-QCD backgrounds from the observed data in the QCD enriched regions. For each analysis region in 0 lepton final state, a $E_{\mathrm{T}}^{\mathrm{miss}}$ shape template for multijet background is generated, including the low purity SRs and the mass sideband VRs.
- **Template normalization**. Following the template generation, the normalization factor of the shape templates are determined via a profile likelihood fit to data. As the data events corresponding to the residual multijet contribution in the SRs are too poor in numbers to provide meaningful fit results, one might consider to relax or drop some of the anti-QCD requirements in the signal event selections, resulting in the so-called 'relaxed signal regions'. In practice, aside from the inverted $\Delta\phi_{\mathrm{min}}(E_{\mathrm{T}}^{\mathrm{miss}}$, small-$R$ jets) cut, two of the remaining SR selections are

modified to form the relaxed SRs. The $E_{\mathrm{T}}^{\mathrm{miss}}$ threshold in the boosted regime is relaxed to 150 GeV, while for events with 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 250 GeV the boosted and the resolved selection criteria are allowed to overlap, preventing the migration effects introduced by the priority-boosted analysis strategy. Meanwhile, the $\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}$, track-based $E_{\mathrm{T}}^{\mathrm{miss}}) < 90°$ requirement is dropped to further increase the contribution of multijet events. However, strong correlation between $\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}$, $p_{\mathrm{T}}^{V_{\mathrm{hadronic}}})$ and $\Delta\phi_{\mathrm{min}}(E_{\mathrm{T}}^{\mathrm{miss}}$, small-$R$ jets) has been observed, which forbids the removal of the $\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}$, $p_{\mathrm{T}}^{V_{\mathrm{hadronic}}}) > 120°$ requirement. This is mainly because that the normalization of multijet templates can be interpreted as the efficiency of the inverted $\Delta\phi_{\mathrm{min}}(E_{\mathrm{T}}^{\mathrm{miss}}$, small-$R$ jets) cut, while adjusting selections on any correlated variables naturally leads to bias on the efficiency measurement.

The final multijet estimation is acquired by normalizing the shape templates using the scale factors obtained in the combined fit. An uncertainty of 100% is assigned to the multijet normalization. Since the contribution of multijet is proven to be small in the SRs, this relatively large uncertainty does not harm the sensitivity of the analysis.

5.7 Systematic uncertainties

Systematic uncertainties arise from biases in the experimental measurements as well as Monte Carlo modeling of the physical processes, including the SM backgrounds, the BSM signals and the interaction between particles and the detector materials. These uncertainties can influence both the overall yield and the shape of the key observable, e.g. $E_{\mathrm{T}}^{\mathrm{miss}}$, which are often referred to as the 'normalization uncertainties' and the 'shape uncertainties', respectively.

A general description of the systematic uncertainties evaluated in the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search is given in this section, with the experimental systematic uncertainties which originate from the reconstruction, identification and calibration of the physical objects being discussed in Section 5.7.1 and the theoretical systematic uncertainties which correspond to theory predictions and MC modelings being discussed in Section 5.7.2.

5.7.1 Experimental systematic uncertainties

The total integrated luminosity is used to normalize the MC prediction to the measured data. In the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search, 36.1 fb^{-1} of pp collision data is analyzed, which corresponds to an uncertainty of 2.1% on the luminosity measurement during 2015 and 2016 data-taking. Additionally, the pile-up conditions of the MC simulated events are reweighted to match those in data, while the pile-up reweighting uncertainty is introduced, accounting for the possible biases in the modeling of pile-up events. The total pile-up uncertainty is at the order of 6%, based on measurements performed using Run 1 data [164].

Aside from the luminosity uncertainty and the pile-up reweighting uncertainty, experimental systematic uncertainties are evaluated for various reconstructed objects using standard tools provided by the ATLAS Combined Performance (CP) groups.

As mentioned in Section 5.2.3, $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers and single-lepton triggers are used in the analysis, while corrections and uncertainties accounting for trigger inefficiencies have to be taken into consideration. The uncertainty on $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger calibration is estimated in two terms: the systematic uncertainty is taken as the largest variation between the scale factors derived using events with different flavor compositions, while the statistical uncertainty is evaluated as the 1σ interval of the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger calibration function fit. The single-lepton trigger uncertainties are derived in a similar manner, including one term accounting for single-electron trigger efficiency factors and two terms accounting for single-muon trigger efficiency factors.

Electrons are used in both the 0 lepton SRs for event veto and the 2 lepton CRs for Z boson reconstruction. The reconstruction, identification and isolation criteria of electrons naturally affect the event yields in various analysis regions, while the uncertainties on electron p_{T} measurements affect the shape of the $E_{\mathrm{T,\,no\,lepton}}^{\mathrm{miss}}$ variable. Experimental uncertainties on electron reconstruction, identification and isolation efficiencies are therefore taken into account, as well as uncertainties on the energy scale and resolution.

Similar to electrons, muons are involved in various SR and CR selections, while the corresponding uncertainties are introduced to account for the efficiency scale factors of muon reconstruction (track-to-vertex association), identification and isolation, as well as measurements of muon p_{T}, including uncertainties on p_{T} scale, p_{T} resolution in the Muon Spectrometer and the Inner Detector.

The full chain of calibration is applied to the small-R jets used in the analysis, consisting of origin correction, pile-up correction, JES and η calibration, GSC and in-situ calibration, which accounts for various detector effects and aims to correct the kinematic properties of jets to truth-level. Multiple sources of systematic uncertainty are evaluated for each procedure, resulting in more than 90 components in total. As these JES related uncertainties are expected to be subdominant in this analysis, a strongly reduced uncertainty set containing four terms only is applied in the statistic model, as mention in Section 3.3.3.1. Apart from the JES uncertainties, jet energy resolution (JER) uncertainty and flavor tagging uncertainties are also taken into consideration, with latter serving as the dominant small-R jet uncertainties. A total of five groups of flavor tagging uncertainties are defined, corresponding to the efficiency scale factors of light-flavor, c-tagged and b-tagged jets, as well as the extrapolation of c- and b-tagging efficiencies to high p_{T} regimes.

For large-R jets, uncertainties on energy, mass and $D_2^{(\beta=1)}$ resolution are considered. In addition to these, three groups of systematic uncertainties, corresponding to large-R jet p_{T}, mass and $D_2^{(\beta=1)}$ modeling, are derived using experimental approaches, with each group consisting of four terms, which account for the difference between data and MC simulation, the fragmentation modeling, the tracking reconstruction efficiency, fake rate and bias in the q/p_{T} distribution, and the total statistical uncertainty.

Track jets are used in the boosted regime to provide b-tagging information, hence the uncertainties on flavor tagging have to be accounted. Similar to small-R jets, five groups of flavor tagging uncertainties are derived for track jets, which stands for the efficiency scale factors of light-flavor, c-tagged and b-tagged jets, as well as the extrapolation of c- and b-tagging efficiencies to high p_{T} regimes.

The systematic uncertainties on $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction can be categorized into two components, which originate from the hard term and the soft term, respectively. The uncertainties introduced by the hard term can be evaluated by propagating the uncertainties on the input objects to $E_{\mathrm{T}}^{\mathrm{miss}}$ computation, while the soft term related uncertainties are taken from experimental measurements. Four $E_{\mathrm{T}}^{\mathrm{miss}}$ uncertainties are considered, including two terms corresponding to the soft term resolution in the direction perpendicular and parallel to the hadronic recoil system, and two terms standing for the scale uncertainty of the soft term and the jet terms.

5.7.2 Theoretical systematic uncertainties

Theoretical systematic uncertainties account for various MC simulation parametrization, e.g. matrix elements, renormalization and factorization scales, ISR and FSR, by means of normalization and shape uncertainties.

The shape uncertainties, derived in terms of p_{T}^{V} and m_{jj}, are estimated by comparing the kinematic distributions with scales normalized to unity area between different samples. For W+jets and Z+jets processes, the modeling uncertainties are obtained using the following samples:

- Sherpa 2.2.1 (default) versus Sherpa 2.1;
- Sherpa 2.2.1 (default) versus MadGraph5 + Pythia;
- Sherpa 2.2.1 (default) versus data in W/Z+jets enrich region.

The largest variation with respect to the default is taken and fitted with an analytical function, which is then symmetrized to provide the $\pm 1\sigma$ uncertainty of the shape distribution. Furthermore, the p_{T}^{V} and m_{jj} uncertainties are derived individually for V+light, $V + cl$, and $V + bb$ processes. This is motivated by the fact that different contributing diagrams dominate events with different flavor compositions, e.g. the 2 b-tag region is dominated by the gluon splitting process $g \to bb$, thus resulting in different event topologies.

Similarly, the p_{T}^{V} and m_{jj} modeling uncertainties for $t\bar{t}$ are estimated via shape comparisons between:

- Powheg + Pythia 8 (default) versus Powheg + Pythia 8 with increased and decreased radiation;
- Powheg + Pythia 8 (default) versus Powheg + Herwig 7;
- Powheg + Pythia 8 (default) versus MadGraph5_aMC@NLO [165] + Pythia 8.

The single top shape uncertainties are obtained by comparing:

- Powheg + Pythia 6 (default) versus Powheg + Pythia 6 with increased and decreased radiation;
- Powheg + Pythia 6 (default) versus Powheg + Herwig++ [166];
- Powheg + Herwig++ versus MadGraph5_aMC@NLO + Herwig++.

And for diboson:

- Powheg + Pythia 8 (default) versus Powheg + Herwig++;
- Powheg + Pythia 8 (default) versus Sherpa 2.1.

For all these backgrounds, the largest variations are symmetrized and taken as the $\pm 1\sigma$ envelopes of the shape uncertainties.

The normalization uncertainties, on the other hand, are implemented in terms of global normalization parameters in the combined fit. While some of the normalization parameters are allowed to float freely, such that they can be constrained purely from the fit and without any prior knowledge, other normalization parameters are assigned with uncertainties extracted from the relative differences in the event yields predicted by the above mentioned variations. Table 5.7.1 shows the background normalization parameters used in the combined fit with assigned uncertainties. $V + hf$ denotes the combination of $V + bb$, $V + bc$, $V + bl$ and $V + cc$.

Process	**Uncertainty [%]**
$V + bc/V + hf$ ratio	30
$V + bl/V + hf$ ratio	30
$V + cc/V + hf$ ratio	30
$V + cl$	30
$V + l$ (resolved)	20
$t\bar{t}$ (resolved)	30
Single top t-channel	4.4
Single top s-channel	4.6
Single top Wt-channel	6.2
WW	25
WZ	26
ZZ	20

Table 5.7.1: Summary of the background theoretical systematic normalization uncertainties included in the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search.

Theoretical uncertainties are estimated for DM signals by modifying parameters in the MadGraph5 and the Pythia generators and comparing the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution at truth-level. For each $E_{\mathrm{T}}^{\mathrm{miss}}$ bin included in the fit, a relative uncertainty is calculated by symmetrizing the up and down variations. In the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search, the theoretical uncertainties for signals consist of three components:

- **PDF uncertainty**. This is estimated by replacing the default NNPDF23LO PDF set with alternative PDF sets, namely MSTW2008LO PDF [167] and CTEQ6L1 PDF [168]. The $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution for vector mediator simplified model signals applying different PDF sets are compared and the largest variation relative to the nominal is taken as the PDF uncertainty.

- **Scale uncertainty**. This is estimated by varying the renormalization and factorization scales in MadGraph5. The scales are changed coherently up and down by a factor of 2 on an event-by-event basis. The $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution for vector mediator simplified model signals applying different renormalization and factorization scales are compared and the largest variation relative to the nominal is taken as the scale uncertainty.

- **Tune uncertainty**. This is estimated by adjusting the amount of initial state radiation, final state radiation and multi-parton interactions. The tune uncertainty covers the impact of underlying events, jet substructure and additional jet production. The $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution for vector mediator simplified model signals applying different A14 tune variations are compared and the largest variation relative to the nominal is taken as the tune uncertainty.

Table 5.7.2 documents the relative PDF, scale and tune uncertainties applied to DM signals in the combined profile likelihood fit in each $E_{\mathrm{T}}^{\mathrm{miss}}$ bin.

$E_{\mathrm{T}}^{\mathrm{miss}}$ bin [GeV]	Uncertainty [%]		
	PDF	Scale	Tune
150 - 250	1.0	1.0	1.5
250 - 350	2.0	2.0	3.0
350 - 500	3.0	2.5	5.0
500 - 800	5.0	3.0	7.0
800 - 1500	6.0	6.0	8.5

Table 5.7.2: Relative PDF, scale and tune uncertainties applied to DM signals in the combined profile likelihood fit in each $E_{\mathrm{T}}^{\mathrm{miss}}$ bin. The numbers are taken from truth-level studies.

5.8 Results

All analysis regions are considered in a statistical model using a combined profile likelihood fit, where the 'background-only' hypothesis is tested against the 'signal-plus-background' hypothesis by assuming the presence of DM signals on top of the SM background processes. A brief introduction to the profile likelihood function, as well as the CL_s method [169] used for limit setting, is given in Section 5.8.1. The impact of various sources of systematic uncertainty is propagated in terms of the nuisance parameters (NPs) in the model and evaluated from a fit to the Asimov dataset [170], as described in Section 5.8.2. Section 5.8.3 shows the results obtained from the fit to the actual data, while the exclusion limits on the vector mediator simplified model are interpolated and presented in the two-dimensional m_χ-m_{med} plane in Section 5.8.4.

5.8.1 Statistical model

A profile likelihood fit [170] is performed on the binned distributions of the discriminating variables in all analysis regions in order to test the existence of DM signals. For each bin included in the fit, the expected bin content can be expressed as:

$$N_i^{\mathrm{exp}}(\mu,\boldsymbol{\theta}) = \mu \cdot S_i(\boldsymbol{\theta}) + B_i(\boldsymbol{\theta}), \tag{5.8.1}$$

where S_i and B_i are the predicted signal and background events in bin i, respectively, μ corresponds to the strength of the signal process, often defined as the ratio of the signal cross section to a reference signal cross section, and $\boldsymbol{\theta}$ represents the nuisance parameters (NPs) accounting for various types of systematic uncertainties, as discussed in Section 5.7. The binned likelihood function $\mathcal{L}$ of N^{obs} observed events given N^{exp} is then constructed as the product of Poisson probability terms, as implemented in the HistFactory package [171]:

$$\begin{aligned}\mathcal{L}(N^{\text{obs}}|\mu,\boldsymbol{\theta}) = &\prod_i \frac{(\mu \cdot S_i(\boldsymbol{\theta}) + B_i(\boldsymbol{\theta}))^{N_i^{\text{obs}}}}{N_i^{\text{obs}}!} e^{\mu \cdot S_i(\boldsymbol{\theta}) + B_i(\boldsymbol{\theta})} \\ &\times \prod_{\theta_i \in \boldsymbol{\theta}} f(\langle \theta_i \rangle, \sigma_{\theta_i}).\end{aligned} \tag{5.8.2}$$

$f(\langle \theta_i \rangle, \sigma_{\theta_i})$ is the probability density function of NP θ_i, which takes the form of Gaussian or Log-normal, dependent on the prior understanding of that systematic uncertainty, with its nominal value and $\pm 1\sigma$ variation denoted by $\langle \theta_i \rangle$ and σ_{θ_i}. Additionally, the normalization of some background processes can be directly constrained in the statistical model without the usage of prior knowledge obtained from auxiliary measurements or modeling studies. In the $E_{\text{T}}^{\text{miss}} + V$ (hadronic) search this applies to:

- W+jets, including $W + bb$, $W + bc$, $W + bl$, $W + cc$ and W+light;
- Z+jets, including $Z + bb$, $Z + bc$, $Z + bl$, $Z + cc$ and Z+light;
- $t\bar{t}$ production.

The fit results are obtained by maximizing the likelihood function $\mathcal{L}(N^{\text{obs}}|\mu,\boldsymbol{\theta})$. Based on the Neyman-Pearson lemma [172], the log-likelihood ratio is applied as the test statistic:

$$q_\mu = -2 \ln \left(\frac{\mathcal{L}(\mu, \hat{\hat{\boldsymbol{\theta}}})}{\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}})} \right), \tag{5.8.3}$$

where $\mathcal{L}(\mu, \hat{\hat{\boldsymbol{\theta}}})$ is maximized over all NPs for a given μ and $\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}})$ is maximized over the full parameter space of $(\mu, \boldsymbol{\theta})$. Hence $\hat{\hat{\boldsymbol{\theta}}}$ is addressed as the 'conditional maximum-likelihood estimator' of $\boldsymbol{\theta}$, while $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ are the 'unconditional maximum-likelihood estimators' of μ and $\boldsymbol{\theta}$, respectively.

The test statistic q_μ can be used to evaluate the validity of hypotheses with the given signal strength μ. For discovery searches where a significant excess over the SM prediction is expected, $\mu = 0$ is chosen to build the null hypothesis (the 'background-only' hypothesis) to be rejected, resulting in:

$$q_0 = \begin{cases} -2 \ln \left(\frac{\mathcal{L}(0, \hat{\hat{\boldsymbol{\theta}}})}{\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}})} \right) & , \text{ if } \quad \hat{\mu} > 0; \\ 0 & , \text{ if } \quad \hat{\mu} \leq 0. \end{cases} \tag{5.8.4}$$

For analyses where no excess of events is found, $\mu > 0$ hypothesis (the 'signal-plus-background' hypothesis) is tested, with q_μ taking the form of:

$$q_\mu = \begin{cases} -2\ln\left(\frac{\mathcal{L}(\mu,\hat{\hat{\boldsymbol{\theta}}})}{\mathcal{L}(0,\hat{\hat{\boldsymbol{\theta}}})}\right) & , \text{ if } \quad \hat{\mu} \leq 0; \\ -2\ln\left(\frac{\mathcal{L}(\mu,\hat{\hat{\boldsymbol{\theta}}})}{\mathcal{L}(\hat{\mu},\hat{\boldsymbol{\theta}})}\right) & , \text{ if } \quad 0 < \hat{\mu} \leq \mu; \\ 0 & , \text{ if } \quad \hat{\mu} > \mu. \end{cases} \tag{5.8.5}$$

Based on the fact that small values of q_μ are expected if the data is in good agreement with the tested hypothesis, the p-value

$$p_\mu = \int_{q_\mu^{\mathrm{obs}}}^{+\infty} f(q_\mu|\mu)\mathrm{d}q_\mu \tag{5.8.6}$$

is defined to quantify the level of incompatibility between the observed data and the hypothesized signal strength μ, with $f(q_\mu|\mu)$ standing for the probability density function of q_μ. To distinguish between the 'background-only' and the 'signal-plus-background' hypotheses, the corresponding p-value for a given q_μ^{obs} can be constructed as:

$$p_b = \int_{-\infty}^{q_\mu^{\mathrm{obs}}} f(q_\mu|b)\mathrm{d}q_\mu, \tag{5.8.7}$$

$$p_{s+b} = \int_{q_\mu^{\mathrm{obs}}}^{+\infty} f(q_\mu|s+b)\mathrm{d}q_\mu, \tag{5.8.8}$$

while the signal model is excluded at 95% confidence level (CL) if $p_{s+b} < 0.05$.

However, if a search is insensitive to the probed signal, e.g. the expected number of signal events is much lower than the background prediction, the probability density function $f(q_\mu|b)$ and $f(q_\mu|s+b)$ will mostly overlap and the signal hypothesis might be rejected when a sufficiently large downward fluctuation in data takes place. To avoid such accidental exclusion of signals, the CL_s probability [169] is introduced:

$$\mathrm{CL}_s = \frac{p_{s+b}}{1-p_b}. \tag{5.8.9}$$

The 95% confidence level upper limit on μ can be calculated as the value of μ corresponding to $\mathrm{CL}_s = 0.05$. As $p_b > 0$, the upper limit calculated using the CL_s method is always more conservative than the limit given by $p_{s+b} < 0.05$.

Table 5.8.1 shows the regions included in the fit and the discriminating observables fitted in each region.

5.8.2 Impact of uncertainties

The Asimov dataset [170] is an artificial dataset built from the predicted distribution of MC backgrounds, with a Poisson error corresponding to the statistical uncertainty of data assumed in each bin. In practice, fit to the Asimov data can be used to validate the

	0 lepton	1 muon	2 lepton
Enriched in	signals	$t\bar{t}$, W+jets	Z+jets
Fitted observables	$E_{\mathrm{T}}^{\mathrm{miss}}$	$E_{\mathrm{T,\,no}\,\mu}^{\mathrm{miss}}$	$E_{\mathrm{T,\,no\,lepton}}^{\mathrm{miss}}$
	5 bins: [150, 250, 350, 500, 800, 1500] GeV		
Kinematic topologies	[150, 250] GeV: resolved		
	above 250 GeV: priority-boosted		
b-tag categories	0, 1 and 2		
m_J/m_{jj} selection	W/Z mass window and upper mass sideband		

Table 5.8.1: Summary of the analysis regions considered in the statistical model of the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search. A combined profile likelihood fit is performed on $E_{\mathrm{T}}^{\mathrm{miss}}$, $E_{\mathrm{T,\,no}\,\mu}^{\mathrm{miss}}$ and $E_{\mathrm{T,\,no\,lepton}}^{\mathrm{miss}}$ in 0 lepton, 1 muon and 2 lepton final states, respectively, with each distribution divided into five bins.

statistical model of the analysis and evaluate the relative impact of different sources of uncertainty.

For a given NP θ_i, its relevance to the sensitivity of the search can be estimated as a fractional uncertainty on the fitted signal strength. This is computed by repeating the fit with the absence of θ_i and subtracting in quadrature the resulting uncertainty on μ from the total uncertainty:

$$\sigma_{\theta_i} = \sqrt{\sigma_{\mathrm{total}}^2 - \sigma_{\mathrm{no}\ \theta_i}^2}. \tag{5.8.10}$$

Table 5.8.2 shows the impact of different uncertainties on σ_{total} for three representative signal models, all with $m_\chi = 1$ GeV and with m_{med} = (a) 200 GeV, (b) 600 GeV and (c) 2000 GeV. Multiple NPs are grouped into categories according to their sources. The total statistical uncertainty is evaluated by neglecting all systematic uncertainties in the fit. The dominant sources of uncertainty include:

- Finite MC events. The number of events for the SM backgrounds, in particular the V+jets production, is strongly limited at $p_{\mathrm{T}} > 0.5$ TeV. The impact of MC statistical uncertainty increases with m_{med}, which correlates to the transverse momentum of the recoiling boson.
- Modeling uncertainties. The modeling of signal and V+jets background have large impact on signals with high m_{med}, also due to the limitations on p_{T}^V modeling at $p_{\mathrm{T}} > 0.5$ TeV.
- Large-R jet reconstruction and calibration. The priority-boosted analysis strategy ensures that most of the sensitivity of the search comes from the boosted regime, where large-R jets are deeply involved in the selection criteria, e.g. identification of W/Z. The impact of large-R jet uncertainties rises with the increasing $E_{\mathrm{T}}^{\mathrm{miss}}$, as the uncertainties generally increase with jet $p_{\mathrm{T}} \sim E_{\mathrm{T}}^{\mathrm{miss}}$.
- Normalization. The major components include Z+jets, diboson and multijet. Z+jets and diboson production both contribute as the irreducible background ($Z \to \nu\nu$) in the SRs, which naturally influences the sensitivity of the analysis. On the other

Source	Uncertainty [%]		
	(a)	(b)	(c)
Luminosity	2.7	3.6	3.7
Pile-up reweighting	2.1	0.7	2.8
Triggers	0.2	1.4	3.0
Electrons	3.8	9.3	12.0
Muons	5.9	7.1	3.9
Small-R jets	3.4	8.3	10.3
Small-R jet b-tagging	1.6	4.1	2.2
Large-R jets	8.5	19.6	18.8
Track jet b-tagging	3.6	3.8	5.9
$E_{\mathrm{T}}^{\mathrm{miss}}$	1.3	4.3	6.5
V+jets modeling	3.8	9.7	13.6
V+jets composition	1.2	2.9	2.9
$t\bar{t}$ modeling	2.2	3.7	3.3
Diboson modeling	0.9	1.9	2.1
Signal modeling	7.1	8.8	10.4
W+jets normalization	2.6	3.5	4.6
Z+jets normalization	4.7	9.0	11.7
$t\bar{t}$ normalization	2.7	1.2	3.2
Diboson normalization	5.1	11.2	12.8
Multijet	7.3	11.3	10.0
MC Stats.	9.8	18.0	24.4
Data Stats.	6.6	20.8	45.3
Total Syst.	20.9	40.1	49.4
Total	21.9	45.2	67.0

Table 5.8.2: Relative impact of different sources of uncertainty on the signal strength for the vector mediator simplified models with $m_\chi = 1$ GeV and $m_{\mathrm{med}} =$ (a) 200 GeV, (b) 600 GeV and (c) 2000 GeV, representing signal event topologies with low, medium and high $E_{\mathrm{T}}^{\mathrm{miss}}$, respectively.

hand, as mentioned in Section 5.6.4, an 100% normalization uncertainty is assigned to the multijet process estimated using the template method. This especially affects region of $E_{\mathrm{T}}^{\mathrm{miss}} < 350$ GeV, while the impact of multijet contamination decreases at high $E_{\mathrm{T}}^{\mathrm{miss}}$.

For all the three signal models, the systematic uncertainties overwhelm the data statistical uncertainty, indicating that the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search is mostly systematically limited.

5.8.3 Observed results

To check the strategy of control region design, the MC predictions are fitted to the observed distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$, $E_{\mathrm{T,\,no}\,\mu}^{\mathrm{miss}}$ and $E_{\mathrm{T,\,no\,lepton}}^{\mathrm{miss}}$ in all analysis regions. Figure 5.8.1 and Figure 5.8.2 show the $E_{\mathrm{T,\,no}\,\mu}^{\mathrm{miss}}$ and the $E_{\mathrm{T,\,no\,lepton}}^{\mathrm{miss}}$ distributions in the 1 muon and the 2 lepton CRs after the combined profile likelihood fit, respectively. The observed results are in good agreement with the SM predictions, thus validates the statistical model of the analysis.

The numbers of observed events entering each signal region are presented in Table 5.8.3 and Table 5.8.4 for the boosted and the resolved regimes. Alongside are the signal predictions for models with different m_{med}, as well as the expected background contributions determined by the combined profile likelihood fit to the full measured data. The normalization of the major background components constrained by the statistical model under the 'background-only' hypothesis is consistent with the SM expectation.

Process	**0 *b*-tag**		**1 *b*-tag**		**2 *b*-tag**
	high purity	low purity	high purity	low purity	
Vector mediator model					
m_χ =1 GeV, m_{med} =200 GeV	814 ± 48	759 ± 45	96 ± 18	99 ± 16	49.5 ± 4.3
m_χ =1 GeV, m_{med} =600 GeV	280.9 ± 9.0	268.5 ± 8.8	34.7 ± 3.6	33.8 ± 3.1	15.38 ± 0.84
W+jets	3170 ± 140	10120 ± 380	218 ± 28	890 ± 110	91 ± 12
Z+jets	4750 ± 200	15590 ± 590	475 ± 52	1640 ± 180	186 ± 12
$t\bar{t}$	775 ± 48	937 ± 60	629 ± 27	702 ± 34	50 ± 11
Single top	159 ± 12	197 ± 13	89.7 ± 6.7	125.5 ± 8.7	16.1 ± 1.7
Diboson	770 ± 110	960 ± 140	88 ± 14	115 ± 18	54 ± 10
Multijet	12 ± 35	49 ± 140	3.7 ± 3.3	15 ± 13	9.3 ± 9.4
Total background	9642 ± 87	27850 ± 150	1502 ± 31	3490 ± 52	407 ± 15
Data	9627	27856	1502	3525	414

Table 5.8.3: Expected and observed numbers of events in the boosted regime with an integrated luminosity of 36.1 fb^{-1} and $\sqrt{s}$ = 13 TeV, shown separately in each signal region [1]. The background yields and uncertainties are shown after the profile likelihood fit to the data (with $\mu = 0$). The quoted background uncertainties include both the statistical and systematic contributions, while the uncertainties in the signals are statistical only.

Figure 5.8.3 and Figure 5.8.4 show the corresponding distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$ in all SRs after the combined profile likelihood fit. No significant excess over the SM prediction is observed.

Process	**0 b-tag**	**1 b-tag**	**2 b-tag**
Vector mediator model			
m_χ =1 GeV, m_{med} =200 GeV	5050 ± 130	342 ± 29	136.7 ± 6.0
m_χ =1 GeV, m_{med} =600 GeV	840 ± 16	59.9 ± 4.6	27.86 ± 0.94
W+jets	117500 ± 4600	5000 ± 680	598 ± 98
Z+jets	135400 ± 5600	7710 ± 780	1219 ± 67
$t\bar{t}$	13800 ± 780	12070 ± 420	2046 ± 70
Single top	2360 ± 140	1148 ± 71	222 ± 14
Diboson	6880 ± 950	514 ± 71	228 ± 34
Multijet	11900 ± 2300	1130 ± 370	290 ± 150
Total background	287770 ± 570	27580 ± 170	4601 ± 90
Data	287722	27586	4642

Table 5.8.4: Expected and observed numbers of events in the resolved regime with an integrated luminosity of 36.1 fb^{-1} and $\sqrt{s}$ = 13 TeV, shown separately in each signal region [1]. The background yields and uncertainties are shown after the profile likelihood fit to the data (with $\mu = 0$). The quoted background uncertainties include both the statistical and systematic contributions, while the uncertainties in the signals are statistical only.

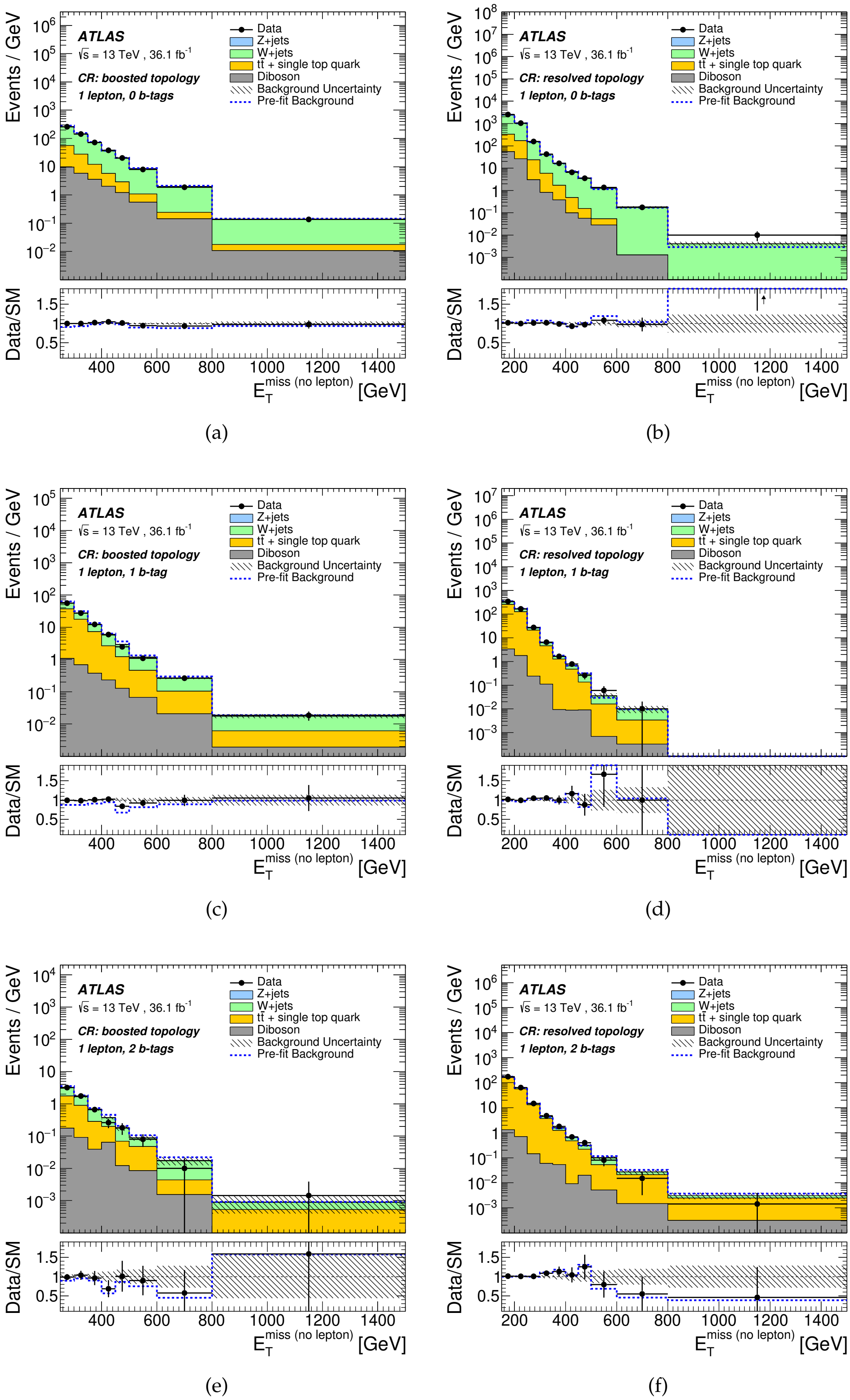


Figure 5.8.1: $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions after the combined profile likelihood fit in the boosted (left) and the resolved (right) 1 muon control regions with 0 b-tag (top), 1 b-tag (middle) and 2 b-tag (bottom) [1].

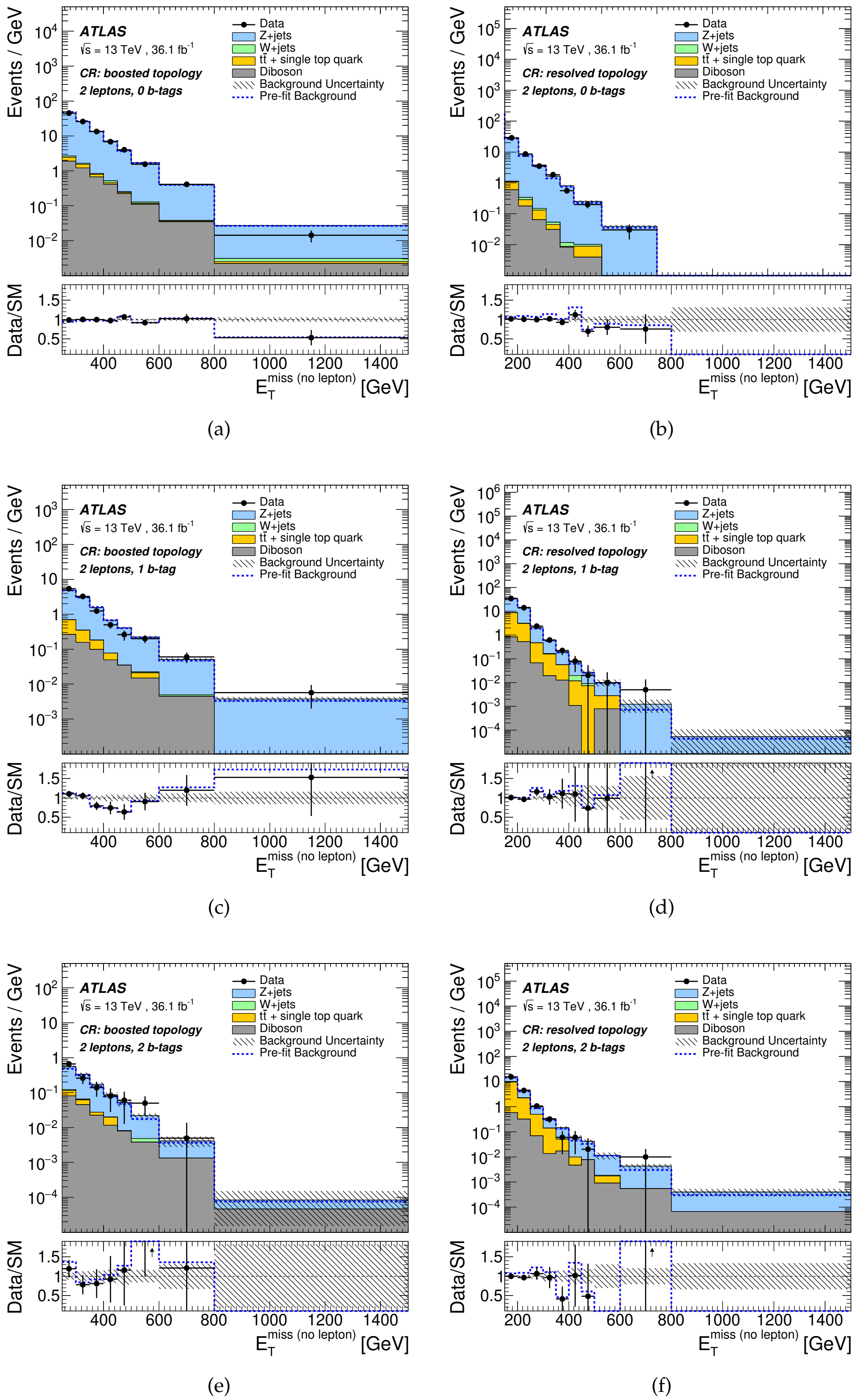
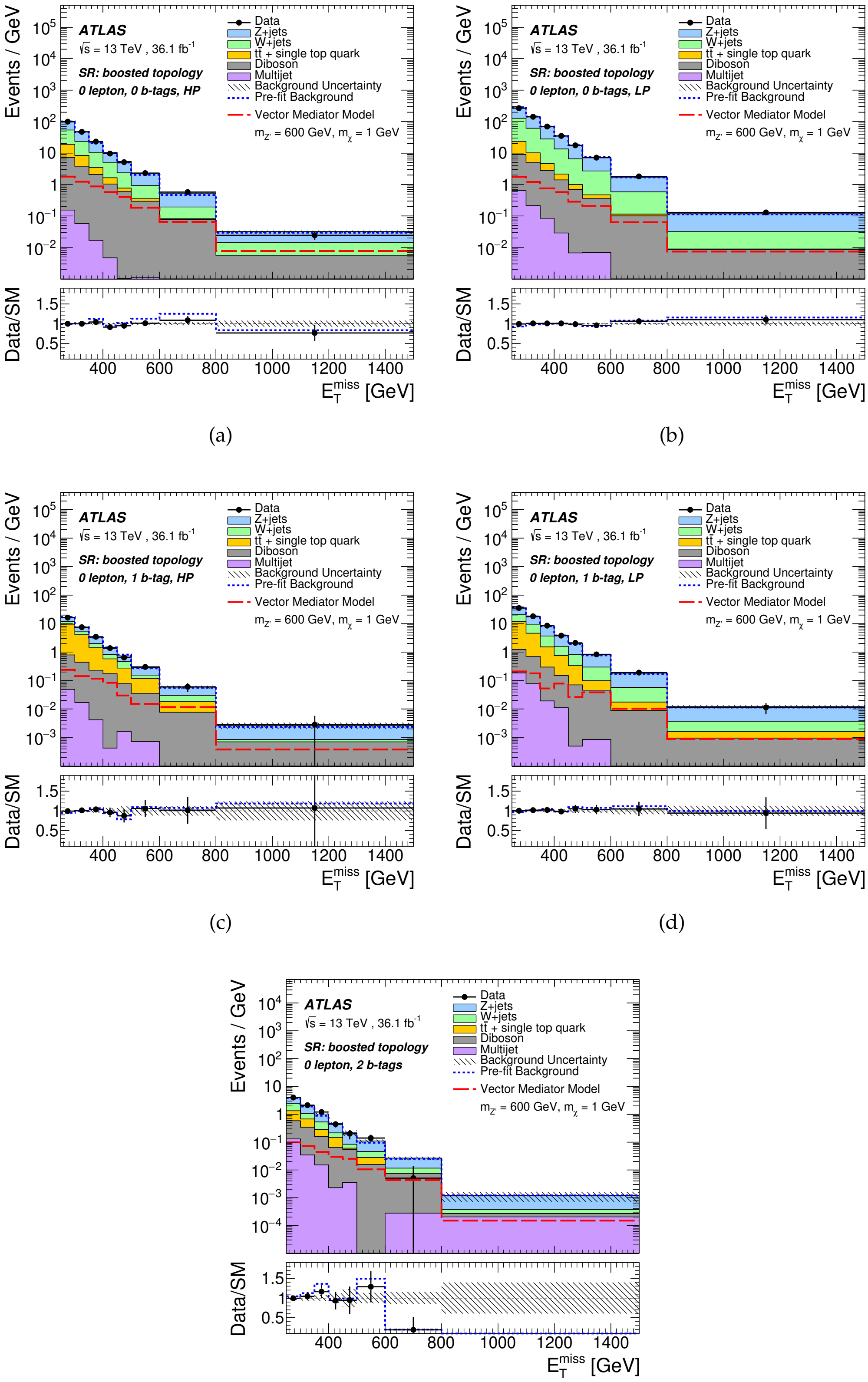
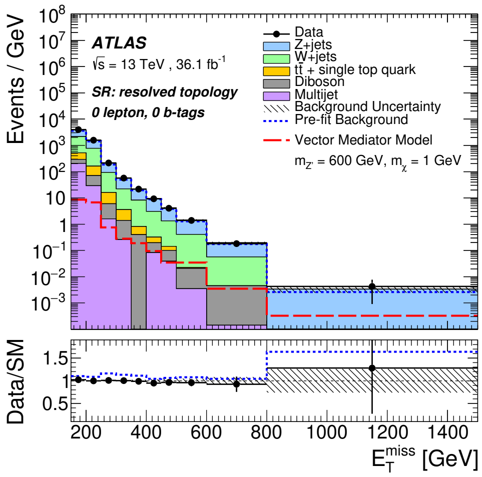


Figure 5.8.2: $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions after the combined profile likelihood fit in the boosted (left) and the resolved (right) 2 lepton control regions with 0 b-tag (top), 1 b-tag (middle) and 2 b-tag (bottom) [1].

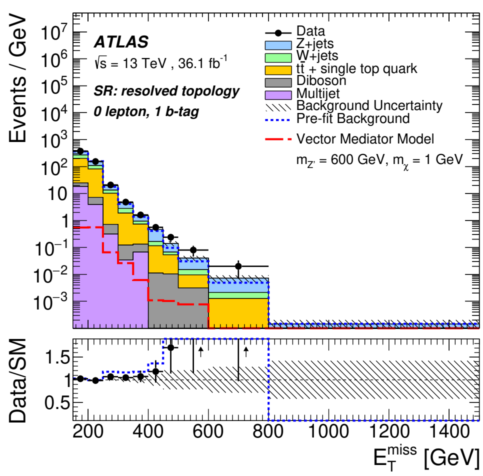


(a) (b) (c) (d) (e)

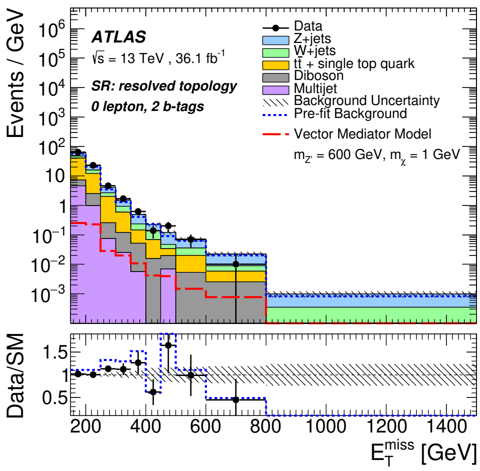
Figure 5.8.3: $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions after the combined profile likelihood fit in the boosted signal regions: (a) 0 b-tag high purity signal region, (b) 0 b-tag low purity signal region, (c) 1 b-tag high purity signal region, (d) 1 b-tag low purity signal region and (e) 2 b-tag high purity signal region [1].



(a)



(b)



(c)

Figure 5.8.4: $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions after the combined profile likelihood fit in the resolved signal regions: (a) 0 b-tag high purity signal region, (b) 1 b-tag high purity signal region and (c) 2 b-tag high purity signal region [1].

5.8.4 Exclusion limits

As no evidence for the production of DM is found in $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) final state, 95% CL$_s$ exclusion limit in the two-dimensional m_χ-m_{med} plane is set in order to constrain the parameter space of the vector mediator simplified model.

However, due to the usage of $D_2^{(\beta=1)}$ variable in the signal event selection, complete detector simulation has to be performed to provide accurate modeling of the jet substructure, which is highly demanding in computational resources. As a result, only a limited set of signal models could be simulated with the full chain of Monte Carlo, while an interpolation procedure is used to estimate the fitted signal strength for ungenerated mass points.

The expected signal yield S at mass point $(m_{\mathrm{med}}, m_\chi)$ is given by:

$$S = \mathcal{L} \cdot (\mathcal{A} \cdot \epsilon)_{\text{total}}^{m_{\text{med}},\, m_\chi} \cdot \sigma_{pp \to Z' \to \chi\chi}^{m_{\text{med}},\, m_\chi}, \tag{5.8.11}$$

where $\mathcal{L}$ is the total integrated luminosity, $(\mathcal{A} \cdot \epsilon)_{\text{total}}^{m_{\text{med}},\, m_\chi}$ represents the detector acceptance multiplied by the selection efficiency of the SRs and $\sigma_{pp \to Z' \to \chi\chi}^{m_{\text{med}},\, m_\chi}$ denotes the theory predicted cross section of $pp \to Z' \to \chi\chi$ process.

The interpolation is based on the assumption that all signal processes with the same mediator mass m_{med} and different m_χ are expected to have similar $(\mathcal{A} \cdot \epsilon)_{\text{total}}$ value to that of the simulated sample with $m_\chi = 1$ GeV. Such approximation is verified to be reliable for the on-shell production of DM particles, as discussed in Section 5.1. Thus, the expected yield of signals with fixed m_{med} only depend on the value of $\sigma_{pp \to Z' \to \chi\chi}^{m_{\text{med}},\, m_\chi}$, which can be expressed in terms of $\sigma_{pp \to Z' \to \chi\chi}^{m_{\text{med}},\, m_\chi = 1\text{ GeV}}$ and the branching ratio $\mathcal{B}_{Z' \to \chi\chi}^{m_\chi}$ under the narrow width approximation:

$$\begin{aligned} \sigma_{pp \to Z' \to \chi\chi}^{m_{\text{med}},\, m_\chi} &= \sigma_{pp \to Z'}^{m_{\text{med}}} \cdot \mathcal{B}_{Z' \to \chi\chi}^{m_\chi} \\ &= \sigma_{pp \to Z' \to \chi\chi}^{m_{\text{med}},\, m_\chi = 1\text{ GeV}} \cdot \frac{\mathcal{B}_{Z' \to \chi\chi}^{m_\chi}}{\mathcal{B}_{Z' \to \chi\chi}^{m_\chi = 1\text{ GeV}}}. \end{aligned} \tag{5.8.12}$$

Moreover, $\mathcal{B}_{Z' \to \chi\chi}^{m_\chi}$ is defined as:

$$\mathcal{B}_{Z' \to \chi\chi}^{m_\chi} = \frac{\Gamma_V^{\chi\chi}}{\Gamma_V^{\chi\chi} + \Gamma_V^{qq}}. \tag{5.8.13}$$

where the math expression of the two partial decay widths $\Gamma_V^{\chi\chi}$ and Γ_V^{qq} have already been given in Equation 2.2.14, analytically determined by the model parameters g_χ, g_q, m_{med} and m_χ. Hence, for each given m_{med}, two signal models are simulated with $m_\chi \sim 1$ GeV and $m_\chi \lesssim m_{\text{med}}/2$, while the expected signal strength for other m_χ values are interpolated as:

$$\mu_{m_\chi} = \mu_{m_\chi = 1\text{ GeV}} \cdot \frac{\mathcal{B}_{Z' \to \chi\chi}^{m_\chi = 1\text{ GeV}}}{\mathcal{B}_{Z' \to \chi\chi}^{m_\chi}}. \tag{5.8.14}$$

The upper limits on the cross section of the vector mediator simplified models are translated into the final exclusion limits in the m_χ-m_{med} plane. Figure 5.8.5 and Figure 5.8.6 show the excluded phase space of signal models obtained from measurements in hadronic $E_{\text{T}}^{\text{miss}} + W$, $E_{\text{T}}^{\text{miss}} + Z$ and $E_{\text{T}}^{\text{miss}} + V$ final states. Vector mediator masses m_{med} of up to 650 GeV are excluded at 95% CL for DM masses m_χ of up to 230 GeV, compatible with the expected exclusion of m_{med} of up to 700 GeV for m_χ of up to 250 GeV. For different mass points, the expected limits are improved by 15% - 30% with respect to the previous $E_{\text{T}}^{\text{miss}} + V$ (hadronic) DM search performed in ATLAS [146]. As illustrated in

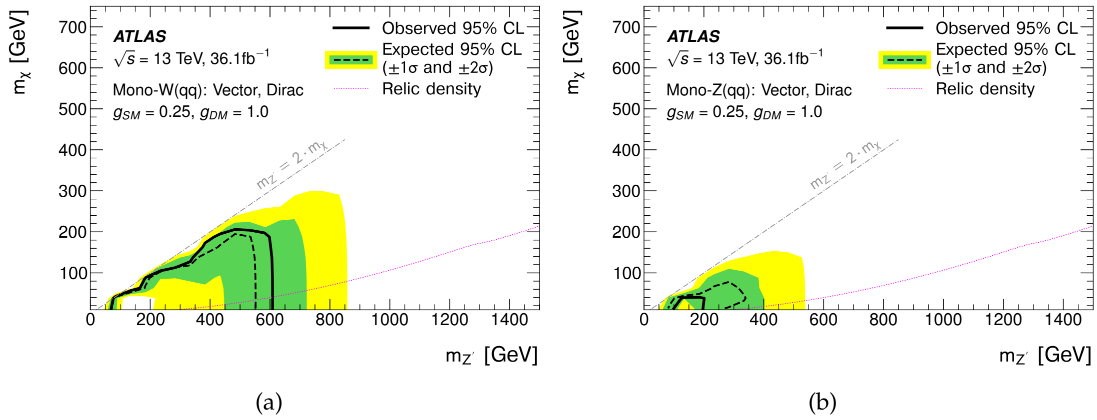


Figure 5.8.5: Exclusion limits on the Dark Matter production in association with (a) a W boson and (b) a Z boson for the vector mediator simplified model in the m_χ-m_{med} plane [1]. The couplings of the mediator to the SM and DM particles are fixed to $g_q = 0.25$ and $g_\chi = 1$, following the recommendations of the dark matter forum [147]. The solid (dashed) black curve shows the median of the observed (expected) exclusion limit, while the filled green (yellow) band denotes the $\pm 1\sigma$ ($\pm 2\sigma$) uncertainties on the expected result. The dotted magenta line represents the mass points for which the expected DM relic density is consistent with the WMAP [173] and Planck [29] measurements ($\Omega_{\text{DM}} h^2 = 0.1186 \pm 0.0020$), computed with MadDM [174]. The region on the right of the curve corresponds to higher predicted relic abundance.

Figure 5.8.5, the sensitivity of this analysis mainly comes from the $E_{\text{T}}^{\text{miss}} + W$ (hadronic) channel.

Figure 5.8.7 shows the exclusion limit of the $E_{\text{T}}^{\text{miss}} + V$ (hadronic) search together with the results from other DM searches in ATLAS, assuming $g_q = 0.25$, $g_l = 0$ and $g_\chi = 1$. Although the sensitivity of the $E_{\text{T}}^{\text{miss}} + X$ searches are dominated by the $E_{\text{T}}^{\text{miss}}$+jet result, it is still crucial to probe other $E_{\text{T}}^{\text{miss}} + X$ final states, given that the Standard Model particle X does not have to come from the initial state radiation as assumed in the simplified models. In the context of more complete models of DM, the recoiling particles may be emitted as part of the new effective vertex coupling DM to the SM and the sensitivity of the different searches will depend on the exact theory realized in nature.

Furthermore, these collider limits can be translated into DM-nucleon cross section limits to be compared to the current results from the direct detection experiments. As shown in Figure 5.8.8, the $E_{\text{T}}^{\text{miss}} + X$ searches present almost constant exclusion on the spin-independent DM-nucleon scattering cross section with respect to the DM mass for the vector mediator simplified model with the given set of couplings, while the direct detection experiments start to loose sensitivity at low m_χ. Therefore, the $E_{\text{T}}^{\text{miss}} + V$ (hadronic) search can provide a complementary approach to DM searches based on the production mechanism at hadron colliders when compared to the direct detection.

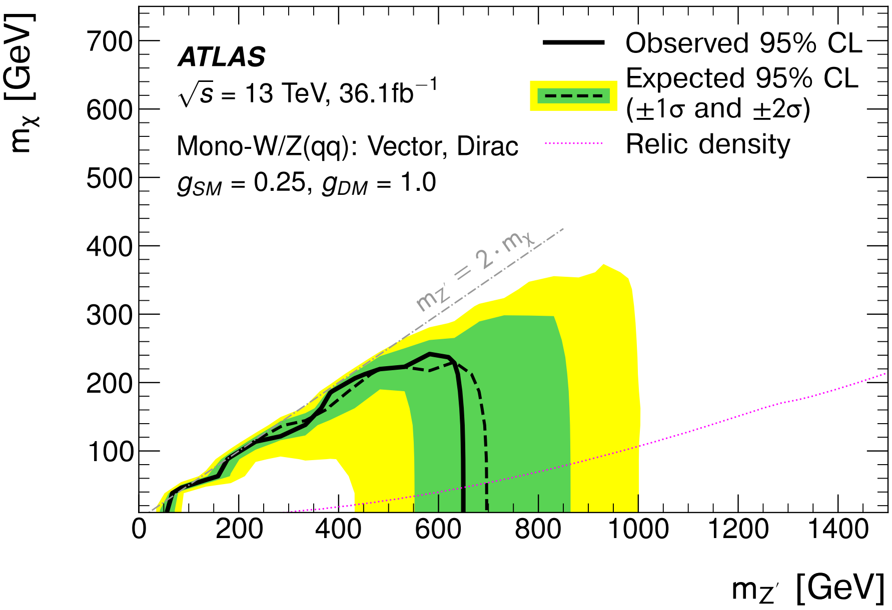


Figure 5.8.6: Exclusion limits on the Dark Matter production in association with a W/Z boson for the vector mediator simplified model in the m_χ-m_{med} plane [1]. The couplings of the mediator to the SM and DM particles are fixed to $g_q = 0.25$ and $g_\chi = 1$, following the recommendations of the dark matter forum [147]. The solid (dashed) black curve shows the median of the observed (expected) exclusion limit, while the filled green (yellow) band denotes the $\pm 1\sigma$ ($\pm 2\sigma$) uncertainties on the expected result. The dotted magenta line represents the mass points for which the expected DM relic density is consistent with the WMAP [173] and Planck [29] measurements ($\Omega_{\text{DM}} h^2 = 0.1186 \pm 0.0020$), computed with MadDM [174]. The region on the right of the curve corresponds to higher predicted relic abundance.

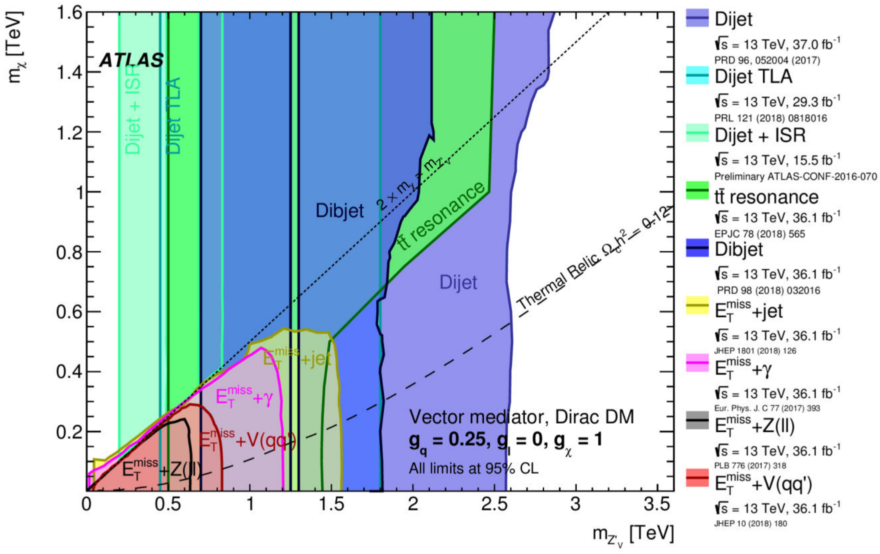


Figure 5.8.7: Exclusion in the m_χ-m_{med} plane at 95% CL by the dijet, dileptonic $t\bar{t}$ resonance searches and the $E_T^{\text{miss}} + X$ searches for the vector mediator simplified model [175]. The couplings of the mediator to the SM and DM particles are fixed to $g_q = 0.25$, $g_l = 0$ and $g_\chi = 1$. The dashed line represents the mass points for which the expected DM relic density is consistent with the WMAP [173] and Planck [29] measurements ($\Omega_{\text{DM}}h^2 = 0.1186 \pm 0.0020$), computed with MadDM [174]. The region on the right of the curve corresponds to higher predicted relic abundance.

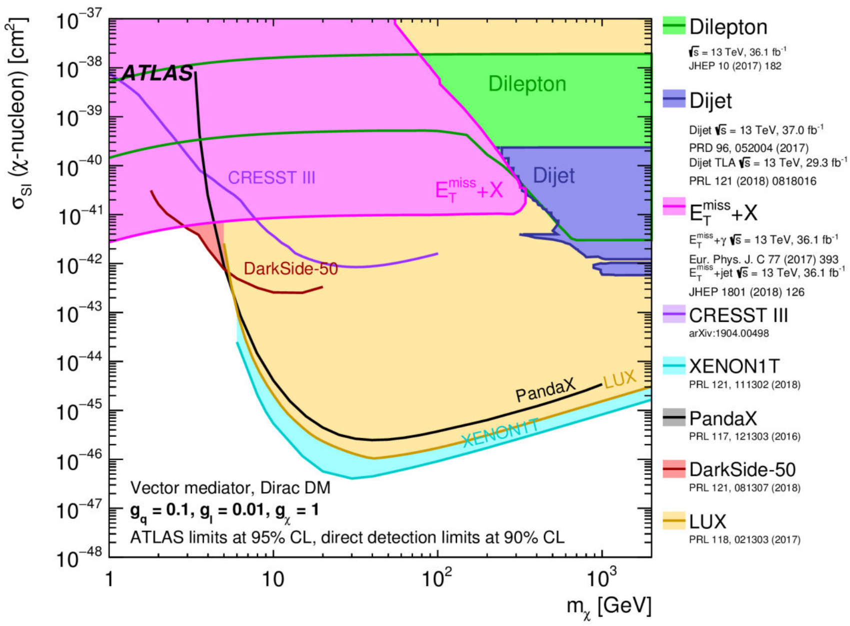


Figure 5.8.8: Exclusion on the spin-independent DM-nucleon scattering cross section as functions of m_χ inferred from the constraints from the direct detection experiments at 90% CL and from the ATLAS limits at 95% CL [175]. Note that the comparison is valid solely in the context of the vector mediator simplified model with the couplings of the mediator to the SM and DM particles fixed to $g_q = 0.1$, $g_l = 0.01$ and $g_\chi = 1$.

Chapter 6

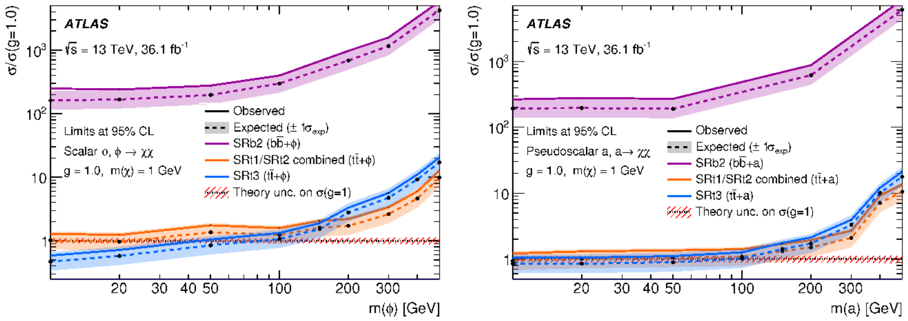
Search for Dark Matter produced in association with a top quark pair and medium missing transverse energy in 0 lepton final state

Search for Dark Matter in $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) final state has been performed by the ATLAS collaboration at the center-of-mass energy of $\sqrt{s} = 13$ TeV using 36.1 fb^{-1} of pp collision data [71] and limits were set on the parameters of simplified DM models which include an s-channel scalar or pseudo-scalar mediator, as shown in Figure 6.0.1.

With the accumulation of $\sqrt{s} = 13$ TeV data and improvements in the detector performance in LHC Run 2, the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search is performed using data recorded from 2016 to 2018, with an integrated luminosity of 126.7 fb^{-1}.

Compared to the previously performed searches for new physics beyond the Standard Model in fully-hadronic $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ final state, the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search intends to exploit the kinematic regions where comparably low missing transverse energy is present, which have never been looked into in the ATLAS collaboration. The b-jet triggers are applied in the event selection for the first time in searches for DM and SUSY particles, where $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers are of top popularity. The usage of $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger allows for full efficiency at $E_{\mathrm{T}}^{\mathrm{miss}} > 250$ GeV but also constrains the lower threshold of the $E_{\mathrm{T}}^{\mathrm{miss}}$ selection in the analyses. In contrast, the b-jet triggers provide best signal acceptance at low $E_{\mathrm{T}}^{\mathrm{miss}}$ regions but become less efficient at lager $E_{\mathrm{T}}^{\mathrm{miss}}$ values. Hence, the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search serves as a good complement to the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger based searches and is expected to add sensitivity to the statistical analysis after the future combination.

Section 6.1 focuses on the scalar and the pseudo-scalar mediator simplified models of Dark Matter, while the data and Monte Carlo samples as well as the triggers applied in the search are described in Section 6.2. Section 6.3 and Section 6.4 summarize the



(a) (b)

Figure 6.0.1: Observed limit on the signal strength, μ, for (a) the scalar mediator and (b) the pseudo-scalar mediator simplified models as functions of the mediator mass, m_{med}, for a fixed DM particle mass, $m_\chi = 1$ GeV [71], obtained using 36.1 fb^{-1} of pp collision data at $\sqrt{s} = 13$ TeV. The result in fully-hadronic (dileptonic) $E_{\text{T}}^{\text{miss}} + t\bar{t}$ final state is shown as the orange (blue) line, while the result in $E_{\text{T}}^{\text{miss}} + b\bar{b}$ final state is also shown in magenta.

analysis strategy and the definitions of the physical objects being considered in the analysis, respectively. The event selection of the signal regions is given in Section 6.5. Control regions are defined to constrain the dominant background compositions, as presented in Section 6.6. Section 6.7 lists all the systematic uncertainties taken into account in the statistical model. Finally, the results of the $E_{\text{T}}^{\text{miss}} + t\bar{t}$ (fully-hadronic) search are shown in Section 6.8.

6.1 Signal models

Similar to the vector models discussed in Section 5.1, the scalar/pseudo-scalar mediator simplified models considered in the $E_{\text{T}}^{\text{miss}} + t\bar{t}$ (fully-hadronic) search consist of six parameters:

- g_χ, coupling of the DM particles to the mediator;
- g_q, coupling of the SM quarks to the mediator;
- g_l, coupling of the SM leptons to the mediator;
- Γ, decay width of the mediator;
- m_{med}, mass of the mediator;
- m_χ, mass of the DM particle.

Among these parameters, the couplings to the SM and DM particles are chosen following the recommendations of the dark matter forum [147]:

$$g \equiv g_\chi = g_q = g_l = 1. \tag{6.1.1}$$

The choice of $g_\chi = g_q = g_l$ is made for the sake of simplicity. Unlike the quark coupling in the vector mediator models which is constrained by the dijet resonance searches, $g_q = 1$ is applied to all signals used in the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search, given that the dijet production from the scalar or the pseudo-scalar mediator decay only contributes at percent level to the total $t\bar{t}$ cross section [176]. Moreover, the leptonic decay of the mediator is estimated to be negligible in most of the parameter space considered, thus the non-trivial lepton coupling will not harm the orthogonality between the fully-hadronic and the dileptonic searches [177].

The minimal total decay width Γ is assumed to allow only for mediator decaying into DM or quark, while its value is determined by the choice of the couplings g_χ and g_q. For $g_\chi = g_q = 1$, $\Gamma_{\mathrm{min}}/m_{\mathrm{med}} \lesssim 0.1$ is expected, fulfilling the narrow width approximation.

Three categories can be defined for the scalar and the pseudo-scalar models based on the mediator mass m_{med} and the DM mass m_χ:

- **On-shell**. When $m_{\mathrm{med}} \gg 2m_\chi$, the acceptance of the analysis is independent of m_χ, allowing for extrapolation of exclusion limits for signals with a same m_{med}. The on-shell region is where this analysis has best sensitivity.

- **Threshold** and **Off-shell**. When $m_{\mathrm{med}} \approx 2m_\chi$ or $m_{\mathrm{med}} \ll 2m_\chi$, the predicted cross section will be much smaller compared to the on-shell production. Based on the total integrated luminosity of Run 2 data, searches for DM production in association with top quark pair are expected to have minimal sensitivity to such signals.

As a result, DM signals with different m_{med} and a fixed m_χ of 1 GeV are simulated and tested in the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search. Additionally, in low m_{med} region, the kinematic dependence of event topology on m_{med} is significantly stronger for the scalar signals with respect to the pseudo-scalar signals. This is mainly due to the fact that for mediators with $m_{\mathrm{med}} \ll 2m_t$, a major component of $t\bar{t}$+DM production is the FSR, where the mediator radiated from the final state top quarks decays to a pair of DM particles. The corresponding radiation pattern is determined by the fragmentation functions [178]:

$$f_{t\to\phi}(x) = \frac{g_t^2}{(4\pi)^2}\left[\frac{4(1-x)}{x} + x\ln\left(\frac{s}{m_t^2}\right)\right], \tag{6.1.2}$$

$$f_{t\to a}(x) = \frac{g_t^2}{(4\pi)^2}\left[x\ln\left(\frac{s}{m_t^2}\right)\right], \tag{6.1.3}$$

with ϕ and a representing the scalar and the pseudo-scalar mediator, g_t the top coupling, m_t the top mass, x the energy fraction of the radiated particle, $x = E_{\phi/a}/E_t$, and $\sqrt{s} = 13$ TeV. For scalar mediators, an additional term of $4(1-x)/x$ is present in the fragmentation function, which naturally leads to enhanced FSR at small x. Hence, the kinematic properties of signals events are shaped by the interference between the s-channel production and the dominating FSR process for light scalars. To exploit the characteristics of such physical effect, a finer scan on m_{med} is performed for the scalar mediator simplified models at $m_{\mathrm{med}} < 100$ GeV, with 13 and 10 mass points generated for the scalar and the pseudo-scalar signals, respectively.

6.2 Data and Monte Carlo simulation

Hard-scatter pp collision data measured by the ATLAS detector together with the MC simulated signal and background samples are used in the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search. A brief overview of the analyzed data samples is given in Section 6.2.1 and the configurations of MC production of various signal and background processes are detailed in Section 6.2.2. L1 triggers and HLTs are applied to stabilize the event rate with the increasing instantaneous luminosity during LHC Run 2 data-taking and to reject the overwhelming QCD background which harms the sensitivity of the analysis. Section 6.2.3 presents the triggers included in the event selection and a truth-level study which evaluates the signal efficiencies of event selections based on the usage of $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers and 2 b-jet + 2 jet triggers is documented in Section 6.2.3.1.

6.2.1 Data samples

The $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search uses pp collision data collected by the ATLAS detector from 2016 and 2018 under the premise of stable beam conditions. The 2015 data, on the other hand, is excluded due to the lack of suitable b-jet triggers. This results in the loss of approximately 2% - 3% of the total integrated luminosity.

All analyzed events are required to meet the standard ATLAS data quality assessment criteria, as described in Section 3.2.5. Special GRLs for b-jet triggers are applied to confirm the normal functionality of all sub-detector systems, including a sub-system requisite for online b-tagging. This leads to the integrated luminosity of 24.6 fb^{-1}, 43.7 fb^{-1} and 58.4 fb^{-1} in 2016, 2017 and 2018 data-taking, respectively, with a total luminosity of 126.7 fb^{-1}.

The average number of pp interactions per bunch crossing $\langle\mu\rangle$ was around 25 in 2016, around 38 in 2017 and around 36 in 2018, as shown in Figure 3.1.3.

6.2.2 Monte Carlo samples

As described in Section 5.2.2, the MC simulation of signal and background processes are performed in two steps, i.e. event generation and detector simulation. Various MC event generators are designed to model the hard-scattering and hadronization processes with fixed-order perturbation theory, while the interactions between the final particles and the detector materials are simulated by the GEANT4 toolkit [148].

In this search, the simplified model DM signals are generated in a grid of mediator mass m_{med}, following the recommendations of the dark matter forum [147], with 23 mass points in total. MADGRAPH5 [149] is used to provide the leading-order (LO) matrix element calculation for the hard-scattering process using NNPDF23LO PDF set [143] and PYTHIA 8 [142] is used for the simulation of parton shower and hadronization using A14 tune [144]. Next-to-leading order (NLO) corrections are applied to signal cross sections calculation.

The major background processes include the production of W and Z bosons in addition to partons (W+jets and Z+jets), top quark production ($t\bar{t}$ and single top), top quark pair production in association with a W/Z boson, as well as diboson production. The W+jets and Z+jets events are simulated with the SHERPA 2.2.1 [150] generator using NNPDF30NNLO PDF set [151] and normalized to the next-to-next-to-leading order (NNLO) cross section. For $t\bar{t}$ and single top processes, the POWHEG [152] generator using NNPDF30NNLO PDF set is applied and interfaced to Pythia 8 using A14 tune, while the NNLO correction and the next-to-next-to-leading-logarithm (NNLL) resummation [179, 180, 181, 182, 183, 184, 185, 186, 187] are considered for the cross section calculation. The $t\bar{t}+V$ process is modeled by MadGraph5_aMC@NLO [165] and Pythia 8 using NNPDF30 PDF set and A14 tune, with its cross section normalized to the next-to-leading order (NLO). Diboson production, including WW, WZ and ZZ, is simulated with Sherpa 2.2.1 with NNPDF30NNLO PDF set.

A summary of the MC generators, PDF and tune sets used for the production of simulated signal and background samples can be found in Table 6.2.1.

Process	**Generator**	**PDF / tune**
Scalar mediator models	MadGraph5 + Pythia 8 (NLO)	NNPDF23LO + A14
Pseudo-scalar mediator models	MadGraph5 + Pythia 8 (NLO)	NNPDF23LO + A14
V + jets		
W + jets	Sherpa 2.2.1 (NNLO)	NNPDF30NNLO
Z + jets	Sherpa 2.2.1 (NNLO)	NNPDF30NNLO
$t\bar{t}$	Powheg + Pythia 8 (NNLO + NNLL)	NNPDF30NNLO + A14
Single top		
s-channel	Powheg + Pythia 8 (NNLO + NNLL)	NNPDF30NNLO + A14
t-channel	Powheg + Pythia 8 (NNLO + NNLL)	NNPDF30NNLO + A14
Wt	Powheg + Pythia 8 (NNLO + NNLL)	NNPDF30NNLO + A14
$t\bar{t}+V$		
$t\bar{t}+W$	MadGraph5_aMC@NLO (NLO)	NNPDF23 + A14
$t\bar{t}+Z$	MadGraph5_aMC@NLO (NLO)	NNPDF23 + A14
Diboson		
$WlvWqq$	Sherpa 2.2.1 (NLO)	NNPDF30NNLO
$WlvZqq$	Sherpa 2.2.1 (NLO)	NNPDF30NNLO
$WqqZll$	Sherpa 2.2.1 (NLO)	NNPDF30NNLO
$WqqZvv$	Sherpa 2.2.1 (NLO)	NNPDF30NNLO
$ZqqZll$	Sherpa 2.2.1 (NLO)	NNPDF30NNLO
$ZqqZvv$	Sherpa 2.2.1 (NLO)	NNPDF30NNLO

Table 6.2.1: List of the MC generators, PDF and tune sets used for the production of simulated signal and background samples in the $E_{\mathrm{T}}^{\mathrm{miss}}+t\bar{t}$ (fully-hadronic) search.

For all MC simulated samples used in the analysis, the amount of pile-up contribution is adjusted by a reweighting procedure to match the actual condition in the corresponding data-taking period.

6.2.3 Triggers

The $E_{\mathrm{T}}^{\mathrm{miss}}+t\bar{t}$ (fully-hadronic) search aims to maximize the acceptance of DM signals with medium missing transverse energy, i.e. $E_{\mathrm{T}}^{\mathrm{miss}} > 160$ GeV, in contrast to the SUSY searches which focus on larger $E_{\mathrm{T}}^{\mathrm{miss}}$ signatures. As shown in Figure 5.2.1, $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers are not fully efficient in region of $E_{\mathrm{T}}^{\mathrm{miss}} < 250$ GeV, excluding a major part of the phase space targeted by the analysis. Consequently, 2 b-jet + 2 jet triggers are used in the event selection instead of the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers, ensuring 100% efficiency in the full kinematic region.

The decision of 2 b-jet + 2 jet trigger is made based on the online multivariate b-tagging algorithms which are trained in a similar way as the offline flavor tagging tools, allowing for coherence between the online and the offline b-jet definitions. Additionally, the time performance of b-jet triggers is optimized by increasing the track p_{T} threshold in vertex finding and reducing the size of the region of interest for b-tagged track matching, with no significant loss in identification performance.

For the selected 2 b-jet + 2 jet triggers, different tagging algorithms with varying fixed-cut working points are used in 2016, 2017 and 2018. During 2016 data-taking, the MV2c20 algorithm which is trained with backgrounds consisting of 20% c-jets and 80% light-flavor jets was employed with a fixed 60% b-tagging efficiency, while the MV2c10 algorithm, trained on b-jet signals with backgrounds consisting of 7% c-jets and 93% light-flavor jets, was used in 2017 and 2018, providing fixed tagging efficiency of 40% and 60%, respectively.

Table 6.2.2 lists the 2 b-jet + 2 jet triggers used in the analysis. Same trigger applies to all regions with different lepton (e/μ) multiplicities in each year, in order to maintain the minimal extrapolation of event topology from the control regions to the signal regions, which is further discussed in Section 6.6.1.1 and Section 6.6.2.1.

Period	Name
2016	HLT_2j35_bmv2c2060_split_2j35_L14J15.0ETA25
2017	HLT_2j15_gsc35_bmv2c1040_split_2j15_gsc35_boffperf_split_L14J15.0ETA25
2018	HLT_2j35_bmv2c1060_split_2j35_L14J15.0ETA25

Table 6.2.2: List of triggers applied in the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search using data recorded from 2016 to 2018. Same trigger applies to all regions with different lepton (e/μ) multiplicities in each year.

6.2.3.1 Truth level study on *b*-jet trigger

The application of 2 b-jet + 2 jet triggers is validated by a truth-level study which compares the signal acceptance for the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger based and the b-jet trigger based event selections. Gaussian smearing is applied to the particle-level kinematic properties according to the detector resolution to mimic the performance of reconstructed objects, while the decision of identification, isolation and flavor tagging algorithms are simulated with respect to the corresponding efficiencies extracted from the dedicated measurements performed by the ATLAS CP groups. Table 6.2.3 summarizes the various truth-level object definitions used in this study.

A minimal event selection is applied at truth-level, which resembles the expected baseline selection at reconstruction-level. Two sets of selection criteria are defined as shown in Table 6.2.4, one assuming the triggering on $E_{\mathrm{T}}^{\mathrm{miss}}$ and the other one on b-jets. For the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger based selection, $E_{\mathrm{T}}^{\mathrm{miss}} > 250$ GeV has to be fulfilled, ensuring the full efficiency of the trigger. The $E_{\mathrm{T}}^{\mathrm{miss}}$ threshold is lowered to 100 GeV when 2 b-jet + 2 jet triggers are applied.

Object	Kinematics	Type, Quality
Electrons	$p_{\mathrm{T}} > 7$ GeV, $\lvert\eta\rvert < 2.47$	VeryLooseLH, GradientLoose
Muons	$p_{\mathrm{T}} > 6$ GeV, $\lvert\eta\rvert < 2.5$	Loose, GradientLoose
Small-R jets	$p_{\mathrm{T}} > 20$ GeV, $\lvert\eta\rvert < 2.8$	anti-k_t $R = 0.4$, JVT requirement
b-jets	loose: $\lvert\eta\rvert < 2.5$	MV2c10, fixed-cut 77%
	medium: $\lvert\eta\rvert < 2.5$	MV2c10, fixed-cut 60%
	tight: $\lvert\eta\rvert < 2.5$	MV2c10, fixed-cut 40%

Table 6.2.3: Summary of the objects used in the truth-level study.

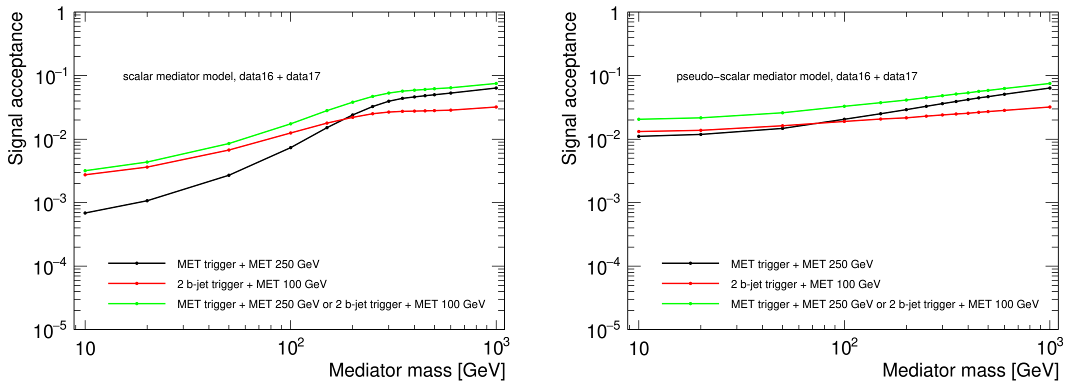
At least four jets with p_{T} above (80, 80, 40, 40) GeV are required, which allows the b-jet triggers to operate fully efficiently and effectively removes the possible multijet contamination. Events with $\Delta\phi_{\min}(E_{\mathrm{T}}^{\mathrm{miss}}, \text{jet}^{1\text{-}4}) \leq 0.4$ are vetoed to further reduce the multijet background with fake $E_{\mathrm{T}}^{\mathrm{miss}}$ arising from mismeasurements. The transverse mass between $E_{\mathrm{T}}^{\mathrm{miss}}$ and the closest b-jet in terms of $\Delta\phi$, also known as $m_{\mathrm{T}}^{b,\,\min}$, provides very good discrimination between the signal events and backgrounds containing tops. A cut-off of $m_{\mathrm{T}}^{b,\,\min}$ at around m_t is expected in case the measured $E_{\mathrm{T}}^{\mathrm{miss}}$ and the b-jet originate from the same top quark decay, while for $t\bar{t}$+DM production $m_{\mathrm{T}}^{b,\,\min}$ is allowed to have much larger values. A very loose cut on $m_{\mathrm{T}}^{b,\,\min}$ at 50 GeV is applied for QCD suppression.

For both selection criteria, at least two b-jets are required, tagged by different identification working points. Loose b-jets are selected in the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger based selection, while medium and tight b-jets are used in the selection based on 2 b-jet + 2 jet triggers, accounting for the different online flavor tagging working points applied in 2016 and 2017, respectively.

$E_{\mathrm{T}}^{\mathrm{miss}}$ **trigger based selection**	***b*-jets trigger based selection**
overlap removal	
0 electron / muon	
at least 4 jets	
jet $p_{\mathrm{T}} > (80, 80, 40, 40)$ GeV	
$\Delta\phi_{\min}(E_{\mathrm{T}}^{\mathrm{miss}}, \text{jet}^{1\text{-}4}) > 0.4$	
$m_{\mathrm{T}}^{b,\,\min} > 50$ GeV	
$E_{\mathrm{T}}^{\mathrm{miss}} > 250$ GeV	$E_{\mathrm{T}}^{\mathrm{miss}} > 100$ GeV
at least 2 loose b-jets	at least 2 medium (tight) b-jets for 2016 (2017)

Table 6.2.4: Summary of the selections applied in the truth-level study.

The signal acceptance for the scalar/pseudo-scalar mediator models with a fixed m_χ of 1 GeV is calculated as functions of the mediator mass, m_{med}, after the above mentioned event selections. As shown in Figure 6.2.1, the 2 b-jet + 2 jet trigger based event selection outperforms the selection based on $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger at low m_{med} region. This result strongly motivates the design of a search focusing on medium transverse energy and using b-jet triggers for the optimal acceptance of signals with low mediator masses.



(a) (b)

Figure 6.2.1: Signal acceptance after truth-level event selections as functions of m_{med} with a fixed m_χ of 1 GeV for (a) the scalar mediator models and (b) the pseudo-scalar mediator models. Two medium (tight) b-jets are required in the b-jet trigger based selection to mimic the performance of the 2 b-jet + 2 jet trigger available in 2016 (2017). The combination of the b-jet trigger based selection in 2016 and 2017 is achieved by randomly assigning simulated signal events to either criterion with respect to the integrated luminosity of each year. Hence, any correlation due to double counting the same event in both selections can be avoided.

6.3 Analysis strategy

A general overview of the analysis strategy of the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search is given in this section, including event categorization, main discriminants for signal event selection and treatment of the dominant SM backgrounds. A brief summary of all signal regions (SRs), control regions (CRs) and validation regions (VRs) used in the analysis is presented in Section 6.3.3.

6.3.1 Signal region strategy

The final state of the analyzed signal scenarios contains DM particles recoiling against a pair of top quarks, which then decay to a bunch of jets including two b-jets from $t \to bW$. Compared to other decay modes of $t\bar{t}$, fully-hadronic final state presents the advantage of including a relatively high cross section and not having neutrinos in the final state such that $E_{\mathrm{T}}^{\mathrm{miss}}$ originates only from the searched invisible particles.

Due to the Lorentz boost of the top system, different event topologies need to be considered. When top p_{T} is rather low, the decay products of top are expected to have good angular separation, forming multiple small-R jets in the detector, often referred to as the resolved topology. With the rising of top p_{T}, the small-R jets from top decays tend to be closer in ΔR and start to merge with each other. In this case, large-R jets are used to study the kinematic properties of the boosted top. By comparing the mass of the large-R jet to the top mass, one can learn whether it contains most of the decay products from a single top quark, which is likely to happen when the top system is highly boosted.

Three kinematic regimes are defined based on such reasoning, namely resolved, boosted and highly boosted, with event selections optimized in each individual regime. Events in the resolved regime are required to have no large-R jets, while for boosted and highly boosted at least one large-R jet has to present. Then, the mass of the highest mass large-R jet, denoted by m_J, is compared to 130 GeV. If $m_J < 130$ GeV, the event is categorized as boosted. Otherwise it enters the highly boosted regime.

The main discriminants of the analysis are $\cosh_{\mathrm{max}}$ and $\chi^2_{t\bar{t},\,\mathrm{had}}$, with detailed descriptions in Section 6.5.3.1 and Section 6.5.3.2, respectively. In general, $\cosh_{\mathrm{max}}$ provides very good discrimination against background processes containing leptonically decaying top quark, while $\chi^2_{t\bar{t},\,\mathrm{had}}$ evaluates the probability of events to have two hadronically decaying tops, thus can be used to select the DM signals.

Three SRs are defined in 0 lepton final state, corresponding to the three kinematic regimes. Additionally, the $\cosh_{\mathrm{max}}$ cuts in the SRs are inverted to select events with leptonic tops, forming the $t\bar{t}$ enriched validation regions, whereas the $\chi^2_{t\bar{t},\,\mathrm{had}}$ cuts are inverted to select events with no hadronic tops, forming the $t\bar{t}$ suppressed validation regions. Six VRs in total are designed to validate the control-to-signal region extrapolation, including three $t\bar{t}$ enrich VRs and three $t\bar{t}$ suppressed VRs. Unlike in the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search, these VRs are not included in the combined profile likelihood fit.

6.3.2 Control region strategy

The common experimental signature searched in this analysis consists of multiple jets and b-jets from $t\bar{t}$ decay, medium $E_{\mathrm{T}}^{\mathrm{miss}}$ and no isolated leptons. The dominant background sources are:

- $Z \to \nu\nu$ plus additional b-jets from QCD gluon splitting or the 'scattering' diagrams;
- $t\bar{t}$ production containing $W \to l\nu$ decay, where the lepton is either lost or misidentified as a jet;
- Wt-channel single top production, where one of the W decays hadronically and the other one decays leptonically, with the lepton being either lost or misidentified as a jet;
- $t\bar{t} + Z$, where both tops decay hadronically and Z decays to $\nu\nu$, the irreducible background of the analysis.

In addition, subdominant background components include:

- $W \to l\nu$ plus additional b-jets from QCD gluon splitting, where the lepton is either lost or misidentified as a jet;
- Diboson decaying fully-hadronically or semi-leptonically, with the lepton being either lost or misidentified as a jet, plus additional b-jets;
- $t\bar{t} + W$, where both tops decay hadronically and W decays to $l\nu$, with the lepton being either lost or misidentified as a jet.

The background contributions are modeled with MC simulations and constrained by the usage of CRs, specifically designed to be enriched in certain background processes. The 1 lepton CRs are used to constrain the $t\bar{t}$ and single top backgrounds, while the 2 lepton CRs are used to constrain the $Z \to \nu\nu$ background using $Z \to ll$ events, since the hard-scattering processes for these two backgrounds are very similar. The $t\bar{t} + Z$ background, however, is irreducible in this analysis and cannot be distinguished from the DM signal.

Fortunately the rather small $t\bar{t}+Z$ cross section prevents it from becoming a major component in the SRs.

Moreover, the $t\bar{t}$ background is divided into two categories in the simulation, based on the number of truth-level b-hadrons exclusively associated to reconstructed jets[1]:

- $t\bar{t}$: at most two associated b-hadrons;
- $t\bar{t}+b$: more than two associated b-hadrons.

Such categorization is motivated by the very different kinematic behaviors observed between $t\bar{t}$ and $t\bar{t}+b$, especially in the distribution of $\cosh_{\max}$, one of the main discriminants of the analysis. Figure 6.3.1 shows the shape of $\cosh_{\max}$ for $t\bar{t}$, $t\bar{t}+b$ and single top backgrounds, together with two DM signals with $m_\chi = 1$ GeV and $m_{\phi/a} = 20$ GeV. A nice Gaussian peak which centers above one is formed by the $t\bar{t}$ events, while $t\bar{t}+b$ and single top both have a lower tail below one. Hence, the ratio between the event yields of $t\bar{t}$ and $t\bar{t}+b$ in the SRs is strongly dependent on the choice of the $\cosh_{\max}$ threshold and constraining these two background components separately appears then to be the most robust and natural solution.

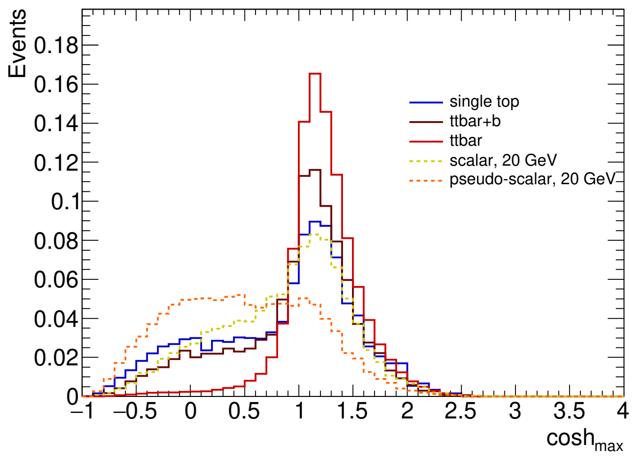


Figure 6.3.1: $\cosh_{\max}$ distribution for $t\bar{t}$ (solid red), $t\bar{t}+b$ (solid brown) and single top (solid blue) backgrounds, normalized to unit area. The scalar mediator (dashed yellow) and the pseudo-scalar mediator (dashed orange) models are plotted alongside. The broad peak visible for the scalar model comes from the semi-leptonically decaying signal events with the lepton being either lost or misidentified as a jet. This is not a part of the phase space expected to be constrained by the $E_{\mathrm{T}}^{\mathrm{miss}}+t\bar{t}$ (fully-hadronic) search.

The QCD multijet events with fake $E_{\mathrm{T}}^{\mathrm{miss}}$ coming from poorly measured jets may contaminate the SRs especially when loose $E_{\mathrm{T}}^{\mathrm{miss}}$ selections are applied. Therefore, anti-QCD cuts are specially designed to suppress this process, as discussed in Section 6.5.2. The remaining multijet contribution is estimated using a data-driven 'Rebalance and Smear' method, where poorly measured or imbalanced data events are recycled and used as templates to generate pseudo multijet background events, as detailed in Section 6.6.4.

Three CRs are defined for each of the major background processes, i.e. $t\bar{t}$, $t\bar{t}+b$, single top and Z+jets, resulting in nine disjoint CRs in 1 lepton final state and three CRs in 2

[1]The number of truth b-hadrons is calculated by counting the weakly-decaying b-hadrons with $p_{\mathrm{T}} > 5$ GeV within $\Delta R = 0.3$ of any reconstructed small-R jet with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 4.5$.

lepton final state. The normalization of $t\bar{t}+Z$ is determined by MC prediction, therefore no dedicated $t\bar{t}+Z$ CR is defined in this analysis.

6.3.3 Summary of analysis regions

Table 6.3.1 provides a summary of all signal regions, control regions and validation regions defined in the $E_{\mathrm{T}}^{\mathrm{miss}}+t\bar{t}$ (fully-hadronic) search.

Kinematic regimes		Resolved	Boosted	Highly boosted
0 lepton	signal enriched	●	●	●
	$t\bar{t}$ enriched	◇	◇	◇
	$t\bar{t}$ suppressed	◇	◇	◇
1 lepton	$t\bar{t}$ enriched	○	○	○
	$t\bar{t}+b$ enriched	○	○	○
	single top enriched	○	○	○
2 lepton	Z+jets enriched	○	○	○

Table 6.3.1: Summary of the signal regions '●', control regions '○' and validation regions '◇' defined in the $E_{\mathrm{T}}^{\mathrm{miss}}+t\bar{t}$ (fully-hadronic) search.

6.4 Object definition

A large variety of physical object definitions are applied in the $E_{\mathrm{T}}^{\mathrm{miss}}+t\bar{t}$ (fully-hadronic) search, including different kinematic requirements, identification and isolation criteria, etc. A brief introduction is given as follows.

Two types of electron definitions are considered:

- **Loose**. Loose electrons are used in the signal region selection and are required to have $p_{\mathrm{T}} > 5$ GeV and $|\eta| < 2.47$. The LooseAndBLayerLH identification criterion and the FixedCutLoose isolation criterion have to be fulfilled, as defined in Section 3.3.1.
- **Tight**. Tight electrons are used in the 1 lepton and the 2 lepton CR selections and are required to have $p_{\mathrm{T}} > 15$ GeV and $|\eta| < 2.47$. The TightLH identification criterion and the FixedCutTight isolation criterion have to be fulfilled.

Two types of muon definitions are considered:

- **Loose**. Loose muons are used in the signal region selection and are required to have $p_{\mathrm{T}} > 5$ GeV and $|\eta| < 2.7$. The Loose identification criterion and the FixedCutLoose isolation criterion have to be fulfilled, as defined in Section 3.3.2.
- **Tight**. Tight muons are used in the 1 lepton and the 2 lepton CR selections and are required to have $p_{\mathrm{T}} > 15$ GeV and $|\eta| < 2.7$. The Medium identification criterion and the FixedCutTight isolation criterion have to be fulfilled.

Aside from electrons and muons, hadronic taus are used for event veto in all analysis regions and are required to have $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 1.37$ or $1.52 \leq |\eta| < 2.5$, ensuring that the tau components fall into the fiducial area of the tracker excluding the transition region between the barrel and the end-caps of the ECal. The VeryLoose identification criterion [188] defined by the ATLAS Tau Performance group has to be fulfilled.

The small-R jets used in the analysis are reconstructed from topo-clusters using the anti-k_{t} algorithm with $R = 0.4$ and calibrated at the EM+JES scale. Furthermore they can be divided into two categories based on their pseudorapidity:

- **Central jets**, with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$;
- **Forward jets**, with $p_{\mathrm{T}} > 20$ GeV and $2.5 \leq |\eta| < 4.5$.

For better pile-up suppression, low p_{T} central jets (i.e. $p_{\mathrm{T}} < 60$ GeV) are required to pass the Jet Vertex Tagger (JVT) [134] selection with the Medium criterion. Based on the more complex pile-up condition in 2017 and 2018, the forward Jet Vertex Tagger (fJVT) [135] procedure functioning with the Tight working point is used to identify pile-up jets in the high pseudorapidity region based on unbalanced central jets identified by the JVT as originating from pile-up interactions. Correspondingly, a JVT (fJVT) efficiency scale factor is applied to the MC event weight for each event passing the selection, which corrects the residual difference in JVT (fJVT) score evaluation between data and MC simulation.

Additionally, the trigger-level jets, also referred to as the 'online jets', are built from HLT objects following the same reconstruction scheme as the small-R jets ('offline jets'). A trigger matching is performed to associate the online and the offline jets if their angular separation ΔR is smaller than 0.2.

As described in Section 6.2.3, the trigger-level jets are assigned b-tagging discriminant values by the online MV2 algorithms, which are trained with different background compositions and operate at different working points:

- **2016**: 60% fixed-cut working point of MV2c20
- **2017**: 40% fixed-cut working point of MV2c10
- **2018**: 60% fixed-cut working point of MV2c10

For the offline jets, the MV2c10 discriminant is used to identify small-R jets originating from b-hadrons with a fixed 60% efficiency. Given that the online and the offline b-tagging scores are highly correlated due to the same tagging algorithm applied, offline b-jets are required to be exclusively matched to online b-jets, which helps to provide consistent b-jet definition over the full analysis chain and accounts for the lack of efficiency calibration for b-jet identification at trigger-level.

The large-R jets used in the analysis are reconstructed from topo-clusters using the anti-k_{t} algorithm with $R = 1.0$ and calibrated at the LCW+JES+JMS scale. The trimming procedure is applied to the large-R jets by reclustering the initial jet constituents using the k_{t} algorithm with $R_{\mathrm{sub}} = 0.2$ and then removing any subjets with p_{T} less than $f_{\mathrm{cut}} = 0.05$ times the p_{T} of the parent jet. Lastly, the trimmed large-R jets are required to have $p_{\mathrm{T}} > 200$ GeV, $m > 40$ GeV and $|\eta| < 2.0$, in order to select high p_{T} central jets with a good overlap between the ID and the calorimeter in a phase space where the p_{T} and the mass calibrations are available.

The missing transverse momentum is reconstructed using well-calibrated physical objects and the track-based soft term (TST), while the Tight working point is applied with the Tight fJVT criterion, considering the rather high $\langle\mu\rangle$ in 2017 and 2018 data-taking. In the 1 lepton and the 2 lepton CRs, a modified version of $E_{\mathrm{T}}^{\mathrm{miss}}$, namely $E_{\mathrm{T,\,no\ lepton}}^{\mathrm{miss}}$, is computed by subtracting the lepton contribution from the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector. This is achieved by removing all components associated to the target lepton, including the calorimeter clusters and ID tracks, from the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction input, which ensures the invisibility of the leptons.

The object-based missing transverse energy significance is used to reject events with $E_{\mathrm{T}}^{\mathrm{miss}}$ originating from resolution effects. Similarly, a modified version of object-based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance, namely object-based $E_{\mathrm{T,\,no\ lepton}}^{\mathrm{miss}}$ significance, is applied in the 1 lepton and the 2 lepton CR selections.

A concise summary of the various object definitions used in this analysis is presented in Table 6.4.1.

Object	**Kinematics**	**Type, Quality**
Electrons	loose: $p_{\mathrm{T}} > 5$ GeV, $\lvert\eta\rvert < 2.47$	LooseAndBLayerLH, FixedCutLoose
	tight: $p_{\mathrm{T}} > 15$ GeV, $\lvert\eta\rvert < 2.47$	TightLH, FixedCutTight
Muons	loose: $p_{\mathrm{T}} > 5$ GeV, $\lvert\eta\rvert < 2.7$	Loose, FixedCutLoose
	tight: $p_{\mathrm{T}} > 15$ GeV, $\lvert\eta\rvert < 2.5$	Medium, FixedCutTight
Hadronic taus	$p_{\mathrm{T}} > 20$ GeV, $\lvert\eta\rvert < 1.37$ or $\lvert\eta\rvert = [1.52, 2.5)$	VeryLoose
Small-R jets	central: $p_{\mathrm{T}} > 20$ GeV, $\lvert\eta\rvert < 2.5$	anti-k_t $R = 0.4$, EMTopo
	forward: $p_{\mathrm{T}} > 20$ GeV, $\lvert\eta\rvert = [2.5, 4.5)$	b-tag: MV2c10, fixed-cut 60%
Large-R jets	$p_{\mathrm{T}} > 200$ GeV, $m > 40$ GeV, $\lvert\eta\rvert < 2.0$	anti-k_t $R = 1.0$, LCTopo trimmed, $R_{\mathrm{sub}} = 0.2$, $f_{\mathrm{cut}} = 0.05$
$E_{\mathrm{T}}^{\mathrm{miss}}$	$E_{\mathrm{T}}^{\mathrm{miss}} > 160$ GeV	track-based soft term Tight working point Tight fJVT
$E_{\mathrm{T}}^{\mathrm{miss}}$ significance	$E_{\mathrm{T}}^{\mathrm{miss}}$ sig. > 8	

Table 6.4.1: Summary of the reconstructed objects used in the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search.

6.4.1 Resolving overlapping objects

Similar to the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search, an overlap removal procedure is applied to resolve the ambiguities of object definition and avoid any double-counting of detector signals with the following order:

- **Electron-muon overlap removal**. Electron candidates are removed if they share the same ID tracks with a CB muon. In the case of CT muons, the muon candidates are removed instead.
- **Electron-jet overlap removal**. This is performed in two steps. First, small-R jets within $\Delta R = 0.2$ of any well-identified electrons are removed. Then, electrons within $\Delta R = 0.4$ of any surviving small-R jets are removed to avoid double-counting of energy.

- **Muon-jet overlap removal**: Similarly, small-R jets within $\Delta R = 0.2$ of any well-identified muons are removed if the jet has fewer than three associated tracks. Then, muons within $\Delta R = 0.4$ of any surviving small-R jets are removed to avoid double-counting of energy. This avoids the inefficiency of high p_{T} muons undergoing significant energy loss in the calorimeter.

Since the hadronic taus are only used to veto events, no particular overlap removal is performed in case they overlap with another object.

Note that the distance metric used to define the overlapping objects is $\Delta R = \sqrt{\Delta y^2 + \Delta\phi^2}$, where the rapidity y is used instead of the pseudorapidity η, as recommended in Ref. [158].

6.5 Event selection

As stated in Section 6.3, the signal events of the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search are characterized by jets with kinematic patterns compatible with hadronically decaying $t\bar{t}$, medium $E_{\mathrm{T}}^{\mathrm{miss}}$ and no isolated leptons. The SR selections are designed and optimized for such final state, while event categorization based on top quark p_{T} is introduced to further enhance the sensitivity to DM signals. Section 6.5.1 discusses the baseline selection applied to all regions in the analysis. Section 6.5.2 details the anti-QCD requirements applied for multijet background suppression. The full list of the SR selections is presented in Section 6.5.3.

6.5.1 Baseline selection

All events are required to satisfy the following basic quality criteria:

- **GRL and event cleaning**: data events failing the standard ATLAS data quality assessment are vetoed.
- **Trigger**: events are required to pass the 2 b-jet + 2 jet trigger selection, see Section 6.2.3.
- **Vertex selection**: at least one reconstructed vertex with at least two associated tracks is required.
- **Jet cleaning**: events containing jets flagged by the BadLoose jet cleaning requirement [159] are vetoed to ensure a good measurement of $E_{\mathrm{T}}^{\mathrm{miss}}$.
- **Muon quality check**: events containing muons flagged as BadMuon with significantly bad momentum resolution are vetoed to ensure a good measurement of $E_{\mathrm{T}}^{\mathrm{miss}}$.

Additionally, to ensure the full efficiency of the 2 b-jet + 2 jet triggers, at least four jets with p_{T} above (80, 80, 40, 40) GeV are required, among which two b-jets are tagged offline and matched to online b-jets.

6.5.2 Anti-QCD requirements

To select signal events in 0 lepton final state, events with loose electrons, muons or hadronic taus are vetoed and $E_{\mathrm{T}}^{\mathrm{miss}}$ greater than 160 GeV is required in all analysis regions. In order to reduce the contamination of QCD multijet events, which may enter the SRs due to the unbalanced momentum in the transverse plane caused by mismeasurement or miscalibration of jet p_{T}, a set of requirements aiming to reject events with $E_{\mathrm{T}}^{\mathrm{miss}}$ originating from mismeasured jets is applied after the baseline selection. These 'anti-QCD' requirements include:

- Object-based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance > 8. Events with large fake $E_{\mathrm{T}}^{\mathrm{miss}}$ originating from mismeasurements, resolution and efficiency effects can be suppressed using object-based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance, which takes into consideration the estimated resolutions of objects contributing to $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction as well as their relative orientations. Furthermore, the $E_{\mathrm{T}}^{\mathrm{miss}}$ significance selection helps to reduce the pile-up contamination by accounting the probability for each jet in the event to originate from pile-up interactions [137].
- $\Delta\phi_{\min}(E_{\mathrm{T}}^{\mathrm{miss}}, \text{central jets}) > 0.5$. The minimal azimuthal angle between $E_{\mathrm{T}}^{\mathrm{miss}}$ and central jets can be used to veto multijet background where one jet is significantly mismeasured, resulting in large fake $E_{\mathrm{T}}^{\mathrm{miss}}$. Under such conditions, the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector will be highly collinear with the mismeasured jet, while for signals the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ and central jets are expected to be well-isolated.
- $\Delta\phi_{\min}(E_{\mathrm{T}}^{\mathrm{miss}}, \text{forward jets}) > 0.3$. Similar argumentation applies to forward jets, which are used in the computation of $\chi^2_{t\bar{t},\,\mathrm{had}}$, one of the main discriminants of the analysis.

The exact selection values are chosen by examining the distribution of these observables and removing regions in which poor MC modeling of data is observed.

6.5.3 Signal region selection

Following the baseline selection and the anti-QCD requirements, events are categorized into the resolved, boosted and highly boosted regimes, while the SR selections are optimized for each kinematic topology. The main discriminating variables used in the analysis as well as the complete SR definition are described below.

6.5.3.1 Leptonically decaying top quark reconstruction

Semi-leptonic $t\bar{t}(+b)$ and Wt-channel single top, two of the dominant background processes in the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search, may enter the SRs due to high $E_{\mathrm{T}}^{\mathrm{miss}}$ originating from $t \to bW \to bl\nu$ decay, with the neutrino escaping the detector and the charged lepton being either lost of misidentified as a jet. To distinguish between such backgrounds and the signal events, a pseudo top reconstruction is performed by associating $E_{\mathrm{T}}^{\mathrm{miss}}$ to the correct b-jet.

In practice, the leptonic top reconstruction is performed assuming that $E_{\mathrm{T}}^{\mathrm{miss}}$ carries the majority of the transverse momentum of the escaping neutrino and the lost/misidentified lepton, so that it corresponds exactly to the p_{T} of the leptonically decaying W boson. A four-vector with p_{T} and ϕ corresponding to the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector and its mass equal to the standard W mass is then built, while its η (or equivalently p_z) remains unknown. Choosing the x-axis to be in the direction of p_{T}^{W}, one can write:

$$\boldsymbol{p}_W = \left(\sqrt{(p_{\mathrm{T}}^W)^2 + (p_z^W)^2 + m_W^2}, p_{\mathrm{T}}^W, 0, p_z^W \right), \tag{6.5.1}$$

$$\boldsymbol{p}_b = \left(\sqrt{(p_{\mathrm{T}}^b)^2 + (p_z^b)^2 + m_b^2}, p_{\mathrm{T}}^b \cdot \cos(\phi_W - \phi_b), p_{\mathrm{T}}^b \cdot \sin(\phi_W - \phi_b), p_z^b \right). \tag{6.5.2}$$

The invariant mass of the top, built from the vector sum of W and b, is given by:

$$m_t^2 = (\boldsymbol{p}_W + \boldsymbol{p}_b)^2. \tag{6.5.3}$$

With $m_b \sim 0$, after some mathematical simplifications and rearrangements:

$$\sqrt{1 + \left(\frac{m_W}{p_{\mathrm{T}}^W \cdot \cosh \eta_W} \right)^2} \cdot \cosh \eta_W \cdot \cosh \eta_b - \sinh \eta_W \cdot \sinh \eta_b = \frac{m_t^2 - m_W^2}{2 p_{\mathrm{T}}^W p_{\mathrm{T}}^b} + \cos(\phi_W - \phi_b), \tag{6.5.4}$$

where η_W is unknown. Given that $E_{\mathrm{T}}^{\mathrm{miss}} > 160$ GeV is required as a part of the baseline selection, $m_W \sim 80$ GeV $\ll E_{\mathrm{T}}^{\mathrm{miss}} \cdot \cosh \eta_W \sim p_{\mathrm{T}}^W \cdot \cosh \eta_W$ can be assumed, i.e.:

$$\sqrt{1 + \left(\frac{m_W}{p_{\mathrm{T}}^W \cdot \cosh \eta_W} \right)^2} \sim 1. \tag{6.5.5}$$

Applying the following general relation:

$$\cosh \eta_W \cdot \cosh \eta_b - \sinh \eta_W \cdot \sinh \eta_b = \cosh(\eta_W - \eta_b), \tag{6.5.6}$$

Equation 6.5.4 can then be simplified to the following form:

$$\cosh(\eta_W - \eta_b) = \frac{m_t^2 - m_W^2}{2 p_{\mathrm{T}}^W p_{\mathrm{T}}^b} + \cos(\phi_W - \phi_b). \tag{6.5.7}$$

By definition, $\cosh(x) \geq 1$. Therefore, the discriminating observable $\cosh_{\mathrm{max}}$ can be defined as:

$$\cosh_{\mathrm{max}} = \max\{\cosh(\eta_W - \eta_{b_1}), \cosh(\eta_W - \eta_{b_2})\}, \tag{6.5.8}$$

while $\cosh_{\mathrm{max}} < 1$ indicates failure of the leptonic top reconstruction.

$\cosh_{\mathrm{max}}$ allows a good separation between signal and background if the two following conditions are fulfilled:

- the background event contains a leptonic top;
- one of the selected b-jets comes from the leptonic top.

Although the majority of $t\bar{t}$, $t\bar{t}+b$ and single top backgrounds in the SRs comply with these criteria, exceptions may take place:

- $t\bar{t}$ with extra b-jet(s) which comes from gluon splitting. It is possible that the b-jet originating from the leptonic top decay is not among the leading and the sub-leading b-jets which are used for $\cosh_{\text{max}}$ calculation. Most of the $t\bar{t}+b$ events entering the SRs fall into this category.
- $t\bar{t}$ with a c-jet originating from hadronic W decay or gluon splitting misidentified as a b-jet. This is expected to be extremely rare due to the very tight b-jet definition applied in this analysis. However, such backgrounds are still not entirely negligible in some of the SRs.
- Wt-channel single top with a neutrino coming from the W produced in association with the top (and not the one originating from the top decay) and an extra b-jet. In this case, $\cosh_{\text{max}}$ is expected to have no rejection power. All single top events entering the signal regions fall into this category.

In addition to these, $t\bar{t}$ events where the measured $E_{\text{T}}^{\text{miss}}$ differs greatly from p_{T}^{W} due to resolution effects may also survive the $\cosh_{\text{max}}$ selection, forming the lower Gaussian tail in Figure 6.3.1. Most of the $t\bar{t}$ background enters the SRs due to this reason. Considering the large cross section of $t\bar{t}$ production, the contribution from these events is actually significant.

The $\cosh_{\text{max}}$ distributions in the resolved, boosted and highly boosted regimes with all SR selections applied except for the $\cosh_{\text{max}}$ cut are shown in Figure 6.5.1.

The expected significance Z_n of observing $\langle n\rangle = s+b$ events given a prediction of $b\pm\sigma$ background is calculated in each bin in Figure 6.5.1, defined as:

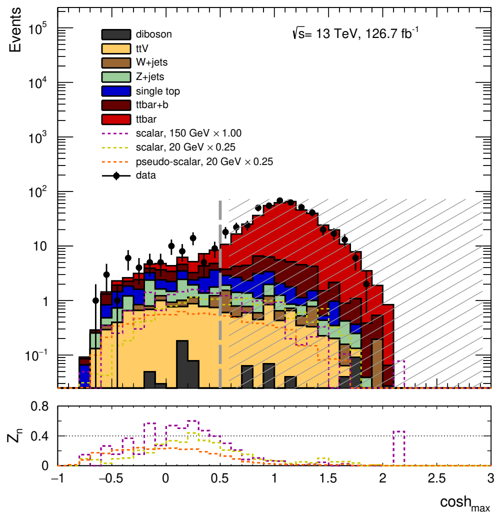
$$Z_n = \begin{cases} +\sqrt{2\left(n\ln\left[\frac{n(b+\sigma^2)}{b^2+n\sigma^2}\right] - \frac{b^2}{\sigma^2}\ln\left[1+\frac{\sigma^2(n-b)}{b(b+\sigma^2)}\right]\right)} & \text{, if} \quad n\geq b; \\ -\sqrt{2\left(n\ln\left[\frac{n(b+\sigma^2)}{b^2+n\sigma^2}\right] - \frac{b^2}{\sigma^2}\ln\left[1+\frac{\sigma^2(n-b)}{b(b+\sigma^2)}\right]\right)} & \text{, if} \quad n< b. \end{cases} \tag{6.5.9}$$

A 15% flat systematic uncertainty ($\sigma = 0.15\cdot b$) is applied for Z_n estimation, while the exact selection value of $\cosh_{\text{max}}$ in each regime is determined by excluding the region with low expected Z_n.

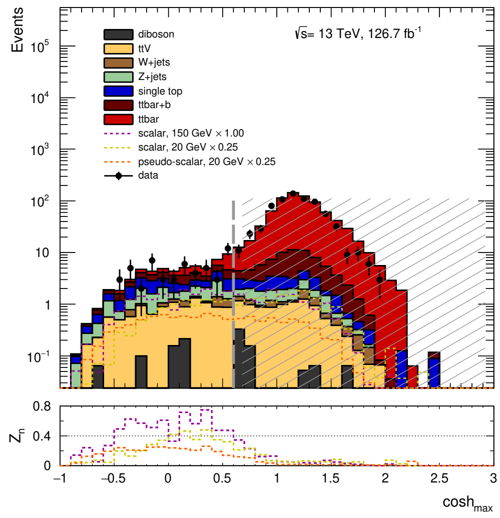
6.5.3.2 Hadronically decaying top quark pair reconstruction

Aside from the leptonic top reconstruction, the hadronic top pair reconstruction provides another perspective on SR design. As the various background components, except $t\bar{t}+Z$, do not have hadronically decaying $t\bar{t}$ in the final state, by reconstructing the top quark pair one can enhance the rejection to backgrounds with no tops, e.g. Z+jets and W+jets, which cannot be effectively suppressed by the leptonic top reconstruction cut.

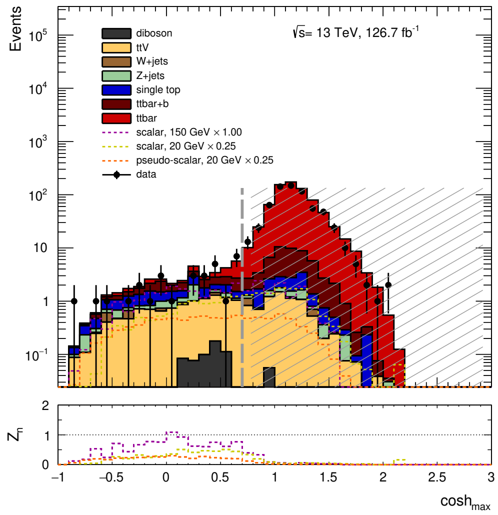
To evaluate the probability of events to contain two hadronically decaying tops, $\chi^2_{t\bar{t},\,\text{had}}$ is defined as:



(a)



(b)



(c)

Figure 6.5.1: $\cosh_{\mathrm{max}}$ distributions in (a) the resolved regime, (b) the boosted regime and (c) the highly boosted regime with all SR selections applied except for the $\cosh_{\mathrm{max}}$ cut. The vertical dashed lines represent the selection value of $\cosh_{\mathrm{max}}$ while the hatching lines indicate the direction of the cut. The distributions are shown before the profiling.

$$\begin{aligned}\chi^2_{t\bar{t},\,\mathrm{had}} = & \left(\frac{m_{W_1} - m_{W_{\mathrm{ref}}}}{\sigma_{m_W}}\right)^2 \\ & + \left(\frac{(m_{t_1} - m_{W_1}) - (m_{t_{\mathrm{ref}}} - m_{W_{\mathrm{ref}}})}{\sigma_{m_t - m_W}}\right)^2 \\ & + \left(\frac{(m_{t_2} - m_{W_2}) - (m_{t_{\mathrm{ref}}} - m_{W_{\mathrm{ref}}})}{\sigma_{m_t - m_W}}\right)^2 .\end{aligned} \tag{6.5.10}$$

Up to five (two) leading p_{T} central (forward) jets, as well as the leading and the sub-leading b-jets are considered for hadronic $t\bar{t}$ reconstruction. W_1 is built from two non b-tagged jets, while t_1 contains the two jets from W_1 and one of the selected b-jets. For W_2, it is assumed that its p_{T}, η and ϕ corresponds roughly to one single jet to which the mass of W boson is attributed, so that t_2 only contains one non b-tagged jet and the other selected b-jet. This assumption is necessary to obtain a sufficiently high reconstruction efficiency as the fraction of events with at least six jets is about 50% only. The values of $m_{W_{\mathrm{ref}}}$, σ_{m_W}, $m_{t_{\mathrm{ref}}}$ and $\sigma_{m_t - m_W}$ are taken as [189]:

- $m_{W_{\mathrm{ref}}} = 80.51$ GeV, $\sigma_{m_W} = 12.07$ GeV,
- $m_{t_{\mathrm{ref}}} - m_{W_{\mathrm{ref}}} = 85.17$ GeV, $\sigma_{m_t - m_W} = 16.05$ GeV.

The first term of $\chi^2_{t\bar{t},\,\mathrm{had}}$ is the constraint from the hadronically decaying W boson, while $m_{W_{\mathrm{ref}}}$ and σ_{m_W} represent, respectively, the mean and the standard deviation of the invariant mass distribution related to the hadronic W. The second and third term corresponds to the two hadronically decaying top quarks. As m_{W_1} (m_{W_2}) and m_{t_1} (m_{t_2}) are strongly correlated, the W boson is subtracted from the top mass to decouple these two terms from the first one.

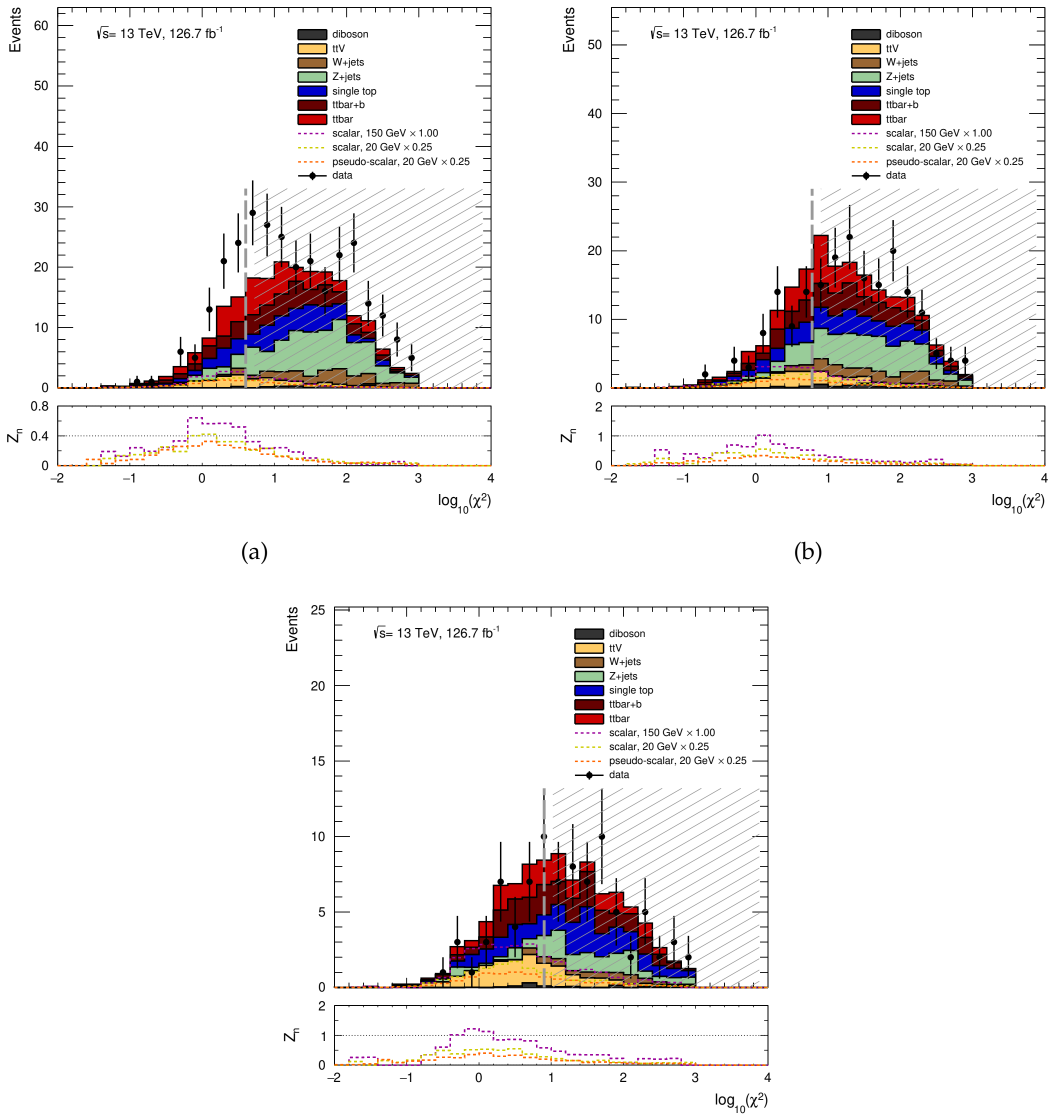
By minimizing $\chi^2_{t\bar{t},\,\mathrm{had}}$ over all combinations of jets, a clear separation between $t\bar{t}$ and W/Z+jets can be observed, as shown in Figure 6.5.2. The distributions are plotted with all SR selections applied except for the $\chi^2_{t\bar{t},\,\mathrm{had}}$ cut. As expected, very little discrimination is observed for $t\bar{t} + V$, which usually includes two hadronically decaying top quarks. Similar to the leptonic top reconstruction cut, the exact selection value of $\chi^2_{t\bar{t},\,\mathrm{had}}$ in each regime is determined by excluding the region with low expected Z_n.

Moreover, in fully-hadronic $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ final state, the ratio between the p_{T} of the reconstructed $t\bar{t}$ system and $E_{\mathrm{T}}^{\mathrm{miss}}$ should be ~ 1 due to the momentum conservation in the transverse plane. Since only the leading five jets are used in $\chi^2_{t\bar{t},\,\mathrm{had}}$ minimization, the transverse momentum of one of the jets originating from the $t\bar{t}$ decay fails to be counted in $p_{\mathrm{T}}^{t\bar{t}}$, leading to underestimation of $p_{\mathrm{T}}^{t\bar{t}}/E_{\mathrm{T}}^{\mathrm{miss}}$. This does not apply to most background processes, where the minimal $\chi^2_{t\bar{t},\,\mathrm{had}}$ is achieved with a random combination of the input jets. Hence, a lower mean of $p_{\mathrm{T}}^{t\bar{t}}/E_{\mathrm{T}}^{\mathrm{miss}}$ is expected for signal events and can be used to further reduce the backgrounds surviving the $\chi^2_{t\bar{t},\,\mathrm{had}}$ selection. In the highly boosted regime, the minimum ΔR between the highest mass large-R jet and b-jets is required to be less than 1.2, ensuring a b-jet close enough to the boosted top quark.

6.5.3.3 Additional cuts

Apart from the top reconstruction requirements, additional cuts are introduced to optimize the signal event selection, including:

- Object-based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance > 10.
- $\Delta R\,(b_1, b_2) > 1.2$. A large fraction of the Z+jets background events in the SRs contain extra b-jets from QCD gluon splitting, where the angular separation between the leading and the sub-leading b-jets is predicted to be small. The selection on $\Delta R\,(b_1, b_2)$ effectively reduce the contribution of such process.
- $\Delta\phi_{\mathrm{min}}(E_{\mathrm{T}}^{\mathrm{miss}}$, central jets$) > 1.0$. This is only applied in resolved regime to further reject the multijet contamination.



(a) (b) (c)

Figure 6.5.2: $\log_{10}(\chi^2_{t\bar{t},\,\text{had}})$ distributions in (a) the resolved regime, (b) the boosted regime and (c) the highly boosted regime with all SR selections applied except for the $\chi^2_{t\bar{t},\,\text{had}}$ cut. The vertical dashed lines represent the selection value of $\chi^2_{t\bar{t},\,\text{had}}$ while the hatching lines indicate the direction of the cut. The distributions are shown before the profiling.

The full list of the signal region event selections used in the $E_{\text{T}}^{\text{miss}} + t\bar{t}$ (fully-hadronic) search can be found in Table 6.5.1, while columns from left to right correspond to the selection criterion of the resolved SR, the boosted SR and the highly boosted SR.

Figure 6.5.3 shows the $E_{\text{T}}^{\text{miss}}$ distributions for data and MC prediction in the SRs after the full event selection and before the profiling. The deviations will be corrected by applying the background normalization factors extracted from the CRs, as defined in Section 6.6.

Resolved regime	Boosted regime	Highly boosted regime
	baseline selection	
	0 loose electron/muon, 0 hadronic tau	
	anti-QCD requirements	
	$E_{\mathrm{T}}^{\mathrm{miss}} > 160$ GeV	
	object-based $E_{\mathrm{T}}^{\mathrm{miss}}$ sig. > 10	
	$\Delta R\,(b_1, b_2) > 1.2$	
0 large-R jet	≥ 1 large-R jet	≥ 1 large-R jet
$\Delta\phi_{\mathrm{min}}(E_{\mathrm{T}}^{\mathrm{miss}}$, central jets) > 1.0	$m_J < 130$ GeV	$m_J \geq 130$ GeV
$\cosh_{\mathrm{max}} < 0.5$	$\cosh_{\mathrm{max}} < 0.6$	$\cosh_{\mathrm{max}} < 0.7$
$\chi^2_{t\bar{t},\,\mathrm{had}} < 4$ $0.7 < p_{\mathrm{T}}^{t\bar{t}}/E_{\mathrm{T}}^{\mathrm{miss}} < 1.2$	$\chi^2_{t\bar{t},\,\mathrm{had}} < 6$ $0.5 < p_{\mathrm{T}}^{t\bar{t}}/E_{\mathrm{T}}^{\mathrm{miss}} < 1.2$	$\chi^2_{t\bar{t},\,\mathrm{had}} < 8$ $0.5 < p_{\mathrm{T}}^{t\bar{t}}/E_{\mathrm{T}}^{\mathrm{miss}} < 1.2$ ΔR_{min}(large-R jet, b-jets) < 1.2

Table 6.5.1: List of the signal region event selections used in the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search.

6.6 Background estimation

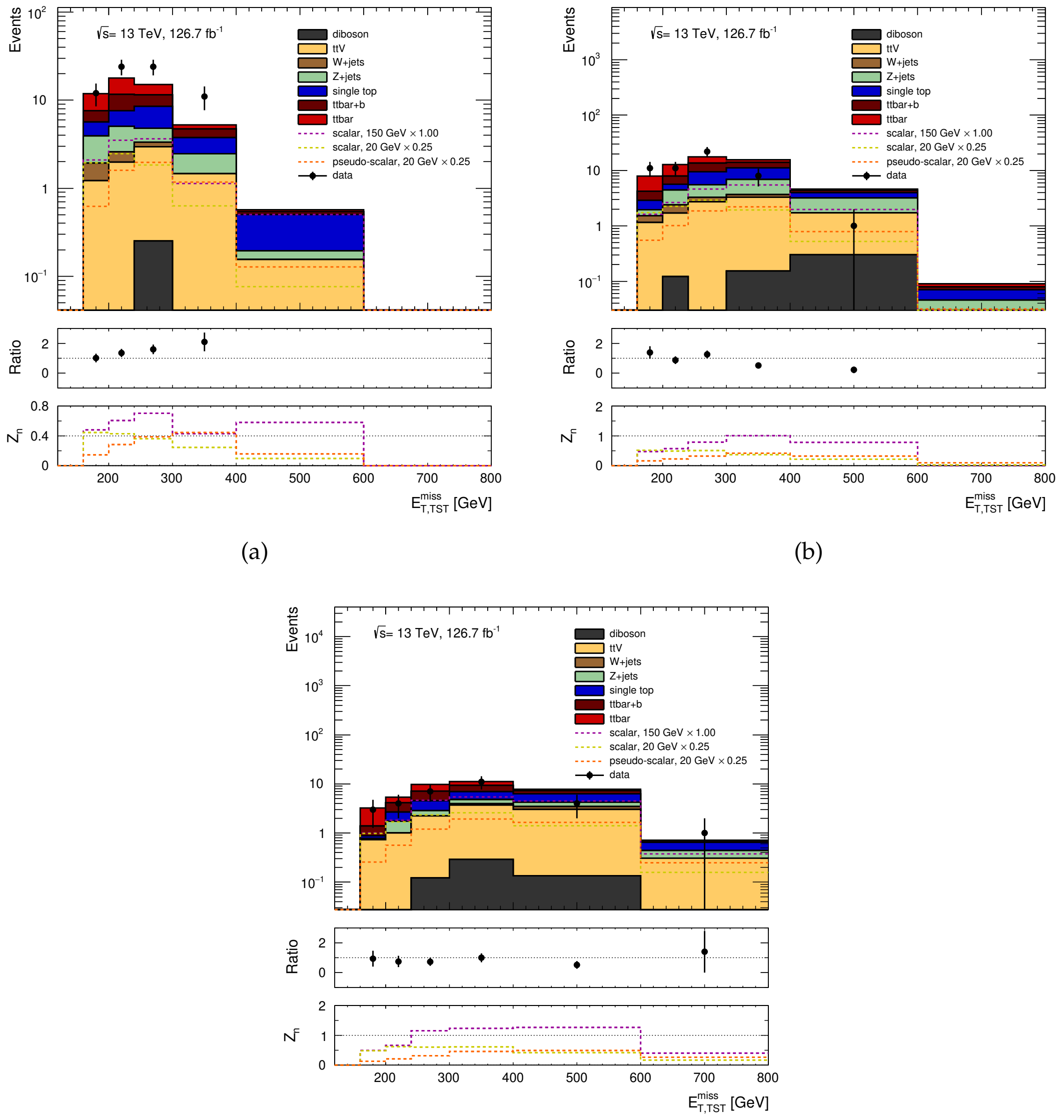
The major background processes of the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search are $t\bar{t}$, $t\bar{t} + b$, single top and Z+jets, while control regions (CRs) are designed to estimate the contribution of these backgrounds by extracting their normalization using high purity data sample in regions enriched in the corresponding processes. Similar to the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search, the baseline selection is applied to all CRs, which are distinguished by the multiplicity of leptons, ensuring the quality of analyzed data and MC simulated events.

As discussed in Section 6.3.2, most of the $t\bar{t}$, $t\bar{t} + b$ and single top backgrounds entering the SRs contain $W \to l\nu$ decay, where the lepton is either lost or misidentified as a jet, and the 1 lepton CRs which select events with exactly one isolated lepton are used to constrain these processes by extrapolating the kinematic properties of this lepton. Hence, the treatment of the well-reconstructed lepton in the 1 lepton CRs should be decided based on the behavior of the misreconstructed lepton in the SRs, so that the extrapolation of the event topology from control to signal regions are kept minimal. Typically, the lepton in the control regions can be treated as either a jet or invisible, which correspond to the situation where the lepton is either misidentified as a jet or totally lost in the signal regions, respectively.

For the case where the lepton is misidentified as a jet, the dominant contribution comes from $W \to \tau\nu \to j\nu\nu$. As a result, the fraction of events in the SRs with ≥ 1 truth-level tau exclusively associated to reconstructed jets[2] provides a good estimation of the fraction of events in which the lepton is misidentified as a jet.

Table 6.6.1 documents the percentage of events in each SR with ≥ 1 truth-level tau lepton. Given the observation that the majority of the background events in the SRs are free from tau contamination, the leptons are treated as invisible in the 1 lepton CRs, with their p_{T} accounted as part of the missing transverse momentum.

[2]The number of truth taus is calculated by counting the tau lepton with $p_{\mathrm{T}} > 5$ GeV within $\Delta R = 0.3$ of any reconstructed small-R jet with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 4.5$.



(a)

(b)

(c)

Figure 6.5.3: $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in the (a) resolved, (b) boosted and (c) highly boosted signal regions. The distributions are shown before the profiling.

Process	Resolved regime	Boosted regime	Highly boosted regime
$t\bar{t}$	45.4%	38.4%	32.4%
$t\bar{t}+b$	40.1%	38.3%	39.0%
Single top	46.3%	48.0%	47.2%

Table 6.6.1: Fraction of events in the SRs with ≥ 1 truth-level tau lepton matched to reconstructed small-R jets for the reducible backgrounds with top quarks.

The 2 lepton CRs are designed to constrain the $Z(\rightarrow \nu\nu)$+jets background by studying the $Z(\rightarrow ll)$+jets events, thus the leptons are treated as invisible in order to mimic the kinematic behavior of the neutrinos.

Therefore, $E_{\mathrm{T,\,no\,lepton}}^{\mathrm{miss}}$, constructed by subtracting the lepton components from $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation, is used in the control region selections. All other $E_{\mathrm{T}}^{\mathrm{miss}}$ related variables, e.g. $E_{\mathrm{T}}^{\mathrm{miss}}$ significance, $\cosh_{\mathrm{max}}$ and $\Delta\phi_{\mathrm{min}}(E_{\mathrm{T}}^{\mathrm{miss}}, \text{jets})$, are recalculated in the same way. The full event selections of the 1 lepton CRs and the 2 lepton CRs are presented in Section 6.6.1 and Section 6.6.2.

Additionally, validations regions (VRs) are defined to check the validity of the control-to-signal region extrapolation. The selection criteria for the VRs are kept orthogonal to the SR selections, featuring a low signal expectation. This is achieved by inverting either the $\cosh_{\mathrm{max}}$ cut or the $\chi^2_{t\bar{t},\,\mathrm{had}}$ cut, with more details to be found in Section 6.6.3.

All background processes are modeled by MC simulation and included in the combined profile likelihood fit (Section 6.8.1), except for the QCD multijet production. A dedicated data-driven approach, namely the 'Rebalance and Smear' method, is applied to evaluate the multijet contamination in the SRs, as documented in Section 6.6.4.

6.6.1 1 lepton control regions

The 1 lepton control regions are defined to constrain $t\bar{t}$, $t\bar{t}+b$ and single top using events containing exactly one tight electron or muon and no additional loose leptons. To avoid any bias on modeling in different kinematic phase space, the event topologies in the SRs and the 1 lepton CRs are kept as similar as possible, with the only extrapolation applied to the lepton. Despite the fact that the multijet contamination is negligible in the 1 lepton CRs, the anti-QCD requirements are included as part of the selection criterion, with all $E_{\mathrm{T}}^{\mathrm{miss}}$ related variables recalculated with $E_{\mathrm{T,\,no\,lepton}}^{\mathrm{miss}}$:

- Object-based $E_{\mathrm{T,\,no\,lepton}}^{\mathrm{miss}}$ significance > 8;
- $\Delta\phi_{\mathrm{min}}(E_{\mathrm{T,\,no\,lepton}}^{\mathrm{miss}}, \text{central jets}) > 0.5$;
- $\Delta\phi_{\mathrm{min}}(E_{\mathrm{T,\,no\,lepton}}^{\mathrm{miss}}, \text{forward jets}) > 0.3$;

However, since very tight event selection is applied in the SRs, applying exactly the same selection in the 1 lepton CRs would result in very low event yield, which then harms the constraining power of the CRs. As a trade-off, part of the signal region selections, including the cuts on $\cosh_{\mathrm{max}}$ and $\chi^2_{t\bar{t},\,\mathrm{had}}$, are loosened in the 1 lepton CRs. The validation of these extra extrapolations is performed in Section 6.6.1.1.

Three types of control regions are designed, each targeting one of the major backgrounds to be constrained.

In $t\bar{t}$ events, the two b-jets are expected to both originate from top decays, which cannot be the case for single top and is usually not the case for $t\bar{t}+b$ surviving the $\cosh_{\mathrm{max}}^{\mathrm{no\,lepton}}$ cut. $t\bar{t}$ can be thus separated from the other backgrounds by reconstructing the semi-leptonically decaying $t\bar{t}$ system using a χ^2 based method [190], which takes into account the kinematic properties of $E_{\mathrm{T}}^{\mathrm{miss}}$, lepton, jets and the b-tagging information.

The $\chi^2_{t\bar{t},\,\mathrm{lep}}$ minimization procedure is based on constraints on the top quark and W boson masses and the kinematic topology of $t\bar{t}$ events. Same jet collection considered for $\chi^2_{t\bar{t},\,\mathrm{had}}$ calculation are used to build the semi-leptonic $t\bar{t}$ system, i.e. up to five (two) leading p_{T} central (forward) jets, as well as the leading and the sub-leading b-jets. The permutation which leads to the minimal value of $\chi^2_{t\bar{t},\,\mathrm{lep}}$ is chosen, with $\chi^2_{t\bar{t},\,\mathrm{lep}}$ taking the form of:

$$\begin{aligned}\chi^2_{t\bar{t},\,\mathrm{lep}} = & \left(\frac{m_{W_{\mathrm{had}}} - m_{W_{\mathrm{ref}}}}{\sigma_{m_W}}\right)^2 \\ & + \left(\frac{(m_{t_{\mathrm{had}}} - m_{W_{\mathrm{had}}}) - (m_{t_{\mathrm{had,\,ref}}} - m_{W_{\mathrm{ref}}})}{\sigma_{m_{t_{\mathrm{had}}} - m_W}}\right)^2 \\ & + \left(\frac{m_{t_{\mathrm{lep}}} - m_{t_{\mathrm{lep,\,ref}}}}{\sigma_{m_{t_{\mathrm{lep}}}}}\right)^2,\end{aligned} \tag{6.6.1}$$

where t_{had} and t_{lep} refer to the hadronically and leptonically decaying top quarks, respectively. The values of $m_{W_{\mathrm{ref}}}$, σ_{m_W}, $m_{t_{\mathrm{had,\,ref}}}$, $\sigma_{m_{t_{\mathrm{had}}} - m_W}$, $m_{t_{\mathrm{lep,\,ref}}}$ and $\sigma_{m_{t_{\mathrm{lep}}}}$ are similar to the ones used in $\chi^2_{t\bar{t},\,\mathrm{had}}$ minimization and are obtained using MC simulation following the prescription described in Ref. [190].

The first term of $\chi^2_{t\bar{t},\,\mathrm{lep}}$ is the constraint from the hadronically decaying W boson, while $m_{W_{\mathrm{ref}}}$ and σ_{m_W} represent, respectively, the mean and the standard deviation of the invariant mass distribution related to the hadronic W. The second term corresponds to the hadronically decaying top quark. As $m_{W_{\mathrm{had}}}$ and $m_{t_{\mathrm{had}}}$ are strongly correlated, the W boson is subtracted from the top mass to decouple this term from the first one. The third term corresponds to the leptonically decaying top quark, which includes the information of the 'b-l-ν' system, $m_{t_{\mathrm{lep}}}$, as well as the mean and the standard deviation of the invariant mass distribution of the leptonic top.

The reconstruction of the neutrino four-momentum, required in the third term of $\chi^2_{t\bar{t},\,\mathrm{lep}}$, is performed beforehand by applying a W mass constraint on the invariant mass of the 'l-ν' system for the z component, while the x and y components of the neutrino momentum are set to the corresponding components of the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector. If the resulting quadratic equation does not have a real solution, the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector is rotated including a parameter in the expression of the 'l-ν' system that is minimized in order to find a real solution. In case of two real solutions of the quadratic equation, both solutions are tested in the reconstruction and the one which minimizes the $\chi^2_{t\bar{t},\,\mathrm{lep}}$ value is chosen.

$\chi^2_{t\bar{t},\,\mathrm{lep}}$ provides very good separation between $t\bar{t}$ and other backgrounds, while the multiplicity of extra offline b-jets serves as the best observable to discriminate between $t\bar{t} + b$ and single top processes. By requesting ≥ 1 extra offline b-jet, i.e. two b-jets tagged online and offline plus one extra b-jet tagged offline, $t\bar{t} + b$ events can be selected with high purity. Additionally, a tight $\cosh_{\mathrm{max}}^{\mathrm{no\ lepton}}$ cut is applied in the single top CRs to reject $t\bar{t}$ events failing the $\chi^2_{t\bar{t},\,\mathrm{lep}}$ reconstruction due to resolution effects.

Table 6.6.2 presents the full list of the 1 lepton control region event selections used in the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search.

Figure 6.6.1, Figure 6.6.2 and Figure 6.6.3 compare the data and MC prediction in the $t\bar{t}$, $t\bar{t} + b$ and single top enriched control regions before the profiling, respectively. In the $t\bar{t}$ enriched CRs the observed and the expected distributions of $E_{\mathrm{T,\,no\ lepton}}^{\mathrm{miss}}$ agree very well, while slight underprediction (overprediction) in total event yield can be observed in the $t\bar{t} + b$ (single top) enriched CRs.

6.6.1.1 1 lepton control region validation

Since part of the signal region selections are loosened in the 1 lepton CRs, an extra step is needed to ensure the associated extrapolation is within the tolerance. The validation is

Resolved regime	Boosted regime	Highly boosted regime
baseline selection		
1 tight electron/muon, 0 loose electron/muon, 0 hadronic tau		
anti-QCD requirements		
$E^{\text{miss}}_{\text{T, no lepton}} > 160$ GeV		
object-based $E^{\text{miss}}_{\text{T, no lepton}}$ sig. > 10		
$\Delta R\,(b_1, b_2) > 1.2$		
0 large-R jet	≥ 1 large-R jet	≥ 1 large-R jet
$\Delta\phi_{\text{min}}(E^{\text{miss}}_{\text{T, no lepton}}$, central jets) > 1.0	$m_J < 130$ GeV	$m_J \geq 130$ GeV
$\cosh^{\text{no lepton}}_{\text{max}} < 0.9$	$\cosh^{\text{no lepton}}_{\text{max}} < 0.95$	$\cosh^{\text{no lepton}}_{\text{max}} < 1.0$
$\chi^2_{t\bar{t},\text{ had}} < 10$ $0.7 < p^{t\bar{t}}_{\text{T}}/E^{\text{miss}}_{\text{T}} < 1.2$	$\chi^2_{t\bar{t},\text{ had}} < 20$ $0.5 < p^{t\bar{t}}_{\text{T}}/E^{\text{miss}}_{\text{T}} < 1.2$	$\chi^2_{t\bar{t},\text{ had}} < 40$ $0.5 < p^{t\bar{t}}_{\text{T}}/E^{\text{miss}}_{\text{T}} < 1.2$ ΔR_{min}(large-R jet, b-jets) < 1.2
$t\bar{t}$ control regions		
$\chi^2_{t\bar{t},\text{ lep}} < 6$		
$t\bar{t}$+b control regions		
$\chi^2_{t\bar{t},\text{ lep}} \geq 6$		
≥ 1 extra offline b-jet		
single top control regions		
$\chi^2_{t\bar{t},\text{ lep}} \geq 30$		
0 extra offline b-jet		
$\cosh^{\text{no lepton}}_{\text{max}} < 0.5$	$\cosh^{\text{no lepton}}_{\text{max}} < 0.6$	$\cosh^{\text{no lepton}}_{\text{max}} < 0.7$

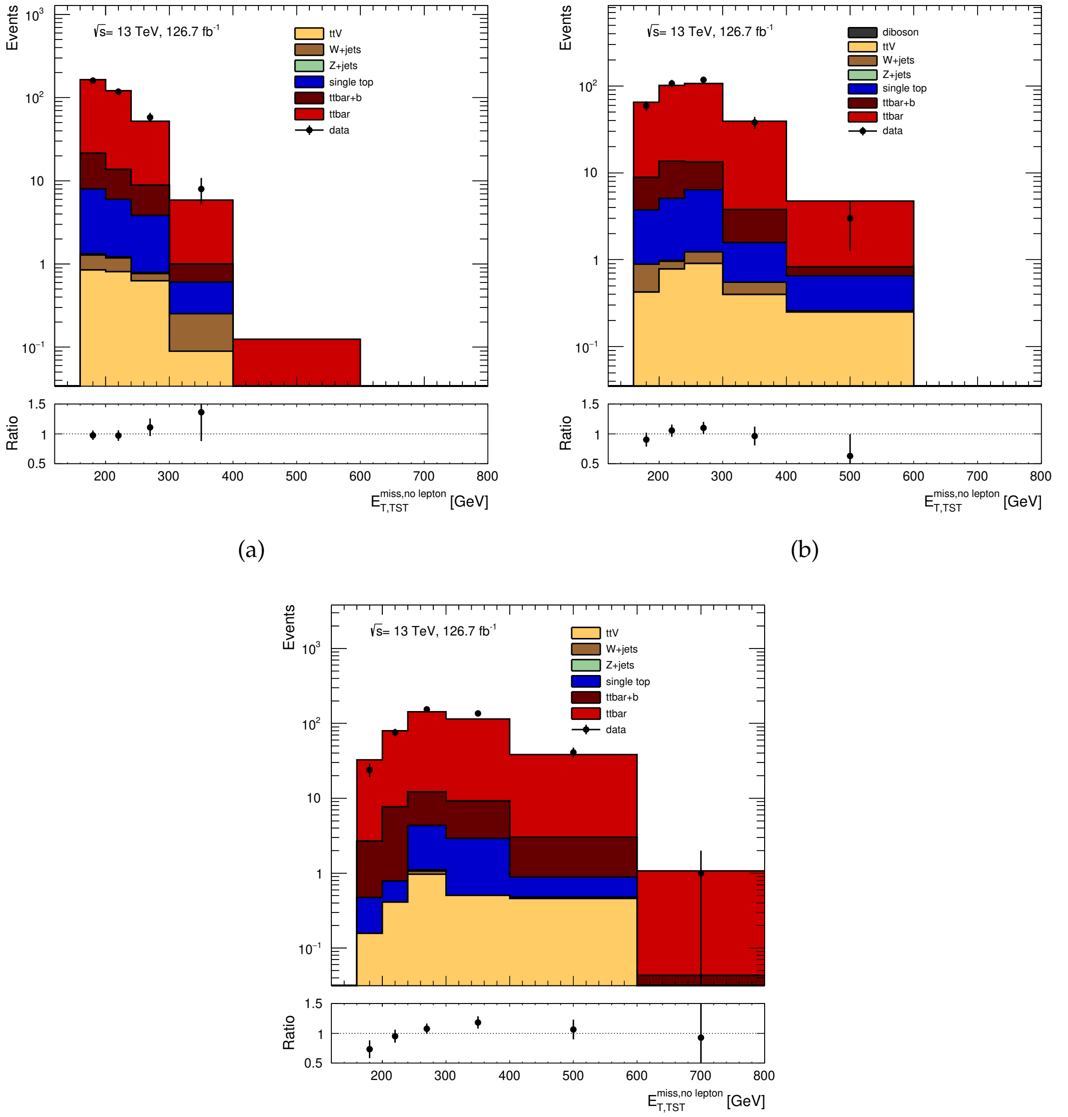
Table 6.6.2: List of the 1 lepton control region event selections used in the $E^{\text{miss}}_{\text{T}} + t\bar{t}$ (fully-hadronic) search.

performed by comparing the observables involving $E^{\text{miss}}_{\text{T}}$ in the SRs to the corresponding 'no lepton' observables in the 1 lepton CRs.

Figure 6.6.4 shows the shape of $\cosh_{\text{max}}$ in 0 lepton final state and $\cosh^{\text{no lepton}}_{\text{max}}$ in 1 lepton final state for $t\bar{t}$, $t\bar{t} + b$ and single top processes. The shape comparisons between $E^{\text{miss}}_{\text{T}}$ and $E^{\text{miss}}_{\text{T, no lepton}}$ are given in Figure 6.6.5. For $t\bar{t}$ and $t\bar{t} + b$, an overall good agreement is reached, while for single top, some deviations between the 0 lepton and the 1 lepton distributions can be observed, mainly due to the significant fraction of hadronic taus in single top events entering the SRs (see Table 6.6.1).

As mentioned in Section 6.6, for events with hadronic tau decays, $\tau \to j\nu$, only part of the transverse momentum of the tau lepton (p^{ν}_{T}) contributes to $E^{\text{miss}}_{\text{T}}$ computation. Hence, treating the lepton as fully invisible leads to an overestimation of the contribution from lepton p_{T} in $E^{\text{miss}}_{\text{T, no lepton}}$, therefore a larger extrapolation from the CRs to the SRs. This is in agreement with the shape comparison in Figure 6.6.4 and Figure 6.6.5, given that the overestimation in $E^{\text{miss}}_{\text{T, no lepton}}$ corresponds to an underestimation in $\cosh^{\text{no lepton}}_{\text{max}}$.

To conclude, the minimal extrapolation requirement for the 1 lepton CR design is generally fulfilled. The deviations in shape between the 0 lepton and the 1 lepton variables originate from the choice of treatment of the lepton and cannot be reduced without adding extra layers of complexity to the analysis.



(a) (b)

(c)

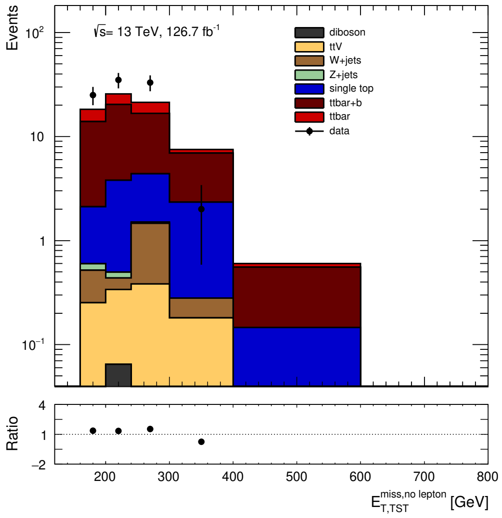
Figure 6.6.1: $E_{\text{T, no lepton}}^{\text{miss}}$ distributions in the (a) resolved, (b) boosted and (c) highly boosted $t\bar{t}$ enriched control regions. The distributions are shown before the profiling.

6.6.2 2 lepton control regions

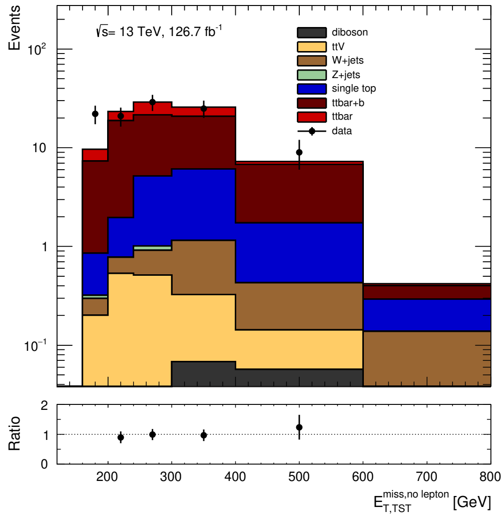
The 2 lepton control regions are defined to constrain $Z(\rightarrow \nu\nu)$+jets using $Z(\rightarrow ll)$+jets events containing exactly two tight electrons or muons and no additional loose leptons. Similar to the 1 lepton CRs, the anti-QCD requirements are applied to the 2 lepton CRs, with all $E_{\text{T}}^{\text{miss}}$ related variables recalculated with $E_{\text{T, no lepton}}^{\text{miss}}$:

- Object-based $E_{\text{T, no lepton}}^{\text{miss}}$ significance > 8;
- $\Delta\phi_{\text{min}}(E_{\text{T, no lepton}}^{\text{miss}}$, central jets) > 0.5;
- $\Delta\phi_{\text{min}}(E_{\text{T, no lepton}}^{\text{miss}}$, forward jets) > 0.3;

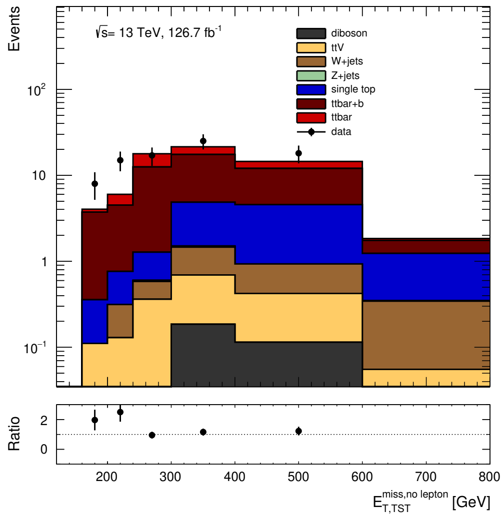
Ideally, the event topologies in the 2 lepton CRs and the SRs should be kept identical,



(a)



(b)

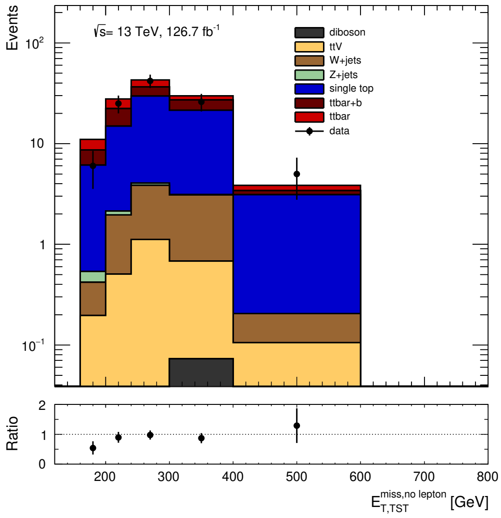


(c)

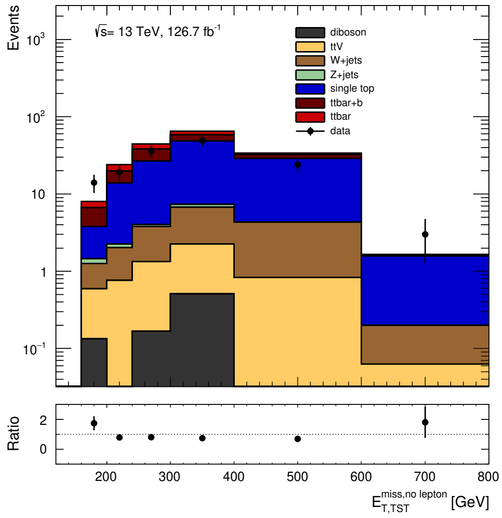
Figure 6.6.2: $E_{\mathrm{T,\,no\;lepton}}^{\mathrm{miss}}$ distributions in the (a) resolved, (b) boosted and (c) highly boosted $t\bar{t}+b$ enriched control regions. The distributions are shown before the profiling.

with the only extrapolation applied to the dilepton system, i.e. the Z boson. As described in Section 6.2.3, all events in the 2 lepton CRs are required to pass the 2 b-jet + 2 jet trigger selection, which allows to avoid extrapolating over the trigger performance but also greatly limits the expected event yield in all kinematic regimes. In order to reduces the impact of the MC statistical uncertainty, leptonic top reconstruction and hadronic top pair reconstruction cuts are dropped in the 2 lepton CRs, as well as all the additional cuts on $E_{\mathrm{T,\,no\;lepton}}^{\mathrm{miss}}$, $E_{\mathrm{T,\,no\;lepton}}^{\mathrm{miss}}$ significance, $\Delta R\,(b_1, b_2)$ and $\Delta\phi_{\mathrm{min}}(E_{\mathrm{T,\,no\;lepton}}^{\mathrm{miss}}$, central jets) applied in the SRs. These discarded selection criteria lead to a larger extrapolation with respect to the 1 lepton CRs on $E_{\mathrm{T}}^{\mathrm{miss}}$ related variables, which can be translated to an extrapolation of p_{T}^{Z}, as zero missing transverse energy is expected in $Z(\rightarrow ll)$+jets events. The validation of the extrapolation is performed in Section 6.6.2.1.

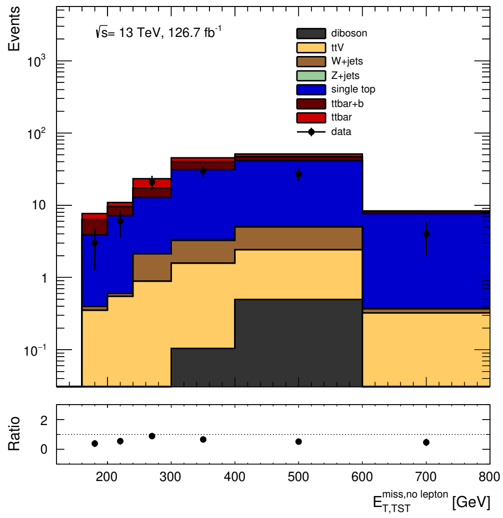
Considering that at least two tight b-jets are required by the triggers, dileptonically



(a)



(b)

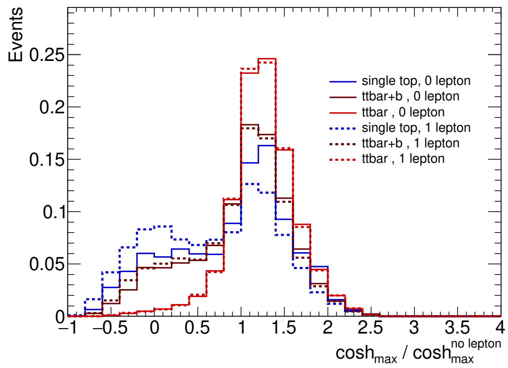


(c)

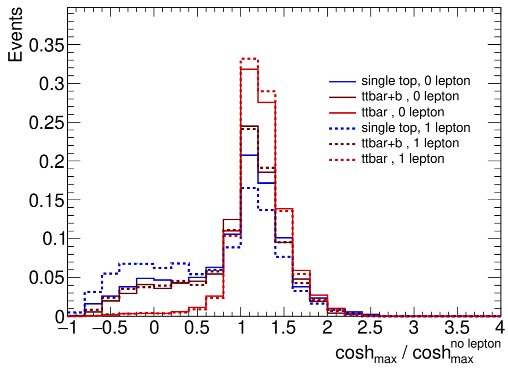
Figure 6.6.3: $E_{\mathrm{T,\,no\ lepton}}^{\mathrm{miss}}$ distributions in the (a) resolved, (b) boosted and (c) highly boosted single top enriched control regions. The distributions are shown before the profiling.

decaying $t\bar{t}$ events dominate the 2 lepton final state after the baseline selection and the anti-QCD requirements. Under such condition, an anti-correlation between $t\bar{t}$ and Z+jets normalization can be formed during the combined fit and thus harms the constraining power of the CRs. The 2 lepton control region event selections are therefore mainly designed to reduce the dileptonic $t\bar{t}$ contamination. The major discriminants used include:

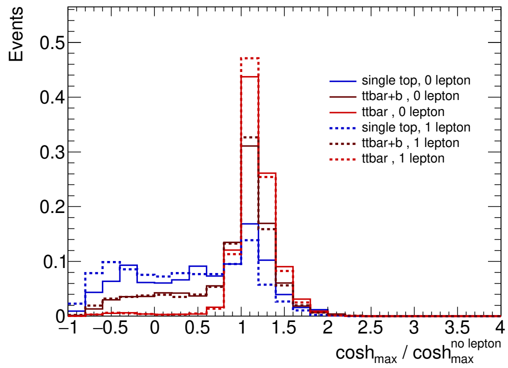
- m_{ll}. The invariant mass of the dilepton system is naturally one of the most straightforward variable which can be used to separate $Z(\rightarrow ll)$+jets from other backgrounds. A mass window cut is applied to m_{ll}, significantly increasing the purity of Z+jets background in the 2 lepton CRs.



(a)



(b)



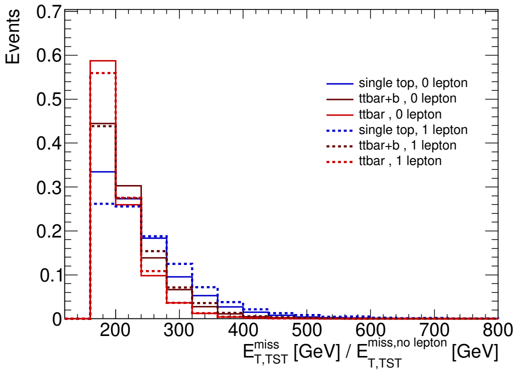
(c)

Figure 6.6.4: Shape of $\cosh_{\text{max}}$ in 0 lepton final state (solid lines) and $\cosh_{\text{max}}^{\text{no lepton}}$ in 1 lepton final state (dashed lines), normalized to unit area, for $t\bar{t}$ (red), $t\bar{t}+b$ (brown) and single top (blue) processes after the baseline selection and the anti-QCD requirements applied in (a) the resolved regime, (b) the boosted regime and (c) the highly boosted regime.

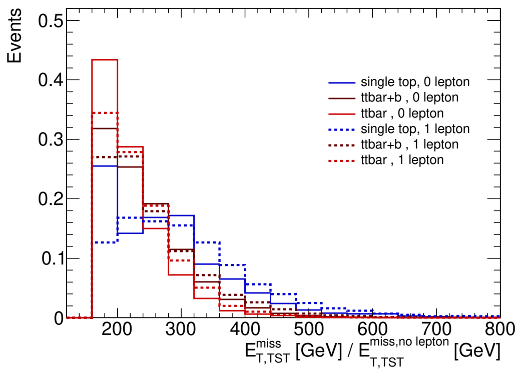
- Object-based $E_{\text{T}}^{\text{miss}}$ significance. Unlike dileptonic $t\bar{t}$, $Z(\rightarrow ll)$+jets events contain no escaping neutrinos, resulting in very low $E_{\text{T}}^{\text{miss}}$ within resolution. An upper cut on $E_{\text{T}}^{\text{miss}}$ significance is applied to reject the $t\bar{t}$ background. Note that compared to the object-based $E_{\text{T, no lepton}}^{\text{miss}}$ significance, the leptons are not treated as invisible in the calculation of $E_{\text{T}}^{\text{miss}}$ significance.
- p_{T}^{ll}. In Z+jets events, the vector sum of the leptons' p_{T}, which corresponds to p_{T}^{Z}, is expected to be close to $E_{\text{T, no lepton}}^{\text{miss}}$ in magnitude. On the other hand, in $t\bar{t}$ events where the two tops recoil against each other, the p_{T} of the two leptons tend to cancel with each other to a large extent, resulting in softer p_{T}^{ll}. The selection of high p_{T}^{ll} events is then useful to further suppress the $t\bar{t}$ contamination.

Table 6.6.3 presents the full list of the 2 lepton control region event selections used in the $E_{\text{T}}^{\text{miss}} + t\bar{t}$ (fully-hadronic) search.

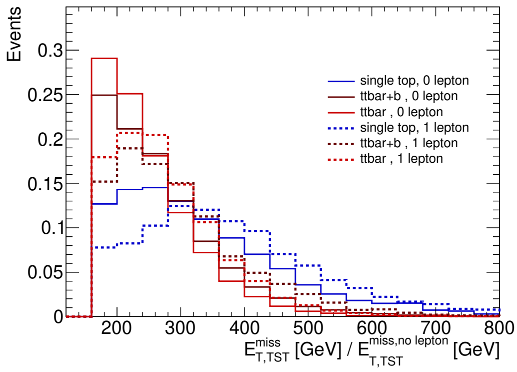
Figure 6.6.6 compares the data and MC prediction in the Z+jets enriched control regions before the profiling. Underprediction in total event yield can be observed in resolved regime, which is mainly due to the mismodeling of hadronic shower where the QCD



(a)



(b)



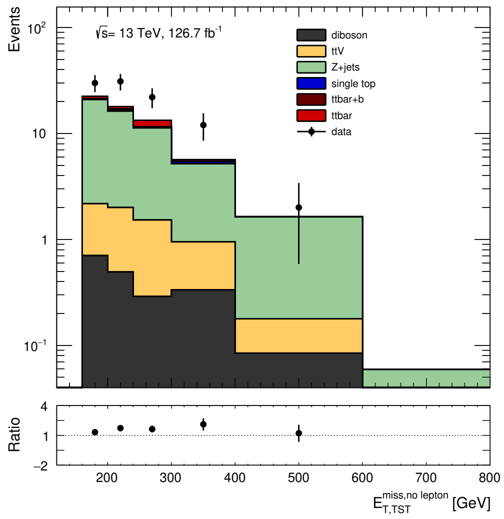
(c)

Figure 6.6.5: Shape of $E_{\mathrm{T}}^{\mathrm{miss}}$ in 0 lepton final state (solid lines) and $E_{\mathrm{T,\,no\ lepton}}^{\mathrm{miss}}$ in 1 lepton final state (dashed lines), normalized to unit area, for $t\bar{t}$ (red), $t\bar{t}+b$ (brown) and single top (blue) processes after the baseline selection, the anti-QCD requirements and the leptonic top reconstruction cut applied in (a) the resolved regime, (b) the boosted regime and (c) the highly boosted regime.

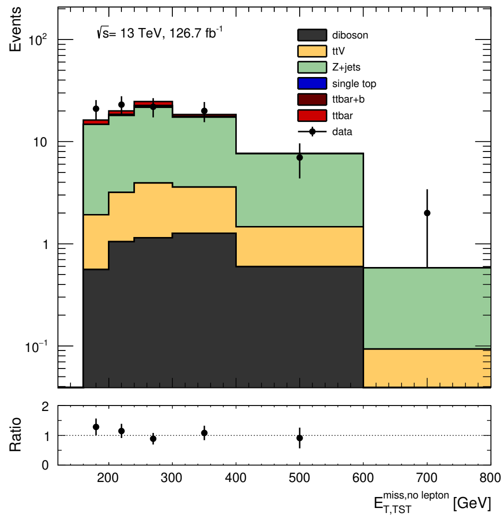
Resolved regime	**Boosted regime**	**Highly boosted regime**
	baseline selection	
	2 tight electron/muon, 0 loose electron/muon, 0 hadronic tau	
	anti-QCD requirements	
	$E_{\mathrm{T,\,no\ lepton}}^{\mathrm{miss}} > 160$ GeV	
	80 GeV $< m_{ll} <$ 100 GeV	
	object-based $E_{\mathrm{T}}^{\mathrm{miss}}$ sig. < 5	
	$p_{\mathrm{T}}^{ll} > 160$ GeV	
0 large-R jet	≥ 1 large-R jet	≥ 1 large-R jet
	$m_J < 130$ GeV	$m_J \geq 130$ GeV

Table 6.6.3: List of the 2 lepton control region event selections used in the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search.

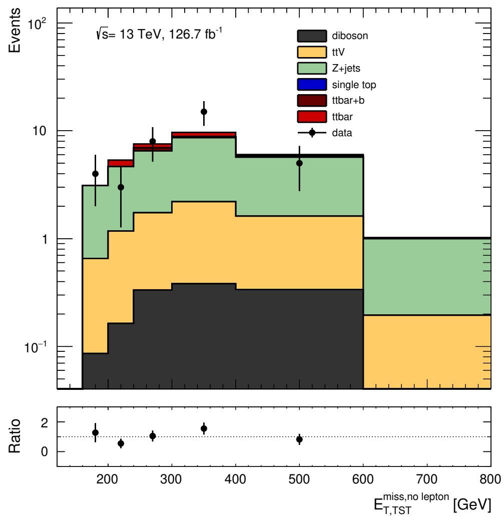
gluon splitting $g \rightarrow bb$ takes place. As shown in Figure 6.6.7, in the resolved regime the deviation between data and MC converges at low $\Delta R\,(b_1, b_2)$, indicating larger contribution from gluon splitting in observation. Such effect is accounted by a theoretical uncertainty specially assigned to Z+jets simulation, as discussed in Section 6.7.2.



(a)



(b)



(c)

Figure 6.6.6: $E_{\mathrm{T,\,no\;lepton}}^{\mathrm{miss}}$ distributions in the (a) resolved, (b) boosted and (c) highly boosted Z+jets enriched control regions. The distributions are shown before the profiling.

6.6.2.1 2 lepton control region validation

Similar to the 1 lepton CRs, part of the signal region selections are loosened in the 2 lepton CRs and the validation is performed to check the impact of these extrapolations by comparing the observables involving $E_{\mathrm{T}}^{\mathrm{miss}}$ in the SRs to the corresponding 'no lepton' observables in the 2 lepton CRs.

Figure 6.6.8 shows the shape of $\cosh_{\mathrm{max}}$ in 0 lepton final state and $\cosh_{\mathrm{max}}^{\mathrm{no\;lepton}}$ in 2 lepton final state for Z+jets background and very good agreement in shape can be observed in

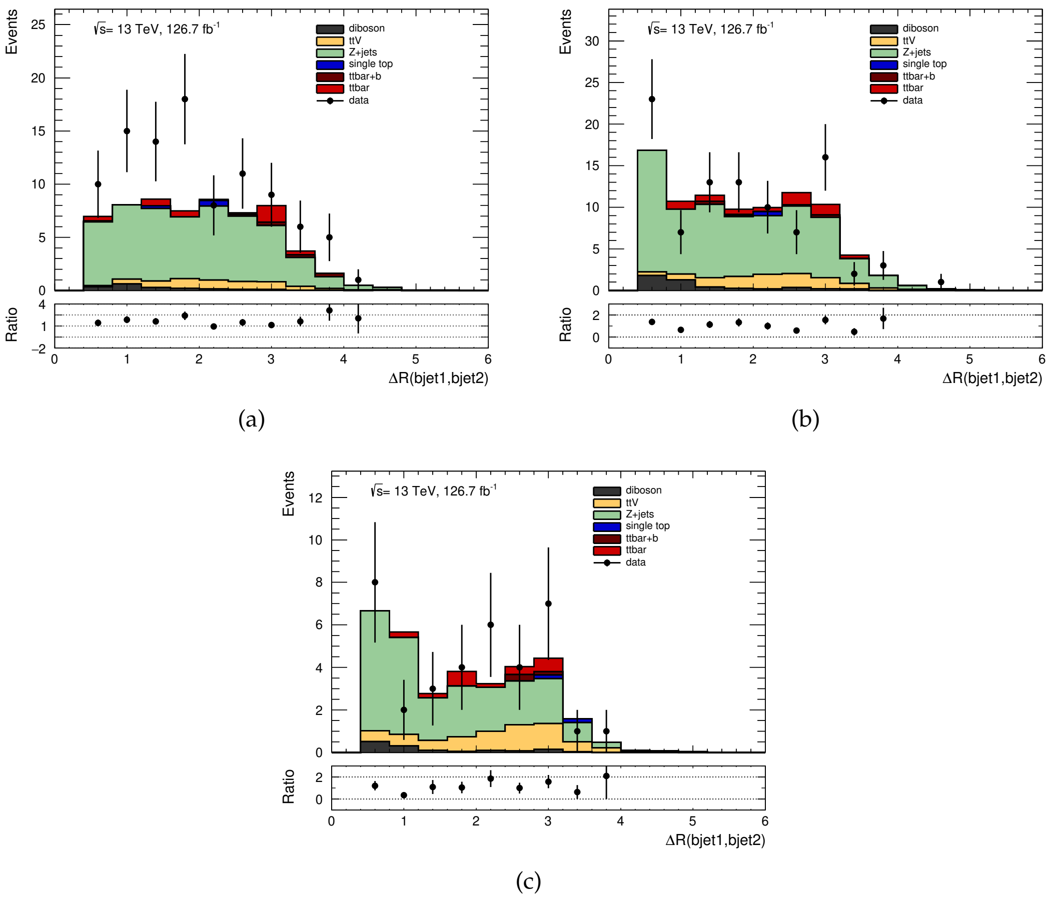


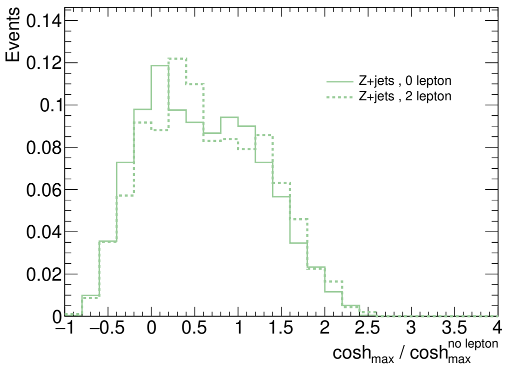
Figure 6.6.7: $\Delta R\,(b_1, b_2)$ distributions in the (a) resolved, (b) boosted and (c) highly boosted Z+jets enriched control regions. The distributions are shown before the profiling.

all kinematic regimes. In Figure 6.6.9, slight deviations between $E_{\mathrm{T}}^{\mathrm{miss}}$ and $E_{\mathrm{T,\,no\ lepton}}^{\mathrm{miss}}$ can be seen in the resolved regime. This is expected because the value of $\cosh_{\max}$ is negatively correlated to $E_{\mathrm{T}}^{\mathrm{miss}}$ (Equation 6.5.7). By cutting on $\cosh_{\max}$ ($\cosh_{\max}^{\mathrm{no\ lepton}}$) only in the SRs and not in the 2 lepton CRs, biases towards events with higher $E_{\mathrm{T}}^{\mathrm{miss}}$ are introduced unavoidably.

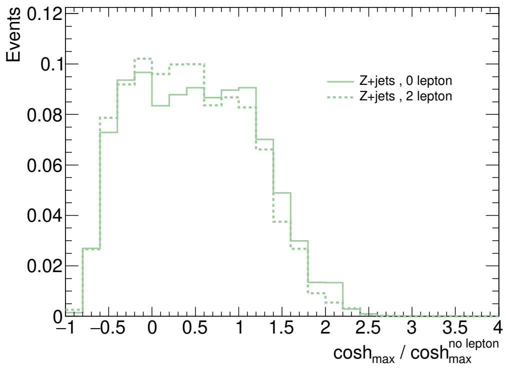
For $Z(\to ll)$+jets background, $E_{\mathrm{T,\,no\ lepton}}^{\mathrm{miss}} \sim p_{\mathrm{T}}^{ll}$, such that extrapolations of the kinematic properties of leptons will inevitably lead to the extrapolation of $E_{\mathrm{T}}^{\mathrm{miss}}$. Hence, the deviations in shape between the 0 lepton and the 2 lepton variables are expected, and somewhat unavoidable due to the statistical limitations, when designing the 2 lepton CRs.

6.6.3 Validation regions

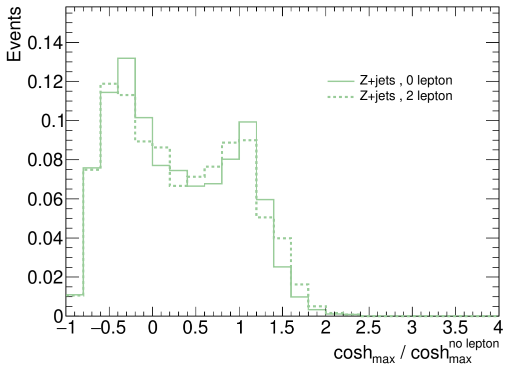
In the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search, VRs are not included in the statistical model to constrain the backgrounds and only serve to validate the extrapolation from the CRs to the SRs. The event selections in the VRs are designed to be orthogonal to the signal region selections, while six disjoint VRs are defined for the four major backgrounds which are expected to be constrained by the CRs:



(a)



(b)



(c)

Figure 6.6.8: Shape of $\cosh_{\max}$ in 0 lepton final state (solid lines) and $\cosh_{\max}^{\text{no lepton}}$ in 2 lepton final state (dashed lines), normalized to unit area, for Z+jets process after the baseline selection and the anti-QCD requirements applied in (a) the resolved regime, (b) the boosted regime and (c) the highly boosted regime.

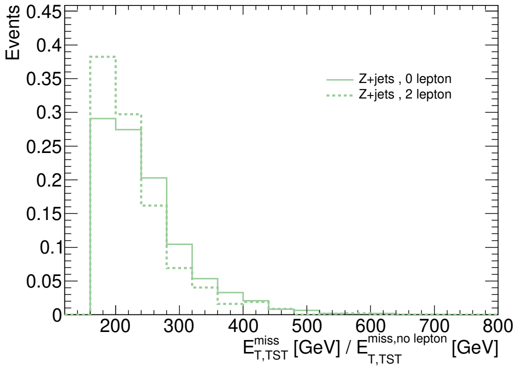
- $t\bar{t}$ enriched VRs: three kinematic regimes, validation of $t\bar{t}$ normalization;
- $t\bar{t}$ suppressed VRs: three kinematic regimes, validation of $t\bar{t}+b$, single top and Z+jets normalization.

6.6.3.1 $t\bar{t}$ enriched validation regions

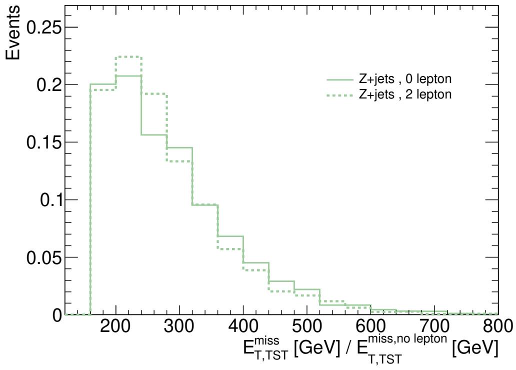
In the $t\bar{t}$ enriched VRs, $t\bar{t}$ events are selected by inverting the tight $\cosh_{\max}$ cut applied in the SRs. In addition, a loose version of $\cosh_{\max}$ requirement has to be fulfilled, with the selection values taken from the $\cosh_{\max}^{\text{no lepton}}$ cut in the 1 lepton CRs. This helps to ensure that the $t\bar{t}$ backgrounds in the VRs share similar kinematic properties to those in the CRs, therefore can be used to validate the control-to-signal region extrapolation. All other selections are kept the same as the SRs.

Table 6.6.4 presents the full list of the $t\bar{t}$ enriched validation region event selections used in the $E_{\mathrm{T}}^{\text{miss}}+t\bar{t}$ (fully-hadronic) search.

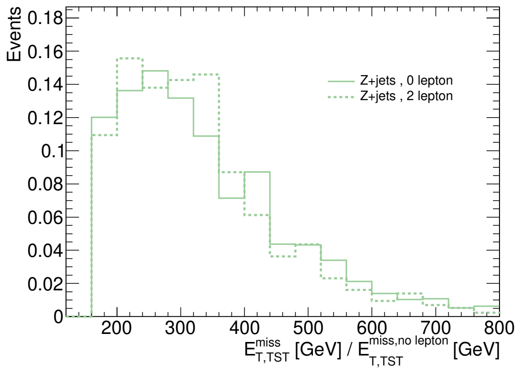
Figure 6.6.10 compares the data and MC prediction in the $t\bar{t}$ enriched VRs before the profiling.



(a)



(b)



(c)

Figure 6.6.9: Shape of $E_{\mathrm{T}}^{\mathrm{miss}}$ in 0 lepton final state (solid lines) and $E_{\mathrm{T,\,no\ lepton}}^{\mathrm{miss}}$ in 2 lepton final state (dashed lines), normalized to unit area, for Z+jets process after the baseline selection, the anti-QCD requirements and the leptonic top reconstruction cut (only in SRs) applied in (a) the resolved regime, (b) the boosted regime and (c) the highly boosted regime.

6.6.3.2 $t\bar{t}$ suppressed validation regions

Due to the limited event yield in the 0 lepton regions, the validation regions for $t\bar{t}+b$, single top and Z+jets are merged into one $t\bar{t}$ suppressed region. In these VRs, $t\bar{t}$ events are rejected by inverting the $\chi^2_{t\bar{t},\,\mathrm{had}}$ cut applied in the SRs. In addition, an upper cut on $\chi^2_{t\bar{t},\,\mathrm{had}}$ is used to veto events which thoroughly fail the hadronic top pair reconstruction or have less than five jets in the final state. The $p_{\mathrm{T}}^{t\bar{t}}/E_{\mathrm{T}}^{\mathrm{miss}}$ requirements are discarded as they become irrelevant when the value of $\chi^2_{t\bar{t},\,\mathrm{had}}$ is too large.

To minimize the W+jets contamination in the $t\bar{t}$ suppressed VRs, a tight $\Delta R\,(b_1, b_2)$ selection is introduced in all kinematic regimes. This is motivated by the fact that a large fraction of the W+jets events entering the VRs contain extra b-jets originating from gluon splitting, where a small angular separation between the leading two b-jets is expected. The values of $\Delta R\,(b_1, b_2)$ cuts are optimized in each regime to provide similar amount of MC prediction with respect to the $t\bar{t}$ enriched VRs.

The additional cuts on $E_{\mathrm{T}}^{\mathrm{miss}}$ and $E_{\mathrm{T}}^{\mathrm{miss}}$ significance applied in the SRs are kept to reduce the multijet background, while the tight $\Delta\phi_{\mathrm{min}}(E_{\mathrm{T}}^{\mathrm{miss}}$, central jets) selection in resolved

Resolved regime	Boosted regime	Highly boosted regime
	baseline selection	
	0 loose electron/muon, 0 hadronic tau	
	anti-QCD requirements	
	$E_{\mathrm{T}}^{\mathrm{miss}} > 160$ GeV	
	object-based $E_{\mathrm{T}}^{\mathrm{miss}}$ sig. > 10	
	$\Delta R\,(b_1, b_2) > 1.2$	
0 large-R jet	≥ 1 large-R jet	≥ 1 large-R jet
$\Delta\phi_{\mathrm{min}}(E_{\mathrm{T}}^{\mathrm{miss}}$, central jets) > 1.0	$m_J < 130$ GeV	$m_J \geq 130$ GeV
$0.5 < \cosh_{\mathrm{max}} < 0.9$	$0.6 < \cosh_{\mathrm{max}} < 0.95$	$0.7 < \cosh_{\mathrm{max}} < 1.0$
$\chi^2_{t\bar{t},\,\mathrm{had}} < 4$ $0.7 < p_{\mathrm{T}}^{t\bar{t}}/E_{\mathrm{T}}^{\mathrm{miss}} < 1.2$	$\chi^2_{t\bar{t},\,\mathrm{had}} < 6$ $0.5 < p_{\mathrm{T}}^{t\bar{t}}/E_{\mathrm{T}}^{\mathrm{miss}} < 1.2$	$\chi^2_{t\bar{t},\,\mathrm{had}} < 8$ $0.5 < p_{\mathrm{T}}^{t\bar{t}}/E_{\mathrm{T}}^{\mathrm{miss}} < 1.2$ ΔR_{min}(large-R jet, b-jets) < 1.2

Table 6.6.4: List of the $t\bar{t}$ enriched validation region event selections used in the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search.

regime is dropped, allowing for more events in the $t\bar{t}$ suppressed VRs.

Table 6.6.5 presents the full list of the $t\bar{t}$ suppressed validation region event selections used in the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search.

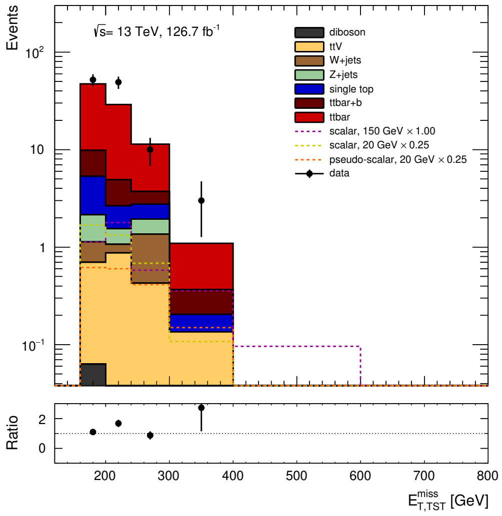
Resolved regime	Boosted regime	Highly boosted regime
	baseline selection	
	0 loose electron/muon, 0 hadronic tau	
	anti-QCD requirements	
	$E_{\mathrm{T}}^{\mathrm{miss}} > 160$ GeV	
	object-based $E_{\mathrm{T}}^{\mathrm{miss}}$ sig. > 10	
$\Delta R\,(b_1, b_2) > 2.2$	$\Delta R\,(b_1, b_2) > 1.6$	$\Delta R\,(b_1, b_2) > 1.2$
0 large-R jet	≥ 1 large-R jet	≥ 1 large-R jet
	$m_J < 130$ GeV	$m_J \geq 130$ GeV
$\cosh_{\mathrm{max}} < 0.5$	$\cosh_{\mathrm{max}} < 0.6$	$\cosh_{\mathrm{max}} < 0.7$
$4 < \chi^2_{t\bar{t},\,\mathrm{had}} < 999$	$6 < \chi^2_{t\bar{t},\,\mathrm{had}} < 999$	$8 < \chi^2_{t\bar{t},\,\mathrm{had}} < 999$ ΔR_{min}(large-R jet, b-jets) < 1.2

Table 6.6.5: List of the $t\bar{t}$ suppressed validation region event selections used in the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search.

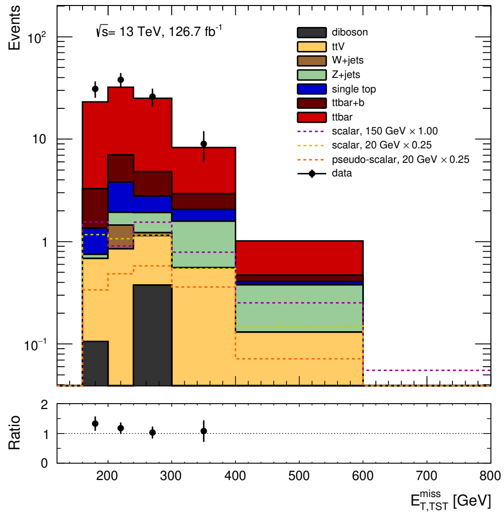
Figure 6.6.11 compares the data and MC prediction in the $t\bar{t}$ suppressed VRs before the profiling.

6.6.4 Multijet estimation

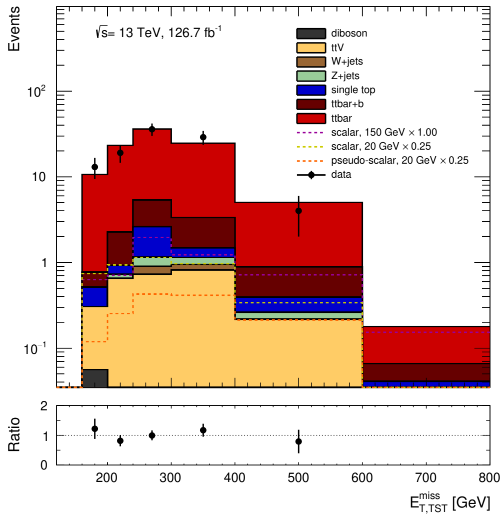
After the anti-QCD requirements, the contribution of multijet process is expected to be subdominant in the SRs compared to other backgrounds, while in the 1 lepton and the 2 lepton CRs it can be assumed as negligible. Hence, the multijet estimation is only performed in 0 lepton final state using a data-driven 'Rebalance and Smear' method [191], which consists of two steps:



(a)



(b)



(c)

Figure 6.6.10: $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in the (a) resolved, (b) boosted and (c) highly boosted $t\bar{t}$ enriched validation regions. The distributions are shown before the profiling.

- **Rebalancing**. Given the assumption that the multijet background entering the SRs have significant $E_{\mathrm{T}}^{\mathrm{miss}}$ mainly due to mismeasurements of the detector response, data events failing the signal region selections are selected to build the 'seed samples', which are then used to generate pseudo multijet events with unbalanced transverse momentum. During the Rebalancing step, a kinematic fit is applied to the seed samples to adjust the p_{T} of each jet, so that the total transverse momentum is balanced. The rectangular coordinates of jets, η and ϕ, on the other hand, are constrained by the corresponding detector resolution and remain mostly unchanged after the fit. These rebalanced data samples are expected to have the same kinematic topologies as truth-level multijet events with $E_{\mathrm{T}}^{\mathrm{miss}} \sim 0$.
- **Smearing**. Next, the rebalanced jets are smeared according to the detector response distributions. The energy of each jet is scaled by a factor randomly picked

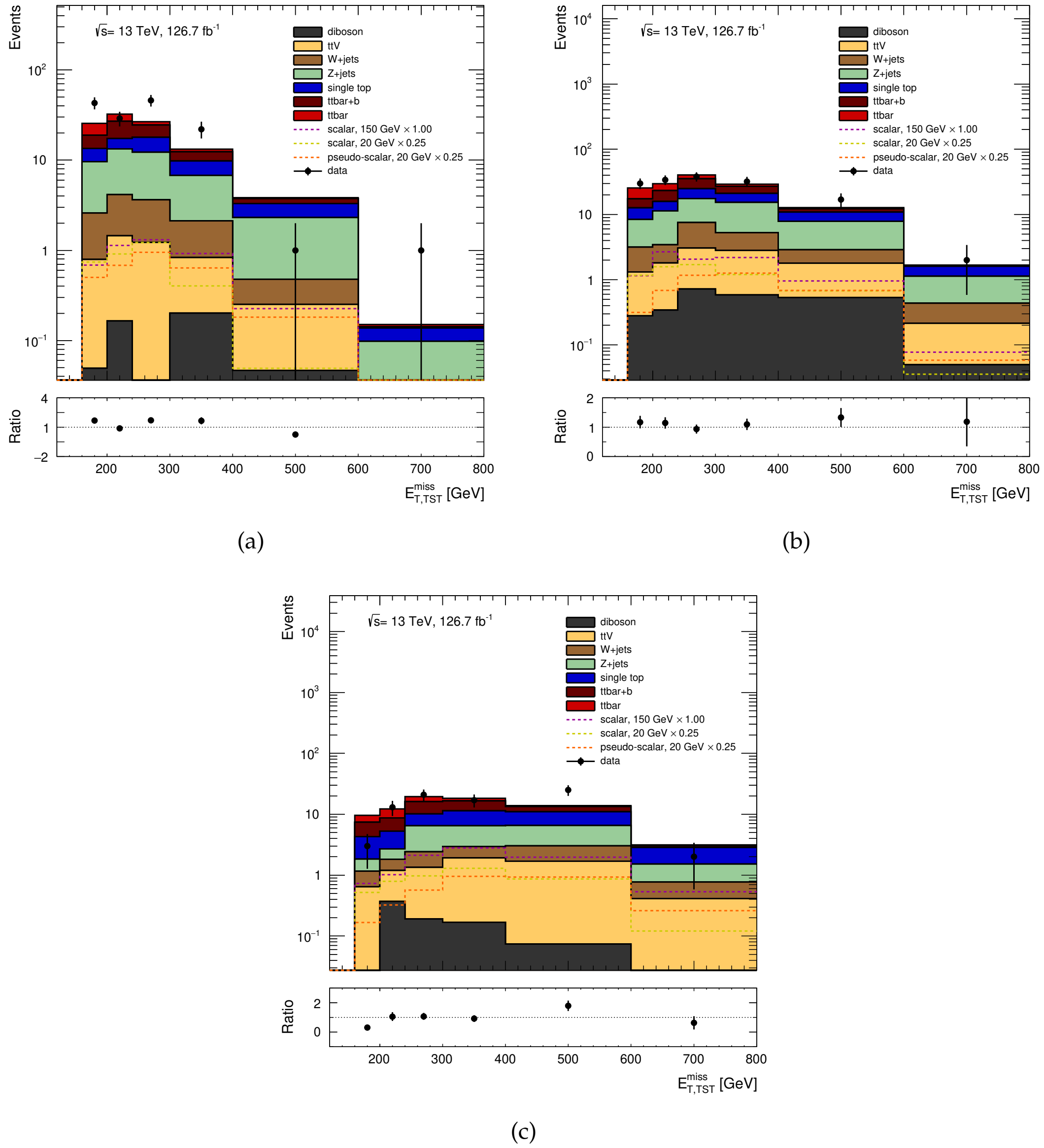


Figure 6.6.11: $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in the (a) resolved, (b) boosted and (c) highly boosted $t\bar{t}$ suppressed validation regions. The distributions are shown before the profiling.

from the non-Gaussian jet energy response distributions, binned in E, $|\eta|$ and jet flavor. These response maps are obtained from MC simulation and then corrected by measurements of dijet data. To increase the size of the pseudo multijet dataset, multiple random smearing is performed for each event. For samples containing both hard-scatter and pile-up components, the kinematic topologies of the smeared events are further enriched by randomly rotating in ϕ the hard-scatter jets with respect to the pile-up jets. Moreover, the JVT discriminants of the jets are recalculated to reflect the substantial changes in the kinematic properties after the smearing procedure, which accounts for the inefficiencies of pile-up suppression.

Figure 6.6.12 gives a schematic overview of the 'Rebalance and Smear' method. The pseudo multijet events generated via this procedure are validated using the measured data in the QCD enriched regions, which are defined by inverting the dominant anti-

QCD cut $\Delta\phi_{\mathrm{min}}(E_{\mathrm{T}}^{\mathrm{miss}}$, central jets) > 0.5 and discarding the requirements on $\Delta\phi_{\mathrm{min}}(E_{\mathrm{T}}^{\mathrm{miss}}$, forward jets). Decent agreement between prediction and observation is reached.

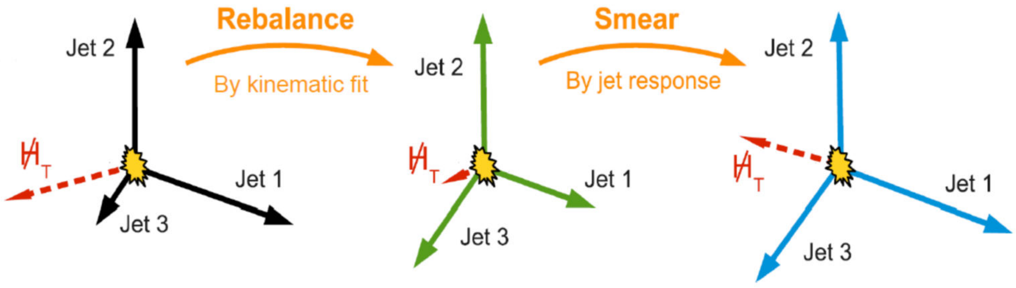


Figure 6.6.12: Schematic illustration of the 'Rebalance and Smear' method used to provide the multijet background estimation.

The estimation of the multijet contamination in each SR is derived following the methodology documented in Ref. [192] and presented in Table 6.6.6. Compared to the major backgrounds, the contribution from the multijet process is proven to be minor, therefore will not be considered in the statistical model.

Process	Resolved regime	Boosted regime	Highly boosted regime
Multijet	0.15 ± 0.43	0.05 ± 0.22	0.32 ± 0.76
Total background	50.61 ± 14.39	58.80 ± 12.38	38.04 ± 6.86

Table 6.6.6: Estimation of the multijet contamination in each signal region derived following the methodology documented in Ref. [192].

6.7 Systematic uncertainties

A general description of the systematic uncertainties evaluated in the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search is given in this section, with the experimental systematic uncertainties which originate from the reconstruction, identification and calibration of the physical objects being discussed in Section 6.7.1 and the theoretical systematic uncertainties which correspond to theory predictions and MC modelings being discussed in Section 6.7.2.

6.7.1 Experimental systematic uncertainties

The total integrated luminosity is used to normalize the MC prediction to the measured data. In the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search, 126.7 fb^{-1} of pp collision data is analyzed, which corresponds to an uncertainty of 1.7% on the luminosity measurement during full Run 2 data-taking. Additionally, the pile-up conditions of the MC simulated events are reweighted to match those in data, while the pile-up reweighting uncertainty is introduced, accounting for the possible biases in the modeling of pile-up events. The total pile-up uncertainty is at the order of 6%, based on measurements performed using Run 1 data [164].

Aside from the luminosity uncertainty and the pile-up reweighting uncertainty, experimental systematic uncertainties are evaluated for various reconstructed objects using standard tools provided by the ATLAS Combined Performance (CP) groups.

Electrons are used in all SR and CR selections. The reconstruction, identification and isolation criteria of electrons naturally affect the event yields in various analysis regions, while the uncertainties on electron p_{T} measurements affect the shape of the $E_{\mathrm{T,\,no\ lepton}}^{\mathrm{miss}}$ variable. Experimental uncertainties on electron reconstruction, identification and isolation efficiencies are therefore taken into account.

Similar to electrons, muons are involved in various SR and CR selections, while the corresponding uncertainties are introduced to account for the efficiency scale factors of muon reconstruction (track-to-vertex association), identification and isolation.

The full chain of calibration is applied to the small-R jets used in the analysis, consisting of origin correction, pile-up correction, JES and η calibration, GSC and in-situ calibration, which accounts for various detector effects and aims to correct the kinematic properties of jets to truth-level. Multiple sources of systematic uncertainty are evaluated for each procedure, resulting in more than 90 components in total. As most of the JES uncertainties are expected to be subdominant in this analysis, a reduced uncertainty set containing 20 terms is applied in the statistic model. Jet energy resolution (JER) uncertainties, on the other hand, are derived using data events in which the jet p_{T} is balanced by a reference object, e.g. γ+jet, Z+jet and dijet processes, and contain 34 terms in total. A reduced set composed of 8 JER uncertainties is taken into consideration. The flavor tagging uncertainties consist of five components, which correspond to the efficiency scale factors of light-flavor, c-tagged and b-tagged jets, as well as the extrapolation of c- and b-tagging efficiencies to high p_{T} regimes. Lastly, JVT and fJVT uncertainties are evaluated on an event-by-event basis to account for the inefficiencies of pile-up rejection in the central and the forward regions, respectively.

As only the invariant mass of large-R jets are used for event categorization, a reduced set of four systematic uncertainties associated to large-R jet mass modeling is included in the model. These terms are derived using experimental approaches, which represent the difference between data and MC simulation, the fragmentation modeling, the tracking reconstruction efficiency, fake rate and bias in the q/p_{T} distribution, and the total statistical uncertainty.

Three sources of $E_{\mathrm{T}}^{\mathrm{miss}}$ uncertainties are considered, including variations on the scale of the soft term as well as its resolution in the direction perpendicular and parallel to the hadronic recoil system, which are extracted from scale and resolution comparisons between 2015 data and MC, between Powheg + Pythia 8 and MadGraph5 generators and between full detector simulation and fast simulation [193].

6.7.2 Theoretical systematic uncertainties

Theoretical systematic uncertainties account for various MC simulation parametrization, e.g. matrix elements, renormalization and factorization scales, ISR and FSR, by means of normalization and shape uncertainties.

Three major sources of uncertainty associated to the modeling of $t\bar{t}$ and single top processes are considered in the analysis:

- **Matrix element calculation**, evaluated by taking the difference between Powheg + Pythia 8 and MadGraph5_aMC@NLO + Pythia 8 predictions.

- **Parton shower and hadronization**, evaluated by taking the difference between Powheg + Pythia 8 and Powheg + Herwig 7 predictions.
- **Scale and tune uncertainty**, evaluated by taking the difference between predictions given by the nominal Powheg + Pythia 8 and Powheg + Pythia 8 with the renormalization and factorization scales and A14 tune varied coherently to increase and decrease the amount of initial state radiation and final state radiation in top events [194];

An additional source of uncertainty which affects single top only corresponds to the interference between single top and $t\bar{t}$ production. This is evaluated by comparing the MC prediction applying the diagram removal (DR) and the diagram subtraction (DS) schemes [195] for single top production.

For W+jets, Z+jets and $t\bar{t}+V$ processes, the theoretical uncertainties are evaluated by varying the renormalization and factorization scales up and down by a factor of 2. Individual renormalization and factorization scale variations are treated as two independent nuisance parameters, while the coherent variation combination of the renormalization and factorization scales are treated as an additional shape uncertainty. The variations are assessed by reweighting the nominal samples based on alternative weights.

Another source of uncertainty is considered for Z+jets events, which accounts for the modeling of matrix element calculation, parton shower and hadronization in the event generators by taking the difference between Sherpa 2.2.1 and MadGraph5 + Pythia 8 predictions. This especially corresponds to the ratio between additional jets originating from gluon splitting and the scattering diagrams.

The PDF uncertainty is evaluated using bootstrap replicas of the NNPDF30NNLO set provided within the simulated samples. The total cross section is also varied by $\pm 6\%$ and $\pm 5\%$ for W+jets and diboson production, respectively. This is not necessary for $t\bar{t}$, single top and Z+jets processes, which have dedicated control regions defined and included in the combined profile likelihood fit to extract the normalization from data.

6.8 Results

All SRs and CRs are considered in a statistical model using a combined profile likelihood fit, where the 'background-only' hypothesis is tested against the 'signal-plus-background' hypothesis by assuming the presence of DM signals on top of the SM background processes. A brief overview of the statistical model applied in the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search is given in Section 6.8.1. The impact of various sources of systematic uncertainty is propagated in terms of the nuisance parameters (NPs) in the model and evaluated from a fit to the Asimov dataset [170], as described in Section 6.8.2. In Section 6.8.3, a 'background-only' fit including only the CRs is performed to validate the control-to-signal region extrapolation with the help of the VRs and the observed results are obtained from the combined fit to the actual data. The exclusion limits on the scalar and the pseudo-scalar mediator simplified models are presented as functions of the mediator mass m_{med} in Section 6.8.4.

6.8.1 Statistical model

Similar to the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search, a profile likelihood fit [170] is performed on the event yields in all analysis regions included in the fit in order to test the existence of DM signals. The fit results are obtained by maximizing the likelihood function $\mathcal{L}(N^{\text{obs}}|\mu,\boldsymbol{\theta})$, while the test statistic q_μ (Equation 5.8.5) is used to evaluate the validity of hypotheses with the given signal strength μ.

To avoid accidental exclusion of signals in regions where limited sensitivity is expected, the CL_s probability [169] is introduced to provide conservative estimation of the exclusion limit. All upper limits given in this section correspond to the 95% CL_s exclusion limit.

As mentioned in Section 6.3.2, the normalization of the major background processes can be directly constrained in the statistical model by the usage of the CRs. Four free floating normalization parameters are included in the model, which correspond to $t\bar{t}$, $t\bar{t}+b$, single top and Z+jets.

Table 6.8.1 shows the regions included in the fit and the physical processes enriched in each region.

	0 lepton	**1 lepton**	**2 lepton**
Enriched in	signals	$t\bar{t}$, $t\bar{t}+b$, single top	Z+jets
Fitted observable	event yield		
Kinematic topologies	0 large-R jet: resolved		
	≥ 1 large-R jet, $m_J < 130$ GeV: boosted		
	≥ 1 large-R jet, $m_J \geq 130$ GeV: highly boosted		

Table 6.8.1: Summary of the analysis regions considered in the statistical model of the $E_{\text{T}}^{\text{miss}} + t\bar{t}$ (fully-hadronic) search. A combined profile likelihood fit is performed on the event yield of all signal regions and control regions.

6.8.2 Impact of uncertainties

To evaluate the relative impact of different sources of uncertainty, the Asimov dataset [170] is built from the predicted distribution of MC backgrounds, with a Poisson error corresponding to the statistical uncertainty of data assumed in each bin.

For a given NP θ_i, its relevance to the sensitivity of the search can be estimated as a fractional uncertainty on the total background estimation. This is computed by repeating the fit with the absence of θ_i and subtracting in quadrature the resulting uncertainty on the background expectation from the total uncertainty:

$$\sigma_{\theta_i} = \sqrt{\sigma^2_{\text{total}} - \sigma^2_{\text{no }\theta_i}}. \tag{6.8.1}$$

Table 6.8.2 shows the impact of different uncertainties on σ_{total} in the three signal regions. Multiple NPs are grouped into categories according to their sources. The total statistical uncertainty is evaluated by neglecting all systematic uncertainties in the fit. The dominant sources of uncertainty include:

- Modeling uncertainties. The modeling of $t\bar{t}$ and single top background have large impact on background estimation, especially in the resolved and the boosted regimes. The largest contribution of $t\bar{t}$ theory uncertainties comes from the modeling of hard-scatter process, while for single top the dominant component corresponds to varying the amount of the final state radiation.

Source	Uncertainty [%]		
	(a)	(b)	(c)
$t\bar{t}$ modeling	4.9	3.4	4.9
Small-R jets	3.0	3.2	2.2
Single top modeling	2.4	2.1	0.6
MC Stats.	1.9	2.1	1.6
Single top normalization	1.1	0.2	0.9
$t\bar{t}$ normalization	1.0	0.9	0.6
$E_{\mathrm{T}}^{\mathrm{miss}}$	0.7	0.8	0.3
Z+jets modeling	0.5	2.6	2.3
Large-R jets	0.3	0.8	0.7
Z+jets normalization	0.2	1.5	1.4
$t\bar{t} + b$ normalization	0.2	1.0	0.3
Pile-up reweighting	0.2	< 0.1	0.3
Flavor tagging	0.2	< 0.1	0.3
$t\bar{t} + V$ modeling	< 0.1	< 0.1	0.3
Electrons	< 0.1	0.2	0.3
W+jets modeling	< 0.1	0.4	0.3
Muons	< 0.1	< 0.1	0.3
Luminosity	< 0.1	< 0.1	0.3
Data Stats.	12.8	11.7	13.8
Total Syst.	10.6	9.5	12.1
Total	16.6	15.1	18.4

Table 6.8.2: Relative impact of different sources of uncertainty on the total background estimation in the (a) resolved, (b) boosted and (c) highly boosted signal region.

- Small-R jet calibration and resolution. In the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search, small-R jets are used for various purposes, including hadronic top pair reconstruction and multijet rejection. As shown, the impact of small-R jet related uncertainties decreases with the rising p_{T} of the top system. This is mainly caused by the loosened $\chi^2_{t\bar{t},\,\mathrm{had}}$ cuts in the boosted and the highly boosted regimes.

- Finite MC events. The number of events for the SM backgrounds, in particular the Z+jets process, is strongly limited due to the two tight b-jets required in the event selection. The impact of MC statistical uncertainty is very stable with respect to the different kinematic topologies, as the selection values applied in the analysis have been adjusted to provide similar amount of event yield in each SR.

- Normalization. The major components include $t\bar{t}$, single top and Z+jets. Uncertainties due to the number of events in the CRs are propagated to the fit result in terms of normalization parameters. In general, the impact of background normalization is only at percent level in all regimes.

In all three signal regions, the data statistical uncertainty overwhelms the systematic uncertainties, indicating that the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search is mostly statistically limited.

6.8.3 Observed results

To check the strategy of control region design, the SM background predictions are fitted to the observed data yield in the CRs, which is often referred to as the 'background-only' fit. Figure 6.8.1 shows the event yield in the VRs and the CRs after the background-only fit. General improvement can be seen over the agreement between data and MC prediction in all bins, while deviations between data and the Standard Model backgrounds in the VRs are all within $\pm 1\sigma$, thus validates the statistical model of the analysis.

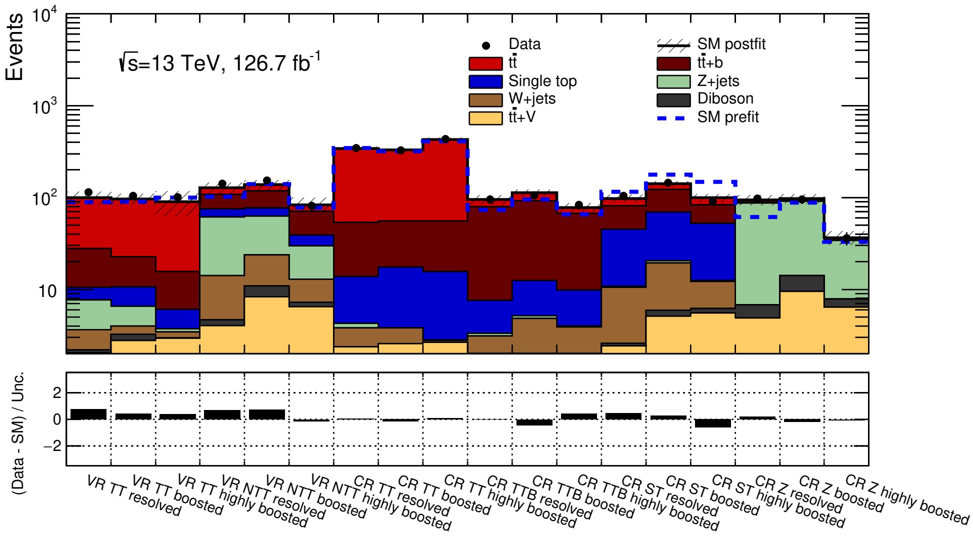


Figure 6.8.1: Event yield in all validation regions and control regions after the 'background-only' fit. The stacked histograms show the SM prediction and the hatched uncertainty band around the SM prediction corresponds to the total MC uncertainty, which consists of the MC statistical uncertainty, experimental systematic uncertainties and theoretical systematic uncertainties. The dotted blue line denotes the MC prediction before the background-only fit. The lower band shows the deviation between observation and prediction with respect to the total uncertainty, which consists of the total MC uncertainty and the data statistical uncertainty.

Then, a combined profile likelihood fit which takes all SRs and CRs into account is performed, while the numbers of observed events entering each signal region are presented in Table 6.8.3. Alongside are the signal predictions for scalar and pseudo-scalar models with $m_{\mathrm{med}} = 200$ GeV, as well as the expected background contributions determined by the fit to the full measured data.

Process	Resolved regime	Boosted regime	Highly boosted regime
Scalar mediator model m_χ =1 GeV, $m_{\rm med}$ =200 GeV	5.89 ± 1.19	9.05 ± 1.42	9.56 ± 2.45
Pseudo-scalar mediator model m_χ =1 GeV, $m_{\rm med}$ =200 GeV	10.68 ± 2.88	13.78 ± 1.96	14.52 ± 1.48
$t\bar{t}$	25.57 ± 6.13	19.75 ± 6.08	8.09 ± 1.18
$t\bar{t}+b$	19.39 ± 4.83	13.89 ± 3.36	9.82 ± 2.42
Single top	4.56 ± 4.08	4.51 ± 2.93	3.11 ± 2.33
Z+jets	6.91 ± 2.84	6.27 ± 3.17	3.84 ± 0.87
W+jets	1.93 ± 1.37	2.60 ± 1.75	0.55 ± 0.42
Diboson	0.39 ± 0.21	0.54 ± 0.15	0.54 ± 0.05
$t\bar{t}+V$	6.97 ± 1.10	9.91 ± 1.18	10.01 ± 1.20
Total background	65.72 ± 6.77	57.47 ± 6.29	35.95 ± 3.60
Data	71	53	30

Table 6.8.3: Expected and observed numbers of events in the resolved, boost and highly boosted signal regions with an integrated luminosity of 126.7 fb^{-1} and $\sqrt{s}$ = 13 TeV. The background yields and uncertainties are shown after the profile likelihood fit to the data (with $\mu = 0$). The quoted background uncertainties include both the statistical and systematic contributions, while the uncertainties in the signals are statistical only.

Table 6.8.4 shows the normalization of the major background components constrained by the statistical model under the 'background-only' hypothesis. For $t\bar{t}$ and Z+jets, the observation is consistent with the SM expectation, while above 1 σ deviation is present for the $t\bar{t}+b$ and single top backgrounds. This can be accounted, to some extent, by the strong anti-correlation between the normalization parameters of $t\bar{t}+b$ and single top, which originates from the non-negligible $t\bar{t}+b$ contamination in the single top enriched CRs and the single top contamination in the $t\bar{t}+b$ enriched CRs. As the $t\bar{t}+b$ and single top processes have very similar kinematic behavior in the SRs (Figure 6.3.1), the sensitivity of the analysis will not be harmed by the anti-correlation. Moreover, positive pull of a NP which corresponds to the modeling of the interference between single top and $t\bar{t}$ processes is observed, implying that the statistical interpretation is more in favor of the diagram subtraction scheme of single top production. Hence, overestimation of single top contribution can be expected in the nominal MC variation, where the diagram removal scheme is applied instead.

Process	Fitted normalization
$t\bar{t}$	0.98 ± 0.07
$t\bar{t}+b$	1.59 ± 0.32
Single top	0.46 ± 0.17
Z+jets	1.18 ± 0.33

Table 6.8.4: Background normalization factors relative to the initial theoretical prediction, extracted from the profile likelihood fit under the background-only hypothesis.

The constraints on the nuisance parameters after the profiling are checked with caution and no pulls beyond 1 standard deviation from the prior knowledge of the systematic uncertainties are observed. Figure 6.8.2 shows the event yield in the SRs and the CRs after the combined profile likelihood fit, where good agreement between data and prediction can be seen in all bins.

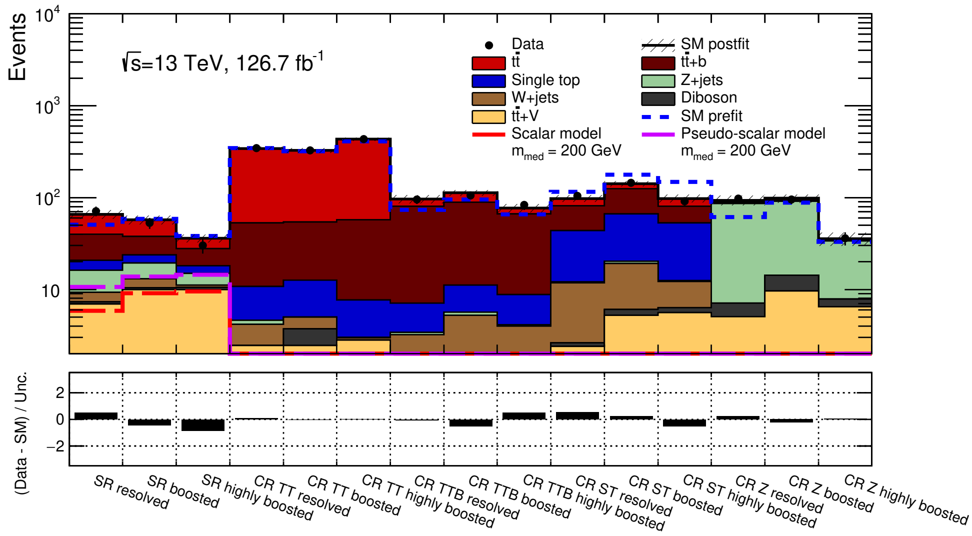


Figure 6.8.2: Event yield in all signal regions and control regions after the combined profile likelihood fit. The stacked histograms show the SM prediction and the hatched uncertainty band around the SM prediction corresponds to the total MC uncertainty, which consists of the MC statistical uncertainty, experimental systematic uncertainties and theoretical systematic uncertainties. The dotted blue line denotes the MC prediction before the combined fit. The dashed lines represent the expected contribution of the scalar mediator simplified model with $m_{\mathrm{med}} = 200$ GeV (violet) and the pseudo-scalar mediator simplified model with $m_{\mathrm{med}} = 200$ GeV (red) assuming $g = 1$. The lower band shows the deviation between observation and prediction with respect to the total uncertainty, which consists of the total MC uncertainty and the data statistical uncertainty.

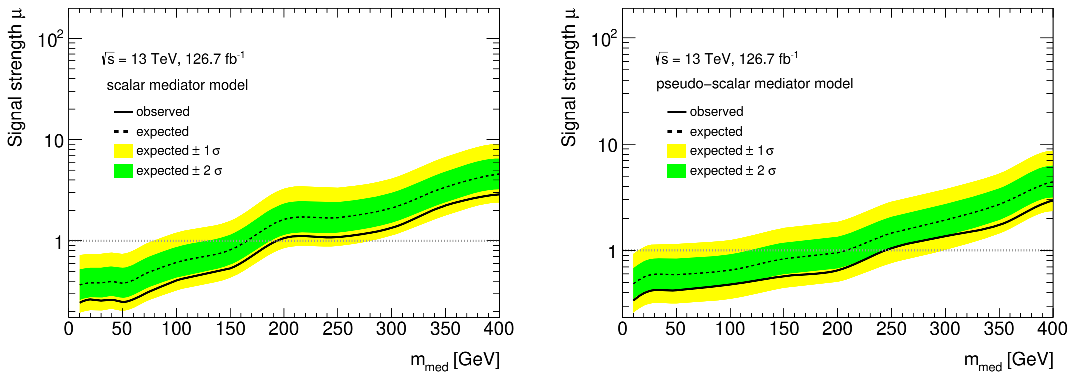
6.8.4 Exclusion limits

As no evidence for the production of DM is found in $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) final state, 95% CL$_s$ exclusion limits on the signal strength μ are set for the scalar and the pseudo-scalar mediator simplified models as functions of the mediator mass m_{med} for a fixed DM mass $m_\chi = 1$ GeV.

Figure 6.8.3 shows the expected and the observed upper limits for scalar mediator and pseudo-scalar mediator signals. For the given choice of the coupling $g = g_\chi = g_q = 1$, scalar mediator simplified models with m_{med} up to 190 GeV and $m_\chi = 1$ GeV are excluded at 95% CL, while pseudo-scalar simplified models with m_{med} up to 240 GeV and $m_\chi = 1$ GeV are excluded at 95% CL.

Comparable exclusion limits on the DM production in fully-hadronic $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ final state are obtained from the DM reinterpretation of the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger based stop (SUSY partner of top quark) search and the b-jet trigger based $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search. Provided that the event selections of these two analyses are mostly orthogonal, significant boost in sensitivity can be expected after the combination. Results from DM searches in semi-leptonic and dileptonic $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ final states, or even the $E_{\mathrm{T}}^{\mathrm{miss}} + b\bar{b}$ and $E_{\mathrm{T}}^{\mathrm{miss}} + t$ searches, can also be integrated using full Run 2 data collected by the ATLAS detector.

Additionally, the scattering between the pseudo-scalar particle and a nucleus is strongly velocity dependent and vanishes in the non-relativistic limit [196]. Therefore, the direct



(a) (b)

Figure 6.8.3: Exclusion limits on the Dark Matter production in association with fully-hadronic $t\bar{t}$ for (a) the scalar mediator simplified model and (b) the pseudo-scalar simplified model as functions of the mediator mass m_{med} for a fixed DM mass $m_{\chi} = 1$ GeV. The couplings of the mediator to the SM and DM particles are fixed to $g = g_{\chi} = g_{q} = 1$, following the recommendations of the dark matter forum [147]. The solid (dashed) black curve shows the median of the observed (expected) exclusion limit, while the filled green (yellow) band denotes the $\pm 1\sigma$ ($\pm 2\sigma$) uncertainties on the expected result.

detection experiments are expected to have no constrain power on pseudo-scalar models, which adds more importance on the $E_{\text{T}}^{\text{miss}} + t\bar{t}$ searches.

Chapter 7

Conclusions

Unraveling the particle nature of Dark Matter has always been one of the main focuses of the LHC program. The accumulation of LHC Run 2 pp collision data at the center-of-mass energy of 13 TeV grants huge potential to searches for DM production in various final states as well as different kinematic phase spaces, while the reconstruction and calibration of various physical objects become crucial to the sensitivity of the statistical analyses.

The variable-radius algorithms are specially designed to reconstruct hadronically decaying objects in extremely boosted event topologies. Compared to the traditional large-R jets, VR jets have effective radius shrinking in proportion to $1/p_{\mathrm{T}}$, resulting in significantly better suppression of pile-up contamination and soft radiation at $p_{\mathrm{T}} > 1$ TeV. Studies on the jet mass scale calibration for variable-radius calorimeter jets are documented as part of the dissertation, while a numerical inversion method is used to calculate the calibration factors as functions of the reconstructed properties, taking MC simulated dijet events as input. To address the validity of the calibration, the closure performance is checked in comparison with two sets of alternative signal and background samples. Moreover, good closure is reached for all jet mass definitions, with the combined mass outperforming both the calorimeter-based mass and the track-assisted mass in terms of response closure and mass resolution over the full p_{T} range. These derived JMS calibration functions enable the usage of VR calorimeter jets in full Run 2 ATLAS analyses.

Aside from the calibration studies, two independent Dark Matter searches are presented in this thesis, one targeting DM production in association with a hadronically decaying Standard Model vector boson, namely the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search, and the other targeting DM production in association with a top quark pair and medium missing transverse energy in 0 lepton final state, often referred to as the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search. In both analyses, the simplified models of DM are used to model the interaction between the Standard Model particles and the dark sector. The results of the statistical analyses are expressed in terms of constraints on the parameter space of the vector mediator or the scalar/pseudo-scalar mediator simplified models.

The $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search is performed in events having at least one large-R jet or a pair of small-R jets with mass (and substructure) compatible with a hadronic W or Z boson decay and large missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$. Collision data recorded by the ATLAS detector in 2015 and 2016 are included in the analysis, corresponding to an integrated luminosity of 36.1 fb^{-1}. The selected events are categorized into non-overlapping regions based on $E_{\mathrm{T}}^{\mathrm{miss}}$, lepton multiplicity and the number of b-tagged jets to enhance the purity of DM signals and the dominant SM background components. All analysis regions are considered in a combined profile likelihood fit on the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution, where the vector mediator simplified models are tested against the null hypothesis by assuming the presence of DM signals on top of the SM prediction. Various sources of experimental and theoretical systematic uncertainties are taken into consideration, with their impacts on the sensitivity of the analysis accounted by the nuisance parameters in the statistical model. Moreover, the inclusion of dedicated control regions assists to constrain the normalization of W+jets, Z+jets and $t\bar{t}$ processes as well as the corresponding shapes of the $E_{\mathrm{T}}^{\mathrm{miss}}$ discriminant using data measurements in regions enriched in these backgrounds, allowing for partial cancellation of the relative uncertainties.

As no significant excess over the Standard Model prediction is observed, exclusion limits on DM pair production are set in the two-dimensional m_χ-m_{med} plane with fixed couplings of $g_\chi = 1$, $g_q = 0.25$ and $g_l = 0$. For the vector mediator simplified model in which the DM is produced via an s-channel exchange of a vector mediator, m_{med} of up to 650 GeV are excluded for Dark Matter masses m_χ of up to 250 GeV at 95% confidence level, in agreement with the expected exclusion of m_{med} of up to 700 GeV for m_χ of up to 250 GeV. The sensitivity is greatly improved compared to the previously performed DM search in the same hadronic $E_{\mathrm{T}}^{\mathrm{miss}} + V$ final state using 3.2 fb^{-1} of pp collision data at $\sqrt{s} = 13$ TeV [146]. These results were published in Ref. [1].

Although the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search does not provide the best exclusion limit among the $E_{\mathrm{T}}^{\mathrm{miss}} + X$ searches in the context of the vector mediator simplified models, its sensitivity might be greatly enhanced when being interpreted in terms of some of the more complete models, e.g. the two-Higgs-doublet models [197] with a vector mediator [198]. Moreover, the $E_{\mathrm{T}}^{\mathrm{miss}} + X$ searches allow to set limits in the kinematic phase space at low m_χ, where current direct detection experiments have no reach.

The $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search, on the other hand, aims to explore a kinematic topology which has never been carefully looked into at the ATLAS collaboration. Collision data recorded from 2016 to 2018 with an integrated luminosity of 126.7 fb^{-1} are analyzed, while the 2 b-jet + 2 jet triggers, instead of the most commonly used $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers, are included as part of the event selection in order to exploit the DM signatures at regions where medium or low $E_{\mathrm{T}}^{\mathrm{miss}}$ is expected. Events containing at least four small-R jets with kinematic patterns compatible with a fully-hadronically decaying top pair are selected and categorized into non-overlapping regions based on the number of large-R jets, the invariant mass of the highest mass large-R jet (if exists) and the lepton multiplicity. These regions are enriched in DM signals and the dominant SM backgrounds. All analysis regions, except for the validation regions, are included in a combined profile likelihood fit on the event yield, where the scalar/pseudo-scalar mediator simplified models are tested against the null hypothesis. Similar to the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) search, the impact of experimental and theoretical systematic uncertainties are propagated in terms of the nuisance parameters and the normalization of $t\bar{t}$, $t\bar{t} + b$, single top and Z+jets processes are extracted from the high purity data samples.

The results of the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search are in agreement with the SM predictions and are translated into exclusion limits on the mediator masses m_{med} for a fixed Dark Matter mass of 1 GeV and with fixed couplings of $g = g_\chi = g_q = g_l = 1$. For the scalar (pseudo-scalar) mediator simplified model in which the DM is produced via an s-channel exchange of a scalar (pseudo-scalar) mediator, m_{med} up to 190 (240) GeV are excluded at 95% confidence level. Slightly weaker exclusion of m_{med} of up to 170 GeV and 220 GeV are expected for the scalar and the pseudo-scalar model, respectively.

Comparable limits are given by the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger based SUSY search reinterpretation and the b-jet trigger based $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search. Since the two analyses focus on different kinematic topologies and their event selections are mostly orthogonal, combination is proposed to provide enhanced sensitivity. Given that the direct detection experiments are expected to have minor exclusion on pseudo-scalar models due to velocity suppression, the DM searches in $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ final states are of great importance also in the future LHC runs.

Despite the fact that no evidence of Dark Matter pair production has been observed at the LHC, foundations of finding this missing piece of the puzzle have been laid with the combined effort of many. At the moment, the sensitivity of many analyses, e.g. the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ (fully-hadronic) search, are still statistically limited, indicating boosts in exclusion power to be expected with higher integrated luminosity. New techniques have been developed to provide more accurate modeling of physical objects and to reduce the systematic uncertainties introduced during the MC simulation, with the VR algorithms being one of these attempts. In the upcoming Run 3, the ATLAS detector will continue to collect pp collision data with an increased center-of-mass energy of 14 TeV and a higher collision rate, allowing for accesses to a large variety of exotic decay channels in search of DM.

The search for Dark Matter at hadron colliders has just commenced and the most intriguing part of it is that nobody knows where the indication of the nature of Dark Matter may hide. Hopefully, the work documented in this thesis can be of service to the future discovery of physics beyond the Standard Model.

Bibliography

[1] M. Aaboud et al. "Search for dark matter in events with a hadronically decaying vector boson and missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *JHEP* 10 (2018), p. 180. DOI: 10.1007/JHEP10(2018)180. arXiv: 1807.11471 [hep-ex].

[2] S. L. Glashow. "Partial Symmetries of Weak Interactions". In: *Nucl. Phys.* 22 (1961), pp. 579–588. DOI: 10.1016/0029-5582(61)90469-2.

[3] Steven Weinberg. "A Model of Leptons". In: *Phys. Rev. Lett.* 19 (1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264.

[4] Abdus Salam. "Weak and Electromagnetic Interactions". In: *Conf. Proc.* C680519 (1968), pp. 367–377. DOI: 10.1142/9789812795915_0034.

[5] Gerard 't Hooft and M. J. G. Veltman. "Regularization and Renormalization of Gauge Fields". In: *Nucl. Phys.* B44 (1972), pp. 189–213. DOI: 10.1016/0550-3213(72)90279-9.

[6] G. Arnison et al. "Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at s**(1/2) = 540-GeV". In: *Phys. Lett.* 122B (1983). [,611(1983)], pp. 103–116. DOI: 10.1016/0370-2693(83)91177-2.

[7] G. Arnison et al. "Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV/c**2 at the CERN SPS Collider". In: *Phys. Lett.* 126B (1983). [,7.55(1983)], pp. 398–410. DOI: 10.1016/0370-2693(83)90188-0.

[8] Christoph Berger et al. "Jet Analysis of the Y (9.46) Decay Into Charged Hadrons". In: *Phys. Lett.* 82B (1979), pp. 449–455. DOI: 10.1016/0370-2693(79)90265-X.

[9] S. Abachi et al. "Observation of the top quark". In: *Phys. Rev. Lett.* 74 (1995), pp. 2632–2637. DOI: 10.1103/PhysRevLett.74.2632. arXiv: hep-ex/9503003 [hep-ex].

[10] F. Abe et al. "Observation of top quark production in $\bar{p}p$ collisions". In: *Phys. Rev. Lett.* 74 (1995), pp. 2626–2631. DOI: 10.1103/PhysRevLett.74.2626. arXiv: hep-ex/9503002 [hep-ex].

[11] J. J. Aubert et al. "Experimental Observation of a Heavy Particle J". In: *Phys. Rev. Lett.* 33 (1974), pp. 1404–1406. DOI: 10.1103/PhysRevLett.33.1404.

[12] J. E. Augustin et al. "Discovery of a Narrow Resonance in e^+e^- Annihilation". In: *Phys. Rev. Lett.* 33 (1974). [Adv. Exp. Phys.5,141(1976)], pp. 1406–1408. DOI: 10.1103/PhysRevLett.33.1406.

[13] Georges Aad et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Phys. Lett.* B716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[14] Serguei Chatrchyan et al. "Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC". In: *Phys. Lett.* B716 (2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].

[15] S. N. Bose. "Planck's law and light quantum hypothesis". In: *Z. Phys.* 26 (1924), pp. 178–181. DOI: 10.1007/BF01327326.

[16] Enrico Fermi. "Sulla quantizzazione del gas perfetto monoatomico". In: *Rendiconti Lincei* 145 (1926).

[17] Paul A. M. Dirac. "On the Theory of quantum mechanics". In: *Proc. Roy. Soc. Lond.* A112 (1926), pp. 661–677. DOI: 10.1098/rspa.1926.0133.

[18] M. Tanabashi et al. "Review of Particle Physics". In: *Phys. Rev.* D98.3 (2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001.

[19] Emmy Noether. "Invariant Variation Problems". In: *Gott. Nachr.* 1918 (1918). [Transp. Theory Statist. Phys.1,186(1971)], pp. 235–257. DOI: 10.1080/00411457108231446. arXiv: physics/0503066 [physics].

[20] Sheldon L. Glashow. "The renormalizability of vector meson interactions". In: *Nucl. Phys.* 10 (1959), pp. 107–117. DOI: 10.1016/0029-5582(59)90196-8.

[21] E. Kearns, T. Kajita, and Y. Totsuka. "Detecting massive neutrinos". In: *Sci. Am.* 281N2 (1999). [Spektrum Wiss. Dossier2003N1,64(2003)], pp. 48–55.

[22] Q. R. Ahmad et al. "Measurement of the rate of $\nu_e + d \to p + p + e^-$ interactions produced by ^8B solar neutrinos at the Sudbury Neutrino Observatory". In: *Phys. Rev. Lett.* 87 (2001), p. 071301. DOI: 10.1103/PhysRevLett.87.071301. arXiv: nucl-ex/0106015 [nucl-ex].

[23] P. A. R. Ade et al. "Planck 2013 results. I. Overview of products and scientific results". In: *Astron. Astrophys.* 571 (2014), A1. DOI: 10.1051/0004-6361/201321529. arXiv: 1303.5062 [astro-ph.CO].

[24] J H Oort. "The force exerted by the stellar system in the direction perpendicular to the galactic plane and some related problems". In: *Bull. Astron. Inst. Netherlands* 6 (1932), pp. 249–287.

[25] F. Zwicky. "Die Rotverschiebung von extragalaktischen Nebeln". In: *Helv. Phys. Acta* 6 (1933). [Gen. Rel. Grav.41,207(2009)], pp. 110–127. DOI: 10.1007/s10714-008-0707-4.

[26] Vera C. Rubin and W. Kent Ford Jr. "Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions". In: *Astrophys. J.* 159 (1970), pp. 379–403. DOI: 10.1086/150317.

[27] K. G. Begeman, A. H. Broeils, and R. H. Sanders. "Extended rotation curves of spiral galaxies: Dark haloes and modified dynamics". In: *Mon. Not. Roy. Astron. Soc.* 249 (1991), p. 523.

[28] Douglas Clowe et al. "A direct empirical proof of the existence of dark matter". In: *Astrophys. J.* 648 (2006), pp. L109–L113. DOI: 10.1086/508162. arXiv: astro-ph/0608407 [astro-ph].

[29] R. Adam et al. "Planck 2015 results. I. Overview of products and scientific results". In: *Astron. Astrophys.* 594 (2016), A1. DOI: 10.1051/0004-6361/201527101. arXiv: 1502.01582 [astro-ph.CO].

[30] N. Aghanim et al. "Planck 2015 results. XI. CMB power spectra, likelihoods, and robustness of parameters". In: *Astron. Astrophys.* 594 (2016), A11. DOI: 10.1051/0004-6361/201526926. arXiv: 1507.02704 [astro-ph.CO].

[31] P. A. R. Ade et al. "Planck 2015 results. XIII. Cosmological parameters". In: *Astron. Astrophys.* 594 (2016), A13. DOI: 10.1051/0004-6361/201525830. arXiv: 1502.01589 [astro-ph.CO].

[32] Albert Einstein. "Lens-Like Action of a Star by the Deviation of Light in the Gravitational Field". In: *Science* 84 (1936), pp. 506–507. DOI: 10.1126/science.84.2188.506.

[33] G. Gamow. "The Evolution of the Universe". In: *Nature* 162.4122 (1948), pp. 680–682. DOI: 10.1038/162680a0.

[34] Ralph A. Alpher and Robert Herman. "Evolution of the Universe". In: *Nature* 162.4124 (1948), pp. 774–775. DOI: 10.1038/162774b0.

[35] Arno A. Penzias and Robert Woodrow Wilson. "A Measurement of excess antenna temperature at 4080-Mc/s". In: *Astrophys. J.* 142 (1965), pp. 419–421. DOI: 10.1086/148307.

[36] C. L. Bennett et al. "Four year COBE DMR cosmic microwave background observations: Maps and basic results". In: *Astrophys. J.* 464 (1996), pp. L1–L4. DOI: 10.1086/310075. arXiv: astro-ph/9601067 [astro-ph].

[37] JR Bond et al. "Dark Matter Shocked Pancakes". In: *Formation and Evolution of Galaxies and Large Structures in the Universe*. Vol. 117. 1984, pp. 87–99.

[38] G Kauffmann, Simon D. M. White, and B. Guiderdoni. "The Formation and Evolution of Galaxies Within Merging Dark Matter Haloes". In: *Mon. Not. Roy. Astron. Soc.* 264 (1993), p. 201.

[39] A. Tikhonov and A. Klypin. "The emptiness of voids: yet another over-abundance problem for the LCDM model". In: *Mon. Not. Roy. Astron. Soc.* 395 (2009), p. 1915. DOI: 10.1111/j.1365-2966.2009.14686.x. arXiv: 0807.0924 [astro-ph].

[40] Fabio Governato et al. "At the heart of the matter: the origin of bulgeless dwarf galaxies and Dark Matter cores". In: *Nature* 463 (2010), pp. 203–206. DOI: 10.1038/nature08640. arXiv: 0911.2237 [astro-ph.CO].

[41] Stephane Colombi, Scott Dodelson, and Lawrence M. Widrow. "Large scale structure tests of warm dark matter". In: *Astrophys. J.* 458 (1996), p. 1. DOI: 10.1086/176788. arXiv: astro-ph/9505029 [astro-ph].

[42] Simon D. M. White, C. S. Frenk, and M. Davis. "Clustering in a Neutrino Dominated Universe". In: *Astrophys. J.* 274 (1983). [,80(1984)], pp. L1–L5. DOI: 10.1086/161425.

[43] R. D. Peccei and Helen R. Quinn. "CP Conservation in the Presence of Instantons". In: *Phys. Rev. Lett.* 38 (1977). [,328(1977)], pp. 1440–1443. DOI: 10.1103/PhysRevLett.38.1440.

[44] O. Erken et al. "Axion Dark Matter and Cosmological Parameters". In: *Phys. Rev. Lett.* 108 (2012), p. 061304. DOI: 10.1103/PhysRevLett.108.061304. arXiv: 1104.4507 [astro-ph.CO].

[45] Peter W. Graham et al. "Experimental Searches for the Axion and Axion-Like Particles". In: *Ann. Rev. Nucl. Part. Sci.* 65 (2015), pp. 485–514. DOI: 10.1146/annurev-nucl-102014-022120. arXiv: 1602.00039 [hep-ex].

[46] N. Du et al. "A Search for Invisible Axion Dark Matter with the Axion Dark Matter Experiment". In: *Phys. Rev. Lett.* 120.15 (2018), p. 151301. DOI: 10.1103/PhysRevLett.120.151301. arXiv: 1804.05750 [hep-ex].

[47] Leanne D. Duffy and Karl van Bibber. "Axions as Dark Matter Particles". In: *New J. Phys.* 11 (2009), p. 105008. DOI: 10.1088/1367-2630/11/10/105008. arXiv: 0904.3346 [hep-ph].

[48] Hans Peter Nilles. "Supersymmetry, Supergravity and Particle Physics". In: *Phys. Rept.* 110 (1984), pp. 1–162. DOI: 10.1016/0370-1573(84)90008-5.

[49] Keith R. Dienes, Emilian Dudas, and Tony Gherghetta. "Grand unification at intermediate mass scales through extra dimensions". In: *Nucl. Phys.* B537 (1999), pp. 47–108. DOI: 10.1016/S0550-3213(98)00669-5. arXiv: hep-ph/9806292 [hep-ph].

[50] Jay Hubisz and Patrick Meade. "Phenomenology of the littlest Higgs with T-parity". In: *Phys. Rev.* D71 (2005), p. 035016. DOI: 10.1103/PhysRevD.71.035016. arXiv: hep-ph/0411264 [hep-ph].

[51] Kim Griest and Marc Kamionkowski. "Unitarity Limits on the Mass and Radius of Dark Matter Particles". In: *Phys. Rev. Lett.* 64 (1990), p. 615. DOI: 10.1103/PhysRevLett.64.615.

[52] R. Agnese et al. "Silicon Detector Dark Matter Results from the Final Exposure of CDMS II". In: *Phys. Rev. Lett.* 111.25 (2013), p. 251301. DOI: 10.1103/PhysRevLett.111.251301. arXiv: 1304.4279 [hep-ex].

[53] G. Angloher et al. "Results from 730 kg days of the CRESST-II Dark Matter Search". In: *Eur. Phys. J.* C72 (2012), p. 1971. DOI: 10.1140/epjc/s10052-012-1971-8. arXiv: 1109.0702 [astro-ph.CO].

[54] R. Bernabei et al. "New results from DAMA/LIBRA". In: *Eur. Phys. J.* C67 (2010), pp. 39–49. DOI: 10.1140/epjc/s10052-010-1303-9. arXiv: 1002.1028 [astro-ph.GA].

[55] D. S. Akerib et al. "First results from the LUX dark matter experiment at the Sanford Underground Research Facility". In: *Phys. Rev. Lett.* 112 (2014), p. 091303. DOI: 10.1103/PhysRevLett.112.091303. arXiv: 1310.8214 [astro-ph.CO].

[56] E. Aprile et al. "First Dark Matter Search Results from the XENON1T Experiment". In: *Phys. Rev. Lett.* 119.18 (2017), p. 181301. DOI: 10.1103/PhysRevLett.119.181301. arXiv: 1705.06655 [astro-ph.CO].

[57] P. Agnes et al. "First Results from the DarkSide-50 Dark Matter Experiment at Laboratori Nazionali del Gran Sasso". In: *Phys. Lett.* B743 (2015), pp. 456–466. DOI: 10.1016/j.physletb.2015.03.012. arXiv: 1410.0653 [astro-ph.CO].

[58] W. B. Atwood et al. "The Large Area Telescope on the Fermi Gamma-ray Space Telescope Mission". In: *Astrophys. J.* 697 (2009), pp. 1071–1102. DOI: 10.1088/0004-637X/697/2/1071. arXiv: 0902.1089 [astro-ph.IM].

[59] F. Donato et al. "Constraints on WIMP Dark Matter from the High Energy PAMELA $\bar{p}/p$ data". In: *Phys. Rev. Lett.* 102 (2009), p. 071301. DOI: 10.1103/PhysRevLett.102.071301. arXiv: 0810.5292 [astro-ph].

[60] F. Aharonian et al. "H.E.S.S. observations of the Galactic Center region and their possible dark matter interpretation". In: *Phys. Rev. Lett.* 97 (2006). [Erratum: Phys. Rev. Lett.97,249901(2006)], p. 221102. DOI: 10.1103/PhysRevLett.97.221102,10.1103/PhysRevLett.97.249901. arXiv: astro-ph/0610509 [astro-ph].

[61] M. G. Aartsen et al. "Search for dark matter annihilations in the Sun with the 79-string IceCube detector". In: *Phys. Rev. Lett.* 110.13 (2013), p. 131302. DOI: 10.1103/PhysRevLett.110.131302. arXiv: 1212.4097 [astro-ph.HE].

[62] Morad Aaboud et al. "Search for large missing transverse momentum in association with one top-quark in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *JHEP* 05 (2019), p. 041. DOI: 10.1007/JHEP05(2019)041. arXiv: 1812.09743 [hep-ex].

[63] Albert M Sirunyan et al. "Search for dark matter produced in association with a single top quark or a top quark pair in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *JHEP* 03 (2019), p. 141. DOI: 10.1007/JHEP03(2019)141. arXiv: 1901.01553 [hep-ex].

[64] Roel Aaij et al. "Search for Dark Photons Produced in 13 TeV pp Collisions". In: *Phys. Rev. Lett.* 120.6 (2018), p. 061801. DOI: 10.1103/PhysRevLett.120.061801. arXiv: 1710.02867 [hep-ex].

[65] Morad Aaboud et al. "Search for dark matter and other new phenomena in events with an energetic jet and large missing transverse momentum using the ATLAS detector". In: *JHEP* 01 (2018), p. 126. DOI: 10.1007/JHEP01(2018)126. arXiv: 1711.03301 [hep-ex].

[66] A. M. Sirunyan et al. "Search for new physics in final states with an energetic jet or a hadronically decaying W or Z boson and transverse momentum imbalance at $\sqrt{s} = 13$ TeV". In: *Phys. Rev.* D97.9 (2018), p. 092005. DOI: 10.1103/PhysRevD.97.092005. arXiv: 1712.02345 [hep-ex].

[67] Morad Aaboud et al. "Search for dark matter at $\sqrt{s} = 13$ TeV in final states containing an energetic photon and large missing transverse momentum with the ATLAS detector". In: *Eur. Phys. J.* C77.6 (2017), p. 393. DOI: 10.1140/epjc/s10052-017-4965-8. arXiv: 1704.03848 [hep-ex].

[68] Albert M Sirunyan et al. "Search for new physics in the monophoton final state in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *JHEP* 10 (2017), p. 073. DOI: 10.1007/JHEP10(2017)073. arXiv: 1706.03794 [hep-ex].

[69] M. Aaboud et al. "Search for an invisibly decaying Higgs boson or dark matter candidates produced in association with a Z boson in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Lett.* B776 (2018), pp. 318–337. DOI: 10.1016/j.physletb.2017.11.049. arXiv: 1708.09624 [hep-ex].

[70] A. M. Sirunyan et al. "Search for new physics in events with a leptonically decaying Z boson and a large transverse momentum imbalance in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J.* C78.4 (2018), p. 291. DOI: 10.1140/epjc/s10052-018-5740-1. arXiv: 1711.00431 [hep-ex].

[71] Morad Aaboud et al. "Search for dark matter produced in association with bottom or top quarks in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector". In: *Eur. Phys. J.* C78.1 (2018), p. 18. DOI: 10.1140/epjc/s10052-017-5486-1. arXiv: 1710.11412 [hep-ex].

[72] Morad Aaboud et al. "Search for Dark Matter Produced in Association with a Higgs Boson Decaying to $b\bar{b}$ using 36 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector". In: *Phys. Rev. Lett.* 119.18 (2017), p. 181804. DOI: 10.1103/PhysRevLett.119.181804. arXiv: 1707.01302 [hep-ex].

[73] Albert M Sirunyan et al. "Search for dark matter particles produced in association with a Higgs boson in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: (2019). arXiv: 1908.01713 [hep-ex].

[74] Giorgio Busoni et al. "On the Validity of the Effective Field Theory for Dark Matter Searches at the LHC". In: *Phys. Lett.* B728 (2014), pp. 412–421. DOI: 10.1016/j.physletb.2013.11.069. arXiv: 1307.2253 [hep-ph].

[75] Nima Arkani-Hamed et al. "Supersymmetry and the LHC inverse problem". In: *JHEP* 08 (2006), p. 070. DOI: 10.1088/1126-6708/2006/08/070. arXiv: hep-ph/0512190 [hep-ph].

[76] Enrico Morgante. "Simplified Dark Matter Models". In: *Adv. High Energy Phys.* 2018 (2018), p. 5012043. DOI: 10.1155/2018/5012043. arXiv: 1804.01245 [hep-ph].

[77] G. D'Ambrosio et al. "Minimal flavor violation: An Effective field theory approach". In: *Nucl. Phys.* B645 (2002), pp. 155–187. DOI: 10.1016/S0550-3213(02)00836-2. arXiv: hep-ph/0207036 [hep-ph].

[78] Morad Aaboud et al. "Search for a scalar partner of the top quark in the jets plus missing transverse momentum final state at $\sqrt{s}$=13 TeV with the ATLAS detector". In: *JHEP* 12 (2017), p. 085. DOI: 10.1007/JHEP12(2017)085. arXiv: 1709.04183 [hep-ex].

[79] Oliver S. Bruning et al. "LHC Design Report Vol.1: The LHC Main Ring". In: (2004).

[80] O. Buning et al. "LHC Design Report. 2. The LHC infrastructure and general services". In: (2004).

[81] M. Benedikt et al. "LHC Design Report. 3. The LHC injector chain". In: (2004).

[82] "LEP Design Report Vol.1". In: (1983).

[83] "LEP Design Report: Vol.2. The LEP Main Ring". In: (1984).

[84] Fabienne Marcastel. "CERN's Accelerator Complex. La chaîne des accélérateurs du CERN". In: (Oct. 2013). URL: https://cds.cern.ch/record/1621583.

[85] Werner Herr, B. J. Holzer, and B. Muratori. "Concept of Luminosity". In: *Elementary Particles - Accelerators and Colliders*. Ed. by S. Myers and H. Schopper. 2013. DOI: 10.1007/978-3-642-23053-0_9.

[86] Morad Aaboud et al. "Luminosity Public Results Run2". In: (2019). URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.

[87] G. Aad et al. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *JINST* 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003.

[88] Joao Pequenao and Paul Schaffner. "How ATLAS detects particles: diagram of particle paths in the detector". Jan. 2013. URL: https://cds.cern.ch/record/1505342.

[89] A. Yamamoto et al. "The ATLAS central solenoid". In: *Nucl. Instrum. Meth.* A584 (2008), pp. 53–74. DOI: 10.1016/j.nima.2007.09.047.

[90] Joao Pequenao. "Computer generated image of the ATLAS inner detector". Mar. 2008. URL: https://cds.cern.ch/record/1095926.

[91] *Track Reconstruction Performance of the ATLAS Inner Detector at $\sqrt{s} = 13$ TeV*. Tech. rep. ATL-PHYS-PUB-2015-018. Geneva: CERN, July 2015. URL: https://cds.cern.ch/record/2037683.

[92] *Expected performance of the ATLAS b-tagging algorithms in Run-2*. Tech. rep. ATL-PHYS-PUB-2015-022. Geneva: CERN, July 2015. URL: http://cds.cern.ch/record/2037697.

[93] Joao Pequenao. "Computer Generated image of the ATLAS calorimeter". Mar. 2008. URL: https://cds.cern.ch/record/1095927.

[94] M. Aharrouche et al. "Energy linearity and resolution of the ATLAS electromagnetic barrel calorimeter in an electron test-beam". In: *Nucl. Instrum. Meth.* A568 (2006), pp. 601–623. DOI: 10.1016/j.nima.2006.07.053. arXiv: physics/0608012 [physics].

[95] T. Barillari. "The ATLAS liquid argon hadronic end-cap calorimeter: Construction and selected beam test results". In: *Nucl. Phys. Proc. Suppl.* 150 (2006). [,102(2004)], pp. 102–105. DOI: 10.1016/j.nuclphysbps.2004.10.087. arXiv: physics/0407026 [physics.ins-det].

[96] Joao Pequenao. "Computer generated image of the ATLAS Muons subsystem". Mar. 2008. URL: https://cds.cern.ch/record/1095929.

[97] Morad Aaboud et al. "Performance of the ATLAS Trigger System in 2015". In: *Eur. Phys. J.* C77.5 (2017), p. 317. DOI: 10.1140/epjc/s10052-017-4852-3. arXiv: 1611.09661 [hep-ex].

[98] Morad Aaboud et al. "Data Quality Public Results Run2". In: (2019). URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/RunStatsPublicResults2010.

[99] *Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment*. Tech. rep. ATLAS-CONF-2012-042. Geneva: CERN, Mar. 2012. URL: http://cds.cern.ch/record/1435196.

[100] *The Optimization of ATLAS Track Reconstruction in Dense Environments*. Tech. rep. ATL-PHYS-PUB-2015-006. Geneva: CERN, Mar. 2015. URL: https://cds.cern.ch/record/2002609.

[101] Morad Aaboud et al. "Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J.* C79.8 (2019), p. 639. DOI: 10.1140/epjc/s10052-019-7140-6. arXiv: 1902.04655 [physics.ins-det].

[102] W Lampl et al. *Calorimeter Clustering Algorithms: Description and Performance*. Tech. rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003. Geneva: CERN, Apr. 2008. URL: https://cds.cern.ch/record/1099735.

[103] *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*. Tech. rep. ATLAS-CONF-2016-024. Geneva: CERN, June 2016. URL: https://cds.cern.ch/record/2157687.

[104] Johannes Josef Junggeburth. *Muon identification and reconstruction efficiencies in full Run-2 dataset*. Tech. rep. ATL-COM-PHYS-2019-1001. Geneva: CERN, Aug. 2019. URL: https://cds.cern.ch/record/2685295.

[105] Georges Aad et al. “Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s} = 13$ TeV”. In: *Eur. Phys. J.* C76.5 (2016), p. 292. DOI: 10.1140/epjc/s10052-016-4120-y. arXiv: 1603.05598 [hep-ex].

[106] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. “FastJet User Manual”. In: *Eur. Phys. J.* C72 (2012), p. 1896. DOI: 10.1140/epjc/s10052-012-1896-2. arXiv: 1111.6097 [hep-ph].

[107] S. Catani et al. “Longitudinally invariant K_t clustering algorithms for hadron hadron collisions”. In: *Nucl. Phys.* B406 (1993), pp. 187–224. DOI: 10.1016/0550-3213(93)90166-M.

[108] Stephen D. Ellis and Davison E. Soper. “Successive combination jet algorithm for hadron collisions”. In: *Phys. Rev.* D48 (1993), pp. 3160–3166. DOI: 10.1103/PhysRevD.48.3160. arXiv: hep-ph/9305266 [hep-ph].

[109] Yuri L. Dokshitzer et al. “Better jet clustering algorithms”. In: *JHEP* 08 (1997), p. 001. DOI: 10.1088/1126-6708/1997/08/001. arXiv: hep-ph/9707323 [hep-ph].

[110] M. Wobisch and T. Wengler. “Hadronization corrections to jet cross-sections in deep inelastic scattering”. In: (1998), pp. 270–279. arXiv: hep-ph/9907280 [hep-ph].

[111] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. “The anti-k_t jet clustering algorithm”. In: *JHEP* 04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph].

[112] M. Aaboud et al. “Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *Phys. Rev.* D96.7 (2017), p. 072002. DOI: 10.1103/PhysRevD.96.072002. arXiv: 1703.09665 [hep-ex].

[113] Georges Aad et al. “Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV”. In: *Eur. Phys. J.* C73.3 (2013), p. 2304. DOI: 10.1140/epjc/s10052-013-2304-2. arXiv: 1112.6426 [hep-ex].

[114] The ATLAS collaboration. “Determination of the jet energy scale and resolution at ATLAS using Z/γ–jet events in data at $\sqrt{s} = 8$ TeV”. In: (2015).

[115] The ATLAS collaboration. “Data-driven determination of the energy scale and resolution of jets reconstructed in the ATLAS calorimeters using dijet and multijet events at $\sqrt{s} = 8$ TeV”. In: (2015).

[116] *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*. Tech. rep. ATL-PHYS-PUB-2017-013. Geneva: CERN, July 2017. URL: https://cds.cern.ch/record/2273281.

[117] Giacinto Piacquadio and Christian Weiser. “A new inclusive secondary vertex algorithm for b-jet tagging in ATLAS”. In: *J. Phys. Conf. Ser.* 119 (2008), p. 032032. DOI: 10.1088/1742-6596/119/3/032032.

[118] R. Fruhwirth. “Application of Kalman filtering to track and vertex fitting”. In: *Nucl. Instrum. Meth.* A262 (1987), pp. 444–450. DOI: 10.1016/0168-9002(87)90887-4.

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

[120] Andreas Hocker et al. "TMVA - Toolkit for Multivariate Data Analysis". In: (2007). arXiv: physics/0703039 [physics.data-an].

[121] *Measurement of b-tagging Efficiency of c-jets in $t\bar{t}$ Events Using a Likelihood Approach with the ATLAS Detector*. Tech. rep. ATLAS-CONF-2018-001. Geneva: CERN, Mar. 2018. URL: https://cds.cern.ch/record/2306649.

[122] *Calibration of light-flavour jet b-tagging rates on ATLAS proton-proton collision data at $\sqrt{s} = 13$ TeV*. Tech. rep. ATLAS-CONF-2018-006. Geneva: CERN, Apr. 2018. URL: https://cds.cern.ch/record/2314418.

[123] Morad Aaboud et al. "ATLAS Flavor-Tagging Calibration Results with 80.5 fb^{-1}". In: (2019). URL: http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/FTAG-2019-003/.

[124] Georges Aad et al. "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1". In: *Eur. Phys. J.* C77 (2017), p. 490. DOI: 10.1140/epjc/s10052-017-5004-5. arXiv: 1603.02934 [hep-ex].

[125] Morad Aaboud et al. "In situ calibration of large-radius jet energy and mass in 13 TeV proton-proton collisions with the ATLAS detector". In: *Eur. Phys. J.* C79.2 (2019), p. 135. DOI: 10.1140/epjc/s10052-019-6632-8. arXiv: 1807.09477 [hep-ex].

[126] David Krohn, Jesse Thaler, and Lian-Tao Wang. "Jet Trimming". In: *JHEP* 02 (2010), p. 084. DOI: 10.1007/JHEP02(2010)084. arXiv: 0912.1342 [hep-ph].

[127] Georges Aad et al. "Performance of jet substructure techniques for large-R jets in proton-proton collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector". In: *JHEP* 09 (2013), p. 076. DOI: 10.1007/JHEP09(2013)076. arXiv: 1306.4945 [hep-ex].

[128] *Jet mass reconstruction with the ATLAS Detector in early Run 2 data*. Tech. rep. ATLAS-CONF-2016-035. Geneva: CERN, July 2016. URL: http://cds.cern.ch/record/2200211.

[129] *Measurement of large radius jet mass reconstruction performance at $\sqrt{s} = 8$ TeV using the ATLAS detector*. Tech. rep. ATLAS-CONF-2016-008. Geneva: CERN, Mar. 2016. URL: https://cds.cern.ch/record/2139642.

[130] *Flavor Tagging with Track Jets in Boosted Topologies with the ATLAS Detector*. Tech. rep. ATL-PHYS-PUB-2014-013. Geneva: CERN, Aug. 2014. URL: https://cds.cern.ch/record/1750681.

[131] Matteo Cacciari and Gavin P. Salam. "Pileup subtraction using jet areas". In: *Phys. Lett.* B659 (2008), pp. 119–126. DOI: 10.1016/j.physletb.2007.09.077. arXiv: 0707.1378 [hep-ph].

[132] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "The Catchment Area of Jets". In: *JHEP* 04 (2008), p. 005. DOI: 10.1088/1126-6708/2008/04/005. arXiv: 0802.1188 [hep-ph].

[133] *Expected performance of missing transverse momentum reconstruction for the ATLAS detector at $\sqrt{s} = 13$ TeV*. Tech. rep. ATL-PHYS-PUB-2015-023. Geneva: CERN, July 2015. URL: http://cds.cern.ch/record/2037700.

[134] *Tagging and suppression of pileup jets with the ATLAS detector*. Tech. rep. ATLAS-CONF-2014-018. Geneva: CERN, May 2014. URL: https://cds.cern.ch/record/1700870.

[135] Morad Aaboud et al. "Identification and rejection of pile-up jets at high pseudorapidity with the ATLAS detector". In: *Eur. Phys. J.* C77.9 (2017). [Erratum: Eur. Phys. J.C77,no.10,712(2017)], p. 580. DOI: 10.1140/epjc/s10052-017-5081-5,10.1140/epjc/s10052-017-5245-3. arXiv: 1705.02211 [hep-ex].

[136] Morad Aaboud et al. "2015-2016 TST Systematics and Forward Pileup Suppression in MET". In: (2017). URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/JETM-2017-001/.

[137] *Object-based missing transverse momentum significance in the ATLAS detector*. Tech. rep. ATLAS-CONF-2018-038. Geneva: CERN, July 2018. URL: http://cds.cern.ch/record/2630948.

[138] David Krohn, Jesse Thaler, and Lian-Tao Wang. "Jets with Variable R". In: *JHEP* 06 (2009), p. 059. DOI: 10.1088/1126-6708/2009/06/059. arXiv: 0903.0392 [hep-ph].

[139] Tobias Lapsien, Roman Kogler, and Johannes Haller. "A new tagger for hadronically decaying heavy particles at the LHC". In: *Eur. Phys. J.* C76.11 (2016), p. 600. DOI: 10.1140/epjc/s10052-016-4443-8. arXiv: 1606.04961 [hep-ph].

[140] *Variable Radius, Exclusive-k_T, and Center-of-Mass Subjet Reconstruction for Higgs($\rightarrow b\bar{b}$) Tagging in ATLAS*. Tech. rep. ATL-PHYS-PUB-2017-010. Geneva: CERN, June 2017. URL: https://cds.cern.ch/record/2268678.

[141] *Boosted Object Tagging with Variable-R Jets in the ATLAS Detector*. Tech. rep. ATL-PHYS-PUB-2016-013. Geneva: CERN, July 2016. URL: https://cds.cern.ch/record/2199360.

[142] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. "A Brief Introduction to PYTHIA 8.1". In: *Comput. Phys. Commun.* 178 (2008), pp. 852–867. DOI: 10.1016/j.cpc.2008.01.036. arXiv: 0710.3820 [hep-ph].

[143] Richard D. Ball et al. "Parton distributions with LHC data". In: *Nucl. Phys.* B867 (2013), pp. 244–289. DOI: 10.1016/j.nuclphysb.2012.10.003. arXiv: 1207.1303 [hep-ph].

[144] *ATLAS Run 1 Pythia8 tunes*. Tech. rep. ATL-PHYS-PUB-2014-021. Geneva: CERN, Nov. 2014. URL: http://cds.cern.ch/record/1966419.

[145] Georges Aad et al. "Search for dark matter in events with a hadronically decaying W or Z boson and missing transverse momentum in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *Phys. Rev. Lett.* 112.4 (2014), p. 041802. DOI: 10.1103/PhysRevLett.112.041802. arXiv: 1309.4017 [hep-ex].

[146] Morad Aaboud et al. "Search for dark matter produced in association with a hadronically decaying vector boson in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Lett.* B763 (2016), pp. 251–268. DOI: 10.1016/j.physletb.2016.10.042. arXiv: 1608.02372 [hep-ex].

[147] Daniel Abercrombie et al. "Dark Matter Benchmark Models for Early LHC Run-2 Searches: Report of the ATLAS/CMS Dark Matter Forum". In: *Phys. Dark Univ.* 26 (2019). Ed. by Antonio Boveia et al., p. 100371. DOI: 10.1016/j.dark.2019.100371. arXiv: 1507.00966 [hep-ex].

[148] S. Agostinelli et al. "GEANT4: A Simulation toolkit". In: *Nucl. Instrum. Meth.* A506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.

[149] Johan Alwall et al. "MadGraph 5 : Going Beyond". In: *JHEP* 06 (2011), p. 128. DOI: 10.1007/JHEP06(2011)128. arXiv: 1106.0522 [hep-ph].

[150] Enrico Bothmann et al. "Event Generation with Sherpa 2.2". In: *SciPost Phys.* 7.3 (2019), p. 034. DOI: 10.21468/SciPostPhys.7.3.034. arXiv: 1905.09127 [hep-ph].

[151] Richard D. Ball et al. "Parton distributions for the LHC Run II". In: *JHEP* 04 (2015), p. 040. DOI: 10.1007/JHEP04(2015)040. arXiv: 1410.8849 [hep-ph].

[152] Simone Alioli et al. "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX". In: *JHEP* 06 (2010), p. 043. DOI: 10.1007/JHEP06(2010)043. arXiv: 1002.2581 [hep-ph].

[153] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. "PYTHIA 6.4 Physics and Manual". In: *JHEP* 05 (2006), p. 026. DOI: 10.1088/1126-6708/2006/05/026. arXiv: hep-ph/0603175 [hep-ph].

[154] Peter Zeiler Skands. "Tuning Monte Carlo Generators: The Perugia Tunes". In: *Phys. Rev.* D82 (2010), p. 074018. DOI: 10.1103/PhysRevD.82.074018. arXiv: 1005.3457 [hep-ph].

[155] T. Gleisberg et al. "Event generation with SHERPA 1.1". In: *JHEP* 02 (2009), p. 007. DOI: 10.1088/1126-6708/2009/02/007. arXiv: 0811.4622 [hep-ph].

[156] Hung-Liang Lai et al. "New parton distributions for collider physics". In: *Phys. Rev.* D82 (2010), p. 074024. DOI: 10.1103/PhysRevD.82.074024. arXiv: 1007.2241 [hep-ph].

[157] Morad Aaboud et al. "Missing Energy Trigger Public Results". In: (2019). URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/MissingEtTriggerPublicResults.

[158] Jason Gallicchio and Yang-Ting Chien. "Quit Using Pseudorapidity, Transverse Energy, and Massless Constituents". In: (2018). arXiv: 1802.05356 [hep-ph].

[159] *Selection of jets produced in 13TeV proton-proton collisions with the ATLAS detector*. Tech. rep. ATLAS-CONF-2015-029. Geneva: CERN, July 2015. URL: http://cds.cern.ch/record/2037702.

[160] *Identification of boosted, hadronically-decaying W and Z bosons in $\sqrt{s} = 13$ TeV Monte Carlo Simulations for ATLAS*. Tech. rep. ATL-PHYS-PUB-2015-033. Geneva: CERN, Aug. 2015. URL: https://cds.cern.ch/record/2041461.

[161] Morad Aaboud et al. "Performance of top-quark and W-boson tagging with ATLAS in Run 2 of the LHC". In: *Eur. Phys. J.* C79.5 (2019), p. 375. DOI: 10.1140/epjc/s10052-019-6847-8. arXiv: 1808.07858 [hep-ex].

[162] Georges Aad et al. "Identification of boosted, hadronically decaying W bosons and comparisons with ATLAS data taken at $\sqrt{s} = 8$ TeV". In: *Eur. Phys. J.* C76.3 (2016), p. 154. DOI: 10.1140/epjc/s10052-016-3978-z. arXiv: 1510.05821 [hep-ex].

[163] Andrew J. Larkoski, Ian Moult, and Duff Neill. "Power Counting to Better Jet Observables". In: *JHEP* 12 (2014), p. 009. DOI: 10.1007/JHEP12(2014)009. arXiv: 1409.6298 [hep-ph].

[164] Morad Aaboud et al. "Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton-proton collisions at the LHC". In: *Eur. Phys. J.* C77.5 (2017), p. 332. DOI: 10.1140/epjc/s10052-017-4887-5. arXiv: 1611.10235 [physics.ins-det].

[165] J. Alwall et al. "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations". In: *JHEP* 07 (2014), p. 079. DOI: 10.1007/JHEP07(2014)079. arXiv: 1405.0301 [hep-ph].

[166] M. Bahr et al. "Herwig++ Physics and Manual". In: *Eur. Phys. J.* C58 (2008), pp. 639–707. DOI: 10.1140/epjc/s10052-008-0798-9. arXiv: 0803.0883 [hep-ph].

[167] A. D. Martin et al. "Parton distributions for the LHC". In: *Eur. Phys. J.* C63 (2009), pp. 189–285. DOI: 10.1140/epjc/s10052-009-1072-5. arXiv: 0901.0002 [hep-ph].

[168] Daniel Stump et al. "Inclusive jet production, parton distributions, and the search for new physics". In: *JHEP* 10 (2003), p. 046. DOI: 10.1088/1126-6708/2003/10/046. arXiv: hep-ph/0303013 [hep-ph].

[169] Alexander L. Read. "Presentation of search results: The CL(s) technique". In: *J. Phys.* G28 (2002). [,11(2002)], pp. 2693–2704. DOI: 10.1088/0954-3899/28/10/313.

[170] Glen Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *Eur. Phys. J.* C71 (2011). [Erratum: Eur. Phys. J.C73,2501(2013)], p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0,10.1140/epjc/s10052-013-2501-z. arXiv: 1007.1727 [physics.data-an].

[171] Kyle Cranmer et al. "HistFactory: A tool for creating statistical models for use with RooFit and RooStats". In: (2012).

[172] Jerzy Neyman and Egon Sharpe Pearson. "On the Problem of the Most Efficient Tests of Statistical Hypotheses". In: *Phil. Trans. Roy. Soc. Lond.* A231.694-706 (1933), pp. 289–337. DOI: 10.1098/rsta.1933.0009.

[173] G. Hinshaw et al. "Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results". In: *Astrophys. J. Suppl.* 208 (2013), p. 19. DOI: 10.1088/0067-0049/208/2/19. arXiv: 1212.5226 [astro-ph.CO].

[174] Mihailo Backović et al. "Direct Detection of Dark Matter with MadDM v.2.0". In: *Phys. Dark Univ.* 9-10 (2015), pp. 37–50. DOI: 10.1016/j.dark.2015.09.001. arXiv: 1505.04190 [hep-ph].

[175] Morad Aaboud et al. "Constraints on mediator-based dark matter and scalar dark energy models using $\sqrt{s} = 13$ TeV pp collision data collected by the ATLAS detector". In: *JHEP* 05 (2019), p. 142. DOI: 10.1007/JHEP05(2019)142. arXiv: 1903.01400 [hep-ex].

[176] Ulrich Haisch, Anthony Hibbs, and Emanuele Re. "Determining the structure of dark-matter couplings at the LHC". In: *Phys. Rev.* D89 (2014), p. 034009. DOI: 10.1103/PhysRevD.89.034009. arXiv: 1311.7131 [hep-ph].

[177] Matthew R. Buckley, David Feld, and Dorival Goncalves. "Scalar Simplified Models for Dark Matter". In: *Phys. Rev.* D91 (2015), p. 015017. DOI: 10.1103/PhysRevD.91.015017. arXiv: 1410.6497 [hep-ph].

[178] Ulrich Haisch, Priscilla Pani, and Giacomo Polesello. "Determining the CP nature of spin-0 mediators in associated production of dark matter and $t\bar{t}$ pairs". In: *JHEP* 02 (2017), p. 131. DOI: 10.1007/JHEP02(2017)131. arXiv: 1611.09841 [hep-ph].

[179] Michał Czakon, Paul Fiedler, and Alexander Mitov. "Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$". In: *Phys. Rev. Lett.* 110 (2013), p. 252004. DOI: 10.1103/PhysRevLett.110.252004. arXiv: 1303.6254 [hep-ph].

[180] Michal Czakon and Alexander Mitov. "NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction". In: *JHEP* 01 (2013), p. 080. DOI: 10.1007/JHEP01(2013)080. arXiv: 1210.6832 [hep-ph].

[181] Michal Czakon and Alexander Mitov. "NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels". In: *JHEP* 12 (2012), p. 054. DOI: 10.1007/JHEP12(2012)054. arXiv: 1207.0236 [hep-ph].

[182] Peter Bärnreuther, Michal Czakon, and Alexander Mitov. "Percent Level Precision Physics at the Tevatron: First Genuine NNLO QCD Corrections to $q\bar{q} \to t\bar{t} + X$". In: *Phys. Rev. Lett.* 109 (2012), p. 132001. DOI: 10.1103/PhysRevLett.109.132001. arXiv: 1204.5201 [hep-ph].

[183] Matteo Cacciari et al. "Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation". In: *Phys. Lett.* B710 (2012), pp. 612–622. DOI: 10.1016/j.physletb.2012.03.013. arXiv: 1111.5869 [hep-ph].

[184] Michal Czakon and Alexander Mitov. "Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders". In: *Comput. Phys. Commun.* 185 (2014), p. 2930. DOI: 10.1016/j.cpc.2014.06.021. arXiv: 1112.5675 [hep-ph].

[185] Nikolaos Kidonakis. "Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production". In: *Phys. Rev.* D83 (2011), p. 091503. DOI: 10.1103/PhysRevD.83.091503. arXiv: 1103.2792 [hep-ph].

[186] Nikolaos Kidonakis. "Two-loop soft anomalous dimensions for single top quark associated production with a W or H boson". In: *Phys. Rev.* D82 (2010), p. 054018. DOI: 10.1103/PhysRevD.82.054018. arXiv: 1005.4451 [hep-ph].

[187] Nikolaos Kidonakis. "NNLL resummation for s-channel single top quark production". In: *Phys. Rev.* D81 (2010), p. 054028. DOI: 10.1103/PhysRevD.81.054028. arXiv: 1001.5034 [hep-ph].

[188] *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC*. Tech. rep. ATL-PHYS-PUB-2015-045. Geneva: CERN, Nov. 2015. URL: https://cds.cern.ch/record/2064383.

[189] The ATLAS collaboration. "Calibration of the ATLAS b-tagging algorithm in $t\bar{t}$ semi-leptonic events". In: (2018).

[190] Morad Aaboud et al. "Search for heavy particles decaying into top-quark pairs using lepton-plus-jets events in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Eur. Phys. J.* C78.7 (2018), p. 565. DOI: 10.1140/epjc/s10052-018-5995-6. arXiv: 1804.10823 [hep-ex].

[191] Georges Aad et al. "Search for squarks and gluinos with the ATLAS detector in final states with jets and missing transverse momentum using 4.7 fb^{-1} of $\sqrt{s} = 7$ TeV proton-proton collision data". In: *Phys. Rev.* D87.1 (2013), p. 012008. DOI: 10.1103/PhysRevD.87.012008. arXiv: 1208.0949 [hep-ex].

[192] Jonas Neundorf. "A Novel Estimate of the Multijet background in a Search for Dark Matter Produced in Association with Top Quarks in the ATLAS experiment". Masterarbeit. Universität Hamburg, 2019, p. 62. URL: https://bib-pubdb1.desy.de/record/427788.

[193] W. Lukas. "Fast Simulation for ATLAS: Atlfast-II and ISF". In: *J. Phys. Conf. Ser.* 396 (2012), p. 022031. DOI: 10.1088/1742-6596/396/2/022031.

[194] *Studies on top-quark Monte Carlo modelling for Top2016*. Tech. rep. ATL-PHYS-PUB-2016-020. Geneva: CERN, Sept. 2016. URL: https://cds.cern.ch/record/2216168.

[195] Andrea Giammanco. "Single top quark production at the LHC". In: *Rev. Phys.* 1 (2016), pp. 1–12. DOI: 10.1016/j.revip.2015.12.001. arXiv: 1511.06748 [hep-ex].

[196] Philip Harris et al. "Constraining Dark Sectors at Colliders: Beyond the Effective Theory Approach". In: *Phys. Rev.* D91 (2015), p. 055009. DOI: 10.1103/PhysRevD.91.055009. arXiv: 1411.0535 [hep-ph].

[197] G. C. Branco et al. "Theory and phenomenology of two-Higgs-doublet models". In: *Phys. Rept.* 516 (2012), pp. 1–102. DOI: 10.1016/j.physrep.2012.02.002. arXiv: 1106.0034 [hep-ph].

[198] Asher Berlin, Tongyan Lin, and Lian-Tao Wang. "Mono-Higgs Detection of Dark Matter at the LHC". In: *JHEP* 06 (2014), p. 078. DOI: 10.1007/JHEP06(2014)078. arXiv: 1402.7074 [hep-ph].

Acknowledgments

Before anything else, I would like to express my thanks to all the people supported me during my Ph.D. studies. This work would not be possible without the help from any of you.

I am wholeheartedly grateful to my supervisor, Dr. Krisztian Peters, who granted me the opportunity to work on experimental high energy physics in the DESY ATLAS group and provided me with inspiring guidance throughout my doctoral program. Even in his busiest weeks he still managed to find time to discuss about the physical projects I was working on and the organization of my thesis. I have learned so much from him as both a physicist and a team leader.

Great thanks to my co-supervisor at the university, Prof. Dr. Johannes Haller, who gave me many valuable instructions during the discussions and full support towards the degree. I would also like to thank Prof. Dr. Dieter Horns, Prof. Dr. Gregor Kasieczka and Dr. Kai Schmidt-Hoberg for being my Ph.D. committee and reviewing my dissertation.

I want to give thanks to Dr. Katharina Behr, Dr. Pierre-Antoine Delsart and Dr. Michael Edward Nelson for their assistance on the derivation of the jet mass scale calibration for variable-radius jets. Thanks Dr. Chris Malena Delitzsch, who made my time in the UK so worthwhile during the Boosted Object Tagging Workshop.

I am thankful to everybody in the $E_{\mathrm{T}}^{\mathrm{miss}} + V$ (hadronic) group for bringing this nice piece of work to public. In particular, great thanks to Dr. Jike Wang, who helped me so much both professionally and personally, Paul Philipp Gadow, who was deeply involved in almost every aspect of the search as the other main analyzer, Dr. Samuel Meehan and Dr. Patrick Rieck for chairing all the analysis meetings. Equally important are the efforts from the members of the Editorial Board, namely Prof. Dr. Ayana Arce, Dr. Paolo Francavilla and Prof. Dr. Koji Terashi, as well as Dr. Sandra Kortner and Dr. Krisztian Peters for their tremendous contributions to the writing of the paper.

For the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ search, I would like to mention the significant contributions from Dr. Matthias Saimpert, without whom the analysis would have been impossible to reach the current status. Thanks Dr. Walter Howard Hopkins who chaired the weekly meetings and all members from the Editorial Board, Prof. Dr. Monica D'Onofrio, Dr. Steve Muanza and Prof. Dr. Xuai Zhuang, for their picky comments which helped to bring this work to a much maturer stage. Additionally, I want to thank Dr. Loic Valery for his technical suggestions on the statistical analysis during my thesis writing.

My sincere gratitude to all my friends and colleagues in the DESY ATLAS group who squeezed out their time (in Christmas!) to read and comment some parts of this thesis, provided language advises as well as shared their experiences in the doctorate proceedings, including Yu-Heng Chen, Judith Höfer, Vincent Kitali, Arthur Linß and Pablo Rivadeneira, and special thanks to Janik von Ahnen, Jonas Neundorf and Alicia Wongel who helped to translate my abstract to German. Moreover, I want to thank Prof. Dr. Ingrid-Maria Gregor as the group leader of the DESY ATLAS group for creating such a great team and family atmosphere. It has always been so enjoyable to be part of the group.

Last but not least, I owe my deepest gratitude to my family for their unconditional love and constant support in all these years, including my dog, LeLe, who has been serving as a replacement of me after I came to Germany. The relaxing and heartwarming moments I spent with them are the greatest encouragement for me in pursuit of the truth of nature.

Eidesstattliche Versicherung / Declaration on oath

Hiermit versichere ich an Eides statt, die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Hilfsmittel und Quellen benutzt zu haben.

Die eingereichte schriftliche Fassung entspricht der auf dem elektronischen Speichermedium.

Die Dissertation wurde in der vorgelegten oder einer ähnlichen Form nicht schon einmal ineinem früheren Promotionsverfahren angenommen oder als ungenügend beurteilt.

Hamburg, den 14.01.2020