Dissertation

submitted to the

Combined Faculties of the Natural Sciences and Mathematics

of the Ruperto-Carola-University of Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

Put forward by

Stanislav Suchek

born in Kyiv

Oral examination on November 21, 2018

Search for Dark Matter produced in association with hadronically decaying Standard Model bosons using 36 fb^{-1} of pp collisions at the ATLAS detector

Referees: Priv.-Doz. Dr. Oleg Brandt
Prof. Dr. André Schöning

Abstract

Dark Matter composes a significant part of the observable Universe. Despite a solid cosmological evidence, its nature, properties and interaction with the Standard Model sector is still to be unraveled. A wide range of different dark matter models beyond the Standard Model can be probed at the Large Hadron Collider. Two searches for the dark matter particles produced in association with the hadronically decaying Standard Model vector bosons and Higgs boson h, decaying to a $b\bar{b}$ pair, are presented. Both searches are performed using 36.1 fb^{-1} of the proton-proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS detector in 2015 and 2016. In preparation for the interpretation of the final search results, the performance of the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers used in both analyses is studied. The scale factors are developed to correct for the discrepancy between the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency in data and Monte Carlo simulation. No significant excess in the observed data over the Standard Model predictions is observed in both searches. The upper limits on the production cross-sections of $W/Z + E_{\mathrm{T}}^{\mathrm{miss}}$ and $h + E_{\mathrm{T}}^{\mathrm{miss}}$ events times the branching ratios of $W/Z \to q\bar{q}$ and $h \to b\bar{b}$ decays for the simplified vector mediator and Z'-2HDM model scenarios are set at 95% confidence level. The limits are significantly improved relative to the previous results using 3.2 fb^{-1} of data. Additionally, the model-independent representation is introduced for the first time for both analyses. The generic upper limits on the visible cross-section of $W/Z + E_{\mathrm{T}}^{\mathrm{miss}}$ and $h + E_{\mathrm{T}}^{\mathrm{miss}}$ events at 95% confidence level are presented in separate regions of the $E_{\mathrm{T}}^{\mathrm{miss}}$ variable at the detector- and particle-levels.

Kurzzusammenfassung

Dunkle Materie bildet einen bedeutenden Teil des beobachtbaren Universums. Trotz einer soliden kosmologischen Evidenz sind ihre Natur, Eigenschaften und Wechselwirkungen mit dem Standardmodellsektor noch nicht aufgeklärt. Eine Reihe verschiedener Dunkler Materie Modelle jenseits des Standardmodells können am Large Hadron Collider getestet werden. In dieser Arbeit werden zwei Suchen nach Modellen vorgestellt, bei denen die Dunklen Materieteilchen in Assoziation mit hadronisch zerfallenden Standardmodell-Vektorbosonen und Higgs-Bosonen h erzeugt werden und zu einem $b\bar{b}$-Paar zerfallen. Beide Suchen verwenden 36.1 fb^{-1} der Proton-Proton-Kollisionsdaten bei $\sqrt{s} = 13$ TeV, die mit dem ATLAS-Detektor in den Jahren 2015 und 2016 aufgezeichnet wurden. In Vorbereitung auf die Interpretation der endgültigen Suchergebnisse wird die Effizienz der $E_{\mathrm{T}}^{\mathrm{miss}}$-Trigger untersucht, die in beiden Analysen verwendet werden. Dabei werden Skalierungsfaktoren entwickelt, um die Diskrepanz zwischen der Triggereffizienz in Daten und der Monte-Carlo-Simulation zu korrigieren. Bei beiden Suchen wurde kein signifikanter Überschuss der beobachteten Daten über den Standardmodell-Vorhersagen beobachtet. Die oberen Ausschlussgrenzen für die Produktionswirkungsquerschnitte von $W/Z + E_{\mathrm{T}}^{\mathrm{miss}}$ und $h + E_{\mathrm{T}}^{\mathrm{miss}}$-Ereignissen multipliziert mit den Verzweigungsverhältnissen der $W/Z \to q\bar{q}$ und $h \to b\bar{b}$ Zerfälle für die vereinfachten Vektormediator- und Z'-2HDM-Modellszenarien werden bei einem Konfidenzniveau von 95% gesetzt. Die Ausschlussgrenzen sind gegenüber den vorherigen Ergebnissen bei Verwendung von 3.2 fb^{-1} signifikant verbessert. Zusätzlich wird für beide Analysen erstmals die modellunabhängige Darstellung eingeführt. Die generischen oberen Auschlussgrenzen für den sichtbaren Wirkungsquerschnitt von $W/Z + E_{\mathrm{T}}^{\mathrm{miss}}$ und $h + E_{\mathrm{T}}^{\mathrm{miss}}$-Ereignissen bei 95%

Konfidenzniveau werden in separaten Bereichen der $E_{\mathrm{T}}^{\mathrm{miss}}$-Variablen auf der Detektor- als auch auf Teilchenebene dargestellt.

Contents

Chapter 1

Introduction

The first idea of matter that is composed of elementary particles came from the atomic hypothesis formulated by a Greek philosopher Democritus around 2500 years ago. At that times the level of technological development was far away from the possibility to check his hypothesis. Only about 100 years ago the technological progress allowed the discoveries of the electron by J. Thompson [1], the proton by E. Rutherford [2] and the neutron by J. Chadwick [3] to happen, putting a solid background under the atomic hypothesis.

Later in 1960s-1980s a huge breakthrough occurred from both theoretical and experimental side. The quark model was independently offered by Gell-Mann [4] and Zweig [5], introducing the strong interaction and new particles as bound states of quarks. And in 1968, deep inelastic scattering experiments at the Stanford Linear Accelerator Center showed that proton was not an elementary particle but consisted of other new particles, which were associated with quarks [6, 7]. In 1967 and 1968 a theory that unified the electromagnetic and weak interactions was postulated by Weinberg [8] and Salam [9], introducing new $W^{\pm}$ and Z boson as mediators of the weak force. In the following years the joint effort of experimental physicists resulted in construction of the more powerful particle accelerators, which provided the discoveries of the gluon, the mediator of the strong interaction, at Deutsches Elektronen-Synchrotron (DESY) in 1979 [10–13] and to the discovery of the W and Z bosons at Conseil Europeen pour la Recherche Nucleaire (CERN) in 1983 [14–17].

All above-mentioned discoveries led to the development of the Standard Model of Particle Physics (SM), one of the most successful theories ever formulated by the mankind. The SM is able to successfully describe most of the experimental data, describing three of the four fundamental interactions, electromagnetic, weak and strong. The discoveries of the top quark in 1995 [18, 19], tau neutrino in 2000 [20], and especially the Higgs boson in 2012 [21, 22] at the Large Hadron Collider put the cherry on the cake of this successful theory.

But the more scientists studied the properties of the Universe, the more questions showed up, which the SM failed to describe. And one of the biggest open questions is an existence of the Dark Matter (DM). First introduced in 1920s-1930s [23–25], the DM has manifests itself in many independent astrophysical measurements. Present measurements indicate that the DM composes more than a quarter of the Universe, which is approximately five times larger than the amount of the visible matter [26, 27]. But despite the large variety of different experiments designed to search for the DM, its nature is still to be understood.

One possible way to study the properties of DM particles is to look for the DM production

in high energy collisions at particle colliders. This includes a variety of different processes, e.g. a direct pair-production or production of additional new particles, which then decay to DM particles. Searches for DM at colliders are super challenging due to the extremely weak interaction between the DM and detector material, making the DM impossible to be directly detected. Thus, indirect signatures of the associated production of undetectable DM particles and detectable SM particles, become extremely important. In this case, the DM particles manifest themselves as an imbalance in the conservation of momentum, which is measured as the missing transverse momentum. This kind of a search signature is denoted as $E_{\mathrm{T}}^{\mathrm{miss}}+X$, where X can be a light quark or gluon [28–31], a heavy quark [32–34], a photon [35–37], W or Z boson [38–41] and the Higgs boson [42–47]. All X particles orinigates either from the initial state radiation (ISR) or the hard interaction, except the Higgs boson, which originates from the latter. This is due to the Yukawa suppression of the ISR of the Higgs boson, meaning that $E_{\mathrm{T}}^{\mathrm{miss}} + h$ process can probe the direct interaction between the SM Higgs sector and the DM sector.

This thesis focuses on the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$, where V stands for the W and Z bosons, and $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ signatures, which imply invisible DM particles, recoiling against hadronically decaying vector bosons and the Higgs boson, decaying into a $b\bar{b}$ pair, respectively. The $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ search [48] probes the DM production when the DM sector couples either directly to the SM sector or indirectly via other additional new particles. Since the radiation of the Higgs boson is higly suppressed, the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ search [46] provide a unique possibility to test the direct coupling between the Higgs sector and the DM sector. A choice of the $W \to q'\bar{q}$, $Z \to q\bar{q}$ and $h \to b\bar{b}$ decays modes, despite a big challenge related to the large SM backgrounds, is motivated by their large branching ratios of $\mathcal{B}_{W\to q'\bar{q}} \approx 69\%$, $\mathcal{B}_{Z\to q\bar{q}} \approx 70\%$ and $\mathcal{B}_{h\to b\bar{b}} \approx 57\%$, respectively.

The $E_{\mathrm{T}}^{\mathrm{miss}}+V(qq)$ search mostly focuses on the DM produced in association with the SM W or Z bosons produced via ISR, where the SM and DM parts are connected via the new BSM mediator. Meanwhile, the $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ search considers the decay of the new BSM mediator to the SM Higgs boson and another new particle, which decays for the pair of the DM particles. These two searches provide an unique chance to probe both indirect and direct possibilities of interactions between the DM and the SM sectors. Despite a big challenge related to the large SM backgrounds, $V \to qq$ and $h \to b\bar{b}$ decays provide the largest branching ratios of $\mathcal{B} = 68\%$ and $\mathcal{B} = 57\%$ respectively.

This thesis starts from introducing the main theoretical aspects of the SM and its open questions in Chapter 2. A detailed discussion about the origins and properties of the DM, as well as the possible candidates and signal models of the DM production at particle colliders, is given in Chapter 3. The general aspects of the ATLAS setup are presented in Chapter 4 with the main focus shifted to the ATLAS trigger system, as one of the main contributions of an author. A discussion about the ATLAS data taking, Monte Carlo simulations and triggers relevant for the searches in the scope of this thesis, performance of which is improved by the author, is given in Chapter 5. The reconstruction and identification of the objects in the ATLAS detector is discussed in Chapter6 The next chapters are aimed to provide a detailed information about the searches considered in this thesis. The event selection used for the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ searches is presented in details in Chapter 6. A discussion on the main backgrounds and their constraints is given in Chapter 8. The main aspects of the statistical model used for the interpretation of the search results are described in Chapter 9. A separate discussion on the systematic uncertainties embedded in the statistical models of $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ searches is given in Chapter 10. Chapter 11 provides an overview of strategies

for the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ searches with a dedicated discussion on the interpretation of the search results. The key results of the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ searches in terms of the different interpretations are presented in Chapter 12. Finally, the overall conclusions are given in Chapter 13.

Contributions

Due to the high complexity of the current experiments in particle physics, a large number of people contribute to all aspects of the experiment, beginning with building and maintaining the detector components and ending with the data analysis. In particular, the ATLAS experiment is a joint effort of around 3000 members from all across the world. The results presented in this thesis contain contribution from other collaborators. The main contributions of the author are as follows.

The author contributed to the different aspects of the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ searches from the early stages to the publications, providing inputs at all stages of searches and helping in maintaining the corresponding software packages.

The contribution of the author to the Monte Carlo $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger calibration for the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ searches is described in Section 5.2. The author studied the difference between performances of the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers in data and Monte Carlo and derived the corresponding scale factors to correct Monte Carlo $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiencies to match those in data for several $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers used in the analyses. The author also derived all statistical and systematic uncertainties related to the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency corrections. Finally, the author provided associated software packages for the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ searches that implemented the corresponding scale factors and their uncertainties.

The author contributed to the final interpretations of $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ search results, presented in Chapter 12. The author developed the model-independent approach and derived the generic limits on the production of W+DM, Z+DM and h+DM events, as discussed in Chapter 11. Such limits are useful for phenomenologists to check the validity of new DM models with a given final state. Limits were derived at the detector and parton levels.

The author also introduced re-scaling procedure to derive the exclusion limits for the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ simplified vector and axial-vector mediator models at next-to-leading order from the existing limits at leading-order, as described in Chapter 12. The exclusion limits were derived for the different model scenarios for the $E_{\mathrm{T}}^{\mathrm{miss}} + W(qq)$, $E_{\mathrm{T}}^{\mathrm{miss}} + Z(qq)$ and combined $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ production.

The author was also the primary person responsible for the re-scaling of the final exclusion limits for the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ simplified vector and axial-vector mediator models from leading-order to next-to-leading-order, as described in Chapter 12. The next-to-leading limits are useful to compare $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ exclusion limits for the simplified DM models to the corresponding limits from other $E_{\mathrm{T}}^{\mathrm{miss}} + X$ and dijet searches, as well as to the limits from the direct detection experiments. These comparisons are included in the summary paper of the DM production at ATLAS [49, 50].

Chapter 2

The Standard Model of particle physics

The Standard Model (SM) of particle physics has been proven to be one of the most successful models in the history of mankind. Formulated in 1960s-1970s and then completed after the Higgs discovery at the LHC, this theory is able to explain many aspects the fundamental physics at the high energy scales. Despite its impressive predictive power and huge success in explaining the various experimental results, it is known to have a limited application area and has some limitations and yet unaddressed questions. In this chapter, various aspects of the SM along with its challenges are briefly discussed.

2.1 Fundamental foundations of the Standard Model

The Standard Model is a locally gauge invariant quantum field theory, which describes all currently known elementary particles and their fundamental interactions, with an exception of gravitational interaction. The model states that all the matter content of our Universe consists of two fundamental types of particles, *fermions* and *bosons*. The fermions have a half-integer internal angular momentum, or *spin*, and follow the Fermi-Dirac statistics. They are the main building blocks of matter. Meanwhile, the bosons have an integer spin and follow the Boson-Einstein statistics. They are known to mediate the fundamental forces within the SM and to be responsible for the mass generation mechanism in case of the Higgs boson.
The SM stays invariant under the symmetry of the local gauge group $SU(3)_C \times SU(2)_L \times U(1)_Y$. Each symmetry corresponds to the fundamental force incorporated in the theory:

- $SU(3)_C$ group has $3^2 - 1 = 8$ generators and thus gives rise to eight gauge boson fields G^a_μ, $a = 1, .., 8$, called *gluons*, the mediators of the strong force. They carry the conserved quantum number C, the *colour charge*. Further details are discussed in Section 2.2.

- $SU(2)_L$ group has $2^2 - 1 = 3$ generators, which represents three gauge boson fields W^i_μ, $i = 1, .., 3$. They are the mediators of the weak force and can be further associated with the observed $W^\pm$ and Z bosons. They carry the conserved quantum number T, called *weak isospin*. Further details are discussed in Section 2.3.

- $U(1)_Y$ generates a single gauge boson B with the conserved quantum number Y, the *weak hypercharge*. It can be further associated with the photon, the mediator of the electromagnetic interaction. Further details are discussed in Section 2.3.

The fermions are divided into different types, depending on the fundamental interactions they are subject to. Fermions, which take part in all fundamental interactions within the SM, are called *quarks*. The quarks form the bound states, called *hadrons*, the most stable of which, protons and neutrons, make up all known atomic nucleus. *Leptons* are governed by the weak and electromagnetic interactions, while *neutrinos* are subjects to only weak interaction. The summary of the particle content of the SM is shown in Figure 2.1.

The local gauge symmetries are known to generate the massless gauge fields, meanwhile all experimental measurements state the opposite. The solution to this contradiction is the *Higgs mass generation mechanism*, which breaks the electroweak symmetry $SU(2)_L \times U(1)_Y$ down to the $U(1)$ symmetry, and generates masses of the particles via the interaction with the new massive *Higgs field*, as described in Sections 2.3.1 and 2.3.2.

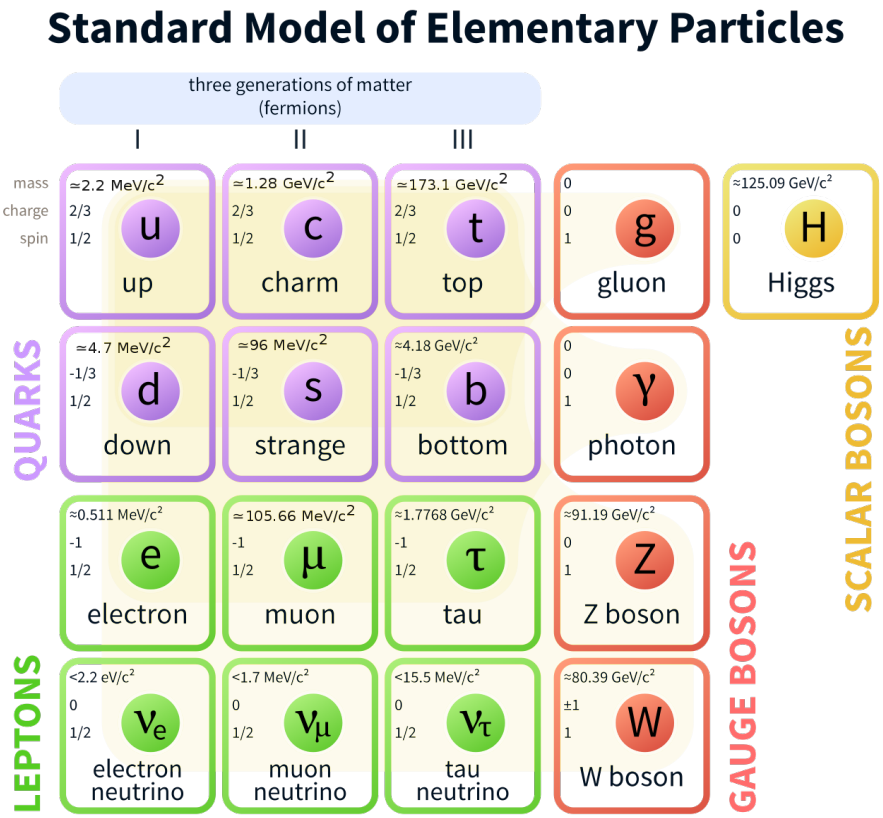


Figure 2.1: The particle content of the Standard Model [51].

2.2 Quantum Chromodynamics

Quantum Chromodynamics (QCD) is the theory of the strong interaction, which is mainly responsible for the processes inside the hadrons, such as the proton and neutron. It is a non-Abelian gauge quantum field theory that describes the interaction of the quarks and gluons. QCD obeys the symmetry of the coloured gauge $SU(3)_C$ Lie group. It has eight independent *generators* T_a, which in the fundamental representation are defined in the

following way:

$$[T_a, T_b] = i \sum_{c=1}^{3^2-1} f_{abc} T_c \,, \tag{2.1}$$

where f_{abc} are the structural constants of the $SU(3)_C$ group. The generators of the group can be expressed via so-called *Gell-Mann matrices* λ_a [52], which are the generalisation of Pauli matrices in three-dimensional space, $T_a = \lambda_a/2$, $a = 1, ..., 8$, which correspond to 8 massless gluon fields. Each quark is represented by the fundamental spinor of the $SU(3)_C$ group. Quarks can carry three types of colour charge states, red, green and blue, known as a colour triplet. Gluons can carry already eight types of colour charge states, known as a color octet.

Now all ingredients are ready to taylor the Lagrangian of the QCD:

$$\mathcal{L}_{\mathcal{QCD}} = -\frac{1}{4} \sum_{a=1}^{8} G^a_{\mu\nu} G^{a\mu\nu} + \sum_{i,j=1}^{N_f} \bar{q}_i \left(i\gamma^\mu D_\mu - m_q \delta_{ij} \right) q_j \,, \tag{2.2}$$

where $G^a_{\mu\nu}$ is the gauge invariant gluon field tensor:

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - g_s f_{abc} G^b_\mu G^c_\nu \,, \tag{2.3}$$

G^a_μ represents eight gluon fields, q_i denotes the quark spinors with the colour indices $i, j = 1, 2, 3$, g_s denotes the coupling strength of the strong interaction and N_f is a number of the quark flavours. The last term of the relation reveals the non-Abelian nature of QCD and corresponds to the self-coupling between the gluons, as a consequnce of the fact that gluons carry colour charge. D_μ in Eq. 2.4 represents the gauge covariant derivative:

$$(D_\mu)_{ij} = \delta_{ij} \partial_\mu + i g_s T^a_{ij} G^a_\mu \,. \tag{2.4}$$

One can use the path integral formulation of the quantum field theory, where the classical trajectory of a system is replaced by the sum over an infinite number of quantum trajectories, the *functional integral*, to obtain the probability amplitude of the transition from the initial state of the system to its final state. One can obtain the amplitude in the form of the perturbative expansion, where each contribution can be associated with the graphical representation, or the *Feynman diagram*. The Feynman diagrams that correspond to the first term in the perturbative expansion, are referred to the *tree-level* diagrams. A representative tree-level Feynman diagram is shown in Figure 2.2.

In the language of these tree-level Feynman diagrams, the QCD Lagrangian can be rep-

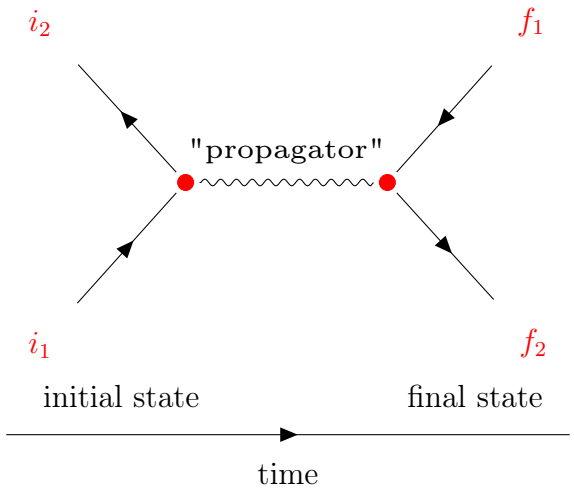
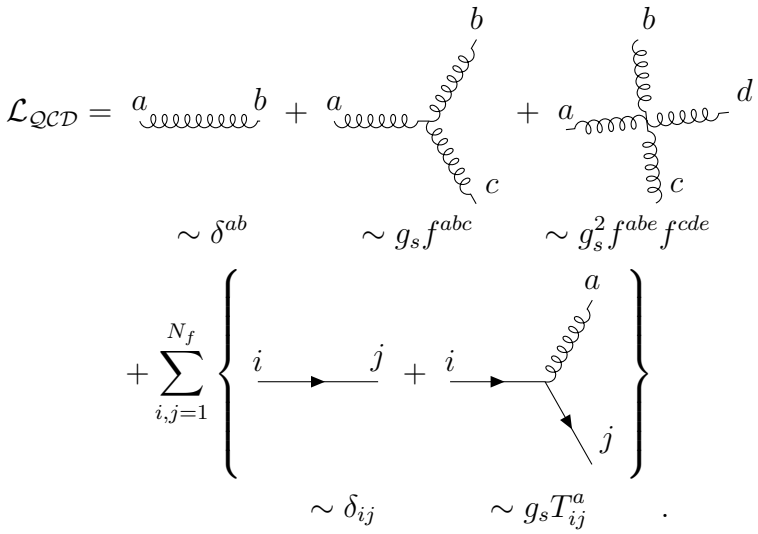


Figure 2.2: Explanatory illustration of the Feynman diagram, where the mediator is exchanged between the initial state particles, i_1 and i_2, and the final state particles, f_1 and f_2. The total momentum of all incoming and outcoming particles is conserved in each vertex (red blobs).

resented graphically as:



(2.5)

The first term describes the propagator of the free gluon field. The second and the third terms describe the three- and four-point self-interaction of gluons, proportional to g_s and g_s^2, respectively. This makes it fundamentally different from Quantum Electrodynamics, where the photon does not carry any electric charge and thus cannot undergo self-interactions. The fourth and the fifth terms describe the propagator of the free quark field and the quark-gluon interaction, respectively.

The self-interaction of the gluons leads to the fundamental behaviour the coupling strength of the strong interaction, $\alpha_s \equiv g_s^2/4\pi$. One can calculate how it evolves with the energy scale Q^2 of the interaction as a solution of the renormalization group equation (RGE) in the 1-loop approximation [53]:

$$\alpha_s(Q^2) \equiv g_s/4\pi = \frac{\alpha_s(\mu^2)}{1 + b_0 \cdot \alpha_s(\mu^2) \cdot \ln(Q^2/\mu^2)} \,. \tag{2.6}$$

Here Q correspond to the energy scale at which α is measured, μ corresponds to the

reference energy scale, where the coupling strength is known. b_0 is the leading order term from the expansion of the beta function $\beta(\alpha_s)$:

$$b_0 = \frac{1}{3}(11N_C - 2N_f)\,, \tag{2.7}$$

where N_C denotes the number of colours, and N_f the number of flavous accessible at the scale Q^2. In our world, where $N_C = 3$ and at the most $N_f = 6$, $b_0 > 0$, meaning that effective coupling strength *decreases* with *increasing* energy scale. It leads to the fact that the quarks and the gluons become asymptotically unbound at high energies. This is exactly the meaning of the *asymptotic freedom*. And vice versa, at low energy scales Q^2, or large distances between quarks or gluons, coloured quarks and gluons form colourless bond states, *mesons* ($q\bar{q}$) and *baryons* (qqq) with the corresponding antihadrons. This process is called *hadronisation*. This corresponds to the phenomenon called *colour confinement*.

Effectively, the interaction between two largely separated quarks can be visualised as an effective tube, *QCD colour flux tube*, of a constant energy density $\kappa \approx 1$ GeV/fm. An effective potential of such interaction is linearly growing with the separation between the quarks, $V(r) \approx \kappa r$. At the large enough separation, it becomes more energetically favourable to produce the quark-antiquark pair out of the QCD vacuum rathen than to further increase the interaction energy. This process goes on until the resulting quarks start to form colourless hadrons, *mesons* ($q\bar{q}$) and *baryons* (qqq) with the corresponding antihadrons.

At hadron colliders, hadronization of highly energetic quarks and gluons leads to a specific observed final state, called *jet*, a collimated spray of hadrons.

At the energy scales below some certain scale $\Lambda_{\mathcal{QCD}}$, the perturbative expression of the α_s starts to diverge, and thus quarks begin to hadronise. It is defined by the *Landau pole* of Eq. 2.6:

$$\Lambda^2_{\mathcal{QCD}} = \mu^2 \exp\left[-\frac{1}{b_0\alpha(\mu^2)}\right]. \tag{2.8}$$

The value of $\Lambda_{\mathcal{QCD}}$ at the energy scales above the mass of the $b-$quark is $\Lambda_{\mathcal{QCD}} \approx 220$ MeV. In the perturbative regime, at the scale of the mass of the Z boson, the value of α_s is $\alpha_s(m_Z^2) = 0.1181 \pm 0.0011$ [54].

2.3 The theory of the electroweak interaction

The theory of the electroweak (EW) interaction provides a joint description of the weak and electromagnetic interactions between the fermions. It is a non-Abelian theory that obeys the $SU(2)_L \times U(1)_Y$ symmetry. The $SU(2)_L$ group represents the group of the weak interaction, where the fermions are considered as the *left-handed* and *right-handed* states, which behave differently under the $SU(2)_L$ transfromations:

$$\psi_{L,R} = \frac{1}{2}(1 \mp \gamma^5)\psi, \qquad \psi = \psi_L + \psi_R\,. \tag{2.9}$$

The left-handed fermions attribute the weak isospin $T = 1/2$, while the right handed $T = 0$. The right-handed fermions are the singlets under the $SU(2)$ transfromations,

while the left-handed states represent the doublet states in the $SU(2)$ group:

$$\text{leptons: } \psi_L = \begin{pmatrix} \nu_{i,L} \\ e_{i,L} \end{pmatrix}, \qquad \text{quarks: } \psi_L = \begin{pmatrix} u_{i,L} \\ d'_{i,L} \end{pmatrix}. \tag{2.10}$$

Here the third component of the weak isospin $T_3 = 1/2$ correspond to neutrinos $\nu_{i,L}$ and *up-type* quarks $u_{i,L}$, while $T_3 = -1/2$ correspond to leptons $e_{i,L}$ and *down-type* quarks $d'_{i,L}$. Note here that in case of the weak interaction the down-type states $d'_{i,L}$ are not the mass eigenstates, but the flavour one. In fact the $d'_{i,L}$ flavour states are the linear combinations of the mass eigenstates with coefficients that are defined by so-called *Cabibbo-Kobayashi-Maskawa* matrix [55, 56]. $SU(2)_L$ group generates three gauge boson fields W^i_μ, which interact only with the left-handed fermions.

The theory is invariant under the local $SU(2)_L$ transformations, which are different in case of the left-handed and right-handed fermions:

$$\psi_L \to e^{-\frac{1}{2}\vec{\tau}\vec{\alpha}(x)}\psi_L, \qquad \psi_R \to \psi_R\,, \tag{2.11}$$

where the 2x2 matrices $\vec{\tau}$ are defined by the Pauli matrices $\vec{\sigma}$, $\vec{\tau} = \vec{\sigma}/2$. The theory is invariant under $SU(2)$ transformations, which are different in case of the left-handed and right-handed fermions:

$$\psi_L \to e^{-\frac{1}{2}\vec{\tau}\vec{\alpha}}\psi_L, \qquad \psi_R \to \psi_R\,. \tag{2.12}$$

An additional global $U(1)_Y$ symmetry is responsible for the electromagnetic sector. It generates an additional gauge field B_ν and corresponds to the conservation of the *weak hypercharge* Y, defined by the following relation [1]:

$$Q = T_3 + \frac{Y}{2}\,, \tag{2.13}$$

where Q represents the eletric charge. Given Eq. 2.13, the left-handed and the right-handed fermions carry different hypercharges Y_L and Y_R, respectively.

The requirement of the locally invariant $SU(2)_L \times U(1)_Y$ gauge symmetry, leads to the following Lagrangian of the theory:

$$\begin{aligned} \mathcal{L} &= \bar{\psi} i\gamma^\mu D_\mu \psi - \frac{1}{4}\vec{W}_{\mu\nu}\vec{W}^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \\ D_\mu &= \partial_\mu - ig\frac{\vec{\sigma}}{2}\vec{A}_\mu - ig'\frac{Y}{2}B_\mu \\ \vec{W}_{\mu\nu} &= \partial_\mu \vec{W}_\nu - \partial_\nu \vec{W}_\mu + g\vec{W}_\mu \times \vec{W}_\nu B_{\mu\nu} \qquad = \partial_\mu B_\nu - \partial_\nu B_\mu\,, \end{aligned} \tag{2.14}$$

where g and g' are the coupling constants, ψ denote both the lepton and quark fields.

The observed physical eigenstates $W^\pm$, Z bosons and the photon A_μ can be obtained as

[1] This relation is analogous to the Gell-Mann-Nishijima relation for the hadrons [54]

the linear combinations of the gauge fields $\vec{W}^{\mu}$ and B_{μ}:

$$\begin{aligned} W_{\mu}^{\pm} &= \frac{1}{\sqrt{2}}\left(W_{\mu}^{1} \mp iW_{\mu}^{2}\right) \\ Z_{\mu}^{0} &= \cos\theta_W W_{\mu}^{3} - \sin\theta_W B_{\mu} \\ A_{\mu} &= \sin\theta_W W_{\mu}^{3} - \cos\theta_W B_{\mu} \\ \cos\theta_W &= \frac{g}{\sqrt{g^2+g'^2}}\,, \end{aligned} \tag{2.15}$$

where the θ_W is the *Weinberg angle of a weak mixing*. Given the physical eigenstates in Eq. 2.15, the corresponding interaction term between the fermion and the gauge fileds in the Lagrangian 2.14 is the following:

$$\mathcal{L} = g\left(W_{\mu}^{+}J_{W}^{\mu+} + W_{\mu}^{-}J_{W}^{\mu-} + Z_{\mu}^{0}J_{Z}^{\mu}\right) + g\sin\theta_W A_{\mu}J_{\text{em}}^{\mu}\,. \tag{2.16}$$

The $W_{\mu}^{\pm}J_{W}^{\mu\pm}$ and $Z_{\mu}^{0}J_{Z}^{\mu}$ terms are responsible for a weak interaction of the fermions via the charged $W^{\pm}$ bosons (*charge current*) and the neutral Z boson (*neutral* currents). The $A_{\mu}J_{\text{em}}^{\mu}$ term corresponds to the electromagnetic interaction of the fermions via the photon and $g\sin\theta_W$ terms can be associated with the electron charge.

2.3.1 Spontaneous symmetry breaking and the Higgs Boson

The W and Z bosons were discovered at CERN in 1983 [14–17] and appeared to be massive. Meanwhile, the physical eigenstates in Eq. 2.14 are massless. Any additional mass term of the gauge field would immediately violate the initial EW symmetry. The mechanism of the mass generation for the gauge bosons and fermions, and a key role of the Higgs boson in it is discussed in this Section. More details about the Higgs mass generation mechanism can be found, e.g. in Ref. [57–59].

2.3.2 The Higgs mass generation mechanism

The Higgs mechanism should give the mass to the W and Z bosons, while keeping the photon massless, and the Higgs field should initially obey the symmetry of the gauge theory, which is $SU(2)_L \times U(1)_Y$. Introduce a complex scalar doublet field ϕ together with the corresponding potential $V(\phi)$:

$$\phi(x) = \begin{pmatrix} \phi^{+}(x) \\ \phi^{0}(x) \end{pmatrix}, \qquad \phi^{*}(x) = \begin{pmatrix} \phi^{-}(x) \\ \phi^{0^*}(x) \end{pmatrix}, \qquad V(\phi) = -\mu^2(\phi^{\dagger}\phi) + \frac{1}{4}\lambda^2(\phi^{\dagger}\phi)^2\,. \tag{2.17}$$

The corresponding Lagrangian of the Higgs field is:

$$\mathcal{L}_{\text{Higgs}} = (D_{\mu}\phi)^{\dagger}(D^{\mu}\phi) + V(\phi^{\dagger}\phi)\,, \tag{2.18}$$

where D_{μ} is the covariant derivate, as defined in Eq. 2.14. Consider the case of $\mu^2 > 0$ and $\lambda^2 > 0$, where the Higgs potential is minimised at the non-zero *vacuum expectation value* (VEV):

$$\phi^{\dagger}\phi = \frac{4\mu^2}{\lambda}\,, \tag{2.19}$$

which is an infinitely degenerate state. Choose the specific VEV:

$$\langle\phi\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v\end{pmatrix}, \tag{2.20}$$

where v is a positive real number. Now the $SU(2)_L \times U(1)_Y$ symmetry is *spontaneously broken* down to the simple $U(1)_Q$ symmetry, which ensures the presence of the photon with zero mass:

$$\hat{Q}\,\langle\phi\rangle = \left(T_3 + \frac{Y}{2}\right)\langle\phi\rangle = 0\,, \tag{2.21}$$

where $e^{i\frac{\vec{\tau}}{2}\cdot g\vec{\omega}(x)}$ is $SU(2)_L$ phase. The fermions fields ψ are the eigenstates of a charge opertor $\hat{Q}$ with the eigenvalues equal to the charges of the corresponding fermions, $\hat{Q}\psi = Q\psi$. Parametrise the Higgs field using the polar coordinates and consider the small excitations of the Higgs field about the VEV:

$$\phi(x) = e^{i\frac{\vec{\tau}}{2}\cdot g\vec{\omega}(x)}\Phi(x) = e^{i\frac{\vec{\tau}}{2}\cdot g\vec{\omega}(x)}\begin{pmatrix}0\\ \rho(x)\end{pmatrix} \approx e^{i\frac{\vec{\tau}}{2}\cdot g\vec{\omega}(x)}\begin{pmatrix}0\\ v+h(x)\end{pmatrix}. \tag{2.22}$$

Then the Lagrangian 2.18 splits into three different parts: the actual Higgs field term $\mathcal{L}_{\text{Higgs}}$, which includes kinetic, mass and self-interaction terms, the mass term of the gauge fields $\mathcal{L}_{mass}$ and the interaction term $\mathcal{L}_{int}$. Given the physical eigenstates in Eq. 2.15, the corresponding terms take the form:

$$\begin{aligned}
\mathcal{L} &= \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{mass}} + \mathcal{L}_{\text{int}}\\
\mathcal{L}_{\text{Higgs}} &= \frac{1}{2}(\partial_\mu h)^2 - \frac{1}{2}m_h^2h^2 - \frac{m_h^2}{2v}h^3 - \frac{m_h^2}{8v^2}h^3\\
\mathcal{L}_{mass} &= \frac{1}{2}\frac{g^2v^2}{4}\left[(A_\mu^1)^2 + (A_\mu^2)^2\right] + \frac{1}{2}\frac{v^2}{4}(-gA_\mu^3 + g'B_\mu)^2\\
&= \frac{1}{2}\frac{g^2v^2}{4}\left[(W_\mu^+)^2 + (W_\mu^-)^2\right] + \frac{1}{2}\frac{(g^2+g'^2)v^2}{4}(Z_\mu^0)^2\\
\mathcal{L}_{int} &= (2vh + h^2)\left[\frac{g^2}{4}(W_\mu)^2 + \frac{g^2+g'^2}{8}(Z_\mu^0)^2\right],
\end{aligned} \tag{2.23}$$

where $m_h = \sqrt{2\mu^2} = \sqrt{\lambda}v$ is the mass of the Higgs boson. Eq. 2.23 shows that after the interaction with the Higgs field with the non-zero VEV, the W and Z fields acquire the non-zero mass terms:

$$\begin{aligned}
m_Z &= \frac{\sqrt{g^2+g'^2}}{2}v\\
m_W &= \frac{gv}{2} = m_Z\cos\theta_W\,.
\end{aligned} \tag{2.24}$$

The masses of the W and Z bosons have been measured with the unprecedented precision by LEP [60], Tevatron [61], and LHC [62], with the current best values $m_W = 80.379 \pm 0.012$ GeV and $m_Z = 91.1876 \pm 0.0021$ GeV [54].

Consider the case of the mass generation of fermions. Introduce the Higgs field and the Higgs potential in the same way as in Eq. 3.20. The Lagrangian of the interaction between

fermions and the Higgs field looks as follows:

$$\mathcal{L}_{\text{int}} = -\sqrt{2}y\left[\bar{\psi}_L\psi_R\phi + \bar{\psi}_R\psi_L\phi^*\right], \tag{2.25}$$

where y is a nondimensional *Yukawa coupling strength*. The expression in brackets in the only scalar constructed from the left-handed and right-handed fields, which satisfies the $SU(2)_L \times U(1)_Y$ symmetry. After spontaneous symmetry breaking, the fermionic mass term Lagrangian 2.25 takes the form:

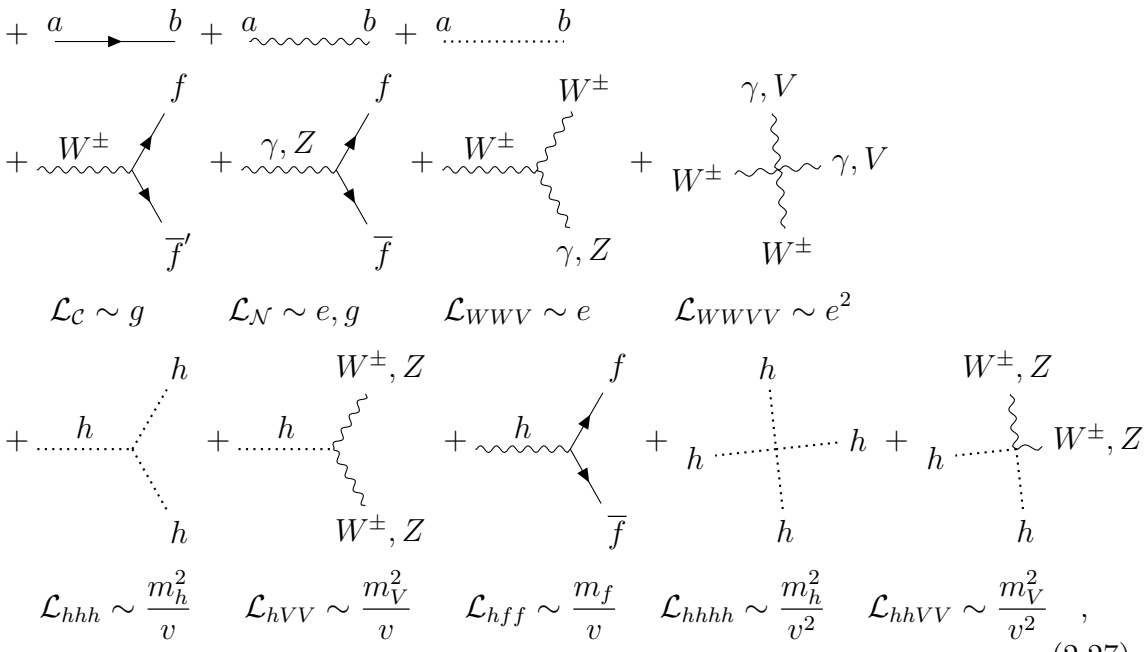
$$\mathcal{L}_{\text{mass}} = yv\psi\bar{\psi}, \tag{2.26}$$

which describes fermions with the mass of $m = yv$.

2.4 Complete Lagrangian of the Standard Model

The Higgs mass generation mechanism turns over the last page of the story of the electroweak theory and the Standard Model. To sum up everything what was discussed in this chapter, the complete Lagrangian of the Standard Model can be schematically illustrated as the tree-level Feynman graphs:

$\mathcal{L}_{\mathcal{SM}} = \mathcal{L}_{\mathcal{QCD}} +$



(2.27)

where $\mathcal{L}_{\mathcal{QCD}}$ is graphically defined in Eq. 2.5. The first line in the equation stands for the free motion of fermions, gauge bosons and the Higgs boson. $\mathcal{L}_{\mathcal{C}}$ and $\mathcal{L}_{\mathcal{N}}$ stand for the interaction of fermions to the neutral bosons (the photon and Z) and to the charged bosons ($W^\pm$) respectively. $\mathcal{L}_{WWV}$ and $\mathcal{L}_{WWVV}$ are responsible for the three- and four-point self-interactions of the gauge bosons. $\mathcal{L}_{hhh}$ stands for the three-point self-interaction of the Higgs boson. $\mathcal{L}_{hVV}$ and $\mathcal{L}_{hff}$ stand for the interaction of the Higgs boson to vector bosons and the fermions respectively. $\mathcal{L}_{hhhh}$ is responsible for the four-point self-interaction of the Higgs boson, while $\mathcal{L}_{hhVV}$ for the interaction between the Higgs boson and vector bosons.

The abovementioned SM Lagrangian has 19 free parameters:

- Nine Yukawa couplings to the fermions.
- Three gauge couplings g, g' and g_s.
- Three mixing angles θ_{12}, θ_{23}, θ_{13} and CP violating phase δ_{CP} of the Cabibbo-Kobayashi-Maskawa matrix (not discussed in the scope of this work).
- The mass of the Higgs m_h and its VEV v.
- QCD CP violating phase δ_{CP} (not discussed in the scope of this work).

The number of the free parameters increases to 26 if the massive neutrino term with three neutrino masses and four parameters of the PMNS matrix are included. The SM Lagrangian becomes complete after all these parameters are defined, and can be used to derive experimental predictions. Table 2.1 summarises our current knowledge about the properties of the fermions within the SM.

Leptons			
Particle		Quantum numbers (C, T, Q/e)	Mass
Electron	e^-	$(3, -1/2, -1)$	0.511 MeV
Muon	μ^-	$(3, -1/2, -1)$	105.7 MeV
Tau	τ^-	$(3, -1/2, -1)$	1777 MeV
Electron neutrino	ν_e	$(0, 1/2, 0)$	< 2 eV
Muon neutrino	ν_μ	$(0, 1/2, 0)$	< 0.19 MeV
Tau neutrino	ν_τ	$(0, 1/2, 0)$	< 18.2 MeV
Light quarks			
Up	u	$(3, 1/2, 2/3)$	$2.2^{+0.5}_{-0.4}$ MeV
Down	d	$(3, -1/2, -1/3)$	$4.7^{+0.5}_{-0.3}$ MeV
Strange	s	$(3, 1/2, 2/3)$	95^{+9}_{-3} MeV
Heavy quarks			
Charm	c	$(3, -1/2, -1/3)$	1275^{+25}_{-35} MeV
Bottom	b	$(3, -1/2, -1/3)$	$4.6^{+0.04}_{-0.03}$ GeV
Top	t	$(3, 1/2, 2/3)$	$173^{+0.4}_{-0.4}$ GeV

Table 2.1: Properties of the quarks and leptons in the SM. Their quantum numbers and experimentally measured masses are listed [54]. The lepton masses are measured with very high precision and thus are given without any uncertainties.

2.5 Limitations and open questions of the Standard Model

Although the SM is able to explain a huge variety of experimental results, it still leaves out several unaddressed phenomena known to exist in nature. Let us go through the most vivid unexplained facts.

Incorporation of gravity

The first obvious unaddressed point is gravity. All models that pretended to be a quantum field theory of the gravitational interaction failed the renormalizability criteria, meaning that they required an infinite number of free parameters to cancel out all divergences. This single fact already indicates that the SM is just an effective field theory, which works up to the gravitational scale of $\mathcal{O}(10^{19}$ GeV).

Dark matter and dark energy

This is the most directly relevant topic in the scope of this work. Dark matter is known to exist in the Universe from various astrophysical observations, such as galaxy rotation curves [63], gravitational lensing [64], cosmic microwave background [65], etc. It does not interact neither strongly, nor electromagnetically, which makes it hard to be explained in the scope of the SM if it has a particle nature. The only possible DM candidate within the SM, the neutrino, can account for only a small fraction of the observed abudance of the DM in the Universe. The most promising candidates to search for at hadron colliders, weakly interacting massive particles (WIMPs), which are not part of the SM, still remain unobserved. More details on the dark matter topic are discussed in Chapter 3.
The knowledge about the nature of the dark energy is even poorer. Astrophysical observations show that the Universe is expanding at accelerating rate. To account for the accelerating expansion, the concept of dark energy that interacts only gravitationally was introduced. The most popular hypothesis states that the dark energy represents very low homogeneous constant energy density filling the entire spacetime. It contributes 68% of the total energy content of the present-day observable Universe. Together with the dark matter, it means that the SM is able to explain only tiny 5% of the Universe.

Matter-antimatter asymmetry

One can also mention a matter-antimatter asymmetry, which is the reason why ordinary matter exists in the modern Universe. The Standard Model predicts equal amount of matter and antimatter produced in the early Universe, but in the modern world matter by far dominates over antimatter. The CP violation, provided by the SM, can absorb only the small part of the asymmetry.

Neutrino mass

Another obvious problem of the SM is the mass of the neutrino. It is known that neutrinos in the Standard Model are left-handed particles. The non-zero mass term in the Dirac Lagrangian automatically leads to the presence of a right-handed fermion term. This looks strange, because right-handed neutrinos do not participate in any interaction. Meanwhile, a bunch of different neutrino experiments confirmed the existence of neutrino oscillations, meaning that the physical eigenstates of neutrinos are rather the mass eigenstates than the flavour eigenstates. The neutrino flavours are mixed by the *Pontecorvo-Maki-Nakagawa-Sakata* (PMNS) matrix, which is similar to the Cabibbo-Kobayashi-Maskawa (CKM) matrix in the QCD, but has a flat structure, rather than the diagonal structure as the CKM matrix. The current measurements of the relative magnitudes of the masses between different generations, Δm^2_{ij}, show that $\Delta m^2_{ij} \sim 10^{-5} - 10^{-3}$ eV. It raises another

question: why the neutrino mass scale is significantly lower than the mass scale of other fermions.

Hierarchy problem

One more open question, more of a theoretical nature, is the *hierarchy problem*. The experimentally measured mass of the Higgs boson is around 125 GeV [21, 22], meaning that the electroweak symmetry breaks at the scale $\mathcal{O}(100\text{ GeV})$. Meanwhile, a fermionic quantum one-loop correction to the Higgs mass is proportional to the scale Λ at which the SM breaks down [66]:

$$(m_h^0)^2 = m_h^2 + \frac{\Lambda^2}{8\pi^2 v^2}\left[3m_h^2 - 12m_t^2 + 6m_W^2 + 3m_Z^2\right] . \tag{2.28}$$

It is known that the SM breaks at the latest at the Planck scale $\mathcal{O}(10^{19}\text{ GeV})$, which naively implies a huge quantum correction $\mathcal{O}(10^{38}\text{ GeV}^2)$. This correction should be cancelled out by the Higgs bare mass to arrive at the known observed value $m_h \approx 125$ GeV. The electroweak breaking mechanism itself fails to explain this difference between the electroweak breaking scale and the gravitational scale. The possible explanation could have been the Higgs propagator loop corrections, which can shift the Higgs mass to the Planck scale, but it has not been confirmed experimentally. This leads to either severe fine-tuning of the Higgs mass or some new kind of physics beyond the SM, which cancels out the divergences in Higgs mass loop corrections and brings it to the observed value.

Chapter 3

Dark Matter

The first observations indicating that the Universe consists not only of ordinary matter have been made in the 1930s [25]. At the end of 1990s it was shown that the contrubution of ordinary matter is only around 5% [26, 27], which leads to the exciting fact that 95% of the Universe is basically "unknown what". This "unknown what" consists of two main components, one of which tends to cluster, meaning that it at least interacts gravitationally, and other is responsible for the expansion of the Universe. The first component is called the *Dark Matter* (DM), which comprises almost a quarter of the mass-energy of the Universe. The second component refers to the *Dark Energy* (DE) and its contribution is about 70%. The word "dark" reflects the fact that they are not visible to us using most of the known experimental techniques. They both do not radiate photons (or in case of DM extremely weakly), do not interact strongly and even weakly (might in case of DM). While the DE is believed to be a homogenous vacuum-like energy spread all over the Universe, everything indicates that DM consists of yet undiscovered (meta)stable particles, which do not fit in the scope of the Standard Model. Moreover, since it was shown that DM can be potentially produced at the energy scales of the hadron colliders [67], it immediately drew and still draws a lot of attention of the particle physics community.

The main observational evidence, in particular the galaxy rotation curves [68], gravitational lensing [69–71] and Cosmic Microwave Background (CMB) [26, 27], are discussed in Section 3.1. The main properties of DM are discussed in Section 3.2, following the discussion in [72]. Possible candidates for DM within and beyond the Standard Model (SM), which can be observed at the particle colliders are discussed in Section 3.3. The certain extensions of the SM to account for DM production are mentioned in details, such as simplified (axial-)vector [73, 74] and two-Higgs-doublet models [75–77], that are relevant for $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ searches covered in the scope of this work.

3.1 Observational evidences of Dark Matter

There are a lot of observational evidence for the existence of DM, but here the main focus will be on three of them: measurement of galaxy rotation curves, gravitational lensing, and measurement of the CMB anisotropy.

3.1.1 Galaxy Rotation Curves

The rotational circular velocity v_c of an object inside the galaxy satisfies the Newton law:

$$\frac{GM(r)}{r^2} = \frac{v_c^2}{r}\,, \tag{3.1}$$

where r is the distance to the galaxy center, $M(R) = 4\pi \int_0^R \rho(r) r^2 dr$ is the mass of the galaxy enclosed by the orbit of the object R. The profile of v_c forms the *galaxy rotation curve*. If the galaxy consisted only of luminous matter, then close to the galaxy center the rotational circular velocity would scale as $v_c \propto \sqrt{r}$, while far away from the center as $v_c \propto 1/\sqrt{r}$. But measurements of the galaxy rotation curves [68] reveal that v_c becomes approximately constant at large distances from the galaxy center, as shown in Figure 3.1. It means that the mass inside the galaxy should scale as $M(r) \propto r$, which contradicts the hypothesis of only luminous matter. This discrepancy can be explained by the existence of a galaxy halo, which consists of matter, which cannot be easily observed. The potential candidate could be extremely densed massive objects, which emit very little to no light, such as Massive Compact Halo Objects, like black holes or neutron stars, and Robust Associations of Massive Baryonic Objects like brown dwarfs. However, these possibilities mostly ruled out by gravitational lensing observations, CMB measurements and Big Bang nucleosynthesis [78].

The most commonly used density profiles of DM and their applications are discussed in [79–82].

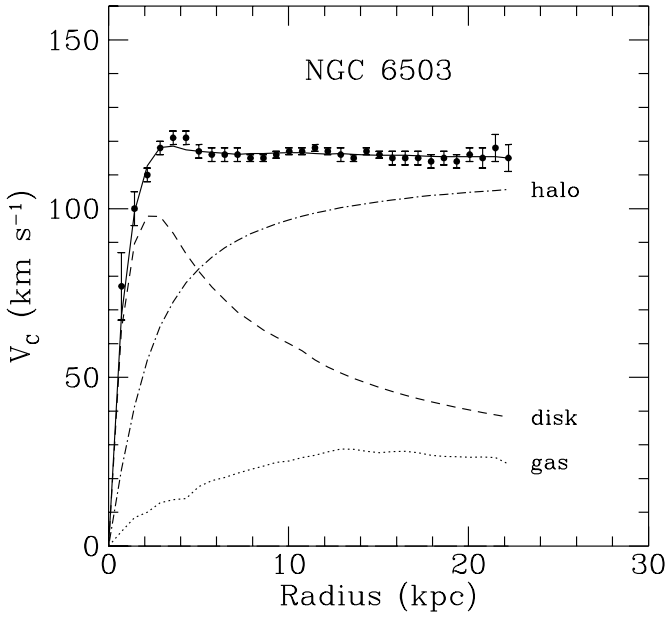


Figure 3.1: Three-parameter dark-halo fit (solid line) to the rotation curve of the galaxy NGC6503 (dots) [68]. The dashed curve shows the rotation curve for the visible matter component. The dotted curve shows the rotational curve for the gas component. The dash-dot curve shows the the dark-halo component. The parameters of the fit are the mass-to-light ratio of the disc, the halo core radius and the halo asymptotic circular velocity.

3.1.2 Gravitational lensing

The theory of the General Relativity states that every massive object curves the spacetime and thus distorts the light path. This effect is the most pronounced if the light travels near a dense concetration of mass, such as galaxy clusters, galaxy cores or the supermassive dense objects like black holes. If such a massive object is on the way between the light source and the observer, the space-time near it becomes so warped that it acts as a focusing gravitational lense. If the source, the gravitational lens, and the observer are aligned, and the lens is circular, then the source is observed in the form of so-called *Einstein ring* [83]. The angular radius of this ring depends on the mass of the lens as $\theta \sim \sqrt{M_{\text{lens}}}$. This is referred to as *strong lensing* [70]. If the alignment is not perfect or the lens is not circular, then the source can be visible from multiple locations.
Far from supermassive objects, the light is deflected weakly, but still enough for the objects to look stretched tangentially around the distorting mass, which is called *shearing*. For distant galaxies, the shearing changes their major-to-minor axis ratio by typically a few percents. It is too small deformation to be measured individually for each of the sources, however it can be obtained statistically from the coherent shearing of the adjacent sources. This is referred to as *weak lensing*. Examples of the strong and weak lensing are shown in Figure 3.2. Such coherent deformations can be used to estimate the mass distribution of the distorting object.

Figure 3.2: Strong gravitational lensing (left) around galaxy cluster CL0024+17 [69] and weak gravitional lensing around the Bullet Cluster (right). The yellow objects in the center of the left image belong to the cluster, all at similar redshift. The blue objects on the left image are distant galaxies behind the cluster, which appear as a series of the tangential arcs around the cluster due to the strong gravitational lensing. On the right image, most of the mass comes from DM (shown in blue), but not from the hot gas (shown in magenta).

One of the most vivid evidence of DM comes from the observation of weak gravitational lensing around the Bullet Cluster, which consists of two colliding galaxy clusters. Theories without DM like the Modified Newtonian Dynamics (MOND) predict the gravitational lensing caused mostly by baryonic matter [84]. In Bullet Cluster, the hot gas of the two colliding clusters comprises the majority of baryonic matter and thus should solely define the magnitude of the gravitational lensing. However, the lensing appears to be strongest in two separate regions coincident with the galaxies, as shown on the right image in Figure 3.2, which proves that the most of the galaxy mass in the cluster comes from the

additional matter component, DM.

3.1.3 Cosmic microwave background anisotropy

Before the Universe reached the age of 380000 years after the Big Bang, its temperature was still higher then the binding energy of electrons in the atoms. The ordinary matter was in the form of a plasma, composed of electrons, photons and baryons in a thermal equillibrium. Baryons at that times consisted mostly of protons ($\sim 75\%$) and ^{4}He nuclei ($\sim 25\%$). At some point the Universe cooled enough to allow protons and electrons form electrically neutral hydrogen atoms. This time in the history of the Universe is called *recombination* [72]. Before the recombination started, photons and electrons actively interacted with each other via the Compton scattering. After the recombination started, the densities of free protons and free electrons started to decrease, lowering the probability of the photon-electron scattering. At some point, when the density electrons degreaded enough and the photodisintegration of hydrogen atoms stopped, photons decoupled from the plasma and started propagating freely in the Universe. The decoupled photons after the last scattering with the matter are called *relic photons* and compose *cosmic microwave background* (CMB), which is basically a radiational footprint of the Universe, as it was around 380000 years old.
The temperature of the Universe at the time of the photon decoupling was $T = 0.27$ eV or around 3000 K. Due to the expansion of the Universe, free photons have been constantly cooling down. The measurements of the relic photon spectrum show that in the modern time it corresponds to the thermal black body spectrum at a temperature of $T_\gamma = 2.7255 \pm 0.0006$ K [26].
After subtracting the dipole moment coming from velocity of the observer with respect to the Universe, the CMB appears to be slightly anisotropic, with the fluctiations at the level of $10^{-6} - 10^{-5}$ K. The anisotropy can be described using the spherical harmonic expansion of the CMB sky and its angular power spectrum:

$$\begin{aligned} T(\theta,\phi) &= \sum_{lm} a_{lm} Y_{lm}(\theta,\phi) \\ C_l^{TT,\mathrm{obs}} &= \frac{1}{(2l+1)} \sum_{m=-l}^{m=+l} |a_{lm}^T|^2 \,, \end{aligned} \tag{3.2}$$

where $C_l^{TT,\mathrm{obs}}$ is an observed angular power spectrum per mode l. The main $l = 0$ mode corresponds to the CMB temperature T_γ itself. The largest anisotropy corresponds to the dipole $l = 1$ mode due to the peculiar velocity of the solar system with respect to the CMB, and is measured to be $3.3645 \pm 0.0020 \times$ K [26]. The higher multipoles are mostly caused by the small fluctuations in the density of the early Universe at the time of the photon decoupling and the photon interaction with the hot gas. The full map of the CMB anistropy is shown in Figure 3.3. The size of the anisotropy cannot be solely explained by the fluctations of baryonic density at time of the recombination and the interaction with the hot gas, but can be if considering the additional fluctuations of the density of DM. The constraints on many cosmological parameters, including DM relic density, can be obtained from the position and the amplitude of the peaks in the CMB power spectrum, as discussed in details in [26]. The measured relic density of DM from the CMB spectrum analysis is $\Omega_\chi h^2 = 0.1186 \pm 0.0020$ [26].

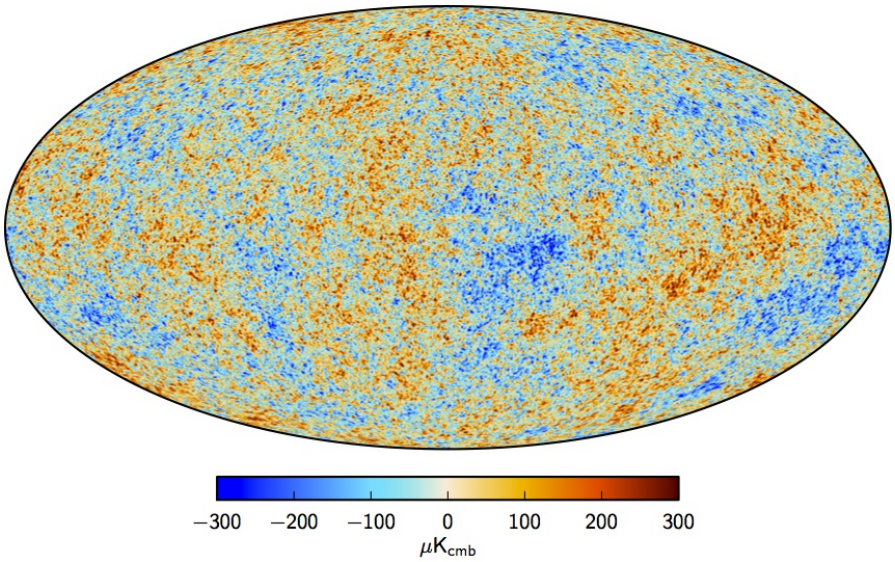


Figure 3.3: CMB intesity map of the sky based on the Planck and WMAP measurements [26].

3.2 Properties of Dark Matter

As mentioned before, all current observations indicate that the Dark Matter consists of non-relativistic particles, which might weakly interact with each other and with photons [1]. The latter means that DM could not radiate photons and cool down as baryonic matter, and its density on the periphery of the galaxy should not be much different from the density in the center. This can explain the rotation velocity profile of luminous matter, immersed in the galaxy dark halo.
Here the relevant DM generation mechanisms with the corresponponding particle candidates are considered. Note that even though the hypotheses discussed in this chapter are able to explain the observed DM relic densities, none of them is able to naturally explain the approximate relation between the relic density of DM and baryonic matter:

$$\rho_{\mathrm{B}} \approx 5\rho_{\mathrm{DM}}\,, \tag{3.3}$$

which has been satisfied since the early stages of the Universe.
When the Universe was hot enough, DM particles interacted strongly enough with ordinary matter and were in the thermal equilibrium with it. At some point due to the cooling of the Universe, DM particles decoupled from ordinary matter, stopped annihilating with each other, and started moving freely in the spacetime. This time of the decoupling is called *freeze-out*, which is characterised by the freeze-out temperature T_f of the Universe, when the decoupling occured. The relation between the freeze-out temperature and the mass of DM defines its main properties:

- $T_f \lesssim m_{\mathrm{DM}}$: so-called *Cold Dark Matter* (CDM). DM particles decoupled when they were non-relativistic and stayed the same until now.
- $T_f \gtrsim m_{\mathrm{DM}} \gtrsim 1$ eV: so-called *Warm Dark Matter* (WDM). DM particles decoupled when they were relativistic, but cooled down due to the Universe expansion and became non-relativistic.

[1] If DM particles self-interacted strongly, it would lead to the formation of the roughly spherical dark halos, which contradicts with observation of mostly elliptical shapes of the halos [85].

- $T_f \gtrsim m_{\rm DM} \lesssim 1$ eV: so-called *Hot Dark Matter* (HDM). DM particles decoupled when they were so relativistic that even the expansion of the Universe could not cool them to the non-relativistic state.

At the moment, the scenarious of the WDM and HDM are still considered, however they have problems explaining the observed large structure formation mechanisms. The most common hypothesis of the CDM is considered to be the case, but also neutrinos as the HDM are briefly mentioned.

In the thermal equilibrium before the decoupling, under the assumption of the dark matter-antimatter symmetry, the number density of the particles χ follows the Bolzmann equation [72]:

$$\frac{dn_\chi}{dt} = -3H(t)n_\chi - \langle\sigma_{\rm ann}v\rangle \left[n_\chi^2 - (n_\chi^{(\rm eq)})^2\right] . \tag{3.4}$$

In the deep equilibrium, the dominant term is $n_\chi^{(\rm eq)}$, which gives the following number density before the freeze-out:

$$n_\chi = n_{\bar\chi} = g_\chi \left(\frac{m_\chi T}{2\pi}\right)^{\frac{3}{2}} e^{-\frac{m_\chi}{T}}, \qquad T > T_f\,. \tag{3.5}$$

Given the stability of DM, its density can only decrease due to the DM self-annihilation. After the freeze-out, the DM density becomes low enough to stop the DM annihilation, and thus its number density stabilises ($dn_\chi/dt = 0$). Then the Bolzmann equation 3.4 after the freeze-out takes the form:

$$n_\chi \langle\sigma_{\rm ann}v\rangle \simeq H(T_f) \equiv \frac{T_f}{M^*_{\rm Pl}}, \quad M^*_{\rm Pl} = \frac{M_{\rm Pl}}{1.66\sqrt{g_f}} \approx 2.435\cdot 10^{18}\ {\rm GeV}\,, \tag{3.6}$$

where $H(T_f)$ is the Hubble parameter at the freeze-out temperature, $M^*_{\rm Pl}$ is an effective Planck mass, v is the relative velocity of DM and $\sigma_{\rm ann}$ is DM annihilation cross-section. Usually, $\chi\bar\chi$ annihilation happens in the $s-$wave, and which implies for the non-relativistic particles that $\sigma_{\rm ann}(v) = \sigma_0/v$ (see, e.g. [86]). Then, given the number density in Eq. 3.5, the freeze-out temperature T_f is defined from Eq. 3.5 as:

$$T_f = \frac{m_\chi}{\ln\left(\frac{g_\chi m_\chi M^*_{\rm Pl}\sigma_0}{(2\pi)^{3/2}}\right)}\,. \tag{3.7}$$

The calculations of the energy density leads to the following relation for the relic density of DM in the modern Universe [72]:

$$\Omega_\chi h^2 = \frac{2m_\chi n_\chi(t_0)}{\rho_c} = 3\cdot 10^{-10}\left(\frac{\rm GeV^{-2}}{\sigma_0}\right)\frac{1}{2\sqrt{g_*(t_f)}}\ln\left(\frac{g_\chi m_\chi M^*_{\rm Pl}\sigma_0}{(2\pi)^{3/2}}\right)\,, \tag{3.8}$$

where h denotes the ratio of the Hubble parameters now and at the time of the freeze-out, $h = H_0/100\ {\rm km\cdot s^{-1}\cdot Mpc^{-1}}$, g_* is the number of degrees of freedom and $\rho_c = 3H^2/8\pi G$ is the critical density of the Universe for which the spatial geometry is flat.

3.3 Dark Matter candidates

Let us summarise all criteria, mentioned in previous sections, which should be satisfied by a successfull DM candidate:

- The DM populates around 25% of total mass-energy composition of the Universe, which is five times more than ordinary matter.
- DM is stable, massive and electrically neutral. Interacts at least gravitaniolly and maybe weakly with the SM sector. The self-interaction is permitted to be at most at the weak scale.
- DM tends to cluster at the large scales, not forming any compact densed objects.
- DM is dynamically cold, meaning that it decoupled when it was non-relativistic and stayed the same until now.

Together with Eq. 3.8 for the relic density, this set of criteria determines a variety of different particles satisfying the properties of DM and its observed relic density.
The only possible candidate within the Standard Model is neutrino. Neutrinos are the hot DM, as they are relativistic before decoupling and stay relativistic until now due to its very low mass and extremely weak interaction with matter. The HDM itself would suppress the local matter perturbations and thus the clustering on the small scales, which would lead to a different galaxy formation than is observed now [87, 88]. On a top of this, they cannot solely account for the observed DM relic density. The neutrino relic density, given the number density $n_\nu = 336\ cm^{-3}$ today, takes the form [54]:

$$\Omega_\nu h^2 = \frac{n_\nu \sum\limits_{e,\mu,\tau} m_\nu}{\rho_c} = \frac{\sum\limits_{e,\mu,\tau} m_\nu}{93.14\ \text{eV}}\,. \tag{3.9}$$

The upper bound on the total mass of neutrinos $\sum\limits_{e,\mu,\tau} m_\nu < 0.68$ measured by the Planck satellite [26] gives the upper bound on the neutrino relic density $\Omega_\nu h^2 \lesssim 0.007$, which is by far smaller than the observed DM relic density.

3.3.1 Weakly interacting massive particles (WIMPs)

The modern DM relic density is known with a very high precision from the Planck satellite measurements, $\Omega_\chi h^2 = 0.1186 \pm 0.0020$ [26]. Considering that DM consists of the new heavy (meta)stable particles [89], it can be shown (see e.g. [72]) that DM particles with the masses around the EW scale that are potentially accessible at the particle colliders can satisfy the observed relic density:

$$10\ \text{GeV} \lesssim m_\chi \lesssim \text{few TeV}\,. \tag{3.10}$$

The logarithm in Eq. 3.8 is a slowly varying function of m_χ. Its value can be estimated using the range of masses in Eq. 3.10 and the simplest dimensional approximation $\sigma_0 \sim 1/m_x^2$:

$$\ln\left(\frac{g_\chi M_{\text{Pl}}^*}{(2\pi)^{3/2} m_\chi}\right) \approx 30 - 40\,. \tag{3.11}$$

Given the $\sqrt{g_*(T_f)} \sim 1 - 10$ in Eq. 3.8 for $T_f \approx m_\chi/30$ [90] and the observed value of DM relic density, the following estimate of DM annihilation cross-section can be obtained from Eq. 3.8:

$$\sigma_0 \sim 10^{-8}\ \text{GeV}^{-2}\,. \tag{3.12}$$

This is quite intriguing result, meaning that the DM particles with masses at the EW scale can annihilate with a characteristic cross-section of the weak interaction, $\sigma_{\mathcal{EW}} \sim \alpha^2_{\mathcal{EW}}(m_W)/m^2_W \sim 10^{-7}$ GeV^{-2}. Thus it makes them potentially observable at particle colliders. Also this rough estimate gives the lower bound on the annihilation cross-section of the hypothetical new particles in the different extensions of the SM, otherwise the abundance of these particles would be higher than the observed abundance of the non-relativistic matter in the Universe. These particles are usually referred to the *weakly interacting massive particles* (WIMPs).
The fact that the WIMPs with masses at the electroweak scale can naturally expain the observed DM relic density via the freeze-out process is called the *WIMP miracle*. The particles with a given mass range, which take part in the weak interaction, are also the subject of particular interest to the supersymmetrical extensions of the Standard Model, where they can be associated with the lightest neutralino.

3.3.2 Other Dark Matter candidates

Only the most relevant DM candidates for the collider searches are discussed in this chapter, but it is worth briefly mentioning a couple of other interesting possibilities:

- **Axions and axion like particles (ALPs):** Originally axions were introduced in terms of PecceiQuinn theory as an extension of the SM sector to solve the strong CP problem. They are non-relativistic, couple extremely weakly to the photons, $g_{a\gamma\gamma} \lesssim \mathcal{O}(10^{-12}\ \text{GeV}^{-1})$, expected to have a tiny mass, $m_a \sim 10^{-6}$ eV $(10^{12}\ \text{GeV}/f_a)$ and can be interesting DM candidates [91, 92].

- **Sterile neutrinos:** right-handed neutrinos, which interact only gravitationally. Significantly heavier than SM neutrinos. With the masses at the keV scale, they are good DM candidates [93].

- **The Kaluza-Klein particle:** the lightest Kaluza-Klein exitation of neutral electroweak gauge bosons is a weakly-interacting particle (WIMP) in models with extra dimensions [94].

- **Lightest Supersymmetric Particle (LSP):** a supersymmetric DM WIMP candidate, the lightest neutralino with the mass range 10 GeV $\lesssim m_{\tilde{\chi}_0} \lesssim$ 10 TeV [95].

3.4 Dark Matter production at particle colliders

Generally, the current experimental methods to search for DM particles can be classified into three big groups, as shown on the Feynman graphs in Figure 3.4:

- *Direct detection* (DD) searches, which probe the direct interaction between DM particles and ordinary matter.

- *Indirect detection* (ID) searches, which look for the possibility to produce the SM particles via the interaction between DM particles.
- *Collider* searches, which look for DM production in the collisions of SM particles.

All three types of experiments are discussed in this chapter with the main focus on the WIMP searches at the particle colliders, possible final states and corresponding DM models.

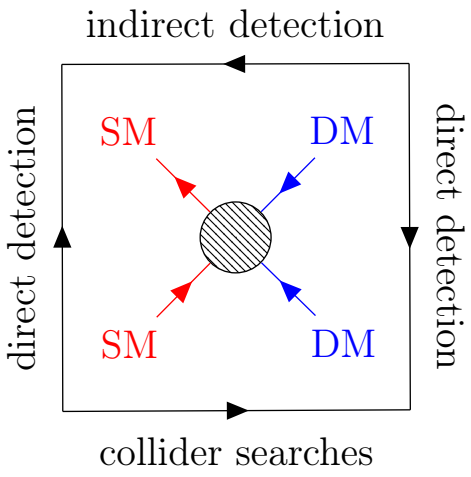


Figure 3.4: The three types of DM searches shown on one Feynman graph. DM-SM scattering is the subject of direct detection. DM annihilation into the SM particles is the subject of indirect detection. SM annihilation into DM particles is the subject of collider searches.

Direct detection

Since DM starts to move freely in the Universe, it rarely interacts with ordinary matter via the elastic scattering on the atomic nuclei. The *direct detection* experiments are built to detect the nuclear recoil of this kind of scattering. The expected event rate, R, is proportional to DM-nucleus scattering cross-section $\sigma_{\chi N}$ [96]:

$$R = \frac{n_\chi N_A}{A} \langle \sigma_{\chi N} v \rangle \sim \sigma_{\chi p} \langle v \rangle A \,, \tag{3.13}$$

where n_χ and v are the number density and the velocity of DM respectively, A is the atomic mass of the material and $\sigma_{\chi p}$ is DM-proton scattering cross-section. For a detector material with a high atomic mass $A \sim 100$, a WIMP mass range in Eq. 3.10, a most probable value for the velocity of DM particles $v \approx 220$ km/s [97] and an expected WIMP-nucleon cross-section for a given material, one can estimate the expected event rate as:

$$R \approx 0.1 \, \frac{\text{events}}{(\text{year} \cdot \text{kg})} \,. \tag{3.14}$$

Usually the detector can be composed of the scintillating crystals like NaI (DAMA/LIBRA [98]), germanium or silicon (CRESST II [99], CDMS [100]) or liquid noble gases like xenon and argon (LUX [101], XENON [102]). The summary of the measurements from the DD experiments and their projected improvements are shown on the top plot

in Figure 3.5 with the current best limit from the XENON1T experiment [103]. The current DD measurements are probing the whole range of the WIMP masses, closely approaching so-called *neutrino floor*, where the coherent elastic neutrino-nucleus scattering substantially limits the sensitivity.

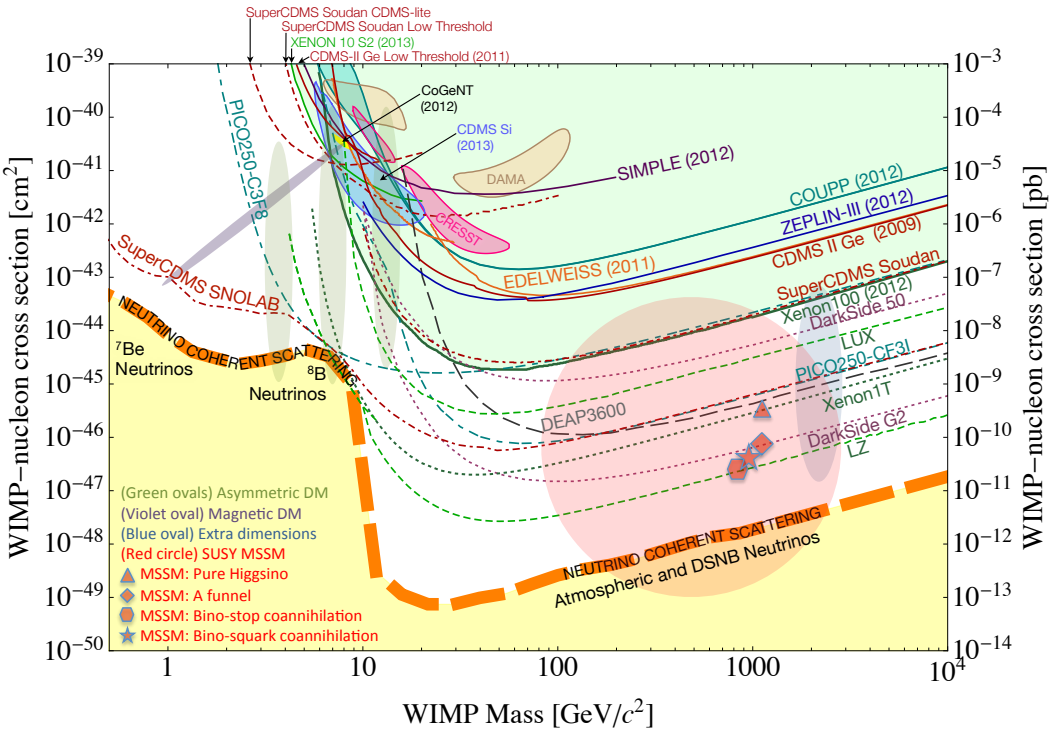


Figure 3.5: Summary plot of the exclusion limits on the WIMP-nucleon spin-independent cross-section is shown [104]. The exclusion limits are shown as solid lines. The possible WIMP signals are shown as shaded closed contours. The projections for the DD experiments to be operated in the future are shown as dotted and dashed lines. The yellow band shows the limit of the sensitivity of the DD experiments due to the coherent neutrino scattering.

Indirect detection

Indirect detection experiments aim to detect the SM products of the annihilation of DM particles, as shown on the Feynman diagram in Fig 3.4. Such annihilation processes are expected to happen mostly in regions with the high DM density, e.g. galaxy cores, where DM density is potentially large enough to initiate the pair interaction. Current experiments are looking for DM annihilation products in the gamma rays, like space-based Fermi Large Area Telescope [105] and ground-based H.E.S.S. telescopes [106], in neutrino fluxes like ANTARES [107] or IceCube [108], and in cosmic rays like AMS [109]. The summary of the measurements from the ID experiments is shown on the bottom plot in Figure 3.6 with the current best limits from the gamma ray detection experiments.

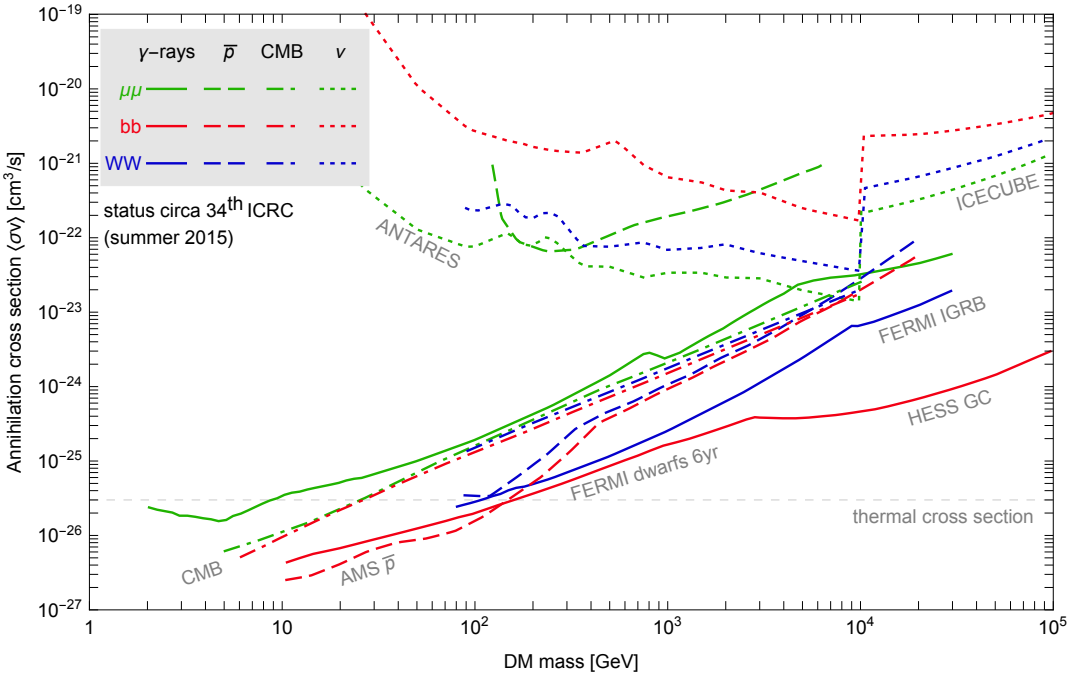


Figure 3.6: Summary plot of the most stringent exclusion limits on the WIMP annihilation cross-section in different channels for various ID experiments is shown as for summer 2015 [110].

3.4.1 Simplified models of Dark Matter production with $E_{\mathrm{T}}^{\mathrm{miss}}+X$ signatures at particle colliders

Simplified models of DM production are the simplest comlete extensions of the SM sector. These models follow the following minimal requirements, which characterise its particle content and corresponding interactions [74]:

- The model should contain (meta)stable DM candidates and a mediator, which couples to the SM and DM sectors. The number of the new particles should be minimal, meaning that the degrees of freedom of all but the lightest particles are integrated out. Such a model can be considered as a limit of a more general theory.

- The model should be renormalizable and Lorentz invariant. The additional interaction terms between the SM and the dark sectors should not violate the exact and approximate accidental global SM symmetries. It means that the baryon and lepton numbers should be conserved, and the flavour symmetries should at most weakly broken to avoid strong constraints from the flavour physics.

In what follows, the simplified models for the searches for DM presented in this thesis are discussed.

3.4.1.1 Vector and axial-mediator simplified models

Consider a model with a new leptophobic spin-1 mediator Z' of the additional $U(1)_{Z'}$ symmetry, which is exchanged in the $s-$channel between the SM and the dark sectors. The mixing between Z' and the SM Z boson is assumed to be negligible. "Leptophobic" means that the mediator couplings to the leptons are considered to be either zero or very small to avoid strong constraints from the di-lepton searches at the LHC. The generic

interaction term between the SM and dark sectors in the Lagrangian of such a model then takes the form [73]:

$$\mathcal{L} \supset \bar{\chi}\gamma^{\mu}(g_{\chi}^{V} - g_{\chi}^{A}\gamma_5)\chi Z'_{\mu} - \sum_{q=u,d,s,c,b,t} \bar{q}\gamma^{\mu}(g_{q}^{V} - g_{q}^{A}\gamma_5)qZ'_{\mu}\,, \tag{3.15}$$

where g_{χ}^{V}, g_{χ}^{A} are the *vector* and *axial-vector* mediator couplings to DM respectively, and g_{q}^{V}, g_{q}^{A} are the universal vector and axial-vector mediator couplings to the quarks respectively. It is important to keep the quark couplings $g_{q}^{V,A}$ universal to satisfy the *Minimal Flavour Violation* (MFV) [111], meaning that all higher-dimensional operators constructed from the SM fields and new $U(1)$ fields should stay CP-invariant. The only source of the CP-violation should come from the CKM phase, otherwise it would lead to the strong constraints from the flavour physics.

Consider two types of simplified models, vector mediator, where $g_{\chi,q}^{A} = 0$, and axial-vector mediator, where $g_{\chi,q}^{V} = 0$. The interaction terms of the Lagrangian of the vector and axial-vector mediator models take the form:

$$\begin{aligned} \mathcal{L}_{\mathcal{V}} &\supset -g_{\chi}\bar{\chi}\gamma^{\mu}\chi Z'_{\mu} - g_{q}\sum_{q=u,d,s,c,b,t}\bar{q}\gamma^{\mu}qZ'_{\mu} \\ \mathcal{L}_{\mathcal{AV}} &\supset -g_{\chi}\bar{\chi}\gamma^{\mu}\gamma_5\chi Z'_{\mu} - g_{q}\sum_{q=u,d,s,c,b,t}\bar{q}\gamma^{\mu}\gamma_5qZ'_{\mu}\,. \end{aligned} \tag{3.16}$$

Each model can be characterised by four parameters, which are the masses of DM particle m_{χ} and the mediator m_{med}, and the mediator couplings to the SM sector g_q and to DM sector g_{χ}.

Define the minimal width of the mediator as a sum of the partial widths of all kinematically possible decays into DM and the quarks. Then for the vector and axial-vector mediator, the minimal widths to DM and to quarks can be obtained as [73]:

$$\begin{aligned} \Gamma_{\mathcal{V}}^{\chi\bar{\chi},q\bar{q}} &= \frac{g_{\chi,q}^{2}m_{\text{med}}}{12\pi}\left(1 - \frac{4m_{\chi,q}^{2}}{m_{\text{med}}^{2}}\right)^{1/2}\left(1 + \frac{2m_{\chi,q}^{2}}{m_{\text{med}}^{2}}\right) \\ \Gamma_{\mathcal{AV}}^{\chi\bar{\chi},q\bar{q}} &= \frac{g_{\chi,q}^{2}m_{\text{med}}}{12\pi}\left(1 - \frac{4m_{\chi,q}^{2}}{m_{\text{med}}^{2}}\right)^{3/2}\,. \end{aligned} \tag{3.17}$$

All minimal widths vanish in the part of the phase space where $m_{\text{med}} < 2m_{\chi,q}$, and only off-shell decays are allowed. The *kinematic limit*, $m_{\text{med}} = 2m_{\text{DM},q}$, separates between the on-shell and the off-shell decays of the mediator. More about the on-shell and off-shell regimes, as well as corresponding constraints on the LHC searches, are described in Chapter 11.

One can consider models, in which the SM final state particles originate from the initial state radiation of gluons or SM vector bosons. The corresponding Feynman diagrams of such processes are shown in Figure 3.7. These models can be probed by the $E_{\text{T}}^{\text{miss}} + j$, $E_{\text{T}}^{\text{miss}} + \gamma$ and $E_{\text{T}}^{\text{miss}} + V$ searches at the LHC. These searches become complementary, when the set of the fixed model parameters is chosen to be the same. More details about the choice of the parameters are discussed in Chapter 11.

It is important to note that the same simplified models can be probed by the measurements from the DD and ID experiments. It means that there is a correspondence between the LHC and the DD/ID cross-section measurements. For instance, for the DD

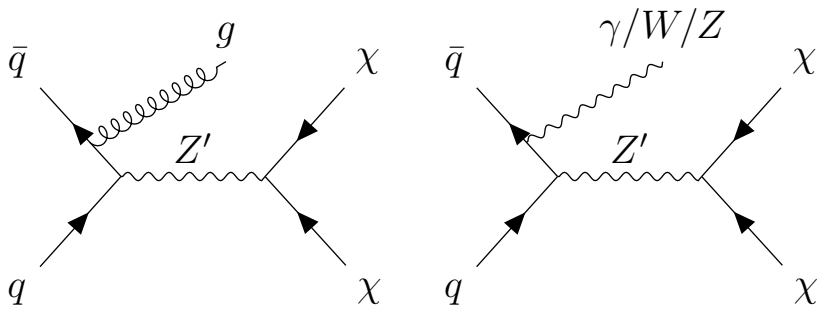


Figure 3.7: Tree-level Feynman diagrams of DM production for the vector and axial-vector simplified models.

searches in case of the vector mediator model, the LHC cross-section corresponds to the *spin-independent* WIMP-nucleon cross-section $\sigma^{\mathrm{SI}}_{\chi N}$, while In case of axial-vector mediator model to the *spin-dependent* WIMP-nucleon cross-section $\sigma^{\mathrm{SD}}_{\chi N}$. Any LHC limit on the parameters of the model can be translated into the corresponding DD/ID limits [112], which makes the collider and the DD/ID measurements complementary.

3.4.1.2 Z'-2HDM simplified model

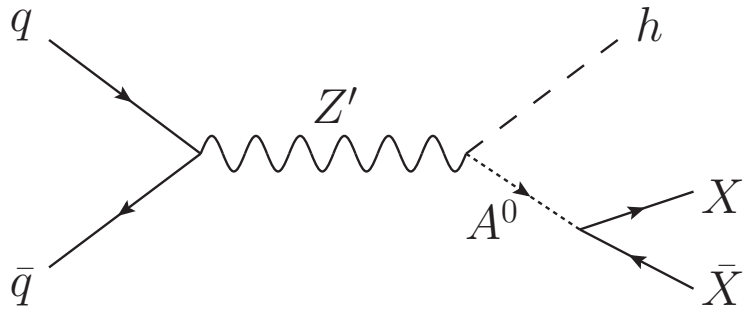


Figure 3.8: The Feynman diagram of DM production in association with the SM Higgs boson in terms of Z'-2HDM model [113].

Since the initial state radiation of the Higgs boson is highly suppressed, one cannot use exactly the simplified models with the initial state radiation that described in Section 3.4.1.1. Instead, assume that the Z' mediator decays to the Higgs boson and to some intermediate particle, which in turn decays primarily to DM pair, $Z' \to hA^0 \to h\chi\bar{\chi}$ [113], as shown on the Feynman diagram in Figure 3.8. The Z' mediator obeys exactly the same $U(1)_{Z'}$ symmetry. Note that Z' vector mediator model described in Eq. 3.15, where the mediator radiates the Higgs boson and no additional intermediate particles are required, can also produce the $E_{\mathrm{T}}^{\mathrm{miss}} + h$ final state, but it is not discussed in the scope of this work. The new particle cannot be added as a singlet state, since the new terms in the Lagrangian would violate gauge invariance. Thus, it is added via a two-Higgs doublet extension, meaning that the dark sector includes a new SM singlet scalar, which spontaneously breaks the $U(1)_{Z'}$ symmetry at the scale above the electroweak symmetry breaking.
This model assumes Type 2 two-Higgs doublets [75], where one of the doublet ϕ_u couples only to the up-type quarks and the other doublet ϕ_d couples only to the down-type quarks

and the leptons. The relevant terms in the Lagrangian then take the form:

$$\mathcal{L} \supset y_u \bar{Q} \phi_u u + y_d \bar{Q} \phi_d d + y_l \bar{L} \phi_d l + V_{\text{2HDM}} \,, \tag{3.18}$$

where y_u, y_d and y_l are the Yukawa couplings of the new Higgs doublets to the fermions and leptons.
The two-Higgs doublet potential V_{2HDM} is required to be Lorentz invariant, CP conserving and to obey all symmetries of the electroweak sector. The following combination of the doublet fields satisfy the abovementioned requirements:

$$\begin{aligned} V_{\text{2HDM}} = {} & \mu_1 \phi_u^\dagger \phi_u + \mu_2 \phi_d^\dagger \phi_d - \mu_3 \left(\phi_u^\dagger \phi_d + \phi_d^\dagger \phi_u \right) + \frac{\lambda_1}{2} \left(\phi_u^\dagger \phi_u \right)^2 + \frac{\lambda_2}{2} \left(\phi_d^\dagger \phi_d \right)^2 \\ & + \lambda_3 \left(\phi_u^\dagger \phi_u \right) \left(\phi_d^\dagger \phi_d \right) + \lambda_4 \left(\phi_u^\dagger \phi_d \right) \left(\phi_d^\dagger \phi_u \right) + \frac{\lambda_5}{2} \left(\phi_u^\dagger \phi_d \right)^2 + \frac{\lambda_5}{2} \left(\phi_d^\dagger \phi_u \right)^2 . \end{aligned} \tag{3.19}$$

This potential includes a Z_2 symmetry under which $H_u \to H_u$ and $H_d \to H_d$ to suppress flavour-changing neutral currents (FCNCs). This symmetry is softly broken by the term $\mu_3 \left(\phi_u^\dagger \phi_d + \phi_d^\dagger \phi_u \right)$. The requirement that all coefficients in the potential as well as the vacuum expectation values (VEVs) are real ensures the CP conservation.
After the symmetry breaking, the new doublets acquire the VEVs v_u and v_d, leading to the following parametrisation:

$$\begin{aligned} \phi_u &= \begin{pmatrix} -H^+ \sin\beta \\ v_u - h \sin\alpha + H \cos\alpha - i A^0 \sin\beta \end{pmatrix} \\ \phi_d &= \begin{pmatrix} H^+ \cos\beta \\ v_d + h \cos\alpha + H \sin\alpha + i A^0 \cos\beta \end{pmatrix} , \end{aligned} \tag{3.20}$$

where h, H are the CP-even scalars with the masses m_h, m_H respectively, H^+ is a heavy charged scalar with the mass m_{H^+} and the corresponding antiparticle H^-, A^0 is the CP-odd with the mass m_{A^0}, the angle β is related to the ratio of VEVs as $\tan\beta \equiv v_u/v_d$ and α is the mixing angle between the CP-even scalars h and H.
The scalar h can be associated with the SM Higgs boson with the mass $m_h \approx 125$ GeV. The other parameters of the new (pseudo)scalars H, $H^\pm$ and A^0 can be constrained from the other measurements:

- $m_{H,H^\pm,A^0} \gtrsim 300$ GeV from the $b \to s\gamma$ process [75].
- $m_{H^\pm} \approx m_H$ or $m_{H^\pm} \approx m_{A^0}$ to protect the relation $\rho \equiv \frac{m_Z}{m_W \cos\theta_W^2} = 1$ [54] against the 1-loop corrections to the masses of the SM W and Z bosons.
- The mixing angles are constrainted to be around the *alignment limit* $\sin(\alpha-\beta) \to 1$ or $\beta \to \alpha + \pi/2$, $\alpha \in (-\pi/2, 0)$ from the fits to the observed Higgs couplings [114].
- $\tan\beta \gtrsim 0.3$ to ensure the perturbativity of the top Yukawa coupling [75].
- The coupling of the Z' to the quarks, g_q, is constrained from the ρ parameter measurements for $\tan\beta \gtrsim 0.6$ and from the dijet measurements for $\tan\beta \lesssim 0.6$ for the mediator masses $m_{Z'} \lesssim 1.3$ TeV [113].
- The additional Higgs VEVs lead to the small $Z - Z'$ mixing, which modifies the SM Z mass. The mixing is strongly constained from the measurements of the ρ_0 parameter and LEP measurements of the Z resonances [113].

The interaction term between DM particle χ and the new extension of the Higgs sector is the following:

$$\mathcal{L} \supset y_\chi \bar{\chi} i\gamma_5 \chi A^0 \,, \tag{3.21}$$

where y_χ is the Yukawa coupling of DM particle to the pseudoscalar A^0. It is assumed to be the only interaction term of the particle A^0 to other particles in the model, meaning large branching ratio $\mathcal{B}_{A^0 \to \chi\bar{\chi}} = 1$.
The width of the decay of the mediator to the SM Higgs boson and the new particle A^0 can be expressed as [113]:

$$\begin{aligned} \Gamma_{Z' \to hA^0} &= [g_{Z'} \cos\alpha \cos\beta]^2 \frac{|p|^3}{24\pi m_{Z'}^2} \\ |p| &= \frac{1}{2m_{Z'}^2} \sqrt{\lambda(m_{Z'}^2, m_h^2, m_{A^0}^2)} \,, \end{aligned} \tag{3.22}$$

where $|p|$ is the center of mass momentum of the decay products and $\lambda(x,y,z) = x^2 + y^2 + z^2 - 2xy - 2xz - 2yz$ is the Källen triangle function. A non-zero decay width $\Gamma_{Z' \to hA^0}$ leads to the Jacobian peak in the momentum distribution of DM pair, or the $E_{\mathrm{T}}^{\mathrm{miss}}$.

3.4.1.3 Two-Higgs doublet model with pseudoscalar mediator

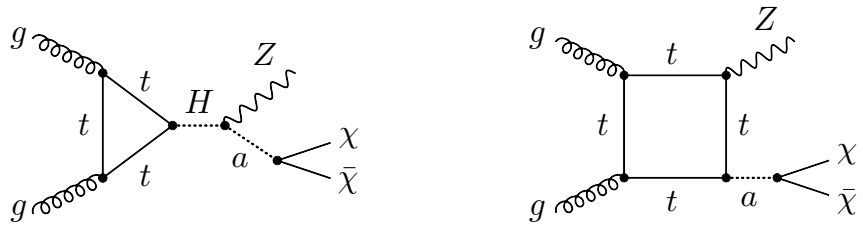


Figure 3.9: The Feynman diagrams of the $E_{\mathrm{T}}^{\mathrm{miss}} + Z$ production in 2HDM+a model. The left diagram is the only triangular diagram which contributes. The box diagram on the right allows both a and A exchange. H is a CP-even scalar, a and A are the CP-odd scalars.

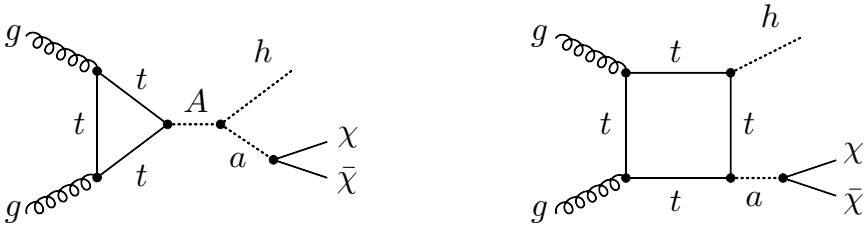


Figure 3.10: The Feynman diagrams of the $E_{\mathrm{T}}^{\mathrm{miss}} + h$ production in 2HDM+a model. The left diagram is the only triangular diagram which contributes. The box diagram on the right allows both a and A exchange. H is a CP-even scalar, a and A are the CP-odd scalars.

This class of models is particularly interesting by its rich $E_{\mathrm{T}}^{\mathrm{miss}} + X$ phenomenology. It provides a wide range of different collider signatures, including $E_{\mathrm{T}}^{\mathrm{miss}} + V$, $E_{\mathrm{T}}^{\mathrm{miss}} + h$, $E_{\mathrm{T}}^{\mathrm{miss}} + j$, $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ and $E_{\mathrm{T}}^{\mathrm{miss}} + b\bar{b}$ [76]. Unlike the Z' simplified models, this model leads to the SM Z production directly from the hard interaction, probing directly the

interaction between the DM and SM sectors. It can lead to stronger constraints on the model parameters for the $E_{\mathrm{T}}^{\mathrm{miss}}+V$ and $E_{\mathrm{T}}^{\mathrm{miss}}+h$ searches, considered in the scope of this thesis, comparing to the $E_{\mathrm{T}}^{\mathrm{miss}}+j$ search, which usually dominates in case of Z' simplified models.
The two-Higgs doublet model with pseudoscalar mediator, or 2HDM+a, assumes the interaction between DM and the SM particles is goverened by the CP-odd spin-0 mediator. The two new Higgs doublets and the corresponding potential are introduced in the same way as described in Eq. 3.19-3.20.
The simplest way to introduce the interaction via the pseudoscalar P is to mix it with the CP-odd scalar from the Higgs doublets [76]:

$$\mathcal{L} \supset P\left(ib_P\phi_u^\dagger\phi_d + h.c.\right) + P^2\left(\lambda_{P1}\phi_u^\dagger\phi_u + \lambda_{P2}\phi_d^\dagger\phi_d\right), \tag{3.23}$$

where b_P is a mass-dimensional CP conserving parameter and $\lambda_{P1,P2}$ are trilinear scalar couplings. With this new term in the Lagrangian, the pseudoscalars P and A^0 are no longer the mass eigenstates. They can be diagonalised using a mixing angle θ:

$$\begin{pmatrix} a \\ A \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} P \\ A^0 \end{pmatrix}. \tag{3.24}$$

Both a and A can couple to DM particles, to the SM sector and to other scalars from the Higgs doublets. Consider the possibilities, which are relevant for the $E_{\mathrm{T}}^{\mathrm{miss}}+h$ and $E_{\mathrm{T}}^{\mathrm{miss}}+Z$ searches.

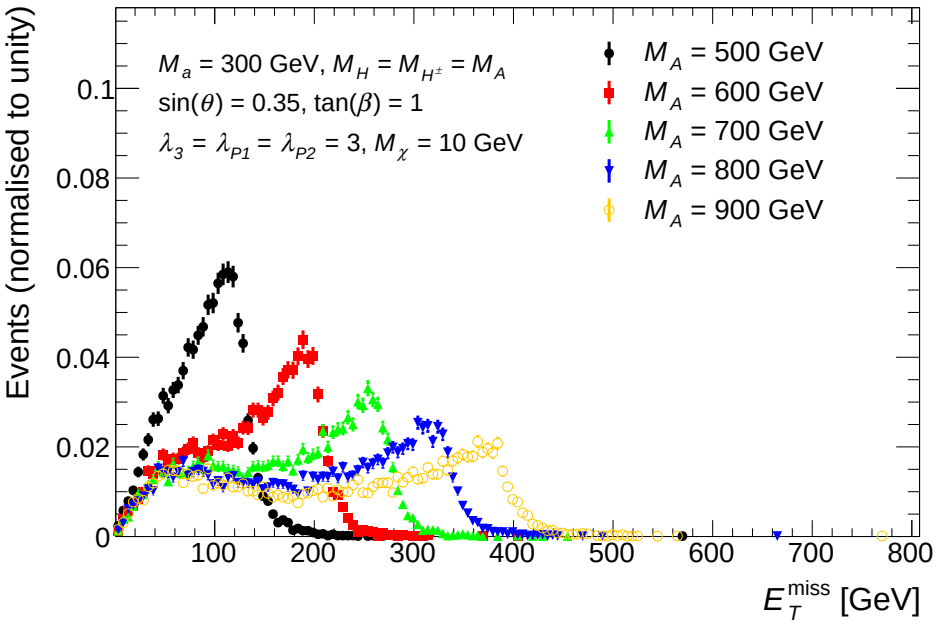


Figure 3.11: $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution at parton level of the $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ signal events for 2HDM+a models with different masses of the $CP-$odd scalar boson A [77].

- **$\mathbf{E_{\mathrm{T}}^{\mathrm{miss}}}$ + Z search**: heavy scalar H couples to the SM and the dark sector, while a couples directly to DM particles. In case of $m_H > m_a + m_Z$, the $H \to aZ$ decay is allowed, where a decays to DM pair, giving rise to the $E_{\mathrm{T}}^{\mathrm{miss}}+Z$ resonant production, as shown in Figure 3.9. The triangle diagram leads to a Jacobian peak

in the momentum distribution of DM pair, or the $E_{\mathrm{T}}^{\mathrm{miss}}$. The position of the peak can be approximated as [115]:

$$(E_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{max}} \approx \frac{\sqrt{(m_H^2 - m_a^2 - m_Z^2)^2 - 4m_a^2 m_Z^2}}{2m_H} \,. \tag{3.25}$$

Higher m_H shifts the peak to the harder $E_{\mathrm{T}}^{\mathrm{miss}}$, while higher m_a shifts it to the softer $E_{\mathrm{T}}^{\mathrm{miss}}$. As soon as certain $E_{\mathrm{T}}^{\mathrm{miss}}$ threshold, $(E_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{cut}}$, is required in the $E_{\mathrm{T}}^{\mathrm{miss}} + Z$ search, the more strict constraint on the m_H is needed to enhance the sensitivity, $m_H \gtrsim m_a + \sqrt{m_Z^2 + (E_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{cut}}^2}$ [76].

The relevant partial decay widths take the form [76]:

$$\begin{aligned} \Gamma_{a\to\chi\bar{\chi}} &= \frac{y_\chi^2 m_a \cos^2\theta}{8\pi}\sqrt{1 - \frac{4m_\chi^2}{m_a^2}} \\ \Gamma_{H\to aZ} &= \frac{\sin^2\theta}{16\pi}\frac{[(m_H^2 - m_a^2 - m_Z^2)^2 - 4m_a^2 m_Z^2]^{3/2}}{m_H^3 v^2} \,. \end{aligned} \tag{3.26}$$

- **$\mathbf{E_{\mathrm{T}}^{\mathrm{miss}} + h}$ search**: heavy pseudoscalar A couples to the SM and the dark sector, while the light pseudoscalar a couples directly to DM particles. In case of $m_A > m_a + m_h$, the $A \to ah$ decay is allowed, giving rise to the $E_{\mathrm{T}}^{\mathrm{miss}} + h$ resonant production, as shown in Figure 3.10. The triangle diagram leads to the same Jacobian peak in $E_{\mathrm{T}}^{\mathrm{miss}}$, where m_H is replaced with m_A in Eq. 3.25, as shown in Figure 3.11.

  The relevant partial decay width takes the form [76]:

$$\begin{aligned} \Gamma_{A\to ah} &= \frac{g_{Aah}^2}{16\pi}\frac{\sqrt{(m_A^2 - m_a^2 - m_h^2)^2 - 4m_a^2 m_h^2}}{m_A} \\ g_{Aah} &= \frac{1}{m_A v}\Big[m_h^2 - m_A^2 - m_a^2 - 2m_H^2 + 4m_{H^\pm}^2 \\ &\quad + 2v^2\left(-\lambda_3 + \lambda_{P1}\cos^2\beta + \lambda_{P2}\sin^2\beta\right)\Big]\sin\theta\cos\theta \,. \end{aligned} \tag{3.27}$$

More details about the different parameter scenarios, kinematic regimes and allowed parameter space for different searches can be found in [76, 77, 115].

Chapter 4

Experimental Setup

4.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is the world's largest and most powerful particle accelerator since it became operational in 2009. It is located at the European Organization for Nuclear Research (CERN [1]) near Geneva, Switzerland. With a circumference of 26.7 km, it is aimed to accelerate protons [2] to the unprecedented center-of-mass energy of up to 14 TeV and collide them at the interaction points.

LHC is the final stage of the complex multi-step accelerator system used to bring protons to the final energy and deliver them to the different experiments, as shown in Figure 4.1. The first accelerator in this system, LINAC 2, increases the energy of the protons to 50 MeV and injects them into the Proton Synchrotron Booster (PSB). The protons are accelerated to the energy of 1.4 GeV in the PSB, filled into the Proton Synchrotron (PS) and brought up to the next energy level of 25 GeV. Finally, the protons are accelerated to the energy of 450 GeV in the Super Proton Synchrotron and injected into the LHC ring as two separate beams running in the opposite directions. Each proton beam reaches the final energy of up to 7 TeV.

The proton beams do not have a continuos structure, but rather consist of the spatially separated *bunches* composed of 1.15×10^{11} protons each. The bunch structure is motivated by the LHC acceleration mechanism that uses the radio frequency cavities, where the protons increase their energy by the electromagnetic field with an effective field strength of 2 MV/m, oscillating at a frequency of 400 MHz. The bunch spacing determines the number of bunches that can be hosted by the proton beam. The bunch spacing in Run 1 of the LHC was 50 ns, resulting in at most 1404 bunches in each beam, and was halved to 25 ns in Run 2, resulting in at most 2808 bunches in each beam. The time spacing between the bunches defines the bunch-crossing frequency, which is 40 MHz for the bunch spacing of 25 ns.

The LHC makes use of the superconducting magnets to keep the proton beams in the LHC trajectory. These include 1232 dipole magnets that bend the beams to follow the circular shape of the LHC ring, and 392 quadrupole magnets that squeeze and focus each beam to eliminate the effect of the beam dispersion. The superconducting magnets made of NbTi operate in the liquid helium at the temperature of 1.9 K, producing the magnetic field of 8.33 T.

[1] The abbreviation comes from the French name *Conseil Europeen pour la Recherche Nucleaire*.

[2] The LHC is also used to accelerate and collide heavy ions, such as lead and xenon. These working regimes are not considered in the scope of this thesis.

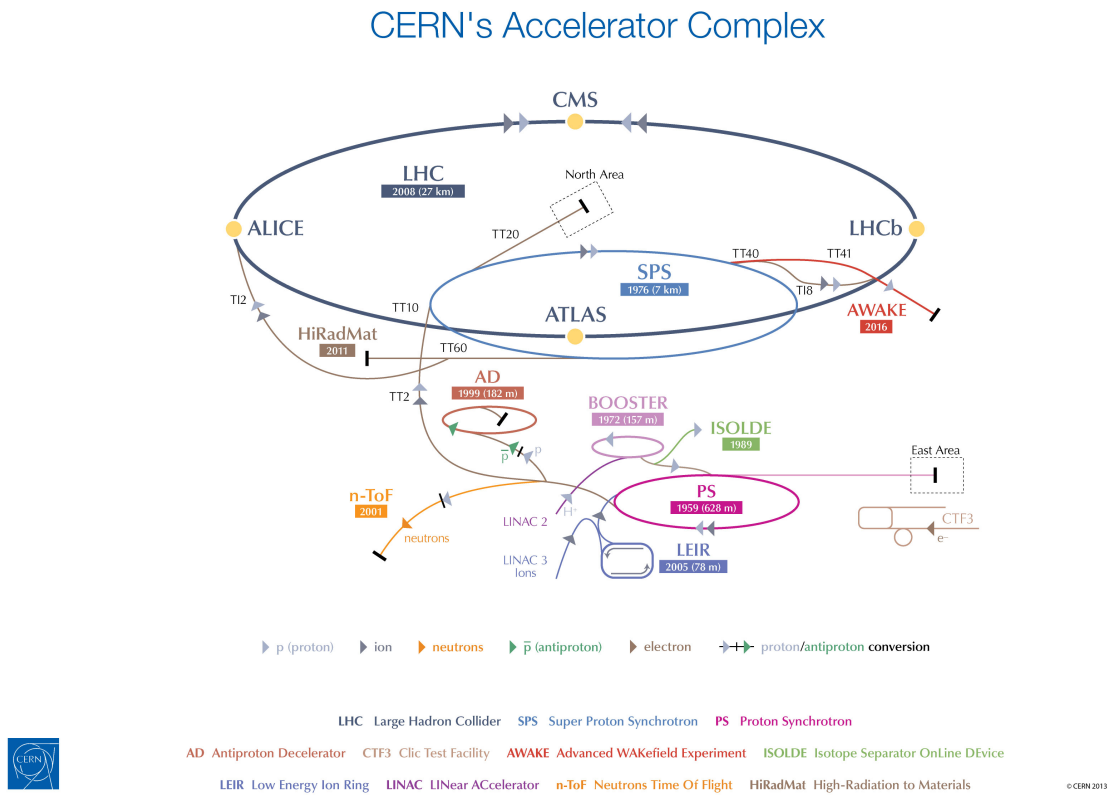


Figure 4.1: Sketch of the CERN accelerator complex, including LHC and experiments supported by this complex [116].

The LHC collides the proton beams at the four interaction points that correspond to the four main experemints, namely ATLAS, CMS, LHCb and ALICE. The ATLAS (A Toroidal LHC ApparatuS) and CMS (Compact Muon Solenoid) are multi-purpose detectors, sensitive to the wide range of different areas of particle physics, including searches for the new physics, studies of the properties of the Higgs boson and the SM precision measurements in the proton-proton and heavy ion collisions. ALICE (A Large Ion Collider Experiment) detector focuses on the studies of the strongly interacting matter at extreme energy densities in the heavy-ion collisions. LHCb (LHC beauty) detector is designed to perform precision measurements of the flavour physics and CP violation in the proton-proton collisions.

The size and bunch-crossing frequency determine a measure of the rate of different physics processes produced in particle collisions. This measure is referred to as *luminosity* $\mathcal{L}$, which is defined as a ratio of the number of events in a given period of time to the cross-section of a given physics process, $\mathcal{L}\sigma = dN/dt$. For the colliding Gaussian beams, the luminosity takes the form [117]:

$$\mathcal{L} = \frac{N_1 N_2 N_b f_r}{\mathcal{A}} \cdot S \cdot W \cdot K \,. \tag{4.1}$$

N_1, N_2 stand for the number of particles per bunch of each beam, N_b denotes the number of bunches in the beam, and $f_r = 11.245$ kHz is the LHC revolution frequency. The luminosity reduction factors S, W and K account for the loss in the lumonisity due to the non-zero crossing angle and transverse offset. $\mathcal{A}$ denotes the area of the overlap of

the beams, which can be expressed as $\mathcal{A} = 4\pi\varepsilon\beta_s$, where ε is the beam emittance and $\beta-$function β_s is related to the transverse size of the beam at the point s along the beam trajectory. Given the value β^* of the $\beta-$function at the interaction point and the Lorentz-inva, the luminosity can be expressed in the following way [3]:

$$\mathcal{L} = \frac{N_1 N_2 N_b f \beta_{\text{rel}}\gamma_{\text{rel}}}{4\pi\varepsilon^*\beta^*} \cdot S \cdot W \cdot K \,. \tag{4.2}$$

However, the high luminosity delivered by the LHC results in a significant background to the interesting physics processes that are associated to the hard scattering process. Around 40% of the bunch crossings results in the diffractive and elastic processes [118]. Even when the hard scattering occurs, it is usually followed by the large number of processes originating from the various secondary interactions. These include additional proton-proton interactions, referred to as *pile-up*, which consists of two main components. The *in-time* component originates from the multiple proton-proton interactions that occur in the same bunch crossing as the collision of interest. The *out-of-time* component corresponds to the proton-proton interactions that happen in the bunch crossing just before and after a given bunch crossing. This is related to the lengths of the electrical signals associated with the collision that are larger than the bunch spacing. This means that the signals from the neighbouring bunch crossings contribute to the signal of a given bunch crossing. Figure 4.2 illustrates the distribution of the pile-up parametrised in terms of the mean number of interactions per bunch crossing μ delivered to ATLAS during the 2015 and 2016 data taking periods. An increase in the luminosity for 2016 data taking comparing to the 2015 data taking results in the increase of the corresponding values of $\langle\mu\rangle$ from 13.7 to 24.9.

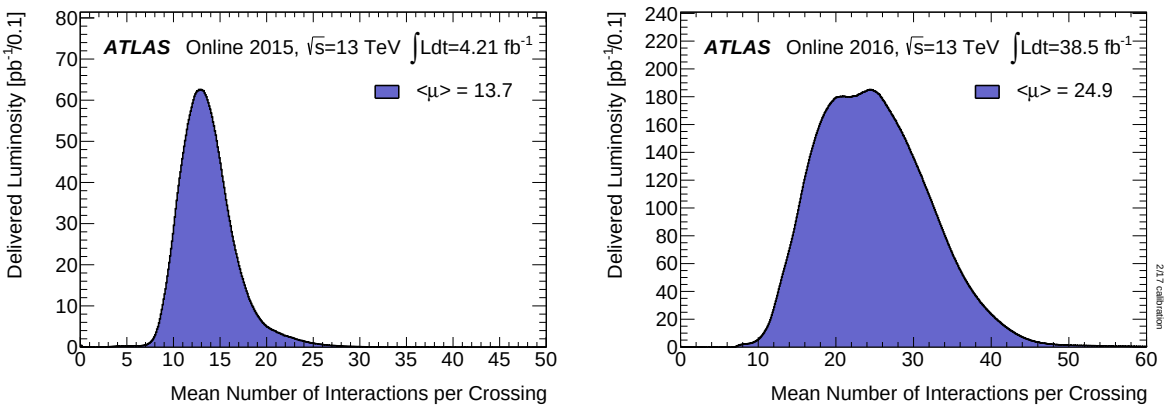


Figure 4.2: Distribution of the mean number of interactions per bunch crossing μdelivered to ATLAS for the pp collisions at $\sqrt{s} = 13$ TeV during the 2015 (left) and 2016 (right) data taking periods [119].

4.2 The ATLAS experiment

The ATLAS detector is a large multi-purpose detector, located at one of the four interaction points of the LHC. It is designed as a hermetic, cylindrically shaped detector,

[3] Since the emittance is inversely proportional to p_T of the particles in the beam, the boost invariant *normalised emittance* is used, $\varepsilon^* = \beta_{\text{rel}}\gamma_{\text{rel}}\varepsilon$, where β_{rel}, γ_{rel} denote the retavistic Lorentz factors.

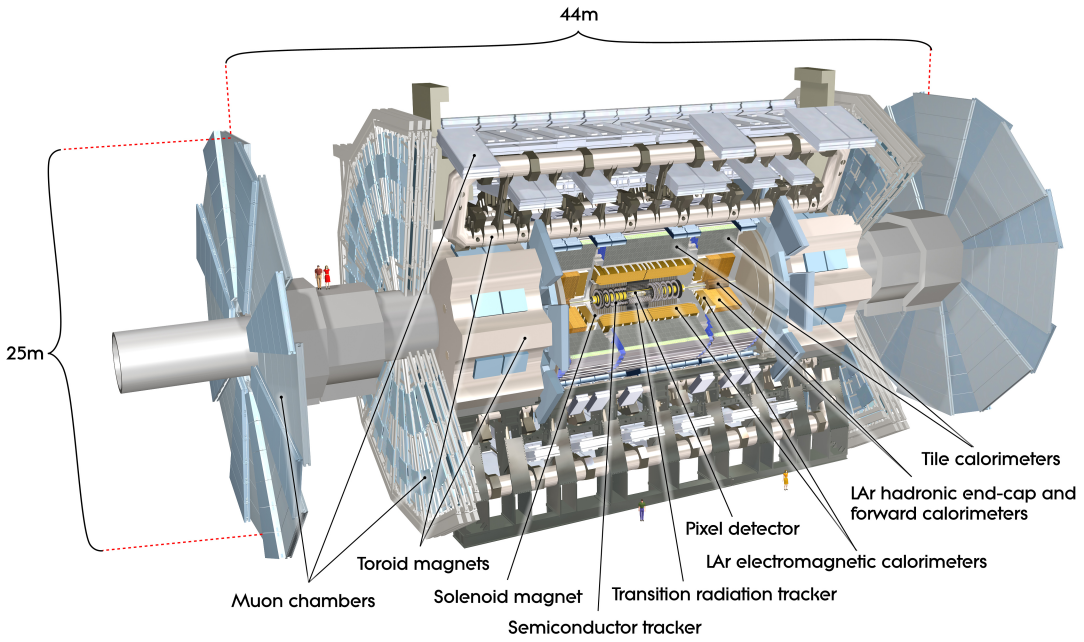


Figure 4.3: Computer-generated schematic representation of the ATLAS detector and its components [120].

enclosing almost 4π in solid angle. This allows to reconstruct and identify nearly every highly energetic particle produced in the proton-proton collisions.

ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) and the $z-$axis that follows the beam direction. The $x-$axis points to the center of the LHC ring, and the $y-$axis points upward. Cylindrical coordinates (ρ, ϕ) are used in the transverse plane, where the azimuthal angle ϕ is measured around the beam pipe. Since the polar angle θ is not Lorentz-invariant, the pseudorapidity is introduced instead as $\eta = -\ln\tan(\theta/2)$. ATLAS covers the pseudorapidity range of $|\eta| < 4.9$, where smaller values correspond to the central region and higher values to the forward regions of the detector. The angular distance between the objects in this coordinate system is defined as $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

ATLAS is designed to be as hermetic as possible, enclosing almost 4π in solid angle. This allows to reconstruct and identify almost every particle produced in the proton-proton collisions using different detector sub-systems. A schematic representation of the ATLAS detector and its sub-systems is shown in Figure 4.3.

The particles produced at the IP initially travel through the innermost layers coated by the solenoid magnet that produces strong homogeneous field of 2 T. These layers serve as tracking detectors, covering the range of $|\eta| < 2.5$. The strong magnetic field bends the trajectory of the charged particles, the curvatures of which determine the corresponding momenta and signs of the charge. The next layers in the detector, the calorimeter layers, are aimed to measure the energy of both charged and neutral particles. The electromagnetic calorimeter is designed to measure the energy of the electromagnetically interacting particles such as electrons and muons that produce electromagnetic showers while interacting with the detector material. The hadronic calorimeter is aimed to provide energy measurements based on the nuclear interaction between the hadrons and detector material. The calorimeters provide the full coverage in $|\eta|$ of up to 4.9. The muons escape

the calorimeter without the significant losses in energy and thus are detected in the outermost layer of the detector, the muon spectrometer. It is bathed in a magnetic field of the toroid magnet, allowing to measure the properties of the muons based on their track information. The muon spectrometer provides the coverage for $|\eta| < 2.7$. Neutrinos and other invisible particles (e.g. DM particles) do not interact with the detector and thus can be revealed from the transverse momentum imbalance measured in terms of the *missing transverse energy*.
All above-mentioned detector sub-systems are described in Sections 4.2.1 to 4.2.3, following . A description of the ATLAS trigger system, including author's contribution in the calibration of the Level-1 Calorimeter Trigger, is given in Section 4.2.4.

4.2.1 Inner detector

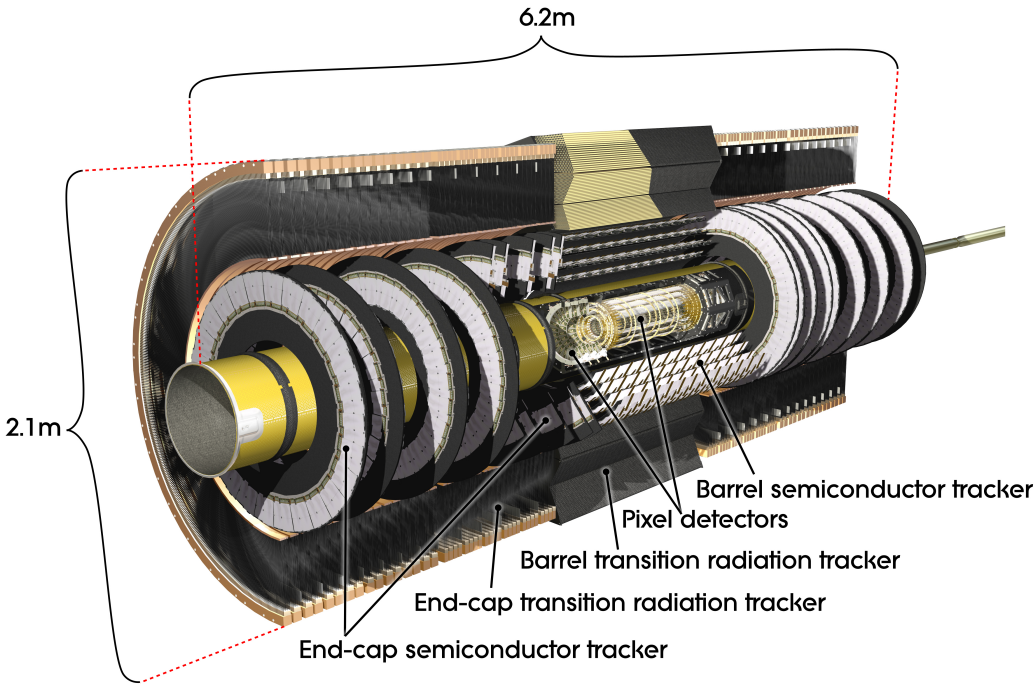


Figure 4.4: Computer generated image of the ATLAS inner detector [121].

The ATLAS inner detector (ID) is designed to provide accurate reconstruction of the charged particle tracks used for the high resolution momentum measurements and particle identification, as well as the reconstruction of the primary and secondary vertices, in the region of $|\eta| < 2.5$. The nominal lower p_T of the reconstructed tracks in 0.5 GeV. As shown in Figure 4.4, the ID consists of three complementary sub-systems, the Pixel Detector (PD), Silicon Semiconductor Tracker (SCT) and Transition Radiation Tracker (TRT). The latter also provides an electron identification for 0.5 GeV $< p_T <$ 150 GeV in the region of $|\eta| < 2.0$. All sub-systems are fully contained within the homogeneous magnetic field of 2 T produced by the ATLAS solenoid magnet.
The PD consists of 1744 modules arranged in three cylindrical layers in the central (*barrel*) region and in three discs layers in each of the two end-caps. The modules host about 80 million pixels with a nominal pixel size of $50 \times 400\mu$m ($50 \times 600\mu$m in the regions at the front-end chips) and nominal thickness of 250μm. To maintain the high density of track

close to the IP, the PD provides an intrinsic accuracy of 10μm per layer in the $R-\phi$ plane transverse to the beam and 115 μm per layer in the $z-$direction ($R-$directron) for the barrel modules (end-cap discs). To sustain the high radiation environment, the pixel modules are operated in the range of temperatures from -10°Cto -5°C.
The SCT contains 4088 two-sided modules arranged in four coaxial cylindrical layers in the barrel region and in nine discs layers in each of the two end-caps. With a surface coverage of 63 m^2 of silicon, it is nearly hermetic with at least four precision space-point measurements over the fiducial coverage of the ID. The SCT modules host 80μm pitch micro-strip sensors, which are connected to about 6.3 million readout chips. Comparing to the PD, each layer in the SCT provides similar resolution of 17μm in the $R-\phi$ plane and much coarser resolution of 580 μm in the $z-$direction ($R-$directron) for the barrel modules (end-cap discs).
The TRT consists of up to 73 layers of polyimide straw tubes interleaved with fibers in the barrel region (parallel to the beam axis) and 160 layers of straw tubes interleaved with foils in the end-cap regions (radially in the wheels). These straw tubes, each 4 mm in diameter and 144 (39) cm long in the barrel (end-cap) region, is filled with a gas mixture of 70% Xe, 27% CO_2, and 3% O_2. Due to the low mass, a highly-boosted electron can emit X-rays as it traverses through the straws and gas mixture. This is referred to as *transition radiation*. As this radiation is suppressed by a factor of $1/m$, where m is the mass of the particle, the heavy particles, such as charged pions, start to emit X-rays only at higher boost. This limits the electron identification to the region of p_T below 150 GeV. The charged tracks with $p_T > 0.5$ GeV leave at least 36 hits in the TRT except the barrel-end-cap transition region of $0.8 < |\eta| < 1.0$, where at least 22 hits are expected. The resulting resolution is much coarser comparing to the PD and SCT, amounting to 130μm in the $R-\phi$ plane.
The ID provides a transverse impact parameter resolution of $\sigma_{d_0}/d_0 = 22.1 \pm 0.9 \mu m$ and a relative momentum resolution of $\sigma_{p_T}/p_T = (4.83 \pm 0.16) \times 10^{-4}$ GeV$^{-1} \times pT$ for high p_T tracks [122].

4.2.2 Calorimetry

The ATLAS calorimeter is designed to provide accurate energy measurements of all particles that interact either electromagnetically or hadronically except muons [4] and neutrinos. It consists of the sampling calorimeters that contain alternating layers of a dense absorber material with high atomic number Z and active material. The particles interact with the absorber material, inducing the cascade of secondary particles, referred to as *showers*. The energy in these particle showers is measured from the interaction with the active material (ionization and scintillation processes). Such a sampling structure makes it possible to design relatively cheap, compact calorimeters at the cost of energy losses in the absorber material. These losses are quantified in terms of the sampling fraction $f_{\text{samp}} = E_{\text{active}}/(E_{\text{active}} + E_{\text{passive}})$, where E_{active}, E_{passive} denote energy absorbed by the active and passive material. This quantity should be accounted for in order to achieve precise measurements of the parton shower energy.
The particles produce two types of showers, depending on the type of interaction with the detector material. The electromagnetic showers, initiated mainly by electrons and photons, are measured in the electromagnetic calorimeter. The hadronic showers, initiated by hadrons, are measured in the hadronic calorimeter. The calorimeter regions closest

[4] The muons are Minimally Ionizing Particles at the LHC energy scale.

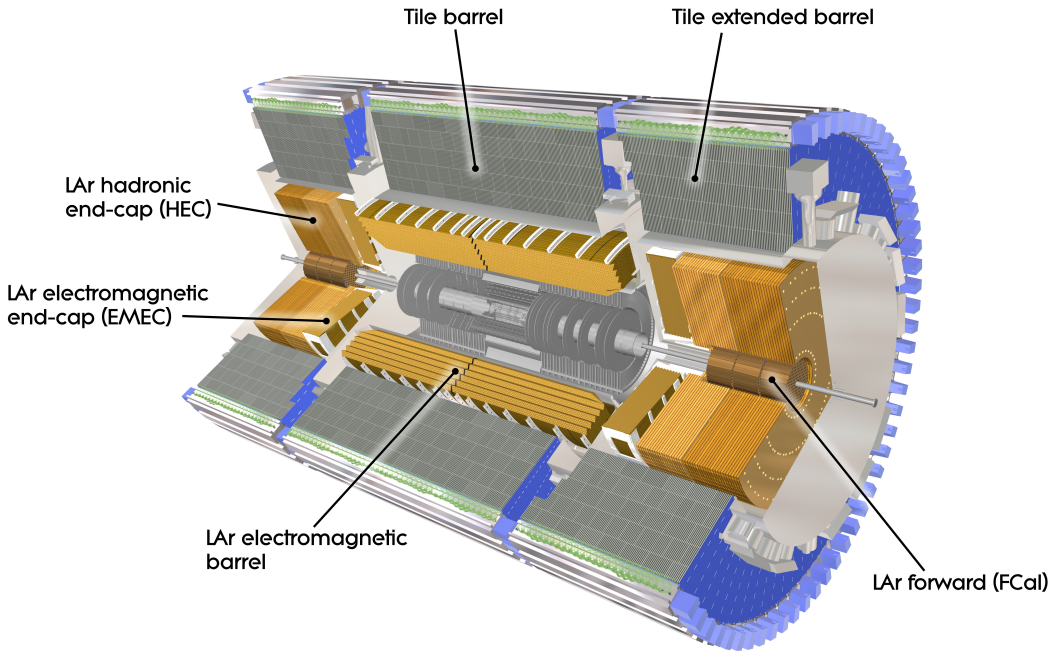


Figure 4.5: Computer generated image of the ATLAS calorimeter [123].

to the beam axis use liquid argon as an active material due to its stable, linear response and radiation hardness. These regions, referred to as *Liquid Argon Calorimeter* (LAr), are the barrel and end-caps of the electromagnetic calorimeter, as well as the end-caps and forward regions of the hadronic calorimeter. The barrel regions are bathed in the same barrel cryostat, while the end-cap and forwards regions are embedded in the two end-cap cryostats. The outer regions of the hadronic calorimeter are located far away from the beam line and thus are exposed to the lower radiation levels. This motivates the least-cost choice of the scintillator tiles as an active material and steel as an absorber material, and use steel as the absorber material, both operated at a room temperature. This region of the hadronic calorimeter is referred to as *Tile Calorimeter*.
Figure 4.5 illustrates a schematic representation of the ATLAS calorimeter and their main components. The calorimeters provide the full $\phi-$symmetry and coverage of $|\eta| < 4.9$ around the beam axis.

4.2.2.1 Electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) is designed to measure the energy of electromagnetically interacting particles, such as electrons and photons. It is arranged in alternating layers of an absorber material and liquid argon as an active material. An accordion geometry is implemented for the absorbers both in the barrel and end-cap regions of the ECAL, providing full coverage in the azimuthal angle ϕ without any energy losses due to the crack regions, as well as the fast signal extraction. The absorbers are made of the lead plates interleaved with the copper read-out electrodes. In the barrel calorimeter, the plates are 1.53 mm thick in the region of $|\eta| < 0.8$ and 1.13 mm thick in the region of $|\eta| > 0.8$. In the end-cap calorimeters, the plates are 1.7 mm thick in the region of $|\eta| < 2.5$ and 2.2 mm thick in the region of $|\eta| > 2.5$. Electrons and photons in the GeVrange loose their energy in the absorbers primarily due to the Bremsstrahlung and

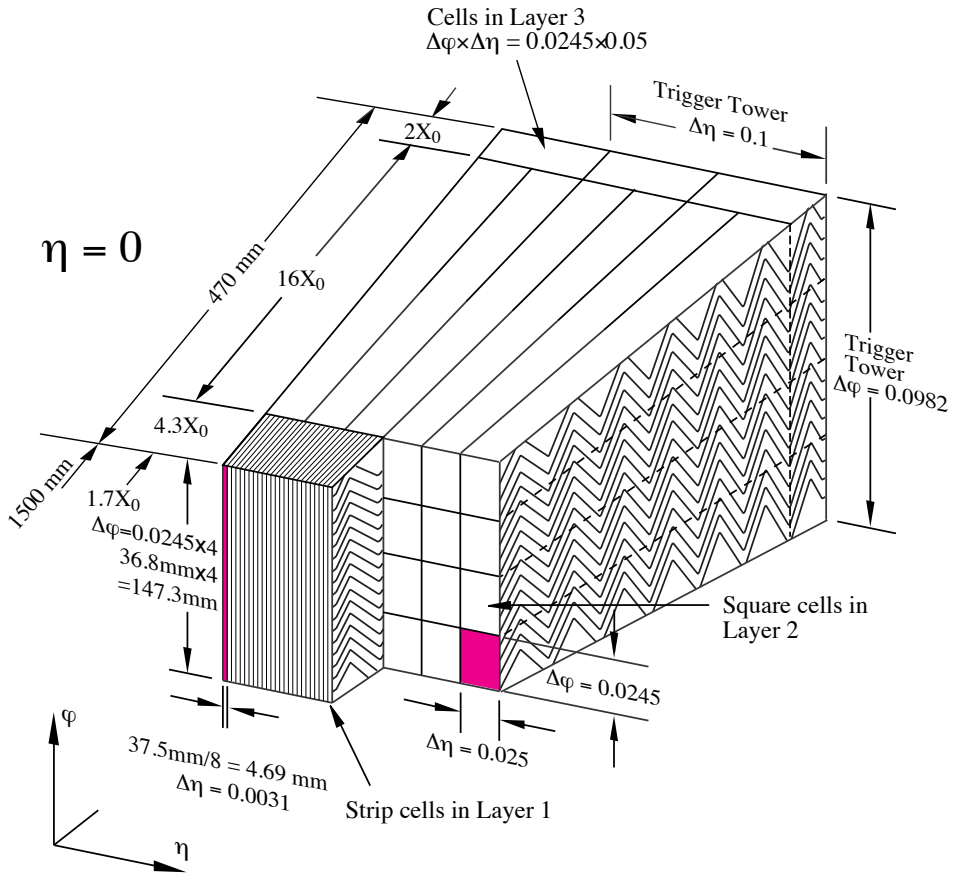


Figure 4.6: Sketch of a section of the electromagnetic barrel calorimeter at $\eta = 0$, showing its accordion geometry and layers of calorimeter cells with the corresponding dimensions. [124].

e^+e^- pair-production.
The ECAL consists of the Electromagnetic Barrel (EMB) calorimeter, covering the region of $|\eta| < 1.475$, and the Electromagnetic Endcap (EMECs) calorimeters, covering the region of $1.325 < |\eta| < 3.2$. The EMB is arranged in three layers with different granularities and depths, as illustrated in Figure 4.6. The first layer provides an extremely fine granularity in η and consists of strip cells with a size of 0.003.1 in (η, ϕ) space and depth of $4.3X_0$, where X_0 stands for the radiation length [5]. This allows to discriminate photons from the $\pi_0 \to \gamma\gamma$ decay. The second layer is aimed to absorb the majority of the energy from the traversing particles and provide accurate (η, ϕ) measurements. It consists of cells with a size of $(0.025, 0.0245)$ in (η, ϕ) space and depth of $16X_0$. The last layer is designed to measure the remnants of the electromagnetic showers and thus does not require such a fine granularity in η. It consists of the larger cells comparing to the other layers with a size of $(0.05, 0.0245)$ in (η, ϕ) space and depth of $2X_0$.
Each of the EMECs consists of the two coaxial wheels, each 63 cm thick, employed symmetrically around the beam axis. The boundary between the wheels is 3 mm thick and located at $|\eta| = 2.5$. Each wheel is subsequently divided into eight wedge-shaped modules. These modules have similar structure as for the EMB, arranged in layers of different granularities. The region of $1.5 < |\eta| < 2.5$ is arranged in three layers that

[5]The radiation length is defined as the average distance of the detector material, which must be traversed by a particle to reduce its energy by a factor 1/e due to electromagnetic interaction with the detector material.

consist of cells of the same size in (η, ϕ) as in the EMB layers. The outermost region of $1.375 < |\eta| < 1.5$ and the innermost region of $2.5 < |\eta| < 3.2$ consist of only two layers with a coarse granularity of up to 0.1×0.1 in (η, ϕ).
The region of $1.375 < |\eta| < 1.52$ is referred to as *crack region*. It corresponds to the transition region between the barrel and end-cap cryostats, which contains a significant fraction of inactive material, several radiation lengths thick. This leads to the energy losses and poor energy resolution in the crack region. Hence, it is not typically used for the electron and photon identification.
Finally, the electromagnetic presampler layers are employed in front of the EMB and EMEC to account for the energy losses due to the inactive material between the IP and the ECAL and in the crack region.
The final energy resolutions in the EMB and EMECs are measured to $\frac{\sigma_E}{E} = \frac{10.1\%}{\sqrt{E}} \oplus 0.17\%$ [125] and $\frac{\sigma_E}{E} = \frac{12.1\%}{\sqrt{E}} \oplus 0.4\%$ [126], respectively, where the symbol $\oplus$ indicates that terms are added in quadrature.

4.2.2.2 Hadronic calorimeter

The hadronic calorimeter (HCAL) is aimed to provide energy measurements of hadronically interacting particles. These particles are mainly protons, neutrons, pions and kaons. Similar to the ECAL, it is organized in alternating layers of an absorber and active material. However, the processes that induce the hadronic showers are totally different from those initiating the electromagnetic showers.
The π^0 mesons are produced in about 1/3 of all hadronic interactions, decaying to a pair of photons with the branching ratio of 98.8% [54]. These decay processes form the electromagnetic component of the hadronic showers, amounting to approximately 20% of the total energy content. The remamining 2/3 of all hadronic interactions contribute to the hadronic component. However, in contrast to the energy measurements in the electromagnetic showers, a significant fraction of the energy in the hadronic showers remains undetectable (*invisible*). The major of this energy, referred to as *nuclear binding energy* is used to break up the nuclear bonds. The remaining part is carried away by the nuclear fragments absorbed before being detected, long-lived particles such as K^0_L, neutrinos and minimally ionizing particles such as muons. The resulting invisible energy fraction amounts to about 30-40% [127]. This leads to the non-linear hadronic responce and in general to the different calorimeter responses to the electromagnetic and hadronic shower components, e and h, respectively. Hence, the ATLAS calorimeter is a *non-compensating* calorimeter due to the response ratio of $e/h < 1$.
The different calorimeter responses imply sophisticated calibration procedures to compensate for the losses in the energy resolution. In particular, the hadronic jets initially measured at the *electromagnetic energy scale* are brought up to the *hadronic scale* by the *jet energy scale* calibration, as will be discussed in details in Section 6.2.3.
The Tile calorimeter is located in the region of $|\eta| < 1.7$ behind the LAr electromagnetic calorimeter. It has an inner radius of 2.28 m and an outer radius of 4.25 m, which corresponds to the radial depth of 7.4λ, where λ denotes the interaction length [6]. The Tile calorimeter consists of a central barrel, 5.8 m in length, placed symmetrically around the IP and two barrel extentions, each 2.5 m in length, installed along the beam axis.

[6] The interaction length is defined as the average distance of the detector material, which must be traversed by a particle to reduce its energy of a particle by a factor 1/e due to electromagnetic and hadronic interactions with the detector material.

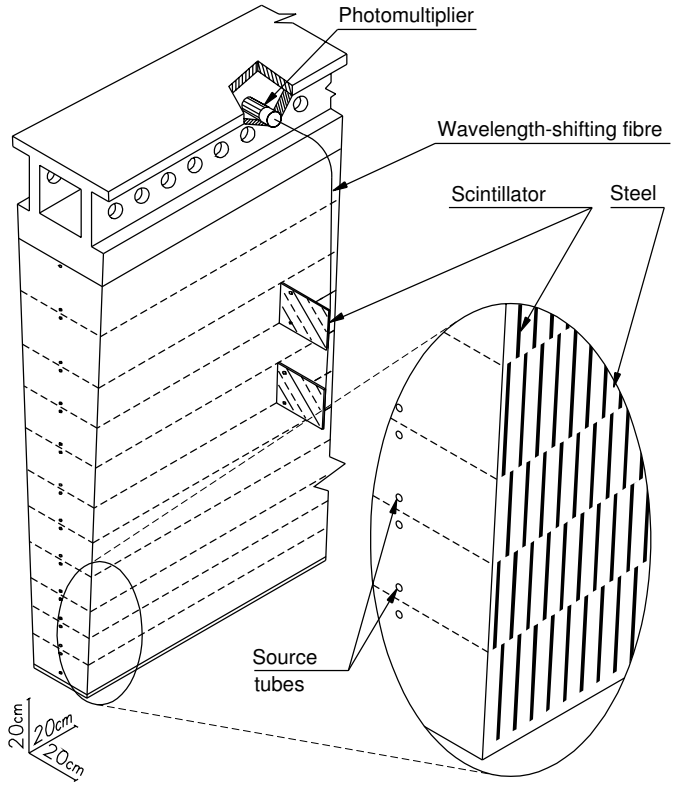


Figure 4.7: Sketch showing the structure of the modules in the Tile Calorimeter [124].

Each part is divided into 64 modules of size $\Delta\phi \sim 0.1$, which form an alternating structure of steel plates as an absorber material and scintillating tiles as an active material with a volume ratio of about $4.7 : 1$, as shown in Figure 4.7.
The hadronic end-cap calorimeter (HEC) is a sampling calorimeter that covers the region of $1.5 < |\eta| < 3.2$. It shares each one of the two end-cap cryostats with the EMEC and the forward calorimeter. In each end-cap cryostat the HEC consists of front and rare cyndrical wheels, each composed of 32 identical wedge-shaped modules. An outer radius of each wheel is 2.03 mm, while an inner radius is 0.372 m for the first nine plates of the front wheel and 0.475 m for the remaining plates of the front wheel and for the whole rare wheel. This provides a combined radial depth of about 12λ. Each wheel composed of 32 identical wedge-shaped modules, arranged in alternating layers of copper plates as absorbers and liquid argon as an active material.
The forward calorimeter (FCAL) is installed between the HEC and the beam pipe, covering the region of $3.2 < |\eta| < 4.9$. Located approximately 4.7 m from the IP, it expriences large particle fluxes, meaning that the densed structure of the absorbers and active material is required. The FCAL consists of three modules, each 45 cm thick, which are the electromagnetic module FCAL1 and two hadronic modules FCAL2 and FCAL3. The FCAL1 module is composed of copper plates, acting as absorbers, stacked one behind the other. 12260 equidistant holes filled with electrodes are drilled through the plates. Each electrode consists of a co-axial copper rod and copper tube filled with liquid argon. The FCAL2 and FCAL3 modules are aimed to counteract the spread of the hadronic showers by a high absorption length. The high absorption is achieved by using tungsten as an absorber. Each module is composed of two copper plates, each 2.35 cm thick, with the same electrode structure employed as in the FCAL1. However, the tungsten rods

are used in electrodes instead of copper rods. The remaining space between the plates and electrodes is also filled with tungsten. Such a structure provides a combined FCAL radiation depth of about 10λ.
The measured energy resolutions in the Tile, end-cap and forward calorimeters are $\frac{\sigma_E}{E} = \frac{52.7\%}{\sqrt{E}} \oplus 5.7\%$ [128], $\frac{\sigma_E}{E} = \frac{70.6\%}{\sqrt{E}} \oplus 5.8\%$ [129] and $\frac{\sigma_E}{E} = \frac{94.2\%}{\sqrt{E}} \oplus 7.5\%$ [130], respectively.

4.2.3 Muon spectrometer

The ATLAS muon spectrometer (MS) is the outermost part of the ATLAS detector, designed to detect muons and provide precise measurements of their momenta in the range of $|\eta| < 2.7$. The muons are deflected by the magnetic fields produced by the large barrel toroidal magnet ($|\eta| < 1.4$) and two smaller end-cap toroidal magnets for ($1.6 < |\eta| < 2.7$). The magnetic field in the transition region of $1.4 < |\eta| < 1.6$ is given by the combination of the barrel and end-cap fields.
In the barrel region the MS consists of the muon chambers arranged in three concentric cylindrical shells around the beam axis at radii of approximately 5 m, 7.5 m, and 10 m. In the end-cap regions the muon chambers are arranged in four large wheels perpendicular to the beam axis at distances of 7 m, 11 m, 14 m and 22 m from the IP. Four different types of the muon chambers are installed in the MS, referred to as Monitored Drift Tubes (MDTs), Cathode Strip Chambers (CSCs), Resistive Plate Chambers (RPCs) and Thin Gap Chambers (TGCs).
The MDTs provide an accurate measurement of the muon momentum for the entire coverage of $|\eta| < 2.7$ except the region of $2.0 < |\eta| < 2.7$ in the innermost end-cap layer, where CSCs take over. These chambers consist of three to eight layers of drift tubes filled with gas, consisting of 93% argon and 7% CO_2, operated at the pressure of 3 bar. An average resolution of about 80 μm per tube and 35 μm per chamber is achieved.
The CSCs are employed in the region of $2.0 < |\eta| < 2.7$ due to their higher rate capability and time resolution. Each chamber provides a resolution of about 40 μm in the bending plane and 5 μm in the transverse plane.
An important is the triggering on the muon tracks. Hence, the RPCs and TGCs are added to the design of the MS in order to provide the fast tracking information within a few tens of nanoseconds after the passage of the muon. The RPCs cover the barrel region of $|\eta| < 1.05$, while TGCs cover the end-cap region of $1.05 < |\eta| < 2.4$.
Since RPCs operate in the barrel region, they are arranged in three cylindrical layers around the beam axis. The chambers from the two inner layers provide the fast triggering for the muons with 6 GeV $< p_T <$ 9 GeV, while the chambers from the outer layer trigger on the muons with 9 GeV $< p_T <$ 35 GeV. The RPCs consist of two resistive parallel electrode-plates separated at a distance of 2 mm. These plates enclose the gas mixture in the strong electric field of 4.9 kV/mm. The muon, passing trought the gas, triggers the avalanche of electrons with respect to the direction of the muon track. The electrical signals from the RPCs are read out via capacitive coupling, leading to a short width of about 5 ns and thus accurate assignment to the correct bunch crossing.
Apart from the fast triggering, the TGCs provide the measurement of the azimuthal coordinate ϕ of the muon, which complements the momentum measurements from the MDTs. The chamber is filled with a gas mixure that surrounds the groups of wires with the distance between the wires of 1.8 mm and the distance between the wire and cathode of 1.4 mm. The high wire potential of 2.9 kV and small distance between wires ensure high time resolution for the majority of the muon tracks.

The MS provides the muon p_T resolution of about 10% for the muon tracks with $p_T =$ 1 TeV.

4.2.4 ATLAS Trigger System

The bunch crossings at LHC occur every 25 ns, which corresponds the rate of 40 millions events per second. A typical size of a recorded event that contains information about the fully reconstructed objects from all sub-systems of ATLAS detector is about 1.5 MB, which corresponds to the data flow of about 60 PB per second. It takes about 15 s for a CPU to reconstruct and record such an event. None of the present-day data acquisition systems and data storages can handle such a huge data flow and store every event from the pp collisions. However, there is no need to reconstruct and record every event. As was mentioned before, about 40% of all bunch crossings lead to the diffractive and elastic processes, and the vast majority of the remaining bunch crossings result in the processes at lower energy scale than for the physics of interest. Figure 4.8 illustrates the cross-section of typical SM processes at hadron colliders. For the center-of-mass energy of $\sqrt{s} = 13$ TeV, the total inelastic cross-section is about nine orders of magnitude larger than, e.g. the Higgs production cross-section.

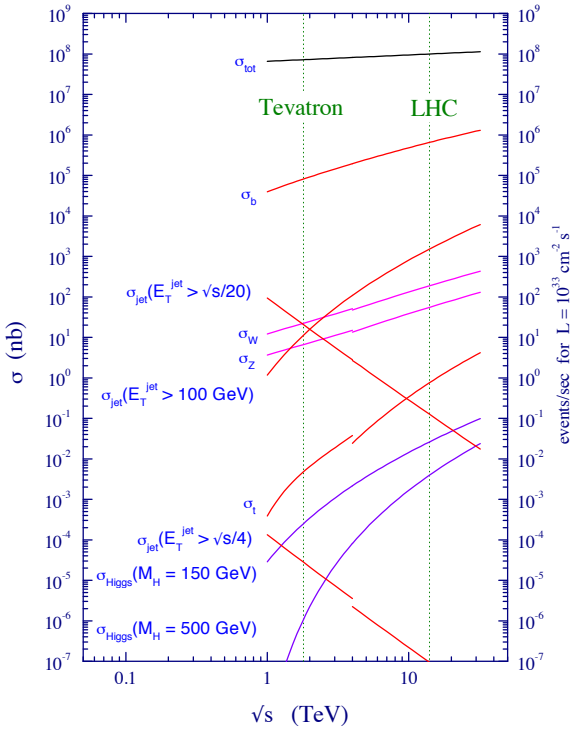


Figure 4.8: Cross-section of various SM processes at the Tevatron and LHC as a function of the center-of-mass energy $\sqrt{s}$ [131]. The cross-section are calculated for the $p\bar{p}$ colliders in the region of $\sqrt{s} < 4$ TeV and for the pp colliders in the region of $\sqrt{s} > 4$ TeV.

Thus, an ATLAS trigger system was introduced to run in parallel to the data taking process and reduce the final event rate by several orders of magnitude by selecting interesting events. These interesting events are defined based on the sets of conditions, which include mainly the particle content requirements and kinematic thresholds for different objects, referred to as *triggers*. The ATLAS trigger system consists of two levels of subsequent event rate reduction, the *Level-1 Trigger* and *High Level Trigger*, as illustrated

in Figure 4.9. A discussion in this Section follows Ref. [132–134].

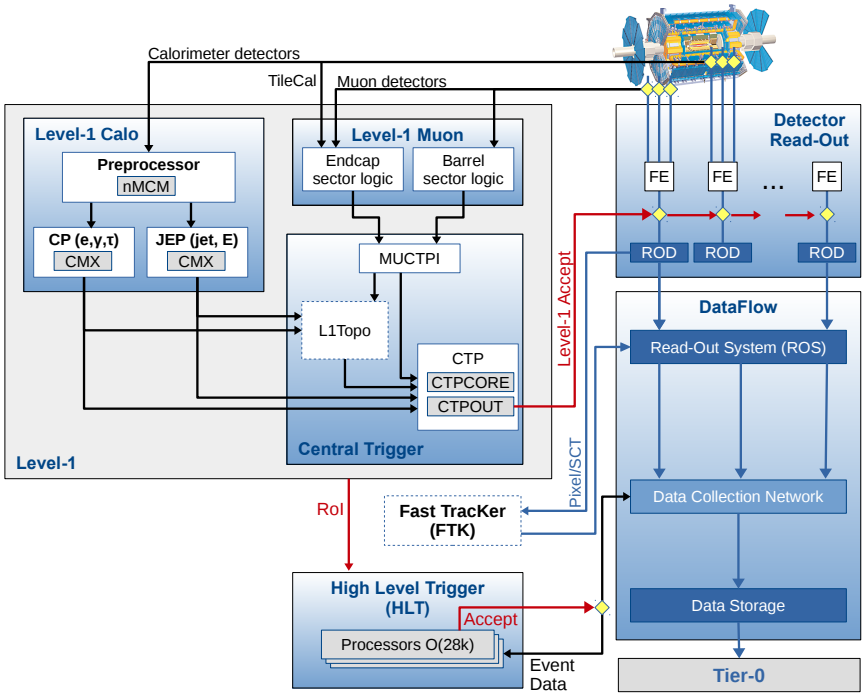


Figure 4.9: Schematic overwiew of the ATLAS trigger system [132]. The L1Topo system was comissioned in 2017, while FastTracker is still in a comissioning state.

4.2.4.1 Level-1 Trigger

The Level-1 (L1) Trigger is a software and hardware based system that process the coarse granularity data from the ATLAS calorimeter and muon spectrometer to select events that contain highly energetic electrons, muons, photons, jets, hadronically decaying $\tau-$leptons, as well as large missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ and total transverse energy $E_{\mathrm{T}}^{\mathrm{total}}$. It consists of four main sub-systems, namely the L1 Calorimeter Trigger (L1Calo), the L1 Muon trigger (L1Muon), the L1 Topological Trigger (L1Topo) and the Central Trigger Processor (CTP).

L1 Calorimeter Trigger

The L1 Calorimeter Trigger is designed to select events that contain all objects but muons based on the information from the ATLAS calorimeter. It analyses signal from 7168 calorimeter segments that are aligned in the radial direction, referred to as *trigger towers* (TTs). The main components that process signals from the TTs are the *PreProcessor Modules* (PPMs). Each PPM contains 16 new *Multichip Modules* (nMCMs) that simultaneously process signals from four TTs. This processing consists of digitisation in Flash Analogue-to-Digital Converters (FADCs), synchronisation and assignment to the correct bunch crossing (Bunch-Crossing Identification), both performed in Application Specific Integrated Circuit (ASIC), and calibration of the digitised signal to the E_T values in (η, ϕ) plane. These E_T values are then transfered to the *Cluster Processor* (CP) and the *Jet Energy Processor* (JEP), which identify the deposits of local energy maxima

using dedicated sliding window algorithms. These deposits are referred to as *regions of interest* (ROIs) The CP is designed to identify candidates for electrons, photons and hadronically decaying $\tau-$leptons and calculate their E_T. The JET is aimed to identify jet candidates and compute their E_T, as well as to calculate values of $E_{\mathrm{T}}^{\mathrm{miss}}$ and $E_{\mathrm{T}}^{\mathrm{total}}$.

L1 Muon Trigger

The L1 Muon Trigger collects data from the RPCs and TGCs in the muon spectrometer that provide required timing resolution for the correct bunch crossing identification, as described in Section 4.2.3. The RPC and TGC consist each of three layers of trigger chambers. The trigger information is based on the coincidences between the hits in different trigger stations from different layers along a trajectory of the traversing muon candidate. Each coincidence is related to a certain deviation from straightness, which is defined by the p_T threshold. This deviation is calculated as deviation of the slope of the track segment between two trigger chambers from the slope of a straight line between the IP and the hit in a reference layer (*pivot plane*).

L1 Topological Trigger

The information about the particle candidates is transfered from L1Calo and L1Muon to the L1 Topological Trigger. It computes complex topological variables, such as angular separation between the particle candidates and invariant mass of particle candidates, to form the topological triggers, which can be used by different physics analyses that require specific topology of the final state.

Central Trigger Processor

The Central Trigger Processor (CTP) collects inputs from the L1Calo, L1Muon and L1Topo and compares them to the trigger items from a programmable L1 *trigger menu*, which consists of up to 512 trigger items. Each trigger is a combination of particle multiplicity requirements, kinematic thresholds and topological requirements, such as angular separation, in case of L1Topo-based triggers. As more luminosity is delivered, the event rate of a single trigger increases and at some point exceeds the allowed bandwidth. In this case a *prescaling factor* P is applied to a given trigger, meaning that it accepts one of P events that pass the trigger requirement. This trigger is referred to as *prescaled*. The triggers that accept every event that passes the trigger requirement are referred to as *unprescaled*. Given the statistical bias to the trigger selection introduced by the prescaling factor, the unprescaled triggers are used as primary triggers for physics analyses.
If at least one of the triggers from the trigger menu is fulfilled, the L1-accept (L1A) is initiated, and the corresponding event is transfered to the HLT. The time that L1 Trigger needs to process the event information and give the L1A is 2.5μs. This reduces the event rate from initial 40 MHz to 100 kHz.

4.2.4.2 High Level Trigger

After the L1A accept, the event is collected by the readout drivers (RODs) and transmitted to the readout system (ROS), which makes it available to the High Level Trigger. The High Level Trigger is a purely software base system implemented for a cluster of

about 40000 commercial CPUs. It processes inputs from L1 Trigger system using finer-granularity information from the calorimeters and MS, as well additional tracking information from the ID. The HLT uses a two-stage chain of event reconstruction algorithms, referred to as *trigger chain*, with the fast partial event reconstruction, such as reconstruction of the parameters of tracks and calorimeter clusters, to reject the majority of events, and the full reconstruction for the remaining events. About 15,%, 30% and 40% of the CPU time are used for the ID track reconstruction, MS reconstruction and calorimeter reconstruction, respectively.
The HLT trigger items uses the basic information about the particle multiplicity and kinematic thresholds, as well as more dedicated information related to the reconstruction and identification of a given object. For instance, HLT_xe100_mht_L1XE50 trigger, used in the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analyses, corresponds to a trigger based on $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstructed using "mht" algorithm [7] with a threshold of $E_{\mathrm{T}}^{\mathrm{miss}} > 100$ GeV seeded by the L1 trigger with a threshold of $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV.
If at least one of the HLT trigger chains is fulfilled, the event is recorded for the offline reconstruction, which will be discussed in Chapter 6. This reduces the event rate to 1 kHz.

[7] $E_{\mathrm{T}}^{\mathrm{miss}}$ of the mht algorithm is calculated as a negative vector sum of transverse momenta of all jets reconstructed by the anti-k_T jet finding algorithm, described in Section 6.2.2, using calorimeter topological clusters.

Chapter 5

Data, event simulation and triggers

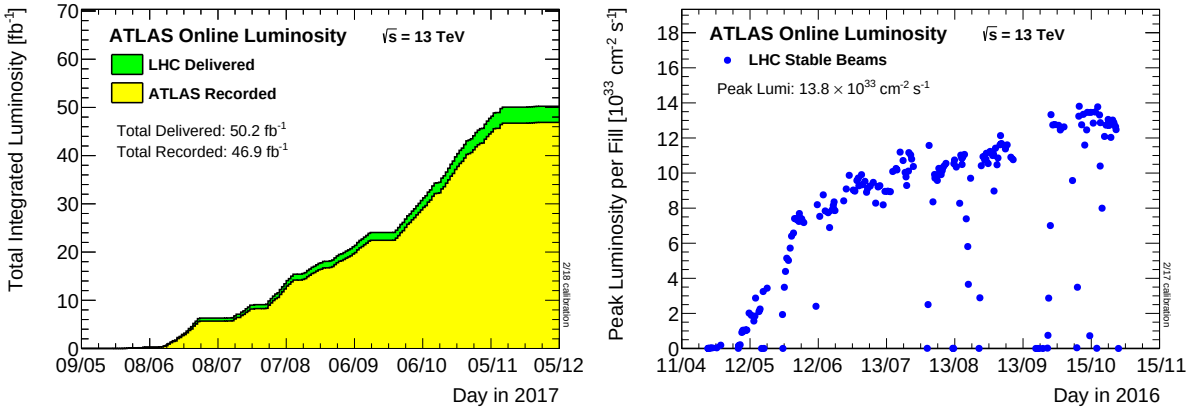


Figure 5.1: Integrated (left) and peak (right) luminosity for the pp collisions at $\sqrt{s} = 13$ TeV as a function of time during the 2016 data taking period [119]. The left plot shows the integrated luminosity delivered to ATLAS (green) and recorded by ATLAS (yellow). The right plot shows the peak luminosity deliverd to ATLAS.

The rate of the physics process measured by a given experiment at the LHC is proporional to the *integrated* luminosity $\mathcal{L}_{int} = \int \mathcal{L} dt$, which corresponds to the total amount of data recorded by this experiment. The amount of recorded data is one of the main driving factor of the sensitivy of a search for new physics. The LHC delivered in total 42.7 fb^{-1} of integrated luminosity during the 2015 and 2016 data taking periods, 39.7 fb^{-1} of which were recorded by ATLAS [119]. This corresponds to the data taking efficiency of about 93%. The maximum instantaneous luminosity provided by the LHC during this period was 13.8×10^{33} cm^{-2}s^{-1}, as shown in Figure 5.1.

The data that are used for physics analysis should pass a set of data quality criteria, referred to as *Data Quality status flags* [135]. These criteria reject corrupted events and incomplete events that do not contain full information about the required objects due to the errors ocurred in the different sub-systems of ATLAS. The *good run list* (GRL) [136] defines the data that passes all data quality requirements and can be used for physics analysis. The data from the GRLs used for $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analyses amounts in total to an integrated luminosity of 36.1 fb^{-1}.

5.1 Event simulation

A full event simulation includes many stages, beginning with the generation of the proton-proton collisions and ending with the object reconstruction using dedicated simulations of ATLAS detector. The Monte Carlo (MC) simulated samples are used for many different purposes, mainly for estimating the Standard Model backgrounds and probing the kinematics of possible models of new physics. Another field of application of MC samples is calibrations of the reconstructed objects, since they contain information about the original particles before the interaction with the detector. The event simulation can performed in two steps, *event generation* and *detector simulation*.
The event generation starts with the hard scattering process between two incoming partons and the corresponding matrix element (ME) calculation at a given order of perturbation theory based on the relevant Feynman diagrams. This stage is referred to as *parton level*. At this stage the parton distribution function (PDF) is used to describe the constituents of the proton and their momentum fractions at a given energy scale. On a top of this, the remormalisation scale is defined to avoid ultraviolet divergencies and fix the running coupling of the strong interaction, while the factorization scale is introduced to avoid infrared divergencies.
The next stage simulates the evolution from the hard scattering scale down to the scale at which hadronisation occurs using different algorithms of parton showering. At this stage partons shower, either radiating off partons or splitting into pairs of other partons. Finally, partons hadronise and form stable colourless bound states. This stage is referred to as *particle level* or *truth level*. The hadronisation is a purely non-perturbative process and is described by different phenomenological models, such as *string model* [137] or *cluster model* [138].
And finally, the last step of the event simulation is to simulate interactions of the final state particles with the ATLAS detector. This is done by using the GEANT4 simulation toolkit [139]. Starting from this stage, the object reconstruction and event selection is performed in the same way as in real data.
The $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analyses make use of the following MC generators:

- Sherpa [140] is a general-purpose MC generator, providing the ME calculation and parton shower evolution. It generates events with up to three final state jets. The cluster model of hadronisation is used [138]. The CKKW method [141] is used to combine calculations of the MEs and parton showers. It is equipped with the CT10 PDF set [142] at next-to-leading-order (NLO) and with the NNPDF3.0NNLO PDF set [143] at next-to-next-to-leading order (NNLO). Sherpa is used to simulate V + jets and diboson background processes.

- Pythia [140] is a general-purpose MC generator, providing the ME calculation at LO and NLO, as well as parton shower evolution. The hadronisation process is modeled using Lund string model [144]. The Pythia 8, equipped with the A14 tune [145] and NNPDF23LO PDF [146] set, is used to simulate SM $Vh(b\bar{b})$ ($gg \to Vh(b\bar{b})$) process for the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis.

- Powheg+Pythia is a general-purpose MC generator, providing the ME calculation and parton shower evolution. Powheg [140] provides ME calculation using an advanced ME reweighting procedure, where the hardest interaction term is replaced by its NLO-weighted correspondent. Pythia 6 version is interfaced to simulate parton showering, hadronisation and underlying events. Powheg+Pythia, equipped

with the A14 tune [145] and CT10 PDF set [142] set, is used to simulate $t\bar{t}$, single top and SM $Vh(b\bar{b})$ ($q\bar{q} \to Vh(b\bar{b})$) background processes.

- MadGraph+Pythia [147] is a leading-order MC generator. MadGraph5 provides the leading-order calculations of the MEs, while Pythia is interfaced to simulate parton showering, hadronisation and underlying events. MadGraph+Pythia is used to produce MC samples for the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ simplified $s-$channel mediator models and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ Z'-2HDM models, described in Section 3.4.1.

- MadGraph_aMC@NLO+Pythia is a next-to-leading-order MC generator. MadGraph5_aMC@NLO [148] provides the next-to-leading-order calculations of the MEs, while Pythia is interfaced to simulate parton showering, hadronisation and underlying events. DMSimp tune is employed to define the particle content and the interactions of the simplified model. MadGraph5_aMC@NLO+Pythia is used to produce MC samples for the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ simplified $s-$channel mediator models at NLO.

A summary of the MC samples used to simulate signal and background processes is given in Table 5.1.

Process	Monte Carlo generator	Tune + PDF set
Simplified $s-$channel mediator	MadGraph5+Pythia 8 (LO)	A14+NNPDF23LO
Simplified $s-$channel mediator	MadGraph5_aMC@NLO+Pythia 8 (NLO)	A14+NNPDF23LO
Z'-2HDM signal	MadGraph5+Pythia 8	A14+ NNPDF30_lo_ as_0130
V+jets		
$W(l,\nu)$+jets	Sherpa 2.2.1 (NNLO)	NNPDF3.0NNLO
$Z(\ell,\nu\nu)$+jets	Sherpa 2.2.1 (NNLO)	NNPDF3.0NNLO
$t\bar{t}$	Powheg+Pythia 6 (NLO)	A14+CT10
Single top		
$s-$channel	Powheg+Pythia 6 (NLO)	CT10
$t-$channel	Powheg+Pythia 6 (NLO)	CT10
$Wt-$channel	Powheg+Pythia 6 (NLO)	CT10
Diboson		
$W(l\nu)W(q'\bar{q})$ Sherpa 2.1 (NLO)	CT10	
$W(l\nu)Z(q\bar{q})$ Sherpa 2.1 (NLO)	CT10	
$W(q'\bar{q})Z(\ell)$ Sherpa 2.1 (NLO)	CT10	
$W(q'\bar{q})Z(\nu\nu)$ Sherpa 2.1 (NLO)	CT10	
$Z(q\bar{q})Z(\nu\nu)$ Sherpa 2.1 (NLO)	CT10	
$Z(q\bar{q})Z(\ell)$ Sherpa 2.1 (NLO)	CT10	
SM Vh ($q\bar{q} \to Vh(bb)$)	Pythia 8 (LO)	A14+NNPDF2.3LO
($gg \to Vh(b\bar{b})$)	Powheg+Pythia 8 (LO)	A14+CT10

Table 5.1: List of the MC generators, corresponding tunes and PDF sets used to produce MC simulated samples for the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analyses.

5.2 Triggers

The *lowest unprescaled* triggers [149] are used both in data and MC throughout $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analyses. Here, the lowest unprescaled trigger is defined as unprescaled trigger (see Section 4.2.4.1) with the lowest trigger threshold. Since events in 0 and 1 lepton region contain non-zero genuine $E_{\mathrm{T}}^{\mathrm{miss}}$ coming from the possible signal

and semileptonic W, Z and $t\bar{t}$ decays, they are selected using $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers. Events in 2 lepton region are selected using different logical combinations of single electron and muon triggers.

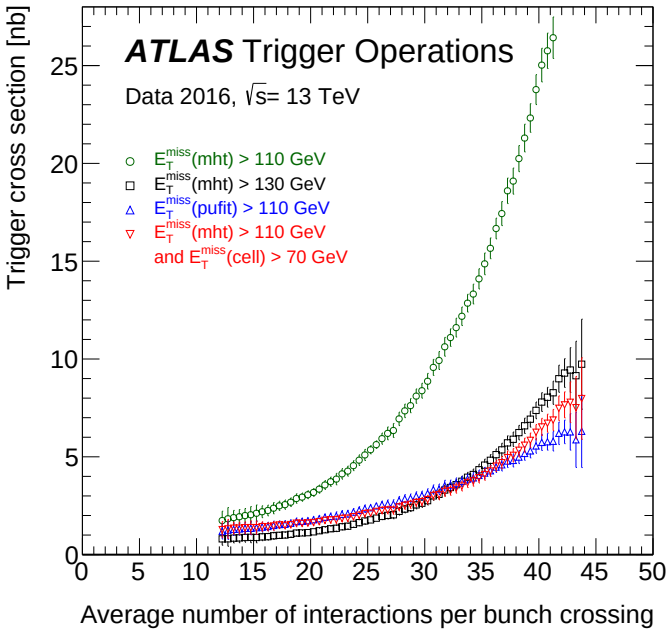


Figure 5.2: Trigger cross section as a function of average number of interactions per LHC bunch crossing for various $E_{\mathrm{T}}^{\mathrm{miss}}$ definitions and thresholds. The $E_{\mathrm{T}}^{\mathrm{miss}}$ of the "mht" algorithm is calculated as the negative vectorial sum of the transverse momenta of all jets reconstructed using the anti-k_T jet finding algorithm (see Section 6.2.2). The pile-up subtraction and JES calibration (see Section 6.2.3) is applied to these jets. The $E_{\mathrm{T}}^{\mathrm{miss}}$ of the "pufit" algorithm is calculated as the negative vectorial sum of the transverse momenta of all calorimeter topological clusters corrected for pile-up.

Jet energy mismeasurements in case of large rate of hadronic jet production and additional pile-up, which deposits significant energy in the calorimeter, lead to significant increase in $E_{\mathrm{T}}^{\mathrm{miss}}$ and the corresponding $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger rates. Different pile-up reduction techniques at the trigger level, in particular different $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction algorithms, can enhance the $E_{\mathrm{T}}^{\mathrm{miss}}$ performance, as shown in Figure 5.2. But in general the problem of large $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger rates remains acute, especially in case when lumonisity increases throughout the data taking period. One can only constantly increase the trigger thresholds of unprescaled triggers with the luminosity to keep the trigger rate under control. Increasing trigger thresholds usually leads to the reduced efficiencies in physics analyses that use corresponding triggers.
Lepton triggers are much less sensitive to the changes in luminosity, but still different combinations of the triggers are used for each data taking period.
Both analyses share the same list of triggers shown in Table 5.2 for each period of 2015 and 2016 data taking.

Period	0 lepton	1 lepton	2 lepton
2015	HLT_xe70	HLT_xe70	HLT_e24_lhmedium_L1EM18VH (MC only)[2] **OR** HLT_e24_lhmedium_L1EM20VH (data only)[2] **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 only)[2] **OR** HLT_mu24_iloose (data only)[2] **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 5.2: Summary table of $E_{\mathrm{T}}^{\mathrm{miss}}$ and lepton triggers implemented in $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis from 2015 and 2016 data taking periods. Triggered labelled as "MC only" and "data only" correspond to $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger implemented only in MC and data, respectively.

5.2.1 $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers in $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analyses

Figure 5.3 illustrates how the luminosity was changing during 2016 data taking period. To reduce pile-up dependence and to keep trigger thresholds relatively low, jet-based $E_{\mathrm{T}}^{\mathrm{miss}}$ is used in the definition of the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger. The $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers used in both analyses for different periods of 2015 and 2016 data taking are the following:

- 2015 (3.2 fb^{-1}): HLT_xe70
- 2016, periods A-D3 (6.1 fb^{-1}): HLT_xe90_mht_L1XE50
- 2016, periods D4-E3 (3.9 fb^{-1}): HLT_xe100_mht_L1XE50 **OR** HLT_xe110_mht_L1XE50
- 2016, periods F1– (23.2 fb^{-1}): HLT_xe110_mht_L1XE50

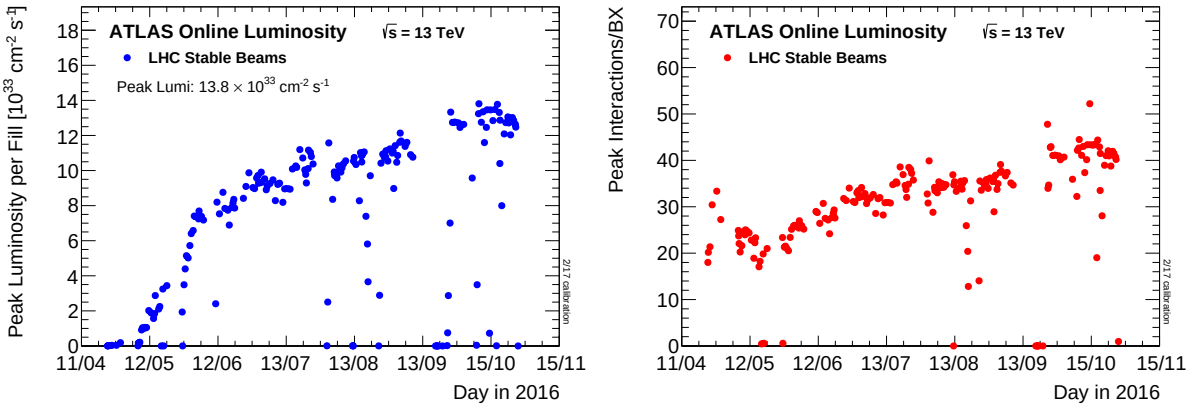


Figure 5.3: The peak instantaneous luminosity delivered to ATLAS (left) and the maximum number of inelastic collisions per bunch crossing (right) during stable beams per each LHC fill as a function of time in 2016.

5.2.2 $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency measurements

The performance of the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger is studied in terms of $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency curve (turn-on), where trigger efficiency is defined as follows:

$$\text{efficiency} = \frac{N_{\text{events}} \text{ passed offline selection } \textbf{AND } E_{\mathrm{T}}^{\mathrm{miss}} \text{ trigger requirement}}{N_{\text{events}} \text{ passed offline selection}}, \qquad (5.1)$$

and is measured as a function of offline reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$. The ideal trigger turn-on looks like a step-function, which immediately switches from the 0% to 100% efficiency state exactly at the trigger threshold value. It corresponds to the identical offline and trigger-level $E_{\mathrm{T}}^{\mathrm{miss}}$ [1]. But in real measurements due to the limited detector resolution, both offline and trigger-level $E_{\mathrm{T}}^{\mathrm{miss}}$ measurements are no longer identical. The corresponding Gaussian distributed $E_{\mathrm{T}}^{\mathrm{miss}}$ response smears the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger turn-on, leading to the trigger inefficiency in the part of the analyses region above the $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement of $E_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV. Thus, it is important to study trigger efficiency both in data and MC exactly in this region. Since MC does not perfectly model the data, the performance of $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers and corresponding efficiency curves are slightly different in data and MC. Thus, the corresponding MC triggers are corrected using scale factors derived from the trigger efficiency measurements in data and MC.

Both $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analyses share same signature consisting of jets and $E_{\mathrm{T}}^{\mathrm{miss}}$. In order to study $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiencies, one needs to consider physics processes, which provide same signal-like signature at the trigger level. One exploits the fact that trigger-level $E_{\mathrm{T}}^{\mathrm{miss}}$ is reconstructed using only calorimeter energy deposits meaning that by definition, muons are not considered in the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation. It means that trigger-level $E_{\mathrm{T}}^{\mathrm{miss}}$ and muon energy, as well as the corresponding triggers, are orthogonal. Thus, one can use muon triggers to select events for the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger measurements. Unlike the trigger-level $E_{\mathrm{T}}^{\mathrm{miss}}$, the offline $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction considers muons, as described in Section 6.4.1. In order to make direct correspondence between the offline and trigger-level $E_{\mathrm{T}}^{\mathrm{miss}}$, the corresponding muon contribution is subtracted from the offline $E_{\mathrm{T}}^{\mathrm{miss}}$. This is referred to as $E_{\mathrm{T,nomu}}^{\mathrm{miss}}$.

A single muon region satisfies all above-mentioned criteria. It is orthogonal to the signal

[1] In other words it means that the corresponding $E_{\mathrm{T}}^{\mathrm{miss}}$ response is a $\delta-$function at any $E_{\mathrm{T}}^{\mathrm{miss}}$ value.

region with its lepton veto and provide required signal-like signature with genuine $E_{\mathrm{T}}^{\mathrm{miss}}$ coming primarily from $W \to \mu\nu$ and semi-leptonic $t\bar{t}$ decays. This region is dominated by $W(\mu,\nu)$+jets and semi-leptonic $t\bar{t}$ events with small contributions from the signle top, $Z(\mu,\mu)$+jets and di-boson production.
All $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger used in the analyses become fully efficient below $E_{\mathrm{T}}^{\mathrm{miss}} > 250$ GeV, and thus are completely covered by resolved regions of both analyses. The corresponding event selections are similar to main 1 lepton selections of both analyses. Events in data and MC are selected using lowest unprescaled single muon triggers. $E_{\mathrm{T}}^{\mathrm{miss}}$ and p_T^{miss} (in case of $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ only) are dropped as they directly affect the shape of the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger turn-ons. The final selection looks as follows:

- Muon triggers:
  - 2015: HLT_MU20_ILOOSE_L1MU15 **OR** HLT_MU50
  - 2016, period A: HLT_MU24_ILOOSE (only used in data)[2]
    OR HLT_MU24_ILOOSE_L1MU15 (only used in MC)[2]
    OR HLT_MU50
  - 2016, period B–D3: HLT_MU24_IVARMEDIUM
    OR HLT_MU50
  - 2016, periods D4–: HLT_MU26_IVARMEDIUM
    OR HLT_MU50
- Exactly one *medium* or *tight* muon candidate with $p_T > 25$ GeV and ISLOOSETRACKONLY isolation working point (see Sections 6.1.2.2 and 6.1.2.3).
- *Loose* electron veto.
- The offline part of the event selection is equivalent to the resolved selection in the 1 lepton control region (see Section 7.10), except $E_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV and $p_T^{\mathrm{miss}} > 30$ GeV (in case of $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ only) cuts that are dropped.
- Sensitivity optimisation cuts are applied in $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis, as described in Section 7.8.

All $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiencies are studied in events with a single muon, which passed the list of above-mentioned selection criteria, iinclusive in 0, 1 and 2 $b-$tags. In case of $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis, as no particular number of $b-$jets in the event is required, trigger efficiencies are studied separately in the events with 0 and ≥ 1 $b-$tags. Events with 0 and 1 $b-$tags are dominated by $W+$jets production, whilst events with 2 $b-$tags by $t\bar{t}$ production.
Trigger efficiency are studied for each $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger in the corresponding data taking period. Since data taking periods are not available in MC samples, a randomized periodization procedure for MC is applied. Each MC event is randomly assigned to the event from any of the 2015 or 2016 data taking period, so that the luminosity of MC periods correspond to the luminosity of the corresponding periods in data.
The resulting $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency curves for $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analyses are shown in Figures 5.4 and 5.5. All comparisons show up to 10% difference between the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiencies in data and MC.

[2]MC does not contain all the triggers used in data taking [150].

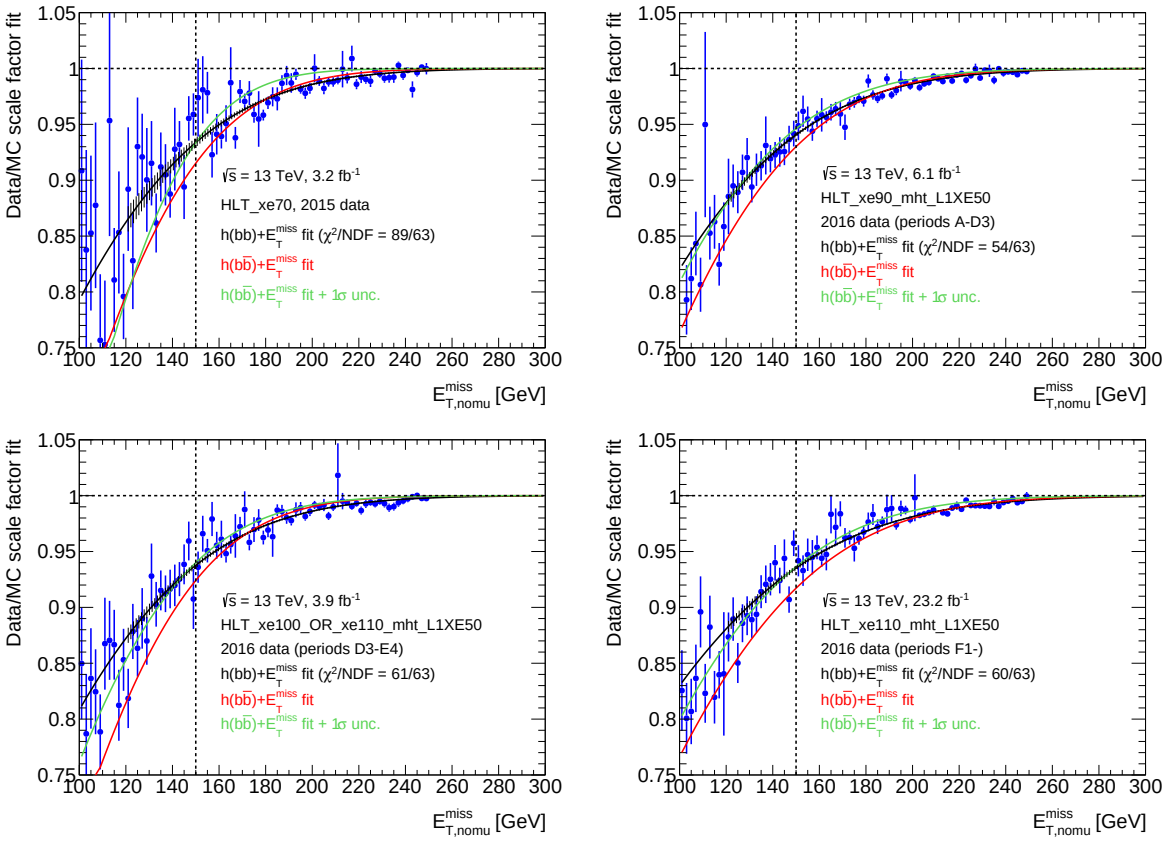


Figure 5.4: Measured $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiencies and corresponding scale factors as function of offline $E_{\mathrm{T,nomu}}^{\mathrm{miss}}$ measured in $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis. Plots for HLT_XE70 (top left), HLT_XE90_MHT_L1XE50 (top right), HLT_XE100_MHT_L1XE50 **OR** HLT_XE110_MHT_L1XE50 (bottom left) and HLT_XE110_MHT_L1XE50 (bottom right) triggers are shown in data and MC for 1 lepton control region. Plots are shown for all $b-$tagged analysis regions together. The MC is dominated by $W+$jets and $t\bar{t}$ events.

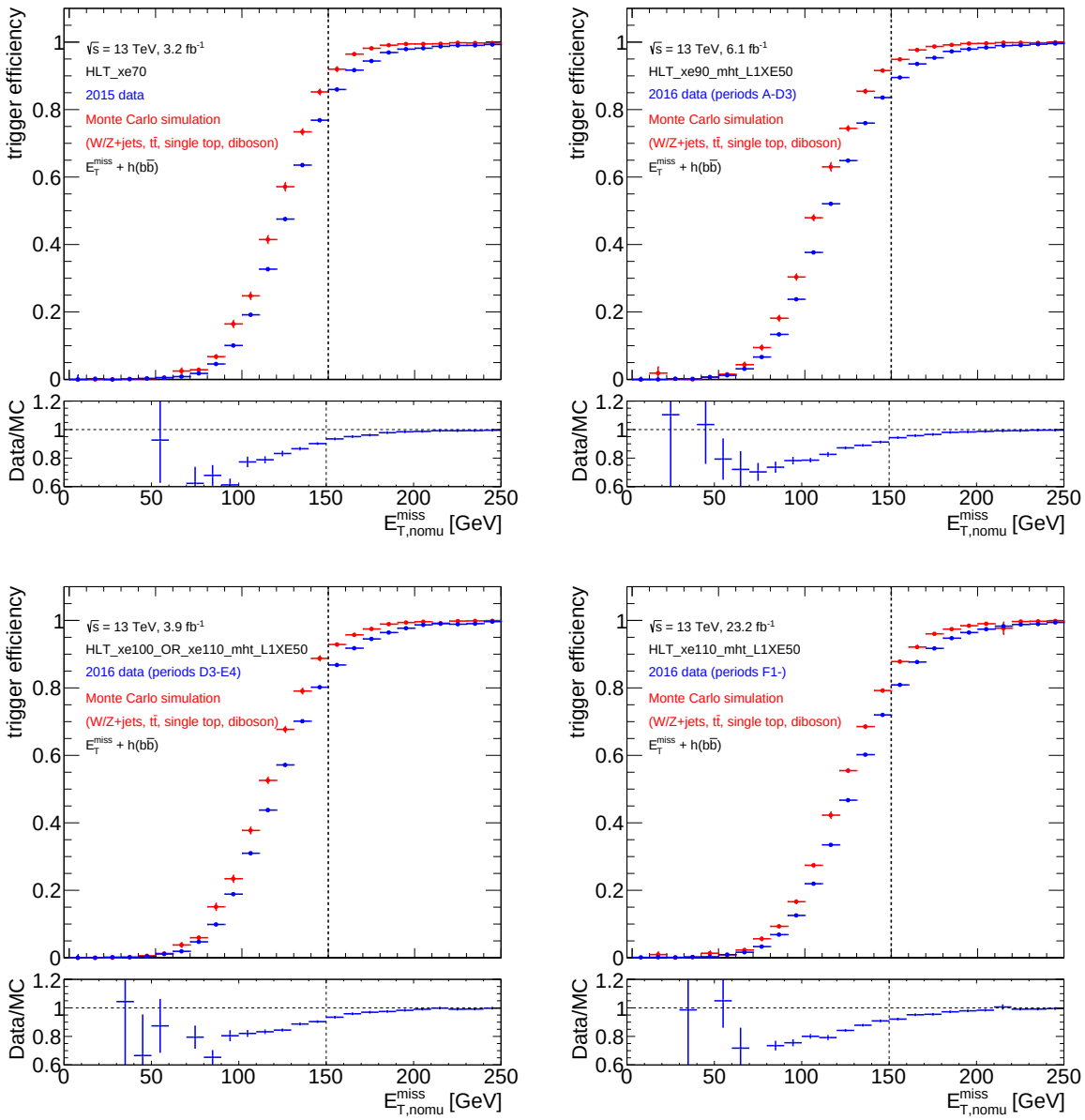


Figure 5.5: Measured $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiencies and corresponding scale factors as function of offline $E_{\mathrm{T,nomu}}^{\mathrm{miss}}$ measured in $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis. Plots for HLT_XE70 (top left), HLT_XE90_MHT_L1XE50 (top right), HLT_XE100_MHT_L1XE50 **OR** HLT_XE110_MHT_L1XE50 (bottom left) and HLT_XE110_MHT_L1XE50 (bottom right) triggers are shown in data and MC for 1 lepton control region. Plots are shown for all $b-$tagged analysis regions together. The MC is dominated by $W+$jets and $t\bar{t}$ events.

5.2.3 Data-driven $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger scale factors

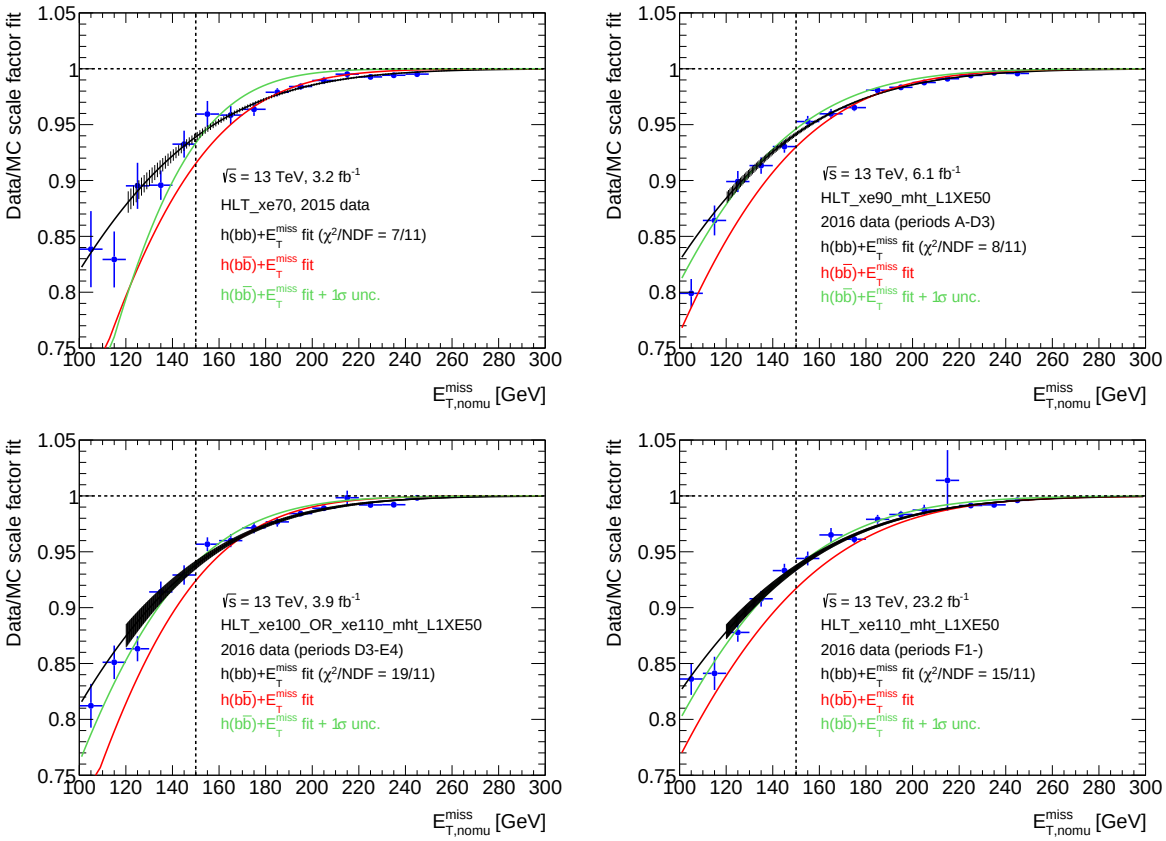


Figure 5.6: $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger scale factors as a function of offline $E_{\mathrm{T,nomu}}^{\mathrm{miss}}$ measured in $E_{\mathrm{T}}^{\mathrm{miss}}+V(qq)$ analysis. Plots for HLT_XE70 (top left), HLT_XE90_MHT_L1XE50 (top right),HLT_XE100_MHT_L1XE50 **OR** HLT_XE110_MHT_L1XE50 (bottom left) and HLT_XE110_MHT_L1XE50 (bottom right) triggers are shown in data and MC for 1 lepton control region. The hatched bands show the 1σ fit uncertainties. The green and red lines labelled as "$E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ fit" show $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger scale factors and 1σ systematic uncertanties measured in $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ analysis [46]. Note that they are not identical to $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ scale factors due to the differences in the final states but agree within systematic uncertainties. Plots are shown for all $b-$tagged analysis regions together. MC is dominated by $W+$jets and $t\bar{t}$ events.

Data-driven MC correction is derived as a function of $E_{\mathrm{T}}^{\mathrm{miss}}$ based on the trigger efficiency scale factors (SF) defined as ratio of trigger efficiencies in data and MC in a single muon region, as described in Section 5.2.2:

$$\mathrm{SF}(E_{\mathrm{T}}^{\mathrm{miss}}) = \frac{\varepsilon_{\text{1-muon}}^{\text{trig,Data}}(E_{\mathrm{T}}^{\mathrm{miss}})}{\varepsilon_{\text{1-muon}}^{\text{trig,MC}}(E_{\mathrm{T}}^{\mathrm{miss}})}. \tag{5.2}$$

In order to obtain a smooth data-driven correction for the MC $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger turn-ons over the entire $E_{\mathrm{T}}^{\mathrm{miss}}$ range in the resolved regime, the scale factors for each $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger are fitted starting from 120 GeV(150 GeV) in $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ ($E_{\mathrm{T}}^{\mathrm{miss}}+V(qq)$) analysis using the following fit function:

$$f\left(E_{\mathrm{T}}^{\mathrm{miss}}\right) = \frac{1}{2}\left[1+\mathrm{erf}\left(\frac{E_{\mathrm{T}}^{\mathrm{miss}}-p_0}{\sqrt{2}p_1}\right)\right]. \tag{5.3}$$

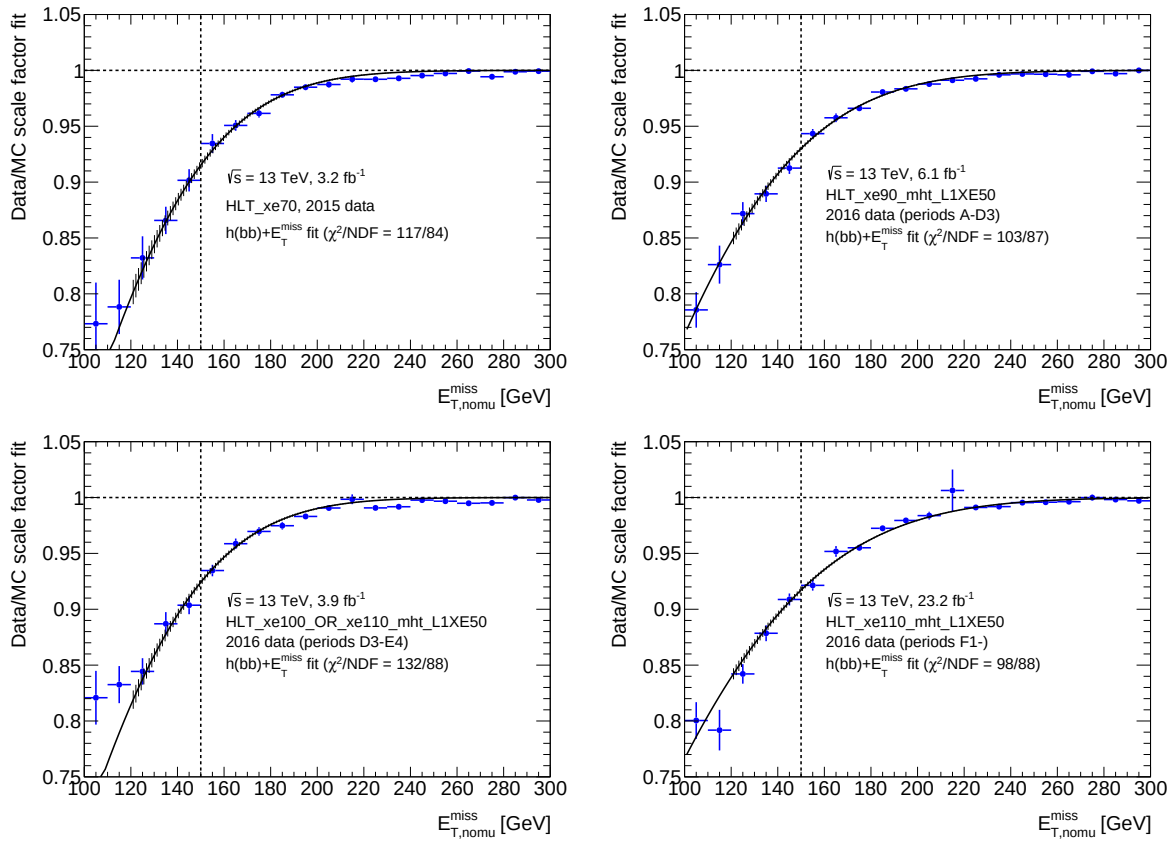


Figure 5.7: $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger scale factors as a function of offline $E_{\mathrm{T,nomu}}^{\mathrm{miss}}$ measured in $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis. Plots for HLT_XE70 (top left), HLT_XE90_MHT_L1XE50 (top right),HLT_XE100_MHT_L1XE50 **OR** HLT_XE110_MHT_L1XE50 (bottom left) and HLT_XE110_MHT_L1XE50 (bottom right) triggers are shown in data and MC for 1 lepton control region. The hatched bands show the 1σ fit uncertainties. Plots are shown for all $b-$tagged analysis regions together. MC is dominated by $W+$jets and $t\bar{t}$ events. The fit corresponds to the $E_{\mathrm{T}}^{\mathrm{miss}}$ bins of 2 GeV.

The results of the SF calculation and fitting are shown in Figure 5.6 and 5.7. The corresponding data-driven correction is applied to the MC for each $E_{\mathrm{T}}^{\mathrm{miss}}$ value using Eq. 5.3 with parameters obtained from the fit. Validation plots of $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiencies after the scale factor corrections are shown in Appendix A.

5.2.4 Uncertainties for data-driven $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger scale factors

The estimation procedure of the systematic uncertainties is initially developed for the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis, and is fully followed by the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis. The following variations are considered as possible sources of systematic uncertainties on the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger scale factors:

- **b−tag** composition: It accounts for the different calorimeter response to the light- and heavy-flavour jets that affects the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation, both offline and trigger-level. Thus, the difference between scale factors obtained for all $b-$tags (dominated by 0 $b-$ tag) and for ≥ 1 $b-$tag is estimated.

- **Sample composition in data**: There is a lack of knowledge about the precise composition of the sample in data, which can be an additional ssystematic uncertainty. The following conservative method is used to estimate this effect. The $E_{\mathrm{T}}^{\mathrm{miss}}$

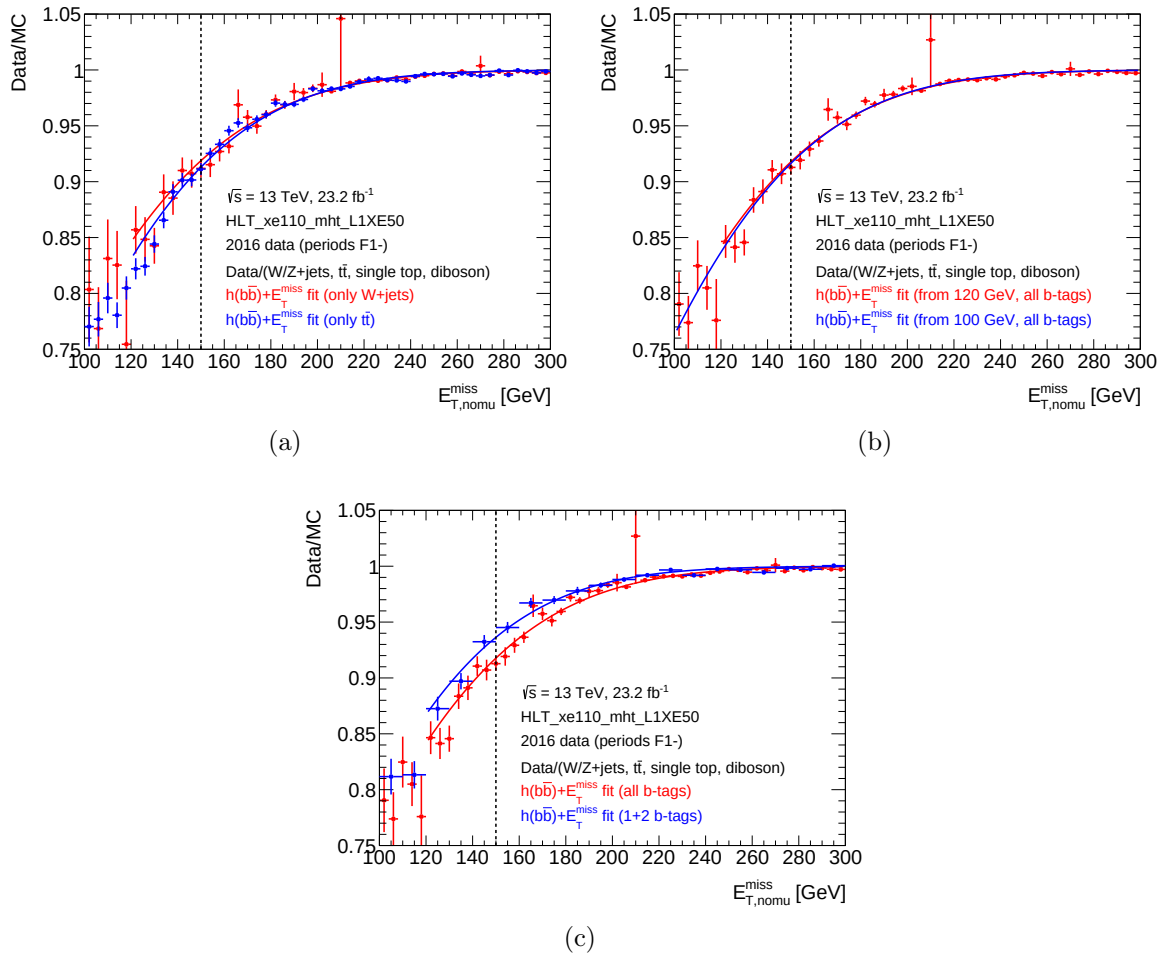


Figure 5.8: Different variations of scale factors for HLT_XE110_MHT_L1XE50 trigger as a function of offline $E_{\mathrm{T,nomu}}^{\mathrm{miss}}$ measured in $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis. Figure (a) shows comparison between scale factors derived for $W+$jets only and $t\bar{t}$ only MC samples. Figure (b) shows comparison between the nominal scale factor fit starting from 120 GeV and the fit starting from 100 GeV. Figure (c) shows comparison between the scale factors derived for all $b-$tagged analysis regions together and for ≥ 1 $b-$tagged analysis regions. The nominal fit corresponds to the $E_{\mathrm{T}}^{\mathrm{miss}}$ bins of 2 GeV. The fits for $W+$jets MC samples, $t\bar{t}$ MC samples and for the extended fit range from 100 GeV corresponds to the $E_{\mathrm{T}}^{\mathrm{miss}}$ bins of 4 GeV, while the fit for ≥ 1 $b-$tagged analysis regions corresponds to the $E_{\mathrm{T}}^{\mathrm{miss}}$ bins of 10 GeV. This is done due to the lack of statistics.

trigger scale factors are calculated separately, once assuming that sample consists of $W+$jets and once of $t\bar{t}$, inclusive in the number of $b-$tags. The absolute difference between derived scale factors in taken as the uncertainty.

- **Extended fit range**: Fit range is extended down to 100 GeVto study the influence of the low $E_{\mathrm{T}}^{\mathrm{miss}}$ tail on the fit.

All variations for HLT_XE110_MHT_L1XE50 are shown in Figure 5.8. Variations for other $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers are shown in Appendix B. A conservative estimation of the systematic uncertainties is considered, when the variation with the largest effect is taken as a systematic uncertainty. The $b-$tag composition uncertainty shows the largest effective and thus is taken as systematic uncertainty. The statistical uncertainty is estimated as the 1σ confidence interval of the fit, as shown in hashed bands in Figure 5.6. The total uncertainty is derived as a sum in quadrature of statistical and systematic uncertainties.

5.2.5 $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiencies in 0 lepton region.

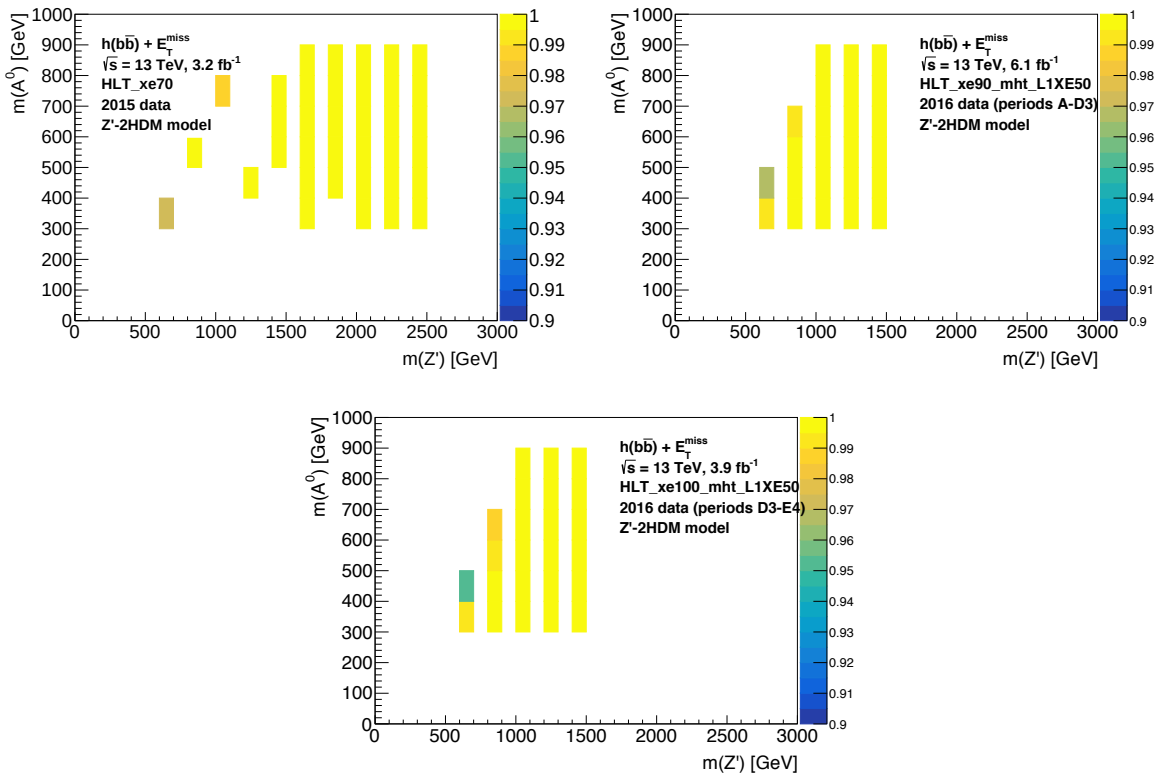


Figure 5.9: Measured $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiencies of HLT_xe70, HLT_xe90_mht_L1XE50 and HLT_xe100_mht_L1XE50 triggers in $(m_{Z'}, m_A)$ plane for Z'-2HDM model after $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ signal selection.

The $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiencies are studied in data and MC, and the same study is performed in 0 lepton region using signal models of the dark matter production. The corresponding study is performed in $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis using simplified Z'-2HDM model described in Section 3.4.1.2 with the mass of the pseudoscalar 300 GeV $< m_A <$ 800 GeV and the mass of Z' 600 GeV $< m_{Z'}$2400 GeV. The $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ 0 lepton selection described in Chapter 7 is applied to select events for the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency study. The effiencies of HLT_xe70, HLT_xe90_mht_L1XE50 and HLT_xe100_mht_L1XE50 triggers for Z'-2HDM model are shown in Figure 5.9. Since Z'-2HDM mass points considered in $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis provide relatively hard $E_{\mathrm{T}}^{\mathrm{miss}}$ spectrum, all $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers are almost fully efficient with the lowest value of 95% for the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger with 100 GeVthreshold.

Chapter 6

Object reconstruction and performance at ATLAS

The topology of every search requires the presence or absence of specific objects in the final state. These objects can be either particles, such as photons, electrons, muons and $\tau-$leptons, or jets of hadrons from parton showers and hadronisation of quarks. All these objects leave traces in the detector, meaning that calibrated energy deposits and tracks can be used to reconstruct and identify each object along with its kinematic properties. This is referred to as *object reconstruction* and *identification* (ID). Due to the different detector effects, objects cannot be reconstructed and identified with 100% efficiency. Another special object is the missing energy, which quantifies the momentum carried away by the invisible particles, such as neutrinos or possible DM particles. The reconstruction and ID criteria as well as the corresponding efficiencies are studied by the different combined performance (CP) groups within the ATLAS colaboration that provide CP recommendations for different data taking periods.

Since the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ searches share very similar final state that consists primarily of jets and the missing transverse momentum, the main focus is put on the reconstruction and performance of these objects described in Section 6.2 and 6.4, respectively. Since both searches expect also $b-$quark jets in the final state, which are treated separately from $c-$ and light-flavour jet, the $b-$jet reconstruction and tagging are discussed in Section 6.2.7. Due to the presence of the regions with different lepton multiplicities in both searches, as shown in Section 7, electron and muon reconstruction and performance are also discussed in Section 6.1.

6.1 Leptons

Even though the primary objects to reconstruct in $E_{\mathrm{T}}^{\mathrm{miss}}+V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ analyses are jets and $E_{\mathrm{T}}^{\mathrm{miss}}$, leptons still play a significant role in definition of the analysis regions. In particular, the lepton veto is used to reject background events in the signal region, while electrons and muons are used to define the 1- and 2-lepton control regions, as discussed in Chapter 8. This Section is aimed to discuss matters related to the electron and muon reconstruction (Sections 6.1.1.1 and 6.1.2.1), identification (Sections 6.1.1.2 and 6.1.2.2) and isolation (Sections 6.1.1.3 and 6.1.2.3).

6.1.1 Electrons

Electrons deposit most of their energy in the EM calorimeter in the form of EM showers. Since electrons are charged particles, they also leave tracks in the inner detector (ID). Thus, the electron reconstruction relies both on the energy deposits in the EM calorimeter and on tracks in the ID. The matters related to the electron reconstruction, identification and isolation are discussed in Sections 6.1.1.1, 6.1.1.2 and 6.1.1.3 respectively, following Ref. [151].

6.1.1.1 Electron reconstruction

Since electron reconstruction requires track information, the available reconstruction region is limited by the angular acceptance of the ID, $|\eta| < 2.47$. The electron reconstruction is performed in several steps:

- **Seed-cluster reconstruction:** reconstruction starts from the fixed-size seeds composed of 3×5 towers of size $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ each, following the granularity of the EM calorimeter middle layer. The total transverse energy E_T of the seed is calculated as a sum of the energy deposited in all cells of the towers in the seed. The *sliding window clustering algorithm* [152] is used to form the cluster around the seed with $E_T > 2.5$ GeV. The efficiency of this clustering algorithm varies from 95% for $E_T = 7$ GeV to $> 99\%$ for $E_T = 15$ GeV.

- **Electron track reconstruction:** consists of two steps, pattern recognition and track fit. The ATLAS pattern recognition uses the pion and electron hypotheses for the energy losses in the detector material. Track seeds, consisting of 3 hits in different layers of the ID, with $p_T > 1$ GeV, are used to obtain the track candidates with at least 7 hits each. Track candidates are then fitted, using the ATLAS Global χ^2 Track Fitter [153].

- **Electron specific track fit:** after tracks are reconstructed, they are matched to the clusters in the EM calorimeter using the angular distance $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ between the position of the track, extrapolated in the middle layer, and the cluster barycenter. If matching fails, the cluster is identified as an unconverted photon from the e^+e^- annihilation. The matching also takes into account the energy losses from the Bremsstrahlung and the number of hits on the track. Successfully matched tracks with at least 4 hits are then refitted using an optimised Gaussian Sum Filter (GSF) [154] to account for the non-linear Bremsstrahlung effects.

- **Electron candidate reconstruction:** once the cluster seed is matched to the track candidate, the cluster is considered to be an electron candidate.

The final energy of the electron candidate is calculated using the energy of the calibrated cluster [155], while η and ϕ are obtained using angular parameters of the track with respect to the beam-line.
In order to further reduce backgrounds from photon conversions and secondary particles, tracks associated with the electron clusters are matched to the PV ($d_0/\sigma_{d_0} < 5$ and $|z_0 \sin\theta| < 0.5$ mm).

6.1.1.2 Electron identification

Electron identification (ID) algorithm is used to separate signal-like electrons from background-like objects like hadronic jets or converted photons. The likelihood-based (LH) approach is used in the baseline ID algorithm. It utilises different cluster- and track-based information, such as calorimeter shower shapes, track properties, matching of tracks to the clusters, information from the Transition Radiation Tracker (TRT), different bremsstrahlung effects, and so on. Using the combination of the LH discriminant and the number of hits in the track, the probability to be a signal or background electron is calculated.

Depending on the LH discrimination power, three operational points (OP) are introduced for the electron ID, *loose*, *medium* and *tight*, ordered by increasing background rejection. All these OPs are inclusive, meaning that *loose* is a sub-set of *medium* and *tight* and *medium* is a sub-set of *tight*, and share the same discriminant but different selection. The performace of the ID algorithm is shown in Figure 6.1. Depending on the OP and electron energy, the signal ID efficiencies vary from $\sim$ 78% to $\sim$ 97%, increasing with energy.

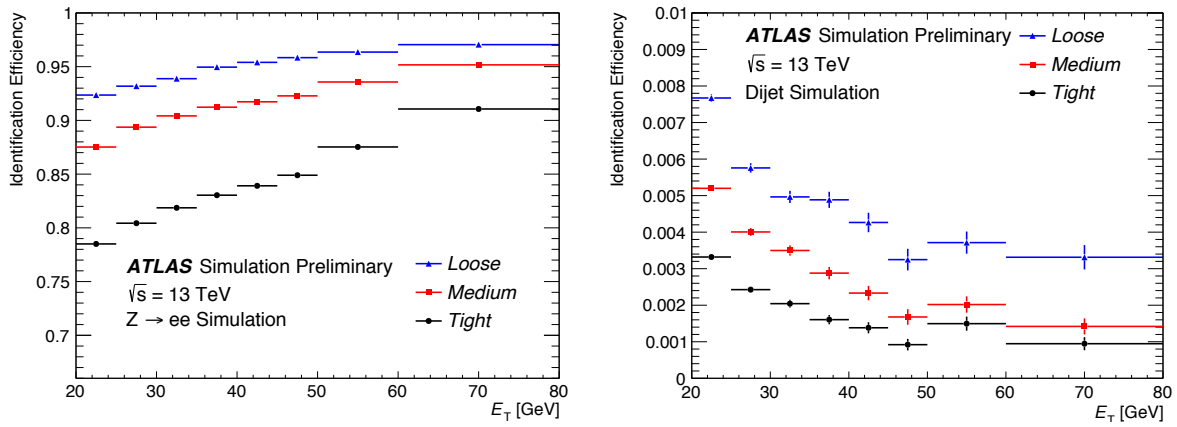


Figure 6.1: Efficiency to identify signal electrons (left) based on $Z \to ee$ MC events and to identify hadrons as electrons (right) based on the dijet MC events [151].

6.1.1.3 Electron isolation

The $E_{\mathrm{T}}^{\mathrm{miss}}+V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ analyses require electrons coming from decays of heavy particles, such as $W \to e\nu_e$ and $Z \to ee$, to be isolated from the background electrons coming from decays of the converted photons and misidentification of the light hadrons. Electron isolation, based on the measurement of the detector activity around the electron candidate, improves background rejection in the $E_{\mathrm{T}}^{\mathrm{miss}}+V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ analyses. Two main discriminating variables are used to define different isolation *working points* (WP). The first one is a calorimeter isolation energy, $E_T^{\mathrm{cone0.2}}$, defined as the sum of the transerve energy of the topological clusers [156] except the electron cluster itself, calibrated at the EM scale, within $\Delta R = 0.2$ of the electron cluster. The second one is a track isolation momentum, $p_T^{\mathrm{varcone0.2}}$, defined as the scalar sum of the transerve momenta of the tracks with $p_T > 1$ GeV originating from the PV (except the electron and converted photon tracks) with $\Delta R = \min(10\ \mathrm{GeV}/E_T, 0.2)$ of the electron track.

The isolation WPs, used in the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analyses, are LooseTrackOnly and FixedCutHighPtCaloOnly [157], which are defined as having fixed 99% efficiency in η and p_T and $E_T^{\mathrm{cone0.2}} < 3.5$ GeV, respectively.

6.1.2 Muons

Muons traverse the ATLAS calorimeter without a significant interaction with the calorimeter material, mostly depositing energy in the inner detector (ID) and muon spectrometers (MS). Thus, the muon reconstruction is primarily based on tracks reconstrusted in these parts of the ATLAS detector. In case when muons fall out of the acceptance of the MS, the calorimeter information is used. The matters related to the muon reconstruction, identification and isolation are discussed in Sections 6.1.2.1, 6.1.2.2 and 6.1.2.3 respectively, following [158, 159].

6.1.2.1 Muon reconstruction

Muon reconstruction starts with identifying the hits from segments in different layers of the MS. Track segments in the monitored drift tube (MDT) are obtained from a straight-line fit to the hits from each layer. Then track segments aligned on a trajectory in the bending plane of the detector are combined into the tracks. Parameters of the tracks are obtained from the global fit to the hits from the corresponding segments of the track. The muon track candidates are then built using the combination of tracks from the MS and ID. Four muon types are defined according to the algorithms used to perform such a combination, depending on which information from which different subsystems of the detector is used. The following muon types are used in the $E_{\mathrm{T}}^{\mathrm{miss}}+V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ analyses:

- **Combined (CB) muon:** Muon tracks reconstructed independently in the ID and MS, and then hits associated with the tracks from both subsystems are globally refitted to obtain a combined muon track.
- **Segment-tagged (ST) muon:** Tracks in the ID are identified as muon candidates if associated with at least one local track segment in the MDT or cathode strip chambers (CSC). It is used to identify muons with low p_T that leave energy deposits only in one layer of MS chambers, or muons that fall in the regions with low MS acceptance.

If ID track is identied as a muon candidate by several algorithms, the preference is given first to the CB muons and then to the ST muons. Finally, muon tracks are associated with the PV ($d_0/\sigma_{d_0} < 3$ and $|z_0 \sin\theta| < 0.5$ mm).

6.1.2.2 Muon identification

Muon identification (ID) provides a selection of reconstructed muon candidates to select signal-like muons coming from decays of heavy particles, such as W and Z bosons, and suppress background-like muons coming from the decays of light hadrons, such as pions and kaons, and from the decays of $B-$hadrons. For CB muon candidate, the main variables used by the ID algorithm are the following:

- q/p significance, $S_{q/p}$, defined as a charge-to-momentum ratio q/p of the muon candidate divided by the sum in quadrature of the corresponding uncertainties.

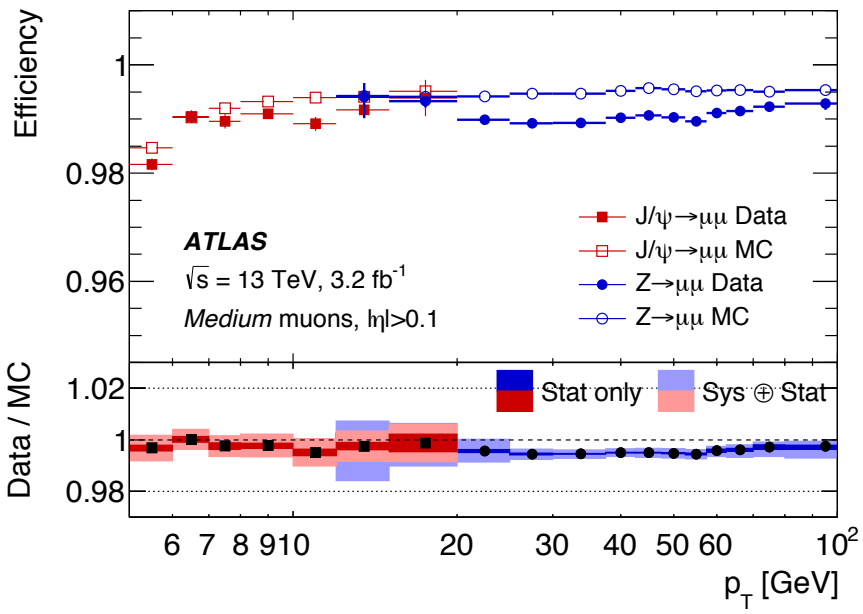


Figure 6.2: Reconstruction efficiency for the *medium* muons as a function of the muon p_T in the $0.1 < |\eta| < 2.5$ region based on $Z \to \mu\mu$ and $J/\psi \to \mu\mu$ events [158].

- $\rho' = \Delta p_T / p_T^{\mu}$, where Δp_T is the absolute value of the difference between the p_T measurements in the ID and MS.
- normalised χ^2 of the combined muon track fit.

Four different operational points, *loose*, *medium*, *tight*, *high-*p_T, are introduced to meet the needs of different physics analyses, where *loose*, *medium* and *tight* are inclusive categories. The *loose* and *medium* muons are used in the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ searches, and defined in the following way:

- ***Medium:*** default muon ID within ATLAS, which minimises systematic uncertainities coming from the muon reconstruction and calibration. Only CB and ME muon candidates are considered. The CB tracks are required to contain at least 3 hits in at least two MDT layers. For the tracks in $|\eta| < 0.1$ at least one MDT layer is required. The ME muon tracks are required to pass through at least three MDT/CSC layers in the region $2.5 < |\eta| < 2.7$. Loose criteria on the compatibility between the MS and ID p_T measurements is required, in particular $S_{q/p} < 7$. The reconstruction efficiency for the *medium* muons depends on the p_T of the muon, as shown in Figure 6.2.
- **Loose:** maximises the reconstruction efficiency and provides good-quality muon tracks. All muon types are involved. CB and ME muon candidates, which satisfy the *medium* criteria, are considered also as *loose* muons. CT and ST muons are considered only in the $|\eta| < 0.1$ region. The *loose* muons within $|\eta| < 2.5$ consist of $\sim 97.5\%$ CB muons.

6.1.2.3 Muon isolation

Muons from the decays of heavy particles, such as $W \to \mu\nu_\mu$ and $Z \to \mu\mu$, are mainly produced isolated from other particles, unlike muons from the semileptonic decays contained in jets. The latter can be accounted for by considering the muon-in-jet correction,

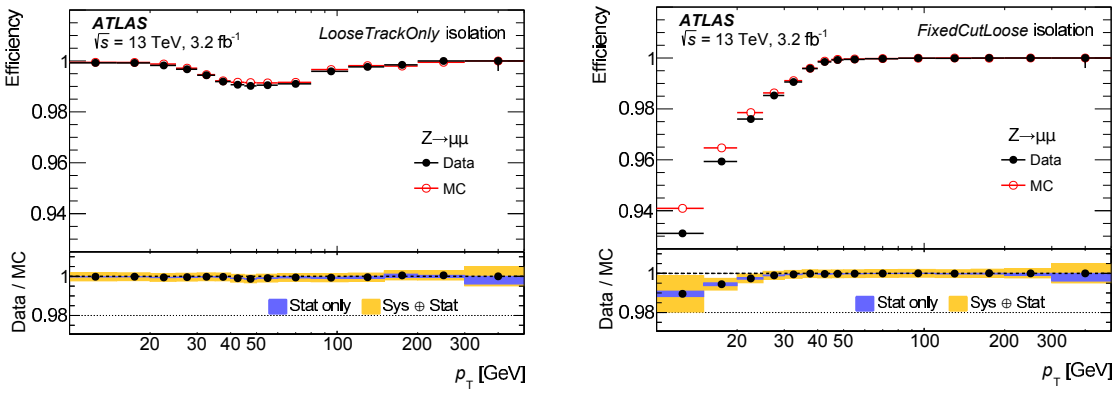


Figure 6.3: Muon isolation efficiencies as a functions of the muon p_T for the LooseTrackOnly (left) and FixedCutTightTrackOnly (right) WPs based on $Z \to \mu\mu$ events [158].

described in Section 6.2.8. Muon isolation based on the measurement of the detector activity around a muon candidate improves background rejection in the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analyses.
Main isolation WPs used in the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analyses are LooseTrackOnly and FixedCutHighPtTrackOnly [157]. Both of them use track-based isolation discriminant, $p_T^{\mathrm{varcone30}}$, defined as the scalar sum of the transverse momenta of the tracks with $p_T > 1$ GeV (except the muon track itself) within $\Delta R = \min(10\ \mathrm{GeV}/p_T^{\mu}, 0.3)$ of the muon track. In case of large muon p_T, ΔR becomes p_T-dependent to improve the perfomance for the high-p_T muons. The LooseTrackOnly WP is defined as having fixed 99% efficiency in η and p_T, and the FixedCutTightTrackOnly WP as having $p_T^{\mathrm{varcone30}} < 1.25$. The corresponding isolation efficiencies based on $Z \to \mu\mu$ events are shown in Figure 6.3.

6.2 Jets

Unlike other physical objects, such as electrons or muons, jets do not directly represent physical particles, but rather the underlying processes of the particle fragmentation and hadronization that form the jet. Charged hadrons associated with the jet leave tracks in the inner detector (ID), and together with the neutral hadrons deposit their energy in the hadronic (HAD) calorimeter as hadronic showers. The jet definition depends on an algorithm used to reconstruct a jet (*jet algorithm*), which utilise information both from energy deposits in the HAD calorimeter and tracks in the ID, as discussed in Section 6.2.2.

6.2.1 Small$-R$ and large$-R$ jets

The spray of particles produced from the hadronization of a parton is collimated due to the kinematic boost of the system. As the center-of-mass energy increases, the partons becomes more energetic, increasing the collimation of a hadronised system. The degree of a collimation drives the jet definition and is characterised by the angular separation in $\eta \times \phi$ space $\Delta R \simeq 2m/p_T$, where m and p_T are the mass and transverse momentum of the particle producing the jet.
Consider an example of a two-prong $h \to b\bar{b}$ decay in the scope of the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis. The jet radius parameter is defined by the mass $m \approx 125$ GeV and transverse

momentum p_T^h of the Higgs boson. For a relatively low boost, $p_T^h < 500$ GeV (or $E_{\mathrm{T}}^{\mathrm{miss}} < 500$ GeV), both $b-$quark jets from the Higgs decay are collimated with $\Delta R < 0.5$, and stay well-separated to be reconstructed as two separate *small*$-R$ jets, typically with similar energies. This kinematic regime is referred to as *resolved*. If the kinematic boost of the Higgs boson is large enough, $p_T^h > 500$ GeV (or $E_{\mathrm{T}}^{\mathrm{miss}} > 500$ GeV), $b-$quark jets tend to overlap and can be reconstructed only as *sub-jets* contained within the *large-R* jet [160, 161], using special jet substructure methods. This kinematic regime is referred to as *merged*. In case of $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis, the transition between the resolved and boosted regimes happens at $p_T^{W,Z} = 250$ GeV ($E_{\mathrm{T}}^{\mathrm{miss}} = 250$ GeV).

6.2.2 Jet reconstruction

Jet reconstruction starts from noise-suppressed, positive-energy topological calorimeter clusters, or *topo-clusters*, calibrated using either the electromagnetic (EM) scale (energy scale at which EM showers are measured) for the small$-R$ jets, or the local calibration weighted (LCW) scale [152] (energy scale depends on whether topo-cluster comes from the EM or hadronic shower) for the large$-R$ jets. To map the topo-clusters to the jets, one needs to define a particular jet algorithm, which satisfies a set of the following criteria [162]:

- Simple to implement both in the theoretical calculations and experimental analyses.
- Defined at any order of perturbation theory.
- Yields finite cross section at any order of perturbation theory.
- Yields a cross section that is relatively insensitive to hadronization.

The last three points are referred to as *collinear* and *infrared* (IR) safety, meaning that the jet multiplicity in every event stays unchanged under the multiple collinear splittings, such as $q \to qg$ and $g \to q\bar{q}$, and additional soft emissions, respectively [163]. Figures 6.4 and 6.5 illustrate how additional collinear emissions and soft gluon radiation can change the jet multiplicity provided by the jet algorithm, i.e. the collinear and IR safety.

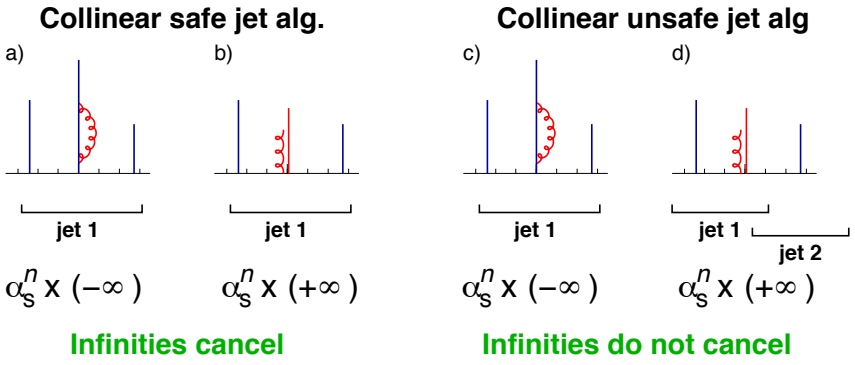


Figure 6.4: An example of the collinear safe (left) and collinear unsafe (right) jet algorithms. The horizontal line represents jet η, while the vertical line jet p_T. Figure (a) and (c) show the stability of the jet algorithms against the virtual loop corrections. Figure (b) shows the stability of the jet algorithms against soft gluon radiation. Figure (d) shows that jet algorithm reconstructs two jets instead of one in case of additional gluon radiation, increasing jet multiplicity of the event.

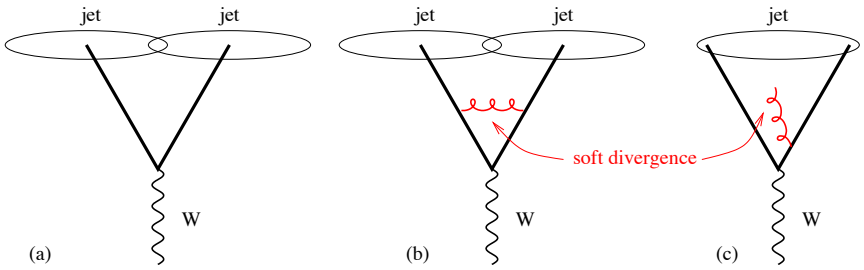


Figure 6.5: An example of the IR unsafety in the event with W boson decaying to two hard partons [163]. Figure (a) shows the jet algorithm finding two separate jets. Figure (b) shows the stability of the jet algorithm against the virtual gluon exchange. Figure (c) shows that the jet algorithm takes an emitted gluon as additional seed and reconstructs a jet around the gluon, containing both hard partons from the W decay. This additional jet overlaps with two other jets from the hard partons, so that only one jet survives at the end, leading to reduced jet multiplicity of the event.

The majority of ATLAS standard jet algorithms, e.g from the k_T-class, are provided by FastJet [164]. Both $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ searches use small$-R$ and large$-R$ jets reconstructed using the anti-k_T clustering algorithm [165] with different values of the radius parameter R.

Algorithms from the k_T-class are *sequential recombination algorithms*, which use topo-clusters as inputs to construct a jet, so that the final momentum of the jet is the sum of the momenta of all topo-clusters associated with the jet. In order to find out if topo-cluster belongs to the jet or not, two distance measures, invariant under longitudinal boosts, are introduced as follows [166]:

$$\begin{aligned} d_{ij} &= \min\left(k_{T,i}^{2p}, k_{T,j}^{2p}\right) \frac{\Delta_{ij}}{R^2} \\ d_{iB} &= k_{T,i}^{2p}, \end{aligned} \tag{6.1}$$

where d_{ij} is the distance between the proto-jet i and topo-cluster j, and d_{iB} between the proto-jet i and the beam-line. Here $\Delta_{ij} = \sqrt{(y_i - y_j)^2 + (\phi_i - \phi_j)^2}$ denotes the angular distance between the proto-jet i and topo-cluster j. The variable R defines the radius parameter of the jet, which is $R = 0.4$ for small$-R$ jets and $R = 1.0$ for large$-R$ jets. The parameter p defines the ordering of inputs in jet reconstruction. The k_T algorithm [167] corresponds to $p = 1$, favouring clusterings of objects with low p_T first. The Cambridge/Aachen (C/A) algorithm [168] corresponds to $p = 0$, providing energy-independent angular clustering. Finally, the anti-k_T algorithm corresponds to $p = -1$, favouring clusterings of objects with large p_T first. All three algorithms are collinear and IR safe.

If $d_{ij} < d_{iB}$, then the objects i and j are combined, and both objects are removed from the pool of available inputs. If $d_{ij} > d_{iB}$ i, object i is declared as a final jet candidate, and is removered from the list of available inputs. Then d_{ij} and d_{iB} are recomputed given the new objects, and procedure repears until no particles remain. At ATLAS, jet candidates with $p_T > 7$ GeV are stored for further analysis.

Since anti-k_T algorithm starts clustering from the hardest object, the direction of the jet axis is mostly driven by the position of the hardest topo-cluster. Thus, the isolated jets coming from the anti-k_T algorithm have circular area shapes defined by the radius parameter R.

Jet trimming

Due to the large radius parameter R, large$-R$ jets are particularly susceptible to the contamination from pile-up, multiple parton interactions and initial-state radiation. All these contributions are generally much softer than the partons inside the large$-R$ coming from the hard scattering or final state radiation. The *trimming algorithm* [169] takes advantage of this fact to remove contamination from underlying processes. The trimming procedure starts from reconstruction of sub-jets of radius R_{sub} from the constituents of the large-R jet using the k_T-algorithm. The sub-jet is removed from the large$-R$ jet, if $p_T^{\text{sub-jet}}/p_T^{\text{large-}R} < f_{\text{cut}}$, where $p_T^{\text{sub-jet}}$ is the sub-jet transverse momentum, and f_{cut} is the threshold parameter defined by the algorithm. The full procedure is illustrated in Figure 6.6. In the scope of this work, $R_{\text{sub}} = 0.2$ and $f_{\text{cut}} = 0.05$.

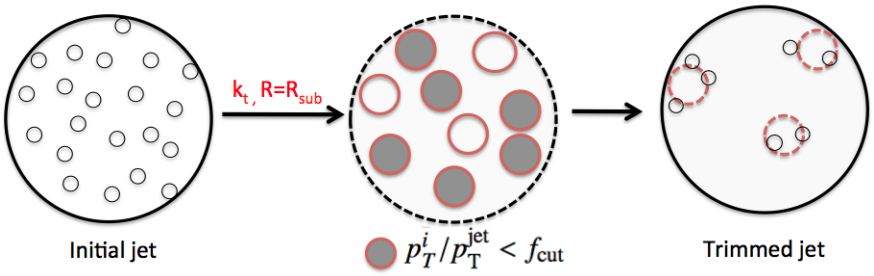


Figure 6.6: Illustration of the trimming algorithm for large$-R$ jets.

Jets in $E_{\text{T}}^{\text{miss}} + V(qq)$ and $E_{\text{T}}^{\text{miss}} + h(b\bar{b})$ searches

Small$-R$ jets utilised in both analyses are reconstructed in the region of $|\eta| < 4.5$. *Central* small$-R$ jets in the region of $|\eta| < 2.5$ are used to reconstruct W, Z and h candidates. Small$-R$ jets in the region of $2.5 < |\eta| < 4.5$ are referred to as *forward* jets. Large$-R$ jets in the region of $|\eta| < 2.0$ are used in both analyses.

6.2.3 Jet energy scale calibration and uncertainty

The reconstructed energy of the jets is determined by the energy measured in the calorimeter, which not necessarily corresponds to the true jet energy. The calorimeter provides correct energy measurements of the EM showers, but at the same time cannot account for the energy losses in the hadronic showers. Primarily these losses come from particles escaping the hadronic showers, but as well from poorly instrumented parts of the detector, such as transition regions between different parts of the calorimeter, energy leakage from the calorimeters, jet constituents falling outside of the reconstructed jet area or not passing the noise thresholds. Thus, the jet energy measured at the EM scale should be calibrated to the hadronic scale, which is provided by so-called *jet energy scale* (JES) calibration [170, 171]. It is applied to the four-momentum of small$-R$ jets, affecting jet energy, transverse momentum and mass.
The reconstructed jet energy is corrected to the energy of *truth* jets, which are reconstructed using the anti-k_T algorithm with $R = 0.4$ radius parameter using stable, final

state particles from Monte Carlo (MC) samples (particle-level jets [1]). Full JES calibration sequence consists of multiple calibration stages, as illustrated in Figure 6.7.

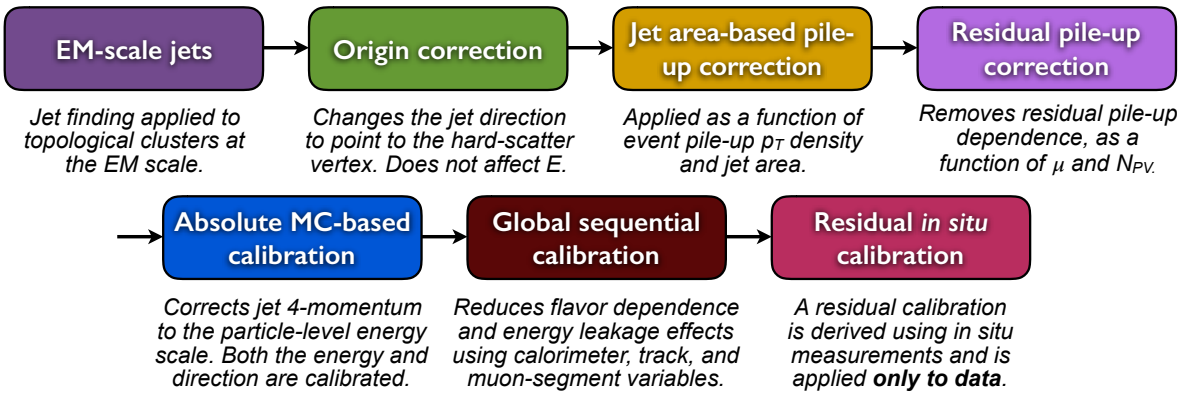
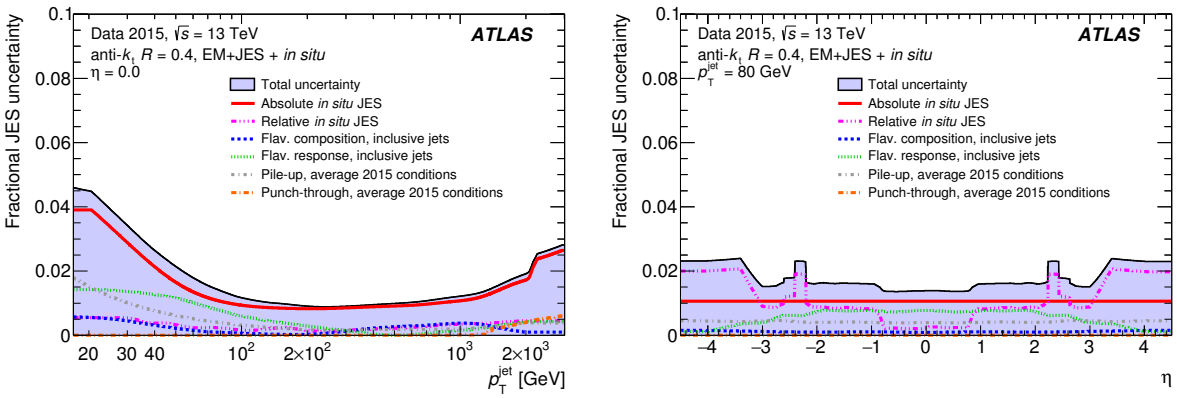


Figure 6.7: An overview of the ATLAS jet energy scale calibration chain, showing the steps needed to bring jets from the EM scale to the final hadronic scale [171]. All calibrations except the origin correction are applied to the four-momentum of the reconstructed jet.

In order to derive JES calibration, the reconstructed jets are matched to the isolated particle-level jets within $\Delta R < 0.3$. The isolation is defined such as no other calorimeter jet with $p_T > 7$ GeV is present within $\Delta R = 0.6$, and only one particle-level jet with $p_T^{\text{truth}} > 7$ GeV is allowed within $\Delta R = 1.0$.
The jet energy response is introduced as $\mathcal{R} = \frac{E^{\text{reco}}}{E^{\text{truth}}}$, where E^{reco}, E^{truth} are energy of the reconstructed and particle-level jets, respectively. An average jet energy response $\langle\mathcal{R}\rangle$, defined as a mean of the Gaussian distributed energy ratio $\frac{E^{\text{reco}}}{E^{\text{truth}}}$, is calculated per $(E^{\text{truth}}, \eta_{det})$ bin, where $|\eta_{det}|$ is the jet η relative to the detector, numerically inverted and applied to the reconstructed jet four-momentum in the corresponding $(E^{\text{reco}}, \eta_{det})$ bin.
The final set of JES calibration uncertanties consists of about 90 uncertainty terms from different calibration steps, shown in Figure 6.7. The majority of uncertainties come from the residual *in situ* calibration stage. The values of different JES uncertainties are summarised in Figure 6.8. The largest uncertainty is observed in low and high p_T regions ($\sim 4.5\%$), while the uncertainty minimises in $p_T \sim 200$ GeV $-$ 2 TeV region. The vivid feature in the $2.0 < |\eta| < 2.6$ region comes from the non-closure in the $\eta-$ calibration between jets extending in the central and forward region.
The majority of the physics analyses are not sensitive to all JES uncertainties, and one only introduces an unnecessary extra level of complexity when all of them are considered. Thus, a reduced set of four JES nuisance parameters is introduced in both $E_{\text{T}}^{\text{miss}} + V(qq)$ and $E_{\text{T}}^{\text{miss}} + h(b\bar{b})$ analyses, representing low-, medium- and high-p_T kinematic regimes, as well as the non-closure uncertanty on η intercalibration, as shown in Figure 6.8. All single components of each nuisance parameter are combined quadratically.

[1] The particle level represents the simulation stage, at which stable particles with a life-time $\tau > 10\mu$s and reconstructed jets after the processes of parton showering and hadronisation are considered.



(a) (b)

Figure 6.8: Breakdown of JES systematic uncertainty of fully calibrated jets [171]. Figure (a) shows JES uncertainties as a function of jet p_T for $|\eta| = 0$. Figure (b) shows JES uncertainties as a function of η for $p_T = 80$ GeV.

6.2.4 Jet mass performace

The standard large$-R$ jet mass is defined as *calorimeter-based* mass m^{calo} using energies E_i and momenta $\vec{p}_i$ ($E_i = |\vec{p}_i|$) of the topo-clusters associated with the jet:

$$m^{\text{calo}} = \sqrt{\left(\sum_{i \in J} E_i\right)^2 - \left(\sum_{i \in J} \vec{p}_i\right)^2}. \tag{6.2}$$

The collimation of jet constituents becomes stronger with a larger boost, scaling as $1/p_T$. With a large enough boost, the angular spread of collimated products is comparable with the calorimeter granularity, limiting the performance of m^{calo}. At the same time, the angular resolution of the ID allows to separate tracks inside the jet in the kinematic regime when calorimeter angular resolution fails. To account for better resolution of the ID, a *track-assisted* jet mass is introduced [172]:

$$m^{\text{TA}} = \frac{p_T^{\text{calo}}}{p_T^{\text{track}}} \times m^{\text{track}}, \tag{6.3}$$

where p_T^{calo} is the calorimeter-based momentum of a large$-R$ jet, p_T^{track} and m^{track} are the transverse momentum and invariant mass of the four-momentum sum of the tracks associated with the jet (each track mass is assumed to be a pion mass). Since the neutral particles do not leave tracks in the ID, the $\frac{p_T^{\text{calo}}}{p_T^{\text{track}}}$ ratio corrects for the corresponding neutral contribution. Both masses m^{calo} and m^{TA}, which are already corrected for the JES calibration, are calibrated to the jet particle-level mass, using JMS calibration procedure described in Section 6.2.5.

The performance of the jet mass is characterised by the width of the corresponding response distribution $\mathcal{R}_{\text{m}} = m^{\text{reco}}/m^{\text{truth}}$, which is defined as a ratio of the half of the 68% interquantile range (IQnR) over the median of the distribution. The corresponding IQnR range is defined as $q_{84\%} - q_{16\%}$, where $q_{84\%}$ and $q_{16\%}$ are 16th and 84th percentiles of the response distribution. Figure 6.9 illustrates the performace of the calorimeter-based and track-assisted jet masses, obtained using multijet MC samples. The track-

assisted mass outperforms the calorimeter-based mass in the $p_T \gtrsim 1$ TeV region, where the calorimeter granularity becomes a limiting factor. Meantime, the calorimeter-based mass outperforms the track-assisted mass in the $p_T \lesssim 1$ TeV region due to the large charged-to-neutral fluctuations coming from the $\frac{p_T^{\text{calo}}}{p_T^{\text{track}}}$ term.
Given the small level of correlation [2] between m^{calo} and m^{TA}, one can make use of best performance of both mass definitions in order to reduce the response resolution over the entire p_T range by introducing the *combined* jet mass as linear combination of m^{calo} and m^{TA} [173]:

$$m^{comb} = \left(\frac{(\sigma^{\text{calo}})^{-2}}{(\sigma^{\text{calo}})^{-2} + (\sigma^{\text{TA}})^{-2}}\right) \times m^{\text{calo}} + \left(\frac{(\sigma^{\text{TA}})^{-2}}{(\sigma^{\text{calo}})^{-2} + (\sigma^{\text{TA}})^{-2}}\right) \times m^{\text{TA}}, \tag{6.4}$$

where σ^{calo}, σ^{TA} denote resolutions of the response distributions of m^{calo} and m^{TA}, respectively. The combined mass resolution, based on simulated $W'/Z' \to WZ \to qqqq$ samples, outperforms both calorimeter-based and track-assisted mass resolutions over the entire p_T range, as shown in Figure 6.9.

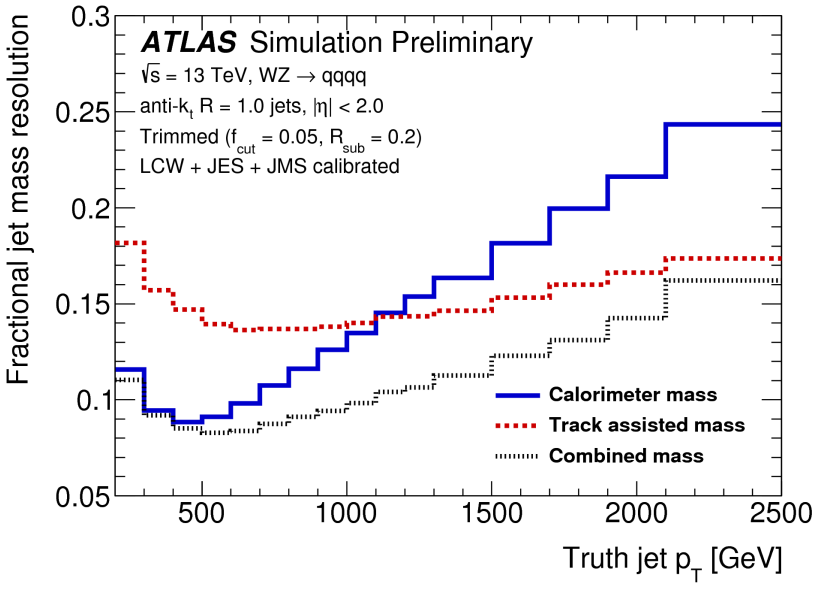


Figure 6.9: Comparison of resolutions of the combined (solid blue), track-assisted (dashed red) and calorimeter-based (dotted black) large$-R$ jet masses binned in truth jet p_T. Results are obtained using $WZ \to qqqq$ MC samples [172]. Large$-R$ jets are reconstructed using anti-k_T algorithm with $R = 1.0$ and trimmed using $f_{\text{cut}} = 0.05$, $R_{\text{sub}} = 0.2$.

6.2.5 Jet mass scale calibration and uncertainty

Similarly to the JES energy calibration of small$-R$ jets described in Section 6.2.3, the large$-R$ jet mass is calibrated using the particle-level large$-R$ jet mass as a reference, using jet mass scale (JMS) calibration procedure [172]. The particle-level large$-R$ jets are reconstructed using anti-k_T algorithm with $R = 1.0$ radius parameter from stable, final state particles using multijet MC samples.

[2] The correlation between m^{calo} and m^{TA} is estimated to be 0.22 and 0.1 for jets coming from the W/Z decay in $|\eta| < 2.0$ region for $p_T > 250$ GeV and $p_T > 1$ TeV, respectively [172]. Results are obtained using $W'/Z' \to WZ \to qqqq$ MC samples.

The JES calibration is applied to the isolated reconstructed jets matched to the isolated particle-level jets within $\Delta R < 0.6$, where isolation is defined such as no other particle-level (reconstructed) large$-R$ jet with $p_T > 100$ GeV is present within $\Delta R = 2.5(1.5)$. The jet mass response variable is defined as $\mathcal{R}_{\mathrm{m}} = \frac{m^{\mathrm{reco}}}{m^{\mathrm{truth}}}$, where the reconstructed jet mass is already calibrated using JES correction factor c^{JES}, $m^{\mathrm{reco}} \in m^{\mathrm{calo,TA}} \cdot c^{\mathrm{JES}}$. An average jet mass response $\langle \mathcal{R}_{\mathrm{m}} \rangle$ is extracted from each $(p_T^{\mathrm{truth}}, |\eta_{det}|, m^{\mathrm{truth}})$ bin, numerically inverted, and applied to the reconstructed jet mass in a corresponding $(p_T^{\mathrm{reco}}, |\eta_{det}|, m^{\mathrm{reco}})$ bin. If the calibration is correctly performed, then $\langle \mathcal{R}_{\mathrm{m}} \rangle \approx 1$.
The corresponding JMS uncertainties are calculated differently for the calorimeter-based and track-assisted masses. The calorimeter-based JMS uncertainty is estimated from data and MC using calorimeter-to-track mass ratio, $r^{\mathrm{m}}_{\mathrm{track}} = m^{\mathrm{calo}}/m^{\mathrm{track}}$, in high-$p_T$ QCD multijet events [174]. The $r^{\mathrm{m}}_{\mathrm{track}}$ can be approximately decomposed as $r^{\mathrm{m}}_{\mathrm{track}} \approx \mathcal{R}_{\mathrm{m}} \times (m^{\mathrm{truth}}/m^{\mathrm{charged\ truth}}) \times (m^{\mathrm{charged\ truth}}/m_{\mathrm{track}})$. If all terms in decomposition are independent, then $\langle r^{\mathrm{m}}_{\mathrm{track}} \rangle \propto \langle \mathcal{R}_{\mathrm{m}} \rangle$, and the calorimeter-based JMS uncertainty can be estimated as $1 - \langle r^{\mathrm{m}}_{\mathrm{track}} \rangle_{\mathrm{data}} / \langle r^{\mathrm{m}}_{\mathrm{track}} \rangle_{\mathrm{MC}}$ using data and Monte Carlo samples.
The track-assisted JMS uncertainty can be estimated by propagating the track reconstruction (primarily from track reconstruction inefficiency and fake track reconstruction) and calorimeter jet p_T uncertainties in the definition of the track-assisted jet mass in Eq. 6.3. The calorimeter jet p_T uncertainties are estimated similarly to the JMS uncertainty for the calorimeter-based jet mass using $r^{p_T}_{\mathrm{track}} = p_T^{\mathrm{calo}}/p_T^{\mathrm{track}}$. A summary of the composition of the calorimeter-based and track-assisted JMS uncertainties is shown in Figure 6.10.

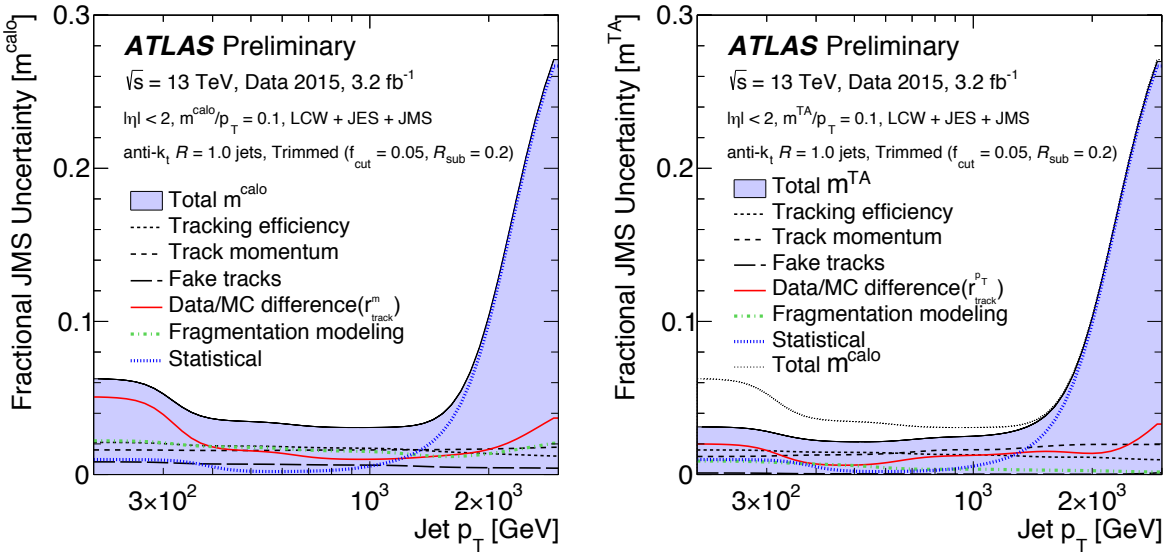


Figure 6.10: Breakdown of JMS systematic uncertainty for the calorimeter-based (left) and track-assisted (right) jet mass as a function of the reconstructed jet p_T in $|\eta| < 2.0$ region and $m^{\mathrm{reco}}/p_T^{\mathrm{reco}} = 0.1$ slice.

6.2.6 Track jets

Track jets [175, 176] are reconstructed from ID tracks with $p_T > 0.4$ GeV in $|\eta| < 2.5$ region using the anti-k_T algorithm with a radius parameter $R = 0.2$. Each track is required to contain at least 6 hits in the silicon strip detector and at least one hit in the pixel detector, at most one of which is shared by different tracks. To significantly

reduce the contribution from the pile-up vertices, tracks are tightly matched to the PV by requiring $|z_0 \sin\theta| < 3$ mm. At the end, track jets with $p_T > 10$ GeV in $|\eta| < 2.5$ region containing at least two tracks are considered in both $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analyses. No JES calibration is applied to track jets.
Ghost association [177, 178] is used to associate track jets with a given large$-R$ jet. The main idea is to add additional *infinitely soft ghost* particles with known (η, ϕ) coordinates to the existing particles in the event and re-cluster particles together with the new ghost particles in jets using the same reconstruction algorithm. Since jet algorithm is IR and collinear safe, the additional soft particles do not change the total number of jets, but rather are clustered in one of the jets. One can consider track jet with infinitesimal small p_T as such a ghost particle and perform re-clustering using anti-k_T algorithm with $R = 1.0$. If after reclustering one of the large$-R$ jets contains the ghost track from the track jet, then the track jet is considered ghost-assosicated with the corresponding large$-R$ jet.

6.2.7 Identification of $b-$jets

Since both $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analyses expect jets coming from $b-$quark ($b-jets$) in the final state, the identification of b-jets together with background suppression of mostly light-flavour jets ($u-$, $d-$, $s-$ and gluon jets) becomes of great importance. The identification of b-jets is usually referred to as $b-tagging$ [179, 180].
The $b-$tagging algorithms exploit such properties of $B-$hadrons inside b-jets as long lifetime, high mass and decay multiplicity. ATLAS standard combined $b-$tagging algorithm is MV2 [180], based on a *Boosted Decision Tree* (BDT) implemented in the TMVA package [181]. It collects information from various algorithms based on the properties of tracks at the primary vertex (PV), reconstruction of displaced vertices and high-quality muon tracks from semi-leptonic $b-$ and $c-$hadron decays [180]:

- **IP3D algorithm:** Impact parameter based algorithm. Exploits the long lifetime of $B-$hadrons, $c\tau \sim 450\mu m$, which leads to at least one vertex being displaced from the PV in the $B-$hadron decay topology. The displacement is characterised by the transverse and longitudinal impact parameters, d_0 and z_0, respectively, with the corresponding uncertainties σ_{d_0} and σ_{z_0} [3]. The log-likelihood ratio (LLR) discriminant is calculated based on probability density functions obtained from a two-dimensional template of the transverse and longitudinal impact parameter significances, $S_{d_0} = d_0/\sigma_{d_0}$ and $S_{z_0} = z_0 \sin\theta/\sigma_{z_0 \sin\theta}$, respectively. The LLR discriminant is calculated for the $b-$, $c-$ and light-flavour jet hypotheses to distinguish between different jet flavours.

- **SV1 algorithm:** Secondary vertex finding algorithm. Explicitely reconstructs a displaced vertex inside the jet, which corresponds to a $B-$hadron decay. Vertex candidates are combined from pairs of displaced tracks with $S_{3D} = d_{3D}/\sigma_{d_{3D}} > 2$, where d_{3D} is the three dimensional distance between the PV and the point of closest approach of the track to the vertex, using vertex fit with $\chi^2 < 4.5$. Two-track vertices are removed, if they come from decays of the long-lived particles, such as K_s and Λ, photon conversions or hadronic interactions with the detector material.

[3] The transverse impact parameter d_0 is the distance of the closest approach of the track to the beam-line. The longitudinal impact parameter z_0 is the distance along the beam-line between the PV and the point of the closest approach of the track to the beam-line.

All remaining tracks from two-track vertices are combined into a single vertex inside the jet using a Kalman-based χ^2 fit [182]. The vertex variables, such as the vertex mass (the invariant mass of all associated tracks, assuming that all tracks are pions), energy fraction of the vertex-associated tracks inside the jet, $\sum_{\text{tracks}} p_T / \sum_{\text{all tracks}} p_T$, and the number of two-track vertices are combined to obtain the LLR discriminant for the $b-$, $c-$ and light-flavour jet hypotheses.

- **JetFitter algorithm:** Exploits the topology of $b-$ and $c-$hadron decays inside the jet [183]. A Kalman-based fit associate the PV, $b-$ and $c-$vertices with a single line, approximating the $B-$hadron line of flight. The algorithm uses six input variables: the number of vertices with at least two associated tracks, the total number of corresponding tracks, the number of one-track vertices on the $B-$hadron line of flight, the vertex mass (the invariant mass of all tracks of the decay chain), the energy fraction of vertex-associated particles inside the jet, $\sum_{\text{tracks}\in\text{vertex}} p_T / \sum_{\text{all tracks in jet}} p_T$, and the flight length significance L/σ_L. These variables are fed into the artificial neural network to produce discriminant for $b-$, $c-$ and light-flavour hypotheses.

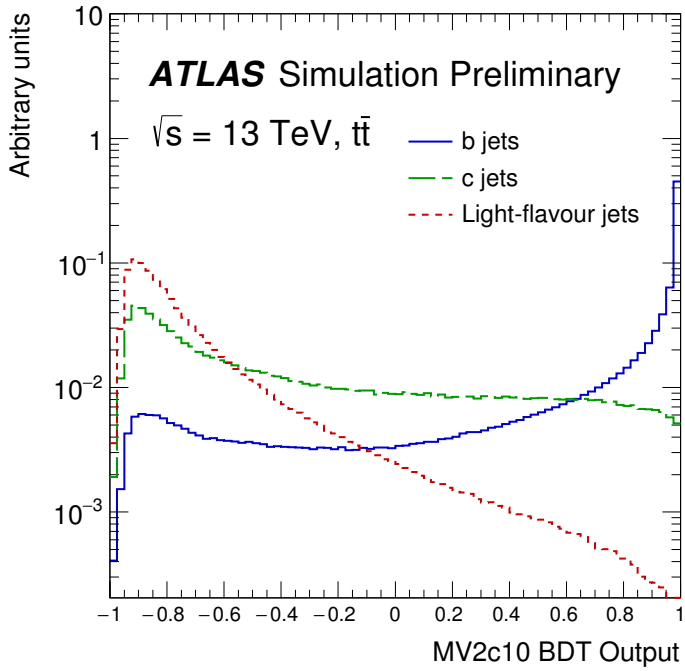


Figure 6.11: MV2c10 BDT output for $b-$ (solid blue), $c-$ (dashed green) and light-flavour (dotted red) based on $t\bar{t}$ events. $b-$jet efficiency corresponds to a working point with a $b-$jet efficiency of 70% [180].

The output of the algorithm is a multivariate discriminant [184], which ranges from 0 to 1, with larger values corresponding to larger probability to originate from a $b-$quark, as illustrated for the case of $t\bar{t}$ events in Figure 6.11. Another important quantity is the $b-$tagging efficiency with which a jet containing b-hadron is tagged, the mistagging rates of $c-$jets and the light-flavour jets as b-jets.
The $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analyses make use of $b-$jets within $|\eta| < 2.5$ tagged by the *MV2c10* algorithm [180]. It is trained based on simulated $t\bar{t}$ events, where signal is composed of b-jets, and background 7% of $c-$jets and 93% of light-flavour jets. A fixed

cut working point of the algorithm corresponds to 70% $b-$tagging efficiency for jet p_T above 20 GeV that is illustracted in Figure 6.11.
Track jets are the key ingredients in identifying the large$-R$ flavour structure coming from the $h \to b\bar{b}$, $W \to cs$ and $Z \to b\bar{b}$ decays. The track jets are $b-$tagged using the same MV2c10 algorithm as for small$-R$ jets. The first two leading ghost-associated track jets in the large$-R$ jet are tested to come from b-jets to associate the large$-R$ jet with a boosted $h \to b\bar{b}$ decay. The $b-$tagged track jets are useful in constraining $t\bar{t}$ background, since they come from separate boosted decays and contained in separate large$-R$ jets. It can be also served as discriminant between $t\bar{t}$ and $W+$jets or as effective veto on $t\bar{t}$ events. The $b-$tagging regions in merged kinematic regimes of both analyses are defined using $b-$tagging for track jets.

6.2.8 Muon-in-jet correction

A significant fraction of $B-$hadrons decays in the jet produce muons, following the $b \to cW \to c\mu\nu_\mu$ or $b \to uW \to c\mu\nu_\mu$ decay chains (W decays to muon with 10% branching ratio). Since muons carry most of their energy away from the calorimeter and thus from the reconstruced jet, jet energy should be corrected for these energy losses coming from the muon-involved decays. This referres to as *muon-in-jet* correction, which corrects the four-momentum of the $b-$tagged jets by adding back the four-momentum contribution of muons coming from the $B-$hadron decays in the corresponding jet.
Reconstructed muons with $p_T > 4$ GeV, which pass the *medium* ID selection (see Section 6.1.2.2), are tested to originate from the $b-$jet. In case of $b-$tagged small$-R$ jets, muons within $\Delta R < 0.4$ of the jet are considered, and the four-momentum of the closest reconstructed muon is added to the four-momentum of the corresponding $b-$tagged jet. In case of large$-R$ jets, the closest reconstructed muons within $\Delta R < 0.2$ of the leading and subleading $b-$tagged ghost-associated track jets are considered, and their four-momenta are added to the four-momentum of the corresponding large$-R$ jet.

6.3 Overlap Removal

The *overlap removal* addresses two main questions of the particle reconstruction: *double counting*, when the object is reconstructed as different physical objects (e.g., electrons reconstructed as jets), and *isolation*, when separately reconstructed objects are close to each other (e.g. muons reconstructed close to jets). The overlap removal is done in the following order [185]:

- **Lepton-lepton overlap removal:** Resolves the muon-electron dublication. If the CB muon radiates the hard photon (final state or bremsstrahlung), it can be reconstructed as an electron, sharing the same ID track. In this case reconstructed electron is removed. In case of CT muons, the electron is kept but the muon is removed.

- **Electron-jet overlap removal:** Removes prompt electrons reconstructed as jets, while keeping heavy-flavour jets with semileptonic decays and light-flavour jets faking a loose electron. The main discriminating variable is $\Delta R(e, j)$ between the jet and electron. All jets in a $\Delta R(e, j) < 0.2$ cone around a well-identified electrons are removed. Electrons in the region $0.2 < \Delta R(e, j) < \min(0.4, 0.04 + 10 \text{ GeV}/p_T^e)$ around each surviving jet are removed to avoid double-counting of energy.

- **Muon-jet overlap removal:** Distinguishes between prompt muons and muons coming from hadron decays inside the jet, and removes jets coming from the collinear final state photons close to the muon tracks and from bremsstrahlung photons. Main discriminating variable is $\Delta R(\mu, j)$ between the jet and muon. All jets in a $\Delta R(\mu, j) < 0.2$ cone around the well-identified muons are removed if at least one of the following conditions is satisfied:
  - < 3 tracks matched to the PV with $p_T > 0.5$ GeV are associated to the jet
  - $p_T^{\mu}/p_T^{j} > 0.5$ and $p_T^{\mu}/\sum_{\text{track}\in\text{jet}} p_T^{\text{track}} > 0.7$, where $\sum_{\text{track}\in\text{jet}} p_T^{\text{track}}$ is the sum of the transverse momenta of tracks with $p_T > 0.5$ GeV associated with the jet and matched to the PV.

6.4 Missing transverse momentum

The missing transverse momentum $E_{\text{T}}^{\text{miss}}$ is an observable, which serves as a measure of the energy that is carried away by undetected particles produced in the proton-proton collisions at the LHC. A non-zero value of $E_{\text{T}}^{\text{miss}}$ may originate not only from SM neutrinos or mismeasurements of the visible objects but also from the production of new particles beyond the Standard Model that escape the ATLAS detector undetected. Since $E_{\text{T}}^{\text{miss}}$ is reconstructed using all detected objects, it requires information from all subsystems of the ATLAS detector. $E_{\text{T}}^{\text{miss}}$ is sensitive to the reconstruction and calibration of objects contrubuting to it, which makes it challenging to perform reconstuction of $E_{\text{T}}^{\text{miss}}$. In this Section the basic ideas behind $E_{\text{T}}^{\text{miss}}$ reconstruction are discussed.

6.4.1 Missing transverse momentum reconstruction

The missing transverse energy consists of two main components [186, 187]. The first component, referred to as *hard term*, comes from the hard scattering and is reconstructed using particles leptons, photons, hadronically decaying $\tau-$leptons, muons and jets originated from the hard partons. The second component, referred to as *soft term*, comes from the the soft activity in the event like the underlying event, and is reconstructed using tracks associated with the PV not coming from the particles contributing to the hard term. All these contributions are reconstructed differently: muons are reconstructed using track information (see Section 6.1.2.1), are reconstructed using both calorimeter and track information, and jets and photons are reconstructed using calorimeter information with track information involved in various calibrations or tagging algorithms. Based on these objects, one can introduce calorimeter-based, track-based or combined $E_{\text{T}}^{\text{miss}}$. The standard $E_{\text{T}}^{\text{miss}}$ definition exploits both calorimeter and track information, while other definition can be useful, e.g. in defining the $E_{\text{T}}^{\text{miss}}$ triggers, as discussed in Section 5.2
The vector missing transverse momentum is calculated as a negative vectorial sum of the transverse momenta of its components:

$$\begin{aligned} \mathbf{E_T^{miss}} &= -\sum_{\text{hard objects}} \mathbf{p_T^i} - \sum_{\text{soft term}} \mathbf{p_T^j} \\ \mathbf{E_T^{miss}} &= (E_{\text{x}}^{\text{miss}}, E_{\text{y}}^{\text{miss}}, 0) \\ E_{\text{T}}^{\text{miss}} &= \sqrt{(E_{\text{x}}^{\text{miss}})^2 + (E_{\text{y}}^{\text{miss}})^2} \end{aligned} \tag{6.5}$$

The missing transverse momentum is given by the magnitude of the $\mathbf{E}_{\mathbf{T}}^{\mathbf{miss}}$ vector. If hard objects used to reconstruct $E_{\mathrm{T}}^{\mathrm{miss}}$ share same calorimeter- or track-based information (see Section 6.3), one needs to consider a rejection algorithm for the overlapping objects. This defines the standard order of the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction sequence. If the lower-priority object share its energy with the higher-priority object, then the lower-priority one is excluded from the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation.

When all hard contributions are considered, ID tracks that are not associated with any of the contributing hard objects but with the PV are used to calculate the *track soft term* (TST) [4] of $E_{\mathrm{T}}^{\mathrm{miss}}$. Finally, $\mathbf{E}_{\mathbf{T}}^{\mathbf{miss}}$ in Eq. 6.5 can be described in terms of individual contributions from all hard objects and ID tracks not associated with them:

$$\mathbf{E}_{\mathbf{T}}^{\mathbf{miss}} = -\Bigg(\underbrace{\sum_{\substack{\text{selected}\\ \text{electrons}}} \mathbf{p}_{\mathbf{T}}^{\mathbf{e}}}_{\mathbf{E}_{\mathbf{T}}^{\mathbf{miss,e}}} + \underbrace{\sum_{\substack{\text{selected}\\ \text{photons}}} \mathbf{p}_{\mathbf{T}}^{\gamma}}_{\mathbf{E}_{\mathbf{T}}^{\mathbf{miss},\gamma}} + \underbrace{\sum_{\substack{\text{selected}\\ \tau_{\mathrm{had}}}} \mathbf{p}_{\mathbf{T}}^{\tau_{\mathrm{had}}}}_{\mathbf{E}_{\mathbf{T}}^{\mathbf{miss},\tau_{\mathrm{had}}}} + \underbrace{\sum_{\substack{\text{selected}\\ \text{jets}}} \mathbf{p}_{\mathbf{T}}^{\mathrm{jet}}}_{\mathbf{E}_{\mathbf{T}}^{\mathbf{miss,jet}}} + \underbrace{\sum_{\substack{\text{selected}\\ \text{muons}}} \mathbf{p}_{\mathbf{T}}^{\mu}}_{\mathbf{E}_{\mathbf{T}}^{\mathbf{miss},\mu}} \Bigg)_{\text{hard term}} - \underbrace{\sum_{\substack{\text{selected}\\ \text{ID tracks}}} \mathbf{p}_{\mathbf{T}}^{\mathrm{track}}}_{\mathbf{E}_{\mathbf{T}}^{\mathbf{miss,TST}}\ \text{soft term}} \tag{6.6}$$

Each object included in the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation is reconstructed, fully calibrated, and selected using set of selection criteria defined in Table 6.1.

To account for a full or partial overlap of the jet with electron or photon, electron(photon)-to-jet energy ratio at the EM scale, $\kappa_E = E_{e,\gamma}^{\mathrm{EM}}/E_{\mathrm{jet}}^{\mathrm{EM}}$, is calculated. If $\kappa_E \leq 0.5$, then theselected jet enters the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction with its p_T scaled by a factor of $1-\kappa_E$. In case of large overlap ($\kappa_E > 0.5$), all tracks associated with the jet and not associated with the overlapping object enter the $E_{\mathrm{T}}^{\mathrm{miss}}$ soft term.

Object	Selection
e	*medium* ID, $\lvert\eta\rvert < 1.37$ OR $1.52 < \lvert\eta\rvert < 2.47$, $p_T > 10$ GeV
γ	*tight* ID, $\lvert\eta\rvert < 1.37$ OR $1.52 < \lvert\eta\rvert < 2.47$, $p_T > 25$ GeV
τ_{had}	*medium* ID, $\lvert\eta\rvert < 1.37$ OR $1.52 < \lvert\eta\rvert < 2.47$, $p_T > 20$ GeV
jets	$\lvert\eta\rvert < 2.4$, 20 GeV $< p_T >$ 60 GeV, JVT > 0.59 [5] OR $2.4 < \lvert\eta\rvert < 4.5$, 20 GeV $< p_T >$ 60 GeV OR $\lvert\eta\rvert < 4.5$, $p_T >$ 60 GeV
μ	*medium* ID, $\lvert\eta\rvert < 2.7$, $p_T > 10$ GeV
ID tracks	$p_T > 0.4$ GeV, $\lvert d_0\rvert < 1.5$ mm $\lvert z_0 \sin\theta\rvert < 1.5$ mm $\Delta R(\text{track}, e/\gamma) > 0.05$, $\Delta R(\text{track}, \tau_{\mathrm{had}}) > 0.2$ no muon ID tracks, no ID tracks ghost-associated with jets contributing to $E_{\mathrm{T}}^{\mathrm{miss}}$

Table 6.1: Summary of objects contributing to the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction and the corresponding selections. Objects in the table are listed according to their priority in $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction sequence. Impact parameters d_0 and z_0 are measured relative to the PV.

[4] TST is a standard algorithm used to reconstruct $E_{\mathrm{T}}^{\mathrm{miss}}$ soft term. Other soft-term definitions, such as the calorimeter soft term, the soft-term vertex-fraction and extrapolated jet area with filter, are discussed in [186].

[5] Jet vertex tagging (JVT) is an algorithm used to identify jets coming from the hard scattering [188]. The JVT discriminant ranges from 0 (likely to come from pile-up) to 1 (like to come from hard-scattering).

6.5 Overview of object reconstruction in $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analyses

A summary of all objects and selections used in $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analyses is given in Table 6.2.

Object	Kinematic selection	Type, Quality
Small$-R$ jets		
cenral	$\lvert\eta\rvert < 2.5$, $p_T > 20$ GeV	
forward	$2.5 \leq \lvert\eta\rvert < 4.5$, $p_T > 30$ GeV	anti-k_T $R = 0.4$, EMTopo
$b-$tagged	MV2c10 > 0.8244273	
Large$-R$ jets	$\lvert\eta\rvert < 2.0$, $p_T > 200$ GeV	anti-k_T $R = 1.0$, LCTopo trimmed ($R_{\mathrm{sub}} = 0.2$, $f_{\mathrm{cut}} = 0.05$)
Track jets	$\lvert\eta\rvert < 2.5$, $p_T > 10$ GeV	anti-k_T $R = 0.2$, $N_{\mathrm{track}} \geq 2$
$E_{\mathrm{T}}^{\mathrm{miss}}$	resolved: $E_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV in both analyses merged: $E_{\mathrm{T}}^{\mathrm{miss}} > 250(500)$ GeV for $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ ($E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$) analysis	MET_TST
Electrons:	$\lvert\eta\rvert < 2.47$	$d_0/\sigma_{d_0} < 5$, $\lvert z_0 \sin\theta\rvert < 0.5$ mm
$V-$loose	$p_T > 7$ GeV	*loose*, LooseTrackOnly
$Z-$signal	$p_T > 25$ GeV	*loose*, LooseTrackOnly
$W-$signal	$p_T > 25$ GeV	*medium*, FixedCutTightTrackOnly
Muons:		$d_0/\sigma_{d_0} < 3$, $\lvert z_0 \sin\theta\rvert < 0.5$ mm
$V-$loose	$\lvert\eta\rvert < 2.7$, $p_T > 7$ GeV	*loose*, LooseTrackOnly
$Z-$signal	$\lvert\eta\rvert < 2.5$, $p_T > 25$ GeV	*loose*, LooseTrackOnly
$W-$signal	$\lvert\eta\rvert < 2.5$, $p_T > 25$ GeV	*medium*, FixedCutHighPtCaloOnly
$\tau-$leptons	$\lvert\eta\rvert < 2.5$, $p_T > 20$ GeV	
Standard $\tau-$leptons [6] Extended $\tau-$leptons [7] ($E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ only)		*loose* $1 \leq N_{\mathrm{tracks} \leq 4}$, $\Delta\phi(\mathbf{E_T^{miss}}, \mathbf{p_T^{\tau}}) \leq \pi/8$

Table 6.2: Summary of recosntructed objects used in $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analyses. The second column shows kinematic selections applied. Third column shows details about reconstruction, ID and isolation algorithms and working points.

[6] Standard $\tau-$leptons are required to have either one or three tracks associated with it. $\tau-$leptons in $1.37 < \lvert\eta\rvert < 1.52$ are not considered.

[7] Extended $\tau-$leptons are also required to meet the same conditions as small$-R$ jets in $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis.

Chapter 7

Event selection

The first measurements of the $E_{\mathrm{T}}^{\mathrm{miss}}+V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ final states were peformed by the ATLAS collaboration at the center-of-mass energy of $\sqrt{s} = 8$ TeV using $\mathcal{L} = 20.3$ fb^{-1} of 2012 data [38, 39, 44]. Despite absence of any evidence for the DM production, both searches were able to set stringent limits on the parameters of the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ effective field theory models and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ Z'-2HDM models, as shown in Figure 7.1. An improvement in the performance of the ATLAS detector and increase in the center-of-mass energy for the LHC Run 2 significantly enhance the sensitivity of both searches. Besides the accelerator and detector updates, a solid improvement from the analyses side is also performed. It includes a selection of the boosted objects in the hard part of the kinematic region of the analysis, improvement in systematic and statistical uncertainties, better backgrounds constraints, better calibration of the relevant objects. All these improvements significantly increase the sensitivity of both searches.
Since both analyses share same final state, the major part of the selection criteria are identical. Difference arise mostly from the different jet flavour composition and different final state Standard Model (SM) bosons decaying into jets. This chapter is aimed to discuss both similarities and differences between the analyses in terms of a signal event selection, topology and analysis regions. The general aspects of event cathegorization and kinematic regimes are discussed in Section 7.1. Basic event selecion is introduced in Sections 7.2 and 7.3. The event selections in different kinematic regimes are given in Sections 7.6 and 7.7. The $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ optimisation selection is discussed separately in Section 7.8. Details about main discriminants used in the statistical model are discussed in Section 7.9. And finally, a concised summary of all analysis regions and corresponding event selections for both analyses is given in Section 7.10.

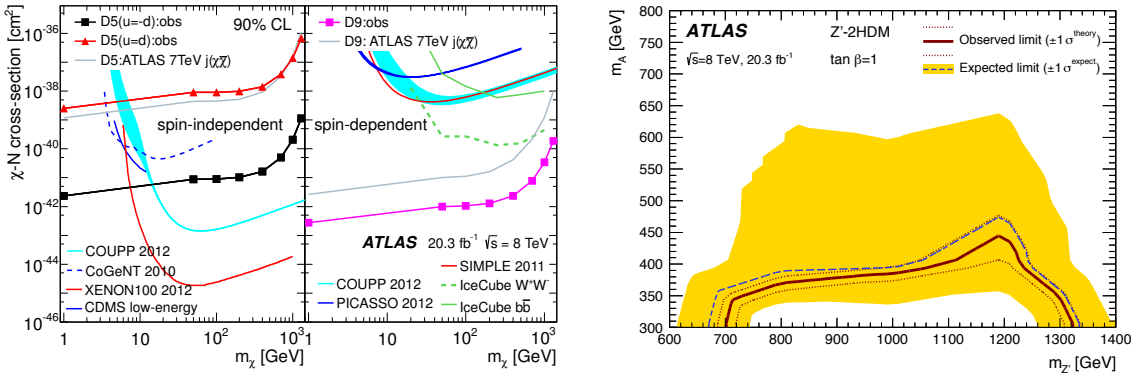


Figure 7.1: Limits placed on the DM production of signal models for $E_{\mathrm{T}}^{\mathrm{miss}}+V(qq)$ (left) and $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ (right) searches based on 2012 data at the center-of-mass energy of $\sqrt{s}=8$ TeV. Left plot shows limits on the DM-nucleon cross-section as a function of the DM mass m_{χ} at 90% CL for spin-independent (left) and spin-dependent (right) effective field theory operators. Right plot shows Z'-2HDM exclusion countour in the $m_{Z'}-m_A$ plane for $\tan\beta=1$.

7.1 Event cathegorization and analysis regions

One of the most important tasks of every analysis is to split the phase space in the analysis regions, which are sensitive to different signal and background processes. Both $E_{\mathrm{T}}^{\mathrm{miss}}+V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ analyses use $E_{\mathrm{T}}^{\mathrm{miss}}$, a number of leptons and $b-$tags in the event as main discriminants to define the main categories of the analysis regions.
The region with no leptons in the event, referred to as *0 lepton* region, corresponds to the signal final state and thus is aimed to study various signal models of the DM production and constrain the irreducible backgrounds coming from the invisible decays, such as $Z(\nu\nu)+$jets. The region with exactly 1 lepton (muon) in the event, referred to as *1 lepton* region, corresponds to the signal-like final state with a real $E_{\mathrm{T}}^{\mathrm{miss}}$, coming from the W leptonic decays. It is aimed to constrain the dominant background processes, such as $W+$jets, $t\bar{t}$ and single top. The region with exactly 2 leptons in the event, referred to as *2 lepton* region, corresponds to the fully leptonic final state with no real $E_{\mathrm{T}}^{\mathrm{miss}}$. It is aimed to study background processes, such as $Z(ll)+$jets, which can be used to constrain $Z(\nu\nu)+$jets background in the 0 lepton region. The 1 and 2 lepton regions are referred to as *control regions*. A concise summary of the lepton regions is given in Table 7.1.

Region	Main goals (both searches)
0 lepton (signal region)	Signal models of DM production. Constraint $Z(\nu\nu)+$jets irreducible background.
1 lepton (control region)	Constrain $W+$jets and $t\bar{t}$ backgrounds.
2 lepton (control region)	Constrain $Z(\nu\nu)+$jets background using $Z(ll)+$jets background.

Table 7.1: Summary of the regions with different number of leptons in the event.

.

The $E_{\mathrm{T}}^{\mathrm{miss}}$ defines the boost of the $h\rightarrow b\bar{b}$ system and therefore serves as the figure of merit to distinguish between the non-boosted and boosted kinematic regimes. In the non-boosted, *resolved*, regime the hadronic decay products of the $h/W/Z$ bosons are reconstructed as well-separated small$-R$ jets. In the boosted, *merged*, regime the decay

products are too collimated to be identified as separate jets, and thus are reconstructed as a single large$-R$ jet. Both definitions of small$-R$ and large$-R$ jets are given in Section 6.2.1. The resolved region is defined as having $E_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV (150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} < 500$ GeV) in the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ ($E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$) analysis. The merged region is defined as having $E_{\mathrm{T}}^{\mathrm{miss}} > 250(500)$ GeV in the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ ($E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$) analysis. Both resolved and merged regimes in the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis share the same part of the phase-space, $E_{\mathrm{T}}^{\mathrm{miss}} > 250$ GeV. Since the most of the sensitivity of the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis comes from the boosted regime, the priority-merged selection is applied, meaning that the event is considered in the merged regime if contains at least one large$-R$ jet, regardless of the presence of absence of the small$-R$ jets in the event.
The regions with different number of $b-$tagged jets ($b-$tags) in the event help to recover sensitivities to the signal processes with different jet flavour content. The summary of the $b-$tag regions is given in Table 7.2.

Region	Main goals	
	$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ search	$E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ search
0 $b-$tag	W/Z bosons decay to light-flavour quarks.	Not used.
1 $b-$tag	W/Z bosons decay via $c-$quark, which is mistagged as $b-jet$.	One of the real $b-$jets is mistagged or track jets inside the large$-R$ jet are reconstructed as a single $b-$tagged track jet due to the very large boost ($\gtrsim 1$ TeV).
2 $b-$tags	$Z \to bb$ decay with both jets successfully tagged as $b-$jets.	$h \to bb$ decay with both jets successfully tagged as $b-$jets.

Table 7.2: Summary of the regions with different number of $b-$tagged jets in the event.

.

Finally, the phase space of both analyses is sub-divided into $E_{\mathrm{T}}^{\mathrm{miss}}$ regions (bins). This is done to account for the different $E_{\mathrm{T}}^{\mathrm{miss}}$ spectra of the signal models of the DM production with the different model parameters. The events in each analysis region are cathegorized according to the following $E_{\mathrm{T}}^{\mathrm{miss}}$ bins:

- $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$: $[150, 200)$ GeV, $[200, 250)$ GeV, $[250, 300)$ GeV, $[350, 400)$ GeV, $[400, 450)$ GeV, $[450, 500)$ GeV, $[500, 600)$ GeV, $[600, 800)$ GeV, $[800, \infty)$ GeV.
- $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$: $[150, 200)$ GeV, $[200, 350)$ GeV, $[350, 500)$ GeV, $[500, \infty)$ GeV.

7.2 Baseline selection

Events in both analysis are required to satisfy the following set of the quality criteria:

- **Data from good-run lists:** Accept 2015 and 2016 data only from the good-run lists, as discussed in Chapter 5.
- **Vertex reconstruction:** Accept events with at least one reconstructed vertex with at least two associated tracks.
- **Event cleaning:** Veto events with corrupted data coming from the detector sub-systems [189], such as Tile/LAr noise bursts or SCT errors.

- **Jet cleaning:** Veto events, which contain jets with LooseBadJet jet cleaning requirement [190].

- $E_{\mathrm{T}}^{\mathrm{miss}}$ **triggers:** Events are required to pass the following $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger requirements (see Section 5.2):

  - 2015 (3.2 fb^{-1}): HLT_xe70
  - 2016, periods A-D3 (6.1 fb^{-1}): HLT_xe90_mht_L1XE50
  - 2016, periods D4-E3 (3.9 fb^{-1}): HLT_xe100_mht_L1XE50 **OR** HLT_xe110_mht_L1XE50
  - 2016, periods F1– (23.2 fb^{-1}): HLT_xe110_mht_L1XE50

- $E_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV: Accept events with significant $E_{\mathrm{T}}^{\mathrm{miss}}$.

7.3 Anti-QCD selection

The multijet events with large fake $E_{\mathrm{T}}^{\mathrm{miss}}$ comprise a significant fraction of the events in the $E_{\mathrm{T}}^{\mathrm{miss}}$+jets final state. It primarily happens when one of the jets has poorly measured jet energy, which leads to the large energy imbalance and thus to the large $E_{\mathrm{T}}^{\mathrm{miss}}$. To remove such type of events, a set of requirements on the $E_{\mathrm{T}}^{\mathrm{miss}}$, p_T^{miss} and the transverse momenta of jets, referred to as *anti-QCD cuts*, are introduced:

- $E_{\mathrm{T}}^{\mathrm{miss}}(p_T^{\mathrm{miss}}) > 30$ GeV: Applied for the events with less than two $b-$tagged jets. This requirement was designed in Run 1 [191] to reduce non-collision background events, which in general do not have tracks in the inner detector, whilst events with a real $E_{\mathrm{T}}^{\mathrm{miss}}$ do have tracks. It also helps to reject events with poorly measured jet energy, which leads to a significant artificial contribution to the $E_{\mathrm{T}}^{\mathrm{miss}}$. It exploits the fact that the tracks in the inner detector associated with the jet with poorly measured energy still have well-measured p_T from the inner detector. It means that in case of the large calorimeter-based $E_{\mathrm{T}}^{\mathrm{miss}}$, the track-based $E_{\mathrm{T}}^{\mathrm{miss}}$ generally stays low.

- $\min\left[\Delta\phi(\mathbf{E_T^{miss}}, \mathbf{p_T^{j_{1,2,3}}})\right] > \pi/9$: Veto multijet events, in which the $E_{\mathrm{T}}^{\mathrm{miss}}$ is aligned along the p_T of the first, second and third small$-R$ jet according to the jet ranking introduced in Section 7.6. It is aimed to reject events, in which poorly measured jet p_T introduces the large artificial $E_{\mathrm{T}}^{\mathrm{miss}}$. Since the $E_{\mathrm{T}}^{\mathrm{miss}}$ is defined as the negative vectorial sum of its components, the resulting $E_{\mathrm{T}}^{\mathrm{miss}}$ points in the same direction as jet p_T. Thus, the jet p_T and $E_{\mathrm{T}}^{\mathrm{miss}}$ is required to be well-separated. This requirement rejects the majority of the events in the signal region.

- $\Delta\phi(\mathbf{E_T^{miss}}, \mathbf{p_T^{miss}}) < \pi/2$: This requirement has similar origin with the previous requirement. In case of the well-measured jet p_T and non-zero real $E_{\mathrm{T}}^{\mathrm{miss}}$, both $E_{\mathrm{T}}^{\mathrm{miss}}$ and p_T^{miss} point in the same direction. In case of the mis-measured jet p_T, $E_{\mathrm{T}}^{\mathrm{miss}}$ points to the same direction as the corresponding jet, while the track-based p_T^{miss} remains almost unaffected and thus does not have preferred direction.

- $\Delta\phi(\mathbf{E_T^{miss}}, \mathbf{p_T^{h/W/Z}}) > 6\pi/9$: This requirement reflects back-to-back topology of the $E_{\mathrm{T}}^{\mathrm{miss}} + h/W/Z$ final state.

7.4 $W/Z/h$ mass window

The invariant mass of the jet(s) from the particle decay is one of the most powerful discriminants to separate between the jets coming from the light and heavy particles. In case of a given analyses, it helps to separate the low-mass QCD jets originating from the gluon and quark decays (except the top quark decay) from the $h/W/Z$ jets. The requirement on the invariant mass is formulated in terms of the mass window around the mass of the decaying particle. In the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis it is used in the resolved regime and in the 2 $b-$tag merged regime, in which the jets come primarily from the $Z \to b\bar{b}$ decay. In the $0,1$ $b-$tag merged regions the special W/Z tagger is used, as discussed later in Section 7.7. Meantime, since no $h \to b\bar{b}$ tagging is available in the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis, the mass window requirement is used in all analysis regions. The following $W/Z/h$ mass windows are defined in the different signal regions:

- $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$**, resolved, 0,1 b−tags:** $65 \text{ GeV} < m_{jj} < 105 \text{ GeV}$.
- $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$**, resolved, 2 b−tags:** $65 \text{ GeV} < m_{jj} < 100 \text{ GeV}$.
- $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$**, merged, 2 b−tags:** $75 \text{ GeV} < m_{J} < 100 \text{ GeV}$.
- $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$**, all regions:** $70 \text{ GeV} < m_{jj/J} < 140 \text{ GeV}$.

The regions outside the mass windows, referred to as *sideband* (SB) regions, are used as the natural control regions to further constrain the backgrounds. In case of $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis, the lower sideband region, which goes down to 50 GeV, is useful to constrain V+jets backgrounds in the 0 lepton region, as they populate the regions of the analysis phase space outside the mass window.
In case of $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis, the sideband regions have two main purposes. The first one is to constrain V+jets backgrounds in the same way as done in the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis. The second one is to wash out the potential effect of the $(E_{\mathrm{T}}^{\mathrm{miss}} + j)$-like signal excess, which can show up in the signal region together with a given final state. Only the upper sideband region above the mass window is considered to avoid difficulties with simultaneous modelling of the low and high mass regions with significantly different statistics.

7.5 W/Z boson tagger

The main tool to associate the large$-R$ jets with the hadronic W/Z boson decay is the boosted W/Z boson tagger [193, 194]. It is based on the specific radiation pattern inside the large$-R$ jet, which separates the boosted $W \to q\bar{q}$ and $Z \to q\bar{q}$ decays from the gluon- and quark-originated jets. The tagging algorithm extensively uses the jet substructure techniques.
A variety of different variables, such as jet masses, jet moments and parton shower shapes, were studied in Run 1 [160] and Run 2 [193] to establish the following variables with the best tagging power:

- **Large$-R$ jet mass:** The most powerful discriminant. The combined jet mass described in Section 6.2.4 after the jet trimming procedure is used. It exploits a significant difference between masses of the gluon- and quark-originated jets, coming mostly from the soft and collinear emissions, and the W/Z-originated jets, as illustrated on Figure 7.3.

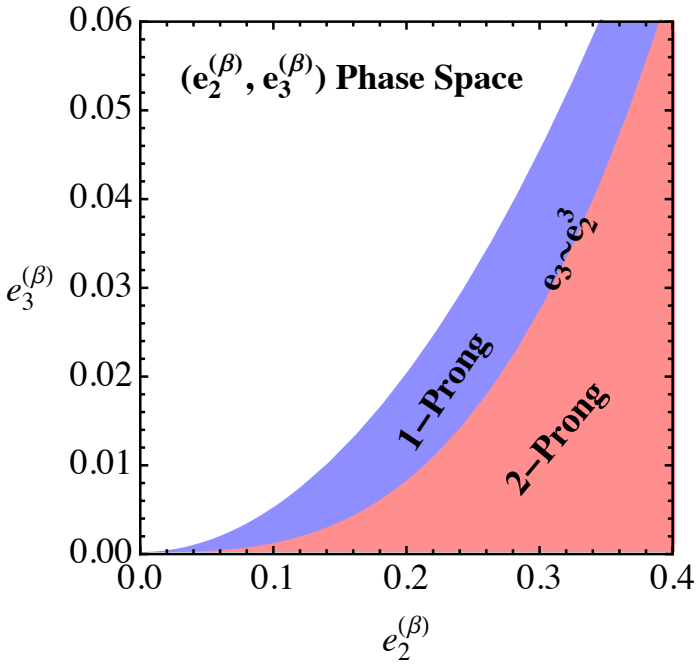


Figure 7.2: The $e_2^{(\beta)} - e_3^{(\beta)}$ phase-space dominated by the 1-prong (blue) and 2-prong (red) decays [192]. The boundary curve approximately corresponds to the $e_3^{(\beta)} \sim (e_2^{(\beta)})^3$ curve.

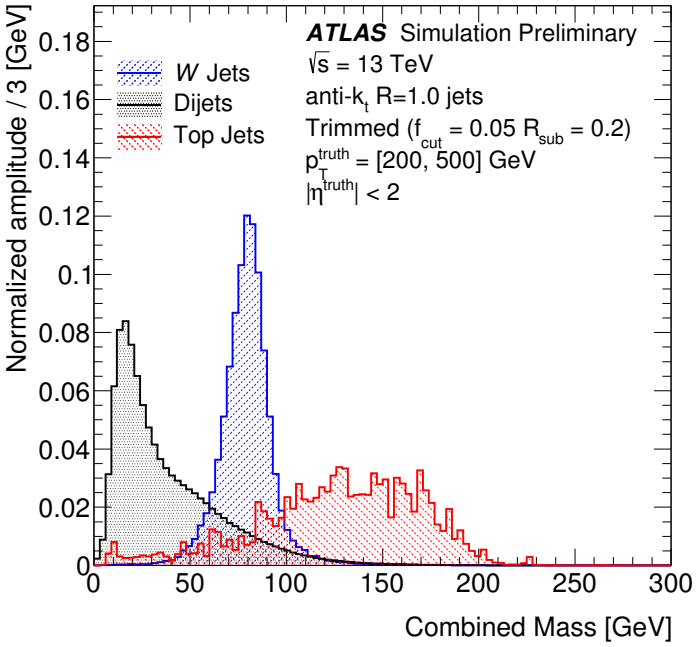


Figure 7.3: The combined jet mass distribution for the jets coming from the QCD processes (black), W decays (blue) and $t-$quark decays (red) for the jet p_T range of $[200, 500]$ GeV [193].

- **Energy correlation ratio** ($D_2^{\beta=1}$): Defined using the 2- and 3-point jet energy

correlation functions, $e_2^{(\beta)}$ and $e_3^{(\beta)}$, respectively [192, 195]:

$$
\begin{aligned}
D_2^{\beta} &\equiv \frac{e_3^{(\beta)}}{(e_2^{(\beta)})^3} \\
e_2^{(\beta)} &= \frac{1}{(p_T^J)^2} \sum_{i<j\in J} p_{T,i} p_{T,j} (R_{ij})^{\beta} \\
e_3^{(\beta)} &= \frac{1}{(p_T^J)^3} \sum_{i<j<k\in J} p_{T,i} p_{T,j} p_{T,k} (R_{ij} R_{jk} R_{ik})^{\beta}.
\end{aligned}
\tag{7.1}
$$

It separates between the 1-prong QCD jets, and 2-prong jets, coming from the $W/Z \to q\bar{q}$ deacys, as illustrated on Figure 7.2. The approximate boundary between the 1- and 2-prong decays, $e_3^{(\beta)} \sim (e_2^{(\beta)})^3$, stays invariant under the Lorentz boost, meaning that $D_2^{\beta=1}$ variable can be successfully used over the entire jet p_T region. The large$-R$ jets associated with the W/Z decays correspond to the low $D_2^{\beta=1}$ values, while the QCD jets correspond to the higher $D_2^{\beta=1}$ values. The upper cut on the $D_2^{\beta=1}$ variable is p_T-dependent, increasing from ~ 1 for $p_T = 200$ GeV to ~ 2 for $p_T = 2500$ GeV [196].

The W/Z tagging algorithm in the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis uses the 50% efficiency working point.

7.6 Resolved regime

In this regime the decay products of the SM $h/W/Z$ bosons are reconstructed as well-separated small-R jets. The main difference consists in the different jet flavour content, meaning that two $b-$tagged jets are expected in the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ final state, while no requirement is put on the flavout content of the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ event.
The small$-R$ jets are ranked according to their angular position in the detector, flavour in case of the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis and p_T. First, jets are divided into cathegories according to jet η coordinate with the higher rank assigned to the central jets with $|\eta| < 2.5$ and the lower rank to the forward jets with $|\eta| \geq 2.5$. In case of $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis, the highest rank is assigned to the central $b-$tagged jets, lower to the central untagged jets and lowest to the forward jets. Within each cathegory jet are ordered in decreasing p_T. The first two jets with the highest ranks, referred to as *leading* and *sub-leading* jets, are associated with the $h \to b\bar{b}$ or $W/Z \to q\bar{q}$ decay. The resulting four-momenta of the $h/W/Z$ candidate is then assigned to the sum of the four-momenta of the corresponding leading and sub-leading jets. The following signal selection criteria are applied to the events in the resolved regime:

- $\geq$ central $2j$: The $h/W/Z$ bosons are produced centrally, and so are the corresponding products of the hadronic decays.

- $p_T^{j_1} > 45$ GeV ($p_T^{j_1}$ **OR** $^{j_2} > 45$ GeV): Since the $h/W/Z$ bosons are required to be relatively boosted with $p_T^{h,W,Z} > 150$ GeV, the hadronic decay products gain a significant boost as well. Thus, the leading jet (or the sub-leading in case of the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis) is required to have $p_T > 45$ GeV.

- $H_T = \sum_{i=1}^{2(3)} p_T^{j_i} > 120(150)$ GeV: Accept events with the scalar sum of the first two (three) leading jet p_T larger than 120(150) GeV. This requirement is aimed to account for the dependence of the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency on the jet activity in the event. In particular, it removes the region of the phase space, in which the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency is poorly modelled in the Monte Carlo simulation [197].

- $\Delta\phi(j_1, j_2) < 7\pi/9$: The hadronic decay products are collimated due to the relatively high boost of the $h/W/Z$ bosons. This requirement reduces the dijet background, in which jets are produced mostly back-to-back.

- $\Delta R(j_1, j_2) < 1.4(1.25)$: Applied only in the $E_{\mathrm{T}}^{\mathrm{miss}}+V(qq)$ selection. $\Delta R(j_1, j_2) < 1.4$ corresponds to the events with 0 $b-$tags, while $\Delta R(j_1, j_2) < 1.25$ to the events with at least one $b-$tag. This requirement reduces multijet and $t\bar{t}$ backgrounds, in which jets are produced rather back-to-back than close to each other.

- **Mass window requirement:** 65 GeV $< m_{jj} <$ 105(100) GeV in 0,1 (2) $b-$tag regions of the $E_{\mathrm{T}}^{\mathrm{miss}}+V(qq)$ analysis (see Section 7.4). No mass window requirement applied in the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ selection.

7.7 Merged regime

In this regime, the decay products of the SM W, Z and h bosons are reconstructed as a single large$-R$ jet. In case of the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis, the $b-$tagging is applied to one (two) highest p_T track jet(s) associated with the leading large$-R$ jet. The 1 $b-$tag region corresponds to the case of strongly collimated decay products, when it is impossible to reconstruct two separate track jets inside the large$-R$ jet. In case of the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis, the priority-merged selection is used to avoid considering the same event in the resolved and merged regimes simultaneously, as discussed in Section 7.1.
The large$-R$ jets are ranked according to their flavour in case of the $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ analysis and p_T. In case of $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ analysis, the highest rank is assigned to the large$-R$ jets with the $b-$tagged track jets inside and the lowest to the large$-R$ jets with no $b-$tagged track jets inside. Within each cathegory, jet are ordered in decreasing p_T. The leading large$-R$ jet with the highest rank is associated with the boosted $h \to b\bar{b}$ or $W/Z \to q\bar{q}$ decay. The resulting four-momenta of the $h/W/Z$ candidate is then assigned to the four-momenta of the corresponding leading jet.
The following signal selection criteria are applied to the events in the merged regime:

- $\geq 1J$: Presence of at least one large$-R$ is required.

- **Mass window requirement:** 75 GeV $< m_{jj} <$ 100 GeV in the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis (see Section 7.4). No mass window requirement applied in the $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ selection.

- W/Z **tagger requirement:** Applied to associate jets with the hadronic W/Z decays (see Section 7.5). Used only in the $0, 1$ $b-$tag regions. In the 2 $b-$tag region the signal events mainly come from the $Z \to b\bar{b}$ decay with very little contamination from the background processes, and thus only the mass window requirement is used. Events, which satisfy both the mass window and tagger requirements, correspond to

the *high purity* signal region. Events, which satisfy the mass window requirements, but fail the tagger requirement, correspond to the *low purity* signal region.

7.8 $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ optimisation selection

The majority of the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ event selection is inherited from the previous version of the analysis [44]. The improvement in the sensitivity of the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ search comes so far mainly from the larger data set, increased center-of-mass energy and improved detector performance. One of the main targets for improvement is constraining the main backgrounds. In case of $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ search, one of the dominant backgrounds is the $t\bar{t}$ process, which accounted for up to 80% of the total background in the resolved regime and 40% in the merged regime in the signal region of the Run 1 search [44]. A set of optimised selection criteria are developed for 2015+2016 analysis in order to reduce the $t\bar{t}$ background and thus increase the sensitivity of the search. The discussion here follows [198].
The optimised selection criteria in the resolved regime are applied in the following order:

- **Additional $b-$jet veto**: Veto events with more than two $b-$tagged small$-R$ jets, since it is expected to have exactly two $b-$jets in the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ final state. The majority of the additional jets, which come mainly from the initial state radiation, originate from gluons and light-flavour quarks. Thus, the presence of additional $b-$jets rather indicates that event comes from the $t\bar{t}$ or single top background. Around 70% of the $t\bar{t}$ and 50% of the single top events contain two real $b-$jets and at least one real $c-$jet. Given that the MV2c10 $b-$tagging algorithm, discussed in Section 6.2.7, mistags around 8% of $c-$jets as the $b-$jets, a significant fraction of the $t\bar{t}$ and single top events will contain more than two $b-$tagged jets.

- **Standard and extended tau veto:** Veto events, which contain any standard or extended tau leptons [198]. It is motivated by the fact that around 70% of the $t\bar{t}$, single top and $W+$jets events contain at least one real hadronically decaying tau lepton and real $E_{\mathrm{T}}^{\mathrm{miss}}$ from the $W \to \tau\nu_\tau$ decay. Since tau lepton is reconstructed as a small$-R$ jet, these events look like signal events with additional light-flavour jets. Since significant fraction of events still contain taus after vetoing standard *loose* taus, a new custom *extended* tau identification is used to veto the remaining taus. The standard and extended taus are defined in Table 6.2 in Section 6.5.

- H_T **ratio requirement:** The main hadronic activity in the signal $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ event is expected from the $b-$jets associated with the $h \to b\bar{b}$ decay, while only a tiny fraction comes from the additional jet activity. Meantime, in the signal-like $t\bar{t}$ event, the jet activity is distributed between the pair of $b-$quarks from the $t\bar{t}$ decay (or large$-R$ jet from the boosted top quark decay) and jets (tau leptons) from the $W \to q\bar{q}$ ($W \to \tau\nu_\tau$) decay. The Higgs candidate in such $t\bar{t}$ event is likely recostructed from the $b\bar{b}$ pair or from the large$-R$ jets with $b-$tagged track jet inside. The measure of the hadronic activity in the event is the scalar sum of jet p_T in the event, H_T. In terms of H_T the above-mentioned means that H_T(non-h jets)/H_T(all jets) $\ll 1$ for a signal $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ event and H_T(non-h jets)/H_T(all jets) < 1 for a signal-like $t\bar{t}$ event. The H_T(non-h jets)/H_T(all jets) < 0.37 requirement shows the best discriminating power in the resolved regime, while rejecting a tiny fraction of the signal events.

- $\Delta R(j_1, j_2) < 1.8$: It takes into account that products of the $h \to b\bar{b}$ decay are collimated due to the boost of the Higgs boson, while products of the $t\bar{t}$ decay are mostly produced back-to-back.

The distributions of the H_T ratio and $\Delta R(j_1, j_2)$ in the resolved regime for various backgrounds and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ signal models in for the DM production are shown in Figure 7.4.

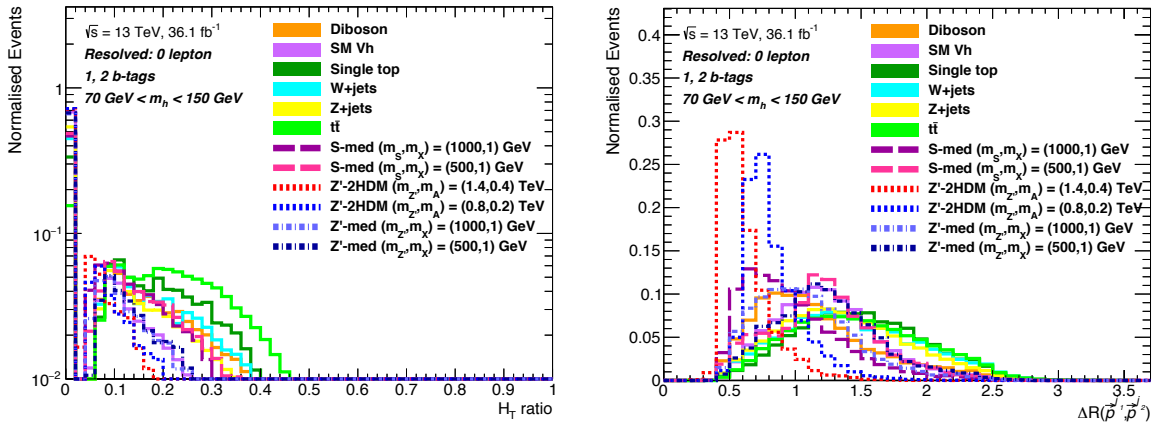


Figure 7.4: H_T ratio (left) and $\Delta R(j_1, j_2)$ (right) distributions normalised to unil area in the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ signal region in the resolved regime within the Higgs candidate mass window of $(70, 140)$ GeV are shown [198]. The main SM backgrounds are shown in solid lines. Z'-2HDM, Z' and scalar mediator signal models are shown in dashed lines. Signal events mostly populate low H_T ratio and $\Delta R(j_1, j_2)$ regions.

The optimised selection criteria in the merged regime are applied in the following order:

- **Non-associated $b-$jet veto**: Veto events with $b-$jets not associated with the large$-R$ jet coming from the boosted $h \to b\bar{b}$ decay. Both $b-$tagged tack jets of the boosted $h \to b\bar{b}$ decay are expected to be associated with the same large$-R$ jet. Meantime, the $b-$tagged track jets from the $t\bar{t}$ decay are contained in the different large$-R$ jets due to the back-to-back topology of the decay. Since the leading large$-R$ jets is reconstructed as the Higgs candidate, the second $b-$tagged track jet cannot be associated with it. Almost all $t\bar{t}$ and single top events contain a real $b-$jet as the leading non-associated $b-$tagged track jet.

- **Non-associated standard and extended tau veto:** Veto events, which contain any standard or extended tau leptons not associated with the large$-R$ jet coming from the boosted $h \to b\bar{b}$ decay. The motivation comes from the back-to-back topology of the $t\bar{t}$ decay, in which tau lepton is contained in the large$-R$ jet, which is not associated with the boosted $h \to b\bar{b}$ decay.

- H_T **ratio requirement:** Same idea as described for the resolved regime. The H_T ratio in the merged regime is defined in the following way:

$$H_T \text{ ratio} \equiv \frac{\sum\limits_{\substack{\text{non-associated} \\ \text{small-}R \text{ jets}}} p_T^j}{p_T^J + \sum\limits_{\substack{\text{non-associated} \\ \text{small-}R \text{ jets}}} p_T^j}, \tag{7.2}$$

where p_T^J stands for the p_T of the leading large$-R$ jet. The H_T ratio is required to be less than 0.57 in the merged regime.

The optimised event selection significantly reduces the $t\bar{t}$ (total) background by 50% (45%) and 70% (35%) in the resolved and merged regimes, respectively, as shown on Fig 7.5.

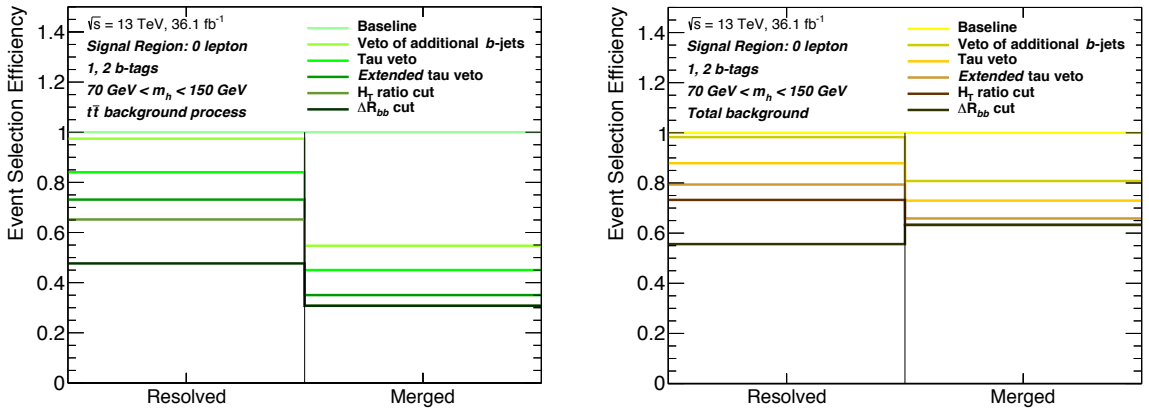


Figure 7.5: Event selection efficiency reduction after applying the optimised selection for the $t\bar{t}$ background process (left) and total background (right) in the 0 lepton region. The coloured lines show the efficiency after consecutively applying the optimised event selections within the Higgs candidate mass window 70 GeV $< m_h <$ 150 GeV.

7.9 Main observables

The main discriminating variable, which enters the statistical models of both analyses, is the invariant mass of the $h/W/Z$ candidates, $m_{h,W,Z} = m_{jj}/m_J$, respectively. However, the invariant mass distribution is treated differently in both analyses. In case of the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis, one includes the full information from the invariant mass distribution in the combined fit (shape and normalisation). Figure 7.6 shows the invariant mass distribution in the $[350, 500)$ GeV and $[500, \infty)$ GeV $E_{\mathrm{T}}^{\mathrm{miss}}$, 0 lepton signal region before the combined fit. The potential DM signal then manifests itself through an excess over the SM background in the invariant mass distribution of the Higgs candidate after the combined fit to the observed data.
Meantime, in case of the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis, one ignores the shape of the mass distribution and considers instead a simple counting experiment for the number of events in the W/Z mass window and mass sidebands of the invariant mass distribution. The potential DM signal then manifests itself through an excess over the SM background in the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution after the combined fit to the observed data.

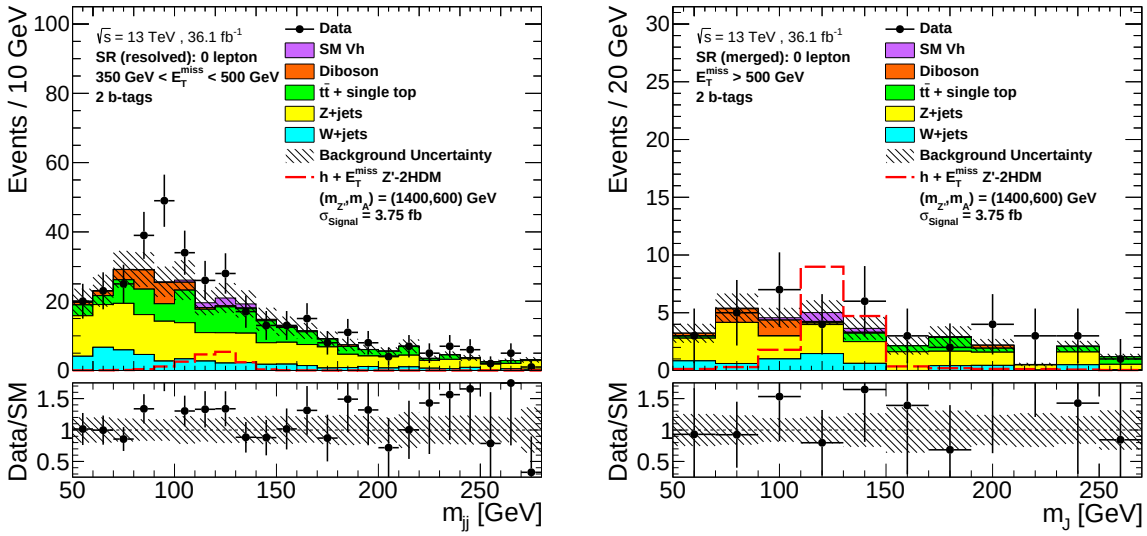


Figure 7.6: Distributions of the mass of the Higgs candidate in the 0 lepton signal region, 2 $b-$tags after the event selection before the fit to data. Plots for the $[350, 500)$ GeV (left) and $[500, \infty)$ GeV (right) $E_{\mathrm{T}}^{\mathrm{miss}}$ regions are shown. The red dashed line shows the expected DM signal coming from the Z'-2HDM model with $(m_{Z'}, m_A) = (1400, 600)$ GeV. The total background uncertainty is shown as a hatched band.

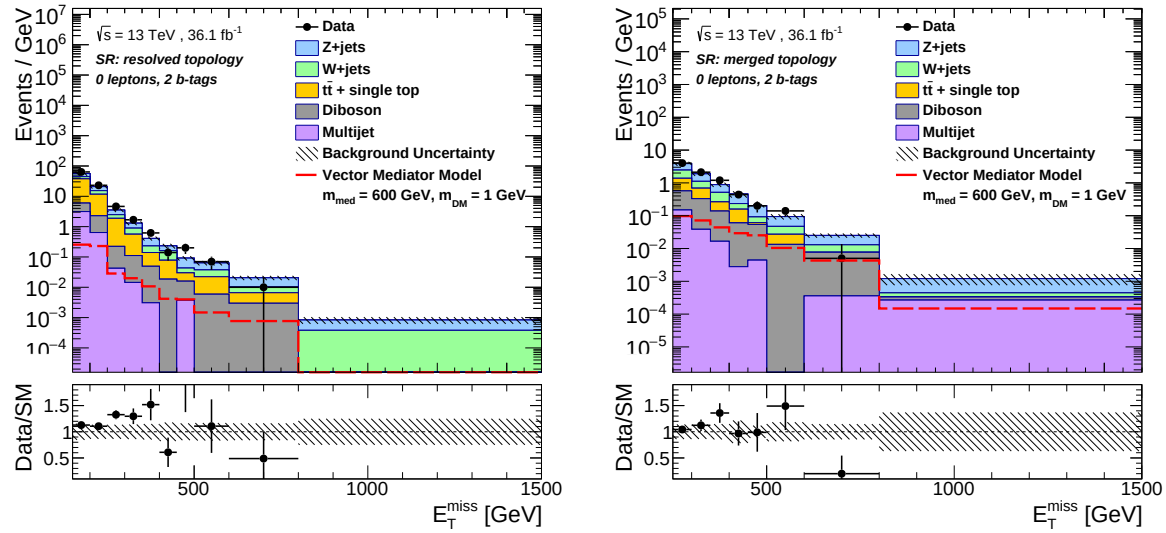


Figure 7.7: Distributions of the $E_{\mathrm{T}}^{\mathrm{miss}}$ in the resolved (left) and merged (right) regimes in the $E_{\mathrm{T}}^{\mathrm{miss}}+V(qq)$ 0 lepton signal region, 2 $b-$tags after the event selection before the fit to data. The red dashed line shows the expected DM signal coming from the simplified vector mediator model with $(m_\chi, m_{Z'}) = (1, 600)$ GeV. The total background uncertainty is shown as a hatched band.

7.10 Summary of event selections and analysis regions

A concised summary of all selection criteria applied in different lepton and $b-$tag regions of the $E_{\mathrm{T}}^{\mathrm{miss}}+V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ analyses is given in Tables 7.3 and 7.4, respectively. The structure of the $E_{\mathrm{T}}^{\mathrm{miss}}+V(qq)$ analysis regions is generally more complex than of the $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ analysis due to the W/Z mass and tagging requirements. The $E_{\mathrm{T}}^{\mathrm{miss}}+V(qq)$ phase space is divided into 40 analysis regions - 6 signal regions, 2 low purity signal regions, 8 zero lepton mass sideband regions and 24 control regions, as shown in Table 7.5.

All analysis regions are fitted simultaneously using the *combined binned profile likelihood fit*, discussed in Section 9.1. The implementation of the nuisance parameters and the coresponding constraints in the combined fit is given in Chapter 10.

Region	Resolved regime	Merged regime
0 lepton	Baseline selection *loose* lepton veto anti-QCD cuts (see Section 7.3) $E_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV $\geq$ central $2j$ [1] $p_T^{j_1} > 45$ GeV $\sum_{i=1}^{2(3)} p_T^{j_i} > 120(150)$ GeV $\Delta\phi(j_1, j_2) < 7\pi/9$ 65 GeV $< m_{jj} <$ 105(100) GeV for $0, 1(2)$ $b-$tags $\Delta R(j_1, j_2) <$ 1.4(1.25) for 0, 1 (2) $b-$tags ≤ 2 $b-$tagged jets	Baseline selection anti-QCD cuts (see Section 7.3) $E_{\mathrm{T}}^{\mathrm{miss}} > 250$ GeV $\geq 1J$ [1] W/Z tagging for $0, 1$ $b-$tags **OR** 75 GeV $< m_J <$ 100 GeV for 2 $b-$tags 0 non-associated $b-$tagged track-jets
1 lepton	Baseline selection One $W-$signal μ, no other $V-$loose e/μ anti-QCD cuts (see Section 7.3) $p_T^V > 150$ GeV [2] $\geq$ central $2j$ [1] $p_T^{j_1} > 45$ GeV $\sum_{i=1}^{2(3)} p_T^{j_i} > 120(150)$ GeV 65 GeV $< m_{jj} <$ 105(100) GeV for 0, 1 (2) $b-$tags $\Delta\phi(j_1, j_2) < 7\pi/9$ ≤ 2 $b-$tagged jets	$p_T^V > 250$ GeV [2] $\geq 1J$ [1] $D_2^{\beta=1}$ tagging **OR** 2 $b-$tagged track jets Mass window requirement in W/Z tagger for $0, 1$ $b-$tags **OR** 75 GeV $< m_J <$ 100 GeV for 2 $b-$tags 0 non-associated $b-$tagged track-jets
2 lepton	Baseline selection Two $V-$loose e/μ, at least one of which is $W-$signal anti-QCD cuts (see Section 7.3) 66 GeV $\leq m_{ee/\mu\mu} \leq$ 116 GeV $p_T^V > 150$ GeV [2] $\geq$ central $2j$ [1] $p_T^{j_1} > 45$ GeV $\sum_{i=1}^{2(3)} p_T^{j_i} > 120(150)$ GeV 65 GeV $< m_{jj} <$ 105(100) GeV for 0, 1 (2) $b-$tags $\Delta\phi(j_1, j_2) < 7\pi/9$ ≤ 2 $b-$tagged jets	$p_T^V > 250$ GeV [2] $\geq 1J$ [1] $D_2^{\beta=1}$ tagging **OR** 2 $b-$tagged track jets Mass window requirement in W/Z tagger for $0, 1$ $b-$tags **OR** 75 GeV $< m_J <$ 100 GeV for 2 $b-$tags 0 non-associated $b-$tagged track-jets

Table 7.3: Summary of the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ event selections in 0,1 and 2 lepton regions. The definitions of the $W-$signal and $V-$loose leptons is given in Section 6.5. Details about the control regions are discussed in Chapter 8.

[1] $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis use merged-priority selection, meaning that event in $E_{\mathrm{T}}^{\mathrm{miss}} > 250$ GeV region falls into the merged regime if contains at least one large$-R$ jet.

[2] p_T^V is defined as $p_T^V = |\mathbf{E_T^{miss}} + \mathbf{p_T^{\mu}}|$ in 1 lepton region and as $p_T^V = |\mathbf{E_T^{miss}} + \mathbf{p_T^{e_1/\mu_1}} + \mathbf{p_T^{e_2/\mu_2}}|$ in 2 lepton region.

Region	Resolved regime	Merged regime
0 lepton	Baseline selection *loose* lepton veto anti-QCD cuts (see Section 7.3) 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} \leq$ 500 GeV	$E_{\mathrm{T}}^{\mathrm{miss}} >$ 500 GeV
	$\geq$ central $2j$	$\geq 1J$
	$p_T^{j_1}$ **OR** $^{j_2} >$ 45 GeV $\sum_{i=1}^{2(3)} p_T^{j_i} >$ 120(150) GeV $\Delta\phi(j_1, j_2) < 7\pi/9$ optimisation selection (see Section 7.8)	
1 lepton	Baseline selection One $W-$signal μ, no other $V-$loose e/μ anti-QCD cuts (see Section 7.3) 150 GeV $< p_T^V \leq$ 500 GeV [2]	$p_T^V >$ 500 GeV [2]
	$\geq$ central $2j$	$\geq 1J$
	$p_T^{j_1}$ **OR** $^{j_2} >$ 45 GeV $\sum_{i=1}^{2(3)} p_T^{j_i} >$ 120(150) GeV 65 GeV $\leq m_{jj} \leq$ 105(100) GeV for 0, 1 (2) $b-$tags $\Delta\phi(j_1, j_2) < 7\pi/9$ optimisation selection (see Section 7.8)	
2 lepton	Baseline selection Two $V-$loose e/μ, at least one of which is $Z-$signal opposite-charged muons anti-QCD cuts (see Section 7.3) 71 GeV $\leq m_{ee} \leq$ 106 GeV **OR** 83 GeV $\leq m_{\mu\mu} \leq$ 99 GeV 150 GeV $< p_T^V \leq$ 500 GeV [2]	$p_T^V >$ 500 GeV [2]
	$\geq$ central $2j$	$\geq 1J$
	$p_T^{j_1}$ **OR** $^{j_2} >$ 45 GeV $\sum_{i=1}^{2(3)} p_T^{j_i} >$ 120(150) GeV $\Delta\phi(j_1, j_2) < 7\pi/9$ optimisation selection (see Section 7.8)	

Table 7.4: Summary of the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ event selections in 0,1 and 2 lepton regions. The definitions of the $W-$signal and $V-$loose leptons is given in Section 6.5. Details about the control regions are discussed in Chapter 8.

$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ **signal regions**								
	Resolved			Merged				
$b-$tags	0	1	2	0 HP	0 LP	1 HP	1 LP	2
$\Delta R(j_1, j_2)$	< 1.4	< 1.4	< 1.25	–	–	–	–	–
$D_2^{\beta=1}$	–	–	–	pass	fail	pass	fail	–
Mass window [GeV]	m_{jj} $[65, 105]$		m_{jj} $[65, 100]$	pass m_J W/Z tagger requirement				m_J $[75, 100]$

$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ **0 lepton control regions**								
	Resolved			Merged				
$b-$tags	0 SB	1 SB	2 SB	0 SB	0 LP SB	1 SB	1 LP SB	2 SB
$\Delta R(j_1, j_2)$	< 1.4	< 1.4	< 1.25	–	–	–	–	–
$D_2^{\beta=1}$	–	–	–	pass	fail	pass	fail	–
Mass window [GeV]	m_{jj} $[105, \infty)$		m_{jj} $[100, \infty)$	fail m_J W/Z tagger requirement				m_J $[100, \infty)$

$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ **1+2 lepton control regions**												
	Resolved						Merged					
$b-$tags	0	1	2	0 SB	1 SB	2 SB	0	1	2	0 SB	1 SB	2 SB
Mass window [GeV]	m_{jj} $[65, 105]$		m_{jj} $[65, 100]$	m_{jj} $[105, \infty)$		m_{jj} $[100, \infty)$	m_J $[75, 100]$			m_J $[100, \infty)$		

Table 7.5: Summary of the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ signal and control regions incorporated in the final combined binned profile likelihood fit. The abbreviations HP and LP correspond to respectively the high- and low-purity analysis regions in the merged topology. The abbreviation SB corresponds to the mass sideband regions.

Chapter 8

Background Estimation

One of the main challenges of every search for the new physics is to constrain the Standard Model (SM) background processes that provide the signatures identical to those expected from the possible new physics processes. The signatures of the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analyses include a pair of jets, coming from the $h \to b\bar{b}$ and $V \to q\bar{q}$ decays characterised by the large branching ratios [1]. However, this implies the large rates of the corresponding SM backgrounds. For the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ searches, the dominant processes are $W+$jets, $Z+$jets and $t\bar{t}$, which comprises more than 90% of the total background in the signal regions (SR). These background processes are modelled using the Monte Carlo (MC) simulations and constrained by designing the dedicated control regions (CR) enriched in these processes. In the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analyses these CRs regions are the 1 lepton CR, which is aimed to constrain $W+$jets and $t\bar{t}$ backgrounds, and 2 lepton CRs, which is designed to constrain $Z+$jets background. As discussed in Section 10.2, the combined fit performed simultaneously in the SRs and CRs makes it possible to constrain the normalisation parameters of the dominant background processes that are initally incorporated as the freely floating parameters in the statistical model. The only background that cannot be estimated with the MC simulated samples is the multijet background due to the lack of sufficient statistics that is required for the MC multijet modelling. Thus, this background is estimates using a dedicated data-driven method.
The detailed description of the main background processes is given in Section 8.1. The control regions used to constrain the backgrounds are introduced in Sections 8.2 and 8.3. The multijet estimation method is briefly discussed in Section 8.4.

8.1 Background processes

$V+$jets production

The $V+$jets background is by far the largest for the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ search, accounting for more than 80% of the total background in the SR, and is one of the largest together with the $t\bar{t}$ background for the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ search. The dominant contribution in the SR comes from the $Z+$jets production, in which Z boson decays to a pair of neutrinos, it provides the same dijet+$E_{\mathrm{T}}^{\mathrm{miss}}$ signature as the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ searches and thus is naturally selected in the SR. The $Z(\nu\nu)+$jets background is complete irre-

[1] The SM value for the branching ratios are $\mathcal{B}_{h\to b\bar{b}} \approx 0.57$ and $\mathcal{B}_{W/Z\to q\bar{q}} \approx 0.7$

ducible in the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis as it provides the identical $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ signature. However, it is semi-irreducible in the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis, meaning that it can be partially distinguishable from the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ signature since its dijet mass spectrum is non-resonant at the Higgs boson mass. The $W+$jets background enters the SR if the W boson decays to tau lepton, $W \to \tau\nu_{\tau}$, which in turn decays hadronically and thus is identified as a QCD-originated jet. For the resolved part of the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis, $\sim 60\%$ of the $W+$jets events within the Higgs boson candidate mass window of $(70, 150)$ GeV in the SR contain a hadronically decaying tau lepton, as shown in Figure 8.1. Figure 8.2 illustrates the tree-level Feynman diagrams of the $V+$jets production.

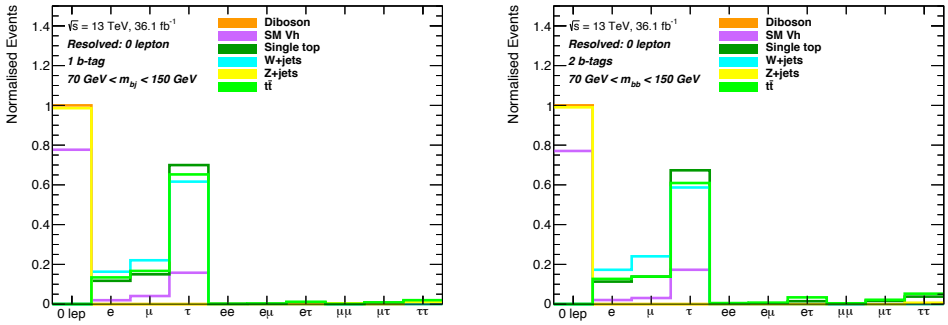


Figure 8.1: The particle-level flavour composition of the final state particles normalised to unity area for the main backgrounds in the SR for the resolved regime within the Higgs boson candidate mass window of $(70, 150)$ GeV are shown [198]. The left plot corresponds to the region with 1 $b-$tag, while the right plot to the region with 2 $b-$tags. Plots are not produced by the author.

Depending on the flavour composition of the final state particles, the $V+$jets background is split in the flavour cathegories to account for the different sensitivities in the different analyses regions. This splitting is also used in the implementation of the theoretical modelling uncertanties, as discussed in Section 10.2. The flavour labelling is assigned by the geometrical matching between the two leading reconstructed jets and particle-level jets, for which the flavour of the originating quark is known from the MC simulation. The $V + ll$ (V+LF, *light flavour*) cathegory stands for the $V+$jets processes, in which the two leading reconstructed jets are associated with the particle-level $u, d, s-$quark jets, referred to as *light jets*. The $V + cl$ cathegory represents the $V+$jets processes, in which the one of the reconstructed jets are associated with the particle-level $c-$quark jet, while other one with the particle-level light jet. And finally, the V+HF (*heavy flavour*) cathegory combines the $V+$jets processes, in which either one of the reconstructed jets is matched to the particle-level $b-$quark jet ($V + bl, bc, bb$) or both leading reconstructed jets are matched to the particle-level $c-$quark jets ($V + cc$).

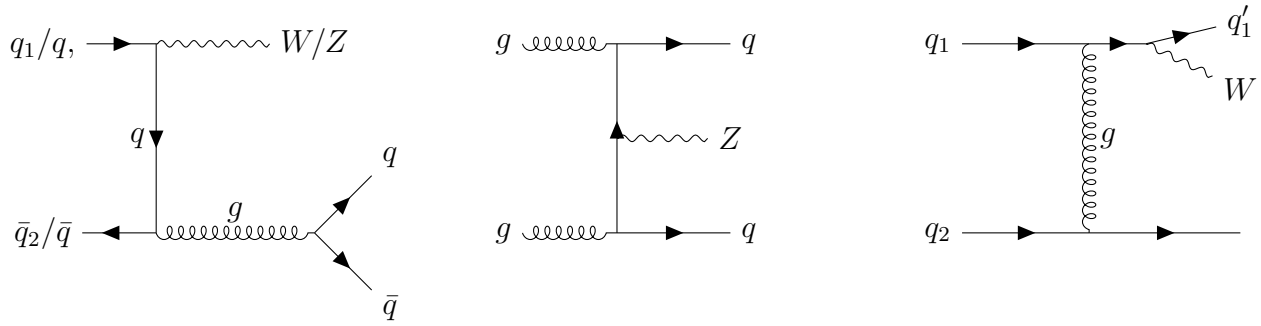


Figure 8.2: Tree-level Feynman diagrams of the $V+$jets production for the $E_{\mathrm{T}}^{\mathrm{miss}}+V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ analyses.

$t\bar{t}$ production

The $t\bar{t}$ background is the second dominant in the SR of the $E_{\mathrm{T}}^{\mathrm{miss}}+V(qq)$ analysis. Before the $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ optimisation selection was introduced (see Section 7.8), the $t\bar{t}$ background was the most dominant for the Run 1 version of the $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ search [44], accounting for the 80% (40%) of the total background in the SR for the resolved (merged) regimes. With the optimisation selection it still provides a sizeable contribution comparable to the $V+$jets contribution in the resolved regime.
The $t\bar{t}$ events are split in three cathegories according to the decay channels of the W bosons coming from the $t \to Wb$ decays. If both W bosons decay leptonically, the corresponding $t\bar{t}$ event is referred to as *dileptonic*. The contribution of the semi-leptonic $t\bar{t}$ events in the SR is negligible due to the lepton veto. If both W bosons decay hadronically, the corresponding $t\bar{t}$ event is referred to as *hadronic*. Since the hadronic $t\bar{t}$ events contain no real $E_{\mathrm{T}}^{\mathrm{miss}}$, they can contribute to the SR only in case of jet energy mismeasurement, which is strongly suppressed by the anti-QCD selection criteria, introduced in Section 7.3. If one W boson decay leptonically and the other hadronically, the corresponding $t\bar{t}$ event is referred to as *semi-leptonic*. The hadronic $t\bar{t}$ events are selected in the SR if the W boson decays as $W \to \tau\nu_\tau$, where the tau lepton decays hadronically. They provide the signal-like signature with two real $b-$quark jets, $E_{\mathrm{T}}^{\mathrm{miss}}$ from the $W \to \tau\nu_\tau$ decay and additional jets coming from the $W \to q\bar{q}$ and hadronic tau lepton decays. About 70% of the $t\bar{t}$ events within the Higgs boson candidate mass window of $(70, 150)$ GeV in the SR of the $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ analysis contain a hadronically decaying tau lepton, as shown in Figure 8.1.

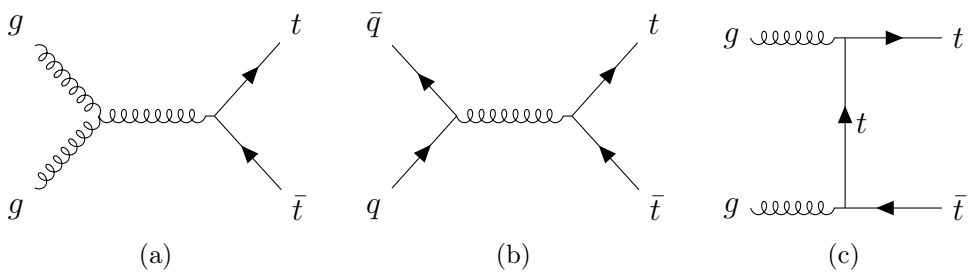


Figure 8.3: Tree-level Feynman diagrams of the $t\bar{t}$ production for the $E_{\mathrm{T}}^{\mathrm{miss}}+V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ analyses.

Single-top production

The single-top production get an access to the SR of the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ via the decay chain $t \to Wb \to \tau\nu_\tau b$, where the hadronically decaying tau lepton is identified as the QCD-originated jet. The tree-level Feynman diagrams of the single-top production are shown in Figure 8.4. Given that the single-top final state contains at least one real $b-$quark jet, it mainly contributes to the analyses regions with 1 and 2 $b-$tags. The associated Wt production mainly contributes to the 2 $b-$tag region via the $c-$jets, coming from the hadronic decay of the one of the two W bosons, that are mistagged as $b-$jets with 8% mistagging efficiency (see Section 6.2.7). For the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis, the one of the W bosons decay to the tau lepton and neutrino in about 70% (75%) of the single-top events in the regions with 1 $b-$tag (2 $b-$tags).

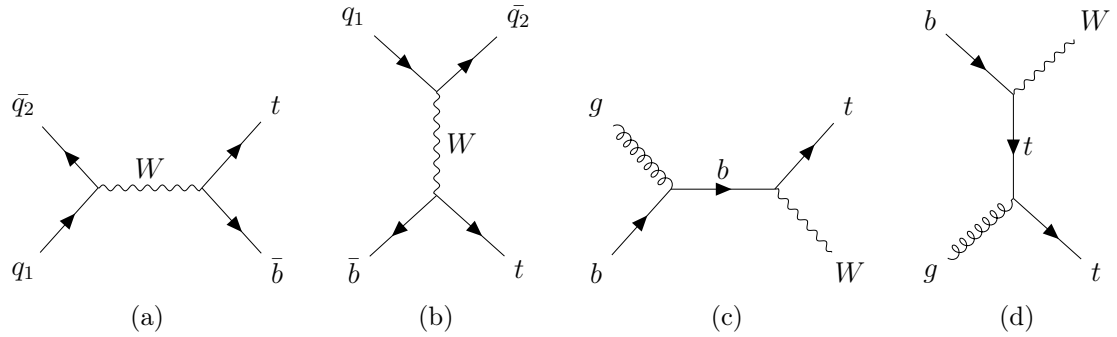


Figure 8.4: Tree-level Feynman diagrams of the single-top production for the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analyses. The (a) and (b) diagrams stand for the $s-$channel and $t-$channel production, respectively. The (c) and (d) diagrams stand for the $s-$channel and $t-$channel Wt associated production, respectively.

SM Vh and diboson production

The SM Vh and diboson backgrounds are another examples of the processes that are resonant at the Higgs and vector boson masses. The corresponding tree-level Feynman diagrams are shown in Figure 8.5. These backgrounds can naturally pass the signal event selection for the Zh and ZV processes when the Z boson decays invisibly to a pair of neutrinos and the remaining h/V bosons decay hadronically [2]. The SM Wh processes also contribute in the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ SR in case of the $W \to \tau\nu_\tau$ decays, where the hadronically decaying τ lepton mimicks the QCD-originated jet. The SM Zh and Wh processes account for the 80% and 20% of the total SM Vh background within the Higgs boson candidate mass window of $(70, 150)$ GeV, as illustrated in Figure 8.1. Both backgrounds also contribute in the 1 and 2 lepton CR for the leptonically decaying W and Z bosons, respectively.

[2]Both $Z(\nu\nu)V(q\bar{q})$ and $Z(\nu\nu)h(b\bar{b})$ backgrounds are irreducible, meaning that one cannot distringuish between these final states and the signal $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ final states, respectively.

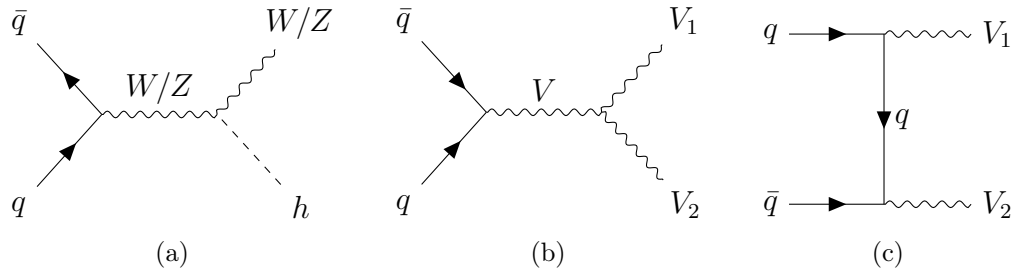


Figure 8.5: Tree-level Feynman diagrams of the (a) SM Vh and (b), (c) diboson productions for the $E_{\mathrm{T}}^{\mathrm{miss}}+V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ analyses. The V stands for the (b), (c) W, Z bosons and the (b) photon. The V_1V_2 final state represents the WW, ZZ and WZ backgrounds.

8.2 1 lepton control region

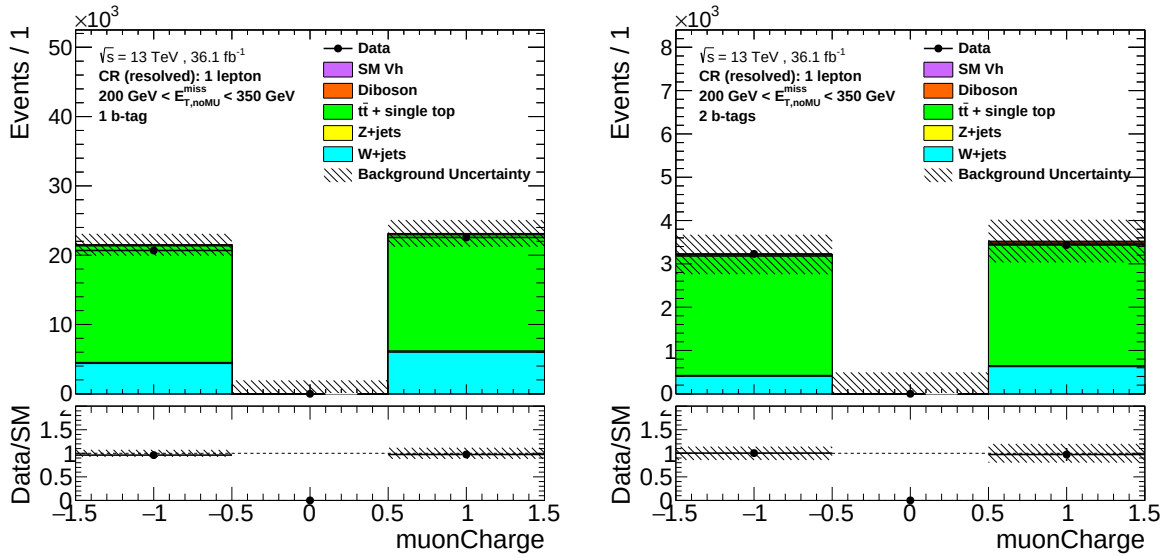


Figure 8.6: Distributions of the muon charge in the 1μCR for the $[200, 350)$ GeV region of $E_{\mathrm{T}}^{\mathrm{miss}}$ with 1 $b-$tag (left) and 2 $b-$tags (right) after the event selection before the fit to data. The total background uncertainty is shown as a hatched band.

The 1 lepton CR is aimed to constrain the normalisations of the $W+$jets and $t\bar{t}$ backgrounds. It is defined by requiring exactly one $W-$signal muon in the event and discarding events with additional $V-$loose muons or electrons. The definitions of the $W-$signal and $V-$loose leptons are given in Section 6.5. The remaining event selection in the CR is similar to that in the SR to keep the events in the SR and CR kinematically and topologically similar. In what follows, this CR is referred to as 1 muon CR or 1μCR. The requirement of the kinematic and topological similarity implies re-definition of the $E_{\mathrm{T}}^{\mathrm{miss}}$ and p_T^{miss} in this CR, since muons are not considered in the $E_{\mathrm{T}}^{\mathrm{miss}}$ and p_T^{miss} calculation in the SR but considered in the 1μCR. Thus, to make $E_{\mathrm{T}}^{\mathrm{miss}}$ and p_T^{miss} in the SR and CR kinematically similar, one introduces the $E_{\mathrm{T,nomu}}^{\mathrm{miss}}$ observable, which is the sum of the four-momenta of the real $E_{\mathrm{T}}^{\mathrm{miss}}$ in the 1μCR and four-momenta of the muon. The $p_{\mathrm{T,nomu}}^{\mathrm{miss}}$ observable is introduced accordingly by summing the four-momenta of the real p_T^{miss} in the 1μCR and four-momenta of the muon.

The main discriminating observable to distinguish between the $W+$jets and $t\bar{t}$ backgrounds in the 1μCR is the electric charge of the muon. The distribution of the muon

charge is expected to be assymetric for the $W+$jets background and symmetric for the $t\bar{t}$ background. The former is related to the charge assymetry in the $W+$jets production in the pp collisions at the LHC, $\sigma(W^+ + \text{jets}) \approx 1.3\sigma(W^- + \text{jets})$, due to the larger number of the valence $u-$quarks relative to the $d-$quarks in the proton, [199]. [3]
Figure 8.6 shows the distributions of the muon charge in the $[350, 500)$ GeV and $[500, \infty)$ regions of $E_{\mathrm{T}}^{\mathrm{miss}}$ with 2 $b-$tags before the fit to data. As expected, the distributions are assymetric for the $W+$jets background and symmetric for the $t\bar{t}$ background. The disagreement between the data and SM prediction originates from the mismodelling of the $W+$jets and $t\bar{t}$ normalisations before the fit to data and is corrected after the fit to the observed data, as will be shown in Section 12.1.3, Appendices D and F.

8.3 2 lepton control region

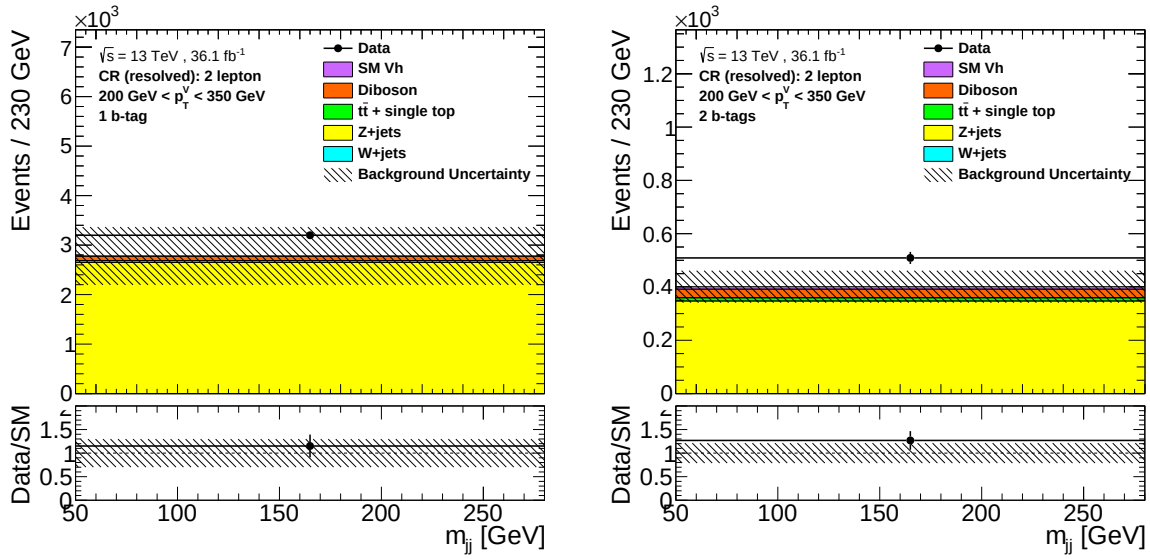


Figure 8.7: Distributions of the muon charge in the 2ℓCR for the $[200, 350)$ GeV region of $E_{\mathrm{T}}^{\mathrm{miss}}$ with 1 $b-$tag (left) and 2 $b-$tags (right) after the event selection before the fit to data. The total background uncertainty is shown as a hatched band.

The 2 lepton CR is aimed to constrain the normalisations of the $Z(\nu\nu)+$jets background in the SR using the $Z(ll)+$jets background process. This CR, denoted as $2l$CR, is defined by requiring two $V-$loose muons or two $V-$loose electrons, at least one of which also passes the $Z-$signal requirement. Since the p_T of the Z boson does not dependent on whether it decays hadronically or leptonically, the event kinematics of the $Z(\nu\nu)+$jets and $Z(ll)+$jets processes are very similar. This means that one can constrain the normalisation of the $Z(\nu\nu)+$jets background in the SR using the normalisation of the $Z(ll)+$jets background in the $2l$CR. The remaining event selection is similar to those in the SR, except the following requirements:

- **Invariant mass of the two leptons:** 71 GeV $\leq m_{ee} \leq$ 106 GeV **OR** 83 GeV $\leq m_{\mu\mu} \leq$ 99 GeV for $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis and 66 GeV $\leq m_{ee/\mu\mu} \leq$ 116 GeV for the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis. Since muons come from the $Z \to \mu\mu$ decay, they are

[3] If this measurement were done at the $p\bar{p}$ collider, such as the Tevatron collider, the charge distribution in the $W+$jets production would be symmetric.

required to have opposite electric charges. This requirement on the invariant mass of the dilepton system is aimed to suppress the contributions of the non-resonant $t\bar{t}$, single-top and multijet background processes.

- $E_{\mathrm{T}}^{\mathrm{miss}}$ **significance:** $E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_T^*} < 3.5$ only for the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis, where $H_T^* = p_T^{ll} + \sum \mathrm{small} - R\ \mathrm{jets} p_T$. This requirement is aimed to reduce the $t\bar{t}$, since the $t\bar{t}$ events contain real $E_{\mathrm{T}}^{\mathrm{miss}}$ from the leptonically decaying W, while $Z(ll) + \mathrm{jets}$ events contain artificial $E_{\mathrm{T}}^{\mathrm{miss}}$ coming mainly from the jet energy mismeasurements.

The $E_{\mathrm{T}}^{\mathrm{miss}}$ and p_T^{miss} in the $2l$CR is also re-defined accordingly to the presence of two leptons in the event by adding the four-momentum of the dilepton system.
Since the $2l$CR is aimed to constrain only the normalisations of the backgrounds, the shape information of the distributions of the main observables is discarded and only the event yield is used as a main discriminant. Figure 8.7 shows the event yields for the regions with different $E_{\mathrm{T}}^{\mathrm{miss}}$ and number of the $b-$tagged jets before the fit to data. The disagreement between the data and SM prediction originates from the mismodelling of the $Z + \mathrm{jets}$ normalisations before the fit to data and is corrected after the fit to the observed data, as will be shown in Section 12.1.3, Appendices D and F.

8.4 Multijet estimate

As was mentioned before, the multijet background is estimated using a dedicated data-driven template method. Due to the requirement of the large $E_{\mathrm{T}}^{\mathrm{miss}}$ and anti-QCD selection, the contribution of the multijet processes in the SR is expected to be sub-dominant relative to other background.
The first step is to define the CR enriched with the multijet events by inverting the dominant anti-QCD requirement in the signal event selection, $\min\left[\Delta\phi(\mathbf{E_T^{miss}}, \mathbf{p_T^{j_{1,2,3}}})\right] > \pi/9$ (see Section 7.3). The unnormalised multijet templates are generated by subtracting the contributions of all non-multijet backgrounds from the data. The number of templates for each analysis follows the number of the corresponding analysis sub-regions in the 0 lepton region.
In the second step, the normalisation of the multijet templates is obtained from the combined fit to data. The selection requirements stay the same as in the SR, except the inverted $\min\left[\Delta\phi(\mathbf{E_T^{miss}}, \mathbf{p_T^{j_{1,2,3}}})\right] > \pi/9$ requirement and those requirements that are dropped or relaxed due to their significant impact on the multijet background. This is referred to as *relaxed* SR. For the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis there requirements are the following:

- The $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement is relaxed to $E_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV in the merged regime. The event selections in the 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 250 GeV region for the resolved and merged regimes are allowed to overlap to eliminate the migration effects due to the priority-merged selection. The priority-merged selection is still used in the $E_{\mathrm{T}}^{\mathrm{miss}} > 250$ GeV region.

- The $\Delta\phi(\mathbf{E_T^{miss}}, \mathbf{p_T^{miss}}) < \pi/2$ requirement is dropped to further increase the multijet contribution. The $\Delta\phi(\mathbf{E_T^{miss}}, \mathbf{p_T^{h/W/Z}}) > 6\pi/9$ requirement is kept due to the strong correlation with the $\min\left[\Delta\phi(\mathbf{E_T^{miss}}, \mathbf{p_T^{j_{1,2,3}}})\right]$ requirement.

For the $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ analysis the anti-QCD $\Delta\phi(\mathbf{E_T^{miss}}, \mathbf{p_T^{miss}}) < \pi/2$ and $\Delta\phi(\mathbf{E_T^{miss}}, \mathbf{p_T^{h/W/Z}}) > 6\pi/9$ requirements, as well as the optimisation requirement of $\Delta R(j_1, j_2) < 1.8$ for the resolved regime, are dropped to further increase the multijet contribution. Additionally, the event selection is applied to the blinded data, meaning that the events in the W/Z and h candidate mass windows are discarded to avoid the contribution from the potential signal in data.

Assuming the same shapes of the multijet templates in the multijet CR and relaxed SR, the templates used in the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ combined fits are constructed from the distributions of the $E_{\mathrm{T}}^{\mathrm{miss}}$ observable and multiplicity of the jets containing a muon, respectively. The choice of the latter is motivated by the fact that the majority of the multijet events in the SR with the real $E_{\mathrm{T}}^{\mathrm{miss}}$ contain also muons coming from the $b \to cW \to c\mu\nu_\mu$ decays inside the jets. Therefore, the multijet processes mainly populate the $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ regions with high multiplicity of the jets containing a muon. The multijet normalisation parameters is introduced as a free parameter. The normalisation parameters of the remaining non-multijet backgrounds are constrained within the uncertanties estimated from the theoretical predictions. The multijet fit for the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis is performed only in the 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 200 GeV and 200 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 350 GeV regions due to the negligible multijet contribution in the higher $E_{\mathrm{T}}^{\mathrm{miss}}$ regions. The final multijet templates of the $W/Z/h$ boson candidate masses in the multijet CR are scaled by the resulting normalisation factors and transferred to the SR.

Chapter 9

Statistical Framework

When it comes to the interpretation of the results of the search for the new physics, two fundamental questions have to be addressed: *"how well a certain model is able to describe the data?"* and *"can or cannot a discovery be claimed?"* The first question refers to the *different model hypothesis testing, evaluation of the confidence intervals and the upper limits*, while the second one refers to the *goodness of the fit of the observed data to the known Standard Model (SM) prediction.* All these questions are answered in the context of the frequentist statistical test, where the SM-only, or *background-only*, hypothesis is tested against the alternative *signal-plus-background* $(s + b)$ hypothesis, which assumes additional signal from the new physics on the top of the SM background. This chapter is aimed to discuss statistical test procedures used in this thesis and address both questions in terms of these models.

9.1 Formalism of binned profile likelihood

Consider an experiment, where independent and identical observations of a certain variable x, e.g. the mass of the $W/Z/h$ candidate, are repeatedly performed, so that the set of measured values is represented as the histogram of the distribution of a given observable. Then the expected event yield n_i in the ith bin is expressed as:

$$n_i^{\text{exp}} = \mu s_i + b_i \,, \tag{9.1}$$

where s_i, b_i correspond to the mean number of the signal and background events in ith bin respectively, μ corresponds to the strength of the signal process, or *signal strength*, which is commonly defined as the ratio of cross-section of the a given signal model over some reference signal cross-section, $\mu = \sigma_{\text{sig}}/\sigma_{\text{ref}}$, which is established to be 10 fb in case of $E_{\text{T}}^{\text{miss}} + V(qq)$ and $E_{\text{T}}^{\text{miss}} + h(b\bar{b})$ searches. Here the background-only hypothesis b corresponds to $\mu = 0$, while the signal-plus-background hypothesis $s + b$ to $\mu = 1$.
The number of the signal and background events can be expressed using the probability density functions (PDFs) of the measured variable x:

$$\begin{aligned} s_i &= s_{\text{tot}} \int\limits_{i\text{th bin}} f_s(x|\boldsymbol{\theta}_s)dx \\ b_i &= b_{\text{tot}} \int\limits_{i\text{th bin}} f_b(x|\boldsymbol{\theta}_b)dx \,. \end{aligned} \tag{9.2}$$

Here it is assumed that the variable x is distributed according to the PDFs $f_s(x|\boldsymbol{\theta}_s)$ and $f_b(x|\boldsymbol{\theta}_b)$ for the number of the signal and background events respectively, where the conditional parameters $\boldsymbol{\theta}_s$, $\boldsymbol{\theta}_b$ define the shape of the PDFs. The parameters s_{tot}, b_{tot} are the total number of signal and background events respectively.
The set of the parameters $\boldsymbol{\theta} = (\boldsymbol{\theta}_s, \boldsymbol{\theta}_b, b_{\text{tot}})$ are usually referred to as the *nuisance parameters* (NPs). Since the event yields are always measured within uncertainties, the NPs account for all possible sources of so-called *systematic uncertanties*. Most of the NPs are determined from additional auxiliary measurements, and some NPs like normalisations of main backgrounds are the free parameters. The remaining NPs are artificially constrained using either Gaussian, Poisson or log-normal distributions. A good example of NPs well-constrained from auxiliary measurements is NPs related to the jet energy scale (JES) components, which will be discussed in Section 6.2.3. In this case, the p_T-dependent JES correction factor is precisely measured with very small total uncertanty, as shown in Figure 9.1 for one representative NP related to JES calibration of the small$-R$ jets.

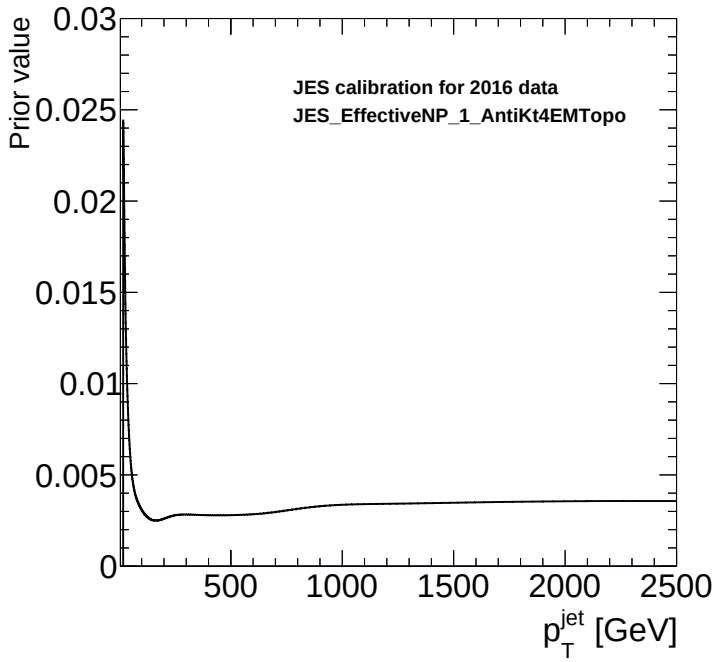


Figure 9.1: Prior value of one representative NP related to the JES calibration of the small$-R$ jets as a function of jet p_T [200].

In the end, the statistical model can be fully parameterised by a set of NPs and the expected signal strength μ, so that one can construct the characteristic function for this model, given the observed data n^{obs}, or the *binned profile likelihood function*:

$$\mathcal{L}(n^{\text{obs}}|\mu, \boldsymbol{\theta}) = \prod_{i=1}^{N} \frac{[\mu s_i(\boldsymbol{\theta}) + b_i(\boldsymbol{\theta})]^{n_i^{\text{obs}}}}{n_i^{\text{obs}}!} e^{-[\mu s_i(\boldsymbol{\theta}) + b_i(\boldsymbol{\theta})]} \prod_{\theta_i \in \boldsymbol{\theta}} f(\theta_i|\theta_i^0, \sigma_{\theta_i}) \,. \tag{9.3}$$

Here the expectation values n_i^{exp} follow the Poisson distribution, $\text{Pois}(n_i^{\text{exp}}|n_i^{\text{obs}}) = \frac{(n_i^{\text{exp}})^{n_i^{\text{obs}}}}{n_i^{\text{obs}!}} e^{-n_i^{\text{exp}}}$. The function $f(\theta_i|\theta_i^0, \sigma_{\theta_i})$ is the distribution of the NPs θ_i, given its estimate θ_i^0 and the standard error σ_{θ_i}. Generally, this likelihood describes how *plausible* is the considered parametrised model, given the observed data.

Often besides the main measurement of the observable x, one needs to perform simultaneous additional measurements in several disjoint regions of the data. These measurements are mostly aimed to better constrain NPs beyond what was done by the auxilary measurements, or to distinguish between different regions with different sensitivies to the signal and background processes. Such measurements imply selecting dedicated signal and control regions (sidebands), enriched with the certain type of events, and selecting the kinematic variable of interest.
One example in the scope of the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ searches is $Z \to \nu\nu$ irreducible background, which normalisation is contrained using 2-lepton control region from the leptonic Z decay, $Z \to \ell\ell$, due to the similar distributions of the kinematic variables. Assume that $n_{Z\to\ell\ell}$ events are observed in the $Z \to \ell\ell$ control region with the mean, which is known to be τ times that of the mean $\mu_{Z\to\nu\nu}$ for the $Z \to \nu\nu$ background in the signal region. In this case one of the constraining terms in the likelihood in Eq. 9.4 takes the following form:

$$\mathcal{L}(n_{Z\to ll}|\mu_{Z\to\nu\nu}, \boldsymbol{\theta}') = \frac{[\tau\mu_{Z\to\nu\nu}(\boldsymbol{\theta}')]^{n_{Z\to ll}}}{n_{Z\to ll}!} e^{-\tau\mu_{Z\to ll}(\boldsymbol{\theta}')} \times G(\tau|\mu_\tau, \sigma_\tau), \tag{9.4}$$

where an additional term $G(\tau|\mu_\tau, \sigma_\tau)$ constrains the parameter τ.

9.1.1 Hypothesis testing and test statistic

When address the question about the discovery of a new signal, one tests two different hypotheses against each other: the main *null hypothesis* H_0, which usually assumes SM background only with $\mu = 0$, against some alternative hypothesis H_1, which usually assumes some signal on a top of the background with $\mu = 1$. After the measurement is performed, the decision about rejecting or accepting the H_0 hypothesis should be made. This decision is based on so-called *test statistic* t, which is a data-based quantity used to estimate how probable is the measured data assuming a given hypothesis. If t is observed in the acceptance region decided upon before the fit, then the H_0 hypothesis is rejected. The corresponding probability α for t to be in this region is called the *significance level* of the test.
The Neyman-Pearson lemma [201] states that the acceptance region with the highest statistical power for a significance level α is defined as:

$$\frac{\mathcal{L}(t|H_0)}{\mathcal{L}(t|H_1)} > c(\alpha), \tag{9.5}$$

where $c(\alpha)$ is the significance-level based constant. Thus, the Neyman-Pearson lemma defines the most powerful test statistics as the *log-likelihood ratio* $\mathcal{L}(n^{\mathrm{obs}}|H_0)/\mathcal{L}(n^{\mathrm{obs}}|H_1)$. Since the model in Eq. 9.4 contains multiple parameters of interest, *profile likelihood ratio* based on the maximum likelihood method is used as the test statistic [202]:

$$\lambda(\mu) = \frac{\mathcal{L}(n^{\mathrm{obs}}|\mu, \hat{\hat{\boldsymbol{\theta}}}(\mu))}{\mathcal{L}(n^{\mathrm{obs}}|\hat{\mu}, \hat{\boldsymbol{\theta}})}, \tag{9.6}$$

Here $\hat{\hat{\boldsymbol{\theta}}}$ denotes so-called *conditional maximum-likelihood* estimator (MLE) of $\boldsymbol{\theta}$ that maximises the likelihood for a fixed value of μ. $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ denote the *unconditional* MLEs of μ and $\boldsymbol{\theta}$ parameters (so-called *true* parameters) respectively. More about the ML method

can be found in [202]. It is important to note that in general the free NPs broaden the $\lambda(\mu)$ as a function of μ comparing to those which are fixed. It means that each free NP reduces the information about μ.

Since $0 \leq \lambda(\mu) \leq 1$, where larger values of λ stand for a better agreement between the observed data and the expected value of μ, it is more convenient to use the logarithm of the test statistic:

$$t_\mu = -2 \ln \lambda(\mu), \tag{9.7}$$

Here, the lower values of t_μ mean the higher level of compatibility between the observed data and μ, meaning that t_μ can be used as a direct figure of merit of the compatibility between the data and a given hypothesis. Thus, instead of choosing to accept or reject the hypothesis, one can calculate so-called $p-value$, which is the probability to find data as compatible or less with the H_0 hypothesis, given the observed value of the test statistic t_μ^{obs}:

$$p_0 = \int_{t_\mu^{\text{obs}}}^{\infty} f(t_\mu|H_0) dt_\mu, \tag{9.8}$$

where $f(t_\mu|H_0)$ is the distribution of t_μ under the H_0 hypothesis. Another equivalent measure of compatibility is the significance Z, defined in the way that the upper-tail probability of a Gaussian distributed variable, located Z standard deviations from its mean, is equal to the $p-$value:

$$Z = \Phi^{-1}(1 - p_0), \tag{9.9}$$

where Φ^{-1} is the inverse of the cumulative distribution of the Gaussian. The H_0 is rejected, when $p_0 \leq \alpha = 2.87 \cdot 10^{-7}$ ($Z \geq 5$). On the other hand, H_0 hypothesis is accepted at *95% confidence level* (CL) if $p_0 < \alpha = 0.05$ ($Z \leq 1.64$). The definitions of the $p-$value and the significance Z are illustrated in Figure 9.2.

Similarly to p_0, the probability to find data as compatible with the H_1 hypothesis or less, is defined as:

$$p_1 = \int_{-\infty}^{t_\mu^{\text{obs}}} f(t_\mu|H_1) dt_\mu, \tag{9.10}$$

Since the presence of a signal is usually expected to appear as an increase in the event yield expected from the background hypothesis, the signal strength is expected be non-negative. Thus, if the observed data leads to the negative estimator $\hat{\mu} < 0$, the best agreement between the data and the signal hypothesis always corresponds to $\mu = 0$. Then the test statistic takes the form:

$$t_\mu = \begin{cases} -2 \ln \lambda(\mu), & \hat{\mu} \geq 0 \\ -2 \ln \left(\frac{\mathcal{L}(n^{\text{obs}}|\mu, \hat{\hat{\boldsymbol{\theta}}})}{\mathcal{L}(n^{\text{obs}}|0, \hat{\hat{\boldsymbol{\theta}}}(0))} \right) & \hat{\mu} < 0 \end{cases}, \tag{9.11}$$

In what follows, the test statistic is used under the assumption that the new physics

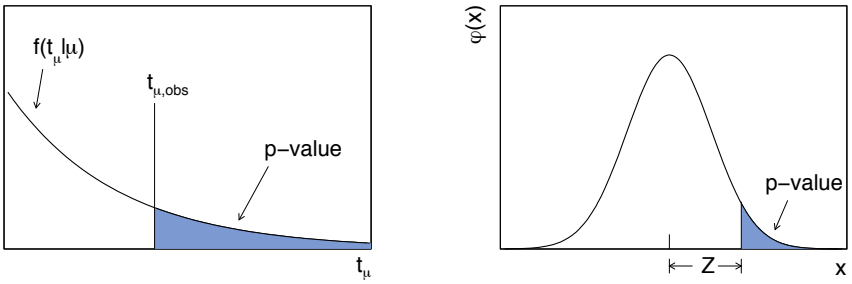


Figure 9.2: Illustration of the definitions of the $p-$value and the significance Z [203]. Left plot shows the relation between t_μ and $p-$value. Right plot show the relation between $p-$value and the significance Z.

leads to an increased event yield. Let us consider two important cases of the test statistic t_μ, the test of the null hypothesis and setting upper limit setting under the alternative hypothesis.

Test statistic under the background-only hypothesis

Consider the class of the models with expected $\mu \geq 0$, which are tested under the background-only (b) hypothesis $\mu = 0$. If the b hypothesis is rejected, it effectively means a signal discovery. The test statistic in this case takes the form:

$$q_0 \equiv t_0 = \begin{cases} -2\ln\lambda(0), & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases}. \tag{9.12}$$

The $\hat{\mu} < 0$ condition rejects the b hypothesis, but due to the decrease in the event yield it rather points to some systematic error than to the presence of the signal. The $p-$value indicates the level of the disagreement between the data and the b hypothesis and is defined from the Eq. 9.8 as:

$$p_0 = \int_{q_0^{\text{obs}}}^{\infty} f(q_0|\mu = 0)dq_0, \tag{9.13}$$

where $f(q_0|\mu = 0)$ denotes the PDF of the test statistic q_0 under the b hypothesis.

Test statistic in case of upper limit setting

Assume that now one wants to set an upper limit on the signal strength μ for a certain signal model under the $s + b$ hypothesis. The corresponding test statistic then takes the form:

$$q_\mu = \begin{cases} -2\ln\lambda(\mu), & \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases}, \tag{9.14}$$

Here, since the μ is the upper limit of an estimator $\hat{\mu}$, the observed data with $\hat{\mu} > \mu$ is not taken into account. This makes it impossible to reduce this defition to the definition of the test statistic in Eq. 9.12 by setting $\mu = 0$. Here the $p-$value indicates the level of

the agreement between the data and the $s + b$ hypothesis, defined identically to Eq. 9.8.

9.1.2 Approximation of test statistic in large sample limit

To calculate the $p-$value for either b or $s + b$ hypothesis, one needs to obtain the corresponding distribution of the test statistic, $f(q_0|\mu = 0)$ and $f(q_\mu|\mu)$. Moreover if the data leads to a different strength parameter than being tested, $\mu 0 \neq \mu$, the corresponding $p-$value and the expected significance are calculated based on the different distribution of the test statistics, $f(q_\mu|\mu_0)$.
Since the statistical model is highly compex due to the large number of parameters with the non-linear dependencies, the exact computation of the above-mentioned distributions for the test statistics become non-triviala and extremely resource-intensive, especially for a large size of the data sample. Thus, in the limit of large sample size, one can use the asymptotic expressions of the test statistic and the correspoding distrubutions, determined by Wilks [204] and Wald [205].
Consider the test of the strength parameter μ under two hypotheses, b with $\mu = 0$ and $s + b$ with $\mu = 1$. Assume that the data correspond to a signal strength μ_0. Then, according to Wald [205], for the case of a single parameter of interest μ in the limit of the large sample size N, the test statistic can be approximated by the following expression:

$$t_\mu = -2 \ln \lambda(\mu) = \frac{(\mu - \hat{\mu})^2}{\sigma^2} + \mathcal{O}(1/\sqrt{N}), \tag{9.15}$$

where $\hat{\mu}$ is Gaussian distributed with the mean μ_0 and the standard deviation σ. Here σ is calculated from the covariance matrix of the estimators of all parameters of the model, $V_{ij} = \text{cov}\left[\hat{\theta}_i, \hat{\theta}_j\right]$, where $\theta_i, \theta_j \in (\mu, \boldsymbol{\theta})$. According to Wilks [204], the test statistic in the form of Eq. 9.15 leads to the PDF of the test stastics for a single parameter μ, which follows the *non-central* χ^2 distribution:

$$f(t_\mu|\Lambda) = \frac{1}{2\sqrt{2\pi t_\mu}} \left[e^{-\frac{1}{2}\left(\sqrt{t_\mu}+\sqrt{\Lambda}\right)^2} + e^{-\frac{1}{2}\left(\sqrt{t_\mu}-\sqrt{\Lambda}\right)^2} \right], \tag{9.16}$$

where $\Lambda = \frac{(\mu-\mu_0)^2}{\sigma^2}$ is the non-centrality parameter. In case of the multiple parameters of interest, one can generalise Eq. 9.15 and 9.16 by taking:

$$\begin{aligned} t_\mu &= (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T V^{-1} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \\ \Lambda &= (\boldsymbol{\theta}_0 - \boldsymbol{\theta})^T V^{-1} (\boldsymbol{\theta}_0 - \boldsymbol{\theta}), \end{aligned} \tag{9.17}$$

where $\hat{\boldsymbol{\theta}}$ are the MLEs of the NPs $\boldsymbol{\theta}$, which are Gaussian distributed with the mean θ_0 along with the inverse covariance matrix $V_{ij}^{-1} = E\left[\frac{\partial^2 \ln \mathcal{L}}{\partial\theta_i \partial\theta_j}\right]$, where E denotes the mathematical expectation.
The search for the new physics often results in a fairly small data sample after requiring all selection criteria. The asymptotic approximation in the large sample limit appears to provide accurate results even in case of small samples. For a very small samples one can always use the Monte Carlo samples to obtain the required distributions.
Eq. 9.17 needs to be solved to estimate the median value of the test statistic t_μ and the covariance matrix V. One can consider the special data set, called the *Asimov data set* [203], where all statistical fluctuations are neglected and all MLEs are assumed to be

the true values of the corresponding parameters, $\hat{\theta}_i = \theta_{0,i}$. Then for the Asimov data set, the likelihood ratio, the test statistic and the standard deviation from Eq. 9.15 take the form:

$$\begin{aligned} \lambda_{\mathrm{A}}(\mu) &= \frac{\mathcal{L}_{\mathrm{A}}(n^{\mathrm{obs}}|\mu, \hat{\hat{\boldsymbol{\theta}}}(\mu))}{\mathcal{L}_{\mathrm{A}}(n^{\mathrm{obs}}|\mu_0, \boldsymbol{\theta}_0)} \\ t_\mu^{\mathrm{A}} &= (\boldsymbol{\theta}_0 - \boldsymbol{\theta})^T V^{-1} (\boldsymbol{\theta}_0 - \boldsymbol{\theta}) = \Lambda \\ \sigma_A^2 &= \frac{(\mu - \mu_0)^2}{t_\mu^{\mathrm{A}}}. \end{aligned} \tag{9.18}$$

The important case of using the Asimov data is when one wants to find the median exclusion significance for the hypothesis $\mu > 0$, given that the data suggests the case of no signal present, $\mu_0 = 0$. In this case then:

$$t_0^{\mathrm{A}} = \frac{\mu^2}{\sigma_A^2}, \qquad Z_{\mathrm{median}} = \sqrt{t_0^{\mathrm{A}}}. \tag{9.19}$$

9.1.3 CL_s test

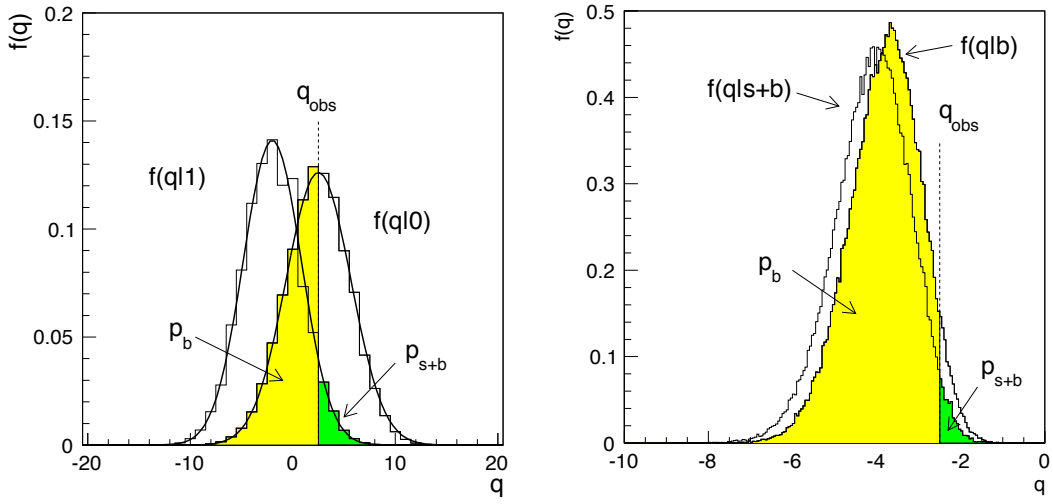
Assume that the test statistic q_μ is constructed in a way to distinguish between the $s+b$ and the b hypotheses. Then, according to the Eq. 9.8 and Eq. 9.10, the corresponding $p-$values for each hypothesis, given the observed value of test statistic q_μ^{obs}, are the following:

$$\begin{aligned} p_{s+b} &= \int_{q_\mu^{\mathrm{obs}}}^{\infty} f(q_\mu|s+b) dq_\mu \\ p_b &= \int_{-\infty}^{q_\mu^{\mathrm{obs}}} f(q_\mu|b) dq_\mu, \end{aligned} \tag{9.20}$$

where example distributions $f(q_\mu|s+b)$ and $f(q_\mu|b)$ are shown in Figure 9.3a. The standard CL_{s+b} test of the $s+b$ hypothesis states that the signal model is excluded at the $1-\alpha = 95\%$ CL if $p_{s+b} < \alpha = 0.05$.
The problems with the CL_{s+b} method appear, when the search probes the signal models to which it has no or very low sensitivity. This corresponds to the case, when the expected number of signal events is much lower that of the background, and the distributions of the test statistic strongly overlap, as shown in Figure 9.3b. Then if the observed data reveals a downward fluctuation with the respect to the b hypothesis, the hypothesis will be rejected. But in the case of low sensitivity one expects to have the rejection probability close to 0. This is achieved by introduces the penalty term $1-p_b$ to the CL_{s+b} test, which accounts for the downward fluctuations that can make data appear not background-like [206]:

$$\mathrm{CL}_s \equiv \frac{p_{s+b}}{1-p_b}. \tag{9.21}$$

In case of widely separated distributions $p_b \ll 1$ and the penalty is small. In opposite case, the penalty term $1-p_b$ decreases, significantly increasing overall $p-$value of the $s+b$ rejection and making it more difficult to reject H_1 hypothesis that one in fact is



(a) (b)

Figure 9.3: Distributions of the test statistic under the $s+b$ and b hypotheses for the different sensitivities to the signal model [206]. Figure (a) corresponds to the high sensitivity case. Figure (b) corresponds to the low sensitivity case. Notations $f(q|0)$ and $f(q|1)$ in Figure (a) correspond to notations $f(q|b)$ and $f(q|s+b)$ in Figure (b), respectively.

not sensitive to. Since the CL_s $p-$value is always larger that of the CL_{s+b}, the CL_s test excludes less signal models, and the corrresponding upper limits are weaker. Thus, CL_s method is conservative.

Chapter 10

Systematic Uncertainties

The $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analyses are sensitive to the different biases coming from the experimental measurements and Monte Carlo modeling of the signal and background processes. These biases are described in terms of the systematic uncertainties. The systematic uncertainties, referred to as *experimental* uncertainties, represent a variety of experimental effects related to the reconstruction, identification and calibration of objects used in $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analyses. The systematic uncertainties, referred to as *theoretical* uncertainties, stands for the theoretical aspects of the signal and background modelling. These uncertainties contribute both to the overall event yield and shape of distributions of the key variables, such as the mass of the Higgs boson candidate in the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis and $E_{\mathrm{T}}^{\mathrm{miss}}$ in the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis. The former referrs to as *normalisation* uncertainties, while the latter to as *shape* uncertainties.
This chapter provides a general description of all sources of the systematic uncertainties considered in both analyses. The experimental and theoretical uncertanties are discussed in Chapters 10.1 and 10.2, respectively. The resulting performance and impact of systematic uncertainties on the search results are discussed in Chapters 12.1.1 and 12.1.2.

10.1 Experimental systematic uncertainties

Experimental systematic uncertainties are estimated using auxiliary measurements performed by the combined performance (CP) groups for all reconstructed objects. A concised summary of all experimental uncertainties is given in Table 10.1.

10.1.1 Luminosity

The measurement of the total integrated luminosity is used to normalise the Monte Carlo (MC) event yields to those measured in data. An uncertainty of 2.1% and 3.4% on the total integrated luminosity is applied to the MC event yields in the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis, respectively.

10.1.2 $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger uncertainty

The $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger uncertainty consists of two components. The statistical component, denoted as *METTrigStat*, is taken as the 1σ uncertainty on the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger scale factor fit. The second component, denoted as *METTrigSyst*, is estimated as the largest variation,

which corresponds to the dependence of the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency on the event topology. A detailed information about the corresponding uncertainties is given in Chapter 5.2.4.

10.1.3 Small$-R$ jet and track jet uncertainties

The uncertainties related to the reconstruction, calibration and $b-$tagging of the small$-R$ jets are one of the dominant in the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis.
Jet energy calibration uncertainties in the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analyses, which include jet energy scale (JES) uncertainties, are sub-dominant comparing to other uncertainties. Therefore, to reduce the complexity of the statistical model, a strongly reduced set of four JES nuisance parameters, which represents all JES uncertainties, is introduced [200]. One NP stands for the non-closure uncertanty on η intercalibration, while the remaining JES uncertainties are combined into three NPs. These NPs represent low-p_T regime, where JES uncertainty is large due to the pile-up, medium-p_T regime with small JET uncertainty and high-p_T regime, where JES uncertainty is very large due to the lack of statistics [1], as illustrated in Figure 6.8 in Chapter 6.2.3. A single NP represents a jet energy resolution (JER) uncertainty [207]. The corresponding analysis tools to derive and implement JES and JER calibration and uncertanties are provided by the JetEtmiss CP group [208].
Another class of the jet uncertanties is related to the $b-$tagging of the small$-R$ and track jets. As will be shown in Chapter 12.1.2, the impact of these uncertainties is dominant comparing to other small$-R$ jet uncertainties and one of the dominant comparing to all other uncertainties in the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis. The uncertainty on the $b-$tagging efficiency is derived separately for small$-R$ jets in the resolved regime and for track-jet ghost-associated to the large$-R$ jet in the merged regime. Each uncertainty is represented by the three NPs, which account for uncertainties on the efficiency scale factors for $b-$tagged, $c-$tagged and light-tagged jets [179, 180]. The corresponding analysis tools to derive and implement the $b-$tagging scale factors and uncertanties are provided by the Flavour Tagging CP group [209].

10.1.4 Large$-R$ jet uncertainties

The uncertainties related to the calibration and kinematics of the large$-R$ jets are treated differently in the $E_{\mathrm{T}}^{\mathrm{miss}}+V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ analyses according to their impact on the total systematic uncertainty. Since the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis uses the priority-merged selection, as described in Chapter 7.1, the corresponding large$-R$ jet uncertainties dominate comparing to other systematic uncertainties. A single group of the three NPs represents the uncertainties related to the energy, mass and substructure variable $D_2^{\beta=1}$ resolutions of jets reconstructed using the anti-k_T algorithm with the $R = 1.0$ radius parameter [210]. Other three groups of four NPs each represent the uncertainties on the combined mass, substructure variable $D_2^{\beta=1}$ and p_T of the large$-R$ jets. These four NPs stand for the difference between the data and MC simulation (*Baseline*), the fragmentation modelling (*Modelling*), the tracking reconstruction efficiency, fake rate and bias in the q/p_T distribution (*Tracking*), and the total statistical uncertainty of the measurement of a given kinematic variable (*TotalStat*) [210]. The jet mass, p_T and substructure vari-

[1] This is illustrated by the impact of the small$-R$ jet uncertainty, which grows to 10% for the models with hard $E_{\mathrm{T}}^{\mathrm{miss}}$ spectrum (see Chapter 12.1.2) and become one of the donimant uncertainties.

able scales are considered as uncorrelated. The JES uncertainties are treated indentically to the small$-R$ jets.
Since no priority-merged selection is used in the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis, large$-R$ uncertainties become sub-dominant comparing to other uncertainties, as will be shown in Chapter 12.1.2. Thus, a simplified set of four NPs, which stands for the combined baseline, modelling, tracking and total statistical uncertainties on the jet mass and p_T, is implemented.
The corresponding analysis tools to derive and implement large$-R$ uncertanties are provided by the JetEtmiss CP group [208].

10.1.5 $E_{\mathrm{T}}^{\mathrm{miss}}$ uncertainties

As discussed in Chapter 6.4.1, $E_{\mathrm{T}}^{\mathrm{miss}}$ is comprised of the hard term, which incorporates the contributions from the reconstructed objects, and the soft term, which incorporates all tracks not associated to any reconstructed object. The uncertainties on the former are estimated by propagating the uncertainties related to the contributing reconstructed objects. The uncertainties related to the $E_{\mathrm{T}}^{\mathrm{miss}}$ soft term are estimated using the analysis tools provided by the JetEtmiss CP group [211]. They are not included in some of the MC samples used in the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis and thus are neglected for the current version of the analysis. For the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis, four independent NPs represent the $E_{\mathrm{T}}^{\mathrm{miss}}$ soft term uncertainty. The two of them are related to the resolution uncertainties in the projection of the $E_{\mathrm{T}}^{\mathrm{miss}}$ track soft term perpendicular and parallel to the hadronic recoil system. The remaining NPs stands for the scale uncertainties on the $E_{\mathrm{T}}^{\mathrm{miss}}$ track soft and jet terms.

10.1.6 Electron uncertainties

Besides the uncertainties on the energy scale and resolution of the reconstructed electrons, the uncertanties on the correction scale factors for the electron reconstruction, identification, isolation and trigger efficiencies are implemented in both analyses. The correction scale factors are calculated as $\mathrm{SF} \equiv \varepsilon^{\mathrm{Data}}/\varepsilon^{\mathrm{MC}}$, where ε represents the efficiency, in the 2ℓCR in bins of p_T and η using $W \to e\nu$, $Z \to ee$ and $J/\psi \to ee$ samples. The analysis tools to calculate and implement the correction scale factors and the corresponding uncertainties are provided by the Electron CP group [212, 213].

10.1.7 Muon uncertainties

Additionally to the uncertainties on the p_T scale and resolution of the reconstructed muons, the uncertanties on the correction scale factors for the muon efficiencies similar to that for the electrons are introduced in both analyses. The scale factors are introduced for the muon reconstruction, identification, isolation, track-to-vertex and trigger efficiencies. The systematic uncertainties on the correction scale factors are derived in the 2ℓCR using $Z \to \mu\mu$ and $J/\psi \to \mu\mu$ samples. The corresponding analysis tools are provided by the Muon CP group [214, 215].

Systematic uncertainty (if different name in $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis)	Description	Analysis
Data		
Luminosity	Total integrated luminosity	both
Large−R jets		
JET_Comb_Baseline_Kin	Jet energy scale (jet mass and p_T scales fully correlated)	$E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$
JET_Comb_Modelling_Kin		$E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$
JET_Comb_TotalStat_Kin		$E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$
JET_Comb_Tracking_Kin		$E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$
FATJET_JER	Jet energy resolution	$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$
FATJET_JMR	Jet mass resolution	$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$
FATJET_D2R	Substructure variable $D_2^{\beta=1}$ resolution	$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$
FATJET_Weak_Comb_Baseline_mass	Jet mass scale (jet mass, p_T and $D_2^{\beta=1}/\tau_{32}$ scales uncorrelated)	$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$
FATJET_Weak_Comb_Modelling_mass		$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$
FATJET_Weak_Comb_Tracking_mass		$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$
FATJET_Weak_Comb_TotalStat_mass		$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$
FATJET_Weak_Rtrk_Baseline_D2	Substructure variable $D_2^{\beta=1}$ scale (jet mass, p_T and $D_2^{\beta=1}/\tau_{32}$ scales uncorrelated)	$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$
FATJET_Weak_Rtrk_Modelling_D2		$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$
FATJET_Weak_Rtrk_Tracking_D2		$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$
FATJET_Weak_Rtrk_TotalStat_D2		$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$
FATJET_Weak_Rtrk_Baseline_pT	Jet p_T scale (jet mass, p_T and $D_2^{\beta=1}/\tau_{32}$ scales uncorrelated)	$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$
FATJET_Weak_Rtrk_Modelling_pT		$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$
FATJET_Weak_Rtrk_Tracking_pT		$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$
FATJET_Weak_Rtrk_TotalStat_pT		$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$
Small−R jets		
JET_SR1_GroupedNP_i	Jet energy scale (split into 3 components)	both
JET_SR1_JET_EtaIntercalibration_NonClosure	Non-closure in jet response for $2.4 < \|\eta\| < 2.5$	both
JET_JER_SINGLE_NP	Jet energy resolution	both
FT_EFF_Eigen_Light_i (FT_EFF_Eigen_L_i)	$b-$tagging efficiency for small$-R$ jets $(i, j, k = 0, 1, 2, 3)$	both
FT_EFF_Eigen_C_j		both
FT_EFF_Eigen_B_k		both
FT_EFF_extrapolation	$b-$tagging efficiency extrapolated to high jet p_T	both
FT_EFF_extrapolation_from_charm		both
Track jets		
FT_EFF_Eigen_Light_i (FT_EFF_Eigen_L_i)	$b-$tagging efficiency for small$-R$ jets $(i, j, k = 0, 1, 2, 3)$	both
FT_EFF_Eigen_C_j		both
FT_EFF_Eigen_B_k		both
FT_EFF_extrapolation	$b-$tagging efficiency extrapolated to high jet p_T	both
FT_EFF_extrapolation_from_charm		both
Electrons		
EG_SCALE_ALL	Energy scale	both
EG_RESOLUTION_ALL	Energy resolution	both
EL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR	Identification efficiency	both
EL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR	Reconstruction efficiency	both
EL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR	Isolation efficiency	both
Muons		
MUON_MS	Muon p_T resolution from the muon system	both
MUON_ID	Muon p_T resolution from the inner detector	both
MUON_SCALE	Muon p_T scale	both
MUON_ISO_SYS	Isolation effiecincy	both
MUON_ISO_STAT		both
MUON_EFF_SYS	Reconstruction and identification efficiency ($p_T > 15$ GeV)	both
MUON_EFF_STAT		both
MUON_EFF_SYS_LOWPT	Reconstruction and identification efficiency ($p_T < 15$ GeV)	both
MUON_EFF_STAT_LOWPT		both
MUON_TTVA_STAT	Track-to-vertex association	both
MUON_TTVA_SYS		both
Missing transverse momentum		
MET_JetTrk_Scale		both
MET_SoftTrk_ResoPerp		both
MET_SoftTrk_ResoPara		both
MET_SoftTrk_Scale		both
Trigger efficiency		
METTrigStat	$E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency	both
METTrigSyst		both
EL_EFF_Trigger_TOTAL_1NPCOR_PLUS_UNCOR	Electron trigger efficiency	both
MUON_EFF_TrigSystUncertainty	Muon trigger efficiency	both
MUON_EFF_TrigStatUncertainty		both

Table 10.1: Summary of the experimental systematic uncertainties considered in the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analyses.

10.2 Theoretical systematic uncertainties

10.2.1 Background shape uncertainties

The background shape uncertainties account for the different modelling parameters in the MC simulation related to the matrix element calculation, fragmenation and hadronisation processes, parton shower evolution, initial and final state radiation, etc. This class of uncertanties is assigned to the main background procceses, such as $W+$jets, $Z+$jets, $t\bar{t}$ and single-top. Both analyses incorporate the background modelling studies, as well as estimation of the corresponding systematic uncertainties, performed for the $Vh(b\bar{b})$ analysis [216, 217]. The uncertainties are provided on the modelling of the p_T distribution of the vectors bosons (p_T^V) and on the mass of the W/Z and h boson candidates (m_{jj}/m_J). The list of the background shape uncertainties is given in Table 10.2.

$V+$jets modelling uncertainties

The uncertainties on the $p_T^V = p_T(E_{\mathrm{T}}^{\mathrm{miss}}, l)$ ($p_T^V = p_T(l,l)$) and m_{jj}/m_J are derived in the 1μCR (2ℓCR), where the latter is enriched with $Z+$jets events at high purity. An estimation procedure relies on the paricle-level comparisons of the distributions with varied scales normalised to unity area to isolate the shape information in the following samples:

- Nominal Sherpa 2.2.1 MC samples, which are default for the $V+$jets processes (see Section 5.1), to the Sherpa 2.1 MC samples to assess the acceptance uncertainties.
- Nominal Sherpa 2.2.1 MC samples to the MadGraph + Pythia MC samples.
- Nominal Sherpa 2.2.1 MC samples to the data in the high purity CR for $Z+$jets.

The largest variation relative to the nominal distribution is then parametrised by an analytical function, symmetrised to get up and down variations and taken as $\pm 1\sigma$ modelling uncertainty on the shape of a given distribution.
For the $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ analysis, the $V+$jets modelling uncertainties are considered independent of the jet flavour content, meaning that a single NP represents a given uncertainty. However, for the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis each $V+$jets modelling uncertainty is described by the six independent NPs, which represent the $V+ll$, $V+cl$ and $V+bb$ components. The motivation for this flavour splitting is driven by the fact that the $V+$jets flavour components can have different p_T^V and m_{jj} distributions due to the different contributing diagrams and corresponding partonic luminosities [2]. The details about the flavour splitting and its impact on the final search results are discussed in Ref. [218] and **[Suchek2017:EB]**.

$t\bar{t}$ and single-top modelling uncertainties

The uncertainties on the $p_T^V \equiv p_T(E_{\mathrm{T}}^{\mathrm{miss}}, l)$ and m_{jj}/m_J of the hadronically decaying W boson candidate are derived in the 1μCR. The particle-level comparisons of the distributions, similar to those for $V+$jets, are used to estimate the $t\bar{t}$ and single top modelling

[2] E.g., the dominant process that contributes to the $V+bb$ production is the gluon splitting, $g \to b\bar{b}$, while the dominant contributing diagrams to the $V+ll$ production are $q \to qg$ and $g \to q\bar{q}$.

uncertainties. To decouple impacts of the matrix element calculation, fragmentation modelling and production of radiation, the following distributions are compared:

- $t\bar{t}$: nominal POWHEG + PYTHIA 8 to the POWHEG + PYTHIA 8 with varied level of radiation, POWHEG + HERWIG 7 and MADGRAPH + PYTHIA 8 [216, 217].
- single-top: nominal POWHEG + PYTHIA 6 to the POWHEG + PYTHIA 6 with varied level of radiation, POWHEG + HERWIG++ and MADGRAPH5_aMC@NLO + HERWIG++ [216, 217].

The largest variation relative to the nominal distribution is parametrised, symmetrised and taken as a $\pm 1\sigma$ modelling uncertainty on the shape of a given distribution.

Systematic uncertainty (if different name in $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis)	Description	Analysis
	Theory	
SysWMbb	The W modeling uncertainty on m_{jj}	$E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$
SysWPtV	The W modeling uncertainty on p_T^V	$E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$
SysZMbb	The W modeling uncertainty on m_{jj}	$E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$
SysZPtV	The Z modeling uncertainty on p_T^V	$E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$
SysMbbWShapeOnly_Wl	The $W + ll$ modeling uncertainty on m_{jj}	$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$
SysMbbWShapeOnly_Wcl	The $W + cl$ modeling uncertainty on m_{jj}	$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$
SysMbbWShapeOnly_Whf	The $W + bb$ modeling uncertainty on m_{jj}	$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$
SysWPtV_Wl	The $W + ll$ modeling uncertainty on p_T^V	$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$
SysWPtV_Wcl	The $W + cl$ modeling uncertainty on p_T^V	$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$
SysWPtV_Whf	The $W + bb$ modeling uncertainty on p_T^V	$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$
SysMbbZShapeOnly_Zl	The $Z + ll$ modeling uncertainty on m_{jj}	$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$
SysMbbZShapeOnly_Zcl	The $Z + cl$ modeling uncertainty on m_{jj}	$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$
SysMbbZShapeOnly_Zhf	The $Z + bb$ modeling uncertainty on m_{jj}	$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$
SysZPtV_Zl	The $Z + ll$ modeling uncertainty on p_T^V	$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$
SysZPtV_Zcl	The $Z + cl$ modeling uncertainty on p_T^V	$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$
SysZPtV_Zhf	The $Z + bb$ modeling uncertainty on p_T^V	$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$
SysTTbarMBB	The top modeling uncertainty on m_{jj}	both
SysTTbarPTV	The top modeling uncertainty on p_T^V	both
SysStoptMBB	The single-top modeling uncertainty on m_{jj}	both
SysStopWtMBB		both
SysStoptPTV	The top modeling uncertainty on m_{jj}	both
SysStopWtPTV		both
SysVVMbbME	The single-top modeling uncertainty on m_{jj}	$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$
SysVVMbbPSUE		$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$
SysVVPTVME	The diboson modeling uncertainty on p_T^V	$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$
SysVVPTVPSUE		$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$

Table 10.2: Summary of the background shape uncertainties for the W+jets, Z+jets, $t\bar{t}$, single-top and diboson processes considered in the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analyses. The V+jets shape uncertainties for the $E_{\mathrm{T}}^{\mathrm{miss}}+V(qq)$ analysis are split into three independent nuisance parameters according to the different jet flavour composition.

10.2.2 Background normalisation and acceptance uncertainties

Background normalisation uncertainties are implemented in terms of global normalisaiton parameters. Some of the background normalisation parameters are freely floating, meaning that the corresponding backgrounds can be constrained from the fit to the observed data without any prior knowledge. These include V+HF ($V + bb, bc, bl, cc$), $t\bar{t}$ for both analyses and V+LF jets for the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis, where HF and LF stands for the heavy and light flavour jets.
Other normalisation parameters are constrainted within uncertainties that are estimated from the theoretical expectations. The background normalisation parameters are allowed to vary within the following normalisation uncertainties [219, 220]:

- **V + ll**: 10% ($E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ only)
- **V**+ HF(**0**ℓ)CR/**V**+ HF(**2**ℓ)CR: 20% ($E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ only)
- **V + cl**: 30%
- **V + bc/V** + HF, **V + bl/V** + HF, **V + cc/V** + HF ratios: 20% (30%) each for the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ ($E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$) analysis.
- **Single-top production:** 4.4%, 4.6% and 6.2% for the $t-$channel, $s-$channel and Wt associated production.
- **Diboson production:** 25%, 26% and 20% for WW, WZ and ZZ processes.
- **V + ll, t$\bar{\text{t}}$ (resolved regime):** 20% each ($E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ only).
- **SM Vh production**: 50% uncertainty on the combined Wh and Zh normalisation ($E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ only).

10.2.3 Signal acceptance uncertainties

These uncertainties are related to the experimental acceptances of the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ signal models due to the modelling of the corresponding signal productions. The uncertainties are estimated by varying the modelling parameters of the MadGraph + Pythia simplified vector mediator and Z'-2HDM samples and comparing the results for a set of the mass points using the signal event selection reproduced at particle level. The following variations of the signal modelling parameters are considered:

- **PDF variations**: These are estimated by replacing the nominal NNPDF2.3LO PDF set with MSTW2008LO PDF and CTEQ6L1 PDF sets and considering the largest variation relative to the nominal PDF [221].
- **Scale variations:** These includes variations of the renormalisation and factorisation scales. The default values in MadGraph MC generator are dynamically set to $m_T^2 + p_T^2$, where m_T and p_T stand for the transverse mass and momentum of the final state particles, respectively. For the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ signal models, an additional initial state radiation scale is introduced. The scales are changed coherently up and down by a factor of 2 on event-by-event basis [222].
- **Tune variations:** These include the variations of the initial state radiation, final state radiaton and multi-parton interactions. These variations are parametrised by a set of the tune variations, which provide the maximal coverage for the effects of underlying events, jet substructure and additional jet production. [223].

Chapter 11

Analysis strategy

In this chapter several ways of presentation and interpretation of the final results of the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ searches are discussed. Generally speaking, there are a couple of approaches considered to search for a Dark Matter (DM) production at the LHC. One approach is to consider a well-motivated theoretical model, which predicts a sizeable amount of DM production for a given final state and is potentially available for observation at the LHC but which is not clearly excluded yet. In case of the excess in data over the SM prediction, indicating for the new physics, this model can be tested in terms of capability to describe a given excess. In case of no observational evidence of the DM production, the exclusion limits on the parameters of a given model can be derived. This is discussed in details in Section 11.1.
In case one wants to avoid any model assumptions in order to probe a variety of different models with a given signature, the generic limits on the production cross-section are derived. All details about the interactions between the SM and DM sectors are dropped and only the generic features of the $E_{\mathrm{T}}^{\mathrm{miss}} + X$ final state are exploited. This allows setting upper limits on the generic quantities, which do not require any knowledge about the specific model, such as the visible cross-section of X+DM events. On a top of this, such limits can be set both on the detector and parton (particle) levels, which makes it convenient for theorists to check the validity of their models for a given final state. All details are discussed in Section 11.2.
The general standards and recommendations for the benchmark models and corresponding interpretations for the DM searches at LHC come from the LHC Dark Matter Working Group [224]. It brings together theorists and experimentalists to discuss and develop well-motivated models of the DM production with a corresponding set of assumptions and requirements. The guidelines of this chapter are Ref. [73, 112, 225].

11.1 Model-dependent approach

When experimentalists talk about setting limits on a particular model, they mostly mean reducing the available phase-space of its parameters through considering experimental search results. Depending on the model of DM production, these parameters can be masses of the DM particles, mediators and their couplings to the SM and DM particles, other possible new particles with their own mixing angles and couplings. In this Section the interpretation of search results is discussed in terms of the simplified $s-$channel mediator models of $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ production and Z'-2HDM model of $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ production. These models are described in details in Section 3.4.

11.1.1 Choice of model parameters for simplified vector and axial-vector mediator models

The simplified $s-$channel mediator models has four degrees of freedom, which are the masses of the mediator m_{med} and DM particles m_χ, as well as the mediator couplings to the SM sector g_q and to the DM sector g_χ. The choice of the couplings for the DM simplified models for the LHC Run-2 is the same for all $E_{\text{T}}^{\text{miss}} + X$ and dijet searches. This was recommended by the LHC DM Working Group to provide a comparison between complementary exclusion limits across all DM searches performed at ATLAS and CMS for a given DM simplified model. The following baseline coupling scenarios are used to produce exclusion limits for the $E_{\text{T}}^{\text{miss}} + X$ and dijet searches depending on the type of the mediator:

- **Vector mediator (V1):** $g_q = 0.25,\ g_l = 0,\ g_\chi = 1.$
- **Axial-vector mediator (A1):** $g_q = 0.25,\ g_l = 0,\ g_\chi = 1.$

The choice of the quark coupling g_q is mainly motivated by the dijet constrains from the LHC and complementarity of the $E_{\text{T}}^{\text{miss}} + X$ and dijet searches. The universality of the vector and axial-vector mediator couplings to the quarks is discussed in Section 3.4.1.1. The lepton couplings g_l are set to zero to avoid any impact of dilepton searches in terms of these models. It would insignificantly increase the mediator width, while bringing in constraints from measurements of the Drell-Yan processes that would unnecessarily restrict the model space.
As more data is collected, the LHC searches for the DM become more sensitive to the models with small production cross-section. This is the motivation to consider other scenarios with smaller quark couplings and non-zero lepton couplings. Smaller quark couplings are used to check the overall impact of the couplings on the final exclusion limits due to the re-scaling of the cross-sections with the coupling strengths and possible changes in kinematics. Non-zero lepton couplings are used to study the impact of dilepton searches in these models.

- **Vector mediator (V2):** $g_q = 0.1,\ g_l = 0.01,\ g_\chi = 1$
  The lepton coupling is chosen to be small, $g_l \ll g_q$. This is the case when the vector mediator couples only to quarks and DM particles at tree level. The couplings to leptons are generated from the mixing between the SM Z boson and vector mediator Z' at loop-level [226]. The exact value $g_l = 0.1 g_q$ is chosen to reach comparable sensitivities of dijet and dilepton searches in terms of a given simplified model.

- **Axial-vector mediator (A2):** $g_q = 0.1,\ g_l = 0.1,\ \ g_\chi = 1$
  A representative scenario with equal quark and lepton couplings for the simplified models with axial-vector Z' boson [227]. It reveals the typical impact of searches for di-lepton resonances in these models. The size of the lepton coupling g_l is chosen to be large enough to obtain a sizeable the relative width $\Gamma_{\text{med}}/m_{\text{med}}$. This helps to assess the magnitude of the interference effects between the possible dilepton resonant signal and the SM Drell-Yan background.

Another important advantage of fixing the couplings of the $s-$channel mediator models is that the final LHC results can be compared to the corresponding limits from the direct and indirect measurements, as well as to the relic density calculations.

11.1.2 Choice of model parameters for Z'-2HDM model

Z'-2HDM model has in total five free parameters, which are the ratio of VEVs of two Higgs doublets $\tan\beta$, the mediator coupling g_z, the mass of the mediator m_{med}, the mass of the DM particles m_χ and the mass of the pseudoscalar scalar m_A. The masses of the heavy scalars H and $H^\pm$ are fixed to 300 GeV, and the aligment limit $\sin(\beta-\alpha)=1$ is considered from the constraints on the parameters of the Z'-2HDM model discussed in Section 3.4.1.2. Following the same constraints, the lower bound on the mass of the pseudoscalar scalar m_A is 300 GeV. The branching ratio of the $A^0 \to \chi\chi$ is taken as 100%.
Additional studies has been performed for the Run 1 $E_{\text{T}}^{\text{miss}} + h(b\bar{b})$ search [228] to reveal the impact of the free parameters on the kinematic distributions. In case of the on-shell production of pseudoscalar scalar A^0, $m_A > 2m_\chi$, variations of $\tan\beta$ and in the DM mass m_χ have no significant impact on the kinematic distributions. According to Eq. 3.22, the width of the $Z' \to hA^0$ decay is suppressed by a factor of $1/\tan^2\beta$ in the limit of large $\tan\beta$. For $\tan\beta < 1$ the decay width start to grow due to the increasing allowed limit on g_z from the electroweak constraints until $\tan\beta \approx 0.6$, where the dijet limits take over, as discussed in [113]. Similarly, changes in g_z do not significantly affect the kinematic distributions.
Given all above-mentioned, the parameters of the model are set to the values of $g_z = 0.8$, $\tan\beta = 1.0$ and $m_\chi = 100$ GeV. The choice of the DM mass m_χ guarantees the on-shell $A^0 \to \chi\chi$ decay.

11.1.3 Exclusion limit setting in mass-mass plane

The results of search for new physics can be interpreted in terms of the exclusion limits at 90% and 95% confidence level (CL) on the signal strength μ, defined a ratio of the cross-section measured in data to the theory cross-section, for a given model of new physics. Such limits are usually introduced in the phase-space of two model parameters, while others are fixed to the well-motivated values. In case of $E_{\text{T}}^{\text{miss}} + V(qq)$ and $E_{\text{T}}^{\text{miss}} + h(b\bar{b})$ analyses, these parameters are the mediator masse m_{med} and DM mass m_χ (mass of pseudoscalar m_A) for simplified $s-$channel mediator models (Z'-2HDM model). Figure 11.1 illustrates 95% exclusion limits for a simplified model with a vector mediator obtained for a hypothetical LHC measurement. Such plot contains observed (with the corresponding uncertainties) and expected limits, as well as the values of the mass parameters where the expected relic density is consistent with the observed density from cosmological observations.
Generally speaking, one can consider three main kinematic regions for such a limit plot:

On-shell region: In this regime $m_{\text{med}} > 2m_\chi$ ($m_{\text{med}} > m_h + m_A$) for the $E_{\text{T}}^{\text{miss}} + V(qq)$ simplified models ($E_{\text{T}}^{\text{miss}} + h(b\bar{b})$ Z'-2HDM model). For the $E_{\text{T}}^{\text{miss}} + V(qq)$ models, the mediator can decay to a pair of the DM particles, giving the $E_{\text{T}}^{\text{miss}} + X$ final state, or to a pair of quarks, which is the subject of the dijet searches. The energy scale of SM particle, which originates from the ISR, is driven by the mass of mediator m_{med}. For Z'-2HDM model, the energy scale of the h boson is driven by the mass of mediator m_{med} and mass of pseudoscalar scalar m_A. The mass of DM particles m_χ has no significant impact on the kinematic distributions, meaning that no fine scan along m_χ is required. In this region LHC provides the most stringent limits for the $E_{\text{T}}^{\text{miss}} + X$ searches.

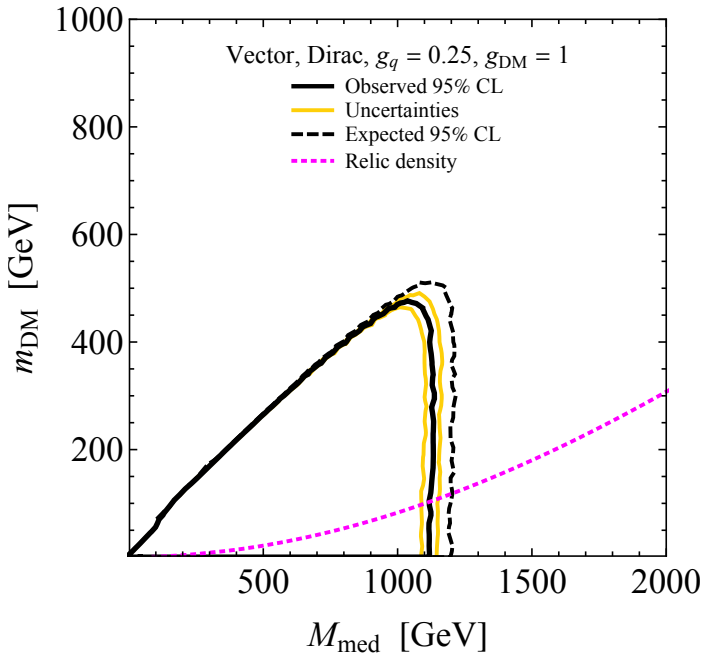


Figure 11.1: Hypothetic 95% exclusion limits in the grid of the DM and mediator masses (m_χ, m_{med}) for a simplified model with a vector mediator. The shown limits serve for the illustrative purpose only. The black solid and dashed curves show the hypothetical observed and expected limits. The yellow curves show the hypothetical uncertainties on the observed limit. The dotted magenta curve stands for the (m_χ, m_{med}) points for which the expected relic density, calculated using MadDM [229], is consistent with the observed density $\Omega h^2 = 0.12$ from the WMAP [27] and Planck [26] measurements. The region on the right of the curve corresponds to higher predicted relic abundance. The chunks of the parameter phase-space covered by the exclusion lines are excluded.

The limits and the signal cross-sections for the $E_{\text{T}}^{\text{miss}} + V(qq)$ simplified models have a complex dependence on all model parameters, $\sigma \sim g_q^2 g_\chi^2 / \Gamma^{\text{total}}$ for the fixed m_{med} and m_χ, where Γ^{total} denotes the total decay width of mediator with partial widths defined in Eq. 3.17. This means that in general it is not trivial to re-scale limits from one set of couplings g_q and g_χ to another. According to Eq. 3.22, the signal cross-sections for Z'-2HDM model scales with the model parameters as $[g_z \cos\alpha \cos\beta]^2$.

Transition region: In this regime $m_{\text{med}} \approx 2m_\chi$ ($m_{\text{med}} \approx m_h + m_A$) for the $E_{\text{T}}^{\text{miss}} + V(qq)$ simplified models ($E_{\text{T}}^{\text{miss}} + h(b\bar{b})$ Z'-2HDM model). The rate of DM production (hA^0 production) is resonantly enhanced. Here, the signal cross-sections and the kinematic distributions depend much stronger on the mass parameters, which requires finer scan along both m_χ (m_A) and m_{med} [230].

Off-shell region: In this regime $m_{\text{med}} < 2m_\chi$ ($m_{\text{med}} < m_h + m_A$) for the $E_{\text{T}}^{\text{miss}} + V(qq)$ simplified models ($E_{\text{T}}^{\text{miss}} + h(b\bar{b})$ Z'-2HDM model). Here, the mediator decays are highly suppressed and $E_{\text{T}}^{\text{miss}} + X$ searches loose their exclusion power. The corresponding production cross-sections are strongly suppressed by a factor of m_{med}^2/Q^2, where Q^2 denotes the momentum transfer of the process. The signal cross-sections of $E_{\text{T}}^{\text{miss}} + V(qq)$ simplified models become proportional to the couplings, $\sigma \sim g_q^2 g_\chi^2$, meaning that the limits for other set of couplings can be easily obtained by simple re-scaling procedure.

11.1.3.1 $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ exclusion limits

95% CLs limits are obtained using the combined profile likelihood fit described in Chapter 9. As was mentioned in Secion 11.1.3, in case of simplified models of the DM production, the kinematic distributions in the on-shell region have no strong dependence on the DM mass m_χ. This means that the kinematic acceptances and efficiencies are similar for the models with different DM mass m_χ. Thus, one just need to re-scale the limit on the signal strength μ as a function of the DM mass m_χ for the fixed mediator mass $m_{Z'}$ by the ratio of the cross-section. An interpolation procedure is used to build smooth continuos exclusion contours. It is based on the following relation:

$$\mu = \frac{N_{pp\to Z'\to\chi\chi}}{\mathcal{L}\times\sigma_{pp\to Z'\to\chi\chi}\times\mathcal{A}\times\varepsilon}\,, \tag{11.1}$$

where $N_{pp\to Z'\to\chi\chi}$ is the limit on the signal yield, $\mathcal{L}$ is the corresponding luminosity, $\sigma_{pp\to Z'\to\chi\chi}$ is the production cross-section of a given simplified model, and $\mathcal{A}\times\varepsilon$ defines the kinematic acceptance and the selection efficiency in the signal region. As was discussed before, for a fixed $m_{Z'}$, $\mathcal{A}\times\varepsilon$ is assumed to be similar for the signal models across different m_χ values. Therefore, the signal strength depends only on the production cross-section for a given mass point $(m_\chi, m_{\mathrm{med}})$. Under the narrow width approximation, the production cross-section for a given m_χ value can be expressed as a product of the total production cross-section $\sigma_{pp\to Z'\to\chi\chi}$ and the branching ratio of the $Z'\to\chi\chi$ decay:

$$\sigma_{pp\to Z'\to\chi\chi}(m_\chi, m_{Z'}) = \sigma_{pp\to Z'\to\chi\chi}(m_{Z'})\times\mathcal{B}_{Z'\to\chi\chi}(m_\chi)\,, \tag{11.2}$$

where the branching ratio depends on m_χ as described in Eq. 3.17. Thus, the signal strength for a given $(m_\chi, m_{Z'})$ can be expressed in terms of the production cross-section and the branching ratio of the $Z'\to\chi\chi$ decay at the mass point $(m_\chi = 1\ \mathrm{GeV}, m_{Z'})$:

$$\mu_{pp\to Z'\to\chi\chi}(m_\chi, m_{Z'}) = \sigma_{pp\to Z'}(m_\chi = 1\ \mathrm{GeV}, m_{Z'})\times\frac{\mathcal{B}_{Z'\to\chi\chi}(m_\chi)}{\mathcal{B}_{Z'\to\chi\chi}(m_\chi = 1\ \mathrm{GeV})}\,. \tag{11.3}$$

This re-scaling is performed in $m_{Z'}$ steps of 100 GeVand in m_χ steps of 50 GeV. The linear interpolation is used between the nominal and re-scaled points to retrieve the upper limit on the signal strength as a function of $m_{Z'}$ for fixed m_χ and as a function of m_χ for fixed $m_{Z'}$. The points on the exclusion limit curve are defined as the crossing points of the upper limit curves for fixed m_χ and for fixed $m_{Z'}$ and th $\mu = 1$ line, which represents the signal strength under the signal-plus-background hypothesis. The model with a set of $(m_\chi, m_{Z'})$ parameters is excluded at 95% confidence level if its upper limit on μ is below the $\mu = 1$ line.

11.1.3.2 $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ Exclusion limits

95% CLs limits are obtained from the combined profile likelihood fit described in Chapter 9. The exclusion limit setting procedure for $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ search is different comparing to the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ search. Limits are calculated in terms of the production cross-section $\sigma_{pp\to h\chi\chi}\times\mathcal{B}(h\to b\bar{b})$ for different $(m_{Z'}, m_A)$ of Z'-2HDM model. The pseudoscalar A^0 is assumed to decay only to a pair of DM particles, $\mathcal{B}(A\to\chi\chi) = 100\%$.
In contrast to the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ search, the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ limits on $\sigma_{pp\to h\chi\chi}\times\mathcal{B}(h\to b\bar{b})$ for one parameter set $(m_{Z'}, m_A)$ cannot be translated to the limits for another parameter set

by simple re-scaling procedure. This is due to the stronger dependence of $E_{\mathrm{T}}^{\mathrm{miss}}$ spectrum on the mass parameters of the model. First, the limits on $\sigma_{pp\to h\chi\chi} \times \mathcal{B}(h \to b\bar{b})$ are translated to the limits on the total production cross-section $\sigma_{pp\to h\chi\chi}$ by dividing out the SM branching ratio of $h \to b\bar{b}$ decay, $\mathcal{B}(h \to b\bar{b}) = 0.571$. The shape of the exclusion limits in the mass-mass plane are derived based on the mass points $(m_{Z'}, m_A)$, at which the theoretical cross-section is equal to the expected limit. The linear interpolation between the fully simulated mass points $(m_{Z'}, m_A)$ is used to reveal an approximate dependence of theoretical cross-section and exclusion limits on the mediator mass $m_{Z'}$ and psedoscalar mass m_A. The crossing points of the theoretical cross-section curve and the expected limit curve as functions of one of the masses, m_A or $m_{Z'}$, when another mass is fixed, are determined. A linear interpolation between the resulting crossing points is used to arrive to the smooth exclusion limit contour [231]. The interpolation curve is defined as a function of one of two masses, m_A or $m_{Z'}$, while another mass is fixed.

11.2 Generic limits

In addition to the exclusion limits for the specific DM model described in Sections 11.1.3.2 and 11.1.3.1, the results are interpreted in terms of generic upper limits on the visible detector-level cross-section σ_{vis} of the production of the W/Z+DM and h+DM events. The interpretation of the search results in terms of the generic limits provides a powerful tool for phenomenologists to check a wide range of the models with a given final state.
In case of the generic limits, any extra model assumptions for the generic limits should be avoided or at least minimised. Only back-to-back topology of the final state is explicitly used. The SM nature of the W, Z and Higgs bosons is assumed [1], meaning that one can use the SM branching ratios of their decays and in particular the distributions of the invariant mass of the products of SM W, Z and Higgs boson decays, m_{jj}/m_J. Since the shape of $E_{\mathrm{T}}^{\mathrm{miss}}$ spectrum is highly dependent on a given signal model, generic limits are set in the regions of $E_{\mathrm{T}}^{\mathrm{miss}}$ individually, after passing all selection requirements described in Chapter 7.

11.2.1 Statistical model for generic limits

Generally, the statistical model used for the generic limit setting is identical to those for the model-dependent interpretation, except one major difference that to set the limits in a given range of $E_{\mathrm{T}}^{\mathrm{miss}}$, this and only this range is analysed in the signal region. The stastical setup for the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ generic limit setting can be summarised as follows:

Signal region: A single $E_{\mathrm{T}}^{\mathrm{miss}}$ range is used at a time in the statistical setup of both analyses. The high purity and low purity regions (see Section 7.7) are considered together for a given $E_{\mathrm{T}}^{\mathrm{miss}}$ range to improve sensitivity of the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis.

Control regions: Since one does not expect any signal contamination in the control regions, all regions of $E_{\mathrm{T}}^{\mathrm{miss}}$ are simultaneously analysed in the 1μCR and 2ℓCR to better constrain backgrounds.

[1] An observation of the $h \to b\bar{b}$ decay was announced in August this year by ATLAS [232] and CMS [233] experiments.

All regions: Regions with 0, 1 and 2 $b-$tags are considered to improve sensitivity to the $W \to q'\bar{q}$ (0/1 $b-$tag regions), $Z \to c\bar{c}$ (1 $b-$tag region) and $Z \to b\bar{b}$ (2 $b-$tag region) decays in the signal region. The fine granularity of regions of $E_{\mathrm{T}}^{\mathrm{miss}}$ for the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ standard setup leads to the limited statistics for both signal and background processes in the regions of high $E_{\mathrm{T}}^{\mathrm{miss}}$. Therefore, the regions above $E_{\mathrm{T}}^{\mathrm{miss}} > 300$ GeV are merged, resulting in the six regions of $E_{\mathrm{T}}^{\mathrm{miss}}$ for the generic limit setting procedure: $[150, 200)$ GeV, $[200, 250)$ GeV, $[250, 300)$ GeV, $[300, 400)$ GeV, $[400, 600)$ GeV, $[600, 1500)$ GeV. This ensures the numerical stability and reduces the magnitude of statistical fluctuations in the observed limits. The $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ setup for generic limit setting shares the same $b-$tag regions and regions of $E_{\mathrm{T}}^{\mathrm{miss}}$ with the standard analysis setup.

11.2.2 Detector-level generic limits

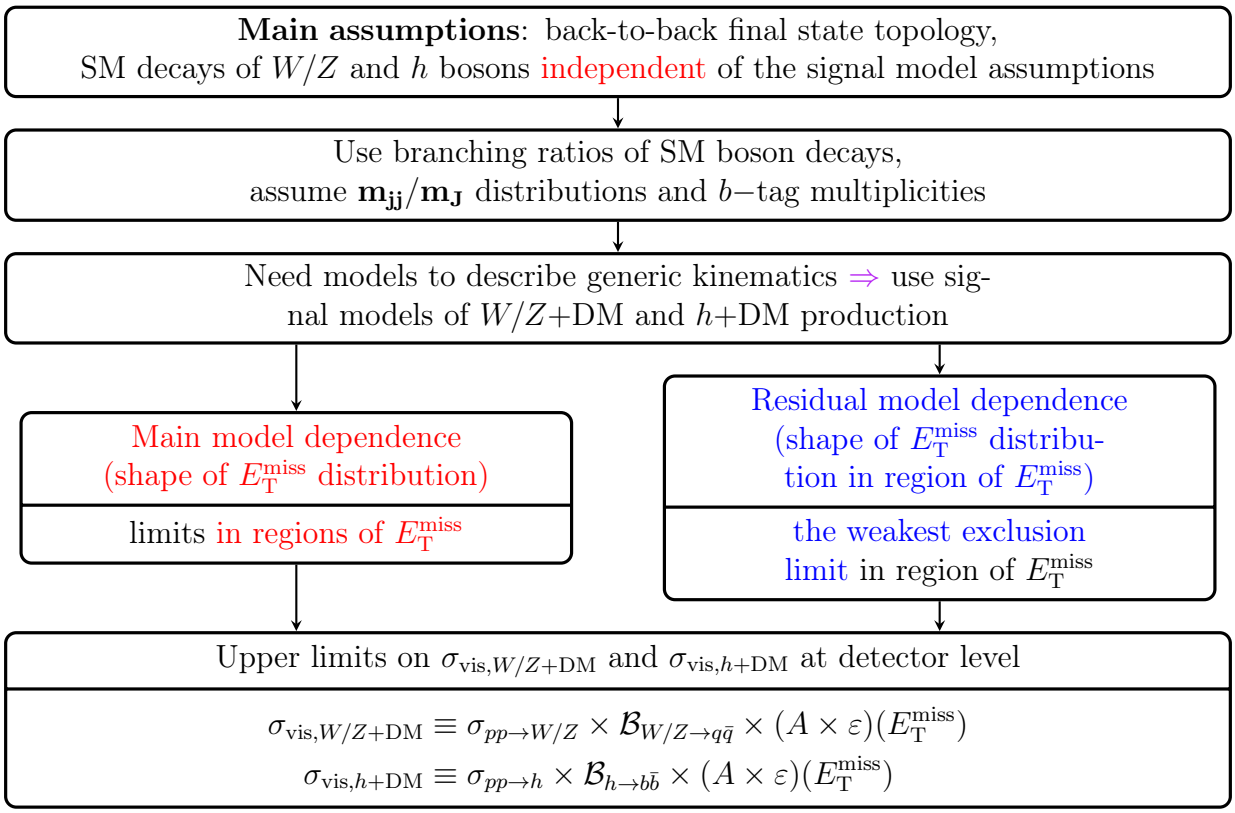


Figure 11.2: Overview of generic limit setting procedure at detector level.

The $E_{\mathrm{T}}^{\mathrm{miss}}+V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ analyses explicitly require the distributions of m_{jj}/m_J and $b-$tagging multiplicity in the statistical model to obtain sensitivities to the W/Z+DM and h+DM events. This means that generic models of W/Z+DM and h+DM production need to be considered for these observables in case of generic limit setting. The Monte Carlo (MC) simulated samples for the simplified vector mediator models and Z'-2HDM models are used in this case to model m_{jj}/m_J distribution and $b-$tagging multiplicity. Despite the fact that limits are set in separate regions of $E_{\mathrm{T}}^{\mathrm{miss}}$, the residual dependence is introduced through the dependence on the event kinematics that originates from the shape of $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in a single region of $E_{\mathrm{T}}^{\mathrm{miss}}$. Figure 11.3 illustrates a strong dependence of $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions on the mass parameters of $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ 2HDM+a model.

The 2HDM+a models with the masses $m_a = 450$ GeV and $m_a = 350$ GeV significantly contribute to a representative region of $E_{\mathrm{T}}^{\mathrm{miss}}$, while having drastically different shapes of $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution. This leads to the different signal acceptances and selection efficiencies, especially close to the edges of the region, where the shape differences are the largest. This can have a sizeable impact on the distributions of key variables, such as m_{jj}/m_J and $b-$tagging multiplicity. To reduce this residual dependence, a range of signal models with different mass parameters, $(m_\chi, m_{Z'})$ or $(m_{Z'}, m_A)$, is considered in a given region of $E_{\mathrm{T}}^{\mathrm{miss}}$, and the parameter set with the weakest expected limit is taken. An overview of detector-level generic limit setting is shown in Figure 11.2. The choice of the mass parameters is motivated by a sizeable contribution of the signal model in a given region of $E_{\mathrm{T}}^{\mathrm{miss}}$ to avoid large statistical fluctuations in the observed limits. This results in the requirement of at least 500 signal events per region of $E_{\mathrm{T}}^{\mathrm{miss}}$ for the $E_{\mathrm{T}}^{\mathrm{miss}}+V(qq)$ analysis, which is relaxed to 200 events in the highest region of $E_{\mathrm{T}}^{\mathrm{miss}}$ due to the lower MC statistics. For the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis this requirement is set to 100 events per region of $E_{\mathrm{T}}^{\mathrm{miss}}$.

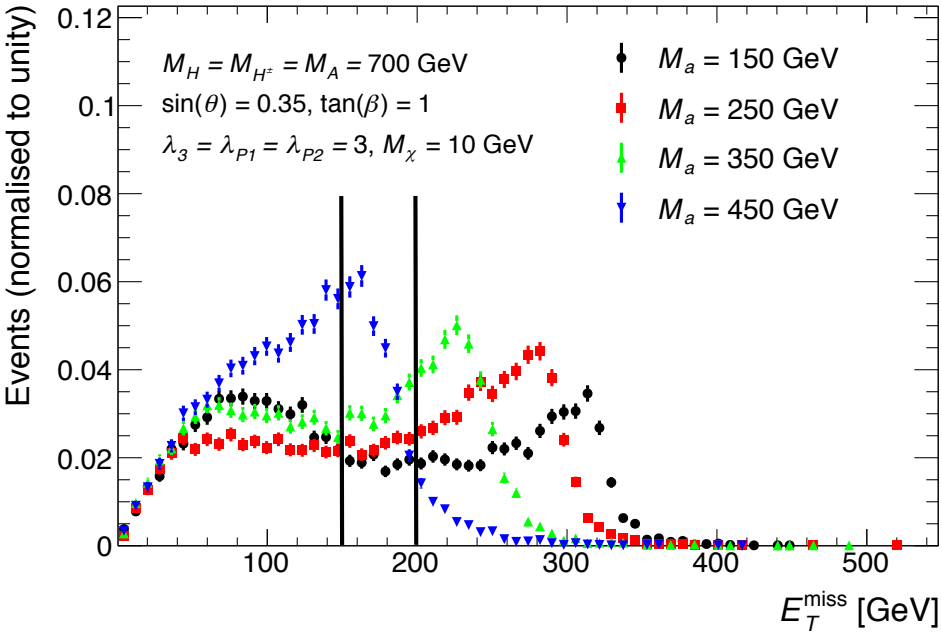


Figure 11.3: Comparison of the normalised $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions for the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ 2HDM+a model for different masses of the light pseudoscalar m_a for a fixed mass of the heavy pseudoscalar $m_A = 700$ GeV. The solid black lines show the edges of representative region of $E_{\mathrm{T}}^{\mathrm{miss}}$.

The final limits are set on the following detector-level observables:

$$\begin{aligned}
\sigma_{\mathrm{vis},W+\mathrm{DM}}(E_{\mathrm{T}}^{\mathrm{miss}}) &\equiv \sigma_{pp\to W} \times \mathcal{B}_{W\to q'\bar{q}} \times (A\times\varepsilon)(E_{\mathrm{T}}^{\mathrm{miss}}) \\
\sigma_{\mathrm{vis},Z+\mathrm{DM}}(E_{\mathrm{T}}^{\mathrm{miss}}) &\equiv \sigma_{pp\to Z} \times \mathcal{B}_{Z\to q\bar{q}} \times (A\times\varepsilon)(E_{\mathrm{T}}^{\mathrm{miss}}) \\
\sigma_{\mathrm{vis},h+\mathrm{DM}}(E_{\mathrm{T}}^{\mathrm{miss}}) &\equiv \sigma_{pp\to h} \times \mathcal{B}_{h\to b\bar{b}} \times (A\times\varepsilon)(E_{\mathrm{T}}^{\mathrm{miss}}) \,,
\end{aligned} \tag{11.4}$$

where $\mathcal{B}_{W\to q'\bar{q}}$, $\mathcal{B}_{Z\to q\bar{q}}$, and $\mathcal{B}_{h\to b\bar{b}}$ denote the branching ratios of the $W\to q'\bar{q}$, $Z\to q\bar{q}$ and $h\to b\bar{b}$ decays, respectively. $A\times\varepsilon$ represents the kinematic acceptance times the experimental efficiencies of the $E_{\mathrm{T}}^{\mathrm{miss}} + W(qq)$, $E_{\mathrm{T}}^{\mathrm{miss}} + Z(qq)$, and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ event selection. $A\times\varepsilon$ is calculated in each region of $E_{\mathrm{T}}^{\mathrm{miss}}$ separately. To facilitate the comparison with the visible detector-level cross-sections calculated by theorists, $b-$tagging multiplicity and m_{jj}/m_J requirements are dropped from the definition of the visible cross-section. This is done because of the poor modelling of the $b-$tagging and m_{jj}/m_J distributions in

parametrised detector simulations. The generic limits are set on the $E_{\mathrm{T}}^{\mathrm{miss}} + W(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + Z(qq)$ production separately to take advantage from the different sensitivities in the difference regions of the phase space.

11.2.3 Parton-level generic limits

Theorists can use different detector simulation packages that exist on the market, e.g. DELPHES [234], to translate their parton-level model prediction to detector level and compare to the generic limits [2]. DELPHES provides a fast simulation of the detector response, but at the same time it poorly models some relevant detector-level observables, such as $b-$tagging multiplicity. Hence, to facilitate comparison of the W+DM, Z+DM and h+DM model predictions to the generic limits, a range of typical values of $A \times \varepsilon$ are provided per region of $E_{\mathrm{T}}^{\mathrm{miss}}$.

The $A \times \varepsilon$ in a given region of $E_{\mathrm{T}}^{\mathrm{miss}}$ is defined as follows:

$$A \times \varepsilon \equiv \frac{N_{\mathrm{evt}} \text{ in given region of } (E_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{truth}} \text{ and same region of } (E_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{reco}} \text{ after reco+selection}}{N_{\mathrm{evt}} \text{ in given region of } (E_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{truth}} \text{ before reco+selection}} . \quad (11.5)$$

The same procedure is used to reduce the possible residual model dependence of $A \times \varepsilon$ as for detector-level limits. The smallest value of $A \times \varepsilon$ is taken in a given region of $E_{\mathrm{T}}^{\mathrm{miss}}$, which results in the weakest exclusion limit at parton level.

There is no access to the exact value of the denominator in the definition of $A \times \varepsilon$ due to the preselection applied at the production stages of the MC signal samples used to set generic limits. Therefore, $A \times \varepsilon$ is factorised into a preselection part $(A \times \varepsilon)_{\mathrm{presel}}$ and into a part after applying the full event selection, $(A \times \varepsilon)_{\mathrm{sel}}$:

$$A \times \varepsilon \equiv (A \times \varepsilon)_{\mathrm{presel}} \times (A \times \varepsilon)_{\mathrm{sel}} . \quad (11.6)$$

$(A \times \varepsilon)_{\mathrm{sel}}$ can be calculated exactly per region of $E_{\mathrm{T}}^{\mathrm{miss}}$ as defined by the following relation:

$$(A \times \varepsilon)_{\mathrm{sel}} \equiv \frac{N_{\mathrm{evt}} \text{ in given region of } (E_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{truth}} \text{ and same region of } (E_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{reco}} \text{ after reco+selection}}{N_{\mathrm{evt}} \text{ in given region of } (E_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{truth}} \text{ after preselection and before reco+selection}} . \quad (11.7)$$

However, the $(A \times \varepsilon)_{\mathrm{presel}}$ term is hard to be derived due to the $E_{\mathrm{T}}^{\mathrm{miss}} > 140$ GeV requirement present at the pre-selection stage. For the models with soft $E_{\mathrm{T}}^{\mathrm{miss}}$ spectra, e.g. Z'-2HDM models with low $m_{Z'}$ and m_A, or simplified $s-$channel mediator models with low $m_{Z'}$ and m_χ, a significant fraction of $E_{\mathrm{T}}^{\mathrm{miss}}$ spectrum fails this requirement, biasing $(A \times \varepsilon)_{\mathrm{presel}}$ to low values. The Z'-2HDM models of the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ production provide $E_{\mathrm{T}}^{\mathrm{miss}}$ spectra, peaking in the different regions of $E_{\mathrm{T}}^{\mathrm{miss}}$ depending on the $(m_{Z'}, m_A)$ parameters. This allows to estimate overall $(A \times \varepsilon)_{\mathrm{presel}}$ using Z'-2HDM models with hard $E_{\mathrm{T}}^{\mathrm{miss}}$ spectrum, which is not strongly affected by the $E_{\mathrm{T}}^{\mathrm{miss}} > 140$ GeV requirement. The simplified vector mediator models provide instead a falling $E_{\mathrm{T}}^{\mathrm{miss}}$ spectrum with the majority of events populating low region of $E_{\mathrm{T}}^{\mathrm{miss}}$, as shown in Figure 11.4. It means that $(A \times \varepsilon)_{\mathrm{presel}}$ is artificially low for this class of models. In this case $(A \times \varepsilon)_{\mathrm{presel}}$ is estimated based on the parton-level cross-sections for the different sets of $(m_\chi, m_{Z'})$ parameters

[2] The parton level represents the simulation stage before the partons (quarks and gluons) undergo the processes of hadronisation to form stable particle states.

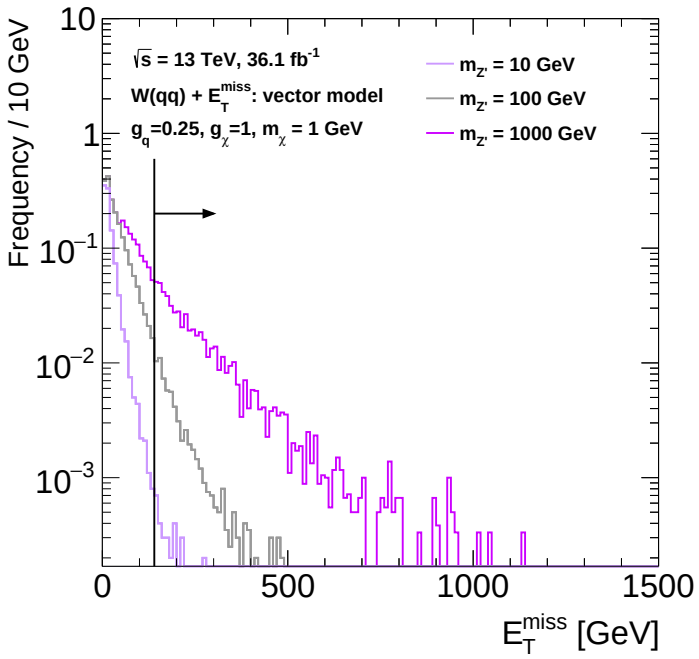


Figure 11.4: Comparison of normalised $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions for the $E_{\mathrm{T}}^{\mathrm{miss}} + W(qq)$ simplified vector mediator model for different mediator masses $m_{Z'}$ for a fixed DM mass $m_{\chi} = 1$ GeV. The region on the right of solid black line shows the region of $E_{\mathrm{T}}^{\mathrm{miss}} > 140$ GeV.

after applying the preselection cut of $(E_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{truth}} > 140$ GeV:

$$(A \times \varepsilon)_{\mathrm{presel}} \equiv \frac{N_{\mathrm{evt}} \text{ before reco+selection}}{\mathcal{L} \cdot \sigma_{\mathrm{signal}}((E_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{truth}} > 140 \text{ GeV})} \,. \tag{11.8}$$

The parton-level cross-sections are derived using MC samples with 20000 events each, generated using MadGraph5+Pythia 8 event generator. $(E_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{truth}}$ is calculated after the parton showering.

Chapter 12

Results

The final stage of every search for the new physics is the interpretation of the results, when one addresses the main questions about the potential presence of new physics and the capability of a given model to describe the observed data. The answers on these two questions are given in terms of the goodness of the fit of the observed data to the known SM prediction and tests of hypotheses of the additional presence of new physics models. The main statistical tools to answer these questions are described in details in Chapter 9. This chapter is aimed to use these statistical tools to provide the exact answers by interpreting the results of the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ searches.
The ability of the SM simulations to describe the observed data is discussed in Sections 12.1.1 and 12.1.2 in terms of the constraints on the sources of the systematic uncertainties represented through the nuisance parameters. A comparison of the key kinematic distributions between the observed data and SM backgrounds after the statistical fit is presented in Section 12.1.3. The results of the model hypothesis testing for the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ simplified vector mediator and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ Z'-2HDM models ar described in Section 12.2.1. And finally, generic limits on the visible cross-section of the production of W/Z+DM events without any extra signal assumptions are discussed in Section 11.2.2 and in Section 12.3.1 for the production of h+DM events.

12.1 Results in data

12.1.1 Constraints from data on systematic uncertainties

An important figure of merit for the ability of the statistical model to describe the observed data is given by the constraints on the nuisance parameters (NP) that represent the systematic uncertainties after the fit to data. This performance can be expressed in terms of the *pull* of the NP, which quantifies how far the fit to data has to shift the value of the NP from its *prior* (expected) value while finding the maximum likelihood estimator (MLE). The prior value is usually based on additional auxiliary measurements or specific Monte Carlo studies like the JES calibration.
The pull of NP is defined as $(\theta_{\mathrm{fit}} - \theta_0)/\Delta\theta$, where $\theta_{\mathrm{fit}} = \hat{\theta}$ is the MLE of a given NP θ, θ_0 is the expected value of a given NP, and $\Delta\theta$ is the corresponding prior uncertainty. The NPs is *constrained* after the fit to the data if $\Delta\theta_{\mathrm{fit}} < \Delta\theta$, where $\Delta\theta_{\mathrm{fit}}$ denotes the uncertainty on the fitted value of NP.

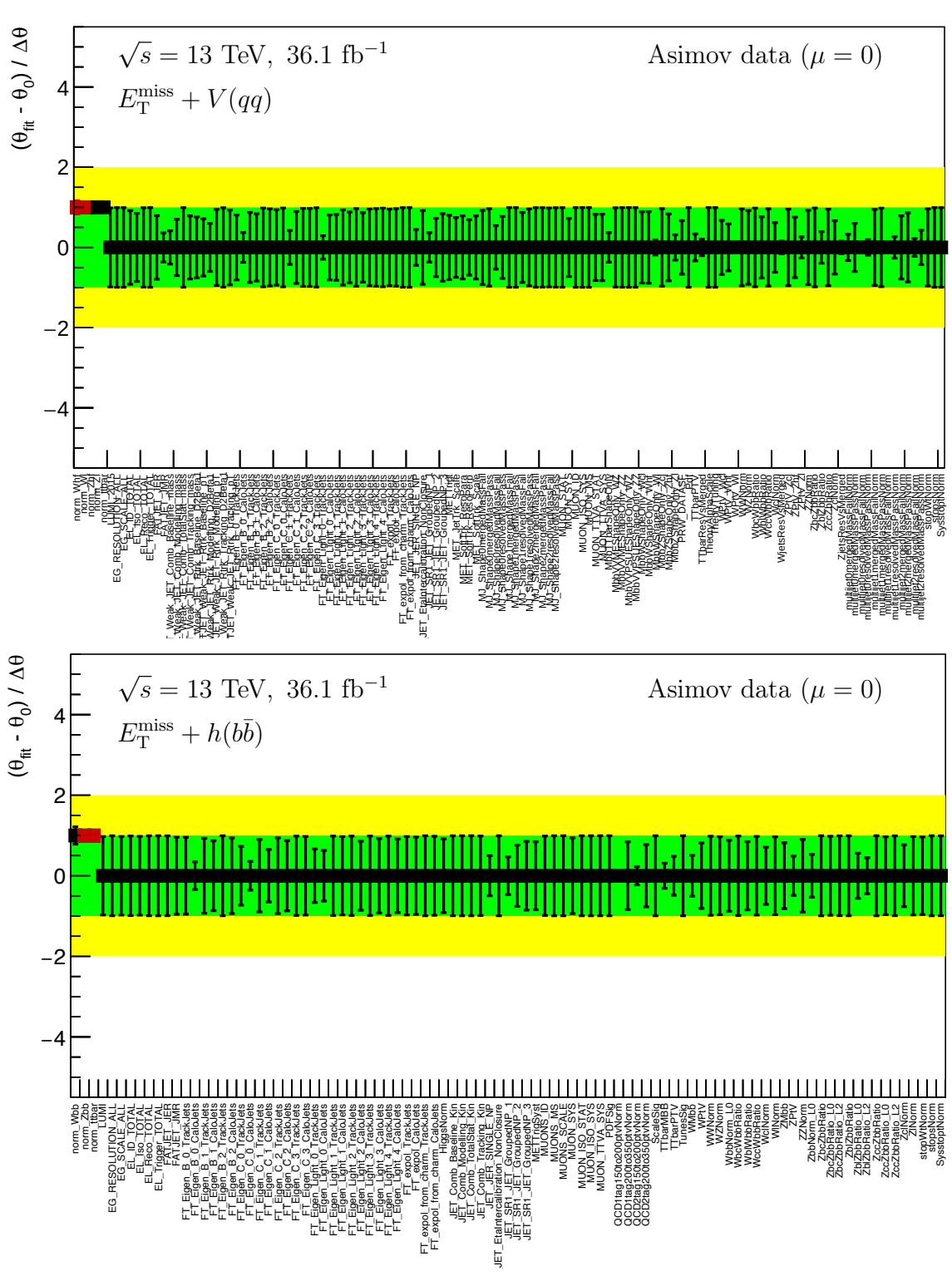


Figure 12.1: The nuisance parameter fitted values after the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ (top) and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ (bottom) analyses profile likelihood fit to Asimov data assuming no signal, $\mu = 0$.

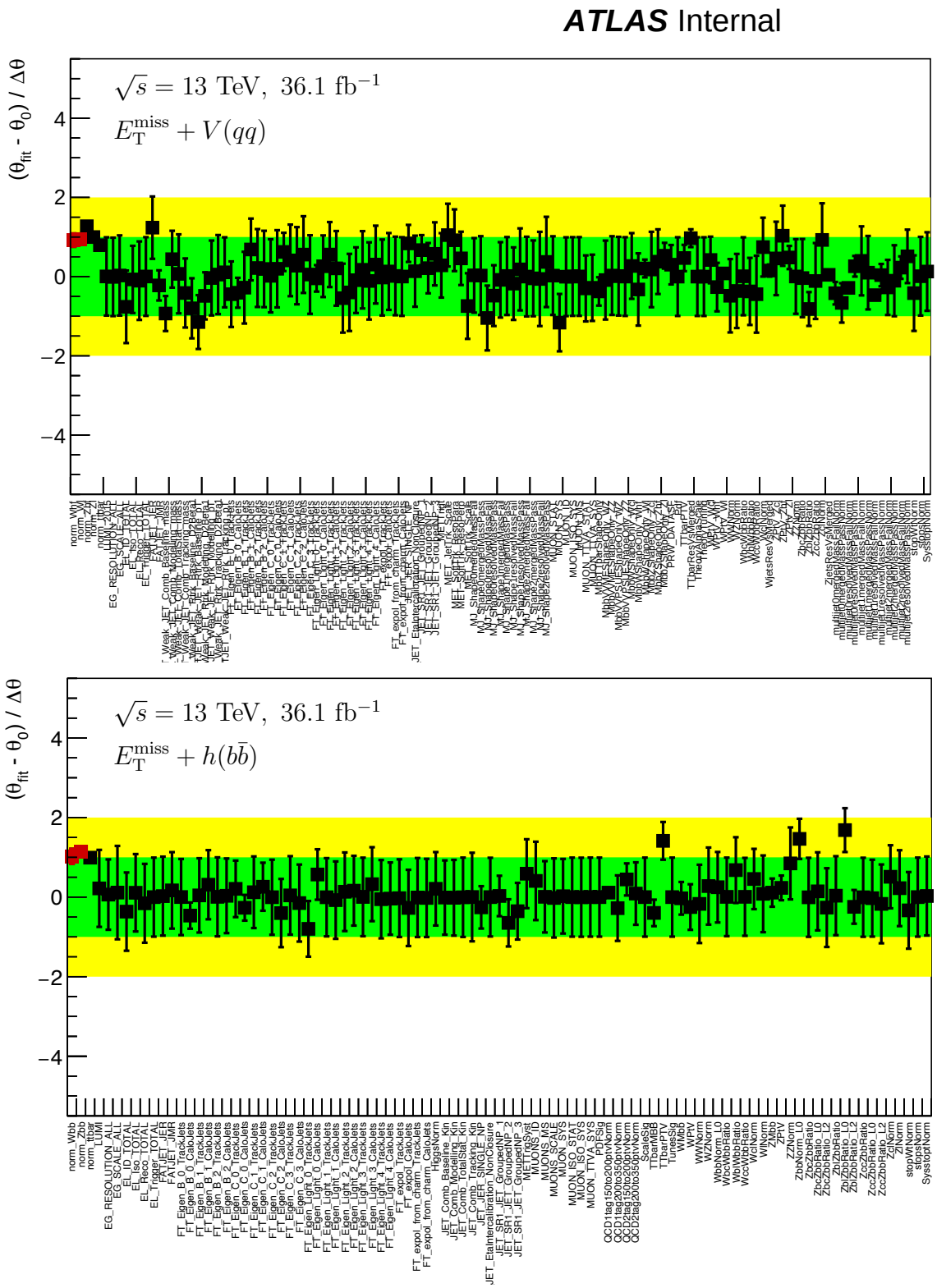


Figure 12.2: The nuisance parameter fitted values after the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ (top) and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ (bottom) analyses profile likelihood fit to observed data in the sidebands of the discriminating variable m_{jj}/m_J. The signal strength μ is allowed to float.

Figure 12.1 illustrates the expected values of NPs and constraints after the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ fit to the Asimov data (see Section 9.1.2). Given the definition of the Asimov data set as having all statistical fluctuations neglected and MLEs of all NPs assumed to be the expected values, the corresponding pulls of NPs are equal to zero by construction. The NPs shown at +1 correspond to the free normalisation parameters of the dominant backgrounds, $W+$jets, $Z+$jets and $t\bar{t}$. Their values before the fit are set to 1, corresponding to the prediction. The fit provides strong constraints on the NPs related to the following systematic uncertainties:

- The modeling uncertainties on the p_T^V and $m_{b\bar{b}}$ for $V+$jets and $t\bar{t}$ backgrounds: These uncertainties can be strongly constrained from the 1 lepton CR enriched with $W+$jets and $t\bar{t}$ processes, and from 2 lepton CR enriched with $Z+$jets processes. For the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis, the fit provides the strongest constraints for the $V+ll$ component compared to the $V+cl$ and $V+$HF components due to the much higher branching ratios of the former. The $V+$jets contributions are considered inclusive in the flavour composition for the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis.

- The multijet normalisation uncertainties in the resolved regime: The multijet background processes mainly populate the regions with low $E_{\mathrm{T}}^{\mathrm{miss}}$. This means that the associated normalisation uncertainties can be strongly constrained from the multijet CRs in the resolved regime of both analyses. For the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis, the constraints in the 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 200 GeV region are considerably stronger due to higher rate of the multijet processes in this region comparing to the higher $E_{\mathrm{T}}^{\mathrm{miss}}$ regions. For the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis, a strong constraint is also put on the multijet normalisation in the region with 0 $b-$tags for the merged regime. This can be explained by the high sensitivity in this analysis region due to the priority-merged selection and significant multijet contribution in the SR for 250 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 350 GeV (see Figure 12.3).

- $Z+bl/Z+bb$ ratio uncertainties: For the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis, the constraints on the uncertainty on the $Z+bl/Z+bb$ ratio are stronger in the 2 lepton CR, which is enriched with the $Z+$jets processes.

- JES uncertainties in the low p_T region: They can be constrained due to high statistics in the low $E_{\mathrm{T}}^{\mathrm{miss}}$ region. The constraints become notably weaker in the medium and higher p_T regions with significantly lower statistics.

The correlations between the NPs after the fit to data are shown in Appendix C.
Figure 12.2 illustrates the observed NP value after the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ fit to observed data. The values of the observed NPs, which correspond to the normalisation parameters of the dominant backgrounds, are given in Table 12.1. The fitted values of the normalisation NPs are consistent with the expectations.
In case of the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis, the largest pulls of the fitted values of the NPs 1σ are of an order of $\pm 1\sigma$. For the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis, similar global picture is observed, however the fitted values of the following three NPs are pulled by approximately 1.5σ [220]:

- **TTbarPTV:** Stands for the shape uncertainty on the reconstructed transverse momentum of the vector boson, p_T^V, for the $t\bar{t}$ background. The large pull of this NP originates from the Monte Carlo modeling of the $E_{\mathrm{T}}^{\mathrm{miss}}$ spectrum for the $t\bar{t}$ backgrounds. It was shown that in case of the SM $Vh(b\bar{b})$ analysis the $E_{\mathrm{T}}^{\mathrm{miss}}$ spectrum

Nuisance parameter	Postfit values	
	$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$	$E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$
$W+$ HF normalisation	0.92 ± 0.17	1.00 ± 0.20
$Z+$ HF normalisation	1.27 ± 0.09	1.14 ± 0.07
$W+$ LF normalisation	0.95 ± 0.04	—
$Z+$ LF normalisation	1.00 ± 0.04	—
$t\bar{t}$ normalisation	0.79 ± 0.03	0.99 ± 0.03

Table 12.1: Values of the free normalisation parameters of the free dominant backgrounds after the fit to the observed data in the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analyses [219, 220]. The abbreviations HF and LF correspond to respectively heavy and light flavour defined in Section 8.1.

for the $t\bar{t}$ samples generated with the default Monte Carlo generator is softer than one in data [217]. The differences in the $t\bar{t}$ shape variations on the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution between the different Monte Carlo generators are small, i.e. they all predict too soft p_T^V spectrum. This results in a small TTbarPTV uncertainty that does not cover the difference to data. The fit compensates this mismodeling with the significant pull of the value of the TTbarPTV NP.

- **ZbbNorm_L0:** Stands for the relative normalisation uncertainty of the $Z + bb$ flavour component in the 0 lepton region with respect to the 2 lepton region. A significant pull is linked to the low statistics of the heavy flavour components of the $Z+$jets backgrounds in the 2 lepton region and an overestimate of the jet activity in the forward region in the $Z+$jets samples generated by the SHERPA Monte Carlo generator.

- **ZblZbbRatio_L0:** Stands for the relative normalisation of the $Z + bl$ flavour component with respect to the $Z + bb$ flavour component, inclusive in all $b-$tag regions. The tension in tge pull originates from the different acceptances of the $Z + bl$ and $Z + bb$ components in the 1 $b-$tag ($Z + bl$ is dominant) and 2 $b-$tag ($Z + bb$ is dominant) regions.

12.1.2 Impact of systematic uncertainties

The relevance of the different sources of systematic uncertainties for a given analysis can be quantified by the evaluating the impact of their corresponding parameters of interest, which is in this case is the signal strength μ. The impact is defined as a fractional uncertainty on the fitted signal strength and estimated for three representative scenarious with low, medium and high $E_{\mathrm{T}}^{\mathrm{miss}}$ spectrum, for each of the two analyses presented in this thesis.
The impact of a given NP θ_i on the fitted signal strength is estimated by performing the fit without the NP of interest, which produces the reduced fitted uncertainty $\sigma_{\mathrm{no}\,\theta_i}$. Given that the total uncertainty is calculated as a squared sum of all uncertainties, the resulting impact of an uncertainty of interest is defined as:

$$\sigma_{\hat{\mu}} = \sqrt{\sigma^2_{\mathrm{total}} - \sigma^2_{\mathrm{no}\,\theta_i}}\,. \tag{12.1}$$

The NPs are cathegorized in different groups according to their source, e.g. small$-R$ jet uncertainty combine all uncertainties related to the calibration of small$-R$ jets, such as jet energy scale and resolution.

Source	Uncertainty on $\mu = 1$ [%]		
	Simplified vector mediator model $(m_\chi, m_{Z'})$ [GeV]		
	$(1, 200)$	$(1, 600)$	$(1, 2000)$
Large$-R$ jets	8.5	19.6	18.8
Small$-R$ jets	3.4	8.3	10.3
Electrons	3.8	9.3	12.0
Muons	5.9	7.1	3.9
$E_{\mathrm{T}}^{\mathrm{miss}}$	1.3	4.3	6.5
$b-$tagging (track jets)	3.6	3.8	5.9
$b-$tagging (small$-R$ jets)	1.6	4.1	2.2
Luminosity	2.7	3.6	3.7
Multijet normalisation	7.3	11.3	10.0
Diboson normalisation	5.1	11.2	12.8
$Z+$jets normalisation	4.7	9.0	11.7
$W+$jets normalisation	2.6	3.5	4.6
$t\bar{t}$ normalisation	2.7	1.2	3.2
Signal modeling	7.1	8.8	10.4
$V+$jets modeling	3.8	9.7	13.6
$t\bar{t}$ modeling	2.2	3.7	3.3
$V+$jets flavour composition	1.2	2.9	2.9
Diboson modeling	0.9	1.9	2.1
Background MC stat.	9.8	18.0	24.4
Data stat.	6.6	20.8	45.3
Total syst.	20.9	40.1	49.4
Total	21.9	45.2	67.0

Table 12.2: Breakdown of expected signal strength uncertainties for representative $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ simplified vector mediator models with a dark matter mass $m_\chi = 1$ GeV and different vector mediator masses, representing kinematic regimes with low, medium and high $E_{\mathrm{T}}^{\mathrm{miss}}$. Only the largest systematic uncertainties are shown. Numbers are taken from [48] and not calculated by the author.

Table 12.2 shows the impact of the dominant sources of uncertainty on the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ signal strength. It can be understood as follows:

- The total systematic uncertainty dominates over the statistical uncertainty, meaning that the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis is systematically limited.
- The uncertainties with the largest impact correspond to finite Monte Carlo statistics and calibration of the large$-R$ jets. The latter is expected since the priority-merged selection is used, reflecting the fact that the most of the sensitivity comes from the boosted regime due to drastically decreasing SM backgrounds. This class of uncertainties has a significant impact on the reconstruction of the W/Z candidate. The impact of the large$-R$ jet uncertainties grows with the increasing $E_{\mathrm{T}}^{\mathrm{miss}}$, as the uncertainties grow with the jet $p_T \sim E_{\mathrm{T}}^{\mathrm{miss}}$. The large impact of the MC statistics comes from the poor statistics of the Standard Model backgrounds, especially $V+$jets, in the region of $p_T > 500$ GeV, and tends to increase with the higher mediator mass $m_{Z'}$.
- Another group of uncertainties with a sizeable impact is related to the normalisation of the main backgrounds, such as multijet, diboson (WW, WZ and ZZ) and

Source	Uncertainty on μ [%]		
	Z'-2HDM model $(m_{Z'}, m_A)$ [GeV]		
	$(600, 300)$	$(1400, 600)$	$(2000, 300)$
Large$-R$ jets	<0.1	1.2	4.7
Small$-R$ jets	1.7	3.8	2.1
$b-$tagging (track jets)	1.4	11	17
$b-$tagging (small$-R$ jets)	5.0	3.4	4.7
Luminosity	3.2	4.5	5.4
SM $Vh(b\bar{b})$ normalizaion	2.2	6.9	6.9
Diboson normalisation	1.1	0.7	2.2
$V+$jets modeling	5.0	5.7	8.2
Signal modeling	3.9	2.9	2.1
$t\bar{t}$, single-top modeling	3.2	3.0	3.9
Background MC stat.	4.9	11.0	22.0
Data stat.	6	37	61
Total syst.	10	21	36
Total	12	43	71

Table 12.3: Breakdown of expected signal strength uncertainties for representative $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ Z'-2HDM simplified model with different mass parameters, representing kinematic regimes with low, medium and high $E_{\mathrm{T}}^{\mathrm{miss}}$. Only the largest systematic uncertainties are shown. Numbers are taken from [46] and not calculated by the author.

$Z+$jets. The large impact of the last two uncertainties comes mainly from the $Z(\nu\nu)V(q\bar{q})$ and $Z(\nu\nu)+$jets processes, which represent irreducible backgrounds with a real $E_{\mathrm{T}}^{\mathrm{miss}}$ and jets in the final state. The $Z(\nu\nu)V(q\bar{q})$ process even provides a dijet system with the invariant mass within the W/Z mass window. The large impact of the multijet background comes from the region of $E_{\mathrm{T}}^{\mathrm{miss}} \lesssim 350$ GeV with no $b-$jets, in which sizeable signal and multijet contributions are expected.

- A relatively large impact of the electron uncertainties is due to the modeling uncertainties on the electron identification and isolation efficiencies. These uncertainties have a large impact on the veto of electrons in the SR and estimation of $W+$jets and $t\bar{t}$ backgrounds in the 1 lepton CR.

- Finally, the last class of uncertainties with a significant impact is related to the signal and $V+$jets modeling, reflecting the fact that it is challenging for MC generators to model the high tail of the p_T^V distribution. This uncertainty also tends to increase with the higher mediator mass $m_{Z'}$.

Table 12.3 show the impact of the dominant sources of uncertainty for the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis that can be understood as follows:

- Unlike the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis, the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis is statistically limited in the region of $E_{\mathrm{T}}^{\mathrm{miss}} \gtrsim 300$ GeV.

- The uncertainty with one of the largest impacts in both resolved and merged regimes is related to finite Monte Carlo statistics. This is again linked to the limited statistics of the Standard Model backgrounds, especially $V+$jets, in the region

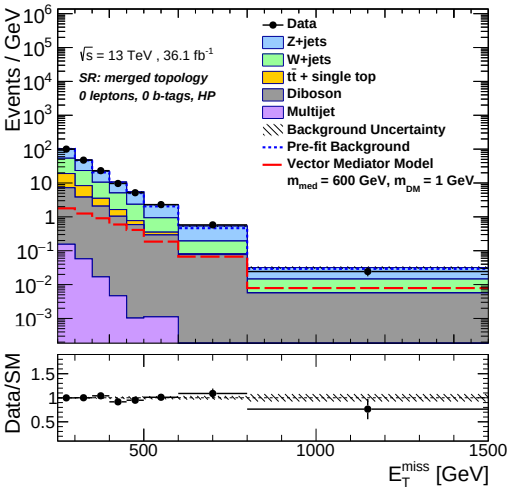
of $p_T > 500$ GeV. This uncertainty tends to increase with the higher mediator mass $m_{Z'}$.

- The uncertainty with the impact comparable to those of MC statistics correspond to $b-$tagging of small$-R$ jets in the resolved regime and track jets in the merged regime. Since two $b-$quark jets are expected in the final state, the uncertainties related to the $b-$tagging have a significant impact on the reconstruction of the h candidate. The impact grows rapidly with the higher $E_{\mathrm{T}}^{\mathrm{miss}}$ due to the large uncertainties on the $b-$tagging data-to-MC efficiency scale factors at the high $p_T \sim E_{\mathrm{T}}^{\mathrm{miss}}$ region [179].

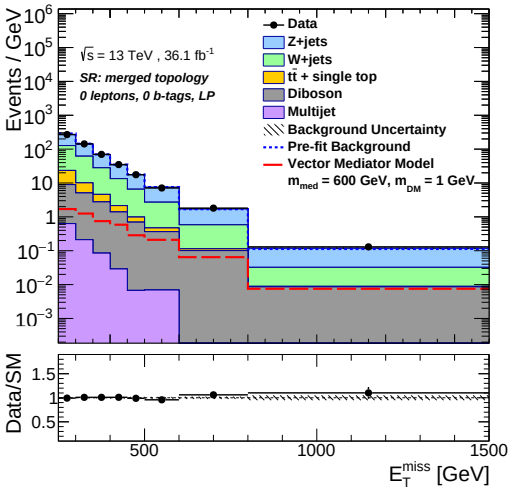
- Another groups of uncertainties with a sizeable impact correspond to the V+jets modeling and SM $Vh(b\bar{b})$ normalisation. The former originates from the MC mismodeling of the high tail of the p_T^V distribution. The latter originates mainly from the irreducible $Z(\nu\nu)h(b\bar{b})$ background that completely mimicks the signal final state. Both uncertainties tend to become larger with the higher mediator mass $m_{Z'}$.

12.1.3 Distributions after the fit to data

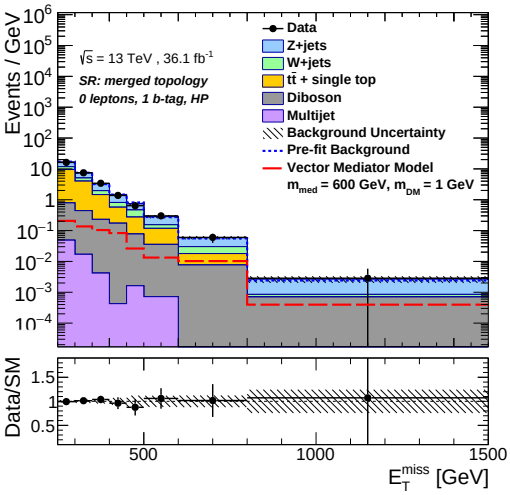
Figure 12.3 and 12.4 show the $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions after the fit to data, referred to as *post-fit* distributions, in the signal regions of the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis with different number of $b-$tags after the simultaneous fit of all regions to observed data. In general, there is a good agreement between the observed data and the SM backgrounds. A small excess in the observed data over the SM backgrounds is observed in the signal region with 1 $b-$tag in the 450 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 800 GeV range for the resolved regime, as shown on the top right plot in Figure 12.4. This excess cannot be assigned to the signal W/Z+DM production, otherwise the same kind of excess would show up in the 0 and 2 $b-$tag regions, which is not seen in the post-fit plots in the corresponding $b-$tag regions. Since no excess is observed in the mass sideband regions, as shown in Figure 12.6, it also cannot be assigned to MC mismodeling. Thus, this excess is rather a statistical fluctuation in data.
Figure 12.5 shows the post-fit distributions of the invariant mass of the Higgs boson candidate in the signal region of the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis with 2 $b-$tags. The post-fit distributions in the signal region of the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis with 1 $b-$tags is shown in Appendix E. The observed data is in the good agreement with the SM backgrounds, except the small statistical fluctuation in the observed data in the $E_{\mathrm{T}}^{\mathrm{miss}}$ region of 350 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 500 GeV for 2 $b-$tags around m_{jj} of 95 GeV.
The post-fit distributions of the muon charge and signal yield in the 1μCR and 2ℓCR are shown in Appendices D and F.



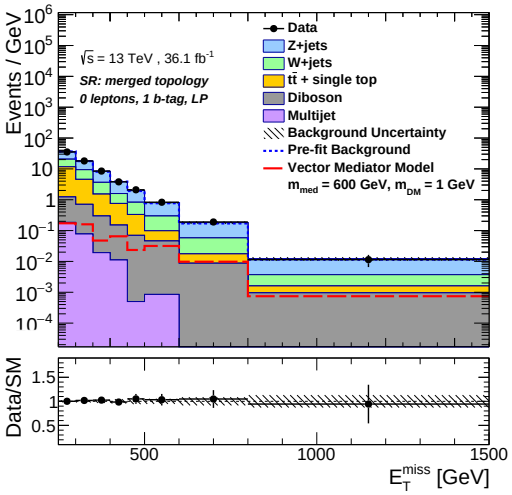
(a)



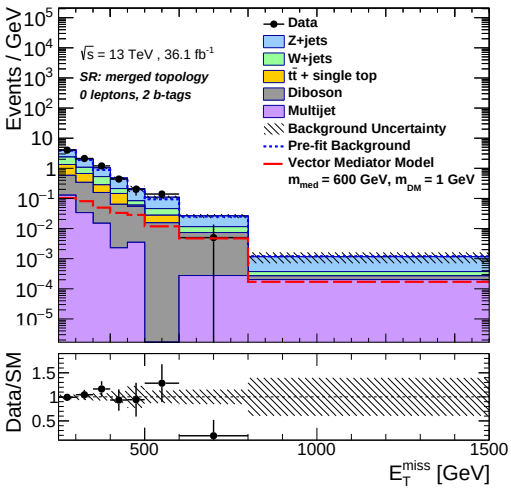
(b)



(c)



(d)



(e)

Figure 12.3: The observed (dots) and expected (histograms) distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$ in the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ signal region for the merged regime after the fit to observed data. Plots are shown separately for the (a) 0b-HP, (b) 0b-LP, (c) 1b-HP, (d) 1b-LP, and (e) 2b-tag event categories. The abbreviations HP and LP correspond to respectively the high- and low-purity analysis regions in the merged topology. The red dashed line shows the expected DM signal coming from the simplified vector mediator model with $(m_{\chi}, m_{Z'}) = (1, 600)$ GeV. The total background uncertainty is shown as a hatched band. The total background contribution before the fit to data is shown as a dotted blue line.

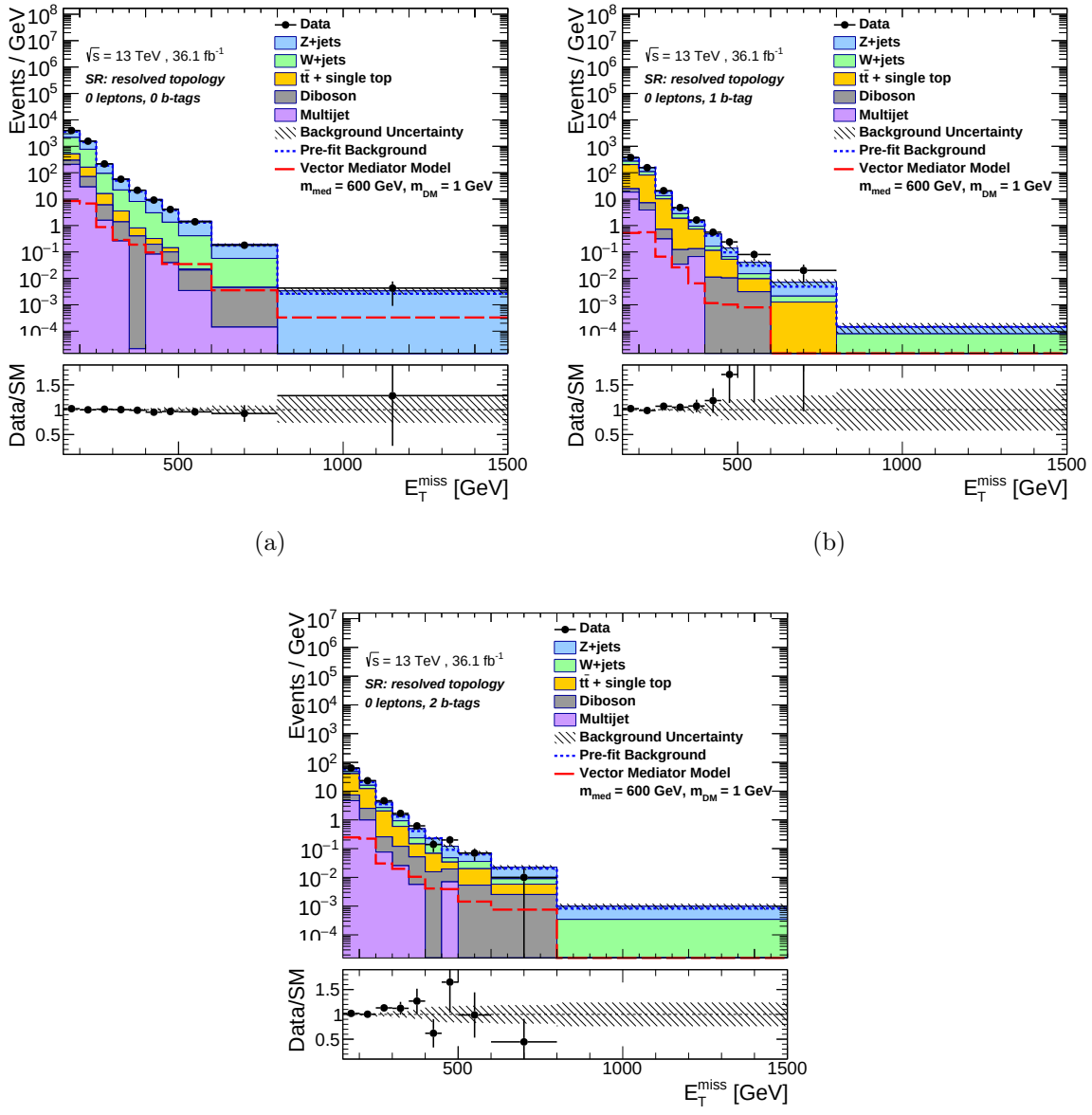


Figure 12.4: The observed (dots) and expected (histograms) distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$ in the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ signal region for the merged regime after the fit to observed data. Plots are shown separately for the (a) 0 $b-$tag, (b) 1 $b-$tag and (c) 2 $b-$tag event categories. The red dashed line shows the expected DM signal coming from the simplified vector mediator model with $(m_{\chi}, m_{Z'}) = (1, 600)$ GeV. The total background uncertainty is shown as a hatched band. The total background contribution before the fit to data is shown as a dotted blue line.

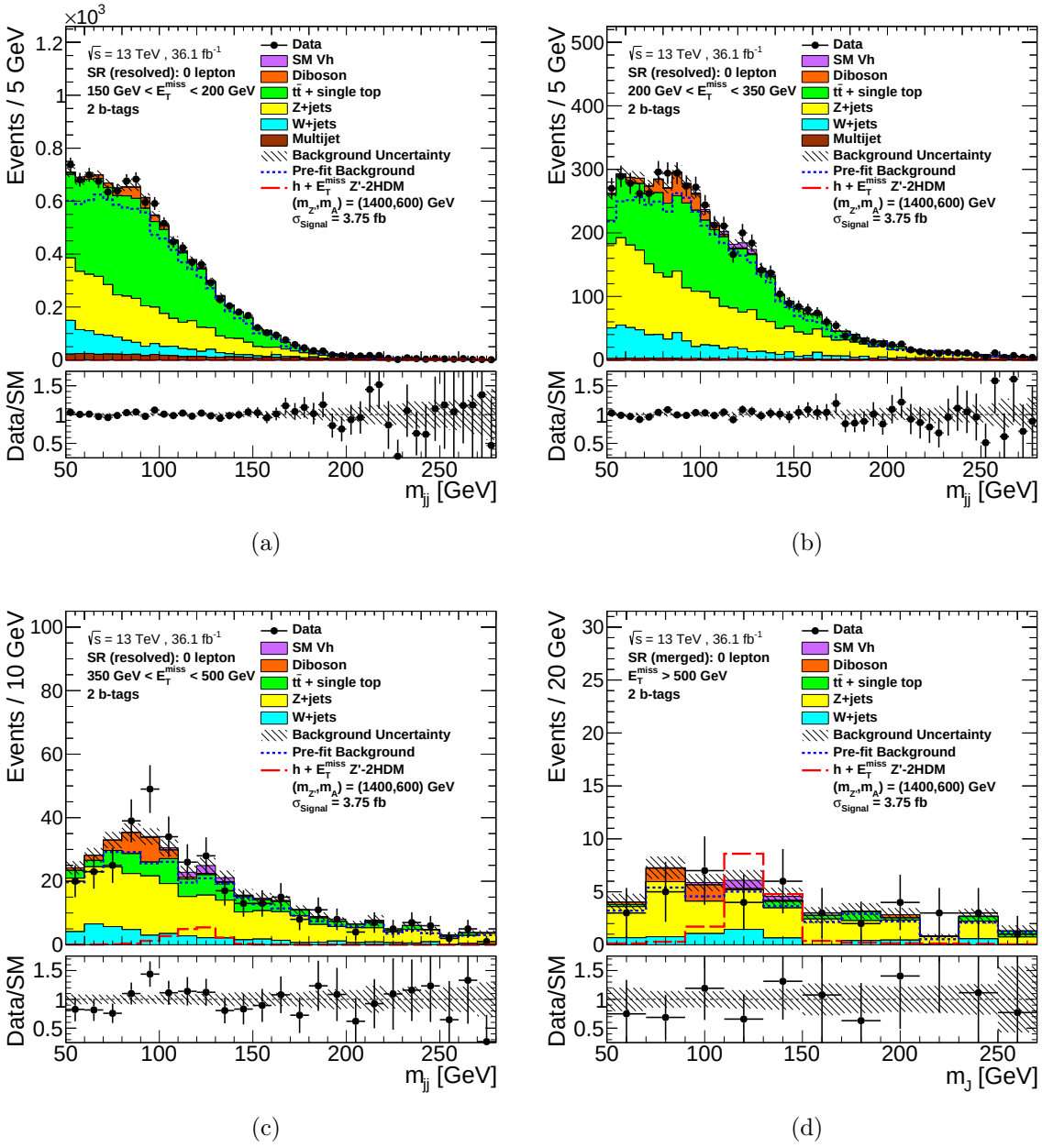


Figure 12.5: The observed (dots) and expected (histograms) distributions of the invariant mass of the Higgs boson candidates in the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ signal region with 2 $b-$tags after the fit to observed data. Plots are shown separately for the (a) $(150, 200)$ GeV, (b) $(200, 350)$ GeV, (c) $(350, 500)$ GeV and (d) $(500, \infty)$ $E_{\mathrm{T}}^{\mathrm{miss}}$ regions. The red dashed line shows the expected DM signal coming from the Z'-2HDM simplified model with $(m_{Z'}, m_A) = (1400, 600)$ GeV. The total background uncertainty is shown as a hatched band. The total background contribution before the fit to data is shown as a dotted blue line.

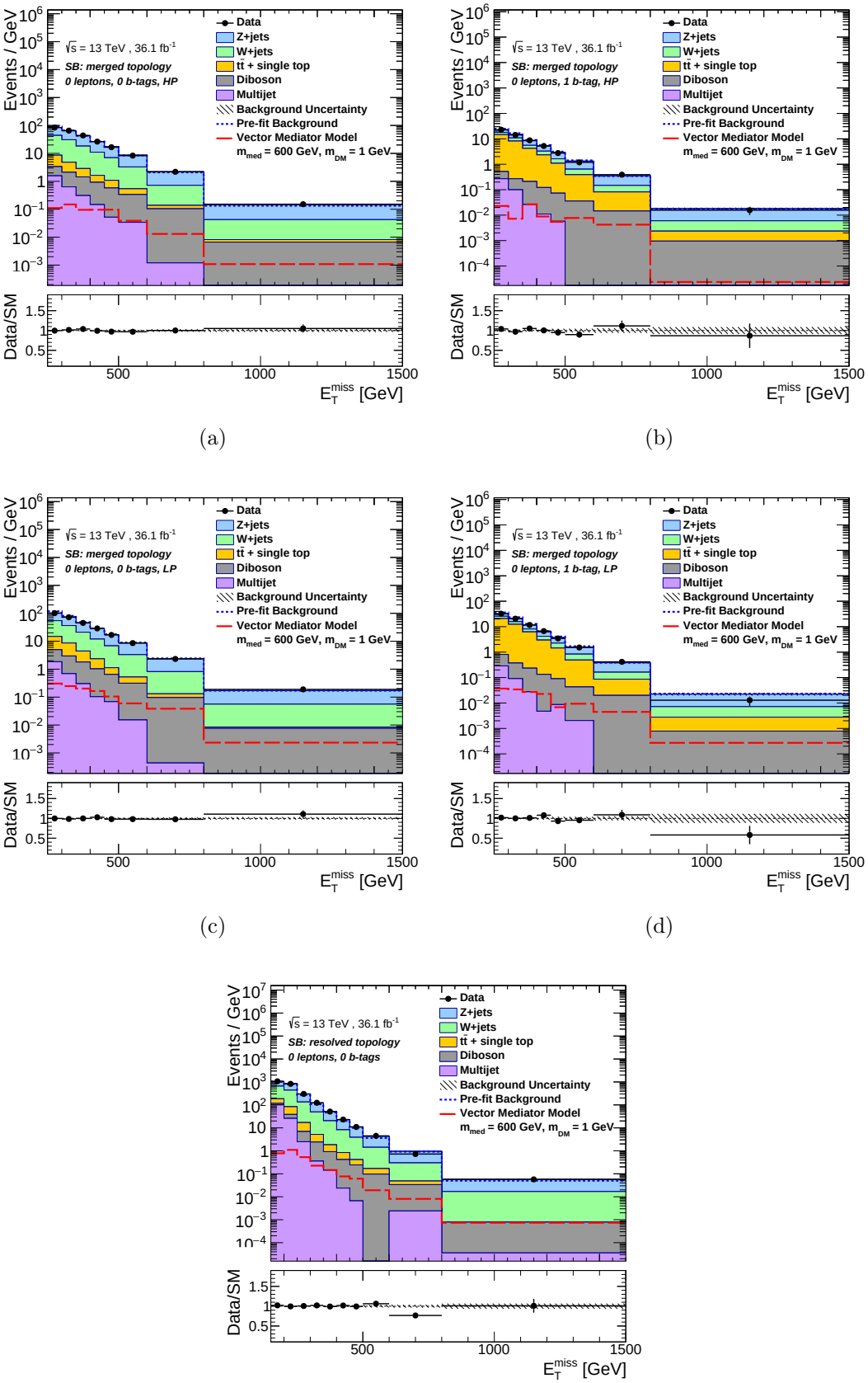


Figure 12.6: The observed (dots) and expected (histograms) distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$ in the mass sidebands of the $E_{\mathrm{T}}^{\mathrm{miss}}+V(qq)$ 0 lepton region after the fit to observed data. Plots are shown for the (a) 0b-D2, (b) 1b, (c) 0b-noD2, (d) 1b-noD2 and (e) 2b-resolved event categories, where D2 (noD2) cathegory includes events that passed (failed) $D_2^{\beta=1}$ requirement. The total background uncertainty is shown as a hatched band. The total background contribution before the fit to data is shown as a dotted blue line.

12.2 Exclusion limits on specific new physics models

As shown in Section 12.1, no evidence of the W/Z+DM and h+DM production is found in the observed data. In case of no evidence for the production of the new physics, the results of the search can be interpreted in terms of either the limits on the specific signal model of the DM production, or generic limits the visible production cross-section, as described in Chapter 11. In this Section, the key results in terms of both interpretations for the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analyses are discussed in details.

12.2.1 Model-dependent limits

A detailed discussion about a exclusion limit setting for a specific physics model is given in Section 11.1. The exclusion limits are set in the mass-mass plane of $(m_\chi, m_{Z'})$ and $(m_{Z'}, m_A)$ parameters for the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ simplified vector mediator and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ Z'-2HDM models, respectively. A given signal model is considered as excluded at 95% confidence level (CL) if the corresponding CL_s value is less than 0.05 (see Section 9.1.3).

$E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ exclusion limits

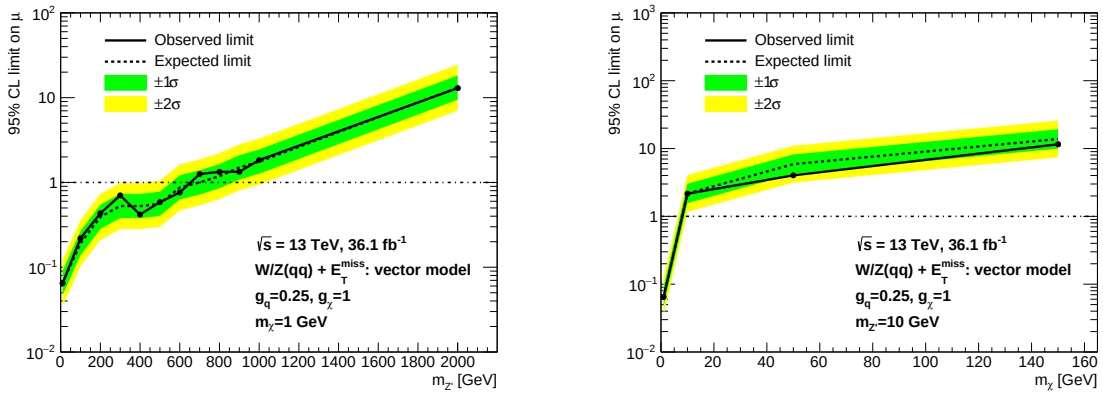


Figure 12.7: Linear interpolation of the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ upper limits on the signal strength as a function of $m_{Z'}$ for fixed $m_\chi = 1$ GeV (left) and as a function of m_χ for fixed $m_{Z'} = 10$ GeV (right). The observed (expected) limit is shown as a black solid (dashed) line. The $\pm 1\sigma$ ($\pm 2\sigma$) expected limits are shown as yellow (green) filled area. The dashed-dotted line denotes the $\mu = 1$ line. The model parameters $(m_\chi, m_{Z'})$ are excluded to the left of the crossing point of the exclusion curve and the $\mu = 1$ line.

Figure 12.7 shows upper limits on $\mu_{W/Z+\mathrm{DM}}$ as a function of the mediator mass $m_{Z'}$ and DM mass m_χ for a fixed $m_\chi = 1$ GeV and $m_{Z'} = 10$ GeV, respectively. The set of $(m_\chi, m_{Z'})$ parameters are excluded at 95% confidence level if the upper limit on the signal strength is below the $\mu = 1$ line.
The final exclusion limits on the W/Z+DM, W+DM and Z+DM production in the $(m_\chi, m_{Z'})$ plane for a simplified vector mediator model are shown in Figures 12.9 and 12.8. The observed limits are consistent with the expected limits within uncertainties. For a given choice of the couplings, the simplified vector mediator models with the mediator masses $m_{Z'}$ up to 650 GeV and the DM masses m_χ up to 250 GeV are excluded at 95% CL, improving the previous 2015 $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ results [235] by $15 - 30\%$. The wiggle shape of the exclusion limit curve near the kinematic limit originates from the large interpolation steps of 100 GeV in the direction of the constant DM mass m_χ. The

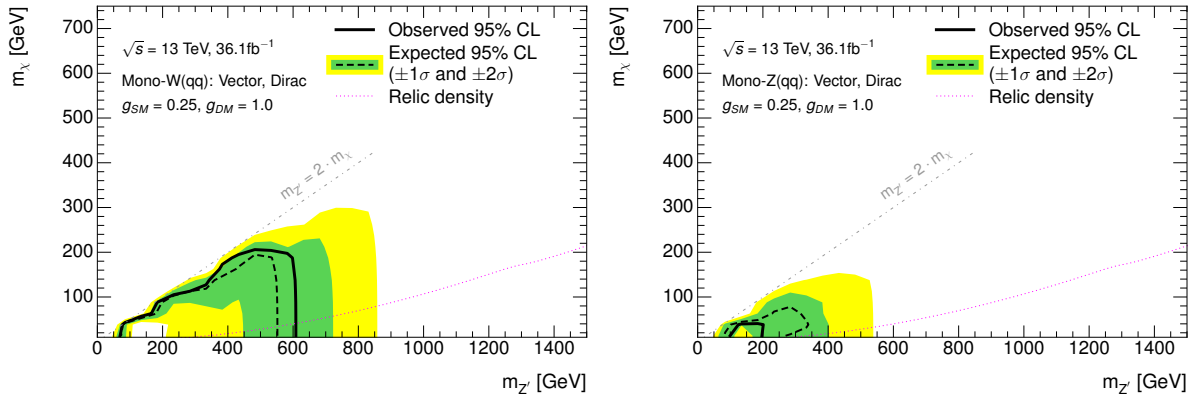


Figure 12.8: Exclusion limits on the W+DM (left) and Z+DM (right) production for the simplified vector mediator model in the $(m_\chi, m_{Z'})$ plane. Other model parameters are fixed to $g_q = 0.25$ and $g_\chi = 1$. The observed (expected) limit curves are shown as a black solid (dashed) line. The $\pm1\sigma$ ($\pm2\sigma$) uncertainties on the expected limits are shown as filled green (yellow) band. All cross-sections are obtained using the leading order level. The dotted magenta curves stands for the $(m_{Z'}, m_A)$ points for which the expected relic density, calculated using MadDM [229], is consistent with the observed density $\Omega h^2 = 0.12$ from the WMAP [27] and Planck [26] measurements. The region on the right of the curve corresponds to higher predicted relic abundance.

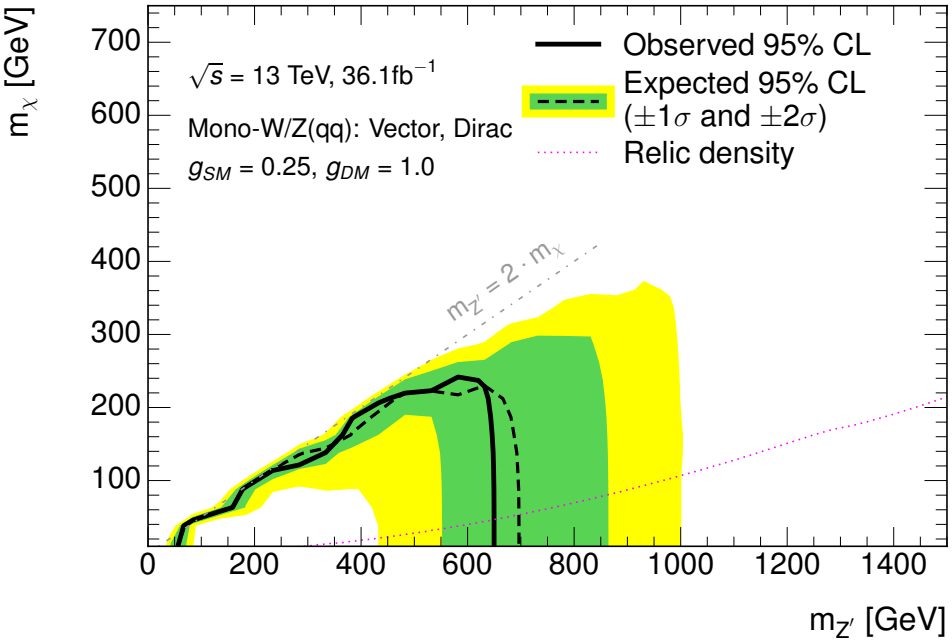


Figure 12.9: Exclusion limits on the W/Z+DM production for the simplified vector mediator model in the $(m_\chi, m_{Z'})$ plane. Other model parameters are fixed to $g_q = 0.25$ and $g_\chi = 1$. The observed (expected) limit curve is shown as a black solid (dashed) line. The $\pm1\sigma$ ($\pm2\sigma$) uncertainty on the expected limits is shown as filled green (yellow) band. All cross-sections are obtained at leading-order. The dotted magenta curve stands for the $(m_{Z'}, m_A)$ points for which the expected relic density, calculated using MadDM [229], is consistent with the observed density $\Omega h^2 = 0.12$ from the WMAP [27] and Planck [26] measurements. The region on the right of the curve corresponds to higher predicted relic abundance.

kinematic acceptances and efficiencies are similar in the direction of increasing DM mass m_χ, meaning the upper limits are also similar in that direction. This explains sharp falling edges of the observed and expected limit curves in the region of $650 - 700$ GeV. The $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ limits are totally dominated by the $E_{\mathrm{T}}^{\mathrm{miss}} + W(qq)$ limits.

$E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ exclusion limits

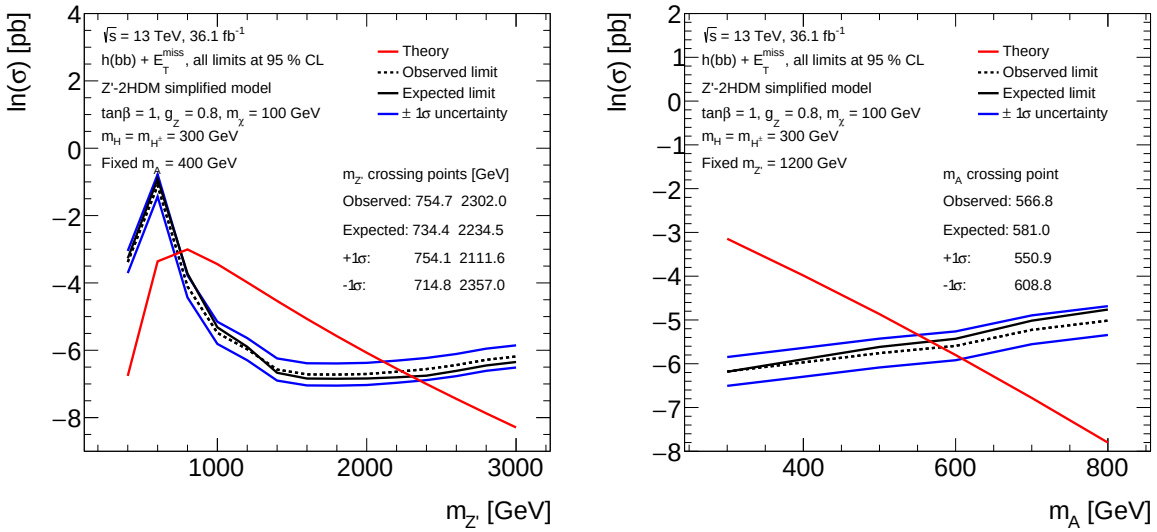


Figure 12.10: Linear interpolation of the upper limits on the Z'-2HDM production cross-section in the logarithmic scale as a function of $m_{Z'}$ for a fixed $m_A = 400$ GeV (left) and m_A for a fixed $m_{Z'} = 1200$ GeV (right). The observed (expected) limit is shown as a black solid (dashed) line. Theoretical cross-section is shown as a red line. The $\pm 1\sigma$ expected limits are shown as the blue lines. The part of the phase space, in which the upper limit curve is below the theory prediction curve, is excluded at 95% CL.

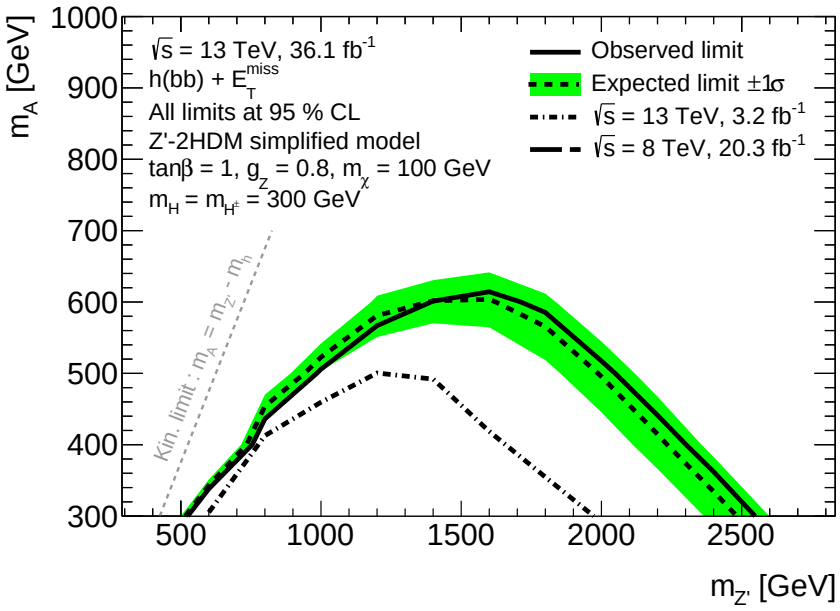


Figure 12.11: Exclusion limits on the h+DM production for the Z'-2HDM model in the $(m_{Z'}, m_A)$ plane. Other model parameters are fixed to $\tan\beta = 1$, $g_{Z'} = 0.8$, $m_\chi = 100$ GeV, and $m_H = m_{H^\pm} = 300$ GeV. The observed (expected) limit curve is shown as a black solid (dashed) line. The $\pm 1\sigma$ uncertainties on the expected limits are shown as a filled green band. The observed exclusion countour from the previous ATLAS results for 2015 data at $\sqrt{s} = 13$ TeV [45] is shown as a dashed-dotted line.

Figure 12.10 illustrates the interpolated theoretical cross-section curves and upper limit curves, as well as the corresponding crossing points, for fixed $m_A = 400$ GeV and for fixed

$m_{Z'} = 1200$ GeV.
The final Z'-2HDM exclusion contour in the $(m_{Z'}, m_A)$ phase space is shown in Figure 12.11. The observed limits are consistent with the expected limits with uncertainties. For a given choice of the Z'-2HDM model parameters, the mediator masses $m_{Z'}$ up to 2.6 TeV and the pseudoscalar masses m_A up to 600 GeV are excluded at 95% CL, significantly improving the exclusion limits from the previous versions of the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis [44, 45]. The exclusion limits degrades for the mediator masses $m_{Z'}$ above 1.5 TeV. One reason is that track jets associated to the large$R$ jet from the $h \to b\bar{b}$ decay tend to merge in the regions of very high boost, which leads to the loss in sensitivity in the regions with 2 $b-$tags. Another reason is the reduced sensitivity in the regions of low $E_{\mathrm{T}}^{\mathrm{miss}}$ due to the hard $E_{\mathrm{T}}^{\mathrm{miss}}$ spectrum of Z'-2HDM models in the region of high masses.

12.2.2 NLO rescaling of $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ model-dependent limits

The $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ exclusion limits for the simplified vector mediator model shown in the previous section are calculated using the MC signal samples generated at leading-order (LO), since no MC simulations were available at next-to-leading-order (NLO). Meantime, other searches, such as $E_{\mathrm{T}}^{\mathrm{miss}} + j$ [30], $E_{\mathrm{T}}^{\mathrm{miss}} + \gamma$ [236], and $E_{\mathrm{T}}^{\mathrm{miss}} + Z$ [237], set limits for the same model but using the signal samples generated at NLO [112]. It means that the current $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ limits cannot be directly compared to other $E_{\mathrm{T}}^{\mathrm{miss}} + X$ and dijet searches, as shown on the summary plot in Ref. [238]. Therefore, a rescaling procedure is introduced to obtain the limits for the simplified vector mediator model at NLO using already exisiting limits at LO. It also allows to obtain the limits at NLO for the different coupling strengths and types of the mediator, such as axial-vector mediator. More details on the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ rescaling procedure are in the summary paper of the ATLAS DM searches at the LHC [49, 50]. The procedure is summarised below.
The MC samples at NLO are generated using MadGraph5_aMC@NLO+Pythia 8 MC generator [148, 239] with the DMsimp [240] implementation of the simlified model at NLO. The parton distribution function set used is NNPDF3.0 at NLO with $\alpha_s = 0.118$ [143]. The vector mediator signal samples at LO that are used to obtain the existing $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ limits are rescaled to NLO for the following four coupling and mediator scenarios, as discussed in Section 11.1.1:

- **Vector mediator (V1):** $g_q = 0.25,\ g_l = 0,\ g_\chi = 1.$
- **Axial-vector mediator (A1):** $g_q = 0.25,\ g_l = 0,\ g_\chi = 1.$
- **Vector mediator (V2):** $g_q = 0.01,\ g_l = 0.1,\ g_\chi = 1.$
- **Axial-vector mediator (A2):** $g_q = 0.1,\ g_l = 0.1,\ g_\chi = 1.$

The impact in change from the LO and NLO and to other coupling strengths is estimated using the particle-level [1] (truth-level) information from the generated MC samples. More precisely, the rescaling is based on the change in the total cross-section, σ, and the kinematic acceptance times efficiency, $A \times \varepsilon$. The latter is estimated using the particle-level distributions of the final state kinematic variables, such as $E_{\mathrm{T}}^{\mathrm{miss}}$.

[1] The particle level represents the simulation stage, at which stable particles with a life-time $\tau > 10\mu$s and reconstructed jets after the processes of parton showering and hadronisation are considered.

The change in the cross-section is expressed by $k-$factor k_σ that is defined as a ratio of the cross-section for a given scenario at NLO to the cross-section for the V1 scenario at LO used to calculate the existing limits:

$$k_\sigma^{\mathrm{V1,V2,A1,A2}} \equiv \sigma_{\mathrm{NLO}}^{\mathrm{V1,V2,A1,A2}} / \sigma_{\mathrm{LO}}^{\mathrm{V1}} \,. \tag{12.2}$$

Figure 12.12 shows the k_σ scale factors for the V1 and A1 scenarios at NLO. The observed cross-sections for both scenarios at NLO are around $30-40\%$ larger than the value for the corresponding scenarios at LO, except the off-shell region for A1 scenario, where the cross-section at NLO decreases. A similar behaviour is observed in the $E_{\mathrm{T}}^{\mathrm{miss}} + Z(ll)$ search [49, 50]. The uncertainty on the k_σ scale factors are from finite MC statistics and shown to be at the sub-percent level (see Figure G.1 in Appendix G). The k_σ plots for other scenarios are quantatively similar, as shown in Appendix G.

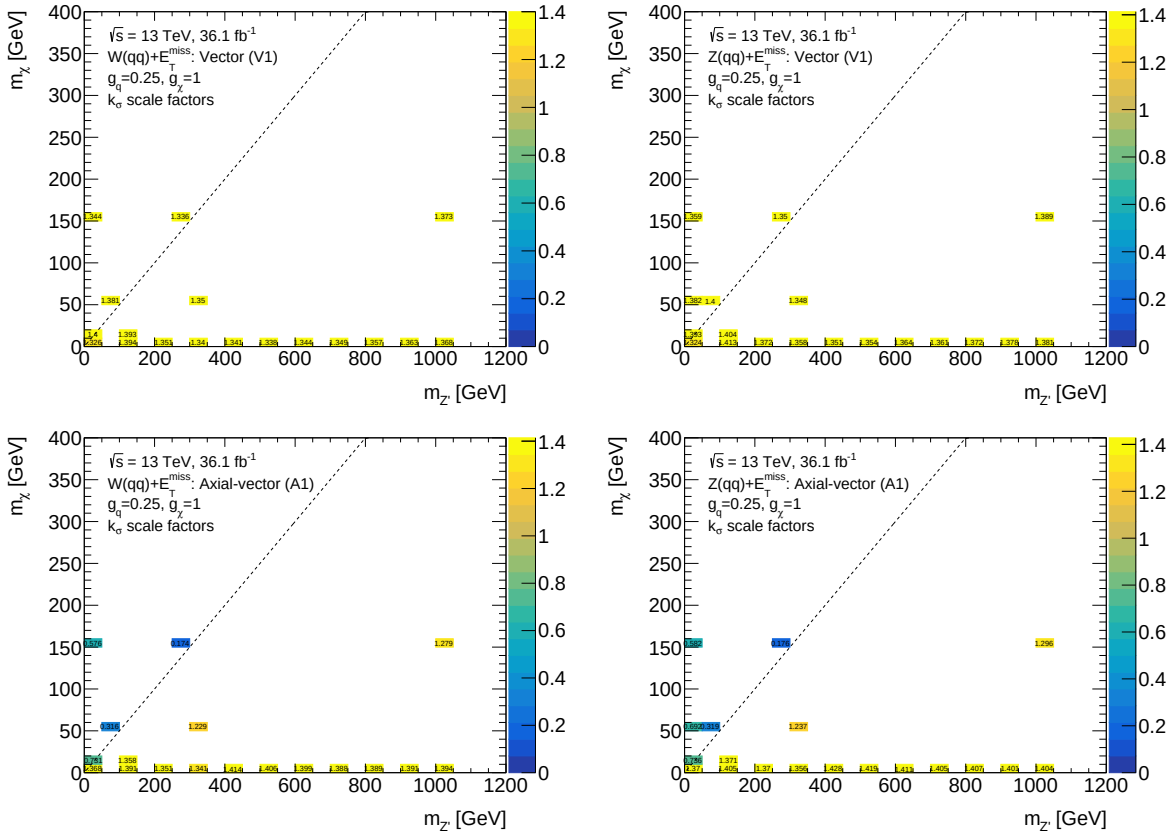


Figure 12.12: The k_σ scale factors for simplified vector (top) and axial-vector (bottom) mediator models of the $E_{\mathrm{T}}^{\mathrm{miss}} + W(qq)$ (left) and $E_{\mathrm{T}}^{\mathrm{miss}} + Z(qq)$ (right) production are shown in the $(m_\chi, m_{Z'})$ plane. The dashed black line shows the kinematic limit $m_{Z'} = 2m_\chi$. Only the mass points where fully simulated MC sampls exist are shown.

The kinematic $A \times \varepsilon$ is defined as a ratio of the number of generated events after the truth-level selection to the total number of generated events:

$$A \times \varepsilon \equiv \frac{N_{\mathrm{evt}}^{\mathrm{truth}} \text{ after truth-level selection}}{\text{total } N_{\mathrm{evt}}^{\mathrm{truth}}} \,, \tag{12.3}$$

where the truth selection is presented in Table 12.4.
The $A \times \varepsilon$ values are calculated at particle level using MC distributions generated for

Region	Resolved regime	Merged regime
0 lepton	Baseline selection	
	Veto electrons with $p_T > 7$ GeV for $\|\eta\| < 2.47$	
	Veto muons with $p_T > 7$ GeV for $\|\eta\| < 2.7$	
	$\Delta\phi(\mathbf{E_T^{miss}}, \mathbf{p_T^{W/Z}}) > 6\pi/9$	
	$E_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV	$E_{\mathrm{T}}^{\mathrm{miss}} > 250$ GeV
	$\geq$ central $2j$	$\geq 1J$
	$p_T^{j_1} > 45$ GeV	
	$\sum_{i=1}^{2(3)} p_T^{j_i} > 120(150)$ GeV	
	$\Delta\phi(j_1, j_2) < 7\pi/9$	

Table 12.4: Summary of the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ truth-level event selection used to derive $A \times \varepsilon$ for the NLO re-scaling procedure. Details about the event selection at detector level are discussed in Chapter 7.

the same $(m_\chi, m_{Z'})$ points as the fully simulated MC samples used for the existing limit setting at LO. Figure 12.13 illustrates the comparison of the V1 $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions of the $E_{\mathrm{T}}^{\mathrm{miss}} + W(qq)$ production at LO and NLO for the $(m_\chi, m_{Z'}) = (1, 100)$ GeV mass point. The difference increases from $\sim 10\%$ in the low $E_{\mathrm{T}}^{\mathrm{miss}}$ region to almost 100% in the $E_{\mathrm{T}}^{\mathrm{miss}}$ region above 200 GeV with the poor MC statistics.

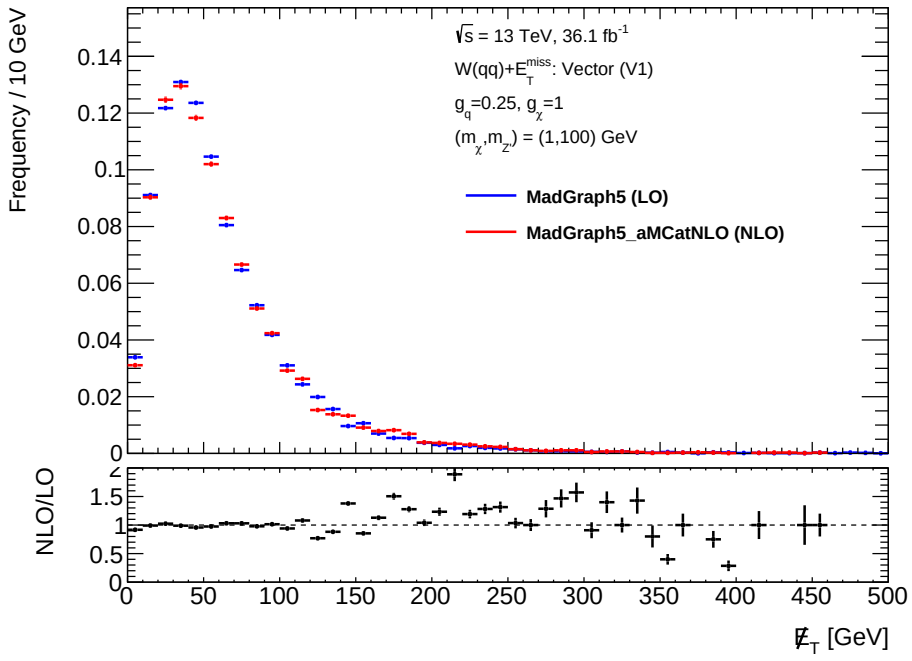


Figure 12.13: Comparison of the normalised $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions for the $E_{\mathrm{T}}^{\mathrm{miss}} + W(qq)$ simplified vector mediator model at LO and NLO for the set of $(m_\chi, m_{Z'}) = (1, 100)$ GeV parameters is shown. The distributions of the models at LO (NLO) are shown in blue (red). The signal model samples at LO and NLO are generated with MadGraph5 and MadGraph5_aMC@NLO, respectively.

The change in the $A \times \varepsilon$ is expressed by the $k-$factor $k_{A\times\varepsilon}$ that is defined as a ratio of the $A \times \varepsilon$ for a given scenario at NLO to the $A \times \varepsilon$ for the V1 scenario at LO used to calculate the existing limits:

$$k_{A\times\varepsilon}^{\mathrm{V1,V2,A1,A2}} \equiv (A \times \varepsilon)_{\mathrm{NLO}}^{\mathrm{V1,V2,A1,A2}} / (A \times \varepsilon)_{\mathrm{LO}}^{\mathrm{V1}} \,. \tag{12.4}$$

Figure 12.14 shows the $k_{A\times\varepsilon}$ scale factors with the corresponding uncertainties for the V1 and A1 scenarios at NLO. The difference between the $A \times \varepsilon$ at LO and NLO is not larger than 15% for the vector scenario and axial-vector scenatio in the on-shell region with an absolute uncertainty of up 4%. The off-shell axial-vector scenarios show large difference in $A \times \varepsilon$ up to 50% in the region of low $(m_\chi, m_{Z'})$ masses. Similar behaviour is observed in the $E_{\mathrm{T}}^{\mathrm{miss}} + Z(ll)$ search [49, 50]. The $k_{A\times\varepsilon}$ plots for other scenarios are shown in Appendix G.

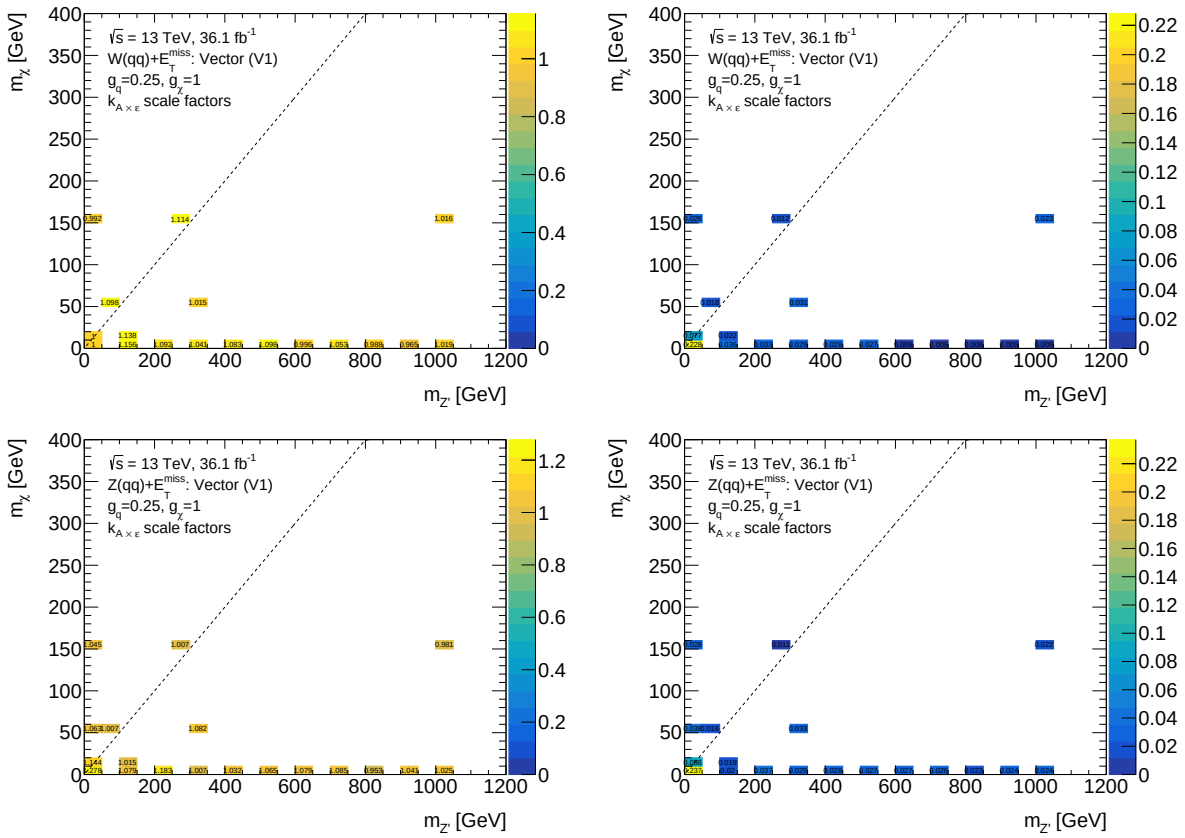


Figure 12.14: The $k_{A\times\varepsilon}$ scale factors (left) and corresponding uncertainties (right) for for the $E_{\mathrm{T}}^{\mathrm{miss}} + W(qq)$ (top) and $E_{\mathrm{T}}^{\mathrm{miss}} + Z(qq)$ (bottom) simplified vector mediator models (V1) are shown in the $(m_\chi, m_{Z'})$ plane. The dashed black line shows the kinematic limit $m_{Z'} = 2m_\chi$. Only the mass points where fully simulated MC samples exist are shown.

The final rescaling factor k_{NLO} that is used to rescale the limits on the signal strength is defined as a product of the cross-section and $A \times \varepsilon$ scale factors:

$$k_{\mathrm{NLO}}^{\mathrm{V1,V2,A1,A2}} = k_{\sigma}^{\mathrm{V1,V2,A1,A2}} \times k_{A\times\varepsilon}^{\mathrm{V1,V2,A1,A2}} . \tag{12.5}$$

The MC signal samples of the $E_{\mathrm{T}}^{\mathrm{miss}}$+$W(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}}$+$Z(qq)$ production can be generated only separately. However, the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ limits are dominated by the $E_{\mathrm{T}}^{\mathrm{miss}} + W(qq)$ limits, as shown in Figures 12.9 and 12.8. Given that the final rescaling factors k_{NLO} for the $E_{\mathrm{T}}^{\mathrm{miss}} + W(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + Z(qq)$ production are consistent within uncertainties, one can use $E_{\mathrm{T}}^{\mathrm{miss}} + W(qq)$ factors to rescale the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ limits from LO to NLO.
Figure 12.15 illustrates the combined $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ limits for the V1 and A1 scenarios at NLO. The limit curve for the V1 scenario almost reaches the kinematic limit $m_{Z'} = 2m_\chi$, excluding the signal models with the mediator mass up to 830 GeV. The alignment of the limit curve along the kinematic limit can be explained by the moderate dependence

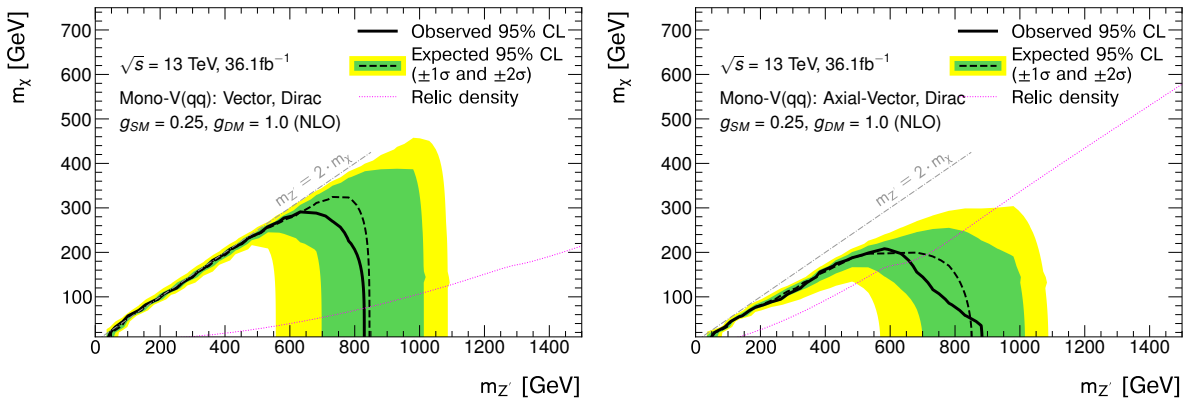


Figure 12.15: Exclusion limits on the V+DM production for the simplified vector (left) and axial-vector (right) mediator models at NLO in the $(m_\chi, m_{Z'})$ plane. The coupling strengths are fixed to $g_q = 0.25$ and $g_\chi = 1$. The observed (expected) limit curves are shown as a black solid (dashed) line. The $\pm1\sigma$ ($\pm2\sigma$) uncertainties on the expected limits are shown as filled green (yellow) band. All cross-sections are obtained at next-to-leading-order. The dotted magenta curves stands for the $(m_{Z'}, m_A)$ points for which the expected relic density, calculated using MadDM [229], is consistent with the observed density $\Omega h^2 = 0.12$ from the WMAP [27] and Planck [26] measurements. The region on the right of the curve corresponds to higher predicted relic abundance.

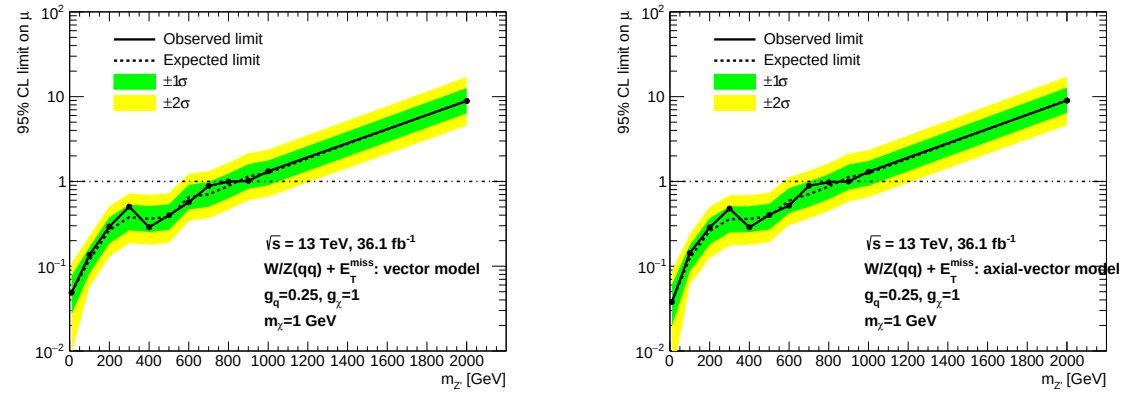


Figure 12.16: Linear interpolation of the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ upper limits on the signal strength for the V1 (left) and A1 (right) scenarios at NLO as a function of $m_{Z'}$ for fixed $m_\chi = 1$ GeV. The observed (expected) limit is shown as a black solid (dashed) line. The $\pm1\sigma$ ($\pm2\sigma$) expected limits are shown as yellow (green) filled area. The dashed-dotted line denotes the $\mu = 1$ line. The model parameters $(m_{Z'}, m_A)$ are excluded in the part of the phase space, in which the theoretical cross-section is larger than the upper limit.

of the decay width of the mediator on the DM mass m_χ (see Eq. 3.17 in Section 3.4.1.1). Given the large branching ratio and $\mu \propto \mathcal{B}_{Z'\to\chi\chi}$ dependence, the signal strength naturally decreases close to the kinematic limit as the decay width goes to zero. However, $\mathcal{B}_{Z'\to\chi\chi}$ for the axial-mediator model decreases more strongly than for the vector mediator model with increasing m_χ (see Eq. 3.17 in Section 3.4.1.1). It manifests itself in much stronger deviation of the limit curve from the kinematic limit relative to the vector mediator model.

The behavior of the observed limit for both scenarios in the range of mediator masses 700 GeV $< m_{Z'} <$ 900 GeV can be explained by the shape of the observed limits for a fixed DM mass $m_\chi = 1$ GeV shown in Figure 12.15. The observed limit curve happens to be almost flat in the range of mediator masses 700 GeV $< m_{Z'} <$ 900 GeV and very close

to $\mu = 1$ due to the statistical fluctuations in data. The branching ratios of the $Z' \to \chi\chi$ decay are very similar for the $m_\chi \lesssim m_{Z'}$. Since the interpolation is performed along the lines of the fixed mediator mass, this flat behaviour translates to the shape of the observed limits in the $(m_\chi, m_{Z'})$ plane in the range of mediator masses 700 GeV $< m_{Z'} <$ 900 GeV. One cannot exclude any models for the V2 and A2 scenarios due to the small signal cross-sections that result from the small coupling strengths.

Figures 12.17 show summary figures of the regions of the $(m_\chi, m_{Z'})$ plane for the simplified vector mediator model with V1 coupling scenario and axial-vector mediator model with A1 coupling scenario excluded at 95% CL by various $E_{\mathrm{T}}^{\mathrm{miss}} + X$ and dijet searches for the DM at ATLAS. The $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ exclusion limits at NLO, shown in Figure 12.15, are included in this figure. The summary figures show larger exclusion power of the $E_{\mathrm{T}}^{\mathrm{miss}} + j$ and $E_{\mathrm{T}}^{\mathrm{miss}} + \gamma$ limits relative to the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ limits for the $s-$channel mediator model. In case of the $E_{\mathrm{T}}^{\mathrm{miss}} + j$ search, this is related to the larger rate of events with gluon radiation comparing to the electroweak radiation of W/Z bosons due to the different coupling strengths, $\alpha_s \gg \alpha_{\mathcal{EW}}$. In case of the $E_{\mathrm{T}}^{\mathrm{miss}} + \gamma$ search, the couplings are similar, $\alpha_{em} \sim \alpha_{\mathcal{EW}}$, but the lower SM backgrounds and absence of the photon decay, meaning that effective branching ratio is 100%, lead to larger exclusion power of $E_{\mathrm{T}}^{\mathrm{miss}} + \gamma$ limits. However, for a model where W or Z boson comes from the hard interaction rather than from the ISR, such as 2HDM+a model, the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ search would provide better sensitivity than the $E_{\mathrm{T}}^{\mathrm{miss}} + j$ search, as discussed in Section 3.4.1.3. At the same time, the mediator can always decay back to a pair of quarks, leading to the dijet signature for the simplified $s-$channel mediator models. The width of the $Z' \to q\bar{q}$ decay does not depend on the DM mass m_χ and mediator coupling strength to the DM particles g_χ. This leads to the limits on the dijet production that are set in the phase space of mediator mass $m_{Z'}$ and coupling strength g_q. Given the fixed $g_q = 0.25$ and no dependence on m_χ, the dijet exclusion limit curves in $(m_\chi, m_{Z'})$ plane take the form of the lines of a constant mediator mass $m_{Z'}$.

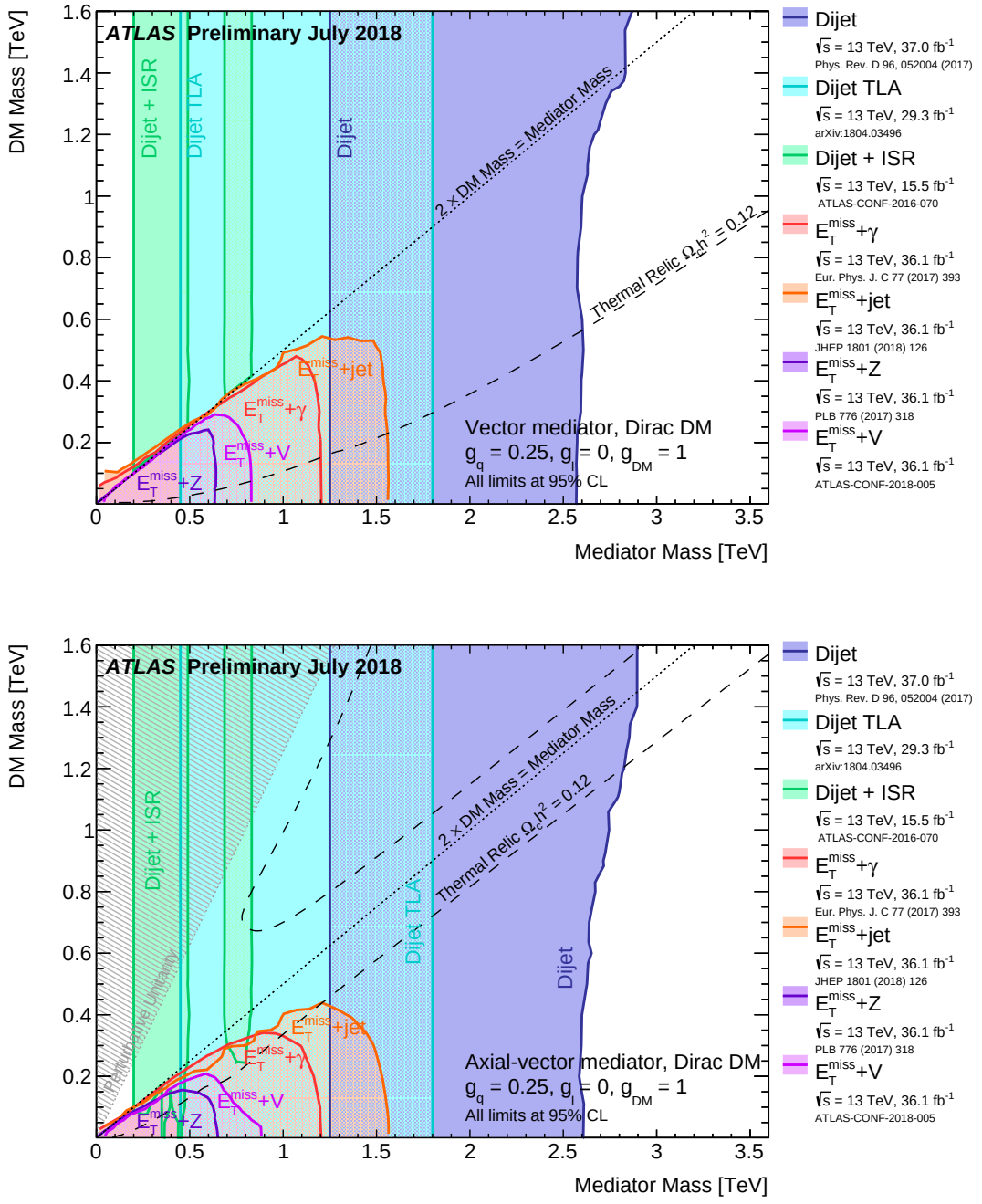


Figure 12.17: Summary plot of the regions in the $(m_\chi, m_{Z'})$ plane for the simplified vector (top) and axial-vector (bottom) mediator models with respectively V1 and A1 coupling scenarios at NLO excluded at 95% CL by a set of ATLAS $E_{\mathrm{T}}^{\mathrm{miss}}+X$ and dijet searches for the DM production [50]. The $E_{\mathrm{T}}^{\mathrm{miss}}+V(qq)$ exclusion limits, shown in Figure 12.15, are represented by the solid magenta contour. The exclusion limits are derived using data collected in 2015 and 2016 at $\sqrt{s} = 13$ TeV. The limits from the dijet searches are obtained from the limits on the Gaussian-shaped resonances, as described in Ref. [241]. The dotted black curve shows the kinematic limit of the on-shell mediator decays into DM. The dashed black curve labeled "Thermal relic" stands for the $(m_{Z'}, m_A)$ points for which the expected relic density, calculated using MadDM [229], is consistent with the observed density $\Omega h^2 = 0.12$ from the WMAP [27] and Planck [26] measurements. The regions on the right of the lower curves and on the top of the upper curve correspond to higher predicted relic abundance. Excluded regions that are in tension with the perturbative unitary considerations [227] are indicated by the shaded area in the upper left part.

12.3 Generic limits

The generic limit setting procedure is discussed in great details in Section 11.2. This Section is aimed to discuss the key results of the generic limit setting.

Generic limits on production of $W + \mathrm{DM}$ and $Z + \mathrm{DM}$ events

The limits on the visible production cross-section $\sigma_{\mathrm{vis},W+\mathrm{DM}}$ and $\sigma_{\mathrm{vis},Z+\mathrm{DM}}$ for a range of the simplified vector mediator models in the lowest and highest $E_{\mathrm{T}}^{\mathrm{miss}}$ regions are shown in Tables 12.5 and 12.6. The limits in other $E_{\mathrm{T}}^{\mathrm{miss}}$ regions are given in Appendix H. The relative variations in the expected limits on the $Z + \mathrm{DM}$ production are on average larger comparing to the variations for $W + \mathrm{DM}$ production, amounting to respectively 50% and 15% in the lowest $E_{\mathrm{T}}^{\mathrm{miss}}$ region. This difference comes from the sensitivity to the regions with different selection efficiency. The highest sensitivity to the $W + \mathrm{DM}$ events comes from the regions with 0 $b-$tags, which are characterised by the high selection effeciency. Meantime, the highest sensitivity to the $Z+\mathrm{DM}$ events comes from the regions with 2 $b-$tags, where the statistical fluctuations are larger due to the small selection efficiency. The difference between relative variations is the largest in the regions of low $E_{\mathrm{T}}^{\mathrm{miss}}$, and decreases with the increasing $E_{\mathrm{T}}^{\mathrm{miss}}$. inrelative variations become comparable in the regions of high $E_{\mathrm{T}}^{\mathrm{miss}}$ due to the generally low MC statistic . The largest relative variations in a given $E_{\mathrm{T}}^{\mathrm{miss}}$ region are in range of 15%-50% (25-50%) for the $W + \mathrm{DM}$ ($Z + \mathrm{DM}$) production.
The resulting generic limits on the visible cross-sections $\sigma_{\mathrm{vis},W+\mathrm{DM}}$ and $\sigma_{\mathrm{vis},Z+\mathrm{DM}}$ in regions of $E_{\mathrm{T}}^{\mathrm{miss}}$ correspond to the weakest expected limits in Tables 12.5, H.1, H.2, H.3, H.4 and 12.6. They are shown in Figures 12.18 and 12.19 and summarised in Tables 12.7 and 12.8. The observed limits are consistent with the expected limits within uncertainties. As a general trend, the generic limits on $Z + \mathrm{DM}$ production are stronger than the limits on the $W + \mathrm{DM}$ production, since the $Z + \mathrm{DM}$ events mostly populate the analysis regions with 2 $b-$tags with the highest sensitivity due to the very low SM backgrounds. The limits become stronger with increasing $E_{\mathrm{T}}^{\mathrm{miss}}$ due to the drastically falling SM backgrounds in the regions of high $E_{\mathrm{T}}^{\mathrm{miss}}$. Small excesses in the generic limits can be linked to the corresponding excesses in the post-fit distributions presented in Section 12.1.3. For instance, the $\sim 1.7\sigma$ excess in limits on $Z + \mathrm{DM}$ production in the region of 300 GeV $\leq E_{\mathrm{T}}^{\mathrm{miss}}$400 GeV can be traced back to the small excesses in observed data in the same $E_{\mathrm{T}}^{\mathrm{miss}}$ region for events in the resolved regime with 1 and 2 $b-$tags (see Figure 12.4) and in the merged regime with 2 $b-$tags (see Figure 12.3).

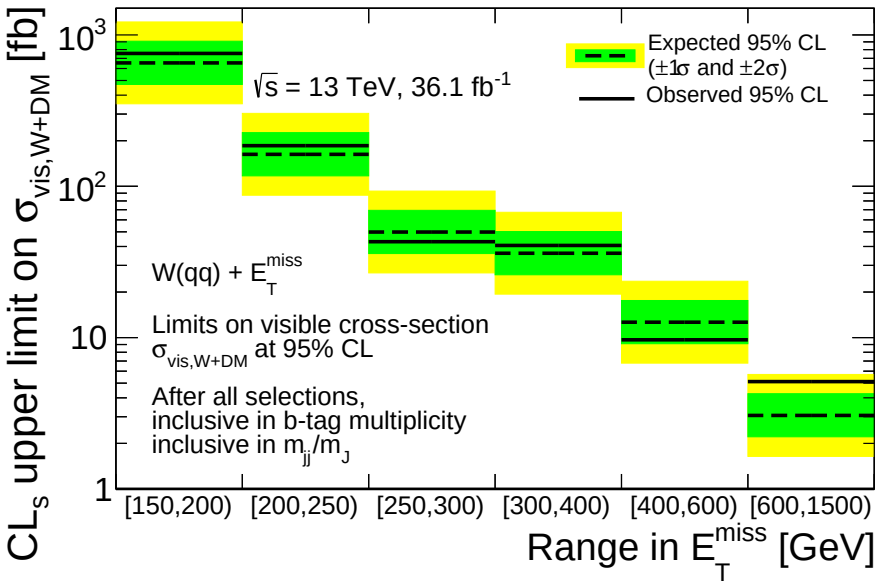


Figure 12.18: Upper limits at 95% CL on the visible cross-section of the W+DM events in regions of $E_{\mathrm{T}}^{\mathrm{miss}}$ after the full $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ event selection except the $b-$tagging and W candidate mass requirements. The observed (expected) limits are shown as solid (dashed) black lines. The $\pm 1\sigma$ ($\pm 2\sigma$) uncertainties on the expected limits are shown as green (yellow) filled bands. The observed limits are consistent with the expected limits within uncertainties.

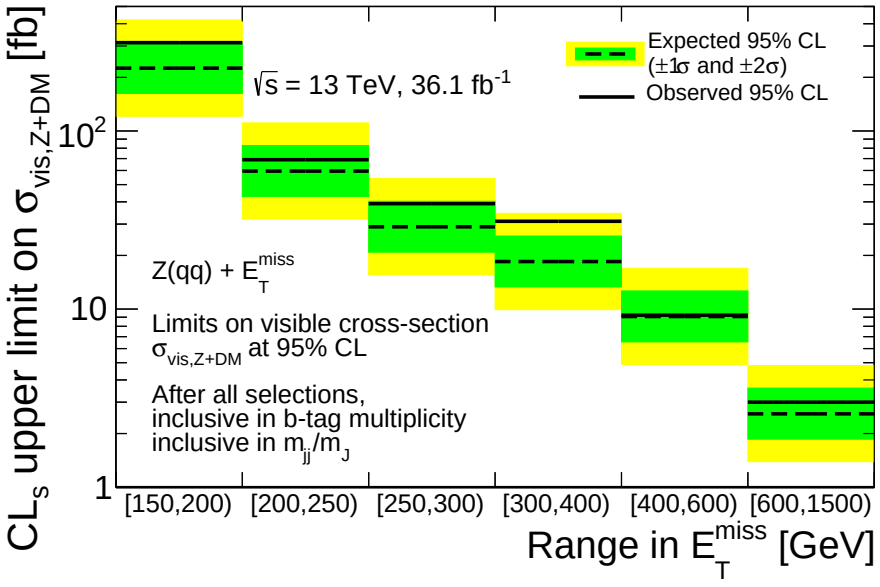


Figure 12.19: Upper limits at 95% CL on the visible cross-section of the Z + DM (right) events in regions of $E_{\mathrm{T}}^{\mathrm{miss}}$ after the full $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ event selection except the $b-$tagging and Z candidate mass requirements. The observed (expected) limits are shown as solid (dashed) black lines. The $\pm 1\sigma$ ($\pm 2\sigma$) uncertainties on the expected limits are shown as green (yellow) filled bands. The observed limits are consistent with the expected limits within uncertainties.

150 GeV $\leq E_{\mathrm{T}}^{\mathrm{miss}} <$ 200 GeV				
	$E_{\mathrm{T}}^{\mathrm{miss}} + W(qq)$		$E_{\mathrm{T}}^{\mathrm{miss}} + Z(qq)$	
$(m_\chi, m_{Z'})$ [GeV]	$\sigma^{\mathrm{exp}}_{\mathrm{vis},W+\mathrm{DM}}$ [fb]	$\sigma^{\mathrm{obs}}_{\mathrm{vis},W+\mathrm{DM}}$ [fb]	$\sigma^{\mathrm{exp}}_{\mathrm{vis},Z+\mathrm{DM}}$ [fb]	$\sigma^{\mathrm{obs}}_{\mathrm{vis},Z+\mathrm{DM}}$ [fb]
1 100	—	—	$155.58^{+60.94}_{-43.48}$	204.11
1 200	—	—	$187.62^{+73.49}_{-52.43}$	250.87
1 300	$580.67^{+227.45}_{-162.26}$	657.68	$159.83^{+62.61}_{-44.66}$	208.73
1 400	$623.57^{+244.26}_{-174.25}$	716.11	$170.44^{+66.76}_{-47.63}$	223.56
1 500	$615.83^{+241.23}_{-172.09}$	705.44	$155.62^{+60.96}_{-43.49}$	204.71
1 600	$647.46^{+253.62}_{-180.93}$	737.28	$168.62^{+66.05}_{-47.12}$	223.42
1 700	$609.29^{+238.66}_{-170.26}$	716.72	$158.43^{+62.06}_{-44.27}$	207.61
1 800	$587.20^{+230.01}_{-164.09}$	676.21	$157.86^{+61.83}_{-44.11}$	204.96
1 900	$625.64^{+245.07}_{-174.83}$	723.59	$173.02^{+67.77}_{-48.35}$	226.84
1 1000	—	—	$163.67^{+64.11}_{-45.74}$	214.67
1 2000	$627.94^{+245.97}_{-175.48}$	713.03	$144.17^{+56.47}_{-40.29}$	182.40
10 10	$601.72^{+235.70}_{-168.15}$	688.74	—	—
10 100	$555.08^{+217.43}_{-155.11}$	632.11	$163.78^{+64.15}_{-45.77}$	215.30
10 10000	$598.46^{+234.42}_{-167.24}$	705.86	$201.14^{+78.79}_{-56.21}$	266.76
1000 10	$651.44^{+255.17}_{-182.04}$	742.92	$147.16^{+57.64}_{-41.12}$	187.96
1000 1000	$635.07^{+248.76}_{-177.47}$	739.16	$154.63^{+60.57}_{-43.21}$	195.58
1000 1995	$634.04^{+248.36}_{-177.18}$	728.61	$172.05^{+67.40}_{-48.08}$	225.51
150 10	$624.47^{+244.61}_{-174.51}$	730.63	$133.20^{+52.18}_{-37.22}$	172.08
150 295	$624.72^{+244.71}_{-174.57}$	717.79	$208.25^{+81.57}_{-58.19}$	276.51
150 1000	$584.57^{+228.98}_{-163.36}$	666.57	$225.28^{+88.24}_{-62.95}$	313.14
50 10	$613.04^{+240.13}_{-171.31}$	697.14	$180.56^{+70.73}_{-50.46}$	241.86
50 95	$599.74^{+234.92}_{-167.59}$	671.86	$200.74^{+78.63}_{-56.10}$	274.73
50 300	$617.83^{+242.01}_{-172.65}$	694.22	$134.61^{+52.73}_{-37.62}$	176.97
500 10	$653.26^{+255.89}_{-182.55}$	754.60	$188.08^{+73.67}_{-52.56}$	261.43
500 995	$609.07^{+238.58}_{-170.20}$	709.43	$152.11^{+59.58}_{-42.51}$	199.82
500 2000	$582.38^{+228.12}_{-162.74}$	663.90	$179.03^{+70.13}_{-50.03}$	240.20
500 10000	$632.49^{+247.75}_{-176.75}$	724.68	$163.52^{+64.05}_{-45.70}$	224.04

Table 12.5: Upper limits at 95% CL on $\sigma_{\mathrm{vis},W+\mathrm{DM}}$ and $\sigma_{\mathrm{vis},Z+\mathrm{DM}}$ in the $[150, 200)$ GeV region of $E_{\mathrm{T}}^{\mathrm{miss}}$. The limits are calculated using a range of simplified vector mediator models with the $(m_\chi, m_{Z'})$ points as indicated in the first column. The largest relative variation in the $E_{\mathrm{T}}^{\mathrm{miss}} + W(qq)$ ($E_{\mathrm{T}}^{\mathrm{miss}} + Z(qq)$) expected limits is 15% (50%). The weakest exclusion limits on the $E_{\mathrm{T}}^{\mathrm{miss}} + W(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + Z(qq)$ production in a given $E_{\mathrm{T}}^{\mathrm{miss}}$ region correspond to the $(m_\chi, m_{Z'}) = (500, 10)$ GeV (red) and $(m_\chi, m_{Z'}) = (1501000)$ GeV (blue) points, respectively.

600 GeV $\leq E_{\mathrm{T}}^{\mathrm{miss}} <$ 1500 GeV				
	$E_{\mathrm{T}}^{\mathrm{miss}} + W(qq)$		$E_{\mathrm{T}}^{\mathrm{miss}} + Z(qq)$	
$(m_\chi, m_{Z'})$ [GeV]	$\sigma^{\mathrm{exp}}_{\mathrm{vis},W+\mathrm{DM}}$ [fb]	$\sigma^{\mathrm{obs}}_{\mathrm{vis},W+\mathrm{DM}}$ [fb]	$\sigma^{\mathrm{exp}}_{\mathrm{vis},Z+\mathrm{DM}}$ [fb]	$\sigma^{\mathrm{obs}}_{\mathrm{vis},Z+\mathrm{DM}}$ [fb]
1 400	$2.54^{+1.00}_{-0.71}$	4.58	—	—
1 500	$1.91^{+0.75}_{-0.53}$	3.14	$1.88^{+0.74}_{-0.52}$	2.60
1 600	$3.06^{+1.20}_{-0.86}$	5.12	$1.82^{+0.71}_{-0.51}$	1.81
1 700	$2.23^{+0.87}_{-0.62}$	4.31	$2.58^{+1.01}_{-0.72}$	3.00
1 800	$2.26^{+0.89}_{-0.63}$	3.98	$2.22^{+0.87}_{-0.62}$	2.47
1 900	$2.52^{+0.99}_{-0.71}$	4.40	$2.08^{+0.81}_{-0.58}$	2.40
1 1000	$2.35^{+0.92}_{-0.66}$	4.57	$1.93^{+0.76}_{-0.54}$	2.45
1 2000	$2.54^{+1.00}_{-0.71}$	5.10	$2.19^{+0.86}_{-0.61}$	2.20
10 10000	$2.35^{+0.92}_{-0.66}$	3.81	$2.08^{+0.82}_{-0.58}$	2.43
1000 10	$2.65^{+1.04}_{-0.74}$	4.63	$2.04^{+0.80}_{-0.57}$	2.19
1000 1000	$2.30^{+0.90}_{-0.64}$	3.95	$2.26^{+0.88}_{-0.63}$	2.73
1000 1995	$2.57^{+1.01}_{-0.72}$	4.66	$2.00^{+0.79}_{-0.56}$	1.95
150 1000	$2.28^{+0.89}_{-0.64}$	3.87	$2.13^{+0.84}_{-0.60}$	2.45
500 10	$2.51^{+0.99}_{-0.70}$	4.42	$1.86^{+0.73}_{-0.52}$	1.75
500 995	$2.22^{+0.87}_{-0.62}$	4.09	$1.99^{+0.78}_{-0.56}$	2.33
500 2000	$2.34^{+0.92}_{-0.65}$	4.04	$2.08^{+0.82}_{-0.58}$	2.13
500 10000	$2.32^{+0.91}_{-0.65}$	4.10	$2.12^{+0.83}_{-0.59}$	2.71

Table 12.6: Upper limits at 95% CL on $\sigma_{\mathrm{vis},W+\mathrm{DM}}$ and $\sigma_{\mathrm{vis},Z+\mathrm{DM}}$ in the $[300, 400)$ GeV region of $E_{\mathrm{T}}^{\mathrm{miss}}$. The limits are calculated using a range of simplified vector mediator models with the $(m_\chi, m_{Z'})$ points as indicated in the first column. The largest relative variation in the $E_{\mathrm{T}}^{\mathrm{miss}} + W(qq)$ ($E_{\mathrm{T}}^{\mathrm{miss}} + Z(qq)$) expected limits is 45% (35%). The weakest exclusion limits on the $E_{\mathrm{T}}^{\mathrm{miss}} + W(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + Z(qq)$ production in a given $E_{\mathrm{T}}^{\mathrm{miss}}$ region correspond to the $(m_\chi, m_{Z'}) = (1, 2000)$ GeV (red) and $(m_\chi, m_{Z'}) = (1, 700)$ GeV (blue) points, respectively.

W+DM production			
$E_{\mathrm{T}}^{\mathrm{miss}}$ region [GeV]	$\sigma^{\mathrm{exp}}_{\mathrm{vis},W+\mathrm{DM}}$ [fb]	$\sigma^{\mathrm{obs}}_{\mathrm{vis},W+\mathrm{DM}}$ [fb]	$A \times \varepsilon$
$[150, 200)$	653^{+256}_{-182}	755	20%
$[200, 250)$	163^{+64}_{-45}	185	20%
$[250, 300)$	50^{+19}_{-14}	43	30%
$[300, 400)$	36^{+14}_{-10}	41	45%
$[400, 600)$	$12.6^{+5.0}_{-3.5}$	9.7	55%
$[600, 1500)$	$3.1^{+1.2}_{-0.9}$	5.1	55%

Table 12.7: Upper limits at 95% CL on the visible detector-level cross-section $\sigma_{\mathrm{vis},W+\mathrm{DM}}$ and $A \times \varepsilon$ from a range of the $E_{\mathrm{T}}^{\mathrm{miss}} + W(qq)$ simplified vector mediator models, as shown in Tables 12.5, H.1, H.2, H.3, H.4, 12.6 and 12.12 per region of $E_{\mathrm{T}}^{\mathrm{miss}}$. The numbers in the last column are rounded to 5% precision.

Z+DM production			
$E_{\mathrm{T}}^{\mathrm{miss}}$ region [GeV]	$\sigma_{\mathrm{vis},Z+\mathrm{DM}}^{\mathrm{exp}}$ [fb]	$\sigma_{\mathrm{vis},Z+\mathrm{DM}}^{\mathrm{obs}}$ [fb]	$A \times \varepsilon$
$[150, 200)$	225^{+88}_{-63}	313	20%
$[200, 250)$	60^{+23}_{-16}	69	20%
$[250, 300)$	29^{+11}_{-8}	39	30%
$[300, 400)$	18^{+7}_{-5}	31	45%
$[400, 600)$	$9.1^{+2.5}_{-2.6}$	9.2	50%
$[600, 1500)$	$2.6^{+1.0}_{-0.7}$	3.0	55%

Table 12.8: Upper limits at 95% CL on the visible detector-level cross-section $\sigma_{\mathrm{vis},Z+\mathrm{DM}}$ and $A \times \varepsilon$ from a range of the $E_{\mathrm{T}}^{\mathrm{miss}} + Z(qq)$ simplified vector mediator models, as shown in Tables 12.5, H.1, H.2, H.3, H.4, 12.6 and 12.12 per region of $E_{\mathrm{T}}^{\mathrm{miss}}$. The numbers in the last column are rounded to 5% precision.

Generic limits on production of h + DM events

The limits on the visible production cross-section $\sigma_{\text{vis},h+\text{DM}}$ for a range of Z'-2HDM models in the lowest and highest $E_{\text{T}}^{\text{miss}}$ regions are shown in Tables 12.9 and 12.10. The limits in other $E_{\text{T}}^{\text{miss}}$ regions are given in Tables I.1 and I.2 in Appendix I. The relative variations in the expected limits on the h+DM production are low in the resolved regime, amounting to 10-15%, and drastically increase up to 70% in the merged regime. This can be explained by the larger regions of $E_{\text{T}}^{\text{miss}}$ comparing to the $E_{\text{T}}^{\text{miss}} + V(qq)$ analysis and very high MC statistics for the signal models that contribute in the region of $E_{\text{T}}^{\text{miss}} < 500$ GeV, which drastically fall for the 1 and 2 $b-$tags in the event in the region of $E_{\text{T}}^{\text{miss}} > 500$ GeV.

$E_{\text{T}}^{\text{miss}} + h(bb)$, 150 GeV $\leq E_{\text{T}}^{\text{miss}} <$ 200 GeV						
$(m_{Z'}, m_A)$ [GeV]	$\sigma^{\text{exp}}_{\text{vis},h+\text{DM}}$ [fb]	$\sigma^{\text{exp}}_{\text{vis},h+\text{DM}}$ [fb]	-2σ [fb]	-1σ [fb]	+1σ [fb]	+2σ [fb]
400 400	18.342	19.124	9.845	13.216	25.527	34.220
600 200	12.931	14.463	6.940	9.317	17.996	24.125
600 400	16.984	17.015	9.116	12.238	23.637	31.687
600 600	12.686	11.380	6.809	9.141	17.656	23.669
800 300	14.436	19.182	7.748	10.402	20.091	26.934
800 500	12.205	12.713	6.551	8.794	16.985	22.770
800 600	12.659	11.985	6.794	9.121	17.618	23.618
800 700	13.295	13.258	7.136	9.580	18.503	24.805
1000 800	13.211	13.731	7.091	9.519	18.386	24.648

Table 12.9: Upper limits at 95% CL on $\sigma_{\text{vis},h+\text{DM}}$ in the $[150, 200)$ GeV region of $E_{\text{T}}^{\text{miss}}$. The limits are calculated using a range of Z'-2HDM models with the $(m_{Z'}, m_A)$ points as indicated in the first column. The largest relative variation in the $E_{\text{T}}^{\text{miss}} + h(b\bar{b})$ expected limits is 10%. The weakest expected limit in a given $E_{\text{T}}^{\text{miss}}$ region corresponds to the $(m_{Z'}, m_A) = (400, 400)$ GeV point (red).

$E_{\text{T}}^{\text{miss}} + h(bb)$, 500 GeV $\leq E_{\text{T}}^{\text{miss}} < \infty$						
$(m_{Z'}, m_A)$ [GeV]	$\sigma^{\text{exp}}_{\text{vis},h+\text{DM}}$ [fb]	$\sigma^{\text{exp}}_{\text{vis},h+\text{DM}}$ [fb]	-2σ [fb]	-1σ [fb]	+1σ [fb]	+2σ [fb]
1200 300	0.706	0.739	0.379	0.508	0.982	1.316
1400 600	0.710	0.745	0.381	0.511	0.988	1.324
1800 800	0.767	0.797	0.412	0.553	1.067	1.431
2000 300	0.942	0.956	0.506	0.679	1.311	1.757
2200 300	1.038	1.055	0.557	0.748	1.445	1.937
2400 600	1.065	1.070	0.571	0.767	1.482	1.986
2600 400	1.335	1.349	0.716	0.962	1.858	2.491
2800 200	1.599	1.607	0.858	1.152	2.225	2.983
3000 300	1.766	1.730	0.948	1.272	2.457	3.294

Table 12.10: Upper limits at 95% CL on $\sigma_{\text{vis},h+\text{DM}}$ in the $[500, \infty)$ GeV region of $E_{\text{T}}^{\text{miss}}$. The limits are calculated using a range of Z'-2HDM models with the $(m_{Z'}, m_A)$ points as indicated in the first column. The largest variation in the $E_{\text{T}}^{\text{miss}} + h(b\bar{b})$ expected limits is 70%. The weakest expected limit in a given $E_{\text{T}}^{\text{miss}}$ region corresponds to the $(m_{Z'}, m_A) = (3000, 300)$ GeV point (red).

The resulting generic limits on the visible cross-section $\sigma_{\text{vis},h+\text{DM}}$ in a given $E_{\text{T}}^{\text{miss}}$ region that correspond to the weakest expected limits in Tables 12.9, I.1, I.2 and 12.10 are shown in Figure 12.20 and summarised in Table 12.11. No significant deviation from the expected limits is observed. The generic limits become stronger with increasing $E_{\text{T}}^{\text{miss}}$ due to the rapidly falling SM backgrounds in the regions of high $E_{\text{T}}^{\text{miss}}$. Generally, the generic limits on the h+DM production are much stronger comparing to the W+DM and

Z+DM limits, since the h + DM events mostly contribute to the analysis regions with 2 $b-$tags with very low SM backgrounds.

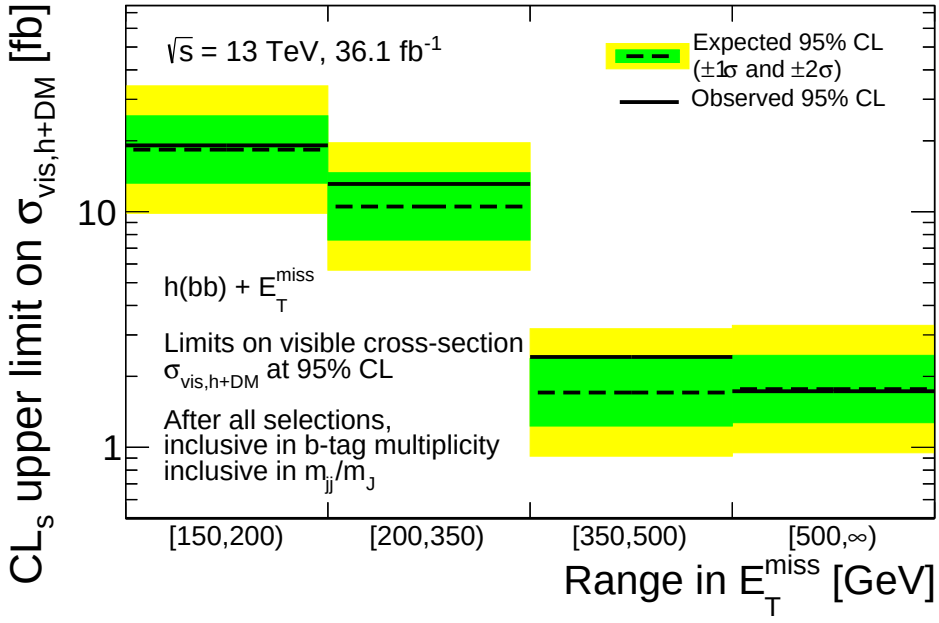


Figure 12.20: Upper limits at 95% CL on the visible cross-section of the h + DM production in regions of $E_{\mathrm{T}}^{\mathrm{miss}}$ after the full $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ event selection except the $b-$tagging and h candidate mass requirements. The observed (expected) limits are shown as solid (dashed) black lines. The $\pm1\sigma$ ($\pm2\sigma$) uncertainties on the expected limits are shown as green (yellow) filled bands. The observed limits are consistent with the expected limits within uncertainties.

h+DM production			
$E_{\mathrm{T}}^{\mathrm{miss}}$ region [GeV]	$\sigma^{\mathrm{exp}}_{\mathrm{vis},h+\mathrm{DM}}$ [fb]	$\sigma^{\mathrm{obs}}_{\mathrm{vis},h+\mathrm{DM}}$ [fb]	$A\times\varepsilon$
$[150,200)$	$18.3^{+7.2}_{-5.1}$	19.1	15%
$[200,350)$	$10.5^{+4.1}_{-2.9}$	13.1	35%
$[350,500)$	$1.7^{+0.7}_{-0.5}$	2.4	40%
$[500,\infty)$	$1.8^{+0.7}_{-0.5}$	1.7	55%

Table 12.11: Upper limits at 95% CL on the visible detector-level cross-section $\sigma_{\mathrm{vis},h+\mathrm{DM}}$ and $A\times\varepsilon$ from a range of Z'-2HDM models, as shown in Tables 12.9, I.1, I.2, 12.10 and 12.14 per region of $E_{\mathrm{T}}^{\mathrm{miss}}$. The numbers in the last column are rounded to 5% precision.

12.3.1 Particle-level limits

A detailed description of the $A\times\varepsilon$ definition and calculation procedure is given in Section 11.2.3. Tables 12.12 and 12.13 show the resulting $(A\times\varepsilon)_{\mathrm{sel}}$ and $(A\times\varepsilon)_{\mathrm{presel}}$ for a range of the $E_{\mathrm{T}}^{\mathrm{miss}}+W/Z(qq)$ simplified vector mediator models. The corresponding numbers for the $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ Z'-2HDM models are given Table 12.14. Since $(A\times\varepsilon)_{\mathrm{presel}}$ is a $E_{\mathrm{T}}^{\mathrm{miss}}$-specific quantity, its value is averaged over all $E_{\mathrm{T}}^{\mathrm{miss}}$ regions. The variations of the $(A\times\varepsilon)_{\mathrm{sel}}$ values in all $E_{\mathrm{T}}^{\mathrm{miss}}$ regions are less than 20%. As discussed in Section 11.2.3, the preselection criteria of $E_{\mathrm{T}}^{\mathrm{miss}}>140$ GeV pushes $(A\times\varepsilon)_{\mathrm{presel}}$ to the artificially low values for the models with the soft $E_{\mathrm{T}}^{\mathrm{miss}}$ spectrum, such as $E_{\mathrm{T}}^{\mathrm{miss}}+V(qq)$ simpliied models with $(m_\chi, m_{Z'})=(10,10)$ GeV and $(m_\chi, m_{Z'})=(10,100)$ GeV mass parameters. The

significant fraction of the $E_{\mathrm{T}}^{\mathrm{miss}}$ is then rejected after requiring $E_{\mathrm{T}}^{\mathrm{miss}} > 140$ GeV. The $(A \times \varepsilon)_{\mathrm{presel}}$ is estimated to be in the range of 82%-93% (74%-92%) for the $E_{\mathrm{T}}^{\mathrm{miss}} + W(qq)$ ($E_{\mathrm{T}}^{\mathrm{miss}} + Z(qq)$) simplified models. The $A \times \varepsilon_{\mathrm{sel}}$ increases with higher $E_{\mathrm{T}}^{\mathrm{miss}}$, as the signal selection efficiency increases as well.
The resulting $A \times \varepsilon$ defined in a given $E_{\mathrm{T}}^{\mathrm{miss}}$ region is defined as the smallest $A \times \varepsilon$ value for the range of signal models that corresponds to the weakest parton-level generic limit. The final values of $A \times \varepsilon$ together with the corresponging generic limits on the visible production cross-section of W+DM, Z+DM and h+DM events are presented in Tables 12.7, 12.8 and 12.11.

$E_{\mathrm{T}}^{\mathrm{miss}} + W(qq)$ simplified vector mediator models							
$(m_\chi, m_{Z'})$	$(A \times \varepsilon)_{\mathrm{sel}}$ in range of $E_{\mathrm{T}}^{\mathrm{miss}}$/GeV						$(A \times \varepsilon)_{\mathrm{presel}}$
	[150, 200)	[200, 250)	[250, 300)	[300, 400)	[400, 600)	[600, 1500)	
1 2000	28.0%	29.1%	35.8%	54.6%	62.4%	66.5%	93.0%
10 10	29.9%	26.9%	–	–	–	–	76.5%
10 100	30.7%	–	–	–	–	–	78.3%
10 10000	27.0%	28.8%	37.2%	53.4%	66.6%	67.3%	86.8%
1000 10	27.0%	28.2%	38.3%	55.7%	64.8%	70.3%	89.3%
1000 1000	27.9%	29.6%	40.0%	54.4%	62.2%	69.5%	91.9%
1000 1995	28.8%	27.9%	37.9%	55.2%	63.9%	69.7%	90.3%
150 10	27.7%	32.3%	–	–	–	–	87.0%
150 295	29.5%	–	–	–	–	–	83.9%
150 1000	28.2%	30.6%	37.0%	52.2%	68.1%	64.1%	87.4%
50 10	28.0%	–	–	–	–	–	83.0%
50 95	29.0%	31.4%	–	–	–	–	82.0%
50 300	29.7%	31.0%	–	–	–	–	84.0%
500 10	29.5%	27.3%	43.2%	53.3%	65.3%	65.3%	89.4%
500 995	28.7%	30.7%	36.9%	53.7%	61.5%	65.1%	89.6%
500 2000	27.6%	28.4%	39.3%	54.8%	64.2%	69.7%	89.2%
500 10000	28.9%	26.5%	38.3%	56.5%	63.5%	68.7%	89.3%

Table 12.12: The $(A \times \varepsilon)_{\mathrm{sel}}$ values per $E_{\mathrm{T}}^{\mathrm{miss}}$ region for the different $(m_\chi, m_{Z'})$ points of the $E_{\mathrm{T}}^{\mathrm{miss}} + W(qq)$ simplified vector mediator model. The $(A \times \varepsilon)_{\mathrm{presel}}$ values averaged over all $E_{\mathrm{T}}^{\mathrm{miss}}$ regions are shown in the last column. The $(A \times \varepsilon)_{\mathrm{presel}}$ is calculated using the cross sections generated with MadGraph5 after applying the preselection requirement of $E_{\mathrm{T}}^{\mathrm{miss}} > 140$ GeV. $(A \times \varepsilon)_{\mathrm{presel}}$ appears to be artificially low for the models with very soft $E_{\mathrm{T}}^{\mathrm{miss}}$, where a significant fraction of the $E_{\mathrm{T}}^{\mathrm{miss}}$ spectrum is rejected with the $E_{\mathrm{T}}^{\mathrm{miss}} > 140$ GeV requirement. Long dashes indicate the models, which have less than 500 events in a given $E_{\mathrm{T}}^{\mathrm{miss}}$ region. For the last $E_{\mathrm{T}}^{\mathrm{miss}}$ region the requirement is relaxed to 200 events. The models with the $(m_\chi, m_{Z'}) = (1, 200 - 1000)$ GeV are excluded from the $A \times \varepsilon$ calculation, as the corresponding MC samples contain additional filters on the numbers of $b-$tags in the event (BVeto and BFilter).

$E_{\mathrm{T}}^{\mathrm{miss}} + Z(qq)$ simplified vector mediator models							
$(m_\chi, m_{Z'})$	$(A \times \varepsilon)_{\mathrm{sel}}$ in range of $E_{\mathrm{T}}^{\mathrm{miss}}$/GeV						$(A \times \varepsilon)_{\mathrm{presel}}$
	[150, 200)	[200, 250)	[250, 300)	[300, 400)	[400, 600)	[600, 1500)	
1 100	26.3%	–	–	–	–	–	73.8%
1 2000	26.4%	27.2%	37.1%	52.7%	59.3%	69.0%	92.1%
10 100	27.0%	–	–	–	–	–	78.1%
10 10000	24.5%	28.4%	38.9%	49.3%	60.6%	65.8%	90.1%
150 10	27.1%	30.8%	–	–	–	–	85.2%
150 295	28.3%	–	–	–	–	–	81.1%
150 1000	27.4%	28.4%	36.8%	50.0%	62.7%	65.2%	88.2%
50 10	28.9%	–	–	–	–	–	81.8%
50 95	26.2%	–	–	–	–	–	74.9%
50 300	26.4%	–	–	–	–	–	76.4%
500 10	25.2%	28.4%	36.5%	52.8%	60.5%	65.8%	92.2%
500 995	24.6%	29.4%	36.4%	52.7%	58.3%	65.7%	87.8%
500 2000	23.8%	27.3%	34.1%	53.8%	61.4%	67.4%	90.9%
500 10000	27.3%	28.3%	39.3%	51.5%	65.2%	65.8%	90.4%
1000 10	25.4%	28.3%	36.1%	55.0%	62.0%	66.2%	88.5%
1000 1000	25.3%	28.1%	35.5%	54.6%	63.0%	64.7%	91.3%
1000 1995	26.2%	26.0%	35.2%	53.9%	61.0%	66.3%	91.1%

Table 12.13: The $(A \times \varepsilon)_{\mathrm{sel}}$ values per $E_{\mathrm{T}}^{\mathrm{miss}}$ region for the different $(m_\chi, m_{Z'})$ points of the $E_{\mathrm{T}}^{\mathrm{miss}} + Z(qq)$ simplified vector mediator model. The $(A \times \varepsilon)_{\mathrm{presel}}$ values averaged over all $E_{\mathrm{T}}^{\mathrm{miss}}$ regions are shown in the last column. The $(A \times \varepsilon)_{\mathrm{presel}}$ is calculated using the cross sections generated with MadGraph5 after applying the preselection requirement of $E_{\mathrm{T}}^{\mathrm{miss}} > 140$ GeV. $(A \times \varepsilon)_{\mathrm{presel}}$ appears to be artificially low for the models with very soft $E_{\mathrm{T}}^{\mathrm{miss}}$, where a significant fraction of the $E_{\mathrm{T}}^{\mathrm{miss}}$ spectrum is rejected with the $E_{\mathrm{T}}^{\mathrm{miss}} > 140$ GeV requirement. Long dashes indicate the models, which have less than 500 events in a given $E_{\mathrm{T}}^{\mathrm{miss}}$ region. For the last $E_{\mathrm{T}}^{\mathrm{miss}}$ region the requirement is relaxed to 200 events. The models with the $(m_\chi, m_{Z'}) = (1, 200 - 1000)$ GeV are excluded from the $A \times \varepsilon$ calculation, as the corresponding MC samples contain additional filters on the numbers of $b-$tags in the event (BVeto and BFilter).

$(m_{Z'}, m_A)$	$(A \times \varepsilon)_{\text{sel}}$ in range of $E_{\text{T}}^{\text{miss}}$/GeV				$(A \times \varepsilon)_{\text{presel}}$
	$[150, 200)$	$[200, 350)$	$[350, 500)$	$[500, \infty)$	
200 200	33%	52.6%	57.4%	69.9%	57.6%
200 300	34.3%	54.6%	57.4%	66.9%	72.9%
200 400	33.5%	54.4%	58.5%	71.4%	79.6%
400 200	13.1%	14.3%	9.8%	4.6%	10.1%
400 300	16.5%	27.4%	38.1%	50.3%	28.1%
400 400	19.4%	42.2%	52.9%	68.8%	36.7%
400 500	16.5%	37.5%	51.4%	71.7%	32.3%
600 200	31.7%	54%	39.6%	16.7%	82.1%
600 300	37.1%	41.3%	25.2%	19.5%	62.2%
600 400	26.4%	26%	12.6%	15.2%	17.1%
600 500	34.7%	39.5%	19.3%	20.4%	43.1%
600 600	34.9%	42.6%	25.1%	25.2%	50.2%
800 200	26.7%	54.2%	54.6%	19.3%	92.3%
800 300	26.7%	59.2%	52%	26.3%	91.4%
800 400	30.8%	59.7%	47%	16%	86.3%
800 500	36.8%	52.9%	37.6%	20.3%	71.6%
800 600	35.9%	52.8%	36.3%	16.7%	52.9%
800 700	35.9%	54.2%	42.9%	19.4%	70.6%
1000 300	24.9%	48.9%	65.3%	44.9%	95.8%
1000 400	25%	51.6%	63.9%	32.3%	93.7%
1000 500	25%	55.2%	59.1%	26.9%	92%
1000 600	28.9%	58.6%	56.7%	21.9%	88.8%
1000 700	34.3%	57.3%	58%	23.2%	81.2%
1000 800	33.6%	56.2%	59.8%	24.7%	78%
1200 300	22.2%	45.8%	53.8%	68.8%	97.4%
1200 400	19%	46.4%	56.4%	64.4%	97.1%
1200 500	23.8%	48%	61.7%	58.7%	95.2%
1200 600	22.8%	50.8%	63.7%	51.4%	95.2%
1200 700	28.1%	54.7%	62.3%	46.1%	93.8%
1200 800	32.4%	56.4%	61.9%	50.3%	90.9%
1400 300	14.5%	40.4%	45.9%	73.9%	98.1%
1400 400	15.4%	42.5%	47.7%	73.3%	98%
1400 500	19.6%	43.1%	50.1%	72.4%	96.4%
1400 600	19.2%	46.2%	54.5%	70.3%	97.5%
1400 700	25%	48.6%	58.6%	67%	95.3%
1400 800	28.1%	51.3%	60.1%	65.3%	95.8%
1600 300	23.5%	38.6%	42.2%	75.9%	98.4%
1600 400	20.6%	41.9%	44.1%	76%	97.1%
1600 500	14.4%	41.2%	44.5%	75.3%	97.2%
1600 600	20.2%	43.1%	46.9%	74.6%	97.3%
1600 700	21.1%	42.8%	50.7%	73.3%	96.8%
1600 800	23.9%	46.6%	54.2%	72.3%	97.6%
1800 300	13.6%	37%	38.3%	76.9%	98.7%
1800 400	18.3%	35.3%	41.9%	77.2%	97.3%
1800 500	18.5%	39.8%	42%	76.7%	98.7%
1800 600	23%	39.1%	44.2%	76.3%	98.5%
1800 700	20.4%	38.5%	44.8%	76.1%	97.2%
1800 800	24.4%	44.6%	48.4%	75.3%	98.3%
2000 300	14.7%	32.2%	41.2%	78.2%	97.5%
2000 400	7.3%	32.3%	38.8%	78.3%	97.4%
2000 500	15.5%	36.7%	39%	77.8%	97.8%
2000 600	17.9%	34.9%	41.7%	77.6%	97.7%
2000 700	18.9%	38.5%	41.3%	77.1%	99.7%
2000 800	24.5%	39.8%	44.4%	76.5%	97.3%
2200 300	42.4%	38.6%	38.6%	78.9%	97.6%

2200 400	7.5%	36.7%	40.9%	78.6%	97.9%
2200 500	34.8%	34.9%	39.8%	78.4%	98.9%
2200 600	16.1%	36%	39.2%	78.4%	97.5%
2200 700	12.4%	37.5%	42%	78.2%	97.7%
2200 800	15.5%	37.3%	42.8%	77.8%	98.7%
2400 200	19.5%	30.9%	36.7%	79.4%	99%
2400 300	11.7%	32.2%	35.1%	78.7%	98%
2400 400	6.8%	27.8%	37.2%	79.3%	97.7%
2400 500	9%	29.9%	34.5%	78.6%	98.9%
2400 600	12.6%	29.8%	39.3%	78.8%	98.3%
2400 700	2.7%	38.5%	40.7%	78.4%	97.6%
2400 800	30.1%	33.7%	41.5%	78.3%	98.2%
2600 200	9%	40.3%	38.9%	79.1%	99%
2600 300	0%	45.6%	40.7%	79.3%	99.1%
2600 400	23.5%	30.6%	34.8%	79.2%	99%
2600 500	15.6%	32.6%	35%	79.3%	97.7%
2600 600	5.2%	37%	36.5%	79.3%	98.5%
2600 700	24.4%	36.2%	44.2%	78.6%	99%
2600 800	12.3%	34.2%	37.3%	79.1%	99%
2800 200	34.9%	28.5%	41.3%	79.4%	99.1%
2800 300	7%	28.1%	36.9%	79.5%	97.8%
2800 400	0%	28.8%	37.8%	79.8%	97.7%
2800 500	42.5%	37.5%	37.6%	79.2%	99%
2800 600	10.3%	29%	36.5%	79.4%	99.1%
3000 200	19.4%	26%	33.1%	80%	97.8%
3000 300	27.5%	26.5%	30.6%	79.5%	99%
3000 400	40%	24.5%	36.2%	79.5%	99.1%
3000 500	0%	17.3%	38.8%	79.2%	99%
3000 600	5.6%	25.2%	37.3%	79.7%	99.1%

Table 12.14: The $(A \times \varepsilon)_{\text{sel}}$ values per $E_{\text{T}}^{\text{miss}}$ region together with the $(A \times \varepsilon)_{\text{presel}}$ averaged over all $E_{\text{T}}^{\text{miss}}$ regions for the different $(m_{Z'}, m_A)$ points of the $E_{\text{T}}^{\text{miss}} + h(b\bar{b})$ Z'-2HDM model. Note that due to the preselection criteria of $E_{\text{T}}^{\text{miss}} > 140$ GeV, the $(A \times \varepsilon)_{\text{presel}}$ appears to be artificially low for the models with very soft $E_{\text{T}}^{\text{miss}}$, where a significant fraction of the $E_{\text{T}}^{\text{miss}}$ spectrum is rejected after the $E_{\text{T}}^{\text{miss}} > 140$ GeV requirement. Thus, the $(A \times \varepsilon)_{\text{presel}}$ for these models is considered to in the 90-99% range as for the models with a hard $E_{\text{T}}^{\text{miss}}$ spectrum away from the 140 GeVthreshold.

Chapter 13

Conclusions

Dark Matter composes a significant part of the visible Universe. Despite a solid cosmological evidence, its nature, properties and interaction with the Standard Model sector is still to be unraveled. A wide range of different models beyond the Standard Model (SM) are able to provide the production of the dark matter (DM) particles that is consisent with the observed DM relic density [26, 27] and can be probed at the LHC. One class of these models predicts the associated production of the DM and SM particles. This thesis presents two searches for the DM particles from this class, one produced in association with the hadronically decaying Standard Model vector bosons ($E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$) and other with the Standard Model Higgs boson h, decaying to a $b\bar{b}$ pair ($E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$).
Both searches are performed using 36.1 fb^{-1} of the proton-proton collision data at $\sqrt{s} =$ 13 TeV collected with the ATLAS detector in 2015 and 2016 [46, 48]. They share similar final state, which contains pair of jets and missing transverse energy, $E_{\mathrm{T}}^{\mathrm{miss}}$, characterised by the boost of the decay products of $E_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV. This enables the joint description of both analyses in the scope of this thesis.
In preparation of the interpretation of the final search results, the performance of the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers used in both analyses for the different periods of data taking has to be studied. Due to the non-ideal $E_{\mathrm{T}}^{\mathrm{miss}}$ response in the detector, the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers are not fully efficient in the 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 250 GeV region covered by the resolved topology. Studies are performed both in data and Monte Carlo (MC) simulations using the 1μ control regions. It is shown that Monte Carlo simulation provides up to 10% lower $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiencies than those in data. To correct for this discrepancy, the scale factors are developed to correct the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency in MC. The scale factors are obtained as functions of the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ from the fit to the ratio of the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency in data to those in MC.
The most important part of every search for the new physics is to interpret the final search results in terms of the discovery of the new physics and capability of a certain model to describe a potential discovery. The first interpretation is addressed by comparing the distributions of the $E_{\mathrm{T}}^{\mathrm{miss}}$ variable and invariant mass of the decay products for the $E_{\mathrm{T}}^{\mathrm{miss}}+V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ searches, respectively, in observed data and SM predictions after the statistical fit to observed data. Both distributions show no significant excess in the observed data over the SM predictions. The second interpretation is given in terms of the upper limits on the production cross-section times the branching ratio of the SM hadronic boson decay [1], $\sigma_{\mathrm{vis},W/Z+\mathrm{DM}} \times \mathcal{B}_{W/Z\to q\bar{q}}$ and $\sigma_{\mathrm{vis},h+\mathrm{DM}} \times \mathcal{B}_{h\to b\bar{b}}$, are set at 95%

[1] The branching ratios of the $W \to q\bar{q}$, $Z \to q\bar{q}$, and $h \to b\bar{b}$ decays are taken as $\mathcal{B}_{W\to q\bar{q}} = 0.674$, $\mathcal{B}_{Z\to q\bar{q}} = 0.699$ and $\mathcal{B}_{h\to b\bar{b}} = 0.571$, respectively

confidence level (CL) in the $(m_\chi, m_{Z'})$ and $(m_{Z'}, m_A)$ plane, respectively. The simplified vector mediator model, in which the DM is produced via the $s-$channel Z' mediator with the coupling parameters $g_q = 0.25$ and $g_\chi = 1.0$, is considered in the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ search. The simplified Z'-2HDM model with the parameters $g_{Z'} = 0.8$, $\tan\beta = 1$, $m_\chi = 100$ GeV, and $m_H = m_{H^\pm} = 300$ GeV is considered in the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ search. The limits are set in the $(m_\chi, m_{Z'})$ and $(m_{Z'}, m_A)$ plane for the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ searches, respectively. The mediator masses $m_{Z'}$ are excluded up to 650 GeVfor the dark matter masses m_χ of up to 250 GeV. This is consistent with uncertainties with the expected exclusion limits on the $m_{Z'}$ mass of up to 700 GeVfor m_χ of up to 230 GeV. The mediator masses $m_{Z'}$ are excluded up to 2.6 TeVfor the pseudoscalar masses m_A of up to 0.6 TeV, which is consistent with the expected limits on the $m_{Z'}$ mass of up to 2.5 TeVfor m_A of up to 0.6 TeV. Both limits are significantly improved relative to the previous versions of the analyses [45, 235].

In addition to the model-dependent upper limits, the model-independent representation is introduced for the first time for both analyses. The generic upper limits on the visible cross-section of $W/Z/h$+DM events at 95% CL (LRMDs) are presented in separate regions of the $E_{\mathrm{T}}^{\mathrm{miss}}$ variable to avoid any model assumptions. Limits are calculated at the detector- and particle-levels, where the latter provides a possibility for phenomenologists to check various models of the DM production with a given final state. The detector-level LRMDs for the W+DM, Z+DM, and h+DM events at 95% CL are in the range of $5.1 - 755$ fb, $3.0 - 313$ fb, and $1.7 - 19.1$ fb respectively, depending on the $E_{\mathrm{T}}^{\mathrm{miss}}$ region. The strongest LRMDs are set in the regions of high $E_{\mathrm{T}}^{\mathrm{miss}}$ due to the drastically falling SM backgrounds. The h+DM and Z+DM final states show the strongest limits as they mainly populate the regions with 2 $b-$tags, which provide the highest sensitivity due to the very low SM backgrounds. The particle-level LRMDs in a given $E_{\mathrm{T}}^{\mathrm{miss}}$ bin are set by dividing out the acceptance times efficiency, $A \times \varepsilon$, from the detector-level limits. The resulting $A \times \varepsilon$ values for the $E_{\mathrm{T}}^{\mathrm{miss}} + W/Z(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ models are in range of $20 - 55\%$ and $15 - 55\%$, respectively, depending on the $E_{\mathrm{T}}^{\mathrm{miss}}$ region.

Appendix A

Validation plots for data-driven $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger scale factors

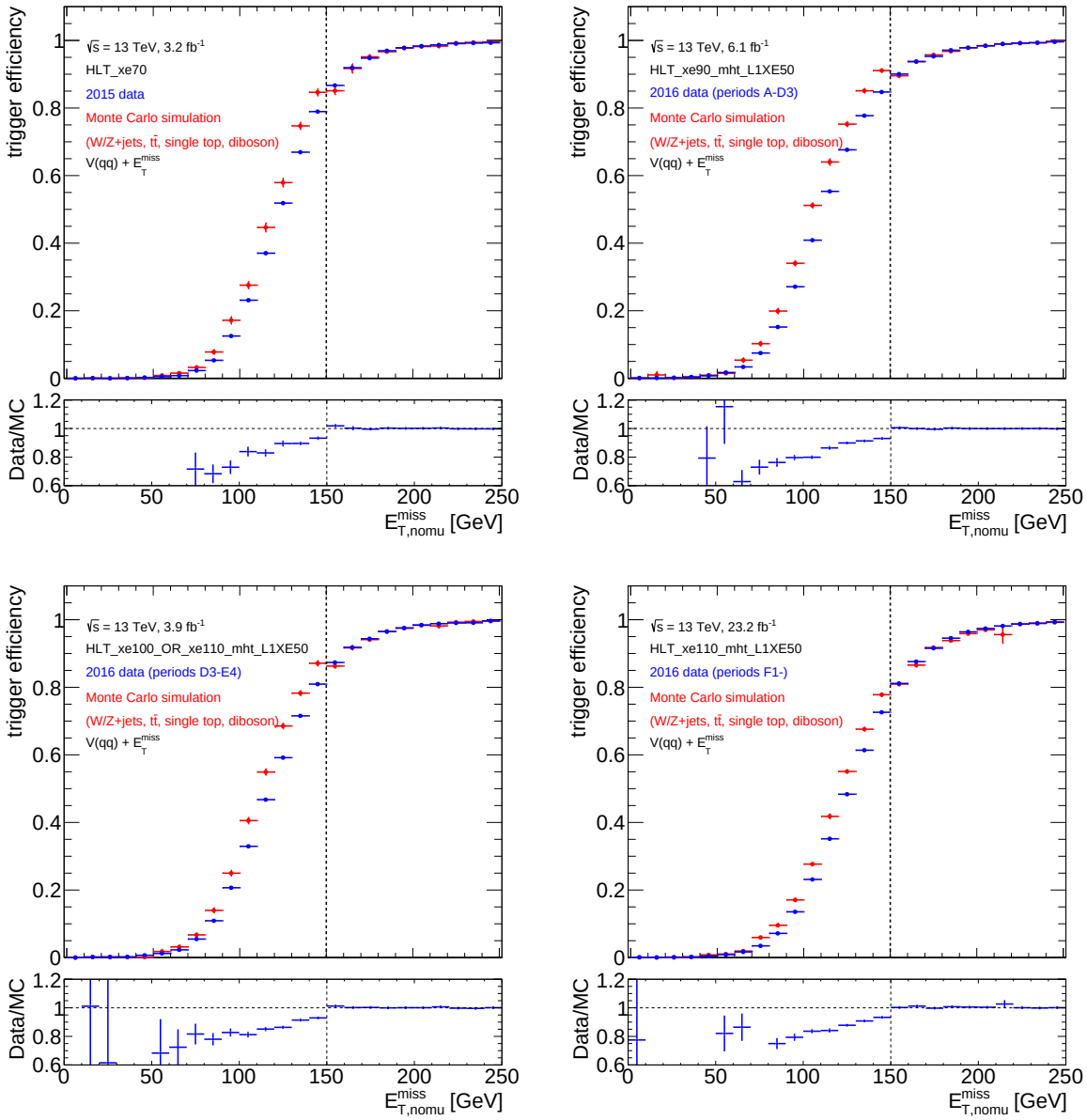


Figure A.1: Validation plots of $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiencies and corresponding scale factors as function of offline $E_{\mathrm{T,nomu}}^{\mathrm{miss}}$ measured in $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis. The data-to-MC scale factor corrections are applied for $E_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV. Plots for HLT_XE70 (top left), HLT_XE90_MHT_L1XE50 (top right), HLT_XE100_MHT_L1XE50 **OR** HLT_XE110_MHT_L1XE50 (bottom left) and HLT_XE110_MHT_L1XE50 (bottom right) triggers are shown in data and MC for 1 lepton control region. Plots are shown for all $b-$tagged analysis regions together. MC is dominated by $W+$jets and $t\bar{t}$ events.

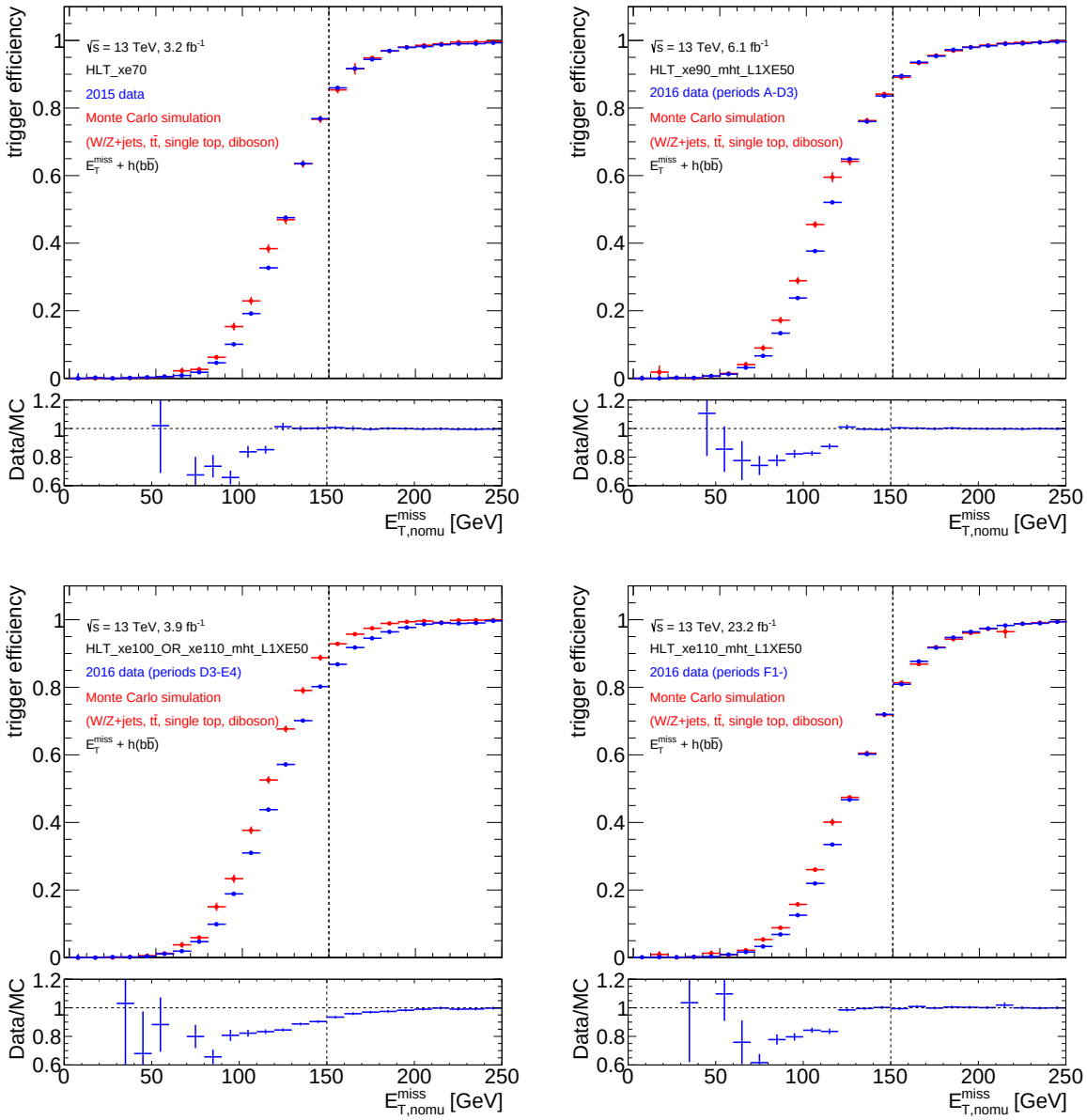
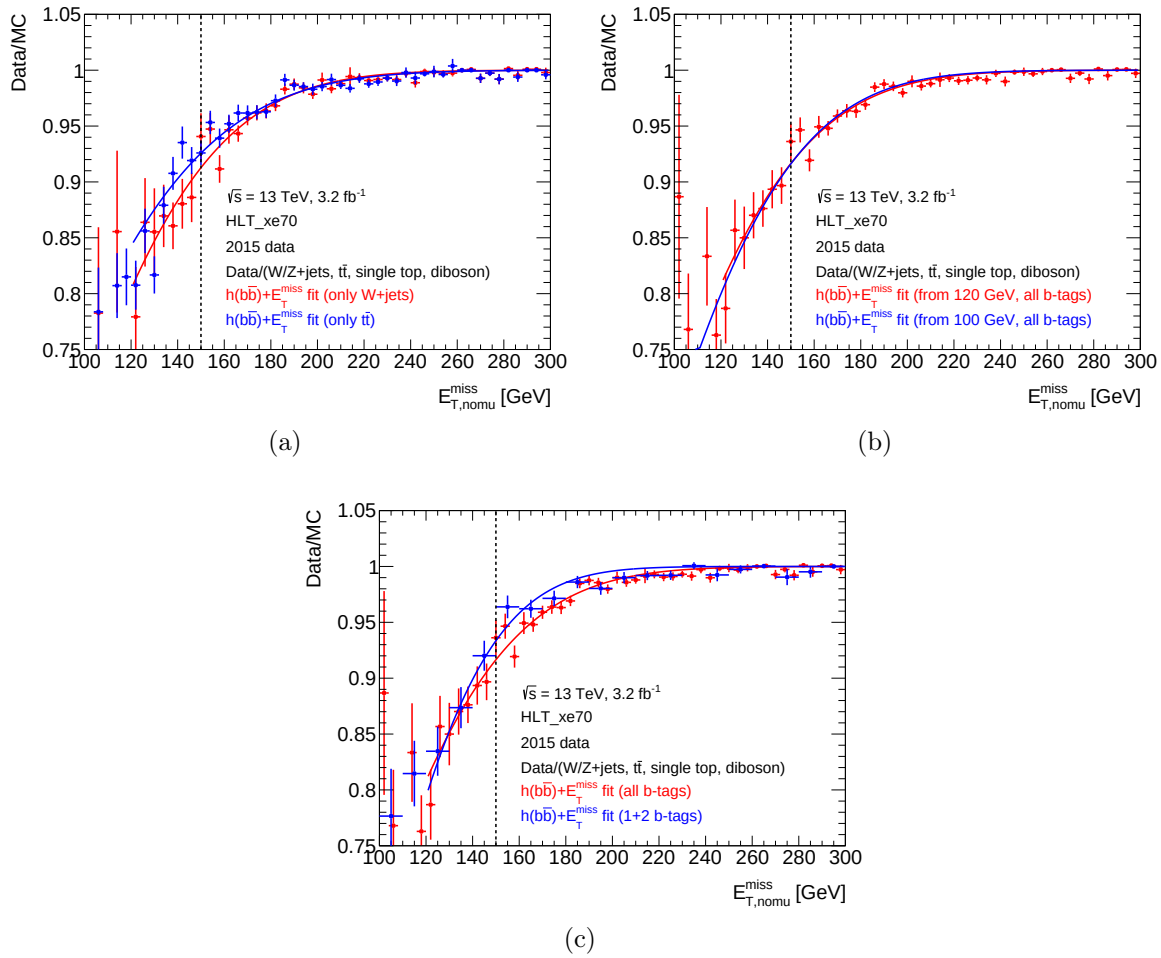


Figure A.2: Validation plots of $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiencies and corresponding scale factors as function of offline $E_{\mathrm{T,nomu}}^{\mathrm{miss}}$ measured in $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis. The data-to-MC scale factor corrections are applied for $E_{\mathrm{T}}^{\mathrm{miss}} > 120$ GeV. Plots for HLT_XE70 (top left), HLT_XE90_MHT_L1XE50 (top right), HLT_XE100_MHT_L1XE50 **OR** HLT_XE110_MHT_L1XE50 (bottom left) and HLT_XE110_MHT_L1XE50 (bottom right) triggers are shown in data and MC for 1 lepton control region. Plots are shown for all $b-$tagged analysis regions together. MC is dominated by $W+$jets and $t\bar{t}$ events.

Appendix B

Systematic variations for data-driven $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger scale factors



(a) (b)

(c)

Figure B.1: Different variations of scale factors for HLT_xe70 trigger as function of offline $E_{\mathrm{T,nomu}}^{\mathrm{miss}}$ measured in $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis. Figure (a) shows comparison between scale factors derived for $W+$jets only and $t\bar{t}$ only MC samples. Figure (b) shows comparison between the nominal scale factor fit starting from 120 GeV and the fit starting from 100 GeV. Figure (c) shows comparison between the scale factors derived for all $b-$tagged analysis regions together and for ≥ 1 $b-$tagged analysis regions. The nominal fit corresponds to the $E_{\mathrm{T}}^{\mathrm{miss}}$ bins of 2 GeV. The fits for $W+$jets MC samples, $t\bar{t}$ MC samples and for the extended fit range from 100 GeV corresponds to the $E_{\mathrm{T}}^{\mathrm{miss}}$ bins of 4 GeV, while the fit for ≥ 1 $b-$tagged analysis regions corresponds to the $E_{\mathrm{T}}^{\mathrm{miss}}$ bins of 10 GeV. This is done due to the lack of statistics.

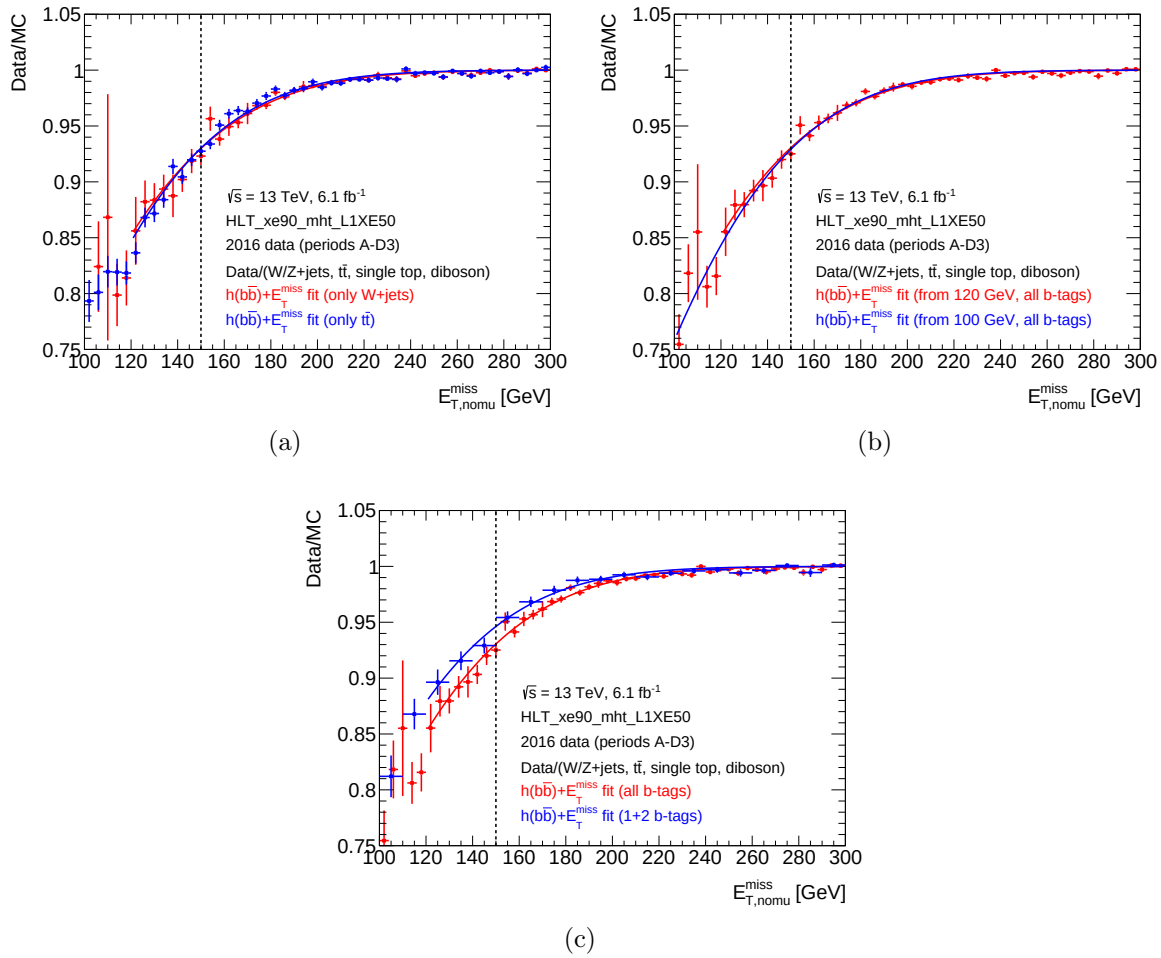


Figure B.2: Different variations of scale factors for HLT_xe90_mht_L1XE50 trigger as function of offline $E^{\text{miss}}_{\text{T,nomu}}$ measured in $E^{\text{miss}}_{\text{T}} + h(b\bar{b})$ analysis. Figure (a) shows comparison between scale factors derived for $W+$jets only and $t\bar{t}$ only MC samples. Figure (b) shows comparison between the nominal scale factor fit starting from 120 GeV and the fit starting from 100 GeV. Figure (c) shows comparison between the scale factors derived for all $b-$tagged analysis regions together and for ≥ 1 $b-$tagged analysis regions. The nominal fit corresponds to the $E^{\text{miss}}_{\text{T}}$ bins of 2 GeV. The fits for $W+$jets MC samples, $t\bar{t}$ MC samples and for the extended fit range from 100 GeV corresponds to the $E^{\text{miss}}_{\text{T}}$ bins of 4 GeV, while the fit for ≥ 1 $b-$tagged analysis regions corresponds to the $E^{\text{miss}}_{\text{T}}$ bins of 10 GeV. This is done due to the lack of statistics.

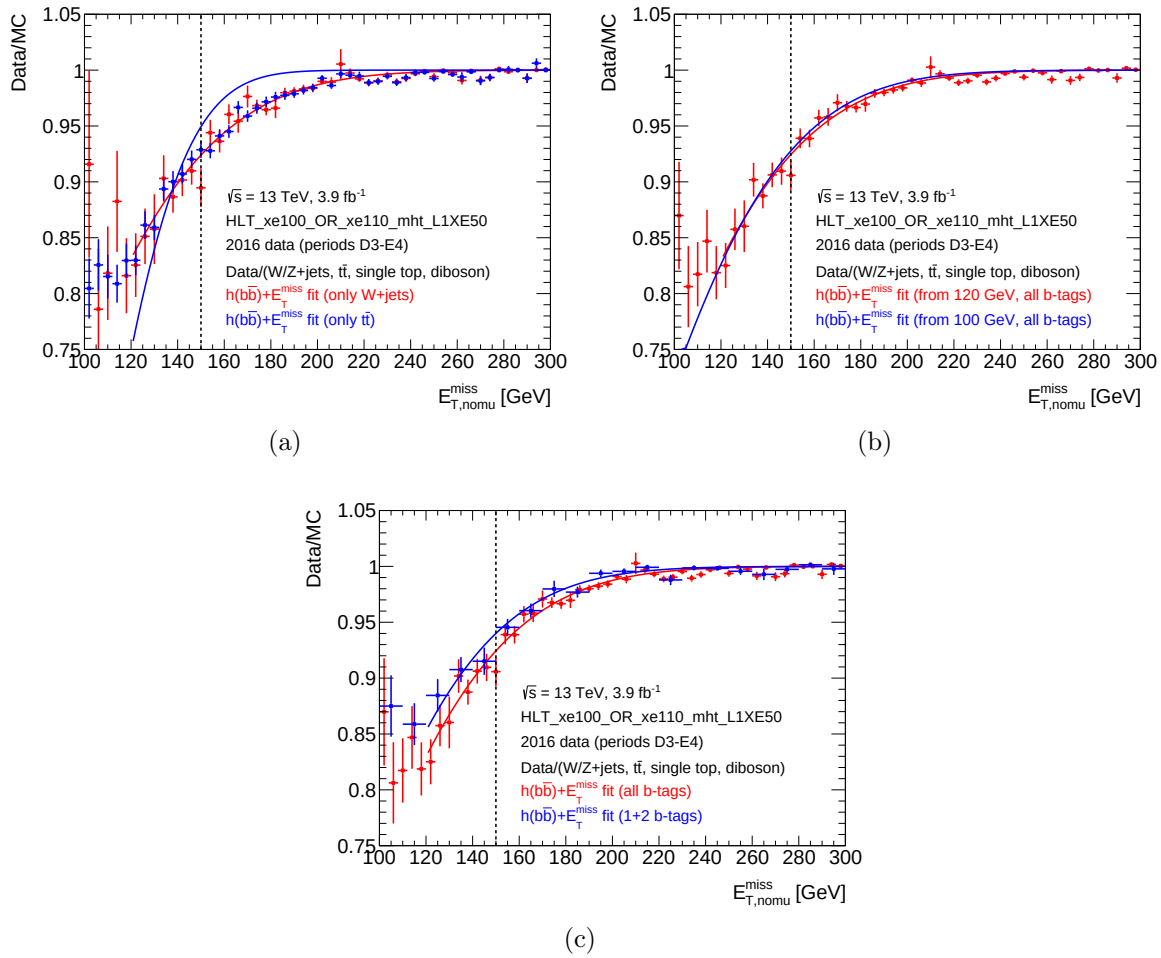


Figure B.3: Different variations of scale factors for HLT_xe100_mht_L1XE50 **OR** HLT_xe110_mht_L1XE50 trigger as function of offline $E_{\mathrm{T,nomu}}^{\mathrm{miss}}$ measured in $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis. Figure (a) shows comparison between scale factors derived for $W+$jets only and $t\bar{t}$ only MC samples. Figure (b) shows comparison between the nominal scale factor fit starting from 120 GeV and the fit starting from 100 GeV. Figure (c) shows comparison between the scale factors derived for all $b-$tagged analysis regions together and for ≥ 1 $b-$tagged analysis regions. The nominal fit corresponds to the $E_{\mathrm{T}}^{\mathrm{miss}}$ bins of 2 GeV. The fits for $W+$jets MC samples, $t\bar{t}$ MC samples and for the extended fit range from 100 GeV corresponds to the $E_{\mathrm{T}}^{\mathrm{miss}}$ bins of 4 GeV, while the fit for ≥ 1 $b-$tagged analysis regions corresponds to the $E_{\mathrm{T}}^{\mathrm{miss}}$ bins of 10 GeV. This is done due to the lack of statistics.

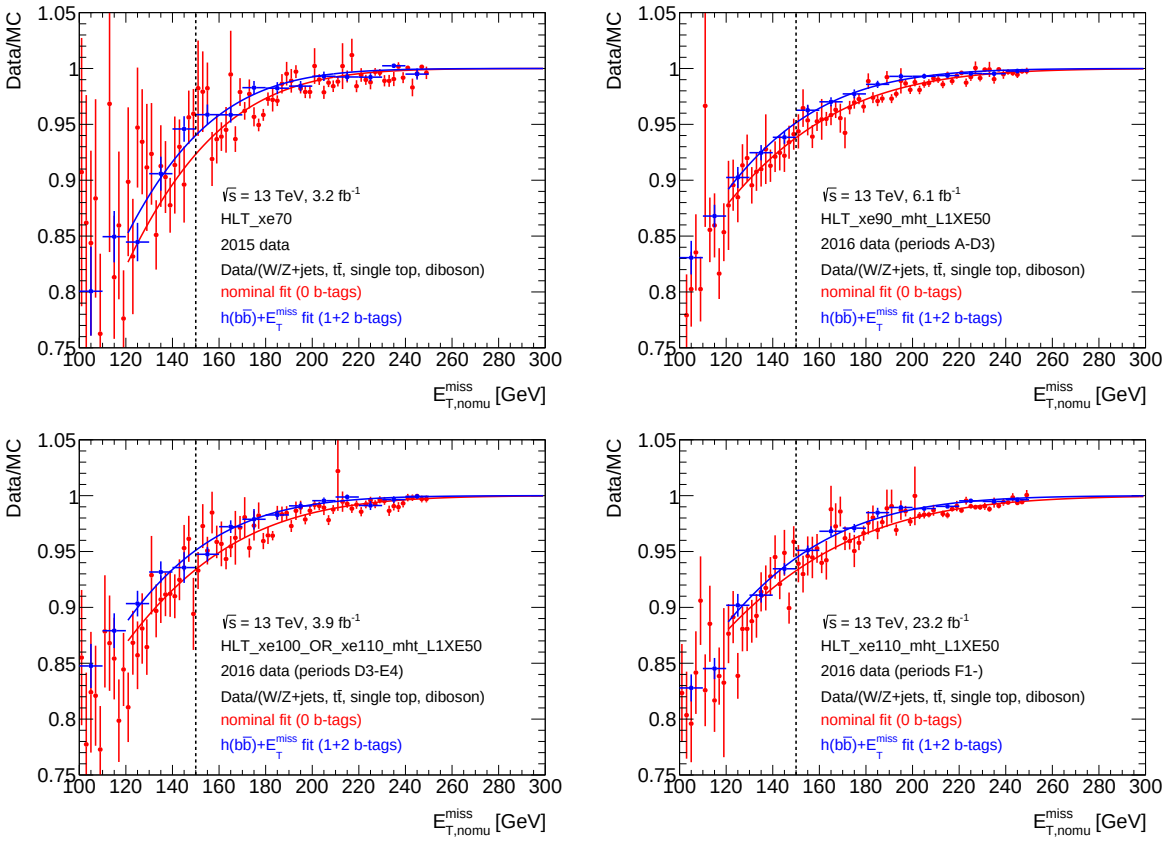


Figure B.4: Comparison between scale factors in regions with 0 $b-$tags and in regions with $\geq$ 1 $b-$tags as function of offline $E_{\mathrm{T,nomu}}^{\mathrm{miss}}$ measured in $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis. Plots for HLT_xe70 (top left), HLT_xe90_mht_L1XE50 (top right),HLT_xe100_mht_L1XE50 **OR** HLT_xe110_mht_L1XE50 (bottom left) and HLT_xe110_mht_L1XE50 (bottom right) triggers are shown in data and MC for 1 lepton control region. Scale factors in regions with 0 $b-$tags are shown in red. Scale factors in regions with ≥ 1 $b-$tags are shown in blue.

Appendix C

Correlation of nuisance parameters

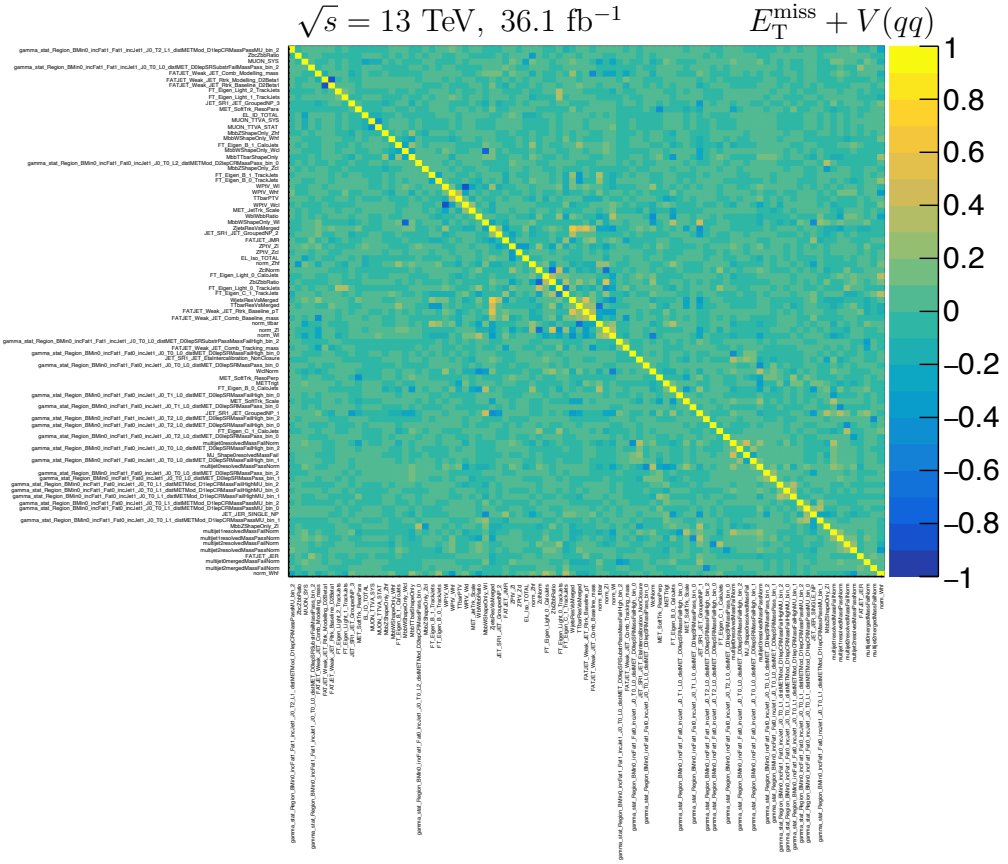


Figure C.1: Correlation of the nuisance parameters after for the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ combined binned profile likelihood fit to the observed data under the background-only hypothesis, $\mu = 0$.

Appendix D

Distributions in $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ control regions after the fit to data

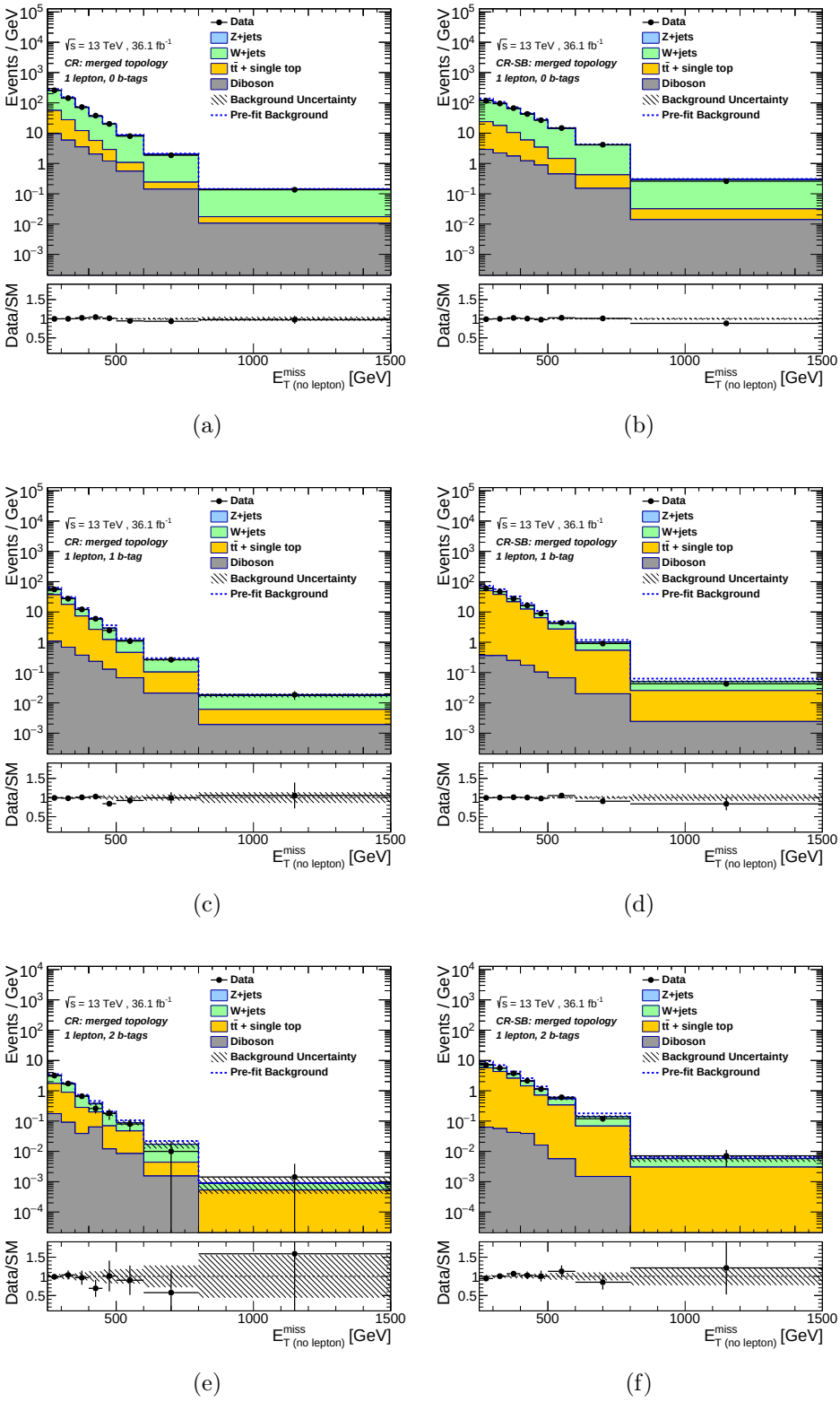


Figure D.1: The observed (dots) and expected (histograms) distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$ in the 1μCR of the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis after the fit to observed data for the merged regime. Plots are shown separately for the (a) 0b-MW, (b) 0b-SB, (c) 1b-MW, (d) 1b-SB, (e) 2b-MW and (f) 2b-SB event categories. Here, the abbrevaitions MW and SB stand for the mass window and mass sideband, respectively. The total background uncertainty is shown as a hatched band. The total background contribution before the fit to data is shown as a dotted blue line.

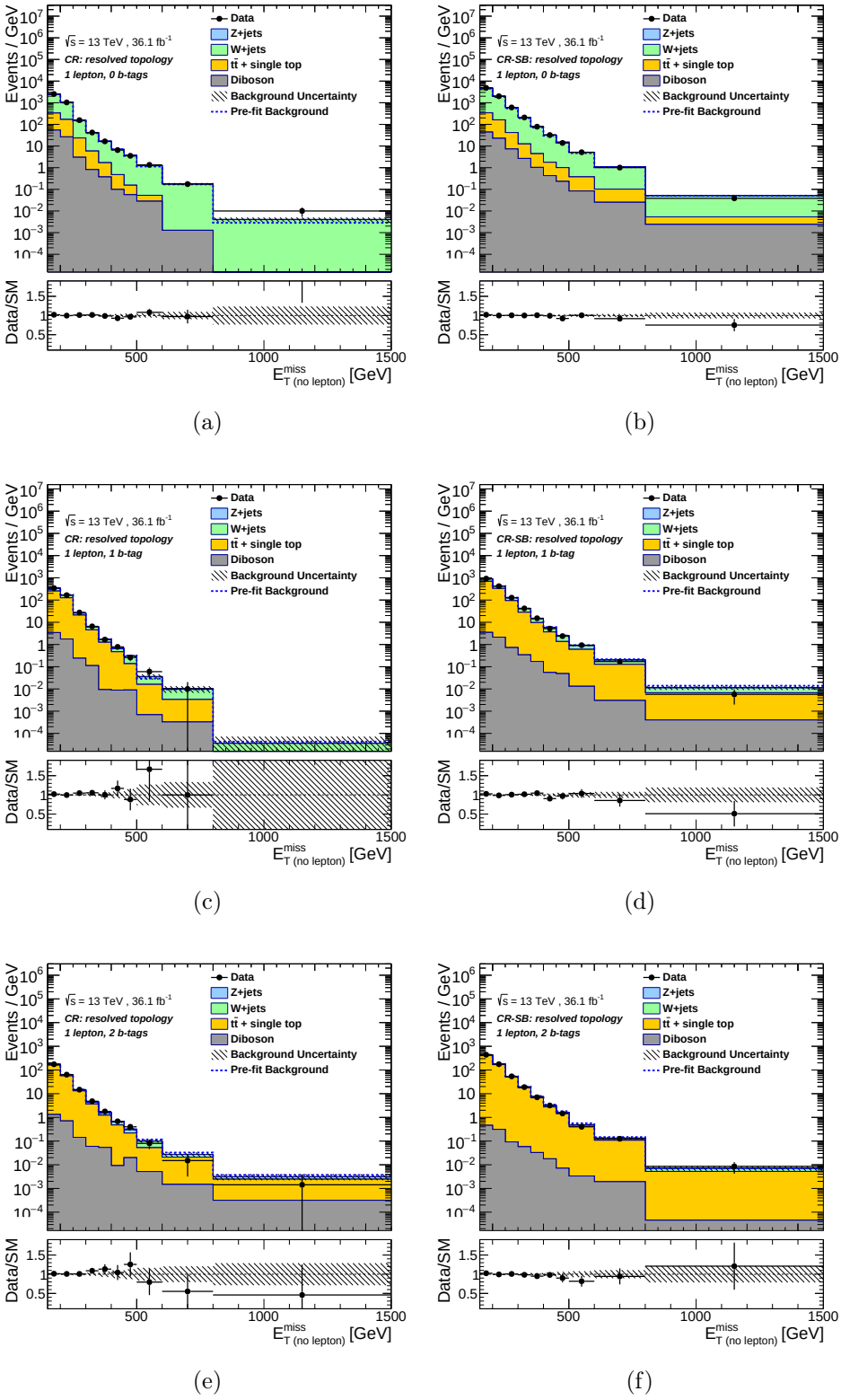


Figure D.2: The observed (dots) and expected (histograms) distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$ in the 1μCR of the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis after the fit to observed data for the resolved regime. Plots are shown separately for the (a) 0b-MW, (b) 0b-SB, (c) 1b-MW, (d) 1b-SB, (e) 2b-MW and (f) 2b-SB event categories. Here, the abbrevaitions MW and SB stand for the mass window and mass sideband, respectively. The total background uncertainty is shown as a hatched band. The total background contribution before the fit to data is shown as a dotted blue line.

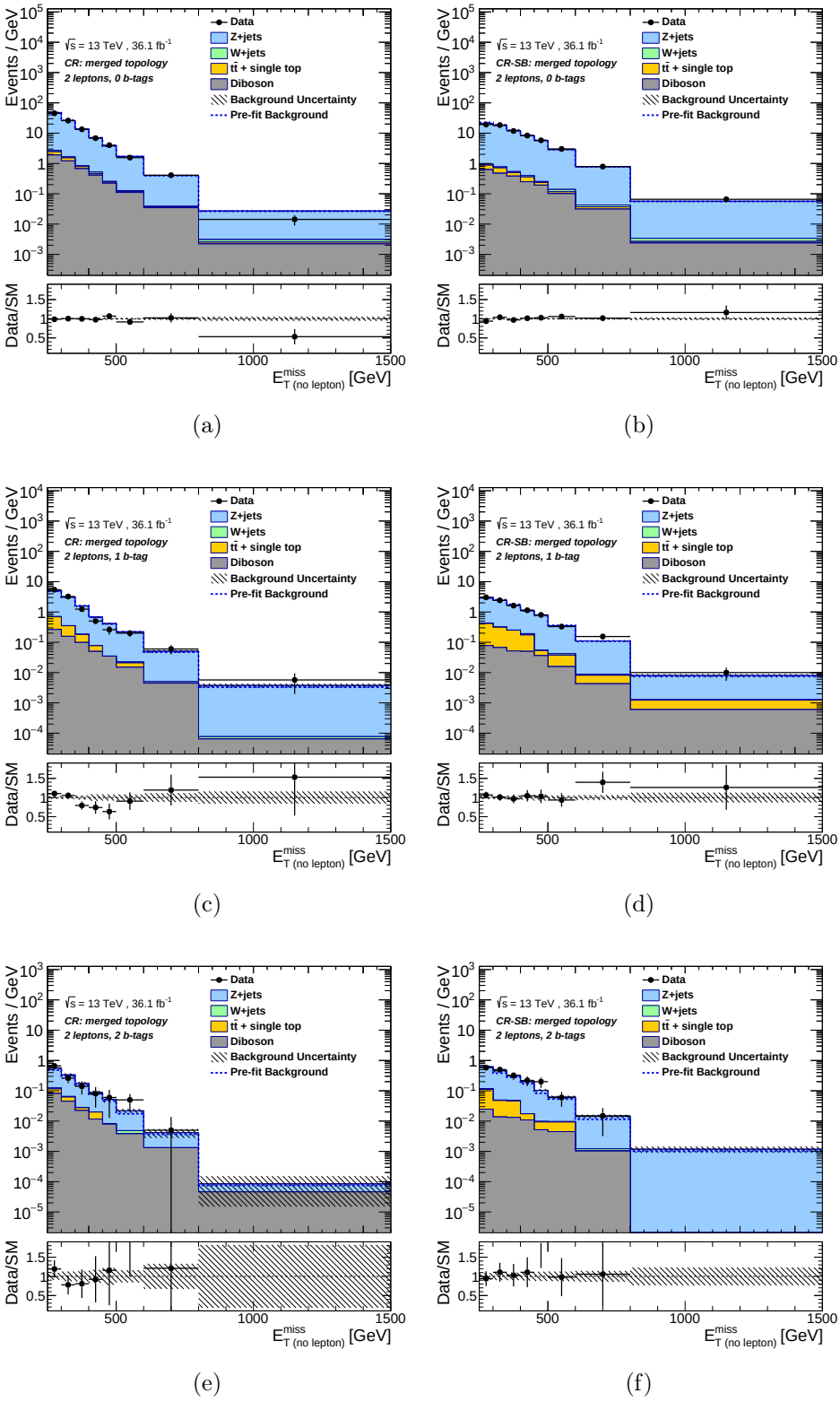


Figure D.3: The observed (dots) and expected (histograms) distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$ in the 2ℓCR of the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis after the fit to observed data for the merged regime. Plots are shown separately for the (a) 0b-MW, (b) 0b-SB, (c) 1b-MW, (d) 1b-SB, (e) 2b-MW and (f) 2b-SB event categories. Here, the abbrevaitions MW and SB stand for the mass window and mass sideband, respectively. The total background uncertainty is shown as a hatched band. The total background contribution before the fit to data is shown as a dotted blue line.

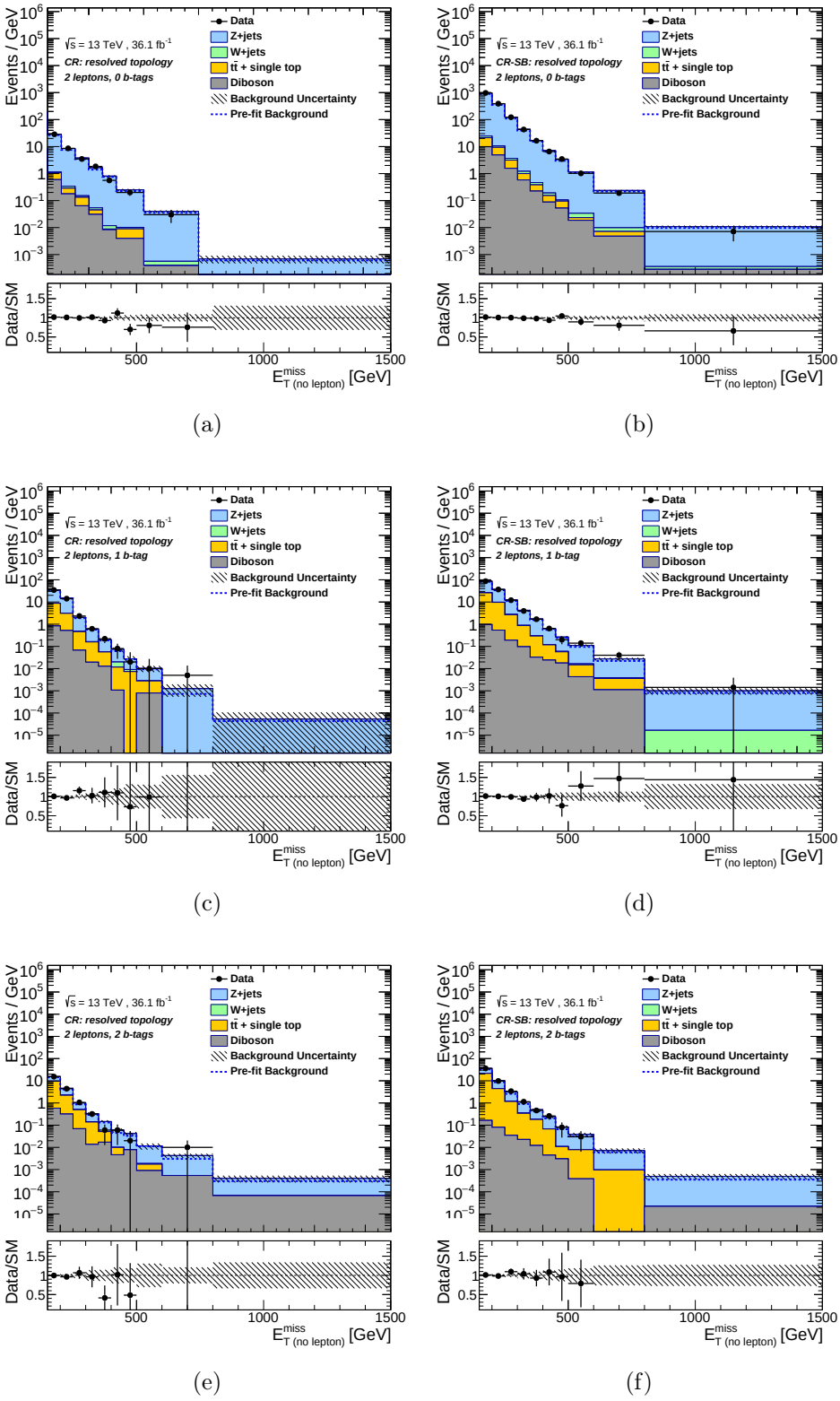


Figure D.4: The observed (dots) and expected (histograms) distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$ in the 2ℓCR of the $E_{\mathrm{T}}^{\mathrm{miss}} + V(qq)$ analysis after the fit to observed data for the resolved regime. Plots are shown separately for the (a) 0b-MW, (b) 0b-SB, (c) 1b-MW, (d) 1b-SB, (e) 2b-MW and (f) 2b-SB event categories. Here, the abbrevaitions MW and SB stand for the mass window and mass sideband, respectively. The total background uncertainty is shown as a hatched band. The total background contribution before the fit to data is shown as a dotted blue line.

Appendix E

Distributions in $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ 0 lepton signal region, 1 $b-$tag, after the fit to data

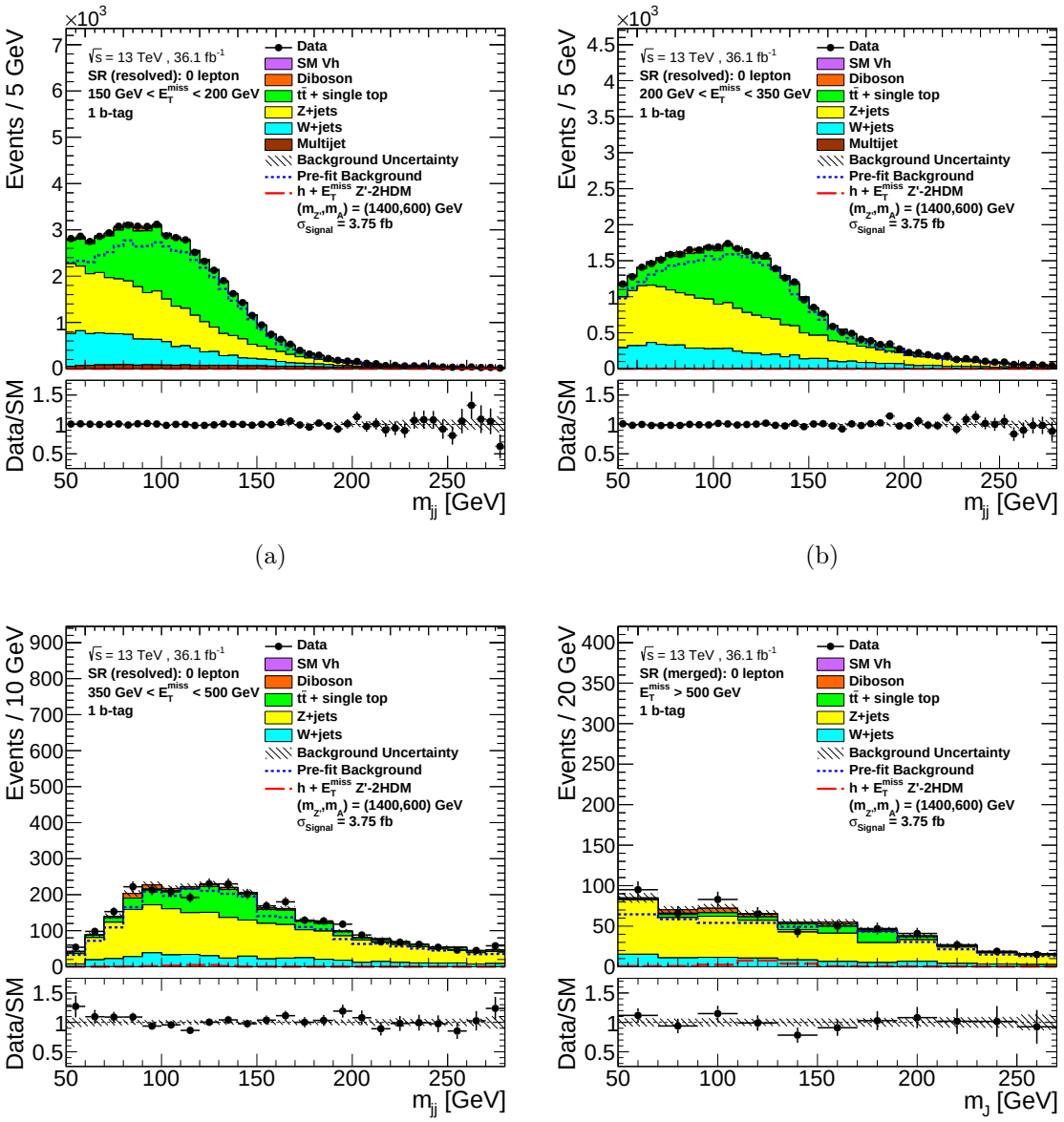


Figure E.1: The observed (dots) and expected (histograms) distributions of the invariant mass of the Higgs boson candidates in the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ signal region with 1 $b-$tag after the fit to observed data. Plots are shown separately for the (a) $(150, 200)$ GeV, (b) $(200, 350)$ GeV, (c) $(350, 500)$ GeV and (d) $(500, \infty)$ $E_{\mathrm{T}}^{\mathrm{miss}}$ regions. The red dashed line shows the expected DM signal coming from the Z'-2HDM simplified model with $(m_{Z'}, m_A) = (1400, 600)$ GeV. The total background uncertainty is shown as a hatched band. The total background contribution before the fit to data is shown as a dotted blue line.

Appendix F

Distributions in $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ control regions after the fit to data

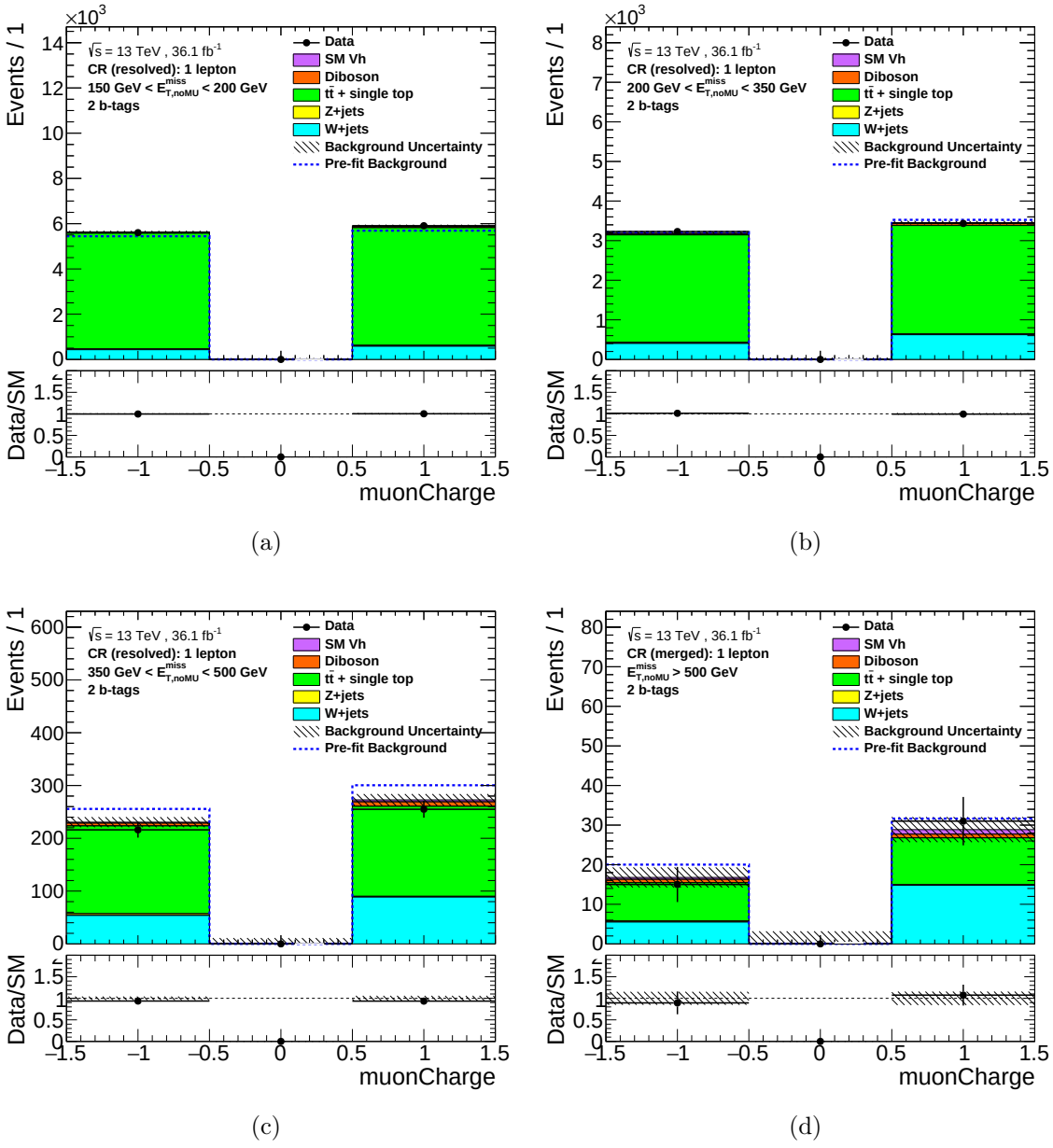


Figure F.1: The observed (dots) and expected (histograms) distributions of the muon charge in the 1μCR of the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis with 2 $b-$tags after the fit to observed data. Plots are shown separately for the (a) $[150, 200)$ GeV, (b) $[200, 350)$ GeV, (c) $[350, 500)$ GeV and (d) $[500, \infty)$ regions of $E_{\mathrm{T}}^{\mathrm{miss}}$. The total background uncertainty is shown as a hatched band. The total background contribution before the fit to data is shown as a dotted blue line.

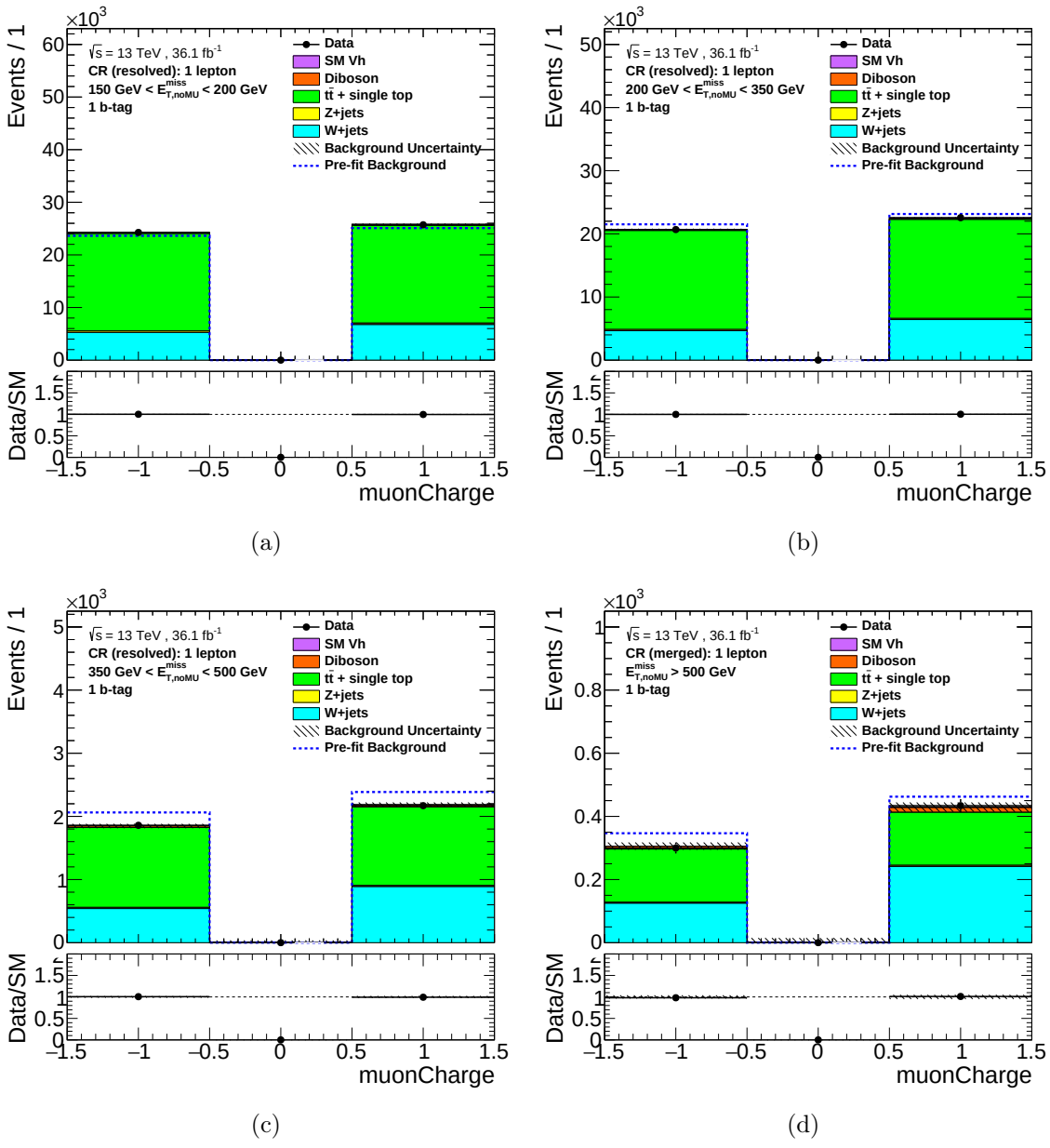


Figure F.2: The observed (dots) and expected (histograms) distributions of the muon charge in the 1μCR of the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis with 1 $b-$tag after the fit to observed data. Plots are shown separately for the (a) $[150, 200)$ GeV, (b) $[200, 350)$ GeV, (c) $[350, 500)$ GeV and (d) $[500, \infty)$ regions of $E_{\mathrm{T}}^{\mathrm{miss}}$. The total background uncertainty is shown as a hatched band. The total background contribution before the fit to data is shown as a dotted blue line.

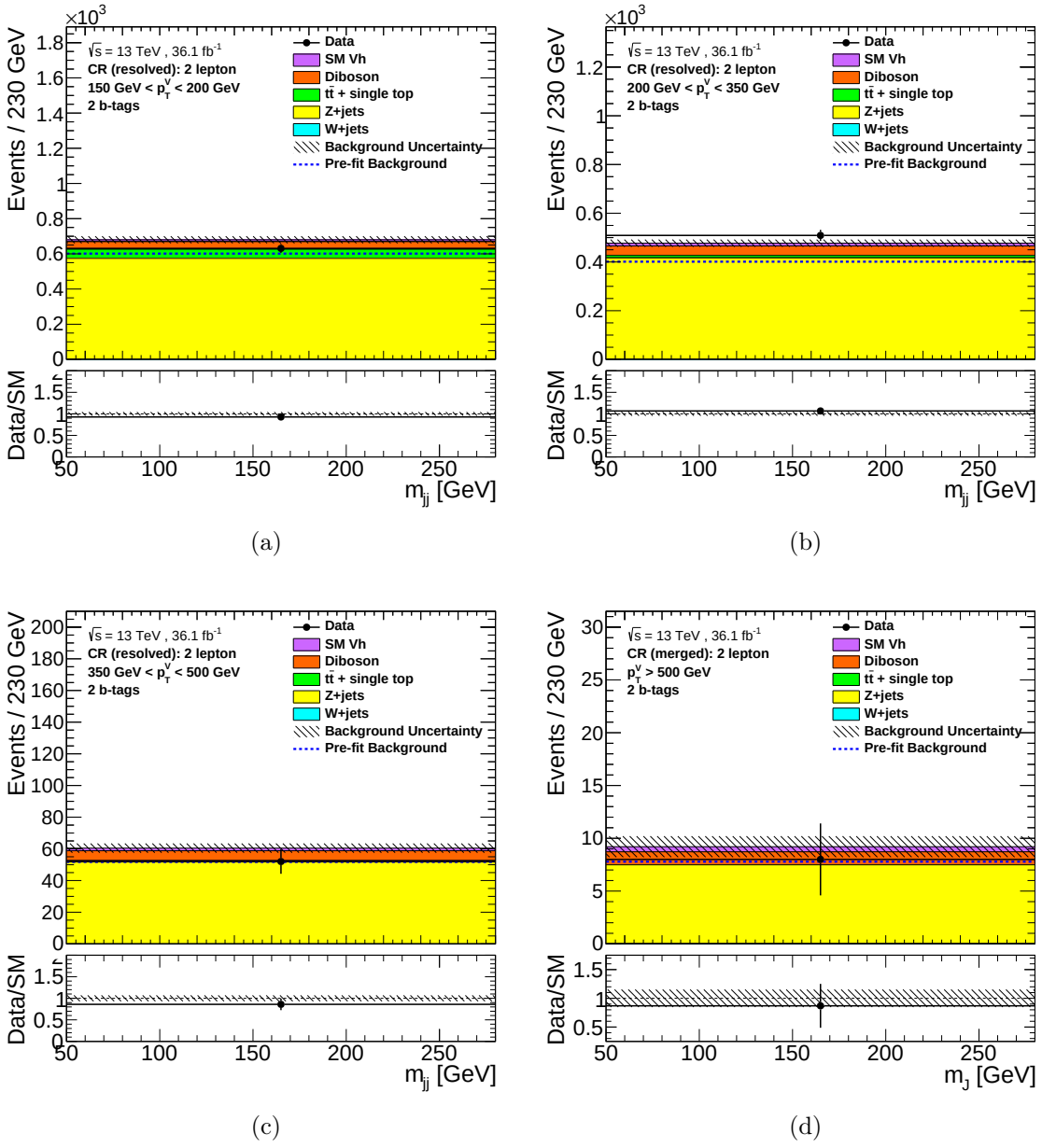


Figure F.3: The observed (dots) and expected (histograms) distributions of the muon charge in the 2ℓCR of the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis with 2 $b-$tags after the fit to observed data. Plots are shown separately for the (a) $[150, 200)$ GeV, (b) $[200, 350)$ GeV, (c) $[350, 500)$ GeV and (d) $[500, \infty)$ regions of $E_{\mathrm{T}}^{\mathrm{miss}}$. The total background uncertainty is shown as a hatched band. The total background contribution before the fit to data is shown as a dotted blue line.

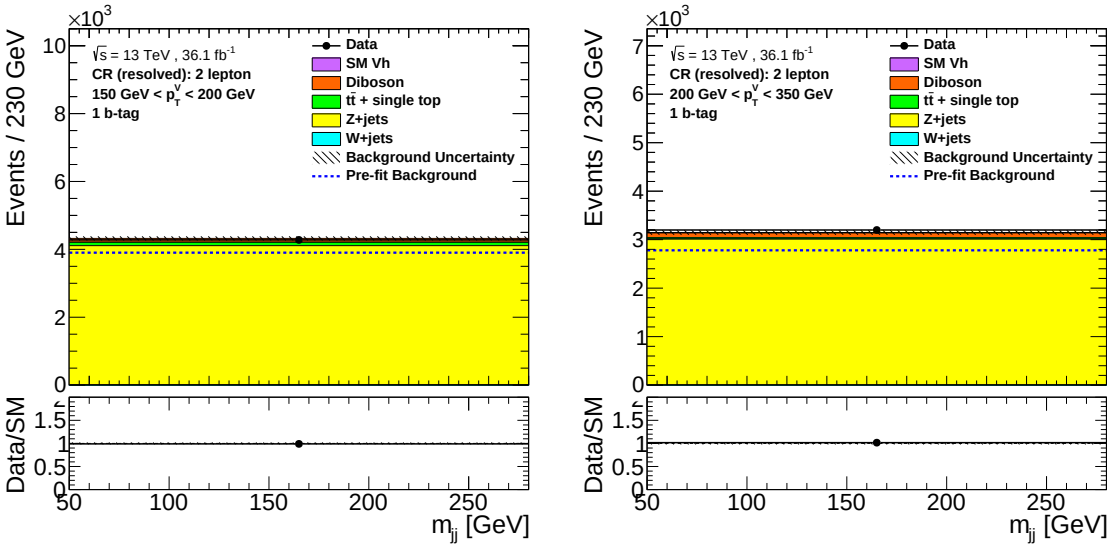


Figure F.4: The observed (dots) and expected (histograms) distributions of the muon charge in the 2ℓCR of the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ analysis with 1 $b-$tag after the fit to observed data. Plots are shown separately for the (a) $[150, 200)$ GeV, (b) $[200, 350)$ GeV, (c) $[350, 500)$ GeV and (d) $[500, \infty)$ regions of $E_{\mathrm{T}}^{\mathrm{miss}}$. The total background uncertainty is shown as a hatched band. The total background contribution before the fit to data is shown as a dotted blue line.

Appendix G

NLO rescaling: $k-$factors and uncertainties

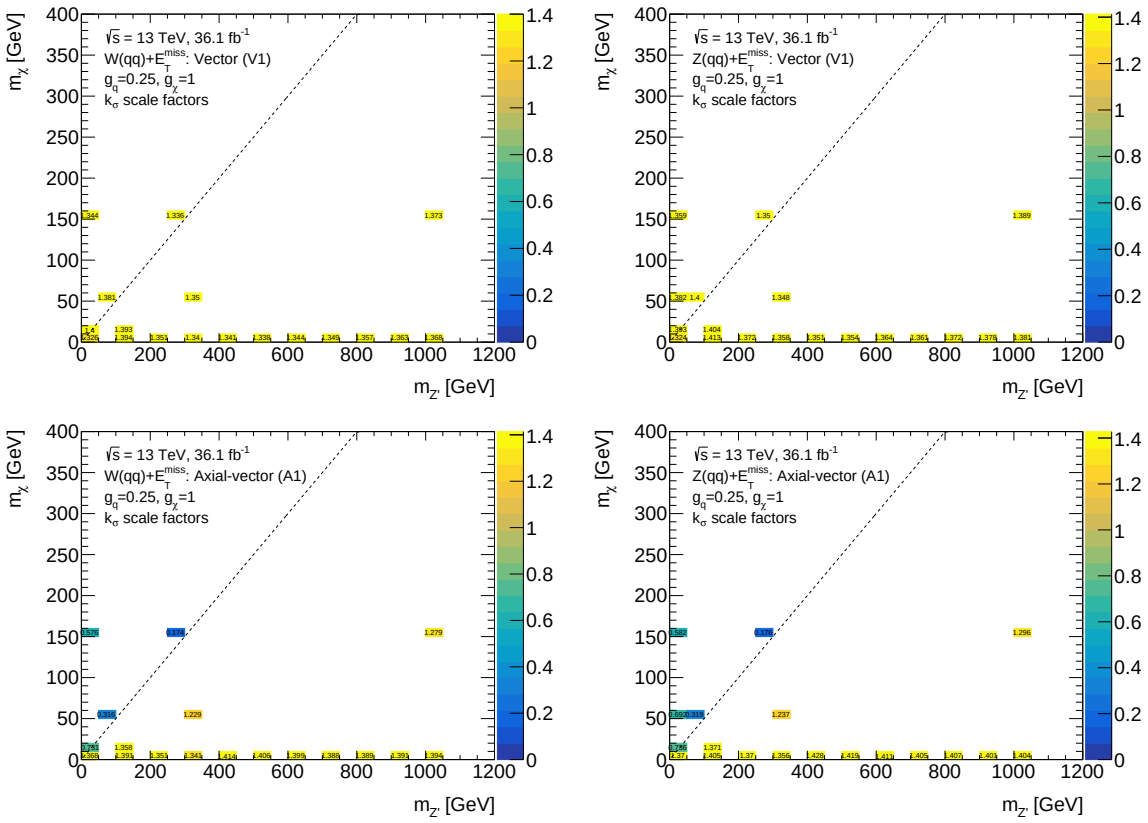


Figure G.1: The uncertainty on the k_σ scale factors for simplified vector (top) and axial-vector (bottom) mediator models of the $E_{\mathrm{T}}^{\mathrm{miss}} + W(qq)$ (left) and $E_{\mathrm{T}}^{\mathrm{miss}} + Z(qq)$ (right) production are shown in the $(m_\chi, m_{Z'})$ plane. The dashed black line shows the kinematic limit $m_{Z'} = 2m_\chi$.

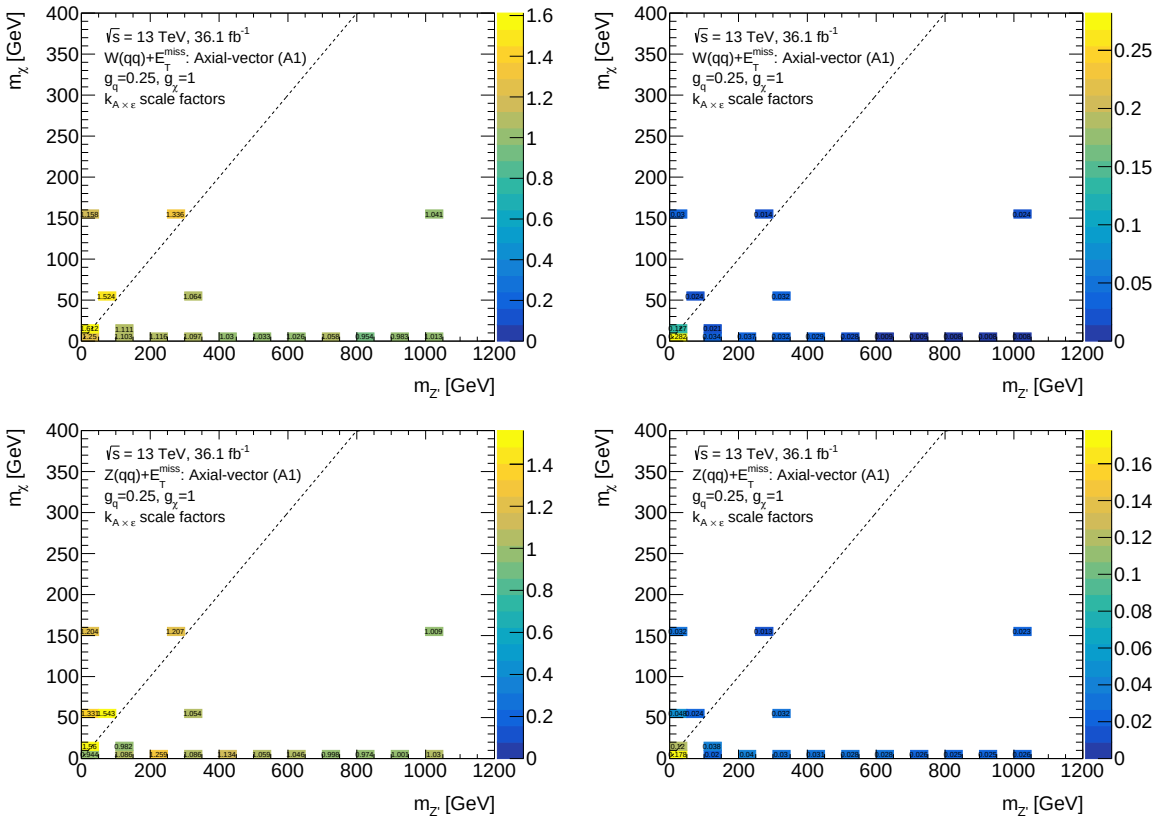


Figure G.2: The $k_{A\times\varepsilon}$ scale factors (left) and corresponding uncertainties (right) for the $E_{\mathrm{T}}^{\mathrm{miss}} + W(qq)$ (top) and $E_{\mathrm{T}}^{\mathrm{miss}}$+$Z(qq)$ (bottom) simplified axial-vector mediator models (A1) are shown in the $(m_{\chi}, m_{Z'})$ plane. The dashed black line shows the kinematic limit $m_{Z'} = 2m_{\chi}$.

Appendix H

Upper limits on $\sigma_{\mathrm{vis},W+\mathrm{DM}}$ and $\sigma_{\mathrm{vis},Z+\mathrm{DM}}$ for simplified vector mediator models in bins of $E_{\mathrm{T}}^{\mathrm{miss}}$

200 GeV $\leq E_{\text{T}}^{\text{miss}} <$ 250 GeV				
	$E_{\text{T}}^{\text{miss}} + W(qq)$		$E_{\text{T}}^{\text{miss}} + Z(qq)$	
$(m_\chi, m_{Z'})$ [GeV]	$\sigma_{\text{vis},W+\text{DM}}^{\text{exp}}$ [fb]	$\sigma_{\text{vis},W+\text{DM}}^{\text{obs}}$ [fb]	$\sigma_{\text{vis},Z+\text{DM}}^{\text{exp}}$ [fb]	$\sigma_{\text{vis},Z+\text{DM}}^{\text{obs}}$ [fb]
1 200	$144.65^{+56.66}_{-40.42}$	167.73	$57.22^{+22.41}_{-15.99}$	65.86
1 300	$142.49^{+55.82}_{-39.82}$	163.11	$59.52^{+23.32}_{-16.63}$	68.95
1 400	$144.35^{+56.55}_{-40.34}$	163.95	$50.65^{+19.84}_{-14.15}$	58.55
1 500	$146.51^{+57.39}_{-40.94}$	168.39	$46.67^{+18.28}_{-13.04}$	53.70
1 600	$149.46^{+58.54}_{-41.76}$	166.51	$45.83^{+17.95}_{-12.81}$	52.90
1 700	$145.04^{+56.81}_{-40.53}$	164.78	$51.85^{+20.31}_{-14.49}$	59.93
1 800	$156.13^{+61.16}_{-43.63}$	174.52	$50.86^{+19.92}_{-14.21}$	58.79
1 900	$151.16^{+59.21}_{-42.24}$	171.53	$51.77^{+20.28}_{-14.47}$	58.87
1 1000	$148.97^{+58.35}_{-41.63}$	164.48	$52.01^{+20.37}_{-14.53}$	59.94
1 2000	$160.40^{+62.83}_{-44.82}$	182.94	$51.73^{+20.26}_{-14.46}$	60.01
10 10	$154.70^{+60.60}_{-43.23}$	177.11	—	—
10 10000	$148.70^{+58.25}_{-41.55}$	169.03	$47.70^{+18.69}_{-13.33}$	56.08
50 95	$143.53^{+56.22}_{-40.11}$	150.73	—	—
50 300	$141.02^{+55.24}_{-39.41}$	159.92	—	—
150 10	$152.45^{+59.71}_{-42.60}$	169.48	$55.51^{+21.74}_{-15.51}$	66.44
150 1000	$156.04^{+61.12}_{-43.60}$	179.30	$40.32^{+15.80}_{-11.27}$	45.80
500 10	$143.95^{+56.39}_{-40.23}$	161.58	$48.07^{+18.83}_{-13.43}$	55.37
500 995	$162.54^{+63.67}_{-45.42}$	185.45	$54.61^{+21.39}_{-15.26}$	61.80
500 2000	$153.68^{+60.20}_{-42.95}$	178.53	$51.69^{+20.25}_{-14.45}$	59.67
500 10000	$156.27^{+61.21}_{-43.67}$	173.07	$55.11^{+21.59}_{-15.40}$	63.25
1000 10	$156.82^{+61.43}_{-43.82}$	171.65	$46.47^{+18.20}_{-12.99}$	53.55
1000 1000	$152.96^{+59.92}_{-42.74}$	178.32	$39.83^{+15.60}_{-11.13}$	46.02
1000 1995	$148.63^{+58.22}_{-41.53}$	174.18	$50.56^{+19.80}_{-14.13}$	58.61

Table H.1: Upper limits at 95% CL on $\sigma_{\text{vis},W+\text{DM}}$ and $\sigma_{\text{vis},Z+\text{DM}}$ in the $[200, 250)$ GeV region of $E_{\text{T}}^{\text{miss}}$. The limits are calculated using a range of simplified vector mediator models with the $(m_\chi, m_{Z'})$ points as indicated in the first column. The largest variation in the $E_{\text{T}}^{\text{miss}} + W(qq)$ ($E_{\text{T}}^{\text{miss}} + Z(qq)$) expected limits is 15% (40%). The weakest limits on the $E_{\text{T}}^{\text{miss}} + W(qq)$ and $E_{\text{T}}^{\text{miss}} + Z(qq)$ production in a given $E_{\text{T}}^{\text{miss}}$ bin correspond to the $(m_\chi, m_{Z'}) = (500, 995)$ GeV (red) and $(m_\chi, m_{Z'}) = (1, 300)$ GeV (blue) points, respectively.

250 GeV $\leq E_{\mathrm{T}}^{\mathrm{miss}} <$ 300 GeV				
	$E_{\mathrm{T}}^{\mathrm{miss}} + W(qq)$		$E_{\mathrm{T}}^{\mathrm{miss}} + Z(qq)$	
$(m_\chi, m_{Z'})$ [GeV]	$\sigma^{\mathrm{exp}}_{\mathrm{vis},W+\mathrm{DM}}$ [fb]	$\sigma^{\mathrm{obs}}_{\mathrm{vis},W+\mathrm{DM}}$ [fb]	$\sigma^{\mathrm{exp}}_{\mathrm{vis},Z+\mathrm{DM}}$ [fb]	$\sigma^{\mathrm{obs}}_{\mathrm{vis},Z+\mathrm{DM}}$ [fb]
1 200	—	—	$25.38^{+9.94}_{-7.09}$	37.49
1 300	$49.83^{+19.52}_{-13.93}$	43.02	$25.41^{+9.95}_{-7.10}$	34.86
1 400	$43.63^{+17.09}_{-12.19}$	41.02	$27.63^{+10.82}_{-7.72}$	41.10
1 500	$43.11^{+16.89}_{-12.05}$	39.92	$23.84^{+9.34}_{-6.66}$	32.38
1 600	$42.49^{+16.64}_{-11.87}$	41.17	$22.86^{+8.95}_{-6.39}$	28.12
1 700	$44.20^{+17.31}_{-12.35}$	42.27	$26.43^{+10.35}_{-7.39}$	36.84
1 800	$40.96^{+16.04}_{-11.45}$	38.11	$25.63^{+10.04}_{-7.16}$	36.56
1 900	$41.86^{+16.40}_{-11.70}$	36.56	$25.31^{+9.91}_{-7.07}$	36.43
1 2000	$38.10^{+14.92}_{-10.65}$	34.85	$28.96^{+11.34}_{-8.09}$	39.06
10 10000	$40.50^{+15.86}_{-11.32}$	36.68	$24.15^{+9.46}_{-6.75}$	32.55
150 1000	$46.31^{+18.14}_{-12.94}$	39.24	$24.41^{+9.56}_{-6.82}$	30.96
500 10	$46.27^{+18.13}_{-12.93}$	47.73	$22.08^{+8.65}_{-6.17}$	30.74
500 995	$49.72^{+19.47}_{-13.89}$	47.74	$27.81^{+10.89}_{-7.77}$	40.57
500 2000	$43.17^{+16.91}_{-12.06}$	41.14	$21.79^{+8.54}_{-6.09}$	28.56
500 10000	$41.04^{+16.07}_{-11.47}$	44.24	$25.64^{+10.04}_{-7.16}$	38.18
1000 10	$42.88^{+16.80}_{-11.98}$	38.48	$23.76^{+9.31}_{-6.64}$	33.39
1000 1000	$44.29^{+17.35}_{-12.38}$	44.10	$22.88^{+8.96}_{-6.39}$	34.82
1000 1995	$46.40^{+18.18}_{-12.97}$	43.86	$22.72^{+8.90}_{-6.35}$	29.24

Table H.2: Upper limits at 95% CL on $\sigma_{\mathrm{vis},W+\mathrm{DM}}$ and $\sigma_{\mathrm{vis},Z+\mathrm{DM}}$ in the $[250, 300)$ GeV region of $E_{\mathrm{T}}^{\mathrm{miss}}$. The limits are calculated using a range of simplified vector mediator models with the $(m_\chi, m_{Z'})$ points as indicated in the first column.The largest variation in the $E_{\mathrm{T}}^{\mathrm{miss}} + W(qq)$ ($E_{\mathrm{T}}^{\mathrm{miss}} + Z(qq)$) expected limits is 25% (25%). The weakest limits on the $E_{\mathrm{T}}^{\mathrm{miss}} + W(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + Z(qq)$ production in a given $E_{\mathrm{T}}^{\mathrm{miss}}$ bin correspond to the $(m_\chi, m_{Z'}) = (1, 300)$ GeV (red) and $(m_\chi, m_{Z'}) = (1, 2000)$ GeV (blue) points, respectively.

300 GeV $\leq E_{\text{T}}^{\text{miss}} <$ 400 GeV				
	$E_{\text{T}}^{\text{miss}} + W(qq)$		$E_{\text{T}}^{\text{miss}} + Z(qq)$	
$(m_\chi, m_{Z'})$ [GeV]	$\sigma_{\text{vis},W+\text{DM}}^{\text{exp}}$ [fb]	$\sigma_{\text{vis},W+\text{DM}}^{\text{obs}}$ [fb]	$\sigma_{\text{vis},Z+\text{DM}}^{\text{exp}}$ [fb]	$\sigma_{\text{vis},Z+\text{DM}}^{\text{obs}}$ [fb]
1 300	$27.80^{+10.89}_{-7.77}$	36.40	$17.69^{+6.93}_{-4.94}$	28.15
1 400	$26.76^{+10.48}_{-7.48}$	27.57	$14.43^{+5.65}_{-4.03}$	24.12
1 500	$29.10^{+11.40}_{-8.13}$	31.54	$15.82^{+6.20}_{-4.42}$	26.46
1 600	$27.30^{+10.70}_{-7.63}$	28.23	$14.32^{+5.61}_{-4.00}$	24.52
1 700	$25.51^{+9.99}_{-7.13}$	28.76	$18.48^{+7.24}_{-5.16}$	31.11
1 800	$30.04^{+11.77}_{-8.39}$	35.04	$16.28^{+6.38}_{-4.55}$	27.49
1 900	$26.81^{+10.50}_{-7.49}$	27.49	$16.68^{+6.53}_{-4.66}$	29.22
1 2000	$36.09^{+14.14}_{-10.08}$	40.64	$17.74^{+6.95}_{-4.96}$	30.20
10 10000	$29.64^{+11.61}_{-8.28}$	31.01	$13.93^{+5.46}_{-3.89}$	23.29
1000 10	$29.53^{+11.57}_{-8.25}$	30.88	$17.28^{+6.77}_{-4.83}$	29.75
1000 1000	$27.97^{+10.96}_{-7.82}$	30.18	$17.23^{+6.75}_{-4.81}$	31.21
1000 1995	$33.10^{+12.97}_{-9.25}$	36.46	$16.12^{+6.31}_{-4.50}$	27.23
150 1000	$30.23^{+11.84}_{-8.45}$	30.01	$17.84^{+6.99}_{-4.99}$	30.19
500 10	$29.95^{+11.73}_{-8.37}$	33.57	$13.51^{+5.29}_{-3.78}$	22.79
500 995	$32.02^{+12.54}_{-8.95}$	36.16	$15.82^{+6.20}_{-4.42}$	25.82
500 2000	$30.86^{+12.09}_{-8.62}$	37.41	$17.62^{+6.90}_{-4.92}$	32.78
500 10000	$27.43^{+10.74}_{-7.66}$	30.02	$17.09^{+6.69}_{-4.78}$	27.71

Table H.3: Upper limits at 95% CL on $\sigma_{\text{vis},W+\text{DM}}$ and $\sigma_{\text{vis},Z+\text{DM}}$ in the $[300, 400)$ GeV region of $E_{\text{T}}^{\text{miss}}$. The limits are calculated using a range of simplified vector mediator models with the $(m_\chi, m_{Z'})$ points as indicated in the first column. The largest variation in the $E_{\text{T}}^{\text{miss}} + W(qq)$ ($E_{\text{T}}^{\text{miss}} + Z(qq)$) expected limits is 35% (30%). The weakest limits on the $E_{\text{T}}^{\text{miss}} + W(qq)$ and $E_{\text{T}}^{\text{miss}} + Z(qq)$ production in a given $E_{\text{T}}^{\text{miss}}$ bin correspond to the $(m_\chi, m_{Z'}) = (1, 2000)$ GeV (red) and $(m_\chi, m_{Z'}) = (1, 700)$ GeV (blue) points, respectively.

400 GeV $\leq E_{\mathrm{T}}^{\mathrm{miss}} <$ 600 GeV				
	$E_{\mathrm{T}}^{\mathrm{miss}} + W(qq)$		$E_{\mathrm{T}}^{\mathrm{miss}} + Z(qq)$	
$(m_\chi, m_{Z'})$ [GeV]	$\sigma^{\mathrm{exp}}_{\mathrm{vis},W+\mathrm{DM}}$ [fb]	$\sigma^{\mathrm{obs}}_{\mathrm{vis},W+\mathrm{DM}}$ [fb]	$\sigma^{\mathrm{exp}}_{\mathrm{vis},Z+\mathrm{DM}}$ [fb]	$\sigma^{\mathrm{obs}}_{\mathrm{vis},Z+\mathrm{DM}}$ [fb]
1 300	$11.76^{+4.61}_{-3.29}$	13.23	$7.13^{+2.79}_{-1.99}$	6.92
1 400	$11.12^{+4.36}_{-3.11}$	8.81	$7.95^{+3.12}_{-2.22}$	7.50
1 500	$10.18^{+3.99}_{-2.85}$	8.39	$8.63^{+3.38}_{-2.41}$	8.17
1 600	$11.05^{+4.33}_{-3.09}$	8.82	$7.65^{+3.00}_{-2.14}$	7.29
1 700	$11.49^{+4.50}_{-3.21}$	9.71	$8.07^{+3.16}_{-2.26}$	8.27
1 800	$11.09^{+4.34}_{-3.10}$	8.91	$8.44^{+3.31}_{-2.36}$	8.30
1 900	$10.97^{+4.30}_{-3.07}$	8.89	$7.38^{+2.89}_{-2.06}$	7.11
1 1000	—	—	$8.63^{+3.38}_{-2.41}$	8.04
1 2000	$11.76^{+4.60}_{-3.28}$	9.89	$8.74^{+3.42}_{-2.44}$	8.44
10 10000	$12.32^{+4.83}_{-3.44}$	9.65	$7.89^{+3.09}_{-2.21}$	7.39
150 1000	$10.58^{+4.14}_{-2.96}$	8.98	$7.76^{+3.04}_{-2.17}$	7.63
500 10	$11.06^{+4.33}_{-3.09}$	8.64	$7.30^{+2.86}_{-2.04}$	6.91
500 995	$12.21^{+4.78}_{-3.41}$	10.25	$7.68^{+3.01}_{-2.14}$	7.04
500 2000	$11.65^{+4.57}_{-3.26}$	9.87	$9.08^{+3.56}_{-2.54}$	9.24
500 10000	$12.38^{+4.85}_{-3.46}$	9.62	$8.69^{+3.40}_{-2.43}$	8.83
1000 10	$11.58^{+4.54}_{-3.24}$	9.14	$7.51^{+2.94}_{-2.10}$	7.44
1000 1000	$10.79^{+4.23}_{-3.02}$	8.94	$8.40^{+3.29}_{-2.35}$	7.55
1000 1995	$12.64^{+4.95}_{-3.53}$	9.67	$8.11^{+3.18}_{-2.27}$	8.42

Table H.4: Upper limits at 95% CL on $\sigma_{\mathrm{vis},W+\mathrm{DM}}$ and $\sigma_{\mathrm{vis},Z+\mathrm{DM}}$ in the $[250, 300)$ GeV region of $E_{\mathrm{T}}^{\mathrm{miss}}$. The limits are calculated using a range of simplified vector mediator models with the $(m_\chi, m_{Z'})$ points as indicated in the first column. The largest variation in the $E_{\mathrm{T}}^{\mathrm{miss}} + W(qq)$ ($E_{\mathrm{T}}^{\mathrm{miss}} + Z(qq)$) expected limits is 20% (25%). The weakest limits on the $E_{\mathrm{T}}^{\mathrm{miss}} + W(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}} + Z(qq)$ production in a given $E_{\mathrm{T}}^{\mathrm{miss}}$ bin correspond to the $(m_\chi, m_{Z'}) = (1000, 1995)$ GeV (red) and $(m_\chi, m_{Z'}) = (500, 2000)$ GeV (blue) points, respectively.

Appendix I

Upper limits on $\sigma_{\text{vis},h+\text{DM}}$ for Z'-2HDM models in bins of $E_{\text{T}}^{\text{miss}}$

$E_{\text{T}}^{\text{miss}} + h(bb)$, 150 GeV $\leq E_{\text{T}}^{\text{miss}} <$ 200 GeV						
$(m_{Z'}, m_A)$ [GeV]	$\sigma^{\text{exp}}_{\text{vis},h+\text{DM}}$ [fb]	$\sigma^{\text{exp}}_{\text{vis},h+\text{DM}}$ [fb]	-2σ [fb]	-1σ [fb]	+1σ [fb]	+2σ [fb]
400 500	10.495	14.640	5.633	7.562	14.606	19.581
600 200	8.182	10.917	4.391	5.895	11.386	15.264
600 500	10.519	13.115	5.646	7.579	14.639	19.625
800 300	7.869	10.468	4.224	5.670	10.952	14.681
800 500	8.666	11.305	4.651	6.244	12.061	16.168
800 700	8.309	10.838	4.459	5.987	11.563	15.501
1000 400	8.563	11.177	4.596	6.170	11.918	15.977
1000 600	7.966	10.593	4.276	5.740	11.086	14.862
1000 800	8.183	10.718	4.392	5.896	11.389	15.267
1200 800	8.225	11.127	4.415	5.927	11.447	15.346

Table I.1: Upper limits at 95% CL on $\sigma_{\text{vis},h+\text{DM}}$ in the $[200, 350)$ GeV region of $E_{\text{T}}^{\text{miss}}$. The limits are calculated using a range of Z'-2HDM models with the $(m_{Z'}, m_A)$ points as indicated in the first column. The largest variation in the $E_{\text{T}}^{\text{miss}} + h(b\bar{b})$ expected limits is 15%. The weakest expected limit in a given $E_{\text{T}}^{\text{miss}}$ bin corresponds to the $(m_{Z'}, m_A) = (600, 500)$ GeV point (red).

$E_{\text{T}}^{\text{miss}} + h(bb)$, 350 GeV $\leq E_{\text{T}}^{\text{miss}} <$ 500 GeV						
$(m_{Z'}, m_A)$ [GeV]	$\sigma^{\text{exp}}_{\text{vis},h+\text{DM}}$ [fb]	$\sigma^{\text{exp}}_{\text{vis},h+\text{DM}}$ [fb]	-2σ [fb]	-1σ [fb]	+1σ [fb]	+2σ [fb]
800 200	1.453	1.910	0.780	1.047	2.022	2.711
1000 400	1.454	1.904	0.780	1.048	2.024	2.713
1000 600	1.525	2.046	0.818	1.099	2.122	2.844
1200 300	1.682	2.406	0.903	1.212	2.341	3.138
1200 500	1.505	2.023	0.808	1.084	2.095	2.808
1200 800	1.484	1.995	0.797	1.070	2.066	2.770
1200 800	1.484	1.995	0.797	1.070	2.066	2.770
1400 600	1.708	2.418	0.917	1.231	2.378	3.187
1400 800	1.562	2.149	0.839	1.126	2.175	2.915
1600 800	1.701	2.431	0.913	1.226	2.368	3.174

Table I.2: Upper limits at 95% CL on $\sigma_{\text{vis},h+\text{DM}}$ in the $[350, 500)$ GeV region of $E_{\text{T}}^{\text{miss}}$. The limits are calculated using a range of Z'-2HDM models with the $(m_{Z'}, m_A)$ points as indicated in the first column. The largest variation in the $E_{\text{T}}^{\text{miss}} + h(b\bar{b})$ expected limits is 10%. The weakest expected limit in a given $E_{\text{T}}^{\text{miss}}$ bin corresponds to the $(m_{Z'}, m_A) = (1400, 600)$ GeV point (red).

Bibliography

[1] J. J. Thomson, *Cathode rays*, Phil. Mag. **44** (1897) pp. 293–316.

[2] E. Rutherford, *The scattering of alpha and beta particles by matter and the structure of the atom*, Phil. Mag. **21** (1911) pp. 669–688.

[3] J. Chadwick, *Possible Existence of a Neutron*, Nature **129** (1923) p. 312.

[4] M. Gell-Mann, *A Schematic Model of Baryons and Mesons*, Phys. Lett. **8** (1964) pp. 214–215.

[5] G. Zweig, *An SU_3 model for strong interaction symmetry and its breaking*, Developments in the Quark Theory of Hadrons **1** (1964) pp. 22–101, URL: https://cds.cern.ch/record/570209.

[6] E. D. B. et al., *High-Energy Inelastic $e-p$ Scattering at 6-Degrees and 10-Degrees*, Phys. Rev. Lett. **23** (1969) 930–934.

[7] M. B. et al., *Observed Behavior of Highly Inelastic Electron-Proton Scattering*, Phys. Rev. Lett. **23** (1969) 935–939.

[8] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) pp. 1264–1266.

[9] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. **C680519** (1968) pp. 367–377.

[10] R. B. et al., *Evidence for Planar Events in e^+e^- Annihilation at High-Energies*, Phys. Lett. B **86** (1979) pp. 243–249.

[11] D. P. B. et al., *Discovery of Three Jet Events and a Test of Quantum Chromodynamics at PETRA Energies*, Phys. Rev. Lett. **43** (1979) p. 830.

[12] C. B. et al., *Evidence for Gluon Bremsstrahlung in e^+e^- Annihilations at High-Energies*, Phys. Lett. B **88** (1979) pp. 417–425.

[13] W. B. et al., *Observation of Planar Three Jet Events in e^+e^- Annihilation and Evidence for Gluon Bremsstrahlung*, Phys. Lett. B **91** (1980) pp. 142–147.

[14] G. A. et al., *Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s} = 540$ GeV*, Phys. Lett. B **122** (1983) pp. 103–106.

[15] M. B. et al., *Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN pp Collider*, Phys. Lett. B **122** (1983) 476–485.

[16] G. A. et al., *Experimental Observation of Lepton Pairs of Invariant Mass Around* $95 GeV/c^2$ *at the CERN SPS Collider*, Phys. Lett. B **126** (1983) 398–410.

[17] P. B. et al., *Evidence for* $Z^0 \to e^+e^-$ *at the CERN pp Collider*, Phys. Lett. B **129** (1983) 130–140.

[18] D0 collaboration, *Observation of the top quark*, Phys. Rev. Lett. **74** (1995) pp. 2632–2637, arXiv: hep-ex/9503003 [hep-ex].

[19] CDF collaboration, *Observation of top quark production in* $\bar{p}p$ *collisions*, Phys. Rev. Lett. **74** (1995) pp. 2626–2631, arXiv: hep-ex/9503002 [hep-ex].

[20] D. collaboration, *Observation of tau neutrino interactions*, Phys. Lett. **B504** (2001) pp. 218–224, arXiv: hep-ex/0012035 [hep-ex].

[21] ATLAS collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) pp. 1–29, arXiv: 1207.7214 [hep-ph].

[22] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012) p. 30, arXiv: 1207.7235 [hep-ph].

[23] J. C. Kapteyn, *First attempt at a theory of the arrangement and motion of the sidereal system*, Astrophysical Journal **55** (1922) 302–327.

[24] J. H. Oort, *The force exerted by the stellar system in the direction perpendicular to the galactic plane and some related problems*, Bulletin of the Astronomical Institutes of the Netherlands **6** (1932) 249–287.

[25] F. Zwicky, *Die rotverschiebung von extragalaktischen nebeln*, Helvetica Physica Acta **6** (1933) 110–127.

[26] Planck Collaboration, *Planck 2015 results. I. Overview of products and scientific results*, Astron. Astrophys. **594** (2016), arXiv: 1502.01582 [astro-ph].

[27] WMAP Collaboration, *Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results*, Astrophys. J. Suppl. **208** (2013), arXiv: 1212.5226 [astro-ph].

[28] ATLAS collaboration, *Search for new phenomena in final states with an energetic jet and large missing transverse momentum in pp collisions at* $\sqrt{s} = 8$ *TeV with the ATLAS detector*, Eur. Phys. J. C. **75** (2015) p. 299, arXiv: 1502.01518 [hep-ex].

[29] ATLAS collaboration, *Search for new phenomena in final states with an energetic jet and large missing transverse momentum in pp collisions at* $\sqrt{s} = 13$ *TeV using the ATLAS detector*, Phys. Rev. D **94** (2016) p. 032005, arXiv: 1604.07773 [hep-ex].

[30] ATLAS collaboration, *Search for dark matter and other new phenomena in events with an energetic jet and large missing transverse momentum using the ATLAS detector*, JHEP **01** (2018) p. 126, arXiv: 1711.03301 [hep-ex].

[31] CMS Collaboration, *Search for dark matter, extra dimensions, and unparticles in monojet events in protonproton collisions at* $\sqrt{s} = 8$ *TeV*, Eur. Phys. J. C. **75** (2015) p. 235, arXiv: 1408.3583 [hep-ex].

[32] ATLAS collaboration, *Search for dark matter in events with heavy quarks and missing transverse momentum in pp collisions with the ATLAS detector*, Eur. Phys. J. C **75** (2015) p. 92, arXiv: 1410.4031 [hep-ex].

[33] CMS Collaboration, *Search for Monotop Signatures in Proton-Proton Collisions at* $\sqrt{s} = 8$ *TeV*, Phys. Rev. Lett. **114** (2015) p. 101801, arXiv: 1410.1149 [hep-ex].

[34] CMS Collaboration, *Search for the production of dark matter in association with top-quark pairs in the single-lepton final state in proton-proton collisions at* $\sqrt{s} = 8$ *TeV*, JHEP **06** (2015) p. 121, arXiv: 1504.03198 [hep-ex].

[35] ATLAS collaboration, *Search for new phenomena in events with a photon and missing transverse momentum in pp collisions at* $\sqrt{s} = 8$ *TeV with the ATLAS detector*, Phys. Rev. D **91** (2015) p. 012008, arXiv: 1411.1559 [hep-ex].

[36] ATLAS collaboration, *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** (2016) p. 059, arXiv: 1604.01306 [hep-ex].

[37] CMS Collaboration, *Search for Dark Matter and Large Extra Dimensions in pp Collisions Yielding a Photon and Missing Transverse Energy*, Phys. Rev. Lett. **108** (2012) p. 261803, arXiv: 1204.0821 [hep-ex].

[38] ATLAS collaboration, *Search for dark matter in events with a Z boson and missing transverse momentum in pp collisions at* $\sqrt{s}$*=8 TeV with the ATLAS detector*, Phys. Rev. D **90** (2014) p. 012004, arXiv: 1404.0051 [hep-ex].

[39] ATLAS collaboration, *Search for dark matter in events with a hadronically decaying W or Z boson and missing transverse momentum in pp collisions at* $\sqrt{s}$*=8 TeV with the ATLAS detector*, Phys. Rev. Lett. **112** (2014) p. 041802, arXiv: 1309.4017 [hep-ex].

[40] CMS Collaboration, *Search for dark matter and unparticles produced in association with a Z boson in proton-proton collisions at* $\sqrt{s} = 8$ *TeV*, Phys. Rev. D **93** (2016) p. 052011, arXiv: 1511.09375 [hep-ex].

[41] ATLAS collaboration, *Search for dark matter in events with a hadronically decaying vector boson and missing transverse momentum in pp collisions at* $\sqrt{s}$ *= TeV with the ATLAS detector*, (2018), URL: http://cds.cern.ch/record/2311941.

[42] A. 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** (2015) p. 131801, arXiv: 1506.01081 [hep-ex].

[43] 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*, Phys. Rev. D **96** (2017) p. 112004, arXiv: 1706.03948 [hep-ex].

[44] 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** (2016) p. 072007, arXiv: 1510.06218 [hep-ex].

[45] ATLAS collaboration, *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. B **765** (2017) p. 11, arXiv: 1609.04572 [hep-ex].

[46] 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** (2017) p. 181804, arXiv: 1707.01302 [hep-ex].

[47] 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*, JHEP **10** (2017) p. 180, arXiv: 1703.05236 [hep-ex].

[48] ATLAS collaboration, *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 [hep-ex].

[49] J. K. Behr et al., "Searches for Dark Matter with the ATLAS Detector at the LHC", tech. rep. ATL-COM-PHYS-2017-1031, CERN, 2017, URL: https://cds.cern.ch/record/2273840.

[50] E. Tolley et al., "Dark Matter Simplified Model Exclusions: Vector and Axial-Vector Mediator : Dark Matter Summary plot for ICHEP 2018", tech. rep. ATL-COM-PHYS-2018-842, CERN, 2018, URL: https://cds.cern.ch/record/2625076.

[51] Wikipedia, "Standard Model of elementary particles", https://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg, Accessed: 26.07.2018, 2018.

[52] M. Gell-Mann, *Symmetries of Baryons and Mesons*, Phys. Rev. **125** (3 1962) pp. 1067–1084, URL: https://link.aps.org/doi/10.1103/PhysRev.125.1067.

[53] L. H. Ryder, *Quantum Field Theory, Second edition*, Cambridge University Press, 1996.

[54] P. D. Group, *2018 Review of Particle Physics*, Phys. Rev. D **98** (2018) p. 030001.

[55] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (12 1963) pp. 531–533, URL: https://link.aps.org/doi/10.1103/PhysRevLett.10.531.

[56] M. Kobayashi and T. Maskawa,
CP-Violation in the Renormalizable Theory of Weak Interaction,
Progress of Theoretical Physics **49** (1973) pp. 652–657, eprint: /oup/backfile/content_public/journal/ptp/49/2/10.1143/ptp.49.652/2/49-2-652.pdf,
URL: http://dx.doi.org/10.1143/PTP.49.652.

[57] F. Halzen and A. D. Martin, *QUARKS AND LEPTONS: AN INTRODUCTORY COURSE IN MODERN PARTICLE PHYSICS*, 1984,
ISBN: 0471887412, 9780471887416.

[58] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*,
Addison-Wesley, 1995, ISBN: 9780201503975, 0201503972,
URL: http://www.slac.stanford.edu/~mpeskin/QFT.html.

[59] C. Quigg,
Gauge Theories of the Strong, Weak, and Electromagnetic Interactions,
Princeton University Press, 2013, ISBN: 9780691135489, 9781400848225,
URL: http://chrisquigg.com/gauge-theories/.

[60] A. S. Schael et al., DELPHI, L3, OPAL, and L. E. Collaborations,
Electroweak Measurements in Electron-Positron Collisions at W-Boson-Pair Energies at LEP, Phys. Rep. **532** (2013) p. 119, arXiv: 1302.3415 [hep-ex].

[61] e. a. T. Aaltonen, *Combination of CDF and D0 W-Boson Mass Measurements*,
Phys.Rev. D **88** (2018) p. 052018, arXiv: 1307.7627 [hep-ex].

[62] ATLAS collaboration, *Measurement of the W-boson mass in pp collisions at* $\sqrt{7}$ TeV *with the ATLAS detector*, Eur. Phys. J. C **78** (2018) p. 110,
arXiv: 1701.07240 [hep-ex].

[63] W. K. F. J. V. C.Rubin, *Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions*, Astrophys. J. **159** (1970) pp. 379–403.

[64] G. H. et al., *Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Data Processing, Sky Maps, and Basic Results*,
Astrophys. J. Suppl. **180** (2009) pp. 225–245, arXiv: 0803.0732 [astro-ph].

[65] D. C. et al., *A direct empirical proof of the existence of dark matter*,
Astrophys. J. **648** (2006) pp. 109–113, arXiv: astro-ph/0608407 [astro-ph].

[66] L. Bian, *Renormalization group equation, the naturalness problem, and the understanding of the Higgs mass term*, Phys. Rev. D **88** (2013) p. 056022,
arXiv: 1308.2783 [hep-ph].

[67] J. L. Feng,
Dark Matter Candidates from Particle Physics and Methods of Detection,
Ann. Rev. Astron. Astrophys. **48** (2010) pp. 495–545,
arXiv: 1003.0904 [astro-ph.CO].

[68] K. G. Begeman, A. H. Broeils, and R. H. Sanders,
Extended rotation curves of spiral galaxies: dark haloes and modified dynamics,
Mon. Not. R. Astr. Soc. **249** (1991) pp. 523–537.

[69] R. Massey, T. Kitching, and J. Richard, *The dark matter of gravitational lensing*,
Rept. Prog. Phys. **73** (2010) p. 086901, arXiv: 1001.1739 [astro-ph.CO].

[70] G. M.Bernstein, J. A. Tyson, and C. S. Kochanek,
A large arc in the gravitational lens system 0957 + 561,
Astronomical Journal **105** (1993) pp. 816–830.

[71] A. Refregier, *Weak gravitational lensing by large scale structure*,
Ann. Rev. Astron. Astrophys. **41** (2003) pp. 645–668,
arXiv: astro-ph/0307212 [astro-ph].

[72] D. G. V. Rubakov,
Introduction to the Theory of the Early Universe. Hot Big Bang Theory,
World Scientific, 2011 p. 488.

[73] A. Boevia et al., *Recommendations on presenting LHC searches for missing transverse energy signals using simplified s−channel models of dark matter*,
(2016), arXiv: 1603.04156 [hep-ph].

[74] J .Abdallah et al., *Simplified Models for Dark Matter Searches at the LHC*,
Phys. Dark Univ. **9-10** (2015) pp. 8–23, arXiv: 1506.03116 [hep-ph].

[75] G. C. Branco et al., *Theory and phenomenology of two-Higgs-doublet models*,
Phys. Rept. **516** (2012) pp. 1–102, arXiv: 1106.0034 [hep-ph].

[76] M. Bauer, U. Haisch and F. Kahlhoefer,
Simplified dark matter models with two Higgs doublets: I. Pseudoscalar mediators,
JHEP **05** (2017) p. 138, arXiv: 1701.07427 [hep-ph].

[77] L. Henkelmann, "Search for Dark Matter using the $h(b\bar{b}) + E_T^{\text{miss}}$ Signature of the New "Two-Higgs-Doublet Model with Pseudoscalar Mediator" Benchmark Model with the ATLAS Detector", Masterarbeit: Universität Heidelberg, 2018.

[78] B.J. Carr, K. Kohri, Y. Sendouda and J. Yokoyama,
New cosmological constraints on primordial black holes, Phys. Rev. D **81** (2010),
arXiv: 0912.5297 [astro-ph].

[79] J. E. Gunn and I. J. R. Gott, *On the Infall of Matter into Clusters of Galaxies and Some Effects on Their Evolution*, Astrophys. J. **176** (1972) pp. 1–19.

[80] J. F. Navarro, C. S. Frenk, and S. D. M. White,
A Universal density profile from hierarchical clustering,
Astrophys. J. **490** (1997) pp. 493–508, arXiv: astro-ph/9611107 [astro-ph].

[81] J. Einasto, *Kinematics and Dynamics of Stellar Systems*,
Trudy Inst. Astrofiz. Alma-Ata. **5** (1965) 87100.

[82] A. W. Graham, D. Merritt, B. Moore, J. Diemand, and B. Terzic,
Empirical models for Dark Matter Halos. I. Nonparametric Construction of Density Profiles and Comparison with Parametric Models,
Astron. J. **132** (2006) pp. 2685–2700, arXiv: astro-ph/0509417 [astro-ph].

[83] A. Einstein,
Lens-Like Action of a Star by the Deviation of Light in the Gravitational Field,
Science **84** (1936) pp. 506–507.

[84] M. Milgrom, *Testing the MOND Paradigm of Modified Dynamics with Galaxy-Galaxy Gravitational Lensing*, Phys. Rev. Lett. **111** (2013) p. 041105,
arXiv: 1305.3516 [astro-ph.CO].

[85] N. Yoshida, V. Springel, S. D. M. White, and G. Tormen,
Weakly Self-interacting Dark Matter and the Structure of Dark Halos,
The Astrophysical Journal Letters **544** (2000) p. L87,
URL: http://stacks.iop.org/1538-4357/544/i=2/a=L87.

[86] L. D. Landau and E. M. Lifshitz, *Quantum mechanics. Non-relativistic theory*,
2nd ed., vol. 3, Pergamon Press, 1965 p. 596.

[87] J. R. Bond and A. S. Szalay,
The collisionless damping of density fluctuations in an expanding universe,
Astrophysical Journal **274** (1983) pp. 443–468.

[88] S. D. M. White, C. S. Frenk, and M. Davis,
Clustering in a neutrino-dominated universe,
Astrophysical Journal **274** (1983) pp. L1–L5.

[89] M. S. Turner and F. Wilczek,
Positron line radiation as a signature of particle dark matter in the halo,
Phys. Rev. D **42** (4 1990) pp. 1001–1007,
URL: https://link.aps.org/doi/10.1103/PhysRevD.42.1001.

[90] L. Husdal, *On Effective Degrees of Freedom in the Early Universe*,
Galaxies **4** (2016) p. 78, arXiv: 1609.04979 [astro-ph.CO].

[91] ADMX Collaboration, *A Search for Invisible Axion Dark Matter with the Axion Dark Matter Experiment*, Phys. Rev. Lett. **120** (2018),
arXiv: 1804.05750 [astro-ph].

[92] R. Daido, F.Takahashi, and W. Yin,
The ALP miracle: unified inflaton and dark matter, JCAP **1705** (2017) p. 044,
arXiv: 1702.03284 [hep-ph].

[93] The Daya Bay Collaboration, *Search for a Light Sterile Neutrino at Daya Bay*,
Phys. Rev. Lett. **113** (2014), arXiv: 1407.7259 [astro-ph].

[94] K. K. T. Flacke D. W.Kang, G. Mohlabeng, and S. C. Park,
Electroweak Kaluza-Klein Dark Matter, JHEP **1704** (2017) p. 041,
arXiv: 1702.02949v2 [astro-ph].

[95] G. Jungman, M. Kamionkowski, and K. Griest, *Supersymmetric dark matter*,
Phys. Rept. **267** (1996) pp. 195–373, arXiv: hep-ph/9506380 [hep-ph].

[96] A. P. Colijn,
"Particle Cosmology: DM Direct Detection, Topical Lectures "Dark Matter"",
2015, URL: https://indico.nikhef.nl/event/869/material/0/1.pdf.

[97] M. C. Smith et al.,
The RAVE Survey: Constraining the Local Galactic Escape Speed,
Mon. Not. Roy. Astron. Soc. **379** (2007) pp. 755–772,
arXiv: astro-ph/0611671 [astro-ph].

[98] R. Bernabei et al., *The DAMA/LIBRA apparatus*,
Nucl. Instrum. Meth. **A592** (2008) pp. 297 –315, arXiv: 0804.2738 [astro-ph].

[99] G. Angloher et al.,
Results on low mass WIMPs using an upgraded CRESST-II detector,
Eur. Phys. J. **C74** (2014) p. 3184, arXiv: 1407.3146 [astro-ph.CO].

[100] R. Agnese et al., *New Results from the Search for Low-Mass Weakly Interacting Massive Particles with the CDMS Low Ionization Threshold Experiment*, Phys. Rev. Lett. **116** (2016) p. 071301, arXiv: 1509.02448 [astro-ph.CO].

[101] D. S. Akerib et al., *The Large Underground Xenon (LUX) Experiment*, Nucl. Instrum. Meth. **A704** (2013) pp. 111–126, arXiv: 1211.3788 [physics.ins-det].

[102] E. Aprile et al., *The XENON100 Dark Matter Experiment*, Astropart. Phys. **35** (2012) pp. 573–590, arXiv: arXiv:1107.2155 [physics.ins-det].

[103] E. Aprile et al., *Dark Matter Search Results from a One Tonne× Year Exposure of XENON1T*, (2018), arXiv: 1805.12562 [astro-ph.CO].

[104] P. Cushman et al., "Working Group Report: WIMP Dark Matter Direct Detection", *Proceedings, 2013 Community Summer Study on the Future of U.S. Particle Physics: Snowmass on the Mississippi (CSS2013): Minneapolis, MN, USA, July 29-August 6, 2013*, 2013, arXiv: 1310.8327 [hep-ex], URL: https://inspirehep.net/record/1262767/files/arXiv:1310.8327.pdf.

[105] M. Ackermann et al., *Constraining Dark Matter Models from a Combined Analysis of Milky Way Satellites with the Fermi Large Area Telescope*, Phys. Rev. Lett. **107** (2011) p. 241302, arXiv: 1108.3546 [astro-ph.HE].

[106] H. Abdallah et al., *Search for dark matter annihilations towards the inner Galactic halo from 10 years of observations with H.E.S.S*, Phys. Rev. Lett. **117** (2016) p. 111301, arXiv: 1607.08142 [astro-ph.HE].

[107] S. Adrian-Martinez et al., *Limits on Dark Matter Annihilation in the Sun using the ANTARES Neutrino Telescope*, Phys. Lett. **B759** (2016) pp. 69–74, arXiv: 1603.02228 [astro-ph.HE].

[108] F. Halzen and R. S. Klein, *IceCube: An Instrument for Neutrino Astronomy*, Rev. Sci. Instrum. **81** (2010) p. 081101, arXiv: 1007.1247 [astro-ph.HE].

[109] 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 0.5350 GeV*, Phys. Rev. Lett. **110** (2013) p. 141102.

[110] M. Cirelli, *Status of Indirect (and Direct) Dark Matter searches*, (2015), arXiv: 1511.02031 [astro-ph.HE].

[111] G. D'Ambrosio, G. F. Giudice, G. Isidori, and A. Strumia, *Minimal flavor violation: An Effective field theory approach*, Nucl. Phys. **B645** (2002) pp. 155–187, arXiv: hep-ph/0207036 [hep-ph].

[112] A. 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 [hep-ex].

[113] A. Berlin, T. Lin, and L.-T. Wang, *Mono-Higgs Detection of Dark Matter at the LHC*, JHEP **06** (2014) p. 078, arXiv: 1402.7074 [hep-ph].

[114] N. Craig, J. Galloway, and S. Thomas,
Searching for Signs of the Second Higgs Doublet, (2013),
arXiv: 1305.2424 [hep-ph].

[115] J. M. No, *Looking through the pseudoscalar portal into dark matter: Novel mono-Higgs and mono-Z signatures at the LHC*,
Phys. Rev. **D93** (2016) p. 031701, arXiv: 1509.01110 [hep-ph].

[116] F. Marcastel,
CERN's Accelerator Complex. La chaîne des accélérateurs du CERN,
(2013), General Photo, URL: https://cds.cern.ch/record/1621583.

[117] W. W. Herr and B. Muratori, *Concept of luminosity*, (2006),
URL: https://cds.cern.ch/record/941318.

[118] M. Bruschi,
"Diffraction and Forward Physics in ATLAS: results and perspectives",
tech. rep. ATL-PHYS-PROC-2014-278, CERN, 2014,
URL: https://cds.cern.ch/record/1975956.

[119] ATLAS collaboration, "TWiki", https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2, Accessed: 24.07.2018.

[120] J. Pequenao, "Computer generated image of the whole ATLAS detector", 2008,
URL: https://cds.cern.ch/record/1095924.

[121] J. Pequenao, "Computer generated image of the ATLAS inner detector", 2008,
URL: https://cds.cern.ch/record/1095926.

[122] ATLAS collaboration,
The ATLAS Inner Detector commissioning and calibration,
Eur. Phys. J. **C70** (2010) pp. 787–821, arXiv: 1004.5293 [physics.ins-det].

[123] J. Pequenao, "Computer Generated image of the ATLAS calorimeter", 2008,
URL: https://cds.cern.ch/record/1095927.

[124] ATLAS collaboration,
The ATLAS Experiment at the CERN Large Hadron Collider,
JINST **3** (2008) S08003.

[125] ATLAS collaboration, *Energy linearity and resolution of the ATLAS electromagnetic barrel calorimeter in an electron test-beam*,
Nucl. Instrum. Meth. A **568** (2006) pp. 601–623, arXiv: 0608012 [physics].

[126] T. Barillari, *The ATLAS liquid argon hadronic end-cap calorimeter: Construction and selected beam test results*,
Nucl. Phys. Proc. Suppl. **150** (2006) pp. 102–105,
arXiv: physics/0407026 [physics.ins-det].

[127] C. Grupen and B. Shwartz, *Particle Detectors*, 2nd ed.,
Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology,
Cambridge University Press, 2008.

[128] A. collaboration,
ATLAS Tile Calorimeter performance for single particles in beam tests,
Journal of Physics: Conference Series **160** (2009) p. 012057,
URL: http://stacks.iop.org/1742-6596/160/i=1/a=012057.

[129] ATLAS collaboration,
Performance of the ATLAS hadronic end-cap calorimeter in beam tests,
Nucl. Instrum. Meth. **A482** (2002) pp. 94–124.

[130] ATLAS collaboration,
Energy calibration of the ATLAS liquid argon forward calorimeter,
JINST **3** (2008) P02002.

[131] J. M. Campbell, J. W. Huston, and W. J. Stirling,
Hard Interactions of Quarks and Gluons: A Primer for LHC Physics,
Rept. Prog. Phys. **70** (2007) p. 89, arXiv: hep-ph/0611148 [hep-ph].

[132] ATLAS collaboration, *Performance of the ATLAS Trigger System in 2015*,
Eur. Phys. J. **C77** (2017) p. 317, arXiv: 1611.09661 [hep-ex].

[133] B. Asman et al., *The ATLAS Level-1 Calorimeter Trigger: PreProcessor implementation and performance*, JINST **7** (2012) P12008.

[134] J. Jongmanns, "The Upgrade of the PreProcessor of the ATLAS Level-1 Calorimeter Trigger for LHC Run-2", PhD Thesis: Universität Heidelberg, 2017.

[135] ATLAS collaboration, "Data Quality Flag Interpretation",
https://twiki.cern.ch/twiki/bin/view/Atlas/DQFlagInterpretation.

[136] ATLAS collaboration, "Good Run Lists For Analysis Run 2", https://twiki.cern.ch/twiki/bin/view/AtlasProtected/GoodRunListsForAnalysisRun2.

[137] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand,
Parton Fragmentation and String Dynamics, Phys. Rept. **97** (1983) pp. 31–145.

[138] J.-C. Winter, F. Krauss, and G. Soff, *A Modified cluster hadronization model*,
Eur. Phys. J. **C36** (2004) pp. 381–395, arXiv: hep-ph/0311085 [hep-ph].

[139] S. Agostinelli et al., *GEANT4: A Simulation toolkit*,
Nucl. Instrum. Meth. **A506** (2003) pp. 250–303.

[140] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) p. 007,
arXiv: 0811.4622 [hep-ph].

[141] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber,
QCD matrix elements + parton showers, JHEP **11** (2001) p. 063,
arXiv: hep-ph/0109231 [hep-ph].

[142] H.-L. Lai et al., *New parton distributions for collider physics*,
Phys. Rev. D **82** (7 2010) p. 074024,
URL: https://link.aps.org/doi/10.1103/PhysRevD.82.074024.

[143] R. D. Ball et al., *Parton distributions for the LHC Run II*,
JHEP **04** (2015) p. 040, arXiv: 1410.8849 [hep-ph].

[144] B. Andersson, S. Mohanty, and F. Soderberg,
"Recent developments in the Lund model", *36th Annual Winter School on Nuclear and Particle Physics (PINP 2002) and 8th St. Petersburg School on Theoretical Physics St. Petersburg, Russia, February 25-March 3, 2002*, 2002,
arXiv: hep-ph/0212122 [hep-ph].

[145] ATLAS collaboration, "ATLAS Run 1 Pythia8 tunes",
tech. rep. ATL-PHYS-PUB-2014-021, CERN, 2014,
URL: https://cds.cern.ch/record/1966419.

[146] S. Carrazza, S. Forte, and J. Rojo, "Parton Distributions and Event Generators", *Proceedings, 43rd International Symposium on Multiparticle Dynamics (ISMD 13)*, 2013 pp. 89–96, arXiv: 1311.5887 [hep-ph].

[147] J. Alwall et al., *MadGraph 5 : Going Beyond*, JHEP **06** (2011) p. 128, arXiv: 1106.0522 [hep-ph].

[148] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) p. 079, arXiv: 1405.0301 [hep-ph].

[149] ATLAS collaboration, "TWiki", https://twiki.cern.ch/twiki/bin/view/Atlas/LowestUnprescaled, Accessed: 12.08.2018.

[150] ATLAS collaboration, "TWiki", https://twiki.cern.ch/twiki/bin/view/Atlas/L34TriggerMenu2016, 2016.

[151] ATLAS collaboration, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*, (2016).

[152] W. Lampl et al., *Calorimeter clustering algorithms: Description and performance*, (2008).

[153] T. Cornelissen et al., *The global χ^2 track fitter in ATLAS*, J. Phys. Conf. Ser. **119** (2008) p. 032013.

[154] ATLAS collaboration, *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*, (2012).

[155] ATLAS collaboration, *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, Eur. Phys. J. **C74** (2014) p. 3071, arXiv: 1407.5063 [hep-ex].

[156] ATLAS collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. **C77** (2017) p. 490, arXiv: 1603.02934 [hep-ex].

[157] ATLAS collaboration, "Combined Performance Tools for Physics Analysis", https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/PhysicsAnalysisWorkBookRel20CPRec, Accessed: 02.08.2018.

[158] ATLAS collaboration, *Muon reconstruction performance of the ATLAS detector in protonproton collision data at $\sqrt{s}$ =13 TeV*, Eur. Phys. J. **C76** (2016) p. 292, arXiv: 1603.05598 [hep-ex].

[159] ATLAS collaboration, "ATLAS Official Isolation Working Points", https://twiki.cern.ch/twiki/bin/view/AtlasProtected/IsolationSelectionTool, Accessed: 02.08.2018.

[160] ATLAS collaboration, *Identification of boosted, hadronically decaying W bosons and comparisons with ATLAS data taken at $\sqrt{s} = 8$ TeV*, Eur. Phys. J. **C76** (2016) p. 154, arXiv: 1510.05821 [hep-ex].

[161] "Identification of boosted, hadronically-decaying W and Z bosons in $\sqrt{s} = 13$ TeV Monte Carlo Simulations for ATLAS", tech. rep. ATL-PHYS-PUB-2015-033, CERN, 2015, URL: https://cds.cern.ch/record/2041461.

[162] J. E. Huth et al., "Toward a standardization of jet definitions", *1990 DPF Summer Study on High-energy Physics: Research Directions for the Decade (Snowmass 90) Snowmass, Colorado, June 25-July 13, 1990*, 1990 pp. 0134–136, URL: http://lss.fnal.gov/cgi-bin/find_paper.pl?conf-90-249.

[163] G. P. Salam, *Towards Jetography*, Eur. Phys. J. **C67** (2010) pp. 637–686, arXiv: 0906.1833 [hep-ph].

[164] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet User Manual*, Eur. Phys. J. **C72** (2012) p. 1896, arXiv: 1111.6097 [hep-ph].

[165] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti-k(t) jet clustering algorithm*, JHEP **04** (2008) p. 063, arXiv: 0802.1189 [hep-ph].

[166] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, *Longitudinally invariant K_t clustering algorithms for hadron hadron collisions*, Nucl. Phys. **B406** (1993) pp. 187–224.

[167] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*, Phys. Rev. D **48** (7 1993) pp. 3160–3166, URL: https://link.aps.org/doi/10.1103/PhysRevD.48.3160.

[168] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, *Better jet clustering algorithms*, Journal of High Energy Physics **1997** (1997) p. 001, URL: http://stacks.iop.org/1126-6708/1997/i=08/a=001.

[169] D. Krohn, J. Thaler, and L. T. Wang, *Jet Trimming*, JHEP **02** (2010) p. 084, arXiv: 0912.1342 [hep-ph].

[170] 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*, (2015).

[171] 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. **D96** (2017) p. 072002, arXiv: 1703.09665 [hep-ex].

[172] ATLAS collaboration, "Jet mass reconstruction with the ATLAS Detector in early Run 2 data", tech. rep. ATLAS-CONF-2016-035, CERN, 2016, URL: https://cds.cern.ch/record/2200211.

[173] ATLAS collaboration, "ATLAS Combined Mass resolution and W mass reconstruction performance", tech. rep. JETM-2017-002, CERN, 2017, URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/JETM-2017-002/.

[174] M. Czakon, P. Fiedler, and A. Mitov, *Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\frac{4}{S})$*, Phys. Rev. Lett. **110** (2013) p. 252004, arXiv: 1303.6254 [hep-ph].

[175] "Flavor Tagging with Track Jets in Boosted Topologies with the ATLAS Detector", tech. rep. ATL-PHYS-PUB-2014-013, CERN, 2014, URL: https://cds.cern.ch/record/1750681.

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

[177] M. Cacciari and G. P. Salam, *Pileup subtraction using jet areas*, Phys. Lett. **B659** (2008) pp. 119–126, arXiv: 0707.1378 [hep-ph].

[178] M. Cacciari, G. P. Salam, and G. Soyez, *The Catchment Area of Jets*, JHEP **04** (2008) p. 005, arXiv: 0802.1188 [hep-ph].

[179] ATLAS collaboration, *Performance of b-Jet Identification in the ATLAS Experiment*, JINST **11** (2016) P04008, arXiv: 1512.01094 [hep-ex].

[180] ATLAS collaboration, "Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run", tech. rep. ATL-PHYS-PUB-2016-012, CERN, 2016, URL: https://cds.cern.ch/record/2160731.

[181] A. Hocker, *TMVA - Toolkit for Multivariate Data Analysis*, (2007), arXiv: physics/0703039 [physics.data-an].

[182] G. Piacquadio and C. Weiser, *A new inclusive secondary vertex algorithm for b−jet tagging in ATLAS*, Journal of Physics: Conference Series **119** (2008) p. 032032, URL: http://stacks.iop.org/1742-6596/119/i=3/a=032032.

[183] ATLAS collaboration, *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, (2009), arXiv: 0901.0512 [hep-ex].

[184] F. A. D. Bello, "Optimisation of the ATLAS $b-$tagging algorithms for the 2017-2018 LHC data-taking", tech. rep. ATL-PHYS-PROC-2017-179, CERN, 2017, URL: https://cds.cern.ch/record/2286993.

[185] D. Adams et al., "Recommendations of the Physics Objects and Analysis Harmonisation Study Groups 2014", tech. rep. ATL-COM-PHYS-2014-451, CERN, 2014, URL: https://cds.cern.ch/record/1700874.

[186] ATLAS collaboration, *Performance of algorithms that reconstruct missing transverse momentum in $\sqrt{s} = 8$ TeV proton-proton collisions in the ATLAS detector*, Eur. Phys. J. **C77** (2017) p. 241, arXiv: 1609.09324 [hep-ex].

[187] ATLAS collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13$ TeV*, (2018), arXiv: 1802.08168 [hep-ex].

[188] ATLAS collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, Eur. Phys. J. **C76** (2016) p. 581, arXiv: 1510.03823 [hep-ex].

[189] ATLAS collaboration, "Twiki", https://twiki.cern.ch/twiki/bin/view/AtlasProtected/PhysicsAnalysisWorkBookRel20CPRec#Event_cleaning, Accessed: 16.08.2018.

[190] ATLAS collaboration, *Selection of jets produced in 13TeV proton-proton collisions with the ATLAS detector*, (2015).

[191] S. W. C. 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", tech. rep. ATL-COM-PHYS-2014-262, CERN, 2014, URL: https://cds.cern.ch/record/1692942.

[192] I. M. A. J. Larkoski and D. Neill, *Power Counting to Better Jet Observables*, JHEP **12** (2014) p. 009, arXiv: 1409.6298 [hep-ph].

[193] ATLAS collaboration, "Performance of Top Quark and W Boson Tagging in Run 2 with ATLAS", tech. rep. ATLAS-CONF-2017-064, CERN, 2017, URL: https://cds.cern.ch/record/2281054.

[194] A. Vaidya et al., "Optimisation procedure for two-variable top, W and Z substructure taggers for the ATLAS detector", tech. rep. ATL-COM-PHYS-2017-033, CERN, 2017, URL: https://cds.cern.ch/record/2242865.

[195] G. P. S. A.J. Larkoski and J. Thaler, *Energy Correlation Functions for Jet Substructure*, JHEP **06** (2013) p. 108, arXiv: 1305.0007 [hep-ph].

[196] ATLAS collaboration, *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 [hep-ex].

[197] ATLAS collaboration, "Search for the bb decay of the Standard Model Higgs boson in associated W/ZH production with the ATLAS detector", tech. rep. ATLAS-CONF-2013-079, CERN, 2013, URL: https://cds.cern.ch/record/1563235.

[198] D. I. N. Villar, "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", PhD thesis: Universität Heidelberg, 2018.

[199] W. J. S. C.-H. Kom, *Charge asymmetry in W + jets production at the LHC*, Eur. Phys. J. **C69** (2010) pp. 67–73, arXiv: 1004.3404 [hep-ph].

[200] ATLAS collaboration, "Uncertainty release for analyses using 2015 and 2016 data in 20.7 (MC15c) releases", https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JetUncertainties20152016Data20p7, 2016.

[201] E. S. P. J. Neyman, *IX. On the problem of the most efficient tests of statistical hypotheses*, Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **231** (1933) pp. 289–337, ISSN: 0264-3952, URL: http://rsta.royalsocietypublishing.org/content/231/694-706/289.

[202] G. Cowan, *Statistical Data Analysis*, Clarendon Press, 1998.

[203] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. **C71** (2011) p. 1554, [Erratum: Eur. Phys. J.C73,2501(2013)], arXiv: 1007.1727 [physics.data-an].

[204] S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*, Annals Math. Statist. **9** (1938) pp. 60–62.

[205] A. Wald, *Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large*, Transactions of the American Mathematical Society **54** (Nov., 1943) pp. 426–482, URL: http://www.jstor.org/stable/1990256.

[206] ATLAS Statistics Forum, "The CL_s method: information for conference speakers", 2011, URL: http://www.jstor.org/stable/1990256.

[207] ATLAS collaboration, "Jet Resolution 2015 Prerecommemdation", https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JetResolution2015Prerecom, 2016.

[208] ATLAS collaboration, "JetEtMiss Recommendations for MC15 (Release 20.7)", https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JetEtmissRecommendations2016, 2016.

[209] ATLAS collaboration, "Flavour tagging recommendations in release 20.7", https://twiki.cern.ch/twiki/bin/view/AtlasProtected/BTagCalib2015, 2016.

[210] ATLAS collaboration, "Recommendations for 2015 or 2015+2016 data analysis aiming for ICHEP2016 or later", https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JetUncertainties2016PrerecLargeR, 2016.

[211] ATLAS collaboration, "Usage of Missing ET in analyses: rebuilding and systematics", https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/METUtilities, 2017.

[212] ATLAS collaboration, "Electron Efficiencies for Run2", https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/ElectronEfficiencyRun2, 2016.

[213] ATLAS collaboration, "Electron and Photon Calibration for Run2 (energy scale and resolution)", https://twiki.cern.ch/twiki/bin/view/AtlasProtected/EGammaCalibrationRun2, 2016.

[214] ATLAS collaboration, "Muon Combined Performance Analysis Guidelines", https://twiki.cern.ch/twiki/bin/view/AtlasProtected/MCPAnalysisGuidelinesMC15, 2016.

[215] N. Köhler and J. Junggeburth, "$Z \to \mu^+\mu^-$ reconstruction and TTVA efficiencies and scale factors", 2016, URL: https://indico.cern.ch/event/574876/contributions/2327421/attachments/1348842/2040671/koehler161005.pdf.

[216] A. Robson, G. Piacquadio, and E. 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", tech. rep. ATL-COM-PHYS-2016-1747, CERN, 2016, URL: https://cds.cern.ch/record/2235887.

[217] E. S. P. Thompson, "Theoretical studies for the Standard Model $Vh, h \to b\bar{b}$ and related searches", tech. rep. ATL-COM-PHYS-2016-475, CERN, 2016, URL: https://cds.cern.ch/record/2151835.

[218] S. Suchek, "Mono-V(had) analysis status", 2017, URL: https://indico.cern.ch/event/678514/contributions/2781934/attachments/1556586/2448157/monoV_closure_101117.pdf.

[219] J. Wang et al., "Search for Dark Matter in association with a hadronically decaying vector boson in pp collisions at $\sqrt{s} = 13$ TeV using 2015+2016 data collected with the ATLAS detector", tech. rep. ATL-COM-PHYS-2016-1499, CERN, 2016, URL: https://cds.cern.ch/record/2226524.

[220] S. Meehan et al., "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", tech. rep. ATL-COM-PHYS-2016-1484, CERN, 2016, URL: https://cds.cern.ch/record/2225941.

[221] ATLAS collaboration, "Recommendation for using PDFs", https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/PdfRecommendations, 2018.

[222] ATLAS collaboration, "Jet Dark Matter Signal Uncertainties", https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JDMSignalUncertainties, 2016.

[223] ATLAS collaboration, "Recommendations for MC tunes: systematic uncertainties and non-pQCD corrections", https://twiki.cern.ch/twiki/bin/view/AtlasProtected/MCTuningRecommendations, 2017.

[224] LHC Physics Centre at CERN, "LHC DM WG: WG on Dark Matter Searches at the LHC", https://lpcc.web.cern.ch/content/lhc-dm-wg-wg-dark-matter-searches-lhc.

[225] D. 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 [hep-ex].

[226] M. Duerr et al., *How to save the WIMP: global analysis of a dark matter model with two s-channel mediators*, JHEP **09** (2016) p. 042, arXiv: 1606.07609 [hep-ph].

[227] F. Kahlhoefer et al., *Implications of unitarity and gauge invariance for simplified dark matter models*, JHEP **1602** (2016) p. 016, arXiv: 1510.02110 [astro-ph].

[228] Y. Cheng, M. Shochet, and B. Penning, "Search for Dark Matter produced in association with a Higgs boson decaying to two bottom quarks in proton-proton collisions at $\sqrt{s} = 8$TeV with the ATLAS detector", tech. rep. ATL-COM-PHYS-2014-1577, CERN, 2014, URL: https://cds.cern.ch/record/1977384.

[229] M. Backovi, A. Martini, O. Mattelaer, K. Kong, and G. Mohlabeng, *Direct Detection of Dark Matter with MadDM v.2.0*, Phys. Dark Univ. **9-10** (2015) pp. 37–50, arXiv: 1505.04190 [hep-ph].

[230] K. N. T. Jacques, *Mapping monojet constraints onto Simplified Dark Matter Models*, JHEP **06** (2015) p. 142, arXiv: 1502.05721 [hep-ex].

[231] D. Guest, "Interpolating Exclusion on the Mono $H \to b\bar{b}$ Grid", 2017, URL: https://indico.cern.ch/event/538483/contributions/2548655/attachments/1441663/2219793/dguest-interpolation.pdf.

[232] ATLAS collaboration, *Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector*, (2018), arXiv: 1808.08238 [hep-ex].

[233] CMS collaboration, *Observation of Higgs boson decay to bottom quarks*, Submitted to: Phys. Rev. Lett. (2018), arXiv: 1808.08242 [hep-ex].

[234] DELPHES 3 Collaboration, *DELPHES 3: a modular framework for fast simulation of a generic collider experiment*, JHEP **02** (2014) p. 057, arXiv: 1307.6346 [hep-ex].

[235] ATLAS collaboration, *Search for dark matter produced in association with a hadronically decaying vector boson in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Lett. **B763** (2016) pp. 251–268, arXiv: 1608.02372 [hep-ex].

[236] ATLAS collaboration, *Search for dark matter at $\sqrt{s} = 13$ TeV in final states containing an energetic photon and large missing transverse momentum with the ATLAS detector*, Eur. Phys. J. **C77** (2017) p. 393, arXiv: 1704.03848 [hep-ex].

[237] ATLAS collaboration, *Search for an invisibly decaying Higgs boson or dark matter candidates produced in association with a Z boson in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Lett. **B776** (2018) pp. 318–337, arXiv: 1708.09624 [hep-ex].

[238] ATLAS collaboration, "Exotics Physics Public Results", 2017, URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/EXOTICS/ATLAS_DarkMatter_Summary_Vector/ATLAS_DarkMatter_Summary_Vector_201707.pdf.

[239] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) pp. 852–867, arXiv: 0710.3820 [hep-ph].

[240] M. Backovi et al., *Higher-order QCD predictions for dark matter production at the LHC in simplified models with s-channel mediators*, Eur. Phys. J. **C75** (2015) p. 482, arXiv: 1508.05327 [hep-ph].

[241] ATLAS collaboration, *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** (2017) p. 052004, arXiv: 1703.09127 [hep-ex].

Acknowledgements

During 4.5 years I spent in Heidelberg, I met and worked with a lot of remarkable people. Without any doubt, all my successful work and this thesis would not be possible without their great help and support.
I would like to express my sincere appreciation to Priv.-Doz. Dr. Oleg Brandt, who gave me an opportunity to work on great topics of particle physics in KIP ATLAS group. Huge thanks to you for supervising me all these years, for your support and great patience. I am thankful for our numerous fruitful discussions, for countless number of stupid question you answered, for advancing my interests and pushing my results, for all your time and efforts you put in me. Thank you for pushing me all the time in a right direction.
I am extremely grateful to Prof. Hans-Christian Schultz-Coulon for his remarkable efforts in making the atmosphere in KIP ATLAS group highly productive and super friendly. His care about the comfort and professional development of all group members is just invaluable.
I am very thankful to Dr. Monica Dunford, Dr. Martin Wessels, Dr. Rainer Stamen and Dr. Pavel Starovoitov for sharing their knowledge and firm support. You all made a valuable contribution to my professional growth.
I also want to give special thanks to Prof. Tilman Plehn, Prof. Matthias Bartelmann, who kindly agreed to join my Ph.D. defense committee, and especially to Prof. Dr. André Schöning, who also agreed to review this thesis.
During my doctoral times, I had a privilege of working with great people in KIP ATLAS group. Many of them I am happy to call friends now. First, I want to thank my former and current office mates Hanno for our endless valuable conversation, Claire, who kindly tolerated maybe the coldest office she ever worked in, Sascha, Nadia, Valerie and Basti. Many thanks for your patience and great time we had together. I was happy to work together with Merve, Fabrizio, Daniel, Sascha, Philipp and Lars under the supervision of Oleg Brandt. Thank you for all discussions we had during our local group meetings, for all refreshing chats during countless coffee breaks and for a great time we spent together off hours. Special thanks to my collegue and friend Fabrizio and his wife Martina for their kind hospitality and for an enjoyable time together at CERN and of course in Heidelberg. My appreciation to Sebastian, Thomas and Veit, who were in charge of our local computer cluster and helped me to get all those results, which are successfully published and presented in this thesis. And beyond doubt I want to thank Jan, Falk, Sebastian, Manuel, Thomas, Alessandra, Julia and Victor. You all made me feel like home in Heidelberg. Additionally, I appreciate great efforts of Perta Pfeifer and HGSFP office, in particular to Gesine Heinzelmann and Elisabeth Miller, in making bureaucracy and tonnes of paperwork look much easier.
I had a fortune to meet a lot of awesome people at CERN. Many thanks to Miha and Joze for a great time we spent together at CERN summer school, and of course afterwards at CERN and various conferences. I want to thank Gabriel Facini for being a convener

of Exotics Group all these years and for your valuable help in pushing and publishing $E_{\mathrm{T}}^{\mathrm{miss}}+V(qq)$ and $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ analyses. Huge thanks to all peoples at CERN, who I had a fortune to meet and have fun with. You all made my time at CERN really amazing.
And last but not the least, I want to thank my dear family, to Liudmila, Stanislav and Sergiy, and all my Ukrainian friends for your warm support and great patience in challenging times, even though we live far away from each other now. I appreciate those rare opportunities when I am back home to a place where I can always relax my body and soul.