Search for Dark Matter Produced in Association with a Higgs Boson Decaying to a pair of bottom quarks with the ATLAS Detector

Dilia María Portillo Quintero

THESE DE DOCTORAT
DE SORBONNE UNIVERISTÉ

présentée par

Dilia María Portillo Quintero

Pour obtenir le grade de

DOCTEUR ES SCIENCES
DE SORBONNE UNIVERISTÉ

Spécialité :

Physique de l'Univers (ED 560)

Recherche de la matière sombre produite en association avec un boson de Higgs se désintégrant dans une paire de quark b avec le detécteur Atlas.

Soutenue le 24 octobre devant le jury compose de :

Mme	F.	Ledroit	Rapporteur
MM	G.	Bernardi	Directeur de thèse
MM	N.	Makovec	Examinateur
MM	R.	Salerno	Rapporteur
MM	M.	Cacciari	Président
MM	S.	De Cecco	co-Directeur de thèse

LPNHE
PARIS

PhD THESIS
OF THE UNIVERSITE PIERRE ET MARIE CURIE

presented by

Dilia María Portillo

Submitted in fulfillment of the recquirements for the degree of

DOCTEUR In SCIENCES
OF UNIVERSITE PIERRE ET MARIE CURIE

Specialty :

Physique de l'Univers (ED 560)

Search for Dark Matter Produced in Association with a Higgs Boson Decaying to a pair of bottom quarks with the ATLAS Detector

Defended on October 24th in front of the comittee :

Mrs.	F.	Ledroit	Referee
M.	G.	Bernardi	Supervisor
M.	N.	Makovec	Examiner
M.	R.	Salerno	Referee
M.	M.	Cacciari	President
M.	S.	De Cecco	co-Supervisor

cnrs
dépasser les frontières

LPNHE
PARIS

SORBONNE UNIVERSITÉ
CRÉATEURS DE FUTURS
DEPUIS 1257

Contents

Introduction

The Standard Model (SM) of particle physics is the theoretical framework that describes the fundamental particles and interactions. It has successfully described and predicted most of the phenomena observed so far. The Higgs boson discovery by ATLAS and CMS collaborations at the Large Hadron Collider (LHC) is one of the main milestones in modern experimental physics, providing the last missing particle in the SM spectrum and a crucial consistency test of the SM electroweak symmetry breaking mechanism.

However, the SM does not successfully explain all the observed phenomena. One of the most pressing open question of the SM is to elucidate the particle nature of the dominant matter component in the Universe, called dark matter. Multiple strands of evidence have shown that not only dark matter exists and have gravitational effects, but also that it represents more than five times of the visible matter content in the Universe. From large scale gravitational and cosmological evidences, it is inferred that this unknown form of matter is non-baryonic, non-luminous, weakly interacting, electrically neutral, massive, not relativistic and stable. The SM does not provide a viable candidate that meets the above properties, posing one of the biggest yet open questions in present-day physics.

Despite the abundance of direct and indirect evidence, the particle nature of the dark matter and its connection to the SM remains a mystery. A compelling hypothesis that accommodates the observed relic abundance of dark matter in the Universe is the existence of an electrically neutral, stable particle, which could interacts very weakly with other SM particles and with a mass between few GeV to several TeV. These hypothetical dark matter candidates are predicted in many beyond the SM theories, are called *WIMPs*, and are thermally created dark matter particles with production cross-section at the electroweak scale.

If the dark matter is made of WIMPs, it could be produced and probed at the LHC energies. Discovering or excluding dark matter production at collider would provide new insights about the particle nature of the overarching composition of the Universe, as well as the connection between physics at very high and small scales. However, the detection of dark matter particles at colliders is challenging since the WIMP is electrically neutral, stable, and interacts very weakly, therefore it escapes the detector once produced without leaving any signal. Because of this, the conservation of momentum is used as a handle to detect dark matter particles: since the total momentum of the collision products in the transverse plane, in the initial state of p-p collisions, is zero, an imbalance in the visible transverse momentum, called missing transverse momentum $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$, can be an indication of the presence of dark matter particles.

A robust reconstruction and understanding of the missing transverse momentum observable is critical in the search for dark matter at colliders. In ATLAS collaboration, this

variable uses the combined inputs from all the reconstructed physics objects exploiting all the different sub-detector signals. An imbalance in the total transverse momentum, $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}} \neq 0$, may be the effect of *real* missing transverse momentum coming from weakly interacting particles as dark matter or neutrinos. But it may also be *fake* $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ arising from mis-reconstructed and mis-identified objects (hadronic jets, leptons, tracks...) or from detector dead regions. In this thesis, an object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance variable is proposed in order to improve the discrimination of events in which the total $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ is consistent with a fake reconstruction from momentum resolutions, against vents with real reconstruction from weakly interacting particles as WIMPs. The performance of the newly designed $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance discriminant is studied in $Z \to ee$ and $ZZ \to ee\nu\nu$ simulation samples and checked with data collected in proton–proton collisions at a centre-of-mass energy of 13 TeV during 2015 and 2016, corresponding to an integrated luminosity of 36 fb^{-1}.

Most collider-based searches for WIMPs focus on signatures characterized by one or more high-momentum and well defined SM particles, named "X", recoiling against large amounts of missing transverse momentum, associated with the dark matter production. This signature is denoted as mono-"X" or "X"+$\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$. Such mono-"X" signatures have been explored by the LHC experiments for the cases where "X" is an hadronic jet, a heavy flavour quark, a photon, or a vector boson (W/Z).

The Higgs boson discovery in 2012 opened a unique probe method for searching for searching dark matter at LHC in the mono-h topology. In most of the previously mentioned mono-"X" searches, the SM particle "X" is originated from initial state radiation. In contrast, the initial state radiation of a Higgs boson is Yukawa suppressed, making the h in the final state closely connected to the dark matter pair production mechanism, providing a direct probe of the coupling between the dark matter and the SM particles.

This thesis presents the search for dark matter produced in association with a Higgs boson, mono-h, with 13 TeV proton-proton collision at the LHC recorded by ATLAS experiment. In particular, this search focuses on the Higgs boson decaying to two bottom quarks ($h \to b\bar{b}$) which corresponds to its dominant decay mode. The $h \to b\bar{b}$ decay mode was recently observed by ATLAS to be consistent with SM expectations.

To first order, the transverse momentum of the Higgs boson has the same magnitude and opposite direction with respect to the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$. For dark matter signals producing low $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$, the angular separation of the jets initiated by the b-quarks is large and can be reconstructed as two standard radius jets in the ATLAS calorimeter. While for signals with high $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$, the Higgs boson tends to have a high Lorentz boost and it is often reconstructed as a single jet with large radius. This search employs both reconstruction techniques to efficiently cover large kinematic range.

A first mono-h analysis is presented using data recorded by ATLAS during 2015 and 2016 corresponding to an integrated luminosity of 36.1 fb^{-1}. An improved mono-h analysis is then presented, using data recorded by ATLAS during 2015, 2016 and 2017 corresponding to an integrated luminosity of 79.8 fb^{-1}. The improved mono-$h(b\bar{b})$ search substantially increases its sensitivity compared with the previous result. In addition to the approximate doubling the dataset size, the sensitivity of this search is enhanced with the inclusion of new object reconstruction techniques. One is the implementation of the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ variable in order to reject more effectively background events main-

taining a high signal significance. The other is the variable radius track jets technique to reconstruct and b-tag jets in boosted topologies with higher efficiency.

The thesis is organized as follows. The general theoretical and experimental context of the SM Higgs boson and dark matter paradigm is introduced in Chapter 1. A general overview of the dark matter searches at LHC is presented in Chapter 2. The LHC and the ATLAS detector are described in Chapter 3. The reconstruction of the physics objects is described in Chapter 4. In Chapter 5 the new object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance variable is defined and its performance studies are detailed. The general description of the mono-h search strategy, selection and backgrounds is described in Chapter 6. The results of the mono-h search from the analysis of data recorded during 2015 and 2016 is presented in Chapter 7. Finally, the improved analysis using data recorded during 2015-2017 is presented in Chapter 8 making special emphasis in the new object reconstruction techniques.

Personal contributions

Since I am part of a large high energy physics collaboration since October 2015, my research activity relies on the results obtained by a large number of researchers and analysis groups. In the analyses and results presented in this thesis, I contributed as part of the analysis team and I am author of the public and internal documentation.

I initiated my participation in the mono-h search with data collected during 2015, corresponding to $3.2\,\mathrm{fb}^{-1}$ [1]. I contributed with dark matter simplified model study and validation with acceptance studies, modelling with data-MC comparisons, analysis of the cross sections with respect to the new particle masses, and kinematic distributions studies for different model parameters. The simplified models were the baryonic vector mediator, the scalar mediator and the $Z' - 2HDM$ [2]. I also did a multivariate analysis optimization for the scalar mediator simplified model with boosted decision trees in the merge regime presented in spring 2016.

I was more involved in the mono-h analysis with data collected during 2015 and 2016, which corresponds to $36.5\,\mathrm{fb}^{-1}$ [3]. I studied the simplified model exclusion contour extrapolation scaling with respect to luminosity and sensitivity gain in order to predict the parameter space that will be excluded with the new data collected during 2016 and after full Run 2. I also studied the mass-mass plane grid for simplified models, and performed signal cross-section and truth acceptance studies. All these results were closely related to the dark matter signal model validation and generation in MadGraph, task that I was in charge of. I also participated in the ATLAS dark matter summary paper that will be soon published which contains the results shown in Chapter 2 with MadGraph signal generator.

In the last years I played the leading role in the definition and algorithm development of the new object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance variable. I worked from the beginning of the variable proposal to its performance. I also studied different potential improvements on the algorithm for a future publication. I was the note editor and contact of the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance conference note [4].

I was fully involved in the mono-h search with data collected during 2015-2017, corresponding to $78.8\,\mathrm{fb}^{-1}$ [5]. I initially studied the 1-lepton (1μ-CR) channel with cut-flow

validation, code maintenance and new offline reconstruction release validation. I participated also in the Signal Region and 1-lepton Control Region data-Monte Carlo comparisons and validation. My main contribution to the analysis was the implementation of a new requirement on the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance in the Signal Region in order to reject multijet background. In order to achieve this, I was involved in the commissioning, validation and modelling studies in all channels of this new variable. I did performance studies and signal significance optimization studies for different object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance definitions. After this, I successfully tested the impact of the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance in rejecting multijet background while keeping a high signal significance. Also, I participated in several data-driven multijet background estimations with and without the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance requirement.

From the detector operation side, during my PhD. I participated as Online Run Control Shifter in the ATLAS control room.

Chapter 1

The quest for Dark Matter

A plethora of new experimental data accumulated in the past decade has confirmed the Standard Model (SM) to an impressively high precision. However, a deep relation between particle physics and cosmology is being revealed, and new questions emerge. What is the quantum nature of dark matter is currently one of the most relevant. This Chapter generally presents the SM and the Higgs boson mechanism and phenomenology at LHC in Section 1.1, and the outstanding questions and limitations of the SM in Section 1.2. Finally, the most important dark matter evidences, candidates and detection strategies are presented in Section 1.3.

1.1 The Standard Model and Higgs boson

The SM of particle physics [6] together with general relativity [7] is the most detailed and successful description of nature so far. The former one, SM, is a gauge field theory of the elementary particles and their interactions, which is based on the symmetry group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$. This is the group of transformations of the fundamental fields of the theory under which the equations of motion are left invariant.

The strong interactions reflect invariance under the local $SU(3)$ gauge group with its colour charge C [8, 9, 10]. The electromagnetic and weak interactions are described by the invariance under local weak isospin and hypercharge gauge transformations described by the $SU(2) \otimes U(1)$ Lie group [11, 12, 13, 14].

The SM symmetry group, $SU(3) \otimes SU(2) \otimes U(1)$, has $8+3+1=12$ generators. In gauge theory, each generator has a vector boson associated, and if the symmetry is unbroken, this gauge boson is massless. In the SM, these massless gauge bosons mediates the fundamental interactions: eight gluon gauge fields G_a^μ for the strong interactions, three gauge fields W_i^μ for the the $SU(2)$ symmetry based on chirality L, and the gauge field B^μ for the hypercharge symmetry.

While the massless gluons can be directly identified with the strong interaction, the massless W_i^μ cannot be identified with the massive mediators of the weak force, $W^\pm$ and Z^0. The masses associated to these mentioned vector boson are dynamically originated via spontaneous symmetry breaking, through the Brout-Englert–Higgs mechanism [15, 16, 17].

The $SU(2) \otimes U(1)$ symmetry is broken by the vacuum expectation value (vev) v of a scalar field ϕ, the Higgs field, which is added to the theory ($\langle\phi\rangle = v_0$). In this case, the currents and charges are as dictated by the symmetry but the fundamental state of minimun energy, the vacuum v, is not unique and there is a continuum of degenerate states that together respect the symmetry. The symmetry is broken due to the system is found in one particular vacuum state, v_0, and this choice makes the symmetry spontaneously violated in the spectrum of states.

The Higgs field is dynamically coupled to gauge bosons though gauge interactions. The gauge bosons masses, as well as the values of the gauge couplings, are determined by the v_0 of the Higgs field. The gauge bosons acquire mass by incorporating the Goldstone degrees of freedom such that no physical Goldstone remains [18], however there is one massive scalar degree of freedom: the Higgs boson.

As a consequence of the Brout-Englert–Higgs mechanism, the $SU(2) \otimes U(1)$ is broken to $U(1)_{\mathrm{EM}}$. This procedure leads to a massless gauge boson A^μ that is identified with the photon. A massive gauge boson Z^0 identified with the mediator of the neutral currents of the weak force, and two charged massive vector bosons $W^\pm$. In addition, the fermions masses are also generated dynamically thought the Higgs field by including new Yukawa interactions which obeys $SU(2) \otimes U(1)$.

The Lagrangian that is invariant under local weak isospin and hypercharge gauge transformations, described under the $SU(2) \otimes U(1)$ group, can be written as [19]:

$$\begin{aligned}
\mathcal{L} \;=\; & -\frac{1}{4}\mathbf{F}^a_{\mu\nu}\mathbf{F}^{a\mu\nu} && (1.1)\\
& + i\psi \not{D} \psi + \mathrm{h.c} && (1.2)\\
& + \psi_i y_{ij} \psi_j \phi + \mathrm{h.c} && (1.3)\\
& + \left| D_\mu \phi^2 \right| - V(\phi)\,. && (1.4)
\end{aligned}$$

The first line, Equation 1.1, contains the kinetic terms for the gauge sector of the theory, where a is running over all the gauge fields. A coupling g_i is associated to each of the gauge groups. The second line, Equation 1.2, describes the interactions between the matter fields (fermions) ψ and the gauge fields. The third line, Equation 1.3, corresponds to the Yukawa interactions and describes the interactions between the fermions ψ and the Higgs field ϕ, which is a scalar $SU(2)$ doublet responsible for giving mass to the fermions when the electroweak symmetry breaking occurs. Finally, the fourth line, Equation 1.4, describes the scalar sector which corresponds to the Higgs kinetic term and its potential. The latter takes the following form in the SM:

$$V(\phi) = -\mu^2 |\phi|^2 + \lambda |\phi|^4 \,. \tag{1.5}$$

The Higgs boson sector is characterized by the "new" parameters μ and λ. Since $v_0 = \mu/\sqrt{\lambda}$ can be measured from the mass of the charged weak boson $W^\pm$ and the Fermi coupling G_F, only one degree of freedom is left and it can be identified with the Higgs boson mass, which can be written as $m_h = \mu\sqrt{2} = \sqrt{2\lambda}\, v_0$.

1.1.1 Higgs boson phenomenology at LHC

On July 4th, 2012, the ATLAS and CMS collaborations of the Large Hadron Collider (LHC) at CERN announced the observation of a new neutral heavy scalar particle consistent with the Higgs boson [20, 21], thus making a big step towards completing the experimental verification of the SM.

The new particle was discovered decaying to pair of photons, Z ($ZZ^{(*)} \to 4\ell$) or W ($WW \to e\nu\mu\nu$) bosons, based on about $10\,\text{fb}^{-1}$ data collected by both ATLAS and CMS at $\sqrt{s} = 7$ TeV in 2011 and at $\sqrt{s} = 8$ TeV in 2012 (Run 1). Figure 1.1(a) shows the observed and expected local p_0 values from the combination of these three decay channels as a function of the new particle mass. The largest local significance for the combination of the 7 and 8 TeV data is found for a SM Higgs boson mass hypothesis of $m_h = 126.5$ GeV, where it reaches 6.0σ for the ATLAS experiment. Combined measurements from both experiments, ATLAS and CMS, with the $\gamma\gamma$ and 4ℓ decay channels, based on $25\,\text{fb}^{-1}$ of the full Run 1 data, show that the mass of the new scalar particle is $m_h = 125.09 \pm 0.24$ GeV [22]. Figure 1.1(b) presents a summary of Higgs boson mass measurements from the individual analyses of ATLAS and CMS and from the combined result.

The Higgs boson acts as a mediator of new interactions, in which, at the tree level, it couples in proportion to the masses of the fermions and to the mass square of the bosons. Figure 1.2 shows a parameterization of the (square root of the) coupling to fermions (bosons) as a function of their mass, from a combined measurement based on Run 2 ATLAS data. The parameterization is chosen so that the behaviour expected in the SM is linear with the same slope for fermions and bosons. It is shown in Figure 1.2 that the combined measurements is in a good agreement with the SM predictions.

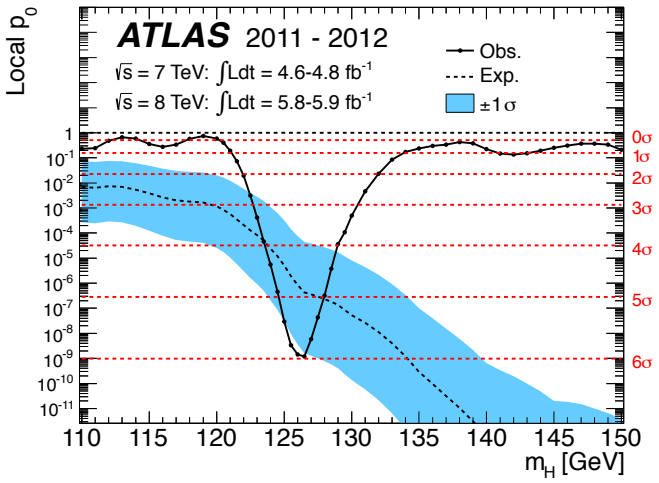
In pp collisions, the largest coupling for Higgs boson production is via the top quark, as can be seen in Figure 1.2. The top content of the proton is effectively zero, however the top quark Yukawa coupling can be accessed through a gluon initiated loop or via $t\bar{t}$ pair production, as shown in Figures 1.3 a) and d). The gluon-initiated top loop, called gluon-gluon fusion or ggF, is the Higgs boson production mode with highest cross section.

Alternatively, the Higgs boson can be produced via its couplings with the massive weak vector bosons. In this case, the dominant production process is the vector boson fusion (VBF), in which the Higgs boson is produced from a t-channel exchange of W and Z between the two quarks, in association with two final-state quarks. The Feynman diagram of the VBF process is shown in Figure 1.3 b). Also, the Higgs boson can be produced via radiation from a s-channel W or Z boson. This is referred to as associated production or Higgs-strahlung, and the diagram is shown in Figure 1.3 c).

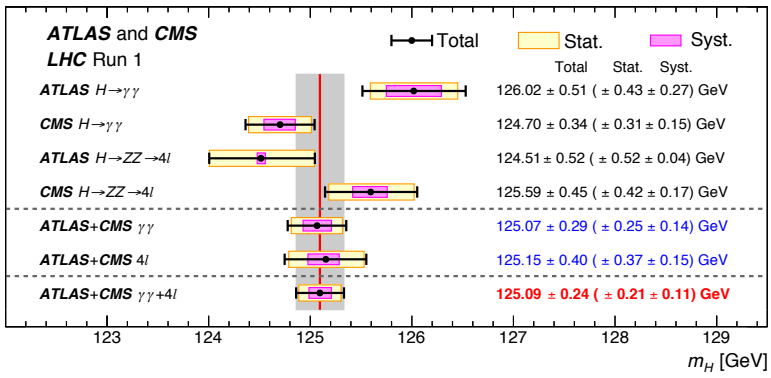
The Higgs boson production cross section is shown in Figure 1.4 as a function of the LHC centre of mass energy. For a Higgs boson mass of $m_h = 125$ GeV, the inclusive production cross section is shown in Table 1.1 for pp collisions at $\sqrt{s} = 13$ TeV.

Since the Higgs boson couplings grow with the masses of the gauge bosons and fermions, it will tend to decay into the pair of particles with highest mass that are kinematically allowed.

Regarding fermions, the decays to $t\bar{t}$ are kinematically forbidden due to the large mass of the top and the off-shell process is heavily suppressed. Thus, the most common decay of the Higgs boson is to bottom quarks. For a Higgs mass of $m_h = 125.09$ GeV, the $h \to b\bar{b}$



(a)



(b)

Figure 1.1 – Higgs boson mass measurement in Run 1. 1.1(a) shows the observed and expected local p_0 values from the combination of channels in ATLAS as a function of the new particle invariant mass m_h [20]. 1.1(b) shows the combined Higgs boson mass measurements from the individual analyses of ATLAS and CMS. The systematic (narrower, magenta-shaded bands), statistical (wider, yellow-shaded bands), and total (black error bars) uncertainties are indicated. The (red) vertical line and corresponding (gray) shaded column indicate the central value and the total uncertainty of the combined measurement, respectively [22].

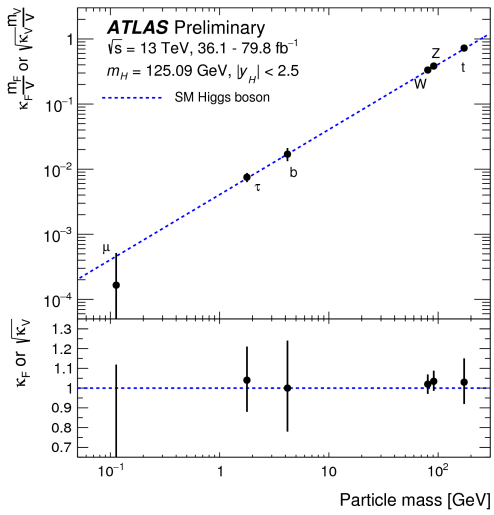


Figure 1.2 – Coupling strength modifiers as a function of particle mass for W, Z, t and b quarks, τ leptons and muons. The modifiers are measured assuming no BSM contributions to the Higgs boson decays. The result combines the analysis: $h \to \gamma\gamma$, $h \to ZZ \to 4\ell$, $h \to \mu\mu$ with $80\,\text{fb}^{-1}$ and $h \to WW$, $h \to \tau\tau$,$h \to bb$ and tth($h \to$multi-leptons).

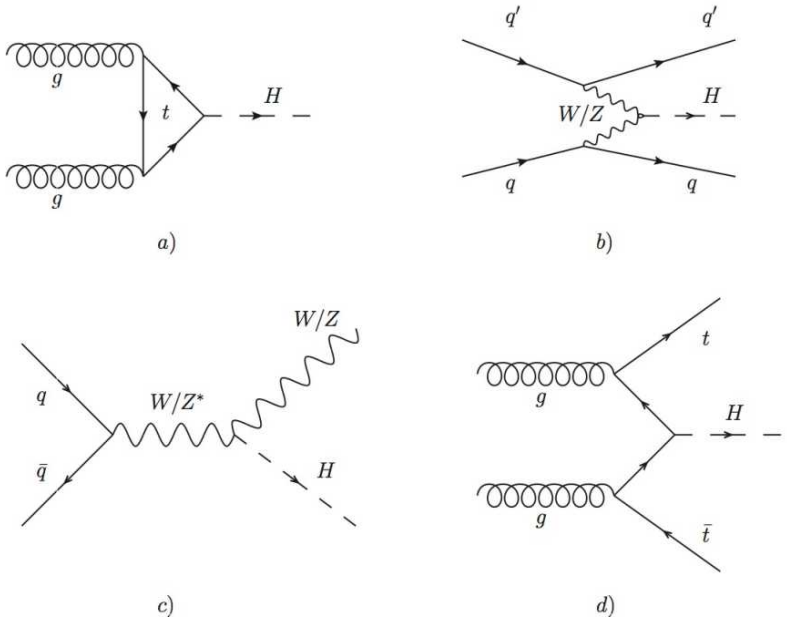


Figure 1.3 – Leading order diagrams for the main Higgs boson production mechanisms at the LHC. a) gluon-gluon fusion, b) vector boson fusion, c) W/Z associated production, and d) $t\bar{t}$ associated production

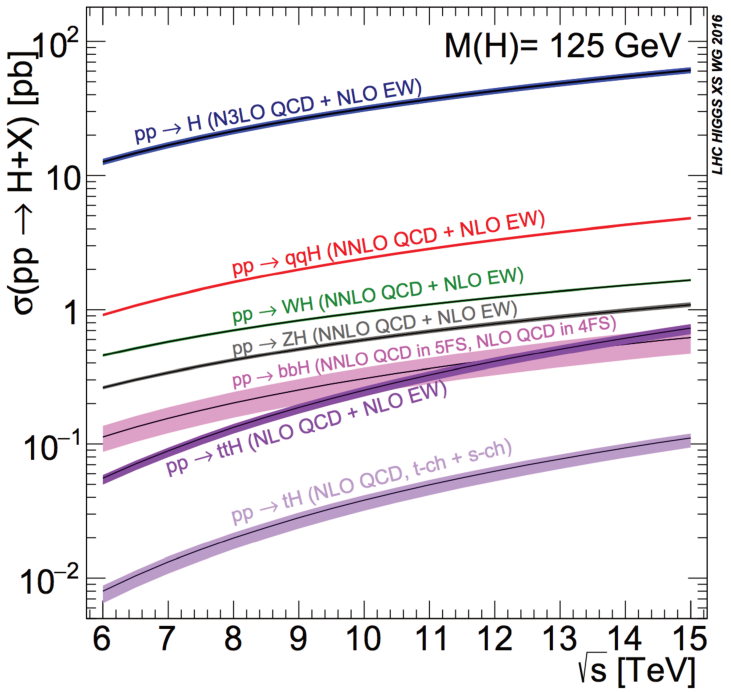


Figure 1.4 – The SM Higgs boson production cross sections, as a function of the LHC centre of mass energy [23].

Production process	Cross section σ_i [pb]	$\frac{\sigma_i}{\sigma_{\text{tot}}}$ [%]
ggF	$48.58^{+2.23}_{-3.25}$	87.2
VBF	3.78 ± 0.08	6.8
ZH	$0.88^{+0.04}_{-0.03}$	1.6
WH	1.37 ± 0.03	2.5
ttH	$0.51^{+0.03}_{-0.05}$	0.9
bbH	$0.49^{+0.10}_{-0.12}$	0.9

Table 1.1 – Inclusive SM cross sections, σ_i, for production of a 125 GeV Higgs boson at $\sqrt{s} = 13$ TeV pp collisions [23].

has a branching ratio of 58%. Despite being the dominant decay mode, the large QCD background with two jets in the final state strongly reduces the Higgs boson observation power in this channel. It was not until 2018 that the Higgs boson was observed decaying to bottom quarks [24, 25]. This was done by combining the results from other searches in Run 1 and in Run 2 for the $h \to b\bar{b}$ decay mode, which yields an observed (expected) significance of 5.4 (5.5) standard deviations for ATLAS, and similar for CMS.

Regarding the bosons, decays to a pair of weak vector bosons (W^+W^- and ZZ) are possible and have a large coupling, however at least one of the vectors is off-shell. Even

though the Higgs boson does not have direct couplings to photons or gluons, decay to these particles can occur through fermion or vector boson diagram mediated loops.

The SM branching ratios for the dominant decay modes are shown in Figure 1.5 as a function of the Higgs boson mass, and in Table 1.2 for a Higgs boson mass of 125 GeV.

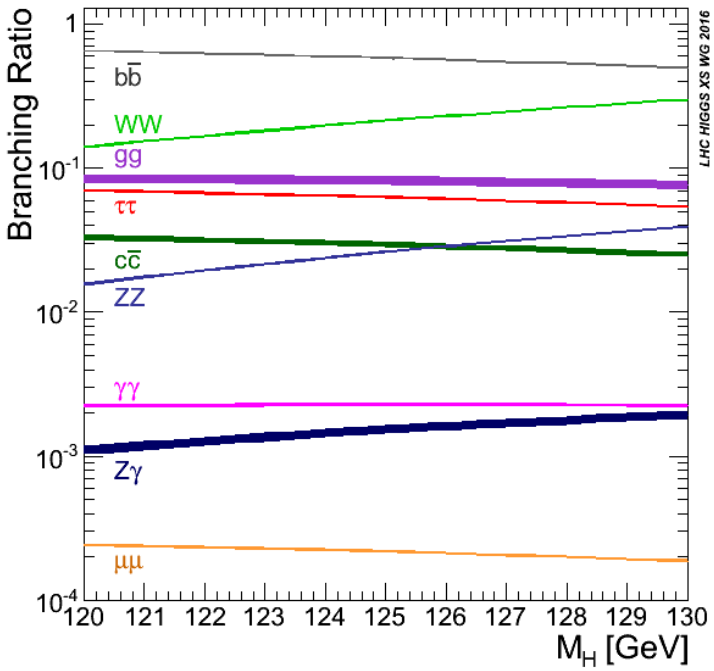


Figure 1.5 – Higgs boson branching ratios and their uncertainties for the mass range around 125 GeV [23].

Decay mode	Branching ratio [%]	Relative Uncertainty [%]
$b\bar{b}$	58.09	± 1.24
W^+W^-	21.52	± 1.53
gg	8.18	± 5.14
$\tau\bar{\tau}$	6.26	± 1.65
$c\bar{c}$	2.88	$^{+5.55}_{-1.98}$
ZZ	2.64	± 1.53
$\gamma\gamma$	0.23	± 2.09
$Z\gamma$	0.15	± 5.81

Table 1.2 – Inclusive SM branching ratios for the main decay modes of a 125.09 GeV Higgs boson [23].

1.2 Standard Model open questions

The SM is the greatest and complete theory of the fundamental constituents and interactions of Nature conceived so far. It is one of the cornerstones of all science and one of the great triumphs of the past century. It has been carefully experimentally verified from many different angles, especially during the past 30 years.

Whereas the gauge sector and its interactions with fermions described in the first two lines of the SM Lagrangian, Equations 1.1 and 1.2, have been confirmed in different experiments, before the start of the LHC, there was no experimental evidence of the Higgs field sector and its Yukawa interactions of Equations 1.3 and 1.4, which are responsible for particles masses and electroweak symmetry breaking. One of the main objectives of the LHC is to study and verify this part of the SM. However, after the SM-like Higgs boson discovery, there are still some open questions.

The intrinsic questions of the SM can be divided in three categories, the problem of *unification*, *flavour* and *mass*, which are directly related to the understanding of the SM parameters, 19 in total [19].

The SM gauge sector is characterized by three gauge couplings and a charge-parity (CP)-violating phase in the strong interactions. The simplification of this sector can be referred to as the *unification* problem. The questions that it raises are: Is it possible to unify the fundamental forces? Can the gravity be included as a quantum field theory?

The Yukawa interactions on Equation 1.3 is characterized by a total of 13 parameters: the masses of six quarks and three charged-leptons, three Cabibbo-Kobayashi-Maskawa angles and a CP-violating phase. The understanding of this sector can be referred to as the the *flavour* problem. The questions that it raises are: Why there are different types of matter particles organized in families? What is the origin of the difference between matter and antimatter, and is it related to the origin of the matter in the Universe? Why are quark mixing angles very small, while lepton mixing angles are not? How neutrinos masses can be generated?

Finally, the minimal Higgs boson sector shown in Equation 1.4 and 1.5 has just two parameters associated to the quadratic and quartic term in the Higgs potential. The understanding of this sector can be referred to as the the *mass* problem. The questions that it raises are: Is the origin of particle masses related to a single elementary Higgs boson? How the Higgs boson acquires its vev, and as a byproduct also its mass? Is it possible to dynamically generate the Higgs potential in order to not add it in ad-hoc way to the SM? How could one stabilize a theory against quantum corrections with a $m_h = 125$ GeV Higgs boson? Why is the Higgs mass so much smaller than the Planck mass?

Even though the SM successfully describes the data from precision experiments of all sorts, it also causes many theoretical challenges, as described above. These problems are mainly of theoretical nature and may be ignored as long as the SM is interpreted as an effective theory or as a low energy realization of a more complete model. An indication that the SM is a effective theory from a more complete one, is that the strong interactions, described by the $SU(3)$ group, is completely independent of the electroweak force, described by the $SU(2) \otimes U(1)$ group.

Popular examples of complete models that addresses one or various of the *unification*, *flavour* and *mass* theoretical problems of the SM are: Grand Unified Theories (GUT) [26,

27, 28], Supersymmetry [29, 30, 31], Supergravity [32], Large extra dimensions [33, 34], kaluza-Klein theories [35], and Superstring [36].

Most of the previously mentioned open questions within the SM are related to the fine-tuning of the theory [37]. The crucial problems are the ones related to experimental evidences which can not be explained in the light of the SM framework and are connected to observations of cosmological/astrophysical nature, strongly related to the composition and history of the Universe. In this sense, the most pressing questions for particle physics definitely include: gravity, inflation, baryon asymmetry, dark energy and dark matter, due to their great relevance for the completeness of the theoretical framework of particle physics and cosmology.

Gravity

The general relativity has not been translated into a quantum field theory and gravity is not explained as a fundamental interaction within the SM. At the electroweak scale the effect of gravity can be neglected, however it is expected to play a role at much high energy scale, the Planck scale.

Inflation

The correct theoretical description of inflation is a unsolved problem in cosmology. The exponential expansion of the early Universe that led to the remarkable spatial flatness observed today may require physics beyond the SM.

Baryon asymmetry

The amount of CP violation predicted in the SM by the mechanism of Kobayashi and Maskawa via a single comples phase in the Cabibbo-Kobayashi-Maskawa quark-mixing matrix is insufficient to generate the amount of matter over antimatter in the Universe today. This requires new physics beyond the SM, such as supersymmetry or extensions with a "baryogenesis" mechanism.

Dark matter and dark energy

Maybe the most pressing problem for the SM plus the general relativity is related to the composition of the Universe. These two theories together should fundamentally describe all know physics, however at scales larger than the Solar System, they fail. Only about 5% of the mass-energy that affects the galaxies dynamics consist of SM matter. The rest is called "dark matter" and its origin and composition is of unknown nature. Furthermore, the Universe appears to be accelerating and this can be explained by postulating the "dark energy". The latter should provide approximately the 70% of the mass-energy density of the Universe.

1.3 The dark matter paradigm

Several astrophysical and cosmological observations have shown that the vast majority of matter in the Universe is non-luminous. Precise measurements, such as the cosmic microwave background (CMB) radiation by WMAP [38], the distance redshift relation of the Type Ia supernovae [39, 40] and the large scale structure (LSS) survey from SDSS [41, 42], show that on large scales, the average dark matter density is about 6 times the baryon density [43, 44].

All these observations are of astrophysical nature, however there is yet no evidence for non-gravitational interactions between dark matter and SM particles. Therefore, to unveil the mystery of the dark matter nature and composition is a fundamental problem of modern cosmology and physics.

1.3.1 Evidence of Dark Matter

Currently, there has not been any evidence of the sub-atomic or quantum nature of a dark matter particle from experiments. However, there is a great evidence which points towards the existence of this type of matter [45]. It is inferred from a wide number of different astrophysical observations which span a large range of scales and redshifts, pointing towards the existence of a non luminous matter component which, with common notation, goes under the name of "dark matter".

The first hint of dark matter was provided by Fritz Zwicky [46] in the 1930's. By using the virial theorem in the Coma cluster, he inferred the dynamical mass of the system and found that it was two orders of magnitude larger than the mass inferred from the luminosity of the cluster. This result could not be explained unless the vast majority of the mass of the Coma cluster was for some reason "missing" or non-luminous.

The most direct way that indicates the existence of dark matter is the study of rotation curve of galaxies [47, 48]. Roughly 40 years after Zwicky, Vera Rubin and collaborators performed an extensive study of the rotation curves of 60 isolated spiral galaxies. A rotation curve shows the rotation velocity of an object around the galaxy center as a function of the radius r, which scales as $\sqrt{M(r)/r}$ with $M(r)$ the mass within the orbit r. The rotation curve should decrease as $1/\sqrt{r}$ if r is beyond most of the visible part of the galaxy. However, they observed that the rotation curves of galaxies become flat for large distances from the galactic centre. The collected data for NGC 3198 is depicted in Figure 1.6, showing that the rotation curves at a certain distance are flat, meaning that the velocities of stars continue to increase with distance from the galactic center [49]. Baryonic matter (gas, dust, stars) could not explain alone the galaxy rotation curves. Because of this, it was concluded that galaxies have a halo of dark matter distributed at large radii.

The observation of X-ray emission of hot gas in the clusters can give a precise measurement of the gravitational potential felt by the gas to keep it in hydrostatic equilibrium [50, 51]. In addition to this, the measurements of weak lensing can give indication of the dark matter components [52, 53]. Maybe the most spectacular dark matter evidence is the analysis of the galaxies in the Bullet Cluster, which consists of two colliding galaxy clusters. The mass reconstruction from gravitational lensing shows two massive

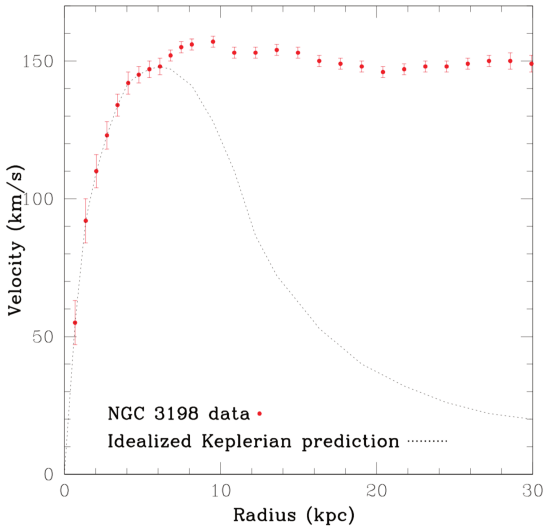


Figure 1.6 – Measured rotational velocities of HI regions in NGC 3198 compared to an idealized Keplerian behavior.

substructures that are in offset with respect to the baryon distribution observed in X-rays by Chandra, as shown in Figure 1.6. In this scenario, the gas is decelerated due to the viscosity, while the dark matter component can pass through each other without collision. This imply that dark matter is collisionless and constraints on the dark matter self-interaction can be derived. Also, other galaxies collisions have been studied, for example in Reference [51]

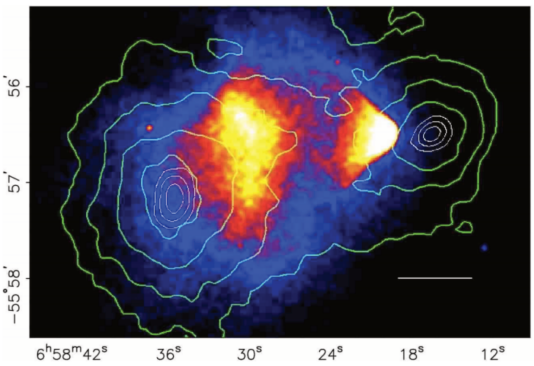


Figure 1.7 – X-ray image of the Bullet Cluster obtained with Chandra. The contours overlaid show the gravitational map reconstructed by the lensing analysis. Both images show the offsets between the distribution of baryons and that of the reconstructed gravitational fields.

The existence of dark matter in the cosmological scale is inferred by a global fit to the Cosmic Microwave Background (CMB), supernovae and LSS data. Fluctuations in the CMB are indications of the initial density perturbations that allowed for the formation of early gravitational wells, as well as dynamics of the photon-baryon fluid. Because of this, the temperature fluctuations of the CMB are dependent on the total amount of baryons

in the Universe at the time of recombination. However, temperature anisotropies in the CMB (only bout $30 \pm 5\,\mu$K) can not account for the formation of large structures in the Universe [54]. From WMAP precise measurements of the power spectrum of the CMB radiation, the total and baryonic matter densities can be computed:

$$\Omega_{\text{Total}}h^2 = 0.1334 \pm 0.0056\,,\ \Omega_{\text{baryonic}}h^2 = 0.02260 \pm 0.0053\,, \tag{1.6}$$

where $\Omega_{\text{Total}}h^2$ is the total matter density and $\Omega_{\text{baryonic}}h^2$ is the baryonic matter density, and h is the Hubble parameter in the unit of 100 km s^{-1} Mpc^{-1}. Thus, baryonic matter is not the only form of matter in the universe and, in fact, dark matter density is around 83% of the total mass density.

Also, the nowadays dark matter density, i.e. relic density, is well established. The measurement reported by PLANCK satellite [55] is

$$\Omega_{\text{dark matter}}h^2 = 0.1199 \pm 0.0027\,, \tag{1.7}$$

This corresponds to what remains of the production production process that happened in the early Universe.

1.3.2 Dark Matter candidates

Up to know, the evidence for dark matter is purely of gravitational origin. This implies that what is called "dark matter" could be either a manifestation of the incomplete understanding of gravity on scales larger than the solar system, or to a new fundamental component of Nature, i.e. a new elementary particle, not yet discovered and not present in the SM. If the latter is the correct solution, it implies that the SM needs to be extended and new physics is required in terms of new particles and interactions.

All the dark matter evidences point towards a particle that is non-baryonic, massive, electrically neutral (dark), stable at the cosmological scales, collisionless (weakly interacting) and cold (non-relativistic during structure formation).

Dark matter candidates are generically considered to as WIMPs (Weakly Interacting Massive Particles), meaning that they are massive particles that are electrically neutral which do not interact very strongly with other matter [45].

In the WIMP scenario, the dark matter particles can reach thermal equilibrium in the early Universe and decouple from the thermal bath when the temperature decreases, this mechanism is called "freeze-out". The relic density of WIMPs can be calculated by Boltzmann equation, and a good approximate solution gives

$$\Omega_{\text{dark matter}}h^2 \sim \frac{3 \times 10^{-27}\text{cm}^3\text{s}^2}{\langle \sigma v \rangle}\,, \tag{1.8}$$

where $\langle \sigma v \rangle$ is the thermal averaged dark matter annihilation cross section times velocity, often called as annihilation cross section. It represents the interaction strength between the WIMPS and the SM particles. The value often considered as benchmark for dark matter annihilation cross section is $\langle \sigma v \rangle = 3 \times 10^{-26}\text{cm}^3\text{s}^2$. It is found that a WIMP

with a mass and interaction strength at the weak scale can easily give the correct relic density. This is referred as to the "WIMP miracle".

Many such dark matter particles have been proposed in the literature, with the prototypical example being the lightest supersymmetric particle (LSP). However, there are other possibilities: the basic requirement is some conserved quantum number to guarantee the stability of the dark matter particle. In the supersymmetric case, the R-parity is conserved, which is a combination of baryon number, lepton number and spin.

One of the simplest and therefore one of the most studied model for dark matter is known as the Inert Doublet Model (IDM) [56]. E. Ma and G. Deshpande showed that by adding an $SU(2)_L$ scalar doublet to the SM and imposing a Z_2 symmetry, it is possible to obtain a viable cold dark matter candidate, which is the lightest particle carrying Z_2 charge. This new set of scalars are called *inert* because of their lack of direct coupling to the SM fermions, of which the lightest is a dark matter candidate. More details on this model are given in Appendix A.

There are other non-WIMP dark matter candidates that can also account for the observed properties, like the axion [57] or the Lightest Kaluza-Klein particle [58]. Unfortunately, the search for these or any other possible candidate has not been conclusive and all the theoretical possibilities are still valid.

1.3.3 Dark Matter searches

In order to determine the nature of the dark matter particles, the interaction between the dark matter and the SM needs to be probed. There are three different directions to study this interaction: Search for the scattering between dark matter and SM particles, detect the annihilation or decay products of dark particles, or search for the dark matter production process. The three ways of dark matter detection are called *direct*, *indirect* and *collider* detection, respectively. Figure 1.8 illustrates these three dark matter detection squemes, which are not independent, but complementary to each other.

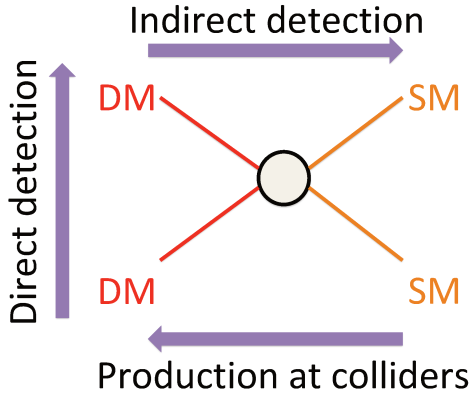


Figure 1.8 – Dark matter detection approaches related to each other by changing the time axis in the search for dark matter particles.

Direct detection

Since the Earth while rotating around the sun is moving through a dark matter halo, it is constantly traversed by dark matter particles. A setup of a very sensitive detector with a large amount of a characteristic element which can detect very small motions and interactions of the atoms, can occasionally sense an energy deposit coming from a dark matter particle scattered against the target nuclei in underground detectors. For dark matter masses of $\sim O(10^2)$ GeV (in the electroweak scale) and velocity of $\sim 10^{-3}c$, the typical energy scale of the recoiling signal is $O(10)$ keV. The interaction between dark matter and nuclei can be divided into two classes in the non-relativistic limit: the spin-independent and spin-dependent. The spin independent interaction couples to the mass of the detector nuclei, while the spin dependent couples to the spin of the nuclei. Different experiment material and techniques are sensitive to search for different dark matter interactions with nucleus, some of the most important experiments are LUX [59] and XENON [60].

Indirect detection

The dark matter particles may decay or annihilate with each other and the final products can be detected. These products include γ-rays, neutrinos and charged anti-particles such as positrons and anti-protons. The search for such signals is performed at the directions where the dark matter density is expected to be high. Some of the most important experiments are Fermi-LAT [61], IceCube [62], Super-Kamiokande [63], AMS [64], PAMELA [65] and HESS [66].

Production in colliders

Since the dark matter mass is usually assumed to be $\lesssim O(10^2)$ GeV in the WIMP scenario, the dark matter particles are expected to be generated at high energy colliders. The particles created in the accelerator may decay or annihilate to the dark matter particle, which escapes the detector and therefore their production can be inferred as missing energy. The most important experiments are ATLAS [67] and CMS [68]. Details on dark matter searches at colliders are given in Chapter 2.

Chapter 2

Dark Matter searches at colliders

Dark matter searches in colliders are complementary to direct and indirect searches, described in Section 1.3.3. As shown in Figure 1.8, the three different detection methods exploit the production, scattering, and annihilation processes for dark matter with respect to SM particles. Direct and indirect detection suffer from large uncertainties in the initial state, mainly of astrophysical nature. On the other hand, in colliders the dark matter is produced and studied in a controlled environment. Also, colliders do not have a lower limit on their sensitivity on dark matter masses (few GeV), a region inaccessible to current direct detection experiments. However, the masses of the dark matter candidate and of other particles that mediate the interaction of dark sector with the SM particles have to be within the energy scale accessible to the experiment.

Searches for dark matter have been accomplished at several high-energy colliders, such as the Tevatron at Fermilab or the Large Electron Positron collider at CERN [69, 70, 71]. It is now the data from the Large Hadron Collider at CERN which provide the greatest sensitivity to new exotic processes and give access to the highest energy scales for new physics involving dark matter.

If the dark matter particle is assumed to be a weakly interacting massive particle, with mass at the weak scale, as for massive SM particles, and coupling strengths to SM particles similar to that of the electroweak interactions, the relic abundance of dark matter particles is correctly predicted [72], see discussion in 1.3.2. Even though the WIMP is not the only viable dark matter candidate, it is a compelling scenario, and testing it is a necessary milestone in the understanding of the dark matter puzzle. This electroweak energy scale is powerfully probed by the LHC and thus its data can help to explore a wide region of the WIMP mass-coupling parameter space.

The freeze-out mechanism in the early Universe is based on the assumption that dark matter particles interact strongly enough with ordinary matter that they enter into thermal equilibrium with the SM bath at high temperatures [73, 45]. The critical temperature is reached when the dark matter annihilation rate drops below the expansion rate of the Universe, and therefore the dark matter interactions become insufficient in order to maintain thermal equilibrium. Since the interactions between the SM and dark matter particles have to be sizable for the freeze-out process, the expectation would be that the inverse process also has a sizable cross section. Then, collider experiments can be used to invert the annihilation process that frequently happened in the early Universe, see Figure 1.8.

Furthermore, many beyond the SM (BSM) theories require new stable particles at the electroweak scale which are either a viable dark matter candidate or might couple to a dark matter particle. This is the case for supersymmetric models which not only fix many know SM problems, like the gauge hierarchy problem, but also provide a viable dark matter candidate in the form of the lightest supersymmetric particle [74]. Many searches for these well-motivated extensions of the SM are therefore ultimately also searches for dark matter. Nevertheless, in these complete models, the dark matter is typically produced with additional exotic particles, implying that there is no direct connection between the annihilation and production process, see Figure 1.8. Also, most of the constraints on these models are unrelated to the actual dark matter properties, and the dark matter particle is therefore a tool in the search rather than the actual object of interest. Because of this, a more narrow definition of dark matter searches is adopted in this work, and it is focused on a search with direct pair-production of dark matter. This is described in the following Section 2.2 regarding the theoretical interpretation and Section 2.3 regarding the experimental signature.

2.1 Overview of dark matter theoretical framework

In order to interpret dark matter signatures at LHC with their respective cross section limits, and finally compare these bound to the constraints derived from direct or indirect detection, a dark matter model is needed.

Given that not only the quantum nature of the dark matter particle is unknown, but also which is the connection or portal between the dark sector and the SM, it is important to explore all possibilities that the dark matter "theory space" can offer. This theory space is populated by a large number of all possible realizations of physics beyond the SM with a viable dark matter WIMP candidate. Figure 2.1 illustrate some representative dark matter models.

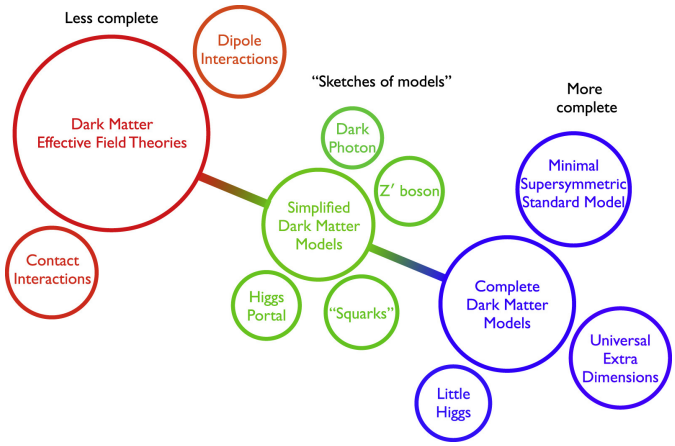


Figure 2.1 – Artistic view of the DM theory space [75]

The models filling this theory space fall into three distinct classes [75] with increasing level of completeness and complexity:

- **Effective Field Theory**
  The Effective Field Theory (EFT) approach allows to describe the interactions between the SM and the dark matter sector mediated by contact/effective operators with higher dimension, > 4, providing a universal description of the dark matter-SM connection. In order to obtain the right dimensionality for the new term(s) in the Lagrangian, a suppression energy scale Λ must be introduced. Since it is a non-renormalizable field theory, Λ is the break down energy scale that represent the masses of those particles which have been integrated out[1]. In this minimal approach, Λ and the dark matter mass are the only free parameters of the model.

  The derived limits are independent of the theory behind the dark matter, yet it has been assumed that other hypothetical particles are too heavy to be produced directly in pp collisions. Some of the possible operators are listed in Table 2.1, for which the dark matter candidate is assumed to be a Dirac fermion [76].

Dimension	Mediator type	Operator
1	Scalar	$\frac{1}{\Lambda^3}\overline{\chi}\chi\overline{q}q$
5	Vector	$\frac{1}{\Lambda^2}\overline{\chi}\gamma^\mu\chi\overline{q}\gamma_\mu q$
8	Axial-vector	$\frac{1}{\Lambda^2}\overline{\chi}\gamma^\mu\gamma^5\chi\overline{q}\gamma_\mu\gamma^\mu q$
9	Tensor	$\frac{1}{\Lambda^2}\overline{\chi}\sigma^{\mu\nu}\chi\overline{q}\sigma_{\mu\nu}q$
11	Scalar	$\frac{1}{\Lambda^3}\overline{\chi}\chi\,\alpha_s\left(G^s_{\mu\nu}\right)^2$

Table 2.1 – Effective interaction operators of WIMP pair production considered in the mono-"X" analyses described in Section 2.3.1, following the formalism of Reference [76].

  EFT was basically the principal dark matter model interpretation for Run 1 data analysis [77, 76, 78, 79, 80]. This approach has proven to be very useful to derive stringent constraints on physics beyond the SM at the scale Λ, however it has an intrinsic limitation: EFT is only valid if the momentum transfer in the process under study is lower than the cut-off scale Λ [81, 82, 83].

- **Simplified models**
  The validity of the EFT approach is questionable at LHC energies where the momentum transfer involved is comparable to the scale of the non-renormalizable operators considered. At such high energies, the degrees-of-freedom (often called *mediators*) that were assumed to generate these operators are important and should be "integrated-in" [84]. The simplified model approach incorporate the effect of these particles in the modeling while considering a simple and inclusive theoretical framework.

  The simplified models can be seen as a step beyond the EFT, for which the level of detail can be expanded by including in a minimal model the extra particles and interactions needed to reproduce the non-renormalizable operators. Including the

[1] In the EFT framework, the mediating particles are experimentally inaccessible and can be "integrated-out".

mediator's propagator allows not only to avoid the energy limitation of the EFT, but also to describe the full kinematics of dark matter production at LHC with an ultra-violet complete model, at the price of an increased number of parameters. The suppression mass scale Λ is traded in for at least the parameters related to the mass of the mediator particle and the respective coupling strengths.

Also, simplified models allow exploiting the complementarity with other searches not related directly to dark matter production such as resonances in the di-jet channel, see Section 2.3.2 for details. This is because the new mediators included in the model can be produced on-shell and contribute significantly to other processes not considered in the EFT.

More details on dark matter simplified models realisations are given in Section 2.2.

- **Complete models**
  At the end of the model spectrum of Figure 2.1, there are the complete BSM theories which include a viable dark matter candidate. This is the case of the theories inspired by the hierarchy or strong CP problems, the grand unification of forces, or neutrino masses. The classical example is the Minimal Supersymmetric SM (MSSM) for which the neutralino is a weakly interacting massive particle that serves as a dark matter candidate [74, 85]. There are other well-motivated extensions of the SM, such as the Large Extra Dimensions (LED) [34], little Higgs models with T-parity [86], etc., all of them predicting dark matter candidates leading to large missing transverse energy at LHC.

  These models may account for all phenomena up to a very high energy scale, typically much higher than the one testable in current colliders. Also, a complete theory framework with full predictive power has a large number of parameters and degrees-of-freedom. This may lead to different dark matter candidates and new interactions, thus it is typically difficult to decorrelate effects and constrain dark matter properties alone.

All the above described theoretical scenarios are interesting, well motivated, and in some cases complementary. They must be considered as potential sources of new physics regarding the dark matter problem. In the following, the attention will be concentrated in the simplified model interpretation which is the focus for the ATLAS Run 2 data analysis regarding dark matter signatures.

2.2 Dark matter simplified model interpretations

In the past years, a lot of progress has been made in testing and understanding the EFT dark matter models at LHC. It was until early Run 2 data taking that simplified models of dark matter started to be in the spotlight of the analysis interpretations, serving as a bridge between the two ends of the spectrum in the theory space illustrated in Figure 2.1.

The simplified models of dark matter aim to provide a realistic and complete scenario within the energy reach of the LHC, restricting as much as possible the number of possible

free parameters. The main characteristics of the simplified models are that they must contain a viable dark matter candidate and a mediator that couple the SM and dark sector, other additional states, if any, should be decoupled. All the renormalizable terms consistent with the SM gauge symmetry and Lorentz invariance should be included in the Lagrangian. Finally, in order to focus on the dark matter constraints, lepton and baryon number conservation is assumed, together with minimal flavour violation (MFV) [87, 75].

One main motivation for simplified models is that they allow the possibility of producing missing transverse energy without other exotic objects, as it is the case for complete models. For example, in LED scenarios particles escape to extra dimensions, or in MSSM a pair of lightest supersymmetric particle can be produced, and both signatures leads to missing transverse energy.

In any case, most simplified models can be understood as the limit of a more general or complete beyond the SM scenario, where the majority of the new particles are integrated out and decoupled from the energy scale reachable at the LHC or because they have no role in dark matter interactions with the SM. Similarly, in the limit where the mass of the mediator is very large, the EFT may be recovered by integrating out the mediator.

2.2.1 Simplified models with vector or axial-vector mediators

The majority of simplified models considered in dark matter searches at LHC are s-channel production, as illustrated in Figure 2.2. The mediator sector between SM and dark matter particles is composed by a single particle, named $Z'_{V/A}$ in Figure 2.2. This mediator is usually assumed to be a boson, describing a vector or axial-vector interaction in case of spin-1 particle [88, 89]. It can be also considered a scalar or pseudo-scalar interaction for spin-0, however they are not gauge invariant [84].

One of the simplest ways to add a new mediator to the SM is by extending its gauge symmetry by a new $U(1)'$, which is spontaneously broken such that the mediator obtains mass [90]. The Dark Matter candidate has charges only under this new group. It is assume that some SM particles has also charges under this group, therefore the new boson can mediate interactions between the SM and the dark matter sectors.

If the dark matter particle χ is assumed to be a fermion (either Dirac or Majorana), and assuming CP-conservation, the Lagragian of the models with a vector or axial-vector couplings between the spin-1 mediator Z' is:

$$\mathcal{L}_{\text{vector}} \supset \frac{1}{2} m^2_{Z'} Z'_\mu Z'^\mu - g_q \sum_{q=u,d,s,c,b,t} Z'_\mu \bar{q}\gamma^\mu q - g_\chi Z'_\mu \bar{\chi}\gamma^\mu \chi \tag{2.1}$$

$$\mathcal{L}_{\text{axial-vector}} \supset \frac{1}{2} m^2_{Z'} Z'_\mu Z'^\mu - g_q \sum_{q=u,d,s,c,b,t} Z'_\mu \bar{q}\gamma^\mu\gamma_5 q - g_\chi Z'_\mu \bar{\chi}\gamma^\mu\gamma_5 \chi \tag{2.2}$$

The mediator couplings to quarks was chosen to be diagonal, since the existence of off-diagonal couplings is tightly constrained by various flavour changing neutral current processes [91]. As a further simplification, the couplings are fixed to be flavour blind, $g^i_d = g^i_u = g_q$ for $i = 1,2,3$. Therefore, the mediators are assumed to couple to all quarks with equal strength, consistent with the hypothesis of minimal flavor violation [92]. Typical

values of the couplings to quarks chosen are $g_q = 0.1, 0.25$, and the vertex between the dark matter particle and mediator is always taken to be $g_\chi = 1.0$ [89].

This model can be expanded to contain charged lepton couplings of the mediator, equal for all lepton flavors, with a typical value of $g_\ell = 0.01$ as described in Reference [93]. Each of these simplified models is fully characterized by five parameters: the masses of the two particles m_χ and $m_{Z'}$, the flavour universal coupling to quarks g_q, the coupling to all lepton flavours g_ℓ, and the coupling to the dark matter g_χ.

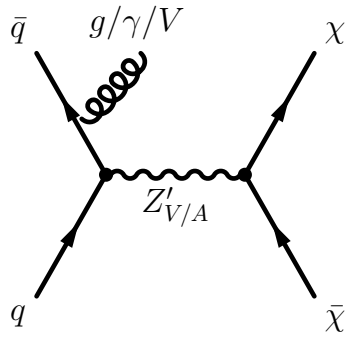


Figure 2.2 – Leading order Feynman diagram of the dominant dark matter production mode for the vector and vector-axial models.

2.2.2 Two-Higgs-Doublet model with a vector mediator Z'

In this model, a resonant production of a new Z' gauge boson is considered, as in the vector model described in Section 2.2.1. This Z' then decays to a Higgs boson plus an intermediate state which decays to a dark matter pair, as depicted in Figure 2.3. Since the decay of a SM particle to dark matter is highly constrained, a heavy pseudoscalar A^0 from a two-Higgs doublet extension to the SM is considered, and this pseudoscalar has a large branching ratio to dark matter [94]. Thus, the phenomenology of this model is extended with respect to the simplified case due to the presence of a new decay mode $Z' \to hA^0$.

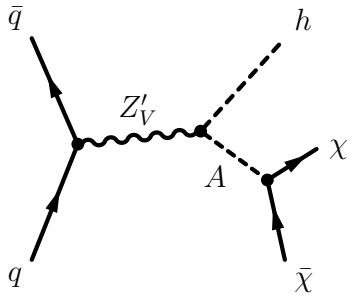


Figure 2.3 – Leading order Feynman diagram of the dominant dark matter production mode for the Z'-2HDM model.

The gauge symmetry of the SM is extended by a $U(1)'$ associated to a new Z' gauge boson. This Z' is massive thanks to an additional scalar singlet ϕ that leads to spontaneous symmetry breaking at a scale above the electroweak symmetry breaking.

For simplicity, it is assumed generation-independent charges under $U(1)'$ for the fermions and that only quarks are charged. In this case, the Z' can be produced at LHC, and also stringent constraints from di-lepton resonances searches are avoided, since leptons are neutral under $U(1)'$.

Regarding the scalar sector, a 2HDM Type 2 is considered, which comprises two scalar doublets: Φ_u couples to the up-type quarks and Φ_d couples to the down-type quarks and leptons [95]:

$$\mathcal{L} \supset y_u Q\tilde{\Phi}_u\overline{u} + y_d Q\Phi_d\overline{d} + y_e L\Phi_d\overline{e} + \text{h.c.} \tag{2.3}$$

After electroweak symmetry breaking, the Higgs doublets obtain vacuum expectation values (vev), v_u and v_d. In unitary gauge, the doubles are parametrized as:

$$\Phi_d = \frac{1}{\sqrt{2}}\begin{pmatrix} -\sin\beta\, H^+ \\ v_d - \sin\alpha\, h + \cos\alpha\, H - i\sin\beta\, A^0 \end{pmatrix}, \tag{2.4}$$

$$\Phi_u = \frac{1}{\sqrt{2}}\begin{pmatrix} \cos\beta\, H^+ \\ v_u - \cos\alpha\, h + \sin\alpha\, H - i\cos\beta\, A^0 \end{pmatrix}, \tag{2.5}$$

where h, H are neutral CP-even scalars, $H^\pm$ is a charged scalar, and A^0 is a neutral CP-odd scalar. Furthermore, the mixing angles that diagonalizes the $h-H$ mass matrix are $\tan\beta = v_u/v_d$ and α.

The scalar h is taken to be the observed Higgs boson with $m_h \sim 125$ GeV. The Type-2 2HDM is tightly constrained around the alignment limit where $\beta \to \alpha + \pi/2$ and $\alpha \in (-\pi/2, 0)$. In this limit, h has SM-like couplings to fermions and gauge bosons [96]. Furthermore, pertubativity of the top yukawa coupling implies $\tan\beta \geq 0.3$ [94]. Because of this, the parameter space is chosen to have $\beta \to \alpha + \pi/2$ and $\tan\beta \geq 0.3$.

The vev of the Higgs lead to $Z - Z'$ mass mixing with the consequence of a modification to the Z mass, which is dependent to a small mixing parameter ϵ. The decay width of $Z' \to hA^0$ can be expressed to leading order in ϵ as [94]:

$$\Gamma_{Z'\to A^0} = (g_z \sin\beta\cos\beta)^2 \frac{|p|}{24\pi}\frac{|p|^2}{M_{Z'}} \tag{2.6}$$

where the center of mass momentum for the decay products is

$$|p| = \frac{1}{2M_{Z'}}\sqrt{\left(M_{Z'}^2 - (m_h + m_{A^0})^2\right)\left(M_{Z'}^2 - (m_h - m_{A^0})^2\right)}. \tag{2.7}$$

The model is described by five parameters: the pseudoescalar mass m_{A^0}, the dark matter mass m_χ, the new boson mass $m_{Z'}$, $\tan\beta = v_u/v_d$ and its coupling strength $g_{Z'}$. The benchmark values used in this model are $g_{Z'} = 0.8$, $\tan\beta = 1$ and $m_\chi = 100$ GeV [89].

2.3 Dark matter searches at LHC

Weakly interacting massive particles are electromagnetically neutral, their interaction with the SM particles is very weak, and have a stability of the order of the lifetime of the Universe. Because of this, when produced in a pp collision, dark matter particles does not interact with the detector and live enough to escape undetected. Thus, they traverse the various detector layers without leaving a trace, just like neutrinos.

However, by using the hermeticity of the detector, the presence of dark matter particles can be inferred via an apparent imbalance in the total momentum. In hadron colliders, since the longitudinal momenta of the colliding partons are unknown, only the transverse momentum can be used in dark matter searches. From the momentum conservation, the net momentum in the transverse plane is zero before and after the collision. An imbalance in this plane, obtained as the negative vector sum of the transverse momenta of all detected particles, is called "missing transverse momentum", or $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$, and it could be a hint of dark matter recoiling against SM particles. Its magnitude $E_{\mathrm{T}}^{\mathrm{miss}}$ is called "missing transverse energy", Section 4.6 details the missing transverse energy reconstruction and performance in ATLAS.

Even though WIMP particles can not be directly observed, its search at the LHC is a thriving research field due to the experimental feasibility of the searches. The SM backgrounds and their respective distributions are now sufficiently understood with very good accuracy that small distortions in their corresponding observables, mainly in the missing transverse energy spectrum (particularly on its high momentum tail), may be observed and used to constraint dark matter models.

Collider searches can be broadly distinguished into two categories: Searches for the dark matter particle in final states, described in Section 2.3.1, and searches for the dark matter-SM sector mediator in final states without dark matter particles, described in Section 2.3.2.

2.3.1 Searches for invisible final states: "Mono-X" searches

The basic diagram of dark matter pair production is shown in Figure 2.4(a), where the incoming partons, or their decay products, witness new exotic interactions and produce a pair of dark matter particles in the final state. Once produced, these particles escape the detector without any energy deposit, as discussed before. If the pair of dark matter particles is the decay product of a heavy mediator produced at rest, momentum conservation implies that the final state would be a pair of back-to-back dark matter particles, with balancing transverse momentum and $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}} = 0$. Such empty event will not be detectable, as depicted in Figure 2.4(b).

Missing transverse energy/momentum is not enough for an event to be recorded by the detector. Then, in order to trigger dark matter events, at least one recoiling SM particle is necessary.

The leading generic diagrams responsible for dark matter production at LHC involve the pair production of WIMPs plus a SM particle, named "X" as shown in Figure 2.4(c). If the dark matter is produced in association with a SM particle "X", the final state is expected to present a well defined "X" signal recoiling against large amount of missing

transverse momentum, $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$, corresponding to the dark matter pair production, as shown in Figure 2.4(d).

The "X" particle can be radiated by the initial- or final-state, and is necessary to balance the two WIMP's momentum, so that they are not produced back-to-back as in Figure 2.4(b) resulting in negligible $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$. These SM recoiling objects can be photons [97], electroweak bosons [98], heavy flavour quarks [99, 100], hadronic jets [101] or even a Higgs boson [5].

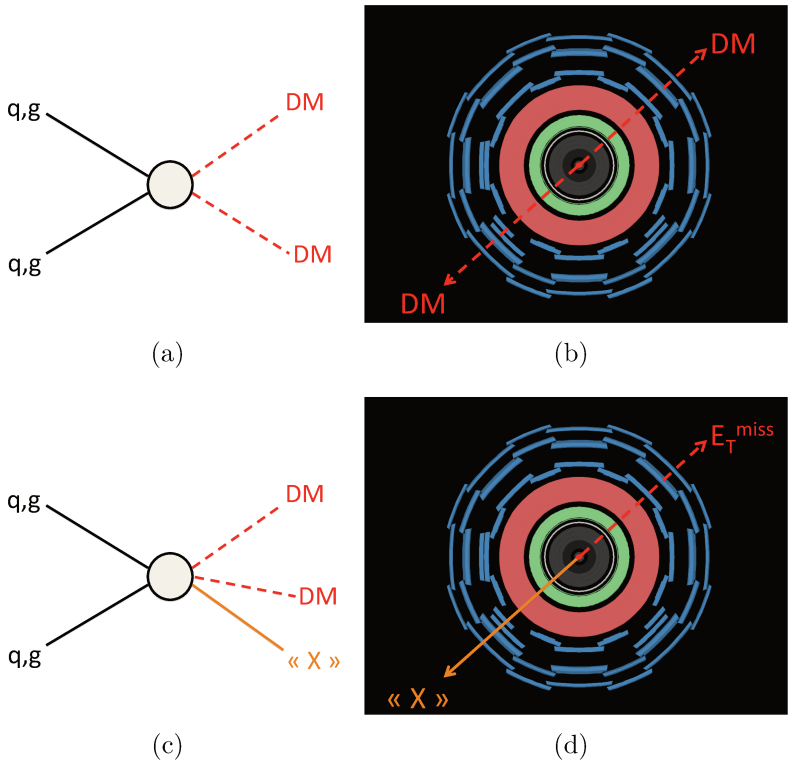


Figure 2.4 – Basic dark matter production diagram 2.4(a) and transverse plane scheme 2.4(b), and generic mono-"X" diagram 2.4(c) and transverse plane scheme 2.4(d).

Mono-"X" analyses are characterized by selecting the "X" signal with high quality criteria and vetoing any other SM activity in the detector, and requiring large $E_{\mathrm{T}}^{\mathrm{miss}}$. Also, requirements in the topology are considered, such as the azimuthal separation between the produced SM "X" and the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$, from the dark matter pair production, which should be large since these two objects are expected to be recoiling against each other in the transverse plane, i.e. $\Delta(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}},\mathrm{X}) \sim \pi$. For a mono-"X" analyses with an hadronic final state, the dominant background sources are $Z(\rightarrow \nu\nu)$ + jets and $W(\ell\nu)$ + jets where the lepton fails the reconstruction, given that both process are characterized by real $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ coming from the SM neutrinos.

In the following, a general overview of the principal mono-"X" searches will be given:

Mono-jet

This search is the canonical mono-X analysis. Dark matter is produced in association with one or more QCD jets from initial state radiation. This search selects events with at least one energetic jet (p_{T} >250 GeV and $|\eta| < 2.4$) that fulfils high jet-reconstruction quality criteria, and impose a strict veto on events containing leptons. Large hadronic activity is expected coming from this high p_{T} jet and from up to four other jets (with p_{T} > 30 GeV and $|\eta|$< 2.8) recoiling against large amount of missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$ >250 GeV), see Figure 2.5 for a representative mono-jet topology event display. Back-to-back dijet events are suppressed by requiring that the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ vector does not point into the azimuthal direction of the leading jets. This requirement also suppresses jet mismeasurement, that leads to an imbalance in the transverse momentum from multijet background.

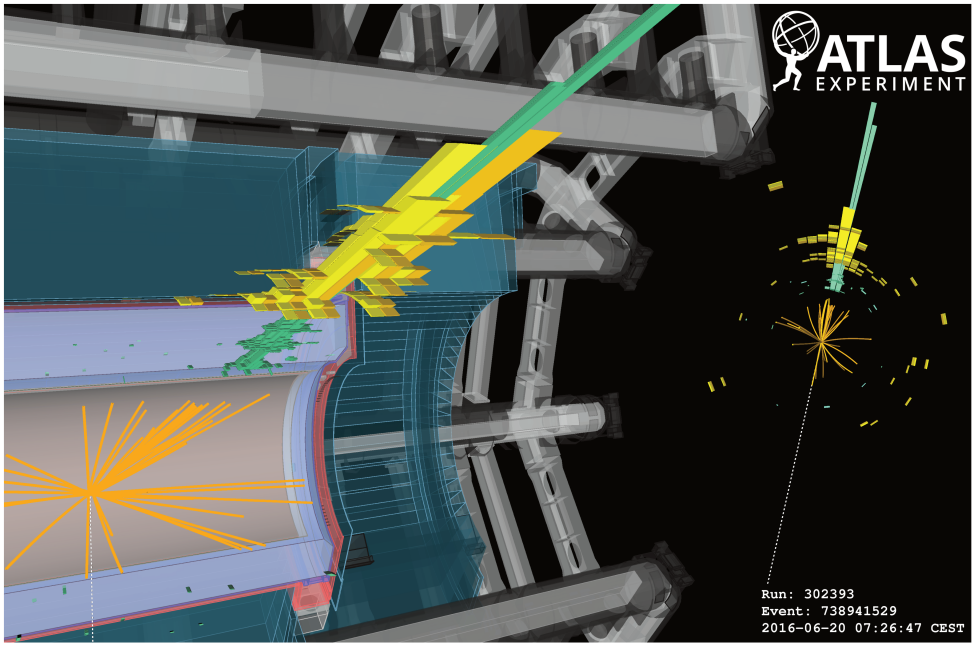


Figure 2.5 – Mono-jet event display recorded on 2016 with a energetic jet with $p_{\mathrm{T}} = 1707$ GeV and a invisible recoil of $E_{\mathrm{T}}^{\mathrm{miss}} = 1735$ GeV [101].

Figure 2.6 shows the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution of selected mono-jet events in 36 fb^{-1}, where the stacked histograms of background processes is overlaid on various signal models depicted as thick lines [101]. It is notable that the signal model for an axial-vector mediator of 1 TeV mass , $(m_{\mathrm{DM}}, M_{\mathrm{med}}) = (400, 1000)$ GeV, is flatter than the backgrounds, thus the original signal-to-background ratio increases with $E_{\mathrm{T}}^{\mathrm{miss}}$. Dark matter signals are expected to appear as an excess in the tail of the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution. Exclusion limits on mono-jet production are derived from the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution that lead to the currently strongest constraints on dark matter from collider searches.

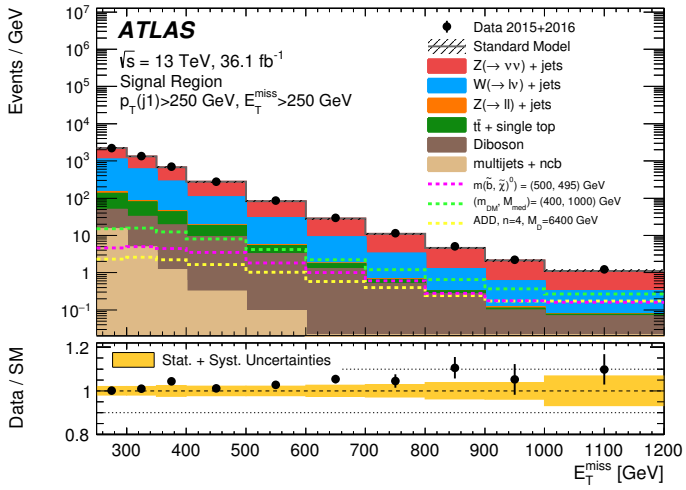


Figure 2.6 – Measured distribution of the $E_{\mathrm{T}}^{\mathrm{miss}}$ compared to SM predictions. The latter are normalized with normalization factors as determined by a global fit that considers exclusive $E_{\mathrm{T}}^{\mathrm{miss}}$ regions [101].

Mono-V

Mono-V searches for dark matter production in association with a vector boson V, $V = W, Z$. Three different analysis can be considered depending on the vector boson decay mode: full hadronic, leptonic and semi-leptonic.

Mono-V(had) searches hadronic final states + $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ with no reconstructed lepton. In this case, jets in the final state comes from the hadronic decay of a vector boson. Even though the production cross section is significantly smaller than for QCD radiation, the process is much cleaner and thus can be searched for with higher sensitivity. At the energies provided by the LHC, their decay products may be boosted, collimating the daughter particles in a large radius jet with an internal substructure due to the two quarks from $V \to qq$, see Section 4.2.2.2 for more details on this reconstruction technique.

Also, leptonically decaying Z bosons can be considered, having a very clean signal [102]. It is know as mono-Z analysis. In this case, it is required that the transverse momentum of the di-lepton system has opposite azimuthal direction with respect to $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$, and that the di-lepton invariant mass is in a narrow window around the Z boson mass.

Another analysis related to vector bosons, consists in a search for dark matter with a W boson that decays leptonically, the produced neutrino is considered in the missing transverse momentum and, therefore, this is called a mono-lepton event [103].

Mono-photon

Mono-γ searches were the first dedicated dark matter searches at colliders. These searches require typically only the presence of a high p_{T} photon and reject isolated leptons [104] or a second jet. Because of the relative strength on the electromagnetic coupling α_{EM} compared to the strong coupling constant, the sensitivity of mono-γ searches is reduced in comparison to hadronic searches. However, background levels are generally very low making this topology a particularly clean channel to study dark matter.

Mono-Higgs

The discovery of Higgs boson at the LHC [20, 105] opens the opportunity of searching the dark matter pair production in association to a Higgs boson. The observed final state are $E_{\mathrm{T}}^{\mathrm{miss}}$ plus the Higgs decay products with an invariant mass of $m_h \sim 125$ GeV.

There is an important difference between mono-Higgs and other mono-"X" searches. In pp collisions a jet, photon or vector boson can be emitted directly from initial state radiation through the usual gauge interactions. In contrast, the initial state radiation of a Higgs boson is highly suppressed due to the small yukawa coupling of the Higgs boson to the quarks. In that sense, a positive mono-Higgs signal would be more closely connected to the dark matter pair production, providing a direct probe of the SM-dark matter sector coupling.

There are two promising Higgs boson decay channels to search for dark matter: $h \to b\bar{b}$ and $h \to \gamma\gamma$. The $b\bar{b}$ channel is the dominant decay mode of a Higgs boson with a mass of $m_h \sim 125$ GeV, corresponding to the largest branching ratio $Br(h \to b\bar{b}) \sim 0.577$ [106], and gives the best statistics for the signal. While the diphoton branching ratio is only $Br(h \to \gamma\gamma) \sim 2.28 \times 10^{-3}$, but is a very clean channel [107].

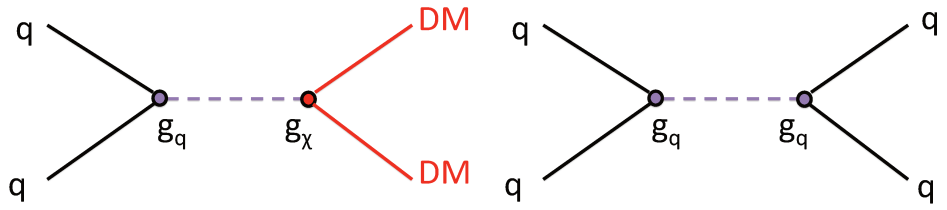
This work is focused in the mono-Higgs search for dark matter with the Higgs boson decaying to a pair of bottom quarks. The strategy, analysis and results are presented in Chapters 6, 7 and 8.

2.3.2 Searches for visible final states: "Di-X" resonance searches

In the mono-"X" topology, described in Section 2.3.1, the mass of the dark matter is small compared with respect to the mediator, see Section 2.2 for an overview on simplified models. As a result, the dark matter production cross section receives a resonant enhancement and also the produced WIMP pair is boosted opposite to the direction of the visible SM particle(s), "X". In the case of dark matter particles heavier than the mediator, this mediator becomes off-shell resulting in weaker constraints due to a kinematically suppressed production cross section. The latter case are searches for the mediator, which look for a narrow peak in the invariant mass of two SM particles. Multiple searches for narrow resonances, for example di-jet or di-lepton resonances, can be interpreted in terms of dark matter models.

The di-jet resonance searches are based on the reconstruction of two high-energy jets in the final state, and were usually used in the search for new Z'-like resonances and for new contact interactions at high energy scale [108]. New physics signals can be visible as a resonant peak in the di-jet invariant mass spectrum over a smoothly falling background from QCD jet production.

These analysis are also interesting to study the coupling of the dark matter mediator with SM particles. In this sense, a new s-channel produced particle, that serves as a mediator to the dark matter sector, can be observed as a resonance in the invariant dijet mass at about the new particle's mass [109]. Instead of a dark matter pair production in the final state via the exotic mediator as depicted in Figure 2.7(a), visible decays of the mediator are searched for, as shown in Figure 2.7(b).



(a) (b)

Figure 2.7

In the light of simplified models, "di-jet" signals are interpreted as a mediator decaying back to SM particles. In the EFT approach, dark matter particles can interact with quarks via an effective operator of the form $\frac{1}{\Lambda}\bar{q}\gamma^\mu q\bar{\chi}\gamma_\mu\chi$ and one would also expect an analogous operator involving only quarks of the form $\frac{1}{\Lambda'}\bar{q}\gamma^\mu q\bar{q}\gamma_\mu q$, with similar energy suppression scales, $\Lambda \sim \Lambda'$.

There are different di-jet analysis strategies. The classic di-jet resonance search is based on a fit on the di-jet invariant mass spectrumin the data using an analytical function in sliding windows [110]. For each event it is required that the highest unprescaled single-jet trigger is fired. This trigger has an efficiency greater than 99.5% for $p_{\mathrm{T}}^{\mathrm{lead}} > 428$ GeV, thus the leading jet is required to have $p_{\mathrm{T}} > 440$ GeV. In addition to that, it is required that the subleading jet has $p_{\mathrm{T}} > 60$ GeV, in order to obtain almost pile-up insensitive distributions.

Searches for dijet resonances with sub-TeV masses are statistically-limited by the bandwidth available to inclusive single-jet triggers. Due to large SM multijet backgrounds, these triggers must be prescaled in order to record full events at a manageable rate. However, one can avoid this limitation and record much higher event rates by recording only partial events, containing just the small subset of the full event that is needed for the search. We call this technique Trigger-Level Analysis (TLA). The "di-jet TLA" analysis search for new physics in the di-jet final state using the TLA strategy targets low-mass dijet resonances between 470 GeV and 1.5 TeV [111].

Sensitivity to light resonances is reduced by immense background rates that would saturate the trigger and data acquisition systems. An alternative strategy to search for low mass di-jet resonances is to select events with jets recoiling against a photon or an additional jet from initial state radiation (ISR). In this case, the trigger threshold limitations are reduced by examining data where the light resonance is boosted in the transverse direction via recoil from high-p_{T} ISR of a photon or jet. This analysis is called "di-jet+ISR" [112]. If the associated photon or jet has large transverse momentum, the di-jet resonance candidate can be reconstructed as a large-R jet, see details on this objects on Section 4.2.2.2.

Finally, resonant high-mass phenomena in "di-lepton" final states can be considered [113]. The pair of same-flavour leptons with highest p_{T} are chosen as the resonance candidate. This experimental signature benefits from a fully reconstructed final state, high signal selection efficiencies and relatively small, well-understood backgrounds, representing a powerful test for a wide range of theories beyond the SM.

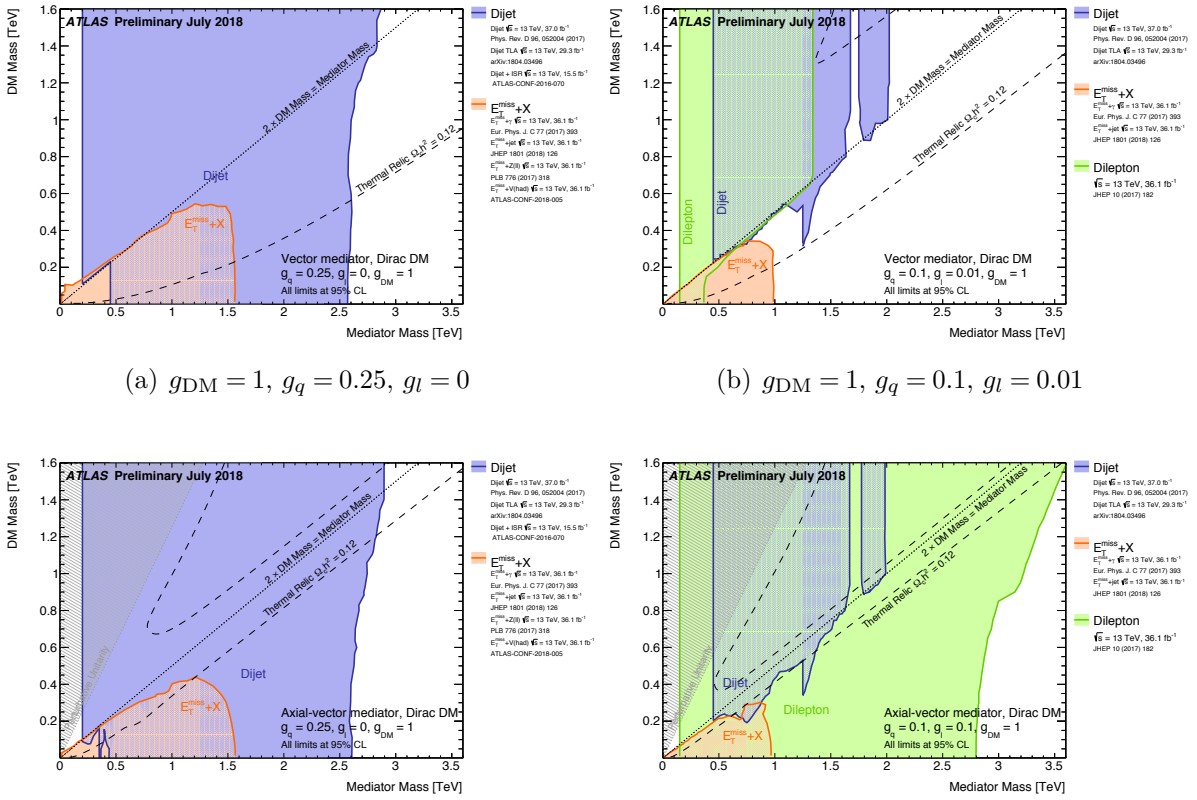
Other dark matter searches with visible final states are: "di-bjet" [114] and "$t\bar{t}$ resonance" [115].

2.3.3 Complementarity and combination of signatures

The complementarity of the various searches is studied in the plane of the mediator mass vs the dark matter candidate mass for fixed couplings for the vector and axial-vector simplified models.

Mono-X and Di-X searches have different different sensitivity reach depending on the simplified model and on the coupling strength choice. For example, Figure 2.8 shows the exclusion regions of the various experimental signatures for the vector mediator (upper plots) and the axial-vector mediator (bottom plots) simplified models. The sensitivity of the searches is highly dependent on the coupling choice mainly for resonance searches. Therefore, in order to highlight the complementarity of the dedicated dark matter searches, different coupling scenarios are considered in the presentation of the results: $g_{\mathrm{DM}} = 1$, $g_q = 0.25$, $g_l = 0$ for a leptophobic mediator; $g_{\mathrm{DM}} = 1$, $g_q = 0.1$, $g_l = 0.01$ for a leptophilic vector mediator; and $g_{\mathrm{DM}} = 1$, $g_q = 0.1$, $g_l = 0.1$ for a leptophilic axial-vector mediator.

Collider experiments represent a complementary model-dependent approach to dark matter searches with respect to direct and indirect detection experiments [118]. It is therefore interesting and informative to compare the models with vector or axial-vector model limits with the results from other dark matter searches. Figure 2.9 show the translation of the vector (left) and axial-vector (right) model limits into limits on the spin-dependent χ-proton and spin-independent χ-nucleon scattering cross-section as a function of the dark matter mass, following the procedure on Reference [93]. As can be seen from Figure 2.9(a) for vector and Figure 2.9(b) for axial-vector mediator models, the LHC searches allow to complement these limits in the low dark matter mass range where the direct dark matter search experiments have less sensitivity due to the very low-energy recoils that such low-mass particles would induce.



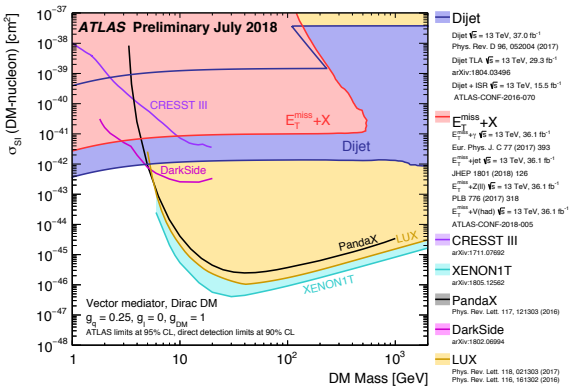
(a) $g_{\mathrm{DM}} = 1$, $g_q = 0.25$, $g_l = 0$

(b) $g_{\mathrm{DM}} = 1$, $g_q = 0.1$, $g_l = 0.01$

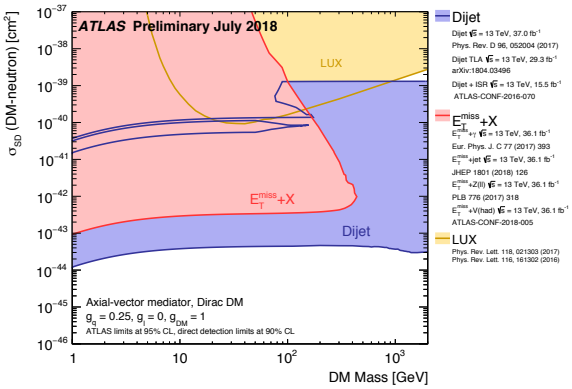
(c) $g_{\mathrm{DM}} = 1$, $g_q = 0.25$, $g_l = 0$

(d) $g_{\mathrm{DM}} = 1$, $g_q = 0.1$, $g_l = 0.1$

Figure 2.8 – Regions in a dark matter mass-mediator mass plane excluded at 95% CL by di-jet, di-lepton and mono-X searches, for vector (top) or axial-vector (bottom) mediator simplified models. The exclusions are computed for a dark matter coupling g_{DM} (also g_χ), quark coupling g_q, universal to all flavors, and lepton coupling gl as indicated in each case. Dashed curves labeled "Thermal Relic" indicate combinations of dark matter and mediator mass that are consistent with a dark matter density of $\Omega_C = 0.12h^2$ and a standard thermal history, as computed in MadDM [116]. Between the two curves, annihilation processes described by the simplified model deplete Ω_C below $0.12h^2$. A dotted curve indicates the kinematic threshold where the mediator can decay on-shell into dark matter. Excluded regions that are in tension with the perturbative unitary considerations of [117] are indicated by shading in the upper left



(a)



(b)

Figure 2.9 – Comparison of the inferred limits to the constraints from direct detection experiments on the spin-dependent WIMP-neutron scattering cross section for the 2.9(a) vector mediator and the 2.9(b) axial-vector mediator. LHC limits are shown at 95% CL and direct detection limits at 90% CL. The comparison is valid solely in the context of these models with the shown specific couplings. LHC searches and direct detection experiments exclude the shaded areas.

Chapter 3

The Large Hadron Collider and the ATLAS experiment

The analyses presented in this work use proton-proton collisionat a center of mass energy of 13 TeV collected with ATLAS (A Toroidal LHC Apparatus) detector at the LHC (Large Hadron Collider) during 2015, 2016 and 2017. This chapter introduces the LHC accelerator in Section 3.1 and the ATLAS detector with its sub-detectors in Section 3.2.

3.1 The Large Hadron Collider

The Large Hadron Collider [119] is a circular proton-proton particle accelerator and collider of the European Organization for Nuclear Research (CERN) [120], located near Geneva at the French-Swiss border.

The LHC is a proton-proton (pp) and also Pb-Pb and p-Pb collider with a circumference of 27 km and is installed at an average depth of 100 m in the tunnel of the Large Electron-Positron collider (LEP) [121] that was operational until 2000.

The accelerator complex at CERN is a succession of machines that accelerate particles to increasingly higher energies. Each machine boosts the energy of a beam of particles, before injecting the beam into the next machine in the sequence. In the LHC - the last element in this chain - particle beams are accelerated up to the record energy of 6.5 TeV per beam. The LHC and the pre-accelerator chain are shown in Figure 3.1. Protons are obtained from hydrogen atoms and are formed into bunches consisting of $\sim 10^{11}$ protons with a nominal bunch spacing of 25 ns. These bunches are accelerated in the LINAC2 up to an energy of 50 MeV, then in the Proton Synchroton Booster (PBS) to 1.4 GeV. The Proton Synchrotron (PS) accelerates the protons up to 25 GeV before being injected into the Super Proton Synchrotron (SPS) which accelerates them to 450 GeV.

Protons in the LHC circulate in two counter-rotating beams bent by large magnetic fields acquiring energy at each round, until they are finally brought to collision. The energy in the centre-of-mass of the pp collision was $\sqrt{s} = 7$ TeV in 2011, reaching 8 TeV in 2012. The first LHC period of collisions, called Run 1, ended at the beginning of 2013. It was followed by a long shutdown (LS1) for planned upgrades. The second LHC period of collisions, Run 2, started in June 2015 with an energy in the centre-of-mass of 13 TeV

and will end by November 2018. It will be followed by a new shutdown (LS2) until early 2021 when LHC will restart operations [122].

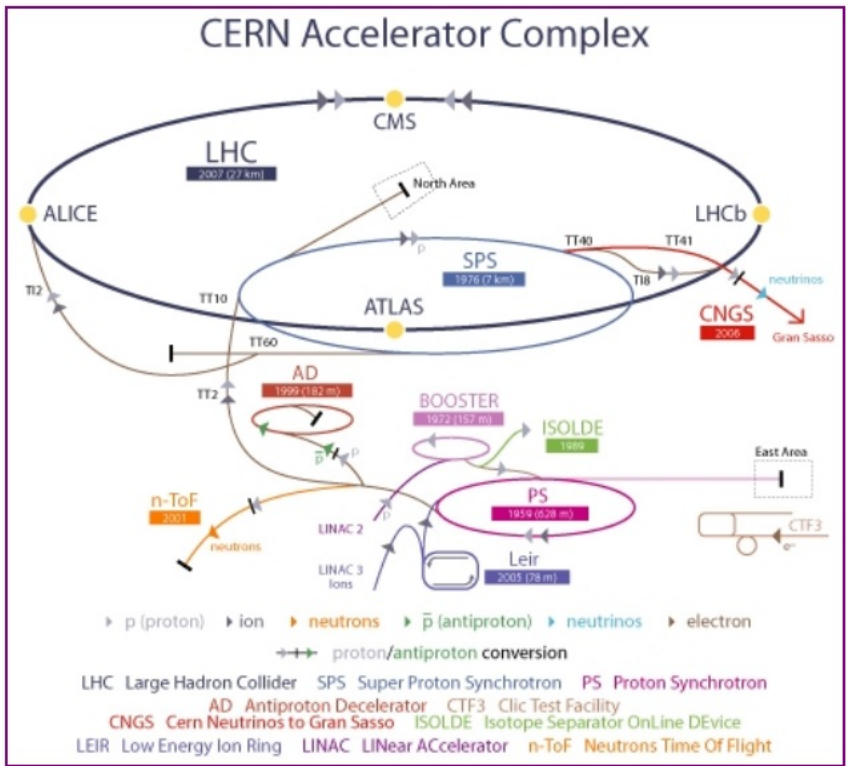


Figure 3.1 – The accelerator complex at CERN [123].

The proton beams intersect each other at four interaction points, where the main particle detectors are located: ALICE [124], ATLAS [125], CMS [126] and LHCb [127]. ATLAS and CMS are multi-purpose devices, optimised for proton-proton interactions in order to perform precision measurements of the SM and to search for new physics BSM. ALICE is optimised for the study of interactions between heavy ions, whereas LHCb is aimed for the particular study of the decays of the b-quarks. Furthermore, two smaller experiments are in operation that are focusing on diffractive physics, namely LHCf [128] and TOTEM [129].

The instantaneous luminosity is the quantity to maximize in order to increase the probability to detect a given process. The luminosity $\mathcal{L}$ depends on the beam parameters and can be written as,

$$\mathcal{L} = \frac{N_b n_b f \gamma}{4\pi \epsilon_n \beta^*} F(\theta_c), \tag{3.1}$$

where N_b is the number of particles per bunch, n_b the number of bunches per beams, f the revolution frequency, γ the relativistic Lorentz factor, ϵ_n the normalised transverse beam emittance, β^* a function which give the amplitude of oscillations of the protons at the collision point and $F(\theta_c)$ a geometric luminosity factor due to the crossing angle at

the intersection point:

$$F(\theta_c) = \left(1 + \left(\frac{\theta_c \sigma_z}{2\sigma_*}\right)^2\right)^{1/2}, \tag{3.2}$$

where θ_c is the crossing angle of the beams, σ_z and σ_* are the RMS of the longitudinal and transversal bunch length.

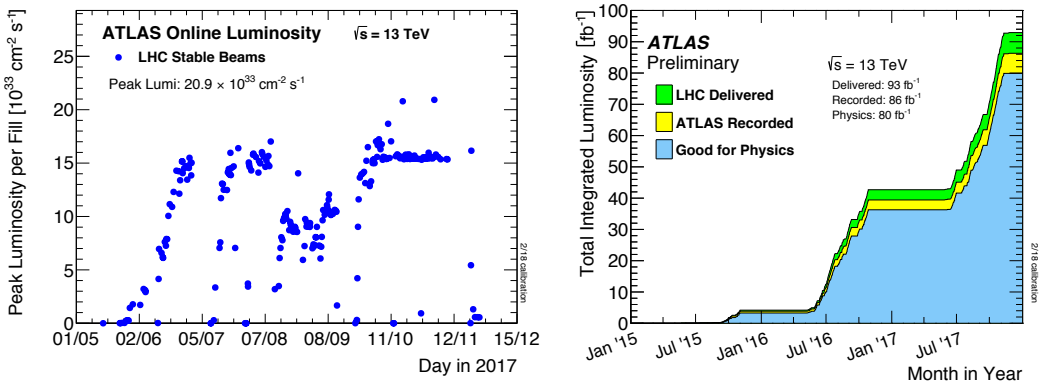


Figure 3.2 – Instantaneous (left) and integrated (right) luminosity for ATLAS Run 2 data taking [130].

Figure 3.2 (left) shows the instantaneous peak luminosity delivered to ATLAS during pp collisions at $\sqrt{s} = 13$ TeV in 2017. The instantaneous luminosity integrated over a specific period of time is referred to as the integrated luminosity and it allows to quantify the amount of data delivered by the LHC,

$$\int \mathcal{L}\, dt = N_{\text{events}}\, \sigma\,, \tag{3.3}$$

where σ is the process cross section. Figure 3.2 (right) shows the time evolution of the integrated luminosity delivered by the LHC (green), recorded by ATLAS (yellow) and the certified runs with good-quality data (blue) during pp collisions in Run 2 until the end of 2017.

The probability to have multiple inelastic pp collisions increases linearly with the luminosity. Other pp collisions, in addition to the collision of interest, are collectively referred as "pile-up" (PU). It presents a serious challenge to physics analyses at the LHC since the reconstructed physics object are affected from additional energy contributions. In-time pile-up corresponds to the multiple pp collisions within the same bunch crossing. The out-of-time PU corresponds to pp interactions occurred in a previous bunch crossing. It is the result of long electronic integration times, becoming significant when the spacing decreases between the bunch crossing. The amount of pileup activity can be estimated in terms of the mean number of interactions per bunch crossing, $< \mu >$, which is an average over a given block of events measured by dedicated ultra-fast detectors, and in terms of the number of reconstructed primary vertices N_{PV}, measured in the tracking detectors on an event-by-event basis. Figure 3.3 shows the mean number of interactions per crossing for the 2015, 2016 and 2017 pp collision data at 13 TeV centre-of-mass energy.

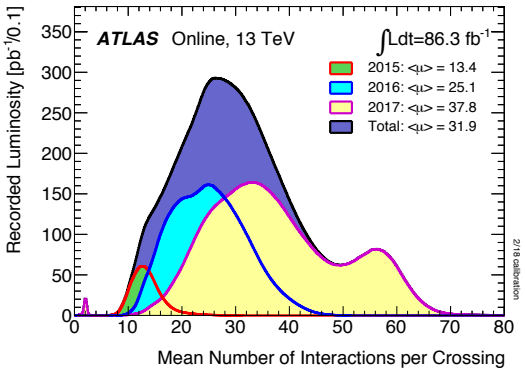


Figure 3.3 – Number of Interactions per Crossing: Luminosity-weighted distribution of the mean number of interactions per crossing for the 2015, 2016 and 2017 pp collision data at 13 TeV centre-of-mass energy. All data recorded by ATLAS during stable beams are shown. The mean number of interactions per crossing corresponds to the mean of the Poisson distribution of the number of interactions per crossing calculated for each bunch.

3.2 The ATLAS detector

The ATLAS detector [125] is a multi purpose detector located at Point 1 on the LHC ring near the main CERN site. It is 44 m long, has a diameter of 25 m and weights around 7,000 tons. It has a typical onion-shell-like structure and covers almost the full 4π solid angle. Figure 3.4 shows a general layout of the ATLAS detector.

The origin of the ATLAS coordinate system lies in the interaction point. ATLAS uses a right-handed system in which the x-component points towards the centre of the LHC ring, the y-component points upwards and the z-component points into the direction of the beampipe. The polar angle θ is defined in the yz-plane and is measured from the z-axis. The azimuthal angle ϕ is defined in the xy-plane. The pseudorapidity is defined using the polar angle θ as $\eta = -\ln\tan\left(\frac{\theta}{2}\right)$. The difference in pseudorapidity of two objects, $\Delta\eta$, is invariant under Lorentz-transformations. For massive objects such as jets, the difference in rapidity Δy is used, where $y = \frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right)$. The distance ΔR in the pseudorapidity-azimuthal angle space is defined as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

ATLAS provides discriminant detection power and identification between electrons, photons, muons and products from hadronic processes leaving wide signals called jets. Combined signals from various subdetectors allow an identification of the various particles together with an estimation of their energy and direction. The tracks of the charged particles are measured by the Inner Detector (ID). It is installed close to the beam pipe and is encompassed in a solenoidal magnetic field of 2 T. The trajectories of charged particles are bent by the magnetic field which allows the determination of the sign of their charge and their transverse momentum. Furthermore, the highly segmented structure provides a good spatial resolution to reconstruct primary and secondary vertices. After the particles traverse the ID, they may induce electromagnetic and hadronic showers in the calorimeter system. Electrons, photons and jets are fully absorbed by the calorimeters. The outer

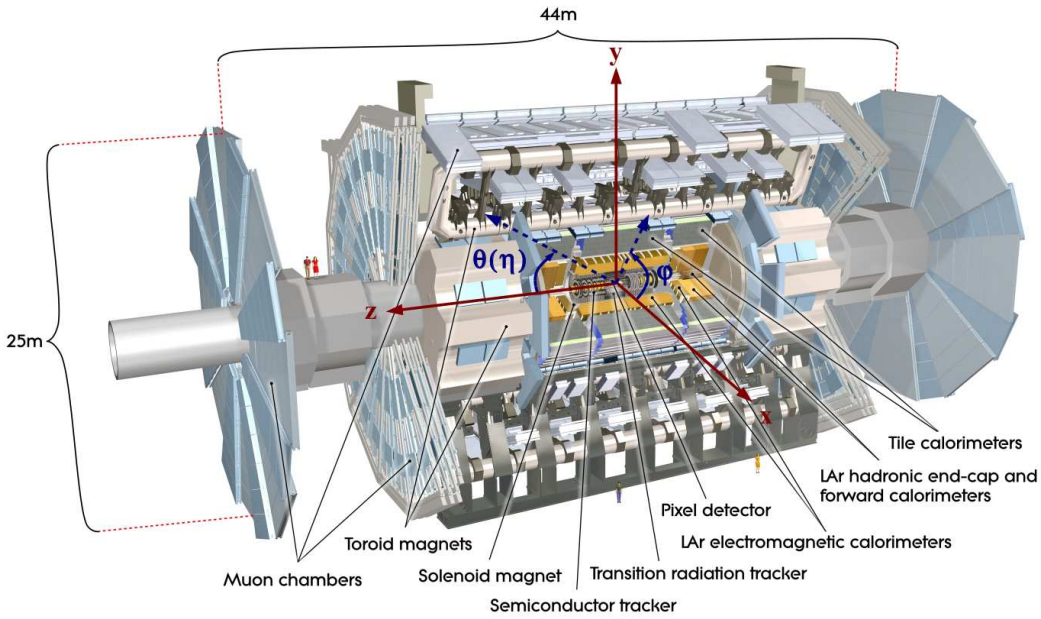


Figure 3.4 – Longitudinal cut-away view of the ATLAS detector, showing the different layers around the LHC beam axis. The collisions occur in the centre of detector. The main detector components are indicated.

part of the detector consists of the Muon Spectrometer (MS), which is embedded in huge toroidal magnets that define the name and the characteristic look of the ATLAS detector. Muons traverse all detector components, while depositing only a small fraction of their energy. Their tracks and momenta are measured in the ID and again in the MS. Figure 3.5 shows an overview of the interactions of the physics objects with the ATLAS detector in the transverse plane.

3.2.1 The magnet system

The ATLAS superconducting magnet system is an arrangement of a central solenoid providing the inner tracking with magnetic field, surrounded by a system of three large air-core toroids generating the magnetic field for the muon spectrometer. The overall dimensions of the magnet system are 26 m in length and 20 m in diameter. The two end-cap toroids are inserted in the barrel toroid at each end and line up with the central one. They have a length of 5 m, an outer diameter of 10.7 m and an inner bore of 1.65 m. The central solenoid provides a central field of 2 T and is responsible for bending charged particle tracks near the interaction point to measure their momentum. The peak magnetic fields on the toroids are about 4 T, bending muons in a different plane than the solenoid field does to improve their momentum measurement.

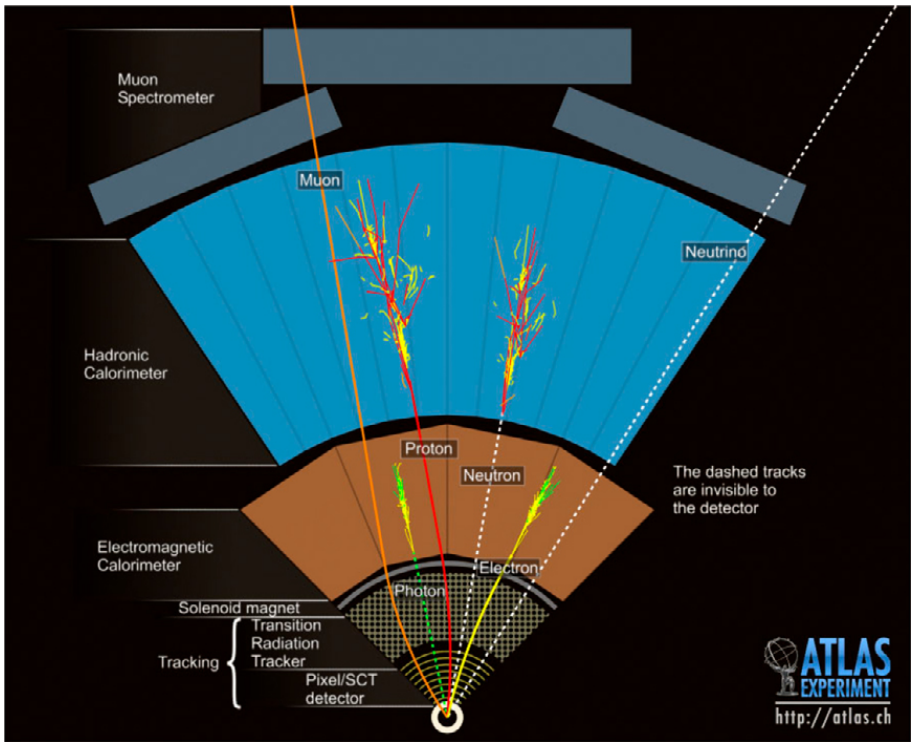


Figure 3.5 – Schematic illustration of the interactions of the physics objects with ATLAS detector in the transverse plane.

3.2.2 The inner detector

The Inner Detector (ID) is installed at a distance of a few cm from the beam pipe and consists of the Pixel Detector, the Semiconductor Tracker (SCT) and the Transition Radiation Tracker (TRT), which have barrel and endcap components as shown in Figure 3.6. The ID has full acceptance in ϕ, and pseudorapidity coverage up to $|\eta| = 2.5$. Its total length is about 6.2 m and its diameter is about 2.1 m.

The charged particles that traverse the ID ionise the detector material, allowing the reconstruction of the particle track as well as primary and secondary vertices, for example from b- or τ-decays products. The momentum resolution of the ID is measured to be $\sigma_{p_T}/p_T = (4.83 \pm 0.16) \times 10^{-4}\ \mathrm{GeV}^{-1} \times p_T$ [131].

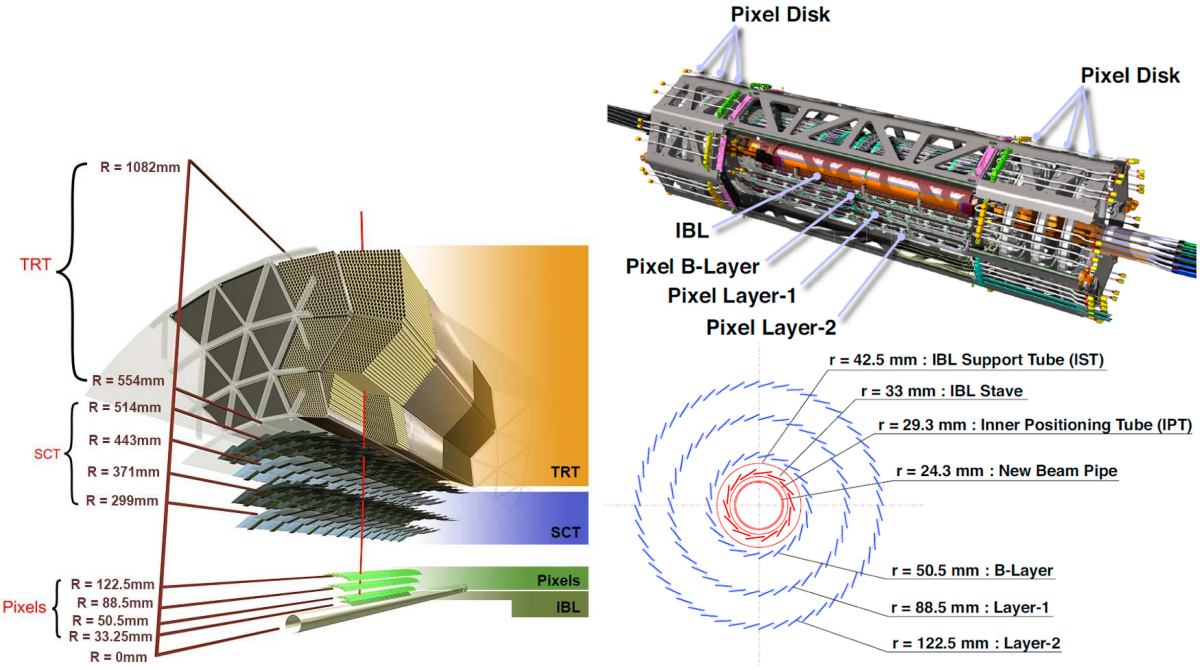


Figure 3.6 – The Inner Detector composed of the Pixel Detector, the Semiconductor Tracker and the Transition Radiation tracker. Left: The barrel and endcap components of the ID. Right: Radial distance of the ID component from the beam pipe.

The Pixel and Semiconductor Tracker detectors

The pixel detector system is composed of four concentric layers of silicon pixels, providing multiple high-resolution hits for charged particles traveling through. The closest layer to the beampipe is the IBL and was added during LS1 before the start of Run 2 operations. At only 3.3 cm away from the beamline, this additional tracking layer greatly improved the efficacy of tracking algorithms by an additional hit for tracks passing through multiple layers, and allowing for better resolution of secondary decay vertices, such as those originating from b-quarks. It provides a resolution of 8 μm in the barrel and 40 μm in the end-caps.

The pixels sensors are semiconductors made of silicon and have a cell size of 50×400 μm^2 (50×250 μm^2 for IBL). The sensors between the FE chips are longer (50×600 μm^2). Bump-bonding is used for the connection to the read-out electronics. 80.4 million read-out channels ($\sim$ 100 million for ATLAS in total) are installed for the pixel detector alone. Since the track density that close to the beam pipe is high, good spatial resolution (10 μm in $R-\phi$ and 115 μm in R) is neccessary to distinguish the different tracks. Three space-points (four with IBL) are provided for the measurement of each track, which enable a precise reconstruction of the vertices and the determination of the impact parameters.

The semiconductor tracker (SCT) consists of four double layers of silicon strips in the barrel (2,112 modules) and nine disks for each endcap (988 modules each). Each strip is 12 cm long and made of two sensors that are connected back-to-back to an other pair of sensors (double-layer). These two layers are rotated by a 40 mrad stereo angle. The strips have a pitch of 80 μm and has 6.3 million read-out channels. It covers a region of up to $|\eta| < 2.5$. The resolution in the barrel is 17 μm in $R-\phi$ direction and 580 μm in z-direction (endcap: 17 μm in $R-\phi$ direction and 580 μm in R-direction).

Transition Radiation Tracker

The transition radiation tracker (TRT) is the outermost ID subsystem, it is installed at a distance between $55-110$ cm from the beam pipe and covers a region of $|\eta| < 2.0$. The barrel section is divided into two parts that meet at $\eta = 0$. In each part, straws of a length of 0.7 m are used. The straws have a diameter of 4 mm and contain a tungsten wire in the middle. The tungsten wire serves as anode while the outer part of the tube serves as cathode. The straws are proportional chambers that are filled with a mixture of 70% xenon, 27% carbon dioxide and 3% oxigen. The barrel contains 52,544 straws, while each endcap (consisting of 18 wheels) contains 122,880, this leads to a total of 351,000 read-out channels. The TRT provides on average 36 hits per track, significantly contributing to the momentum measurement, with the lower spacial "hit" precision being compensated by the large number of space-points and longer track length. It is also used to distinguish electrons from pions based on the different amount of transition radiation that they emit when traversing material with different dielectric constants, due to the presence of propylene fibers and foils interleaved with the straws tubes. The resolution amounts to 130 μm in $R-\phi$ direction.

3.2.3 Calorimeters

The calorimeters are designed to measure the energy of traversing particles and the position of the resulting showers within a region of $|\eta| < 4.9$. They measure particles that mainly interact electromagnetically, producing electromagnetic showers, such as electrons and photons, hadrons which produce hadronic showers, with larger coverage compared to the ID, an important feature for an optimal reconstruction of the missing transverse energy defined in Section 4.6. An overview of the calorimeter system is given in Figure 3.7. ATLAS uses sampling calorimeters for the electromagnetic and the hadronic calorimeters. A sampling calorimeter consists of alternating layers of active and passive detector ma-

terial. The particles interact with the passive material such as lead and induce a shower of secondary particles. The active material is ionised by the created particle shower and used for the signal read-out.

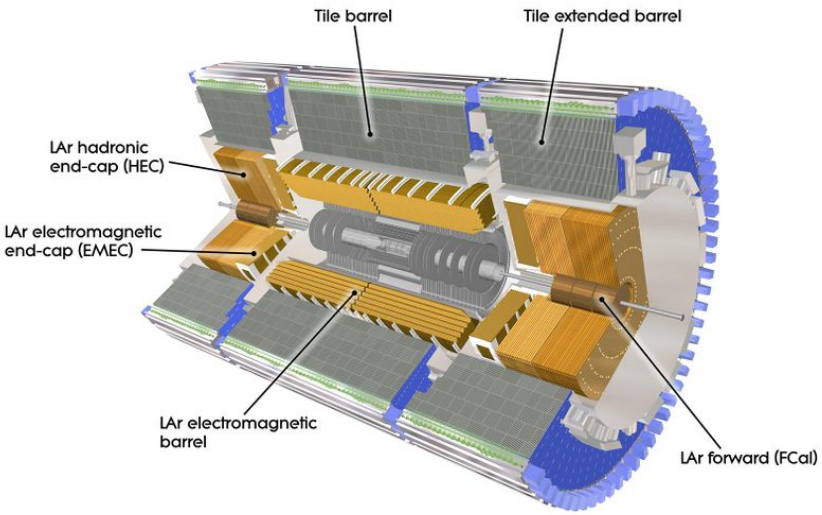


Figure 3.7 – The calorimeter system with central and forward calorimeters.

A good spatial resolution is achieved by segmenting the calorimeter in small cells, whose deposits of energy can be individually read out. The energy resolution of a calorimeter can be parameterised as:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c, \tag{3.4}$$

where a is the stochastic term, b the noise term and c a constant term that accounts for non-uniformities and miscalibrations. The expected energy resolution and η coverage of the ATLAS calorimeters are summarized in Table 3.1.

The innermost subsystem is the electromagnetic calorimeter (Liquid Argon Calorimeter), optimised to measure electromagnetic showers. The outermost calorimeter is the hadronic calorimeter (Tile Calorimeter), optimised to measure hadronic showers.

The depth is an important design parameter since the calorimeters must provide good containment for both electromagnetic and hadronic showers and limit punch through to the muon detectors. For the electromagnetic calorimeter, it is measured in terms of the radiation length X_0 which is defined as the average distance after which the energy of a traversing particle is reduced by a factor of $1/e$ due to interactions with the detector. For the hadronic calorimeter, it is measured in terms of the hadronic interaction length λ_I, defined as the average distance which the energy of a traversing particle is reduced by a $1/e$ factor due to electromagnetic and strong interactions with the detector material.

Another important characteristic of the ATLAS hadronic calorimeter is that it is non-compensating, meaning that it has a lower response to hadronic energy deposition compared to the electromagnetic one. This mainly happens due to the fact that hadronic

interaction with the calorimeter may result in invisible energy which is not detected (such as slow nuclear fragments).

3.2.3.1 The electromagnetic calorimeter

The electromagnetic (EM) calorimeter is a sampling calorimeter using liquid argon (LAr) as active medium and lead as passive material. Is has an accordeon-shape structure with Kapton electrodes that provide full coverage in ϕ. A presampling calorimeter is installed between the electromagnetic calorimeter and the ID up to $|\eta| < 1.8$ to estimate energy loss in front of the EM calorimeter due to interaction with the solenoid and the support structure of the detector, but also with the beampipe and the ID.

The barrel part is divided into two parts with a gap of 4 mm between them. Each part has a length of 3.2 m and covers the range of $|\eta| < 1.475$.

It has a minimal depth of 22 X_0 at $\eta = 0$ and a maximum depth of 33 X_0 for larger η. The two endcaps cover the region of $1.385 < |\eta| < 3.2$ (with a minimum depth of 24 X_0 at $\eta = 1.475$ and a maximum of 38 X_0 at $\eta = 2.5$) and consist of an inner and an outer wheel. The calorimeter has a high granularity with $\Delta\eta \times \Delta\phi = 0.025$ in the barrel part and a bit coarser granularity in the endcaps. This leads to a total amount of 226,176 read-out channels for the barrel and the two endcaps.

The design energy resolution of the electromagnetic calorimeter is $\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 0.7\%$ [125].

3.2.3.2 The hadronic calorimeter

The hadronic calorimeter consists of three parts: the tile calorimeter, the liquid-argon hadronic end-cap calorimeter (HEC) and the liquid-argon forward calorimeter (FCal).

Hadronic jets originate from quarks and gluons, and their energy is measured both by the electromagnetic and the hadronic calorimeter, while they leave the largest part of their energy in the latter.

The hadronic calorimeter consists of a tile calorimeter in the barrel ($|\eta| < 1.0$) and two extended barrels ($0.8 < |\eta| < 1.7$) as well as LAr calorimeters in the endcap and forward calorimeters. The tile calorimeter is a sampling calorimeter using steel as the absorber and scintillating tiles as the active material, located outside the EM calorimeter envelope with a central barrel and two extended barrels, in the region of $|\eta| < 1.7$. It is segmented in depth in three layers, approximately 1.5, 4.1, and 1.8 interaction lengths (λ_I) thick for the central barrel and 1.5, 2.6, and 3.3 λ_I for the extended barrel. The total detector thickness at the outer edge of the tile-instrumented region is 9.7 λ_I at $\eta = 0$. Two sides of the scintillating tiles are read out by wavelength shifting fibers into two separate photomultiplier tubes. The hadronic calorimeter has lower granularity compared with respect to the electromagnetic calorimeter, it is of $\Delta\eta \times \Delta\phi = 0.1$ for the Tile calorimeter and the LAr end-cap for $1.5 < |\eta| < 2.5$. The design energy resolution of the tile calorimeter is $\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\%$ [125].

The hadronic endcap (HEC, $1.5 < |\eta| < 3.2$) and the forward calorimeter (FCal, $3.1 < |\eta| < 4.9$) are overlapping. The HEC is a sampling calorimeter using copper as passive and LAr as active material. Since the forward calorimeter has to cope with a high particle flux, a radiation-hard construction consisting of one layer with copper/LAr and two layers

of tungsten/LAr has been chosen. The FCals have a thickness of about 10 λ_I and the design energy resolution is $\frac{\sigma_E}{E} = \frac{100\%}{\sqrt{E}} \oplus 10\%$ [125].

3.2.4 The muon spectrometer

At the LHC energies, the muons are effectively minimum ionising particles, therefore they cross the ID and calorimeters without being stopped and losing a relatively small amount of energy. This allows to identify muons by reconstructing them in a tracking outside the calorimeters, the muon spectrometer (MS), located in the outermost part of ATLAS.

The MS is the largest subsystem of the ATLAS detector, designed to measure the muon momentum based on the magnetic deflection of muon tracks in the large superconducting air-core toroid magnets, and cover the muon measurement up to $|\eta| = 2.7$. It consists of three concentric layers at a radius of 5 m, 7.5 m and 10 m. For each endcap, four wheels are installed at $|z| = 7.4$ m, 10.8 m, 14 m and 21.5 m and cover a region of $1.0 < |\eta| < 2.7$. The MS does not cover the full solid angle due to a gap at $|\eta| < 0.05$ for support material. The detector is embedded in a magnetic field created by the toroid magnets in the barrel ($|\eta| < 1.0$) and the two endcap magnets ($1.4 < |\eta| < 2.7$). It is instrumented with separate trigger and high-precision tracking chambers. Monitoring Drift Tubes (MDT) in the barrel and MDTs and Cathode Strip Chambers (CSCs) in the forward region allow for a precise measurement of muons. To be able to distinguish between muons from different bunch crossings, Resistive Plate Chambers (barrel) and Thin Gap Chambers (endcap) provide fast trigger information.

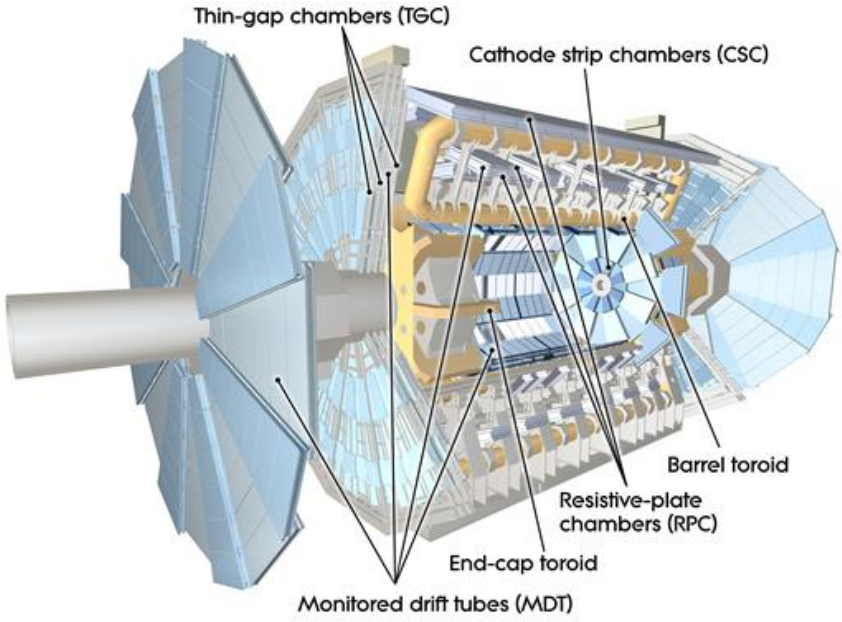


Figure 3.8 – Illustration of the ATLAS Muons subsystem [132].

Monitoring drift tubes (MDTs) provide the precision measurement of the track coordinates in the principal bending direction of the magnetic field. In the barrel region ($|\eta| < 1.3$), the MDTs are positioned in three concentric layers around the beam axis, at an approximate radius of 5, 8 and 10 m. To avoid holes in the acceptance, the chambers are partly overlapping. In the end-cap region ($1.0 < |\eta| < 2.7$), MDT chambers are assembled onto three wheels, inner, middle and outer layers positioned at z= 7.5, 14 and 22.5 m respectively.

The resistive plate chambers (RPCs) have a good timing resolution and are therefore used for triggering in the barrel part ($|\eta| < 1.05$) of the MS. Like the MDT chambers, the RPCs are positioned in three concentric layers around the beam axis. The two inner chambers are assembled together with the middle MDT chambers, and the outer layer is assembled on the outer MDT chambers. Due to the large lever arm between inner and outer RPCs, the trigger is able to select high momentum muons with thresholds ranging from 9 to 35 GeV. The inner RPCs deliver the low momentum trigger with thresholds from 6 to 9 GeV.

Cathode strip chambers (CSCs) are multi-wire proportional chambers that are only installed in the first wheel of the muon spectrometer which is installed at $|z| = 7$ m in the very forward region ($2.0 < |\eta| < 2.7$). Being placed in the forward region of the spectrometer, the detector has to be radiation hard and is filled with Ar/CO_2 (80/20%). To provide a good coverage in the forward region, each disk is made of eight large and eight small modules that overlap.

Thin gap chambers (TGCs) provide two functions in the end-cap MS ($1.05 < |\eta| < 2.4$, with trigger coverage $|\eta| < 2.7$): the muon trigger capability (with good time resolution and high rate) and the determination of the second, azimuthal coordinate to complement the measurement of the MDT's in the bending (radial) direction.

The design muon momentum resolution of the MS is $\frac{\sigma_{p_T}}{p_T[\text{ GeV}]} = 10\%$ at $p_T = 1$ TeV.

3.2.5 Resolution requirements

Requirements for the ATLAS detector system have been defined using a set of processes covering much of the new phenomena which one can hope to observe at the TeV scale [125]. These benchmark physics goals can be turned into a set of general requirements for the detector which are summarized in Table 3.1.

Detector component	Required resolution	η coverage	
		Measurement	Trigger
Tracking	$\sigma_{p_T}/p_T = 0.05\% \times p_T \oplus 1\%$	± 2.5	
EM calorimetry	$\sigma_E/E = 10\%/\sqrt{E} \oplus 30\%/E \oplus 0.7\%$	± 3.2	± 2.5
Hadronic calorimetry (jets)			
barrel and end-cap	$\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$	± 3.2	± 3.2
forward	$\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$	$3.1 < \|\eta\| < 4.9$	$3.1 < \|\eta\| < 4.9$
Muon spectrometer	$\sigma_{p_T}/p_T = 10\%$ at $p_T = 1$ TeV	± 2.7	± 2.4

Table 3.1 – General performance goals of the ATLAS detector [125]. Energies and momenta are given in GeV.

3.3 Trigger, acquisition and data taking

3.3.1 Trigger system

The ATLAS trigger system is responsible for deciding whether a given beam crossing is recorded or not. The selection at the trigger level must provide enough rejection to reduce the event recording rate according to the offline computing power and storage capacity, recording only "interesting events" coming from the 40 MHz rate LHC collisions. Figure 3.9 shows a diagram that illustrates the components of the Trigger and Data acquisition system (TDAQ).

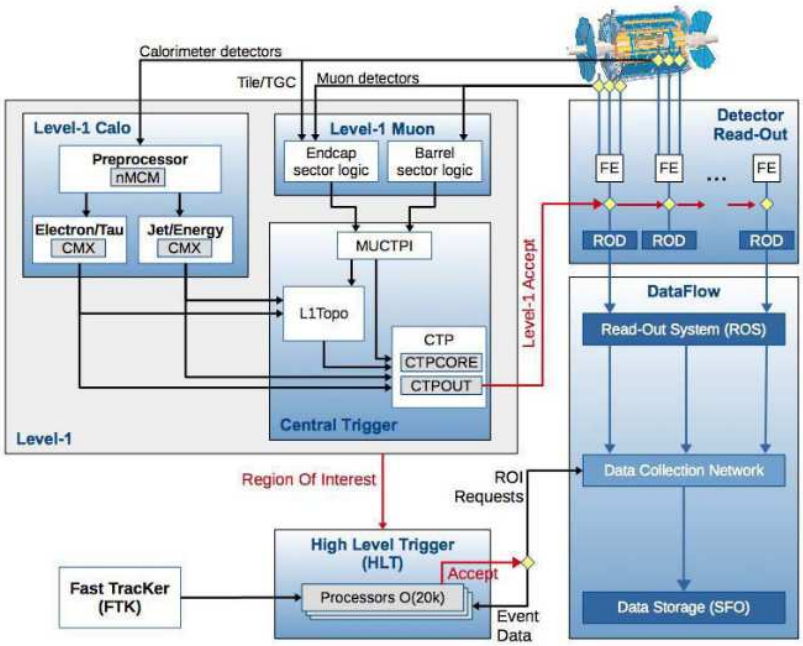


Figure 3.9 – The ATLAS TDAQ system in Run 2. FTK is being commissioned and is not used for the results of this work.

ATLAS uses a two-level trigger system, consisting of the hardware-based Level 1 trigger (L1), reducing the accepted rate to a maximum of 100 kHz, followed by a software-based High Level Trigger (HLT), which reduces the rate of events recorded to 1 kHz.

The Level-1 trigger performs a hardware-based online event selection using information of reduced granularity from all the calorimeters and the muon trigger system (RPCs and TGCs), to select events with high-E_T particles (electrons, taus, jets, muons) as well as high $E_{\mathrm{T}}^{\mathrm{miss}}$ at trigger level. So-called Regions of Interest (RoIs) are defined in the detector, then the central trigger processor (CPT) takes a decision within about 2.5 μs. and passed to the next trigger stage. Since events with large missing transverse energy (defined in Section 4.6) are interesting for the analyses presented in this work, this is an important trigger object that defines an RoI for the L1 trigger.

The event satisfying the RoI multiplicity requirements and/or the E_T thresholds are passed to the HLT trigger. The HLT trigger takes decisions based on complex algorithms,

using the full granularity and precision of the ATLAS detector (in association with the RoI defined at the Level-1 trigger). The trigger selections are optimized to minimize differences between the HLT and the offline analysis selections.

The chain of algorithms used to define a trigger selection is called "trigger chain", and its name follows this convention order:

[LEVEL]_[N]_[TYPE(S)]_[THRESHOLD]_[ISOLATION]_[QUALITY]

where "LEVEL" corresponds to the trigger level used (L1 or HLT), "N TYPE(S)" indicates the type and multiplicity of object candidates, "THRESHOLD" corresponds to the transverse momentum threshold for a trigger selection, "ISOLATION" indicates the isolation working point implemented and "QUALITY" indicates the rigour of requirement in the algorithms.

For example, HLT_e60_medium trigger chain selects events with a single electron of medium quality and $p_T > 60$ GeV using the HLT algorithm, without any isolation requirement.

In the case of the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger chains, the HLT_xe110_mht_L1XE50 trigger for example, selects events with a seed requirement of $E_{\mathrm{T}}^{\mathrm{miss}}$ >50 GeV in the Level 1 trigger. "mht" refers to the algorithm that reconstructs an offline-like $E_{\mathrm{T}}^{\mathrm{miss}}$ using trigger jets as input rather than calorimeter cells or topoclusters, and it is reconstructed as the transverse momentum vectorial sum of all jets reconstructed from calorimeter topological clusters at the HLT with $p_{\mathrm{T}} > 7$ GeV, an then a threshold of this mht $E_{\mathrm{T}}^{\mathrm{miss}}$ >50 GeV is required. Figure 3.10 shows the trigger efficiency turn-on for this particular $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger chain, which starts to be fully efficient after $E_{\mathrm{T}}^{\mathrm{miss}}$ >200 GeV.

3.3.2 Data quality

With about 100 million electronic channels and an event rate of 10^5 Hz it is essential to monitor the status of the ATLAS hardware and determine the quality of the data being taken in an efficient manner. In the online environment, this information can flag the shifter to take action to prevent taking faulty data. In the offline environment, one can perform more complex checks of the data to determine the quality of the data for various physics groups. Data quality (DQ) status flags are determined by the DQ representatives from each sub-detector system, which are used by the combined performance and trigger groups to declare the data as good/flawed/bad for the different physics objects. These flags are set for each luminosity block (approximately 2 minutes long fraction of runs), where the most low-level flags are based on detector control conditions, flagging possible hardware and data-taking problems such as nominal voltages, temperature, humidity, etc. Finally, a "Good Run List" (GRL) are built using all the luminosity blocks for which the sub-detectors useful for a given analysis were operational.

As an example, the relative fraction of luminosity associated to data of good quality delivered by the various sub-detector of ATLAS during the 2017 data-taking is shown in Figure 3.11. The total fraction of luminosity which is good for physics analysis correspond to 93.6% of the total luminosity delivered by the LHC.

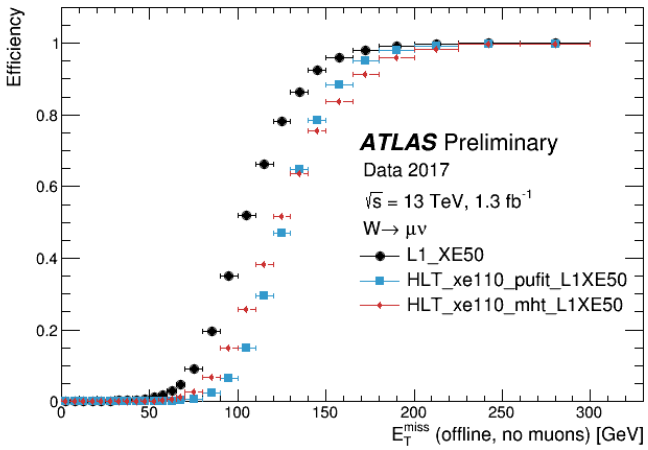
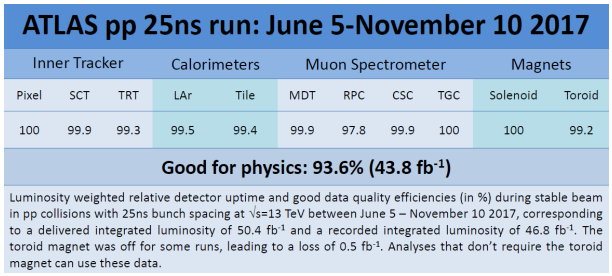


Figure 3.10 – The combined L1 and HLT efficiency of the missing transverse energy triggers HLT_xe110_pufit_L1XE50 and HLT_xe110_mht_L1XE50 as well as the efficiency of the corresponding L1 trigger (L1_XE50) are shown as a function of the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ (modified to count muons as invisible). The events shown are taken from data with a $W \to \mu\nu$ selection to provide a sample enriched in real $E_{\mathrm{T}}^{\mathrm{miss}}$. The HLT $E_{\mathrm{T}}^{\mathrm{miss}}$ of the "pufit" algorithm is calculated as the negative of the transverse momentum vector sum of all calorimeter topological clusters corrected for pileup. The pileup correction is done by grouping the clusters into coarser "towers" which are then marked as pileup if their E_T falls below a pileup-dependent threshold. A simultaneous fit to both classes of towers is performed, taking into account resolutions, making the assumption that the contribution of the pileup to $E_{\mathrm{T}}^{\mathrm{miss}}$ is zero. The fitted pileup E_{T} density is used to correct the above-threshold towers. The HLT $E_{\mathrm{T}}^{\mathrm{miss}}$ of the "mht" algorithm is calculated as the negative of the transverse momentum vector sum of all jets reconstructed by the anti-k_{t} jet finding algorithm from calorimeter topological clusters. These jets have pileup subtraction and JES calibration applied.



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.

Figure 3.11 – Performance of the ATLAS detector for the 2017 data-taking. Runs with a bunch spacing of 25 ns are taken between June and October 2017, corresponding to a recorded integrated luminosity of 43.8 fb^{-1} [133].

Chapter 4

Compound physics objects

3

In pp collisions, ATLAS sub-detectors record information of each event in terms of energy deposits in calorimeter cells and hits in the trackers. In order to convert this information into physics observables, advanced algorithms are used to reconstruct particle candidates in the event.

These physics objects are used by the physics analyses in order to select events with a given experimental signature. The mono-$h(b\bar{b})$ analysis described in Chapter 6 primarily uses jets, missing transverse momentum, and tracking for b-jet tagging. The performance analysis regarding the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ and $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ Significance in Chapter 5 use the combined performance from all the objects but for photons and τ-leptons.

An overview for the reconstruction and identification of these physics objects is given in the following. Section 4.1 introduces the fundamental objects for the physics observables reconstruction: charged tracks, vertices and clusters of energy deposits in the detector; Section 4.2 describes the hadronic jets; the succeeding Sections 4.3, 4.4, 4.5 describe the reconstruction of electrons, photons, muons, and tau-leptons, respectively. Finally, Section 4.6 focus on the missing transverse momentum, which plays an important role in the analysis described on Chapters 5 and 6.

4.1 Basic objects: charged tracks, vertices and energy deposits

Charged tracks

Tracks are reconstructed in the ID with a sequence of algorithms [134]. Track reconstruction starts with three hits in the silicon detectors and adds hits by moving away the interaction point using a combinatorial Kalman filtering algorithm [135]. Then, ambiguities in the tracks are resolved and the remaining tracks candidates are extrapolated into the TRT. Primary charged particles are reconstructed with this procedure, while the secondaries are reconstructed with a back-tracking method [134]. Tracks are required to satisfy the following quality criteria for identification [136]: they must have at least 7 hits in the SCT and pixel detectors, no more than one hit shared by multiple tracks in the pixel detector, no more than one missing hit in the pixel detector and no more than two

missing hits in the SCT detector. Tracks are also required to have $p_{\mathrm{T}} > 0.4$ GeV and $|\eta| < 2.5$ due to the ID coverage, see Section 3.2.2.

Interaction Vertices

An interaction vertex is the reconstructed location of an individual particle collision or decay, and is reconstructed from tracks using an iterative vertex finding algorithm [134]. Vertices are required to contain at least two tracks, and the primary vertex (PV) is selected as the one with the largest sum of squared transverse momenta of the tracks associated to it [137].

Clustering of calorimeter energy deposits

Incoming particles usually deposit energy in different calorimeter cells and these cells are grouped into clusters in order to provide inputs for particle reconstruction and identification. The total energies deposited within each cluster are then calibrated to different scales depending on the physics object type: the electromagnetic scale (EM) [1] [138] or the local cell weighting (LCW) [138]. Two types of clustering algorithms are used: one is the "sliding-window", which clusters calorimeter cells within a fixed-size rectangular window, positioned so that its contained transverse energy is a local maximum [139]. The fact that the cluster size is fixed allows for a very precise cluster energy calibration. It is used for reconstructing electron, photon, and tau lepton. The other one is the "topological" algorithm [138], which clusters together neighboring cells, as long as the signal in the cells is significant compared to noise. It is efficient at suppressing noise in clusters with large amount of cells, and it is used to reconstruct jets and the missing transverse momentum.

4.2 Hadronic Jets

Quarks and gluons can not be directly observed in a physics experiment due to the properties of the strong interaction: quarks and gluons fragment and hadronise quickly after production. Therefore, the observable, which is called *jet* [140], is a spray of collimated hadrons with their momenta aligned in the direction of the initial parton.

A jet is defined with an algorithm that groups its constituents into jets and computes the momentum of the resulting object. The constituents of the jets can be clusters of topologically connected calorimeter cell signals, "topo-clusters", or tracks in the ID. Various jet collections can be formed targeting different topologies and kinematics. In this work, jets are built using the anti-k_{t} algorithm [141], briefly summarized in Section 4.2.1, and the main differences among them are the radius parameter and the inputs of the algorithm: topo-clusters, ID tracks or Monte Carlo (MC) generator-level (truth-level) jets.

[1] This scale reconstructs the energy deposited by electrons and photons correctly but does not include any corrections for the loss of signal for hadrons due to the non-compensating character of the ATLAS calorimeters.

This section presents and describes the calorimeter jets which are reconstructed from noise-suppressed topological clusters in the calorimeter [142] in Section 4.2.2. Jets reconstructed from tracks are described in Section 4.2.4.

4.2.1 Jet reconstruction: the anti-k_t algorithm

The anti-k_t algorithm is the standard jet algorithm used in ATLAS. It has the advantage of being collinear- and infrared-safe, meaning that the structure of the output jet is independent of soft radiation and collinear splitting [141].

It is a sequential clustering algorithm which use the distance d_{ij} between entities i and j, and d_{iB} between entity i and the beam (B), defined as follows:

$$d_{ij} = \min\left(\frac{1}{\mathrm{p}_{\mathrm{T_i}}^2}, \frac{1}{\mathrm{p}_{\mathrm{T_j}}^2}\right) \frac{\Delta \mathrm{R}_{\mathrm{ij}}^2}{\mathrm{R}^2} \text{ and} \tag{4.1}$$

$$d_{iB} = \frac{1}{p_{\mathrm{T_i}}^2}, \tag{4.2}$$

where $\Delta R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$ and the radius parameter R controls the size of the cone for the resulting jet, which turns out to be circular in the $(\eta - \phi)$ plane.

First, the algorithm evaluates all the distance values for the input particles. If $d_{ij} < d_{iB}$, the objects i and j are combined and removed from the inputs and the combined object is added as a new input. If instead $d_{ij} > d_{iB}$, the object i is selected as a jet and removed from the input collection. At each step the distance parameters d_{ij} and d_{iB} are recomputed and the procedure repeats iteratively until no further combination is possible.

There are other sequential clustering jet algorithms such as the k_t [143] and the Cambridge/Aachen [144], however the anti-k_t algorithm was chosen as the standard jet reconstruction algorithm in ATLAS since it showed a better performance regarding jet reconstruction efficiency and purity and under pile-up, also has the advantage of a simpler underling event energy subtraction due to the circular $\eta - \phi$ plane shape for the reconstructed jets [145]. To perform the jet clustering for these algorithms, the FastJet [146] program is used.

4.2.2 Calorimeter jets

The shower of hadrons, originated from a quark or gluon, deposits energy in the calorimeter cells, and they are grouped in topological clusters ("topo-clusters") in order to suppress calorimeter noise by excluding cells which are unlikely to have energy coming from the actual jet. The basic observable controlling this cluster formation is the cell signal significance, defined as

$$\zeta = \frac{E_{\mathrm{cell}}}{\sigma_{\mathrm{cell}}}, \tag{4.3}$$

where E_{cell} is the cell signal and σ_{cell} the average expected noise (electronic and from pile-up) measured on the electromagnetic EM energy scale [138].

Topo-clusters are formed by a growing-volume algorithm starting from a calorimeter cell with a highly significant seed signal, $\zeta > t_{\text{seed}}$. Then, the cells neighbouring the seed and satisfying $\zeta > t_{\text{neighbor}}$ are collected into the corresponding cluster. The cluster grows until all direct neighbor cells on the outer perimeter, $\zeta > t_{\text{cell}}$, are included. The values of the threshold parameters are optimized to find clusters efficiently without being overwhelmed by noise. This formation is a sequence of seed and collect steps, which are repeated iteratively until all topologically connected cells passing the criteria are found. These topo-clusters can grow and merge together, therefore an additional splitting procedure is performed in order to separate nearby showers. All cells having a local energy maximum above a threshold are designated as seeds and then the clusters are split spatially between them [138].

The final clusters are used as inputs to the anti-k_{t} algorithm described in Section 4.2.1 to provide reconstructed jets. Two categories can be defined depending on the distance parameter used for the anti-k_{t} algorithm: small-R and large-R jets.

4.2.2.1 Small-R Jets

They are reconstructed from topo-clusters, calibrated to the electromagnetic scale [139] using the anti-k_{t} algorithm [147] if they pass a p_{T} threshold of 7 GeV. The distance parameter is $R = 0.4$ in Equation 4.1, which is the commonly used value in ATLAS to reconstruct gluon, light quark and b-tagged jets. This jets are reconstructed for $|\eta| < 4.5$, driven by the spatial extent of the detector at high $|\eta|$.

Jet calibration

After the clustering procedure outlined in Section 4.2.2, the energies of the constituent cells are summed to obtain the total energy of the topo-cluster. This energy is calibrated at the EM scale, which correctly estimates the energy of the electromagnetic showers, but does not reflect well the energy for hadronic showers. This is due to the different responses of the calorimeters to the electromagnetic and hadronic energy deposition.

The calibration of the reconstructed jets restores the jet energy scale to that of truth jets reconstructed at the particle-level energy scale. Figure 4.1 presents an overview of the calibration scheme for EM-scale calorimeter jets and the steps are described in the following [148]:

1. Origin correction
   This correction recalculates the four-momentum of jets in order to point to the hard-scatter PV instead of to the center of the detector, while keeping the jet energy constant.

2. Pile-up correction
   The excess energy due to pile-up is removed. It consists of two components: subtracting the pile-up p_{T} density integrated over the jet area [149], and a residual correction derived from the MC simulation [150].

3. MC-based calibration
   A re-scaling is applied to adjust the detector-level energy scale to that of the actual

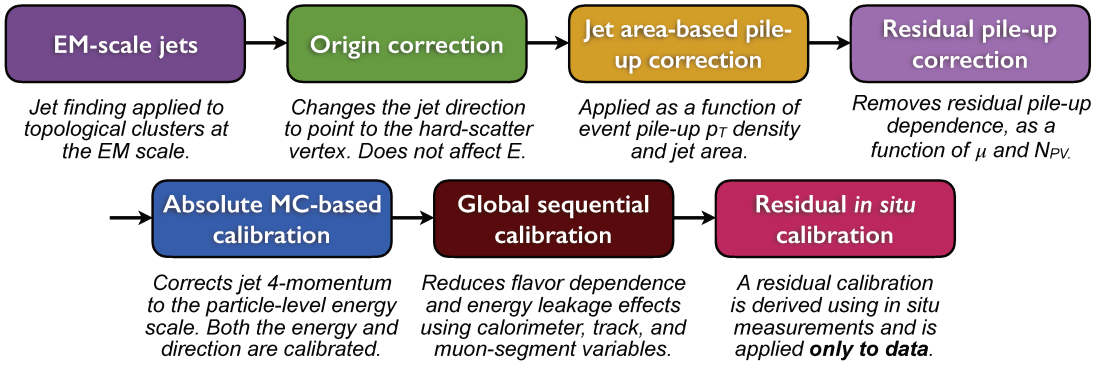


Figure 4.1 – Calibration stages for EM-scale jets. Other than the origin correction, each stage of the calibration is applied to the four-momentum of the jet [148].

particle-level energy scale. This jet energy scale (JES) factor is calculated from MC as a function of η(jet) and p_{T}(jet) [151].

4. Global sequential calibration
   Corrects sequentially the dependence of the jet response considering five observables related to the topology of calorimeter deposits and tracks associated with the jet as described in [148]. This does not affect the mean energy response, but improves the resolution of the JES against fluctuations in the jet particle composition and in the energy distribution within the jet.

5. Residual in-situ calibration
   A final data-driven correction is applied to cover residual differences between data and MC. Differences between data and MC are quantified by balancing the p_{T} of a jet against other well-measured reference objects, such as the Z boson (in $Z(\ell\ell)$+jets events) or γ (in γ + jets events) [148].

Jet selection

Pile-up suppression via JVT:
Pile-up jets can arise from hard QCD jets originating from a pile-up vertex, and can be tagged and rejected using the vertex-pointing information of charged-particle tracks. Also, pile-up jets can arise from local fluctuations that are a superposition of random combination of particles from multiple pile-up vertices. Tracking information plays an important role in providing information about the pile-up structure and vertex composition of jets within the tracking detector acceptance $|\eta| < 2.5$, therefore track-based observables can be used for discrimination of pile-up jets [150].

The jet vertex tagger (JVT) is a track-based tagging algorithm that helps to identify pile-up jets. It uses two variables with information about the primary vertex and tracks associated with the jet, which has a good separation power with respect to pile-up jets, and combine them in a two-dimensional likelihood [152].

Figure 4.2 shows the JVT discriminant, which is the result of a multivariate combination of the two track-based variables, for hard-scatter and pile-up jets with $20 <$

p_{T}/ GeV < 30. For values of the JVT close to zero, the reconstructed jet is likely to come from pile-up interactions, whereas for a JVT value close to one, the jet is more likely to be a hard scatter jet. A value of JVT$=-1$ is assigned to jets with no associated tracks.

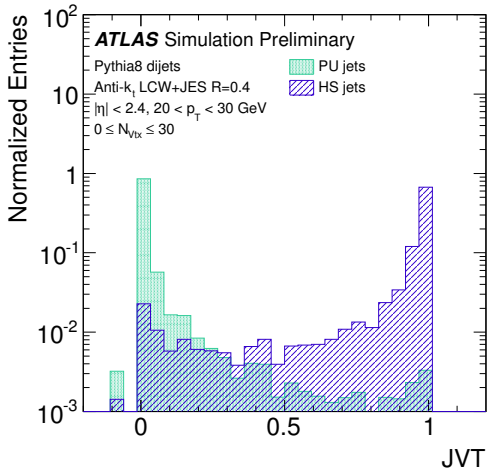


Figure 4.2 – Distribution of JVT for pileup and hard-scatter jets with $20 < p_{\mathrm{T}} < 30$ GeV [152].

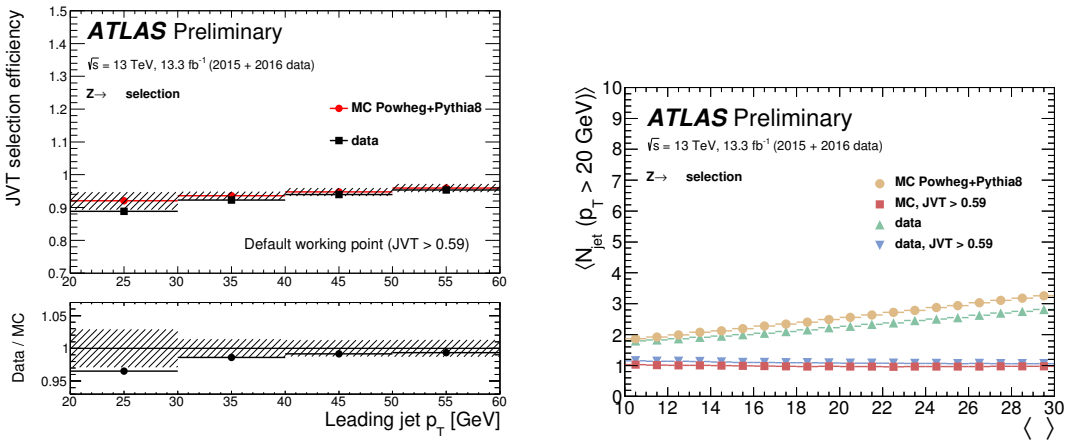
For the Run 2 data-taking period, three JVT working points are defined: *loose*, *medium* and *tight*, targeting different signal efficiencies [152]. This JVT score requirement is applied to jets with $20 < p_{\mathrm{T}}/GeV < 60$ and $|\eta| < 2.4$.

The medium working point is commonly used and it corresponds to a requirement of JVT> 0.59, which corresponds close to 95% efficiency for hard-scatter jets as it can be seen in Figure 4.3(a). Figure 4.3(b) shows that with the medium JVT working point, the multiplicity of jets looses it dependency on the average number of interactions per bunch crossing μ.

Jet cleaning:
Jets at high p_{T} produced in proton-proton collisions must be distinguished from misidentified jets of non-collision origin. The latter include fake jets from:

- Beam induced background due to proton losses upstream of the interaction point.
- Cosmic-ray showers produced in the atmosphere overlapping with collision events.
- Calorimeter noise from large scale coherent noise or isolated pathological cells.

Multiple jet cleaning criteria are defined in order to suppress these fake jets. These criteria are related to variables that quantify the quality of the signal pulse shape for the LAr calorimeter (calorimeter noise discriminating variables), or measure the fraction of energy deposited in the electromagnetic calorimeters or in the HEC to the total energy of the jet (energy ratio variables), or the fraction between the scalar sum of the jet tracks coming from the primary vertex to the total jet p_{T} (track-based variables) [153].



(a) (b)

Figure 4.3 – Medium JVT working point. 4.3(a) Hard-scatter jet selection efficiency, in Powheg+Pythia8 MC and in 2015+2016 data, of a JVT $>$ 0.59 cut on a jet balanced against a Z boson decaying to muons. The uncertainties shown are the statistical uncertainty summed in quadrature with the systematic uncertainty, evaluated varying the residual contamination from pileup jets by 20%. 4.3(b) The average number of jets with $p_{\mathrm{T}} > 20$ GeV in Powheg+Pythia8 MC and in 2015+2016 data before and after a cut of JVT $>$ 0.59, as a function of the average number of interactions per bunch crossing.

If a jet fails any of these cleaning requirement, it is designated as a "bad jet", and the event is completely removed from consideration. Two working points are defined: *BadLoose* and *BadTight*. The BadLoose selection is designed for high good jet efficiency (close to 100%), while maintining as high fake jet rejection as possible. While BadTight selection is designed to provide a much higher fake jet rejection with an inefficiency for good jets of up to a few percent.

4.2.2.2 Large-R Jets

When heavy particles are produced with a significant Lorentz-boost, their decay products can be so collimated that they can be reconstructed as one jet with a large-radius cone. For example, when sufficiently boosted, the decay products of vector bosons, top quarks, and Higgs bosons can become collimated. The angular separation of the decay products is approximately

$$\Delta R \approx \frac{2m}{p_{\mathrm{T}}}, \tag{4.4}$$

where $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$, and p_{T} and m are the transverse momentum and the mass, respectively, of the decaying particle.

Therefore, for higher p_{T} the ability to resolve the individual hadronic decay products using narrow-radius jets algorithms begins to degrade to the point that the standard reconstruction techniques fail. This happens when the separation of the quarks becomes smaller than the radius parameter of the jets, described in Section 4.2.2. Techniques are

designed to recover sensitivity in such cases considering large-radius jets with substructure in order to maximize efficiency [154].

- Jet Reconstruction

  These jets are built from topo-clusters, calibrated using the local cell signal weighting (LCW) method [138, 154]. This calibration method classifies topo-clusters along a continuous scale as being electromagnetic or hadronic [2], using shower shapes and energy densities. Energy corrections are applied to hadronic clusters that account for the effects of non-compensation [3], signal losses due to noise suppression due to clustering, and energy lost in non-instrumented regions of the calorimeters. The LCW clusters are the input to the jet algorithm, which is the anti-k_T algorithm [141] with a distance parameter of $R = 1.0$, see Section 4.2.2.

  Following this jet finding, the jets are "groomed" using the trimming procedure [155] to reduce the energy of clusters originating from initial state radiation, pile-up interactions or underlying event. The trimming algorithm takes advantage of the fact that contamination from pile-up, multiple parton interactions and initial-state radiation in the reconstructed jet is often much softer than the outgoing partons associated with the hard scatter and their final-state radiation [154]. This procedure starts with reclustering the constituents of the initial jet, using the k_T algorithm [143], into subjets of radius R_{sub} and then removing any subjet that has a transverse momentum that is less than f_{cut} times the transverse momentum of the parent jet. The values used in ATLAS for the two parameters are: $R_{\text{sub}} = 0.2$ and $f_{\text{cut}} = 0.05$ [156]. Figure 4.4 shows a scheme of the trimming procedure. The total four-vector of the jet is obtained by summing the four-vectors of the remaining constituents.

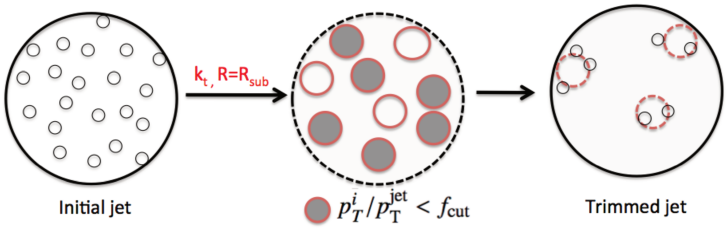


Figure 4.4 – Diagram depicting the jet trimming procedure [154].

- Jet Mass Calibration

[2] The main differences in the hadronic and electromagnetic calibration of topo-clusters are the magnitudes of the applied corrections, which in the EM case are significantly smaller than for hadronic.

[3] Electromagnetic calorimeter signal for hadrons is smaller than the one for electrons and photons depositing the same energy.

The large-R jet mass is a weighed linear combination of two jet mass observables: the calorimeter-based and the track-assisted jet mass.

The calorimeter jet mass is defined using the collection of topo-clusters, J, where they are assumed to be massless, as in Equation 4.5:

$$m^{\text{calo}} = \sqrt{\left(\sum_{i \in J} E_i\right)^2 - \left(\sum_{i \in J} \vec{p}_i\right)^2}, \tag{4.5}$$

with an MC-based calibration that corrects, on average, the reconstructed jet mass to the particle level. This correction is determined in the same way as the MC-based correction to the jet energy scale correction.

At a sufficiently high Lorentz-boost, the angular separation of the decay products can be comparable with the calorimeter granularity, which start to be limited by its angular resolution. A track-assisted mass can improve the performance at high jet p_{T} since the ID tracks have a better angular resolution than calorimeters. The track-assisted mass is the mass of the tracks, m^{track}, reconstructed by the inner detector and weighted by the ratio of the measured p_{T} of the calorimeter and the inner detector, it is given by:

$$m^{\text{TA}} = \frac{p_{\text{T}}^{\text{calo}}}{p_{\text{T}}^{\text{track}}} \cdot m^{\text{track}}. \tag{4.6}$$

The combined mass is a simple linear combination of the previous two mass definitions where the weight for a given mass definition is inversely proportional to the respective jet mass resolution, taken from MC simulation as in Equation 4.7.

$$m^{\text{comb}} = \left(\frac{\sigma_{\text{calo}}^{-2}}{\sigma_{\text{calo}}^{-2} + \sigma_{\text{TA}}^{-2}}\right) \cdot m^{\text{calo}} + \left(\frac{\sigma_{\text{TA}}^{-2}}{\sigma_{\text{calo}}^{-2} + \sigma_{\text{TA}}^{-2}}\right) \cdot m^{\text{TA}}, \tag{4.7}$$

where σ_{TA} and σ_{calo} are the expected resolution functions for the track-assisted and calorimeter-based jet mass, respectively.

This mass reconstruction is expected to have a better mass resolution and a reduction of the systematic uncertainties when using the combined mass instead of the calorimeter mass [157]. Figure 4.5 shows the jet mass resolution vs. the truth jet mass transverse momentum, for three different mass definitions.

4.2.3 *b*-jet tagging

Hadronic jets coming from bottom quarks are called b-jets, and the identification procedure of such jets is referred to as b-tagging. This identification can be done in both track or calorimeter small-R jets.

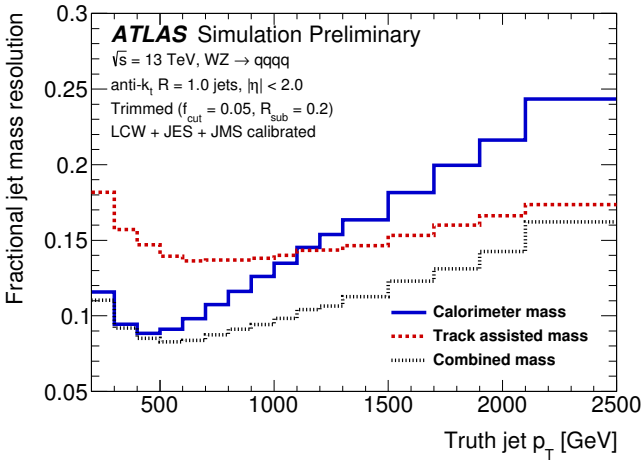


Figure 4.5 – The fractional jet mass resolution, vs. the truth jet mass transverse momentum, for three different mass definitions. The resolutions used as input to the combined mass definition are determined using a sample of jets produced via QCD dijet processes and generated with Pythia 8, where truth-jet matching is applied. The observed fractional resolutions in this figure are for jets from boosted W or Z bosons produced via $WZ \to qqqq$ processes generated using Pythia 8, with matching to particle-level W and Z.

Hadrons containing a b-quark have a long lifetime (~ 1.5 ps, $c\tau \sim 450\,\mu$m). Therefore, the typical b-hadron topology is characterized by at least one vertex displaced from the point where the hard-scatter collision takes place. Also, tracks generated from b-hadron decays tend to have large impact parameters [4] enabling their contribution to be separated from the contribution of tracks from the primary vertex [158, 159, 160].

The b-tagging algorithm uses the output variables from three algorithms, also called low level taggers [161, 162]:

- Impact Parameter based algorithm (IP3D): It uses a log-likelihood ratio discriminant of the impact parameter significances (d_0/σ_{d_0} and $z_0 \sin(\theta)/\sigma_{z_0 \sin(\theta)}$) probability distributions for b-, c- and light-flavour jet hypothesis.

- Secondary vertex finding algorithm (SV1): Reconstructs an inclusive displaced secondary vertex within the jet, corresponding to the decays of heavy hadrons (c or b). All displaced track pairs within the jet are tested for a two-track vertex hypothesis. Any two track vertices which are likely to be originated from the decay of a long-lived particle or photon conversion, are rejected. A new vertex is then fitted with all tracks from the accepted two-track vertices.

[4] The transverse impact parameter (d_0) is defined as the distance of closest approach in the $r-\phi$ plane of the track to the primary vertex while the longitudinal impact parameter ($|z_0 \sin\theta|$) is defined as the distance of the track to the primary vertex in the longitudinal plane at the point of closest approach in $r-\phi$.

- Decay chain multi-vertex reconstruction algorithm (JetFitter): Exploits the topological structure of weak b- and c-hadron decays inside the jet to reconstruct the full b-hadron decay chain. It tries to find a common line between the PV and the vertices of the b- and c- hadron decays, which corresponds to the b-hadron flight path [158].

The outputs of these b-tagging algorithms are combined in a boosted decision tree (BDT) multivariate analysis. The current BDT algorithm used in ATLAS is called the MV2c10 and its output is the discriminant variable used to identify jets containing a b-hadron [161]. The algorithm trains the MV2c10 with b-jets as signal, while the background sample is composed of 7% (93%) c- (light-flavour) jets in $t\bar{t}$ events using small-R jets [163], see Section 4.2.2.1. Figure 4.6 shows the MV2c10 BDT output, for which b-jets have high values, whereas c-jets have intermediate values and light-flavour jets have low values.

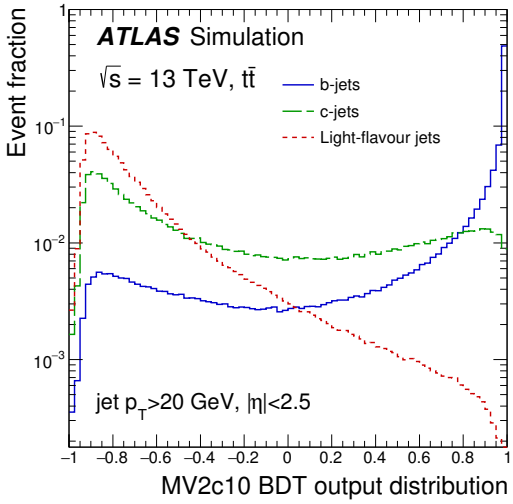


Figure 4.6 – MV2c10 BDT output for b- (solid blue), c- (dashed green) and light-flavour (dotted red) jets evaluated with $t\bar{t}$ events [163].

Different MV2c10 discriminant requirements provide specific b-jet signal efficiencies on a $t\bar{t}$ sample. A requirement of MV2c10> 0.8244 (MV2c10> 0.6459) gives a 70% (77%) b-jet efficiency with a 12 (6) c-jet and 381 (134) light-jet rejection [161].

The b-tagging efficiency and mistag rate are compared between data and MC simulation and the data-to-simulation ratios are applied as scale factors in simulated events with selected or vetoed b-jets [163].

4.2.4 Fixed radius track jets

The fat-jet with $R = 1.0$ radius parameter in the mono-$h(b\bar{b})$ analysis described in Chapter 7 is associated with one or two b-tagged anti-k_t jets reconstructed only from charged particle tracks, called *track-jets*. The radius parameter for these track jets is fixed to $R = 0.2$. The use of track-jets with a smaller R parameter allows Higgs bosons

with higher transverse momentum (p_{T}) to be better reconstructed. Firstly, this decreased radius, from $R = 0.4$ to $R = 0.2$, is preferable because it prevents the merging of the constituents within the jet to occur as the Higgs jet becomes more boosted. Secondly, the choice of using ID tracks is made because it both improves the granularity of the jet finding, thereby assisting with the aforementioned effect of merging in the calorimeter due to its poorer angular resolution, while also improving the identification of the flight direction of the b hadron [164]. These track jets are then uniquely matched to fat-jets in the event using the method of ghost association [147], which uniquely associates objects to parent objects by rerunning the entire jet finding procedure. Figure 4.7 shows a cartoon to illustrate the small-R track jet reconstruction.

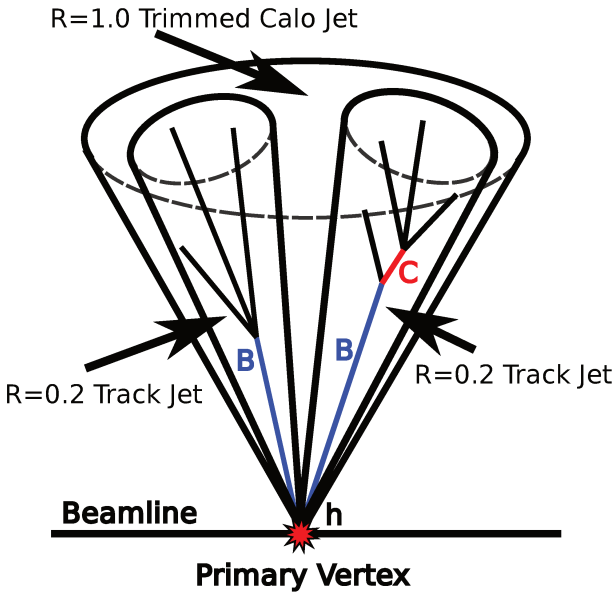


Figure 4.7 – Cartoon illustrating subjet reconstruction using fixed radius R=0.2 track jets [165].

4.3 Electrons and Photons

Electron and photon candidates have very similar signatures in the electromagnetic calorimeter, however they can be distinguished depending on the activity in the inner detector [166].

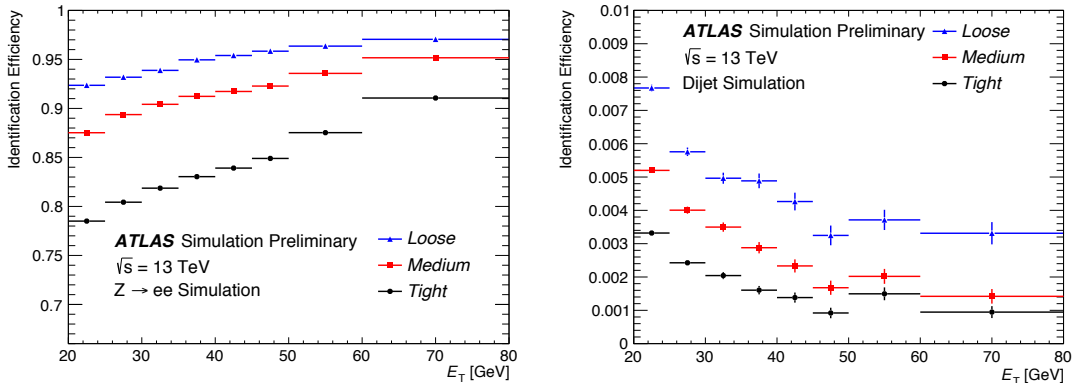
The reconstruction of electrons and photons starts from the energy deposits in the electromagnetic calorimeters. These energy deposits are grouped in clusters seeded by the sliding window algorithm [167]. Then, the track information is used to classify the cluster as coming from electrons or photons.

Electron reconstruction requires the association of an inner detector track with a cluster of energy deposits in the electromagnetic calorimeter. The reconstruction of a photon candidate depends if the photon converts into an electron-positron pair before the electromagnetic calorimeter (converted) or if it arrives undisturbed into the calorimeter

(unconverted). Therefore, its reconstruction may involve the matching of an electromagnetic cluster to two or zero ID tracks compatible to be produced in a displaced conversion vertex in the ID material. Some late photon conversions may occur in the space between the tracker and calorimeter thus could only be tagged using information or the electromagnetic presampler. The photon and electron reconstruction algorithm then proceeds in parallel and a final algorithm decides if the resulting particle candidate is an electron, or an unconverted or a converted photon [168, 169].

Identification algorithms are applied to discriminate signal-like electron from background-like electrons such as hadronic jets or converted photons. The baseline identification algorithm is performed using a likelihood-based (LH) method using variables characterizing the tracks and shower shapes, and with the help of a multivariate analysis simultaneously evaluates several properties of the electron candidates when making a selection decision [168]. This method defines several efficiency working points with increasing background rejection: *loose*, *medium* and *tight*, optimized in regions of $|\eta|$ and E_{T}. These operating points are inclusive subsets: the electrons selected by the Medium operating point are all selected by the Loose operating point, and the Tight electrons are all selected by the Medium operating point. Figure 4.8 shows the performance of the LH identification algorithm, for which the efficiencies for electron candidate identification at the three operating points increases with E_{T}. The efficiencies in MC simulations are compared to the one measured in data with tag-and-probe methods using $Z \to ee$ and $J/\psi \to ee$ decays and the extracted data to MC ratios, as a function of E_{T} and η, are applied as scale factors at analysis level.

For photons, the MC identification efficiencies are cross-checked in data with $Z \to e\gamma$ radiative decays at low momentum.



(a) (b)

Figure 4.8 – Efficiency to identify electrons from $Z \to ee$ decays 4.8(a) and the efficiency to identify hadrons as electrons estimated using simulated dijet samples 4.8(b). The efficiencies are obtained using MC simulations, and are measured with respect to reconstructed electrons [168].

Also, electron candidates are required to fulfil isolation criteria in order to further discriminate prompt electrons (mainly coming from W and Z decays) from other not

isolated electrons, like electrons originating from photon conversions, heavy flavor hadron decays and light hadrons misidentified as electrons.

The isolation variables quantify the energy of the particles produced around the electron [168], and they are:

- Calorimeter isolation energy, $E_{\mathrm{T}}^{\mathrm{cone0.2}}(E_{\mathrm{T}},\eta)$, defined as the sum of E_{T} of topological clusters, calibrated at the electromagnetic scale, within a cone of $\Delta R = 0.2$ around the candidate electron cluster, excluding the detector energy.

- Track isolation, $p_{\mathrm{T}}^{\mathrm{cone0.2}}(E_{\mathrm{T}},\eta)$, defined as the sum of the transverse momenta of all tracks, satisfying quality requirements and originating from the primary vertex within a $\Delta R = 0.2$ cone around the electron candidate direction and excluding the electron associated track.

A variety of selection requirements on the quantities $E_{\mathrm{T}}^{\mathrm{cone0.2}}(E_{\mathrm{T}},\eta)$ and $p_{\mathrm{T}}^{\mathrm{cone0.2}}(E_{\mathrm{T}},\eta)$ are defined to select isolated electron candidates [168]. For example, "LooseTrackOnly" working point targets a constant isolation efficiency of 99% varying $p_{\mathrm{T}}^{\mathrm{cone0.2}}(E_{\mathrm{T}},\eta)$, and "FixedCutHighPtCaloOnly" working point targets a constant upper threshold on $E_{\mathrm{T}}^{\mathrm{cone0.2}}(E_{\mathrm{T}},\eta)$<3.5.

Similar isolation requirements are applied also to photon candidates in order to better discriminate real photons from pions in jets. Additional discrimination between pions (leaving two collimated photons) and photons is coming from the electromagnetic calorimeter shower shape variables (essentially transverse direction shower distribution) exploiting the LAr calorimeter fine segmentation.

4.4 Muons

Muons leaves signatures in the ID and MS. Muon reconstruction is first performed independently in the ID and MS. Then, algorithms combine the sub-detectors information to form muon candidates in the region $|\eta| < 2.7$. In ATLAS, there are four muon reconstruction algorithms using ID track and/or muon spectrometer track and deposited energy in the calorimeter [170]:

- Combined (CB): Independent tracks in the ID and MS are globally refitted to form a combined muon.

- Standalone (SA): Reconstructed from a track in the MS if compatible with the originating interaction point and with at least two layers of MS.

- SegmentTagged (ST): The reconstructed muon track in the ID is associated with at last one local track segment in the CSC or MDT chambers of the MS.

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

Muon identification is applied to discriminate signal-like muons from background-like muons such as muons from in-flight hadron decays. Four muon identification working point are provided: *Loose*, *Medium*, *Tight* and *High-*p_{T} [170]. The default selection is the medium identification criteria and its combined reconstruction and identification efficiencies as measured in $Z \to \mu\mu$ and $J/\psi \to \mu\mu$ decays are shown in Figure 4.9 as a function of p_{T}.

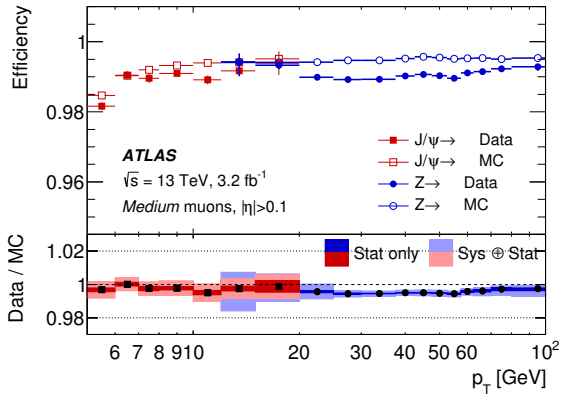


Figure 4.9 – Reconstruction efficiency for the Medium muon selection as a function of the p_{T} of the muon, in the region $0.1 < |\eta| < 2.5$. The error bars on the efficiencies indicate the statistical uncertainty. The panel at the bottom shows the ratio of the measured to predicted efficiencies, with statistical and systematic uncertainties [170].

Muon isolation, that measure the detector activity around a muon candidate, helps to further discriminate prompt muons from muons from semileptonic decays. As is done for electrons, combinations of track- and calorimeter-based variables are used to define several working points [170]. Also, the isolation variables used for muons are the same than those introduced for electrons in Section 4.3.

A calibration of the muon momentum scale is applied to muon candidates in MC simulation in order to account for differences in the muon momentum scale and resolution between data and simulation. This calibration procedure is described in detail in Reference [170].

4.5 Tau

Tau leptons decay either leptonically ($\tau \to \ell\nu_\ell\nu_\tau$) or hadronically ($\tau_{\mathrm{had}} \to$hadrons ν_τ) typically before reaching ATLAS active regions. The latter represent 65% of all decay modes, and the hadronic products contain one or three charged pions in 72% and 22% of the cases, respectively [171]. Leptonically decaying taus are not explicitly selected since their decay products are included if reconstructed as isolated leptons. Only hadronic tau decays are considered for reconstructing tau objects.

Tau candidates are seeded with small-R jets, described in Section 4.2.2.1. They are then required to pass the baseline selections of p_{T} >20 GeV, $|\eta|$<2.5 (excluding the transition region between barrel and forward calorimeters $1.37 < \eta < 1.52$). Also, they are required to have 1 or 3 matching charged tracks within a cone around the jet axis of $R = 0.2$ (1-prong or 3-prong τ candidates) [172, 173].

To reject fake τ-leptons from quark- and gluon-initiated jets, a multivariate discriminator is built with boosted decision trees that combine both tracking and calorimeter based information.

Figure 4.10 shows the τ_{had} reconstruction and identification efficiency defined as the fraction of 1-prong (3-prong) hadronic tau decays that are reconstructed as 1-track (3-track) tau candidates passing a given BDT criterion. Three working points provided are Loose, Medium and Tight, corresponding to different identification efficiency values, as shown in Figure 4.10.

Dedicated energy corrections are applied to calibrate the hadronically decaying τ-lepton candidates at the τ-lepton energy scale (TES) [174, 175] using the Local Cluster Weighting (LCW) calibration [172].

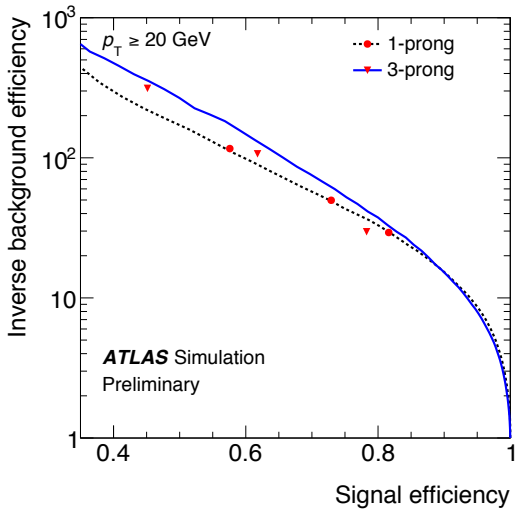


Figure 4.10 – Inverse of the efficiency for mis-tagging QCD jets as a function of the identification efficiency for τ_{had} candidates. The two lines refer to 1-track and 3-track candidates. The Loose, Medium and Tight working points are shown on these lines with decreasing signal efficiency. These working points are not exactly on the line because they implement variable requirements to achieve a reduced p_{T}-dependentcy of the efficiency [172].

4.6 Missing transverse momentum reconstruction

In pp collisions, the incoming protons can be seen as beams of partons, quarks or gluons. These partons come with a longitudinal momentum distribution but with a transverse momenta close to zero. In inelastic collisions, the hard scattering events are characterized by large amount of momentum transfers and, therefore, the final state particles can be produced at large angles with respect to the beam line.

Since the initial transverse momentum of the incident partons is close to zero, the momentum conservation implies that the vectorial sum in the transverse plane of the produced particles momenta should be negligible, meaning that all transverse momenta are balancing each other. An imbalance in the sum of visible transverse momenta is known as "missing transverse momentum" [176, 177], or $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$. Figure 4.11 shows the relationship between the missing transverse momentum to the physics objects in the transverse plane.

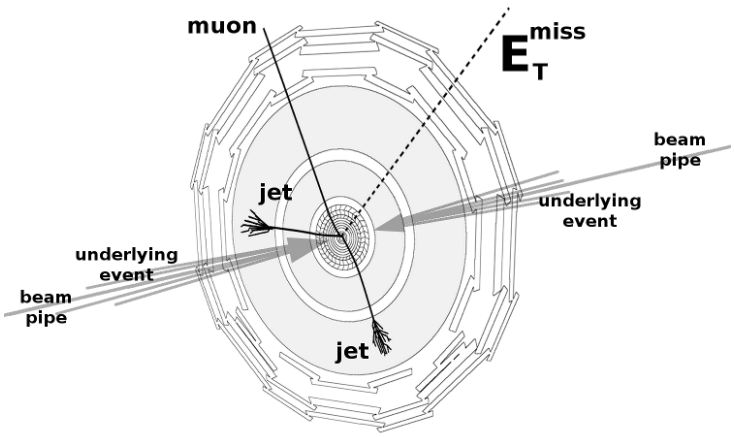


Figure 4.11 – Illustrative diagram of a hypothetical $W \to \mu\nu$ + jets event, for which the W boson candidate is balanced against jets. Taken from [178].

Two effects imply an imbalance in the total transverse momentum, i.e. $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}} \neq 0$:

- It may be the indicative of weakly- or non-interacting stable particles [5] in the final state. Within the Standard Model, these particles are the neutrinos. However, new physics beyond the Standard Model scenarios predict such particles, like dark matter candidates in the simplified models described in Section 2.2, or the neutralino or lightest stable particle in Supersymmetry models. Because of this, the study of the reconstruction and performance of the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ variable is critical in new physics searches. In this case, the reconstructed missing transverse momentum is referred to as real $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$.
- The $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ is a combined event-level quantity constructed using different signals from all sub-detectors. Therefore, it is susceptible to miscalibration, momentum mismeasurements and detector effects like dead regions. Also, it may be contaminated by additional momentum contributions coming from pile-up interactions. In this case, the reconstructed missing transverse momentum is referred to as fake $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$, and can serve as an important measure of the overall reconstruction performance.

[5] Stable particles are defined as those with an expected laboratory lifetime τ corresponding to $c\tau > 10$ mm.

If an imbalance in the sum of visible transverse momenta is present, momentum conservation implies that the total reconstructed $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ is equal to the negative vectorial sum of the visible transverse momenta, i.e:

$$\sum_{i\in\text{visible}} \boldsymbol{p}_{\mathrm{T}}{}^{i} + \boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}} = \mathbf{o} \implies \boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}} = -\sum_{i\in\text{visible}} \boldsymbol{p}_{\mathrm{T}}{}^{i}, \tag{4.8}$$

where $\boldsymbol{p}_{\mathrm{T}}{}^{i}$ represent each of the measured momentum in the transverse plane.

The reconstructed $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ in ATLAS [179, 180] is characterised by two main contributions. The first one comes from the *hard objects* comprising fully reconstructed and calibrated objects: muons, electrons, photons, τ-leptons, and jets. The second one is from the *soft term*, consisting of additional signals which are not associated with any of the reconstructed hard objects.

The $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ is calculated as the negative vectorial sum of selected and calibrated hard objects, and the soft term:

$$\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}} = -\Bigg(\underbrace{\sum_{i\in\text{muons}} \boldsymbol{p}_{\mathrm{T}}{}^{i} + \sum_{i\in\text{electrons}} \boldsymbol{p}_{\mathrm{T}}{}^{i} + \sum_{i\in\text{photons}} \boldsymbol{p}_{\mathrm{T}}{}^{i} + \sum_{i\in\text{hadronic }\tau} \boldsymbol{p}_{\mathrm{T}}{}^{i} + \sum_{i\in\text{jets}} \boldsymbol{p}_{\mathrm{T}}{}^{i}}_{\text{hard term}} + \underbrace{\sum_{i\in\text{Soft Term}} \boldsymbol{p}_{\mathrm{T}}{}^{i}}_{\text{soft term}} \Bigg) \tag{4.9}$$

where $\boldsymbol{p}_{\mathrm{T}}{}^{i}$ is the transverse momentum of each identified and calibrated object, described in more detail in the following Section 4.6.1.

To avoid double counting, clusters of topologically-connected calorimeter cells or topoclusters, see Section 4.1, are associated with the reconstructed hard objects in the following order: muons (μ), electrons (e), photons (γ), hadronically decaying τ-leptons, and finally jets [180].

The soft term is reconstructed from detector signals not associated with any hard object passing the selection cuts. These can be ID tracks, from which the track-based soft term (TST) is built, or topoclusters, from which the calorimeter-based soft term (CST) is built. For details on the criteria of the soft term definition, refer to Section 4.6.1.2.

From the vectorial components $E_{x,y}^{\mathrm{miss}}$, the magnitude, $|\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}| = E_{\mathrm{T}}^{\mathrm{miss}}$, and the azimuthal angle, ϕ^{miss}, of the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ can be calculated as:

$$E_{\mathrm{T}}^{\mathrm{miss}} = \sqrt{\left(E_x^{\mathrm{miss}}\right)^2 + \left(E_y^{\mathrm{miss}}\right)^2}, \tag{4.10}$$

$$\phi^{\mathrm{miss}} = \tan^{-1}\left(\frac{E_y^{\mathrm{miss}}}{E_x^{\mathrm{miss}}}\right) \tag{4.11}$$

An important quantity to estimate the event activity is $\sum E_{\mathrm{T}}$, which is defined as the scalar sum of the transverse momenta of the hard objects and soft term contributions to

the $E_{\mathrm{T}}^{\mathrm{miss}}$:

$$\sum E_{\mathrm{T}} = \sum p_{\mathrm{T}}^{\mu} + \sum p_{\mathrm{T}}^{e} + \sum p_{\mathrm{T}}^{\gamma} + \sum p_{\mathrm{T}}^{\tau} + \sum p_{\mathrm{T}}^{\mathrm{jets}} + \sum p_{\mathrm{T}}^{\mathrm{soft}}, \tag{4.12}$$

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

$$HT = \sum p_{\mathrm{T}}^{\mu} + \sum p_{\mathrm{T}}^{e} + \sum p_{\mathrm{T}}^{\gamma} + \sum p_{\mathrm{T}}^{\tau} + \sum p_{\mathrm{T}}^{\mathrm{jets}} \tag{4.13}$$

Both $\sum E_{\mathrm{T}}$ and HT provide a useful overall scale for evaluating the hardness of the hard-scatter event in the transverse plane, also they are important observables for understanding the resolution of $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$.

4.6.1 Physics Objects Selection

In order to correctly reconstruct $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$, all physics objects must be reconstructed, and correctly identified and calibrated to their corresponding scales. Also, since achieving a good momentum balance or closure of the event is an important goal, the bias from low-p_{T} particles not passing the reconstruction threshold is reduced by adding to the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ calculation transverse energy measured by tracks or clusters as will be addressed in Section 4.6.1.2.

This section describes the selections that are usually applied to the reconstructed objects which are considered for the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ calculation. This same selection criteria is considered in Chapter 5. Generally, the object selections for electrons, muons, tau-jets and photons require refinements to achieve an optimal $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction performance in the context of a given physics analysis, as the criteria used in the mono-$h(b\bar{b})$ analysis described in Section 6.2.

4.6.1.1 Hard Objects

All the hard objects have its own dedicated calibration, translating detector signals into a fully corrected four-momentum, and a subset of these reconstructed objects are selected using quality and isolation criteria, as introduced in Sections 4.2-4.5.

Due to the event selection used in the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ performance studies, photons and hadronically decaying τ-leptons have almost no impact on the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ calculation, but they are documented in this Section to give a comprehensive description.

- Electrons
  The electrons considered in the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction pass the reconstruction quality requirements described in Section 4.3 with "Loose" isolation and "Medium" likelihood identification criteria. In addition, electrons must have $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 1.37$ or $1.52 < |\eta| < 2.47$, to avoid the transition region between the barrel and end-cap electromagnetic calorimeters.

- Muons
  Muon candidates are reconstructed primarily by matching tracks in the ID and the MS [181], i.e. *Combined* muons. In detector regions outside the ID coverage ($2.5 < |\eta| < 2.7$) the muon momentum is measured from the MS tracks alone, i.e. *Standalone* muons. Muons are selected with "Medium" identification quality and "Loose" isolation criteria. See Section 4.4 for details on muon reconstruction, identification and isolation. The muons included in the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction are required to have $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.7$.

- Jets
  Jets are reconstructed as small-R jets described in Section 4.2.2.1. After the calibration, they are required to have $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 4.5$. To further reduce the effects of pile-up, jets with $p_{\mathrm{T}} < 60\,\mathrm{GeV}$ and $|\eta| < 2.4$ are required to have JVT > 0.59. Finally, Jets must pass "LooseBad" cleaning criteria. For details on jet reconstruction, calibration, cleaning and pile-up suppression via JVT discriminant, refer to Section 4.2.2.1.

- Photons
  Photons are selected and calibrated using the tight selection criteria [182]. In addition to the reconstruction quality requirements, photons are normally selected with $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 1.37$ or $1.52 < |\eta| < 2.47$ in order to be included in the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction.

- Hadronically decaying τ-leptons
  Hadronically decaying τ-leptons are reconstructed as described in Section 4.5 with $|\eta| < 1.37$ or $1.52 < |\eta| < 2.47$ and $p_{\mathrm{T}} > 20$ GeV. Also, candidates must pass the medium quality selection [174].

4.6.1.2 Soft Terms reconstruction

The soft term introduced in Equation 4.9 is a significant contribution to the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction, in particular in events with low jet activity. It is comprised of all detector signals not matched to the reconstructed objects defined previously and may contain contributions from the hard scatter as well as the underlying event and pile-up interactions. Two main algorithms have been developed to reconstruct the soft term: CST and TST [183].

Calorimeter-based soft-term (CST)

The calorimeter-based soft term is reconstructed from energy deposits in calorimeter cells, grouped into topo-clusters and calibrated at the LCW scale, which are not associated to any of the hard objects described in Section 4.6.1.1. This reconstruction has the advantage to take into account the contribution from charged and neutral particles.

However, since there is no association of these topo-clusters to the hardest vertex due to the small accuracy of the calorimeter cells pointing, all deposits are included. In consequence, pile-up contributions are also included in the final $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ calculation. CST was the standard soft term reconstruction algorithm in most ATLAS Run 1 analyses.

Track-based soft-term (TST)

The track-based soft term is reconstructed from ID tracks that are not associated with the hard objects [179]. Only those tracks associated with the hard scatter primary vertex are included. They are required to have:

- $p_{\mathrm{T}} > 400$ MeV;
- $|d_0|/\sigma(d_0) < 2$ and $|z_0 \sin(\theta)| < 3.0$ mm, where d_0 ($z_0 \sin(\theta)$) is the transverse (longitudinal) impact parameter with respect to the primary vertex.
- $\Delta\phi(\mathrm{track}, e/\gamma) > 0.05$ or $\Delta\eta(\mathrm{track}, e/\gamma) > 0.2$;
- $\Delta R(\mathrm{track}, \tau) > 0.2$.

Mismeasured tracks are removed following the strategy described in Reference [179].

These criteria provide excellent track-to-vertex matching for the TST. By not including the neutral contributions and associating tracks to the PV, the track-based reconstruction is more pile-up resilient than CST. Also, given the poorer angular resolution provided by the calorimeters compared to the tracking system, the energy resolution provided by TST is considerably improved, see Section 4.6.3 for details on the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ resolution. Because of this, TST is the standard algorithm for reconstructing the soft term contribution to the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ in ATLAS Run 2 analyses.

4.6.2 Track missing transverse momentum $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$

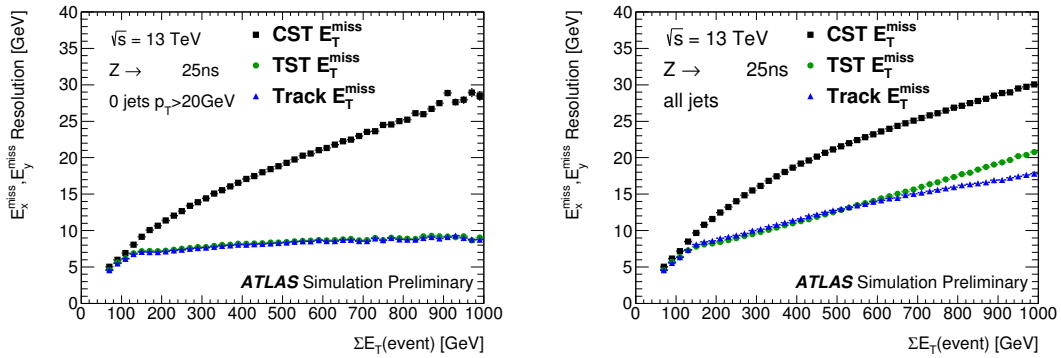
The Track $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$, or $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$, is reconstructed entirely from ID tracks alone, reducing the pile-up contamination and dependence. It is calculated by taking the negative vectorial sum of the $\mathbf{p}_{\mathrm{T}}$ of tracks satisfying the track quality criteria described in Section 4.1. Due to interactions within the ID, the p_{T} of the electron is more precisely measured with the calorimeter, therefore the p_{T} of an electron track is replaced by the calorimeter cluster measurement [177].

Even though $p_{\mathrm{T}}^{\mathrm{miss}}$ gives a pile-up robust $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ estimation, it is insensitive to neutral particles, which do not form tracks in the ID. Also, the η coverage of the $p_{\mathrm{T}}^{\mathrm{miss}}$ is limited to the ID acceptance of $|\eta| < 2.5$, substantially smaller than the calorimeter coverage which extends to $|\eta| = 4.9$

4.6.3 $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ resolution

The vectorial components, $E_{x,y}^{\mathrm{miss}}$, of the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ are expected to behave approximately as Gaussian distributions for $Z \to \ell\ell$ events [176], with deviations coming from noise at large $\sum E_{\mathrm{T}}$ values. Given that these distributions have non-Gaussian tails, the root-mean-square (RMS) is used to estimate the resolution, in order to include information of the $E_{\mathrm{T}}^{\mathrm{miss}}$ tails. For processes with real $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$, such us $ZZ \to \ell\ell\nu\nu$, the true $E_{x,y}^{\mathrm{miss}}$ are subtracted from the reconstructed quantity in simulation.

Figure 4.12 shows the resolution comparison between $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ calculated with TST or CST and Track missing transverse momentum $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$, as a function of $\sum E_{\mathrm{T}}$. The TST reconstruction shows a better resolution at high values of $\sum E_{\mathrm{T}}$, which is related to high event activity, with respect to the CST prescription for the soft terms.



(a) (b)

Figure 4.12 – Comparison of the performance of $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ built from TST and CST, and the $p_{\mathrm{T}}^{\mathrm{miss}}$, as quantified by the resolution, as a function of the CST $\sum E_{\mathrm{T}}$. Powheg+Pythia $Z \to \ell\ell$ simulation is shown [183].

Figure 4.13 shows the resolution of the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ reconstructed with TST or CST and the the track-only-based variant of the missing transverse momentum $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$ as a function of the pile-up activity measured in terms of the number of reconstructed vertices $\mathrm{N_{PV}}$.

More details on the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction and performance are given in Chapter 5.

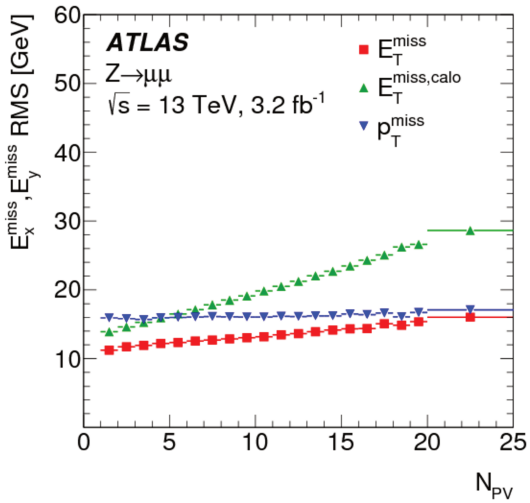


Figure 4.13 – Comparison of the resolution of $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ reconstructed with TST ($E_{\mathrm{T}}^{\mathrm{miss}}$) or with CST ($E_{\mathrm{T}}^{\mathrm{miss,calo}}$), and the track-only-based variant $p_{\mathrm{T}}^{\mathrm{miss}}$. Resolutions are calculated for an inclusive $Z \to \mu\mu$ + jets sample.

Chapter 5

Missing Transverse Momentum Significance

In collider experiments like ATLAS, conservation of momentum implies that the vectorial sum of the transverse momentum of all the collision by-products should be zero in the x-y plane transverse to the beam axis[1], as introduced in Section 4.6. Any imbalance is known as missing transverse momentum, $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$, and may be indicative of weakly interacting, stable particles in the final state, such as neutrinos. There are also prospects for such particles in theories beyond the Standard Model, such as dark matter particles or neutralinos, making the 2-vector $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ an important variable in searches for new physics. However, fake $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ can arise from interacting particles which escape the acceptance of the detector, are inaccurately reconstructed, or fail to be reconstructed altogether.

The degree to which the reconstructed $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ is consistent with momentum resolution and particle identification efficiencies can be identified by evaluating the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance, $\mathcal{S}$. On an event-by-event basis, $\mathcal{S}$ is determined from the log-likelihood ratio that the reconstructed $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ is consistent with the null hypothesis of having zero real $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$, given the full event composition. A high value of $\mathcal{S}$ is an indication that the observed $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ in the event cannot be explained from momentum resolution effects, suggesting that the event may contain undetected objects such as neutrinos or more exotic weakly interacting particles.

ATLAS and other experiments have previously defined an event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance $\mathcal{S}$ as:

$$\mathcal{S} = \frac{E_{\mathrm{T}}^{\mathrm{miss}}}{\sqrt{H_{\mathrm{T}}}} \text{ or } \mathcal{S} = \frac{E_{\mathrm{T}}^{\mathrm{miss}}}{\sqrt{\sum E_{\mathrm{T}}}}; \tag{5.1}$$

where $E_{\mathrm{T}}^{\mathrm{miss}}$ is the reconstructed magnitude of the missing transverse momentum, and $\sqrt{\sum E_{\mathrm{T}}}$ and $\sqrt{H_{\mathrm{T}}}$ are event-based approximations to the total $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution, see Section 4.6 for their definition. These definitions are based on the assumption that the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ is purely calculated using calorimeter signals. From Figure 4.12, it can be seen that

[1] ATLAS uses a right-handed coordinate system with origin at the nominal interaction point (IP) in the center of the detector and the z-axis along the beam pipe. The x-axis points from the IP to the center of the LHC ring, and the y-axis points upwards. Cylindrical coordinates (r, ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the beam pipe. The pseudorapidity is defined in terms of the polar angle θ as $\eta = -\ln\tan(\theta/2)$.

the resolution of $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ reconstructed with CST follows approximately a square root behaviour as a function of $\sum E_{\mathrm{T}}$, while for the TST $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ this assumption is not longer true. The use of measurements from all the ATLAS sub-detectors to improve the performance of the reconstructed objects requires to reconsider this approch. To go beyond these definitions, an object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance variable was developed, that is calculated event by event considering the expected resolutions and likelihood of mismeasurement of all the objects that enter the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction. This definition takes into account directional correlations between measurements. Similar variables used by the CDF and CMS collaborations are described in References [184] and [185, 186], respectively. Also, in some supersymmetry searches in ATLAS experiment, an object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance variable was previously developed by smearing jets according to the measured jet energy resolution [187]. In this Chapter, the performance of the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance in rejecting backgrounds with no real $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ is compared to that of the event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance defined as $E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{\sum E_{\mathrm{T}}}$, as well as to that of the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ itself.

The work presented in this Chapter have been published, and can be find in Reference [4]. This chapter is organised as follows. Section 5.1 describes the algorithm and definition of the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance, while Section 5.3 describes the samples and event selection used in the studies described in this Chapter. Section 5.4 presents the performance of the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance, evaluating its separation power between $Z \to ee$ and $ZZ \to ee\nu\nu$ events. Section 5.5 discusses methods used to estimate two sources of fluctuations in the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction: bias and soft term variance. The Chapter concludes with a summary and outlook in Section 5.6.

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

The $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance is defined to test the hypothesis that the total transverse momentum carried by invisible particles (named $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{inv}}$ in the following) is equal to zero against the hypothesis that $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{inv}}$ is different from zero. If $\mathcal{L}(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}|\boldsymbol{p}_{\mathrm{T}}^{\mathrm{inv}})$ is the likelihood function of the 2D parameter $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{inv}}$ for a given value of measured $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$, the log-likelihood ratio and a significance $\mathcal{S}$ are defined as:

$$\mathcal{S}^2 = 2 \ln \left(\frac{\max_{\boldsymbol{p}_{\mathrm{T}}^{\mathrm{inv}} \neq \mathbf{0}} \mathcal{L}(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}|\boldsymbol{p}_{\mathrm{T}}^{\mathrm{inv}})}{\max_{\boldsymbol{p}_{\mathrm{T}}^{\mathrm{inv}} = \mathbf{0}} \mathcal{L}(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}|\boldsymbol{p}_{\mathrm{T}}^{\mathrm{inv}})} \right) . \tag{5.2}$$

where numerator and denominator are calculated maximising on the parameters of the likelihood with the constraints $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{inv}} \neq \mathbf{0}$ and $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{inv}} = \mathbf{0}$, respectively. The likelihood function depends on the multiplicities, types, and kinematics of the objects measured in each event. As described in Section 4.6, the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ is defined as a vectorial sum over all the reconstructed hard objects plus the soft term. An event-by-event likelihood function is calculated assuming that:

- the measurement of each reconstructed object i is independent from others;

- for each object the probability distribution of measuring $\boldsymbol{p}_{\mathrm{T}}{}^{i}$ given the true transverse momentum value associated with the hard physics process $\pi_{\mathrm{T}}{}^{i}$ has the form $f(\boldsymbol{p}_{\mathrm{T}}{}^{i} - \pi_{\mathrm{T}}{}^{i})$, and in particular it has a Gaussian probability distribution with covariance matrix $\mathbf{V}^{i}$;

- thanks to the conservation of the momentum in the x-y plane, $\sum_i \pi_{\mathrm{T}}{}^{i} = -\boldsymbol{p}_{\mathrm{T}}^{\mathrm{inv}}$.

Under these assumptions, the likelihood function has the form of a two dimensional Gaussian:

$$\mathcal{L}(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}|\boldsymbol{p}_{\mathrm{T}}^{\mathrm{inv}}) \propto \exp\left[-1/2\left(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}} - \boldsymbol{p}_{\mathrm{T}}^{\mathrm{inv}}\right)^{\intercal}\left(\sum_i \mathbf{V}^{i}\right)^{-1}\left(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}} - \boldsymbol{p}_{\mathrm{T}}^{\mathrm{inv}}\right)\right] \tag{5.3}$$

and the log-likelihood ratio becomes a chi square variable, χ^2, with two degrees of freedom:

$$\mathcal{S}^2 = 2\ln\left(\frac{\mathcal{L}(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}|\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}})}{\mathcal{L}(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}|\mathbf{o})}\right) = \left(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}\right)^{\intercal}\left(\sum_i \mathbf{V}^{i}\right)^{-1}\left(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}\right), \tag{5.4}$$

where the sub-index i indicates each reconstructed objects that enters the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ calculation and $\mathbf{V}^{i}$ is the corresponding covariance matrix.

This χ^2 definition of the missing transverse momentum significance $\mathcal{S}$ in Equation 5.4 stresses the dependencies between the reconstructed $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ and the resolution effects propagated in the total covariance matrix $\sum_i \mathbf{V}^{i}$. If $\mathcal{S}$ has small values it is more likely that the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ is fake, or consistent with resolution effects, whereas a large value of $\mathcal{S}$ indicates that it is likely to be real $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$.

For each of the hard objects defined in Section 4.6.1, the covariance matrix, $\mathbf{V}_i$, can be defined in a natural coordinate system having one axis aligned with the measured $\boldsymbol{p}_{\mathrm{T}}{}^{i}$ of each object, with $\boldsymbol{p}_{\mathrm{T}}{}^{i} \equiv \left(p_{\mathrm{T}}{}^{i}\cos\phi^{i}, p_{\mathrm{T}}{}^{i}\sin\phi^{i}\right)^{\intercal}$:

$$\mathbf{V}^{i} = \begin{pmatrix} \sigma^2_{p_{\mathrm{T}}{}^{i}} & 0 \\ 0 & p_{\mathrm{T}}{}^{i\,2}\sigma^2_{\phi^{i}} \end{pmatrix}, \tag{5.5}$$

where the measurements of the transverse energy $p_{\mathrm{T}}{}^{i}$ and the azimuthal angle ϕ^{i} are considered uncorrelated.

An additional covariance matrix of signals associated to the soft term defined in Section 5.2, $\mathbf{V}^{\mathrm{soft}}$, is also considered. It is defined as

$$\mathbf{V}^{\mathrm{soft}} = \begin{pmatrix} \sigma^2_{\mathrm{soft}} & 0 \\ 0 & \sigma^2_{\mathrm{soft}} \end{pmatrix}, \tag{5.6}$$

without correlation between parallel and transverse measurements.

The summation of the covariance matrices is performed in the standard x-y coordinate system of the ATLAS detector, which corresponds to the transverse plane. Therefore, the total covariance matrix can be written considering the matrix rotation in x-y for each of

the objects as:

$$\mathbf{V}_{xy} = \sum_i R^{-1}\left(\phi^i\right) \mathbf{V}^i R\left(\phi^i\right) + \mathbf{V}^{\text{soft}} = \begin{pmatrix} \sigma_x^2 & \sigma_{xy}^2 \\ \sigma_{xy}^2 & \sigma_y^2 \end{pmatrix}, \tag{5.7}$$

where the index i runs on each of the objects considered in the $\boldsymbol{E}_{\text{T}}^{\text{miss}}$ calculation and $R\left(\phi^i\right)$ is the two dimensional rotation matrix in the azimuthal direction for each object.

Finally, the $\boldsymbol{E}_{\text{T}}^{\text{miss}}$ significance definition in Equation 5.4 can be written in a synthetic form in which the x-y coordinate system is rotated parallel (longitudinal L) and perpendicular (transverse T) to the direction of the total missing transverse momentum $\boldsymbol{E}_{\text{T}}^{\text{miss}}$:

$$\mathcal{S} = \left(E_{\text{T}}^{\text{miss}}, 0\right) \begin{pmatrix} \sigma_{\text{L}}^2 & \rho_{\text{LT}}\sigma_{\text{L}}\sigma_{\text{T}} \\ \rho_{\text{LT}}\sigma_{\text{L}}\sigma_{\text{T}} & \sigma_{\text{T}}^2 \end{pmatrix}^{-1} \begin{pmatrix} E_{\text{T}}^{\text{miss}} \\ 0 \end{pmatrix}, \tag{5.8}$$

where σ_{L}^2, σ_{T}^2 are the total variances in the longitudinal and transverse directions to the $\boldsymbol{E}_{\text{T}}^{\text{miss}}$ respectively, ρ_{LT} is the correlation factor of the longitudinal L and transverse T measurements, and the total covariance matrix is given by the rotation in the angle of the total reconstructed $\boldsymbol{E}_{\text{T}}^{\text{miss}}$, $\phi(E_{\text{T}}^{\text{miss}})$:

$$\mathbf{V}_{LT} = \begin{pmatrix} \sigma_{\text{L}}^2 & \rho_{\text{LT}}\,\sigma_{\text{L}}\,\sigma_{\text{T}} \\ \rho_{\text{LT}}\,\sigma_{\text{L}}\,\sigma_{\text{T}} & \sigma_{\text{T}}^2 \end{pmatrix} = R\left(\phi(E_{\text{T}}^{\text{miss}})\right) \mathbf{V}_{xy}\, R^{-1}\left(\phi(E_{\text{T}}^{\text{miss}})\right), \tag{5.9}$$

In this L-T basis, the $\boldsymbol{E}_{\text{T}}^{\text{miss}}$ significance in Equation (5.8) can be written as:

$$\mathcal{S}^2 = \frac{\left|\boldsymbol{E}_{\text{T}}^{\text{miss}}\right|^2}{\sigma_{\text{L}}^2\left(1-\rho_{\text{LT}}^2\right)}. \tag{5.10}$$

This form of the $\boldsymbol{E}_{\text{T}}^{\text{miss}}$ significance illustrates the essential meaning of a significance, in which the measured variable, $\boldsymbol{E}_{\text{T}}^{\text{miss}}$ is in the numerator, and the information of the variance is embedded in the denominator. Also, this definition has a close relationship with the previously used approximated definition of the event-based $\boldsymbol{E}_{\text{T}}^{\text{miss}}$ significance, where $\sqrt{\sum E_{\text{T}}}$ is an approximation of the $\boldsymbol{E}_{\text{T}}^{\text{miss}}$ resolution

Each of the three assumptions used to derive Equation 5.4 deserves a detailed study which could bring an improved definition of the $\boldsymbol{E}_{\text{T}}^{\text{miss}}$ significance. The last part of this note will discuss in more details how the performance of the $\boldsymbol{E}_{\text{T}}^{\text{miss}}$ significance changes by assuming $\boldsymbol{p}_{\text{T}}^{\text{inv}} = -\sum_i {\pi_{\text{T}}}^i - \boldsymbol{bias}$. The extra term $\boldsymbol{bias}$ is introduced to take into account systematic contributions from particles which interact with the detector, but do not enter the $\boldsymbol{E}_{\text{T}}^{\text{miss}}$ calculation. Possible contributions for this systematic effect are particles with a p_{T} lower than the threshold, neutral particles not considered in the track-based soft term reconstruction, or particles lost for any other reason (e.g. not reconstructed). In this case, re-deriving all the equations starting from Equation 5.2, $\mathcal{S}^2$ becomes :

$$\mathcal{S}^2 = \frac{\left|\boldsymbol{E}_{\text{T}}^{\text{miss}} - \boldsymbol{bias}\right|^2}{\sigma_{L'}^2\left(1-\rho_{L'T'}^2\right)}, \tag{5.11}$$

where the longitudinal L' and transverse T' components are defined with respect of the direction of the vector $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}} - \boldsymbol{bias}$, and the $\boldsymbol{bias}$ estimation is addressed in Section 5.5.

Finally, it is straightforward to extend the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance definition to test the hypothesis that $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{inv}}$ is equal to a certain vector $\boldsymbol{\Lambda}$ against the hypothesis that $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{inv}}$ is different from $\boldsymbol{\Lambda}$:

$$\mathcal{S}^2(\boldsymbol{\Lambda}) = 2\ln\left(\frac{\max_{\boldsymbol{p}_{\mathrm{T}}^{\mathrm{inv}}\neq\boldsymbol{\Lambda}}\mathcal{L}(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}|\boldsymbol{p}_{\mathrm{T}}^{\mathrm{inv}})}{\max_{\boldsymbol{p}_{\mathrm{T}}^{\mathrm{inv}}=\boldsymbol{\Lambda}}\mathcal{L}(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}|\boldsymbol{p}_{\mathrm{T}}^{\mathrm{inv}})}\right) = \frac{\left|\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}} - \boldsymbol{\Lambda} - \boldsymbol{bias}\right|^2}{\sigma_{L''}^2\left(1-\rho_{L''T''}^2\right)}. \tag{5.12}$$

in which the longitudinal and transverse components are defined with respect of the direction of the vector $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}} - \boldsymbol{\Lambda} - \boldsymbol{bias}$ and the presence of the term $\boldsymbol{bias}$ is kept explicit. A similar variable that considers a scale of energy was also defined in [188].

5.2 Object resolution

An important ingredient for the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance definition is the momentum resolution in the transverse plane for the objects entering the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ calculation. For each object, one can define a transverse momentum resolution parallel to the direction of the object, σ_{p_T}, and a transverse momentum resolution perpendicular to the direction of the object. The latter can be estimated as the angular resolution on the transverse plane σ_ϕ, multiplied by the p_{T} of the object, $p_{\mathrm{T}} \cdot \sigma_\phi$. The resolutions used to calculate the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance are parametrized using Monte Carlo (MC) simulations, which are well reproducing the resolutions measured on data.

The estimation of the electron and photon transverse momentum resolution depends on the η and p_{T} of the particles [182]. Typical resolutions for electron and photon with $p_{\mathrm{T}} = 100$ GeV in the central region are 1.6–1.9% in the parallel direction and 0.4% in the perpendicular direction. In the calculation of the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance, dedicated parametrizations are used for electrons, photons converting to electron-positron pairs, and photons not converting to electron-positron pairs.

The resolution of the transverse momentum of the muons depends on the momentum, detector region, and whether the muon is measured solely by the muon spectrometer, or by a combination of the inner detector and muon spectrometer [181]. A typical resolution for $p_{\mathrm{T}} = 100$ GeV is 2% in the parallel direction, and 0.1% in the perpendicular direction.

The transverse momentum resolution for hadronically decaying τ-leptons depends on the momentum and detector region. The typical resolutions in the parallel direction for hadronically decaying τ-leptons with $p_{\mathrm{T}} = 100$ GeV in the central region are 5.5-6.7%, depending on the τ decay mode. The typical resolution in the perpendicular direction is 1%.

The transverse momentum resolution for jets depends on the momentum and detector region [189, 148, 190, 191, 192]. The typical resolutions for a central jet with $p_{\mathrm{T}} = 20$ GeV are around 22% in the parallel direction, and 5–7% in the perpendicular direction. For a central jet with $p_{\mathrm{T}} = 100$ GeV, these resolutions decrease to 7% in the parallel direction and 1.1-1.6% in the perpendicular direction.

An additional source of $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ resolution comes from the additional jets produced by pile-up unsuccessfully suppressed by the jet vertex tagger requirement. The inclusion of these jets in the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ calculation may generate an imbalance, resulting in a fake $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$. Given the fact that the presence of these jets has a certain probability, which depends on the value of the JVT discriminant, one can estimate an additional term in the resolution for the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance calculation to be equal to the p_{T} of the jet, multiplied by the probability for that jet to have originated by pile-up. In the calculation of the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance, this source is treated in Gaussian approximation. This probability has been estimated in Monte Carlo simulations in two ranges of pseudo-rapidity, central ($|\eta| < 2.7$) and forward ($|\eta| \geq 2.7$).

For the central region, three ranges of JVT are used: for high values of JVT (JVT $>$ 0.59), the fraction of jets originating from pile-up interactions is negligible, while it becomes 40% for $0.05 \leq \mathrm{JVT} \leq 0.59$ and $p_{\mathrm{T}} < 100$ GeV, and 95% for JVT < 0.05 and $p_{\mathrm{T}} <$ 150 GeV. In the forward region, where no coverage from the inner detector is available, the fraction of jets originating from pile-up interactions is 20% for jets with 20 GeV $\leq p_{\mathrm{T}} <$ 30 GeV, and 7% for jets with 30 GeV $\leq p_{\mathrm{T}} <$ 40 GeV. The full scheme is reported in Figure 5.1. The values are derived per bin, and no smoothing or interpolation is used.

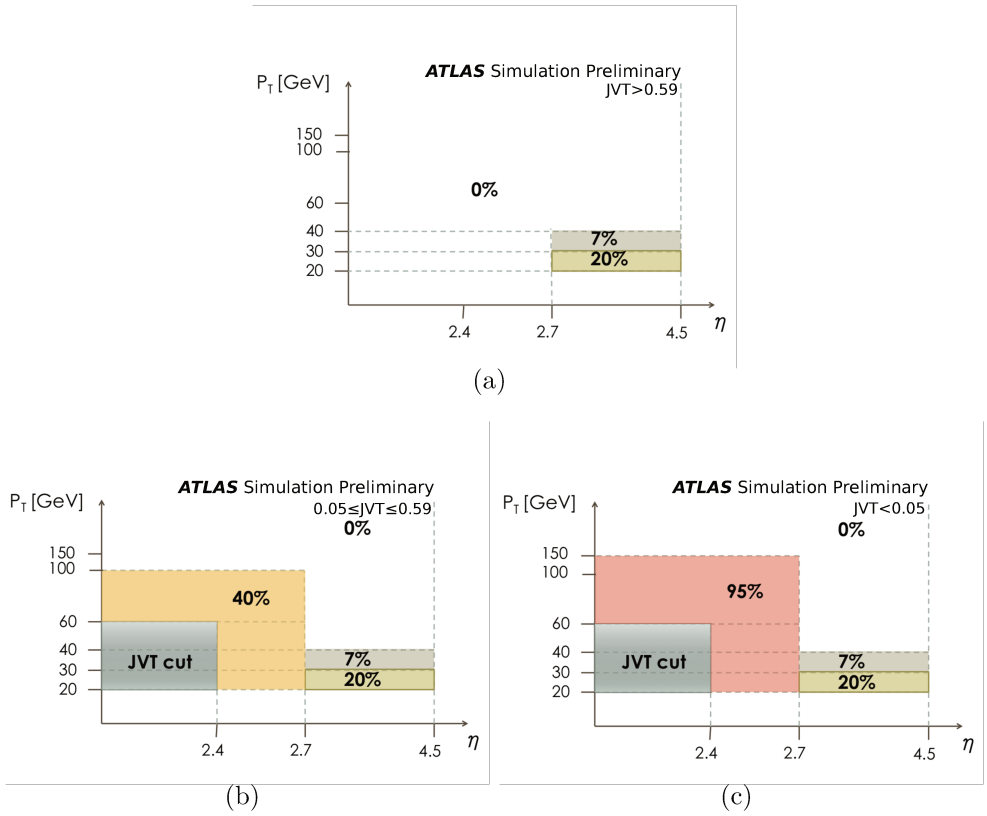


Figure 5.1 – Scheme of the regions in p_{T} and η of the jet with the corresponding percentage representing the fraction of jets originating from pile-up interactions. The regions shown are for (a) JVT > 0.59, (b) $0.05 \leq \mathrm{JVT} \leq 0.59$, and for (c) JVT < 0.05. Bin values are used as-is with no interpolation performed across boundaries.

A first estimation of the track-based soft term resolution can be extracted from the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in a $Z \to \mu\mu$ sample with a jet veto. Apart from sub-leading contributions from hadronic decays involving neutrinos (as in the semileptonic decays of heavy flavor hadrons), no real $E_{\mathrm{T}}^{\mathrm{miss}}$ is expected in this sample. The distribution of the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ projected onto the x- or y-axis resembles a Gaussian distribution, but with larger tails. The width of this distribution, estimated by the RMS, is mostly due to the soft term resolution, with a negligible contribution from the muon resolution. Therefore the resolution of the soft term is set equal to the RMS of the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ projected on the x- or y-axis in this final state with a jet veto. The distribution of this projection is shown in Figure 5.2, and the RMS extracted is 8.9 GeV. This value is used in the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance calculation as estimate of the soft term resolution, independently of the axis of projection. Figure B.1 in Appendix B.1 shows the distribution of this projection for both lepton channels, ee and $\mu\mu$ selections.

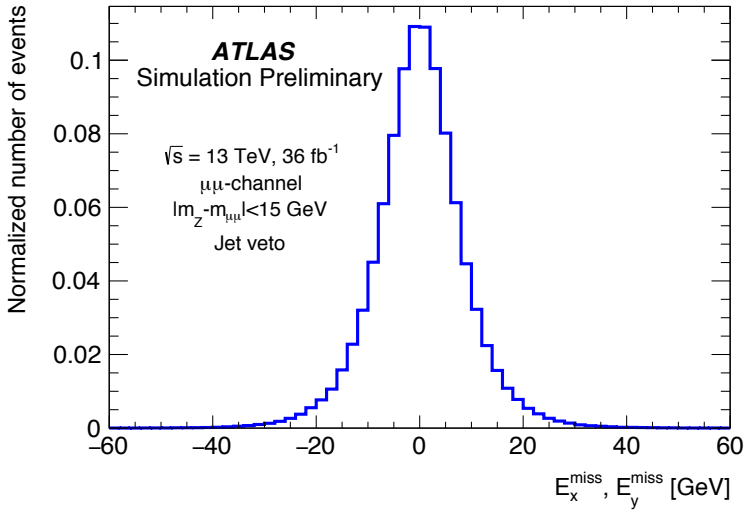


Figure 5.2 – Components of the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ ($E_{\mathrm{x}}^{\mathrm{miss}}$, $E_{\mathrm{y}}^{\mathrm{miss}}$), in the $Z \to \mu\mu$ Monte Carlo simulation with a jet veto selection. The RMS of the distribution provides an estimate of the soft term resolution since no real $E_{\mathrm{T}}^{\mathrm{miss}}$ is expected for this region.

Table 5.1 summarises the relative transverse momentum resolutions for the different objects entering the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ calculation, given for a representative kinematic value for each object. The relative resolutions are shown as a function of p_T in Figure 5.3. Resolution curves for jets include the contribution from the additional jets produced by pile-up, which gives the staggered shape for $|\eta| = 4.5$.

Table 5.1 – Resolution for the objects entering the calculation of the $E_{\mathrm{T}}^{\mathrm{miss}}$ for a representative p_{T} for each object as measured in the centre of the detector. For other η values, refer to Figure 5.3.

Object	Kinematic	Relative resolution	
		Parallel	Perpendicular
Electrons	$p_{\mathrm{T}} = 100$ GeV, $\eta = 0$	1.7%	0.4%
Photons	$p_{\mathrm{T}} = 100$ GeV, $\eta = 0$	1.9%	0.4%
Hadronic τ	$p_{\mathrm{T}} = 100$ GeV, $\eta = 0$	5.5% – 6.7%	1%
Jets	$p_{\mathrm{T}} = 20$ GeV, $\eta = 0$	22%	4.6%–7.1%
	$p_{\mathrm{T}} = 100$ GeV, $\eta = 0$	7%	1.1%–1.6%
Muons	$p_{\mathrm{T}} = 100$ GeV, $\eta = 0$	2%	0.1%
Track Soft Term		8.9 GeV	8.9 GeV

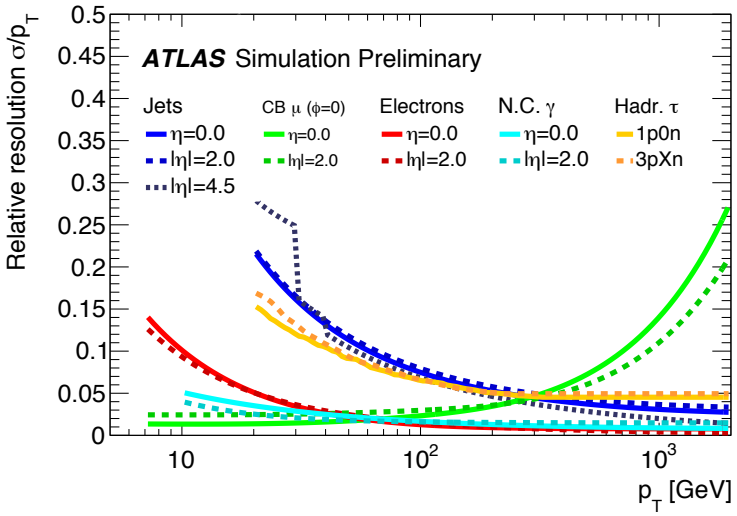


Figure 5.3 – Resolutions for the objects entering the calculation of the $E_{\mathrm{T}}^{\mathrm{miss}}$ for $|\eta| = 0.0$, $|\eta| = 2.0$ and, for jets, $|\eta| = 4.5$. The plot shows the resolutions for jets, combined muons (CB μ), electrons, photons not converted to electron-positron pairs (N.C. γ), and hadronically decaying τ-leptons (Hadr.τ). The resolution for the hadronically decaying τ-leptons denoted by $1p0n$ ($3pXn$) corresponds to the decay mode with one (three) charged hadron, zero (one or more) neutral hadrons and a non detectable τ-neutrino. The curves for jets include the contribution from pile-up, which is in bins of p_{T} and is not smoothed/interpolated, and gives the staggered shape for $|\eta| = 4.5$.

5.3 Physics samples and event selection

The $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance performance is evaluated using Standard Model processes with final states with $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{inv}} = \mathbf{0}$ and with real missing transverse momentum from neutrinos ($\boldsymbol{p}_{\mathrm{T}}^{\mathrm{inv}} = \sum \boldsymbol{p}_{\mathrm{T}}^{\boldsymbol{\nu}}$). Two di-leptonic event selections are designed to capture the leptonic Z boson decays ($Z \to \ell\ell + X$) in the electron and muon channels, respectively. They pro-

duce samples dominated by the $Z \to ee + \text{jets}$ and $Z \to \mu\mu + \text{jets}$ processes, for which $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{inv}} \sim \mathbf{o}$. These samples contain other sub-dominant processes including the diboson production process $ZZ \to ee\nu\nu$ and $ZZ \to \mu\mu\nu\nu$ for which $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{inv}} = \sum \boldsymbol{p}_{\mathrm{T}}^{\boldsymbol{\nu}}$. For this reason, these di-leptonic event selections are used to study the agreement of the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance distribution between data and Monte Carlo simulations, and the performance in separating $Z \to \ell\ell + \text{jets}$ events from $ZZ \to \ell\ell\nu\nu$ events. In the following sections, only the electron channel is shown and discussed in detail.

5.3.1 Data samples

The analysis described in this Chapter uses pp collision data collected in 2015 and 2016 by the ATLAS detector at the center-of-mass energy of 13 TeV. Data are analyzed only if they have been collected during stable proton beams with nominal magnetic field conditions, and if they satisfy the standard ATLAS data quality assessment criteria, which include checks that the inner detector, calorimeters and MS are efficiently functioning. The resulting integrated luminosity is 36 fb^{-1} [193].

5.3.2 Monte Carlo samples

Monte Carlo simulated events are used to model Standard Model processes which are normalized using the currently available predictions for their cross-sections. These simulated events were passed through the ATLAS detector simulation [194] based on GEANT 4 [195]. The effects of pile-up were modeled by overlaying minimum-bias events, simulated using the soft QCD processes of Pythia 8.1 [196] with the A2 [197] set of tuned parameters and MSTW2008LO [198] parton distribution functions (PDF). The simulated samples are weighted to reproduce the observed distribution of the mean number of interactions per bunch crossing in the data. For all samples of simulated events, except for those generated using Sherpa [199], the EvtGen v1.2.0 program [200] was used to describe the decays of bottom and charm hadrons.

Events with W or Z bosons decaying into leptons produced in association with jets are simulated using the Sherpa 2.2.1 [199] generator. Matrix elements are calculated for up to 2 partons at NLO and 4 partons at LO using the Comix [201] and OpenLoops [202] matrix element generators and merged with the Sherpa parton shower [203] using the ME+PS@NLO prescription [204]. The W/Z + jets events are normalized to the NNLO cross-sections.

Diboson processes (WW, WZ and ZZ) are simulated using the Powheg-Box v2 generator [205] to the next-to-leading order (NLO) in QCD and after that interfaced with the Pythia 8 generator. Next-to-next-leading order (NNLO) in QCD and NLO electroweak corrections (EW) are added as a function of the invariant mass, m_{ZZ}, of the ZZ process.

For the generation of $t\bar{t}$ at NLO, the Powheg-Box v2 generator was used. Single top quark events in the s-, t- and Wt-channels were generated using the Powheg-Box v1 generator [206, 207]. The top quark mass was set to 172.5 GeV. The overall yield predicted for the $t\bar{t}$ process is rescaled according to the NNLO cross-section, including the resummation of soft gluon emission at next-to-next-to-leading-logarithm accuracy

(NNLL) as available in TOP++2.0 [208]. The overall yields predicted for single top quark production in the s-, t-, and Wt-channels are rescaled according to their respective NLO cross-sections [209, 210, 211].

5.3.3 Event selection

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

Events are required to have two opposite-charge leptons of the same flavor and pass the "medium" identification criteria, see Sections 4.3 and 4.4 for details, with the more energetic lepton having $p_{\mathrm{T}} > 30$ GeV and the subleading lepton having $p_{\mathrm{T}} > 20$ GeV. If the event has any additional lepton that meets the "loose" identification criteria and having $p_{\mathrm{T}} > 7$ GeV, the event is rejected to reduce the contamination from the WZ process.

The invariant mass of the lepton pair is required to be in the mass window $|m_Z - m_{\ell\ell}| < 15$ GeV, where m_Z is equal to the value of the Z boson mass obtained by the Particle Data Group, and reported in Reference [212]. This condition significantly reduces contributions from the $t\bar{t}$, Wt, and WW processes.

To study the performance of the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance in different topologies, additional selections based on the number of reconstructed jets and on the magnitude of the measured $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ are applied.

5.3.4 Data and Monte Carlo comparisons

The study of the distributions of the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance and other associated kinematic quantities is important to characterize this variable, to investigate its properties, and to assess the capability of the Monte Carlo simulations to reproduce several kinematic distributions observed in data. The distributions for events passing the di-leptonic selections in the electron channel are shown in Figure 5.4 and Figure 5.5.

The $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution, shown in Figure 5.4, is dominated by events with low value of $E_{\mathrm{T}}^{\mathrm{miss}}$ produced by $Z \to ee$ + jets process. At higher values of $E_{\mathrm{T}}^{\mathrm{miss}}$, events are mostly produced by processes with real $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$, such as $t\bar{t}$ or diboson processes. Despite this transition, the plot shows the capability of the Monte Carlo simulation to reproduce the distribution observed in data, with the data/MC ratio being covered by the statistical and experimental systematic uncertainties of the Monte Carlo expectation (band in gray). The experimental systematic uncertainties include uncertainties on lepton and jet reconstruction, identification, isolation and trigger efficiencies, uncertainties on lepton, jet, and soft term energy resolution and scale, uncertainty on the total integrated luminosity, and uncertainty on the mean number of interactions per bunch crossing.

The $E_{\mathrm{T}}^{\mathrm{miss}}$ corresponds to the numerator of the event-based and object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significances of Equations (5.1) and (5.10), respectively. Figure 5.5(a) and Figure 5.5(b) show the respective denominators. The Monte Carlo simulation for $\sqrt{\sum E_{\mathrm{T}}}$ in Figure 5.5(a) reproduces the features of the observed distribution in data with a non-perfect agreement.

This difference, already reported in Reference [179], reflects the level of mismodelling of the final state mostly in terms of the hard-object composition in Monte Carlo simulations. The distribution for $\sigma_L\sqrt{1-\rho_{LT}^2}$ in Figure 5.5(b) shows a good agreement below 20 GeV. For values above 20 GeV, the Monte Carlo simulation overshoots the distribution observed in data. The two visible structures around 20 and 60 GeV are present in both data and Monte Carlo simulations. These structures are a result of the contribution to the covariance matrix of the parametrization used for additional jets from pile-up events. They show up at the jet reconstruction threshold (20 GeV) and at the upper limit for suppression of the jet from pile-up (60 GeV). The data-MC difference for $\sigma_L\sqrt{1-\rho_{LT}^2} > 20$ GeV, reflects the level of accuracy in modeling the rate of jets produced by pile-up in Monte Carlo simulations [150]. Given a certain value of the $E_{\mathrm{T}}^{\mathrm{miss}}$, a higher value of $\sigma_L\sqrt{1-\rho_{LT}^2}$ produces a lower value for the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance. As a result, the value of the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance in events containing jets likely produced by pile-up is smaller than events without these jets, as one would expect.

Figure 5.5(c) and Figure 5.5(d) show the distributions for the event-based and object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significances, respectively. Both distributions are dominated by events with low significance values, as expected from the $Z \to ee +$ jets process. The distribution of the event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance shows a compression of the horizontal axis compared to the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$. This is consistent with the studies reported in Reference [176] in which the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution $\sigma(E_{\mathrm{T}}^{\mathrm{miss}})$ has been parametrized as $\sigma(E_{\mathrm{T}}^{\mathrm{miss}}) = k \cdot \sqrt{\sum E_{\mathrm{T}}}$, where $k \sim 0.4 - 0.5$ GeV$^{1/2}$. For values above 6, events are mostly produced by processes with real $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$, with a longer tail for the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance in Figure 5.5(d). The Monte Carlo simulation reproduces the distributions observed in data, with a very good matching for the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance, which seems to be less affected by mismodelling than the event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance.

Given the change in physics process composition in the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in Figure 5.4, it is important to characterize the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance in different conditions, and at different values of $E_{\mathrm{T}}^{\mathrm{miss}}$. Figure 5.6 shows the same variables shown in Figure 5.5 with the additional requirement of $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV.

The distribution for the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance in Figure 5.6(d) shows some interesting features. The difference between processes with and without real $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ is already evident, with some first indication of an improved separation power with respect to the event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance (confirmed in the following Section). For the $Z \to ee +$ jets process, the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance has values significantly smaller than the other processes, with a shoulder for values below 4. For these events, the jets produced by pile-up play an important role, as a consequence of their contribution to $\sigma_L\sqrt{1-\rho_{LT}^2}$ already discussed above. Not surprisingly, for low values of the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance, the level of data-MC agreement is similar to the one above 20 GeV in Figure 5.5(b) and Figure 5.6(b). This confirms the level of mismodelling of the rate of jets produced by pile-up in Monte Carlo simulations. The capability to isolate these events without real $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ is a clear advantage for the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance.

Distributions for data and MC in the $\mu\mu$-channel can be found in Appendix B.2.

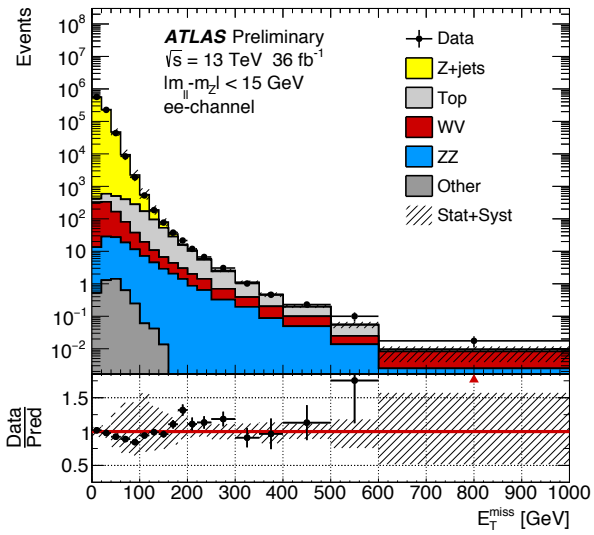


Figure 5.4 – $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in data compared to MC predictions including all relevant backgrounds for events satisfying the $Z \to ee$ selection. The ratio between data and MC predictions is shown below the distribution, with the shaded band which corresponds to the combined experimental systematic and MC statistical uncertainties. The experimental systematic uncertainties include uncertainties on lepton and jet reconstruction, identification, isolation and trigger efficiencies, uncertainties on lepton, jet, and soft term energy resolution and scale, uncertainty on the total integrated luminosity, and uncertainty on the mean number of interactions per bunch crossing. The last bin of the distribution includes overflows.

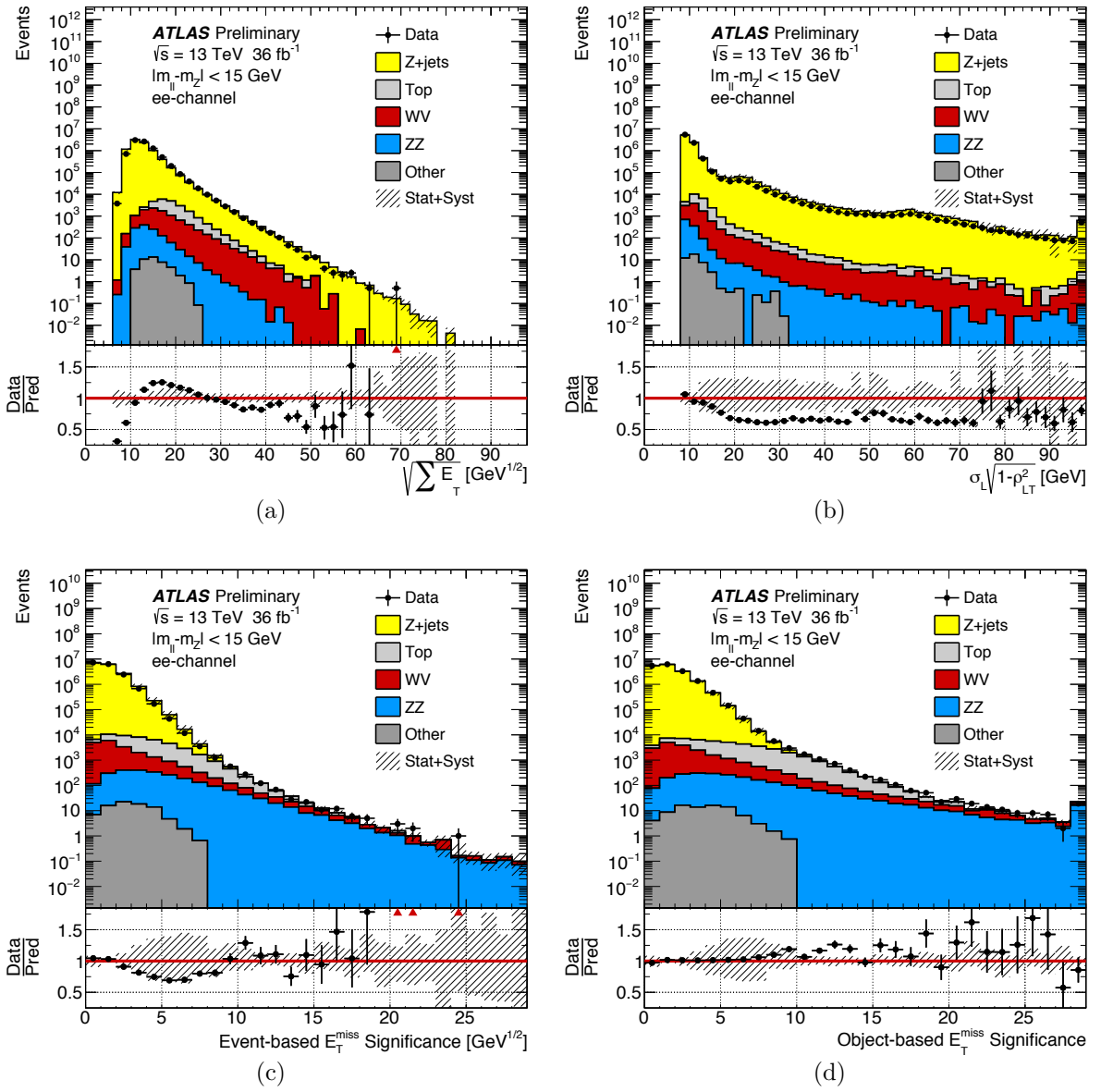


Figure 5.5 – Distributions in data compared to MC predictions including all relevant backgrounds for events satisfying the $Z \to ee$ selection for: (a) $\sqrt{\sum E_{\mathrm{T}}}$, (b) $\sigma_L\sqrt{1-\rho^2_{LT}}$, (c) event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance (Equation 5.1), and (d) object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance (Equation 5.10). The respective ratios between data and MC predictions are shown below the distributions, with the shaded bands which correspond to the combined experimental systematic and MC statistical uncertainties. The experimental systematic uncertainties include uncertainties on lepton and jet reconstruction, identification, isolation and trigger efficiencies, uncertainties on lepton, jet, and soft term energy resolution and scale, uncertainty on the total integrated luminosity, and uncertainty on the mean number of interactions per bunch crossing. The last bin of the distribution includes overflows.

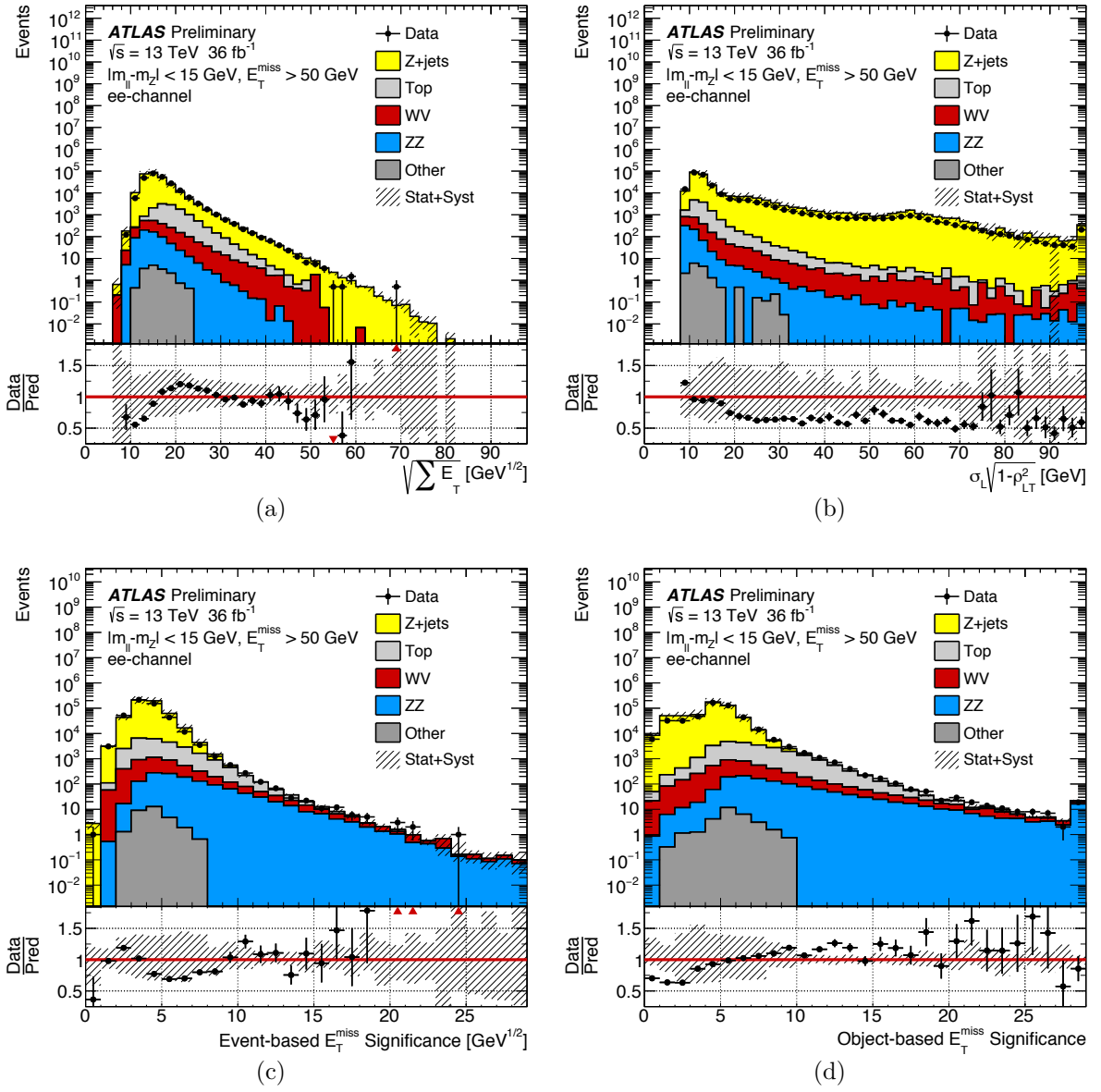


Figure 5.6 – Distributions in data compared to MC predictions including all relevant backgrounds for events satisfying the $Z \to ee$ selection and $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV for: (a) $\sqrt{\sum E_{\mathrm{T}}}$, (b) $\sigma_L\sqrt{1-\rho_{LT}^2}$, (c) event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance (Equation 5.1), and (d) object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance (Equation 5.10). The respective ratios between data and MC predictions are shown below the distributions, with the shaded bands which correspond to the combined systematic and MC statistical uncertainties. The experimental systematic uncertainties include uncertainties on lepton and jet reconstruction, identification, isolation and trigger efficiencies, uncertainties on lepton, jet, and soft term energy resolution and scale, uncertainty on the total integrated luminosity, and uncertainty on the mean number of interactions per bunch crossing. The last bin of the distribution includes overflows.

5.3.5 Pile-up jet resolution impact in total variance

The effect of additional pile-up jets on the calculation of the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ is suppressed by using the jet vertex tagger (JVT) algorithm, where the relevant discriminant ranges from 0 (pile-up-like) to 1 (hard-scatter-like), for each jet with matched tracks, and is set to be -0.2 if no tracks are found inside the jet.

Jets with JVT ≤ 0.59, $p_{\mathrm{T}} < 60\,\mathrm{GeV}$, and $|\eta| < 2.4$ are removed from the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ calculation. The remaining jets have a small but non-negligible probability of originating from pile-up interactions. For these, an additional resolution source, important for the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance, is set to be equal to the p_{T} of the jet, multiplied by the probability for that jet to originate from pile-up interactions. This probability has been estimated in Monte Carlo simulations in two ranges of pseudo-rapidity, central ($|\eta| < 2.7$) and forward ($|\eta| > 2.7$). To highlight the effect of this extra resolution source on the calculation of the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance, Figure 5.7 shows the impact in events with one jet. This plot, produced using the $Z \to ee$ selection on a $Z+$ jets Monte Carlo simulation, presents the distribution of σ_L, as introduced in Equation 5.8, and the contribution of events with a large relative jet resolution due to the effect of pile-up ($\sigma_L^{PU}/p_{\mathrm{T}} > 0.5$). By dividing the sample with a cut on the jet p_{T}, it is possible to isolate the two structures for σ_L equal to 20 and 60 GeV.

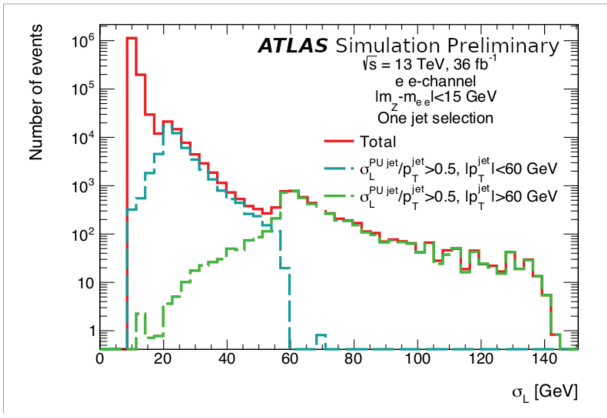


Figure 5.7 – Distribution of σ_L in $Z+$ jets Monte Carlo simulation for the $Z \to ee$ event selection, in events with one jet. Events with large relative jet resolution due to the estimation of the effect of pile-up are isolated, and divided in two groups, depending on the p_{T} of the jet.

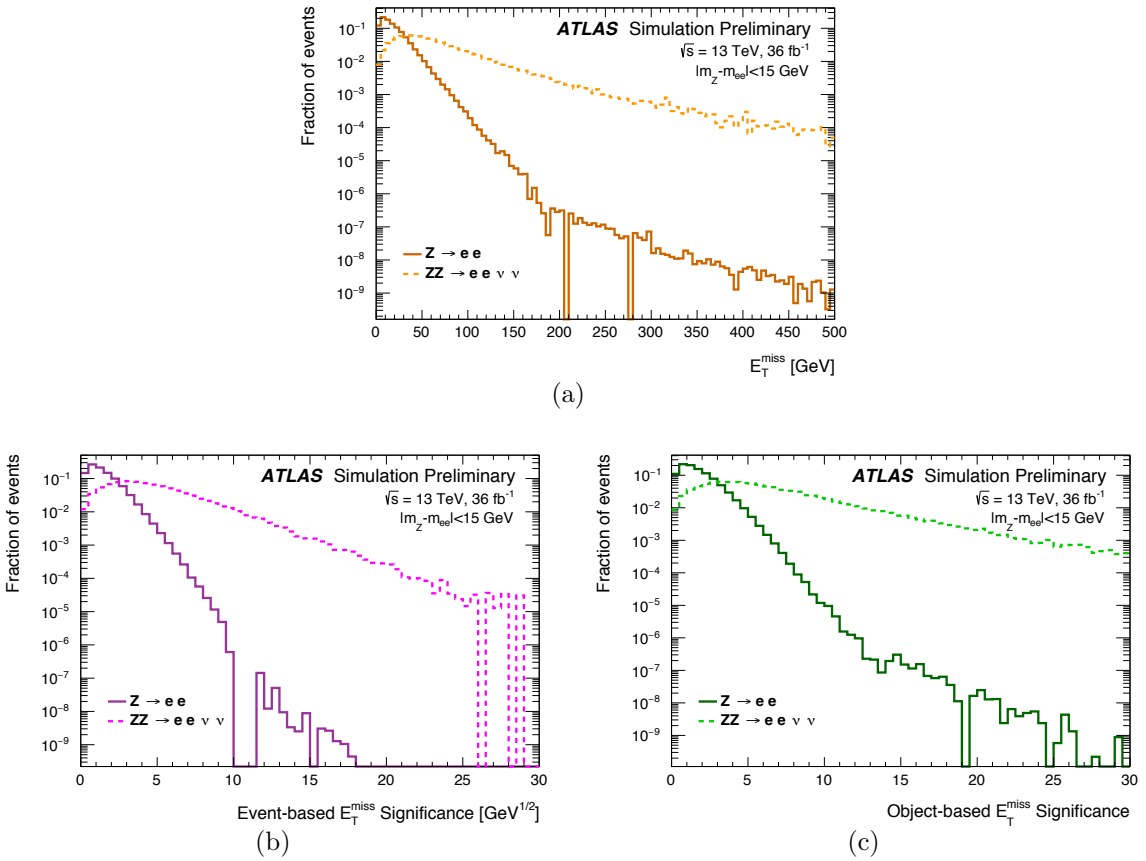
5.4 Characterization and performance of $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance

This section examines the potential gain of introducing the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance variable into the selection criteria for $Z \to ee$ events described in Section 5.3.3. These events provide an ideal final state for evaluating the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance performance,

since they can be selected with high signal to background ratio and because Z-boson kinematics can be measured with high precision. This process is considered to have no real $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$, apart from sub-leading contributions from hadronic decays involving neutrinos, like in the semileptonic decays of the heavy flavor hadrons as stated above (see Figure 5.8(a)). As a result, the bulk of the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance distribution peaks near zero, and it rapidly decreases at higher values as is shown in Figure 5.8(b) for the event-based and 5.8(c) for the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance. It is interesting to notice the presence of a tail at high values of the variables shown in Figure 5.8(a) and Figure 5.8(c), partially due to the presence of the sub-leading contributions of events with real $E_{\mathrm{T}}^{\mathrm{miss}}$. This tail is not discernable for the event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance in Figure 5.8(b).

However, in $ZZ \to ee\nu\nu$ events, the presence of real $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ increases the value of the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance per event as is shown in Figure 5.8. The $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance variable can therefore be used to separate events with real $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ in $ZZ \to ee\nu\nu$ coming from the neutrinos with respect to events with only resolution-induced $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ as the ones present in $Z \to ee$. In the following, the $ZZ \to ee\nu\nu$ sample is considered as signal and the $Z \to ee$ sample as background. Refer to Appendix B.3 for signal and background MC distributions in different kinematic regions related to the performance studies presented in this Section, and to Appendix B.4 for the equivalent plot of Figure 5.8 in the $\mu\mu$-channel.

In the following sections, the separation power between $Z \to ee$ and $ZZ \to ee\nu\nu$ simulated events of the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance is discussed for different jet multiplicities and $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ regions, and compared to the separation power of the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ alone, and of the event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance.



(a)

(b)

(c)

Figure 5.8 – MC simulated distributions for $Z \to ee$ and $ZZ \to ee\nu\nu$ samples with a $Z \to ee$ selection, which corresponds to background and signal respectively in the performance study. The plots show (a) $E_{\mathrm{T}}^{\mathrm{miss}}$, (b) the event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance (Equation 5.1), and (c) the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance (Equation 5.10).

5.4.1 Separation performance in inclusive jet events

The $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$, the event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance, and the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance show a good separation power between $ZZ \to ee\nu\nu$ and $Z \to ee$ events. To estimate the potential gain of using the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance, the separation power of $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$, object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance, and event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance are compared for events passing increasing thresholds on the value of $E_{\mathrm{T}}^{\mathrm{miss}}$. The plots in Figure 5.9 have different pre-selection thresholds on the total reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$: 0, 50, and 100 GeV, without any requirement on the presence of jets.

A small gain in separation power is reached by the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance if compared to the other two variables for the inclusive selection, in which the bulk of the events have little hard activity and the denominator of the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance is dominated by the variance of the soft-term. A clear improvement is observed for events with higher $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$. For a 50 GeV $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ threshold, and choosing a working point at 80% signal efficiency, object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance gives almost 90% background rejection, while $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ and event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance give 55% and 70% background rejection, respectively, as is shown in Figure 5.9(b). This indicates that selection requirements based on both $E_{\mathrm{T}}^{\mathrm{miss}}$ and object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance may be beneficial. This is shown in Section 5.4.2.

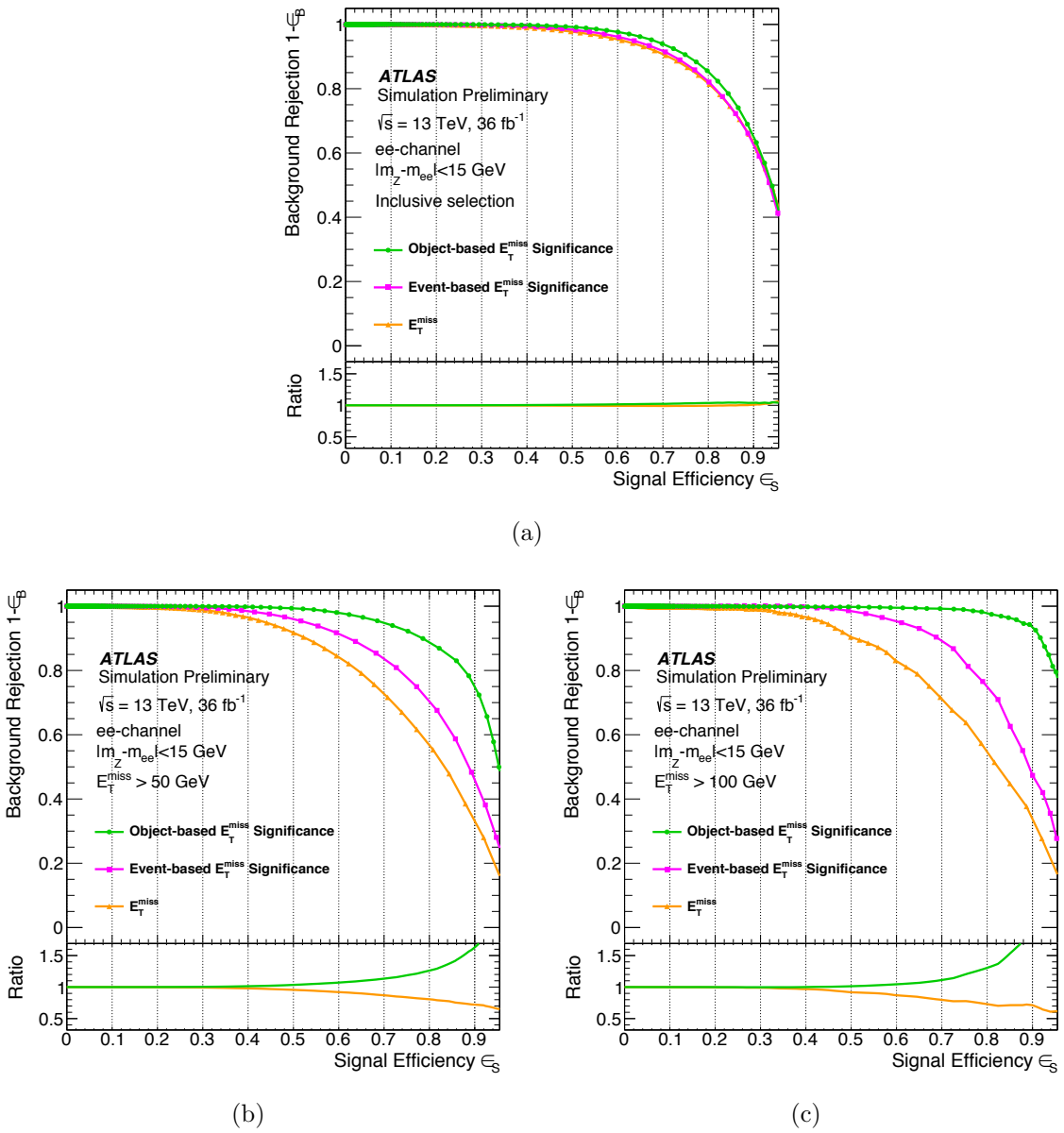
At higher $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ thresholds, the performance gain for the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance with respect to the other two variables is even more evident. In the case of a $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}} > 100$ GeV pre-selection, and for a working point at 80% signal efficiency, object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance gives almost 98% background rejection, while $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ and event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance give approximately 55% and 75%, respectively, as is shown in Figure 5.9(c).

Therefore, for both 50 and 100 GeV $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ thresholds, using object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance results in a background rejection improvement of approximately 30% with respect to the event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance. This improvement is guided by the contribution to the covariance matrix from the jets produced by pile-up interactions. By using this contribution, the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance can more effectively separate events with fake $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ due to the presence of jets produced by pile-up interactions, from events with real $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$.

5.4.2 Separation performance in jet bins for $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV

From Section 4.6.1, it is clear that the primary source of resolution in the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance calculation comes from the energy resolution of low p_{T} jets. To further study the impact of the presence of jets in the final state, the performance with a $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ threshold of $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV (recall Figure 5.9(b)) has been investigated more in detail. Events have been divided in four categories according to the number of jets: jet veto, exactly one jet, exactly two jets, three or more jets. Figure B.6 shows the separation power of the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$, event- and object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significances in these four cases.

In the case of a jet veto selection, the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance considers primarily the effect of the soft term and its corresponding resolution. Since the soft term variance is estimated with a constant number, as explained in Section 5.2, the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$



(a)

(b)

(c)

Figure 5.9 – Background rejection versus signal efficiency in simulated $Z \to ee$ and $ZZ \to ee\nu\nu$ samples with a $Z \to ee$ selection. The performance is shown for $E_{\mathrm{T}}^{\mathrm{miss}}$, event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance, and object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance as discriminants in events with (a) the entire $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ range, (b) $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV, and (c) $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}} > 100$ GeV. The lower panel of the figures shows the ratio of other definitions/event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance.

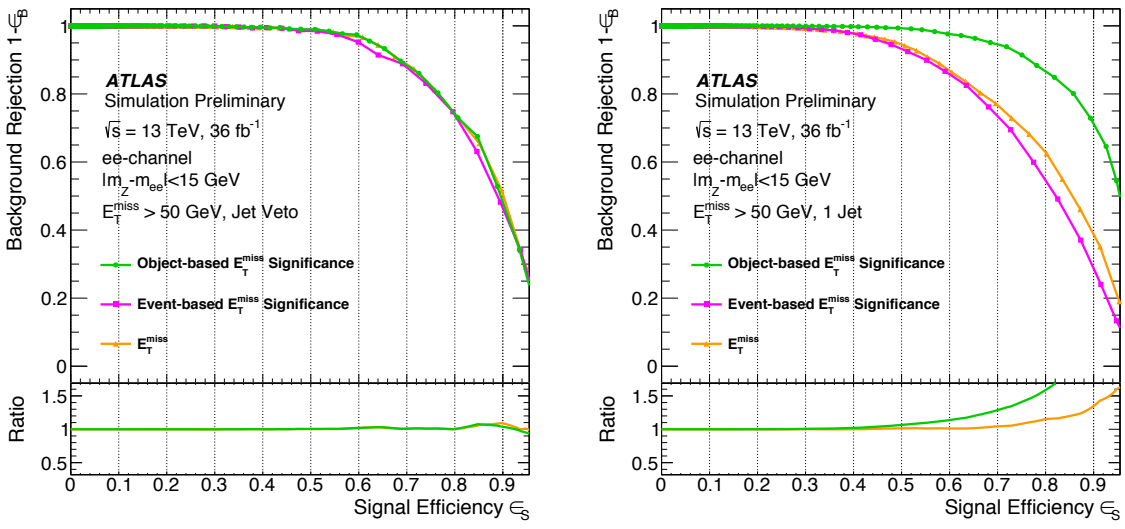
significance is de-facto a re-scaling of the total $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$. Therefore, no gain in performance is expected for the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance compared to $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ alone. Also, given that the inclusive jet selection is dominated by events with zero jets for $Z \to ee$ events, in Figure B.6(a) no significant gain is expected for the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance.

The separation power for $Z \to ee$ simulated events with respect to $ZZ \to ee\nu\nu$ simulated events of the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance starts to improve upon the event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance at higher jet multiplicities, given that the jet resolutions begin to have an important impact in the total variance.

When selecting one single jet, the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance shows a clear performance gain compared to $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ and event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance. Choosing a working point at 80% signal efficiency, object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance gives a 85% background rejection, while $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ and event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance give 60% and 55% background rejection, respectively, as is shown in Figure B.6(b). This gives an improvement of approximately 55% in background rejection with the object-based compared to the event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance.

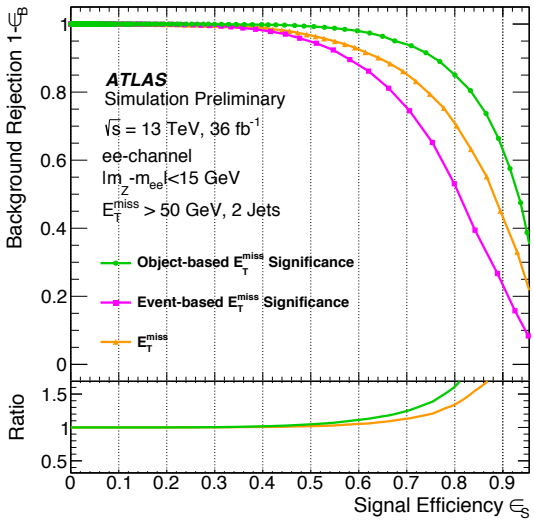
For higher jet multiplicities in Figure B.6(c) and Figure B.6(d), the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance, with a signal efficiency of 80%, shows a gain in background rejection of approximately 70% and 40% with respect to the event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance in the case of two and three or more jets selection, respectively.

Regarding the $\mu\mu$-channel, the background rejection versus signal efficiency plots can be found on Appendix B.5.

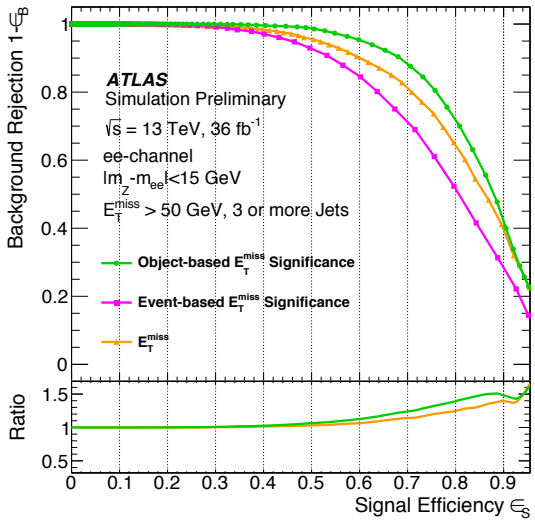


(a)

(b)



(c)



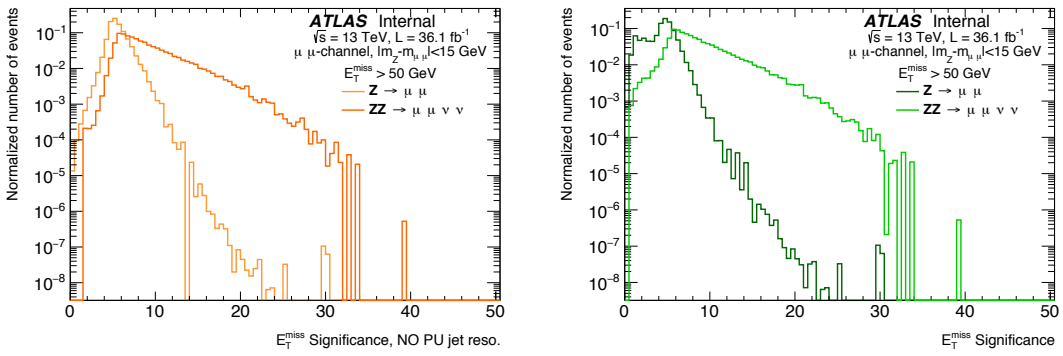
(d)

Figure 5.10 – Background rejection versus signal efficiency in simulated $Z \to ee$ and $ZZ \to ee\nu\nu$ samples with a $Z \to ee$ selection and $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV. The performance is shown for $E_{\mathrm{T}}^{\mathrm{miss}}$, event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance, and object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance as discriminants in events with (a) jet veto, (b) one jet, (c) two jets, and (d) three or more jets. The lower panel of the figures shows the ratio of other definitions/event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance.

5.4.3 Pile-up jet resolution impact in the separation performance

To estimate the potential gain in performance of the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance using an additional pile-up jet resolution described in Section 5.2, the signal efficiency and background rejection is calculated for object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance with and without this additional resolution.

Figure 5.11 shows the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance with a $Z \to \mu\mu$ selection for signal and background events with a $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ >50 GeV threshold in the $\mu\mu$-channel. An improved separation power between signal an background on Figure 5.11(b) is clear with respect to Figure 5.11(a). Events that are more likely to have jets coming from pile-up interactions tends to accumulate at low values of the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance as can be seen on Figure 5.11(b).



(a) (b)

Figure 5.11 – Object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance with a $Z \to \mu\mu$ selection for signal and background events with a $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ >50 GeV threshold.

Figure 5.12 shows the impact on the performance between the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance with and without an additional pile-up jet resolution. For a fixed signal efficiency working point of 80%, the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance with pile-up jet resolution has a background rejection gain up to 20%.

5.5 Bias estimation and soft term resolution

The study of $\boldsymbol{\Delta E}_{\mathrm{T}}^{\mathrm{miss}} = \boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}} - \boldsymbol{p}_{\mathrm{T}}^{\mathrm{inv}}$ is particularly important for the definition of the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance. Under the hypotheses in Section 5.1, this quantity should follow a 2-D Gaussian distribution, with $\langle \boldsymbol{\Delta E}_{\mathrm{T}}^{\mathrm{miss}} \rangle = \mathbf{0}$ and variance $\mathbf{Var}(\boldsymbol{\Delta E}_{\mathrm{T}}^{\mathrm{miss}}) = \mathbf{V}_{L,T}$ correctly rotated. Consequently, the projection of $\boldsymbol{\Delta E}_{\mathrm{T}}^{\mathrm{miss}}$ in any non-trivial direction should be distributed as a Gaussian distribution with a mean of zero. This is not true for the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ definition used in this note, as already reported in Reference [179]. For the di-leptonic event selection used in this note, the projection of the $\boldsymbol{\Delta E}_{\mathrm{T}}^{\mathrm{miss}}$ in the direction of the di-lepton pair forming the Z boson candidate $\boldsymbol{A_Z}$ in events passing a jet veto shows a clear bias (Figure 5.13).

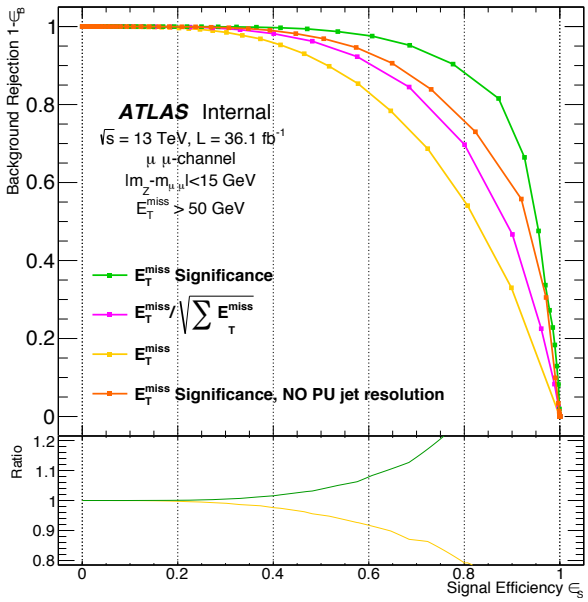


Figure 5.12 – Background rejection versus signal efficiency in simulated $Z \to \mu\mu$ and $ZZ \to \mu\mu\nu\nu$ samples with a $Z \to \mu\mu$ selection and $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV.

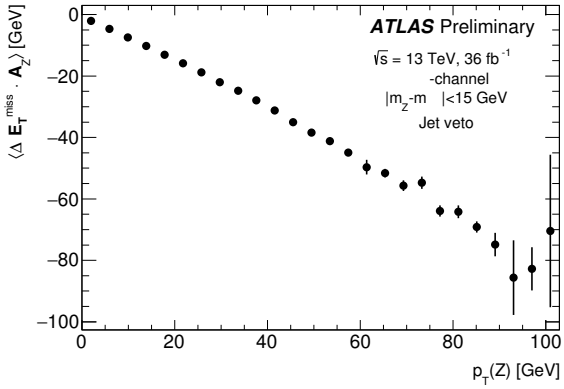


Figure 5.13 – The average projection of $\boldsymbol{\Delta E}_{\mathrm{T}}^{\mathrm{miss}}$ in the direction $\boldsymbol{A_Z}$ of the di-lepton pair forming the Z boson candidate with transverse momentum vector $\boldsymbol{p}_{\mathrm{T}}^{\boldsymbol{Z}}$, is shown as function of p_{T}^{Z} in $Z \to \mu\mu$ events from the sample with a jet veto.

As already anticipated in Section 5.1, this bias is due to particles which interact with the detector, but fail to contribute to the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ measurement either because they have a p_{T} lower than the threshold or because they are neutral particles not considered in the track-based soft term reconstruction, or lost for any other reason. As a consequence, the estimation of 8.9 GeV for the resolution for the track-based soft term in Section 5.2 mixes together the effect of this bias, with the real track-based soft term resolution. An improvement of the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance requires a replacement of the track-based soft term variance described in Section 5.2 with a correction for the bias on the numerator, and an improvement of the estimation of the variances on the denominator.

These two quantities can be parametrized from Monte Carlo simulations as follows:

$$\begin{aligned} \boldsymbol{bias} &= \langle \boldsymbol{\Delta E}_{\mathrm{T}}^{\mathrm{miss}} \rangle && = \boldsymbol{bias}(obs_1, obs_2, ...) \\ \mathbf{V} &= \mathbf{Var}(\boldsymbol{\Delta E}_{\mathrm{T}}^{\mathrm{miss}} - \boldsymbol{bias}(obs_1, obs_2, ...)) && = \mathbf{V}(obs_1, obs_2, ...) \end{aligned} \tag{5.13}$$

where $\boldsymbol{bias}(obs_1, obs_2, ...)$ and $\mathbf{V}(obs_1, obs_2, ...)$ are parametrizations of the bias vector and covariance matrix as a function of event-by-event observables $obs_1, obs_2, ...$

The $Z \to \mu\mu$ sample with a jet veto has been used to derive these parametrizations, thanks to the negligible contributions from the hard objects resolution. Many parametrizations with several options for the variables *obs* have been tested: the magnitude of the track-based soft term, the magnitude of the vectorial sum of the hard objects, the $E_{\mathrm{T}}^{\mathrm{miss}}$, the projections of these three variables on specific directions, the H_{T}, the $\sum E_{\mathrm{T}}$ and the scalar p_{T} sum of the tracks forming the soft term. Figure 5.14 shows an example of these parametrizations. In this example, the variable $\boldsymbol{\Delta E}_{\mathrm{T}}^{\mathrm{miss}}$ is projected in the direction of the TST to derive the parametrizations of the bias and of the variance which depend on the magnitude of the track-based soft term. As a consequence, the $\boldsymbol{bias}$ correction is applied in the direction of the track-based soft term.

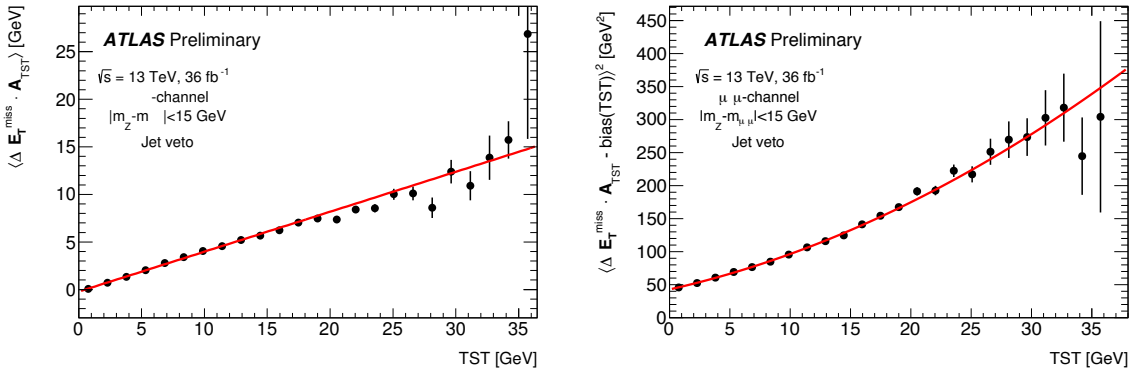
A similar strategy has been adopted to derive a parametrization using the variable $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{HS}}$, already used to study the performance of the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ [179, 177]:

$$\begin{aligned} \boldsymbol{p}_{\mathrm{T}}^{\mathrm{HS}} &= \boldsymbol{p}_{\mathrm{T}}^{\mathrm{inv}} + \sum_{i \in \mathrm{muons}} \boldsymbol{p}_{\mathrm{T}}{}^{i} + \sum_{i \in \mathrm{electrons}} \boldsymbol{p}_{\mathrm{T}}{}^{i} + \sum_{i \in \mathrm{photons}} \boldsymbol{p}_{\mathrm{T}}{}^{i} + \sum_{i \in \mathrm{hadronic}\ \tau} \boldsymbol{p}_{\mathrm{T}}{}^{i} + \sum_{i \in \mathrm{jets}} \boldsymbol{p}_{\mathrm{T}}{}^{i} \\ &= -\boldsymbol{\Delta E}_{\mathrm{T}}^{\mathrm{miss}} - \sum_{i \in \mathrm{Soft\ Term}} \boldsymbol{p}_{\mathrm{T}}{}^{i} \end{aligned} \tag{5.14}$$

This variable shows a strong anti-correlation with $\boldsymbol{\Delta E}_{\mathrm{T}}^{\mathrm{miss}}$. It is therefore a natural candidate to derive a parametrization, but it uses $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{inv}}$ which cannot be calculated on data. A parametrization using $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{HS}}$ is anyhow interesting to interpret the results. In this case, the variable $\boldsymbol{\Delta E}_{\mathrm{T}}^{\mathrm{miss}}$ is projected in the direction of the $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{HS}}$ to derive the parametrization of the bias which is proportional to the magnitude of $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{HS}}$. The $\boldsymbol{bias}$ correction is applied in the direction of $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{HS}}$.

A last parametrization is derived to investigate the effect of using just information on the direction of $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{HS}}$ and not its magnitude. In this last parametrization, the $\boldsymbol{\Delta E}_{\mathrm{T}}^{\mathrm{miss}}$ is projected in the direction of the $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{HS}}$, parametrized as a function of TST, and applied back as a $\boldsymbol{bias}$ in the direction of $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{HS}}$.

These three parametrizations are shown in Figure 5.15. Comparing the different parametrizations highlights the relevance of the correction for the bias in the numerator, while the parameterizations of the soft term variance were found to have a minor impact on the performance. The first parametrization (Object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ Significance, magnitude and direction from TST in Figure 5.15) has a separation power similar to the case where the bias is neglected (Object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ Significance in Figure 5.15), while the other two parametrizations present some gain in performance. The improvement in the separation power shown by the second parametrization (Object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ Significance, magnitude and direction from $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{HS}}$ in Figure 5.15), depends heavily on the use of $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{inv}}$, and it represents a limit case in the gain in performance one could expect from these



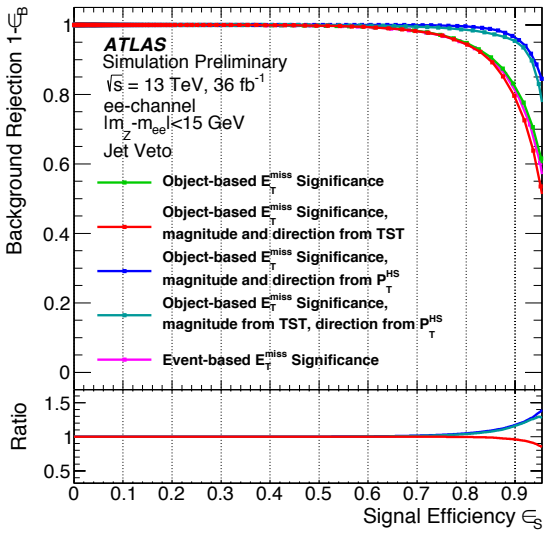
(a) (b)

Figure 5.14 – Examples of the parametrizations for the ***bias*** and the variance **V**, in $Z \to \mu\mu$ events from the sample with a jet veto. (a) The average projection of $\boldsymbol{\Delta E}_{\mathrm{T}}^{\mathrm{miss}}$ in the direction of the track-based soft term $\boldsymbol{A}_{\mathrm{TST}}$ shown as function of the magnitude of the track-based soft term. (b) Variance of $(\boldsymbol{\Delta E}_{\mathrm{T}}^{\mathrm{miss}} - \boldsymbol{bias}(obs_1, obs_2, ...))$ projected in the direction of the track-based soft term shown as function of the magnitude of the track-based soft term. The red curves are the results of a linear fit for (a) and a quadratic fit for (b). They are used to parametrize the bias and the variance of the soft term.

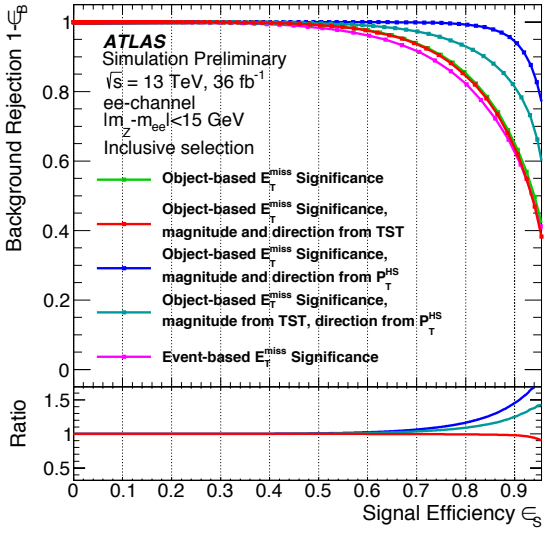
parametrizations. The third parametrization (Object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ Significance, magnitude from TST, direction from $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{HS}}$ in Figure 5.15), which uses $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{inv}}$ just to determine the direction of the bias correction, improves the suppression of background events by 10%–20% for a signal efficiency working point in the range 80%–90% in events without jets, where the impact of the soft term resolution plays a leading role in the calculation of the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance. This suggests that an improvement in the determination of the direction of the bias correction could guide a potential gain in the performance of the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance. This is a promising indication for future work. However, at this stage, this is not included in the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance definition as the direction in which the bias correction should be applied cannot be accurately determined yet.

5.6 Summary and outlook

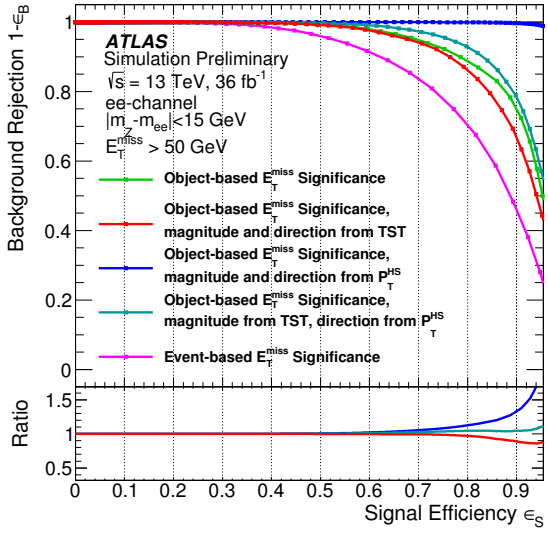
A new object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance variable has been derived in the context of the ATLAS experiment and has been illustrated in this Chapter. This variable makes use of the estimated resolutions of objects contributing to the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ and their relative orientations to the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ to differentiate between events in which the reconstructed missing transverse momentum is real, coming from weakly interacting particles, or is fake, consistent with contributions coming from particle mis-measurement, resolutions and efficiencies. Accounting for the probabilities that jets are induced by pileup is a crucial aspect of this new definition. The performance of the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance has been evaluated using event selections with two leptons designed to contain the leptonic



(a)



(b)



(c)

Figure 5.15 – Background rejection versus signal efficiency in simulated $Z \to ee$ and $ZZ \to ee\nu\nu$ samples with a $Z \to ee$ selection (a) with a jet veto and no $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ cut, (b) inclusive in $E_{\mathrm{T}}^{\mathrm{miss}}$ and number of jets, and (c) without any jet requirements and with $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV. The performance is shown for event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance, standard object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance, and three augmented object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significances with the improved bias and soft term variance parametrizations as discriminants. The lower panel of the figures shows the ratio of the background rejection for other object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ definitions to the one considering a constant soft term resolution.

Z boson decays ($Z \to \ell\ell + X$) in the electron and muon channels. The separation power

of the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance between events with a real $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ produced by di-boson processes and events produced by $Z \to \ell\ell+$ jets process shows a clear improvement compared to the use of the $E_{\mathrm{T}}^{\mathrm{miss}}$ or the event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance.

Chapter 6

Mono-$h(b\bar{b})$: search for Dark Matter in association with $h \to bb$, analysis selection and backgrounds

This Chapter presents the search for dark matter in association with a Higgs boson decaying into a pair of b quarks, mono-$h(b\bar{b})$. The data analysed was recorded during 2015 and 2016, corresponding to 36 fb^{-1}. This analysis updates and improves the previous analysis result with data collected during 2015, which corresponds to 3.2 fb^{-1} [213].

The Chapter is organised as follow. Section 6.1 generally introduces the mono-$h(b\bar{b})$ search and strategy. Section 6.2 describes the physics object selection which is inherited from the search for the $b\bar{b}$ decay of the SM Higgs boson in associated $(W/Z)H$ production [214], referred as $VH(bb)$. Section 6.3 describes the data-set analysed and the trigger used. Section 6.4 details the event selection, and Section 6.5 describes the main backgrounds in the mono-$h(b\bar{b})$ dark matter search. Subsequently, next Chapter 7 presents the statistical treatment and the results for data recorded during 2015 and 2016. Chapter 8 presents the results corresponding to the mono-$h(b\bar{b})$ analysis for data recorded from 2015 to 2017.

6.1 Mono-$h(b\bar{b})$ analysis overview

This analysis focuses on using the discovered Higgs boson in Run1 [20, 21, 215] to search for dark matter produced in proton-proton collisions at the LHC. In particular, this search considers the decay channel of the Higgs boson to two bottom quarks ($h \to b\bar{b}$) which corresponds to the decay channel with the largest branching fraction, $\mathcal{B} = 57\%$ [23].

The search for dark matter production in association with a Higgs boson opens an opportunity to directly probe the hard interaction involving the dark matter particles since the Higgs boson is not likely to originate from initial state radiation given that such process is Yukawa-suppressed.

The Higgs boson is typically produced with large transverse momentum, recoiling from the dark matter particles produced on the opposite side of the event, via $h + E_{\mathrm{T}}^{\mathrm{miss}}$ signature. This leads to two distinguishing signatures in the final state. The first is due

to the presence of the undetected dark matter particles, which produces large amounts of missing transverse momentum, $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$. The second is due to the presence of a Higgs boson decaying to a pair of bottom quarks.

The angular separation, $\Delta R_{b\bar{b}}$, of the two b quarks produced in the Higgs boson decay depends on the Lorentz boost of the Higgs boson, which vary significantly with the Higgs p_{T} and decay orientation, roughly

$$\Delta R_{b\bar{b}} = \frac{2m_h}{p_{\mathrm{T}}(h)}, \tag{6.1}$$

if the two quarks have the same momenta [216].

To first order, the amount of transverse momentum carried by the Higgs boson is the same as the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$, often making the boson highly boosted and causing the final state topology to change when the amount of $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$, $E_{\mathrm{T}}^{\mathrm{miss}}$, reaches a few hundred GeV. The manner in which the pair of bottom quarks is reconstructed to form the Higgs boson leads to two categories of events. For dark matter signals producing low $E_{\mathrm{T}}^{\mathrm{miss}}$, the separation of the jets formed from the b quarks is large and can be reconstructed as two small-radius jets in the calorimeter, this category is called the resolved regime. For high $E_{\mathrm{T}}^{\mathrm{miss}}$ events, the decay products of the two b quarks are collimated to the extent that they can be reconstructed as a single jet with large radius [217], described in Section 4.2.2.2. This category is called the merged regime.

The main SM backgrounds with a signature of two b-jets and large $E_{\mathrm{T}}^{\mathrm{miss}}$ are the production of top-quark pairs, $t\bar{t}$, and of vector bosons, W and Z, with additional b-jets. This analysis uses data control regions to determine the normalisation of these backgrounds. The control region definitions are made orthogonal to the signal region by different requirements on the number of charged leptons in the events considered (1-lepton and 2-leptons). Sub-dominant backgrounds, as the Higgs boson production in association with a vector boson VH, are estimated purely based on MC simulation with no constraints from dedicated control samples, except for QCD multijet events, which is estimated in a data-driven way.

The search is performed considering four disjoint $E_{\mathrm{T}}^{\mathrm{miss}}$ categories, three in the resolved regime and one for the merged regime. A simultaneous profile-likelihood fit [218, 219] to the control and signal regions is performed to constrain the backgrounds and extract information about the potential presence of a signal. The final discriminant variable is the mass of the Higgs candidate reconstructed from the pair of b quarks.

As introduced before, two Higgs boson reconstruction techniques are considered in the analysis, related to the resolved and merge categories, and they are complementary in their acceptance. These two analysis channels are described here below in more detail.

6.1.1 Resolved Regime

The Higgs boson candidate is reconstructed from pairs of close by anti-k_{t} jets [141], each reconstructed with radius parameter $R = 0.4$, and identified as having a b-hadron within the jet, using a multivariate b-tagging algorithm in order to assess the flavour [220], as described in Section 4.2.3. See Figure 6.1 for a representative event display recorded

in 2015. This resolved technique offers good efficiency over a wide kinematic range with the Higgs boson transverse momentum between 150 and 500 GeV.

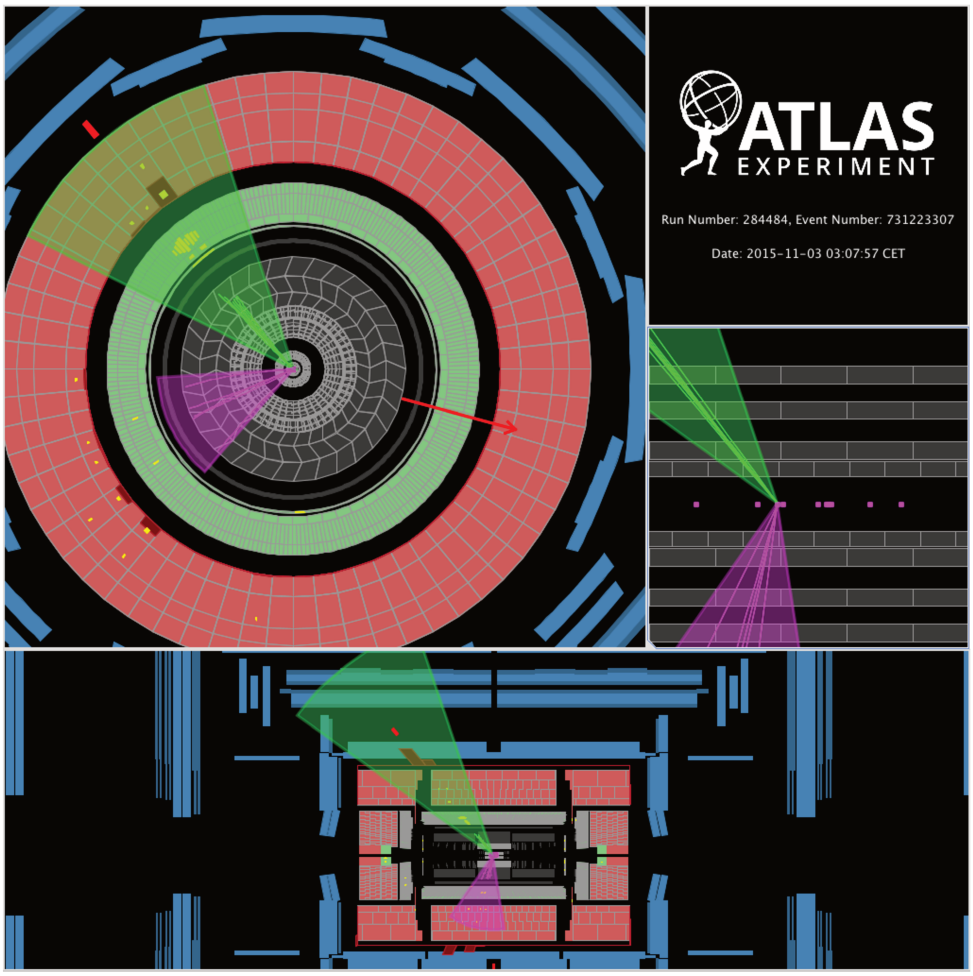


Figure 6.1 – An event display of a signal event in the resolved signal region. This event is characterized by $E_{\mathrm{T}}^{\mathrm{miss}} = 213$ GeV, red arrow, and two b-tagged small-R calorimeter jets, green and purple cones, that form a dijet system with with $m_{\mathrm{jj}} = 120$ GeV.

6.1.2 Merged Regime

At higher Higgs boson Lorentz boost, the average separation between the b- and $\bar{b}$-quarks from the Higgs boson decays, $\Delta R = \sqrt{(\Delta\eta)^2 + \Delta\phi)^2}$, reduces to values below 0.4. Therefore, for a Higgs boson with significant boost, the high momentum of this boson cause the two jets cones containing the b quarks to overlap. This behaviour leads to a decrease in the reconstruction efficiency of the two b-tagged Anti-kT jets with $R = 0.4$ parameter. Because of this, the Higgs is reconstructed with a single, trimmed [221], large radius Anti-kT jet of $R = 1.0$ with substructure [222], usually called large-R jet or fat-jet, see Section 4.2.2.2 for more details. Large-R jets are used in the analysis for events where

the total $E_{\mathrm{T}}^{\mathrm{miss}}$ is large: $E_{\mathrm{T}}^{\mathrm{miss}} > 500$ GeV. See Figure 6.2 for a representative event display recorded in 2015.

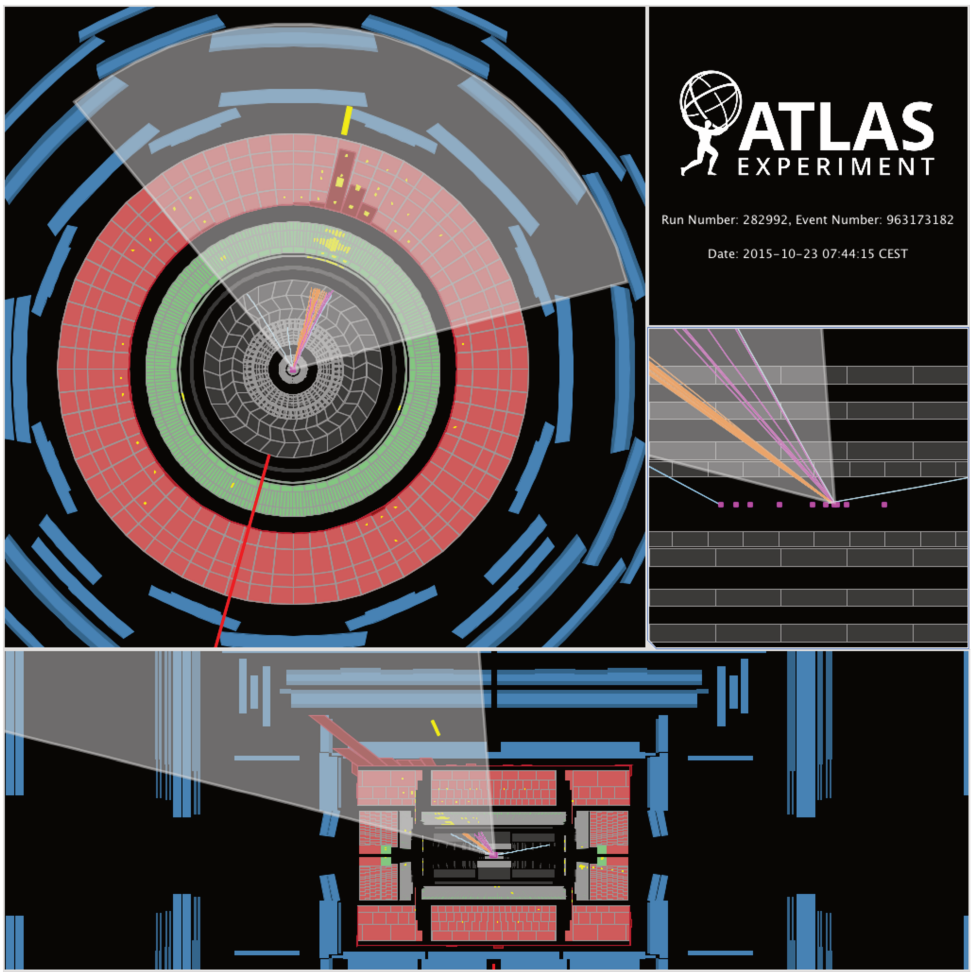


Figure 6.2 – An event display of a signal event in the merged signal region. This event is characterized by $E_{\mathrm{T}}^{\mathrm{miss}} = 694$ GeV , red line, and a large-R jet with $m_{\mathrm{J}} = 106$ GeV , white large cone, and two b-tagged track jets, orange and purple lines.

6.2 Physics objects selection

The analysis uses a wide variety of physics objects that are reconstructed off-line: electrons, muons, photons, jets which may be identified for the quarks flavor it contains, and $E_{\mathrm{T}}^{\mathrm{miss}}$. Object reconstruction efficiencies in simulated events are corrected to reproduce the performance measured in the data, and their systematic uncertainties are detailed in Section 7.2.

The identification criteria are designed to maximise the acceptance of objects as expected in the signal topology while reducing the contamination from background processes such as multijet production. The object selection criteria presented in this section and are based on studies done by SM $VH(bb)$ analysis group reported in [223], and similar notation is used. The physics objects reconstruction and calibration was already generally introduced in Section 4, therefore this section focuses on the selection criteria.

6.2.1 Small-R Jets

Small-radius (small-R) jets are used in the analysis to reconstruct the $h \to b\bar{b}$ candidate with two separated (resolved) jets associated to intermediate $E_{\mathrm{T}}^{\mathrm{miss}}$ ($150 \leq E_{\mathrm{T}}^{\mathrm{miss}}/\ \mathrm{GeV} < 500$) in the resolved regime. Also, small-R jets are used to reconstruct the hadronic activity in background and signal processes.

The constituents of the small-R jets are topoclusters calibrated at the electromagnetic scale using the anti-kt algorithm with an angular coverage of $\Delta R = 0.4$, see Section 4.2 for more details. A jet is categorized as either *central* or *forward* depending on its η value. The *central jets* are defined with $|\eta| < 2.5$ and $p_{\mathrm{T}} > 20$ GeV and are used to reconstruct the $h \to b\bar{b}$ candidate in the resolved regime. The jets in the remaining η range are called *forward jets* and must have $p_{\mathrm{T}} > 30$ GeV.

Beyond the default calibration applied to all jets described in Section 4.2, an additional correction to the full jet four vector is applied to improve the jet energy response to b-jets, including a correction applied when a muon is present within the jet [224]. This correction, however, is only applied to the b-jets that enter into the reconstruction of the dijet system.

For jets with $20 < p_{\mathrm{T}} < 60$ GeV and $|\eta| < 2.4$, jets coming from pileup interactions are suppressed by using tracking information via a selection on the jet vertex tagger observable (JVT) [152] to require JVT > 0.59, which corresponds to the medium working point. Furthermore, the standard overlap removal procedure is followed, which vetoes jets falling within a cone of radii of 0.2 of an electron [225].

The MV2c10 discriminant is used to identify jets containing a b hadron, with a 70% signal efficiency working point, see Section 4.2.3 for more details. These jets are referred to as *'b-jets'*.

Additional selection criteria, including "cleaning" requirements to remove fake jets and jets deposited through areas of the detector that are not functioning properly, are applied as detailed in Section 6.4.

6.2.2 Large-R Jets

Large-radius (large-R) jets, or fat-jets, are used in the Merged Regime analysis for events in which the $h \to b\bar{b}$ candidate has a high Lorentz boost. They are build from topo-clusters calibrated to the hadronic scale using the local cluster weighting (LCW) calibration and reconstructed using the anti-k_t algorithm with a distance parameter of $R = 1$. For more details, refer to Sections 4.2.2 and 4.2.2.2.

Large-R jets in the mono-$h(b\bar{b})$ analysis are required to have $p_T > 200$ GeV and $|\eta| < 2.0$ in order to select high-pT, central jets with a good overlap between the ID and the calorimeter, which is necessary for the derivation of systematic uncertainties.

Fat-jets originating from the decays of boosted Higgs bosons have some peculiar features that make them different from QCD-originated jets [156]. The two most powerful distinguishing features of the $h \to b\bar{b}$ jet are the jet mass and the flavor content, as it contains two b-hadrons originating from the two b-quarks. This is analogous to the identification of a resolved $h \to b\bar{b}$ decay, but the observables are modified to accommodate the boosted topology. To identify the flavor content of the fat-jet, the leading two track jets ghost-associated to the parent of the groomed fat-jet, described in Section 4.2.4, are used to decide if the jet is having either 0, 1, or 2 b-tags. This b-tagging is performed using the MV2c10 discriminant with a 70% b-tagging signal efficiency working point. The b-tagging algorithm is described in Section 4.2.3.

In addition, in an analogous way to the reconstruction of small-R jets, a correction is made to the four vector of the fat-jet in the case that muons are found to be within $\Delta R < 0.2$ of the two leading track jets, where at most two muons are used to correct the large-R jet four-vector.

6.2.3 Track Jets

The constituents of track jets are ID tracks with $p_T > 0.4$ GeVand $\eta < 2.5$. They are required to have at least seven hits in the silicon detectors, no more than one hit in the pixel detector that is shared by multiple tracks, no more than one missing expected hit in the pixel detector and no more than two missing expected hits in the silicon detectors. Finally, the track is either a constituent of the hard scatter primary vertex or has $|z_0 \cdot \sin\theta| < 3$ mm [1].

The identification of b-jets in dense environments is performed by forming jets from ID tracks using the anti-k_T algorithm with $R = 0.2$. This set of track jets is formed for each fat-jet, and only those that have $p_T > 10$ GeV and $|\eta| < 2.5$, and at least two ID tracks are used to identify b-jets with the MV2c10 algorithm.

6.2.4 Leptons

The selection of leptons is important for two main reasons. The first is for the veto of events containing leptons in the dark matter search, and the second is for the selection of events used in the control regions to constrain backgrounds. More details are given in Section 6.4 on the signal region event selection, and on the one and two lepton control regions in Section 6.5.

Different selection criteria for leptons are defined in the mono-$h(b\bar{b})$ analysis, associated to three different lepton channels: 0-lepton in the signal region, and 1- and 2-lepton control regions, which is the same categorisation as the one studied by the $Vh(b\bar{b})$ analysis [223]. In this Section, the selection criteria for these lepton categories is presented.

- Electrons
  Beyond the initial electron reconstruction and selection criteria discussed in Section 4.3, two categories of analysis specific electron categories, "VH-loose" and "VH-signal" electron, are defined depending on the analysis channel: signal region or control regions, defined in Sections 6.4 and 6.5, respectively. These different electron criteria are defined due to the different background composition of fakes in zero, one and two lepton selection and to ensure that each selection regions is mutually orthogonal via requirements on the number on VH-loose and VH-signal electrons.

  In the signal region (SR) and single muon control region (1μ-CR), VH-loose electrons are vetoed. They are identified using the *Loose* identification criterion [168, 226, 227], see Section 4.3, which allows for the highest selection efficiency for real electrons. Isolation requirements are also applied using the LooseTrackOnly working point to reduce jets faking electrons [228, 168]. This isolation working point targets a constant isolation efficiency of 99% varying just $p_{\mathrm{T}}^{\mathrm{cone0.2}}(E_{\mathrm{T}}, \eta)$, see Section 4.3.

  For the two lepton control region (2ℓ-CR), a tight electron definition allows to have a high purity electron selection in the dielectron channel. These electrons are referred to as VH-signal and has $p_{\mathrm{T}} > 27$ GeV in addition to VH-loose electron criteria.

  All electrons are required to be central, $|\eta| < 2.47$, and their impact parameters are required to be consistent with the primary vertex, to have d_0 significance less than 5 s.d. and $|z_0 \sin\theta| < 0.5$ mm to reject tracks from pile-up [229] [1].

  Table 6.1 summarize the selection criteria of the electron categories used in the analysis.

Electron Type	p_{T} [GeV]	$\lvert\eta\rvert$	Identification	Isolation
VH-loose	> 7 GeV	< 2.47	Loose	LooseTrackOnly
VH-signal	> 27 GeV	< 2.47	Loose	LooseTrackOnly

Table 6.1 – A summary of the electron object selection used in the analysis. The VH-loose definition is used in the 0 lepton and 1 muon channels to veto electrons. The VH-signal definition is used in the dielectron channel.

[1] The transverse impact parameter (d_0) is defined as the distance of closest approach in the $r-\phi$ plane of the track to the primary vertex while the longitudinal impact parameter ($|z_0 \sin\theta|$) is defined as the distance of the track to the primary vertex in the longitudinal plane at the point of closest approach in $r-\phi$.

- Muons
  By combining muon identification and isolation criteria described in Section 4.4, three analysis-level muon candidates are defined: "VH-loose", "WH-signal" and "ZH-signal", that are related to the three analysis channels: SR, 1μ-CR and 2ℓ-CR defined in Sections 6.4 and 6.5.

  In the SR, as for electrons, VH-loose muons are vetoed. This category is designed to reject muon as much as possible, thus are selected with loose identification criteria. Also, by using Standalone muons reconstructed without track, the coverage is extended to the forward region ($|\eta| < 2.7$) [170]. Loose track isolation (LooseTrackOnly) is applied to reduce jet-faking muon and is chosen to keep 99% efficiency constant in η and p_{T} [228], see Section 4.4.

  For the two lepton control region 2ℓ-CR, a tight muon definition allows to have a high purity in the dimuon channel. In this case, ZH-signal muon criteria requires muon object with $p_{\mathrm{T}} > 27$ GeV and $|\eta| < 2.5$ in addition to the VH-loose muon criteria.

  In the one lepton control region 1μ-CR, even tighter muon selections are required to suppress multijet background using the WH-signal selection with Medium identification criteria. A tighter track based isolation is also required, FixedCutTTTight, with $p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}}^{\mu} < 0.06$ [181].

  The impact parameters of the muon must be consistent with the primary vertex. It is required that the d_0 significance to be less than 3 s.d. and $|z_0 \cdot \sin\theta| < 0.5$ mm to reject muons from pile-up and cosmics [229].

  Table 6.2 summarize the selection criteria of the muon categories.

Muon Typ	p_{T} [GeV]	$\lvert\eta\rvert$	Identification	Isolation
VH-loose	> 7 GeV	< 2.7	Loose	LooseTrackOnly
ZH-signal	> 25 GeV	< 2.5	Loose	LooseTrackOnly
WH-signal	> 25 GeV	< 2.5	Medium	FixedCutTTTight

Table 6.2 – A summary of the muon object selection used in the analysis. The VH-loose definition is used in the 0 lepton channel to veto muons. The WH-signal and ZH-signal definitions are used in the 1 muon and dimuon channels, respectively.

- Tau Leptons
  In the signal region, a veto on the presence of τ leptons is employed in order to reduce $t\bar{t}$ and W + jets backgrounds. This veto is based on two tau definitions: one is referred as "standard tau" described in Section 4.5 with a Loose BDT identification, and another one is a custom-build definition called "extended tau".

  The extended tau category starts from standard small-R jets as described in Section 4.2.2.1 and requires that the jet contain one to four charged tracks. These tracks

are within a $\Delta R < 0.2$ from the jet axis and associated to the primary vertex. Unlike the standard taus that has exactly 1 or 3 tracks, the extended ones can account for underestimation of track multiplicity due to tracking inefficiencies or photon conversion tracks that pass the standard track selection criteria. Furthermore, due to the topology that this veto is intended to target (leptonic decays of W bosons), a requirement of $\Delta\phi(\mathbf{p}_{\mathrm{T}}(jet^{\tau}), \boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}})$<22.5° is required. This is because the tau candidate is expected to be azimuthally close to the missing transverse momentum for the $W \to \tau\nu$. The threshold value of this requirement is chosen by evaluating the signal gain of tau veto and the signal loss of dark matter models.

Tau Type	p_{T} [GeV]	$\lvert\eta\rvert$	Identification
Standard	> 20 GeV	< 2.5	Loose
Extended	> 20 GeV	< 2.5	$\Delta\phi(\mathbf{p}_{\mathrm{T}}(jet^{\tau}), \boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}})$<22.5°, $1 \leq N_{\mathrm{track}} \leq 4$

Table 6.3 – A summary of the tau object selection used in the analysis.

6.2.5 Overlap Removal

Ambiguities in the object identification which arise during reconstruction, i.e. when a reconstructed object matches multiple object hypotheses, are resolved in several steps that give different priority to the object types. First, electron candidates are removed if they share a track with another electron candidate that has a higher p_{T}. Next, τ candidates are removed if they are within $\Delta R = 0.2$ of an electron or muon candidate. Electron candidates are then rejected if they share a track with a muon candidate. In the next step, small-R jets are removed that lie within $\Delta R = 0.2$ of any remaining electron. Electrons are then removed if they are within $\Delta R = \min(0.4, 0.04 + 10\ \mathrm{GeV}/\mathrm{p}_{\mathrm{T}}^{\mathrm{electron}})$ of a small-R jet. Small-R jets are further removed if they are within $\Delta R = 0.2$ of any surviving muon if the jets have fewer than three tracks or if the muon p_{T} is greater than half the jet p_{T} and greater than 70% of the p_{T} sum of the tracks associated to the jet. Moreover, muons are removed if they are within $\Delta R = \min(0.4, 0.04 + 10\ \mathrm{GeV}/\mathrm{p}_{\mathrm{T}}^{\mathrm{muon}})$ of a small-R jet. Finally, large-R jets are rejected if they are separated from an electron by less than $\Delta R = 0.1$.

6.2.6 Muon in jet correction

There is at least a muon in 44% B-hadron semileptonic decays [223]. These muons are typically removed by the overlap removal procedure, however they influence the jet energy measurement. Contrarily to an electron which deposits energy in the calorimeter, the muon is a minimum ionising particle, depositing very little energy in the calorimeter and this contribution is not taken into account in the b-jet energy. Therefore, the 4-momentum of a jet with a muon inside is corrected by removing the 4-momentum of the

muon energy loss deposited in the calorimeters and then adding the 4-momentum of the reconstructed muon. If more than one muon is found, the closer muon to the jet is used. This procedure improves the energy scale and response resulting in an improved resolution of the Higgs mass candidate [164, 214]. Systematic uncertainties on this correction are found to be negligible.

Typically, muons inside jets have low p_{T} and are not isolated. Therefore, muons inside small-R jets are required to have $p_{\mathrm{T}} > 4$ GeV and medium identification quality in order to reduce fake muons. The correction is applied using the closest reconstructed muon within $\Delta R < 0.4$ of the reconstructed b-jet. In the case of large-R jets, the closest reconstructed muon within $\Delta R < 0.2$ of a b-tagged ghost-associated track jet with $p_{\mathrm{T}} > 10$ GeV is considered.

6.2.7 Missing Transverse Momentum $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ and track $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$

The transverse momentum imbalance $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ is calculated as the negative vector transverse momentum sum of the reconstructed and calibrated physics objects, and an additional soft term [230]. The soft term is built from ID tracks (refered as TST) that are not associated with any reconstructed electron, muon or jet, but which are associated with the primary vertex. For a detailed description of the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction, refer to Section 4.6.

The input objects to this calculation are:

1. Electrons : Calibrated and selected with the VH-loose selection
2. Muons : Calibrated and selected with the VH-loose selection
3. Small-R jets : Calibrated and selected as described before in Section 4.2.2.1.
4. ID tracks : Not attributed to any reconstructed object and selected with p_{T} >0.5 GeV

In addition to the $E_{\mathrm{T}}^{\mathrm{miss}}$ described above, a purely track-based missing transverse momentum vector, $p_{\mathrm{T}}^{\mathrm{miss}}$, is used as well. This track-based $p_{\mathrm{T}}^{\mathrm{miss}}$ observable is calculated as the negative vector sum of transverse momenta of reconstructed tracks associated with the primary vertex with $p_{\mathrm{T}} > 500$ MeV and $|\eta| < 2.5$. The primary usage of this is in the removal of beam-induced and non-collision background events [231].

Object	Kinematics	Type	Additional
Small-R Jets *central*	$p_{\mathrm{T}} > 20$ GeV $\lvert\eta\rvert < 2.5$	anti-k_t R=0.4 EMTopo	if $p_{\mathrm{T}} < 60$ GeV,$\lvert\eta\rvert < 2.4$ then JVT> 0.59 b-tagging: MV2c10 at 70% WP
Small-R Jets *forward*	$p_{\mathrm{T}} > 30$ GeV $2.5 \leq \lvert\eta\rvert < 4.5$	anti-k_t $R = 0.4$ EMTopo	
Large-R Jets	$p_{\mathrm{T}} > 200$ GeV $\lvert\eta\rvert < 2.0$	anti-k_t $R = 1.0$ LCWTopo Trimmed ($R_{\text{subjet}} = 0.2$, $f_{\text{cut}} = 5\%$)	
Track Jets	$p_{\mathrm{T}} > 10$ GeV $\lvert\eta\rvert < 2.5$	anti-k_t variable-R ID Tracks $n_{\text{tracks}} \geq 2$	b-tagging: MV2c10 at 70% WP Ghost-associated $p_{\mathrm{T}}^{track} > 0.4$ GeV, $\lvert\eta^{track}\rvert < 2.5$
τ leptons	$p_{\mathrm{T}} > 20$ GeV $\lvert\eta\rvert = [0, 1.37]\lvert\lvert[1.52, 2.5]$ $N_{\text{tracks}} = 1$ or 3	BDT ID Loose WP	Extended Tau Selection : Small-R jets with $1 \leq N_{\text{tracks}} \leq 4$ and $\Delta\phi(\tau, E_{\mathrm{T}}^{\text{miss}}) \leq 22.5^{\circ}$
$E_{\mathrm{T}}^{\text{miss}}$	$E_{\mathrm{T}}^{\text{miss}} \geq 150$ GeV	TST $E_{\mathrm{T}}^{\text{miss}}$ algorithm $p_{\mathrm{T}}^{\text{miss}} = -\Sigma p_{\mathrm{T}}^{\text{track}}$	Soft term: ID tracks matched to PV $p_{\mathrm{T}}^{\text{track}} > 0.5$ GeV,$\lvert\eta^{track}\rvert < 2.5$ matched to PV

Table 6.4 – A summary of the jet and $E_{\mathrm{T}}^{\text{miss}}$ object selections used in the analysis.

6.3 Data, trigger and simulation samples

This section describes the samples from data in Section 6.3.1, and simulated signal and background processes in Section 6.3.3. All samples were processed with CxAOD framework using ATLAS software off-line reconstruction release 20.7.

6.3.1 Data

This search uses proton-proton collision data recorded by the ATLAS detector at a centre-of-mass energy of 13 TeV, collected during 2015 and 2016 with a bunch crossing interval of 25 ns.

The LHC delivered 4.2 fb^{-1} in 2015 and 38.5 fb^{-1} in 2016 of integrated luminosity in stable beam conditions. For the ATLAS detector, the data quality efficiency was 93% and 92% in 2015 and 2016 respectively. This was due to inefficiencies in the data acquisition system, detector dead times and the ramping up of the ATLAS detector until being fully operating in stable beam condition. In total, ATLAS recorded 3.9 fb^{-1} and 35.6 fb^{-1} in 2015 and 2016 respectively. The mean number of proton-proton interactions per bunch crossing was 13.4 in 2015 and 25.1 in 2016.

Events are used in the analysis only when all relevant parts of the detector are working nominally, therefore all sub-detector components are required to be fully operational. Only data passing quality criteria are considered, and the corresponding runs are listed in the so-called Good-Run lists (GRL), see Section C.1 for details. The resulting dataset used in the analysis corresponds to an integrated luminosity of 36.5 fb^{-1}, of which 3.2 fb^{-1} and 32.9 fb^{-1} were collected in 2015 and 2016 respectively. In the analysis update presented in Chapter 8 data collected during 2017 is included.

6.3.2 Trigger

The analysis signal region uses a $E_{\text{T}}^{\text{miss}}$ trigger in order to select fully hadronic states. Different unprescaled $E_{\text{T}}^{\text{miss}}$ triggers are used, taking into account the increasing instantaneous luminosity in 2015 and in the 2016 periods, adjusting the nominal event rate for the mass storage. The different $E_{\text{T}}^{\text{miss}}$ trigger items considered are:

- 2015 data (3.2 fb^{-1}): HLT_xe70, L1_XE50.
- 2016 data for periods A-D3 (6.1 fb^{-1}): HLT_xe90_mht_L1XE50
- 2016 data for periods D4-E (3.9 fb^{-1}): HLT_xe100_mht_L1XE50

  or HLT_xe110_mht_L1XE50
- 2016 data from period F1 (23.2 fb^{-1}): HLT_xe110_mht_L1XE50

Where the value after "xe" represents the online missing energy thresholds considered for HLT and L1. For example, "L1XE50" indicates a seed corresponding to a Level 1

requirement on $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV. "mht" refers to the algorithm that reconstructs the $E_{\mathrm{T}}^{\mathrm{miss}}$, which is defined as the transverse momentum (p_{T}) vectorial sum of all jets reconstructed from calorimeter topological clusters at the HLT with $p_{\mathrm{T}} > 7$ GeV [232].

6.3.2.1 Data-driven $E_{\mathrm{T}}^{\mathrm{miss}}$ Trigger Scale Factors

The analysis requires a lower threshold in the event selection of $E_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV, see Section 6.4, for which the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers are not fully efficient. This part of the phase space is included to further improve the sensitivity of the analysis. Therefore, it is necessary to correct the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiencies curves in MC, which are not perfectly modelled, with scale factors to be consistent with data.

These trigger efficiencies are calculated in a single muon region for both data and MC, and scale factors are calculated to correct turn-ons in MC to those in data. This 1-muon control region, defined in Section 6.4, is orthogonal to the 0 lepton signal region where the scale factors will be applied.

The $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstructed in the HLT trigger only uses the calorimeter system, therefore, muons are not considered. Because of this, the $E_{\mathrm{T}}^{\mathrm{miss}}$ for HLT for one muon events corresponds to an offline $E_{\mathrm{T}}^{\mathrm{miss}}$ after subtracting the muon contribution, this modified $E_{\mathrm{T}}^{\mathrm{miss}}$ is called $E_{\mathrm{T,no}\mu}^{\mathrm{miss}}$, so it would be kinematically similar to the HLT $E_{\mathrm{T}}^{\mathrm{miss}}$. The offline event selection is consistent with the resolved signal region described in Section 6.4, with the use of muon triggers (see Table C.1 for details) and no low $E_{\mathrm{T}}^{\mathrm{miss}}$ threshold.

The $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency, $\varepsilon_{\mathrm{Trig}}$, is defined as:

$$\varepsilon_{\mathrm{Trig}} = \frac{\#\text{Events passed selection AND } E_{\mathrm{T}}^{\mathrm{miss}} \text{ trigger requirement}}{\#\text{Events passed selection}} \tag{6.2}$$

The efficiency for each trigger corresponding to each data period was studied as function of $E_{\mathrm{T,no}\mu}^{\mathrm{miss}}$, and this is shown in Figure 6.3. Good data and MC agreement is observed, and any residual difference is corrected by a scale factor defined as:

$$\mathrm{SF} = \frac{\mathrm{Efficiency}_{\mu}^{\mathrm{Data}}}{\mathrm{Efficiency}_{\mu}^{\mathrm{MC}}} \tag{6.3}$$

In order to calculate the data-driven corrections for the MC $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers turn-ons, the scale factors are fitted for each $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger in the range 120 GeV $< \mathrm{E}_{\mathrm{T,no}}^{\mathrm{miss}} <$ 300 GeV with the following function:

$$f\left(E_{\mathrm{T,no}\mu}^{\mathrm{miss}}\right) = 0.5 \cdot \left[1 + Erf\left(\frac{E_{\mathrm{T,no}\mu}^{\mathrm{miss}} - p_0}{\sqrt{2}p_1}\right)\right] \tag{6.4}$$

6.3.3 Monte Carlo Simulated Samples

Monte Carlo (MC) simulated samples are produced for both the signal and background processes, using the full GEANT 4-based ATLAS simulation [194]. A summary of the MC samples is provided in Table 6.5, giving explicit information on the MC generators,

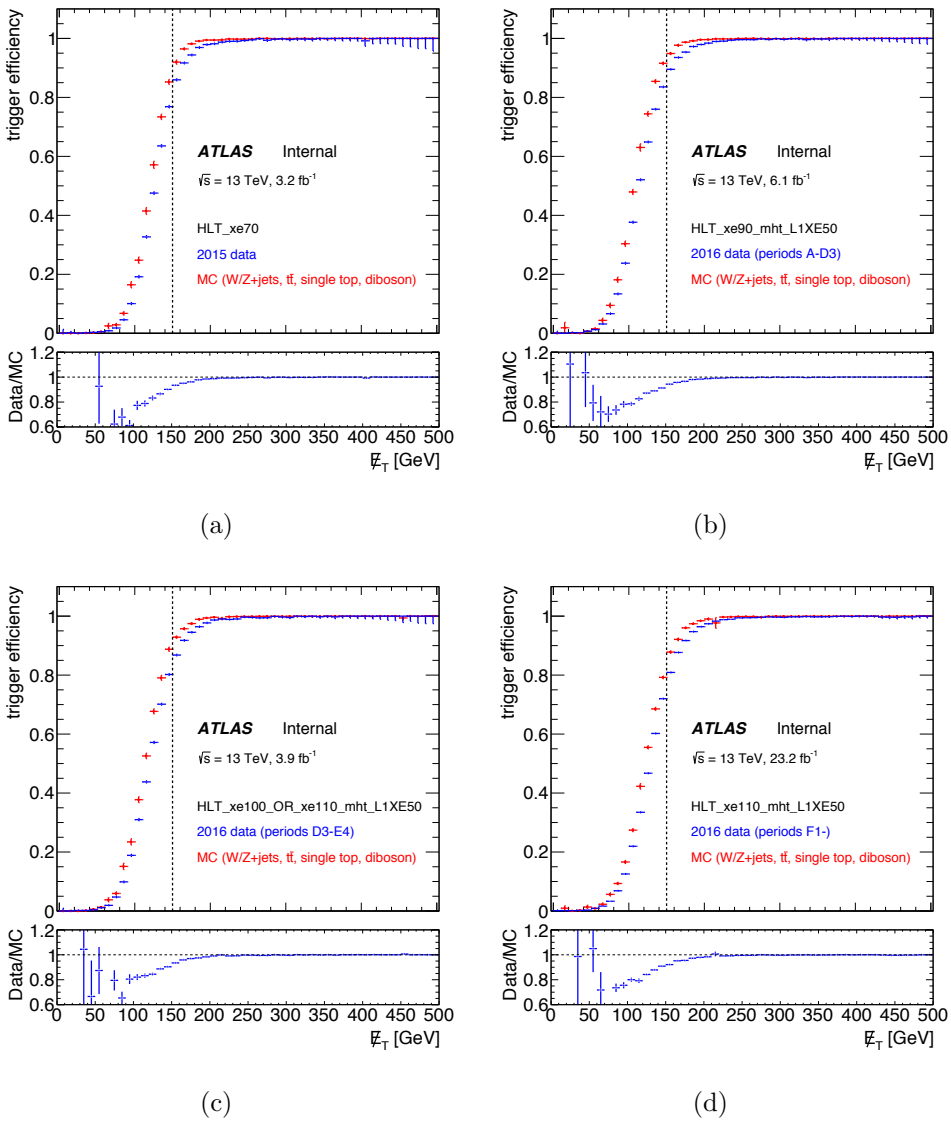


Figure 6.3 – Measured $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiencies and scale factors as function of offline $E_{\mathrm{T,no}\mu}^{\mathrm{miss}}$ in data and Monte-Carlo simulation for 6.3(a) HLT_xe70, 6.3(b) HLT_xe90_mht_L1XE50, and 6.3(c) HLT_xe100_mht_L1XE50 OR HLT_xe110_mht_L1XE50, and 6.3(d) HLT_xe110_mht_L1XE50 triggers are shown. The plots are shown for 0,1 and 2 tags together. The MC is dominated by W+jet and $t\bar{t}$ events [233].

parton distribution functions (PDFs) and the production cross-sections. More extensive descriptions is provided in Sections 6.3.3.1 and 6.3.3.2.

6.3.3.1 Signal Monte Carlo Samples

Mono-H signals events from the Z'-2HDM simplified model, see Section 2.2.2, are generated using MadGraph5 at tree-level in QCD using NNPDF3.0LO parton distribu-

Process	Generator	PDFs [234]	σ_{norm} [pb]
$t\bar{t}$	Powheg + Pythia 6 (NLO)	CT10	831.76
V+jets			
$W \to \ell\nu$+jets	Sherpa 2.2.1 (NNLO)	NNPDF3.0NNLO	20080
$Z \to \ell\ell, \nu\nu$+jets	Sherpa 2.2.1 (NNLO)	NNPDF3.0NNLO	1906 ($66 < m_{ll} < 116$ GeV)
Single top			
t-channel	Powheg + Pythia 6 (NLO)	CT10	136.02+80.95
s-channel	Powheg + Pythia 6 (NLO)	CT10	6.35+3.97
Wt-channel	Powheg + Pythia 6 (NLO)	CT10	71.7
Diboson			
$Wl\nu Wqq$	Sherpa 2.1 (NLO)	CT10	22.65
$Wl\nu Zqq$	Sherpa 2.1 (NLO)	CT10	10.47
$WqqZll$	Sherpa 2.1 (NLO)	CT10	3.12
$WqqZ\nu\nu$	Sherpa 2.1 (NLO)	CT10	6.17
$ZqqZ\ell\ell$	Sherpa 2.1 (NLO)	CT10	2.15
$ZqqZ\nu\nu$	Sherpa 2.1 (NLO)	CT10	4.22
SM $qq \to VH(\to b\bar{b})$	Pythia 8 (LO)	NNPDF2.3LO	0.7639
SM $gg \to VH(\to b\bar{b})$	Powheg + Pythia 8 (LO)	CT10	0.1057
mono-H signals	MadGraph + Pythia 8	NNPDF30_lo_as_0130	-

Table 6.5 – List of MC generators, parton distribution functions (PDFs) and production cross-section used for the signal and background processes.

tion function (PDF) sets [235]. Parton showering and hadronization are simulated with Pythia 8 [196] generator with the A14 tune [236] set and using the NNPDF2.3LO PDF set [237].

Following the notation described in Section 2.2.2, the fixed parameters used are: $\tan\beta$=1.0, g_Z=0.8, and m_χ=100 GeV. The mass of the Z' boson is scanned from 400 to 3000 GeV in 200 GeV steps, and the mass of the pseudo-scalar A is scanned from 200 to 800 GeV in 100 GeV steps. These are the relevant regions of the parameter space where the analysis is sensitive with the present dataset.

All simulated samples include the effect of multiple proton-proton interactions in the same or neighbouring bunch crossings, i.e. pile-up. Simulated events are corrected using per-event weights to describe the distribution of the average number of primary vertices as observed in data.

The cross sections for the different considered signal points are listed in Table C.3 of Section C.2.

6.3.3.2 Background Monte Carlo Samples

The main source of background in the mono-$h(b\bar{b})$ analysis signal region is the V+jets, with $V = W$ or Z, and top-pair production. V+jets events with massive c- and b-quarks are generated with Sherpa 2.2.1 event generator [199] at leading order in QCD, interfaced with the NNPDF3.0NNLO PDF set [237]. They are normalised to cross-sections calculated at next-to-next-to-leading order (NNLO) in QCD. In order to obtain a good

statistical size in regions with heavy flavor production or boosted vector bosons, dedicated filters are used: one filter allows to select events containing light, charm or bottom flavored hadrons, a second filter selects event in regions of the p_{T} of the vector boson V. Top-pair production, $t\bar{t}$, is generated using POWHEG, and PYTHIA 6 [238] is used to simulated parton showering and hadronization. This sample is then normalised using cross-sections calculated NNLO plus next-to-next-to-leading logarithms (NNLL) introducing soft gluon radiation. The diboson (WW, WZ and ZZ) samples are produced with Sherpa 2.1 interfaced with CT10 NLO PDFs and normalised to cross-sections calculated at NLO. PYTHIA 6 is used to simulate the SM VH process, both in the W and Z channels. The single top-quark events are produced using the POWHEG generator. Cross-sections of the three single top-quark processes are based on NLO calculations [239].

All simulated samples include the effect of in-time and out-of-time pile-up from multiple interactions on the same and neighbouring bunch crossing. Pile-up samples are simulated with the *minimum bias events* obtained from soft QCD processes using PYTHIA 8 with A2 tune, and interfaced with MSTW2008LO PDFs. The amount of pile-up in MC simulations is described by the average number of interaction per bunch crossing of the simulated minimum bias events which is reweighted in order to reproduce the observed distribution on data.

6.4 Event selection

This section presents a detailed description of the selection criteria used in the $mono-h(b\bar{b})$ analysis to reject background and enhance signal sensitivity.

6.4.1 Event Preselection

The selections discussed in this section are applied to events in all signal and control regions to ensure high-quality data.

- GRL and Event Cleaning: Only runs with all sub-detectors nominally working are considered. Events with LAr, Tile, SCT, or Core error bits set due to noise bursts or data corruption are vetoed. See Sections 6.3.1 and C.1. Also, data events for which the data quality requirements indicate possible problem with one or more of the ATLAS sub-detectors are vetoed.

- Vertex Selection: At least one reconstructed vertex with at least two associated tracks is required, see Section 4.1.

- Jet cleaning: Veto events if any jet fails the loose requirement of the jet cleaning algorithm, *BadLoose* [153], to ensure a good measurement of $E_{\mathrm{T}}^{\mathrm{miss}}$. See Section 4.2.2.1 for details.

6.4.2 Signal Region Event Selection

The signal is characterized by high $E_{\mathrm{T}}^{\mathrm{miss}}$, no isolated leptons, and an invariant mass of the Higgs boson h_{jj} candidate m_{jj} compatible with the observed Higgs boson mass. The signal region (SR) event selection is designed to enrich this topology and phase space, while reducing features from background processes that could mimic the signal. In the event selection, only objects passing the selection criteria described in Section 6.2 are considered.

6.4.2.1 Baseline Selection

The SR is defined to have $E_{\mathrm{T}}^{\mathrm{miss}}$ >150 GeV and veto in VH-loose electrons or muons. In order to account for changes in the background composition and to benefit from a higher signal sensitivity with increasing $E_{\mathrm{T}}^{\mathrm{miss}}$ and b-tag multiplicity, the events are split into eight orthogonal regions based on the b-tag multiplicities (1 and 2 b-tag) and on the $E_{\mathrm{T}}^{\mathrm{miss}}$ in the event (4 categories).

Also, $p_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV is required only in the 1 b-tag categories (see below for more information about the b-jet categories). This requirement is designed to remove non-collision background events [231]. This can be understood as coming from the fact that beam induced backgrounds do not have, inner detector tracks, whereas real collision events

with real $E_{\mathrm{T}}^{\mathrm{miss}}$ do have ID tracks. In the 2 b-tag category, where the non-collisional background contribution is already quite low due to the requirement of 2 b-tagged jets, no such requirement is applied in order to recover signal efficiency.

After these selection, there are still a large number of multijet events in which one of the jets is poorly measured, causing an energy imbalance and therefore fake $E_{\mathrm{T}}^{\mathrm{miss}}$. Two primary observables described below, referred to as "anti-QCD cuts" are designed to remove such events:

- Minimum $\Delta\Phi(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}, \mathbf{p}_{\mathrm{T}}(jets_{1,2,3})) > 20°$
  In order to veto events with mismeasured jet momenta, the smallest azimuthal angle between $E_{\mathrm{T}}^{\mathrm{miss}}$ and the three first small-R jets is required to be greater than 20°. These $jets_{1,2,3}$ belong to a set of *Central+Forward* jets. This set of jets is formed starting with the *Central* jets, followed by the *Forward* jets, both described in Section 6.2.1. Within each category, jets are ordered in decreasing transverse momentum. Only the first three jets of the full set of *Central+Forward* jets are considered in this calculation. Therefore, $\mathbf{p}_{\mathrm{T}}(jets_{1,2,3})$ refers to the momentum $\mathbf{p}_{\mathrm{T}}$ of the first, second or third sorted jet.

  In the case of a multijet event in which one jet is mis-measured, it will create fake $E_{\mathrm{T}}^{\mathrm{miss}}$ in the direction of the jet, making the total reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ to point closer to the mismeasured jet direction. Therefore, multijet events are expected to reside at low values of this observable, for which any of the three leading jets is angularly close to the $E_{\mathrm{T}}^{\mathrm{miss}}$.

- $\Delta\Phi\ (\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}, \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}) < 90°$
  In the case of an event with real $E_{\mathrm{T}}^{\mathrm{miss}}$ produced by unobserved particles, $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ and $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$ should be aligned, see Section 6.2.7 for details on the reconstruction of these observables. However, in the case of a dijet event with a mis-measured jet, the $E_{\mathrm{T}}^{\mathrm{miss}}$ will align with one of the jets, and the $p_{\mathrm{T}}^{\mathrm{miss}}$ will be small and without a preferred direction.

Following these baseline selections, events are divided into the resolved and merged regions using a single selection on $E_{\mathrm{T}}^{\mathrm{miss}}$, as introduced in Sections 6.1.1 and 6.1.2. For events with $E_{\mathrm{T}}^{\mathrm{miss}} < 500$ GeV, events are considered to be in the resolved regime, while for $E_{\mathrm{T}}^{\mathrm{miss}} > 500$ GeV, events are considered to be in the merged regime. The selection criteria in both regimes are discussed in the following.

6.4.2.2 Resolved Regime Selection

The final state in resolved regime is characterized by moderate $E_{\mathrm{T}}^{\mathrm{miss}}$ and a Higgs reconstructed with two small-R jets associated to the Higgs decaying into bottom quarks, as described in Sections 6.1.1. Therefore, only events with $E_{\mathrm{T}}^{\mathrm{miss}} < 500$ GeV are selected in order to ensure that the $h \to b\bar{b}$ is reconstructed with two small-R jets.

In order to asses the flavor of the hadronic decay of the Higgs boson, b-tagging is used, see Section 4.2.3 for details on this technique. The small-R jets are first sorted by central

b-tagged jets and then by central non-b-tagged jets, and finally ordered in decreasing transverse momentum. The first two jets in this sequence are used to reconstruct the Higgs boson decay, $h \to b\bar{b}$. Therefore, it is required a minimum jet multiplicity of $N_{\text{central jets}} \geq 2$ in order to reconstruct the Higgs boson. Thus, it may be a bb, bj, or jj candidate to reconstruct the Higgs boson in the case of 2, 1, or 0 b-tagged jets, respectively. Events without b-jets, 0 b-tag, are discarded.

Also, in the resolved regime, the events are split in three $E_{\mathrm{T}}^{\mathrm{miss}}$ categories: 150 GeV $\leq$ $\mathrm{E}_{\mathrm{T}}^{\mathrm{miss}} < 200$ GeV, 200 GeV $\leq \mathrm{E}_{\mathrm{T}}^{\mathrm{miss}} < 350$ GeV and 350 GeV $\leq \mathrm{E}_{\mathrm{T}}^{\mathrm{miss}} < 500$ GeV.

Further requirements are applied on the scalar sum of the p_{T} for up to the three leading jets, using the *Central+Forward jets* collection. For 2-jet events, Σp_{T} must be larger than 120 GeV and for 3-jet events, Σp_{T} must be larger than 150 GeV. This criterion is designed to remove a region of phase space which is mismodelled in simulation due to a non-trivial dependence of the trigger efficiency on the jet activity. In addition, the leading jet from the central jets, which is used to reconstruct the $h \to b\bar{b}$ candidate, is required have a transverse momentum of above 45 GeV.

On top of the common anti-QCD cuts for both regimes described before, two additional cuts are applied to reduce the amount of multijet events. These cuts are specific to the resolved regime, because they account for the decay topology of the Higgs boson with two well separated small-R jets, which is different for background events. Also, since the $E_{\mathrm{T}}^{\mathrm{miss}}$ spectrum of the multijet background is exponentially decreasing, higher multijet contamination is expected in the resolved regime. These two additional anti-QCD for the resolved regime are:

- $\Delta\Phi(\mathbf{p}_{\mathrm{T}}(j_1), \mathbf{p}_{\mathrm{T}}(j_2)) < 140^{\circ}$
  Multi-jet background often has a back to back topology unlike the signal in which the di-jet system which reconstruct the Higgs boson is more boosted leading to a pair of collimated jets. Therefore, it is required that the two leading small-R jets to be azimuthally close to each other.

- $\Delta\Phi(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}},\ \mathbf{p}_{\mathrm{T}}(h_{jj})\ ) > 120^{\circ}$
  In the final state, $h_{jj} + E_{\mathrm{T}}^{\mathrm{miss}}$, it is expected to have $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ recoiling against the reconstructed Higgs boson and this is not necessarily true for multi-jet events.

The values chosen for these selections are based on examining the spectra of these observables and removing the region in which there is a data excess over the electroweak backgrounds attributed to multijet background.

The pre-fit [2] distributions of data and MC backgrounds of the observables that helps to reject multi-jet background are shown in Figure 6.4. The distributions showed are inclusive in $E_{\mathrm{T}}^{\mathrm{miss}}$ categories with a pre-selection threshold of $E_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV.

The plots on the left of Figure 6.4 show the data and MC comparison before the selections of any of these observables are made, for which the multijet contribution is

[2] Until Section 7.3, all comparisons of data to MC simulation are made prior to any combined likelihood fit that changes the normalizations of the various background components.

sizable. The multi-jet simulation in these plots corresponds to a di-jet MC generated with PYTHIA8 using the A14 set of tuned parameters and the NNPDF2.3 LO PDF. However, these simulated events does not completely account for all the QCD processes involved in the multij-jet background. Therefore, a purely MC simulation is used to demonstrate the generic shape of the expected multi-jet being consistent with the region where the non-multijet background alone would not describe the data. Because of this, data is expected to exceeds the background simulation due to the absence of a complete multi-jet simulation/estimation. Further details on the multi-jet background estimation will be given in Section 6.5.4.

The plots on the right of Figure 6.4 shows the same distributions after applying these selection requirements on the signal region for $E_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV. It can clearly be seen that the combination of these requirements ($\min(\Delta\Phi(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}, \mathbf{p}_{\mathrm{T}}(jets_{1,2,3}))) > 20°$, $\Delta\Phi\,(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}, \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}) < 90°$, $\Delta\Phi(\mathbf{p}_{\mathrm{T}}(j_1), \mathbf{p}_{\mathrm{T}}(j_2)) < 140°$ and $\Delta\Phi(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}, \mathbf{p}_{\mathrm{T}}(h_{jj})) > 120°$) remove the regions either where the disagreement between data and MC is more prominent and/or where the multi-jet contribution is more important.

6.4.2.3 Merged Regime

In this regime, the leading large-R jet represents the Higgs candidate, h_{jj}. Therefore, at least one fat jet is needed in order to reconstruct h_{jj}, i.e. $N_{\mathrm{large-R\,jets}} \geq 1$, and having at least one ghost-associated track jet. And since the amount of Higgs boson Lorentz boost is correlated to the $E_{\mathrm{T}}^{\mathrm{miss}}$, this events are selected by increasing the $E_{\mathrm{T}}^{\mathrm{miss}}$ threshold to 500 GeV. This selection ensures the set of events to be orthogonal to the resolved region and furthermore reduce the contamination of multijet events to a negligible level. The events are divided in exclusive categories based on the number of b-tagged track-jets. The two highest p_{T} track-jets ghost-associated to the large-R jet are used to assess the flavor content of the Higgs candidate. When both, one or none of them are b-tagged, the event is assign to the 0, 1 or 2 b-tag category, respectively. Events without b-tagged track-jets, 0 b-tag, are discarded. See Section 4.2.3 for details on the b-tagging.

6.4.2.4 Optimised Event Selection

In the previous result of the $mono-h(b\bar{b})$ analysis using 3.2 fb^{-1} of collision data at $\sqrt{13}$ TeV, top pair production, $t\bar{t}$, was the dominant background process [1]. It represent about 80% and 40% of the total background in the resolved and merged regime respectively in the 2 b-tag, 0 lepton region (SR), where most of the sensitivity is expected (see table 2 of Reference [1]). It is therefore crucial to develop selections aiming the reduction of the $t\bar{t}$ background process in order to improve the sensitivity of the search. The selections described below provides an additional background reduction up to 60% relative to previous result, for a small signal lost.

- Veto on taus
  Signal events are not expected to contain any tau leptons in the final state, whereas in background processes for which a quark top decays as $t \to Wb \to \tau\nu b$ the event

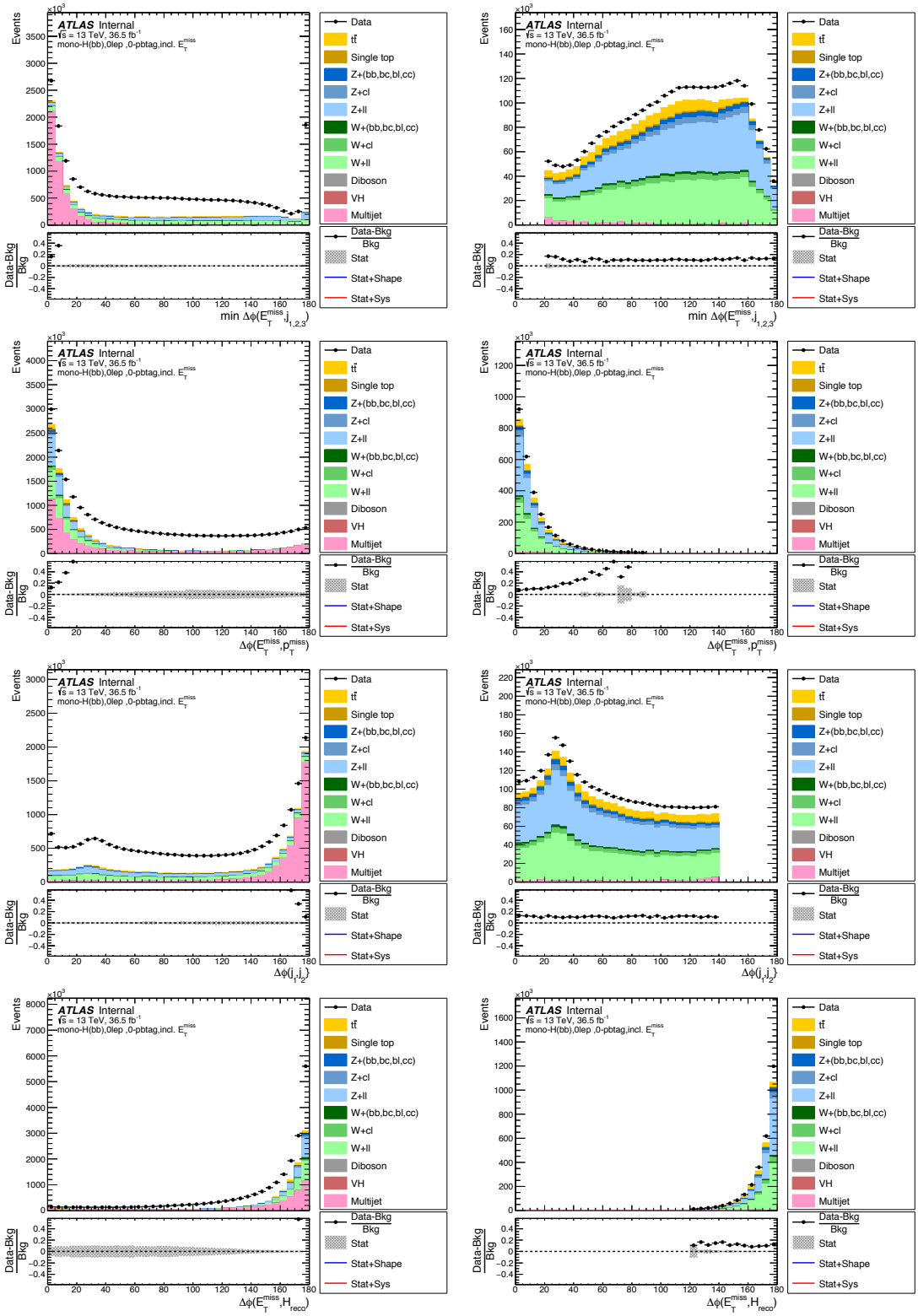


Figure 6.4 – Distributions of variables used to reduce and study the multijet contribution in the signal region. These are the minimum $\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}}, jets_{1,2,3})$, $\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}})$, $\Delta\Phi(j_1, j_2)$ and $\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}}, h_{jj})$ (labeled as $\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}}, \mathrm{H_{reco}})$ in the axis). On the left column are shown the distributions before any anti-QCD cut is applied. While on the right column, the distributions shown are after application of the anti-QCD selection cuts [233].

can be rejected with a tau veto. The veto is performed for hadronic τ-lepton candidate, identified either by an algorithm based on a boosted decision tree [240] (*Standard tau*) or as small-R jets containing one to four tracks within the jet cone and $\Delta\phi(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}, \mathbf{p}_{\mathrm{T}}(\mathrm{jet}))<\pi/8$ (*Extended tau*). See Table 6.3 for details on these tau selections. ranges between 2% and 7%.

- Veto of additional b-jets
  Events with more than two b-tagged central jets are rejected, which typically happens for $t\bar{t}$ events.

- H_{T} ratio requirement
  The most of the hadronic activity in a signal event is expected from the $h \to b\bar{b}$ decay. Because of this, it is required that a certain fraction of the $H_{\mathrm{T}}^{\mathrm{all\,jets}}$ to come from the h_{jj}, where $H_{\mathrm{T}}^{\mathrm{all\,jets}}$ is the scalar sum od the p_{T} of all the jets present in the event. It is required that the scalar sum of the p_{T} of the two jets forming the Higgs candidate h_{jj} and, if present, the highest-p_{T} additional jet must be larger than $0.63 \times H_{\mathrm{T}}^{\mathrm{all\,jets}}$.

- b-tagged jets separation
  The two b-tagged jets forming the Higgs candidate, h_{jj}, are expected to be geometrically close, whereas the b-jets from $t\bar{t}$ events have in average a larger angular separation. Therefore it is required that $\Delta R(\mathbf{p}_{\mathrm{T}}(j_1), \mathbf{p}_{\mathrm{T}}(j_2)) < 1.8$

Tu summarize, the 0-lepton signal region, SR, have eight event categories in total. Two categories related to the b-tag multiplicity: 1 and 2 b-tag. Four categories of missing transverse momentum: three regions in the resolved regime and one for the merge regime. The event selection criteria in the signal region for the resolved and merged region is summarized in Table 6.6.

6.4.2.5 Event Selection Signal Efficiency

Figures 6.5 and 6.6 show the cutflows for three Z'-2HDM simplified model mass points. Every bin shows the number of events selected after all requirements up to the one indicated by the bin label. The last three bins show the b-tag multiplicity distributions after all requirements. These efficiencies refer to the total number of events for a particular model after a primary vertex requirement. The latter has a high efficiency. Figure 6.7 shows the acceptance times efficiency for the different model mass points. The dashed line indicates the transition between the off-shell and on-shell regions. In the on-shell region the efficiency increases as the Z' mass increases for a given A mass. This increase is due to the requirement of $E_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV. Only for a finite gap between the Z' and A masses there is large $E_{\mathrm{T}}^{\mathrm{miss}}$.

Resolved	**Merged**	String Reference
lowest unprescaled $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger		trigger
veto on loose leptons		lepton
$E_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV		met_gt150
$P_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV only for 0 or 1 b-tag		mpt_gt30
min $\Delta\Phi(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}, \mathbf{p}_{\mathrm{T}}(jets_{1,2,3})) > 20°$		mindphi_gt20
$\Delta\Phi\ (\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}, \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}) < 90°$		dphimetmpt_lt90
$E_{\mathrm{T}}^{\mathrm{miss}} < 500$ GeV	$E_{\mathrm{T}}^{\mathrm{miss}} > 500$ GeV	met_lt500, met_gt500
N(central small-R jets) ≥ 2	N(central large-R jets) ≥ 1	njetsCentral_gt2, nfatjet_gt1
$p_{\mathrm{T}}(j_1) > 45$ || $p_{\mathrm{T}}(j_2) > 45$	—	j0orj1_pt45
$\sum_{i=1}^{2(3)} p_{\mathrm{T}}(j_i)$ GeV	—	jets012_pt120or150
$\Delta\Phi(\mathbf{p}_{\mathrm{T}}(j_1), \mathbf{p}_{\mathrm{T}}(j_2)) < 140°$	—	dphijj_lt140
$\Delta\Phi(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}, \mathbf{p}_{\mathrm{T}}(h_{jj})\ ) > 120°$	—	dphimetjj_gt120
τ-veto	τ-veto	taus
b-jet	b-jet	bJetVeto
H_T ratio requirement	H_T ratio requirement	HTcut
$\Delta R(\mathbf{p}_{\mathrm{T}}(j_1), \mathbf{p}_{\mathrm{T}}(j_2)) < 1.8$	—	dRjj_lt18
b-tag requirement on small-R jets	track-jets	Nbtags_2pjet, Nbtags_0pjet

Table 6.6 – Summary of the resolved and merged event selection applied in the 0 lepton channel.

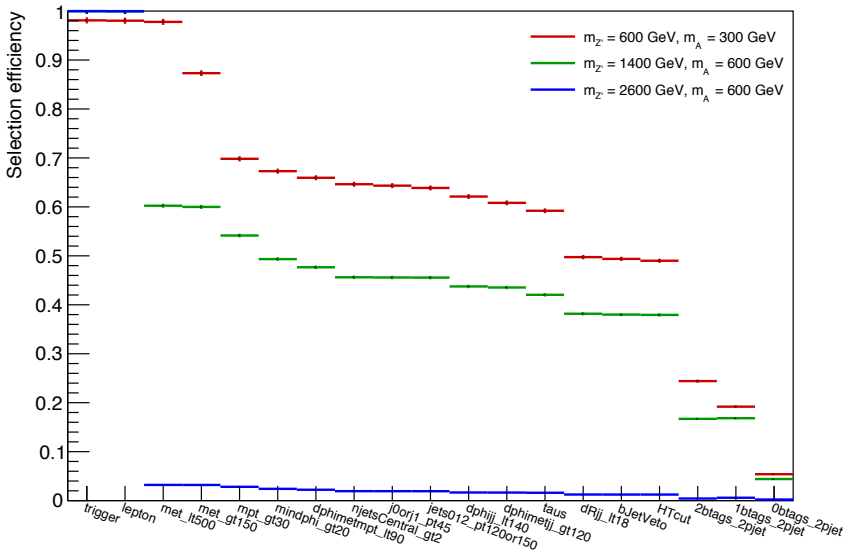


Figure 6.5 – Cutflow for three representative simplified model mass points in case of the resolved event selection [233].

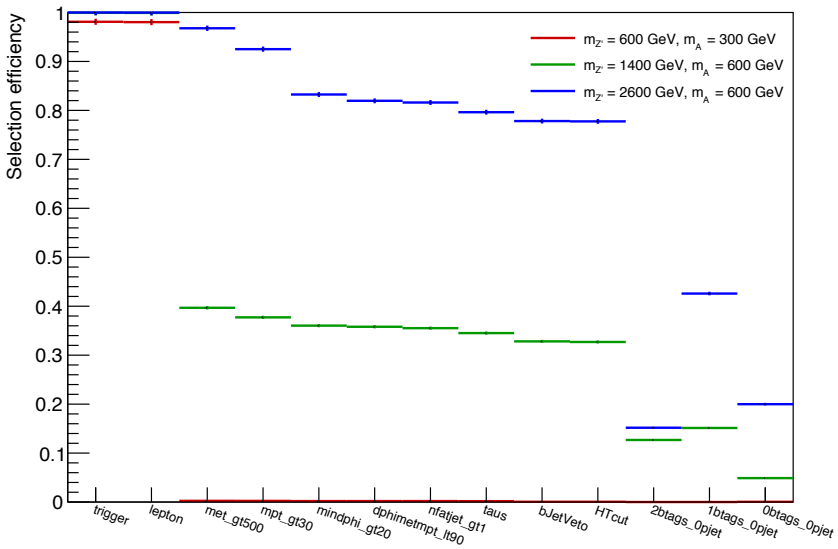


Figure 6.6 – Cutflow for three representative simplified model mass points in case of the merged event selection [233].

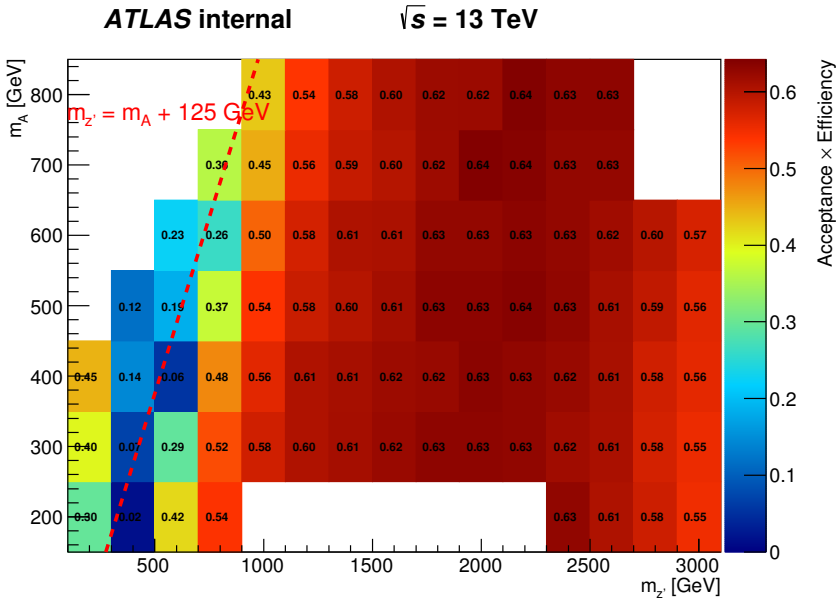


Figure 6.7 – Acceptance times efficiency for the different simplified model mass parameters [233].

6.5 Backgrounds

6.5.1 Background processes

The Signal Region (SR) described in Section 6.4 is characterized by high $E_{\mathrm{T}}^{\mathrm{miss}}$, no isolated leptons, and two jets (small-R or track jets) that reconstructs the SM Higgs boson. The dominant backgrounds, which account for more than 90% of the total background in the SR, are the following:

- $Z(\nu\nu)+\mathrm{jets}$
  This is the dominant background in the SR which contributes $30\%-60\%$ of the total background depending on the $E_{\mathrm{T}}^{\mathrm{miss}}$ and b-tag category. The tree-level Feynman diagrams for $Z+\mathrm{jets}$ with b-jets are shown in Figure 6.8. $Z+\mathrm{jets}$ can have real $E_{\mathrm{T}}^{\mathrm{miss}}$ from the Z boson decay to neutrinos, $Z \to \nu\nu$, one or two b-tagged jets, and no leptons. $Z(\nu\nu)+\mathrm{jets}$ is a semi-irreducible background in the SR since it has the same signature as $h(b\bar{b}) + E_{\mathrm{T}}^{\mathrm{miss}}$ but it is not resonant in the h_{jj} candidate mass.

- $W+\mathrm{jets}$
  This background contributes $10\%-25\%$ of the total background in the SR. The tree-level Feynman diagrams for $W+\mathrm{jets}$ with b-jets are shown in Figure 6.9. A $W+\mathrm{jets}$ event can enters into the SR region when there is real $E_{\mathrm{T}}^{\mathrm{miss}}$ and b-jets. This can happen when $W \to \tau\nu_\tau$, leaving real $E_{\mathrm{T}}^{\mathrm{miss}}$ from the neutrino, b-jets and a hadronic τ mimicking a QCD originated jet. $W+\mathrm{jets}$ can also contributes to SR when $W \to \ell_{e,\mu}\nu_{e,\mu}$ and the lepton is miss-identified of lost.

- $t\bar{t}$ production
  Top-pair production is a dominant background in the SR especially since b jets are considered. It contributes $15\%-50\%$ of the total background depending on $E_{\mathrm{T}}^{\mathrm{miss}}$ and b-tag category. The tree-level Feynman diagrams for $t\bar{t}$ with b-jets are shown in Figure 6.10. The signature of top events is dictated by the decay mode of the W in top decays, $t \to Wb$. $t\bar{t}$ events are divided in *dileptonic* if both W bosons decays leptonically, *semileptonic* if only one W decays leptonically, and *all hadronic* if both W bosons decays hadronically. In the case when the W boson decays as $W \to \tau\nu_\tau$, the $t\bar{t}$ events can be selected in the SR where the tau decay hadronically and the neutrino leads to real $E_{\mathrm{T}}^{\mathrm{miss}}$. Like for the $W+\mathrm{jets}$, this hadronic decay of the tau is reconstructed as a jet which can mimic a QCD-originated jet. Also, semileptonic $t\bar{t}$ days can contributes to SR when the lepton is misreconstructed or misidentified. On the other side, all hadronic and dileptonic decays of the $t\bar{t}$ production are excluded since no real $E_{\mathrm{T}}^{\mathrm{miss}}$ is expected. For the all hadronic $t\bar{t}$ decay, there can be jet-p_{T} mismeasurement leading to fake $E_{\mathrm{T}}^{\mathrm{miss}}$ which is strongly suppressed by the anti-QCD cuts mentioned in Section 6.4. In the case of the $t\bar{t}$ dileptonic decay, it strongly reduced by the lepton veto.

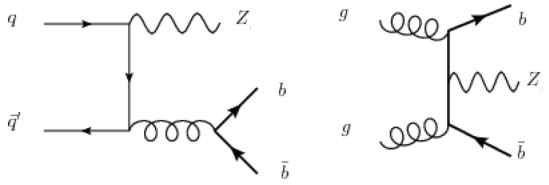


Figure 6.8 – Tree-level Feynman diagram of Z+jets.

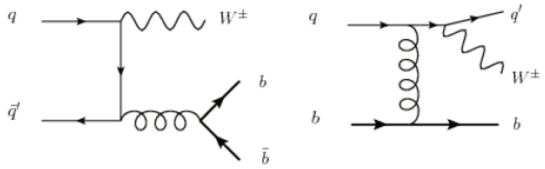


Figure 6.9 – Tree-level Feynman diagram of W+jets.

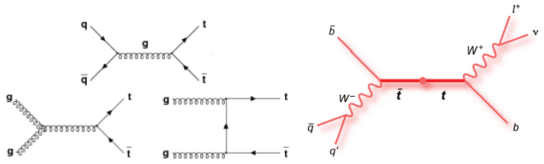


Figure 6.10 – Tree-level Feynman diagrams of top-pair production and decays.

These dominant backgrounds, $V+$jets and $t\bar{t}$, are constrained using two Control Regions (CR) orthogonal to the SR. One region is the single-muon control region, 1μ-CR, which is designed to constrain the $t\bar{t}$ and $W+$jets backgrounds, while the two-lepton Control Region, 2ℓ-CR, constraints the $Z+$jets background contribution. More details on the CR definition will follow on Sections 6.5.2 and 6.5.3. The normalizations of the main backgrounds are determined from a simultaneous fit of the CRs and the SR to data as will be described in Section 7.3. In particular, their modelling is constrained by the sidebands of the fit variable in the SR.

Subdominant backgrounds in the SR are the single top production, the Higgs boson production in association with a vector boson, VH, and the production of two weak gauge bosons, i.e. diboson production. They contribute less than 10% of the total background.

For the single top production, the top quark decays as $t \to Wb$ or can be produced in association with a W, and, as before, the W can produce an hadronically decaying tau, $W \to \tau\nu_\tau$, or a lepton can be lost, leaving real $E_{\mathrm{T}}^{\mathrm{miss}}$ and no lepton to be vetoed.

The production of a SM Higgs in association with a vector boson, WH or ZH, is resonant in the h_{jj} mass spectrum. Most of the VH events in the SR comes from $Z \to \nu\nu$, which has real $E_{\mathrm{T}}^{\mathrm{miss}}$ from the neutrinos production, and since the Higgs can decay to b-quarks $H \to b\bar{b}$, this is an irreducible background process.

Finally, the diboson production can be ZZ, WW or ZW. A Diboson event can be selected in the SR if it leads to real $E_{\mathrm{T}}^{\mathrm{miss}}$, which corresponds to the ZV production where $Z \to \nu\nu$. In most of the cases, the V vector boson of the ZV production in the SR, decays

hadronically. The diboson production is also resonant in the h_{jj} mass spectrum, but at the vector boson mass.

6.5.2 One lepton control region: $W+$jets and $t\bar{t}$

This region, 1μ-CR, is defined in order to normalize both the W+jets and $t\bar{t}$ backgrounds in the combined fit that will be described in Section 7.3. It is characterised for the selection of events with exactly one WH-signal muon and vetoing events containing VH-loose electrons, as defined in Section 6.2.4.

Beyond this requirement, identical selection criteria to those applied in the SR, described in Section 6.4, are imposed to ensure that the set of events selected is topologically and kinematically similar to the events in the 0 lepton SR.

The SR has a lepton veto and thus no muons are considered. Therefore, in order to 1μ-CR be kinematically close to the SR, the $E_{\mathrm{T}}^{\mathrm{miss}}$ has to be modified. Because of this, a new observable called $E_{\mathrm{T,no}\mu}^{\mathrm{miss}}$ is defined to mimic the $E_{\mathrm{T}}^{\mathrm{miss}}$ in the SR. For the $E_{\mathrm{T,no}\mu}^{\mathrm{miss}}$ one considers the selected muon as invisible, therefore, this variable is defined as the magnitude of the vectorial sum of the total missing transverse momentum, $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$, plus the muon four-momentum. Correspondingly, the $p_{\mathrm{T}}^{\mathrm{miss}}$ variable, defined in Section 6.2.7, has to be modified adding the p_{T} of the muon. Those modified missing transverse energy/momentum vectors are then used for defining the $E_{\mathrm{T}}^{\mathrm{miss}}$ regions and applying the anti-QCD cuts described in Section 6.4.

These mentioned $E_{\mathrm{T}}^{\mathrm{miss}}$ regions and the b-tag regions are defined in the exact same way as in the zero lepton SR. Thus, events are categorized in two b-tag regions: 1- and 2-btag, and in four modified $E_{\mathrm{T}}^{\mathrm{miss}}$ bins: 150 GeV $\leq E_{\mathrm{T,no}\mu}^{\mathrm{miss}} <$ 200 GeV, 200 GeV $\leq E_{\mathrm{T,no}\mu}^{\mathrm{miss}} <$ 350 GeV, 350 GeV $\leq E_{\mathrm{T,no}\mu}^{\mathrm{miss}} <$ 500 GeV in the resolved regime, and 500 GeV $\leq E_{\mathrm{T,no}\mu}^{\mathrm{miss}}$ in the merged regime.

6.5.3 Two lepton control region: $Z+$jets

This region, 2ℓ-CR, is defined in order to estimate the Z+jets background normalization. As discussed in Section 6.5.1, $Z(\to \nu\nu)$ + jets production is a dominant background in the SR. Given that the momentum of the Z boson does not depends on its decay products, $Z(\to \nu\nu)$ + jets have the same kinematics and decay topology as $Z(\to \ell\ell)$ + jets. Because of this, the normalisation of $Z(\to \nu\nu)$ + jets events can be estimated with the help of a $Z(\to \ell\ell)$ + jets control region. This allows to constrain the normalization of the $Z(\to \nu\nu)$ + jets background in the SR from a combined fit that will be described in Section 7.3.

The selection criteria in the 2ℓ-CR is the same as for the SR described in Section 6.4 with three major differences.

Firstly, in order to mainly select $Z(\to \ell\ell)$ + jets events, exactly two leptons are required and events with additional VH-Loose leptons are discarded. Also, the events are triggered by the lowest unprescaled single lepton triggers, see Table C.1 for details. In the case of electrons, both must fulfill the ZH-signal electrons described in Section 6.2.4, with one having p_{T} >10 GeV, and a di-electron invariant mass of 83 GeV $< \mathrm{m_{ee}} <$ 99 GeV. In the case of muons, both must fulfill the ZH-signal muons described in Section 6.2.4, with one

muon having p_{T} >10 GeV, and a di-muon invariant mass of 71 GeV<$m_{\mu\mu}$<106 GeV. In addition, both muons must have opposite electric charge, this requirement is not applied to the electron channel due to higher rate of charge misidentification. The di-lepton invariant mass requirement helps to suppress backgrounds having a non-resonant lepton-pair, such as $t\bar{t}$ and multijet events.

Secondly, in order to reduce the $t\bar{t}$ and single top contributions in the 2ℓ-CR, an event-based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance is applied. It is defined as the ratio of the $E_{\mathrm{T}}^{\mathrm{miss}}$ to the square root of the scalar sum of the p_{T} of the two leptons and all the small-R jets in the event. Events are accepted if $E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}} < 3.5\,\sqrt{\mathrm{GeV}}$, where $H_{\mathrm{T}} = \sum_{i=0}^{2} p_{\mathrm{T}}^{\ell_i} + \sum_i p_{\mathrm{T}}^{\mathrm{jet_i}}$. This requirement separates $Z \to \ell\ell$ + jets events processes from $t\bar{t}$ production, since $E_{\mathrm{T}}^{\mathrm{miss}}$ originates from finite detector resolution for the former and mainly from neutrinos for the latter.

Finally, given that events in 2ℓ-CR considers exactly two signal lepton, multijet contribution is expected to be negligible. Because of this, the $p_{\mathrm{T}}^{\mathrm{miss}}$ lower threshold and the anti-QCD cuts are not applied.

As it is done in the SR, the events are separated into regions based on the b-tagging multiplicity and in variable consistent to $E_{\mathrm{T}}^{\mathrm{miss}}$ in SR. Similarly to the 1μ-CR, this regions are not defined in $E_{\mathrm{T}}^{\mathrm{miss}}$ categories, and instead in $p_{\mathrm{T}}^{\ell\ell}$ which mimics the $E_{\mathrm{T}}^{\mathrm{miss}}$ in the SR. Assuming that neutrinos in $Z \to \nu\nu$ + jets events are the main part of $E_{\mathrm{T}}^{\mathrm{miss}}$ in the SR, the transverse momentum of the Z boson ($p_{\mathrm{T}}^{Z} = p_{\mathrm{T}}^{\ell\ell}$) is the equivalent variable in the context of $Z \to \ell\ell$ + jets events. Therefore, events are categorised in four categories related to the ones in the SR: 150 GeV $\leq \mathrm{p}_{\mathrm{T}}^{\ell\ell} <$ 200 GeV, 200 GeV $\leq \mathrm{p}_{\mathrm{T}}^{\ell\ell} <$ 350 GeV, 350 GeV $\leq \mathrm{p}_{\mathrm{T}}^{\ell\ell} <$ 500 GeV in the resolved regime, and 500 GeV $\leq \mathrm{p}_{\mathrm{T}}^{\ell\ell}$ in the merged regime.

6.5.4 Multijet background estimation

Multijet events comes from strong interactions and when produced from high energy pp collisions, usually have large production cross-section, however rarely present large $E_{\mathrm{T}}^{\mathrm{miss}}$. In the mono-$h(b\bar{b})$ dark matter search, high $E_{\mathrm{T}}^{\mathrm{miss}}$ is required given that the threshold is $E_{\mathrm{T}}^{\mathrm{miss}}$ >150 GeV. Multijet events with high $E_{\mathrm{T}}^{\mathrm{miss}}$, as the one considered here, do not occur very often since substantial fluctuations in the calorimeter jet energy measurement are the root cause. And indeed, for $E_{\mathrm{T}}^{\mathrm{miss}}$ >350 GeV the multijet contribution is negligible, as it is going to be shown in the following. This makes a estimation fully coming from MC simulation very difficult due to the high statistics which would be required.

Because of this, the multijet background is estimated in a dedicated data-driven method in the zero lepton SR. Since high $E_{\mathrm{T}}^{\mathrm{miss}}$ is required, this background is a minor contribution to the overall set of selected events. This can be seen on the right of Figure 6.4 after selecting events passing the anti-QCD requirements which act as multijet cleaning cuts. Therefore, the estimation of the multijet contribution can come with a relatively large uncertainty, without diminishing the sensitivity of the analysis.

The strategy used to estimate this background is described in the three following steps:

1. Derive multijet template from a multijet enriched region:
   In order to derive a multijet template for different variables, a multijet enriched region is defined by inverting the cut $\min(\Delta\Phi(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}, \mathbf{p}_{\mathrm{T}}(\mathrm{jets}_{1,2,3}))) > 20°$. From Figure 6.4 one can see that the multijet background has a markedly different dis-

tribution from the electro-weak backgrounds, for which below 20° multijet events dominate.

Therefore, the multijet enriched region is defined orthogonal to the SR with minimum ($\Delta\Phi(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}, \mathbf{p}_{\mathrm{T}}(jets_{1,2,3}))) < 20°$, however the rest of the event selection criteria from the SR is applied. In this region, also named multijet CR, the difference between data and the simulated MC is used as a model of the shape of multijet distributions.

Figure 6.11 shows a scheme of the regions considered in this data-driven estimation, for which the left upper region corresponds to the multijet CR, where the multijet template is extracted.

2. Determine the template fake factor from a fit using a multijet sensitive variable in a selection similar to SR with loosened anti-QCD cuts:
The previous template method describes the shape of the multijet background but does not determine the normalization in SR. Because of this, the template must be correctly normalized in order to take into account the difference in efficiency between the two selections in the multijet CR and in SR.

   In order to obtain the normalisation factor, the multijet template together with the simulated backgrounds are fit to data in the region with minimum ($\Delta\Phi(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}, \mathbf{p}_{\mathrm{T}}(jets_{1,2,3}))) > 20°$. The parameter of interest in the fit is the normalisation of the multijet contribution.

   However, the population of data events corresponding to the multijet contribution after the full SR event selection is too small to produce a reasonable fit. Because of this, in order to have a robust fitting result regarding the multijet normalisation, the event selection is relaxed compared to the full selection. More precisely, the following anti-QCD cuts are dropped in order to increase the amount of multijet events: $\Delta\Phi\ (\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}, \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}) < 90°$, $\Delta\Phi(\mathbf{p}_{\mathrm{T}}(j_1), \mathbf{p}_{\mathrm{T}}(j_2)) < 140°$, $\Delta\Phi(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}, \mathbf{p}_{\mathrm{T}}(h_{jj})) > 120°$ and $\Delta R(j,j) < 1.8$.

   Therefore, instead of fitting the multijet template from CR to SR, it is done from QCDCR to QCDSR showed in Figure 6.11. Also, since the fitting region with loose anti-QCD cuts, QCDSR, is similar to the SR, the fitting procedure just considers events with m_{jj} outside the Higgs boson mass window [3], i.e. the procedure is blinded, in order to remove any possibility of fitting away a dark matter signal.

3. Normalise the multijet template obtained in the step 1. with the fake factor obtained in the step 2.
The template shape constructed from the template region, CR, is multiplied with the fake factor acquired from the fit performed in the QCDSR in order to obtain the prediction after full selection in SR.

Figure 6.12 [4] shows the minimum azimuthal angle between three leading jets in the Central+Forward set (see Section 6.4.2.1 for details) and the $E_{\mathrm{T}}^{\mathrm{miss}}$, i.e. minimum ($\Delta\Phi(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$,

[3] The Higgs boson mass window is defined as $70 < m_{jj}/\ \mathrm{GeV} < 140$

[4] Plots shown in this section uses a newer ATLAS offline reconstruction software release (Athena Release 21) than the one used for the public result (Athena Release 20.7), see Appendix D for details.

$\mathbf{p}_{\mathrm{T}}(jets_{1,2,3})))$ for the four regions considered in the data-driven multijet estimation, and in the same order as displayed in scheme of Figure 6.11. Plots are shown in the 2-btag category, where most of the sensitivity is expected, and in the first $E_{\mathrm{T}}^{\mathrm{miss}}$ category, where most of the multijet background contribution is expected. Since no complete multijet estimation is considered in these plots, apart form the dijet simulation, MC simulation is supposed to underestimate data in regions more contaminated with multijet events, i.e. at low values of $(\Delta\Phi(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}, \mathbf{p}_{\mathrm{T}}(jets_{1,2,3})))$. This is clear for the multijet enriched region in Figure 6.12(a), CR, and even more when the anti-QCD cuts are relaxed in Figure 6.12(c), QCDCR.

The multijet sensitive variable which is fitted in order to obtain the multijet scale factor is the multiplicity of small-R jets that contain a muon within the jet cone. Figure 6.13 shows the distributions in data and simulation of this variable after full SR event selection in the 2-btag category and in the first $E_{\mathrm{T}}^{\mathrm{miss}}$ region, 150 GeV $\leq \mathrm{E}_{\mathrm{T}}^{\mathrm{miss}} <$ 200 GeV. This choice is motivated by the fact that many of the multijet events come with real $E_{\mathrm{T}}^{\mathrm{miss}}$ caused by neutrinos originated from heavy flavour hadron decays, and these decays occur within jets. In addition to the neutrino, they can also result in muons. Thus, the multijet events are likely to accumulate at higher multiplicities of jets which contain muons.

A multijet template shape of the number of jets with muons is obtained in the region with loose anti-QCD cuts and inverting $(\Delta\Phi(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}, \mathbf{p}_{\mathrm{T}}(jets_{1,2,3}))) > 20°$, QCDCR. Then, this template and MC simulated backgrounds are fit to data in SR with loose anti-QCD cuts, QCDSR. Left plot in Figure 6.14 shows data and MC simulated electroweak backgrounds in the QCDSR, and the multijet template obtained by subtracting MC simulation to data in the QCDCR. Right plot in Figure 6.14 shows data, MC simulated electroweak backgrounds and multijet template after fitting to data. In these profile likelihood fits, all simulated backgrounds are allowed to float independently within theoretical uncertainties with one overall normalization factor that controls the normalization of non-multijet components in the fit. Finally, statistical uncertainties are taken into account separately for each component in the fit.

This fitting procedure is performed for each b-tagging category and for each $E_{\mathrm{T}}^{\mathrm{miss}}$ region independently. Figure 6.15 shows the post-fit distributions for data, simulation and multijet template for the first two $E_{\mathrm{T}}^{\mathrm{miss}}$ regions where most of the multijet contribution is expected. The corresponding pulls for these fits are shown in Figure 6.16.

The normalisation or fake factor of the multijet template is left to float freely and is the parameter of interest which provides estimate of the normalisation of the multijet contribution in SR. The fake factors that will scale the multijet template in the SR, obtained by these fits, are listed in Table 6.7. It is observed that these normalisations decrease considerably as the $E_{\mathrm{T}}^{\mathrm{miss}}$ increases. The normalisations decrease by 33% for 0 b-tagged region and more than 50% for 1 and 2 b-tagged regions. Considering this decrease from the $E_{\mathrm{T}}^{\mathrm{miss}}$ interval [150, 200] GeV to the next one [200, 350] GeV, it is clear that the multijet contribution can be neglected in the higher $E_{\mathrm{T}}^{\mathrm{miss}}$ interval [350, 500] GeV.

In order to check that indeed the multijet contribution is negligible at high $E_{\mathrm{T}}^{\mathrm{miss}}$, and in particular in the merged regime, the estimation of this background is performed using

Plots, fitting results, and multijet normalisation estimation are consistent between the two releases. These plots corresponds to a cross-check and validation.

the $E_{\mathrm{T}}^{\mathrm{miss}}$ as multijet sensitive variable instead of the number of jets with muons. The event selection of the merged regime is applied in the same way as it is done in the final search for dark matter. Inverting the cut mentioned above, so that min($\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}}$,small-R jets)) $< 20°$, a template of the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution of multijet events is derived. In order to obtain a useful multijet template, values below $E_{\mathrm{T}}^{\mathrm{miss}} = 500$ GeV are also taken into account. Then a $E_{\mathrm{T}}^{\mathrm{miss}}$ fit of the model to the data is performed in the region min($\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}}$,small-R jets)) $> 20°$. The resulting distributions are shown together with the data in Figure 6.17.

As can be seen from Fig. 6.17, multijet contribution is exponentially falling and is vanishingly small for $E_{\mathrm{T}}^{\mathrm{miss}} > 400$ GeV. For all b-tag multiplicities the fraction of multijet events in merged regime calculated from these fits is less than 1%. Because of this, multijet is estimated only in the resolved regime.

The final discriminant in this search is the invariant mass of the Higgs candidate h_{jj}. Therefore, the multijet template of interest is for the di-jet invariant mass distribution, which is derived from the CR, as explained in step 1. Then, this multijet template is transferred from the multijet CR to the SR by scaling it with the derived fake factors from the fits, as explained in step 2. Finally, this data-driven multijet estimation is included in the total background estimation of the final fit, see Section 7.3. Figure 6.18 shows the pre-fit h_{jj} candidate invariant mass considering the multijet estimation derived with this data-driven method.

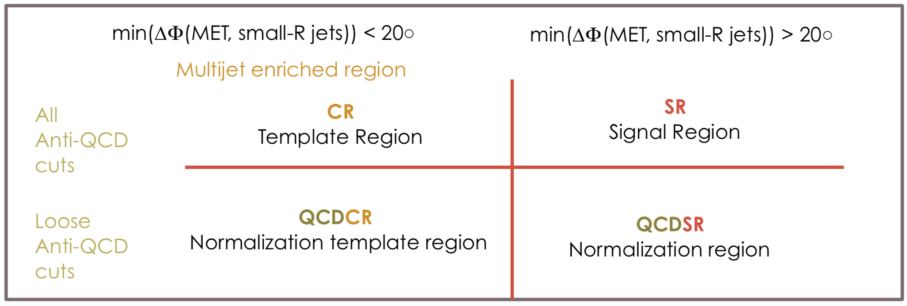


Figure 6.11 – Multijet fitting regions scheme

	$E_{\mathrm{T}}^{\mathrm{miss}} \in [120, 200]$	$E_{\mathrm{T}}^{\mathrm{miss}} \in [200, 350]$
0 b-tag	0.244 ± 0.020	0.108 ± 0.034
1 b-tag	0.128 ± 0.007	0.055 ± 0.010
2 b-tag	0.129 ± 0.004	0.072 ± 0.005

Table 6.7 – Table of scale factors obtained by fitting multiplicity of jets which contain a muon variable in QCDCR selection region as described above.

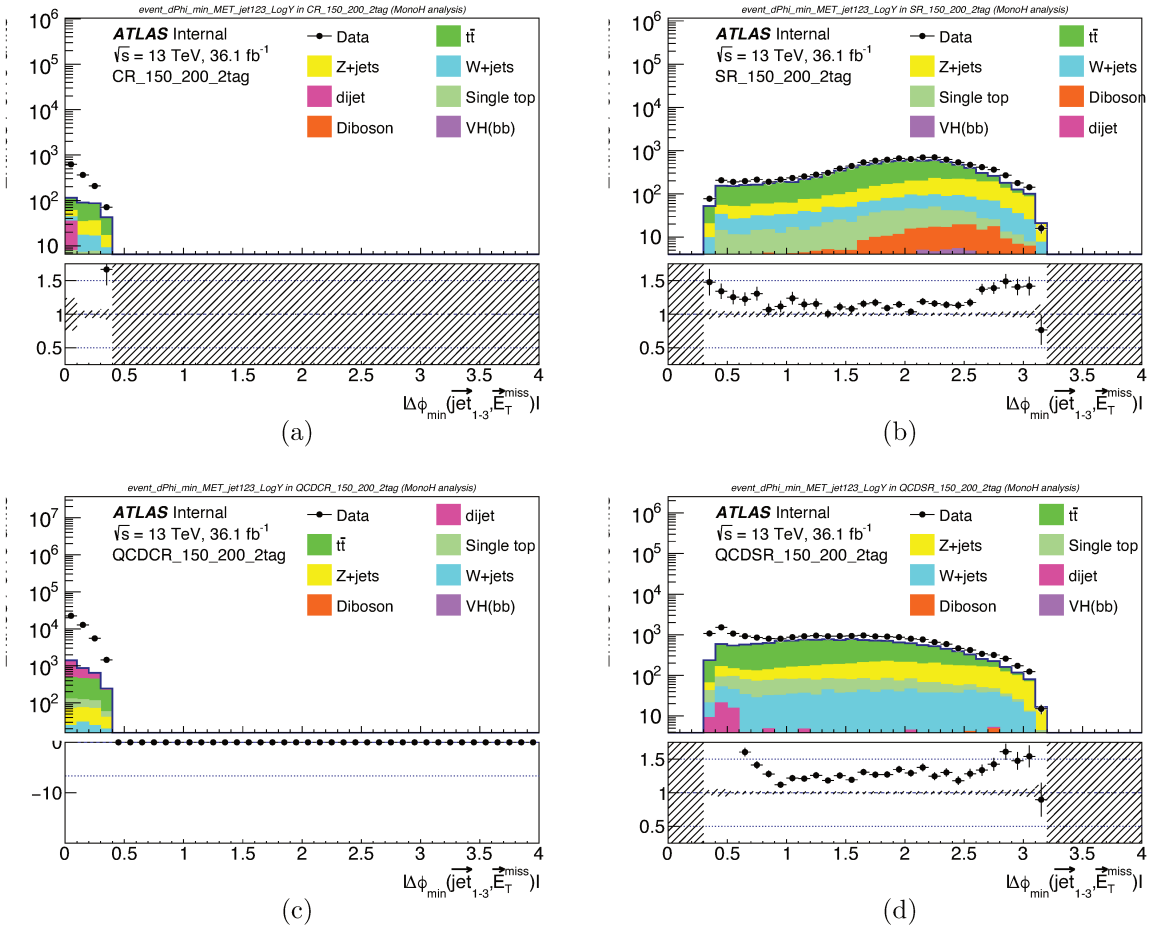


Figure 6.12 – Data and MC simulation plots for the minimum azimuthal angular separation between the first three leading jets and the $E_{\mathrm{T}}^{\mathrm{miss}}$ in the four regions considered in the multijet background data-driven estimation: (a) Multijet control region CR, (b) Signal region SR, (c) region for template to fit QCDCR (d) template normalisation region QCDSR. Plots are shown for the first $E_{\mathrm{T}}^{\mathrm{miss}}$ region, 150 GeV $\leq \mathrm{E}_{\mathrm{T}}^{\mathrm{miss}} <$ 200 GeV in the 2-b tag category.

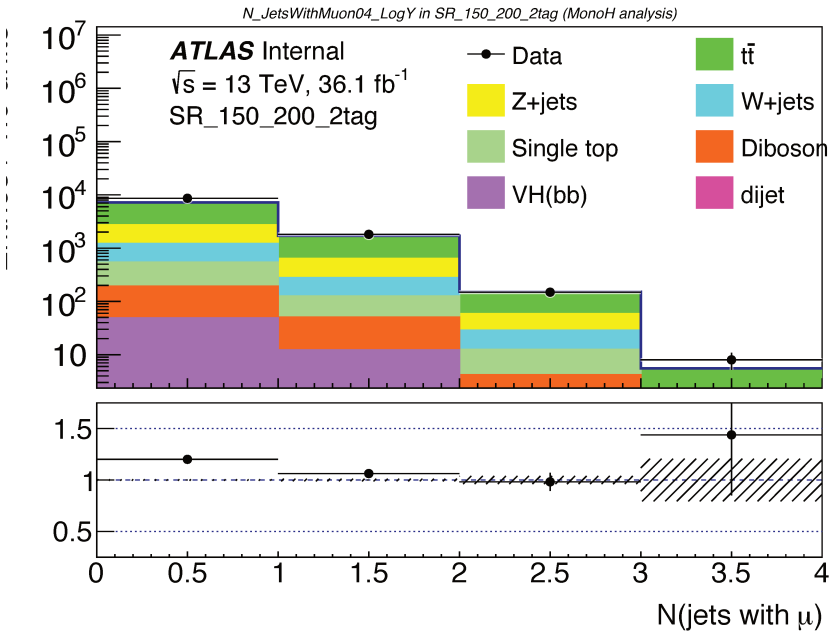


Figure 6.13 – Number of jets which contain a muon in the SR with 2-btag in the first $E_{\mathrm{T}}^{\mathrm{miss}}$ region, 150 GeV $\leq \mathrm{E}_{\mathrm{T}}^{\mathrm{miss}} <$ 200 GeV.

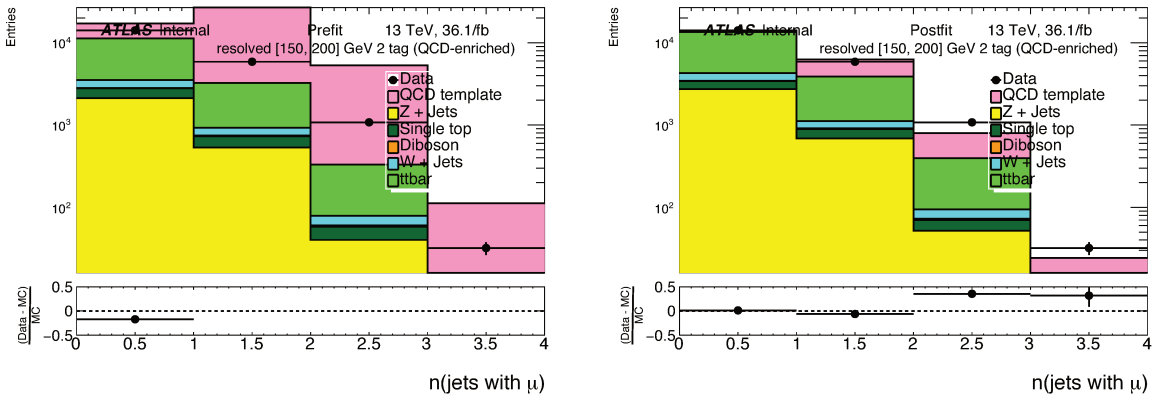


Figure 6.14 – Number of jets which contain a muon distributions in data, MC simulation and multijet template estimation (QCD template). On the left plot, the multijet template is extracted from QCDCR, and data and MC simulation are shown in the QCDSR. The right plot shows the result of fitting simultaneously the multijet template and MC simulation to data.

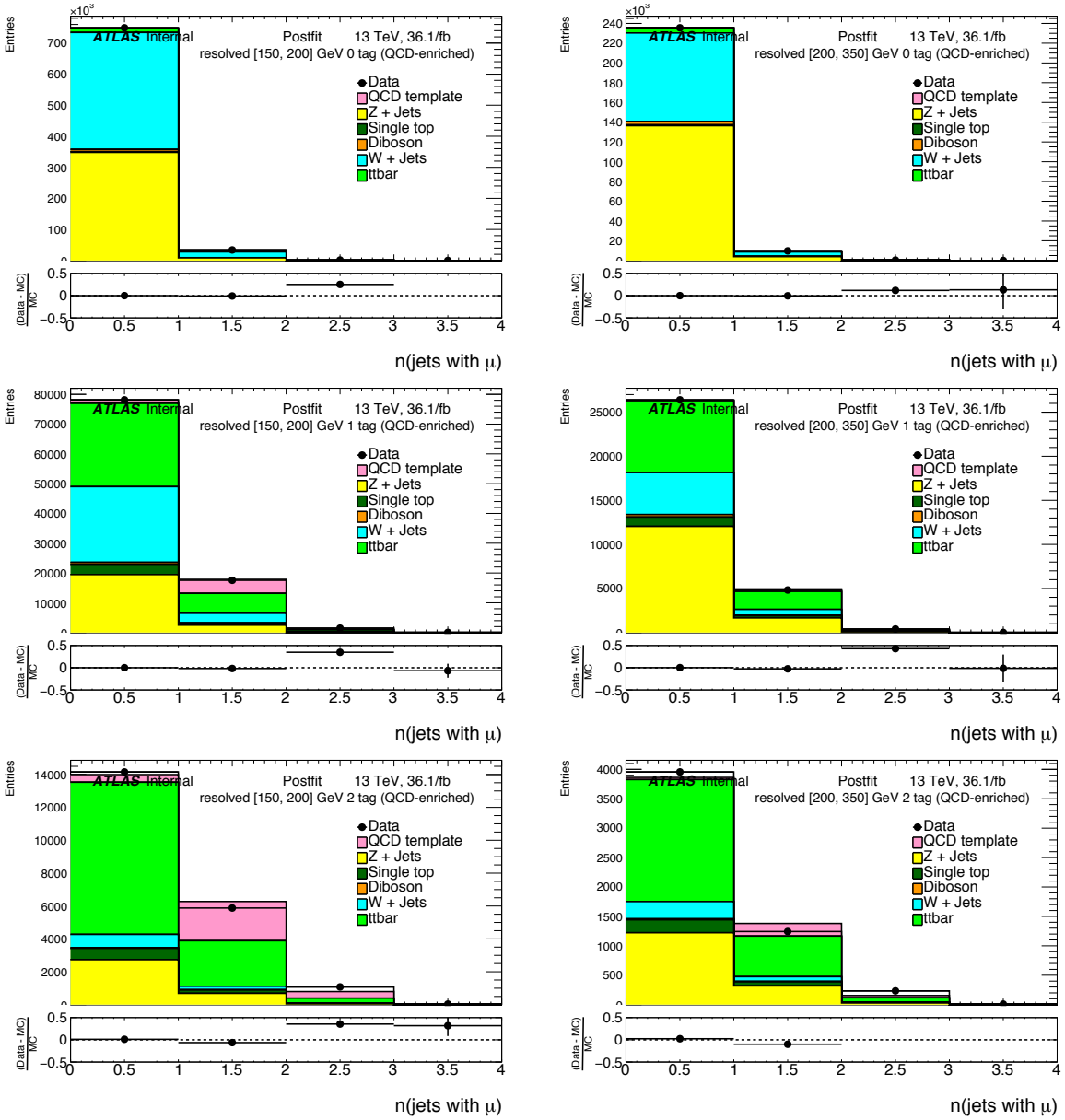


Figure 6.15 – Post-fit distributions of the number of jets which contain a muon distributions in data, MC simulation and multijet template estimation.

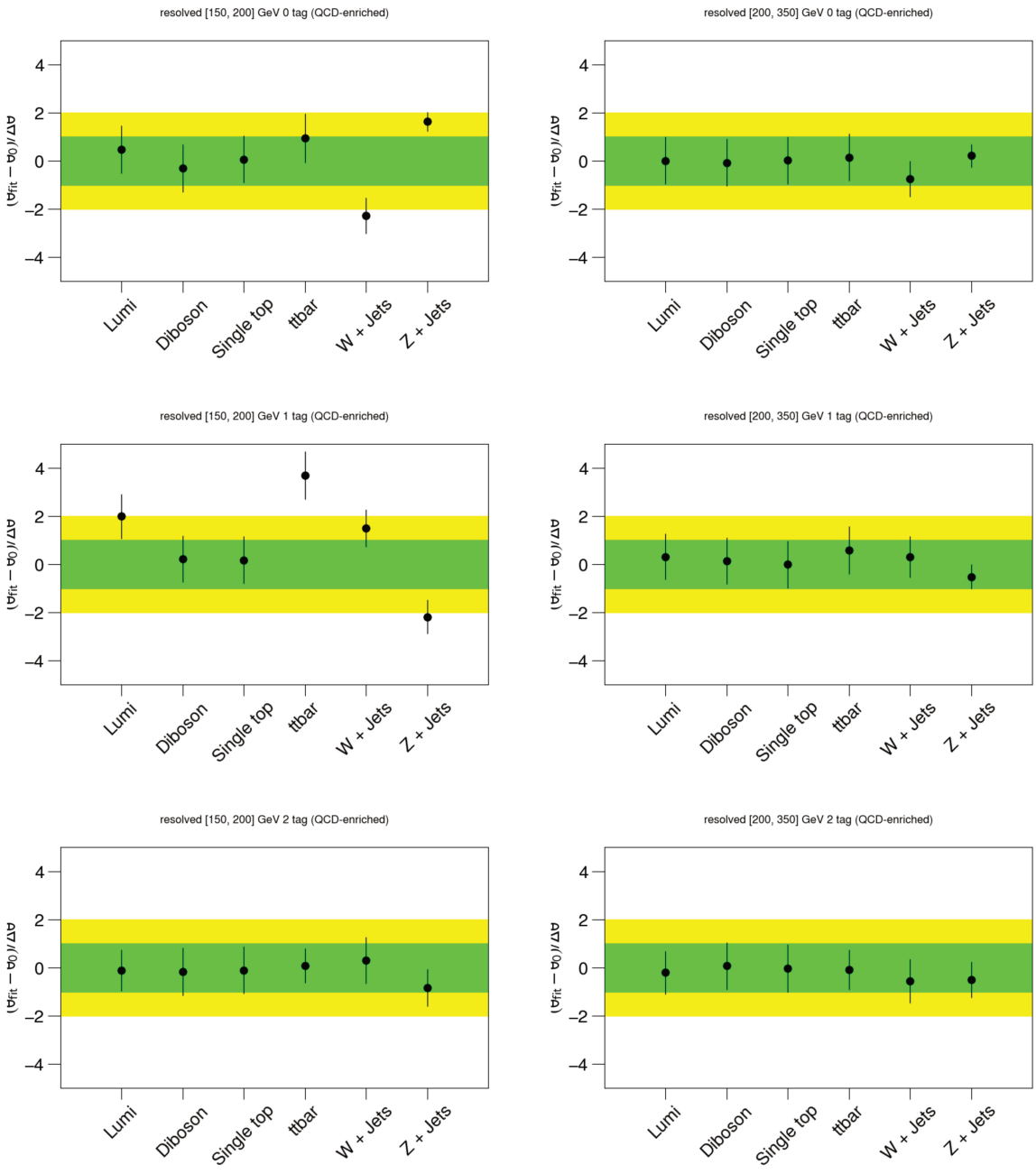
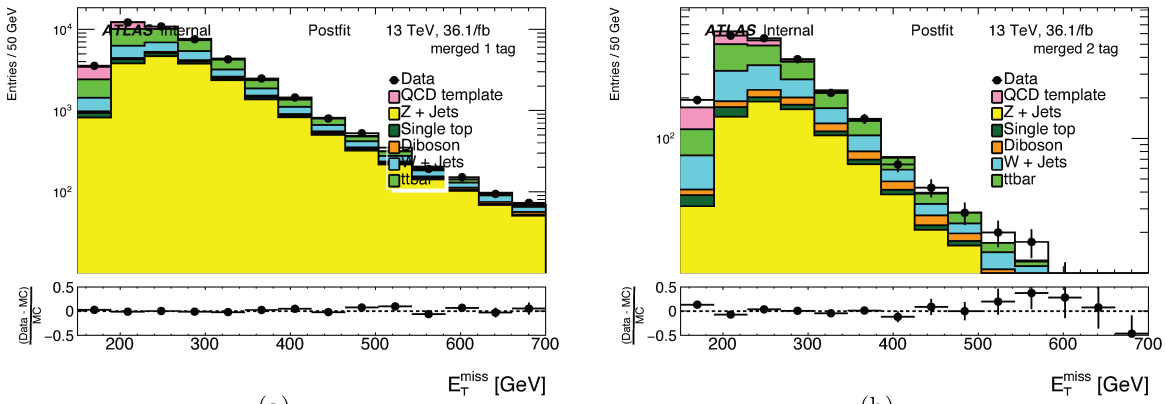


Figure 6.16 – Pulls of the profile likelihood fit for each b-tag multiplicity and first two $E_{\mathrm{T}}^{\mathrm{miss}}$ regions.



(a) (b)

Figure 6.17 – Post-fit $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions after the profile likelihood fit for 1 and 2 b-tag multiplicities merged categories (starting already at 150 GeV): (a) merged 1-btag, (b) merged 2-btag, Multijet background contribution is negligible for events with high $E_{\mathrm{T}}^{\mathrm{miss}}$.

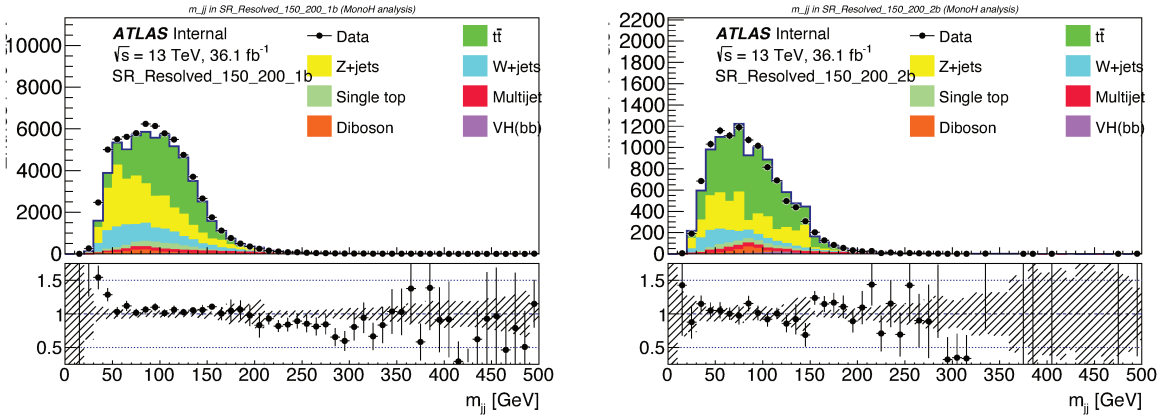


Figure 6.18 – Higgs candidate mass spectrum distributions for data, MC simulations and data-driven multijet estimation in the first two $E_{\mathrm{T}}^{\mathrm{miss}}$ regions with 2b-tag jets.

Chapter 7

Mono-$h(b\bar{b})$: statistical analysis, systematic uncertainties and results with $36\,fb^{-1}$ of 2015 and 2016 data at $\sqrt{s} = 13$ TeV

7.1 Statistical analysis

In this search of new physics, the physics model considers the simulated backgrounds, the data-driven multijet background estimate, and non-SM processes, dark matter signals, in order to statistically test it against data. Under the hypothesis of this model, the statistical test allows to either claim a discovery or exclude regions of the parameter space of new physics at a certain confidence level.

The interpretation of the final results is performed in terms of a profile likelihood fit of a physics model to the data.

The signal model considered is the Z'-2HDM described in Section 2.2.2 with different signal mass points of the massive mediator and pseudoescalar $(m_{Z'}, m_{A^0})$ that are tested.

The fit considers simultaneously different channels. The zero lepton channel is the signal region (SR), while the one muon (1μ-CR) and two lepton channels (2ℓ-CR) are used to constrain the main background contributions. Therefore, the combined fit of signal and control regions, determines not only the signal strength but also the normalisation of the dominant backgrounds. Details of main backgrounds and control regions can be found in Section 6.5.1.

This section presents the generalities of the statistical model, followed by the results of this search and the statistical interpretation.

7.1.1 Binned Profile Likelihood Overview

The statistical analysis of the data uses a binned likelihood function $\mathscr{L}$ constructed as the product of Poisson probability terms [241, 242],

$$\mathscr{L}\left(n^{obs} \mid \mu\right) = \prod_{i=1}^{N} \frac{\left(\mu S_i + B_i\right)^{n_i^{obs}}}{n_i^{obs}!} e^{-(\mu S_i + B_i)} \tag{7.1}$$

where n_i^{obs} denotes the number of observed events in the i-th bin of N total bins. S_i and B_i are the number of signal and background events in the i-th bin, where the expected event yield in the i-th bin is given by $\mu S_i + B_i$. μ is the *signal strength* parameter defined as the ratio of the signal cross-section to the reference signal cross-section. $\mu = 0$ corresponds to the background-only hypothesis, i.e. null hypothesis, and $\mu = 1$ corresponds to the nominal signal hypothesis plus background.

The dependence of the signal and background predictions on the systematic uncertainties is described by a set of *nuisance parameters*, NP, $\boldsymbol{\theta} = (\theta_i)$. These NP are fitted to data by including constraint terms in the likelihood function on Equation 7.1 which represent each source of uncertainty. The systematic uncertainties considered in the statistical model are described in Section 7.2.

Each θ_i is parameterized by Gaussian or log-normal priors, with central value θ_i^0 and width σ_{θ_i}. The log-normal priors are used for normalization uncertainties in order to maintain a positive likelihood. The priors act to constrain the NPs to their nominal values within their assigned uncertainties.

The expected numbers of signal and background events in each bin i are functions of the nuisance parameters θ and are fitted to data in the likelihood function:

$$\mathscr{L} = \prod_{i=1}^{N} \frac{\left(\mu S_i(\boldsymbol{\theta}) + B_i(\boldsymbol{\theta})\right)^{n_i^{obs}}}{n_i^{obs}!} e^{-(\mu S_i(\boldsymbol{\theta}) + B_i(\boldsymbol{\theta}))}. \tag{7.2}$$

$$\prod_{\theta_j, \theta_k \in \boldsymbol{\theta}}^{N} \text{Gaussian}\left(\theta_j \mid \theta_j^0, \sigma_{\theta_j}\right) \text{Lognormal}\left(\theta_k \mid \theta_k^0, \sigma_k\right) \tag{7.3}$$

The nominal fit result is obtained by maximising the likelihood function with respect to all parameters. This is referred to as the maximised log-likelihood value, MLL.

The profile likelihood ratio is used in order to test a signal strength μ hypothesis. It can be written as:

$$\lambda(\mu) = \frac{\mathscr{L}\left(\mu, \hat{\hat{\boldsymbol{\theta}}}(\mu)\right)}{\mathscr{L}\left(\hat{\mu}, \hat{\boldsymbol{\theta}}\right)} \tag{7.4}$$

where $\hat{\hat{\boldsymbol{\theta}}}$ is value for $\boldsymbol{\theta}$ which maximises the likelihood function $\mathscr{L}$ for a given μ, it is called the conditional maximum-likelihood estimator of $\boldsymbol{\theta}$. On the other hand, $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ are the unconditional maximum-likelihood estimators of μ and $\boldsymbol{\theta}$ respectively. Therefore, the

denominator is the "likelihood of best fit", and the numerator is the "likelihood assuming μ signal strength". Then, different μ values are tested.

The profile likelihood ratio on Equation 7.4 takes values $0 \leq \lambda \leq 1$, for which $\lambda \sim 1$ implies a good agreement between data and the hypothesized value of μ, and $\mu = 0$ corresponds to the background-only hypothesis.

The test statistic q_μ under the signal plus background hypothesis is constructed according to the profile likelihood ratio as

$$q_\mu = -2 \ln \lambda(\mu) . \tag{7.5}$$

Higher values of q_μ correspond to increasing incompatibility between the data and μ. The null hypothesis q_0 describes the background only, while the alternative test hypothesis describes signal plus background.

This test statistic is used to measure the compatibility of the background-only model with the observed data and to derive exclusion intervals derived with the CL_s method [243, 244].

7.1.2 Statistical Fit Model

The Higgs boson candidate h_{jj} invariant mass is used as the final discriminating variable for dark matter signal against backgrounds. Therefore, in the SR, which is the zero lepton channel, the fit of the model to the data is based on the mass of the two small-R jets in the resolved regime or on the mass of the single large-R jet in the merged regime. Only events within a h_{jj} candidate signal window $(50, 280)$ GeV are considered in the fit.

Furthermore, different categories distinguished by the lepton multiplicity (0 lepton SR, 1μ-CR, 2ℓ-CR), the b-tag multiplicity (1 and 2) and the range in $E_{\mathrm{T}}^{\mathrm{miss}}$ or $E_{\mathrm{T}}^{\mathrm{miss}}$ proxies ($[150,200)$, $[200,350)$, $[350,500)$, and $[500,\infty)$(GeV)) are fitted simultaneously in order to gain signal sensitivity and to better constrain background contributions.

Signal contributions accumulate in the zero lepton channel, SR. The Z+jets background is constrained by the two lepton channel, 2ℓ-CR. The one lepton channel 1μ-CR is used to constrain the W+jets and top-quark backgrounds. An equivalent observable to $E_{\mathrm{T}}^{\mathrm{miss}}$ in SR is constructed for 1μ-CR and 2ℓ-CR. These $E_{\mathrm{T}}^{\mathrm{miss}}$ proxies variables are $E_{\mathrm{T,no}\mu}^{\mathrm{miss}}$ in the one muon channel and $p_{\mathrm{T}}^{\ell\ell}$ in the two lepton channel. See Section 6.5 for details on the background processes and control regions.

In the 1μ-CR, the electric charge of the μ is used in the fit as discriminating variable to separate $t\bar{t}$ from $W+$ jets processes, since $t\bar{t}$ provides an equal number of μ^+ and μ^-, while a prevalence of μ^+ is expected from $W+$ jets process. In the 2ℓ-CR, due to limited data statistics, only the total number of events, event yield, is considered in the fit. Table 7.1 summarises the event categories used in the fit.

The combined fit considers four free normalisation parameters which are determined from data. One is the signal strength μ, which is the parameter of interest of the statistical signal model. The other three freely-floating parameters are the normalisation of the dominant background processes which are constrained by the signal and control regions. These three background normalisation parameters are $t\bar{t}$, W+jets and Z+jets.

	SR	**1μ-CR**	**2ℓ-CR**
Aim	signal region	$t\bar{t}$ and W+jets control region	Z+jets control region
Lepton selection	0 lepton	1 muon	2 leptons
Discriminant	m_{bb} distribution	muon charge	m_{bb}, single bin
b-tag multiplicities	1 and 2		
$E_{\mathrm{T}}^{\mathrm{miss}}$ or $E_{\mathrm{T}}^{\mathrm{miss}}$ pry	$E_{\mathrm{T}}^{\mathrm{miss}}$	$E_{\mathrm{T}}^{\mathrm{miss,\,no\,mu}}$	$p_{\mathrm{T}}^{\ell\ell}$
	resolved: [150,200], [200,350] and [350,500] GeV merged: more than 500 GeV		

Table 7.1 – Event categories entering the combined fit of the model to the data. The discriminant m_{bb} denotes the mass of the light Higgs boson candidate.

The V+jets background in the SR is dominated by jets containing b- or c-quarks and this is denoted as HF jets, for which the flavour composition of the two leading jets can be $b-b$, $b-c$, $c-c$ or $b-$light. The introduction of different Z + HF compositions in zero and two lepton channels is motivated by a study presented in Reference [245] (table 12 therein), where different event generator predictions are compared.

In addition to the freely floating normalisation parameters, the fit contains nuisance parameters which are constrained by prior knowledge. They are related to further background normalisations, instrumental uncertainties and modeling uncertainties, see section 7.2 for details. The normalisations of other background process different than $t\bar{t}$ and HF V + jets are set to their expectations from theory with an assigned theoretical uncertainties [1]. These background normalisations follow a log-normal distribution, while the other nuisance parameters follow a Gaussian distribution. All of these parameters are individually correlated across all channels.

Partial decorrelation between the fitting categories and part of the backgrounds components is introduced for some nuisance parameters following studies from References [246, 245]. The overall W + Heavy Flavour normalisation is decorrelated among the zero lepton SR and one lepton 1μ-CR channels within an uncertainty of 20%. In the 1μ-CR, the normalisation parameter is fitted to data. While in the 0 lepton SR it is given by the overall normalisation which is simultaneously fitted to data in the 0 lepton, multiplied to an additional nuisance parameter which allows this normalisation to change constrained within an uncertainty of 20% following a Gaussian distribution. This additional scaling parameter accounts for the relative variations in acceptance of the HF W + jets in the SR with respect to the 1μ-CR. Same procedure is done for HF Z + jets background between the 0 lepton SR and the 2ℓ-CR.

Furthermore, the flavour composition of the HF W/Z + jets backgrounds is decorrelated considering relative variations on the jet flavours $b-$light, $b-c$, $c-c$ with respect

to the total HF V + jets. This normalisation parameter is allowed to vary within uncertainties of 20%.

The HF W + jets flavour decorrelation is considered across all channels, while for HF Z + jets is treated individually in the SR and in 2ℓ-CR. The uncertainties for the decorrelations are consistent to those in References [246, 245].

7.1.2.1 Signal Significance

In order to search for new physics, the test statistic introduced in Equation 7.5 is used.

When the presence of a new signal can only increase the mean event rate beyond what is expected from background alone, the signal has $\mu \geq 0$. Rejecting the $\mu = 0$ hypothesis effectively leads to the discovery of a new signal. The background-only hypothesis with $\mu = 0$ is defined as:

$$q_0 = \begin{cases} -2\ln\lambda(0) = -2\ln\left(\dfrac{\mathscr{L}\left(0,\hat{\hat{\boldsymbol{\theta}}}(0)\right)}{\mathscr{L}\left(\hat{\mu},\hat{\boldsymbol{\theta}}\right)}\right) & , \text{if} \quad \hat{\mu} \geq 0 \\ 0 & , \text{if} \quad \hat{\mu} < 0 \end{cases} \tag{7.6}$$

where $\hat{\mu}$ denotes the maximum likelihood estimator of the signal strength, following notation on Section 7.1.1. In order to quantify the level of disagreement between the data and the hypothesis of $\mu = 0$ using the observed value of q_0 from data $q_{0,\text{obs}}$, the p-value is computed as

$$p_0 = \int_{q_{0,\text{obs}}}^{\infty} f(q_0 \mid 0)\, dq_0, \tag{7.7}$$

where $f(q_0 \mid 0)$ denotes the probability density function (PDF) of the statistic under the assumption of the background-only hypothesis. On the other hand, in order to define upper limits on the signal strength, the test statistic under the signal plus background hypothesis is defined as:

$$q_\mu = \begin{cases} -2\ln\lambda(\mu) = -2\ln\left(\dfrac{\mathscr{L}\left(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu)\right)}{\mathscr{L}\left(\hat{\mu},\hat{\boldsymbol{\theta}}\right)}\right) & , \text{if} \quad \hat{\mu} \leq \mu \\ 0 & , \text{if} \quad \hat{\mu} > \mu \end{cases} \tag{7.8}$$

It is clear that higher values of q_μ represent greater incompatibility between the data and the hypothesized value of μ. This definition, Equation 7.8, does not considers data for $\hat{\mu} > \mu$, given that this is not taken as part of the rejection of the test.

Also in this case, the level of agreement between data and the hypothesized μ can be quantified with the p-value as

$$p_\mu = \int_{q_{\mu,\text{obs}}}^{\infty} f(q_\mu \mid \mu)\, dq_\mu, \tag{7.9}$$

The distribution of this test statistic in the large sample limit, called the Wald approximation [247], is known analytically, and is used to derive the p-values [243].

These p-values can be expressed as significances. The significance, Z, is defined such that a Gaussian variable found Z standard deviations above its mean has an upper tail probability equal to p,

$$Z = \Phi^{-1}(1-p), \tag{7.10}$$

where Φ^{-1} is the inverse of the cumulative distribution of the standard Gaussian. In particle physics, the rejection of the background only hypothesis is for a significance of $Z \geq 5$, which corresponds to a p-value of $p = 2.87 \times 10^{-7}$. In the case of excluding a signal hypothesis, it is used a threshold of $p = 0.05$, i.e. 95% of confidence level, corresponding to $Z = 1.64$.

Sometimes searches are dominated by background and no significant signal is expected. In this case, the distributions of the test statistic for the $s+b$ hypothesis and the background-only hypothesis are overlapping. If the observed number of events has a sufficiently large fluctuation below the expected background, p_{s+b} decreases, so the search may exclude a signal model which has no or low sensitivity. To avoid this, the CL_s probability is defined as

$$CL_s = \frac{p_{s+b}}{1-p_b}, \tag{7.11}$$

where the $s+b$ hypothesis is rejected at 95% confidence level if $CL_s < 0.05$ [244].

The sensitivity of the analysis is given by the expected significance given a certain signal, where "expected" means that fits to the Asimov dataset [1] with $\mu = 1$ are performed. All available Z'-2HDM simplified models were tested with a common cross-sections of 10 fb. This allows to estimate the sensitivtiy of the analysis to signals with different kinematic distributions without the need to considers the mass-dependence of signal-cross sections at the same time. Figure 7.1 presents these expected significances. The significance is largest for models with large mediator masses $m_{Z'}$, because these signals result in large $E_{\mathrm{T}}^{\mathrm{miss}}$. Backgrounds rarely result in large $E_{\mathrm{T}}^{\mathrm{miss}}$ values, so that for high Z' masses signal over background ratios are large. However, if the mediator mass becomes too high, the sensitivity goes down because then the Higgs-boson is highly boosted and its decay products overlap. The current analysis strategy requires these decay products (b-tagged track jets in the case of the merged event selection ($E_{\mathrm{T}}^{\mathrm{miss}} > 500$ GeV)) to be resolved, which results in a loss of signal acceptance in these high-mass scenarios. In Section 8.3.1 a new reconstruction technique is presented, Variable Radius track jets, which improves the b-tagging efficiency in boosted topologies.

7.2 Systematic uncertainities

Systematic uncertainties arise from sources related to the reconstruction of the various physics objects used in the mono-$h(b\bar{b})$ search, and additionally from theoretical uncertainties related to predictions of both the backgrounds and signals. These uncertainties manifest themselves as uncertainties both in the overall yield and shape of the final observable, the dijet invariant mass m_{bb}, used to perform the final search.

[1] Artificial dataset defined such that when it is used to evaluate the maximum-likelihood estimators of the likelihood parameters, the results are the input parameters themselves.

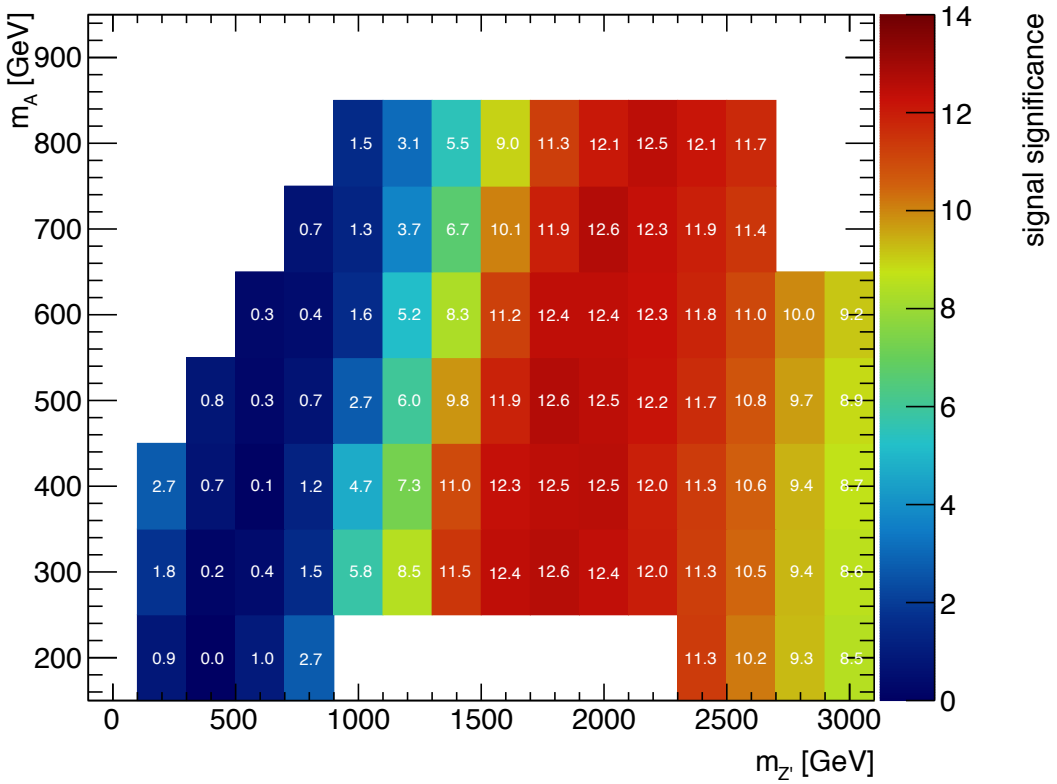


Figure 7.1 – Expected signal significances for the Z'-2HDM simplified model with different mediator masses. For each mass-point, a signal cross-section of 10 fb is used to construct the Asimov dataset [233].

A full summary of all uncertainties applied in this analysis is given in Table 7.2, along with the shorthand name of the systematic used throughout the analysis.

Systematic uncertainty	Short description	Reference
	Event	
Luminosity	uncertainty on the total integrated luminosity	
	Electrons	
EL_EFF_Trigger_TOTAL_1NPCOR_PLUS_UNCOR	trigger efficiency uncertainty	
EL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR	reconstruction efficiency uncertainty	
EL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR	ID efficiency uncertainty	
EL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR	isolation efficiency uncertainty	
EG_SCALE_ALL	energy scale uncertainty	
EG_RESOLUTION_ALL	energy resolution uncertainty	
	Muons	
MUON_EFF_TrigSystUncertainty MUON_EFF_TrigStatUncertainty	trigger efficiency uncertainties	
MUON_EFF_STAT MUON_EFF_SYS	reconstruction and ID efficiency uncertainty for $p_{\mathrm{T}} > 15$ GeV	
MUON_EFF_STAT_LOWPT MUON_EFF_SYS _LOWPT	reconstruction and ID efficiency uncertainty for $p_{\mathrm{T}} < 15$ GeV	
MUON_ISO_STAT MUON_ISO_SYS	isolation efficiency uncertainty	
MUON_TTVA_STAT MUON_TTVA_SYS	track-to-vertex association efficiency uncertainty	
MUONS_SCALE	energy scale uncertainty	
MUONS_ID	energy resolution uncertainty from inner detector	
MUONS_MS	energy resolution uncertainty from muon system	
	Small-R Jets	
JET_GroupedNP	energy scale uncertainty split into 3 components	
JET_SR1_JET_EtaIntercalibration_NonClosure	non-closure in the jet response at $2.4 < \|\eta\| < 2.5$	
JET_SR1_JER_SINGLE_NP	energy resolution uncertainty	
FT_EFF_EIGEN_B	b-tagging efficiency uncertainties ("BTAG_MEDIUM):	
FT_EFF_EIGEN_C FT_EFF_EIGEN_L	3 components for b-jets, 4 for c-jets and 5 for light jets	
FT_EFF_EIGEN_extrapolation	b-tagging efficiency uncertainty on the extrapolation on high p_{T}-jets	
FT_EFF_EIGEN_extrapolation_from_charm	b-tagging efficiency uncertainty on τ-jets	
	Large-R Jets	
JET_Comb_Baseline_Kin JET_Comb_Modelling_Kin JET_Comb_TotalStat_Kin JET_Comb_Tracking_Kin	energy scale uncertainties (p_{T} and mass scales fully correlated)	
	Track-Jets	
FT_EFF_EIGEN_B	b-tagging efficiency uncertainties ("BTAG_MEDIUM):	
FT_EFF_EIGEN_C FT_EFF_EIGEN_L	3 components for b-jets, 4 for c-jets and 5 for light jets	
FT_EFF_EIGEN_extrapolation	b-tagging efficiency uncertainty on the extrapolation on high p_{T}-jets	
FT_EFF_EIGEN_extrapolation_from_charm	b-tagging efficiency uncertainty on τ-jets	
	$E_{\mathrm{T}}^{\mathrm{miss}}$-Trigger and $E_{\mathrm{T}}^{\mathrm{miss}}$-Terms	
METTrigStat METTrigSyst	trigger efficiency uncertainty	
MET_SoftTrk_ResoPerp	track-based soft term related to transversal resolution uncertainty	
MET_SoftTrk_ResoPara	track-based soft term related to longitudinal resolution uncertainty	
MET_SoftTrk_Scale	track-based soft term related to longitudinal scale uncertainty	
MET_JetTrk_Scale	track MET scale uncertainty due to tracks in jets	

Table 7.2 – Qualitative summary of the experimental systematic uncertainties considered in this analysis. The $E_{\mathrm{T}}^{\mathrm{miss}}$ uncertainty information is broken in some of the available simulation samples and must be updated in the next iteration of the analysis.

7.2.1 Experimental systematic uncertainties

This systematic uncertainties arise from sources related to the reconstruction, identification and calibration of objects in the mono-$h(b\bar{b})$ search. The integrated luminosity corresponds to 36.1 fb^{-1} with an uncertainty of 3.4%

7.2.1.1 $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger

The uncertainty on the modelling of the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger uncertainty in MC simulation is taken as a statistical component from the scale factor, SF, fit and a systematic uncertainty from the topology dependence. The data-driven $E_{\mathrm{T}}^{\mathrm{miss}}$ Trigger SF is discussed in Section 6.3.2.1. Three sources for variations are considered in the systematic uncertainty: the b-tag multiplicity, the fit range of the SF fit (starting from 100 GeV to 120 GeV in $E_{\mathrm{T}}^{\mathrm{miss}}$), and the background composition (either W + jets or $t\bar{t}$). The trigger systematic uncertainty is chosen as the variation with the largest SF fit absolute difference.

7.2.1.2 Electron

Scale factors are used to take into account electron efficiencies in MC simulations with respect to data, similar to the procedure used for $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger difference of efficiencies. This SF and its associated uncertainties are evaluated for reconstruction, isolation, identification and trigger (refer to Table C.1 for the 2ℓ-CR).

Uncertainties in the electron reconstruction and identification scale, and the energy scale and resolution are considered.

7.2.1.3 Muon

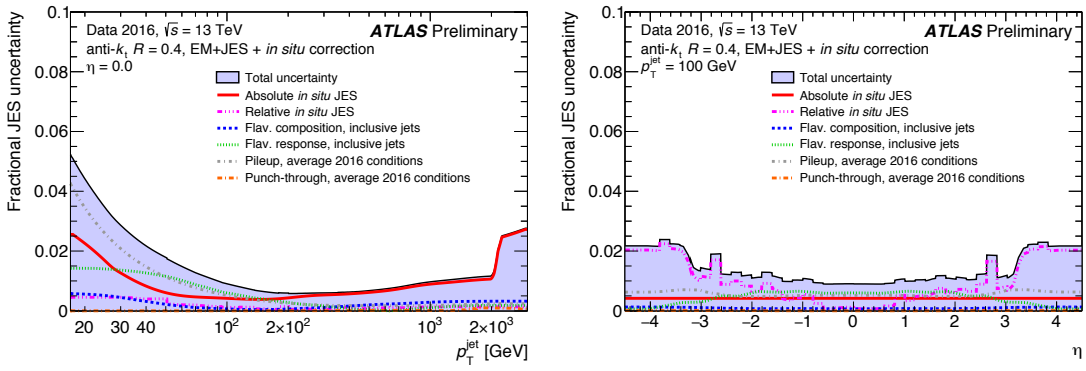
In the same way a it is done for electrons, scale factors correction are considered in the analysis. These SFs and uncertainties are determined for muon reconstruction, identification, track-to-vertex association efficiencies and trigger (refer to Table C.1 for the 2ℓ-CR).

7.2.1.4 Small-R jet

The full jet energy scale correction, described in Section 4.2.2.1, is applied, which corrects for the offset of the primary vertex from the detector origin, the non-compensation of the calorimetry, and dead material, with the final result being a jet calibrated to the energy scale of stable truth hadrons.

For every calibration technique, several sources of systematic uncertainties are considered and propagated to the calibration factors. These uncertainties are calculated using different MC generators by varying the calibration and reconstruction efficiency of the physics objects considering different event selection [148]. Figure 7.2 shows the combination of all uncertainties as a function of p_{T} for $\eta = 0$, and as a function of η for $p_{\mathrm{T}} = 100$ GeV.

The analysis considers nuisance parameters associated to the jet energy calibration which evaluate the jet energy scale (JES) and jet energy resolution (JER) uncertainties.



(a) (b)

Figure 7.2 – Combined uncertainty and its various components in the jet energy scale (JES) of fully calibrated jets as a function of 7.2(a) jet p_{T} at $\eta = 0$ and 7.2(b) η at $p_{\mathrm{T}} =$ 100 GeV. The flavor composition and response uncertainties assume a quark and gluon composition taken from PYTHIA dijet MC simulation (inclusive jets) [148].

They are obtained from in-situ calibration, eta inter-calibration, calibration for high-p_{T} jets and pile-up effects.

Also, the analysis includes uncertainties on the flavor tagging efficiency scale factors. This uncertainties are obtained separately for small-R jets in the resolved regime and for ghost-associated track jets in the merged regime, considering separated nuisance parameters for light, c- and b-quarks initiated jets [160, 161].

7.2.1.5 Large-R jet

The reconstruction uncertainties pertaining to large R jets are quantified as uncertainties on the scale and resolution of the jet p_{T} and mass. The scale uncertainties are derived using an in-situ technique, typically referred to as the "R_{track} double ratio" method, that uses the corresponding track jet as a reference by which to study the reconstructed calorimeter jet, more details of the method can be found in [154]. There are four main uncertainties obtained by this method. These are named as Baseline, Modelling, Tracking, and TotalStat. Baseline uncertainty is derived from the difference between data and Pythia 8, Modelling uncertainty comes from differences between Pythia and Herwig, Tracking is the uncertainties of the tracks that are being used, and TotalStat is the statistical uncertainty of the measurements.

In addition to scale uncertainties, uncertainties on the p_{T} and mass resolution are taken into account. This is done by performing a jet by jet smearing such that the nominal resolution is increased by a fractional amount that is equivalent to the estimated resolution uncertainty. This uncertainty is taken from Run 1 and is taken as 20%.

7.2.1.6 Tau

Since a τ lepton veto is employed in the 0 lepton selection to remove $t\bar{t}$ and W+jets background, systematics associated with τ identification and energy scale are necessary.

This comes in two forms, the first is for the standard TauID, the systematics for which are documented in Reference [173] and include uncertainties on the BDT TauID and the tau energy scale. Beyond this, the extended tau veto that is based on the number of tracks associated to a small-R jet and comes with tracking systematics.

7.2.1.7 Missing transverse momentum

Uncertainties on the reconstructed hard objects that enters into the reconstruction of $E_{\mathrm{T}}^{\mathrm{miss}}$,as described in Section 4.6, are propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation. The systematics associated to $E_{\mathrm{T}}^{\mathrm{miss}}$ come from the propagation of uncertainties on the soft terms of the $E_{\mathrm{T}}^{\mathrm{miss}}$.

7.2.2 Theoretical systematic uncertainties

This systematic uncertainties arise from sources related to theoretical predictions and modelling of background and signal.

7.2.2.1 $t\bar{t}$, W, Z modelling

A combination of methods is used to estimate the effect of the MC modelling choices. These methods are markedly different for W+jets/Z+jets and $t\bar{t}$ and are inherited from the SM VHbb analysis, since final states very similar, more details can be found in Reference [248] and Reference[245].

These uncertainties are important given that backgrounds are modeled using MC simulations and normalized to the data in the likelihood fit by using dedicated control regions. In some cases, this normalization is constrained with a prior that is dictated by the theory uncertainties on the overall cross section of the background. However, there are also shape uncertainties that affect the p_{T} distribution of the W/Z boson in the case of V+jets and the leptonic W boson in the case of $t\bar{t}$ as well as the reconstructed dijet mass spectra.

The uncertainty estimation is described as follow:

$t\bar{t}$ production

The uncertainties on the shape of the backgrounds is evaluated by the comparisons of different MC generators with various ME and PS/hadronization models as described in [249]. These comparisons are performed by comparing the shape of the $p_T(W) = p_T(\ell, E_{\mathrm{T}}^{\mathrm{miss}})$ spectrum between various MC generators as well as the $m(bb)$ spectrum. These two observables are chosen on account of the fact that the final combined fit is performed in the two dimensional space parameterized by these two observables. For the comparisons, the various generators are normalized to unity area to isolate shape differences and the maximal deviation is parameterized by an analytical function and taken as the $\pm 1\sigma$ variation of the $t\bar{t}$ shape. These functions are then used to provide event-by-event scale factors on the $t\bar{t}$ background to derive systematically varied templates for the final combined fit. $t\bar{t}$ background is studied in the dedicated 1 lepton control region which is described in section 6.5.2.

Z+jets

In this case, the two lepton control region is used as a standard candle that is pure in Z+jet events that are qualified to be Z+HF according to the labelling scheme defined in this analysis and where we assume that all detector effects have been corrected by calibrations derived by dedicated combined performance studies. Details on the control regions are discussed in Section 6.5. In this case, any deviation that is observed is attributed to a mismodeling of the underlying Monte Carlo simulation model and taken as the $\pm 1\sigma$ variation of the Z+jets shape. As in the case with $t\bar{t}$, these comparisons are performed in the $p_T(V)$=$p_T(\ell,\ell)$ and $m(bb)$ spectra and the difference between the shape of the data and MC is parameterized by an analytical function. These functions are then used to provide event-by-event scale factors on the Z+jets background to derive systematically varied templates for the final combined fit.

W+jets

Given that the purity of the control regions is not as high as in $Z \to ee$ +jets, it is necessary to again use comparisons between various Monte Carlo generators. As in the case with $t\bar{t}$, these comparisons are performed in the $p_T(V)$=$p_T(\ell, E_{\mathrm{T}}^{\mathrm{miss}})$ and $m(bb)$ spectra and the difference between the shape of the data and MC is parameterized by an analytical function. These functions are then used to provide event-by-event scale factors on the W+jets background to derive systematically varied templates for the final combined fit. The above mentioned calculation of the modeling systematics is based on [246].

7.2.2.2 Signal acceptance uncertainties

The uncertainty on the experimental acceptance for the Mono-H signal due to the modelling of the production is evaluated by varying the parameters of the MADGRAPH + PYTHIA samples and comparing the results by applying the analysis selection at generator level. The following variations are considered:

Renormalisation and factorisation scales:

The default renormalisation and factorisation scales in MadGraph are dynamically set to $M_T^2 + p_T^2$ where M_T and p_T are the transverse mass and transverse momentum of the final state particles, respectively. The event-by-event scales are changed by a factor of 2 coherently

Eigentune variations:

Uncertainty sources on the final state radiation, initial state radiation and multi-parton interactions are parameterized as a subset of tune variations providing maximal variation coverage for underlying event effects, jet structure effects and additional jet production.

Parton showering and hadronisation PDFs:

The nominal NNPDF2.3LO PDF [237] is replaced by the MSTW2008lo68cl PDF [250] and the CTEQ6L1 PDF [251]. The uncertainty is evaluated by taking the envelope of the largest deviation from the two PDFs. The differences between the PDFs from different groups are generally larger than the set of uncertainty eigenvectors from a single PDF set.

The signal samples with the above variations were generated for several mass points of the Z'+2HDM model. Requirements were made on the truth level quantities such as lepton veto, MET and the jets to replicate the cuts of the analysis. The change in signal acceptance due to each source of uncertainty was calculated for the 4 MET bins of the analysis for each of the considered mass points. The upward and downward changes in acceptance are then symmetrized.

The change in signal acceptances are listed in Tables C.4 to C.9 for variations of the scales, parton-shower tunes and PDFs. The variations generally each lead to changes in the acceptance of $< 10\%$, although they can be larger for regions with low acceptance at either low or high $E_{\mathrm{T}}^{\mathrm{miss}}$ depending on the type of model and masses. The final fit described in the next chapter makes use of the uncertainties listed in this section.

7.3 Mono-$h(b\bar{b})$, search for Dark Matter results

7.3.1 Nuisance Parameters in the fit

After fitting the statistical model to data, the nuisance parameters considered can differ from the prior value and one says that the parameter is "pulled". This pull of a nuisance parameter is defined as $(\theta_{\mathrm{fit}} - \theta_0)/\Delta\theta$, where θ_0 is the prior value, $\Delta\theta$ is the uncertainty on the nuisance parameter, and θ_{fit} is the value maximising the profile likelihood function ($\hat{\theta}$ in Section 7.1.1 notation).

The pulls and constraints of the nuisance parameters after fitting to the observed data are shown in Figure 7.3. In particular, the postfit values for the nuisance parameters associated to the dominant background normalisations are shown in Table 7.3.

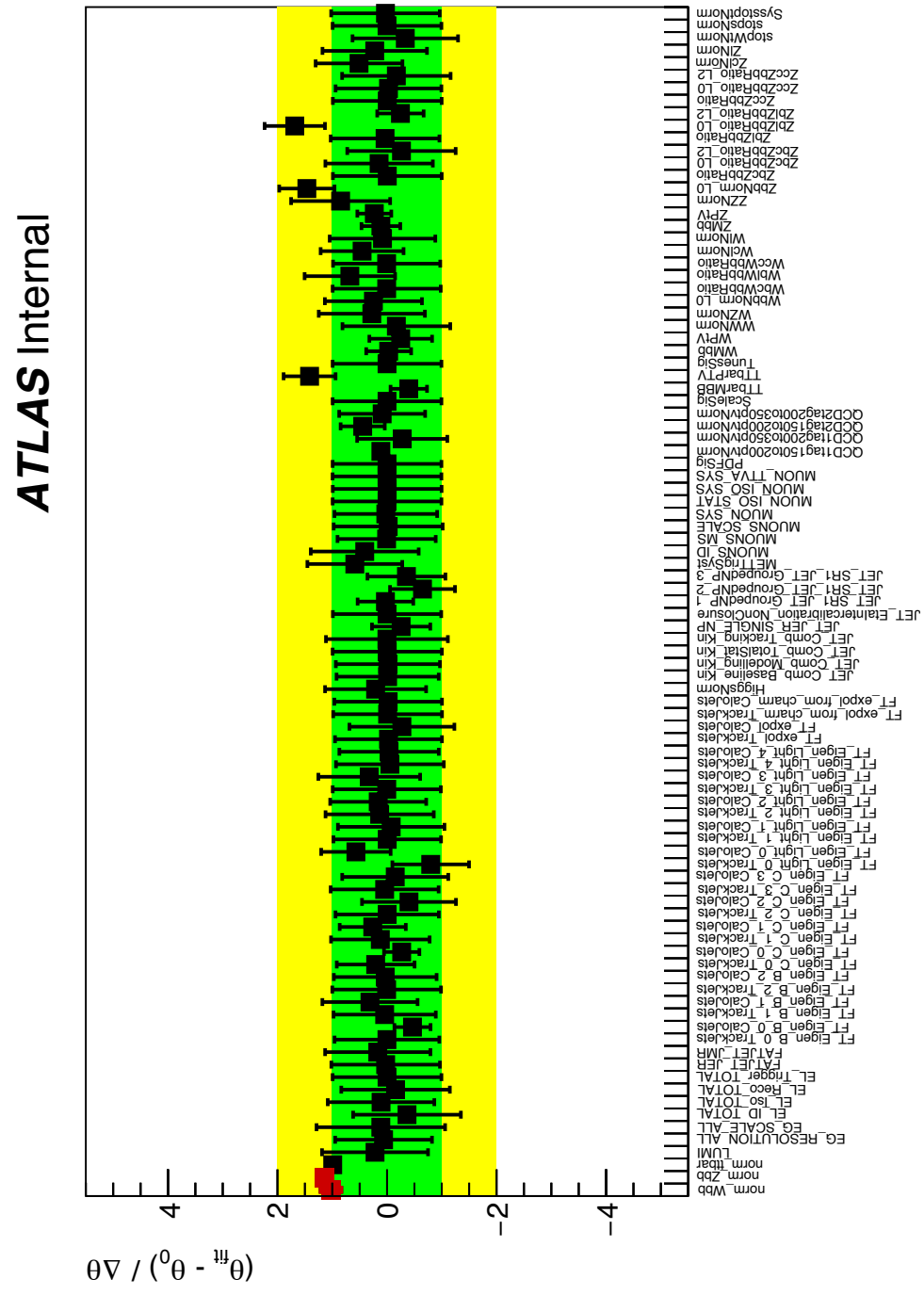


Figure 7.3 – The nuisance parameter pulls for the fit to the observed data in the sidebands of the discriminating variable, assuming $\mu = 0$. The major background normalisations are summarized in Table 7.3 [233].

Large pulls with more than one standard deviation, σ, are observed associated to two sources of uncertainty.

One is related to the uncertainty on the shape of the p_{T}^{V} spectrum for $t\bar{t}$ events. This nuisance parameter is named "TTbarPTV" in Figure 7.3. In the mono-$h(b\bar{b})$ search, the $E_{\mathrm{T}}^{\mathrm{miss}}$ or modified $E_{\mathrm{T}}^{\mathrm{miss}}$ spectrum for $t\bar{t}$ samples, using POWHEG+PYTHIA6, deviates from the expectation by an amount of about 1.5σ. Compared to theory, the data prefers a softer spectrum. Therefore, shape deviations from the nominal in $t\bar{t}$ events for p_{T}^{V} ($E_{\mathrm{T}}^{\mathrm{miss}}$ or modified $E_{\mathrm{T}}^{\mathrm{miss}}$) are not covered within "TTbarPTV" uncertainties for the event selection used in this analysis, thus the uncertainty derived for the search of a SM Higgs boson decaying into $b\bar{b}$ when produced in association with a vector boson [246] is underestimated for mono-$h(b\bar{b})$ analysis.

The other large pulls are related to the normalisation of HF Z + jets contributions in the zero lepton SR channel. They are called "ZbbNorm_L0" and "ZblZbbRatio_L0" in Figure 7.3. The first one comes from the relative uncertainty of HF Z + jets between 0 lepton SR and 2ℓ-CR, which is again underestimated with respect to the study in References [246, 245]. The other one, "ZblZbbRatio_L0", corresponds to the uncertainty on the relative flavour components of Z + bl to HF Z + jets in the 0 lepton SR, which is considered inclusively across 1 and 2 b-tag categories. This is expected since the Z + bl component dominates in the 1 b-tag, while is subdominant in the 2 b-tag category. Because of this, the acceptance for Z + bl differs between 1 and 2 b-tags with respect to HF Z + jets. This tension can be solved with a decorrelation between the 1 and 2 b-tag categories for the ratio of Z + bl to HF Z + jets.

Normalisation NP	Postfit value
HF W + jets	1.07±0.20
HF Z + jets	1.14±0.07
$t\bar{t}$	0.99±0.03

Table 7.3 – Selected nuisance parameters derived from the combined fit that includes the pulls summarized in 7.3. Note that by construction the background normalisation prefit-values are 1.0 (no constraint).

7.3.2 Impact of Uncertainties

The final impact of uncertainties can be evaluated with the fractional uncertainty effect on the signal strength, $\sigma_{\mu}/\hat{\mu}$. In order to estimate the impact of different uncertainties, the statistical model, including a particular signal, is fitted to the Asimov dataset. The postfit fractional uncertainty on the signal strength is evaluated for three representative mass parameters of the Z'+2HDM model in the low, medium and high $E_{\mathrm{T}}^{\mathrm{miss}}$ range: $(m_A, m_{Z'})$/ TeV $= (0.6, 0.3), (1.4, 0.6), (2.6, 0.3)$. Also, these signals are interesting since they are at the edge of the expected exclusion region of this search, as will be shown in Section 7.3.4.

The total uncertainty is derived from a fit including all uncertainties. The statistical uncertainty is estimated in another fit where all systematic uncertainties are neglected.

Then, the impact of each systematic uncertainty, θ_i is estimated by repeating the fit excluding the systematic uncertainty in question. This results in a reduced postfit signal yield uncertainty $\sigma_{\mathrm{No}\theta_i}$, which is subtracted quadratically from the total uncertainty: $\sigma_{\theta_i} = \sqrt{\sigma^2_{\mathrm{tot}} - \sigma^2_{\mathrm{No}\theta_i}}$. In general, multiple related sources of uncertainty are grouped for each fit, e.g. all the jet energy resolution JER eigenvectors components are grouped.

Table 7.4 presents the dominant sources of uncertainty with the corresponding fractional uncertainty on the postfit signal yield after fitting to Asimov data for the three benchmark signals.

The dominant sources of experimental systematic uncertainty arise from: the Monte Carlo simulation statistics, the calibration of the b-tagging efficiency and integrated luminosity, as well as the energy scale and resolution, and the mass of jets. The dominant theoretical systematic uncertainty comes from the V + jets background modelling. Also, Table 7.4 shows that this search for dark matter is limited by the finite amount of data statistics for $E_{\mathrm{T}}^{\mathrm{miss}} \gtrsim 300$ GeV.

7.3.3 Post-fit Distributions

Figure 7.4 shows the data and MC simulation for the muon charge in the 1μ-CR for the 2-b tag category in the four $E_{\mathrm{T,no}\mu}^{\mathrm{miss}}$ regions after fitting to the observed data. As expected, the muon charge distribution is symmetric for $t\bar{t}$ events, while it is not for W + jets due to a predominance of μ^+ as expected in proton-proton collisions. Figure 7.5 shows data and MC simulation for the event yield, showing one bin, in the 2ℓ-CR, 2-$b$tag, in the four $P_{\mathrm{T}}^{\ell\ell}$ regions. Distributions in both Figures 7.4, 7.5 are within the Higgs candidate mass window of (50,280) GeV.

Finally, postfit distributions in the 0 lepton SR, 2-$b$tags, of the Higgs candidate invariant mass is shown in Figure 7.6. Good agreement is observed between the data and the background simulation in all channels.

7.3.4 Exclusion contour limits

Since no significant deviation from SM prediction is observed, the results are interpreted as exclusion limits in light the of the Z'-2HDM simplified model described in Section 2.2.2.

The limits are set using the $\mathrm{CL_s}$ formalism [244]. If the p-value is found below a specific threshold α, which is taken to be $\alpha = 0.05$, then the value of μ is excluded at a confidence level (CL) of $1-\alpha$. The set of points not excluded form a confidence level interval with $\mathrm{CL} = 1 - \mathrm{ff}$. The test statistic q_μ for upper limits of Equation 7.8 can be solved analytical since it follows a χ^2 distribution in the asymptotic limit of large sample [247].

The signal plus background hypothesis for the Z'-2HDM simplified model is excluded at 95% confidence level. The most informative presentation of results is the two dimensional space of ($m(Z')$,$m(A)$) with the mass of the dark matter particle fixed to 100 GeV.

The limits are obtained in terms of the upper limit on the production cross section of h+dark matter times the $h \to b\bar{b}$ branching ratio, $\sigma(pp \to h\chi\chi)Br(h \to bb)$, for all of the different signal points. This is done by determining the mass point at which the

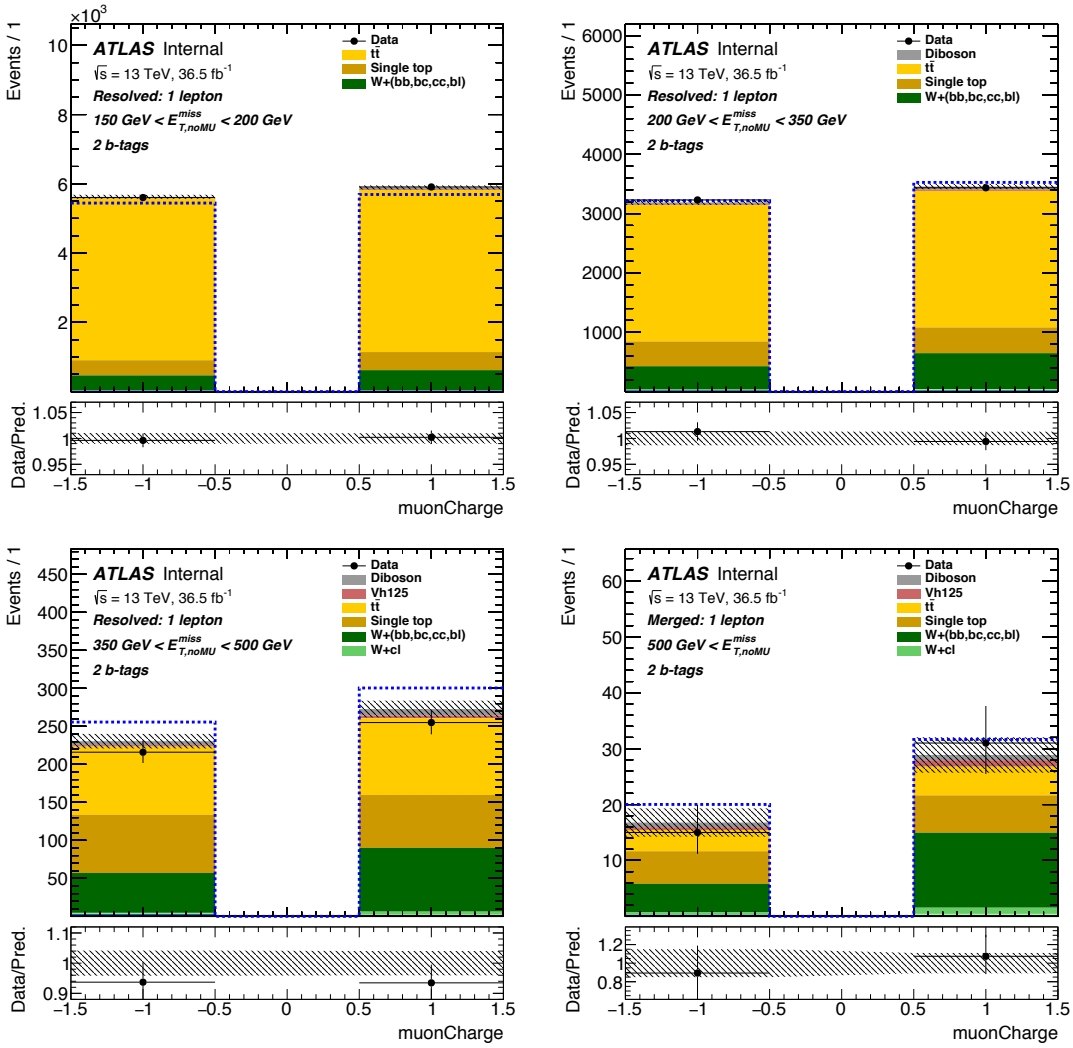


Figure 7.4 – Distributions of the muon charge for the 1 muon control region for 2 tag events. The upper panels show a comparison of data to the SM expectation before (dashed lines) and after the fit (solid histograms). The lower panels display the ratio of data to SM expectations after the fit, with its systematic uncertainty [233].

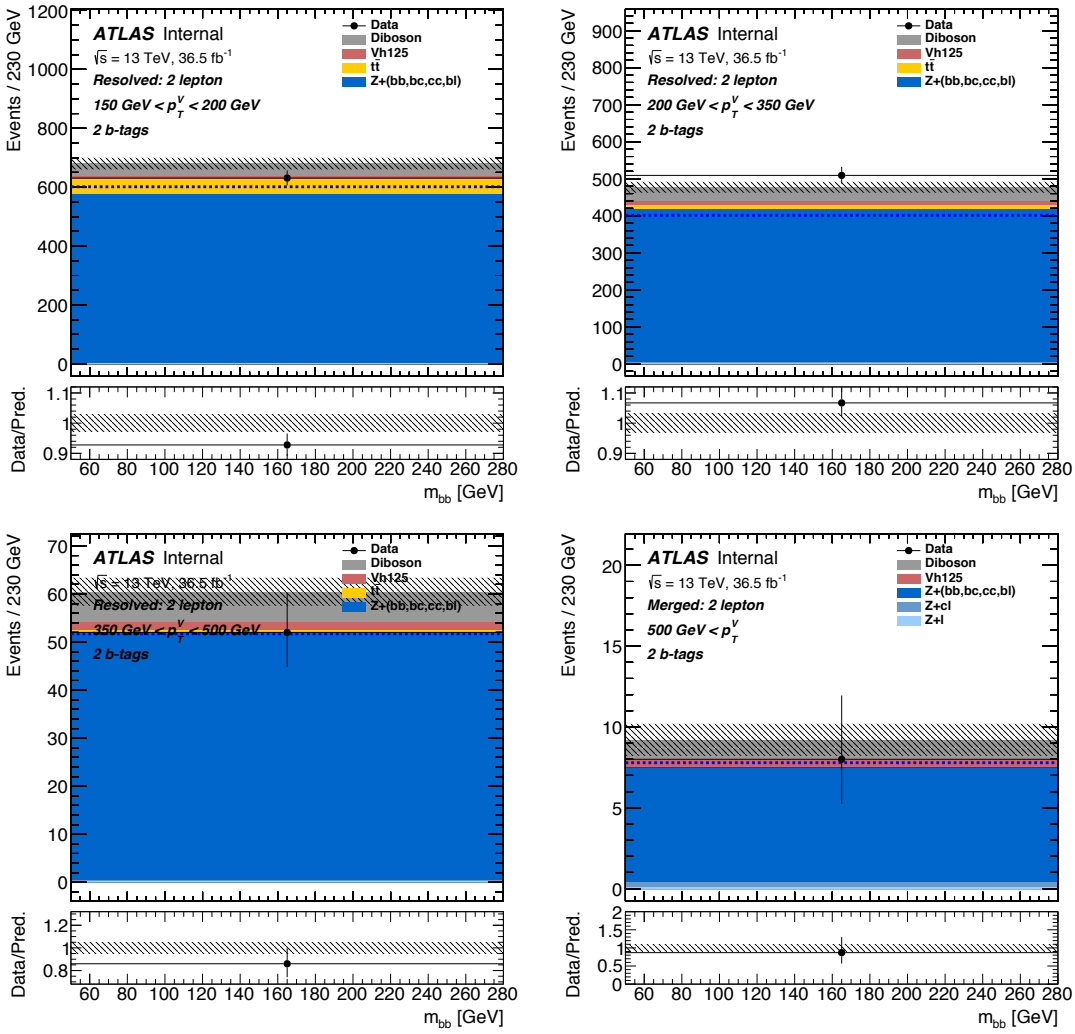


Figure 7.5 – One bin distribution of the invariant mass of the two signal jets for the 2 lepton control region for 2 tag events. The upper panels show a comparison of data to the SM expectation before (dashed lines) and after the fit (solid histograms). The lower panels display the ratio of data to SM expectations after the fit, with its systematic uncertainty [233].

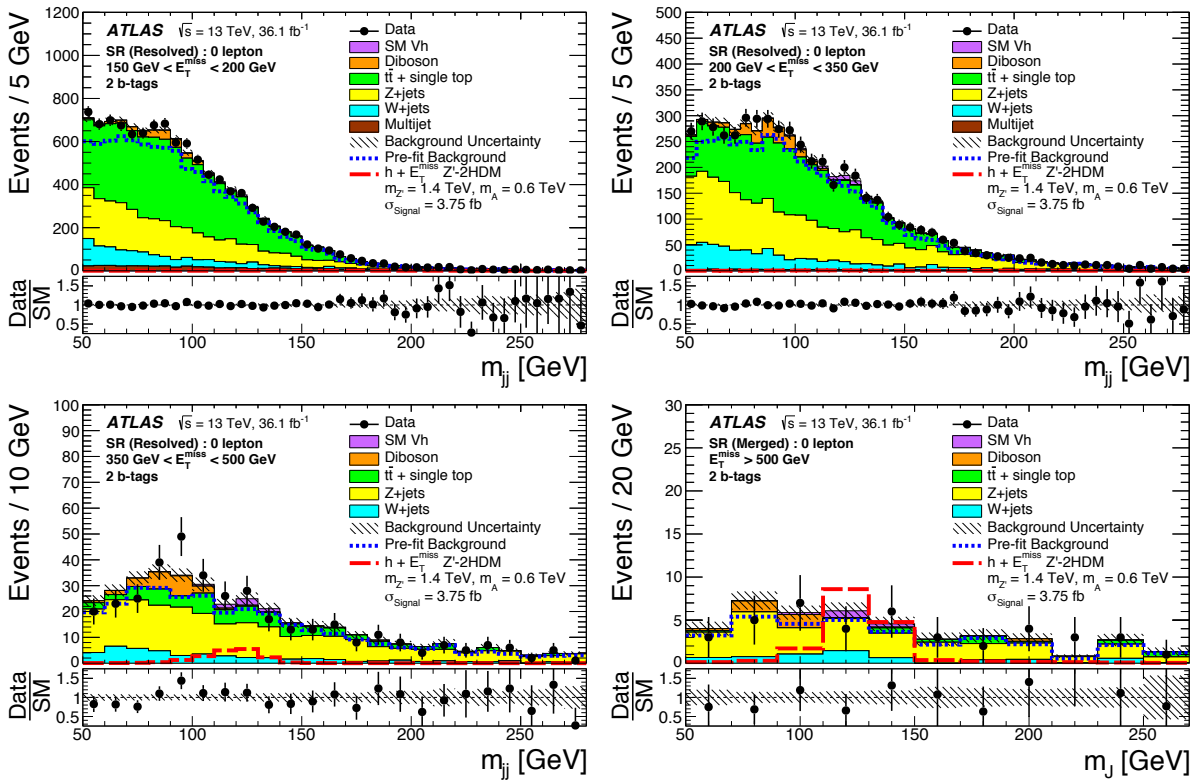


Figure 7.6 – Distributions of the invariant mass of the Higgs boson candidates $m_H = m_{jj}, m_J$ with 2 b-tags in the SR for the four $E_{\mathrm{T}}^{\mathrm{miss}}$ categories that are used as inputs to the fit. The upper panels show a comparison of data to the SM expectation before (dashed lines) and after the fit (solid histograms) with no signal included. The lower panels display the ratio of data to SM expectations after the fit, with its systematic uncertainty considering correlations between individual contributions indicated by the hatched band. The expected signal from a representative Z'-2HDM model is also shown (long-dashed line) [252].

Source of uncert.	Impact [%]		
	(a)	(b)	(c)
V + jets modeling	5.0	5.7	8.2
$t\bar{t}$, single-t modeling	3.2	3.0	3.9
SM $Vh(b\bar{b})$ norm.	2.2	6.9	6.9
Signal modeling	3.9	2.9	2.1
MC statistics	4.9	11	22
Luminosity	3.2	4.5	5.4
b-tagging, track-jets	1.4	11	17
b-tagging, calo jets	5.0	3.4	4.7
Jets with $R = 0.4$	1.7	3.8	2.1
Jets with $R = 1.0$	<0.1	1.2	4.7
Total syst. uncert.	10	21	36
Statistical uncert.	6	38	62
Total uncert.	12	43	71

Table 7.4 – Dominant sources of uncertainty for three representative scenarios after the fit to data: (a) with $(m_{Z'}, m_A) = (0.6\ \text{TeV}, 0.3\ \text{TeV})$, (b) with $(m_{Z'}, m_A) = (1.4\ \text{TeV}, 0.6\ \text{TeV})$, and (c) with $(m_{Z'}, m_A) = (2.6\ \text{TeV}, 0.3\ \text{TeV})$. The effect is expressed as the fractional uncertainty on the signal yield. Total is the quadrature sum of statistical and total systematic uncertainties.

theoretical cross section is crossed by the expected cross section upper limit. This is shown in Figure 7.7 for $m_{Z'} = 800$ GeV on the left or $m_A = 500$ GeV on the right. From these plots for all the signal points, one can determine lower mass limits to be translated into a single point in the $(m_{Z'}, m_\chi)$ plane.

This analysis leads to the set of exclusion points which are then translated into the final result in the $(m_{Z'}, m_A)$ space in Figure 7.8. Limits exclude $m_{Z'}$ up to 2.6 TeV and m_A up to 0.6 TeV, substantially extending previous limits [213, 107, 1, 253, 254].

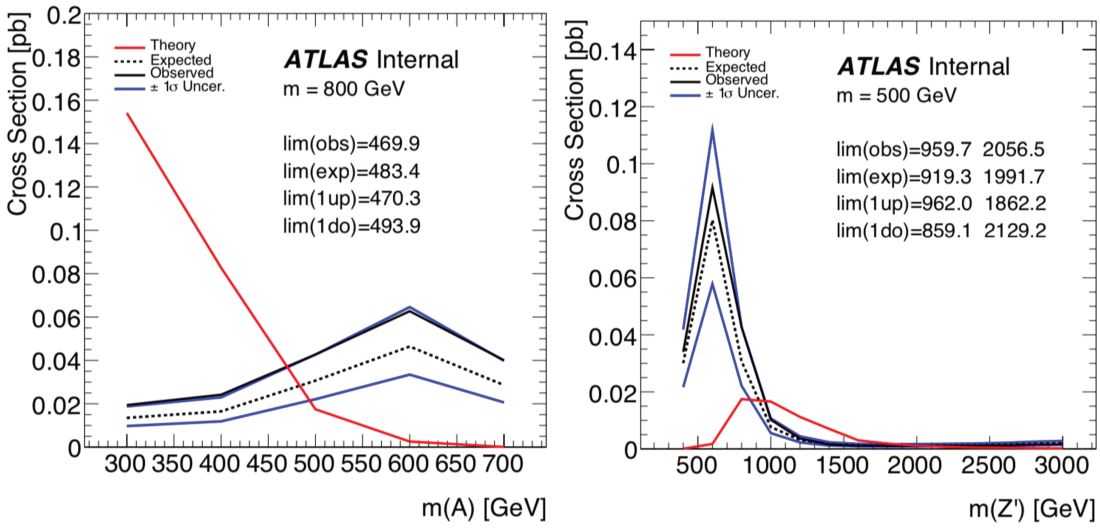


Figure 7.7 – Cross section limits determined for the Z'-2HDM model shown in logarithmic scale upon fixing $m_{Z'} = 800$ GeV on the left or $m_A = 500$ GeV on the right. Expected and observed limits are shown with dashed and solid black line respectively. Theory cross section is shown as a red solid line.

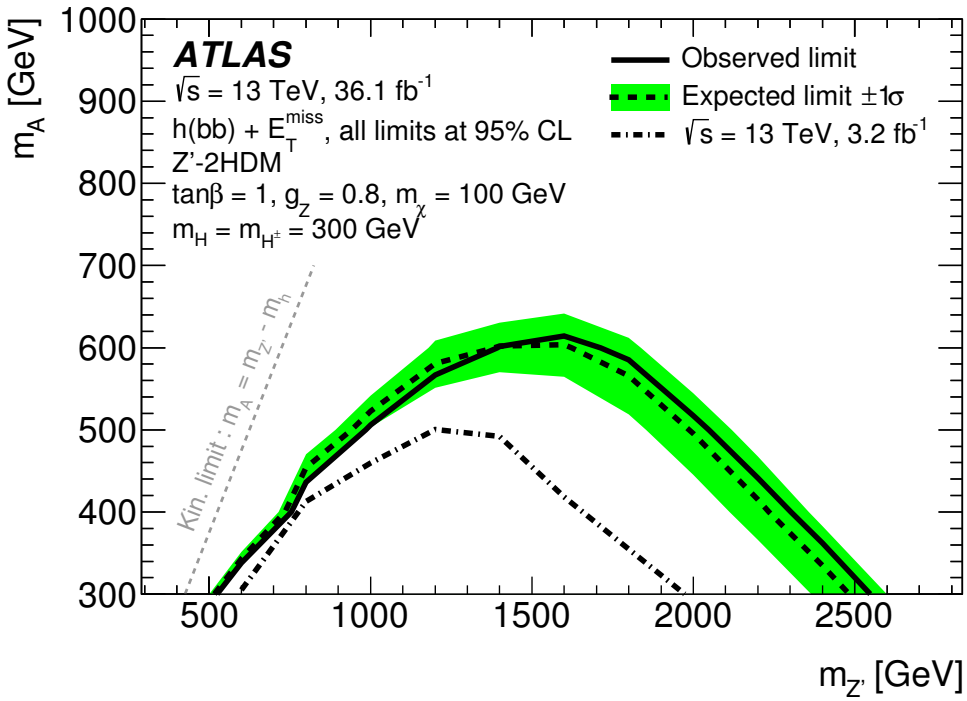


Figure 7.8 – Exclusion contours for the Z'-2HDM scenario in the $(m_{Z'}, m_A)$ plane for $\tan\beta = 1$, $g_{Z'} = 0.8$, and $m_\chi = 100$ GeV. The observed limits (solid line) are consistent with the expectation under the SM-only hypothesis (dashed line) within uncertainties (filled band). Observed limits from previous ATLAS results at $\sqrt{s} = 13$ TeV (dash-dotted line) [1] are also shown [252].

Chapter 8

Mono-$h(b\bar{b})$: search for Dark Matter with improved analysis and $79.8\,fb^{-1}$ of 2015, 2016 and 2017 data at $\sqrt{s} = 13$ TeV

8.1 New analysis improvements

The analysis described in this Chapter provides an improved search for dark matter production in association with a Higgs boson decaying into pairs of bottom quarks with respect to the one presented in Chapter 7.

This new analysis exploits improved analysis techniques in the area of object reconstruction and performance. In addition, a better sensitivity is expected from the extension of the amount of collision data from $36\,fb^{-1}$ to $79.8\,fb^{-1}$, with the addition of data collected during 2017.

The signal model used in the interpretation of the results is the Z'-2HDM model described in Section 2.2.2 as in the previous analysis, presented on Chapter 7, making it easier to quantify the improvement of the analysis sensitivity.

The sensitivity of the search strongly relies on the ability of identifying the b-jets coming from the Higgs decay, since this is a powerful way to suppress background from QCD and V + light jets production. In this search of dark matter production, high energies are probed for which the Higgs boson is typically boosted and the hadronisation products of the b-quarks from its decay may overlap. This is introduced in more detail on Section 6.1.

As described in Section 4.2.4, jets reconstructed from tracks are used for the b-jet identification in the merged regime. However, when these track jets are reconstructed with standard algorithms using a fixed radius parameter [164], they may not always be separated resulting in an inefficiency of the double b-tagging requirement.

This problem is overcome by the use of a new reconstruction technique: Variable-Radius (VR) track jets, i.e. track jets reconstructed with an algorithm in which the radius parameter decreases with increasing jet transverse momentum [255, 256]. This allows better separation and individual b-tagging of the two jets coming from a highly

boosted Higgs boson and thus increases the sensitivity for large masses of the mediator, $m_{Z'}$, as shown in Figure 7.8.

On the other hand, in events with low $E_{\mathrm{T}}^{\mathrm{miss}}$, typically produced if $m_{Z'}$ is low, multijet background has a sizable contribution, see Section 6.5.4 for details. Due to the increase of integrated luminosity, the size of multijet background is not be negligible and a more robust way to evaluate its uncertainty is needed.

An important characteristic of this background, apart that is difficult to model in MC simulation due to its large cross-section and rejection, is that it leads to fake $E_{\mathrm{T}}^{\mathrm{miss}}$ coming from jet p_{T} mismeasurement. The object-based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance [4], presented in detail on Chapter 5, helps to identify events in which the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ is more consistent with fake detection for the multijet background or is more likely to come from weakly interacting particles like the dark matter signal associated to real $E_{\mathrm{T}}^{\mathrm{miss}}$. The object-based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance is used in this analysis to efficiently reduce the background contribution from multijet events.

8.2 Data samples and trigger

This section describes the samples from data in Section 8.2.1, and simulated signal and background processes in Section 8.2.2 [1].

8.2.1 New dataset and trigger

This search uses proton-proton collision data recorded by the ATLAS detector at a centre-of-mass energy of 13 TeV, collected during 2015, 2016 and 2017 with a bunch crossing spacing of 25 ns.

Events are used in the analysis only when all relevant parts of the detector are in good operating conditions, therefore only data passing quality criteria are considered, and the corresponding runs are listed in the GRL, see Appendix C.1 for details. The resulting dataset corresponds to an integrated luminosity of 3.2 fb^{-1}, 32.9 fb^{-1}, and 43.6 $fb^{-1}$for 2015, 2016 and 2017, respectively, giving a total integrated luminosity of 79.8 fb^{-1}.

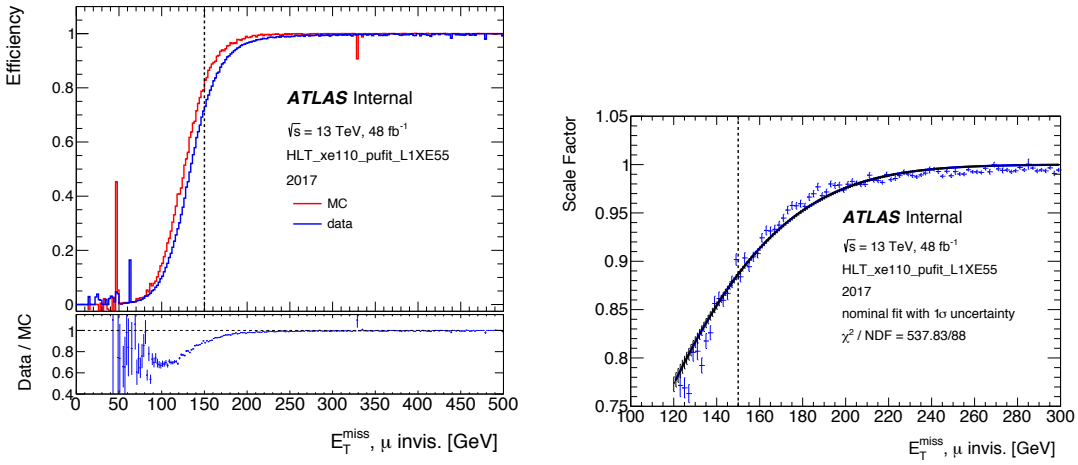
This analysis considers the same event categories distinguished in the lepton multiplicity (0 lepton SR, 1μ-CR, 2ℓ-CR) described in Section 6.5, and the same trigger strategy of Section 6.3.2. Events for the regions containing less than two signal leptons (0 lepton SR and 1μ-CR) are selected by triggering on signatures with large $E_{\mathrm{T}}^{\mathrm{miss}}$ in the calorimeters. The online $E_{\mathrm{T}}^{\mathrm{miss}}$ threshold increased from 70 GeV for 2015 data to 110 GeV in 2017 data [257, 258].Events for regions with two leptons (2ℓ-CR) are selected by single lepton triggers with the lowest unprescaled thresholds at different levels of identification and isolation requirements [257, 258]. Details on the triggers keys used can be found in Table C.2.

This analysis uses a lower offline $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement of 150 GeV, where some of the triggers are not fully efficient. As it is done for previous results in Section 6.3.2.1, data-

[1] Data and simulation were reconstructed with the same version (Release 21) of the ATLAS reconstruction software (Athena).

driven correction factors for simulated samples are derived in order to match efficiencies from data.

The efficiency measurements for 2015 and 2016 triggers are similar to the ones shown in Figure 6.3. Figure 8.1 shows the measured trigger efficiencies in data and MC for the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger used in 2017 on the left. The obtained scale factors, as defined in Equation 6.2, along with the ratio of the functions fitted to the individual data and MC turn-on curves are shown on the right.



(a) (b)

Figure 8.1 – 8.1(a): Measured trigger efficiency as a function of offline $E_{\mathrm{T,no}\mu}^{\mathrm{miss}}$ in data and MC for the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger used in 2017. The lower panels provide the ratio of data and MC events (the scale factor). 8.1(b): Measured scale factors as a function of offline $E_{\mathrm{T,no}\mu}^{\mathrm{miss}}$ for the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger used in 2017. The hatched band shows the 1σ fit uncertainty. Both plots are shown for 0,1 and 2 b-tags together [259].

8.2.2 Monte Carlo samples for signal and electroweak backgrounds

Monte Carlo simulated samples are produced for both the signal and background processes processed with the GEANT 4 [260] based ATLAS detector simulation [194].

Standard Model background simulated samples are described in more details in Section 6.3.3.2. The dominant backgrounds for the search presented here are V+jets (with $V = W$ or Z) and $t\bar{t}$ events as explained in Section 6.5. W/Z+jets events are generated using the SHERPA 2.2.1 event generator [199] with the NNPDF3.0 PDF set [235] and filters to separate those with light, charm, and bottom jets in the final state. Details of the generator configurations can be found in Reference [261]. The difference with respect to the V+jets samples used in Chapter 7 is that, in order to obtain a larger number of events in regions with a boosted vector boson, samples are split not only according to the p_{T} of the vector boson (W,Z) but also on the H_{T} of the event, introducing a cut at

generation level, and producing samples for different slices in $\max(\mathrm{H_T}, \mathrm{p_T^V})$. Here H_T is defined as the sum p_T of central small-R jets with $p_\mathrm{T} > 35$ GeV.

The diboson background samples use the same generator setup (except for the flavour filtering) as for the V+jets samples [262].

Top-quark pair and single production events are generated using the POWHEG [263, 264, 265, 266] generator interfaced with PYTHIA 8 [238, 267] (it was PYTHIA 6 in Section 6.3.3.2) Details on the configurations can be found in Reference [249].

For the signal, the same simulation as for the previous result [252], described in more detail on Section 6.3.3.1, is used: Signal events based on the Z'-2HDM simplified model are produced at tree-level using the MadGraph event generator interfaced to PYTHIA 8.

8.3 Reconstructed objects improvements

Various physics objects are used to categorise events and sort them into different regions used in the analysis. The reconstruction and definition of the objects used in this search is outlined in Section 6.2 and summarized in Table 6.4.

The main difference with respect to previous analysis result in Chapter 7, is the use of VR track jets, instead of Fixed Radius (FR) track jets introduced in Section 4.2.4 with a selection criteria described in Section 6.2.3.

Another difference with respect to previous result is that the working point at which the b-tagging efficiency is 77% for small-R jets originating from a b-quark is chosen. While the working point used before was 70%.

Compared to the previous result of the mono-$h(b\bar{b})$ search described in Chapter 7, there are two significant improvements enhancing the sensitivity: VR track jets and the object-based $\boldsymbol{E}_\mathrm{T}^\mathrm{miss}$ significance described in Section 8.3.1 and Section 8.3.2, respectively.

8.3.1 Including Variable Radius track jets

This new technique is very helpful when it comes to the identification of boosted Higgs boson decays. At high Higgs boson Lorentz boost in the merged regime, introduced in Section 6.1.2, jets resulting from the two primary b-quarks tend to overlap if they are reconstructed on the basis of ordinary, fixed radius jet algorithms, see Section 4.2.4 for more details. This leads to some of the associated tracks of this sub-jets to not be correctly identified and this effect is translated into a drop in b-tagging efficiency, as shown in the following. In order to overcome this problem, the sub-jets inside the fat-jet are reconstructed with a dynamic radius parameter.

This technique reconstructs jets from tracks where the radius parameter of the anti-k_t algorithm, described in Section 4.2.1, scales with the inverse of the jet p_T [255]:

$$R_\mathrm{eff} = \frac{\rho}{p_\mathrm{T}}, \tag{8.1}$$

where the parameter ρ estimates how fast the effective jet size must decrease with the p_T of the jet. By studying the efficiency of sub-jet double b-tagging at the truth level of a Higgs jet, it was found that the optimal values are $\rho = 30$ GeV with $0.02 < R_\mathrm{eff} < 0.4$ [165].

With VR track jets, boosted Higgs boson decays can still be identified properly as each track jet area in the η-φ plane is small if the p_{T} of the decaying Higgs boson is large. Figure 8.2 shows a cartoon that conceptually illustrates the subject reconstruction with a fixed radius jet on the left and with a variable radius on the right, and for the latter the tracks associated to each sub-jet are better identified.

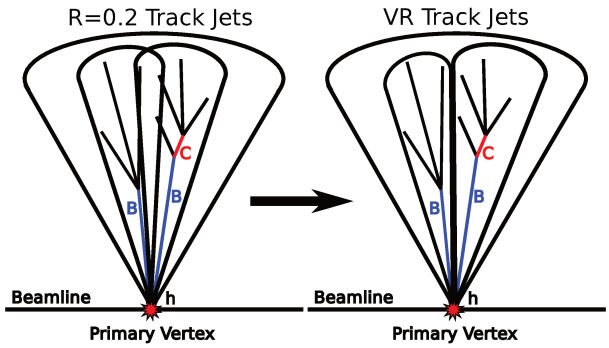


Figure 8.2 – Cartoon illustrating sub-jet reconstruction using variable radius track jets. [165].

The b quarks identification via the b-tagging can be applied to VR track jets in order to discriminate signal events from background due to light jets in a powerful way. The performance of VR track jets in terms of $h \to b\bar{b}$ tagging efficiency, as well as modeling of the angular separation between the sub-jets and the b-hadrons, was found to outperform the $R = 0.2$ fixed jet radius reconstruction technique, detailed in Section 4.2.4, as shown in Figure 8.3.

In light of the mono-$h(b\bar{b})$ search, the improvement of VR track jets with respect to fixed radius FR track jets is compared in order to estimate the sensitivity gain in the merged region, which is related to high masses of the mediator $m_{Z'}$. The signal significance is evaluated as [268]:

$$Z = \sqrt{2\left((s+b)\ln\left(1+\frac{s}{b}\right)-s\right)}, \tag{8.2}$$

where s and b represent, respectively, the number of signal and background events in a given phase space region.

Figure 8.4 shows the significance ratio of using VR track jets to FR track jets for the merged regime event selection in the 2 b-tag category. By using VR track jets, the sensitivity increases by up to 30-75% at large values of $m_{Z'}$. No gain is expected in the intermediate and low masses of the mediator Z' since this region is related to intermediate and low $E_{\mathrm{T}}^{\mathrm{miss}}$, corresponding to the resolved regime event selection.

For a particular point in the parameter space, the gain in selection efficiency can be studied in detail. Figure 8.5 shows the selection efficiency for a signal with $m_A = 500\,\mathrm{GeV}$ from using VR track jets. Figure 8.5(a) shows the overall efficiency for the full signal region selection described in Section 8.4 for events in the merged regime, and with the additional requirement that the mass of the Higgs boson candidate lies between 70 and 140 GeV, as a function of $m_{Z'}$. In particular, this selection requires at least two track

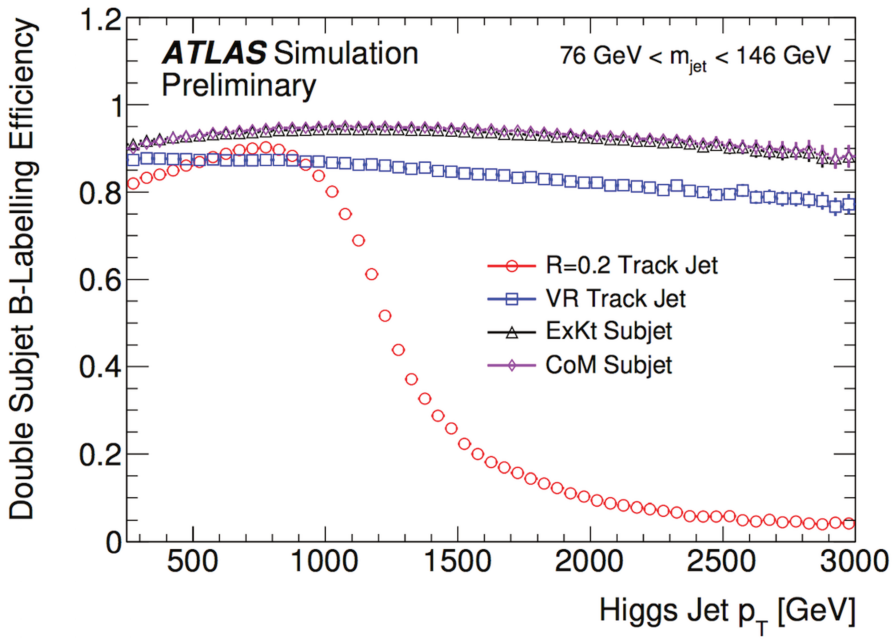


Figure 8.3 – Efficiency of sub-jet double b-labelling at the truth level of a Higgs jet (defined as a large-R matching a truth Higgs boson) as a function of the Higgs jet p_{T}. The error bars include statistical uncertainties only. Other b-labelling efficiencies for Higgs boson decaying to $b\bar{b}$ sub-jet reconstruction techniques are shown: Exclusive-k_{t} and the center-of-mass [165].

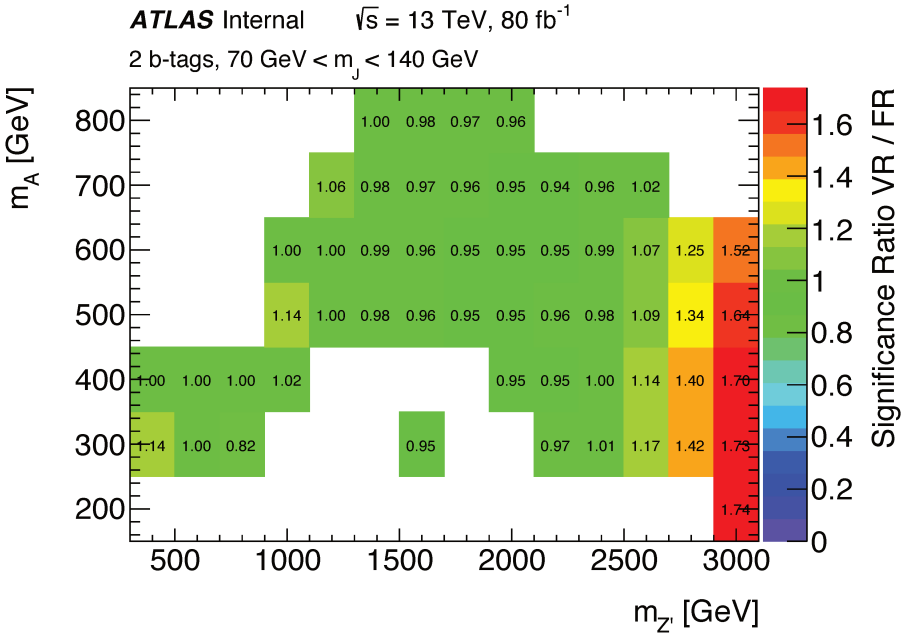


Figure 8.4 – Signal significance ratio between VR and R=0.2 track jets in the merged region with 2 b-tags in the $m_{Z'}$-m_A phase space [259].

jets associated to the large-R jet, and a minimum angular separation of the two leading track jets as described in Appendix E. The efficiency is calculated only considering Higgs boson decays to b-quarks. The effect is shown separately for b-tagged jet multiplicities of 1 (square markers) and 2 (triangular markers). In addition, the combined acceptance times efficiency from events with 0, 1 or 2 b-tagged jets is drawn with circular markers. In the region above 2.5 TeV, it is considerably higher when using VR track jets, as events more often fail the requirement of more than one track jet in the case of FR jets. The increase in acceptance for both 1 and 2 b-tag events is also most pronounced for $m_{Z'}$ above 2.5 TeV, corresponding to the largest boost of the Higgs boson. At $m_{Z'} = 3$ TeV, it amounts to a factor of about 1.7 for events with two b-tags and 1.8 for events with one b-tag. Figure 8.5(b) shows the relative efficiencies of 1 and 2 b-tagged events normalised by the combined acceptance for the two types of track jets, which shows similar b-tagging efficiencies for VR and FR track jets. Therefore the enhancement in signal acceptance originates from the improved sub-jet recovery ability with VR track jets, as shown by the combined acceptance (line with filled circular markers) in Figure 8.5(a).

Below 2.5 TeV, a higher efficiency for identifying two b-jets is obtained with FR track jets. This is due to the fact that in this regime, often more than two jets are reconstructed with the VR jet algorithm, and the two highest-p_{T} jets are not always the b-jets [165]. Thus, when considering only the two highest-p_{T} track jets, the efficiencyis smaller for VR compared to FR track jets. It is possible to recover this signal efficiency by considering the three highest-p_{T} jets [165], but in the search presented here this also increase the background contamination to a level that might lead to an overall decrease in sensitivity. Therefore, only the two highest-p_{T} VR track jets are used in this analysis. Even though the signal efficiency is higher for FR track jets below 2.5 TeV, the use of the VR algorithm still provides a better signal to background ratio.

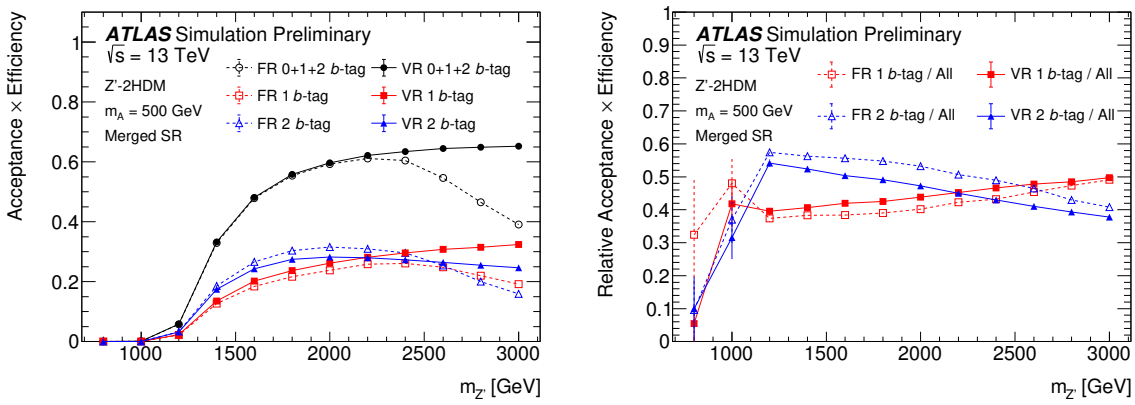
Including events containing only one b-tagged jet leads only to a marginal improvement in sensitivity when using VR track jets. Since the background modelling in such events is more challenging, they are not included in the new result, but are shown for illustration in Figure 8.5.

The performance of VR track jet Higgs tagging was first described in [165]. The present mono-$h(b\bar{b})$ analysis serves the commissioning of this new technique, being the first use case.

For the b-tagging of VR track jets, a working point with average b-tagging efficiency of 77% for a fixed cut in the `MV2c10` discriminant is used, see Section 4.2.3 for detail in b-tagging procedure. The efficiency of the b-tagging algorithm is calibrated in data and scale Factors (SF) to correct the b-tagging performance in simulation to that observed in data are measured. Appendix E details the VR track jet b-tagging efficiency calibration.

8.3.2 Including missing transverse momentum significance

The $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance can be used to identify events in which the reconstructed $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ is likely to come from weakly interacting particles, like dark matter, or whether the reconstructed $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ is caused by object mismeasurements, resolutions or efficiencies.



(a) (b)

Figure 8.5 – (a) Acceptance × efficiency ($A \times \varepsilon$) for different b-tag multiplicities as a function of $m_{Z'}$ in a Z'-2HDM model with $m_A = 500$ GeV, $\tan\beta = 1.0$, $g_Z = 0.8$, $m_\chi = 100$ GeV, and $m_H = m_{H^\pm} = 300$ GeV, for events with $E_{\mathrm{T}}^{\mathrm{miss}} > 500$ GeV (merged SR) when using FR track jets (open symbols, dashed lines) and VR track jets (filled symbols, solid lines). The selection includes a requirement to have at least two track jets associated to the large-R jet, and requires a minimum angular separation of the two leading track jets, as described in Appendix E. The combined $A \times \varepsilon$ for events with either 0, 1 or 2 b-tagged jets is drawn with circular markers, while triangular and square markers correspond to the individual $A \times \varepsilon$ for 1 and 2 b-tag(s), respectively. (b) Relative $A \times \varepsilon$ for different b-tag multiplicities with the same Z'-2HDM model and selections. The $A \times \varepsilon$ for 1 b-tagged and 2 b-tagged events are normalised to the sum of events with either 0, 1 or 2 b-tagged jets ("All"), as obtained for each track jet choice individually [259].

The mono-$h(b\bar{b})$ search can take advantage of the new defined object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance, $\mathcal{S}$, extensively described in Chapter 5. This novel definition depends on the multiplicities, types, and kinematics of the objects measured in each event [269].

Multijet background is originated from strong interactions and introduces fake $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ coming from mis-measured jet momenta. The object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance can help to identify and separate multijet background with respect to the standard model electroweak background and dark matter signals.

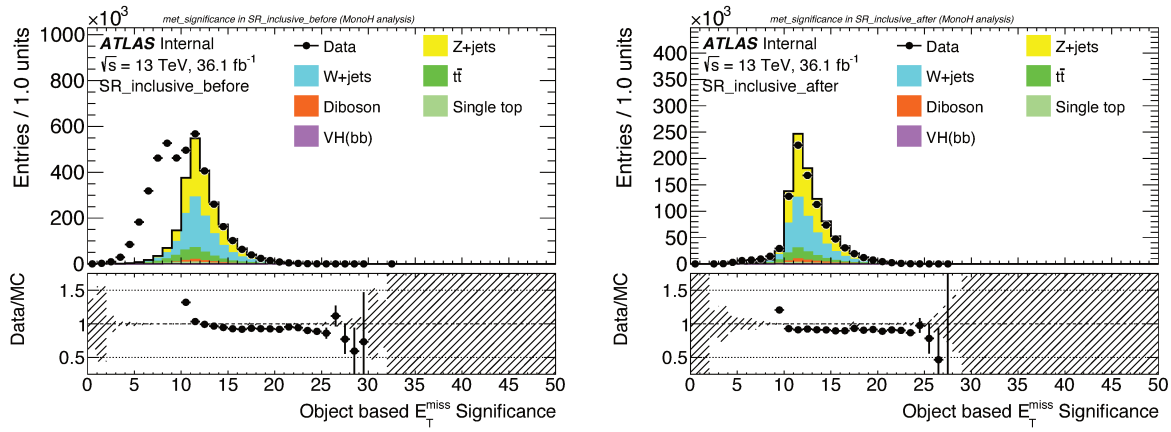
Since the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ associated to multijet events mainly comes from jet resolutions, i.e. fake $E_{\mathrm{T}}^{\mathrm{miss}}$, these events are more likely to have lower values of the $E_{\mathrm{T}}^{\mathrm{miss}}$ significance compared to other processes with real $E_{\mathrm{T}}^{\mathrm{miss}}$. Such processes with real $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ are not only the dark matter signals, but also the dominant backgrounds of the mono-$h(b\bar{b})$ analysis as $t\bar{t}$ production, $Z(\nu\nu)$+jets and $W(\ell\nu)$+jets.

Even though a fraction of the multijet events comes with real $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ caused by neutrinos originated from heavy flavour hadron decays, this is a second order effect, and multijet events with considerable high $E_{\mathrm{T}}^{\mathrm{miss}}$ ($E_{\mathrm{T}}^{\mathrm{miss}}$ >150 GeV for this analysis) do not occur very often since substantial fluctuations in calorimeter jet energy measurements are the root cause. Therefore, multijet events are expected to contribute in the resolved regime, introduced in Section 6.1.1, in the low $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ categories of the analysis. Since $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$

in multijet events arises mostly from fluctuations of poorly reconstructed jets, multijet events are expected to have low $\mathcal{S}$ values. Jet energy resolutions corresponds to the highest source of fluctuation in the total $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ variance, as can be seen in Figure 5.3, which again translates to low $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance values at a fixed $E_{\mathrm{T}}^{\mathrm{miss}}$.

Figure 8.6 shows the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance distributions for data collected in 2015 and 2016, and MC simulation, inclusively in b-tag multiplicities and $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ categories, propagating the object resolutions described in Section 5.2. On the right, Figure 8.6(b) considers the mono-$h(b\bar{b})$ signal region event selection described in Section 6.4, while on the left, Figure 8.6(a), anti-QCD cuts are removed (equivalent to Figure 6.4). Explicitly, these mentioned anti-QCD cuts in the resolved regime are: $\min(\Delta\Phi(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}, \mathbf{p}_{\mathrm{T}}(jets_{1,2,3})))>20°$, $\Delta\Phi\ (\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}, \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}) < 90°$, $\Delta\Phi(\mathbf{p}_{\mathrm{T}}(j_1), \mathbf{p}_{\mathrm{T}}(j_2))< 140°$ and $\Delta\Phi(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}, \mathbf{p}_{\mathrm{T}}(h_{jj})\ ) > 120°$, see Section 6.4.2.2 for details.

The multijet background is not included in Figure 8.6, therefore data is expected to overshoot MC simulation in regions with multijet contamination, corresponding to low values of the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance. This behaviour is very clear in Figure 8.6(a) for which no anti-QCD (multijet cleaning) cuts are applied.



(a) (b)

Figure 8.6 – Object-based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance before and after anti-QCD cuts, inclusive in b-jet multiplicity. The left-hand side figure shows that this new variable is a powerful means to separate multijet backgrounds from genuine $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ processes. Only statistical uncertainties are taken into account for displaying the uncertainty bands.

Even after the full signal region event selection, which considers anti-QCD cuts, from Figure 8.6(b) it is evident that some residual mismodelling at low values of the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance is left if the multijet contribution is neglected. In the previous mono-$h(b\bar{b})$ analysis result with data collected in 2015 and 2016, this residual multijet background was estimated in a data-driven way described in Section 6.5.4. In the energies presented here, it is exploit instead the new object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance variable to further suppress the multi-jet background.

The object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance definition, described in detail on Chapter 5, considers a constant track-based soft term (TST) resolution, and an additional jet resolution which estimates the effect of pile-up jets which unbalance the total $E_{\mathrm{T}}^{\mathrm{miss}}$. These object

resolutions are described in Section 5.2. Both resolutions are estimated from a $Z(\mu\mu)$ + jets sample, with the consequence of an improved $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance performance in that particular topology.

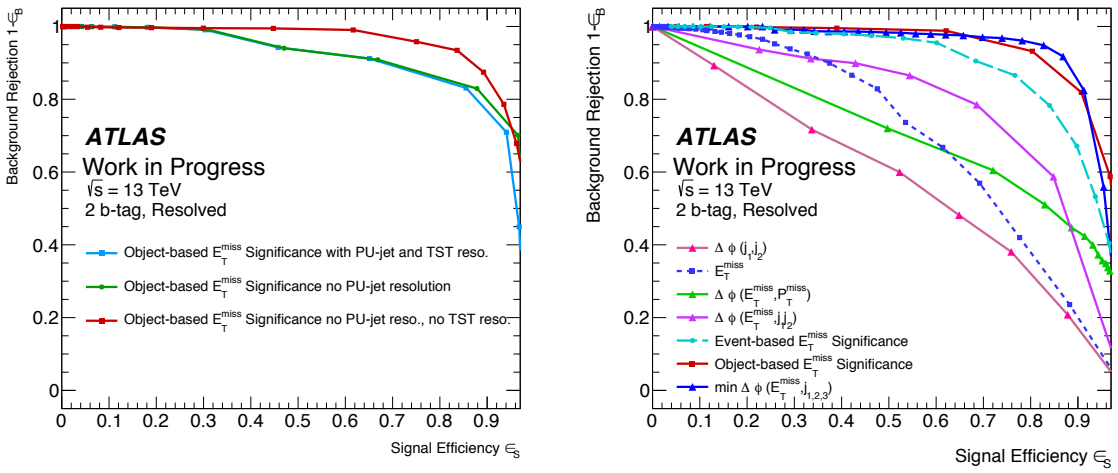
However, when considering the mono-$h(b\bar{b})$ event selection and topology, the performance is degraded when including in the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ variance the estimation of soft term and pile-up jet resolutions. Firstly, since events with high $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ are selected in the analysis, high hadronic activity is expected in the recoil with respect to the $E_{\mathrm{T}}^{\mathrm{miss}}$, and thus the soft term contribution is minor. The soft term resolution is estimated as a constant number extracted from the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ distribution in a $Z \to \mu\mu$ sample with a jet veto, however this first order estimation seems to be suboptimal for the mono-$h(b\bar{b})$ topology and kinematics, and a more robust approach is needed, similar to the one introduced in Section 5.5. Secondly, the estimation of the pile-up jet resolution, described in Section 5.2, is parameterized inclusively in b-tagging. However, requesting one or two b-tagged hadrons, considerably reduces the jets coming from pile-up interactions. This motivates a pile-up jet resolution parameterized also in b-tagging regions for future studies. In this mono-$h(b\bar{b})$ search, only events with 2-b tags are considered and thus the additional pile-up jet resolution is overestimating the total $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ variance, impacting negatively the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance performance.

The performance of the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance in terms of background rejection and signal efficiency is studied using simulated event samples. For the background, dijet events generated with PYTHIA 8 using the A14 set of tuned parameters and the NNPDF2.3 LO PDF set are used. The signal efficiency is determined for a simulation with $m_{Z'} = 400\,\mathrm{GeV}$ and $m_A = 300$ GeV, resulting in relatively low $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$, and therefore more challenging to discriminate against multijet background.

The separation power of the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance considering all the object resolutions, described in Section 5.2, is shown in Figure 8.7(a). It considers the 2-$b$tag category for the resolved regime event selection ($150 \leq E_{\mathrm{T}}^{\mathrm{miss}}/\,\mathrm{GeV} < 500$) without the anti-QCD cuts. Also, Figure 8.7(a) shows the performance of the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance when the Pile-up (PU) jet resolution is not considered and when neither the PU-jet nor TST resolutions are propagated in the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ variance. The latter has a clear separation power improvement with respect the other two. For this reason, the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance without PU-jet resolution and TST resolution is used in the mono-$h(b\bar{b})$ search for dark matter, and in the following it will be called object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance for simplicity.

Figure 8.7(b) shows the performance of the optimal object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance in comparison with: the anti-QCD cuts, event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance, defined in Equation 5.1, and $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ itself. It considers the 2-$b$tag category with the resolved regime event selection but without the anti-QCD cuts. The most powerful variable in discriminating dijet versus the dark matter signal is the minimum azimuthal angle between the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ and the three leading small-R jets. The object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance shows a slightly worse but comparable performance with respect to this mentioned observable. It is also interesting to note that the relative performance between object- and event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significances, and $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ variables reported in Section 5.4 stays valid for the mono-$h(b\bar{b})$ topology. For example, for rejecting almost 99% of the dijet events, the signal

efficiency increases from about 55% for the event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance to close to 80% for the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance



(a) (b)

Figure 8.7 – Performance of different multijet discriminating observables in terms of the signal efficiency and background rejection as estimated from a signal simulation with $m_{Z'} = 400$ GeV and $m_A = 300$ GeV and a dijet simulation, respectively. These observables are 8.7(a) different object based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance definitions on the left, and 8.7(b) the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance without PU-jet and TST resolution, other observables used in the analysis to reject multijet background defined in 6.4, and the event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance and $E_{\mathrm{T}}^{\mathrm{miss}}$ itself.

The object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance is, therefore, a promising observable to reject multijet background at low $E_{\mathrm{T}}^{\mathrm{miss}}$. Even after applying the most discriminating anti-QCD cut, $\min(\Delta\Phi(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}, \mathbf{p}_{\mathrm{T}}(jets_{1,2,3})))>20°$, the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance is still showing a good dijet rejection power. Figure 8.8 shows the performance of the object- and event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significances, and $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ for which the efficiencies for signal and background are estimated with respect to a selection that requires that the minimum angular separation between the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ vector and the leading jets in an event is greater than 20°. Figure 8.8, illustrates the significant improvement obtained with the object-based definition employed in this search. For example, for signal efficiencies above 80%, the background rejection factor $(1/\varepsilon_B)$ for the object-based significance is approximately three to four times higher than for the event-based one. Refer to Figure F.3 for an alternative way of presenting this information with the background rejection factor.

Figure 8.9 shows the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance distributions in data and MC simulation for events with 2 b-tagged jets in the three resolved $E_{\mathrm{T}}^{\mathrm{miss}}$ categories after full event selection. The dijet MC sample suffers from low statistics, which is likely to contribute to the mismodeling at low values of the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance, however it can illustrate the approximate multijet contribution in the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance distribution for the first $E_{\mathrm{T}}^{\mathrm{miss}}$ categories. Since $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ and object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance, $\mathcal{S}$, are highly correlated variables, the distribution is shifted at higher values of $\mathcal{S}$ for higher

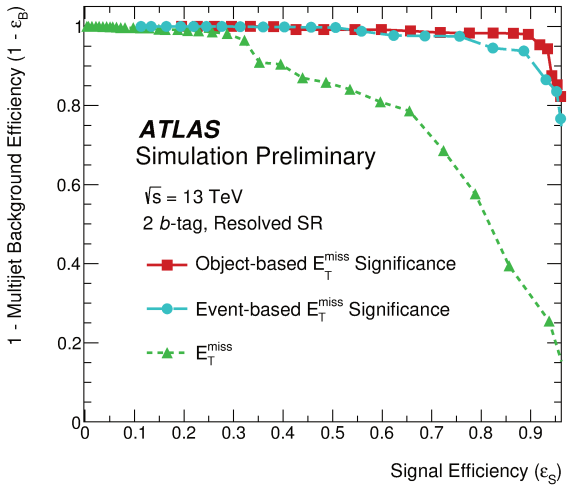


Figure 8.8 – Performance of the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance (line with square markers) in terms of the signal efficiency and background rejection as estimated from a signal simulation with $m_{Z'} = 400\,\mathrm{GeV}$ and $m_A = 300$ GeV and a dijet simulation, respectively, in comparison to an alternative definition of the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance (dashed line with circular markers) and $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ itself (densely-dashed line with triangular markers). Selections on these three variables are applied in addition to a requirement on the angular separation between the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ vector and the leading jets in an event.

$E_{\mathrm{T}}^{\mathrm{miss}}$ thresholds. Also, it can be seen that the dijet contribution is sizable for the first $E_{\mathrm{T}}^{\mathrm{miss}}$ category, whereas is less for the second, and likely negligible for the third category. Distributions for events with 1 b-tagged jet can be found in Appendix F.

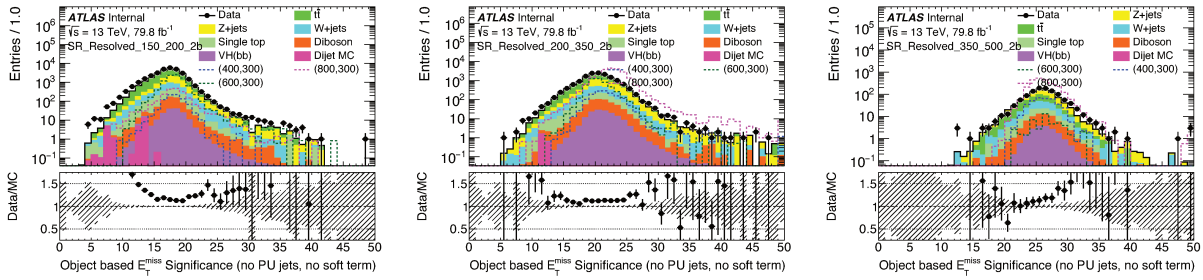


Figure 8.9 – $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance distribution for events with 2 b-tagged jets in the regions with $150\,\mathrm{GeV} < E_{\mathrm{T}}^{\mathrm{miss}} < 200\,\mathrm{GeV}$ (left), $200\,\mathrm{GeV} < E_{\mathrm{T}}^{\mathrm{miss}} < 350\,\mathrm{GeV}$ (middle), $350\,\mathrm{GeV} < E_{\mathrm{T}}^{\mathrm{miss}} < 500\,\mathrm{GeV}$ (right). Only statistical uncertainties are taken into account for displaying the uncertainty bands. The signal distributions overlayed are normalised to 1 pb.

Introducing in the resolved regime event selection a requirement on the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance variable removes mismodeling at low values by suppressing multijet background. This requirement is not introduced in the merged regime event selection, described in Section 6.4.2.3, since the multijet contribution is shown to be negligible due to the much higher $E_{\mathrm{T}}^{\mathrm{miss}}$ threshold, $E_{\mathrm{T}}^{\mathrm{miss}} > 500\,\mathrm{GeV}$, as is illustrated in Figure 6.17.

The expected signal significance is calculated from MC simulation by scanning different values of the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance $\mathcal{S}$ with the resolved event selection. Figure 8.10 shows the expected signal significance for three low $m_{Z'}$ signals using the simple definition of $S/\sqrt{S+B}$, where S is the dark matter model signal yield and B is the total background yield after each cut in $\mathcal{S}$. This signal significances are calculated after the resolved event selection and for the 2 b-tag category, for other b-tag categories refer to Figure F.2. The expected signal significance in the resolved regime with a 2 b-tag selection shows an approximated signal significance gain from 2% to 18% for cut values of $16 < \mathcal{S} < 20$. For lower cut values, $\mathcal{S} < 16$, the signal significance remains approximately constant.

In Figure 8.11, the signal significance curves are divided in the three $E_{\mathrm{T}}^{\mathrm{miss}}$ categories of the resolved regime: $150 \leq E_{\mathrm{T}}^{\mathrm{miss}}/\,\mathrm{GeV} < 200$, $200 \leq E_{\mathrm{T}}^{\mathrm{miss}}/\,\mathrm{GeV} < 350$ and $350 \leq E_{\mathrm{T}}^{\mathrm{miss}}/\,\mathrm{GeV} < 500$, shown in Figures 8.11(a), 8.11(b) and 8.11(c), respectively. They show that higher $E_{\mathrm{T}}^{\mathrm{miss}}$ categories, a higher object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance threshold is needed in order to achieve an optimal signal significance value. However, the baseline for this first analysis adding a $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance requirement is to reject multijet background, and thus, the first $E_{\mathrm{T}}^{\mathrm{miss}}$ category is our region of interest. From Figure 8.11(a), a gain in signal significance can obtained for cut values between $14 < \mathcal{S} < 18$.

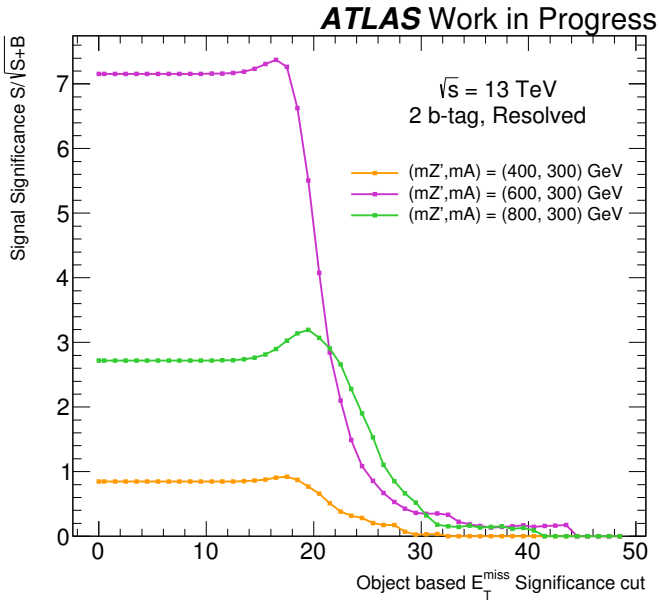
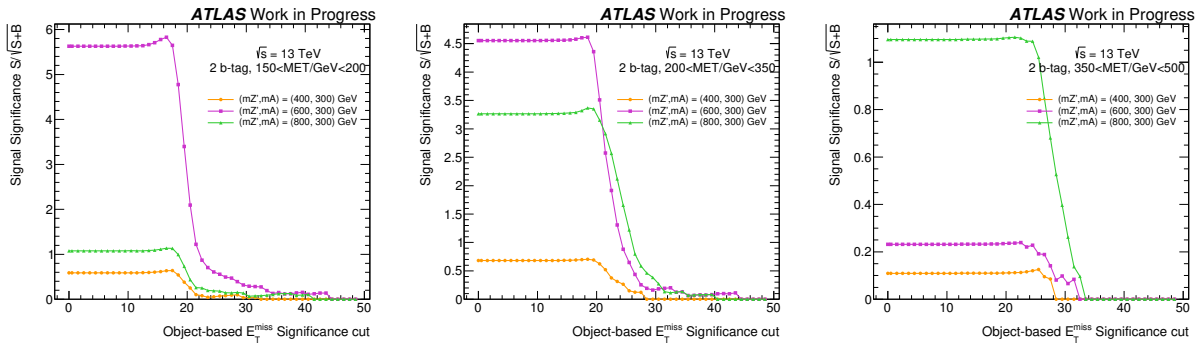


Figure 8.10 – Expected signal significance versus object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance thresholds for three signal models with different $m_{Z'}$. Events shown are required to have 2 b-tagged jets after the full event selection in the resolved regime, which is described in Section 6.4.

The overall signal significance impact of the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance cut can be illustrated in the $(m_{Z'}, m_A)$ plane, where the exclusion contour limits are set, see Figure 7.8 for the limits from the previous analysis result. Figure 8.12 shows the ratio of expected signal significance, with the resolved event selection, of adding a given $\mathcal{S}$ thresh-



(a) (b) (c)

Figure 8.11 – Expected signal significance versus object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance thresholds for three signal models with different $m_{Z'}$, in the regions with 8.11(a) $150\,\mathrm{GeV} < E_{\mathrm{T}}^{\mathrm{miss}} < 200\,\mathrm{GeV}$, 8.11(b) $200\,\mathrm{GeV} < E_{\mathrm{T}}^{\mathrm{miss}} < 350\,\mathrm{GeV}$, 8.11(c) $350\,\mathrm{GeV} < E_{\mathrm{T}}^{\mathrm{miss}} < 500\,\mathrm{GeV}$. Events shown are required to have 2 b-tagged jets after the full event selection in the resolved regime, described in Section 6.4.

old requirement over expected signal significance without this cut. The expected signal significance is calculated as in Equation 8.2 and each mass-mass plane shows a different object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance cut, considering $13 \leq \mathcal{S} \leq 18$.

Since Figure 8.12 has the resolved regime event selection, the relevant part of the parameter space is related to low masses of the mediator, $m_{Z'} \lesssim 1.2$ TeV. A requirement of $\mathcal{S} > 16$ gives improvements on signal significances of 5-8%, while a cult of $\mathcal{S} > 17$ gives a gain in signal significance of 2-15%, depending on the masses of the pseudoescalar m_A and vector mediator $m_{Z'}$.

Figure 8.13 shows the expected signal significance for only the event selection targeted at the resolved topology. The upper plot shows the expected signal significance for the resolved regime event selection, while the bottom plots shows the expected signal significance adding $\mathcal{S} > 16$ or $\mathcal{S} > 17$ cut to the event selection. The absolute values of the signal significance are sufficiently large to be relevant for being considered when deciding on the value of the cut only for low values $m_{Z'}$ and m_A.

It was decided to have a conservative approach for this analysis result, and a loose cut in the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance of $\mathcal{S} > 16$ was chosen. This cut has a good compromise of not changing significantly the background composition, as is going to be shown in the following, while having the mismodeling under control, but the expected signal significance is not enhanced as it potentially could. And, even more important, strongly suppress any remaining multijet background, as will be discussed in Section 8.5.1. Based on the performance study on Figure 8.8, more than 95% of dijet events can be rejected by introducing a requirement of $\mathcal{S} > 16$, while retaining a signal efficiency close to 90%.

Also, with a requirement on the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance variable of $\mathcal{S} > 16$, the data and MC simulation disagreement at low values reported in Figure 8.9 is removed. This mismodeling is associated to the multijet contribution affecting mainly the first $E_{\mathrm{T}}^{\mathrm{miss}}$ category, as discussed before. If same data-driven strategy described in Sec-

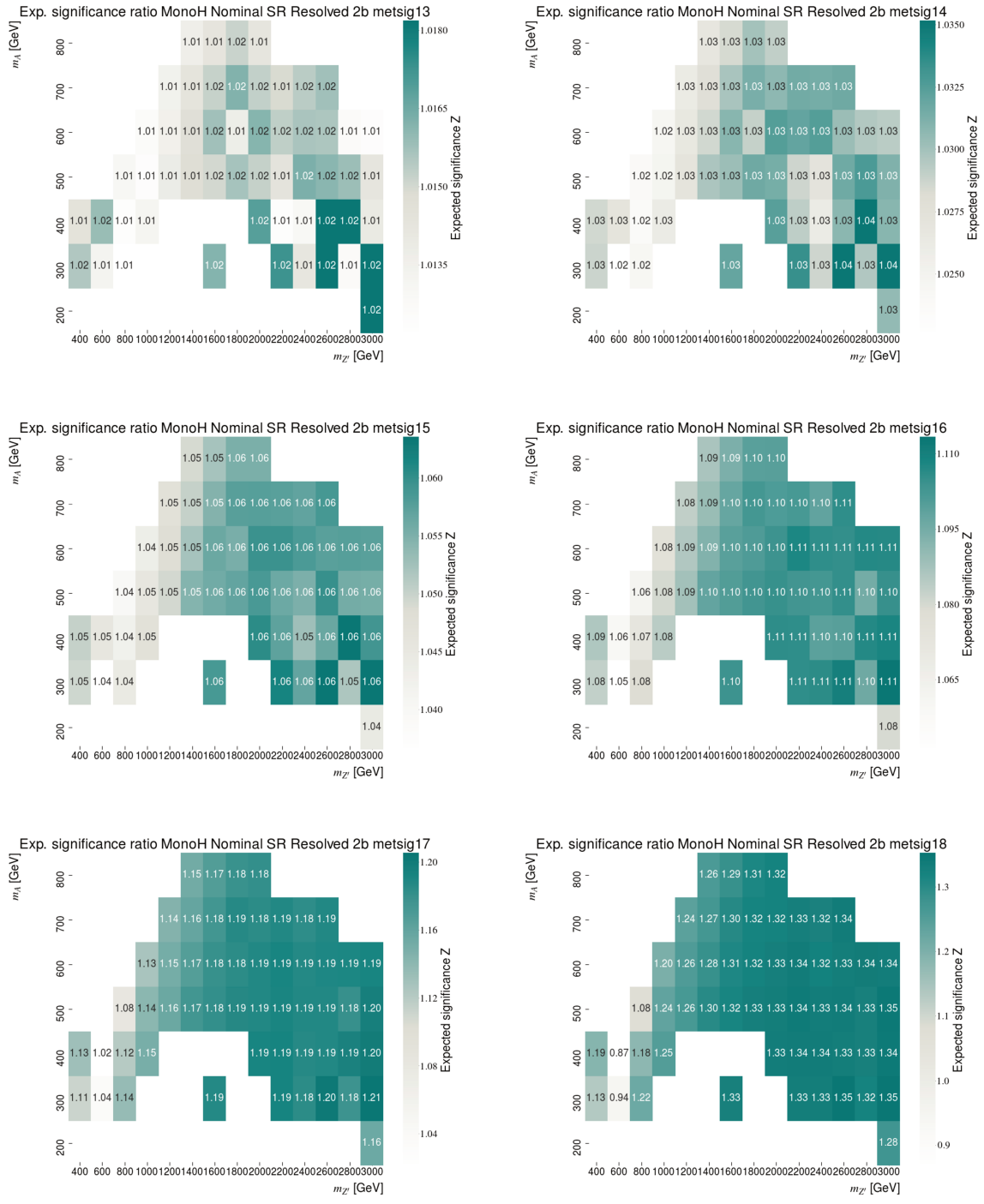


Figure 8.12 – Ratio of expected signal significance (taking into account the limited background MC statistics) with $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance cut varied from 13 to 18 with respect to no cut for events with 2 b-tagged jets.

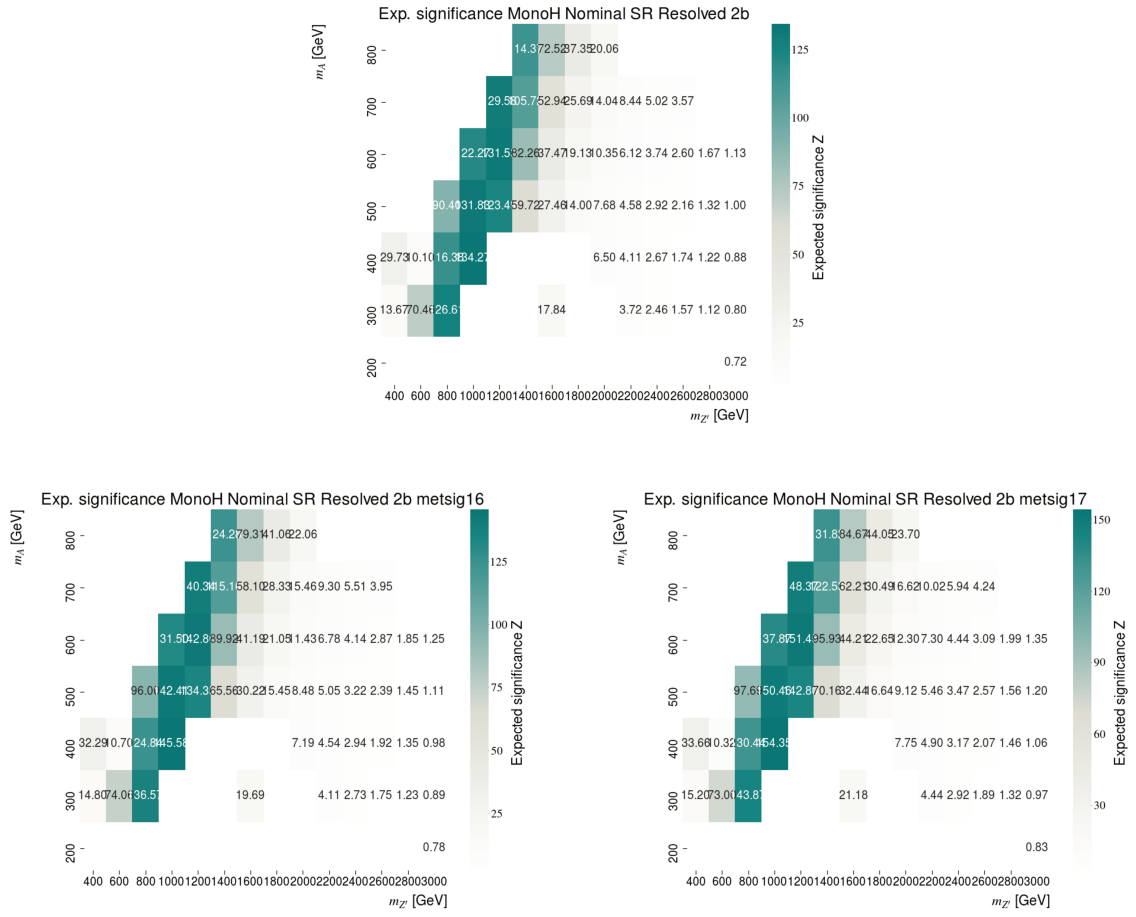


Figure 8.13 – Expected signal significance taking into account the limited background MC statistics for events with 2 b-tagged jets, without (top), with $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance > 16 requirement (bottom left) and with $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance > 17 requirement (bottom right).

tion 6.5.4 is performed in order to estimate the multijet contribution in the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance distribution, it is clearer that the bulk part of the multijet distribution has $\mathcal{S} < 16$. Figure 8.14 shows the pre- and post-fit distributions for data, MC simulation and multijet template in the first $E_{\mathrm{T}}^{\mathrm{miss}}$ category with 2 b-tag jets following the procedure outlined in Section 6.5.4, being the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance the multijet sensitive variable and not the multiplicity of jets with muons.

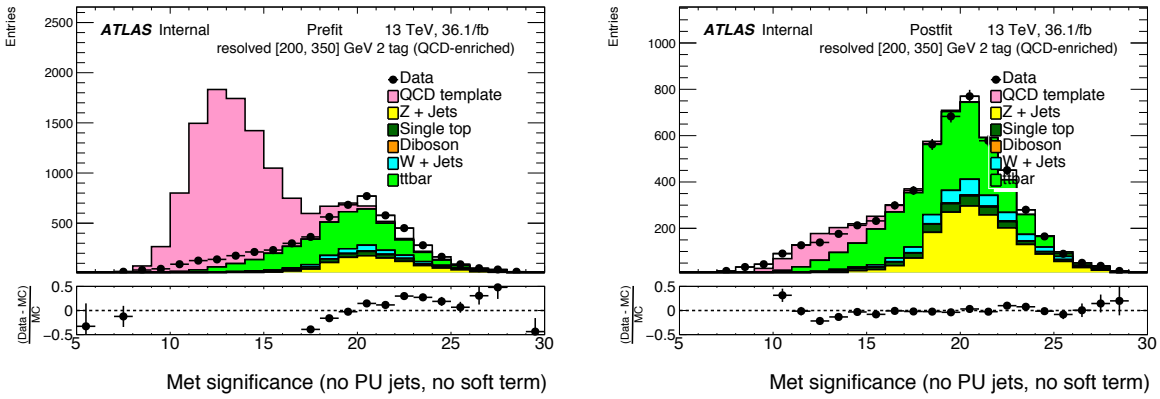


Figure 8.14 – Object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance distributions in data, MC simulation and multijet template estimation (QCD template). On the left plot, the multijet template is extracted from QCDCR, and data and MC simulation are shown in the QCDSR. The right plot shows the result of fitting simultaneously the multijet template and MC simulation to data for $5 \leq \mathcal{S} \leq 30$.

Figure 8.15 shows the distributions of the invariant mass of the Higgs boson candidate dijet system that are used as inputs to the fit as described in Section 7.1.2. The distributions are shown in the three resolved event selection $E_{\mathrm{T}}^{\mathrm{miss}}$ categories for events with 2 b-tagged jets. The plots on the left corresponds to the resolved event selection described in Section 6.4.2.2, while the distributions on the right also include the requirement on the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance $\mathcal{S} > 16$.

In a logarithmic scale one can see that the dijet MC has a negligible contribution after the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance cut. For the same distributions in linear scale, refer to Figure F.4. This dijet MC reduction is evident in the $150 \leq E_{\mathrm{T}}^{\mathrm{miss}} / \text{ GeV} < 200$ category for which this cut has a more relevant impact as expected.

The introduction of the $\mathcal{S} > 16$ cut in order to suppress multijet background has an insignificant effect on the general shape of the m_{jj} distributions in Figure 8.15 used as input to the statistical analysis.

Figure 8.16 shows the backgrounds yields in the three resolved event selection $E_{\mathrm{T}}^{\mathrm{miss}}$ categories for events with 2 b-tagged jets. The plots on the left correspond to the resolved event selection, while the plots on the right also have the requirement on the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance selection, $\mathcal{S} > 16$. No dijet events remain after the cut.

Figure 8.17 shows the relative background composition in the three resolved event selection $E_{\mathrm{T}}^{\mathrm{miss}}$ categories for events with 2 b-tagged jets. Again, the plots on the left correspond to the resolved event selection, while those on the right the $\mathcal{S} > 16$ cut is added.

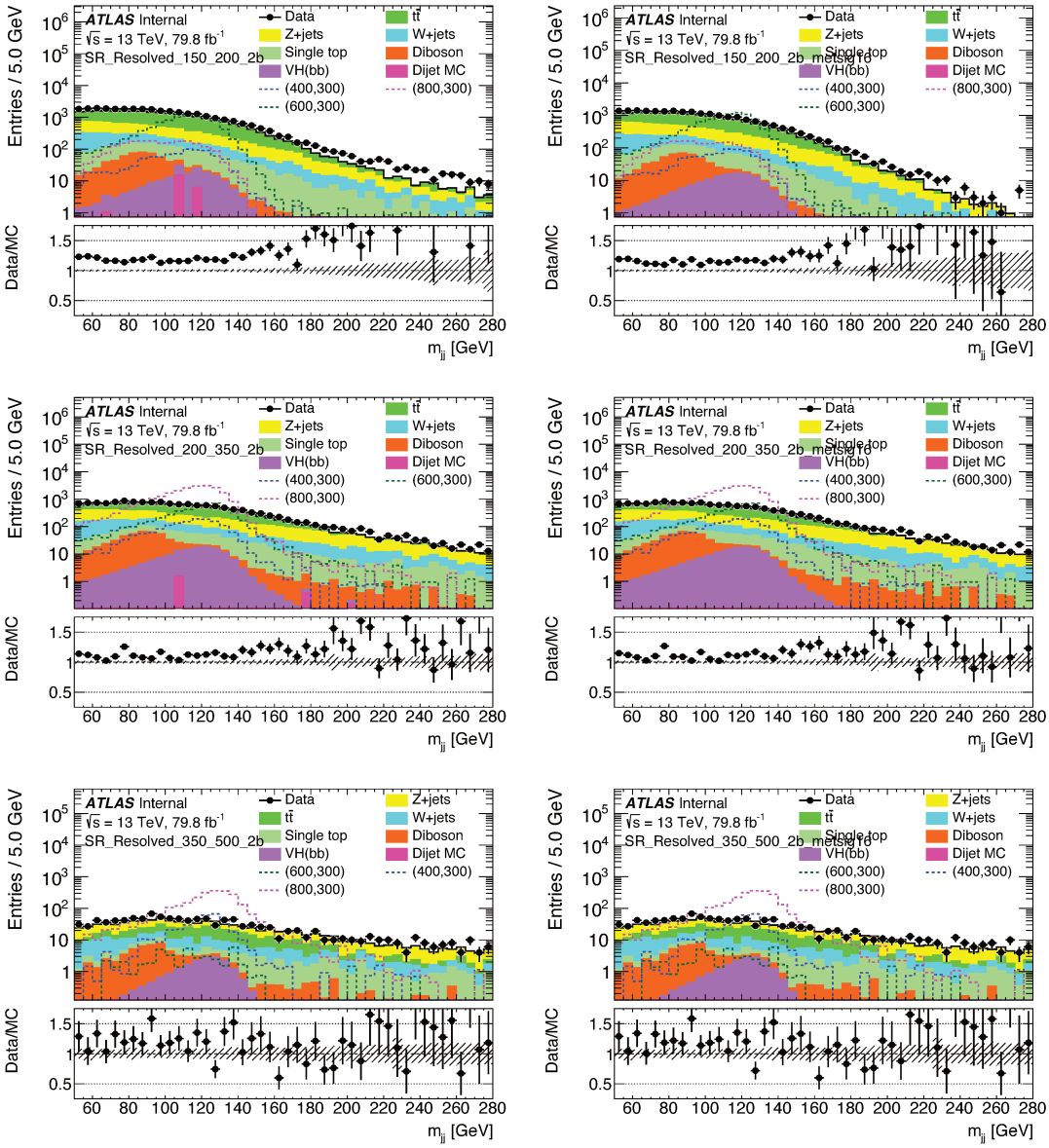
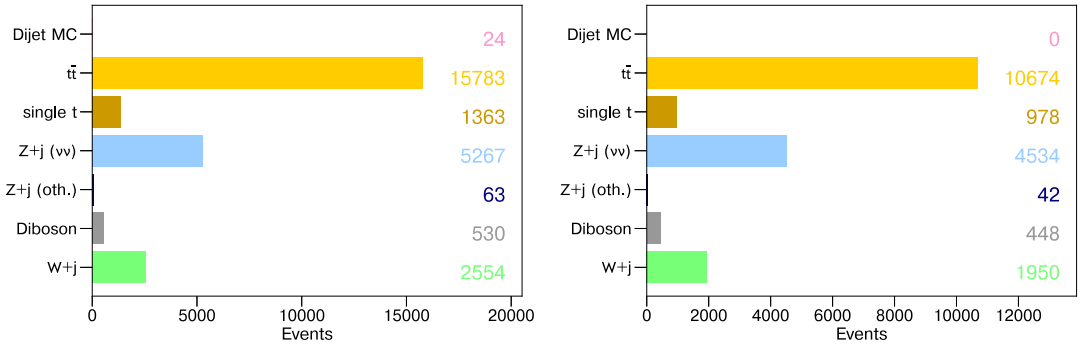
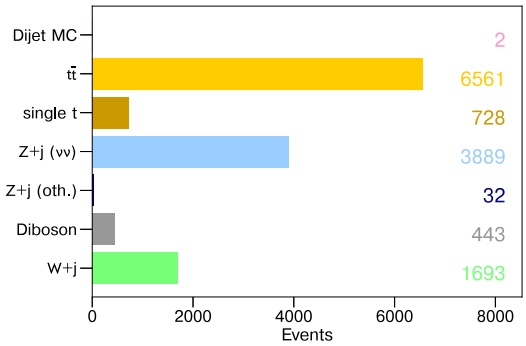


Figure 8.15 – Distributions of the invariant mass of the Higgs boson candidate dijet system in the three resolved event selection $E_{\mathrm{T}}^{\mathrm{miss}}$ bins for events with 2 b-tagged jets, without (left) and with $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance > 16 requirement (right).

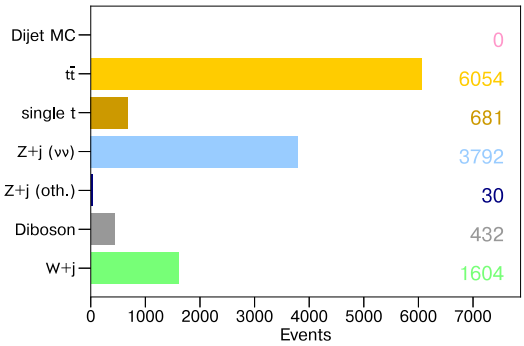


(a) $150\,\text{GeV} < E_{\text{T}}^{\text{miss}} < 200\,\text{GeV}$ (no cut)

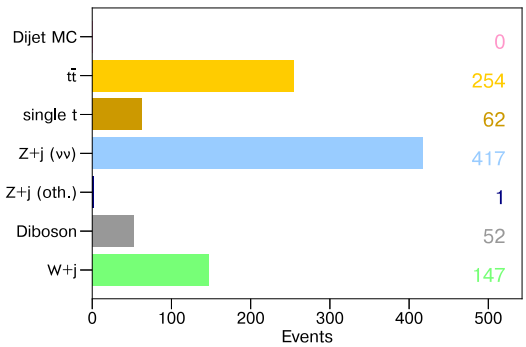
(b) $150\,\text{GeV} < E_{\text{T}}^{\text{miss}} < 200\,\text{GeV}$ ($\mathcal{S} > 16$)



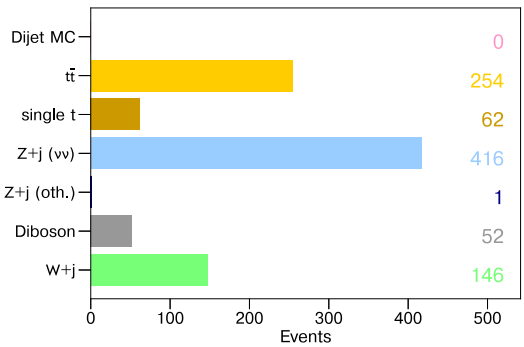
(c) $200\,\text{GeV} < E_{\text{T}}^{\text{miss}} < 350\,\text{GeV}$ (no cut)



(d) $200\,\text{GeV} < E_{\text{T}}^{\text{miss}} < 350\,\text{GeV}$ ($\mathcal{S} > 16$)



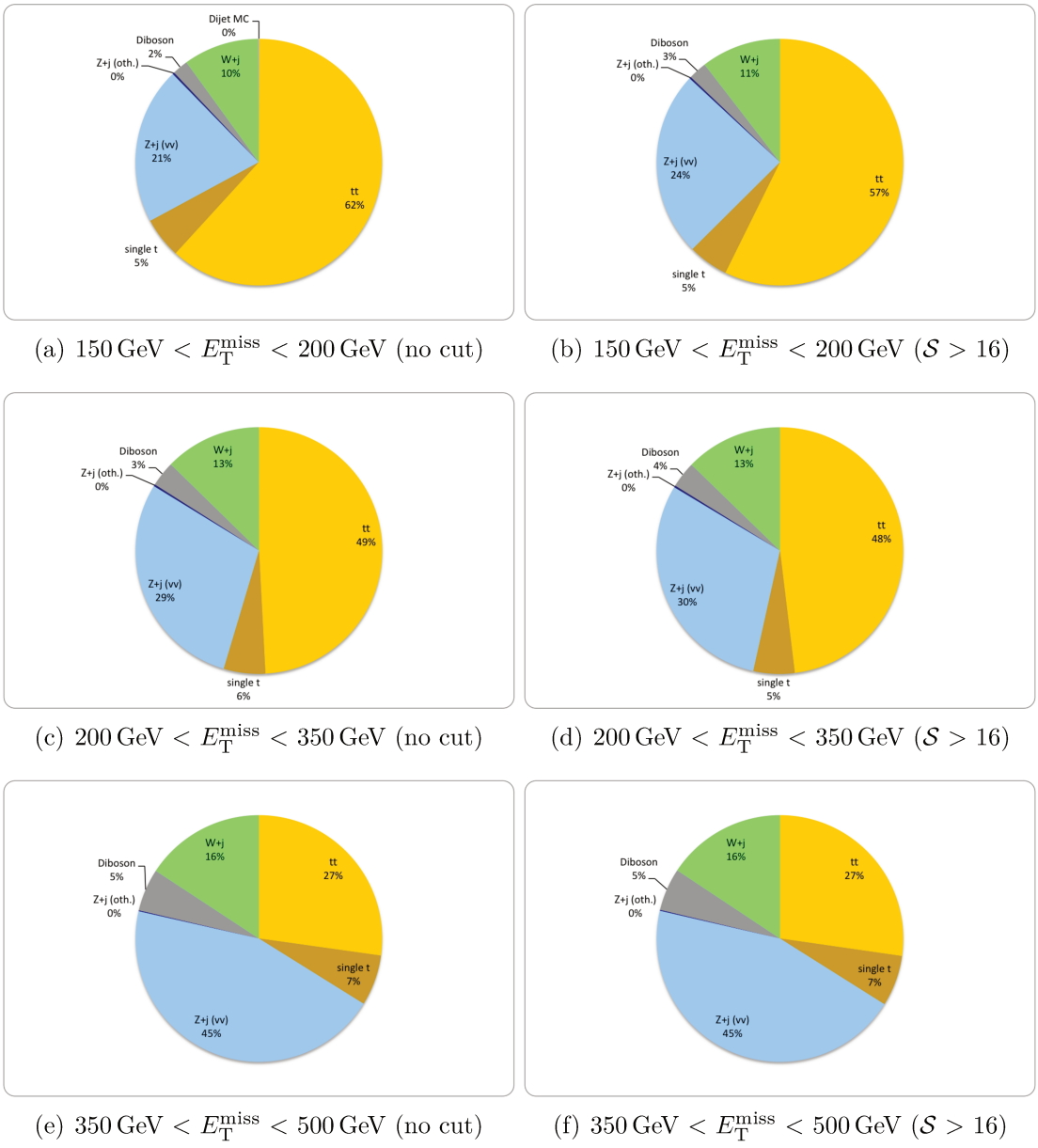
(e) $350\,\text{GeV} < E_{\text{T}}^{\text{miss}} < 500\,\text{GeV}$ (no cut)



(f) $350\,\text{GeV} < E_{\text{T}}^{\text{miss}} < 500\,\text{GeV}$ ($\mathcal{S} > 16$)

Figure 8.16 – Background composition in the three resolved event selection $E_{\text{T}}^{\text{miss}}$ bins for events with 2 b-tagged jets, without (left) and with $\boldsymbol{E}_{\text{T}}^{\text{miss}}$ significance > 16 requirement (right) in the signal region.

Background processes with real $E_{\mathrm{T}}^{\mathrm{miss}}$, such as $Z(\nu\nu)$ + jets are enhanced, while other background processes such as $t\bar{t}$ production are reduced when $\mathcal{S} > 16$ is required. However, the background relative composition is almost unaffected by this cut, with a maximum change of 5% in the first $E_{\mathrm{T}}^{\mathrm{miss}}$ category. Also, in the first $E_{\mathrm{T}}^{\mathrm{miss}}$ category, a reduction of about 27% of the total background is found. In the second and third $E_{\mathrm{T}}^{\mathrm{miss}}$ categories, the overall background reduction is not as significant and is of about 6% and less 1% respectively.



(a) 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 200 GeV (no cut)

(b) 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 200 GeV ($\mathcal{S} > 16$)

(c) 200 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 350 GeV (no cut)

(d) 200 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 350 GeV ($\mathcal{S} > 16$)

(e) 350 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 500 GeV (no cut)

(f) 350 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 500 GeV ($\mathcal{S} > 16$)

Figure 8.17 – Background composition in the three resolved event selection $E_{\mathrm{T}}^{\mathrm{miss}}$ bins for events with 2 b-tagged jets, without (left) and with $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance > 16 requirement (right) in the signal region.

The modelling of the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance variable is studied in a control region where neither multijet nor signal is expected. Figures 8.18 and 8.19 shows the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance variable and, for comparison purposes, the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ for data, recorded in 2015 and 2016, and simulation in the two lepton control region (2ℓ-CR) for the $\mu\mu$-channel, see Section 6.5.3 for details. Also, the event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance $E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}} < 3.5\ \sqrt{\mathrm{GeV}}$ requirement was excluded from the selection. The band of the bottom panels includes both systematic and statistical uncertainties which are propagated in the calculation.

Figure 8.18 considers events with 2 b-tagged jets. The disagreement observed in the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance is coherent with the one observed for the $E_{\mathrm{T}}^{\mathrm{miss}}$ variable. The MC simulation of the background is missing a b-filtered low p_{T}^{V} $Z(\mu\mu)$ + jets MC sample. [2] Therefore, data and MC simulation disagreement is expected at low values of $E_{\mathrm{T}}^{\mathrm{miss}}$ and object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance observables.

The modeling of this variable is more under control in the 1-b tag category, for which this missing sample is not affecting the MC prediction. Figure 8.19 considers events with 1 b-tagged jet. In this case, MC simulation and data agree within statistical and systematic uncertainties. For data and MC simulation comparisons inclusive in the $E_{\mathrm{T}}^{\mathrm{miss}}$ categories of the resolved regime and b-tagging multiplicities, refer to Appendix F.1.

[2] At the moment in which these plots were produced, a $Z(\mu\mu)$ + jets sample with $70 < \max(\mathrm{HT}, \mathrm{p_T^V})/\ \mathrm{GeV} < 140$ with b-jet filter was missing, which impacts at low $E_{\mathrm{T}}^{\mathrm{miss}}$ and $\mathcal{S}$ values. In particular, the following MC samples were missing: ZmumuMAXHTPTV70140BFilter, WenuMAXHTPTV70140BFilter.

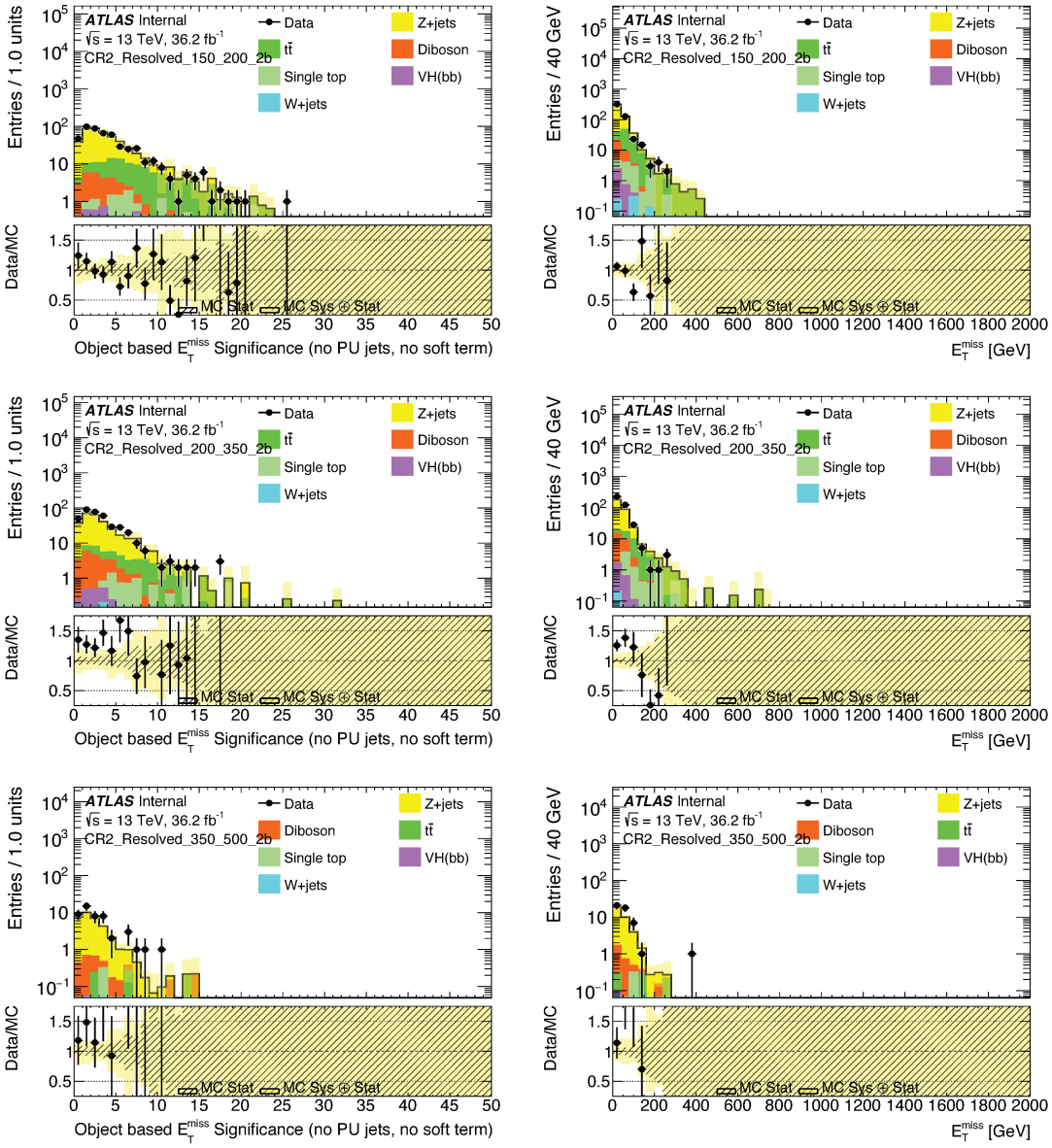


Figure 8.18 – Object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance (left) and $E_{\mathrm{T}}^{\mathrm{miss}}$ (right) distributions for data and MC simulation in the 2 lepton control region, without the event-based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance applied in event selection, for events with 2 b-tagged jets in different $E_{\mathrm{T}}^{\mathrm{miss}}$ categories in the resolved regime. Band of the bottom panels includes both systematic and statistical uncertainties. The mismodeling at low $E_{\mathrm{T}}^{\mathrm{miss}}$ values is due to a missing $Z(\mu\mu)$ + jets MC sample.

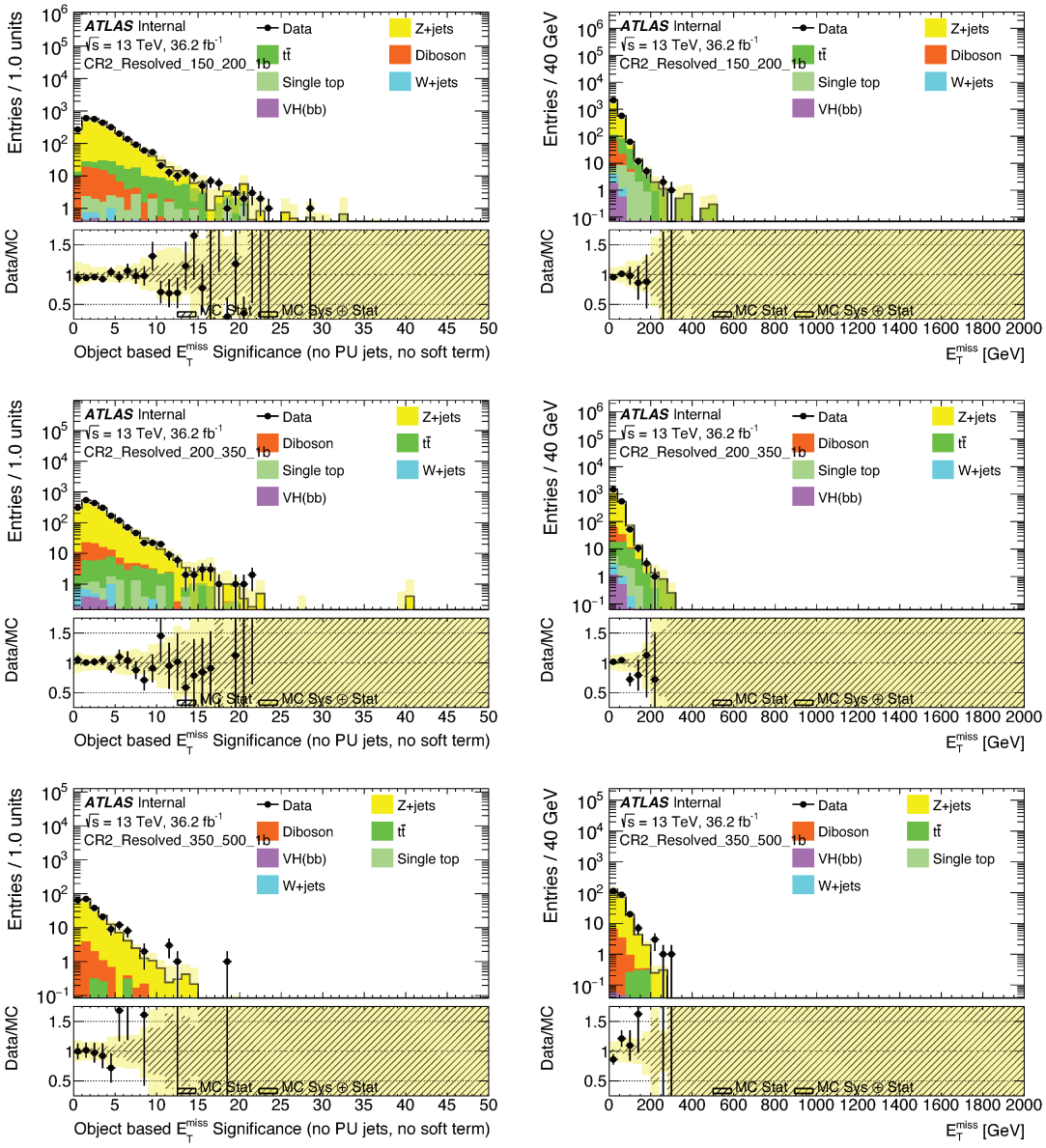


Figure 8.19 – Object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance (left) and $E_{\mathrm{T}}^{\mathrm{miss}}$ (right) distributions for data and MC simulation in the 2 lepton control region, without the event-based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance applied in event selection, for events with 1 b-tagged jets in different $E_{\mathrm{T}}^{\mathrm{miss}}$ categories in the resolved regime. Band of the bottom panels includes both systematic and statistical uncertainties.

8.4 Event selection

The same event selection strategy described in detail on Section 6.4 is used for the new analysis result with data recorded during 2015, 2016 and 2017. The main motivation for this analysis is the commissioning of new object reconstruction and performance techniques as it is introduced in Section 8.1. Because of this, the changes in the event selection are related to the implementation of the Variable Radius (VR) track jets in the merged regime and to the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance $\mathcal{S}$ requirement in the resolved regime. Also, only events with exactly 2 b-tag jets are considered since the gain in sensitivity provided by the 1 b-tag category is quite small.

In the following a quick overview of the event selection will be described. For more details refer to Section 6.4.

The reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ (or an equivalent quantity, depending on the lepton multiplicity) is required to be larger than 150 GeV. Multijet background events can pass this selection if the energy of a jet is poorly measured or if semi-leptonic hadron decays occur inside a jet, both leading to sizeable $E_{\mathrm{T}}^{\mathrm{miss}}$. To further suppress this background, additional cuts are introduced:

- $\min(\Delta\phi(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}, (\mathrm{central})\mathrm{jets}_{1,2,3})) > 20°$.
- $\Delta\Phi(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}, \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}) < 90°$

Events containing a τ-candidate are also rejected.

Events passing the above selections are divided into resolved regions with $E_{\mathrm{T}}^{\mathrm{miss}} <$ 500 GeV and a merged region with $E_{\mathrm{T}}^{\mathrm{miss}} \geq 500$ GeV. The resolved region is further divided into three exclusive regions according to the $E_{\mathrm{T}}^{\mathrm{miss}}$ with [150,200) GeV, [200,350) GeV and [350,500) GeV.

8.4.1 Resolved regime event selection

In the resolved regime, events with at least two small-R jets are considered. Exactly two of the jets are required to be b-tagged and the Higgs boson candidate is reconstructed from the these b-jets. The scalar sum of the p_{T} of these two jets and of the highest-p_{T} additional jet, if present, has to be larger than 63% of the scalar p_{T} sum of all jets in the event to reject $t\bar{t}$ events. This cut is called the "H_T ratio".

In order to further reduce background from multijet events, the following selection criteria are applied:

- $\Delta\Phi(\mathbf{p}_{\mathrm{T}}(j_1), \mathbf{p}_{\mathrm{T}}(j_2)) < 140°$
- $\Delta\Phi(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}, \ \mathbf{p}_{\mathrm{T}}(h_{jj})\ ) > 120°$
- $\Delta R(j_1, j_2) < 1.8$

- $\mathcal{S} > 16$ (detailed in Section 8.3.2)

For events with two (three or more) jets, the scalar p_{T} sum of the first two (three) jets has to be greater than 120 (150) GeV.

8.4.2 Merged regime event selection

The merged regime selection requires the presence of at least one large-R jet. The two leading VR track jets associated with the leading p_{T} large-R jet are required to be b-tagged and pass the minimum angular separation cut, $\Delta R(\mathrm{VR}_1, \mathrm{VR}_2) > R_{\mathrm{min}}$ with $R_{\mathrm{min}} = 0.2$ as described in Section 8.3.1. Events that contain one or more b-tagged VR track jets outside the large-R jet are rejected.

The p_{T} of the leading large-R jet is required to be greater than 43% of the scalar sum of the transverse momenta of the leading large-R jet and all small-R jets outside it, to reduce the contribution from $t\bar{t}$ events. This cut is the "H_T ratio" cut for the merged regime.

8.4.3 Signal Region and Control Regions

The signal region (SR) selection retains only events that do not contain any loose electrons or muons. Figure 8.20 shows the selection efficiency for a signal with $m_A = 500\,\mathrm{GeV}$ as a function of $m_{Z'}$ for events with two b-jets. The selections applied are the same as in Figure 8.5, i.e. the full event selection with the additional Higgs-candidate mass window cut of [70, 140] GeV. The branching ratio for the decay into $b\bar{b}$ is not included. Figure 8.20 illustrates the complementarity of the resolved (dashed line) and merged (solid line) regimes, leading to an overall efficiency between approximately 20% and 35% in the 2b-tag category over most of the mass range, as shown in black.

Events containing leptons, on the other hand, are used to define control regions (CRs) to constrain the main background contributions, as described in Section 6.5.1. To estimate the backgrounds from W + jets and $t\bar{t}$ production, a control sample is defined by requiring exactly one signal muon and no further loose muons or electrons, 1μ-CR. In order to mimic the way these processes can contribute in the signal region, the muon momentum is added to the missing momentum vector. The almost purely calorimeter-based missing transverse energy obtained in this way is labelled $E_{\mathrm{T,no}\mu}^{\mathrm{miss}}$.

To constrain the $Z(\nu\nu)$ + jets background contribution, events containing exactly two same-flavour signal leptons and no further loose electrons or muons are used both in the muon and in the electron channel. These regions are referred to as the 2e-CR, the 2μ-CR or, collectively as the 2ℓ-CR. In the muon channel, the two muons are required to have opposite charge. The electron (muon) pair must have an invariant mass of $83\,\mathrm{GeV} < m_{ll} < 99\,\mathrm{GeV}$ ($71\,\mathrm{GeV} < m_{ll} < 106\,\mathrm{GeV}$). In order to reduce the $t\bar{t}$ and single top background, the event-based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance is required to be less than $3.5\sqrt{\mathrm{GeV}}$. The p_{T} of the dilepton system, $p_{\mathrm{T}}^{\ell\ell}$, corresponds to the p_{T} of the Z boson and as such serves as a proxy for the $E_{\mathrm{T}}^{\mathrm{miss}}$, in analogy to $Z(\nu\nu)$ + jets events where the $E_{\mathrm{T}}^{\mathrm{miss}}$ from the neutrinos also corresponds to the boson p_{T}.

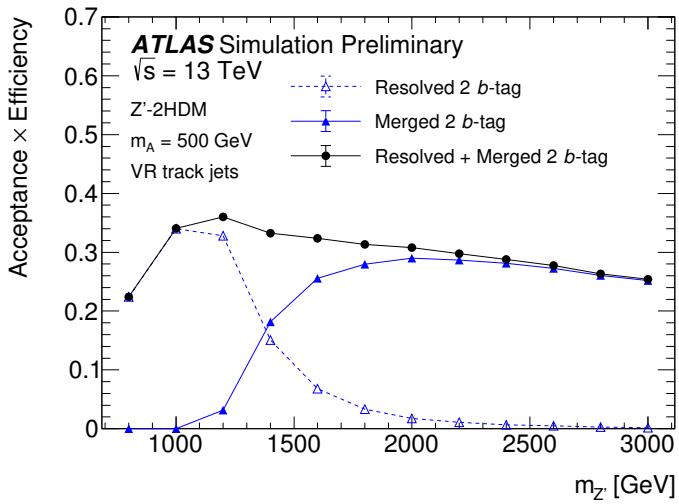


Figure 8.20 – Acceptance × efficiency for events with 2 b-tagged jets as a function of $m_{Z'}$, for a fixed mass $m_A = 500$ GeV, $\tan\beta = 1.0$, $g_Z = 0.8$, $m_\chi = 100$ GeV, $m_H = m_{H^\pm} = 300$ GeV. The values obtained for the resolved regime are shown in blue with open triangles and a dashed line, the ones for the merged regime with filled triangles and a solid line. The combined selection efficiency is shown in black [5].

Figure 8.21 shows the distribution of the variables used to mimic $E_{\mathrm{T}}^{\mathrm{miss}}$ in the control regions. The $E_{\mathrm{T,no}\mu}^{\mathrm{miss}}$ variable is shown at the top for events in the 1μ-CR with a positively charged muon on the left and with a negatively charged muon on the right. The bottom figure shows the $p_{\mathrm{T}}^{\ell\ell}$ distribution in the 2ℓ-CR, combining the electron and muon channel. The filled histograms correspond to the background prediction after the combined fit described in Section 8.7, while the dashed blue line indicates the total background expectation before the fit. The post-fit predictions agree with the data (black points) within the uncertainties (grey shaded areas). The various sources of uncertainties considered are discussed in Section 8.6.

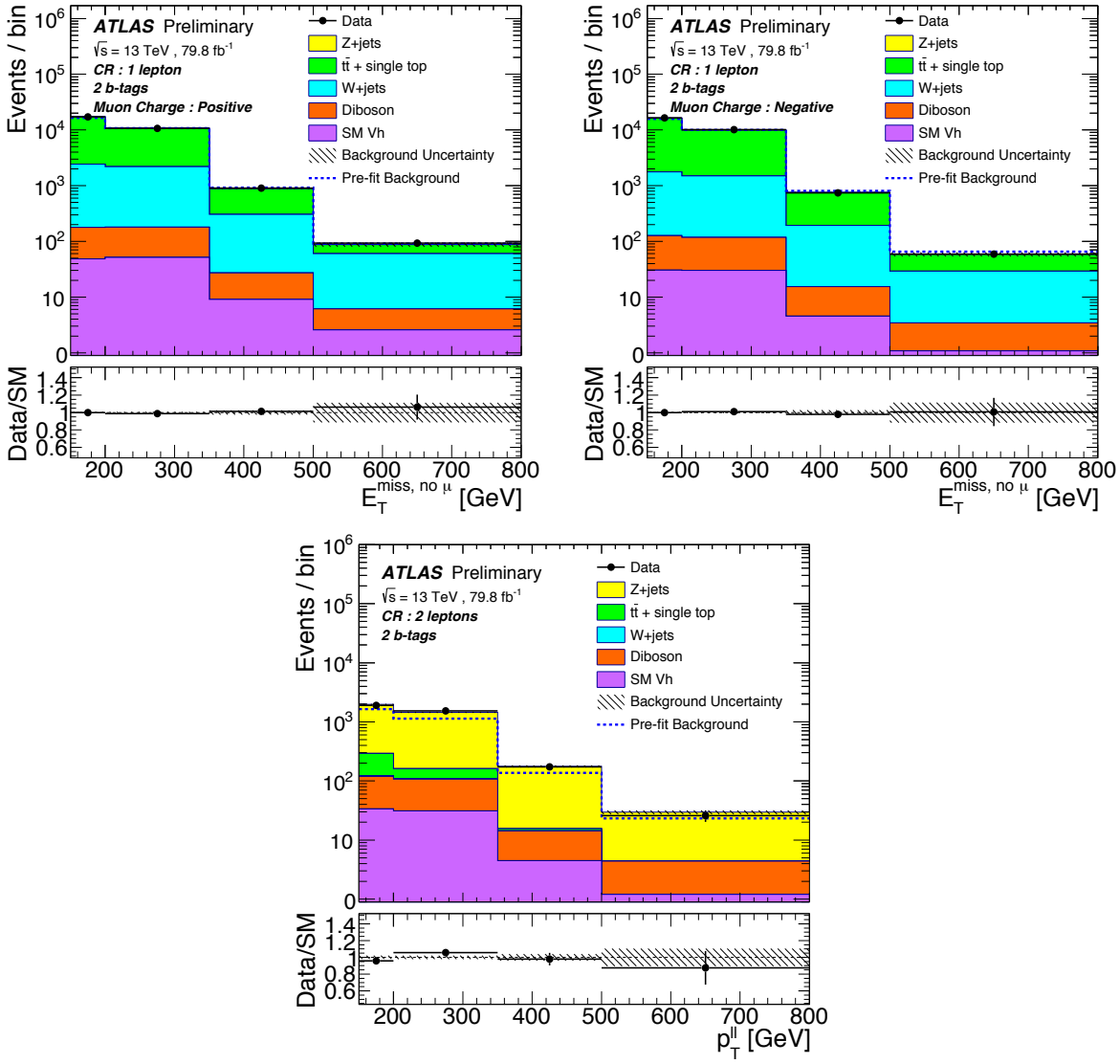


Figure 8.21 – Distributions of $E_{\mathrm{T,no}\mu}^{\mathrm{miss}}$ in the 1μ-CR (top) and $p_{\mathrm{T}}^{\ell\ell}$ in the 2ℓ-CR (bottom). The distribution of the 1μ-CR is shown for events with positively (left) and negatively (right) charged muons separately. The distributions in the the 1μ-CR are separated by the muon charge because the fit uses the muon charge as the discriminating variable to separate the $t\bar{t}$ process from the W + jets process. The upper panels show a comparison of data to the SM expectation before (dashed lines) and after the background-only fit (solid histograms). The bottom panel displays the ratio of data to SM expectations after the background-only fit, with its systematic uncertainty considering correlations between individual contributions indicated by the hatched band. The rightmost bin includes overflows [5].

8.5 Background estimation

Main SM backgrounds to the signature of two b-jets and large $E_{\mathrm{T}}^{\mathrm{miss}}$ are the production of top-quark pairs, $t\bar{t}$, and of heavy vector bosons with additional b-jets, $Z(\nu\nu)$ +jets and W + jets, as described in Section 6.5.1. The acceptance of these backgrounds is estimated using MC simulation, and the analysis uses data control samples to determine the normalisation of these backgrounds.

The signal region definition includes a veto on the presence of a charged lepton. The control region definition is made orthogonal to the signal region by different requirements on the number of charged leptons in the events considered. The search is performed considering four disjoint $E_{\mathrm{T}}^{\mathrm{miss}}$ regions. The event selection for signal region and control regions is defined in Section 8.4.

Sub-dominant backgrounds are estimated from simulations, with the exception of QCD multijet events. This contribution is determined in a data-driven way based on ratios between several control samples and employing the object-based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance, as it is detailed in the following Section 8.5.1.

A simultaneous profile-likelihood fit to the control and signal regions is performed to constrain the backgrounds and extract information about the potential presence of a signal, respectively. This is detailed in Section 8.7.

8.5.1 Multijet background estimation

The multijet processes originating from pure strong interactions is one of the sub-dominant backgrounds in the resolved SR due to the requirement of large $E_{\mathrm{T}}^{\mathrm{miss}}$. This background is challenging to model by means of MC simulation due to the high statistics which would be required. Therefore, this background is estimated in a data-driven way, which accounts not only for the total predicted number of multijet background events, but also for the multijet shape of the h_{jj} invariant mass.

For the merged SR and the two CRs (1μ-CR and 2ℓ-CR), the multijet background has been found to be negligible [3], see Section 6.5.4 for details in the multijet estimation in the previous result. This section describes a different data-driven strategy from the one detailed in Section 6.5.4, making use of the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance variable in order to estimate the multijet background contribution in the SR. Appendix G describes other alternative multi-jet estimations that where studied before considering the $\mathcal{S} > 16$ requirement in the resolved SR event selection, as described in Section 8.3.2.

In order to derive template shapes of multijet distributions for different variables, a multijet enriched region, QCDCR, is defined by inverting the cuts associated to the most powerful observables in rejecting multijet events: $\min(\Delta\Phi(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}},\text{small-R jets})) > 20°$ and the cut on the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance > 16, as can be seen in Figure 8.7(b). In the resulting selection multijet events dominate. For the previous analysis iteration, the multijet enriched region was defined by inverting just the $\min(\Delta\Phi(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}},\text{small-R jets}))$ cut, see Figure 6.11 for the scheme used.

Therefore, in the region of $\min(\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}},\text{small-R jets}))<20°$ and $\mathcal{S} \leq 16$, the difference between the data and the simulated non-multijet backgrounds is used as a model of the shape of multijet distributions.

In a second step, these shapes need to be scaled with a normalisation factor to account for the differences in efficiency of the min($\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}}$,small-R jets)) and the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance cuts in the QCDCR and SR. The normalisation factor is determined with the so called "ABCD method".

The strategy of determining this normalisation factor via a fake factor from a fit using the multiplicity of jets containing a muon as multijet sensitive variable with loosened anti-QCD cuts is not considered here because it was found that this variable is not shape invariant under QCDCR and QCDSR selections. This strategy is described in Section 6.5.4 and Figure G.3 on Appendix G shows the dependencies of multijet template shapes under QCDCR and QCDSR selections.

For illustration purposes, the distributions for data recorded during 2015 and 2016 and MC simulation in the QCDCR, from which the shapes can be derived, are shown in Figure 8.22.

The template shapes extracted by subtracting MC simulation to the complete data-set, without normalisation factor, are shown in Figure 8.23. These shapes of m_{jj} distributions needs to be scaled to model the multijet background in the SR (min($\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}$,jets))>20°, $\mathcal{S} > 16$), assuming min($\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}$,jets)) and $\mathcal{S}$ distributions are independent.

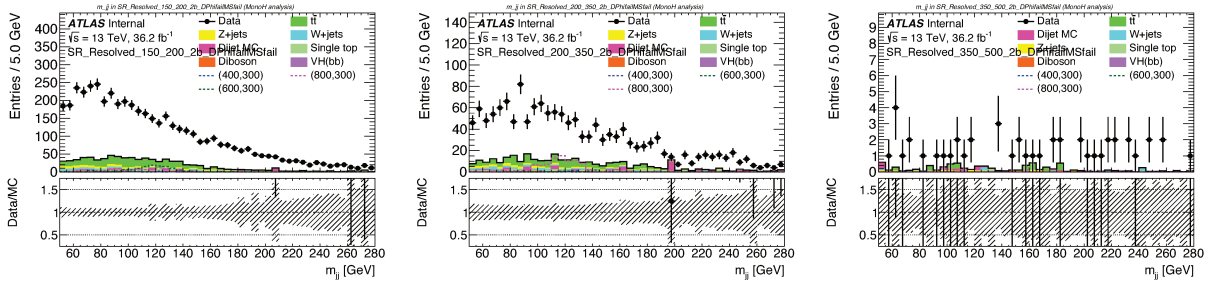


Figure 8.22 – Higgs boson candidate mass distribution for data data recorded during 2015 and 2016 and MC simulation in the QCDCR. Selected events has exactly 2 b-tagged jets and 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 200 GeV (left), 200 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 350 GeV (middle), and 350 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 500 GeV (right) in the QCD CR.

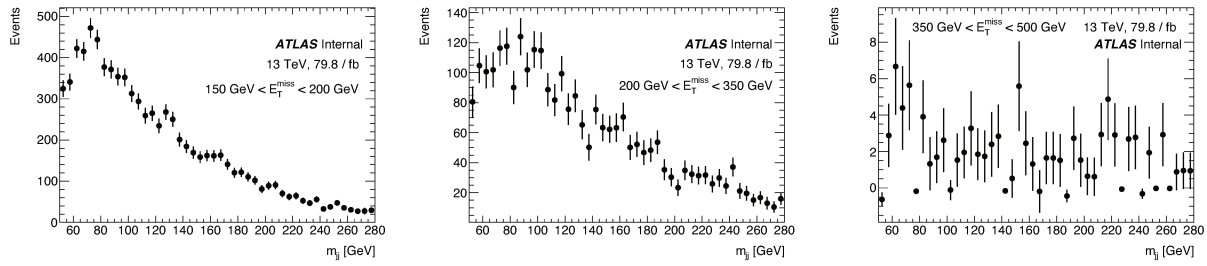


Figure 8.23 – Multijet m_{jj} shape obtained by subtracting simulated backgrounds from data in QCDCR for events with 2 b-tagged jets and 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 200 GeV (left), 200 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 350 GeV (middle), and 350 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 500 GeV (right).

The ABDC method considers four disjointed regions (A, B, C and D) where the number of events of the reconstructed Higgs boson mass template are extracted. These regions

are defined by failing and/or passing the min($\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}$,jets))>20° and $\mathcal{S} > 16$ cuts, i.e. fail-fail, fail-pass, pass-pass.

If the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance and the min $\Delta\phi$ cut are uncorrelated, then the predicted number of multijet events in the signal region is calculated as:

$$N_{\mathcal{S}>16,\Delta\phi>20^\circ} = \frac{N_{\mathcal{S}>16,\Delta\phi\leq 20^\circ} \cdot N_{\mathcal{S}\leq 16,\Delta\phi>20^\circ}}{N_{\mathcal{S}\leq 16,\Delta\phi\leq 20^\circ}}. \tag{8.3}$$

Figure 8.24 shows the min($\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}$,jets)) and $\mathcal{S}$ distributions for the three resolved $E_{\mathrm{T}}^{\mathrm{miss}}$ categories with 2-b tagged jets.

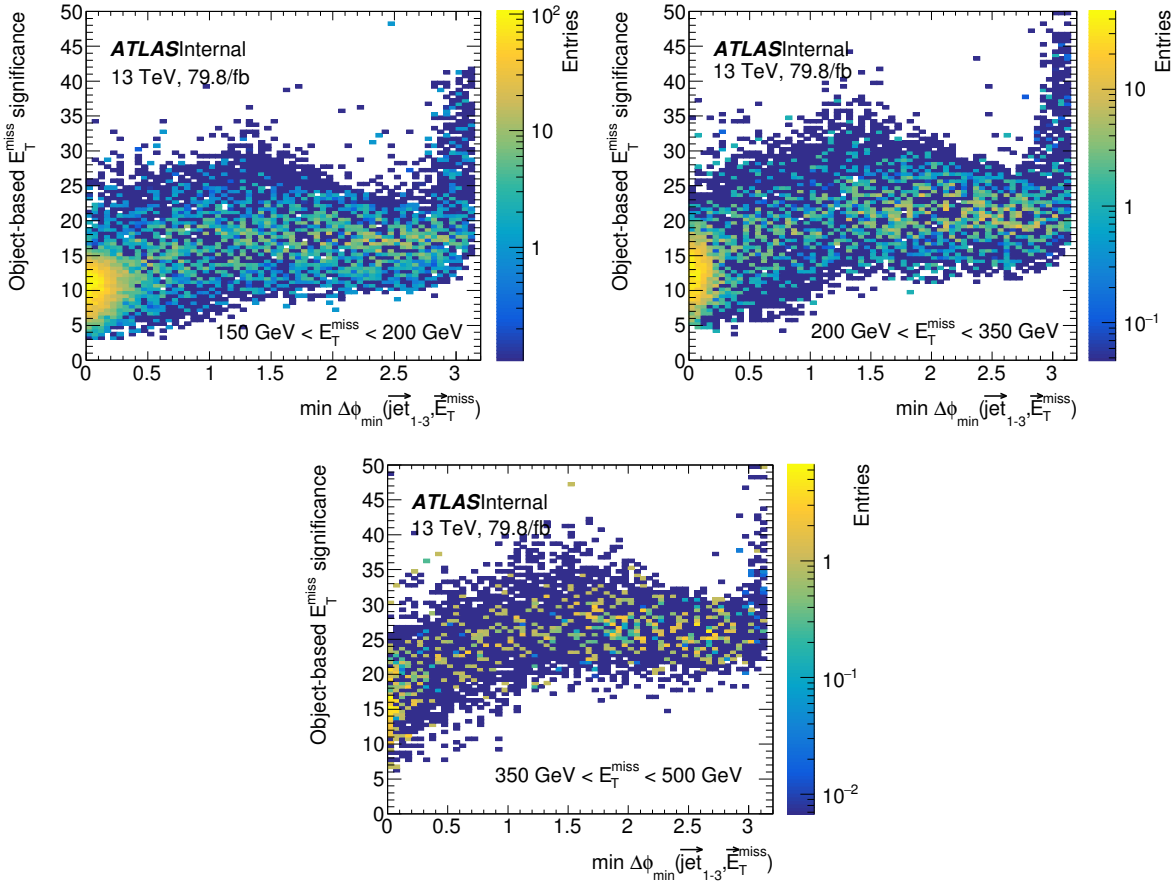


Figure 8.24 – 2D distribution of min $\Delta\phi$ variable vs. object based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance with all analysis cuts applied except requirements on $E_{\mathrm{T}}^{\mathrm{miss}}$ significance and min $\Delta\phi$ in the mass side-bands of the signal region (50 GeV $< m_{jj} <$ 70 GeV and 140 GeV $< m_{jj} <$ 280 GeV) for events with 2 b-tagged jets and respectively 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 200 GeV (top left), 200 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 350 GeV (top right), and 350 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 500 GeV (bottom).

In order for the ABCD method to hold, the variables defining the regions need to be uncorrelated. Figure 8.25 shows the correlation between the $E_{\mathrm{T}}^{\mathrm{miss}}$ significance and the min $\Delta\phi$ variable The $E_{\mathrm{T}}^{\mathrm{miss}}$ significance (normalised to unity) is shown for different ranges of min $\Delta\phi$. The distribution is invariant within the statistical uncertainty.

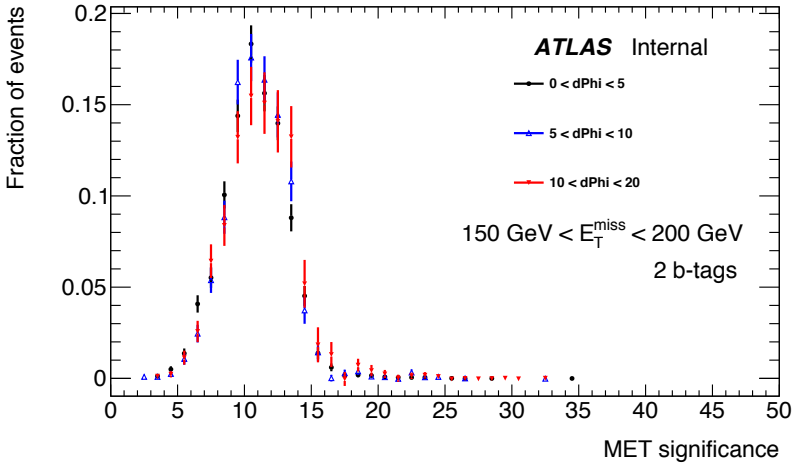


Figure 8.25 – $E_{\mathrm{T}}^{\mathrm{miss}}$ significance distribution (without the $\mathcal{S} > 16$ requirement) for different selections of min $\Delta\phi$, showing events with 2 b-tagged jets and 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 200 GeV.

Also, the extrapolation of the dijet invariant mass shapes from the QCDCR to the SR is only possible if the Higgs boson candidate mass distribution is invariant under the ABCD method regions related to the min($\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}$,jets)) and $\mathcal{S}$ observables. Figure 8.26 shows the Higgs candidate mass distribution for different values of min $\Delta\phi$ with 2 b-tagged jets in the first $E_{\mathrm{T}}^{\mathrm{miss}}$ category.

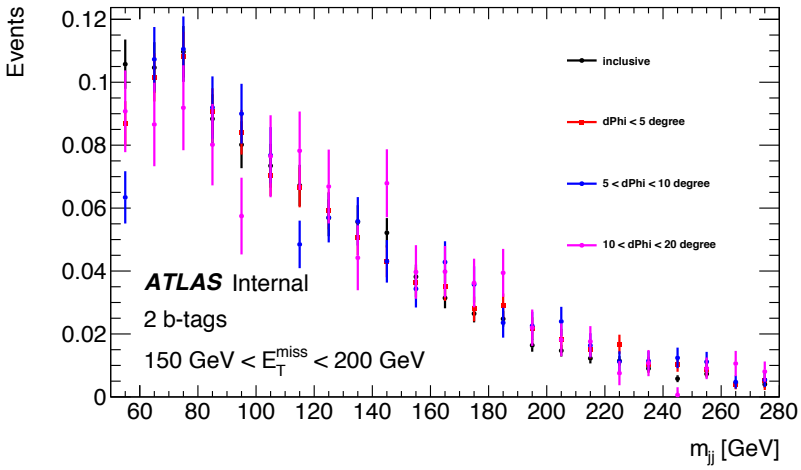


Figure 8.26 – Higgs candidate mass distribution for different values of min $\Delta\phi$, showing events with 2 b-tagged jets and 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 200 GeV.

The mass shapes for the different regions with either both or only one of the two cuts passed/failed are shown in Figure 8.27 for events with 2 b-tagged jets and 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 200 GeV. Within large statistical uncertainties, the template shapes of the Higgs candidate mass are similar in the regions where min($\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}$,jets))>20° ($\phi$ label), $\mathcal{S} > 16$

(MS label) cuts fail-fail, fail-pass, pass-pass. In order not to compromise the estimation

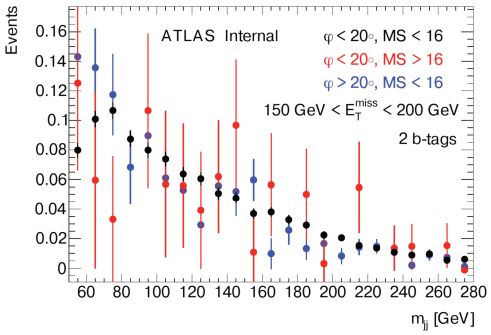


Figure 8.27 – Higgs candidate mass distribution for different values of min $\Delta\phi$ and $\mathcal{S}$, showing events with 2 b-tagged jets and 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 200 GeV.

given by Equation 8.3 with signal contamination, the ABCD method is carried out with all regions blinded, i.e. with the requirement 50 GeV $< m_{jj} <$ 70 GeV or 140 GeV $< m_{jj} <$ 250 GeV. The distributions of the Higgs candidate mass in the regions used for determining the normalisation with the ABCD method are shown in Figure 8.28, 8.29 and 8.30.

The results are shown in Table 8.1.

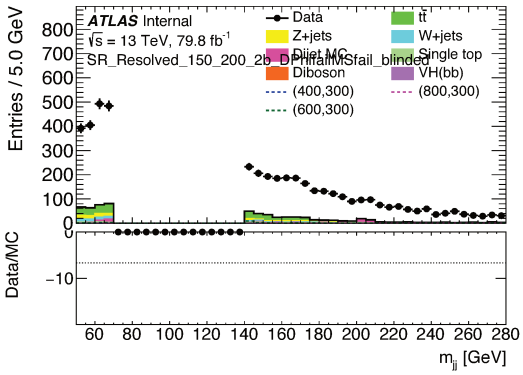
Sample	Region	$N_{\mathrm{MetSig}<16,\Delta\phi<20^\circ}$	$N_{\mathrm{MetSig}<16,\Delta\phi>20^\circ}$	$N_{\mathrm{MetSig}>16,\Delta\phi<20^\circ}$	ABCD Prediction	$N_{\mathrm{MetSig}>16,\Delta\phi>20^\circ}$
Data	150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 200 GeV	4573 ± 68	3452 ± 59	129 ± 11	–	7698 ± 88
Total background	150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 200 GeV	595 ± 13	2435 ± 28	50 ± 3	–	6397 ± 45
Multijet shape	150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 200 GeV	3978 ± 69	1017 ± 65	79 ± 12	20 ± 3.3	–
Data	200 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 350 GeV	1705 ± 41	381 ± 20	355 ± 19	–	5429 ± 74
Total background	200 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 350 GeV	282 ± 9	317 ± 10	174 ± 7	–	4731 ± 36
Multijet shape	200 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 350 GeV	1423 ± 42	64 ± 22	181 ± 20	8 ± 2.9	–
Data	350 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 500 GeV	66 ± 8	1 ± 1	94 ± 10	–	484 ± 22
Total background	350 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 500 GeV	9 ± 1	0.9 ± 0.2	52 ± 4	–	434 ± 9
Multijet shape	350 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 500 GeV	57 ± 8	0.1 ± 1	42 ± 11	0.1 ± 0.76	–

Table 8.1 – Results of the ABCD study for events with 2 b-tagged jets and $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement as specified in the Table for data from 2015, 2016 and 2017 data-taking and MC from the mc16a and mc16d campaigns. The entries relate to the blinded (i.e. mass window outside 70 to 140 GeV) signal regions.

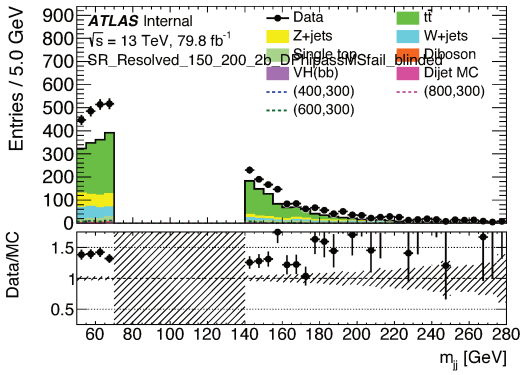
To get a multijet yield estimate for the complete mass range, the results from the ABCD method are scaled by the ratio of the full mass range over the blinded mass sideband of templates shown in Figure 8.23.

In the 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 200 GeV region the predicted multijet yield is 20 ± 3.3 events, in the 200 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 350 GeV region the predicted multijet yield is 8 ± 2.29 events and in the 350 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 500 GeV region it is 0.1 ± 0.76 events.

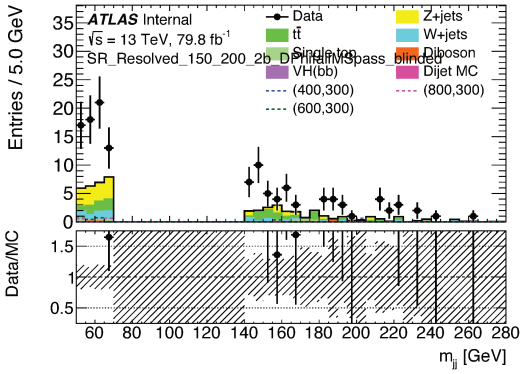
The predicted multijet yield according to the ABCD method is substantially smaller than the statistical uncertainty on the data. Since in the two lowest $E_{\mathrm{T}}^{\mathrm{miss}}$ bins the multijet background estimate is substantially smaller than the data statistical uncertainty, and since the multijet background has a $E_{\mathrm{T}}^{\mathrm{miss}}$ spectrum which is exponentially falling steeper than other backgrounds, it is expected that the multijet background is also negligible in the highest $E_{\mathrm{T}}^{\mathrm{miss}}$ bin of the resolved regime (350 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 500 GeV). Therefore, this background is neglected in the combined fit of the analysis.



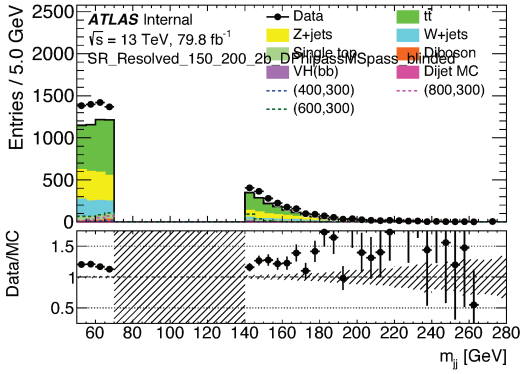
(a) MetSig < 16, $\Delta\phi < 20°$



(b) MetSig < 16, $\Delta\phi > 20°$



(c) MetSig > 16, $\Delta\phi < 20°$



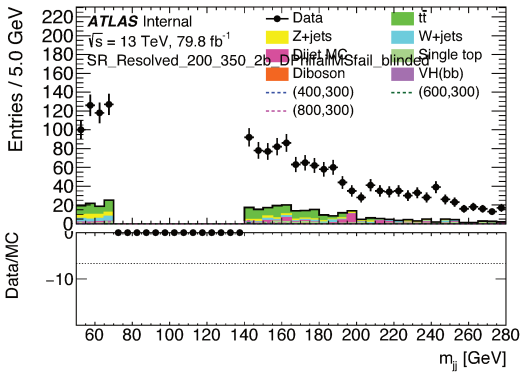
(d) MetSig > 16, $\Delta\phi > 20°$

Figure 8.28 – Regions used for ABCD method, showing events with 2 b-tagged jets and 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 200 GeV.

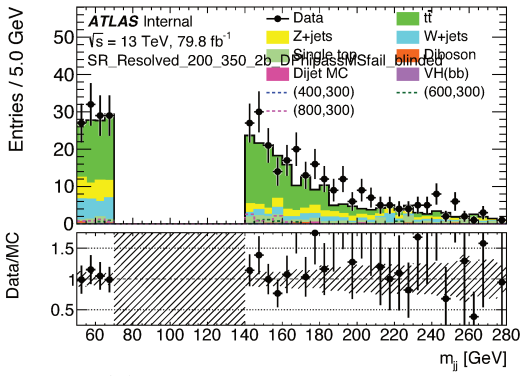
In order to check check if the multijet contribution really is negligible in the combined fit, a data-driven multijet estimate was still performed to determine pre-normalised templates for the m_{jj} shape of the multijet background to be used in the combined fit with an uncertainty on the normalisation of 100%.

The shapes from the region of min($\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}}$,small-R jets))$<20°$ and $E_{\mathrm{T}}^{\mathrm{miss}}$ significance ≤ 16 are normalised to the yield according to the ABCD method prediction, taking into account that only the mass side-band has been used for the ABCD method.

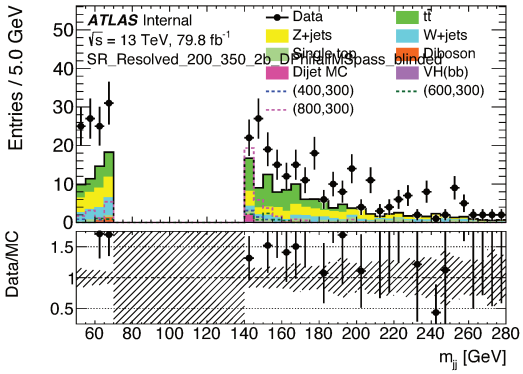
The study showed that if included, the nuisance parameters related to the multijet background normalisation get only very slightly pulled. There is almost no correlation of the multijet background normalisation nuisance parameters to other nuisance parameters in the combined fit. Therefore, It is concluded that the multijet background can be neglected in the final combined fit for all the $E_{\mathrm{T}}^{\mathrm{miss}}$ categories.



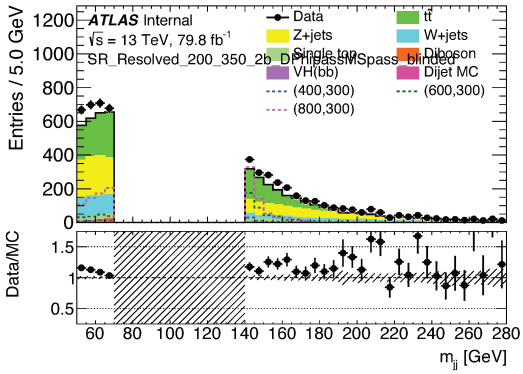
(a) MetSig < 16, $\Delta\phi < 20°$



(b) MetSig < 16, $\Delta\phi > 20°$

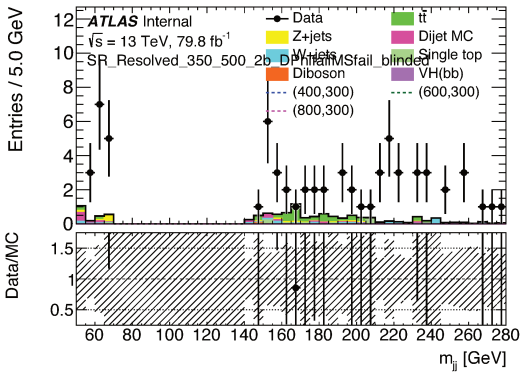


(c) MetSig > 16, $\Delta\phi < 20°$

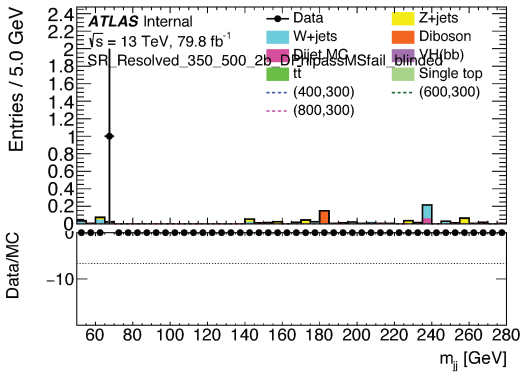


(d) MetSig > 16, $\Delta\phi > 20°$

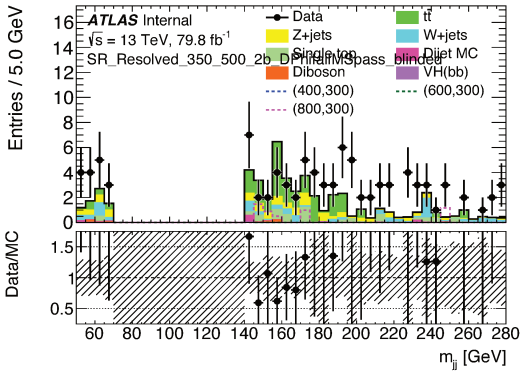
Figure 8.29 – Regions used for ABCD method, showing events with 2 b-tagged jets and 200 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 350 GeV.



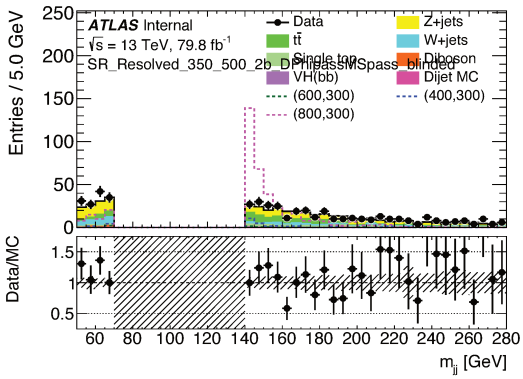
(a) MetSig < 16, $\Delta\phi < 20°$



(b) MetSig < 16, $\Delta\phi > 20°$



(c) MetSig > 16, $\Delta\phi < 20°$



(d) MetSig > 16, $\Delta\phi > 20°$

Figure 8.30 – Regions used for ABCD method, showing events with 2 b-tagged jets and 350 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 500 GeV.

8.6 Systematic uncertainties

Dominant sources of experimental systematic uncertainty arise from the calibration of the b-tagging efficiency, the integrated luminosity, as well as the scale and resolution of the energy and the mass of the jets. The uncertainties in the b-tagging efficiency originate mainly from the uncertainties on the measurement of the flavour tagging efficiency in $t\bar{t}$ events [270, 162], and from the extrapolation to high-p_{T} jets as described in Section 8.3.1. The uncertainty in the combined 2015-2017 integrated luminosity is 2.0%. It is derived following a methodology similar to that detailed in Reference [271] from calibrations of the luminosity scale. The small-R and the large-R jets are calibrated separately, with a series of simulation-based corrections and in situ techniques, taking into account differences between MC and real data [272, 157]. Another main source is the uncertainties on $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiencies, which was determined inclusively for all b-tagged jet multiplicities. An uncertainty on the $E_{\mathrm{T}}^{\mathrm{miss}}$ soft-term resolution and scale is also taken into account [179]. Other uncertainties which have a smaller impact are the uncertainties on the lepton energy and momentum scales, and the lepton identification and trigger efficiencies [166, 168, 170].

Dominant sources of theoretical systematic uncertainty originate from the number of MC events, the modelling of the signal and the background processes listed in the following: $t\bar{t}$, V+jets, associated production of the SM Higgs boson decaying to $b\bar{b}$($Vh(b\bar{b})$), and diboson production. The evaluation of the theory uncertainties follows the same strategy described in Reference [273].Signal acceptance uncertainties are evaluated by varying the parameters of the MC samples, and comparing the results by applying the analysis selection at generator level. Generator tuning, variations of parton distribution functions, and scale uncertainties are taken into account. Leading background (V+jets and $t\bar{t}$) processes are normalised to the data in the final likelihood fit by using dedicated CRs, and other backgrounds are constrained by the theory uncertainties on the overall cross-section of the background [252]. Shape uncertainties on the p_{T} distributions of the W/Z boson and the reconstructed dijet mass spectra are evaluated by comparing samples obtained with different MC generator settings, as well as comparison to dedicated Z control region data. An overall normalisation uncertainty is used for the SM $Vh(b\bar{b})$ background [273]. Table 8.2 quantifies the dominant sources of uncertainties in terms of the relative uncertainty on the signal strength, defined as the ratio of the measured signal yield to the theory prediction, after the fit to simulated data including the signal assuming three representative Z'-2HDM-scenarios. This search is statistically limited for highly-boosted signatures.

Table 8.2 – Dominant sources of uncertainty for three representative Z'-2HDM scenarios after the fit to simulated data including the signal: (a) $(m_{Z'}, m_A) = (0.6\ \text{TeV}, 0.3\ \text{TeV})$, (b) $(m_{Z'}, m_A) = (1.4\ \text{TeV}, 0.6\ \text{TeV})$, and (c) $(m_{Z'}, m_A) = (2.6\ \text{TeV}, 0.3\ \text{TeV})$. The effect is expressed as the relative uncertainty on the signal strength, assuming total cross-sections of (a) 452 fb, (b) 3.75 fb, and (c) 2.03 fb. The three Z'-2HDM scenarios are chosen to represent kinematics from resolved to merged regions, and they are close to the expected exclusion limit. The total uncertainty is the quadrature sum of statistical and total systematic uncertainties.

Source of uncert.	Impact on signal strength [%]		
	(a)	(b)	(c)
b-tagging	4.0	8.0	10
V+jets modeling	3.5	6.0	5.0
Top modeling	3.7	4.8	4.5
MC statistics	1.8	5.4	4.9
SM $Vh(b\bar{b})$	0.8	3.2	2.1
Diboson modeling	0.8	1.5	1.1
Signal modeling	3.0	2.5	1.5
Luminosity	2.0	2.5	2.5
Small-R jets	1.4	3.0	2.0
Large-R jets	0.2	1.0	2.0
$E_{\text{T}}^{\text{miss}}$	1.2	1.7	1.1
Leptons	0.2	0.8	0.7
Total syst. uncert.	6.5	13	13
Statistical uncert.	2.3	20	22
Total uncertainty	7	24	25

8.7 Results

A fit to the invariant mass of the Higgs boson candidate m_{jj} is used to search for a signal, where m_{jj} is represented by the dijet invariant mass of the two leading small-R jets in the resolved SR, and the leading large-R jet mass m_J in the merged SR. The fit is based on a binned likelihood approach, and systematic uncertainties are included in the likelihood function as nuisance parameters with Gaussian or log-normal constraints and profiled [273]. To account for changes in the background composition and to benefit from a higher signal sensitivity with increasing $E_{\mathrm{T}}^{\mathrm{miss}}$, the data in the SR and the two CRs are split into four ranges in $E_{\mathrm{T}}^{\mathrm{miss}}$ ($E_{\mathrm{T}}^{\mathrm{miss,no}^-}$ in 1μ-CR, and $p_{\mathrm{T}}^{\ell\ell}$ in 2ℓ-CR) that are fit simultaneously: [150 GeV, 200 GeV), [200 GeV, 350 GeV), [350 GeV, 500 GeV), and [500 GeV, ∞). In the 1μ-CR, the fit uses the electric charge of the μ as the discriminating variable to separate the $t\bar{t}$ process from the W+jets process, since $t\bar{t}$ provides an equal number of μ^+ and μ^-, while a prevalence of μ^+ is expected from W+jets. In the 2ℓ-CR, the fit uses only the total event yield due to limited data statistics. Table 8.3 summarises the event categories used in the fit.

The V+jets background in the SR is dominated by jets containing b- or c-quarks, denoted as heavy flavour (HF) jets. The normalisations of $t\bar{t}$, W+HF, and Z+HF processes are free parameters in the fit. For V+jets, the flavour composition of the two jets being bb, bc, b and LF (LF = light-flavour quark), and cc are subject to systematic uncertainties. The post-fit normalisation factors of $t\bar{t}$, W+HF, and Z+HF are 1.10 ± 0.08, 1.51 ± 0.22, and 1.42 ± 0.10, respectively. The normalisations of other backgrounds modelled using MC simulations are constrained to the theory predictions within uncertainties, as detailed in Reference [3].

The distributions of m_{jj} or m_J for the SR are shown in the four $E_{\mathrm{T}}^{\mathrm{miss}}$ regions in Figure 8.31. The $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution for all SR regions combined is shown in Figure 8.32. The number of expected background events for each process after the fit, as well as the observed number of data events in each $E_{\mathrm{T}}^{\mathrm{miss}}$ region, are summarised in Table 8.4. No significant deviation from the SM expectation is found. The statistical model is further validated by replacing the signal with the $ZZ \to \nu\nu b\bar{b}$ process, and the signal strength of ZZ production extracted from the fit agrees with the SM prediction.

The results are interpreted as exclusion limits at 95% confidence level (CL) on the production cross-section of $h + \mathrm{DM}$ events $\sigma_{h+\mathrm{DM}}$ times the branching ratio $\mathcal{B}(h \to b\bar{b})$ with the $\mathrm{CL_s}$ formalism [244] using a profile likelihood ratio test statistic. Exclusion contours in the $(m_{Z'}, m_A)$ space in the Z'-2HDM scenario are presented in Figure 8.33. The improvement from using VR track jets instead of jets with a fixed cone size of 0.2 (FR) is shown in Figure 8.34, which compares the expected upper limits on the signal strength μ, using the same luminosity in the 2 b-tagged region. Other differences between the two analyses include the suppression of the multijet background using the object-based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance, and reduced uncertainties from the MC statistics and the calibration of the b-tagging efficiency in the VR analysis. For signals with highly boosted Higgs bosons, the results with VR track jets show significant improvement of up to 200%. Figure 8.35 compares the expected upper limits on μ against the previous iteration (scaled to 79.8 fb^{-1}) performed with FR track jets and both one and two b-tagged events included in the SR [3]. The analysis based on VR track jets outperforms the predecessor based on

Table 8.3 – Event categories used in the fit.

	0 lepton	**1 muon**	**2 leptons**
Region	SR	$t\bar{t}$ and W+jets CR	Z+jets CR
$E_{\mathrm{T}}^{\mathrm{miss}}$ or $E_{\mathrm{T}}^{\mathrm{miss}}$ proxy	$E_{\mathrm{T}}^{\mathrm{miss}}$	$E_{\mathrm{T}}^{\mathrm{miss,\,no\,\mu}}$	$p_{\mathrm{T}}^{\ell\ell}$
	Resolved: [150,200), [200,350) and [350,500) GeV Merged: Larger than 500 GeV		
Fit variable in each $E_{\mathrm{T}}^{\mathrm{miss}}$ bin	m_{jj}	muon charge	Event yield

FR track jets considerably, justifying the strategy of using only events with two b-tagged jets as described in Section 8.3.1.

Table 8.4 – Numbers of expected background events for each background process after the background-only profile likelihood fit, the sum of all background components after the fit, and observed data yields for events with two b-tags in the resolved and merged channels for each $E_{\mathrm{T}}^{\mathrm{miss}}$ region. The multijet background is negligible and not included in the fit. Statistical and systematic uncertainties are combined. The uncertainties in the total background take into account the correlation of systematic uncertainties among different background processes. The uncertainties on the total background can be smaller than those on individual components due to anti-correlations between nuisance parameters.

Category	Range in $E_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]			
	$[150,200)$	$[200,350)$	$[350,500)$	$[500,\infty)$
W+jets	3020 ± 530	2240 ± 360	184 ± 32	26.4 ± 5.7
Z+jets	6330 ± 450	5180 ± 340	565 ± 37	80.5 ± 6.3
$t\bar{t}$ + single top quark	11800 ± 350	6450 ± 200	308 ± 25	10.8 ± 2.5
Diboson	438 ± 67	400 ± 59	49.0 ± 11	9.37 ± 1.7
Vh	136 ± 39	129 ± 37	17.3 ± 5.0	3.86 ± 1.1
Bkg	21700 ± 140	14400 ± 110	1120 ± 25	131 ± 7.2
Data	21818	14350	1128	119

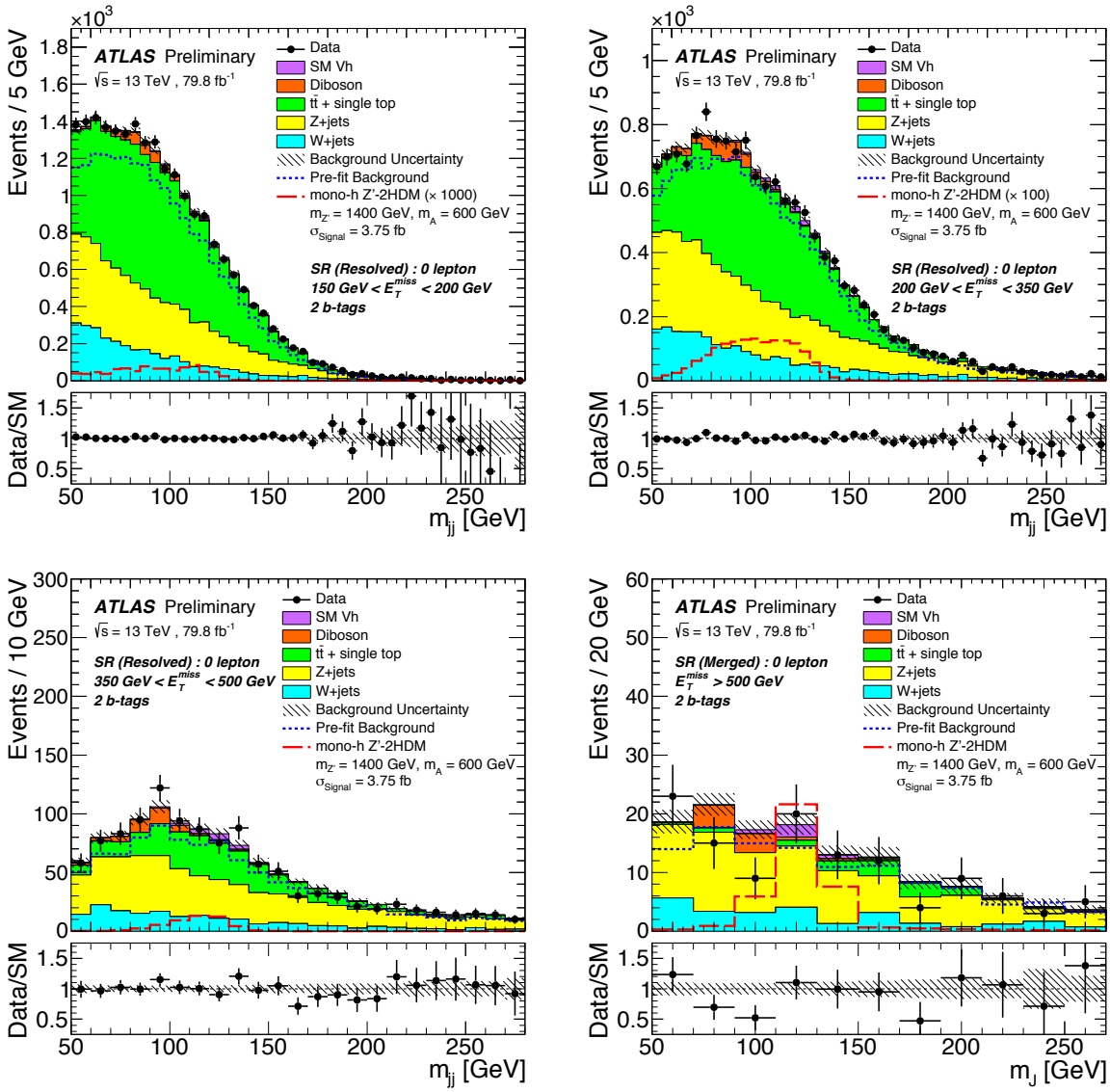


Figure 8.31 – Distributions of the invariant mass of the Higgs boson candidates $m_{jj} = m_{jj}, m_J$ with two b-tagged jets in the SR for the four $E_{\mathrm{T}}^{\mathrm{miss}}$ categories that are used as inputs to the fit. The upper panels show a comparison of data to the SM expectation before (dashed lines) and after the fit (solid histograms) with no signal included. The lower panels display the ratio of data to SM expectations after the background-only fit, with its systematic uncertainty considering correlations between individual contributions indicated by the hatched band. The expected signal from a representative Z'-2HDM model is also shown (long-dashed line), and it is scaled up by a factor of 1000 and 100 for the lowest two $E_{\mathrm{T}}^{\mathrm{miss}}$ bins [150 GeV, 200 GeV) and [200 GeV, 350 GeV), respectively [5].

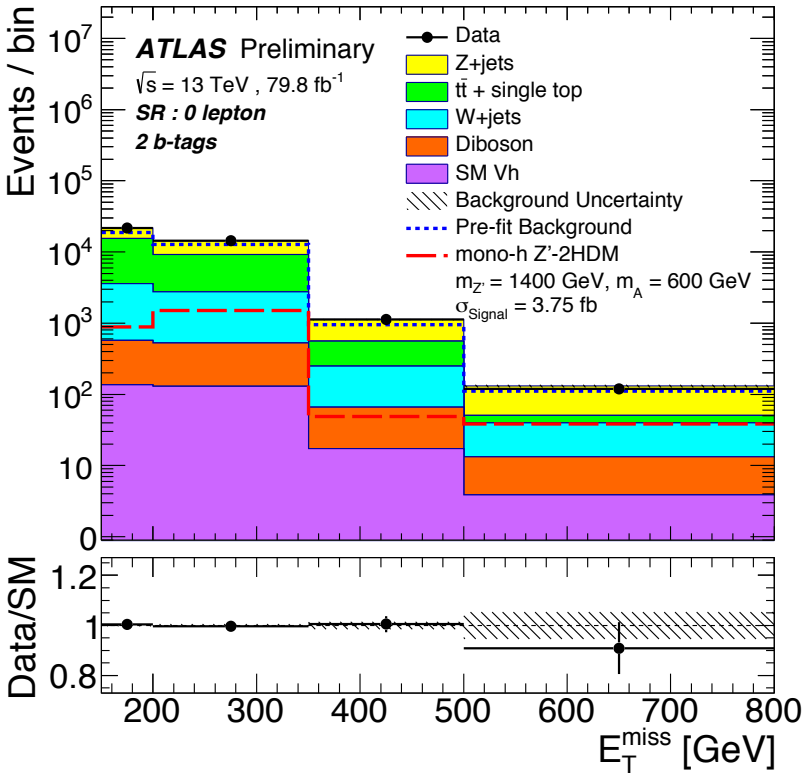


Figure 8.32 – $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution for the resolved and the merged signal regions combined. The upper panel shows a comparison of data to the SM expectation before (dashed lines) and after the fit (solid histograms) with no signal included. The lower panels display the ratio of data to SM expectations after the background-only fit, with its systematic uncertainty considering correlations between individual contributions indicated by the hatched band. The expected signal from a representative Z'-2HDM model is also shown (long-dashed line) [5].

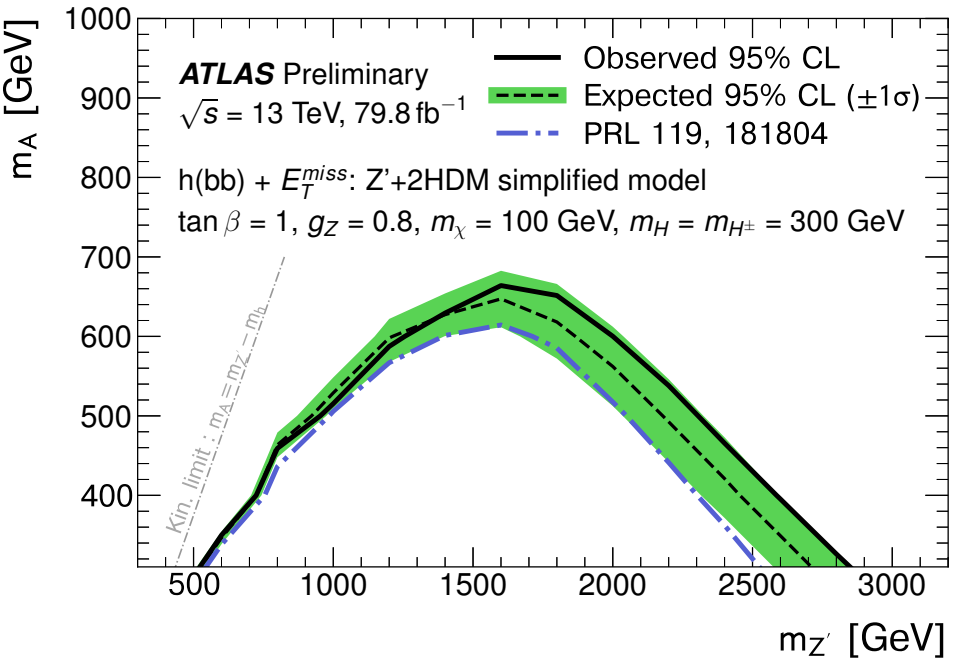


Figure 8.33 – Exclusion contours for the Z'-2HDM scenario in the $(m_{Z'}, m_A)$ plane for $\tan\beta = 1$, $g_{Z'} = 0.8$, and $m_\chi = 100$ GeV. The observed limits (solid line) are consistent with the expectation under the SM-only hypothesis (densely dashed line) within uncertainties (filled band). Observed limits from previous ATLAS results at $\sqrt{s} = 13\,\text{TeV}$ (dash-dotted line [3]) are also shown [5].

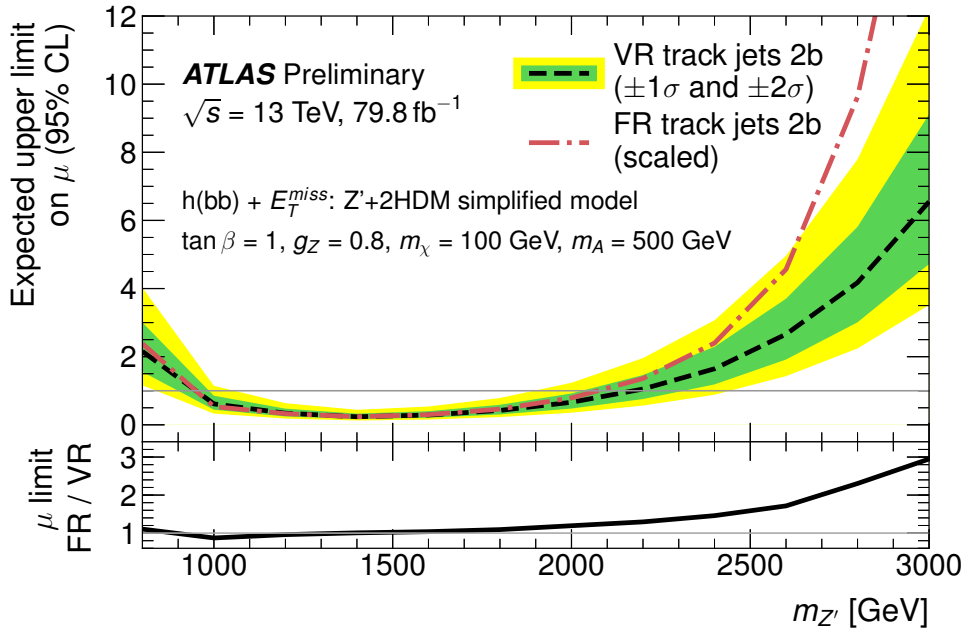


Figure 8.34 – Comparison of the expected upper limits on the signal strength μ for the analysis using variable-radius (VR) track-jets (dashed line) against the previous iteration of the analysis performed with fixed-radius (FR) track-jets (dash-dotted line) with two b-tagged jet and scaled to 79.8 fb^{-1}, for fixed $m_A = 500$ GeV and different values of $m_{Z'}$ of the Z'-2HDM benchmark model. Other differences between the two analyses include the suppression of the multijet background using the object-based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance, reduced uncertainties from the MC statistics, and the improve calibration of the b-tagging efficiency in the VR analysis. The lower panel is the ratio of the upper limits, showing a significant improvement in the high $m_{Z'}$ region [5].

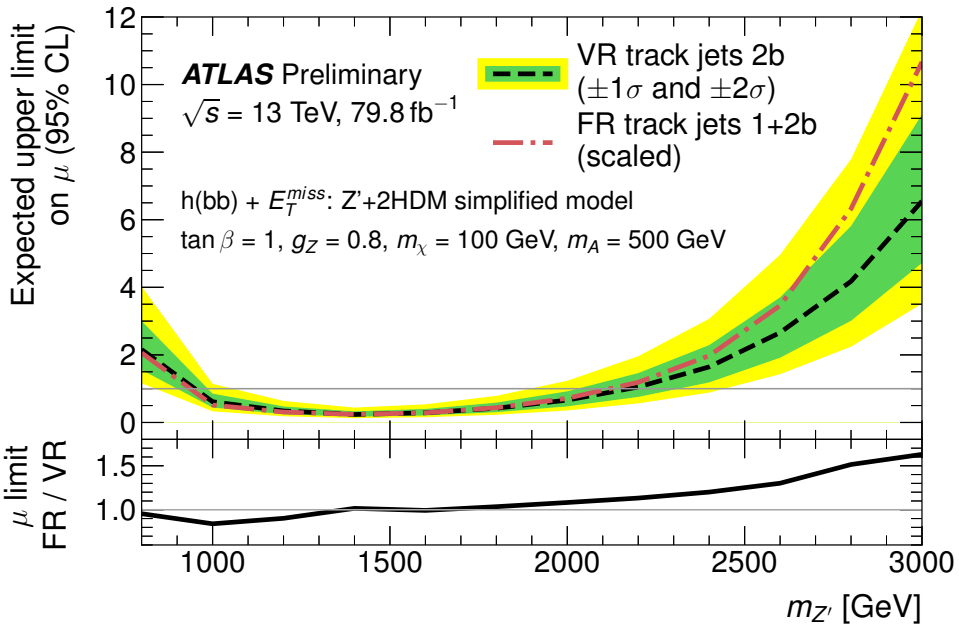


Figure 8.35 – Comparison of the expected upper limits on the signal strength μ for the analysis using variable-radius (VR) track-jets (dashed line) against the previous iteration of the analysis performed with fixed-radius (FR) track jets (dash-dotted line) with both one and two b-tagged signal regions (as described in [3]) and scaled to 79.8 fb^{-1}, for fixed $m_A = 500$ GeV and different values of $m_{Z'}$ of the Z'-2HDM benchmark model. The lower panel is the ratio of the upper limits from different track jets, showing a significant improvement from VR track jets in the high $m_{Z'}$ region [5].

Summary

The evidence for the existence of dark matter points to the current limits of the SM in describing the observed phenomenology in astro-particle physics and as a consequence to the need of a new physics extension beyond the SM. Also, the still recent discovery of a Higgs boson in Run 1, opens the possibility of searching for dark matter in the mono-h signature at the LHC. Indeed, many beyond the SM theories predict associated production of dark matter particles with a Higgs boson at energies that can be probed at collider experiments.

This thesis work aims at investigating a possible connection between dark matter and the Higgs boson by searching for dark matter pair production in association with a Higgs boson decaying into a pair of bottom quarks, referred as mono-$h(b\bar{b})$, using pp collisions collected at $\sqrt{13}$ TeV with the ATLAS detector at the LHC. Selected collision events comprise large missing transverse momentum and either two b-tagged small radius jets or a single large radius jet containing two b-tagged subjets. The results are interpreted in the framework of a simplified model with an exotic Z' gauge boson and two Higgs doublets, referred as $Z' - 2HDM$, where the dark matter χ is coupled to a heavy pseudoscalar A, $Z' \to hA \to b\bar{b}\chi\chi$, and both Z' and A are produced on-shell.

The first analysis presented is performed using data collected during 2015 and 2016 with an integrated luminosity of $36.1\,\text{fb}^{-1}$. The observed data are found to be consistent with the SM expectations in the signal region. Therefore, upper limits on the production cross-section of $pp \to h + \chi\chi$ times the $h \to b\bar{b}$ branching ratio, with $Br(h \to b\bar{b}) = 0.571$, are derived at 95% confidence level with the CL_s formalism using a profile likelihood fit. The excluded parameter space is presented in the 2-D mass space of $m_{Z'} - m_A$ with fixed parameters $\tan(\beta) = 1$, $g_{z'} = 0.8$, $m_\chi = 100$ GeV and $m_{H^\pm} = 300$ GeV, using a combined fit of signal and control regions that exploits the shape of the reconstructed m_{bb} spectrum to search for a resonant excess near m_h. Masses of the vector mediator $m_{Z'}$ are excluded up to 2.6 TeV for low masses of the pseudoscalar, excluding up to $m_A = 0.6$.

In searches for dark matter particles in the final state, the missing transverse momentum $\boldsymbol{E}_{\text{T}}^{\text{miss}}$ is a crucial quantity, since the presence of a pair of WIMP candidates would be signaled by an imbalance in the transverse momentum. Therefore, the study of the reconstruction and performance of this variable is critical in the sear for dark matter signatures as mono-$h(b\bar{b})$. In this work, a new object-based $\boldsymbol{E}_{\text{T}}^{\text{miss}}$ significance $\mathcal{S}$ variable was defined, which helps to separate events in which the reconstructed $\boldsymbol{E}_{\text{T}}^{\text{miss}}$ originates from weakly interacting particles, as the dark matter, from those in which $\boldsymbol{E}_{\text{T}}^{\text{miss}}$ is consistent with contributions coming from mis-measurements, resolutions and inefficiencies. On an event-by-event basis, $\mathcal{S}$ evaluates the p-value that the observed $\boldsymbol{E}_{\text{T}}^{\text{miss}}$ is consistent with the null hypothesis of zero real $\boldsymbol{E}_{\text{T}}^{\text{miss}}$, given the full event composition. A high value of $\mathcal{S}$

is an indication that the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ observed in the event is not well explained by resolution smearing alone, implying that the event is more likely to contain unseen objects such as neutrinos or more exotic weakly interacting particles as the WIMPs. This novel definition depends on the multiplicities, types, and kinematics of the objects measured in each event, providing an improved performance than the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ variable and it is now in use in several new physics searches in ATLAS collaboration.

An improved mono-$h(b\bar{b})$ analysis was presented in this work, exploiting two new reconstruction techniques: the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance and the Variable-Radius track jets, which impacts low and high regions of the $E_{\mathrm{T}}^{\mathrm{miss}}$, respectively. With a new requirement on the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance, the multijet background originated from pure strong interactions was successfully rejected to a negligible level and no estimation was needed from this background in the signal region. By selecting events with high values of the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance, $\mathcal{S} > 16$, more than 95% of dijet events can be rejected while retaining a signal efficiency higher than 90%.

The analysis also improved by adopting a new jet definition, the Variable-Radius track jets, to identify b-jets from highly-boosted Higgs bosons. The reconstruction of these subjets inside a large radius jet is based on a jet algorithm that uses inner detector tracks as inputs and a radius parameter that decreases as the subjet transverse momentum increases. This new reconstruction technique improves up to three times the expected upper limit for high boosted topologies related to high masses of the vector mediator Z'.

The improved analysis is performed using data collected during 2015, 2016 and 2017 with an integrated luminosity of $79.8\,\mathrm{fb}^{-1}$. The observed data are in agreement with the SM predictions and the results are interpreted in the same simplified model $Z' - 2HDM$. The masses of the Z' are excluded up to 2.8 TeV depending on the choices of the model parameters and the pseudoscalar mass is excluded up to 0.65 GeV.

Appendix A

The Inert Doublet Model

One of the best known and simple extensions of the standard model is the called Two Higgs Doublet Model [95], where there are two $SU(2)$ scalar doublets with weak hypercharge $Y = 1$ and give masses to the W and Z boson. Fermion masses are generated via Yukawa interactions, for which various models are considered: Model I, II, III, IV, ... [95]. In this models, five scalars appear: two charged and three neutral ones.

The IDM [56] addresses the DM issue and is a particular case of the Two Higgs Doublet Model. In this model, only one doublet H is involved in the spontaneous symmetry breaking. The second doublet, Φ, is inert (it has VEV=0) and contains four scalars which have limited interactions with the SM particles. Yukawa interactions are as in the Two Higgs Doublet Model I: the new doublet does not interact with fermions at tree level and in that way Z_2 symmetry is exact. Therefore, the neutral scalar, called H^0 (or A^0), of the inert doublet may be a viable DM candidate. The particle content is shown in Table A.1.

The scalar potential is Z_2 symmetric with respect to the transformation $H \to H$ and $\Phi \to -\Phi$ that forbids the inert doublet to develop VEV and all the inert particles in this doublet can only appear in pairs in their interaction vertices. Therefore, most general scalar potential that conserves Z_2 symmetry is [1]:

$$V = \mu_1^2 H^\dagger H + \mu_2^2 \Phi^\dagger \Phi + \lambda_1 \left(H^\dagger H\right)^2 + \lambda_2 \left(\Phi^\dagger \Phi\right)^2 + \lambda_3 \left(H^\dagger H\right)\left(\Phi^\dagger \Phi\right) + \lambda_4 \left(H^\dagger \Phi\right)\left(\Phi^\dagger H\right) + \left\{\frac{\lambda_5}{2}\left(H^\dagger \Phi\right)^2 + \text{h.c}\right\}, \tag{A.1}$$

where H is the SM Higgs doublet and Φ is the inert doublet. All the couplings in equation (A.1) are assumed to be real [2]. The electroweak gauge symmetry is broken when H gets its VEV, $\langle H^T \rangle = \left(0, v/\sqrt{2}\right)$ while $\langle \Phi^T \rangle = 0$. This spontaneous symmetry breaking results in a unbroken Z_2 symmetry and in a CP-even neutral scalar H^0, one CP-odd neutral scalar A^0, a pair of charged scalars $H^\pm$ in addition to the SM CP-even scalar Higgs h. These

[1] The scalar potential with Z_2 symmetry forbids the mass term $-\mu\left(H^\dagger \Phi + \text{h.c.}\right)$ which mixes H and Φ.

[2] All the parameters are necessarily real except λ_5, however we can rotate the relative phase between H and Φ to make λ_5 real.

scalar doublets can be parameterized as:

$$\Phi = \begin{pmatrix} H^+ \\ \frac{1}{\sqrt{2}}\left(H^0 + iA^0\right) \end{pmatrix} , \tag{A.2}$$

$$H = \begin{pmatrix} G^+ \\ \frac{1}{\sqrt{2}}\left(h + v + iG^0\right) \end{pmatrix} , \tag{A.3}$$

where $G^\pm$ and G^0 are the Goldstone bosons, which will be absorbed by the $W^\pm$ and Z to acquire their masses.

This potential has 8 real parameters: λ_i, μ_j and the VEV v, with $i = 1, 2, ...5$ and $j = 1, 2$. Minimization condition eliminates μ_1^2 due to the Higgs mass and the VEV is already known ($v = 246$ GeV), so there are 6 remain free parameters: μ_2^2, λ_1, λ_2, λ_3, λ_4, and λ_5.

After spontaneous symmetry breaking, the masses of the physical scalars can be written as:

$$m_h^2 = -2\mu_1^2 = 2\lambda_1 v^2 , \tag{A.4}$$

$$m_{H^\pm}^2 = \mu_2^2 + \frac{1}{2}\lambda_3 v^2 , \tag{A.5}$$

$$m_{H^0}^2 = \mu_2^2 + \frac{1}{2}\left(\lambda_3 + \lambda_4 + \lambda_5\right) v^2 = \mu_2^2 + \lambda_L v^2 , \tag{A.6}$$

$$m_{A^0}^2 = \mu_2^2 + \frac{1}{2}\left(\lambda_3 + \lambda_4 - \lambda_5\right) v^2 , \tag{A.7}$$

where $\lambda_L = 1/2\left(\lambda_3 + \lambda_4 + \lambda_5\right)$. The Lightest Odd Particle (LOP) can be either H^0 or A^0 and will play the role of DM candidate, if $\lambda_5 < 0$ the LOP in the IDM is H^0. λ_L is related to the cubic and quartic coupling between the SM-Higgs h and the DM candidate H^0. λ_2 gives the quartic DM self-coupling, while λ_3 describes the Higgs particle interaction with charged scalars.

The parameters on equation (A.1) have different restrictions. The perturbative condition require all quartic couplings to obey $|\lambda_i| < 8\pi$. Also, the potential must be bounded from below and the corresponding vacuum stability conditions are given as [95]:

$$\lambda_1 > 0 \, , \, \lambda_2 > 0 \, , \tag{A.8}$$

$$\lambda_3 + \lambda_4 - |\lambda_5| > \sqrt{\lambda_1 \lambda_2}, \;\; \lambda_3 + 2\sqrt{\lambda_1 \lambda_2} > 0 \, . \tag{A.9}$$

Four of the five λ coupling can be written in terms of the physical masses and μ_2^2 and therefore we are free to take these 6 free independent parameters [274]:

$$m_h, m_{H^0}, m_{A^0}, m_{H^\pm}, \lambda_2, \lambda_L \tag{A.10}$$

There are two regions of m_{H^0} in which the relic density is in agreement with the data [55]:

- Medium mass regime of about $50-90$ GeV.

- High mass regime of $m_{H^0} > 500$ GeV. In this region all Z_2-odd particles have almost degenerate masses and coannihilation processes are crucial to have the correct measured relic density.

The relic density data constrain the λ_L parameter for chosen values of m_{H^0} but the λ_2 parameter is no limited. If the mass splitting between the LOP H^0 and the next-lightest odd particle A^0 is, approximately, < 10 GeV, thus the number densities of A^0 have only a slight Boltzmann suppression respect to the LOP number density. Therefore, the contributions to the relic density from the scattering of $H^0 - A^0$ and $A^0 - A^0$ have to be taken into account in order to have a more precise relic abundance prediction (1.7), these are called coannihilation [275].

The Higgs particle provides a portal between the inert dark sector and the Z_2-even SM sector. Thus the SM Higgs may decay into a pair of DM, if the kinematics allows it, and this will contribute to the invisible Higgs width which is now constrained by the LHC. Also, annihilation of the DM into SM particles will provide thermal relic density and the scattering of the DM particle onto nucleons will lead to direct detection signatures.

Within IDM it is possible to obtain the correct DM relic density and simultaneously be consistent with the current limits on direct detection [274]. Therefore, the IDM could be considered as a simple but appealing model in the framework of Weakly Interacting Massive Particle -WIMP-.

	L_i	e^i_R	H	Φ
$SU(2)_L$	2	1	2	2
$U(1)_Y$	-1	-2	1	1
Z_2	+	+	+	-

Table A.1 – Fermions and scalar fields with their charges under $SU(2)_L \otimes U(1)_Y \otimes Z_2$ in the IDM

Appendix B

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

B.1 Constant soft term resolution estimation in the $\mu\mu$ and ee-channels

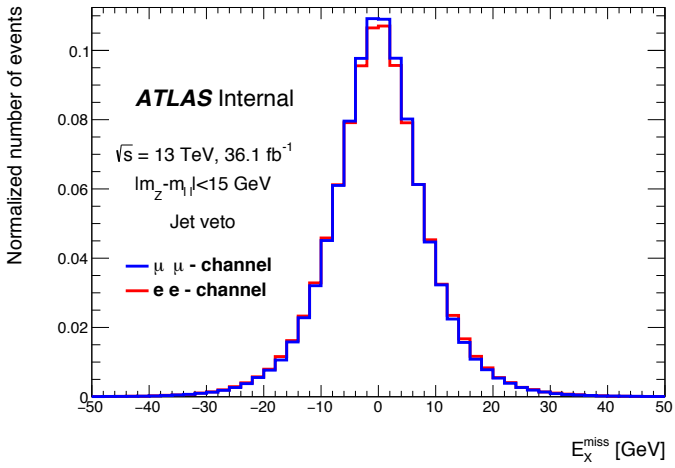


Figure B.1 – One of the components of the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$, E_x^{miss}, in the $Z \to \mu\mu$ sample with a jet veto selection. The RMS of the distribution is a first order estimation of the soft term resolution since no genuine $E_{\mathrm{T}}^{\mathrm{miss}}$ is expected for this region.

B.2 data-MC comparisons in the $\mu\mu$-channel

Muons resolutions are currently outdated and not appropriate in several regions (mainly for CB muons with eta>|1.9|).

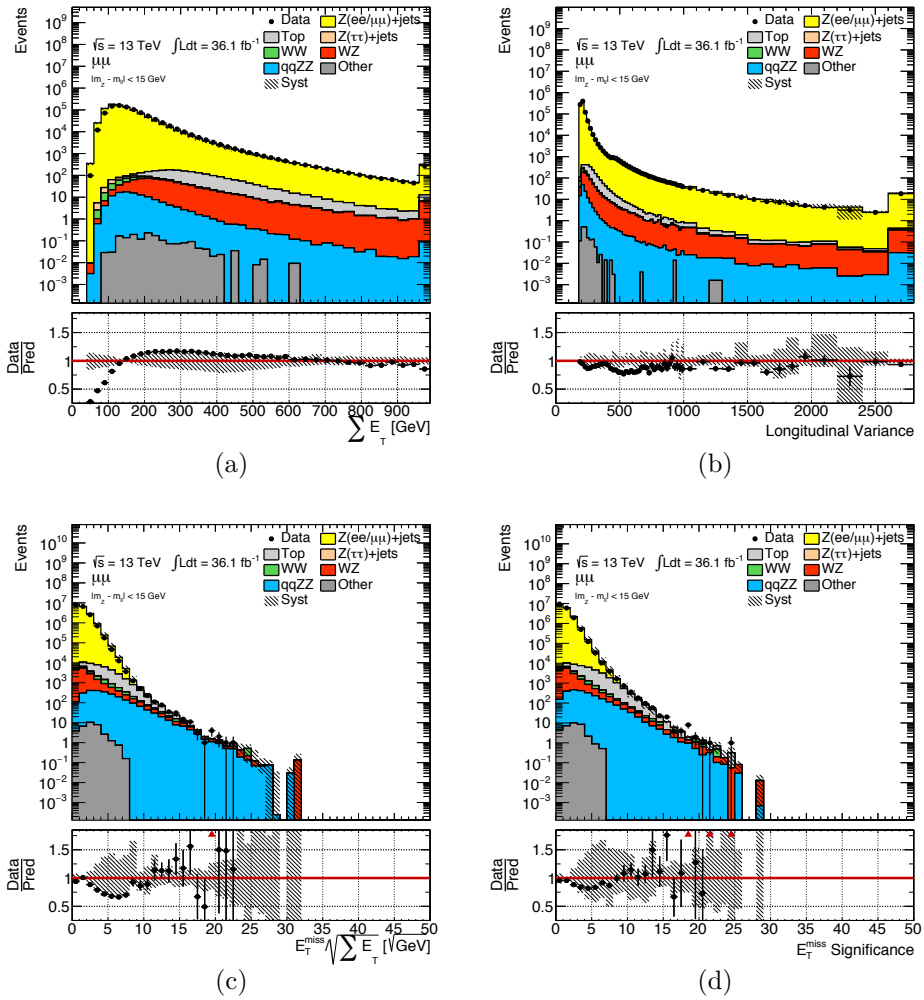


Figure B.2 – Distributions in data compared to MC simulations including all relevant backgrounds for events satisfying the $Z \to \mu\mu$ selection for: (a) $\sum E_{\mathrm{T}}^{\mathrm{miss}}$, (b) longitudinal variance σ_L^2, (c) event based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ Significance $E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{\sum E_{\mathrm{T}}^{\mathrm{miss}}}$ and (d) object based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ Significance. The respective ratios between data and MC simulations are shown below the distributions, with the shaded bands which correspond to the combined systematic and MC statistical uncertainties.

B.3 Signal and background distributions in the ee-channel for the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance

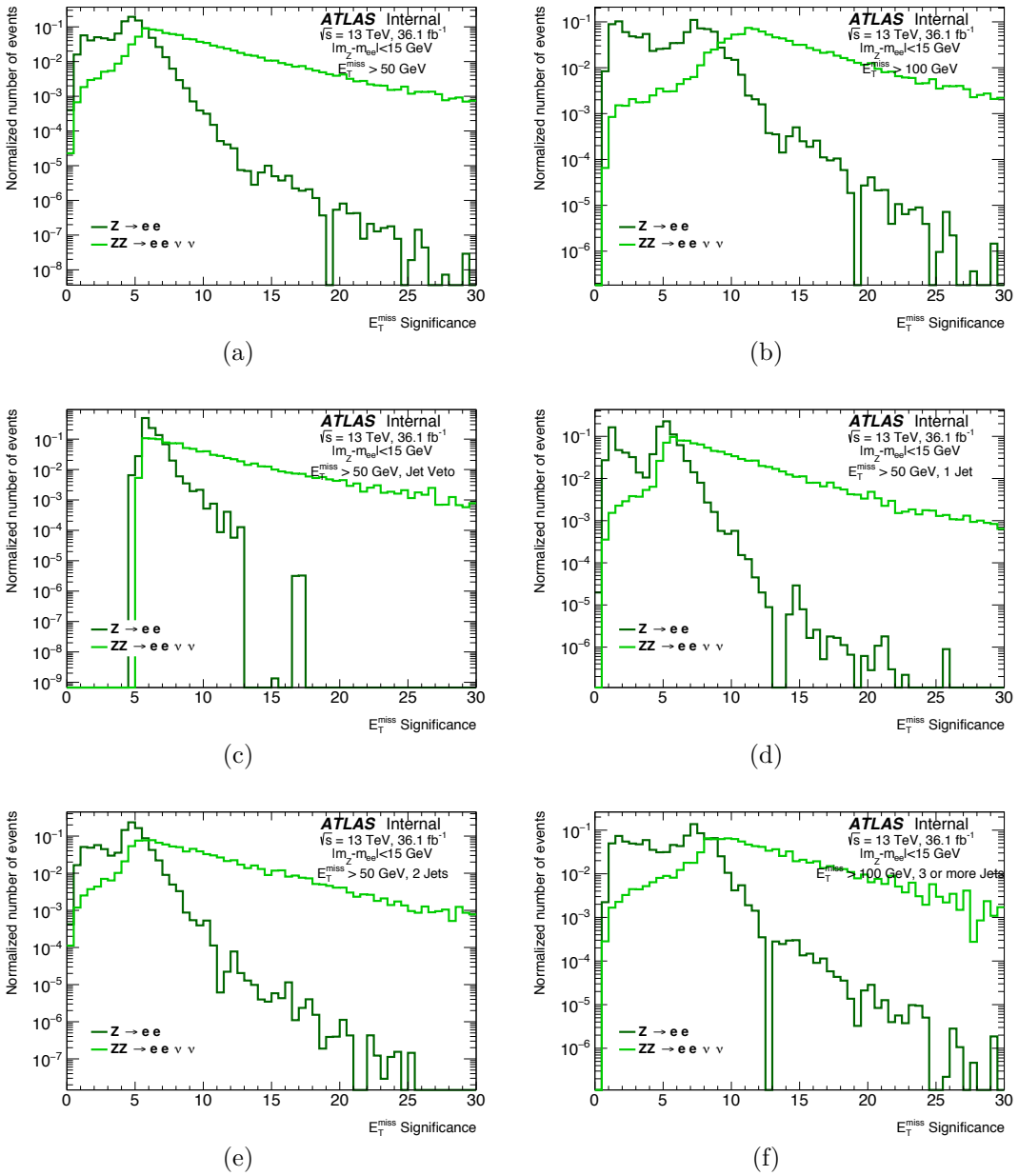
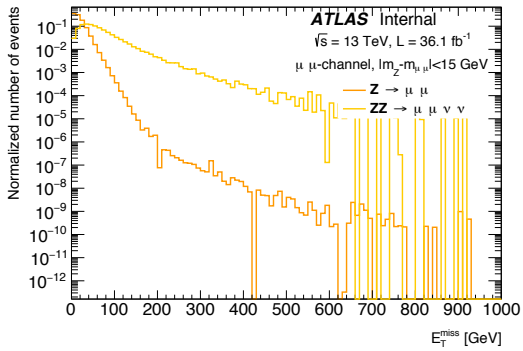
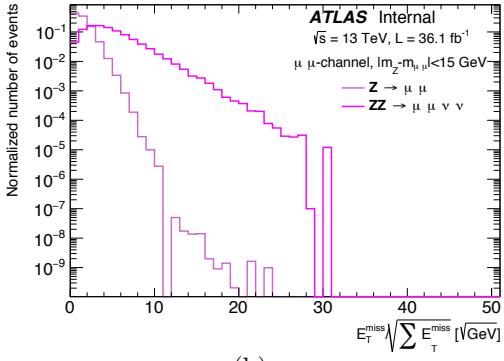


Figure B.3 – MC simulated distributions for $Z \to ee$ and $ZZ \to ee\nu\nu$ samples with a $Z \to ee$ selection, which corresponds to background and signal respectively in the performance study for the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance.

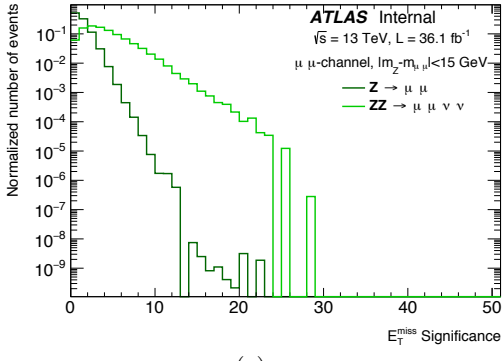
B.4 Signal and background distributions in the $\mu\mu$-channel



(a)



(b)

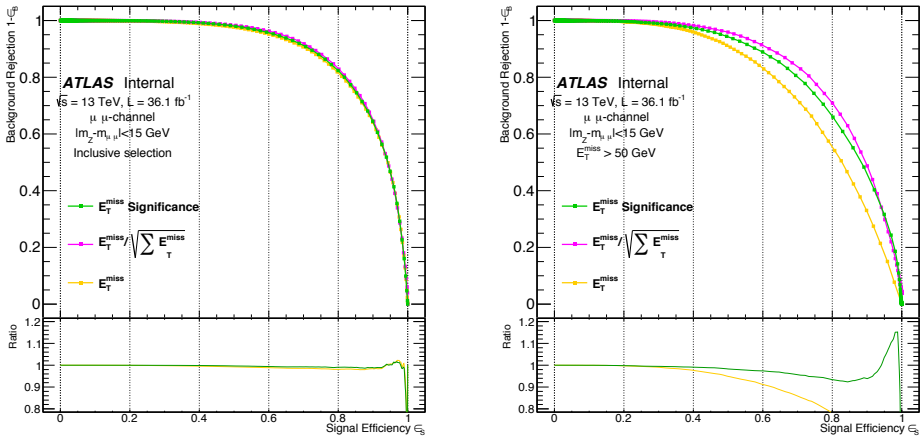


(c)

Figure B.4 – MC simulated distributions for $Z \to \mu\mu$ and $ZZ \to \mu\mu\nu\nu$ samples with a $Z \to \mu\mu$ selection, which corresponds to background and signal respectively in the performance study. The upper plot is the (a) $E_{\mathrm{T}}^{\mathrm{miss}}$, the lower plots corresponds to (b) the event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance $E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{\sum E_{\mathrm{T}}}$ (Equation 5.1), and (d) the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance (Equation 5.10).

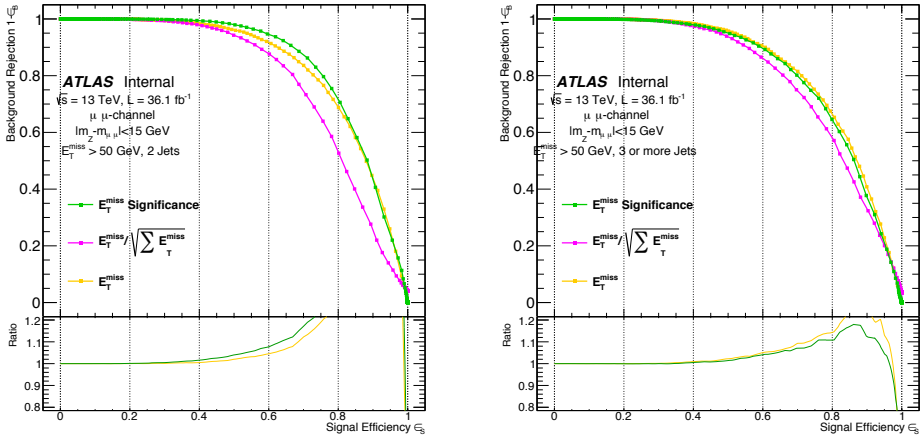
B.5 Signal efficiency versus background rejection in the $\mu\mu$-channel

Muons resolutions are currently outdated and not appropriate in several regions (mainly for CB muons with eta>|1.9|).



(a) (b)

Figure B.5 – Background rejection versus signal efficiency in simulated $Z \to \mu\mu$ and $ZZ \to \mu\mu\nu\nu$ samples with a $Z \to \mu\mu$ selection. The performance is shown for $E_{\mathrm{T}}^{\mathrm{miss}}$, $E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{\sum E_{\mathrm{T}}}$ and object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance as discriminants for (a) all the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ range, and (b) $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV. The lower panel of the figures shows the ratio of other definitions/event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance.



(a) (b)

Figure B.6 – Background rejection versus signal efficiency in simulated $Z \to \mu\mu$ and $ZZ \to \mu\mu\nu\nu$ samples with a $Z \to \mu\mu$ selection and $E_{\mathrm{T}}^{\mathrm{miss}} > 50$GeV. The performance is shown for $E_{\mathrm{T}}^{\mathrm{miss}}$, $E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{\sum E_{\mathrm{T}}}$ and object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance as discriminants with (a)two jets and (b) three or more jets. The lower panel of the figures shows the ratio of other definitions/event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance.

Appendix C

Data and Monte Carlo Simulation Samples

C.1 Data Samples

Only runs recorded with stable beam and optimal conditions of the ATLAS detector are used. The runs and luminosity blocks whose high quality has been studied by the Data Quality ATLAS group, considering the performance of all su-detectors, beam conditions and object reconstruction performance are listed in the *Good-Run list* (GRL).

For data recorded during 2015 and 2016, the GRL used in the analysis in Chapter 7 are:

- 2015 data: data15_13TeV.periodAllYear_DetStatus-v79-repro20-02_DQDefects-00-02-02_PHYS_StandardGRL_All_Good_25ns.xml
- 2016 data: data16_13TeV.periodAllYear_DetStatus-v83-pro20-15_DQDefects-00-02-04_PHYS_StandardGRL_All_Good_25ns.xml

Events in the 0 and 1 lepton regions are selected using $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers. Events in the 2 lepton region are selected with unprescaled single lepton triggers. The full list of the triggers used in all channels can be found in Table C.1 for the analysis in Chapter 7. [1]

For the search described in Chapter 8, the following GRL are used:

- `data15_13TeV.periodAllYear_DetStatus-v89-pro21-02_Unknown_PHYS_StandardGRL_All_Good_25ns.xml`
- `data16_13TeV.periodAllYear_DetStatus-v89-pro21-01_DQDefects-00-02-04_PHYS_StandardGRL_All_Good_25ns.xml`

[1] MC does not contain all the triggers used in data taking

Period	0 lepton	1 lepton	2 lepton
2015	HLT_xe70	HLT_xe70	HLT_e24_lhmedium_L1EM18VH (MC)[1] **OR** HLT_e24_lhmedium_L1EM20VH (data)[1] **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)[1] **OR** HLT_mu24_iloose (data)[1] **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 (F1)	HLT_xe110_mht_L1XE50	HLT_xe110_mht_L1XE50	HLT_mu26_ivarmedium
2016 (F2-)	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 C.1 – Summary table of triggers used in 2015 and 2016 data.

- `data17_13TeV.periodAllYear_DetStatus-v97-pro21-13_Unknown_PHYS_StandardGRL_All_Good_25ns_Triggerno17e33prim.xml`

The full list of the triggers used in all channels in the search described in Chapter 8 can be found in Table C.2. The trigger thresholds are determined by the lowest unprescaled triggers defined for the Run 2 luminosity according to the trigger menu which has been updated for 2017 data taking [276]. The trigger HLT_xe70 in Table C.1 was replaced by HLT_xe70_mht for the 2015 runs, as the latter one is more efficient.

C.2 Signal Monte Carlo Samples

The cross sections for the $Z' - 2HDM$ simplified model signals used in the search of Dark Matter in association with H→bb using Using $36\,fb^{-1}$ with data from 2015 and 2016 at $\sqrt{s} = 13$ TeV are in C.3.

DSID	Sample Name	XS [pb]
304102	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp600_mA300	0.0024148
304103	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp600_mA400	0.00054273
304104	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp800_mA300	0.0015131
304105	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp800_mA400	0.0008099
304106	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp800_mA500	0.00017132
304107	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp800_mA600	2.5334e-05
304108	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp1000_mA300	0.00078839
304109	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp1000_mA400	0.0005233
304110	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp1000_mA500	0.00016296
304111	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp1000_mA600	4.5002e-05
304112	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp1000_mA700	1.1755e-05
304113	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp1000_mA800	3.2736e-06
304114	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp1200_mA300	0.00041488
304115	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp1200_mA400	0.00030282
304116	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp1200_mA500	0.00011017
304117	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp1200_mA600	3.8347e-05
304118	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp1200_mA700	1.343e-05
304119	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp1200_mA800	4.594e-06
304120	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp1400_mA300	0.0002265
304121	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp1400_mA400	0.00017377
304122	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp1400_mA500	6.8921e-05
304123	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp1400_mA600	2.6999e-05
304124	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp1400_mA700	1.1072e-05
304125	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp1400_mA800	4.5057e-06
305202	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp1600_mA300	0.00012809

305203	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp1600_mA400	0.00010136
305204	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp1600_mA500	4.2387e-05
305205	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp1600_mA600	1.7876e-05
305206	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp1600_mA700	8.0716e-06
305207	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp1600_mA800	3.6847e-06
305208	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp1800_mA300	7.4755e-05
305209	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp1800_mA400	6.0409e-05
305210	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp1800_mA500	2.617e-05
305211	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp1800_mA600	1.1607e-05
305212	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp1800_mA700	5.5902e-06
305213	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp1800_mA800	2.7635e-06
305214	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp2000_mA300	4.4844e-05
305215	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp2000_mA400	3.6754e-05
305216	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp2000_mA500	1.6317e-05
305217	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp2000_mA600	7.5135e-06
305218	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp2000_mA700	3.7894e-06
305219	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp2000_mA800	1.9847e-06
305220	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp2200_mA300	2.7502e-05
305221	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp2200_mA400	2.2776e-05
305222	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp2200_mA500	1.0302e-05
305223	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp2200_mA600	4.8822e-06
305224	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp2200_mA700	2.5494e-06
305225	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp2200_mA800	1.3947e-06
305226	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp2400_mA300	1.7185e-05
305227	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp2400_mA400	1.4344e-05
305228	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp2400_mA500	6.5765e-06
305229	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp2400_mA600	3.1912e-06
305230	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp2400_mA700	1.7121e-06
305231	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp2400_mA800	9.6915e-07
305822	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp200_mA200	1.9351e-06
305823	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp200_mA300	4.6802e-07
305824	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp200_mA400	1.5668e-07
305825	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp400_mA200	0.0082302
305826	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp400_mA300	2.3341e-05
305827	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp400_mA400	2.4673e-06
305828	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp400_mA500	6.8447e-07
305829	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp600_mA200	0.0056062
305830	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp600_mA500	1.7191e-05
305831	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp600_mA600	3.9803e-06
305832	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp800_mA200	0.0024483
305833	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp800_mA700	4.8377e-06
305834	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp2400_mA200	2.1232e-05

305835	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp2600_mA200	1.3417e-05
305836	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp2600_mA300	1.0912e-05
305837	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp2600_mA400	9.1631e-06
305838	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp2600_mA500	4.242e-06
305839	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp2600_mA600	2.099e-06
305840	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp2600_mA700	1.1515e-06
305841	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp2600_mA800	6.6995e-07
305842	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp2800_mA200	8.5958e-06
305843	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp2800_mA300	7.0163e-06
305844	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp2800_mA400	5.9199e-06
305845	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp2800_mA500	2.764e-06
305846	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp2800_mA600	1.3891e-06
305847	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp3000_mA200	5.5703e-06
305848	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp3000_mA300	4.5589e-06
305849	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp3000_mA400	3.861e-06
305850	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp3000_mA500	1.8158e-06
305851	MadGraphPythia8EvtGen_A14NNPDF23LO_zp2hdm_bb_mzp3000_mA600	9.2473e-07

Table C.3 – Cross sections for the Z'-2HDM models used as benchmark signals.

C.3 Signal Acceptance Uncertainties

Period	0 lepton	1 lepton	2 lepton
2015	HLT_xe70_mht	HLT_xe70_mht	HLT_e24_lhmedium_L1EM20VH **OR** HLT_e60_lhmedium **OR** HLT_e120_lhloose **OR** HLT_mu20_iloose_L1MU15 **OR** HLT_mu50 **OR** HLT_mu60_0eta105_msonly
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 **OR** HLT_mu24_ivarloose **OR** HLT_mu40 **OR** HLT_mu50
2016 (B-D3)	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_ivarmedium **OR** HLT_mu24_imedium **OR** HLT_mu50
2016 (D4-E3)	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_mu24_ivarmedium **OR** HLT_mu24_imedium **OR** HLT_mu26_ivarmedium **OR** HLT_mu26_imedium **OR** HLT_mu50
2016 (F1)	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_mu26_imedium **OR** HLT_mu50
2016 (F2-)	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_mu26_imedium **OR** HLT_mu50
2017 (period B onwards)	HLT_xe110_pufit_L1XE55	HLT_xe110_pufit_L1XE55	HLT_e26_lhtight_nod0_ivarloose **OR** HLT_e60_lhmedium_nod0 **OR** HLT_e140_lhloose_nod0 **OR** HLT_e300_etcut **OR** HLT_mu26_ivarmedium **OR** HLT_mu50 **OR** HLT_mu60_0eta105_msonly

Table C.2 – Summary table of triggers used in 2015, 2016 and 2017 data.

Uncertainty on Mono-H Signals due to scales (%)				
2HDM Model	$E_{\mathrm{T}}^{\mathrm{miss}}$ GeV			
m(prop)_m(A)	150–200	200–350	350–500	>500
mzp400_mA200	0.881	1.93	-	-
mzp600_mA200	0.791	0.277	-	-
mzp600_mA300	0.333	1.14	1.22	8
mzp600_mA400	1.18	1.98	8.99	8.37
mzp600_mA500	0.435	2.06	-	-
mzp600_mA600	0.289	0.955	-	-
mzp800_mA200	2.23	0.860	0.944	-
mzp800_mA300	2.09	0.313	1.31	4.8
mzp800_mA400	1.48	0.46	9.01	3.7
mzp800_mA500	0.475	0.474	7.18	23.2
mzp800_mA600	0.576	0.471	1.02	10.9
mzp800_mA700	1.23	0.465	8.62	-
mzp1000_mA300	3.48	0.382	0.317	1.44
mzp1000_mA400	5.78	0.469	0.519	2.08
mzp1000_mA500	1.2	0.267	0.476	1.38
mzp1000_mA600	2.68	0.451	0.576	0.482
mzp1000_mA700	0.356	0.603	0.747	10.3
mzp1000_mA800	0.595	0.726	2.04	4.21
mzp1200_mA300	8.84	3.8	1.27	0.685
mzp1200_mA400	15.6	2.4	0.382	0.846
mzp1200_mA500	4.76	0.653	0.291	1.62
mzp1200_mA600	2.94	0.971	0.628	2.15
mzp1200_mA700	1.14	0.798	1.14	1.62
mzp1200_mA800	0.988	0.608	1.03	1.08
mzp1400_mA300	3.81	2.84	1.2	0.318
mzp1400_mA400	2.49	2.53	0.145	0.353
mzp1400_mA500	6.8	3.46	0.557	0.326
mzp1400_mA600	4.64	0.884	0.854	0.378
mzp1400_mA700	3.43	1.01	0.454	0.648
mzp1400_mA800	0.196	0.918	0.671	0.949
mzp1600_mA300	-	-	2.06	0.153
mzp1600_mA400	-	8.54	3.72	0.235
mzp1600_mA500	-	4.16	4.91	0.271
mzp1600_mA600	-	3.25	1.93	0.518
mzp1600_mA700	-	1.54	2.33	0.172
mzp1600_mA800	-	1.30	1.74	0.292

Table C.4 – The mono-H signal uncertainties due to variation of the scale for the 2HDM model in intervals of $E_{\mathrm{T}}^{\mathrm{miss}}$ for values of the mass of the $Z' \leq 1600$ GeV. In the case of the acceptance being less than 1% the errors are neglected and denoted by "-".

Uncertainty on Mono-H Signals due to scales (%)				
2HDM Model	$E_{\mathrm{T}}^{\mathrm{miss}}$ GeV			
m(prop)_m(A)	150–200	200–350	350–500	>500
mzp1800_mA300	-	-	-	0.132
mzp1800_mA400	-	-	-	0.444
mzp1800_mA500	-	-	4.69	0.175
mzp1800_mA600	-	2.20	5.57	0.176
mzp1800_mA700	-	1.97	5.04	0.124
mzp1800_mA800	-	2.00	3.63	0.169
mzp2000_mA300	-	-	-	0.071
mzp2000_mA400	-	-	-	0.278
mzp2000_mA500	-	-	-	0.171
mzp2000_mA600	-	-	-	0.0856
mzp2000_mA700	-	-	-	0.118
mzp2000_mA800	-	-	-	0.412
mzp2200_mA300	-	-	-	0.103
mzp2200_mA400	-	-	-	0.309
mzp2200_mA500	-	-	-	0.0830
mzp2200_mA600	-	-	-	0.375
mzp2200_mA700	-	-	-	0.169
mzp2200_mA800	-	-	-	0.236
mzp2400_mA200	-	-	-	0.146
mzp2400_mA300	-	-	-	0.0126
mzp2400_mA400	-	-	-	0.202
mzp2400_mA500	-	-	-	0.234
mzp2400_mA600	-	-	-	0.205
mzp2400_mA700	-	-	-	0.127
mzp2400_mA800	-	-	-	0.129
mzp2600_mA200	-	-	-	0.158
mzp2600_mA300	-	-	-	0.189
mzp2600_mA400	-	-	-	0.113
mzp2600_mA500	-	-	-	0.103
mzp2600_mA600	-	-	-	0.0516
mzp2600_mA700	-	-	-	0.248
mzp2600_mA800	-	-	-	0.0588
mzp2800_mA200	-	-	-	0.228
mzp2800_mA300	-	-	-	0.303
mzp2800_mA400	-	-	-	0.108
mzp2800_mA500	-	-	-	0.139
mzp2800_mA600	-	-	-	0.255
mzp3000_mA200	-	-	-	0.144
mzp3000_mA300	-	-	-	0.370
mzp3000_mA400	-	-	-	0.212
mzp3000_mA500	-	-	-	0.164
mzp3000_mA600	-	-	-	0.224

Table C.5 – The mono-H signal uncertainties due to variation of the scale for the 2HDM model in intervals of $E_{\mathrm{T}}^{\mathrm{miss}}$ for values of the mass of the $Z' \geq 1800$ GeV. In the case of the acceptance being less than 1% the errors are neglected and denoted by "-".

Uncertainty on Mono-H Signals due to shower tunes (%)				
2HDM Model	$E_{\mathrm{T}}^{\mathrm{miss}}$ GeV			
m(prop)_m(A)	150–200	200–350	350–500	>500
mzp400_mA200	9.24	10.7	-	-
mzp600_mA200	3.51	1.33	-	-
mzp600_mA300	2.77	4.64	8.99	53.3
mzp600_mA400	6.17	7.97	24.9	25.3
mzp600_mA500	1.61	4.80	-	-
mzp600_mA600	2.40	3.37	-	-
mzp800_mA200	6.74	4.46	6.29	-
mzp800_mA300	6.51	1.31	10.1	17.1
mzp800_mA400	5.43	1.21	16	25.4
mzp800_mA500	2.6	1.04	12.3	55.8
mzp800_mA600	2.9	1.59	12.1	16.6
mzp800_mA700	3.69	2.79	13.5	-
mzp1000_mA300	10.4	3.71	1.81	7.57
mzp1000_mA400	10.9	2.39	0.61	10.6
mzp1000_mA500	7.64	1.66	1.57	16.5
mzp1000_mA600	4.82	1.2	3.08	13.7
mzp1000_mA700	4.21	1.53	5.03	16
mzp1000_mA800	2.63	1.24	2.12	11.7
mzp1200_mA300	16.3	8.46	3.7	1.53
mzp1200_mA400	15.7	6.33	2.28	1.91
mzp1200_mA500	7.75	4.95	1.62	5.5
mzp1200_mA600	10.1	2.88	1.57	8.25
mzp1200_mA700	6.35	2.7	1.15	5.64
mzp1200_mA800	4.48	1.87	2.19	6.11
mzp1400_mA300	18.3	8.79	4.76	1.3
mzp1400_mA400	18.2	7.52	4.73	1.49
mzp1400_mA500	26.8	9.24	3.41	1.44
mzp1400_mA600	15.5	4.78	2.74	1.75
mzp1400_mA700	9.82	5.35	2.1	2.24
mzp1400_mA800	4.74	3.48	1.71	2.55
mzp1600_mA300	-	-	8.94	1.08
mzp1600_mA400	-	10.3	7.20	1.48
mzp1600_mA500	-	6.96	5.54	1.30
mzp1600_mA600	-	6.21	7.28	1.35
mzp1600_mA700	-	6.25	5.44	1.44
mzp1600_mA800	-	4.70	5.37	1.22

Table C.6 – The mono-H signal uncertainties due to variation of the shower tune for the 2HDM model in intervals of $E_{\mathrm{T}}^{\mathrm{miss}}$ for values of the mass of the $Z' \leq 1600$ GeV. In the case of the acceptance being less than 1% the errors are neglected and denoted by "-".

Uncertainty on Mono-H Signals due to shower tunes (%)				
2HDM Model	$E_{\mathrm{T}}^{\mathrm{miss}}$ GeV			
m(prop)_m(A)	150–200	200–350	350–500	>500
mzp1800_mA300	-	-	-	1.18
mzp1800_mA400	-	-	-	1.08
mzp1800_mA500	-	-	9.32	1.14
mzp1800_mA600	-	11.8	7.88	1.33
mzp1800_mA700	-	8.51	9.08	1.30
mzp1800_mA800	-	6.07	7.89	1.36
mzp2000_mA300	-	-	-	1.04
mzp2000_mA400	-	-	-	1.02
mzp2000_mA500	-	-	-	1.14
mzp2000_mA600	-	-	-	0.922
mzp2000_mA700	-	-	-	0.997
mzp2000_mA800	-	-	-	1.03
mzp2200_mA300	-	-	-	1.00
mzp2200_mA400	-	-	-	0.868
mzp2200_mA500	-	-	-	0.776
mzp2200_mA600	-	-	-	0.993
mzp2200_mA700	-	-	-	0.713
mzp2200_mA800	-	-	-	1.27
mzp2400_mA200	-	-	-	1.08
mzp2400_mA300	-	-	-	1.00
mzp2400_mA400	-	-	-	1.24
mzp2400_mA500	-	-	-	0.978
mzp2400_mA600	-	-	-	1.16
mzp2400_mA700	-	-	-	1.11
mzp2400_mA800	-	-	-	0.941
mzp2600_mA200	-	-	-	1.20
mzp2600_mA300	-	-	-	1.13
mzp2600_mA400	-	-	-	1.24
mzp2600_mA500	-	-	-	1.14
mzp2600_mA600	-	-	-	1.05
mzp2600_mA700	-	-	-	1.12
mzp2600_mA800	-	-	-	0.949
mzp2800_mA200	-	-	-	1.06
mzp2800_mA300	-	-	-	0.961
mzp2800_mA400	-	-	-	1.20
mzp2800_mA500	-	-	-	1.05
mzp2800_mA600	-	-	-	1.08
mzp3000_mA200	-	-	-	1.04
mzp3000_mA300	-	-	-	0.94
mzp3000_mA400	-	-	-	1.30
mzp3000_mA500	-	-	-	1.02
mzp3000_mA600	-	-	-	1.03

Table C.7 – The mono-H signal uncertainties due to variation of the shower tune for the 2HDM model in intervals of $E_{\mathrm{T}}^{\mathrm{miss}}$ for values of the mass of the $Z' \geq 1800$ GeV. In the case of the acceptance being less than 1% the errors are neglected and denoted by "-".

Uncertainty on Mono-H Signals due to PDF (%)				
2HDM Model	$E_{\mathrm{T}}^{\mathrm{miss}}$ GeV			
m(prop)_m(A)	150–200	200–350	350–500	>500
mzp400_mA200	1.84	0.639	-	-
mzp600_mA200	1.08	0.194	-	-
mzp600_mA300	0.0317	1.82	7.23	76
mzp600_mA400	3.04	2.79	16	6.9
mzp600_mA500	1.50	1.25	-	-
mzp600_mA600	1.51	1.16	-	-
mzp800_mA200	2.25	1.14	2.97	-
mzp800_mA300	0.949	0.842	0.809	3.52
mzp800_mA400	1.24	0.43	11.4	7.41
mzp800_mA500	1.51	0.608	3.28	44.7
mzp800_mA600	1.25	0.0292	7.22	16.8
mzp800_mA700	0.904	0.468	8.44	-
mzp1000_mA300	5.54	0.93	1.11	3.96
mzp1000_mA400	6.57	0.975	0.851	4.72
mzp1000_mA500	2.25	0.354	1.38	5.19
mzp1000_mA600	3.01	0.0691	0.145	2.14
mzp1000_mA700	1.98	0.53	1.79	16.5
mzp1000_mA800	0	0	0	0
mzp1200_mA300	8.42	5.34	2.12	0.488
mzp1200_mA400	20.1	3.24	1.38	2.15
mzp1200_mA500	5.29	1.42	1.06	3.46
mzp1200_mA600	5.73	0.946	0.841	4.42
mzp1200_mA700	2.16	0.58	0.648	1.88
mzp1200_mA800	4.11	0.556	1.3	2.61
mzp1400_mA300	3.57	4.23	0.236	0.0739
mzp1400_mA400	14.1	3.36	1.14	0.629
mzp1400_mA500	11.7	2.18	0.702	0.741
mzp1400_mA600	4.27	3.54	0.87	1.13
mzp1400_mA700	4.34	3.03	0.729	1.65
mzp1400_mA800	1.27	0.714	0.082	1.5
mzp1600_mA300	-	-	1.04	0.390
mzp1600_mA400	-	8.39	2.40	0.303
mzp1600_mA500	-	0.681	4.61	0.630
mzp1600_mA600	-	5.94	3.62	0.519
mzp1600_mA700	-	1.23	5.24	0.597
mzp1600_mA800	-	4.47	2.72	0.502

Table C.8 – The mono-H signal uncertainties due to variation of the PDFs for the 2HDM model in intervals of $E_{\mathrm{T}}^{\mathrm{miss}}$ for values of the mass of the $Z' \leq 1600$ GeV. In the case of the acceptance being less than 1% the errors are neglected and denoted by "-".

Uncertainty on Mono-H Signals due to PDF (%)				
2HDM Model	$E_{\mathrm{T}}^{\mathrm{miss}}$ GeV			
m(prop)_m(A)	150–200	200–350	350–500	>500
mzp1800_mA300	-	-	-	0.348
mzp1800_mA400	-	-	-	0.774
mzp1800_mA500	-	-	9.91	0.0299
mzp1800_mA600	-	6.97	5.82	0.475
mzp1800_mA700	-	5.00	4.99	0.505
mzp1800_mA800	-	3.51	8.38	1.28
mzp2000_mA300	-	-	-	0.413
mzp2000_mA400	-	-	-	0.213
mzp2000_mA500	-	-	-	0.540
mzp2000_mA600	-	-	-	0.359
mzp2000_mA700	-	-	-	0.142
mzp2000_mA800	-	-	-	0.341
mzp2200_mA300	-	-	-	0.288
mzp2200_mA400	-	-	-	0.575
mzp2200_mA500	-	-	-	0.100
mzp2200_mA600	-	-	-	0.478
mzp2200_mA700	-	-	-	0.238
mzp2200_mA800	-	-	-	0.349
mzp2400_mA200	-	-	-	0.463
mzp2400_mA300	-	-	-	0.320
mzp2400_mA400	-	-	-	0.389
mzp2400_mA500	-	-	-	0.504
mzp2400_mA600	-	-	-	0.488
mzp2400_mA700	-	-	-	0.105
mzp2400_mA800	-	-	-	0.245
mzp2600_mA200	-	-	-	0.302
mzp2600_mA300	-	-	-	0.626
mzp2600_mA400	-	-	-	0.226
mzp2600_mA500	-	-	-	0.306
mzp2600_mA600	-	-	-	0.280
mzp2600_mA700	-	-	-	0.374
mzp2600_mA800	-	-	-	0.390
mzp2800_mA200	-	-	-	0.176
mzp2800_mA300	-	-	-	0.252
mzp2800_mA400	-	-	-	0.244
mzp2800_mA500	-	-	-	0.277
mzp2800_mA600	-	-	-	0.245
mzp3000_mA200	-	-	-	0.303
mzp3000_mA300	-	-	-	0.438
mzp3000_mA400	-	-	-	0.434
mzp3000_mA500	-	-	-	0.162
mzp3000_mA600	-	-	-	0.242

Table C.9 – The mono-H signal uncertainties due to variation of the PDFs for the 2HDM model in intervals of $E_{\mathrm{T}}^{\mathrm{miss}}$ for values of the mass of the $Z' \geq 1800$ GeV. In the case of the acceptance being less than 1% the errors are neglected and denoted by "-".

Appendix D

QCD multi-jet background estimation in older Atlas offline release

The plots shown in Section 6.5.4 were part of a validation performed by the author of a new ATLAS software reconstruction release, i.e. Athena Release 21. In this Appendix, plots in the release consistent with the analysis result presented in Chapter 7 are shown for consistency. However, conclusions stays the same.

Fits are performed for each $E_{\mathrm{T}}^{\mathrm{miss}}$ interval and for each b-tag multplicity individually, Figure D.1, D.2, D.3.

Pulls can be found in Figure D.4.

In the merged regime where there is a large-R jet and $E_{\mathrm{T}}^{\mathrm{miss}} > 500$ GeV, the the contribution of multijet events to the background is negligible. Here the $E_{\mathrm{T}}^{\mathrm{miss}}$ shape of multijet events is used. The event selection of the merged regime is applied in the same way as it is done in the final search for dark matter. Inverting the cut mentioned above, so that $\min(\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}},\text{small-R jets})) < 20°$, a template of the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution of multijet events is derived. In order to obtain a useful multijet template, values below $E_{\mathrm{T}}^{\mathrm{miss}} = 500$ GeV are also taken into account. Then a $E_{\mathrm{T}}^{\mathrm{miss}}$ fit of the model to the data is performed in the region $\min(\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}},\text{small-R jets})) > 20°$. The resulting distributions are shown together with the data in Fig. D.5. As can be seen multijet contribution above $E_{\mathrm{T}}^{\mathrm{miss}} > 500$ GeV is vanishingly small. For all b-tag multiplicities the fraction of multijet events in merged regime calculated from these fits is less than 1%.

The fake factors that will scale the multijet template in the SR, obtained by these fits, are listed in Table D.1.

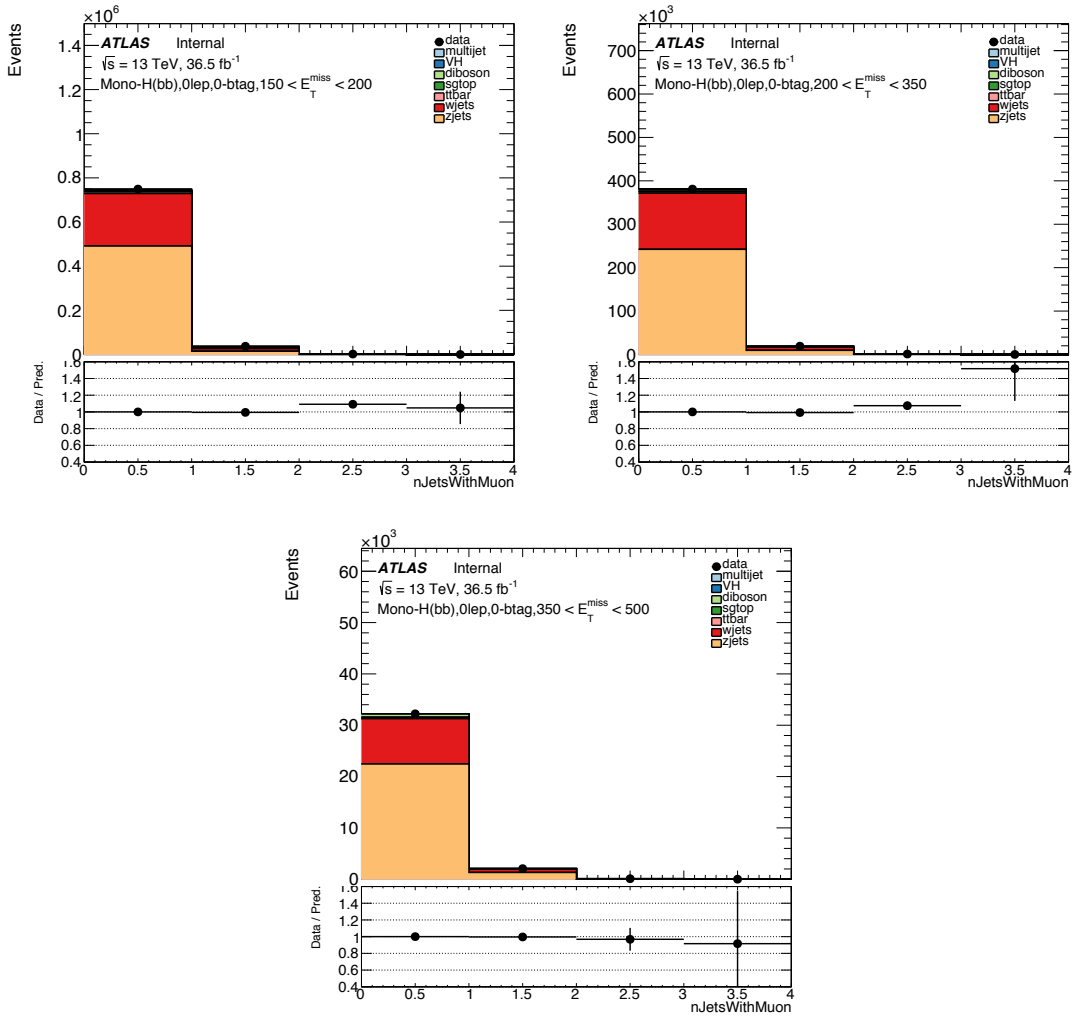


Figure D.1 – Distributions of the number of small-R jets that contain a muon for 0 tag small R jets shown for all three resolved $E_{\mathrm{T}}^{\mathrm{miss}}$ bins. The small-R data is fit with templates of the multijet and MC.

		$E_{\mathrm{T}}^{\mathrm{miss}} \in [150, 200]$			$E_{\mathrm{T}}^{\mathrm{miss}} \in [200, 350]$		
	Parameters	0-tag	1-tag	2-tag	0-tag	1-tag	2-tag
Normalizations	Multijet	0.26 ± 0.03	0.13 ± 0.01	0.12 ± 0.01	0.11 ± 0.05	0.02 ± 0.01	0.06 ± 0.01
NP Pulls	Luminosity	0.45 ± 1.02	2.30 ± 0.92	2.24 ± 0.84	0.30 ± 0.95	1.45 ± 0.94	1.37 ± 0.88
	$t\bar{t}$	0.33 ± 0.99	2.32 ± 1.01	2.87 ± 0.99	0.29 ± 0.99	1.41 ± 1.02	1.32 ± 0.94
	single top	0.04 ± 0.99	0.13 ± 0.99	0.01 ± 0.99	0.01 ± 0.99	0.05 ± 0.99	0.06 ± 0.99
	diboson	-0.03 ± 0.99	0.05 ± 0.99	0.02 ± 0.99	-0.01 ± 0.99	0.10 ± 0.99	0.05 ± 0.99
	W+jets	-1.05 ± 0.83	3.78 ± 0.78	1.35 ± 1.10	-0.31 ± 0.78	0.66 ± 1.03	0.78 ± 1.11
	Z+jets	1.95 ± 0.49	0.69 ± 0.94	1.01 ± 0.92	0.88 ± 0.57	2.09 ± 0.53	1.82 ± 0.77

Table D.1 – Table of scale factors and pulls obtained by fitting multiplicity of jets which contain a muon variable in fit selection region as described above. In the table scale factors for multijet normalization and pulls for luminosity and other backgrounds are shown. Luminosity uncertainty is taken as 3.7%. The other backgrounds are constrained by theoretical uncertainties which are taken as: $t\bar{t}$: 6%, single top: 5%, diboson: 10%, W+jets: 20%, Z+jets: 20%.

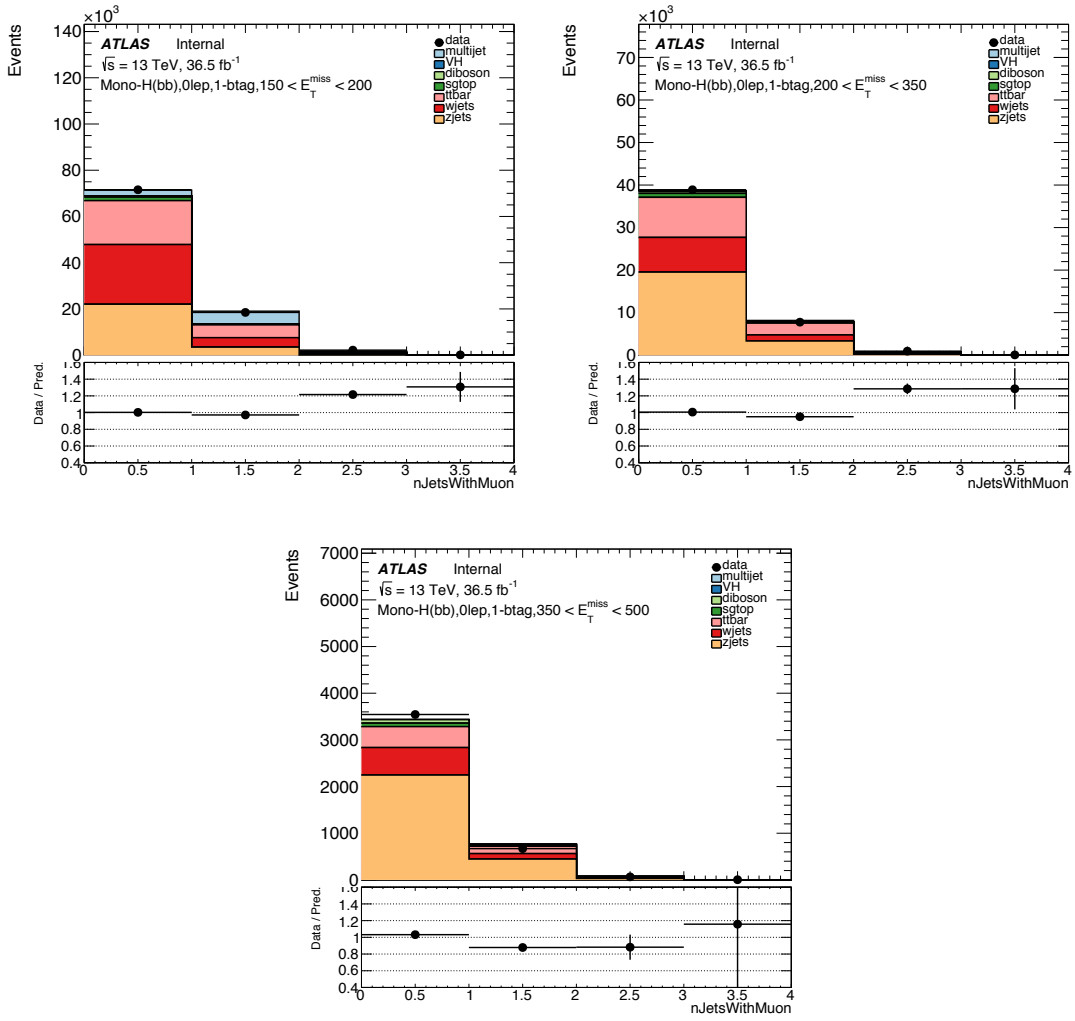


Figure D.2 – Distributions of the number of small-R jets that contain a muon for 1 tag small R jets shown for all three resolved $E_{\mathrm{T}}^{\mathrm{miss}}$ bins. The small-R data is fit with templates of the multijet and MC.

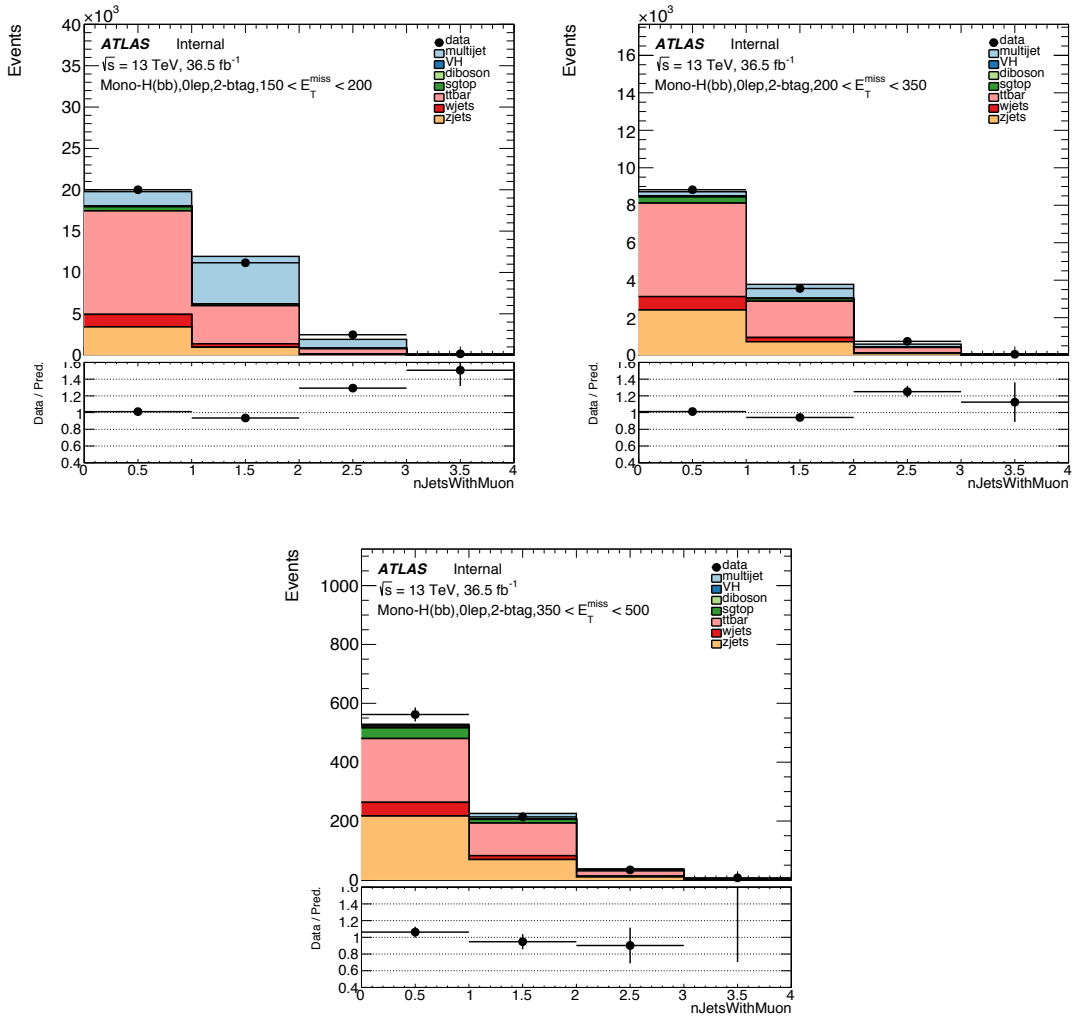
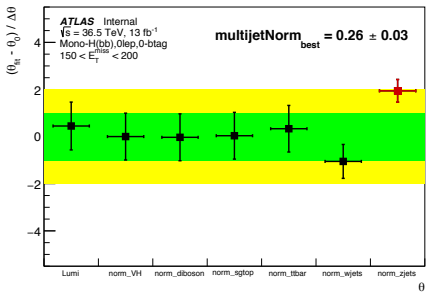
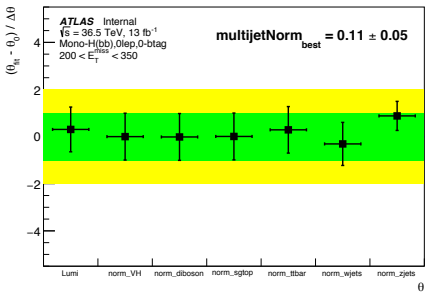


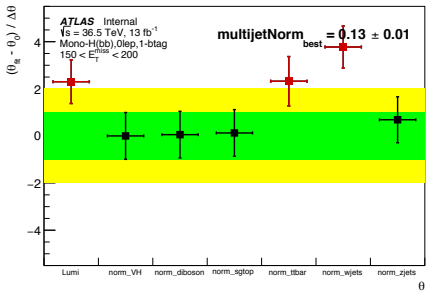
Figure D.3 – Distributions of the number of small-R jets that contain a muon for 2 tag small R jets shown for all three resolved $E_{\mathrm{T}}^{\mathrm{miss}}$ bins. The small-R data is fit with templates of the multijet and MC.



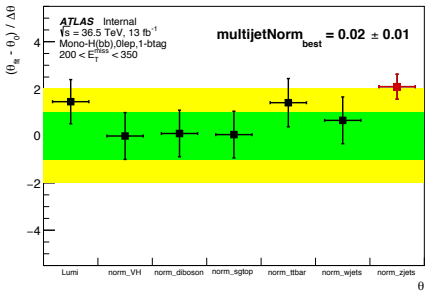
(a) 0-btag, 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 200 GeV



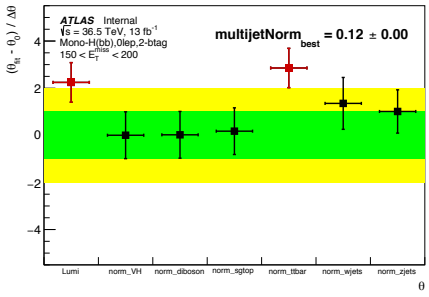
(b) 0-btag, 200 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 350 GeV



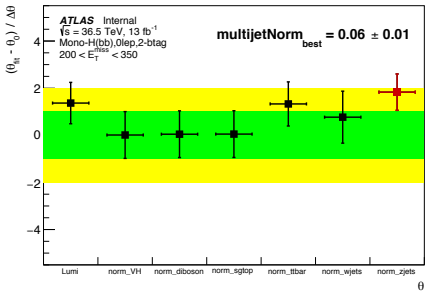
(c) 1-btag, 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 200 GeV



(d) 1-btag, 200 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 350 GeV



(e) 2-btag, 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 200 GeV



(f) 2-btag, 200 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 350 GeV

Figure D.4 – The pulls of the profile likelihood fit for each b-tag multiplicity and first two $E_{\mathrm{T}}^{\mathrm{miss}}$ bins.

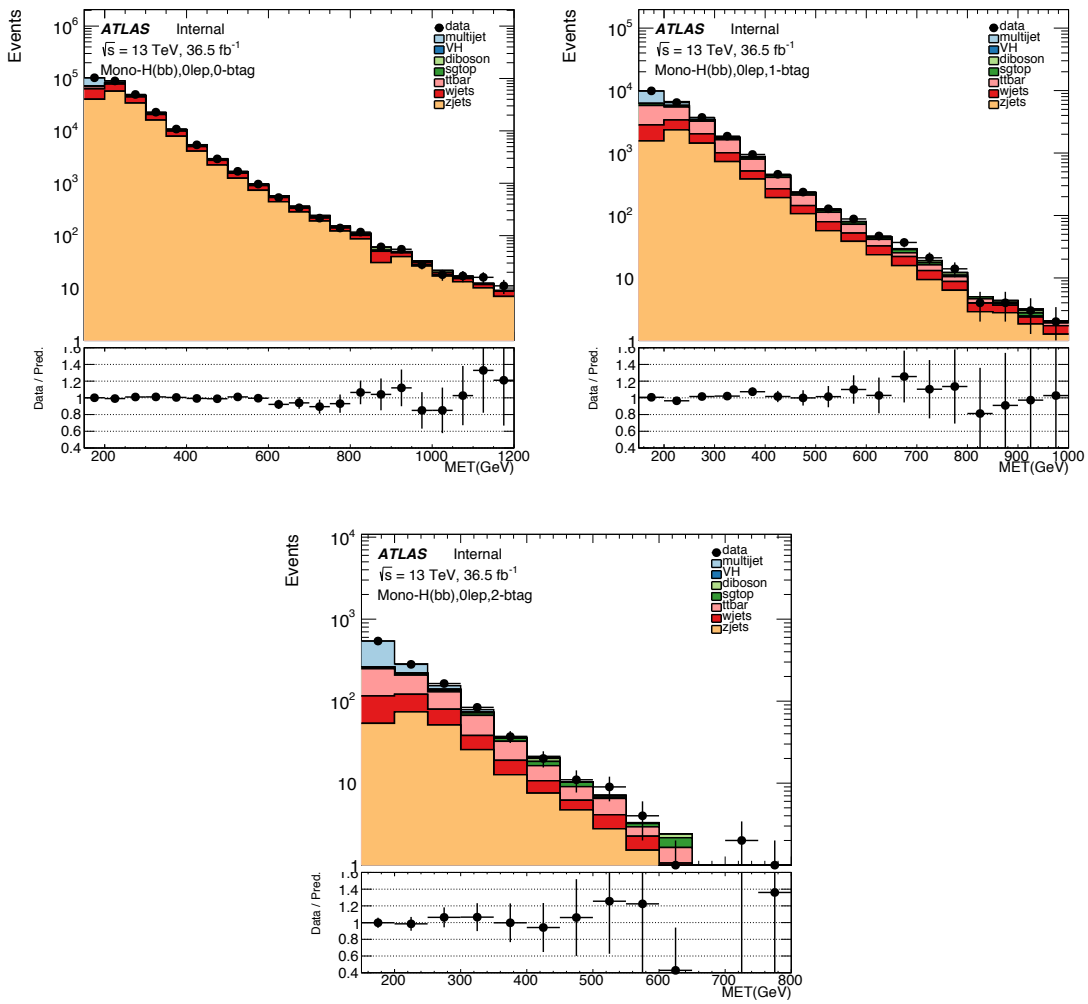


Figure D.5 – Distribution of the $E_{\mathrm{T}}^{\mathrm{miss}}$ showing that the multijet background contribution is negligible for events with high $E_{\mathrm{T}}^{\mathrm{miss}}$.

Appendix E

b-tagging calibration for variable radius jets

For the VR track jets, a working point with average b-tagging efficiency of 77% for a fixed cut in the `MV2c10` discriminant is used, see Section 4.2.3 for detail in b-tagging procedure. The b-tagging algorithm needs to be calibrated in data with the use of scale Factors (SF) in order to correct the b-tagging performance in simulation to that observed in data. These SF are defined as usual:

$$SF \equiv \frac{\varepsilon_b^{\text{data}}}{\varepsilon_b^{\text{MC}}}, \tag{E.1}$$

where $\varepsilon_b^{\text{data}}$ and $\varepsilon_b^{\text{MC}}$ correspond to the b-tagging efficiency in data and simulation, respectively.

The efficiency is estimated from the data set collected in 2015 and 2016 analysing a $t\bar{t}$ sample, in which both the W bosons from the top quarks decay leptonically, using a combinatorial likelihood approach [**?**]. This procedure is equivalent to the one carried out for the small-R calorimeter jets.

The VR track jets considered for the calibration are required to have $p_{\text{T}} > 10$ GeV and $|\eta| < 2.5$. Moreover, an additional cut on the angular separation between the two leading associated VR track jets $\Delta R(\text{VR}_1, \text{VR}_2)$ is applied. This is motivated given that a high-p_{T} VR jet with small radius can be enclosed in a low-p_{T} VR jet jet with a large radius, an in a limit case the two can be concentric. In these critical cases, the truth flavour labelling and the track-to-jet association, which rely on the angular separation of the jet with respect to the truth hadron and the tracks, can be compromised given that the two jet axes are approximately collinear. In order to avoid these pathological cases with an ambiguous b-tagging, events for which the $\Delta R(\text{VR}_1, \text{VR}_2)$ is lower than the radius of the smaller VR track jet are rejected. This cut is also applied in the analysis event selection.

The SFs are measured as a function of the jet p_{T} and derived up to p_{T} =250 GeV, above which the event statistics is limited. The uncertainty in the SFs for the p_{T} range below 250 GeV is dominated by uncertainties on the modelling of $t\bar{t}$ events in the simulation. At high p_{T}, the SFs are fixed to those in the last explicitly calibrated bin, and

an additional high-p_{T} uncertainty is assigned to account for effects on the b-tagging performance from possible reconstruction imperfections in the simulation. This is inherited from the procedure applied for calorimeter jets as described in Reference [163].

Figure E.1 shows in the top panel the b-tagging efficiency in data (black markers) and the $t\bar{t}$ simulation (dashed red line). The scale factors resulting as the ratio between data and simulation are shown in the bottom panel. The band is the total uncertainty including statistical and systematic uncertainties.

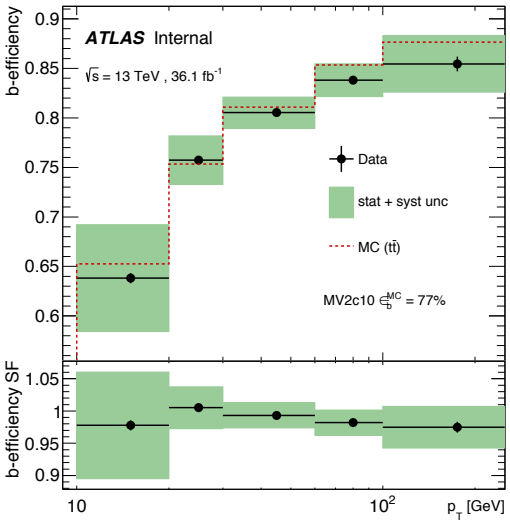


Figure E.1 – b-tagging efficiency (top) estimated in data (black symbols) and a $t\bar{t}$ simulation (dashed red line) and the resulting b-tagging scale factors (bottom) obtained for the 77% efficiency working point. The error band (green) includes both statistical and systematic uncertainty.

Similarly to the efficiency for tagging jets originating from b-quarks, the rate of incorrectly tagging jets originating from charm or light-flavour quarks as b-jets has to be calibrated. Since the effect of the mistag uncertainties is subdominant, calibration results of these mistag efficiencies based on calorimeter jets are extrapolated to the case of track jets.

Appendix F

Object-based transverse momentum significance in mono-$h(b\bar{b})$ analysis

Figure F.1 shows the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance distributions for data recorded in 2015 and 2016, and MC simulation for events with 1 b-tagged jet in the three resolved $E_{\mathrm{T}}^{\mathrm{miss}}$ categories after full event selection.

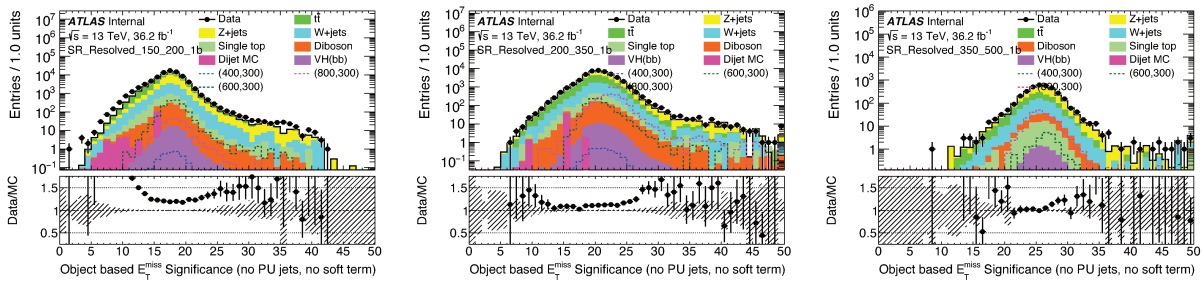
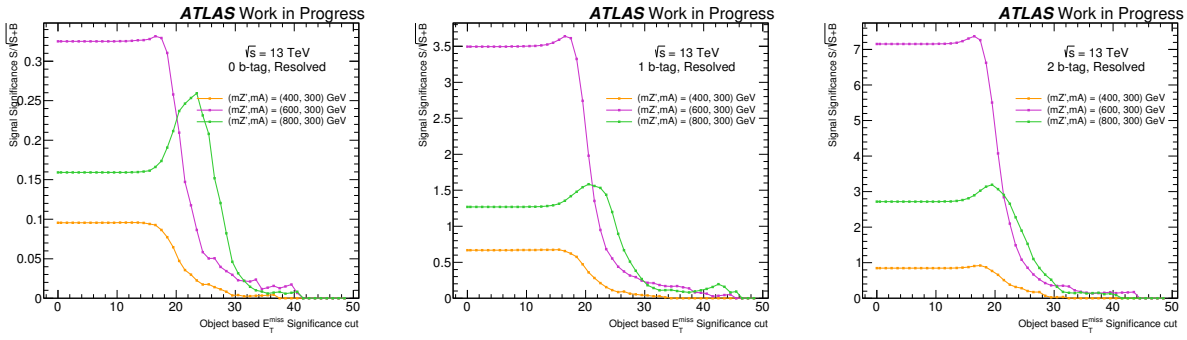


Figure F.1 – Met significance distribution for events with 1 b-tagged jet in the regions with 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 200 GeV (left), 200 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 350 GeV (middle), 350 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 500 GeV (right). Only statistical uncertainties are taken into account for displaying the uncertainty bands.

Figure F.2 shows the expected signal significance for different cut values in $\mathcal{S}$, but for all the resolved regime $E_{\mathrm{T}}^{\mathrm{miss}}$ range, in the case of 0, 1 and 2 b-tag selection in Figures F.2(a), F.2(b) and F.2(c), respectively.

Figure F.3 shows the performance of the object- and event-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significances, and $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ for which the efficiencies for signal and background are estimated with respect to a selection that requires that the minimum angular separation between the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ vector and the leading jets in an event is greater than 20°. Figure F.3, illustrates the significant improvement obtained with the object-based definition employed in this search. For example, for signal efficiencies above 80%, the background rejection factor $(1/\varepsilon_B)$ for the object-based significance is approximately three to four times as high as for the event-based one.

Figure F.4 shows the distributions of the invariant mass of the Higgs candidate dijet system in the three resolved event selection $E_{\mathrm{T}}^{\mathrm{miss}}$ categories for events with 2 b-tagged



(a) (b) (c)

Figure F.2 – Met significance distribution for events with 2 b-tagged jets in the regions with $150\,\text{GeV} < E_{\text{T}}^{\text{miss}} < 200\,\text{GeV}$ (left), $200\,\text{GeV} < E_{\text{T}}^{\text{miss}} < 350\,\text{GeV}$ (middle), $350\,\text{GeV} < E_{\text{T}}^{\text{miss}} < 500\,\text{GeV}$ (right). Only statistical uncertainties are taken into account for displaying the uncertainty bands. The signal distributions overlayed are normalised to 1 pb.

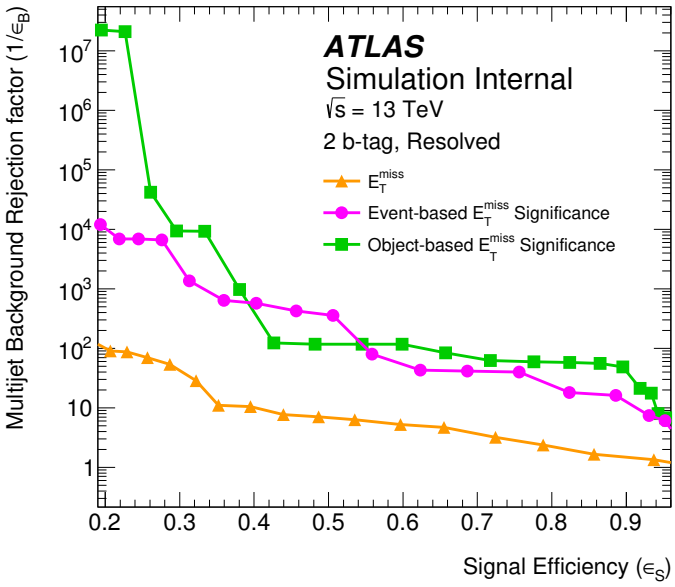


Figure F.3 – Performance of the object-based $\boldsymbol{E}_{\text{T}}^{\text{miss}}$ significance (line with square markers) in terms of the signal efficiency and background rejection as estimated from a signal simulation with $m_{Z'} = 400\,\text{GeV}$ and $m_A = 300$ GeV and a dijet simulation, respectively, in comparison to an alternative definition of the $\boldsymbol{E}_{\text{T}}^{\text{miss}}$ significance (dashed line with circular markers) and $\boldsymbol{E}_{\text{T}}^{\text{miss}}$ itself (densely-dashed line with triangular markers). Selections on these three variables are applied in addition to a requirement on the angular separation between the $\boldsymbol{E}_{\text{T}}^{\text{miss}}$ vector and the leading jets in an event.

jets in linear scale. Other than supressing the multijet background, introducing the cut has little effect on the general shape of the distributions used as input to the statistical analysis and to the background composition of the resolved signal regions.

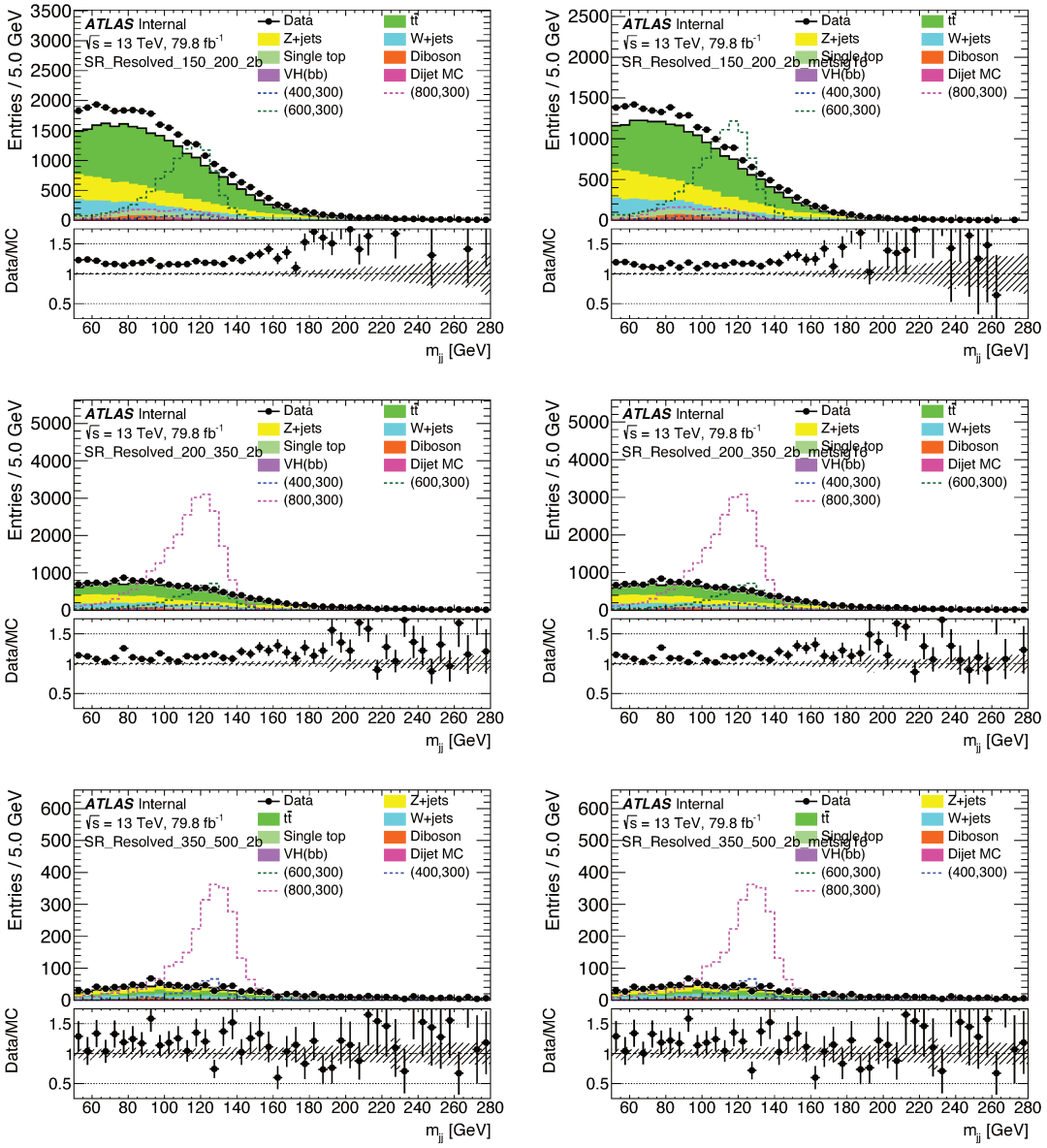


Figure F.4 – Distributions of the invariant mass of the Higgs candidate dijet system in the three resolved event selection $E_{\mathrm{T}}^{\mathrm{miss}}$ bins for events with 2 b-tagged jets, without (left) and with $E_{\mathrm{T}}^{\mathrm{miss}}$ significance > 16 requirement (right).

F.1 Object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance modelling

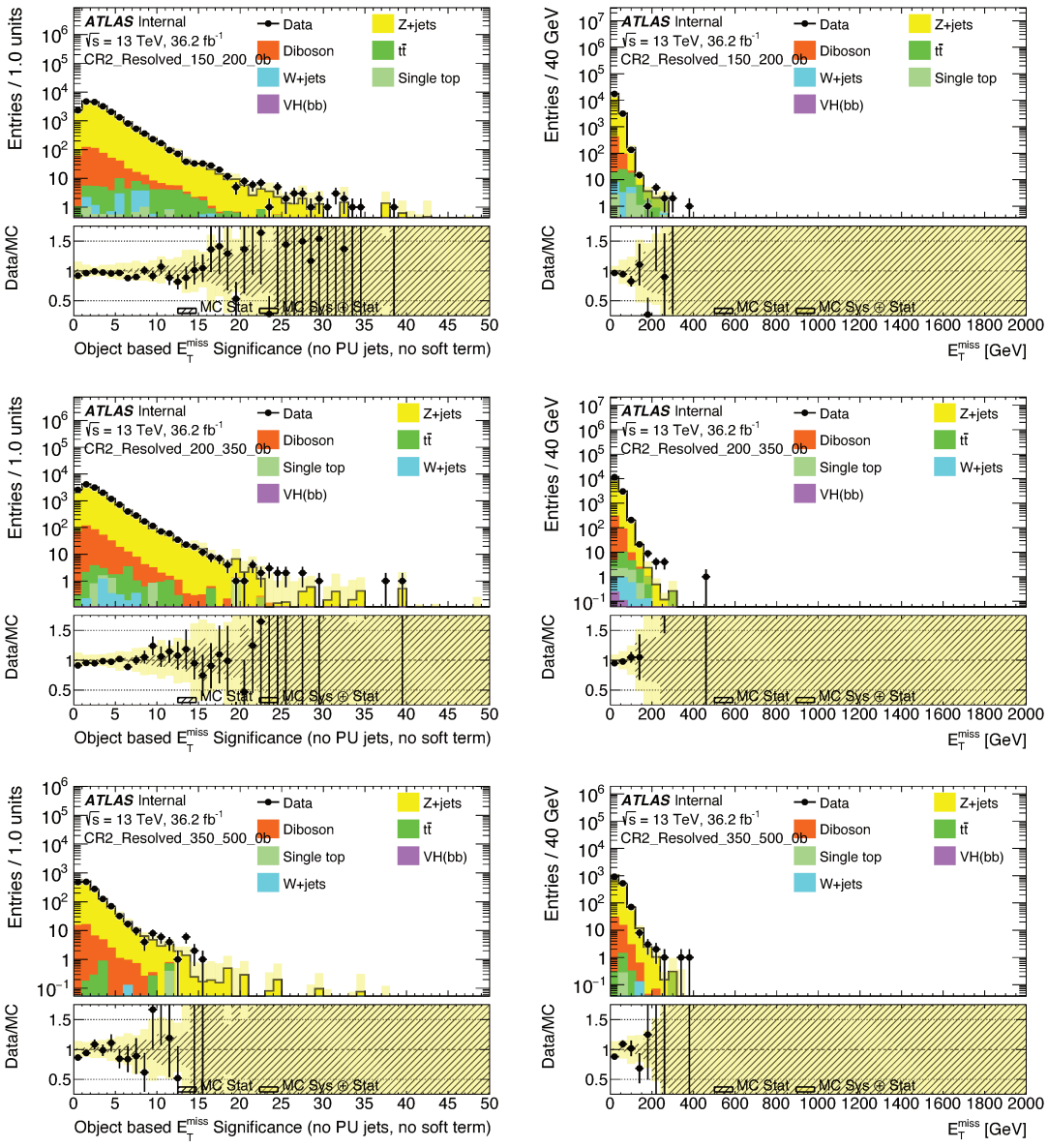


Figure F.5 – Object based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance (left) and $E_{\mathrm{T}}^{\mathrm{miss}}$ (right) distributions in the 2 lepton control region (without the event-based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance applied in event selection) for events with 0 b-tagged jets in different $E_{\mathrm{T}}^{\mathrm{miss}}$ bins.

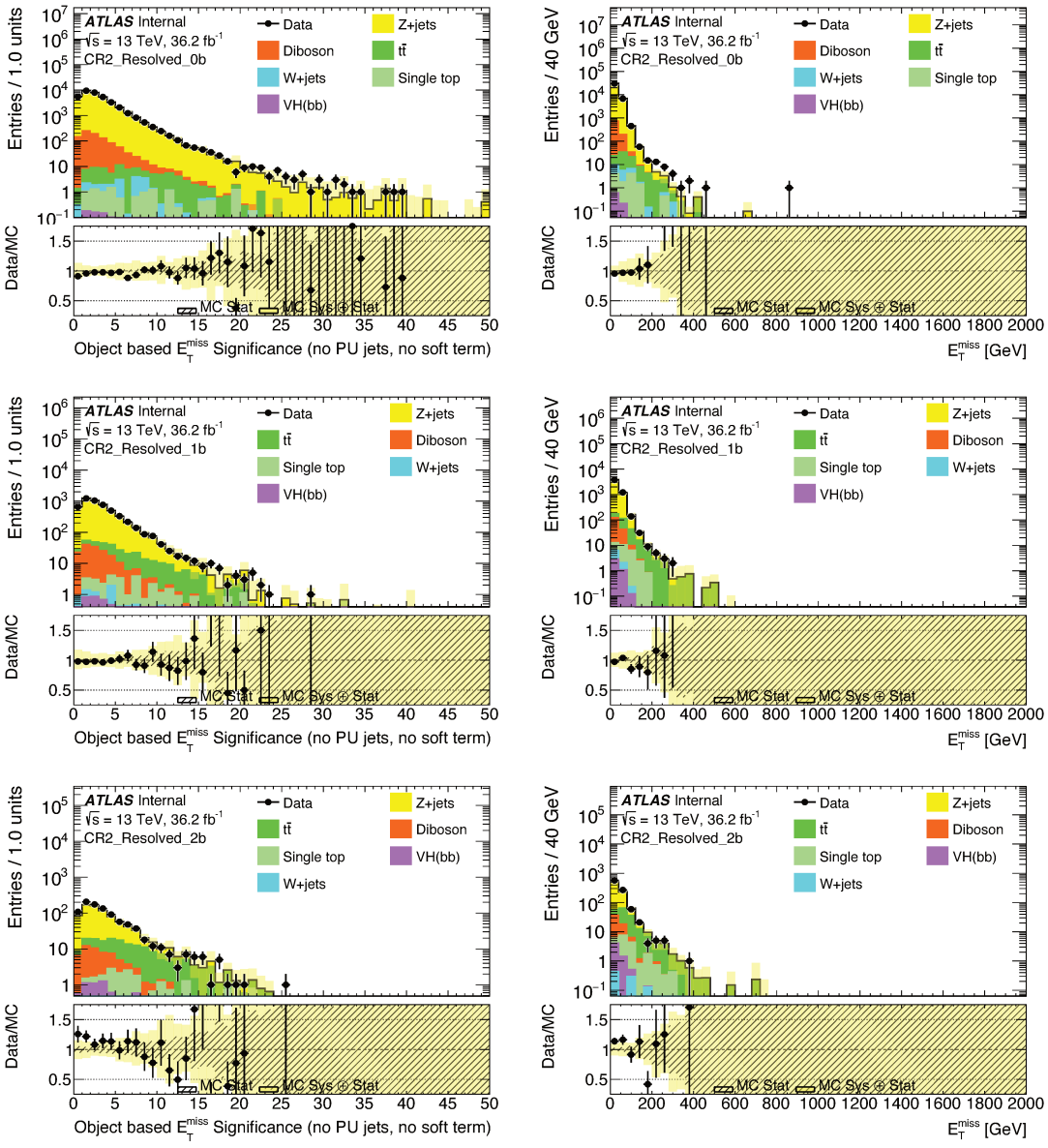


Figure F.6 – Object based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance (left) and $E_{\mathrm{T}}^{\mathrm{miss}}$ (right) distributions in the 2 lepton control region (without the event-based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance applied in event selection) for events with 0, 1 and 2 b-tagged jets.

Appendix G

QCD multi-jet background data-driven alternative estimations

Besides the method for estimating the multijet background contribution to the signal region described in Section 8.5.1 alternative strategies are listed here.

For the methods outlined in this Appendix, the requirement of $\mathcal{S} > 16$ is not considered. It was shown in Section 8.5.1 that after consideration of a object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance cut in the resolved signal region event selection, no multijet background estimation is needed.

G.1 Template method using the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance as fitting variable

The object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance, introduced in Chapter 5, can help to identify and separate multijet background with respect to the other SM backgrounds. Even though many of the multijet events comes with real $E_{\mathrm{T}}^{\mathrm{miss}}$ caused by neutrinos originated from heavy flavour hadron decays, they are more likely to have lower values of the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance compared to other SM processes and the dark matter signals.

The same data-driven multijet background estimation strategy presented in Section 6.5.4 is considered, however, instead of using the multiplicity of jets which contain a muon as a fitting variable, the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance is used in order to obtain the multijet template and it's corresponding normalisation.

The likelihood fits are performed for the first two $E_{\mathrm{T}}^{\mathrm{miss}}$ intervals and for each b-tag multiplicity independently and postfit plots for the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance variable are shown in Figure G.1

In the fit, the simulated MC backgrounds are free to float independently within theoretical uncertainties with one overall normalization factor that controls the normalization of non-multijet components in the fit. This overall normalization, or luminosity, has an uncertainty taken to be 2.1%. The other backgrounds are constrained by theoretical uncertainties which are taken as: $t\bar{t}$: 6%, single top: 5%, diboson: 10%, W+jets: 20%, Z+jets: 20%. The corresponding pulls of the profile likelihood fit for the fitting categories are shown in Figure G.2.

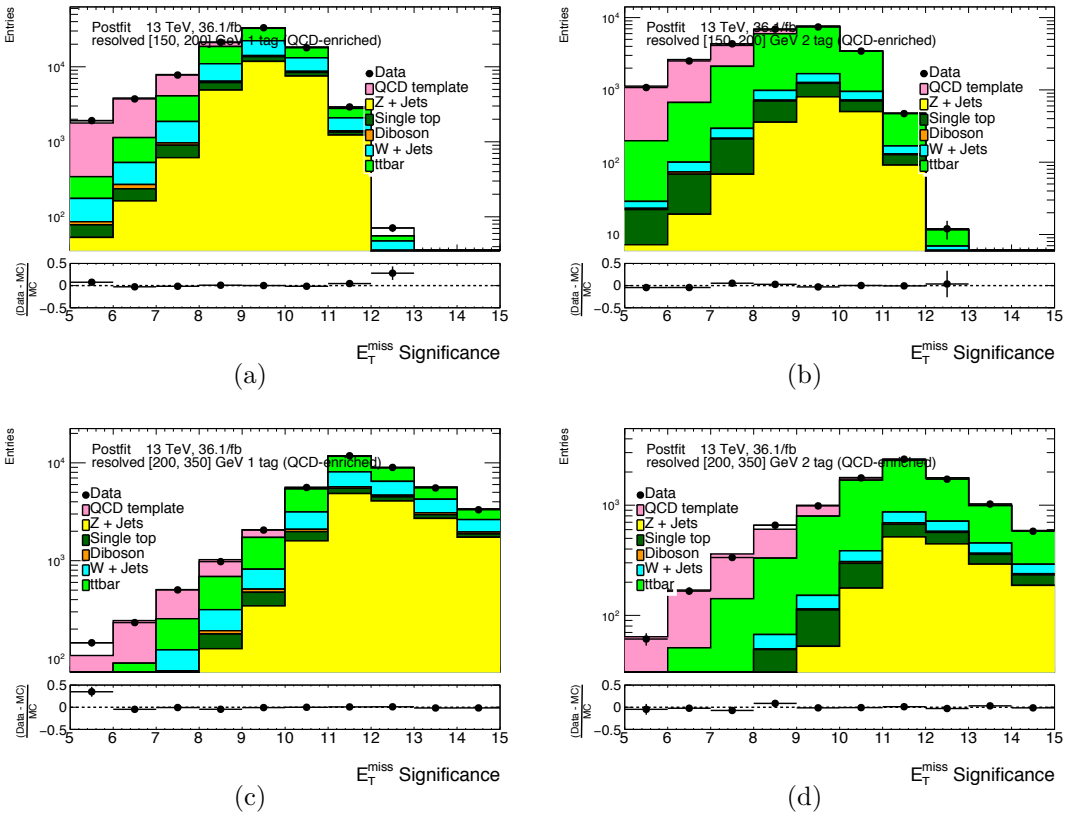
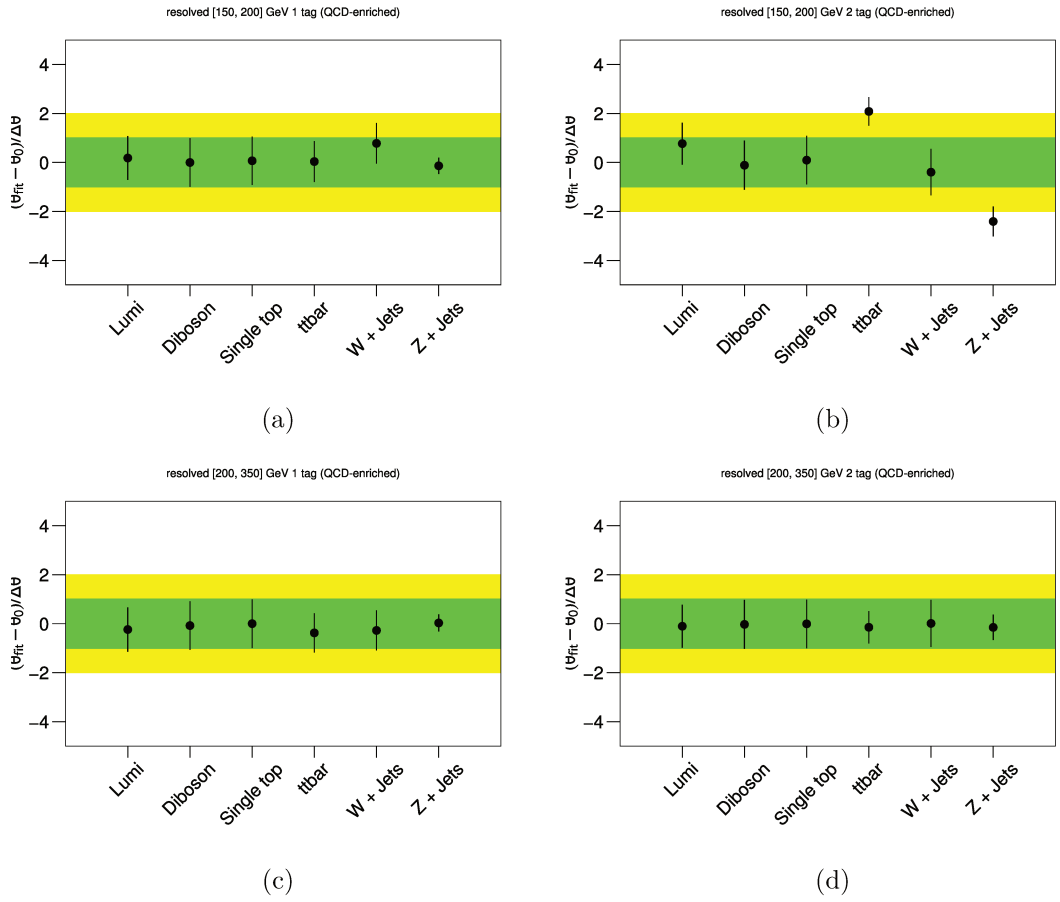


Figure G.1 – Distributions of the object based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance for 1 and 2 tagged small radius jets shown for the first two resolved $E_{\mathrm{T}}^{\mathrm{miss}}$ bins. The data is fit with templates of the multijet background and modeled MC in the regions with (a) 150 GeV $< \mathrm{E}_{\mathrm{T}}^{\mathrm{miss}} <$ 200 GeV, 1-btag, (b) 150 GeV $< \mathrm{E}_{\mathrm{T}}^{\mathrm{miss}} <$ 200 GeV, 2-btag, (c) 200 GeV $< \mathrm{E}_{\mathrm{T}}^{\mathrm{miss}} <$ 350 GeV, 1-btag, (d) 200 GeV $< \mathrm{E}_{\mathrm{T}}^{\mathrm{miss}} <$ 350 GeV, 2-btag,



(a) (b)

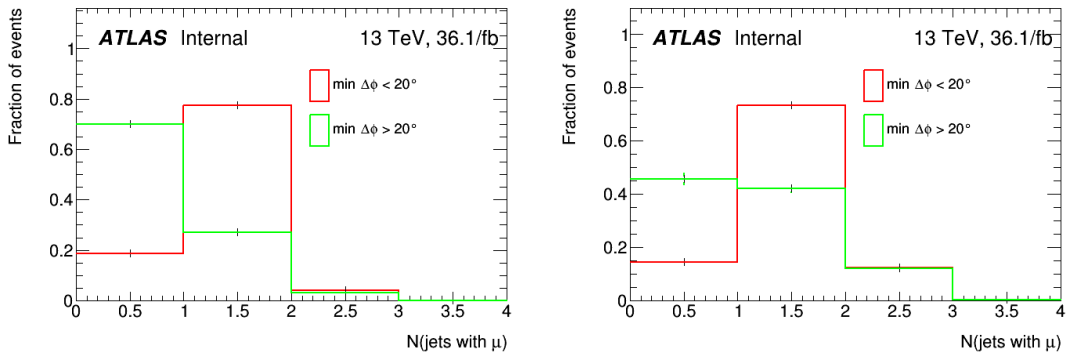
(c) (d)

Figure G.2 – Pulls of the profile likelihood fit for 1 and 2 b-tag multiplicities and the first two resolved $E_{\mathrm{T}}^{\mathrm{miss}}$ bins, using the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance variable: (a) 150 GeV $< \mathrm{E}_{\mathrm{T}}^{\mathrm{miss}} <$ 200 GeV, 1-btag, (b) 150 GeV $< \mathrm{E}_{\mathrm{T}}^{\mathrm{miss}} <$ 200 GeV, 2-btag, (c) 200 GeV $< \mathrm{E}_{\mathrm{T}}^{\mathrm{miss}} <$ 350 GeV, 1-btag, (d) 200 GeV $< \mathrm{E}_{\mathrm{T}}^{\mathrm{miss}} <$ 350 GeV, 2-btag,

The normalisations obtained by these fits in each of the categories are listed in Table G.1. It is observed that these normalisations decrease as the $E_{\mathrm{T}}^{\mathrm{miss}}$ increases. For $E_{\mathrm{T}}^{\mathrm{miss}} > 300$ GeV, the multijet contribution can be neglected.

G.2 Using the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance to directly construct a template

In the multijet estimation presented in Section 6.5.4 a fit of the multijet templates and simulated backgrounds to data in the mass-sidebands of the signal region was performed using the multiplicity of jets containing a muon as fit variable. In the improved analysis described in Chapter 8, it was observed that the multiplicity of jets containing a muon is correlated with min($\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}}$,small-R jets)). Figure G.3 shows a comparison of the shapes in the region to the left and to the right of the cut, which differ substantially.



(a) (b)

Figure G.3 – Shape comparison of the distributions of the multiplicity of jets containing a muon betwen the multijet-enriched region and the signal region: (a) resolved 1-btag 150 - 200 GeV, (b) resolved 2-btag 150 - 200 GeV,

Siminarly, the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance shows to be correlated with respect the min($\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}}$,small-R jets)) observable. Therefore, the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance template shape slightly changes under inversion of the min($\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}}$,small-R jets)) requirement. Because

Table G.1 – Multijet normalisation factors (SR/CR normalisation ratios) obtained by fitting the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance variable in fit selection region as described above. The uncertainty stated is the relative uncertainty on the normalisation.

Multijet Normalisation				
	150 GeV < $E_{\mathrm{T}}^{\mathrm{miss}}$ < 200 GeV		200 GeV < $E_{\mathrm{T}}^{\mathrm{miss}}$ < 350 GeV	
	1-btag	2-btag	1-btag	2-btag
normalisation	0.191	0.102	0.079	0.058
relative uncertainty	0.035	0.034	0.065	0.064

of this, the approach in Section G.1, in which the fit is performed with the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance instead of the number of jets with muons, is not completely appropriate.

A new approach based on the object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance for estimating the normalisation factor is presented below. This particular definition of the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance is based on the uncertainties for individual objects, which are propagated to the $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance, excluding the resolution of pile-up jets and the $E_{\mathrm{T}}^{\mathrm{miss}}$ soft term resolution as described in Chapter 5.

The normalisation factor ε for the multijet templates obtained from the multijet enriched region can be written as a ratio of the multijet events in the (blinded) signal region (SR) over the multijet events in (blinded) blinded multijet enriched region (CRMJ)

$$\varepsilon = \frac{\int_{SR,50\,\mathrm{GeV}}^{70\,\mathrm{GeV}} \frac{\mathrm{d}N}{\mathrm{d}m_{jj}} + \int_{SR,140\,\mathrm{GeV}}^{280\,\mathrm{GeV}} \frac{\mathrm{d}N}{\mathrm{d}m_{jj}}}{\int_{CRMJ,50\,\mathrm{GeV}}^{70\,\mathrm{GeV}} \frac{\mathrm{d}N}{\mathrm{d}m_{jj}} + \int_{CRMJ,140\,\mathrm{GeV}}^{280\,\mathrm{GeV}} \frac{\mathrm{d}N}{\mathrm{d}m_{jj}}}.$$

In order to avoid compromisation of the background estimate by signal contributions only the blinded regions (i.e. the mass side-bands) are considered.

The denominator of this ratio can be directly obtained from the event yield in the blinded multijet m_{jj}-template from the min($\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}}$,small-R jets))$<20^{\circ}$ region.

The nominator can be estimated in the mass-side bands of the signal region by subtracting all simulated backgrounds from the data using a distributions with strong separation between multijet backgroundground and other backgrounds. Then the integral of the difference gives an estimate for the events due to the multijet background in the blinded signal region. The multijet background in the signal region is expected to be small, so that being difference between the data and simulated backgrounds it will mostly depend on the correct description of the simulated backgrounds. Therefore a simplified fit of the simulated backgrounds to the data in a region with no expected multijet contribution is performed to obtain scale factors for the backgrounds which are applied for the multijet estimation method.

G.2.1 Step 1: Multijet templates from multijet-enriched control region

The template distributions obtained from the multijet enriched region defined as the difference between simulated backgrounds and the data are shown in Figure G.4.

G.2.2 Step 2: Estimation of the multijet template normalisation factor

Since estimating the number of multijet events in the blinded signal region is dependent on the normalisations of the simulated backgrounds, a simplified fit of those to the data in a multijet-depleted region is performed in order to obtain scale factors for each region using the met significance distributions as a fit variable.

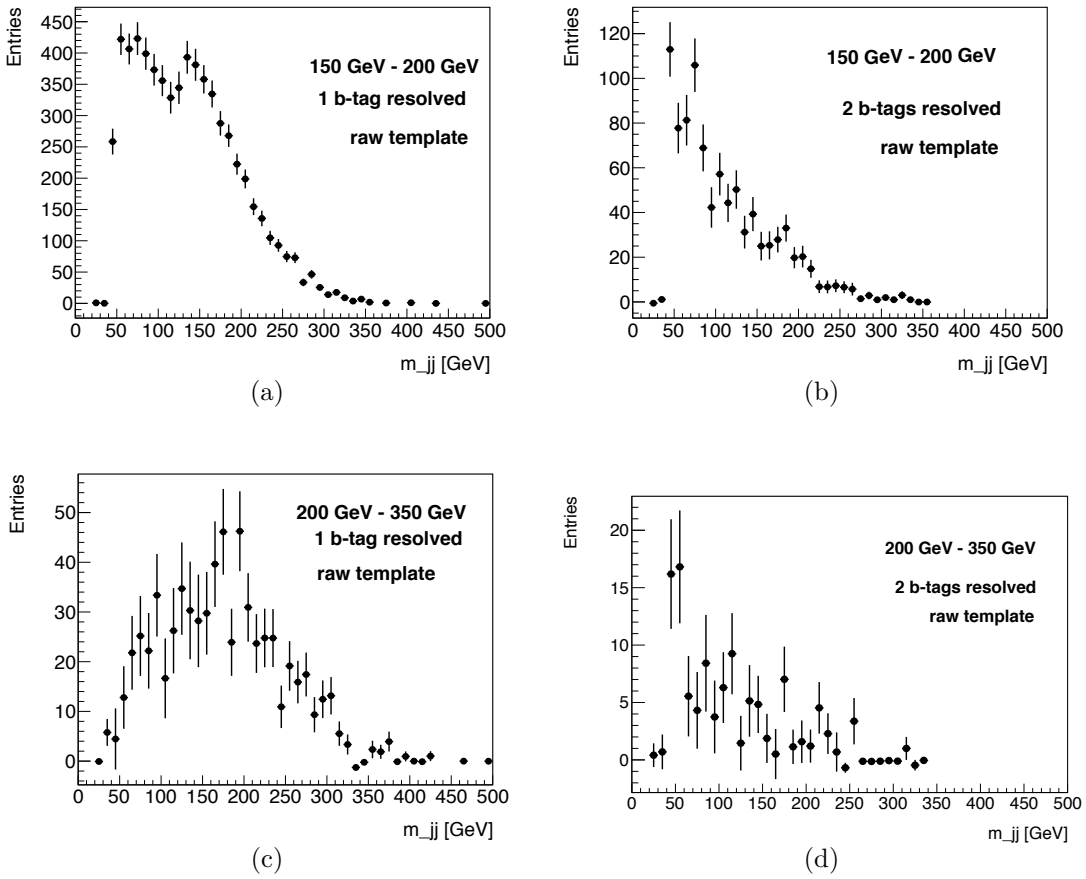


Figure G.4 – Templates for the m_{jj} distribution obtained from the region defined by inverting the min($\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}}$,small-R jets)) cut of the signal region. (a) resolved 1-btag 150 - 200 GeV, (b) resolved 2-btag 150 - 200 GeV, (c) resolved 1-btag 200 - 350 GeV, (d) resolved 2-btag 200 - 350 GeV,

In order to cut away the multijet contribution, only for this fit a cut of object based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance > 15 is applied. The pre-fit distributions and the fit results are shown in Figures G.5, G.6, G.7 and G.8.

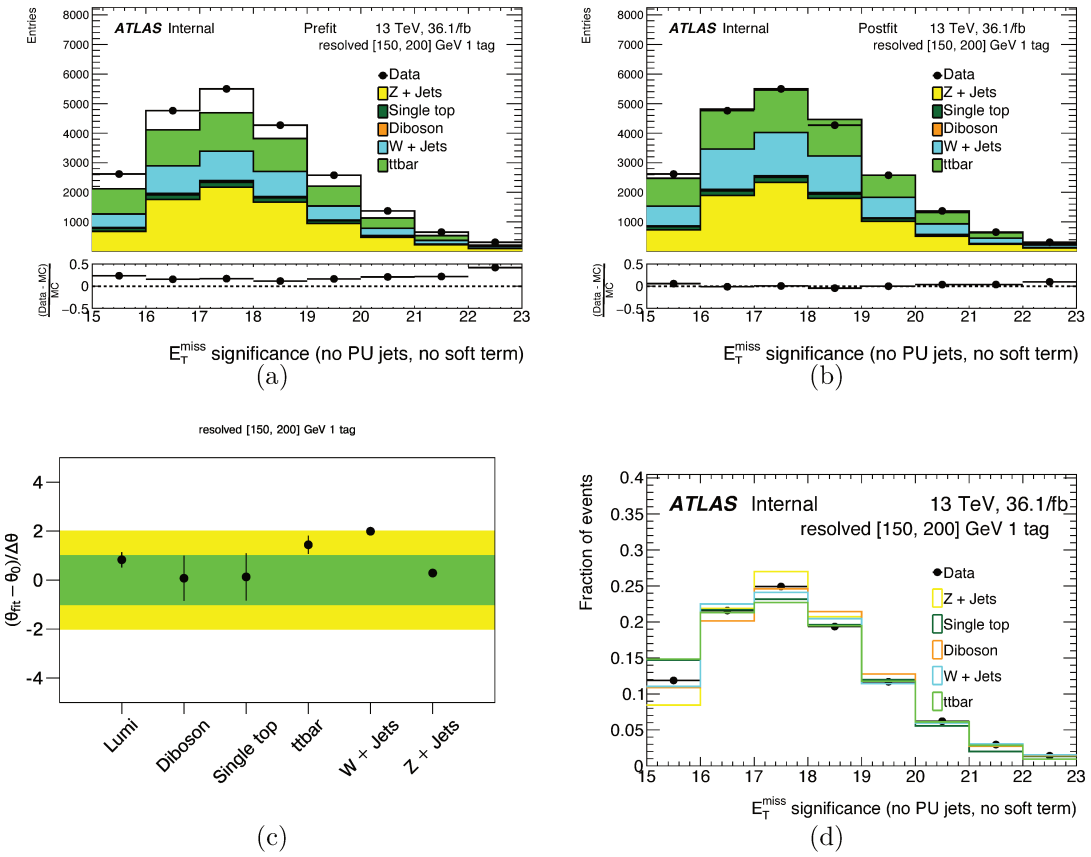


Figure G.5 – Fit results for the 150 GeV $< \mathrm{E}_{\mathrm{T}}^{\mathrm{miss}} <$ 200 GeV, 1-btag region (a) Pre-fit distribution (b) Post-fit distribution (c) Pulls and constraints for background normalisations (d) Comparison of shapes for different simulated backgrounds

The theoretical background uncertanties are listed in Table G.2. They are adapted from the theoretical background uncertainties used in the statistical interpretation of the previous iteration of the analysis [252].

The background scale factors obtained from fitting each region independently are listed in Table G.3. Only a very simple fit setup was used since the multijet templates will be used with a large normalisation uncertainty in the final statistical interpretation of the analysis result.

Table G.2 – Table of assumed theoretical background uncertainties for simulated backgrounds in profile-likelihood fit for obtaining the background scale factors.

background	$t\bar{t}$	W+jets	Z+jets	single top quark	diboson	Luminosity
uncertainty	0.06	0.20	0.20	0.05	0.25	0.021

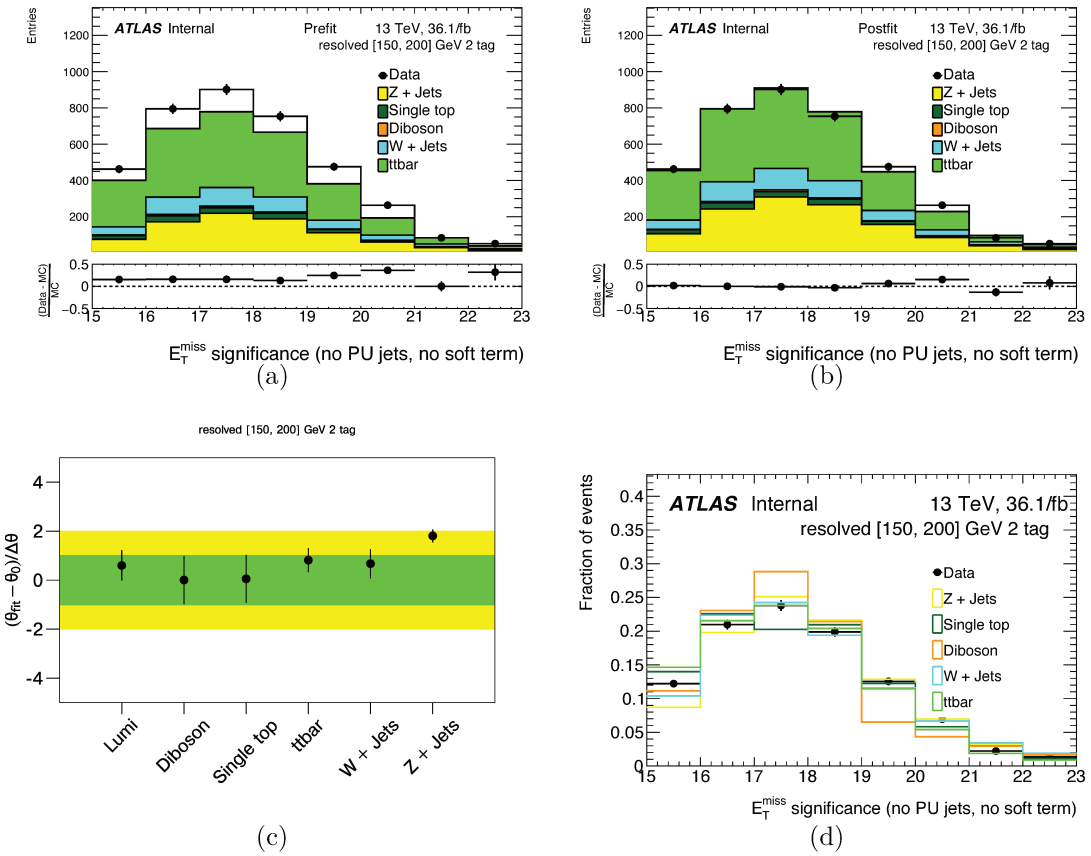


Figure G.6 – Fit results for the 150 GeV $< E_T^{miss} <$ 200 GeV, 2-btag region (a) Pre-fit distribution (b) Post-fit distribution (c) Pulls and constraints for background normalisations (d) Comparison of shapes for different simulated backgrounds

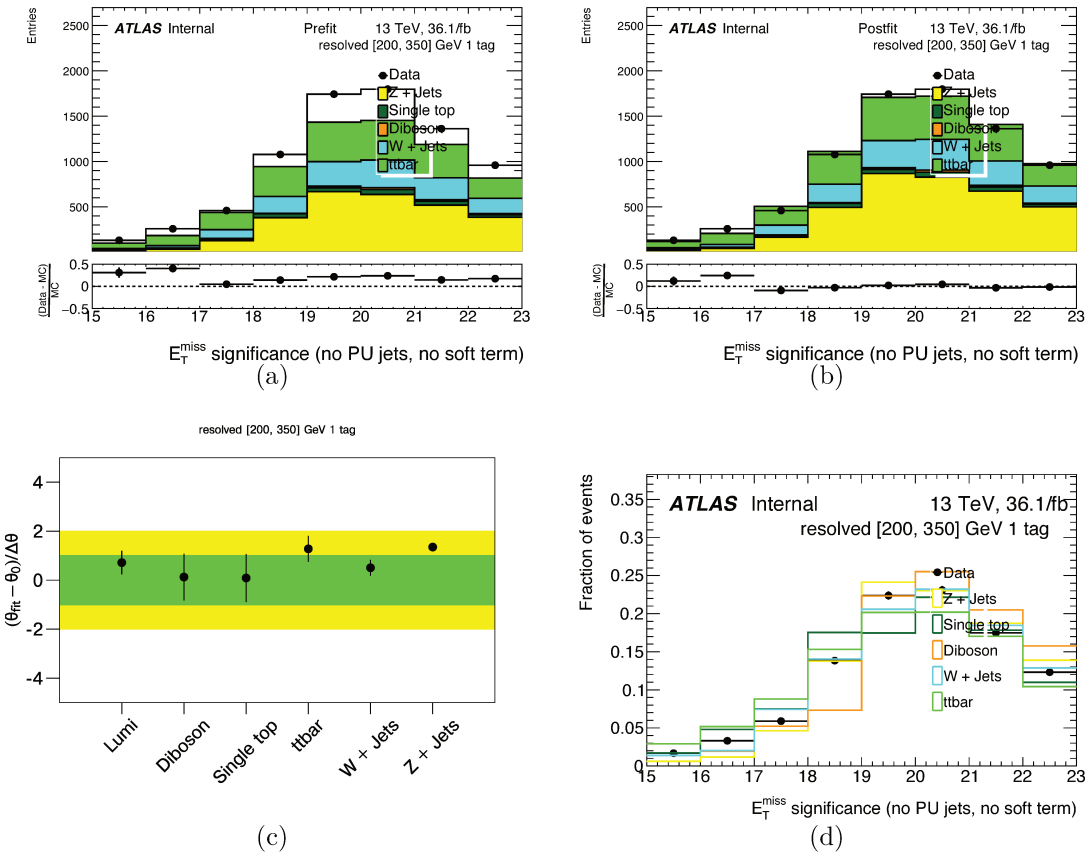


Figure G.7 – Fit results for the 200 GeV $< \mathrm{E}_{\mathrm{T}}^{\mathrm{miss}} <$ 350 GeV, 1-btag region (a) Pre-fit distribution (b) Post-fit distribution (c) Pulls and constraints for background normalisations (d) Comparison of shapes for different simulated backgrounds

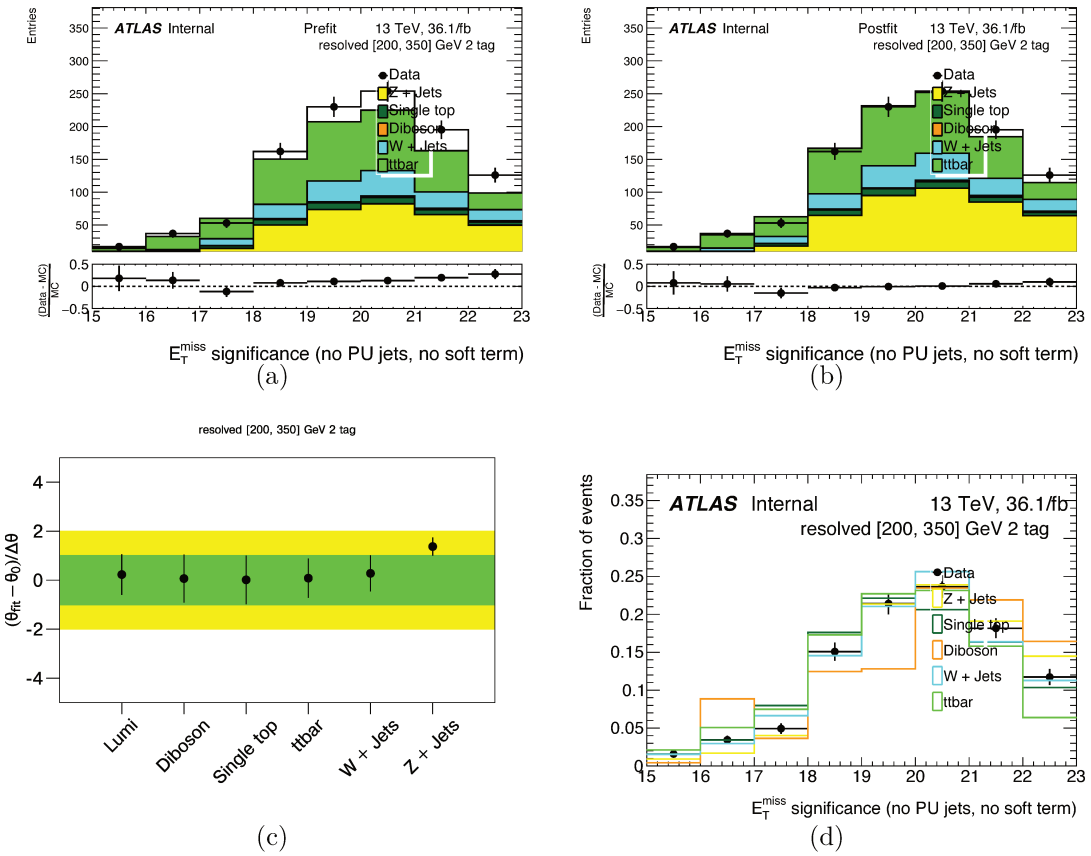


Figure G.8 – Fit results for the 200 GeV $< E_T^{miss} <$ 350 GeV, 2-btag region (a) Pre-fit distribution (b) Post-fit distribution (c) Pulls and constraints for background normalisations (d) Comparison of shapes for different simulated backgrounds

	150 GeV - 200 GeV 1b	150 GeV - 200 GeV 2b	200 GeV - 350 GeV 1b	200 GeV - 300 GeV 2b
Z + jets	1.08	1.41	1.30	1.29
single t quark	1.03	1.01	1.02	1.00
diboson	1.04	1.01	1.05	1.02
W + jets	1.47	1.15	1.12	1.06
$t\bar{t}$	1.10	1.06	1.09	1.01

Table G.3 – Background scale factors for different regions obtained with simplified profile likelihood fit, regions are fitted independently from each other.

After applying the scale factors to the background in each respective region, for the full object based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance distribution, the difference between data and simulated backgrounds is computed in the blinded signal region. Taking the integral over this distribution gives the respective multijet contribution in the blinded signal region. The distributions are shown in Figure G.9.

To compute the normalisation factor, the event yield of the blinded mass templates obtained from the multijet enriched region needs to be also computed. Thise distributions are shown in Figure G.10.

The resulting normalisation factors for the multijet templates are listed in Table G.4. The contra-intuitive value larger than 1 for 200 GeV - 350 GeV 2b is because of the large statistical uncertainty in deriving the number of multijet events in the blinded signal region. A possibility that could be investigated if required is to fit the difference between data and simulated backgrounds in the blinded signal region with a function (e.g. Chi2 function as natural functional form of a object-based $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ significance variable) and compute the integral of the function to obtain a more reliable event yield estimate.

To account for the rough method of estimating the template normalisation, the templates are used with a large normalisation uncertainty in the final statistical interpretation of the analysis result. This method provides a estimation of the multijet background normalisation and a template shape for the m_{jj} distribution in the resolved signal region. These are the inputs to the final combined fit with a 100% normalisation uncertainty.

region	150 GeV - 200 GeV 1b	150 GeV - 200 GeV 2b	200 GeV - 350 GeV 1b	200 GeV - 350 GeV 2b
template norm factor	0.39	0.42	0.87	1.27

Table G.4 – Normalisation factors for multijet templates

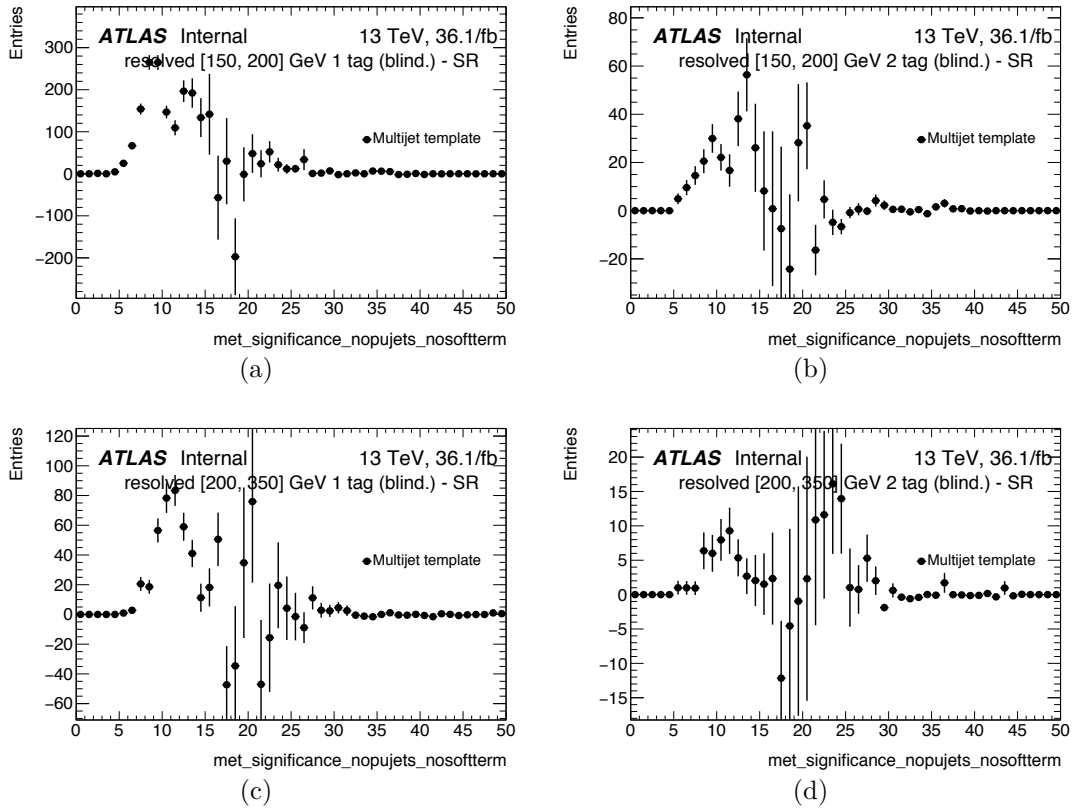


Figure G.9 – Difference between data and simulated backgrounds in the blinded signal region for the object based $E_{\mathrm{T}}^{\mathrm{miss}}$ significnace distribution. (a) resolved 1-btag 150 - 200 GeV, (b) resolved 2-btag 150 - 200 GeV, (c) resolved 1-btag 200 - 350 GeV, (d) resolved 2-btag 200 - 350 GeV,

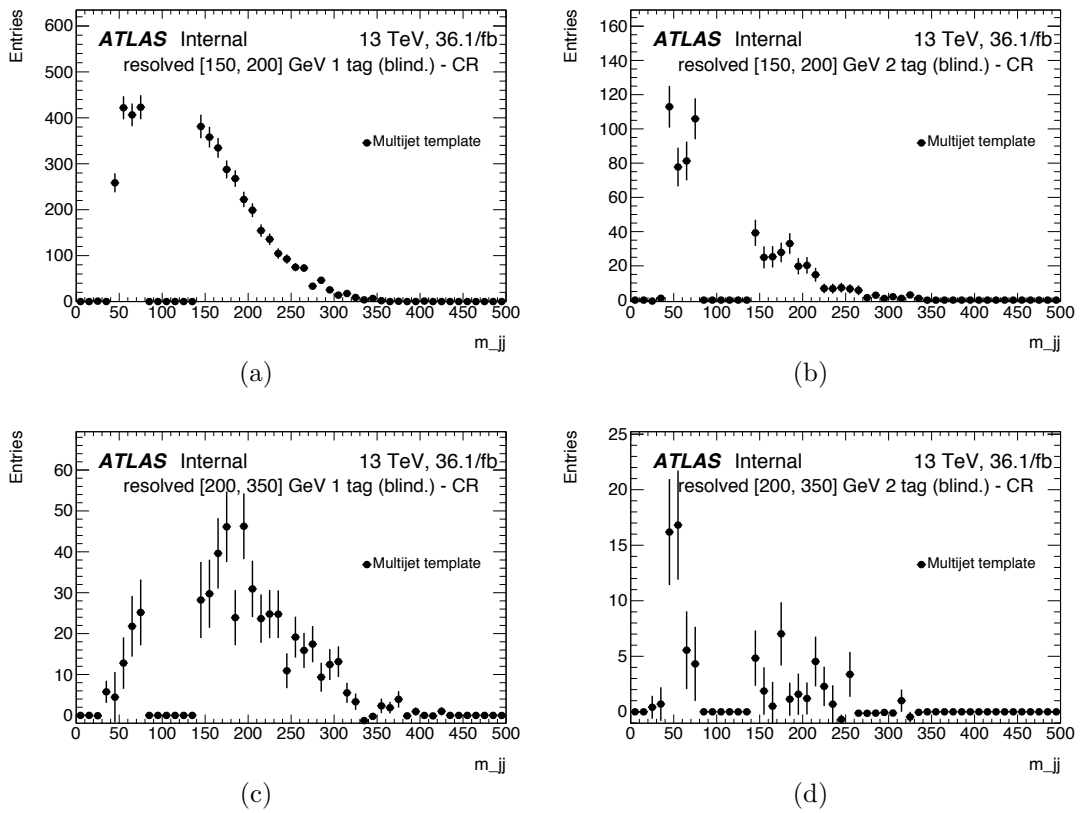


Figure G.10 – Blinded multijet templates for m_{jj} obtained from the multijet enriched region in the regions (a) resolved 1-btag 150 - 200 GeV, (b) resolved 2-btag 150 - 200 GeV, (c) resolved 1-btag 200 - 350 GeV, (d) resolved 2-btag 200 - 350 GeV,

List of figures

List of tables

Bibliography

[1] Morad Aaboud et al. Search for dark matter in association with a Higgs boson decaying to b-quarks in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Lett.*, B765:11–31, 2017. arXiv:1609.04572, doi:10.1016/j.physletb.2016.11.035.

[2] Linda Carpenter, Anthony DiFranzo, Michael Mulhearn, Chase Shimmin, Sean Tulin, and Daniel Whiteson. Mono-Higgs-boson: A new collider probe of dark matter. *Phys. Rev.*, D89(7):075017, 2014. arXiv:1312.2592, doi:10.1103/PhysRevD.89.075017.

[3] ATLAS Collaboration. 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. *Phys. Rev. Lett.*, 119:181804, 2017. arXiv:1707.01302, doi:10.1103/PhysRevLett.119.181804.

[4] Object-based missing transverse momentum significance in the ATLAS detector. Technical Report ATLAS-CONF-2018-038, CERN, Geneva, Jul 2018. URL: https://cds.cern.ch/record/2630948.

[5] Search for Dark Matter Produced in Association with a Higgs Boson decaying to $b\bar{b}$ at $\sqrt{s} = 13$ TeV with the ATLAS Detector using 79.8 fb^{-1} of proton-proton collision data. Technical Report ATLAS-CONF-2018-039, CERN, Geneva, Jul 2018. URL: http://cds.cern.ch/record/2632344.

[6] C. P. Burgess and G. D. Moore. *The standard model: A primer*. Cambridge University Press, 2006.

[7] Albert Einstein and Zugangsinformation. Grundlage der allgemeinen relativitätstheorie. *Annalen der Physik, Vol. 40 (1916), pp. 284-337*, 1916.

[8] H. Fritzsch, Murray Gell-Mann, and H. Leutwyler. Advantages of the Color Octet Gluon Picture. *Phys. Lett.*, 47B:365–368, 1973. doi:10.1016/0370-2693(73)90625-4.

[9] David J. Gross and Frank Wilczek. Ultraviolet Behavior of Nonabelian Gauge Theories. *Phys. Rev. Lett.*, 30:1343–1346, 1973. [,271(1973)]. doi:10.1103/PhysRevLett.30.1343.

[10] Steven Weinberg. Nonabelian Gauge Theories of the Strong Interactions. *Phys. Rev. Lett.*, 31:494–497, 1973. doi:10.1103/PhysRevLett.31.494.

[11] S. L. Glashow. Partial Symmetries of Weak Interactions. *Nucl. Phys.*, 22:579–588, 1961. doi:10.1016/0029-5582(61)90469-2.

[12] Steven Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19:1264–1266, 1967. doi:10.1103/PhysRevLett.19.1264.

[13] Abdus Salam. Weak and Electromagnetic Interactions. *Conf. Proc.*, C680519:367–377, 1968.

[14] Steven Weinberg. Elementary particle theory of composite particles. *Phys. Rev.*, 130:776–783, 1963. doi:10.1103/PhysRev.130.776.

[15] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global Conservation Laws and Massless Particles. *Phys. Rev. Lett.*, 13:585–587, 1964. [,162(1964)]. doi:10.1103/PhysRevLett.13.585.

[16] F. Englert and R. Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Phys. Rev. Lett.*, 13:321–323, 1964. [,157(1964)]. doi:10.1103/PhysRevLett.13.321.

[17] Peter W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13:508–509, 1964. [,160(1964)]. doi:10.1103/PhysRevLett.13.508.

[18] Jeffrey Goldstone, Abdus Salam, and Steven Weinberg. Broken Symmetries. *Phys. Rev.*, 127:965–970, 1962. doi:10.1103/PhysRev.127.965.

[19] John Ellis. Outstanding questions: Physics beyond the Standard Model. *Phil. Trans. Roy. Soc. Lond.*, A370:818–830, 2012. doi:10.1098/rsta.2011.0452.

[20] 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. *Phys.Lett.*, B716:1–29, 2012. arXiv:1207.7214, doi:10.1016/j.physletb.2012.08.020.

[21] Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys.Lett.*, B716:30–61, 2012. arXiv:1207.7235, doi:10.1016/j.physletb.2012.08.021.

[22] Georges Aad et al. Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments. *Phys. Rev. Lett.*, 114:191803, 2015. URL: https://cds.cern.ch/record/2004386, arXiv:1503.07589, doi:10.1103/PhysRevLett.114.191803.

[23] D. de Florian et al. Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector. 2016. arXiv:1610.07922, doi:10.23731/CYRM-2017-002.

[24] Morad Aaboud et al. Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector. 2018. arXiv:1808.08238.

[25] Albert M Sirunyan et al. Observation of Higgs boson decay to bottom quarks. *Submitted to: Phys. Rev. Lett.*, 2018. arXiv:1808.08242.

[26] H. Georgi and S. L. Glashow. Unity of All Elementary Particle Forces. *Phys. Rev. Lett.*, 32:438–441, 1974. doi:10.1103/PhysRevLett.32.438.

[27] H. Georgi, Helen R. Quinn, and Steven Weinberg. Hierarchy of Interactions in Unified Gauge Theories. *Phys. Rev. Lett.*, 33:451–454, 1974. doi:10.1103/PhysRevLett.33.451.

[28] Paul Langacker. Grand Unified Theories and Proton Decay. *Phys. Rept.*, 72:185, 1981. doi:10.1016/0370-1573(81)90059-4.

[29] Howard E. Haber. Introductory low-energy supersymmetry. In *Proceedings, Theoretical Advanced Study Institute (TASI 92): From Black Holes and Strings to Particles: Boulder, USA, June 1-26, 1992*, pages 589–686, 1993. arXiv:hep-ph/9306207.

[30] Howard E. Haber and Gordon L. Kane. The Search for Supersymmetry: Probing Physics Beyond the Standard Model. *Phys. Rept.*, 117:75–263, 1985. doi:10.1016/0370-1573(85)90051-1.

[31] Hans Peter Nilles. Supersymmetry, Supergravity and Particle Physics. *Phys. Rept.*, 110:1–162, 1984. doi:10.1016/0370-1573(84)90008-5.

[32] P. Van Nieuwenhuizen. Supergravity. *Phys. Rept.*, 68:189–398, 1981. doi:10.1016/0370-1573(81)90157-5.

[33] Nima Arkani-Hamed, Savas Dimopoulos, and G. R. Dvali. Phenomenology, astrophysics and cosmology of theories with submillimeter dimensions and TeV scale quantum gravity. *Phys. Rev.*, D59:086004, 1999. arXiv:hep-ph/9807344, doi:10.1103/PhysRevD.59.086004.

[34] Nima Arkani-Hamed, Savas Dimopoulos, and G. R. Dvali. The Hierarchy problem and new dimensions at a millimeter. *Phys. Lett.*, B429:263–272, 1998. arXiv:hep-ph/9803315, doi:10.1016/S0370-2693(98)00466-3.

[35] D. Bailin and A. Love. kaluza-Klein Theories. *Rept. Prog. Phys.*, 50:1087–1170, 1987. doi:10.1088/0034-4885/50/9/001.

[36] Michael B. Green. Unification of forces and particles in superstring theories. *Nature*, 314:409 EP –, 04 1985. URL: http://dx.doi.org/10.1038/314409a0.

[37] John F. Donoghue. The Fine-tuning problems of particle physics and anthropic mechanisms. pages 231–246, 2007. arXiv:0710.4080.

[38] Erminia Calabrese et al. Cosmological Parameters from pre-Planck CMB Measurements: a 2017 Update. *Phys. Rev.*, D95(6):063525, 2017. arXiv:1702.03272, doi:10.1103/PhysRevD.95.063525.

[39] Joshua Frieman, Michael Turner, and Dragan Huterer. Dark Energy and the Accelerating Universe. *Ann. Rev. Astron. Astrophys.*, 46:385–432, 2008. arXiv:0803.0982, doi:10.1146/annurev.astro.46.060407.145243.

[40] S. Perlmutter et al. Cosmology from Type Ia supernovae. *Bull. Am. Astron. Soc.*, 29:1351, 1997. arXiv:astro-ph/9812473.

[41] Beth A. Reid et al. Cosmological Constraints from the Clustering of the Sloan Digital Sky Survey DR7 Luminous Red Galaxies. *Mon. Not. Roy. Astron. Soc.*, 404:60–85, 2010. arXiv:0907.1659, doi:10.1111/j.1365-2966.2010.16276.x.

[42] Lauren Anderson et al. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: baryon acoustic oscillations in the Data Releases 10 and 11 Galaxy samples. *Mon. Not. Roy. Astron. Soc.*, 441(1):24–62, 2014. arXiv:1312.4877, doi:10.1093/mnras/stu523.

[43] J. L. Weiland et al. Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Planets and Celestial Calibration Sources. *Astrophys. J. Suppl.*, 192:19, 2011. arXiv:1001.4731, doi:10.1088/0067-0049/192/2/19.

[44] M. Kowalski et al. Improved Cosmological Constraints from New, Old and Combined Supernova Datasets. *Astrophys. J.*, 686:749–778, 2008. arXiv:0804.4142, doi:10.1086/589937.

[45] Gianfranco Bertone, Dan Hooper, and Joseph Silk. Particle dark matter: Evidence, candidates and constraints. *Phys.Rept.*, 405:279–390, 2005. arXiv:hep-ph/0404175, doi:10.1016/j.physrep.2004.08.031.

[46] F. Zwicky. On the Masses of Nebulae and of Clusters of Nebulae. *Astrophys. J.*, 86:217–246, 1937. doi:10.1086/143864.

[47] K.G. Begeman. H I rotation curves of spiral galaxies. I - NGC 3198. *Astron.Astrophys.*, 223:47–60, 1989.

[48] R.H. Sanders and M.A.W. Verheijen. Rotation curves of uma galaxies in the context of modified newtonian dynamics. *Astrophys.J.*, 503:97, 1998. arXiv:astro-ph/9802240, doi:10.1086/305986.

[49] V. C. Rubin, N. Thonnard, and W. K. Ford, Jr. Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 /R = 4kpc/ to UGC 2885 /R = 122 kpc/. *Astrophys. J.*, 238:471, 1980. doi:10.1086/158003.

[50] Douglas Clowe, Marusa Bradac, Anthony H. Gonzalez, Maxim Markevitch, Scott W. Randall, et al. A direct empirical proof of the existence of dark matter. *Astrophys.J.*, 648:L109–L113, 2006. arXiv:astro-ph/0608407, doi:10.1086/508162.

[51] Myungkook James Jee, H.C. Ford, G.D. Illingworth, R.L. White, T.J. Broadhurst, et al. Discovery of a Ringlike Dark Matter Structure in the Core of the Galaxy Cluster Cl 0024+17. *Astrophys.J.*, 661:728–749, 2007. arXiv:0705.2171, doi:10.1086/517498.

[52] C. Alcock et al. The MACHO project: Microlensing results from 5.7 years of LMC observations. *Astrophys.J.*, 542:281–307, 2000. arXiv:astro-ph/0001272, doi:10.1086/309512.

[53] P. Tisserand et al. Limits on the Macho Content of the Galactic Halo from the EROS-2 Survey of the Magellanic Clouds. *Astron.Astrophys.*, 469:387–404, 2007. arXiv:astro-ph/0607207, doi:10.1051/0004-6361:20066017.

[54] George F. Smoot, C.L. Bennett, A. Kogut, E.L. Wright, J. Aymon, et al. Structure in the COBE differential microwave radiometer first year maps. *Astrophys.J.*, 396:L1–L5, 1992. doi:10.1086/186504.

[55] D. Larson, J. Dunkley, G. Hinshaw, E. Komatsu, M.R. Nolta, et al. Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Power Spectra and WMAP-Derived Parameters. *Astrophys.J.Suppl.*, 192:16, 2011. arXiv:1001.4635, doi:10.1088/0067-0049/192/2/16.

[56] Nilendra G. Deshpande and Ernest Ma. Pattern of Symmetry Breaking with Two Higgs Doublets. *Phys.Rev.*, D18:2574, 1978. doi:10.1103/PhysRevD.18.2574.

[57] Laura Covi, Hang-Bae Kim, Jihn E. Kim, and Leszek Roszkowski. Axinos as dark matter. *JHEP*, 0105:033, 2001. arXiv:hep-ph/0101009, doi:10.1088/1126-6708/2001/05/033.

[58] Geraldine Servant and Timothy M.P. Tait. Is the lightest Kaluza-Klein particle a viable dark matter candidate? *Nucl.Phys.*, B650:391–419, 2003. arXiv:hep-ph/0206071, doi:10.1016/S0550-3213(02)01012-X.

[59] D.S. Akerib et al. First results from the LUX dark matter experiment at the Sanford Underground Research Facility. *Phys.Rev.Lett.*, 112:091303, 2014. arXiv:1310.8214, doi:10.1103/PhysRevLett.112.091303.

[60] S.E.A. Orrigo. Direct Dark Matter Search with XENON100. 2015. arXiv:1501.03492.

[61] Andrea Albert. Indirect Searches for Dark Matter with the Fermi Large Area Telescope 1. *Phys.Procedia*, 61:6–12, 2015. doi:10.1016/j.phpro.2014.12.004.

[62] Claudio Kopper. Searches for Dark Matter and High-Energy Neutrinos with IceCube. pages 189–194, 2014.

[63] K. Choi et al. Search for neutrinos from annihilation of captured low-mass dark matter particles in the Sun by Super-Kamiokande. *Phys.Rev.Lett.*, 114(14):141301, 2015. arXiv:1503.04858, doi:10.1103/PhysRevLett.114.141301.

[64] M. Aguilar et al. First Result from the Alpha Magnetic Spectrometer on the International Space Station: Precision Measurement of the Positron Fraction in Primary Cosmic Rays of 0.5–350 GeV. *Phys.Rev.Lett.*, 110:141102, 2013. doi:10.1103/PhysRevLett.110.141102.

[65] Piergiorgio Picozza et al. The Physics of Pamela Space Mission. 2:19–22, 2007.

[66] A. Abramowski et al. Constraints on an Annihilation Signal from a Core of Constant Dark Matter Density around the Milky Way Center with H.E.S.S. *Phys.Rev.Lett.*, 114(8):081301, 2015. arXiv:1502.03244, doi:10.1103/PhysRevLett.114.081301.

[67] Georges Aad et al. Search for dark matter candidates and large extra dimensions in events with a jet and missing transverse momentum with the ATLAS detector. *JHEP*, 1304:075, 2013. arXiv:1210.4491, doi:10.1007/JHEP04(2013)075.

[68] Serguei Chatrchyan et al. Search for Dark Matter and Large Extra Dimensions in pp Collisions Yielding a Photon and Missing Transverse Energy. *Phys.Rev.Lett.*, 108:261803, 2012. arXiv:1204.0821, doi:10.1103/PhysRevLett.108.261803.

[69] Patrick J. Fox, Roni Harnik, Joachim Kopp, and Yuhsin Tsai. Lep shines light on dark matter. *Phys. Rev. D*, 84:014028, Jul 2011. URL: https://link.aps.org/doi/10.1103/PhysRevD.84.014028, doi:10.1103/PhysRevD.84.014028.

[70] Yang Bai, Patrick J. Fox, and Roni Harnik. The Tevatron at the Frontier of Dark Matter Direct Detection. *JHEP*, 12:048, 2010. arXiv:1005.3797, doi:10.1007/JHEP12(2010)048.

[71] John Ellis, D.V. Nanopoulos, L Roszkowski, and D.N. Schramm. Dark matter in the light of lep. 245:251–257, 08 1990.

[72] Oliver Buchmueller, Caterina Doglioni, and Lian-Tao Wang. Search for dark matter at colliders. *Nature Physics*, 13:217 EP –, 03 2017. URL: http://dx.doi.org/10.1038/nphys4054.

[73] Mark Srednicki, Richard Watkins, and Keith A. Olive. Calculations of Relic Densities in the Early Universe. *Nucl. Phys.*, B310:693, 1988. [,247(1988)]. doi:10.1016/0550-3213(88)90099-5.

[74] Stephen P. Martin. A Supersymmetry primer. pages 1–98, 1997. [Adv. Ser. Direct. High Energy Phys.18,1(1998)]. arXiv:hep-ph/9709356, doi:10.1142/9789812839657_0001,10.1142/9789814307505_0001.

[75] Jalal Abdallah et al. Simplified Models for Dark Matter Searches at the LHC. *Phys. Dark Univ.*, 9-10:8–23, 2015. arXiv:1506.03116, doi:10.1016/j.dark.2015.08.001.

[76] Jessica Goodman, Masahiro Ibe, Arvind Rajaraman, William Shepherd, Tim M. P. Tait, and Hai-Bo Yu. Constraints on Dark Matter from Colliders. *Phys. Rev.*, D82:116010, 2010. arXiv:1008.1783, doi:10.1103/PhysRevD.82.116010.

[77] Arvind Rajaraman, William Shepherd, Tim M. P. Tait, and Alexander M. Wijangco. LHC Bounds on Interactions of Dark Matter. *Phys. Rev.*, D84:095013, 2011. arXiv:1108.1196, doi:10.1103/PhysRevD.84.095013.

[78] Jessica Goodman, Masahiro Ibe, Arvind Rajaraman, William Shepherd, Tim M. P. Tait, and Hai-Bo Yu. Constraints on Light Majorana dark Matter from Colliders. *Phys. Lett.*, B695:185–188, 2011. arXiv:1005.1286, doi:10.1016/j.physletb.2010.11.009.

[79] Qing-Hong Cao, Chuan-Ren Chen, Chong Sheng Li, and Hao Zhang. Effective Dark Matter Model: Relic density, CDMS II, Fermi LAT and LHC. *JHEP*, 08:018, 2011. arXiv:0912.4511, doi:10.1007/JHEP08(2011)018.

[80] Patrick J. Fox, Roni Harnik, Joachim Kopp, and Yuhsin Tsai. Missing Energy Signatures of Dark Matter at the LHC. *Phys. Rev.*, D85:056011, 2012. arXiv:1109.4398, doi:10.1103/PhysRevD.85.056011.

[81] Giorgio Busoni, Andrea De Simone, Enrico Morgante, and Antonio Riotto. On the Validity of the Effective Field Theory for Dark Matter Searches at the LHC. *Phys. Lett.*, B728:412–421, 2014. arXiv:1307.2253, doi:10.1016/j.physletb.2013.11.069.

[82] Herbi Dreiner, Daniel Schmeier, and Jamie Tattersall. Contact Interactions Probe Effective Dark Matter Models at the LHC. *EPL*, 102(5):51001, 2013. arXiv:1303.3348, doi:10.1209/0295-5075/102/51001.

[83] O. Buchmueller, Matthew J. Dolan, and Christopher McCabe. Beyond Effective Field Theory for Dark Matter Searches at the LHC. *JHEP*, 01:025, 2014. arXiv:1308.6799, doi:10.1007/JHEP01(2014)025.

[84] Jalal Abdallah et al. Simplified Models for Dark Matter and Missing Energy Searches at the LHC. 2014. arXiv:1409.2893.

[85] Philip Bechtle, Tilman Plehn, and Christian Sander. Supersymmetry. In Thomas Schorner-Sadenius, editor, *The Large Hadron Collider: Harvest of Run 1*, pages 421–462. 2015. arXiv:1506.03091, doi:10.1007/978-3-319-15001-7_10.

[86] Martin Schmaltz and David Tucker-Smith. Little Higgs review. *Ann. Rev. Nucl. Part. Sci.*, 55:229–270, 2005. arXiv:hep-ph/0502182, doi:10.1146/annurev.nucl.55.090704.151502.

[87] G. D'Ambrosio, G. F. Giudice, G. Isidori, and A. Strumia. Minimal flavor violation: An Effective field theory approach. *Nucl. Phys.*, B645:155–187, 2002. arXiv:hep-ph/0207036, doi:10.1016/S0550-3213(02)00836-2.

[88] Mads T. Frandsen, Felix Kahlhoefer, Anthony Preston, Subir Sarkar, and Kai Schmidt-Hoberg. LHC and Tevatron Bounds on the Dark Matter Direct Detection Cross-Section for Vector Mediators. *JHEP*, 07:123, 2012. arXiv:1204.3839, doi:10.1007/JHEP07(2012)123.

[89] Daniel Abercrombie et al. Dark Matter Benchmark Models for Early LHC Run-2 Searches: Report of the ATLAS/CMS Dark Matter Forum. 2015. arXiv:1507.00966.

[90] E. Dudas, Y. Mambrini, S. Pokorski, and A. Romagnoni. (In)visible Z-prime and dark matter. *JHEP*, 08:014, 2009. arXiv:0904.1745, doi:10.1088/1126-6708/2009/08/014.

[91] Gino Isidori, Yosef Nir, and Gilad Perez. Flavor Physics Constraints for Physics Beyond the Standard Model. *Ann. Rev. Nucl. Part. Sci.*, 60:355, 2010. arXiv:1002.0900, doi:10.1146/annurev.nucl.012809.104534.

[92] Felix Kahlhoefer. Review of LHC Dark Matter Searches. *Int. J. Mod. Phys.*, A32(13):1730006, 2017. arXiv:1702.02430, doi:10.1142/S0217751X1730006X.

[93] Andreas Albert et al. Recommendations of the LHC Dark Matter Working Group: Comparing LHC searches for heavy mediators of dark matter production in visible and invisible decay channels. 2017. arXiv:1703.05703.

[94] Asher Berlin, Tongyan Lin, and Lian-Tao Wang. Mono-Higgs Detection of Dark Matter at the LHC. *JHEP*, 06:078, 2014. arXiv:1402.7074, doi:10.1007/JHEP06(2014)078.

[95] G.C. Branco, P.M. Ferreira, L. Lavoura, M.N. Rebelo, Marc Sher, et al. Theory and phenomenology of two-Higgs-doublet models. *Phys.Rept.*, 516:1–102, 2012. arXiv:1106.0034, doi:10.1016/j.physrep.2012.02.002.

[96] Marcela Carena, Alejandro Daleo, Bogdan A. Dobrescu, and Timothy M. P. Tait. Z' gauge bosons at the Tevatron. *Phys. Rev.*, D70:093009, 2004. arXiv:hep-ph/0408098, doi:10.1103/PhysRevD.70.093009.

[97] Albert M Sirunyan et al. Search for new physics in the monophoton final state in proton-proton collisions at $\sqrt{s} = 13$ TeV. *JHEP*, 10:073, 2017. arXiv:1706.03794, doi:10.1007/JHEP10(2017)073.

[98] Morad 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. *Submitted to: JHEP*, 2018. arXiv:1807.11471.

[99] 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. *Eur. Phys. J.*, C78(1):18, 2018. arXiv:1710.11412, doi:10.1140/epjc/s10052-017-5486-1.

[100] A. M. Sirunyan et al. Search for dark matter produced in association with heavy-flavor quark pairs in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Eur. Phys. J.*, C77(12):845, 2017. arXiv:1706.02581, doi:10.1140/epjc/s10052-017-5317-4.

[101] Search for dark matter and other new phenomena in events with an energetic jet and large missing transverse momentum using the ATLAS detector. Technical Report ATLAS-CONF-2017-060, CERN, Geneva, Jul 2017. URL: http://cds.cern.ch/record/2273876.

[102] Search for new phenomena in the $Z(\to \ell\ell) + E_{\mathrm{T}}^{\mathrm{miss}}$ final state at $\sqrt{s} = 13$ TeV with the ATLAS detector. Technical Report ATLAS-CONF-2016-056, CERN, Geneva, Aug 2016. URL: https://cds.cern.ch/record/2206138.

[103] Georges Aad et al. Search for new particles in events with one lepton and missing transverse momentum in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. *JHEP*, 09:037, 2014. arXiv:1407.7494, doi:10.1007/JHEP09(2014)037.

[104] Morad Aaboud et al. Search for new phenomena in events with a photon and missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *JHEP*, 06:059, 2016. arXiv:1604.01306, doi:10.1007/JHEP06(2016)059.

[105] Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett.*, B716:30–61, 2012. arXiv:1207.7235, doi:10.1016/j.physletb.2012.08.021.

[106] J R Andersen et al. Handbook of LHC Higgs Cross Sections: 3. Higgs Properties. 2013. arXiv:1307.1347, doi:10.5170/CERN-2013-004.

[107] ATLAS Collaboration. Search for Dark Matter in Events with Missing Transverse Momentum and a Higgs Boson Decaying to Two Photons in pp Collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector. *Phys. Rev. Lett.*, 115:131801, 2015. arXiv:1506.01081, doi:10.1103/PhysRevLett.115.131801.

[108] Robert M. Harris and Konstantinos Kousouris. Searches for Dijet Resonances at Hadron Colliders. *Int. J. Mod. Phys.*, A26:5005–5055, 2011. arXiv:1110.5302, doi:10.1142/S0217751X11054905.

[109] Bogdan A. Dobrescu and Felix Yu. Coupling-mass mapping of dijet peak searches. *Phys. Rev. D*, 88:035021, Aug 2013. URL: https://link.aps.org/doi/10.1103/PhysRevD.88.035021, doi:10.1103/PhysRevD.88.035021.

[110] Morad Aaboud et al. Search for new phenomena in dijet events using 37 fb^{-1} of pp collision data collected at $\sqrt{s}$ =13 TeV with the ATLAS detector. *Phys. Rev.*, D96(5):052004, 2017. arXiv:1703.09127, doi:10.1103/PhysRevD.96.052004.

[111] Morad Aaboud et al. Search for low-mass dijet resonances using trigger-level jets with the ATLAS detector in pp collisions at $\sqrt{s} = 13$ TeV. *Phys. Rev. Lett.*, 121(8):081801, 2018. arXiv:1804.03496, doi:10.1103/PhysRevLett.121.081801.

[112] Morad Aaboud et al. Search for light resonances decaying to boosted quark pairs and produced in association with a photon or a jet in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. 2018. arXiv:1801.08769.

[113] Search for new high-mass phenomena in the dilepton final state using 36 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV with the ATLAS Detector. *JHEP*, 10(CERN-EP-2017-119):182. 42 p, Jul 2017. URL: https://cds.cern.ch/record/2273892.

[114] Morad Aaboud et al. Search for resonances in the mass distribution of jet pairs with one or two jets identified as b-jets in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. 2018. arXiv:1805.09299.

[115] M. 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. *Eur. Phys. J.*, C78(7):565, 2018. arXiv:1804.10823, doi:10.1140/epjc/s10052-018-5995-6.

[116] Mihailo Backovic, Antony Martini, Kyoungchul Kong, Olivier Mattelaer, and Gopolang Mohlabeng. MadDM: New dark matter tool in the LHC era. *AIP Conf. Proc.*, 1743:060001, 2016. arXiv:1509.03683, doi:10.1063/1.4953318.

[117] Felix Kahlhoefer, Kai Schmidt-Hoberg, Thomas Schwetz, and Stefan Vogl. Implications of unitarity and gauge invariance for simplified dark matter models. *JHEP*, 02:016, 2016. arXiv:1510.02110, doi:10.1007/JHEP02(2016)016.

[118] Giorgio Busoni et al. Recommendations on presenting LHC searches for missing transverse energy signals using simplified s-channel models of dark matter. 2016. arXiv:1603.04156.

[119] L. Evans and Ph. Bryant. LHC Machine. *JINST*, 3:S08001, 2008. doi:doi:10.1088/1748-0221/3/08/S08001.

[120] CERN web site. URL: https://home.cern.

[121] LEP Injector Study Group. LEP design report. CERN-LEP-84-01, 1984. URL: https://cds.cern.ch/record/102083.

[122] LHC Commissioning with beam. URL: https://lhc-commissioning.web.cern.ch/lhc-commissioning/schedule/LHC-long-term.htm.

[123] The CERN accelerator complex. URL: https://cds.cern.ch/record/2197559.

[124] ALICE Collaboration. The ALICE Experiment at the CERN LHC. *JINST*, 3:S08002, 2008. doi:10.1088/1748-0221/3/08/S08002.

[125] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008. doi:10.1088/1748-0221/3/08/S08003.

[126] CMS Collaboration. The CMS experiment at the CERN LHC. *JINST*, 3:S08004, 2008. doi:10.1088/1748-0221/3/08/S08004.

[127] LHCb Collaboration. The LHCb experiment at the CERN LHC. *JINST*, 3:S08005, 2008. doi:10.1088/1748-0221/3/08/S08005.

[128] LHCf Collaboration. The LHCf detector at the CERN Large Hadron Collider. *JINST*, 3:S08006, 2008. doi:10.1088/1748-0221/3/08/S08006.

[129] TOTEM Collaboration. The TOTEM Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08007, 2008. doi:10.1088/1748-0221/3/08/S08007.

[130] Luminosity public results. URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.

[131] G. Aad et al. The ATLAS Inner Detector commissioning and calibration. *Eur. Phys. J.*, C70:787–821, 2010. arXiv:1004.5293, doi:10.1140/epjc/s10052-010-1366-7.

[132] Computer generated image of the ATLAS Muons subsystem. URL: https://cds.cern.ch/record/1095929.

[133] Data quality public results. URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/RunStatsPublicResults2010.

[134] ATLAS Collaboration. Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment. *ATLAS-CONF-2012-042*, Mar 2012. URL: http://cds.cern.ch/record/1435196.

[135] Thijs G. Cornelissen, M. Elsing, I. Gavrilenko, J. F. Laporte, W. Liebig, M. Limper, K. Nikolopoulos, A. Poppleton, and A. Salzburger. The global χ^2 track fitter in ATLAS. *J. Phys. Conf. Ser.*, 119:032013, 2008. doi:10.1088/1742-6596/119/3/032013.

[136] M. Aaboud et al. Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC run 2. Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC run 2. *Eur. Phys. J. C*, 77(CERN-EP-2017-045. 10):673. 44 p, Apr 2017. URL: https://cds.cern.ch/record/2261156, arXiv:1704.07983, doi:10.1140/epjc/s10052-017-5225-7.

[137] ATLAS Collaboration. Performance of primary vertex reconstruction in proton-proton collisions at $\sqrt{s}$ =7 TeV in the ATLAS experiment. *ATLAS-CONF-2010-069*, Jul 2010. arXiv:http://cds.cern.ch/record/1281344.

[138] ATLAS Collaboration. Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1. *Eur. Phys. J.*, C77:490, 2017. URL: https://cds.cern.ch/record/2138166, arXiv:1603.02934, doi:10.1140/epjc/s10052-017-5004-5.

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

[140] Gavin P. Salam. Towards Jetography. *Eur. Phys. J.*, C67:637–686, 2010. arXiv:0906.1833, doi:10.1140/epjc/s10052-010-1314-6.

[141] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The Anti-k(t) jet clustering algorithm. *JHEP*, 04:063, 2008. arXiv:0802.1189, doi:10.1088/1126-6708/2008/04/063.

[142] W. Lampl et al. Calorimeter clustering algorithms: description and performance. *ATL-LARG-PUB-2008-002*, April 2008. URL: http://cdsweb.cern.ch/record/1099735.

[143] S. Catani, Yuri L. Dokshitzer, M. H. Seymour, and B. R. Webber. Longitudinally invariant K_t clustering algorithms for hadron hadron collisions. *Nucl. Phys.*, B406:187–224, 1993. doi:10.1016/0550-3213(93)90166-M.

[144] Yuri L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber. Better jet clustering algorithms. *JHEP*, 08:001, 1997. arXiv:hep-ph/9707323, doi:10.1088/1126-6708/1997/08/001.

[145] L Asquith, B Brelier, J M Butterworth, M Campanelli, T Carli, G Choudalakis, P A Delsart, S De Cecco, P O Deviveiros, M D'Onofrio, S Eckweiler, E Feng, P Francavilla, S Grinstein, I La Plante, J Huston, N Ghodbane, D Lopez Mateos, B Martin, N Makovec, S Majewsky, M Martinez, D W Miller, J Monk, K Perez, C Roda, J Robinson, A Schwartzmann, F Spano, K Terashi, F Vives, P Weber, and S Zenz. Performance of Jet Algorithms in the ATLAS Detector. Technical Report ATL-PHYS-INT-2010-129, CERN, Geneva, Dec 2010. URL: https://cds.cern.ch/record/1311867.

[146] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet User Manual. *Eur. Phys. J.*, C72:1896, 2012. arXiv:1111.6097, doi:10.1140/epjc/s10052-012-1896-2.

[147] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The Catchment Area of Jets. *JHEP*, 04:005, 2008. arXiv:0802.1188, doi:10.1088/1126-6708/2008/04/005.

[148] 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. *Phys. Rev.*, D96(7):072002, 2017. arXiv:1703.09665, doi:10.1103/PhysRevD.96.072002.

[149] Matteo Cacciari and Gavin P. Salam. Pileup subtraction using jet areas. *Phys. Lett. B*, 659:119–126, 2008. arXiv:0707.1378, doi:10.1016/j.physletb.2007.09.077.

[150] Georges Aad et al. Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector. *Eur. Phys. J.*, C76(11):581, 2016. arXiv:1510.03823, doi:10.1140/epjc/s10052-016-4395-z.

[151] ATLAS Collaboration. Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV. *Eur. Phys. J. C*, 73:2304, 2013. arXiv:1112.6426, doi:10.1140/epjc/s10052-013-2304-2.

[152] Tagging and suppression of pileup jets with the ATLAS detector. *ATLAS-CONF-2014-018*, May 2014. arXiv:http://cds.cern.ch/record/1700870.

[153] ATLAS Collaboration. Selection of jets produced in 13TeV proton-proton collisions with the ATLAS detector. Technical Report ATLAS-COM-CONF-2015-024, CERN, Geneva, May 2015. URL: https://cds.cern.ch/record/2016323.

[154] ATLAS Collaboration. Performance of jet substructure techniques for large-R jets in proton–proton collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector. *JHEP*, 09:076, 2013. arXiv:1306.4945, doi:10.1007/JHEP09(2013)076.

[155] D. Krohn, J. Thaler, and L.-T. Wang. Jet Trimming. *JHEP*, 02:084, 2010. arXiv:0912.1342, doi:10.1007/JHEP02(2010)084.

[156] Identification of boosted, hadronically-decaying W and Z bosons in $\sqrt{s} = 13$ TeV Monte Carlo Simulations for ATLAS. Technical Report ATL-PHYS-PUB-2015-033, CERN, Geneva, Aug 2015. URL: https://cds.cern.ch/record/2041461.

[157] ATLAS Collaboration. Jet mass reconstruction with the ATLAS Detector in early Run 2 data. ATLAS-CONF-2016-035, 2016. URL: https://cds.cern.ch/record/2200211.

[158] ATLAS Collaboration. Expected performance of the ATLAS b-tagging algorithms in Run-2. Technical Report ATL-PHYS-PUB-2015-022, CERN, Geneva, Jul 2016. URL: https://cds.cern.ch/record/2037697.

[159] ATLAS Collaboration. Commissioning of the ATLAS b-tagging algorithms using $t\bar{t}$ events in early Run-2 data. (ATL-PHYS-PUB-2015-039/), 2016. URL: http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2015-039/.

[160] Georges Aad et al. Performance of b-Jet Identification in the ATLAS Experiment. 2015. arXiv:1512.01094.

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

[162] Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run. Technical Report ATL-PHYS-PUB-2017-013, CERN, Geneva, Jul 2017. URL: https://cds.cern.ch/record/2273281.

[163] Morad Aaboud et al. Measurements of b-jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s} = 13$ TeV. 2018. arXiv:1805.01845.

[164] Boosted Higgs ($\to b\bar{b}$) Boson Identification with the ATLAS Detector at $\sqrt{s} = 13$ TeV. Technical Report ATLAS-CONF-2016-039, CERN, Geneva, Aug 2016. URL: https://cds.cern.ch/record/2206038.

[165] Variable Radius, Exclusive-k$_T$, and Center-of-Mass Subjet Reconstruction for Higgs($\to b\bar{b}$) Tagging in ATLAS. Technical Report ATL-PHYS-PUB-2017-010, CERN, Geneva, Jun 2017. URL: http://cds.cern.ch/record/2268678.

[166] Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung. Technical Report ATLAS-CONF-2012-047, CERN, Geneva, May 2012. URL: https://cds.cern.ch/record/1449796.

[167] M Delmastro, S Gleyzer, C Hengler, M Jimenez, T Koffas, M Kuna, K Liu, Y Liu, G Marchiori, E Petit, M Pitt, E Soldatov, and K Tackmann. Photon identification efficiency measurements with the ATLAS detector using LHC Run 1 data. Technical Report ATL-COM-PHYS-2014-949, CERN, Geneva, Aug 2014. URL: https://cds.cern.ch/record/1747242.

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

[169] Photon identification in 2015 ATLAS data. Technical Report ATL-PHYS-PUB-2016-014, CERN, Geneva, Aug 2016. URL: https://cds.cern.ch/record/2203125.

[170] ATLAS Collaboration. Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13$ TeV. *Eur. Phys. J. C*, 76:292, 2016. arXiv:1603.05598, doi:10.1140/epjc/s10052-016-4120-y.

[171] Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at $\sqrt{s} = 13$ TeV. Technical Report ATLAS-CONF-2017-029, CERN, Geneva, May 2017. URL: http://cds.cern.ch/record/2261772.

[172] ATLAS collaboration. Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC. Technical Report ATL-PHYS-PUB-2015-045, CERN, Geneva, Nov 2015. URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2015-045/ATL-PHYS-PUB-2015-045.pdf.

[173] Tau systematic recommendations Moriond 2017, February 2017. URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TauRecommendationsMoriond2017.

[174] ATLAS Collaboration. Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment in pp collisions at $\sqrt{s}$=8 TeV. *Eur. Phys. J. C*, 75:303, 2015. arXiv:1412.7086, doi:10.1140/epjc/s10052-015-3500-z.

[175] T Barillari, E Bergeaas Kuutmann, T Carli, J Erdmann, P Giovannini, K J Grahn, C Issever, A Jantsch, A Kiryunin, K Lohwasser, A Maslennikov, S Menke, H Oberlack, G Pospelov, E Rauter, P Schacht, F Spanó, P Speckmayer, P Stavina, and P Strízenec. Local Hadronic Calibration. Technical Report ATL-LARG-PUB-2009-001-2. ATL-COM-LARG-2008-006. ATL-LARG-PUB-2009-001, CERN, Geneva, Jun 2008. Due to a report-number conflict with another document,

the report-number ATL-LARG-PUB-2009-001-2 has been assigned. URL: https://cds.cern.ch/record/1112035.

[176] ATLAS Collaboration. Performance of Missing Transverse Momentum Reconstruction in Proton-Proton Collisions at 7 TeV with ATLAS. *Eur. Phys. J. C*, 72:1844, 2012. arXiv:1108.5602, doi:10.1140/epjc/s10052-011-1844-6.

[177] G. Aad et al. Performance of algorithms that reconstruct missing transverse momentum in $\sqrt{s} = 8$ TeV proton-proton collisions in the ATLAS detector. *Eur. Phys. J.*, C77(4):241, 2017. arXiv:1609.09324, doi:10.1140/epjc/s10052-017-4780-2.

[178] Jeremiah Jet Goodson and Robert McCarthy. Search for Supersymmetry in States with Large Missing Transverse Momentum and Three Leptons including a Z-Boson, May 2012. Presented 17 Apr 2012. URL: http://cds.cern.ch/record/1449722.

[179] Morad Aaboud et al. Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13$ TeV. 2018. arXiv:1802.08168.

[180] $E_{\mathrm{T}}^{\mathrm{miss}}$ performance in the ATLAS detector using 2015-2016 LHC p-p collisions. Technical Report ATLAS-CONF-2018-023, CERN, Geneva, Jun 2018. URL: https://cds.cern.ch/record/2625233.

[181] ATLAS Collaboration. Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$ =13 TeV. *Eur. Phys. J.*, C76(5):292, 2016. arXiv:1603.05598, doi:10.1140/epjc/s10052-016-4120-y.

[182] ATLAS Collaboration. Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data. *Eur. Phys. J. C*, 74:3071, 2014. arXiv:1407.5063, doi:10.1140/epjc/s10052-014-3071-4.

[183] Expected performance of missing transverse momentum reconstruction for the ATLAS detector at $\sqrt{s} = 13$ TeV. Technical Report ATL-PHYS-PUB-2015-023, CERN, Geneva, Jul 2015. URL: http://cds.cern.ch/record/2037700.

[184] T. Aaltonen et al. Search for Anomalous Production of Events with Two Photons and Additional Energetic Objects at CDF. *Phys. Rev.*, D82:052005, 2010. arXiv:0910.5170, doi:10.1103/PhysRevD.82.052005.

[185] The CMS collaboration. Missing transverse energy performance of the cms detector. *Journal of Instrumentation*, 6(09):P09001, 2011. URL: http://stacks.iop.org/1748-0221/6/i=09/a=P09001.

[186] Vardan Khachatryan et al. Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s} = 8$ TeV. *JINST*, 10(02):P02006, 2015. arXiv:1411.0511, doi:10.1088/1748-0221/10/02/P02006.

[187] ATLAS Collaboration. Search for top squark pair production in final states with one isolated lepton, jets, and missing transverse momentum in $\sqrt{s}$ =8 TeV pp collisions with the ATLAS detector. *JHEP*, 11:118, 2014. arXiv:1407.0583, doi:10.1007/JHEP11(2014)118.

[188] Benjamin Nachman and Christopher G. Lester. Significance Variables. *Phys. Rev.*, D88(7):075013, 2013. arXiv:1303.7009, doi:10.1103/PhysRevD.88.075013.

[189] ATLAS Collaboration. Jet energy resolution in proton-proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector. *Eur. Phys. J.*, C73(3):2306, 2013. arXiv:1210.6210, doi:10.1140/epjc/s10052-013-2306-0.

[190] ATLAS Collaboration. Monte Carlo Calibration and Combination of In-situ Measurements of Jet Energy Scale, Jet Energy Resolution and Jet Mass in ATLAS. *ATLAS-CONF-2015-037*, 2015. URL: http://cds.cern.ch/record/2044941.

[191] ATLAS Collaboration. Determination of the jet energy scale and resolution at ATLAS using Z/γ-jet events in data at $\sqrt{s} = 8$ TeV. ATLAS-CONF-2015-057, 2015. URL: https://cds.cern.ch/record/2059846.

[192] 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. *ATLAS-CONF-2015-017*, Apr 2015. URL: http://cds.cern.ch/record/2008678.

[193] Morad Aaboud et al. Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC. *Eur. Phys. J.*, C76(12):653, 2016. arXiv:1608.03953, doi:10.1140/epjc/s10052-016-4466-1.

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

[195] D Costanzo, A Dell'Acqua, M Gallas, A Nairz, N Benekos, A Rimoldi, J Boudreau, and V Tsulaia. Validation of the GEANT4-Based Full Simulation Program for the ATLAS Detector: An Overview of Performance and Robustness. Technical Report ATL-SOFT-PUB-2005-002. ATL-COM-SOFT-2005-004. CERN-ATL-SOFT-PUB-2005-002, CERN, Geneva, Mar 2005. URL: https://cds.cern.ch/record/830149.

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

[197] ATLAS Collaboration. Summary of ATLAS Pythia 8 tunes. (ATL-PHYS-PUB-2012-003), Aug 2012. arXiv:http://cds.cern.ch/record/1474107.

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

[199] T. Gleisberg, Stefan. Höche, F. Krauss, M. Schönherr, S. Schumann, et al. Event generation with SHERPA 1.1. *JHEP*, 02:007, 2009. arXiv:0811.4622, doi:10.1088/1126-6708/2009/02/007.

[200] D. J. Lange. The EvtGen particle decay simulation package. *Nucl. Instrum. Meth.*, A462:152–155, 2001. doi:10.1016/S0168-9002(01)00089-4.

[201] Tanju Gleisberg and Stefan Höche. Comix, a new matrix element generator. *JHEP*, 12:039, 2008. arXiv:0808.3674, doi:10.1088/1126-6708/2008/12/039.

[202] Fabio Cascioli, Philipp Maierhofer, and Stefano Pozzorini. Scattering Amplitudes with Open Loops. *Phys. Rev. Lett.*, 108:111601, 2012. arXiv:1111.5206, doi:10.1103/PhysRevLett.108.111601.

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

[204] Stefan Höche, Frank Krauss, Marek Schönherr, and Frank Siegert. QCD matrix elements + parton showers: The NLO case. *JHEP*, 04:027, 2013. arXiv:1207.5030, doi:10.1007/JHEP04(2013)027.

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

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

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

[208] Michal Czakon and Alexander Mitov. Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders. *Comput. Phys. Commun.*, 185:2930, 2014. arXiv:1112.5675, doi:10.1016/j.cpc.2014.06.021.

[209] Nikolaos Kidonakis. NNLL resummation for s-channel single top quark production. *Phys. Rev.*, D81:054028, 2010. arXiv:1001.5034, doi:10.1103/PhysRevD.81.054028.

[210] Nikolaos Kidonakis. Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production. *Phys. Rev.*, D83:091503, 2011. arXiv:1103.2792, doi:10.1103/PhysRevD.83.091503.

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

[212] Particle Data Group. Review of Particle Physics. *Chin. Phys.*, C40(10):100001, 2016. doi:10.1088/1674-1137/40/10/100001.

[213] ATLAS Collaboration. Search for dark matter produced in association with a Higgs boson decaying to two bottom quarks in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. *Phys. Rev. D*, 93(7):072007, 2016. arXiv:1510.06218, doi:10.1103/PhysRevD.93.072007.

[214] Georges Aad et al. Search for the $b\bar{b}$ decay of the Standard Model Higgs boson in associated $(W/Z)H$ production with the ATLAS detector. *JHEP*, 01:069, 2015. arXiv:1409.6212, doi:10.1007/JHEP01(2015)069.

[215] Serguei Chatrchyan et al. Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s} = 7$ and 8 TeV. *JHEP*, 06:081, 2013. arXiv:1303.4571, doi:10.1007/JHEP06(2013)081.

[216] Jonathan M. Butterworth, Adam R. Davison, Mathieu Rubin, and Gavin P. Salam. Jet substructure as a new Higgs search channel at the LHC. In *Proceedings, 34th International Conference on High Energy Physics (ICHEP 2008): Philadelphia, Pennsylvania, July 30-August 5, 2008*, 2008. arXiv:0810.0409.

[217] Mrinal Dasgupta, Alessandro Fregoso, Simone Marzani, and Gavin P. Salam. Towards an understanding of jet substructure. *JHEP*, 09:029, 2013. arXiv:1307.0007, doi:10.1007/JHEP09(2013)029.

[218] Lorenzo Moneta, Kevin Belasco, Kyle S. Cranmer, S. Kreiss, Alfio Lazzaro, Danilo Piparo, Gregory Schott, Wouter Verkerke, and Matthias Wolf. The RooStats Project. *PoS*, ACAT2010:057, 2010. arXiv:1009.1003, doi:10.22323/1.093.0057.

[219] Wouter Verkerke and David P. Kirkby. The RooFit toolkit for data modeling. *eConf*, C0303241:MOLT007, 2003. [,186(2003)]. arXiv:physics/0306116.

[220] Calibration of the performance of b-tagging for c and light-flavour jets in the 2012 ATLAS data. Technical Report ATLAS-CONF-2014-046, CERN, Geneva, Jul 2014. URL: https://cds.cern.ch/record/1741020.

[221] David Krohn, Jesse Thaler, and Lian-Tao Wang. Jet Trimming. *JHEP*, 02:084, 2010. arXiv:0912.1342, doi:10.1007/JHEP02(2010)084.

[222] Georges Aad et al. Jet mass and substructure of inclusive jets in $\sqrt{s} = 7$ TeV pp collisions with the ATLAS experiment. *JHEP*, 05:128, 2012. arXiv:1203.4606, doi:10.1007/JHEP05(2012)128.

[223] Adrian Buzatu and Wei Wang. Object selections for SM Higgs boson produced in association with a vector boson in which $H \to b\bar{b}$ and V decays leptonically with Run-2 data: Object support note for VH(bb) 2015+2016 dataset publication. Technical Report ATL-COM-PHYS-2016-1674, CERN, Geneva, Nov 2016. This is

a support note for the VH(bb) SM publication using the 2015+2016 datasets. URL: https://cds.cern.ch/record/2233686.

[224] Adrian Buzatu. Decoupled b-jet-energy corrections (muon-in-jet, neutrino-in-jet, and transverse momentum) for $H \to b\bar{b}$ searches for ATLAS in early Run II. Technical Report ATL-COM-PHYS-2016-035, CERN, Geneva, Jan 2016. URL: https://cds.cern.ch/record/2124442.

[225] D. Adams et al. Recommendations of the Physics Objects and Analysis Harmonisation Study Groups 2014. 2014.

[226] Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data. Technical Report ATLAS-CONF-2014-032, CERN, Geneva, Jun 2014. URL: http://cds.cern.ch/record/1706245.

[227] ATLAS Collaboration. Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton-proton collision data. *Eur. Phys. J. C*, 74:2941, 2014. arXiv:1404.2240, doi:10.1140/epjc/s10052-014-2941-0.

[228] Isolation Recommendations for ICHEP 2016. Technical report, CERN, Geneva, May 2016. URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/IsolationSF2016DS1.

[229] Tracking CP Recommendations for ICHEP 2016. Technical report, CERN, Geneva, May 2016. URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TrackingCPICHEP2016.

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

[231] C Wang, S Wang, and L Zhang. Search for a high mass Higgs boson in the $H \to ZZ \to \nu\nu qq$ decay channel with the ATLAS Detector. Technical Report ATL-COM-PHYS-2014-262, CERN, Geneva, Apr 2014. URL: https://cds.cern.ch/record/1692942.

[232] Morad Aaboud et al. Performance of the ATLAS Trigger System in 2015. *Eur. Phys. J.*, C77(5):317, 2017. arXiv:1611.09661, doi:10.1140/epjc/s10052-017-4852-3.

[233] Samuel Meehan, Rainer Rohrig, Patrick Rieck, Andrew Mehta, Carl Gwilliam, Paul Thompson, Laser Seymour Kaplan, Chen Zhou, Dilia Maria Portillo Quintero, Oleg Brandt, Daniel Isaac Narrias Villar, Stanislav Suchek, Efe Yigitbasi, John Butler, Sau Lan Wu, Sandra Kortner, Sandro De Cecco, Pai-hsien Jennifer Hsu, Yuan-tang Chou, Daniel Guest, Arturos Sanchez Pineda, Shih-Chieh Hsu, Nikola Lazar Whallon, Yun-ju Lu, Jike Wang, Veronica Fabiani, Russell James Turner, Konstantinos

Nikolopoulos, and Manuel Silva Jr. Search for Dark Matter Produced in Association with a Higgs Boson Decaying to $b\bar{b}$ at $\sqrt{s} = 13$ TeV with the ATLAS Detector. Technical Report ATL-COM-PHYS-2016-1484, CERN, Geneva, Oct 2016. URL: https://cds.cern.ch/record/2225941.

[234] Hung-Liang Lai, Marco Guzzi, Joey Huston, Zhao Li, Pavel M. Nadolsky, Jon Pumplin, and C.-P. Yuan. New parton distributions for collider physics. *Phys. Rev. D*, 82:074024, Oct 2010. URL: http://link.aps.org/doi/10.1103/PhysRevD.82.074024, doi:10.1103/PhysRevD.82.074024.

[235] Richard D. Ball et al. Parton distributions for the LHC Run II. *JHEP*, 04:040, 2015. arXiv:1410.8849, doi:10.1007/JHEP04(2015)040.

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

[237] Richard D. Ball et al. Parton distributions with LHC data. *Nucl. Phys.*, B867:244–289, 2013. arXiv:1207.1303, doi:10.1016/j.nuclphysb.2012.10.003.

[238] Torbjorn Sjöstrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 05:026, 2006. arXiv:hep-ph/0603175, doi:10.1088/1126-6708/2006/05/026.

[239] ATLAS-CMS recommended predictions for single-top cross sections using the Hathor v2.1 program, June 2016. URL: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/SingleTopRefXsec#Predictions_at_7_8_13_and_14_TeV.

[240] Georges Aad et al. Reconstruction of hadronic decay products of tau leptons with the ATLAS experiment. *Eur. Phys. J.*, C76(5):295, 2016. arXiv:1512.05955, doi:10.1140/epjc/s10052-016-4110-0.

[241] L. Moneta, K. Cranmer, G. Schott, and W. Verkerke. The RooStats project. In *Proceedings of the 13th International Workshop on Advanced Computing and Analysis Techniques in Physics Research. February 22-27, 2010, Jaipur, India. http://acat2010.cern.ch/.*, page 57, 2010. URL: http://adsabs.harvard.edu/abs/2010acat.confE..57M, arXiv:1009.1003.

[242] W. Verkerke and D. Kirkby. The RooFit toolkit for data modeling. *ArXiv Physics e-prints*, June 2003. arXiv:physics/0306116.

[243] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J. C*, 71:1554, 2011. arXiv:1007.1727, doi:10.1140/epjc/s10052-011-1554-0.

[244] Alexander L. Read. Presentation of search results: The CL(s) technique. *J. Phys.*, G28:2693–2704, 2002. [,11(2002)]. doi:10.1088/0954-3899/28/10/313.

[245] Carlo Enrico Pandini and Paul Thompson. Theoretical studies for the 2HDM $A \to Zh \to, h \to b\bar{b}$ search and the SM $Vh \to b\bar{b}$ search
- Supporting Document. Technical Report ATL-COM-PHYS-2015-1474, CERN, Geneva, Dec 2015. URL: https://cds.cern.ch/record/2111370.

[246] Paul Thompson and Elisabeth Schopf. Theoretical studies for the Standard Model $VH, H \to b\bar{b}$ and related searches
- Supporting Document. Technical Report ATL-COM-PHYS-2016-475, CERN, Geneva, May 2016. URL: https://cds.cern.ch/record/2151835.

[247] Abraham Wald. Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society*, 54(3):426–482, 1943. URL: http://www.jstor.org/stable/1990256.

[248] Aidan Robson, Giacinto Piacquadio, and Elisabeth Schopf. Signal and Background Modelling Studies for the Standard Model $VH, H \to b\bar{b}$ and Related Searches: Modelling support note for VH(bb) 2015+2016 dataset publication. Technical Report ATL-COM-PHYS-2016-1747, CERN, Geneva, Dec 2016. This is a support note for the VH(bb) SM publication using the 2015+2016 datasets. URL: https://cds.cern.ch/record/2235887.

[249] ATLAS Collaboration. Studies on top-quark Monte Carlo modelling for Top2016. ATL-PHYS-PUB-2016-020, 2016. URL: https://cds.cern.ch/record/2216168.

[250] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Parton distributions for the LHC. *Eur. Phys. J. C*, 63:189, 2009. arXiv:0901.0002, doi:10.1140/epjc/s10052-009-1072-5.

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

[252] 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. *Phys. Rev. Lett.*, 119(18):181804, 2017. arXiv:1707.01302, doi:10.1103/PhysRevLett.119.181804.

[253] ATLAS Collaboration. Search for dark matter in association with a Higgs boson decaying to two photons at $\sqrt{s}$= 13 TeV with the ATLAS detector. 2017. arXiv:1706.03948.

[254] CMS Collaboration. Search for associated production of dark matter with a Higgs boson decaying to $b\bar{b}$ or $\gamma\gamma$ at $\sqrt{s} = 13$ TeV. 2017. arXiv:1703.05236.

[255] David Krohn, Jesse Thaler, and Lian-Tao Wang. Jets with Variable R. *JHEP*, 06:059, 2009. arXiv:0903.0392, doi:10.1088/1126-6708/2009/06/059.

[256] Boosted Object Tagging with Variable-R Jets in the ATLAS Detector. Technical Report ATL-PHYS-PUB-2016-013, CERN, Geneva, Jul 2016. URL: http://cds.cern.ch/record/2199360.

[257] ATLAS Collaboration. 2015 start-up trigger menu and initial performance assessment of the ATLAS trigger using Run-2 data. ATL-DAQ-PUB-2016-001, 2016. URL: https://cds.cern.ch/record/2136007.

[258] ATLAS Collaboration. Trigger Menu in 2016. ATL-DAQ-PUB-2017-001, 2017. URL: https://cds.cern.ch/record/2242069.

[259] John Butler, Pai-hsien Jennifer Hsu, Shih-Chieh Hsu, Sandra Kortner, Yun-Ju Lu, Samuel Meehan, Dilia Maria Portillo Quintero, Patrick Rieck, Rainer Rohrig, Nikola Lazar Whallon, Veronica Fabiani, Frank Filthaut, Cheng-hsin Han, Po-shan Shih, Yu-hsin Chan, Andrea Matic, Jeanette Lorenz, Efe Yigitbasi, Ruth Pottgen, Eleni Skorda, and K. Technical report.

[260] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl. Instrum. Meth. A*, 506:250–303, 2003. doi:10.1016/S0168-9002(03)01368-8.

[261] ATLAS Collaboration. ATLAS simulation of boson plus jets processes in Run 2. ATL-PHYS-PUB-2017-006, 2017. URL: https://cds.cern.ch/record/2261937.

[262] ATLAS Collaboration. Multi-Boson Simulation for 13 TeV ATLAS Analyses. ATL-PHYS-PUB-2017-005, 2017. URL: https://cds.cern.ch/record/2261933.

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

[264] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching nlo qcd computations with parton shower simulations: the powheg method. *JHEP*, 11:070, 2007. arXiv:0709.2092, doi:10.1088/1126-6708/2007/11/070.

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

[266] Simone Alioli, Sven-Olaf Moch, and Peter Uwer. Hadronic top-quark pair-production with one jet and parton showering. *JHEP*, 01:137, 2012. arXiv:1110.5251, doi:10.1007/JHEP01(2012)137.

[267] Torbjorn Sjostrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. An Introduction to PYTHIA 8.2. *Comput. Phys. Commun.*, 191:159–177, 2015. arXiv:1410.3012, doi:10.1016/j.cpc.2015.01.024.

[268] Glen Cowan and Eilam Gross. Discovery significance with statistical uncertainty in the background. Technical report, ATLAS Statistics Forum, 2008. URL: https://www.pp.rhul.ac.uk/~cowan/stat/notes/SigCalcNote.pdf.

[269] The ATLAS collaboration. Object-based missing transverse momentum significance in the ATLAS detector. 2018.

[270] ATLAS Collaboration. Performance of b-jet identification in the ATLAS experiment. *JINST*, 11:P04008, 2016. arXiv:1512.01094, doi:10.1088/1748-0221/11/04/P04008.

[271] ATLAS Collaboration. Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC. *Eur. Phys. J. C*, 76:653, 2016. arXiv:1608.03953, doi:10.1140/epjc/s10052-016-4466-1.

[272] ATLAS Collaboration. Jet energy scale measurements and their systematic uncertainties in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Rev. D*, 96:072002, 2017. arXiv:1703.09665, doi:10.1103/PhysRevD.96.072002.

[273] Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector. Technical Report ATLAS-CONF-2018-036, CERN, Geneva, Jul 2018. URL: http://cds.cern.ch/record/2630338.

[274] Abdesslam Arhrib, Yue-Lin Sming Tsai, Qiang Yuan, and Tzu-Chiang Yuan. An Updated Analysis of Inert Higgs Doublet Model in light of the Recent Results from LUX, PLANCK, AMS-02 and LHC. *JCAP*, 1406:030, 2014. arXiv:1310.0358, doi:10.1088/1475-7516/2014/06/030.

[275] Kim Griest and David Seckel. Three exceptions in the calculation of relic abundances. *Phys.Rev.*, D43:3191–3203, 1991. doi:10.1103/PhysRevD.43.3191.

[276] ATLAS Collaboration. TWiki. URL: https://twiki.cern.ch/twiki/bin/viewauth/Atlas/LowestUnprescaled.