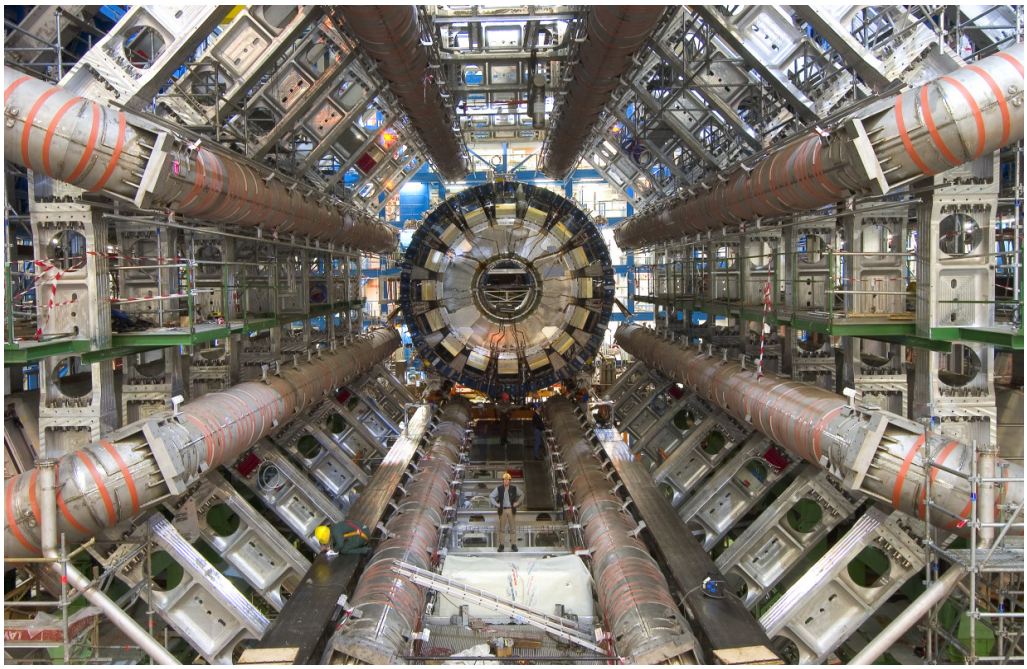
Università degli Studi di Pavia
Dottorato di Ricerca in Fisica – XXXIII Ciclo

Searches for Dark Matter production in events with top quarks in the final state with the ATLAS detector at the LHC

Giulia Rovelli



Tesi per il conseguimento del titolo

Università degli Studi di Pavia
Dottorato di Ricerca in Fisica - XXXIII Ciclo

Searches for Dark Matter production in events with top quarks in the final state with the ATLAS detector at the LHC

Giulia Rovelli

Submitted to the Graduate School of Physics in partial fulfillment
of the requirements for the degree of
Dottore di Ricerca in Fisica
Doctor of Philosophy in Physics
at the
University of Pavia

Supervisor: Dott. Giacomo Polesello
Internal Supervisor: Prof. Daniela Rebuzzi

Cover: the ATLAS detector under construction in 2004. The eight barrel toroid magnets can be clearly seen. Note the figure at bottom center for size comparison.

Searches for Dark Matter production in events with top quarks in the final state with the ATLAS detector at the LHC

Giulia Rovelli

PhD thesis - University of Pavia
Pavia, Italy, April 2021

Contents

Introduction

The Standard Model (SM) of high-energy physics provides a remarkably successful description of presently known subatomic particles and their interactions (strong, weak and electromagnetic). Its predictions have been broadly tested, and the 2012 Higgs boson observation at the Large Hadron Collider (LHC) has completed the picture. Nevertheless, the SM is an incomplete theory, as it is unable to provide an explanation to several phenomena, like gravity, the matter-antimatter asymmetry, neutrino masses and the mass hierarchy and mixing origins. A pragmatic approach to expand the SM is to consider the strongest piece of evidence of physics beyond the SM: the existence of Dark Matter (DM), assessed from cosmological studies, which accounts for around 26% of the Universe.

There is no scientific consensus on the nature of DM. A possibility which has been explored at great length in literature is that is made of new particles non-present in the SM. A general class of very promising particle DM candidates are the so-called Weakly Interacting Massive Particles (WIMPs), particles that do not interact with the SM particles with the known forces but only through a new interaction, of the same magnitude of the weak one. There is a huge number of models containing dark matter candidates in the form of a WIMP, and this work will focus in particular on the so-called Two Higgs Doublets Model + Pseudoscalar Mediator (2HDMa), in which a pseudoscalar particle acting as a "mediator" between SM and DM is introduced together with a second Higgs doublet. After the breaking of the electroweak symmetry, the mediator and the Higgs doublets give rise to six new bosons, either scalar, pseudoscalar or charged. Thus the model presents many experimental signatures, usually involving missing energy from undetected DM along with SM particles. These signatures can be looked for also at particle colliders, like the LHC, since the high centre-of-mass energy of the proton-proton collision allows for their production.

This thesis addresses two different searches for DM production in the context of the 2HDMa model, performed using proton-proton collisions data recorded at 13 TeV by the ATLAS experiment at the Large Hadron Collider (LHC) between 2015 and 2018.

The first part of the work consisted in the study of the production of DM in association with a pair of top quarks, trying to assess the sensitivity in the

parameters space of 2HDMa of already existing ATLAS analyses. These analyses were performed in the framework of a different model, similar to the 2HDMa model since it included a pseudoscalar mediator. The production of DM in association with a pair of top quarks was chosen because it is directly sensitive to the nature of the mediator through the polarization of the two top quarks, which can be reconstructed from their decays products. The model employed in the existing analyses and the 2HDMa model were thus compared in detail, and after that a recasting strategy to translate the existing results in the parameter space of 2HDMa was developed and validated. The results of this work, for which the author of this thesis was the main contributor, were included in the ATLAS DM summary paper, published in 2019 [1].

The second part of the present study focused instead on a new search channel for the 2HDMa model, never explored before, including the production of dark matter associated with a single top quark. This signature was identified as the only one sensitive to the production of charged Higgs bosons, and thus holds a key role in the ATLAS research program dedicated to the 2HDMa model. Single top production can be studied in different final states, explored in a publication by the ATLAS Collaboration [2]. Being a new and challenging signature, dedicated strategies were developed to maximise the sensitivity of the analyses focusing on the different final states. In particular, the author of this thesis was the main analyser of the two-lepton final state, in which the single top is produced in association with a W-boson and they both decay leptonically, developing its entire analysis strategy and handling the interpretation of the results.

This thesis is organised as follows. Chapter 1 provides an introduction to the dark matter, presenting the evidence of it and analysing possible particle candidates, focusing then on WIMPs. Chapter 2 presents the 2HDMa model in details, together with an introduction on the DM models used before as a benchmark for LHC searches, included the searches later reinterpreted in this work in the context of the 2HDMa model. In Chapter 3 an overview of the LHC and the ATLAS detector is given. Chapter 4 describes how both Standard Model processes and DM production are simulated, which is a key step to compare theoretical prediction with the data collected by ATLAS, while Chapter 5 presents the techniques used for reconstructing the different physics objects starting from the signals recorded in the detector after proton-proton collisions inside the LHC. Finally, Chapter 6 introduces the analysis techniques employed in the 2HDMa searches presented in the subsequent Chapters. Chapter 7 describes the search for DM in association with a pair of top quarks, obtained recasting the existing ATLAS analyses, while Chapter 8 presents the novel search for dark matter produced in association with a single top quark.

Chapter 1

Dark Matter

The existence of Dark Matter (DM), i.e. matter which is non-luminous and that interacts only gravitationally with ordinary matter, is an accepted fact by the scientific community. The investigation about its nature is one of the main challenges of contemporary physics, since about 26% of the Universe consists of DM, while ordinary matter only composes 5% of it. In the following, the evidence for DM will be discussed in details, focusing then on possible particle explanations.

1.1 Evidence for Dark Matter

The presence of DM in our Universe has been confirmed by many independent observations on different astrophysical scales, even if its nature remains unknown.

The first and oldest evidence of DM, dating back to 1933, comes from the study of the rotation curves of galaxies. The rotation curves represent the circular velocities of stars and gas inside a galaxy as a function of their distance from the galactic centre. They are obtained combining optical surface photometry with observations of the 21-cm hydrogen line. Since from Newton laws the rotation velocity is given by

$$v(r) = \sqrt{\frac{GM(r)}{r}}, \tag{1.1}$$

where $M(r) = 4\pi \int \rho(r) r^2 dr$, with $\rho(r)$ being the mass density function, at distances greater than the radius of the optical disk, these rotation curves should decrease as $\propto r^{-1/2}$. A characteristic flat behaviour at large distances is instead observed [3], as shown in Fig. 1.1, suggesting the presence of a halo of non-luminous matter with $M(r) \propto r$ and $\rho(r) = r^{-2}$.

Another strong argument for dark matter comes from the observations of the galaxy cluster 1E0657-56 [4], better know as Bullet Cluster. Fig. 1.2 shows

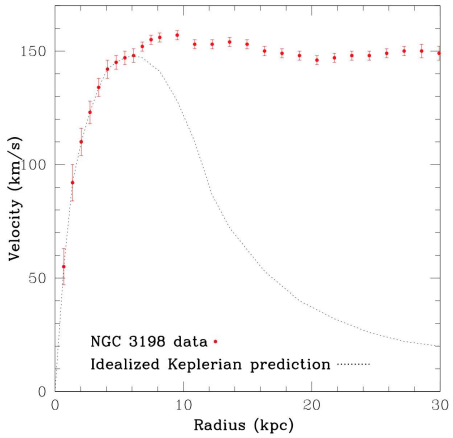


Figure 1.1: Measured rotational velocities of galaxy NGC 3198 [3] compared to an idealized Keplerian behavior.

a combination of images of it obtained with different techniques. The optical image is superimposed with pink shading obtained from the study of hot gas X-ray emission and highlighting the ordinary baryonic mass distribution. The blue region is inferred from gravitational lensing and marks where the highest mass concentration is found. It can be observed that the baryonic matter (pink) is clearly shifted with respect to the centre of the mass distribution (blue), making it possible to interprete the blue shading as the dark matter distribution. The spatial offset of the centers of total and baryonic mass corresponds to an 8σ significance [5], proving that the largest fraction of matter in the cluster is unseen.

Additional evidence for dark matter existence is given by distant galaxies measurements, which show that they are subjected to weak gravitational lensing by foreground structures not fully accounted for by the visible objects [6]. Moreover, also the velocity dispersions of dwarf spheroidal galaxies and spiral galaxy satellites cannot be explained by visible matter only [7].

A further indirect proof of the existence of dark matter is derived from the study of structure formation in the early Universe. The N-body simulations performed to study this phenomenon show indeed that the formation of structures cannot be modeled correctly without assuming the presence of dark matter. The standard model of cosmology is thus called ΛCDM, since it contains both the cosmological constant Λ and Cold Dark Matter (i.e. dark matter made up of particles moving slowly compared to the speed of light, see Sec. 1.2) in order to explain the evolution of the Universe to its present state.

The amount of DM in the present Universe is usually expressed in term of

Figure 1.2: Combination of different images of galaxy cluster 1E0657-56 [4].

relic density Ω, which is defined as the density relative to the critical density ρ_c

$$\Omega = \frac{\rho}{\rho_c}, \tag{1.2}$$

where the critical density is the value of density at which the Universe is at balance and the expansion rate goes asymptotically to zero. The relic density of DM and baryonic matter can be measured from the analysis of the cosmic microwave background (CMB), which is the background radiation from the propagation of photons in the early universe, as soon as they were decoupled from matter. The CMB exhibits the spectrum of a black body corresponding to a temperature of 2.726 K, with anisotropies at the 10^{-5} level [8] (Fig. 1.3). These anisotropies originate from temperature and density fluctuations in the early Universe and can be used to constrain the relative abundance of the different components in it. In fact, the power spectrum of the fluctuations, which shows their strength as a function of their angular scale, is very sensitive to the total amount of dark matter in the Universe, as shown in Fig. 1.4. From Planck satellite measurements of the anisotropies [8], it is then possible to obtain the relic density of dark matter (Ω_χ) and baryons (Ω_b) in the Universe, resulting in:

$$\Omega_\chi h^2 = 0.1200(12), \qquad \Omega_b h^2 = 0.02237(15), \tag{1.3}$$

where the number in parenthesis is the 1-sigma uncertainty in the last two digits and h is the reduced Hubble constant $h \approx 0.674$. This means that dark matter constitutes 26.4% of the Universe, while baryonic matter only 4.9% of it.

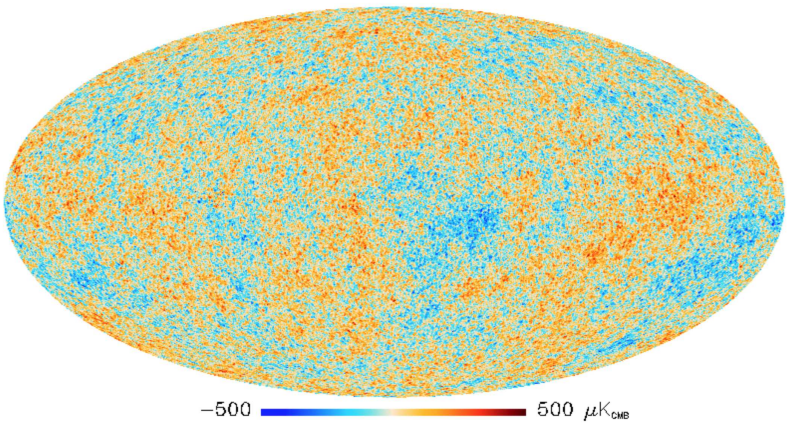


Figure 1.3: Map of the temperature fluctuations of cosmic microwave background measured by the Planck satellite [8].

1.2 Dark Matter Models

There are several models trying to explain DM evidence. One of the most well-known models assumes that DM is composed of ordinary matter clumped together in MAssive Compact Halo Objects (MACHOs), which prevent it from interacting with other baryonic matter. Some possible MACHOs candidates are (primordial) black holes, neutron stars, and brown dwarfs. However, this hypothesis is severely challenged by the hints of the non-baryonic nature of DM coming from the CMB analysis and big-bang nucleosynthesis. Moreover, searches for MACHOs have already excluded almost the entire mass range allowed for DM.

Another possibility is to suppose a modification of general relativity. Even if there are some relativistic theories able to reproduce some of the observed phenomena like galaxy rotation curves, other measurements, for example those relating to properties of galaxy clusters, cannot be easily explained by such models. Besides, it is very hard to formulate a theory of modified gravity that can be incorporated in any cosmological model.

Making the assumption that DM is made up of non-baryonic particles, all observations previously presented can be easily explained. For the rest of this thesis, particle DM is then assumed. From the different experimental observations, DM particles should have only gravitational interactions, with possible small weak interactions in addition. Moreover, they have to be stable on cosmological time scales, otherwise they would have decayed by now. Finally, following the ΛCDM model, DM is assumed to be non-relativistic when the formation of galaxies started, which, as introduced in the previous section, is referred to as

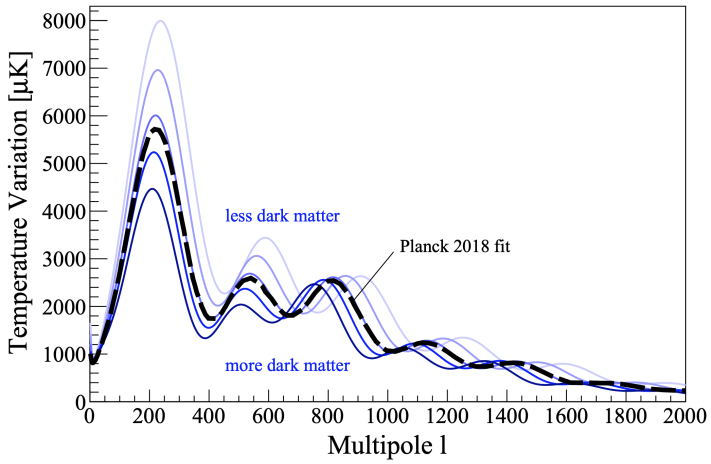


Figure 1.4: Temperature power spectrum of CMB for a dark matter density varying between 0.11 and and 0.43 (blue lines). The dashed black line shows the best fit to the Planck 2018 data [9].

cold dark matter - in contrast to hot relativistic or warm dark matter. Cold, warm and hot dark matter differ with respect to their mass and hence their background velocity dispersion: hot dark matter particles have usually masses of a few eV, warm ones of the order of keV-MeV, while cold dark matter particles can reach hundreds of GeV. For the sake of completeness, it is worth noting the the assumption of cold dark matter, although successfully describing the universe at large scales, encounters problems in describing structures at small scales [10]. As an example, the observed abundance of dwarf satellite in the Milky Way seems to be order of magnitude smaller than what predicted by simulations of the ΛCDM model.

1.3 Dark Matter Particle Candidates

In order to yield the observed relic density (Eq. 1.3), the particles composing dark matter need to have the right mass and abundance. Within the Standard Model, there is only one particle partially fulfilling the required DM properties: the neutrino. However, since the upper limit on neutrino masses is ∼2 eV [11], the neutrino is a candidate for hot dark matter, and in any case its abundance is not large enough to account for the amount of dark matter. It is then necessary to introduce new particles beyond the Standard Model.

The first possible candidates are sterile neutrinos [12]. Such particles should

be similar to SM neutrinos but right-handed, and this would prevent them from taking part in the weak interaction. They could, however, mix with the SM neutrinos, and this would allow them to decay into SM neutrinos and photons. The study of the possible decays involving sterile neutrinos, together with the analysis of their contribution to the total density, yields stringent constraints on their mass, restricted to ∼1 keV-10 MeV [13], making them candidate for warm dark matter. Nonetheless, the ranges of couplings and masses left open by these constraints still allow sterile neutrinos with an abundance sufficient to account for all the dark matter.

Another class of hypothetical DM candidates are axions, originally proposed as a possible solution for the strong CP problem [14]. Data from laboratory searches, stellar cooling and supernova 1987A illustrate that axion masses have to be below 0.01 eV. Their interactions with SM particles are expected to be very weak, implying they were not in thermal equilibrium in the early universe. The assumptions made about their production mechanism strongly affect the calculation of their relic density, yielding a large uncertainty on the prediction. However, there exist ranges of model parameters such that axions may satisfy all existing astrophysical constraints.

Both sterile neutrinos and axions are then viable candidates for dark matter under certain conditions. However, they would be very difficult to detect due to their almost non-existing interaction with SM particles. In particular, present accelerator would not be able to produce them in a sizable quantity, limiting the possibilities for discovery to other detection techniques. Since the purpose of the present work is to study DM at the Large Hadron Collider (LHC), it is necessary to study new candidates, with couplings to SM particles strong enough to allow them to be produced at the LHC. Nonetheless, it is worth noting that efforts are proposed or underway to search also for sterile neutrinos and axion DM, like SHiP [15] for heavy neutrinos and haloscopes such as ADMX [16] and HAYSTAC [17] for axions.

A general class of very promising cold DM candidates are the so-called Weakly Interacting Massive Particles, WIMPs, particles with masses of the order of a few GeV to TeV and with weak interactions with SM particles. Their appeal lies in the so-called WIMP miracle, which calculates that a particle with such a mass and weak coupling to the SM can match the observed relic density from CMB analysis in a natural way (see next section). There is a huge number of models for physics beyond the Standard Model that contain dark matter candidates in the form of a WIMP. One of the most popular is R-conserving supersymmetry (SUSY), in which the decay of the lightest super-symmetric particle (LSP) into SM and SUSY particles is vetoed respectively by the conservation of a quantum number and by kinematics, making the LSP a viable dark matter candidate if it

is neutral.

1.4 WIMP Miracle

It is normally assumed that the DM observed today is a thermal relic of the evolution of the Universe. In the early universe DM was in thermal equilibrium, i.e. DM particles annihilated into SM at the same rate at which SM particles annihilated into DM; when the interaction rate of DM decreased below the expansion rate of the Universe, no more interaction took place and it decoupled. The decoupling time and temperature, known as freeze-out temperature, determine the relic abundance of DM today, as explained in the following.

The starting point for deriving DM relic density is the Boltzmann equation

$$\frac{dn}{dt} + 3H_0 n = -\langle \sigma v \rangle (n^2 - n_{eq}^2), \tag{1.4}$$

where $\langle \sigma v \rangle$ is the thermal average of the DM annihilation cross-section multiplied by its velocity, H_0 is the Hubble constant, n is the number density (i.e. number of particles per unit volume) and n_{eq} is the number density in thermal equilibrium. The term $3H_0 n$ describes the decrease of the number density with the expansion of the Universe at the Hubble rate H_0, while the term on the right-hand side is the annihilation term. In the non-relativistic limit, n_{eq} is given by

$$n_{eq} = g \left(\frac{mT}{2\pi} \right)^{\frac{3}{2}} e^{-\frac{m}{T}}, \tag{1.5}$$

with g denoting the number of degrees of freedom, m the DM particle mass and T the temperature.

The Boltzmann equation in Eq. 1.4 cannot be solved analytically. Numerical solutions can nonetheless be extracted, leading to the behaviour depicted in Fig. 1.5, which shows the evolution of the comoving number density[1] with time. As long as thermal equilibrium is maintained, the DM density decreases following Eq. 1.4. When the annihilation rate becomes smaller than the expansion rate of the Universe, thermal equilibrium is lost and the DM density in the comoving volume becomes approximately constant.

Solving Eq. 1.4 in the approximation of long time after the freeze-out, an order-of-magnitude estimation for the relic density can be derived

$$\Omega_\chi h^2 \approx \frac{3 \times 10^{-27} \mathrm{cm}^3 \mathrm{s}^{-1}}{\langle \sigma v \rangle}, \tag{1.6}$$

[1] Since the universe is expanding, the density has to be considered w.r.t. to the "expanding volume".

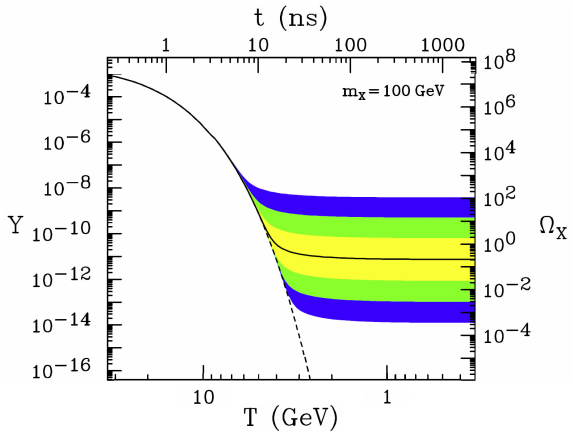


Figure 1.5: Illustration of the co-moving number density of DM Y (left) and resulting thermal relic density (right) evolving through the process of thermal freeze-out [18].

where h is the reduced Hubble constant $h \approx 0.674$. This equation shows that the present abundance of DM is determined by the annihilation cross-section at the time of freeze-out: smaller cross-sections correspond to earlier freeze-out temperatures and higher relic abundances and vice versa. This is also illustrated in Fig. 1.5 by the different coloured bands.

From Eq. 1.6, the measured DM relic abundance of Eq. 1.3 can thus be obtained assuming DM to be a WIMP, i.e. a particle with masses between 1 GeV and 1 TeV and interactions with the Standard Model particles of the order of the weak interaction, as explained above. It then looks like there is a connection between the relic abundance of DM and weak scale parameters, and this is what is usually called WIMP miracle.

1.5 WIMP Detection

If dark matter is made up of WIMPs, there are three different way to observe it, illustrated in Fig. 1.6:

- Direct detection experiments, which look for evidence of the scattering of dark matter particles off atomic nuclei within a detector
- Indirect detection, which search for the products of dark matter particle annihilations or decays in the space
- Accelerator production, when DM particles are produced at experiments.

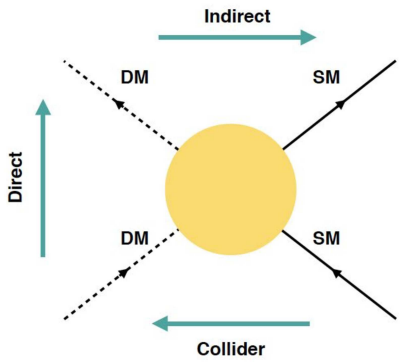


Figure 1.6: The three different possible methods for DM searches.

The three approaches of DM searches are complementary to one another and it is then important to combine them all in order to comprehensively investigate the nature of DM. As anticipated, this thesis will focus in particular on the last approach, studying DM production at the Large Hadron Collider.

1.5.1 Direct Detection

Direct detection searches for DM particles can probe a wide range of WIMP masses. They are based on a simple principle: if the Universe is filled with DM particles, some of them should reach Earth, where they would interact with matter. The signal of such an interaction consists of a nuclear recoil and could be detected at dedicated experiments. The expected event rate is proportional to

$$R \propto N \frac{\rho_\chi}{m_\chi} \langle \sigma_{\chi N} \rangle, \tag{1.7}$$

where N is the number of target nuclei, ρ_χ and m_χ are DM density and mass and $\langle \sigma_{\chi N} \rangle$ denotes the average DM-nucleon scattering cross-section. Since this rate has been calculated to be very small for WIMPs - below one event per year and per kg of material - dedicated experiments usually make use of a large amount of material. Furthermore, in order to reduce the background coming from cosmic rays and natural radioactivity, detectors are built deep underground and employ special detector shielding.

A possible DM signal observed in direct detection is expected to display two main features, which will help in confirming the DM nature of the observation. One is the annual modulation of the recoil rate, originating from the movement of the Earth around the Sun, whereas the other is the change of the recoil direction within the course of a day due to the changing direction of the Earth within the

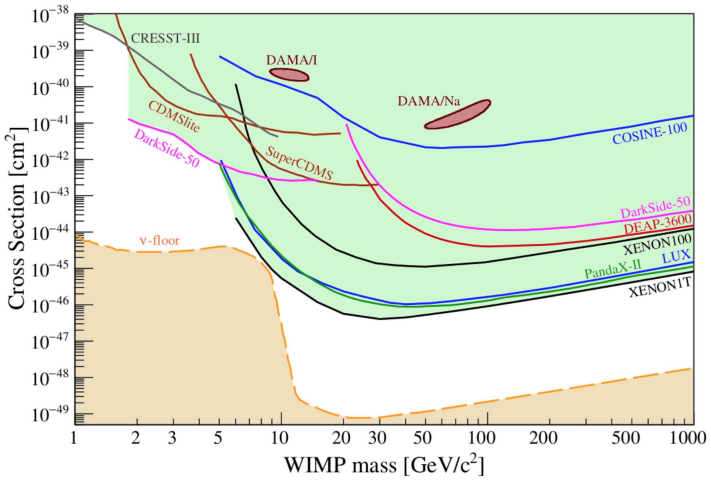


Figure 1.7: Excluded cross-sections at 90% C.L. for spin-independent coupling as a function of DM mass [9]. The dashed line limiting the parameter space from below represents the projection of the neutrino irreducible background.

DM cloud. The annual modulation is a ∼10% effect, so it would be possible to study it in detail after a discovery with a subsequent generation of detectors, while exploitation of directionality is considerably more challenging.

Direct detection experiments can be divided into two types, depending on the nature of DM couplings to SM. The first category looks for spin-independent (SI) interactions, where the cross-section depends only on the number of nucleons squared and motivates the usage of massive target particles like germanium or xenon. The other one hunts instead spin-dependent (SD) interactions, where the coupling between the DM and the target SM particle depends on the nuclear spin, motivating the usage of highly polarized target material such as fluorine. Some of the leading SI experiments include CDMS-II [19], CRESST-II [20], CoGeNT [21], DarkSide-50 [22], LUX [23], Xenon100 [24], XENON1T [25] and PANDAX-II [26], while among SD experiments there are IceCube [27], PICASSO [28], and SIMPLE [29]. The results of these experiments are normally expressed as allowed/excluded areas in the plane defined by the DM mass and its cross-section with nucleons, and the latest results are shown in Figs. 1.7 and 1.8.

1.5.2 Indirect Detection

Indirect detection experiments look for the decay products resulting from DM annihilation. Since the annihilation rate of DM is proportional to its density,

the searches are concentrated in regions where DM abundance is expected to be sizeable, like galaxy centres or within heavy objects, such as the Sun or the Earth. The latter act indeed as gravitational traps for DM particles, leading to an increase DM local density.

The products of annihilation may be various, but the most studied are gamma rays, neutrinos, positrons, anti-protons or anti-nuclei. A signal for the presence of DM would consist of an excess of observations of these particles, not accounted for by any other source.

Leading indirect detection experiments include space-based experiments, such as AMS [30] and Fermi-LAT [31], or experiments using large telescope arrays, such as HESS [32], which can be used to constrain DM annihilation rate through measurements of the photon spectrum. Telescope experiments have shown to be typically more sensitive than space-based experiments for high-mass DM particles.

1.5.3 Collider Production

WIMPs can also be directly produced at collider experiments, where the weakly interacting DM particles escape the detector, leaving as signature particles or jets recoiling against an invisible state and missing momentum. By searching for an excess of such events, it is possible to place constraints on WIMP production in association with other visible particles. The results can be translated into limits on the cross-sections of DM-nucleon scattering and DM annihilation, making the collider results comparable to direct and indirect detection experiments, respectively.

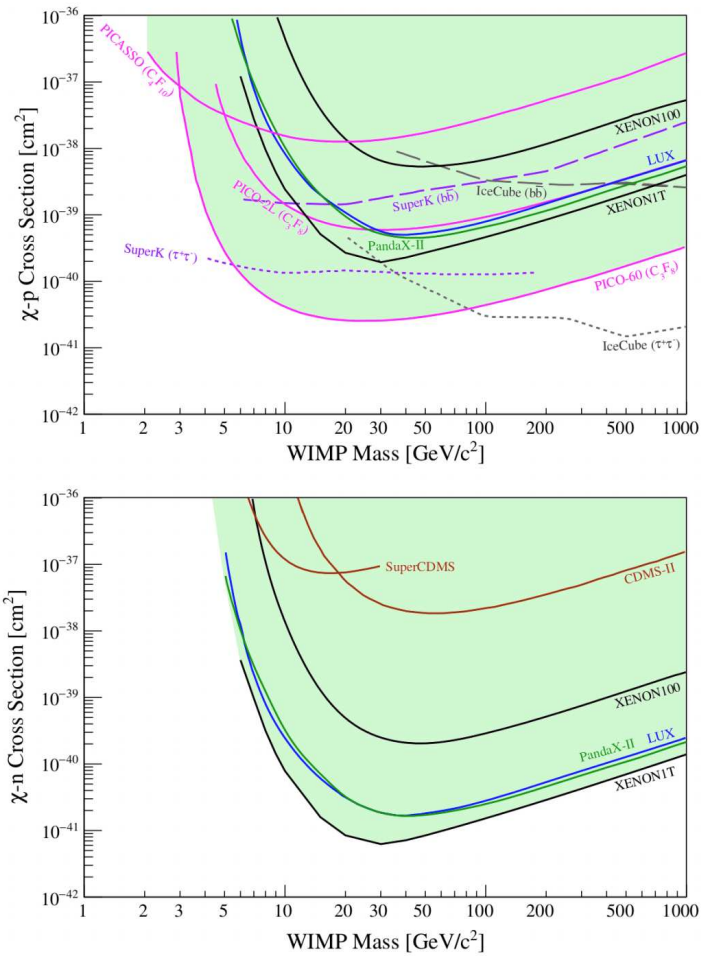


Figure 1.8: Excluded cross-sections at 90% C.L. for spin-dependent coupling as a function of DM mass. Top: DM-proton interactions. Bottom: DM-neutron interactions [9].

Chapter 2

The Two Higgs Doublet Model+Pseudoscalar Mediator

In order to search for dark matter at the LHC, a model for the production of DM in the collision of two quarks/gluons is needed. Any such model will have free parameters, such as, for instance, the mass of the DM, its spin and its couplings to SM particles, and any search will be sensitive to specific ranges of these parameters. Models of different level of complexity have been used as a template for DM searches, starting from simple effective theories, to full fledged UV-complete theories such as supersymmetry. The present thesis is focused on the Two Higgs Doublets Model + Pseudoscalar Mediator (2HDMa), in which a pseudoscalar particle acting as a "mediator" between SM and DM is introduced together with a second Higgs doublet. This model will be presented in the second part of this Chapter, while the first part will focus on the so-called simplified models, which share some similarities with 2HDMa and were first used as benchmarks for LHC searches.

2.1 Simplified Models

Due to the numerous theoretical models providing a WIMP candidate, it would be desirable to study possible DM signatures in a model-independent way. During LHC Run-1, this was done using effective field theories (EFTs), which assume a contact interaction between the DM (made of a single particle) and the SM particles that can be described by only two parameters, the effective scale Λ and the DM mass m_χ. However, the EFT approach is not applicable when the effective scale is comparable to the momentum transferred during the partonic interactions, as in the case of LHC Run-2, where it is of the order of few TeV. To overcome this problem, a new class of simplified models has been developed.

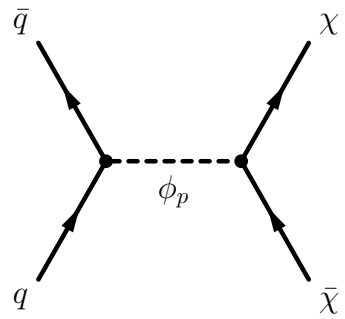


Figure 2.1: Dark matter production in a simplified model with pseudoscalar mediator exchanged through s-channel.

The simplified models introduce a second particle beside the DM one, a mediator, which is responsible for the interaction between the SM and the DM particles.

The majority of simplified models can be thought as the limit of a more general scenario, where most of new particles are integrated out because they have a mass larger than the energy reachable at the LHC or because they do not participate in the interactions of DM and SM particles. With a similar reasoning, the EFT framework can be recovered in the limit where the mass of the mediator of simplified models is very large. Nonetheless, there are new DM models which cannot be recast in terms of simplified models, due to more than one mediator active at the same time or interfering with each other.

There are many possible choices of mediator in simplified models, and this section will focus in particular on a pseudoscalar mediator exchanged through s-channel (Fig. 2.1). A possible Lagrangian for such a model is given by [33]

$$\mathcal{L} \supset g_f m_a \sum_f \frac{y_f}{\sqrt{2}} \bar{f}\gamma^5 f + g_\chi m_{\phi_p} \bar{\chi}\gamma^5 \chi, \tag{2.1}$$

where m_{ϕ_p} is the mass of the pseudoscalar mediator with coupling g_f and g_χ to SM fermions and dark matter and χ is a spin-1/2 DM particle.

The inclusion of two new particles in the SM lead to the addition of new parameters, on which kinematic distributions depend in a non-trivial way. Simplified models of this type are fully characterized by five parameters: the masses of the two additional particles m_χ and m_{ϕ_p}, the two couplings g_f and g_χ and the width of the mediator Γ. However, keeping Γ as a free parameter may lead to the unphysical situation where it is smaller than the sum of the partial widths of the decays into DM and SM fermions, typically for large m_{ϕ_p}. In order to avoid this, it is usually assumed that the mediator does not couple to any other light

particle, so its width can be computed as:

$$\Gamma = \Gamma_{\chi\bar{\chi}} + \sum_f \Gamma_{f\bar{f}} + \Gamma_{gg}. \tag{2.2}$$

This ansatz is known as Minimal Width Assumption [33]. It removes one free parameter of the model but, as a drawback, the dependence of the cross-sections on the other parameters becomes more complex.

Another model assumption is Minimal Flavor Violation (MFV), which means that the flavor structure of the couplings between Dark Matter and ordinary particles follows the same structure of the Standard Model. This choice ensures that simplified models do not violate flavor constraints coming from experimental measurements. As a consequence, the pseudoscalar mediator must have couplings to fermions proportional to the SM Higgs couplings, i.e. to the fermion masses. This means that the couplings to third generation quarks is larger than the couplings to the first two, providing an enhanced cross-section for DM production in association with top quarks.

2.2 Two Higgs Doublets Model + Pseudoscalar Mediator

Unfortunately, the operators in the Lagrangian of simplified models (Eq. 2.1) violate gauge invariance, because the left- and right-handed SM fermions belong to different representations of the SM gauge group. This leads to unitarity-violating amplitudes. In order to overcome this issue, the model proposed in [34] adds to the SM not only a pseudoscalar mediator and a dark matter particle, but also a second Higgs doublet. This model is know as Two Higgs Doublets Models + Pseudoscalar Mediator, shortened in 2HDMa, and it is the simplest gauge-invariant and renormalisable extension of simplified pseudoscalar models.

2.2.1 Scalar Potential

The 2HDMa tree-level scalar potential for the two Higgs doublets is given by:

$$\begin{aligned} V_H =& \mu_1 H_1^\dagger H_1 + \mu_2 H_2^\dagger H_2 + (\mu_3 H_1^\dagger H_2 + h.c.) + \lambda_1 (H_1^\dagger H_1)^2 + \lambda_2 (H_2^\dagger H_2)^2 \\ &+ \lambda_3 (H_1^\dagger H_1)(H_2^\dagger H_2) + \lambda_4 (H_1^\dagger H_2)(H_2^\dagger H_1) + \left[\lambda_5 (H_1^\dagger H_2)^2 + h.c.\right]. \end{aligned} \tag{2.3}$$

The vacuum expectation values (VEVs) of the Higgs doublets are $\langle H_i \rangle = (0, v_i/\sqrt{2})^T$, with $v = v_1^2 + v_2^2 \sim 246$ GeV and $\tan\beta = v_2/v_1$. The mass-squared terms μ_j, the quartic couplings λ_k and the VEVs are all real to avoid problems with electric

dipole moments. The three physical neutral Higgs that emerge from V_H are in such a case both mass and CP eigenstates.

In order to couple fermionic DM to the SM through pseudoscalar exchange, a CP-odd mediator P is mixed with the CP-odd Higgs arising from the scalar potential with the interaction terms

$$V_P = \frac{1}{2}m_P^2 P^2 + P(ib_P H_1^\dagger H_2 + h.c.) + P^2(\lambda_{P1} H_1^\dagger H_1 + \lambda_{P2} H_2^\dagger H_2), \tag{2.4}$$

with m_P and b_P having mass dimensions. These interactions mix the neutral CP-even eigenstates with a mixing angle α. The portal coupling b_P mixes instead the two neutral CP-odd weak eigenstates, with θ representing the associated mixing angle. The resulting CP-even mass eigenstates are denoted by h and H, while the odd ones are called A and a and are responsible of the interaction between SM particles and DM. The scalar spectrum also contains two charged mass eigenstates $H^\pm$ of identical mass.

Diagonalizing the mass-squared matrices of the scalar states leads to relations between the fundamental parameters entering V_H and V_P. These relations allow the elimination of some of the parameters in favour of sines and cosines of mixing angles, VEVs and the masses of the physical Higgses. The phenomenology of the model is thus fully described by the angles α, β, θ, the EW VEV v, the quartic couplings λ_3, λ_{P1}, λ_{P2} and the masses m_h, m_H, m_A, $m_{H^\pm}$, m_a.

2.2.2 Alignment/Decoupling Limit

After EW symmetry breaking, the Higgs kinetic terms give rise to interactions between the CP-even mass eigenstates h and H and the massive EW gauge bosons. These interactions are of the form

$$\mathcal{L} \supset (\sin(\beta-\alpha)h + \cos(\beta-\alpha)H)\left(\frac{2M_W^2}{v}W_\mu^+ W^{-\mu} + \frac{M_Z^2}{v}Z_\mu Z^\mu\right). \tag{2.5}$$

It is usually assumed the well-motivated alignment/decoupling limit of the 2HDMa, where $\alpha = \beta - \pi/2$. In this case $\cos(\beta-\alpha) = 0$, meaning that the boson h has SM-like EW gauge boson couplings. It can therefore be identified with the boson of mass $m_h \approx 125$ GeV discovered at the LHC [35, 36] and the constraints from the Run-1 combination of the ATLAS and CMS measurements are automatically fulfilled.

2.2.3 Yukawa Assignments

The Yukawa couplings are given by:

$$\mathcal{L} = -\sum_{i=1,2} \left(\bar{Q} Y_u^i \tilde{H}_i u_R + \bar{Q} Y_d^i H_i d_R + \bar{L} Y_\ell^i H_i \ell_R + h.c. \right), \tag{2.6}$$

where Y_f^i are Yukawa matrices, $\tilde{H}_i = \epsilon H_i^*$, Q and L are the left-handed quark and lepton doublets and u_R, d_R and ℓ_R are the right-handed up-type and down-type quarks and charged lepton singlets. In order to satisfy the stringent experimental limits on FCNCs (Flavour-Changing Neutral Currents), it is necessary to require that no more than one of the Higgs doublets couples to fermions of a given charge. This natural flavour conservation hypothesis can be satisfied with four different choices of Yukawa matrices, assuming that up-type quarks couple to H_2 by convention:

$$\begin{aligned} &Y_u^1 = Y_d^1 = Y_\ell^1 = 0 \quad \text{(type I)}, \\ &Y_u^1 = Y_d^2 = Y_\ell^2 = 0 \quad \text{(type II)}, \\ &Y_u^1 = Y_d^1 = Y_\ell^2 = 0 \quad \text{(type III)}, \\ &Y_u^1 = Y_d^2 = Y_\ell^1 = 0 \quad \text{(type IV)}. \end{aligned} \tag{2.7}$$

In the alignment/decoupling limit, after EW symmetry breaking, the most relevant fermion-boson interactions are given by

$$\begin{aligned} \mathcal{L} \supset &- \frac{y_t}{\sqrt{2}} \bar{t} \Big[h + \xi_f^M H - i \xi_f^M (\cos\theta A - \sin\theta a) \gamma_5 \Big] t \\ &- \sum_{f=b,\tau} \frac{y_f}{\sqrt{2}} \bar{f} \Big[h + \xi_f^M H + i \xi_f^M (\cos\theta A - \sin\theta a) \gamma_5 \Big] f \\ &- \frac{y_t}{\sqrt{2}} V_{tb} \xi_t^M H^+ \bar{t}_R b_L + \frac{y_b}{\sqrt{2}} V_{tb} \xi_b^M H^+ \bar{t}_L b_R + h.c. \\ &- i y_\chi (\sin\theta A + \cos\theta a) \bar{\chi} \gamma_5 \chi, \end{aligned} \tag{2.8}$$

where $y_f = \sqrt{2} m_f / v$ denote the SM Yukawa couplings, y_χ is the DM coupling and V_{ij} are the elements of the CKM matrix. The couplings ξ_f^M contain the dependence on the choice of Yukawa sector:

$$\begin{aligned} \xi_t^I = \xi_b^I = \xi_\tau^I = -\cot\beta \quad &\text{(type I)}, \\ \xi_t^{II} = -\cot\beta, \quad \xi_b^{II} = \xi_\tau^{II} = \tan\beta \quad &\text{(type II)}, \\ \xi_t^{III} = \xi_b^{III} = -\cot\beta, \quad \xi_\tau^{III} = \tan\beta \quad &\text{(type III)}, \\ \xi_t^{IV} = \xi_\tau^{IV} = -\cot\beta, \quad \xi_b^{IV} = \tan\beta \quad &\text{(type IV)}. \end{aligned} \tag{2.9}$$

In this thesis, the study of 2HDMa will be restricted to a Yukawa sector of type

II. This choice is due to the fact that type II 2HDM are historically the most studied, since others 2HDM models like supersymmetry and the Peccei-Quinn model are all of type II.

2.2.4 Electroweak Precision Constraints

Electroweak precision measurements constrain the differences between the masses of the additional bosons H, A, $H^{\pm}$ and a, because these new particles introduce loop corrections to W and Z propagators. For $m_H = m_{H^{\pm}}$ and $\cos(\beta - \alpha) = 0$, these corrections disappear thanks to a custodial symmetry in the potential of Eq. 2.3. Also for $m_A = m_{H^{\pm}}$ and $\cos(\beta - \alpha) = 0$ this custodial symmetry is present, but the interaction potential of Eq. 2.4 breaks such symmetry softly in this case. This leads to stringent constraints on the pseudoscalar mixing angle θ and the mass splitting between m_H, m_A and m_a. Therefore, to keep $\sin\theta$ and m_a as free parameters, it is usually assumed that the masses of H, A and $H^{\pm}$ are equal, $m_H = m_A = m_{H^{\pm}}$.

2.2.5 Bounded from Below Potential Constraints

Requiring that the scalar potential of the 2HDMa is bounded from below (BFB) imposes limits on the possible choices of the additional boson masses, mixing angles and quartic couplings. Taking $m_H = m_A = m_{H^{\pm}}$, the BFB conditions can be cast into two inequalities. The first one requires

$$\lambda_3 > 2\lambda, \tag{2.10}$$

where λ is the cubic SM Higgs self-coupling, $\lambda \approx 0.13$. The second BFB condition states instead

$$\lambda_3 > \frac{m_H^2 - m_a^2}{v^2}\sin^2\theta - 2\lambda\cot^2(2\beta). \tag{2.11}$$

This formula implies that large values of $(m_H^2 - m_a^2)/v^2 \sin^2\theta$ are only compatible with BFB if the quartic coupling λ_3 is sufficiently large. Since due to perturbativity constraints $\lambda_3 < 4\pi$, the choice $\lambda_3 = 3$ is usually imposed. Moreover, in order to keep the total widths of H and A small, $\lambda_3 = \lambda_{P1} = \lambda_{P2}$ is assumed.

2.2.6 Recommended Parameter Choice

Summarizing the discussion above, in this work the following parameter choices are adopted:

$$m_H = m_A = m_{H^\pm}, \tag{2.12}$$
$$\cos(\beta - \alpha) = 0, \tag{2.13}$$
$$\lambda_3 = \lambda_{P1} = \lambda_{P2} = 3, \tag{2.14}$$
$$Y_u^1 = Y_d^2 = Y_\ell^2 = 0. \tag{2.15}$$

In addition, the DM mass and coupling are set to

$$m_\chi = 10 \text{ GeV}, \ y_\chi = 1. \tag{2.16}$$

The h particle of the model is identified with the boson observed by the ATLAS and CMS Collaborations, and its mass is set at 125 GeV [35] [36].

Under these assumptions, the free parameters of the model are only four: $m_H = m_A = m_{H^\pm}$, m_a, $\tan\beta$ and $\sin\theta$.

2.3 Experimental signatures

The experimental searches performed at the LHC aim to observe DM particles mostly when they are produced in association with visible SM particles. Such processes, which are observed as particles or jets recoiling against an invisible state, give raise to the so-called $E_{\text{T}}^{\text{miss}} + X$ signatures, where $E_{\text{T}}^{\text{miss}}$ is the missing transverse momentum observable in the detector due to DM production and X can be a light quark or gluon, a W or Z boson, a b or t quark, a photon or a Higgs boson [33]. In the 2HDMa model there are two different classes of $E_{\text{T}}^{\text{miss}} + X$ signatures: resonant and non-resonant.

In the first case, according to the Higgs sector's mass hierarchy, one of the heavy Higgs bosons ($H, A, H^\pm$) introduced in the model can be produced on-shell and subsequently decay to a SM particle and the pseudoscalar a, which then decays to a pair of DM particles. Channels such as $E_{\text{T}}^{\text{miss}} + h$, $E_{\text{T}}^{\text{miss}} + Z$ and $E_{\text{T}}^{\text{miss}} + tW$ are part of this class, as shown by the Feynman diagrams on the left-hand side of Fig. 2.2. Thanks to the on-shell resonant production of a massive particle, the cross-section of these signals can be greatly enhanced. The most important channels in the non-resonant class are instead $E_{\text{T}}^{\text{miss}} + t\bar{t}$, $E_{\text{T}}^{\text{miss}} + b\bar{b}$ and $E_{\text{T}}^{\text{miss}} + j$ production, presented in Fig. 2.3.

In addition to $E_{\text{T}}^{\text{miss}} + X$ signatures, also some visible final states can be used to explore the parameter space of the 2HDMa model at the LHC. Within these signatures, the most interesting is the four-top final state, presented in Fig. 2.4. The possibility to observe 2HDMa realisations simultaneously in a number of different channels, some of which are not limited by systematic errors and can be improved by statistics, optimises the chances of discovery.

All the signatures introduced above, with the exception of $E_{\text{T}}^{\text{miss}} + tW$, were

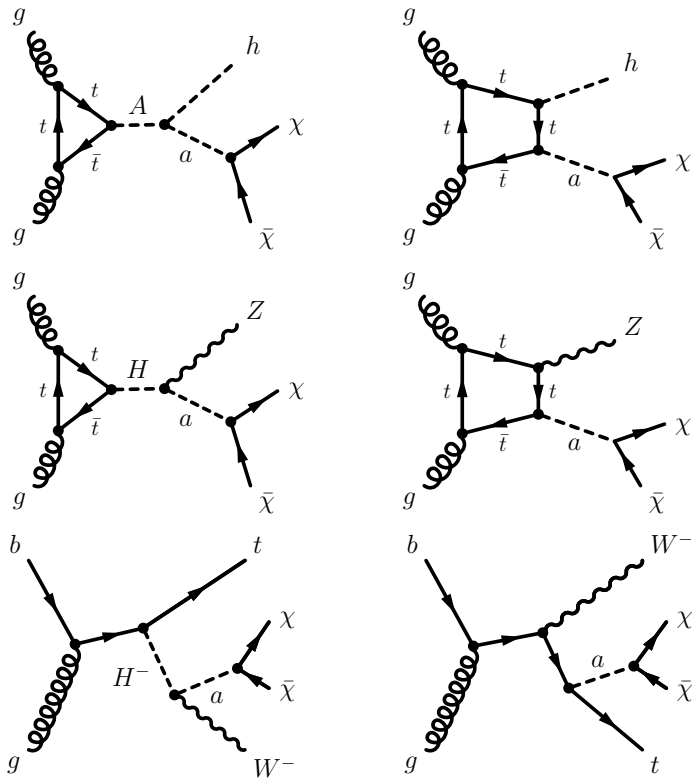


Figure 2.2: Example diagrams giving rise to $E_{\mathrm{T}}^{\mathrm{miss}} + h$ (top), $E_{\mathrm{T}}^{\mathrm{miss}} + Z$ (centre), and $E_{\mathrm{T}}^{\mathrm{miss}} + tW$ (bottom) signals in the 2HDMa model.

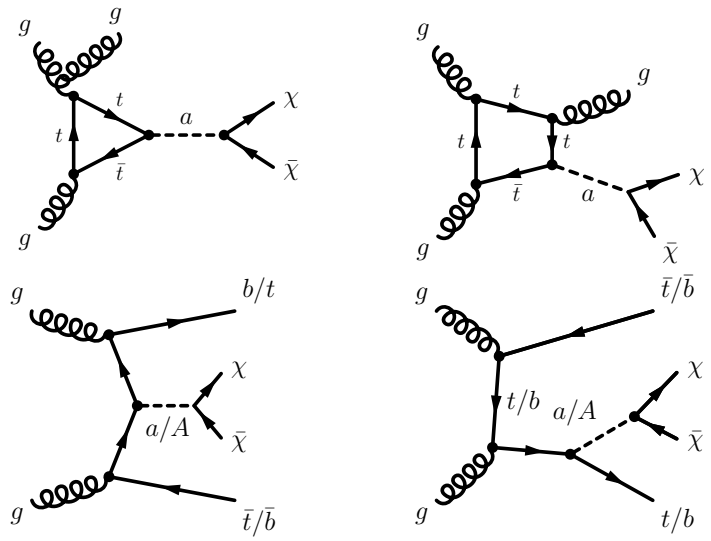


Figure 2.3: Example diagrams giving rise to $E_{\mathrm{T}}^{\mathrm{miss}}+j$ (top) and $E_{\mathrm{T}}^{\mathrm{miss}}+t\bar{t}/E_{\mathrm{T}}^{\mathrm{miss}}+b\bar{b}$ (bottom) signals in the 2HDMa model.

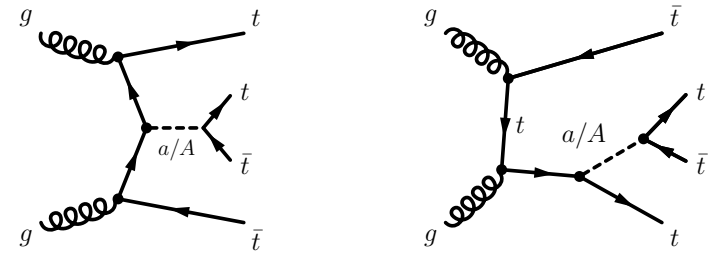


Figure 2.4: Example diagrams giving rise to 4-top signal in the 2HDMa model.

investigated by the ATLAS Collaboration with the first 36.1 fb^{-1} Run-2 data in different dark matter theoretical framework, and then reinterpreted in the context of 2HDMa in a dedicated paper [1] (with the exception of $j + E_{\mathrm{T}}^{\mathrm{miss}}$). The $E_{\mathrm{T}}^{\mathrm{miss}} + tW$ has been instead addressed in a dedicated analysis, which will be described in detail in Chapter 8.

The $E_{\mathrm{T}}^{\mathrm{miss}} + h$ signal has been looked for in the $h \to \gamma\gamma$ [37] and $h \to b\bar{b}$ [38] channels. Both searches use $E_{\mathrm{T}}^{\mathrm{miss}}$ as main selection variable to discriminate the signal from the background, with the $E_{\mathrm{T}}^{\mathrm{miss}} + h(\gamma\gamma)$ channel being more sensitive to lower $E_{\mathrm{T}}^{\mathrm{miss}}$ values than the $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ one thanks to the presence of photons, which can be exploited to select interesting events also at lower $E_{\mathrm{T}}^{\mathrm{miss}}$ values. The $E_{\mathrm{T}}^{\mathrm{miss}}+h(b\bar{b})$ channel is instead more sensitive to smaller production cross-sections.

The two channels are thus complementary: if the masses of A and a are similar, the $E_{\mathrm{T}}^{\mathrm{miss}} + h(\gamma\gamma)$ channel sensitivity dominates, while 2HDMa realisations with a larger mass hierarchy can be better probed via the $E_{\mathrm{T}}^{\mathrm{miss}} + h(b\bar{b})$ final state.

The $E_{\mathrm{T}}^{\mathrm{miss}}+Z$ searches have focused on both the leptonic [39] and hadronic [40] channels, targeting events consistent with a boosted Z boson produced in the direction opposite to the missing energy. The hadronic and leptonic signatures are complementary, since hadronic decays of the Z boson are more frequent than leptonic ones, but suffer from larger backgrounds.

The $E_{\mathrm{T}}^{\mathrm{miss}}+t\bar{t}$ signal has been investigates in three different channels, according to the decays of the W bosons from the top-quark decays and thus the number of leptons in the final states [41, 42]. The characteristics of this signature and the reinterpretation of existing results in the context of 2HDMa will be described in detail in Chapter 7, since the author of this thesis is the main contributor of this work. The reinterpretation strategy developed for the $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ signal has also been successfully applied to $E_{\mathrm{T}}^{\mathrm{miss}} + b\bar{b}$ [41].

Finally, the four-top final state has been searched for in events characterised by a single lepton and high jet multiplicity [43].

Since none of the existing analyses found any significant data excess compatible with a DM signal, their reinterpretation allowed to set exclusion limits on the 2HDMa parameter space [1]. The exclusion power of existing analyses have been investigated in four benchmark scenarios, consistent with bounds from electroweak precision, flavour and Higgs observables, chosen by a dedicated LHC Dark Matter Working Group [44]. As expected, the $E_{\mathrm{T}}^{\mathrm{miss}} + Z$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h$ channels proved to be the most constraining ones in the most part of the chosen benchmarks, thanks to the resonant enhancement of their cross-section. In order not to anticipate some of the main results of this thesis, the final exclusion limits coming from the reinterpretation of existing analyses will be presented at the end of Chapter 7, where the recast of the existing $E_{\mathrm{T}}^{\mathrm{miss}} + t\bar{t}$ analyses is described in detail.

Chapter 3

The ATLAS Experiment

The present study performs dark matter searches using proton-proton collisions data collected between 2015 and 2018 by the ATLAS experiment, during the so-called Run-2. ATLAS (A Toroidal LHC ApparatuS) is one of the four main experiments at the Large Hadron Collider (LHC), the largest particle accelerator in the world. This chapter will provide a brief overview of the LHC and present an outline of the ATLAS experiment, paying particular attention to the subsystems relevant for the analyses.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [45] is a 27-kilometre-long circular particle accelerator and collider situated at CERN, on the border of Switzerland and France near Geneva. It is capable of accelerating two beams of either protons or heavy nuclei in opposite directions and makes them collide head-on at unprecedented centre-of-mass-energies in four beam-crossing points. Up to now, the LHC delivered proton-proton (pp) collisions at centre-of-mass energies of 7, 8 and 13 TeV. Besides, it has also collided heavy ions at centre-of-mass energies per nucleon of 2.76 TeV, 5.02 TeV ($Pb-Pb$), and 5.44 TeV ($Xe-Xe$), as well as $p-Pb$ with 5.02 TeV and 8.16 TeV centre-of-mass energy per nucleon. The LHC first run, Run-1, took place between 2009 and 2013, while the second one, Run-2, started in 2015 and lasted until the end of 2018.

Before being injected into the LHC, protons (or heavy ions) need to be accelerated. This pre-acceleration proceeds in several steps, represented in Fig. 3.1. The acceleration cycle for proton collisions starts with the extraction of protons from hydrogen atoms through ionization. The protons are first fed into a linear accelerator (LINAC2), where they reach an energy of 50 MeV, and are then inserted into a chain of injectors: the Proton Synchrotron Booster (PSB), the Proton

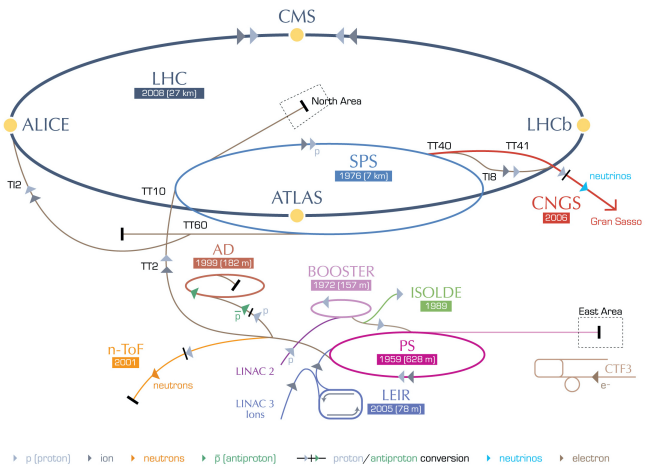


Figure 3.1: The CERN accelerator complex [46].

Synchrotron (PS), and the Super Proton Synchrotron (SPS). This complex accelerates the protons to 450 GeV before finally injecting them into the LHC ring via two transfer lines. Inside the LHC, the protons are grouped in distinct bunches 25 ns apart, which are further accelerated in both directions in separate vacuum pipes up to an energy of 6.5 TeV per beam. They are then brought to collide in four interaction points, where the four LHC main experiments (ATLAS, CMS, LHCb and ALICE) are situated. Heavy ions follow a slightly different path: they are accelerated by the LINAC3 and by the Low Energy Ion Ring (LEIR) before entering in the PS.

3.1.1 Key Parameters

The expected event rate of a certain process X is determined by the product of its cross-section, which depends on the centre-of-mass energy of the collider $\sqrt{s}$, and the so-called instantaneous luminosity $\mathcal{L}$:

$$\frac{dN_X}{dt} = \mathcal{L} \cdot \sigma_X(\sqrt{s}). \tag{3.1}$$

The luminosity and the centre-of-mass energy are therefore the most important parameters of a particle collider.

The instantaneous luminosity $\mathcal{L}$ indicates the number of collisions that can be produced in a particle accelerator per cm^2 and per second. It depends on the number of protons per bunch, the number of bunches, and the configuration of the magnets (quadrupoles) in the proximity of the experiments, which have to

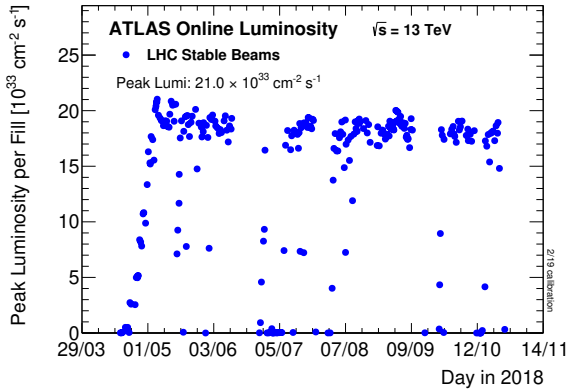


Figure 3.2: Peak instantaneous luminosity delivered to ATLAS for each LHC fill during stable beams for *pp* collisions at 13 TeV as a function of time in 2018.

focus the beams into the interaction point. The instantaneous luminosity can be approximated by:

$$\mathcal{L} = \frac{n_b n_p^2 f}{4\pi\sigma_T^2} F, \tag{3.2}$$

where n_b denotes the number of bunches, f is the revolution frequency and n_p the number of protons per bunch. The numerator gives hence the number of interactions per time interval. In the denominator, σ_T indicates the beam transverse size at the interaction point, while the factor F accounts for a possible beam crossing angle.

The LHC was designed to reach peak instantaneous luminosities of $\mathcal{L} = 10^{34}$ cm^{-2} s^{-1}, but during Run-2 the design luminosity was exceeded, reaching peaks of about $\mathcal{L} = 2 \cdot 10^{34}$ cm^{-2} s^{-1} (Fig. 3.2). The instantaneous luminosity at LHC is so high that for each bunch crossing several *pp* interactions may occur. This phenomenon is called pile-up and it requires the experiments to be able to distinguish the different interaction vertices.

The amount of collisions produced over a certain time is given by the integrated luminosity, $L = \int \mathcal{L}\, dt$, measured in units of inverse cross-section. From Eq. 3.1, the integrated luminosity is directly proportional to the number of events N_X of a certain process produced over a given period, $N_X = \sigma_X(\sqrt{s})L$, which means the higher the integrated luminosity, the greater the discovery potential. The integrated luminosity provided by LHC and recorded by the ATLAS experiment for physics analyses during Run-2 is 139 fb^{-1} (3.2 fb^{-1} in 2015, 33.0 fb^{-1} in 2016, 44.3 fb^{-1} in 2017 and 58.5 fb^{-1} in 2018 - Fig. 3.3).

The centre-of-mass energy, which expresses the energy available in the collisions, is the parameter that defines the phase-space accessible to the final state

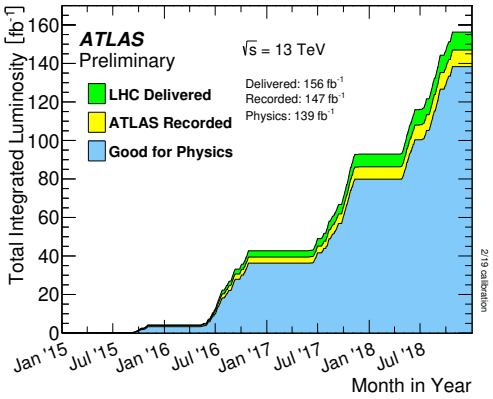


Figure 3.3: Integrated luminosity versus time delivered to ATLAS (green), recorded by ATLAS (yellow), and certified to be good quality data (blue) during stable beams for *pp* collisions at 13 TeV centre-of-mass energy in 2015-2018.

products of the interaction and the particles that can be created. As shown in Fig. 3.4, the production cross-section of a process is a function of this parameter. The maximum centre-of-mass energy reached by the LHC during Run-2 is 13 TeV, the highest one ever achieved in a particle collider, allowing to study rare processes inaccessible before.

3.2 The ATLAS Detector

ATLAS [49, 48] is a general-purpose experiment located at the Interaction Point 1 on the LHC ring, designed to detect and record all the particles emerging from the collisions inside the LHC. It is the largest particle detector ever built: it is a 44-meter-long cylinder with a radius of about 12 meters, weighing more than 7000 tons. The two LHC beams collide in the centre of the detector, giving rise to different particles with a broad range of energies. The design of the experiment was driven by the necessity to precisely measure and identify as many of the products of the interactions as possible, aiming at observing rare processes and performing precision measurements.

The ATLAS structure consists of a series of concentrical layers with separate functions surrounding the interaction point. The detector closest to the interaction point - known as Inner Detector (ID) - is a tracking detector, made to measure with high precision the trajectory of charged particles and reconstruct the interaction vertex. The ID is enclosed by a thin solenoid, which provides an axial magnetic field of 2 T that bends the trajectory of charged particles, allowing the precise measurement of their momenta. The solenoid is followed by

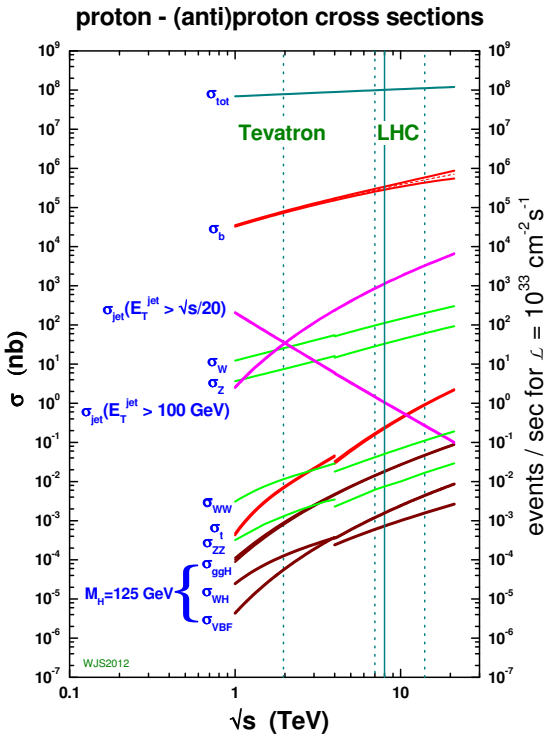


Figure 3.4: Production cross-section for different process as a function of the centre-of-mass energy [47].

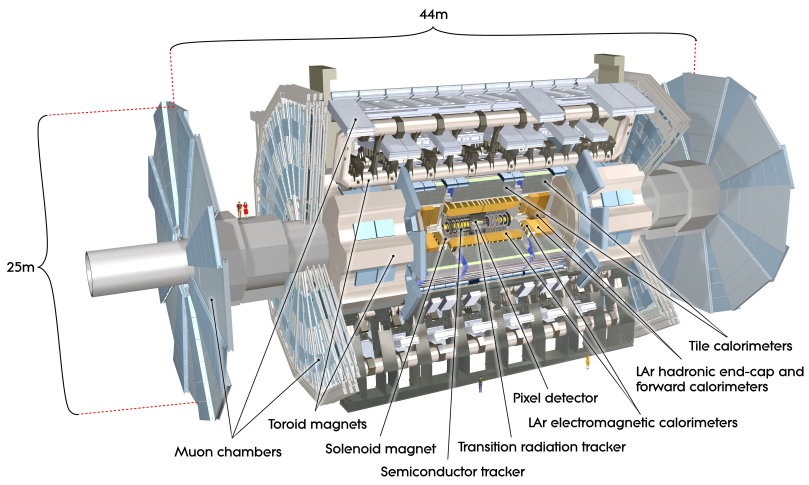


Figure 3.5: Drawing of the ATLAS detector showing the different subdetectors and the magnet systems [48].

the calorimeters, where the majority of particles are absorbed and their energy measured. Finally, the muon system surrounds the calorimeters, reconstructing muons tracks and measuring their momentum thanks to three huge air-core toroid magnets. A cut-away view of the ATLAS detector layout is shown in Fig. 3.5. The cylindrical part in the central region is called barrel, while the detectors at both ends form the end-caps.

3.3 Coordinate System and Useful Variables

Given the intrinsic symmetry of the detector, the ATLAS experiment makes use of a modified cylindrical coordinate system (ϕ, η, z) instead of the classical cartesian one (x, y, z), where ϕ and η are defined in the following. The origin of the system is set in the interaction point. The x-axis is pointed radially towards the centre of the LHC ring, while the y-axis points upwards and the z-axis points along the beam pipe, with direction defined by the requirement of a right-handed system. The azimuthal angle ϕ is the angle in the x-y plane from the x-axis, while the pseudo-rapidity η is defined as

$$\eta = -\ln(\tan\frac{\theta}{2}), \tag{3.3}$$

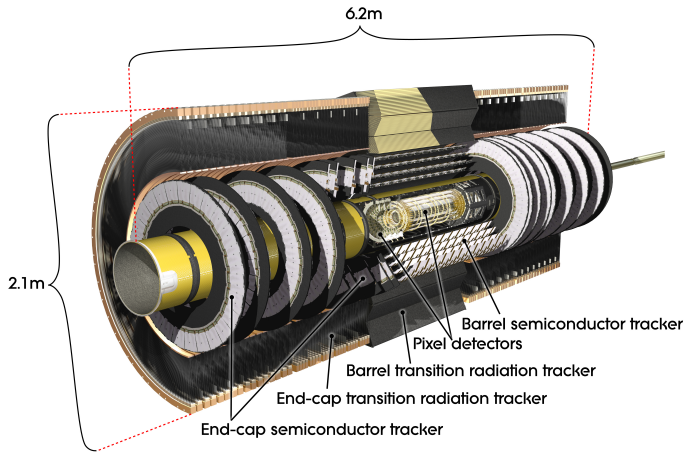


Figure 3.6: Schematic view of the ATLAS inner detector. The IBL (not shown here) is located between the beam pipe and the inner-most layer of the Pixel detector [48].

where θ is the polar angle formed with the z-axis. The pseudo-rapidity η has been chosen over θ because differences $\Delta\theta$ are not Lorentz-invariant. A large absolute value of η indicates a point close to the beam line, while a small value a point near the x and y axes.

The angular distance between two particles is often measured in terms of ΔR, which is defined as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$. Another crucial quantity for each reconstructed object is the transverse momentum measured in the $x - y$ plane, i.e. $p_T = \sqrt{p_x^2 + p_y^2}$. The importance of this variable comes from the fact that the colliding protons have momenta approximately parallel to the beam axis, setting the total transverse momentum to be zero in the final state due to momentum conservation. This means that summing all objects transverse momenta it is possible to reconstruct the momentum carried away by invisible particles like neutrinos or BSM particles, and this is usually one of the most important discriminating variables of analyses (the so-called missing transverse momentum $E_{\mathrm{T}}^{\mathrm{miss}}$).

3.4 Inner Detector

The Inner Detector (ID) [48] is ATLAS innermost system. It is 6.2 m long and has a diameter of 2.1 m, covering the pseudorapidity region with $|\eta| < 2.5$. The ID offers a precise measurement of track impact parameters, with resolution down to 80 μm along the z-axis and 20 μm in the transverse plane. It is made up of four complementary sub-systems, located at different radii from the beam pipe: three

high-resolution semiconductor detectors - the new Insertable B-Layer (IBL), the Pixel Detector and the Semi-Conductor Tracker (SCT) - and a gaseous straw-tube tracking device, the Transition Radiation Tracker (TRT). Each sub-detector consists of cylindrical concentric barrel modules in the central region and disk-shaped end-cap modules in the forward ones, as shown in Fig. 3.6. All these systems are contained in the central superconducting solenoid, which provides a magnetic field of 2 T that bends charged particles within the ID, allowing transverse momentum measurement in the range 0.5–150 GeV within $|\eta| < 2.5$ and electron identification within $|\eta| < 2.0$.

Insertable B-Layer The Insertable B-Layer (IBL) [50, 51] is the tracking detector closest to the beam pipe, at a radial distance of only 3.3 cm from the beam axis. It was installed during the long shutdown between Run-1 and Run-2, due to radiation damage to the innermost Pixel Detector Layer, which was expected to spoil vertex reconstruction precision in Run-2 and 3. It is a very-high-resolution semiconductor pixel detector, with a spatial hit resolution of 8 μm in the $R-\phi$ plane and 40 μm along the z-axis, and covers the barrel region up to $|\eta| < 2.9$. The installation of the IBL improved the track reconstruction, providing more precise vertex and impact parameter measurements. The latter is extremely important for efficient b-jet identification, the so-called b-tagging, to which the detector owes its name.

Pixel Detector The IBL is surrounded by the Pixel Detector, consisting of three concentric cylinders around the beam axis - at average radii of 5.05, 8.85 and 12.2 cm - and three discs in each end-cap between 9 and 15 cm from the interaction point (Fig. 3.7).

The pixel sensors are made of oxygenated 250-μm-thick n-type wafers with read-out pixels on the n^+ implanted side of the detector. Each pixel has a size of 50×400 μm^2 - limited by the size of the corresponding read-out cell on the read-out chip - which determines the intrinsic resolution of the detector. The high granularity of the Pixel system is the crucial element to achieve a robust pattern recognition in the challenging environment of ATLAS, dominated by high occupancy and pile-up. Thanks to more than 80 million read-out channels, a spatial resolution of 10 μm in the $R-\phi$ plane and 115 μm along z is obtained in the barrel, and of 10 μm in the $R-\phi$ plane and 115 μm along z in the endcaps.

Semi-Conductor Tracker The Semi-Conductor Tracker (SCT) is the second system of the Inner Detector. It consists of four concentric barrel layers positioned between radii of 30 cm and 52 cm and nine disks in each end-cap, providing coverage up to $|\eta| < 2.5$. Like the Pixel Detector, it uses semiconducting silicon

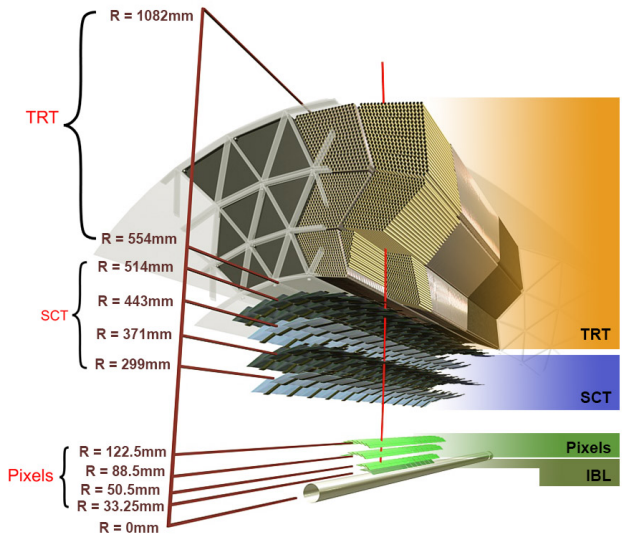


Figure 3.7: Cut-away view of the ATLAS Inner Detector [48, 50, 51].

technology, but its sensors are segmented into strips with 80 μm pitch due to the bigger area to cover. In the end-cap disks, the strip pitch ranges from 56.9 to 90.4 μm to accommodate the more complex geometry.

Each SCT module consists of two layers of strips with a relative rotation (stereo angle) of 40 mrad, allowing for two-dimensional position measurements. The SCT has around 6.3 million readout channels, reaching a resolution of 17 μm in the azimuthal direction and 580 μm along the z-axis.

Transition Radiation Tracker The outermost element of the ID is the Transition Radiation Tracker (TRT). It is composed of straw tubes filled with a Xe-based gas mixture and interleaved with transition radiation material - polypropylene fibres in the barrel and foils in the end-caps. It enables electron identification from the detection of transition radiation photons emitted during the passage of the electrons in the transition radiation material.

The detector consists of a barrel region for $|\eta| \leq 0.8$ and two end-caps sections, made of multi-plane wheels, up to $|\eta| < 2.0$. The barrel has a total of 73 layers of 144 cm long straw tubes, parallel to the beam axis, extending from 56 to 107 cm in the radial direction. The two end-caps consist of 18 wheels of radially oriented straw tubes, positioned between 0.8 m and 2.7 m on the z-axis. The 14 nearest the interaction point cover a radius between 64 and 102 cm, while the last four wheels extend down to a radius of 48 cm to provide coverage of the full pseudorapidity

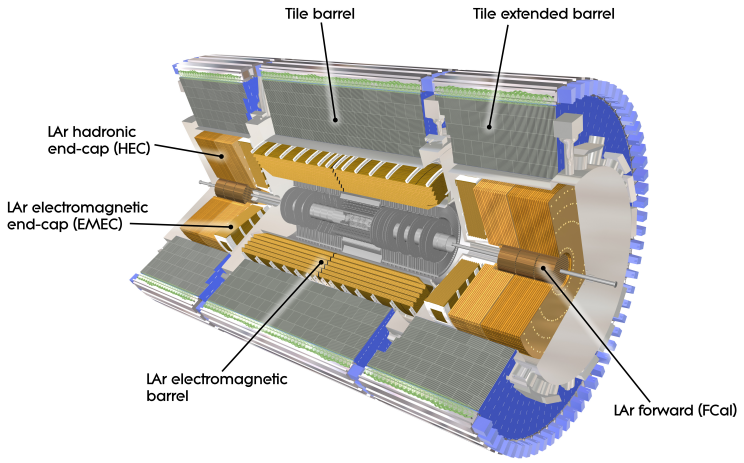


Figure 3.8: Schematic illustration of the calorimeter system of the ATLAS detector [48].

range. The TRT detector provides typically 36 measurements per track (except in the barrel-end-cap transition region, where this number decreases to a minimum of 22 crossed straws) with a spatial resolution of 130 μm, compensating the lower resolution compared to silicon detectors with a large number of hits per track.

3.5 Calorimeter System

The calorimeter system (Fig. 3.8) [48] is the second layer of the ATLAS detector, dedicated to measuring the energy of electrons, photons and hadrons. It makes use of different technologies across the different regions of pseudorapidity, guaranteeing a good energy resolution up to $|\eta| = 4.9$. There are two different types of calorimeters: the Electromagnetic Calorimeter (ECAL), for electrons and photons, is placed right after the solenoid, while the Hadronic Calorimeter (HCAL), for protons, neutrons, pions and kaons, surrounds the ECAL. Moreover, a three-layer Forward Calorimeter (FCAL) is located on both sides at the highest values of $|\eta|$, performing both electromagnetic and hadronic energy measurements.

All calorimeters are sampling calorimeters, which means that they alternate layers of absorbing and active materials. Most of the energy of the particles produced in the collisions is absorbed by high-density metals, while the energy measurement is performed in the active material. In this way, the energy resolution is worsened with respect to a homogeneous device due to fluctuations in the energy release processes, but it is possible to build more compact calorimeters,

allowing also position measurements thanks to the segmentation.

In order to avoid particles escaping, which would spoil the energy resolution, all the calorimeters are designed to contain showers of particles up to the TeV scale. For this reason, the electromagnetic calorimeter thickness is greater than 22 radiation lengths[1] (X_0) for the Barrel and 24 for the End-caps, while the total thickness of electromagnetic and hadronic calorimeter combined amounts to 11 interaction lengths[2] (λ). This efficiently reduces punch-through into the muon system and allows a good resolution for highly energetic jets, assuring a precise measurement of the missing transverse energy, which has a pivotal role in many searches for new physics.

Electromagnetic Calorimeter The Electromagnetic Calorimeter (ECAL) is divided into a barrel part ($|\eta| < 1.475$) and two end-caps ($1.375 < |\eta| < 3.2$). The barrel extends in radial direction from 2.8 m to 4 m and is made of 32 modules. The end-caps are constituted by two coaxial wheels: the outer one covers the region from $|\eta| = 1.375$ to $|\eta| = 2.5$, while the inner wheel extends the coverage up to $|\eta| = 3.2$. In total, they extend over radii from about 0.3 m to 2.1 m, and are segmented into eight wedge-shaped modules each.

The ECAL is composed of different layers of lead, used as absorber, interleaved with liquid argon (LAr) as active material. LAr is chosen because of its intrinsic radiation-resistant properties and its intrinsically linear response. It is held in a liquid state at 89 K thanks to cryostats (one for the barrel and one for each end-caps). The absorber plates and electrodes are arranged in an accordion geometry, which allows a full azimuthal coverage without cracks and a good energy and space resolution.

The ATLAS Electromagnetic Calorimeter is longitudinally segmented. In the region $|\eta| < 2.5$, dedicated to precision physics, the calorimeter is divided into three sections. The strip section, $\sim 6X_0$ thick (upstream material included), is composed of narrow strips with a pitch of ~ 4 mm in the η direction. This high segmentation allows particle identification and precise position measurements in η. The middle section extends up to $24X_0$ and is transversally segmented into square towers of size $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ ($\sim 4 \times 4$ cm^2 at $\eta = 0$). It is designed to contain most of the energy of the showers created by photon and electrons with energy up to 50 GeV. The back section has a granularity of 0.05 in η and a thickness varying between $2X_0$ and $12X_0$, and it is used to estimate the energy leakage in the hadronic cells. Outside the precision physics region ($2.5 < |\eta| < 3.2$) the calorimeter is segmented in only two longitudinal sections

[1]The radiation length is defined as the mean distance over which the energy of a high-energy electron is reduced to $1/e$ of its initial energy by bremsstrahlung.

[2]The interaction length is the mean distance travelled by a hadron before undergoing an inelastic nuclear interaction.

and has a coarser granularity of 0.1×0.1 in $\Delta\eta \times \Delta\phi$. Due to transitions between different detector layouts, some regions of the ECAL have worst performances, in particular for $\eta = 0$, between 1.37 and 1.52, and at $|\eta| = 2.5$.

In the region $0 < |\eta| < 1.8$, a pre-sampler detector is placed inside the solenoid in order to study the shower development before the calorimeter and estimate the energy loss inside the ID. It consists of a liquid argon active layer of thickness 1.1 cm in the barrel region and 0.5 cm in the end-caps.

Hadronic Calorimeter The Hadronic Calorimeter (HCAL) measures hadronic showers up to $|\eta| = 3.2$. It is composed of two different detectors: the Tile Hadronic Calorimeter (TileCal) and the liquid-argon Hadronic End-cap Calorimeter (HEC).

The TileCal uses plastic polystyrene scintillator tiles as active material and steel as absorber. It is divided into three parts: a barrel, covering the region up to $|\eta| < 1.0$, and two extended barrels, which cover the region between $|\eta| > 0.8$ and $|\eta| < 1.7$. The TileCal has an inner radius of 2.28 m and an outer radius of 4.25 m, corresponding approximately to 7.4λ, and is longitudinally segmented in three layers.

The HEC, which covers the region $1.5 < |\eta| < 3.2$, is a copper-LAr sampling calorimeter with a flat-plate design. It consists of two wheels per end-cap with an outer radius of 2.03 m, each of them containing two longitudinal sections, for a total of four layers per end-cap.

Forward Calorimeter The Forward Calorimeter (FCal) covers the $3.1 < |\eta| < 4.9$ region. It is made of three calorimeter modules per end-cap, one electromagnetic and two hadronic. All modules consist of a metal matrix containing the electrode structures in longitudinal channels, with liquid argon as active medium. The electromagnetic layer uses copper as absorber, whereas tungsten is used in the hadronic ones.

The FCal was not designed for precision measurements, but only to provide information for the reconstruction of missing transverse energy and very forward jets. It has a high resistance to radiation, which is extremely important due to the high particle flux in the forward region.

3.6 Muon Spectrometer

The Muon Spectrometer (MS) [48] is the outermost detector of the ATLAS experiment. It is designed to trigger on high-p_T muons and di-muon pairs and to measure the muon transverse momentum with a resolution of 10% at $p_T = 1$ TeV up to $|\eta| < 2.7$.

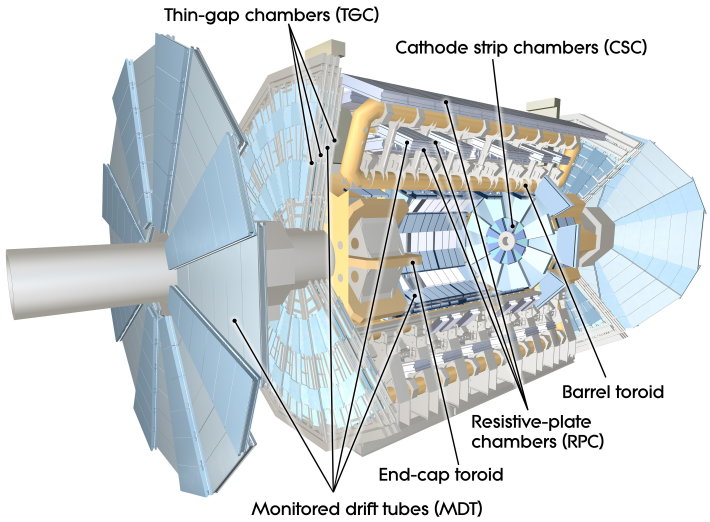


Figure 3.9: Layout of ATLAS Muon Spectrometer [48].

Since muons are minimum ionising particles, their energy loss in the calorimeters is negligible and they reach the MS almost unperturbed.

A magnetic field perpendicular to the muon tracks bends them in the $R - z$ plane and allows momenta measurement. The magnet system is composed of three sets of eight flat superconducting coils: the large barrel toroid in the region $|\eta| < 1.4$ and two end-cap toroids at $1.6 < |\eta| < 2.7$, one on each side. The latter are inserted into the central toroid with a rotation of 22.5°, in order to optimise the bending power in the transition region between the two systems. The system provides a magnetic field of 0.5 T in the barrel and 1 T in the end-caps.

The layout of the MS is presented in Fig. 3.9. The barrel region is formed by three concentric cylinders equipped with Monitor Drift Tube chambers (MDTs) at radii of about 5, 8 and 10 meters. MDT chambers consist of six layers of drift tubes in the middle and outer stations and eight layers in the inner one. The middle and the outer stations also include Resistive Plate Chamber (RPC) detectors for triggering purpose. In the end-cap regions, the MS consists of three independent wheels per side, at a distance of 7.5, 14 and 22.5 meters from the interaction point. The first wheel, called Small Wheel, makes use of Cathode Strip Chambers (CSCs) close to the beam pipe and MDTs in the outer part for tracking purpose, while Thin Gap Chambers (TGCs) provide trigger signals. The two outermost wheels are instead equipped with MDTs only for track reconstruction, with the central wheel, known as Big Wheel, including also two (TGCs) doublets and one triplet for the trigger.

Monitor Drift Tube Chambers Monitor Drift Tubes (MDTs) are 3-centimeter cylindrical aluminum tubes, with a 30 μm tungsten-rhenium wire at a high electric potential placed in the centre. They are filled with Ar/CO_2 (93:7) gas mixture and are operated at 3 bar overpressure. The length of the tubes is between 0.9 m and 6.2 m. Three or four layers of tubes form a multi-layer and two multi-layers separated by aluminium spacers form a chamber. The average resolution of each tube is 80 μm, providing a resolution of 35 μm per chamber. MDTs can be operated up to a rate of 150 Hz/cm^2, making necessary the use of CSCs in the region $|\eta| > 2$ of the innermost end-cap layer, where this rate is exceeded.

Cathode Strip Chambers Cathode Strip Chambers (CSCs) are multi-wire proportional chambers made of four planes, which provide independent measurements of η and ϕ. Each plane consists of a layer of anode wires placed between two layers of cathode strips, one with strips parallel to the wires and the other one perpendicular. The strip segmentation in η is finer than in ϕ, giving a resolution of about 60 μm in the precise coordinate and 5 mm in the other one. An $Ar/CO_2/CF_4$ gas mixture (30:50:20) is employed in the chambers.

Resistive Plate Chambers Resistive Plate Chambers (RPCs) are detectors designed to provide trigger signal. A single module is composed of two parallel plastic resistive plates, with a 2 mm gap filled with a freon-based gas mixture. The signal is read out via capacitive coupling to metallic strips, which are placed on the outer faces of the resistive plates.

A RPC trigger chamber is composed of two detectors, next to each others, called units. Each of these units consists of two independent detector layers, which are each read out by two perpendicular sets of strips.

Thin Gap Chambers Thin Gap Chambers (TGCs) are multi-wire proportional chambers with a wire-to-cathode distance smaller than the one between the wires (1.4 mm vs 1.8 mm). The signal read from the wires provides a position resolution in the η coordinate of a few mm, while copper strips provide a measurement of the ϕ coordinate with a similar precision. The chambers are operated in quasi-saturated mode, leading to a typical signal rising time of 5 ns. The gas used is a mixture of 45% n-Pentane and 55% CO_2.

3.7 Trigger

Taking into account a bunch-spacing of 25 ns and an average of 34 collisions per bunch crossing, the collision rate inside the ATLAS detector is around 1.5 GHz. Due to the limitations in data storage capacity and rates, only a small fraction

of this data can be recorded. The ATLAS Trigger System has been designed to perform run-time event selection, recognising and saving only events interesting for the analyses. It has to reduce the number of events collected by the experiment down to approximately 1 kHz.

The trigger system of Run-2 [52] consists of two different trigger levels: a hardware-based level trigger (L1) and a software-based high-level trigger (HLT).

The L1 trigger makes use of the measurements from the calorimeters and muon spectrometer to determine Regions-of-Interest (RoIs) in the detector, searching for high energy leptons, photons, and jets. Electrons and photons are triggered using energy deposits in the EM calorimeter, while jet candidates are constructed from HCAL coarse calorimeter towers. Muons are instead triggered based on a coincidence of hits among multiple layers of the Muon Spectrometer. The L1 trigger reduces the event rate down to 100 kHz.

The HLT consists of a Level 2 (L2) trigger and an event filter (EF), with an output rate of 1 kHz. The L2 trigger is similar to the L1, but it uses a finer granularity and includes measurements from the ID for the RoI. The EF, instead, fully reconstructs offline the event, using the ATLAS Athena framework. The majority of events that pass the EF selection requirement are directly written to the main analysis stream, while events requiring longer processing time are saved to a debug stream for later reprocessing.

Chapter 4

Event Simulation

In order to compare theoretical prediction with the data collected by ATLAS, it is essential to have simulated events for both Standard Model processes and DM production. The goal of these simulated events is to describe, as accurately as possible, the characteristics of the different processes and how they appear in the detector. The simulation includes several steps: first, the generation of the final state particles produced during the *pp* collisions, then their passage through the detector and finally the reconstruction, performed using the same algorithms applied to recorded data.

4.1 Proton-Proton Collisions Overview

Since protons are composite objects, the description of what happens in proton-proton collision is quite complicated [54]. Usually, there are mainly two partons that interact in a hard scattering process, while the others give rise to softer interactions. This is illustrated in Fig. 4.1, where the incoming protons are indicated by a green ellipse with three green lines that represent the valence quarks, while the hard process is depicted by the big red circle. Both before and after the hard scattering interaction, partons may produce additional QCD radiation which is called, respectively, initial state radiation (ISR) and final state radiation (FSR). The products of the hard scattering (represented by the small red circles) then keep emitting additional QCD radiation and a parton shower (PS) evolves, as shown in red. The partons resulting from the parton shower are subsequently combined into colourless hadrons thanks to a process called hadronisation, indicated by the light green blobs. Finally, these hadrons decay into stable particles. The lower half of Fig. 4.1 illustrates also a secondary interaction happening between the partons not involved in the hard scattering process (purple ellipse). Also in this case a parton shower is produced (purple lines), which then hadro-

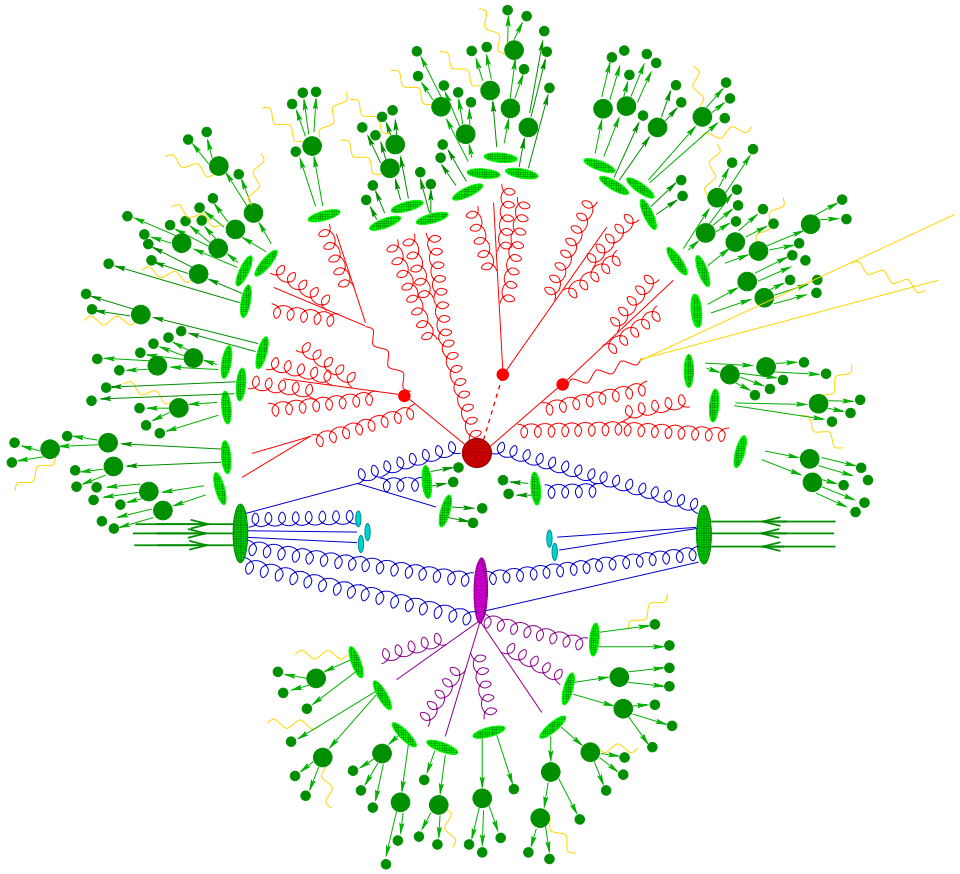


Figure 4.1: Representation of a proton-proton collision event, containing all steps of the event generation chain [53] See text for details.

nises and decays into stable particles. Secondary interactions are usually softer than the primary one and, together with the beam remnants (blue ellipses), constitutes the underlying event. The yellow lines in Fig. 4.1 remind than in any stage electromagnetic radiation can be emitted by charged particles.

The hard-scattering interaction between partons, which happen at high energy, can be calculated in perturbative Quantum Chromodynamics (QCD), while the evolution of partons into stable hadrons cannot, since it takes place at lower energy. For this reason, it is essential for simulation of pp collisions to adopt different approaches to simulate processes happening at different energy scales [55].

4.2 Cross-Section Calculation

Event simulation starts with the computation of the cross-section. The inclusive cross-section for a $pp \to X$ process can be connected with the cross-section for the hard scattering process $\hat{\sigma}_{ab\to X}$ thanks to the factorisation theorem [56]:

$$\sigma_{pp\to X} = \sum_{a,b} \int dx_a dx_b f_a(x_a, \mu_F^2) f_b(x_b, \mu_F^2) \hat{\sigma}_{ab\to X}(x_a, p_a, x_b, p_b, \mu_R^2, \mu_F^2), \quad (4.1)$$

where the sum runs over the partons types a, b that can initiate the process. The parton distribution functions (PDFs), $f_i(x_i, \mu_F^2)$, describe the momentum distribution of partons within a proton. In particular, PDFs give the probability density of finding a parton of type i at energy μ_F with a fraction of the proton's momentum x_i. PDFs cannot be predicted from theory calculations due to the non-perturbative nature of QCD at low energy. They are instead extrapolated fitting experimental data coming from hadron colliders and deep inelastic scattering experiments. The measurements are only feasible for certain energy scales and needs to be extrapolated to the scale of interest. This is made possible thanks to the DGLAP evolution equations [57, 58, 59], which describe the energy dependence of the PDFs. Several collaborations constantly work to improve the PDF fits using the most recent data, like NNPDF [60], CTEQ [61] and MSTW [62].

In Eq. 4.1, μ_R is the renormalisation scale for the strong running coupling constant, while μ_F is the factorisation scale, which marks the boundary between the kinematic region where emissions are treated as part of the hard scatter and the region where emissions are included in the PDF. The real cross-section should be independent of the choices of μ_R and μ_F, but this is true only if all orders in perturbation theory are considered. Otherwise, they have to be fixed to reasonable values, typically a scale characteristic for the process. Usually the two scales are assumed equal.

The cross-section for the hard scattering process $\hat{\sigma}_{ab\to X}$, where a, b are the

interacting partons, is derived as a consequence of Fermi's Golden Rule, which states that the transition rate of a process is given by the absolute square of the quantum mechanical amplitude $\mathcal{M}$ - also called Matrix Element - integrated over the available phase space:

$$\hat{\sigma}_{ab\to X} = \frac{S}{4\sqrt{(p_a \cdot p_b)^2 - (m_a m_b)^2}} \times \tag{4.2}$$

$$\int |\mathcal{M}|^2 (2\pi)^4 \delta^4(p_a + p_b - p_1 ... - p_n) \times \prod_{i=1}^{n} 2\pi\delta(p_i^2 - m_i^2)\Theta(p_i^0)\frac{d^4 p_i}{(2\pi)^4}, \tag{4.3}$$

where $1...n$ indicate the outgoing particles and S is a factor correcting for double counting if identical particles are present in the final state. The δ- and Θ-functions ensure four-momentum conservation and that the outgoing particles are on mass shell and with positive energy. The matrix element $\mathcal{M}$ can be calculated using the so-called Feynman rules. These rules are illustrated by Feynman diagrams, which describe the process using lines for each particle and vertices for the couplings between them. Each diagram corresponds to a specific term in the amplitude for the depicted process and their sum gives the total amplitude, taking also into account interference effects. The matrix element may be calculated at different orders in perturbation theory. In particular, if only diagrams without loops or ISR/FSR are considered, the calculation is called Leading Order (LO), while if also diagrams with a single loop/emission are included it is referred to as Next to Leading Order (NLO).

4.3 Event Simulation

As introduced in Sec. 4.1, a collision event includes several processes happening at different energies, which can be simulated separately. All processes are simulated through Monte Carlo (MC) techniques, which are based on repeated random sampling process-specific variables with known probability distributions in order to obtain a sample of predictions for that process.

Matrix Element The matrix element (ME) of the hard-scattering process can be calculated in perturbation theory to some fixed order using Feynman rules, as explained in the previous section. The PDF set to be used is chosen within the available ones via the LHAPDF library. It is recommended to use a PDF of the same perturbative order as the matrix element calculation. Usually, renormalisation and factorisation scales are fixed to a scale characteristic for the process, for example the mass of the resonant particle in a s-channel resonance.

Parton Shower Parton shower (PS) algorithms describe the evolution from the scale of the hard process down to hadronisation scale at which the partons become confined in hadrons, typically $\sim$ 1 GeV. PS consists of the initial-state parton shower or radiation and the final-state parton shower or radiation, describing respectively emissions off the ingoing and outgoing partons, accounting for higher order effects in perturbation theory not covered by the fixed order matrix element calculation. The ME calculation cannot account for higher order effects due to the fact that collinear and soft emissions give rise to divergencies in it, which can be factorised and dealt with separately by the PS. Final-state parton shower algorithms make use of a Markov chain, which adds probabilistically one emission at a time, where the probability for the parton to evolve from a higher to a lower scale without radiating a gluon or splitting is given by the so-called Sudakov form factor. The description of the initial state parton shower follows instead a backward evolution, dressing the incoming partons with further radiation, in order to avoid wasting computing time simulating showers that ends with partons unsuitable for the hard scattering due to low momentum or wrong flavour. Since the evolution starts with the partons in the initial state of the hard scattering and evolves them back to the constituent partons in the proton, PDFs have to be taken into account in this case. A Markov chain based on Sudakov form factor including PDFs and using DGLAP equations is then employed. PS algorithms are applied iteratively until all particles reach the hadronisation scale, below which the perturbative approach no longer applies.

Matching Matrix element calculation and parton shower describe different types of processes: the ME describes processes with a given number of external partons, accounting correctly for hard emissions, while PS describes events with a variable number of additional particles, giving a correct treatment of soft and collinear emissions. It is then fundamental to combine these two approaches in a coherent way, avoiding possible double-counting of contributions and ensuring a smooth transition between the two. The combination procedure is called matching. Different matching schemes exist, like for example CKKW [63], CKKW-L [64] and MLM [65], usually based on the definition of some matching scale that separates the regimes treated by the ME and the PS.

Hadronisation Due to QCD confinement, the partons produced in the PS need to be combined into colourless states, taking into account also the colour connections with the beam remnants. Since hadronisation happens at energies beyond the perturbativity of QCD, it can be described only through phenomenological models. These models introduce more free parameters than the previous steps, but they can be tuned on one dataset and then be implemented in other simu-

lations. Mainly two hadronisation models are used: the cluster model [66, 67], in which the partons in a shower cluster together into colourless groups, and the Lund [68, 69] string model, which treats quark-antiquark pairs as the ends of a string, with a potential energy between the quarks that is proportional to its length, such that when the distance increases it is energetically-favourable for the string to break and form an additional quark-antiquark pair. Some of the hadrons produced in the hadronisation process may decay before reaching the detector. These decays are usually simulated implementing in the simulation the decay tables from the Particle Data Group (PDG) [11].

Underlying Event During proton collisions, also particles not associated with the leading hard scattering process may be produced, forming the underlying event (UE). One of the most used model employed to simulate these additional particles is the multiple parton interaction model, which assumes additional interactions taking place between partons not participating in the hard scattering process. The number of these additional interactions is usually small thanks to the fact that most interactions would be too soft to resolve the hadrons internal substructure. All the simulation steps presented above (ME, PS and hadronisation) also apply to additional interactions, and are performed in a similar way. Different MC tunings to data are used to improve the modelling of the UE, like the A14 tune [70] and the A3 tune [71].

Event Filter During simulation, the generated events can be filtered such that only events fulfilling certain properties are saved. These cuts can be exploited to obtain enough statistics in different regions of phase space without consuming too much computing power.

4.4 Monte Carlo Generators

The software tools used to produce simulated events are generally referred to as Monte Carlo (MC) generators. Generator can be divided in two categories: general purpose generators, which can carry out all the steps of event simulation, and generators specialised on a single simulation step, usually the matrix element calculation. There are different generators available for LHC physics, and this section will introduce the most relevant ones for the analyses presented in the following Chapters.

4.4.1 Matrix Element Generators

Madgraph5_aMC@NLO MADGRAPH5_AMC@NLO [72] is a generator that automatically computes matrix elements at LO and NLO. The NLO calculation uses the MC@NLO method to match the matrix element calculation with the parton shower.

Powheg POWHEG [73, 74, 75] is a NLO parton-level event generator that provides matrix element calculation in perturbative QCD for the majority of SM processes, using the POWHEG method to match the matrix element calculation with the parton shower.

4.4.2 General Purpose Generators

Pythia PYTHIA [76] is a multi-purpose MC generator with more than 200 hard-coded LO subprocesses, lacking of automated code generation for new processes. It is used mainly interfaced with a ME generator for the parton shower, hadronisation, and the underlying event modeling. For the hadronisation, it uses the Lund string model.

Sherpa SHERPA [77] is a general-purpose generator, with automated generation of LO matrix elements. It can be used to model all the steps of the simulation of all SM processes. For hadronisation, it makes use of the cluster model.

Herwig++ HERWIG++ [78, 79] is a multi-purpose MC generator with a library of hard-coded LO process, unable to automatic calculate matrix element for new processes. For this reason, it is usually interfaced with ME generators for the other steps of the simulation. Its PS includes colour coherence effects, with special emphasis on the correct description of radiation from heavy particles. It models the hadronisation based on the cluster model.

4.5 ATLAS Detector Simulation

In order to compare generated event samples with recorded data, a detailed simulation of the detector is necessary. The simulation software of the ATLAS detector is based on the GEANT4 framework [80], which simulates the passage of particles through matter. In GEANT4, the geometry structure is based on three different type of volumes: solids, basic shapes without a position in the detector, logical volumes, solids with additional properties (e.g. material), and physical volumes, individual placements of logical volumes.

Any "stable" particle produced by MC generators which is expected to propagate through a part of the detector is fed into the full ATLAS detector simulation [81]. The energies deposited in the sensitive portions of the detector are recorded as hits, saving the total energy deposition, position, and time. These hits are then digitalised into electronic signals, which are used as input for the reconstruction algorithm used also for recorded data (see next Chapter).

The detector simulation is highly CPU intensive. For example, the full detector simulation, called FullSim, of a single $t\bar{t}$ event requires a time of the order of 10 minutes. For this reason, a faster and less refined simulation, known as AtlFast-II (AFII) [81], is available. The AFII simulation parametrises the description of the particle showers in the calorimeters, which accounts for about 75% of the CPU time spent in the FullSim, making the simulation ten times faster. Fully simulated samples are however usually preferred for the main background processes because they provide higher precision. For phenomenological studies, the detector simulation can be also completely parameterised, applying only some smearing and efficiency functions to generator outputs. Studies performed in this way are usually referred to as TRUTH studies. It may also be useful to perform some studies directly on the event generator output, without applying any simulation of the detector. These studies can be carried out both at "parton level", if they use only the matrix element calculation output, or at "particle level", if they make use of the final stable particles.

Chapter 5

Event Reconstruction

Particles produced in the pp collisions inside ATLAS interact with the different detectors of the experiment, giving rise to electronic signals. These signals have to be processed, grouped and combined to reconstruct the physics objects of each event, i.e. leptons, jets, photons and $E_{\mathrm{T}}^{\mathrm{miss}}$. The aim of this Chapter is to describe the physical objects used in the analyses presented in this thesis, giving an overview of the algorithms used for their definition and their identification criteria.

5.1 Tracks and Primary Vertex

The reconstruction of charged particles trajectories [82, 83], usually called tracks, is performed using the signals from the Inner Detector. Hits recorded by the Pixel Detector are used as seeds for a reconstruction algorithm, which combines them first with the information coming from the SCT and then with TRT measurements, following either a Gaussian sum filter or a global χ^2 fit. If no hit is found inside the Pixel Detector, the reconstruction algorithm can also be applied backwards, in order to reconstruct also tracks coming from photon conversion. Tracks with $p_T < 400$ MeV and $|\eta| < 2.5$ are discarded. Fake tracks coming from instrumental effects or pile-up are rejected requiring a hit in the IBL or in the innermost pixel layer and no missing hits in intermediate layers.

Reconstructed tracks are used as input in the vertex identification algorithm [84, 85]. This algorithm uses an iterative χ^2 procedure, starting with a seed position selected using the mode of z-coordinates of tracks at their points of closest approach to the reconstructed beam spot. This seed and all reconstructed tracks are then used to estimate the best vertex position with a recursive fit, which in each iteration down-weights less compatible tracks and recompute the vertex position. After the vertex position is found, tracks incompatible with it are

used to find another vertex, repeating this procedure until all tracks are assigned. The position of the different vertices is then used to better reconstruct the tracks assigned to them. Since interesting physics processes usually have multiple high-p_T tracks, the vertex with the highest sum of the squared-tracks-p_T is chosen as the primary one.

5.2 Electrons

Electrons are reconstructed in the region with $|\eta| < 2.47$ using information from the ID and calorimeters, excluding the transition region ($1.37 < |\eta| < 1.52$) between the barrel and the end-cap EM calorimeter. An electron usually has 12 silicon hits: one in the IBL pixel layer, three in the pixel detector and eight in the double-sided silicon strips. In addition, it has approximately 35 TRT hits. After the ID, the electron deposits the majority of its energy in the electromagnetic calorimeter, leaving only about 2% to reach the hadronic calorimeter. An electron is then defined as an object consisting of a cluster built from energy deposits in the calorimeters and a matched track.

The first step in the reconstruction of an electron consists in building the clusters from the calorimeter energy deposits. This was performed in two different way during Run-2.

For early Run-2 searches [86], like the ones presented in Chapter 7, electron reconstruction uses fixed-size clusters of size $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$, called towers, which are assigned as energy the sum of the energy collected in the three ECAL layers and in the presampler. Cluster candidates are seeded from localised energy deposits using a sliding-window algorithm of size 3×5 towers. If the transverse energy within this window is above 2.5 GeV, the region is marked as a seed. Once the seed clusters are selected, they are attempted to be matched to ID tracks, extrapolating the track to the second layer of the ECAL, using either the measured track momentum or rescaling the latter to match the cluster energy. The momentum rescaling is used to improve the matching for electrons that loose a significant amount of energy in the ID due to bremsstrahlung radiation. A track is considered to match a cluster if its extrapolation satisfies $|\Delta\eta| < 0.05$ and $0.10 < q(\phi_{track} - \phi_{cluster}) < 0.05$ with either momentum magnitude, where q is the charge reconstructed from the track. If matching is possible with a track coming from the primary vertex, then the cluster is tagged as an electron. After that, reconstructed electron clusters are built starting from the seed clusters using an extended window of size 3×7 in the barrel region and 5×5 in the endcap by simply expanding the cluster size on either side of the original seed cluster.

For full-statistics Run-2 searches [87], included the one presented in Chapter 8, electron reconstruction makes use of dynamic, variable-size clusters, called

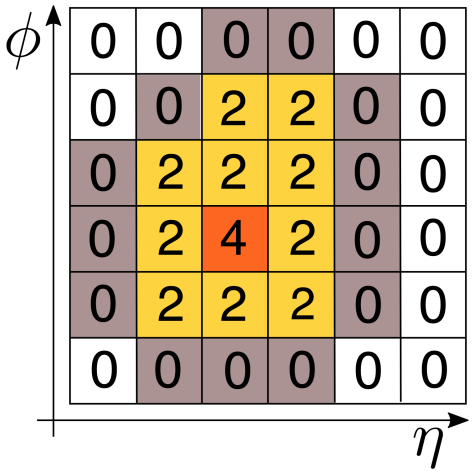


Figure 5.1: Simplified representation of a topo-cluster. The numbers inside the cells indicate the signal-to-noise ratio threshold.

superclusters, which are needed to recover energy loss due to bremsstrahlung photons. The topo-cluster reconstruction algorithm [88] starts by forming proto-clusters in the electromagnetic and hadronic calorimeters, using as seeds cells with a signal-to-noise significance $|\varsigma_{cell}^{EM}| \geq 4$, where

$$\varsigma_{cell}^{EM} = \frac{E_{cell}^{EM}}{\sigma_{noise,cell}^{EM}}, \tag{5.1}$$

and the expected cell noise includes the electronic noise and an estimate of the pile-up noise for Run-2. Neighbouring cells with significance $|\varsigma_{cell}^{EM}| \geq 2$ are then added to the proto-cluster, followed by a ring of guard cells added independently of their energy. This set of thresholds is commonly known as "4-2-0" topo-cluster reconstruction, and is represented schematically in Fig. 5.1. Proto-clusters coming from pile-up are rejected with a 60% efficiency asking the ratio of the energy deposited in the ECAL with respect to the total cluster energy to be greater than 0.5. This cut does not affect the efficiency of true electron topo-clusters selection because, as explained previously, electrons deposit the majority of their energy in the ECAL. The topo-clusters selected in this way are matched to ID tracks in the same way used for early Run-2 searches and, if the matching is successful, superclusters construction starts. The previously selected topo-clusters are tested as seed candidates for the superclusters, following a descending E_T order. Topo-clusters near the seed candidates are then identified as satellite cluster candidates, which may form from electrons due to bremsstrahlung radiation, if they fall within a window of $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ around the seed cluster barycentre, or if

they are within a window $\Delta\eta \times \Delta\phi = 0.125 \times 0.300$ and their "best-matched" track is the same of the seed cluster. The union of seed clusters with their associate satellite clusters forms the superclusters, which, after an initial energy calibration and position correction, are finally matched with tracks as explained before, forming a candidate electron.

After the analysis-level candidate electrons are finally constructed, identification algorithms are used to quantify the quality of reconstruction and remove potential misidentified particles. A set of calorimeter- and track-based variables [87], describing the shape of the electromagnetic shower, track-cluster matching and the quality of the track, is employed to discriminate prompt electrons coming from the *pp* hard interaction from other particles with similar properties, such as charged pions or e^+e^- pairs coming from photon conversion. These variables are used as input to a likelihood (LH) discriminant, developed to reject non-prompt electrons. Four identification operating points are available to identify prompt electrons, all using the same LH discriminant but with increasing thresholds: VeryLoose, Loose, Medium, and Tight [87].

To further suppress the contribution from non-signal electrons, additional requests on the isolation of the electron are used. In fact, in the case of prompt electrons there is little activity surrounding the particle, contrary to what happens in semileptonic decays of heavy-flavour quarks. Two sets of isolation variables are defined [87], track-based and calorimeter-based, which evaluate how much activity is contained within a cone around the direction of the electron. Using these variables, nine electron isolation working points are defined [87], providing different choices of efficiency and background rejection.

The efficiencies of reconstruction, identification and isolation algorithms are all evaluated using the tag-and-probe method, which is a data-driven technique that exploits the Z and J/ψ resonances as a source of electron-positron pairs [87]. In particular, it selects events with a Z or J/ψ candidate using tighter requirements on the so-called "tag" electron and looser requirements on the "probe" electron, and calculates the efficiency of a given selection as the fraction of probe electrons which pass it. In Fig. 5.2 the performance of the identification algorithms are shown for the different working points, underlining the dependency of the efficiency on η and E_T.

Due to effects like energy losses in passive materials, electromagnetic shower leakages and fluctuations in the energy deposited in the calorimeters, the detector simulation is expected to only approximately reproduce the real efficiencies of electron reconstruction. For this reason, the simulated samples are corrected to reproduce the measured data efficiencies making use of scale factors [87], given by the ratio of the efficiencies estimated in both data and simulation for well know SM processes such as $Z \to e^+e^-$ and $J/\psi \to e^+e^-$. The scale factors are provided

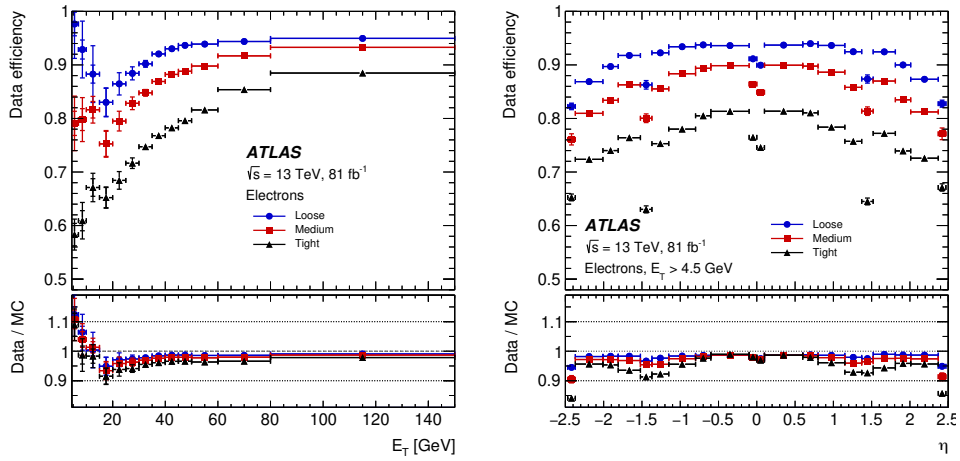


Figure 5.2: Electron identification efficiency measured in $Z \to e^+e^-$ events as a function of E_T (left) and η (right) for different identification working points. The bottom panels show the data-to-simulation ratios [87].

as function of both η and E_T and they are usually close to unity. The agreement between data and Monte Carlo simulations after the calibration is presented in Fig. 5.3.

5.3 Muons

Muons are reconstructed from information provided by the Muon Spectrometer, the Inner Detector and, in minor part, from the calorimeters. Each of the three MDT layers provides from 6 to 8 hits for muons within $|\eta| < 2.7$. Signals from each layer are combined to form local track segments, which are then combined to form an overall MS track. The inner detector provides an independent measurement of the muon trajectory near the interaction point for muons within $|\eta| < 2.5$. A typical muon has one IBL hit, three pixel hits, eight SCT hits and, on average, 30 TRT hits.

Four different "types" of muons are available, depending on the reconstruction algorithm used [89]:

- Combined (CB) muons, based on a fully reconstructed track in the MS matched with an ID one

- Segment-Tagged (ST) muons, which are a combination of an ID track with a Muon Spectrometer segment. They allow to recover low momentum muons

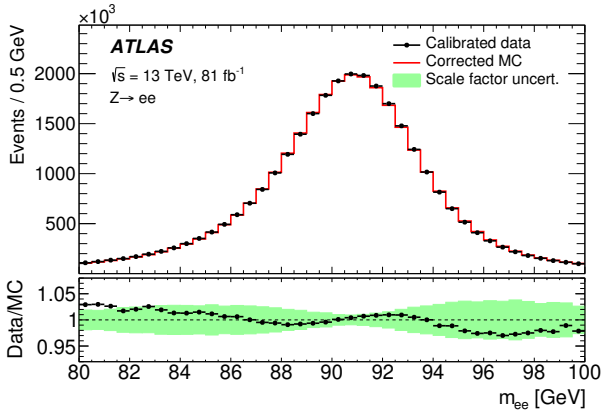


Figure 5.3: Comparison between data and Monte Carlo simulation after calibration for the invariant mass distribution of the two electrons coming from $Z \to e^+e^-$ event candidates [87].

that do not have enough momentum to traverse the whole spectrometer due to energy loss in the calorimeter

- Calorimeter-Tagged (CaloTag) muons, which are a combination of an ID track and a calorimeter energy deposit compatible with a minimum ionizing particle, allowing to recover efficiency at $\eta \approx 0$, uncovered by the Muon Spectrometer

- Extrapolated muons (ME), reconstructed based only on MS measurements with a loose requirement on compatibility with an origin from the primary vertex.

Muon identification is necessary in order to distinguish prompt muons from background events, in particular muons from pion and kaon decays. For CB muons, which will be used in this thesis, specific requirements on the following set of variables are used (cut-based identification):

- $|q/p|$ significance, the absolute value of the difference between the measured charge-to-momentum ratio in the ID and MS, divided by the quadratic sum of the uncertainties on the measurements

- ρ', the absolute value of the difference between the ratio of the transverse momentum measurements in the ID and MS divided by the p_T of the combined track

- the normalized χ^2 of the combined track fit.

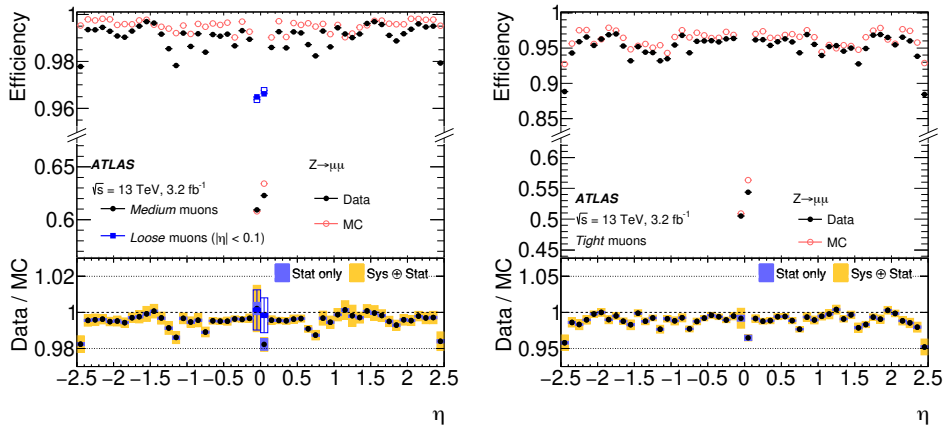


Figure 5.4: Muon identification efficiency measured in $Z \to \mu^+\mu^-$ events as a function of and η for different identification working points. The bottom panels show the data-to-simulation ratios [89].

In addition, a minimum number of hits in the ID and MS is requested in order to guarantee a robust momentum measurement. Four muon identification working points are available: Medium, Loose, Tight, and High-p_T [89].

To improve the rejection of muons from heavy-flavor hadron semi-leptonic decays, additional isolation requirements are imposed similar to the electrons, using both track-based and calorimeter-based variables. Seven isolation working points are provided [89], each optimized for different physics analyses.

Reconstruction, isolation, and identification efficiencies are measured also for muons in data and simulation using the tag-and-probe method on $Z \to \mu^+\mu^-$ and $J/\psi \to \mu^+\mu^-$ events [89]. Muon reconstruction efficiency for three different identification working points is shown in Fig. 5.4. Scale factors to compensate for differences between data and simulation performances are derived in the same way of electrons, analysing $Z \to \mu^+\mu^-$ and $J/\psi \to \mu^+\mu^-$ processes. The comparison of data with simulation before and after the application of scale factors is shown in Fig. 5.5.

5.4 Jets

Quarks and gluons produced in the pp collisions are coloured particles which, due to QCD colour confinement, cannot exist as free states. For this reason each parton goes through the hadronisation process, which results in a collimated spray of hadrons with null colour charge, approximately in the direction of the original

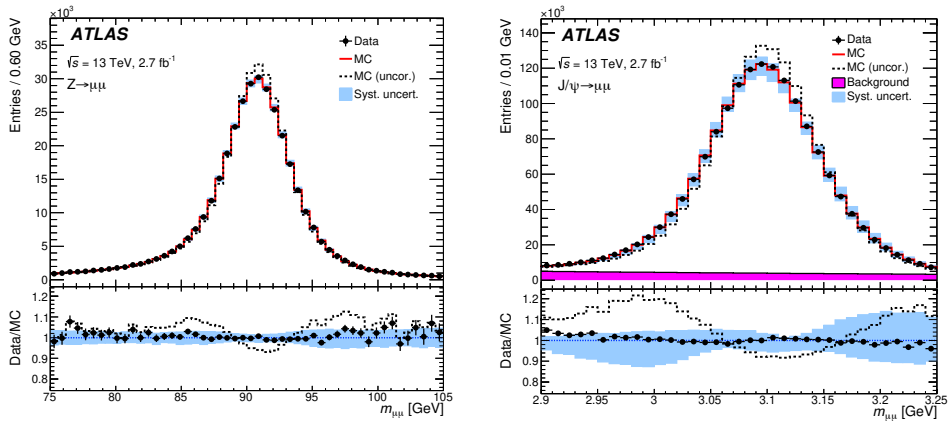


Figure 5.5: Comparison between data and Monte Carlo simulation after calibration for the invariant mass distribution of the two muons coming from $Z \to \mu^+\mu^-$ (left) and $J/\psi \to \mu^+\mu^-$ (right) event candidates [89].

parton. This highly-collimated collection of particles is called jet, and it is what is observed in the detector. Jets are reconstructed from the energy deposits in the calorimeters cells, aiming to produce physics objects with kinematics and characteristics as close as possible to those of the initial partons.

Jet reconstruction starts with the identification of the clusters of energy deposited in the calorimeter by the particles in the jet. A "4-2-0" topo-cluster reconstruction algorithm [88] similar to the one illustrated for electron reconstruction for the full Run-2 statistics (Sec. 5.2) is used for this purpose. After the initial clusters are built, local energy maxima are identified and if more than one maximum is found in a cluster it is split. At this point, for each cluster a four-vector with zero mass is built using the direction of the cluster with respect to the centre of the detector and the sum of the energy in its cells.

After all clusters have been associated with a four-vector, a complex algorithm is exploited to gather them in jets. In ATLAS, jet candidates are reconstructed using the anti-kT jet clustering algorithm described in [90]. The first step of this algorithm calculates the distance between each pair of clusters (d_{ij}) within a distance R and the distance between each cluster and the beam (d_{iB}) as follows:

$$\begin{aligned} d_{ij} &= \min(p_{Ti}^{2n}, p_{Tj}^{2n}) \frac{\Delta R_{ij}^2}{R^2}, \\ d_{iB} &= p_{Ti}^{2n}, \end{aligned} \tag{5.2}$$

where $\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, with y_i being the rapidity of the four-momentum associated to the ith cluster, ϕ_i its azimuth angle and p_T its transverse

momentum. n is chosen to be -1 for the anti-kT algorithm. All searches presented in the present work make use of jets reconstructed fixing $R = 0.4$ [91].

The algorithm proceeds then by identifying the smallest of the distances d_{ij} and d_{iB}. If it is a distance between two clusters d_{ij}, the two clusters four-momenta are replaced by a single four-vector combination and the distances of this new object are calculated. Otherwise, if the smallest distance is a d_{iB}, the ith cluster four-momentum is classified as a final jet and removed from the list of candidates. This procedure is then repeated until all clusters are classified as jets. The choice of $n = -1$ assures that all softer clusters are merged with the harder ones in order of their closeness in terms of ΔR.

Once jets are clustered, they need to undergo an energy calibration [92], which relates the jet energy measured within the ATLAS calorimeter to the true energy of the corresponding initial parton. Calibration accounts for a number of detector effects that affect the jet energy measurement. The Jet Energy Scale (JES) calibration corrects for the non-compensating behaviour of the ATLAS calorimeters, the energy losses in inactive material, noise and particle reconstruction efficiency. This calibration is derived from the ratio of the true jet energy and the energy reconstructed in Monte Carlo simulations, and it is a function of the reconstructed jet p_T and η.

Also the origin of the jets has to be corrected [93], making it pointing to the hard-scatter rather than to the geometrical center of the detector as it is initially constructed. This correction has a small effect on the jet p_T, but it improves the angular resolution.

Finally, jets arising from pile-up events are suppressed using the Jet Vertex Tagger (JVT) discriminant [94]. The JVT uses information about the fraction of the p_T of a given jet that originate from the primary vertex, providing as output the likelihood that the jet is from the hard scatter event. The efficiency and the corresponding scale factors for JVT are derived from data and MC using $Z \to \mu^+\mu^-$ events with an additional hard-scatter jet.

5.5 b-jets

The identification of jets originating from bottom quark fragmentation (b-jets) is called b-tagging. It takes advantage of the relatively long lifetime of b-quark (of the order of 1.5 ps) which, thanks to relativistic effects, results in an observable displacement of the decay vertex of the b-quark from the primary vertex (2-3 mm for hadrons with $p_T \approx 30$ GeV). This peculiarity is exploited by different types of b-jet identification algorithms:

- the IP3D algorithm (Impact Parameter based algorithm) [95] uses both the transverse and longitudinal impact parameter significances, considered in a

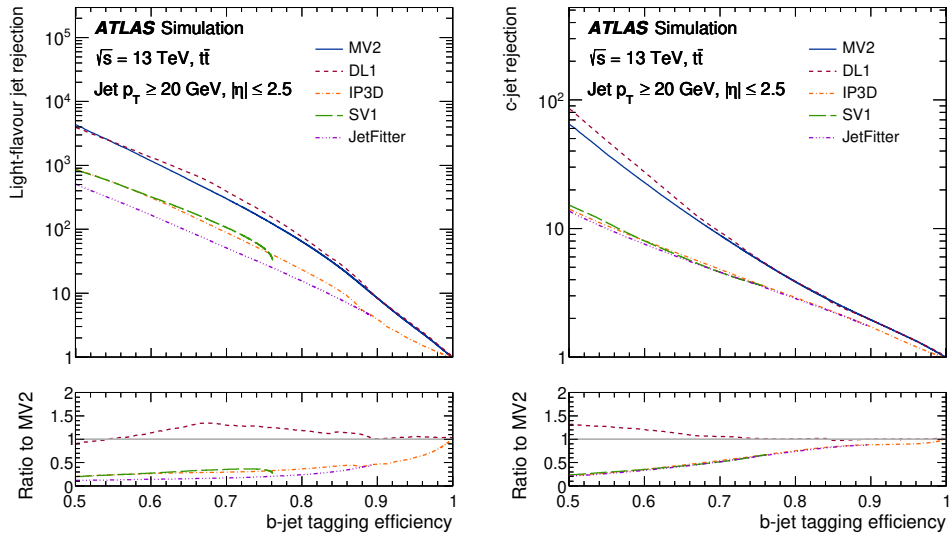


Figure 5.6: Light-flavour jet (left) and c-jet (right) rejections versus the b-jet tagging efficiency for different b-tagging algorithms, evaluated on simulated $t\bar{t}$ events [99].

two-dimensional template to account for their correlation

- the SV1 algorithm (Secondary Vertex Finding algorithm) [96] tries to reconstruct the displaced secondary vertex of the jet testing all tracks for a two-track vertex hypotesis
- the JetFitter algorithm (Decay Chain Multi-Vertex algorithm) [97] takes advantage of the topological structure of b- and c-hadron decays inside the jet to try to reconstruct the decay chain
- the MV2c10 algorithm (Multivariate algorithm) [98] combines the output of the above algorithms in a Boosted Decision Tree (BDT) to achieve a better discriminant power. The output of the BDT is the MV2c10 score, which is trained using b-jets as signal and a mixture of 93% light-flavor jets and 7% c-jets as background
- the DL1 algorithm (Deep Learning algorithm) [98] uses as input the results of the first three algorithms above and feeds them in a deep feed-forward neural network that provides an improved discrimination.

The performance of a b-tagging algorithm is given by its ability to correctly tag jets coming from real b-quarks (b-tagging efficiency) compared to the probability

of mistakenly identifying as a b-jet (mistag rate). Fig. 5.6 shows the light-flavour jet rejection and c-jet rejection as a function of the b-tagging efficiency for the different algorithms.

The b-tagging performance exhibits some difference in data and Monte Carlo simulations. For this reason, simulated events need to be corrected with scale factors derived from the ratio of the efficiency in data and in MC [99], obtained as a function of jet p_T, η and truth flavour in $t\bar{t}$ events.

The b-tagging algorithm used in the analyses presented in this thesis is the MV2c10 tagger. Four operating points are available [99], with different average b-tagging efficiency and the one with 77% efficiency is chosen.

5.6 Missing Transverse Momentum

Since the colliding proton beams have zero transverse momenta, the sum of the transverse momenta of the collision products should be zero in each event due to momentum conservation. Any measured imbalance is known as missing transverse momentum $E_{\mathrm{T}}^{\mathrm{miss}}$ and allows to reconstruct the momentum carried away by invisible particles like neutrinos or BSM particles.

The x and y component of $E_{\mathrm{T}}^{\mathrm{miss}}$ are given by

$$E_{x(y)}^{\mathrm{miss}} = E_{x(y)}^{\mathrm{miss},e} + E_{x(y)}^{\mathrm{miss},\mu} + E_{x(y)}^{\mathrm{miss,\ jets}} + E_{x(y)}^{\mathrm{miss,\ soft}}, \tag{5.3}$$

where each term is calculated as the negative vectorial sum of the momenta of the respective particles in the x- and y-direction. The magnitude and direction of $E_{\mathrm{T}}^{\mathrm{miss}}$ can be calculated as

$$\begin{aligned} E_{\mathrm{T}}^{\mathrm{miss}} &= \sqrt{(E_x^{\mathrm{miss}})^2 + (E_y^{\mathrm{miss}})^2}, \\ \phi^{\mathrm{miss}} &= \arctan(E_y^{\mathrm{miss}}/E_x^{\mathrm{miss}}). \end{aligned} \tag{5.4}$$

The soft term in Eq. 5.3 includes all the energy deposits in detector not associated with any final state object. These deposits can consist of tracks in the ID (track-based soft term) or signals from the calorimeters (calorimeter-based soft term). The track-based soft term is more robust in high pile-up environments, because tracks can be matched to the primary vertex, and for this reason it is the default choice for $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation in Run-2 [100, 101].

5.7 Overlap Removal

Different final state objects can be reconstructed using the same detector signals, giving raise to some overlap between the objects. For this reason an overlap

removal (OR) algorithm needs to be applied to remove such ambiguities. The algorithm is performed on objects satisfying loose "baseline" identification criteria and proceeds as follows:

- If two electrons share a track in the ID, the lower p_T electron is rejected.
- If an electron shares an ID track with a muon, the electron is removed.
- Any non b-tagged jets within $\Delta R = 0.2$ of an electron are rejected, or within $\Delta R = 0.2$ of a muon if the jet has < 3 tracks or if the muon and jet are ghost-associated.
- For the remaining jets, any electron or muon satisfying $\Delta R < 0.4$ is rejected.

Objects remaining after the overlap removal procedure and that pass the analysis-level reconstruction criteria form the final analysis objects.

Chapter 6

Analysis Strategy

This Chapter aims to introduce the methods and tools employed in the searches for dark matter presented in the subsequent Chapters. The analysis strategy used in the searches is first presented, together with an overview of the statistical methods used to draw conclusions. The main SM background processes contributing in analyses focusing on leptons+$E_{\mathrm{T}}^{\mathrm{miss}}$ signatures are then introduced, followed by a description of the key discriminant variables exploited to select candidate signal events. Finally, an overview of systematic uncertainties and their evaluation is given.

6.1 Analysis Strategy

6.1.1 Event Selection

In order to be able to observe a potential DM signal above the SM background, it is necessary to identify a region of the phase-space where the signal over background ratio is enhanced. Despite the generally small cross-section of the DM production, this is achievable exploiting selections over particular kinematic variables, called discriminant variables, which show different distributions for the signal and the background. Fig. 6.1 shows an example of such a variable. Requirements over many discriminant variables are set at the same time, and the area of the phase-space selected in this way is called Signal Region (SR).

The optimisation of cuts over the discriminant variables is performed by maximising the statistical sensitivity of the analysis. Among the various definitions in literature for such a quantity, the variable Z, defined as

$$Z = \sqrt{2\left(n\ln\left(\frac{n(b+\sigma^2)}{b^2+n\sigma^2}\right) - \frac{b^2}{\sigma^2}\ln\left(1+\frac{\sigma^2(n-b)}{b(b+\sigma^2)}\right)\right)}, \tag{6.1}$$

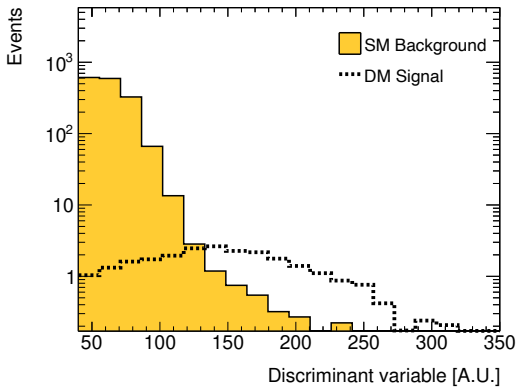


Figure 6.1: Example of a discriminant variable having different distribution for the signal and the SM background.

was adopted, following the recommendations in [102]. In the equation, b is the number of expected background events, σ is its uncertainty and n is the sum of the expected background and signal events. The higher is Z, the better the SR is able to discriminate the signal over the background.

Two different analysis techniques are employed in the next Chapters in the definition of SRs, referred to as cut-and-count and shape-fit. The first one is based on designing a SR optimised to maximise the discovery significance and counting the events in it. The latter employs instead SRs split into multiple bins in a specific discriminating kinematic variable. In this case the individual discovery power of each SR is smaller than in the case of cut-and-count, but by utilising different signal-to-background ratios in the various bins the search sensitivity is enhanced.

6.1.2 Background Estimation

After defining the SR(s), it is necessary to estimate the SM backgrounds there with as much accuracy as possible. This can be done with three different methods:

- Using simulated MC events
- Employing data-driven techniques
- Exploiting simulated MC events and constraining them with data in kinematic regions close to the SR.

The first method is usually employed for minor background processes, since it provides estimates with high systematics uncertainties. The second one is instead

mainly used for backgrounds that would require a very large number of simulated events to obtain a reliable prediction, like the contribution from non-prompt or misidentified leptons. Finally, the third method is used for backgrounds with a signature identical to the signal (called irreducible backgrounds), which give the main contributions to the SR. This method is further explored in the following.

The SR may have a very different kinematic from the regions where the MC simulations have been tuned and validated. The MC might thus poorly model the data in the SR and lead to inaccurate predictions. In order to improve the modelling and to reduce the related uncertainties, the MC events of a specific process can be normalised to the data in a region of the phase space close to the SR and enriched with the process considered, called Control Region (CR). The normalisation factor obtained in this way can thus be extrapolated to the SR for a better background estimate:

$$N_p(\text{SR, est}) = \frac{N_p(\text{CR, obs})}{N_p(\text{CR, MC})} \cdot N_p(\text{SR, MC}), \tag{6.2}$$

where N_p(SR, est) is the SR final estimate for the background process p, N_p(CR, obs) is the observed number of data events in the CR for p, and N_p(SR, MC) and N_p(CR, MC) are the expected number of events of the process p given by MC simulation in the SR and CR respectively. Thanks to presence of the ratio of MC yields in the SR and CR, Eq. 6.2 also allows to absorb some systematics uncertainties in the extrapolation and have a more precise estimate in the SR. CRs are required to have high purity for the background process of interest in order to prevent fluctuation of the normalisation factor, and very low signal contamination to avoid over-estimating the background normalisation factor and hide potential DM signals in the SR. Usually, multiple CRs are defined, one for each main background in the SR, and they are all used in a simultaneous fit to obtain the final estimates for each process in the SR.

The normalisation factors obtained in the CRs are validated in additional regions of the phase space, called Validation Regions (VRs). VRs are defined in a similar way of the CRs, with kinematic cuts closer to the ones of the SRs and low signal contamination. VRs do not take part as contraints in the fit, they are only used to check the validity of the extrapolation of the normalisation factors from the CRs to other regions of the phase space and thus to the SR.

6.1.3 Statistical Treatment

The statistical analysis is one of the key step of each analysis, allowing to draw conclusions about the viability of the model under study. For the searches presented in this thesis, the statistical analysis is performed using the HistFitter

framework [103], a statistical framework developed by the ATLAS collaboration that helps implementing the different analysis regions introduced above, while taking care of the correct treatment of all statistical and systematic uncertainties.

The results of the analyses are extracted using hypothesis tests based on a likelihood function L, given by the product of the Poisson distributions of event counts in the SR(s) and/or CR(s) and of additional distributions that implement the constraints on systematic uncertainties. It can be written as [103]:

$$\begin{aligned} L(n, \theta^0 | \mu_{\text{sig}}, b, \theta) &= P_{\text{SR}} \times P_{\text{CR}} \times C_{\text{syst}} \\ &= P(n_S | \lambda_S(\mu_{\text{sig}}, b, \theta) \times \prod_{i \in \text{CR}} P(n_i | \lambda_i(\mu_{\text{sig}}, b, \theta)) \times C_{\text{syst}}(\theta^0, \theta). \end{aligned} \tag{6.3}$$

The first two factors represent the Poisson observation of n_S and n_i events in the signal region and each control region i. The Poisson expectation values λ_i depend on the expected number of background events b, the nuisance parameters that model the systematic uncertainties θ, the normalisation factors for background processes and also the signal strength parameter μ_{sig}. For $\mu_{\text{sig}} = 0$ the signal component is turned off, while for $\mu_{\text{sig}} = 1$ the signal expectation is taken to be equal to the expected value from the model considered. The systematic term C_{syst} is defined as a product of Gaussian constraints, each representing a systematic variation, with θ^0 being the nominal values around which θ can be varied when maximising the likelihood function.

From the likelihood it is possible to build a test statistic, which can be used to quantify the agreement/disagreement between the observed data and the prediction in a given region. The test statistic employed in this work is the profile log likelihood ratio, defined as

$$q = \begin{cases} -2 \log \frac{L(\mu_{\text{sig}}, \hat{\hat{\theta}})}{L(\hat{\mu}_{\text{sig}}, \hat{\theta})} & \text{if } \hat{\mu}_{\text{sig}} \geq \mu_{\text{sig}} \\ 0 & \text{if } \hat{\mu}_{\text{sig}} < \mu_{\text{sig}} \end{cases} \tag{6.4}$$

where $\hat{\mu}_{\text{sig}}$ and $\hat{\theta}$ are the parameter choices that maximise the likelihood, while $\hat{\hat{\theta}}$ maximises the likelihood for the specific signal strength under test.

Using the profile log likelihood ratio as test statistic, one can calculate the p-value, which gives the probability of having a result at least as extreme as the one observed given a certain hypothesis H_0. The p-value is defined as

$$p\text{-value} = \int_{q_{\text{obs}}}^{\infty} f(q|H_0) dq, \tag{6.5}$$

where f is the probability density function of the test statistic given the hypothesis H_0 and q_{obs} is the value of the test statistic obtained with the observed data. H_0 is usually identified with the background-only hypothesis, i.e. the hypothesis that only SM events are present in the SR. The hypothesis H_0 can be considered excluded if its observed p-value is below a specified threshold. The significance Z corresponding to a given p-value is defined as the number of standard deviations at which a normal distributed random variable, with mean zero, would give a one-sided tail area equal to the p-value. The threshold for the rejection of the background-only hypothesis is typically a significance of $Z = 5$, corresponding to a p-value of $2.8 \cdot 10^{-7}$. For the exclusion of a DM signal model, in Eq. 6.5 the hypothesis H_0 is substituted with H_1, i.e. the hypothesis that the DM predicted by the model is present in the SR, and the threshold p-value used is usually 0.05, which corresponds to $Z = 1.64$.

Using the p-value to perform exclusion may lead to reject the hypothesis of new physics H_1 also in the case where the expected estimate of signal events is very small or when there is a downwards fluctuation of data. To avoid this problem, the modified confidence level CL_s [104] is defined as:

$$CL_s = \frac{p_1}{1 - p_0} = \frac{\int_{q_{\text{obs}}}^{\infty} f(q|H_1)dq}{1 - \int_{q_{\text{obs}}}^{\infty} f(q|H_0)dq}, \tag{6.6}$$

where the denominator is often referred to as CL_b, which indicates the confidence in the background-only hypothesis (values close to 1 indicate poor compatibility with it). With the CL_s method, the hypothesis of new physics is excluded at 95% C.L. if $CL_s < 0.05$. The CL_s method is the default procedure used in the searches presented in the present study to decide on the exclusion of a DM signal model.

Background-Only Fit

The background-only fit aims to estimate the total background in SR(s) and VRs, extrapolating from the CRs the normalisation factors that need to be applied to the background processes to improve the MC estimate. The extraction of the normalisation factors is performed by maximising the probability that the observed data in the CRs would agree with the MC expectations. The likelihood function of Eq. 6.3 is employed, including only the CRs and neglecting any signal contamination fixing $\mu_{\text{sig}} = 0$.

Model Dependent Results

The results of the background-only fit in the SR(s) can be used to calculate the p-value of Eq. 6.5 and check if there is a good agreement with the background-

only hypothesis H_0 or if there is an excess of observed events. In case of excess, a model dependent fit is used to measure signal properties such as the signal strength μ_{sig}, whereas in the absence of a significant excess limits are set on the signal model under study. In particular, it is possible to assess which regions of the parameter space of the model are not compatible with data and are thus excluded using Eq. 6.6. The fit is performed using simultaneously the CRs and SRs, evaluating background and signal in all regions in order to correctly account for possible signal contamination in the CRs. The model dependent fit is repeated for different signal points, obtained with different parameter choices, leading to exclusion limits in the phase space of the model.

Model Independent Results

If no excess of data is observed in the SR(s), analyses searching for new physics usually provide model independent upper limits on the number of possible Beyond the Standard Model (BSM) events compatible with the observed data in each SR, obtained from Eq. 6.6 fixing $CL_s = 0.05$. In this way, it is possible to estimate for any BSM model the number of signal events predicted in each SR and check if the model is already excluded by the current measurements. The model independent fit is performed including both SRs and CRs in the fit, similarly to the model dependent fit. Signal contamination is not allowed in the CRs, but no other assumptions are made for the signal model.

6.2 Background Processes

The analyses described in this thesis focus on two different types of final state signatures, including one or two leptons together with b-tagged jets and $E_{\text{T}}^{\text{miss}}$, with the latter due to the production of DM. In analyses focusing on such signatures, the main backgrounds come from the production of $t\bar{t}$, W+jets (for the one-lepton case) and $t\bar{t}V$ (for the two-lepton one). Sub-dominant contributions come from single top, $t\bar{t}h$, tWZ and diboson productions, while Z+jets events are easily suppressed. This section aims to briefly describe these processes, together with the event generators used for their simulation, based on ATLAS Physics and Modeling Group recommendation [105].

$\mathbf{t\bar{t}}$ Top quarks pair production is the dominant production mechanism for top quarks at the LHC. The process is mediated by the strong interaction, and it can occur through quark-antiquark annihilation or through gluon-gluon fusion (Fig. 6.2). The cross-section of $t\bar{t}$ production is ∼832 pb, as calculated with the Top++2.0 program to next-to-next-to-leading order in perturbative QCD,

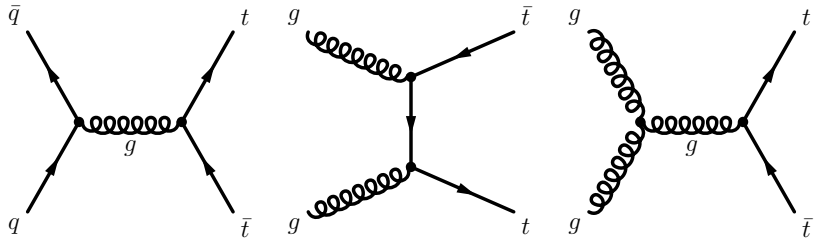


Figure 6.2: Representative Feynman diagrams contributing to $t\bar{t}$ production.

including soft-gluon resummation to next-to-next-to-leading-log order [106], and assuming a top-quark mass of 172.5 GeV. Due to its mass, the top quark decays before hadronising in a W-boson and a b-quark ($t \to Wb$). Roughly 33% of the W-bosons decay leptonically ($W \to \ell\nu_\ell$), while the remaining 67% decay hadronically ($W \to qq'$). According to the decay modes of the W-bosons produced in the top decays, $t\bar{t}$ events can thus result in three different final states, which consist of:

- no leptons, two b-jets and four additional jets from W decays
- one lepton and $E_{\mathrm{T}}^{\mathrm{miss}}$ from the leptonic decay of only one of the W-boson, two b-jets, two additional jets from the decay of the other W boson
- two leptons and $E_{\mathrm{T}}^{\mathrm{miss}}$ from the leptonic decay of both W-bosons and two b-jets.

The production of $t\bar{t}$ events is modelled using the PowhegBox v2 generator at NLO [73, 74, 75, 107]. The events are interfaced to Pythia8 [76] with parameters set according to the A14 tune [70] to model the parton shower, hadronisation, and underlying event.

$t\bar{t}$V(V = Z/W) The associate productions of a top quark pair and a vector boson, either a W or a Z, are quite rare processes. These processes are indeed characterised by small cross-sections of ∼0.88 pb and ∼0.60 pb at 13 TeV for $t\bar{t}Z$ and $t\bar{t}W$, respectively, calculated at NLO [108]. Despite their small cross-section, they represent one of the most important backgrounds for BSM searches, because invisible decays of the Z-boson or W leptonic decays where the lepton gets lost might mimic the signal. Representative Feynman diagrams contributing to $t\bar{t}V$ production are shown in Fig. 6.3.

The production of $t\bar{t}V$ events is modelled using the MadGraph5_aMC@NLO v2.3.3 generator [72] at NLO, interfaced to Pythia8 [76] using the A14 tune [70] for parton shower, hadronisation, and underlying event simulation.

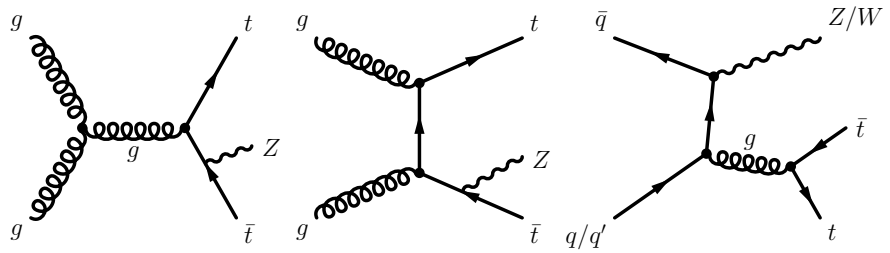


Figure 6.3: Representative Feynman diagrams contributing to $t\bar{t}V$ production with $V = Z, W$.

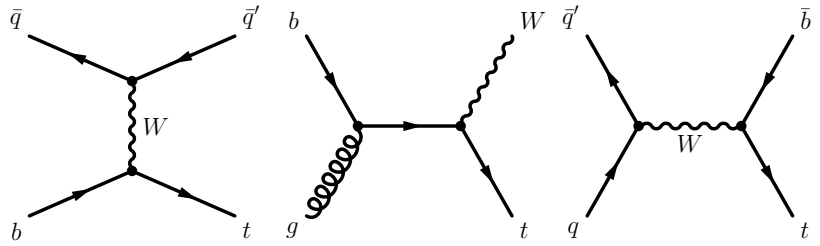


Figure 6.4: Representative Feynman diagrams contributing to single top production.

Single top Top quarks can be produced alone through electroweak processes. The dominant production mode of single top is through t-channel (∼217 pb), followed by the associate production with a W-boson (∼71.7 pb) and finally s-channel production (∼10.3 pb). Another extremely rare production channel is the associated production of a single top quark with a Z boson and a light quark (∼0.6 pb). The cross-sections are calculated for a top quark mass of 172.5 GeV at next-to-leading order (NLO) in QCD with Hathor v2.1 [109, 110]. The Feynman diagrams for the main production modes are shown in Fig. 6.4.

Single-top production in the Wt-channel, s-channel and t-channel is modelled using the PowhegBox v2 generator [73, 74, 75, 111, 112, 113], which provides matrix elements at next-to-leading order (NLO) in QCD. The events are interfaced with Pythia8 [76] for parton shower, hadronisation, and underlying event using the A14 tune [70]. Some of the NLO diagrams contributing to Wt associate production include the same final state as the $t\bar{t}$ process, namely $WWbb$. Considering the two processes separately leads thus to an improper treatment of the interference between them. Two different strategies have been developed to handle this problem: diagram removal (DR) and diagram subtraction (DS) [114].

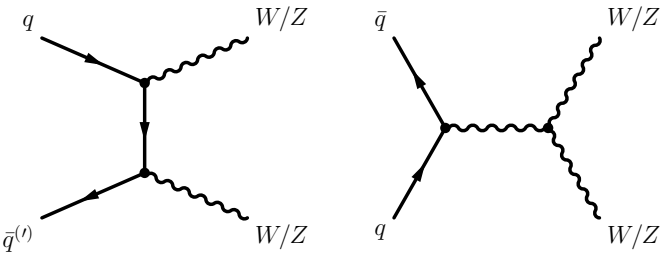


Figure 6.5: Representative Feynman diagrams contributing to diboson production.

None of them correctly reproduces the interference, but they allow to mitigate the size of the effect. The DR scheme is employed for the modelling of Wt-channel in the following.

Diboson Pairs of vector bosons are produced in the SM through the electroweak interaction, as shown in Fig. 6.5. The dominant processes are WW ($\sim$112 pb [115]), WZ ($\sim$48 pb [116]) and ZZ($\sim$16 pb [115]), where the cross-sections are calculated at NLO.

Samples of diboson final states (VV) are simulated with the Sherpa v2.2.1 or v2.2.2 generator [77, 117] depending on the process. Fully leptonic final states and semileptonic final states, where one boson decays leptonically and the other hadronically, are generated using matrix elements calculated at NLO in QCD for up to one additional parton and at LO accuracy for up to three additional parton emissions. Samples for the loop-induced processes $gg \to VV$ are generated using LO-accurate matrix elements for up to one additional parton emission for both cases of fully leptonic and semileptonic final states.

V+jets The production of a W or Z boson in association with jets is shown in Fig. 6.6. For each extra radiated jet, the production cross section decreases by roughly half an order of magnitude, and the total cross-section is respectively $\sim$20 nb and $\sim$1.9 nb for W and Z production [118].

The production of V+jets is simulated with the Sherpa v2.2.1 generator [77, 117] using next-to-leading order (NLO) matrix elements (ME) for up to two partons, and leading order (LO) matrix elements for up to four partons.

t$\bar{\text{t}}$H A pair of top quarks can be produced also in association with a Higgs boson, with a NLO cross-section at 13 TeV of $\sim$507 fb [108]. The corresponding Feynman diagram is shown in Fig. 6.7 (left).

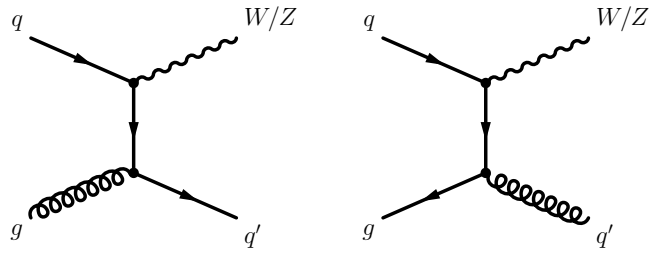


Figure 6.6: Representative Feynman diagrams contributing to V+jets production.

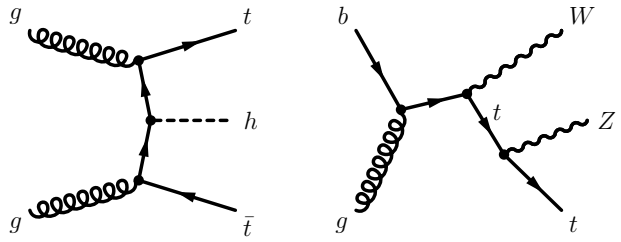


Figure 6.7: Representative Feynman diagram contributing to $t\bar{t}h$ (left) and tWZ (right) productions.

The production of $tt\bar{H}$ events is modelled using the PowhegBox v2 generator [73, 74, 75] at NLO, interfaced to Pythia8 [76] using the A14 tune [70].

tWZ The associated production of a top quark with both a Z and W bosons is a very rare process, with a cross-section at NLO of ∼15.6 fb [119]. It is illustrated in Fig. 6.7 (right).

The production of tWZ events is modelled using the MadGraph5_aMC@NLO v2.3.3 generator [72] at NLO interfaced with Pythia8 [76], making use of the A14 tune [70]. Similarly to the Wt single top production, also in the case of tWZ at NLO there are diagrams with the same final state of $t\bar{t}Z$, which give rise to interference problems. The DR scheme [114] is employed to mitigate them.

6.3 Discriminant Variables

As explained in Sec. 6.1.1, signal region(s) are defined exploiting selections over discriminant variables that allow to reduce the background contamination while retaining the signal. This section will summarise the main discriminant variables used in analyses focusing on final states with one or two leptons, b-jets and $E_{\mathrm{T}}^{\mathrm{miss}}$

coming from DM particles.

6.3.1 Common Discriminant Variables

$\mathbf{E_T^{miss}}$: missing transverse momentum, introduced in Sec. 5.6. The searches presented in the following look for signals characterised by the production of two dark matter particles that escape the detector without interacting. The presence of a large quantity of $E_{\mathrm{T}}^{\mathrm{miss}}$ is thus one of the main discriminant for the signal.

$\boldsymbol{\Delta\phi_{\mathrm{min}}}$: the smallest azimuthal angular distance between the $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ vector and the four leading signal jets. It helps to discriminate against events with $E_{\mathrm{T}}^{\mathrm{miss}}$ coming from jet mismeasurements.

$\mathbf{H_T^{T,sig}}$: object-based missing transverse momentum, divided by the per-event resolution of the jets. It is defined as

$$H_{\mathrm{T,sig}}^{\mathrm{miss}} = \frac{|\vec{H}_T^{\mathrm{miss}}| - M}{\sigma_{\vec{H}_T^{\mathrm{miss}}}}, \tag{6.7}$$

where $\vec{H}_T^{\mathrm{miss}}$ is the negative sum of the jets and lepton vectors. The denominator is computed from the per-event jet energy resolution, while the lepton is assumed to be well-measured. The parameter M is chosen to be a characteristic scale of the background, and is fixed at 100 GeV for the analyses presented in this thesis. This variables assumes higher values in the case of DM production, thanks to the additional invisible momentum carried by DM particles.

$\boldsymbol{\Delta\phi(j_i, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})}$: the azimuthal angle between the ith jet and $E_{\mathrm{T}}^{\mathrm{miss}}$, used to suppress multijet events.

$\mathbf{\Delta R(b_1, b_2)}$: the opening angle in $\eta - \phi$ plane between the two leading b-tagged jets.

6.3.2 One-Lepton Discriminant Variables

$\mathbf{m_T}$: tranverse mass, defined as

$$m_T = \sqrt{2 \cdot p_T(\ell) \cdot E_{\mathrm{T}}^{\mathrm{miss}}(1 - \cos\Delta\phi(\vec{\ell}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}}))}, \tag{6.8}$$

where $p_T(\ell)$ is the lepton p_T, and $\Delta\phi(\vec{\ell}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})$ is the azimuthal angle between the lepton and the $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ direction. If the lepton comes from the decay of a particle p,

m_T is bounded from above by the mass of p. This can be exploited in particular to reduce the background coming from leptonically decaying W-bosons, since DM signals are not kinematically bounded thanks to the additional $E_{\mathrm{T}}^{\mathrm{miss}}$ coming from DM particles.

$\mathbf{am_{T2}}$: asymmetric stransverse mass [120, 121], a variation of the m_{T2} variable described in the next section for two-lepton events. It is built to reconstruct dileptonic $t\bar{t}$ event with one of the lepton going undetected. The two visible decay products needed for the calculation are the system made of the lepton and one b-jet and the other b-jet in the event. In this way, one of the missing particles is assumed to have the W-boson mass, so that the distribution has a lower bound at the W mass, while the upper endpoint is at the top mass. In the case of the DM signal instead am_{T2} is not bounded, since the $E_{\mathrm{T}}^{\mathrm{miss}}$ does not come only from top decays.

$\mathbf{m_{T2}^{\tau}}$: another variation of the two-lepton m_{T2} variable [120, 121], designed to suppress dilepton $t\bar{t}$ events where one of the two leptons is a hadronically decaying τ. In this case, the visible decay products are chosen to be the leading light-flavour jet, which is likely to originate from the τ lepton decay, and the lepton. For dileptonic $t\bar{t}$ events decaying into a τ the distribution presents an endpoint at the W mass, whereas for the signal is has a bigger tail.

$\boldsymbol{\Delta\phi(\ell, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})}$: the azimuthal angular distance between the leading lepton and the $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ vector.

$\mathbf{m_{top}^{reclustered}}$: top-quark candidate reconstructed mass [122]. It is built by considering all small-radius jets in the event and clustering them into large-radius jets using the anti-kT algorithm with a radius parameter $R = 3.0$. The radius of each jet is then iteratively reduced to an optimal radius $R(p_T) = 2 \times m_{\mathrm{top}}/p_T$ that matches their p_T. If a candidate loses a large fraction of its p_T in the shrinking process, it is discarded. In events where two or more top-quark candidates are found, the one with the mass closest to the top-quark mass is taken. This variable is used to identify events containing a top quark decaying hadronically.

$\mathbf{m_{W}^{reclustered}}$: W candidate reconstructed mass [122]. It is constructed in the same way of $m_{\mathrm{top}}^{\mathrm{reclustered}}$, using the W-boson mass instead of the top mass in the optimal radius calculation. It is used to target events with a hadronically decaying W-boson.

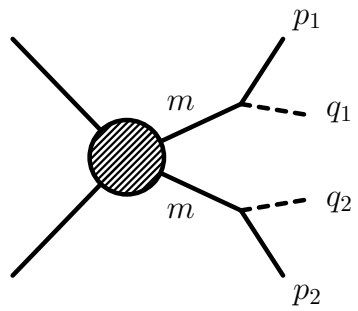


Figure 6.8: Diagram of a generic process for which the variable m_{T2} ha been built.

$\mathbf{\Delta R(b,\ell)}$: the opening angle in $\eta - \phi$ plane between the leading b-tagged jet and the selected lepton.

6.3.3 Two-Lepton Discriminant Variables

$\mathbf{m_{T2}}$: stranverse mass [123, 124]. It is built taking into account a process like the one shown in Fig. 6.8, that leads to the production of a pair of equal particles m, each decaying into one particle that is observed p_i and one particle that is not directly detected q_i, whose presence can only be inferred from the missing transverse momentum. If the momenta of the two invisible particles were known, one could reconstruct for each branch of the diagram the transverse mass m_T, defined in the previous paragraph, which would satisfy $m_T \leq m_m$ for both branches. However, it is impossible to measure separately the missing momenta of the two branches, one can only measure the total $\vec{E}_{\mathrm{T}}^{\mathrm{miss}} = \vec{q}_{T1} + \vec{q}_{T2}$. Hence, event by event, the best one can say is that

$$\min_{\vec{q}_{T,1}+\vec{q}_{T,2}=\vec{E}_{\mathrm{T}}^{\mathrm{miss}}} \{\max[m_T(\vec{p}_{T,1}, \vec{q}_{T,1}), m_T(\vec{p}_{T,2}, \vec{q}_{T,2})]\} < m_m, \tag{6.9}$$

where the minimisation is done over all the possible decompositions of $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$. The expression on the left of Eq. 6.9 is by definition the stranverse mass $m_{T2}(\vec{p}_1, \vec{p}_2, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})$. If $\vec{p}_{T,1}$ and $\vec{p}_{T,2}$ are taken to be the momenta of leptons coming from $t\bar{t}$ or tW decays, the $m_{T2}(\vec{\ell}_1, \vec{\ell}_2, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})$ distribution is thus bounded sharply from above by the mass of the W boson. For DM signals, the additional $E_{\mathrm{T}}^{\mathrm{miss}}$ coming from dark matter generates instead a tail at high m_{T2}, which can be exploited to discriminate the signal from the background.

$\mathbf{\Delta\phi_b}$: the azimuthal angular distance between the $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ vector and the $\vec{p}_{Tb}^{\,\ell\ell} = \vec{E}_{\mathrm{T}}^{\mathrm{miss}} + \vec{p}_T^{\,\ell_1} + \vec{p}_T^{\,\ell_2}$ vector. In case of jet mismeasurement, the first vector points towards the hadronic activity, while the second one is the opposite of the sum of all the transverse hadronic activity. Small angles of this variable reject thus events with $E_{\mathrm{T}}^{\mathrm{miss}}$ arising from mismeasurements.

$\mathbf{\min m_{b\ell}}$: the smallest invariant mass obtained from the leading b-jet and the two leptons. When there are two top quarks both decaying leptonically, at least one of the two reconstructed invariant mass must be below $\sqrt{m_t^2 - m_W^2} \sim 150$ GeV, and this allows to drastically eliminate backgrounds without top quark decay in the final state.

$\mathbf{c_{EM}}$: linear combination of m_{T2} and $E_{\mathrm{T}}^{\mathrm{miss}}$ defined as

$$c_{EM} = m_{T2} + 0.2 \cdot E_{\mathrm{T}}^{\mathrm{miss}}, \tag{6.10}$$

that exploits the correlation between m_{T2} and $E_{\mathrm{T}}^{\mathrm{miss}}$ identified for the signal considered in Chapter 7 and absent in the $t\bar{t}$ background in order to separate them.

$\mathbf{m_{\ell\ell}}$: invariant mass of the dilepton system in the event.

$\mathbf{m_{b\ell}^t}$: variable capable of discriminating against events containing two semileptonic top quark decays, recently introduced in [125]. Similarly to $\min m_{b\ell}$, it exploits the fact that if a b-jet and a lepton are produced in a top-quark decay, their invariant mass is bounded from above by $\sqrt{m_t^2 - m_W^2} \sim 150$ GeV. If a pair of leptonically decaying top quarks is produced in the event, like in $t\bar{t}$ and $t\bar{t}V$ production (Figs. 6.2-6.3), on both legs of the decay the invariant mass of the b-jet and the lepton is thus smaller than ~ 150 GeV. However, it is not possible to know which pair of lepton and b-jet comes from the same leg, one can only define the variable

$$m_{b\ell}^t = \min(\max(m_{\ell_1 j_1}, m_{\ell_2 j_2}), \max(m_{\ell_1 j_2}, m_{\ell_2 j_1})), \tag{6.11}$$

which tries all possible combinations of b-jets and leptons, where $m_{\ell_i j_m}$ is the invariant mass of lepton l_i and jet j_m and j_1, j_2 are the two jets with highest b-tag discriminator value. For $t\bar{t}$ and $t\bar{t}V$ events, $m_{b\ell}^t$ has to satisfy $m_{b\ell}^t < 150$ GeV, while in events with only a single leptonic top decay it is unbounded.

6.4 Systematic Uncertainties

There are two broad classes of systematic uncertainties that have to be considered when performing an ATLAS analysis: experimental uncertainties, related to uncertainties in the reconstruction of physics objects, and modelling uncertainties, related to the Monte Carlo simulations. All systematic uncertainties are implemented in the fits as nuisance parameters, each of them described by a normalised gaussian centred on zero and of width one. Zero corresponds to the nominal value in all regions, while ± 1 correspond to varying the systematic "up" and "down" of one sigma.

6.4.1 Experimental Uncertainties

Experimental uncertainties include all the effects relative to the reconstruction of the physics objects described in Chapter 5 and can affect both the normalisation and shape of kinematics distributions. As a consequence, the number of reconstructed objects that pass the analysis regions requirements may vary, introducing uncertainties in the results. This Section briefly describes the experimental systematics evaluated in the analyses detailed in the next Chapters.

Luminosity The uncertainty on the total luminosity affects the normalisation of Monte Carlo predictions, having thus a direct effect on the final limits of an analysis. For the analyses detailed in Chapter 7, which use data collected in 2015-2016, the luminosity uncertainty is determined to be 2.1%, while for the analyses described in Chapter 8, using the full Run-2 dataset, the luminosity uncertainty is measured to be 1.7%.

Leptons Uncertainties in the reconstruction of leptons arise from the lepton energy resolution and scale calibrations, as well as the efficiencies corresponding to the reconstruction, identification and isolation. They are computed varying the scale factors applied to match MC simulations and data [87, 89].

Jets As previously described in Sec. 5.4, the energy of reconstructed jets needs to be calibrated to the Jet Energy Scale (JES) to account for the non-compensating behaviour of the ATLAS calorimeters. This JES calibration is derived from Monte Carlo simulation, and its uncertainties arise from choices made when developing the calibration. The JES uncertainties [92] are derived as a function of the p_T and η of the jet, as well as of the pile-up conditions and the jet flavour composition. The full JES uncertainty set includes around 80 nuisance parameters, but a reduced set of JES uncertainties is also provided [92] for analyses which are

insensitive to small variations in the calibration, like the ones presented in the following.

Despite the JES calibration, the jet energy cannot be measured exactly due to noise, stochastic fluctuations in the calorimeter response, and detector calibration effects. The Jet Energy Resolution (JER) is thus defined as the width of the Gaussian distribution describing the measurements of the same truth jet energy at the calibrated jet energy scale. The JER uncertainty is determined from both data and MC as a function of p_T and η.

Finally, also an uncertainty on the efficiency of the JVT has to be considered. It is estimated comparing the efficiency obtained using different MC generator to simulate the process $Z \to \mu^+\mu^-$ with an additional hard scatter jet.

Flavour Tagging Uncertainties on the b-tagging are due to the scale factors used to correct the different efficiencies observed in data and Monte Carlo. These uncertainties depend on the p_T, η and flavour of the jet, and are estimated by varying the scale factors within a range that reflects the uncertainty in the measured tagging efficiency and mis-tag rates in data [99].

Missing Transverse Momentum The uncertainties on $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction are evaluated separately for each term in Eq. 5.3. The uncertainties on the object-based $E_{\mathrm{T}}^{\mathrm{miss}}$ terms are derived by propagating the individual energy scale and resolution uncertainties on the physics objects. Correlations between systematic uncertainties for the same type of object are taken into account, while it is assumed that systematic uncertainties of the different object types entering $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction are uncorrelated. The energy scale and resolution of the soft term are instead evaluated comparing data and MC for the $Z \to \mu^+\mu^-$ process, where no $E_{\mathrm{T}}^{\mathrm{miss}}$ is expected [100].

6.4.2 Modelling Uncertainties

Modelling uncertainties arise from the Monte Carlo simulations of backgrounds and signal. Also in this case, the effect can be both on the normalisation and on shape of kinematics distributions, leading to different yields in the analysis regions.

Modelling systematics are evaluated by simulating the same physical process with two different MC generators or configurations and comparing the results. Some MC generators are also able to associate weights to the events such that the impact of different parameter choices can be evaluated directly using different weights. The observed difference in the results in each analysis region is then used in the statistical analysis as a systematic affecting the nominal distributions of

the corresponding process. In the following, a general overview of the sources of modelling uncertainties is presented.

Matrix Element Uncertainties The nominal generator used to model the hard scatter is typically chosen based on which one provides the best agreement between MC predictions and data. The uncertainty on the calculation of the ME is evaluated by generating a MC sample with a different generator for the hard scatter, while using the same PS algorithm.

PDF Uncertainties The PDF uncertainties are estimated comparing predictions obtained with the nominal PDF set and with alternative ones. Usually they are found to be of the order of $\sim 1\%$, and are thus neglected.

Factorisation and Renormalisation Scales Uncertainties As explained in Sec. 4.2, when not all orders in perturbation theory are considered the cross-section for a hard scattering process depends on the choice of the factorisation and renormalisation scales (μ_F and μ_R). The default scales are thus varied using the nominal generator, and the differences obtained in the cross-section are taken as an uncertainty.

Parton Showering Uncertainties Similar to the case of ME, the nominal choice of PS generator is based on the best possible agreement between data and MC simulations. To estimate the PS uncertainties, the nominal configuration is compared with a sample produced using the same generator for the hard scatter but a different PS algorithm.

Chapter 7

Search for Dark Matter Produced in Association with a Top Pair

In Chapter 2 the 2HDMa model was introduced as a well-motivated template for the production of dark matter at the LHC. Based on this model, a coherent research program for the discovery (or the exclusion) of this model with the LHC pp collision data collected at 13 TeV has been developed. The program involved different steps: first of all, the interesting final-state signatures at the LHC were identified, and the coverage on the parameter space of the model of existing analyses was assessed, taking into account in particular analyses based on the very similar simplified models for dark matter production. After that, an analysis program to extend the coverage of the model was developed, including both the refinement of the already published searches thanks to the increased integrated luminosity and the identification of new search channels specific to the 2HDMa model. The work presented in this thesis has developed within this program, starting with the master thesis of the author and continuing during the PhD years, focusing on 2HDMa DM signatures including top quarks.

The first part of the work consisted in the study of the production of DM in association with a pair of top quarks ($pp \to t\bar{t}\chi\bar{\chi}$, called $DMt\bar{t}$ in the following) in the framework of 2HDMa, identifying which parts of the parameter space of the model were already covered by published analyses. The main motivation behind the choice of this final state is the fact that the $DMt\bar{t}$ signature, although less sensitive than $E_{\mathrm{T}}^{\mathrm{miss}} + h$ and $E_{\mathrm{T}}^{\mathrm{miss}} + Z$ on most of the parameters space, offers unique handles for unravelling the nature of the model through the detection of the decay products of the two top quarks in the final state, as described in [126]. Three different analyses published by the ATLAS Collaboration focusing on the same signature in the context of a simplified model were identified [42, 41]. Both the simplified model used for the ATLAS analyses and the 2HDMa model

were then studied in detail to understand their similarities and differences, and a recasting strategy was developed to allow the reinterpretation of results based on the simplified model, as documented in [44]. Subsequently, this strategy was applied to the ATLAS data in the framework of a paper [1] which reinterpreted all the existing 36.1 fb^{-1} results in the context of 2HDMa. All the different steps of this work, for which the author of this thesis is the main contributor, are presented in this Chapter. In the beginning of the Chapter, the phenomenology of the production of $DMt\bar{t}$ in the 2HDMa model is analysed, exploring the dependency on the parameters of the model. After that, the ATLAS analyses investigating the $DMt\bar{t}$ final state in the context of the simplified model are briefly summarised. Finally, the development of the recasting strategy is illustrated in detail, followed by the presentation of the resulting limits into the parameter space of the 2HDMa model.

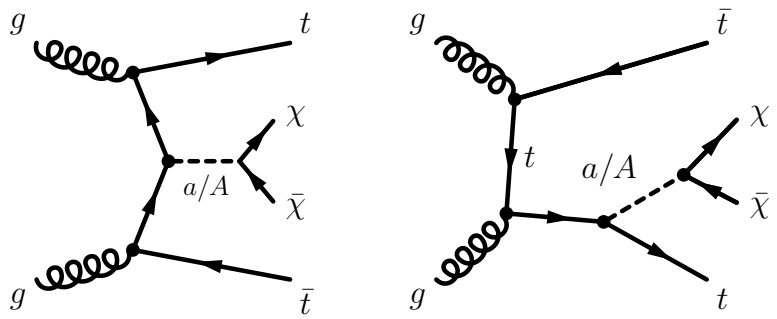
The second part of the work focused instead on exploring a new search channel for the 2HDMa model, not yet addressed by existing analyses. The identified channel was the production of dark matter associated with a single top quark, recently proposed in [127]. This signature has the unique feature of being sensitive to the production of charged Higgs bosons, to which none of the previously proposed analysis were sensitive. The analysis was carried out on the full Run-2 data recorded by the ATLAS experiment, and will be the focus of the next Chapter.

7.1 Signal Simulation

Signal samples of $DMt\bar{t}$ production in the context of 2HDMa were simulated using the MadGraph5_aMC@NLO v2.3.3 [72] generator at NLO. The authors of the 2HDMa model [34] have provided the model Lagrangian in the format accepted by the FeynRules 2 [128] packages and NLOCT packages, which calculate the Feynman rules for the relevant processes, and outputs them in the form of an UFO (Universal FeynRules output), which was used as input for MadGraph5_aMC@NLO. The generation was performed using NNPDF3.0 PDFs [60]. The events were interfaced to Pythia8 [76] using the A14 tune for showering, hadronisation and underlying event simulation.

In order to perform a rapid evaluation of the Branching Fractions (BR) of the particles of the model and to understand the interplay of the model parameters in determining the observable final state, a dedicated library based on the code included in the 2HDMa UFO was developed on purpose for the present study. The library allows the calculation of the 2-body branching fractions of all the scalars of 2HDMa, while ignoring 3-body decays. The impact of the latter, however, is negligible, except in extreme regions of the model parameter space.

Throughout this Chapter, if not stated otherwise, the parameter choices pre-



[tb]

Figure 7.1: Feynman diagrams contributing to $pp \to t\bar{t}\chi\bar{\chi}$ in the 2HDMa model.

sented in Sec. 2.2.6 are adopted:

$$\begin{gathered} m_H = m_A = m_{H^\pm}, \\ \cos(\beta - \alpha) = 0, \\ \lambda_3 = \lambda_{P1} = \lambda_{P2} = 3, \\ Y_u^1 = Y_d^2 = Y_\ell^2 = 0, \\ m_\chi = 10 \text{ GeV}, \\ y_\chi = 1. \end{gathered} \tag{7.1}$$

7.2 Phenomenology of $DMt\bar{t}$ Production in the 2HDMa Model

7.2.1 Process Composition

From an inspection of the Feynman diagrams contributing to the $DMt\bar{t}$ production in the 2HDMa model, automatically produced by MadGraph5_aMC@NLO, the process is dominated by a and A bosons on-shell production, followed by their decays into DM particles (Fig. 7.1):

$$\begin{aligned} pp \to t\bar{t}a \to t\bar{t}\chi\bar{\chi}, \\ pp \to t\bar{t}A \to t\bar{t}\chi\bar{\chi}. \end{aligned} \tag{7.2}$$

This is confirmed by the study at parton level of the invariant mass of the two DM particles produced, shown in Fig. 7.2, where the two peaks correspond to the resonant production of the two pseudoscalar bosons a and A. One can thus work

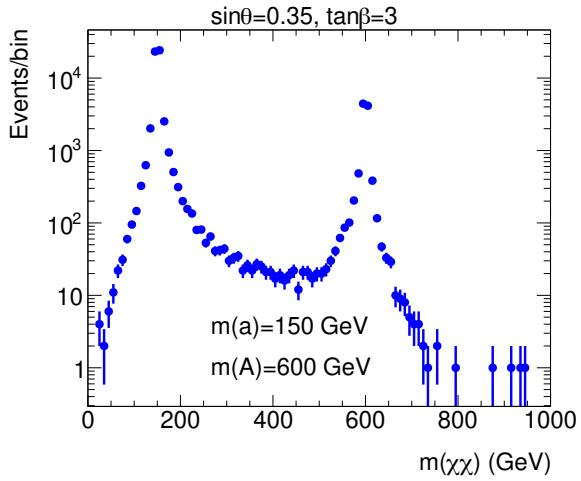


Figure 7.2: Invariant mass of the two DM particles coming from $pp \to t\bar{t}\chi\bar{\chi}$ in 2HDMa.

in the narrow width approximation and separately study the processes

$$pp \to t\bar{t}a, \qquad pp \to t\bar{t}A, \tag{7.3}$$

and the subsequent decays

$$a \to \chi\bar{\chi}, \qquad A \to \chi\bar{\chi}. \tag{7.4}$$

In the regions of parameter space where the narrow width approximation holds, the total cross-section of $pp \to t\bar{t}\chi\bar{\chi}$ can thus be written as

$$\sigma(pp \to t\bar{t}\chi\bar{\chi}) = \sigma(pp \to t\bar{t}A) \cdot BR(A \to \chi\bar{\chi}) + \sigma(pp \to t\bar{t}a) \cdot BR(a \to \chi\bar{\chi}). \tag{7.5}$$

Based on this ansatz, the interplay of cross-sections and branching fractions in determining the total rate of $DMt\bar{t}$ events for different choices of the model parameters can be investigated.

However, Eq. 7.5 does not take into account one important effect: the interference between a and A, which results in a LO matrix element of the form

$$\mathcal{M}(pp \to t\bar{t}\chi\bar{\chi}) \propto \frac{1}{m^2_{\chi\bar{\chi}} - m^2_a - im_a\Gamma_a} - \frac{1}{m^2_{\chi\bar{\chi}} - m^2_A - im_A\Gamma_A}, \tag{7.6}$$

where $m_{\chi\bar{\chi}}$ denotes the invariant mass of the DM pair and Γ_a and Γ_A are the total

decay widths of the two pseudoscalar mass eigenstates. The contributions from virtual a and A exchange have opposite signs, meaning that the interference is destructive. In Fig. 7.3, as an example, it is shown the cross-section of $pp \to t\bar{t}\chi\bar{\chi}$ as a function of m_a for $\tan\beta = 1$, $\sin\theta = 0.5$, $y_\chi = 1$, $m_A = 300$ GeV, $m_h = 125$ GeV and $m_{H^\pm} = m_H = 750$ GeV. The blue line represents the cross-section for $pp \to t\bar{t}\chi\bar{\chi}$ at 13 TeV directly generated with MadGraph5_aMC@NLO, the green one is obtained generating the cross-sections of the processes in Eq. 7.2 with MadGraph5_aMC@NLO and summing them, while the red one is obtained producing the processes in Eq. 7.3 and employing Eq. 7.5, with the BR calculated with the library developed for the purpose. The blue line differs in particular for $m_a \approx m_A = 300$ GeV, where the destructive interference described by Eq. 7.6 is maximal, while the red one shows a different behaviour around 170 GeV, where a new decay channel opens for A ($A \to ah$), which becomes very broad and breaks the narrow width approximation. When trying to separate the different contribution to the final cross-section it is thus important to keep in mind that Eq. 7.5 does not hold in regions where interference is important or where the narrow width approximation is not valid.

7.2.2 Dependency on $\tan\beta$ and $\sin\theta$

The parameters that control the couplings of the top quark to a and A are $\sin\theta$ and $\tan\beta$. As it can be seen from Sec. 2.2.3, in the Yukawa sector of type II under study one has:

$$\sigma(pp \to t\bar{t}a) \propto \frac{\sin^2\theta}{\tan^2\beta}, \qquad \sigma(pp \to t\bar{t}A) \propto \frac{\cos^2\theta}{\tan^2\beta}, \tag{7.7}$$

At this point, it is interesting to investigate if the scaling described by Eq. 7.7 can be exploited to rescale also the $DMt\bar{t}$ cross-section on the basis of the factorisation of production cross-sections and BR as written in Eq. 7.5, or if in some region of parameter space the decay widths become big enough to destroy the narrow width approximation and the validity of this approach.

In order to verify the factorisation, the processes in Eq. 7.3 were generated with MadGraph5_aMC@NLO v2.3.3 with fixed values of $\sin\theta$ and $\tan\beta$ and then their cross-sections was rescaled using Eq. 7.7 for new values of the two parameters. After that, the rescaled cross-sections was multiplied for the corresponding branching fractions and the two contributions were summed (Eq. 7.5). The input chosen for the simulation were: $\tan\beta = 1$, $\sin\theta = 0.5$, $m_h = 125$ GeV, $m_a = 200$ GeV, $m_A = m_H = m_H^\pm = 600$ GeV. The rescaled cross-section was then compared with the one directly obtained generating the process $pp \to t\bar{t}\chi\bar{\chi}$ for the different $\tan\beta$ and $\sin\theta$ values. The results are presented in Table 7.1,

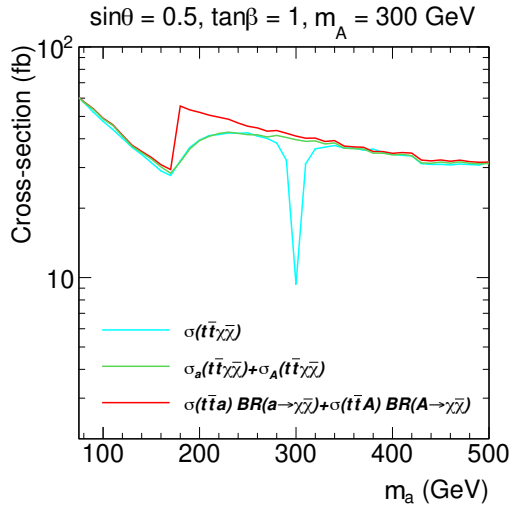


Figure 7.3: Predictions for $DMt\bar{t}$ cross-section as a function of m_a calculated directly with MadGraph5_aMC@NLO (blue line), obtained generating the cross-sections of the processes in Eq. 7.2 with MadGraph5_aMC@NLO and summing them (green line) and calculated producing the processes in Eq. 7.2 and employing Eq. 7.5, with the BR calculated with the library developed for the purpose (red line).

and show that the scaling is valid for every value of $\sin\theta$ and $\tan\beta$ considered to better than 10%.

The dependency on $\sin\theta$ and $\tan\beta$ of the branching fraction of a and A into $\chi\bar{\chi}$ was also investigated. The results in the plane $m_a - m_A$ are shown in Fig. 7.4, with $m_H = m_{H^\pm} = 600$ GeV and $m_h = 125$ GeV. Analysing first the behaviour of a, from Fig. 7.4 (top) one sees that as long as the decay in $t\bar{t}$ is closed, $BR(a \to \chi\bar{\chi}) \sim 1$ except for very high $\sin\theta$ and $\tan\beta$ values, where the $a \to b\bar{b}$ decays starts to be competitive. If the $t\bar{t}$ channel is open, it will dominate for low values of $\tan\beta$ or higher values of $\sin\theta$. The same pattern is visible for the decay of A, but with $\cos\theta$ replacing $\sin\theta$.

The dependency on $\tan\beta$ described by Eq. 7.7 and the consideration of Fig. 7.4 suggest that for moderate light a the $DMt\bar{t}$ searches are more sensitive to small values of $\tan\beta$, where the production cross-section is enhanced.

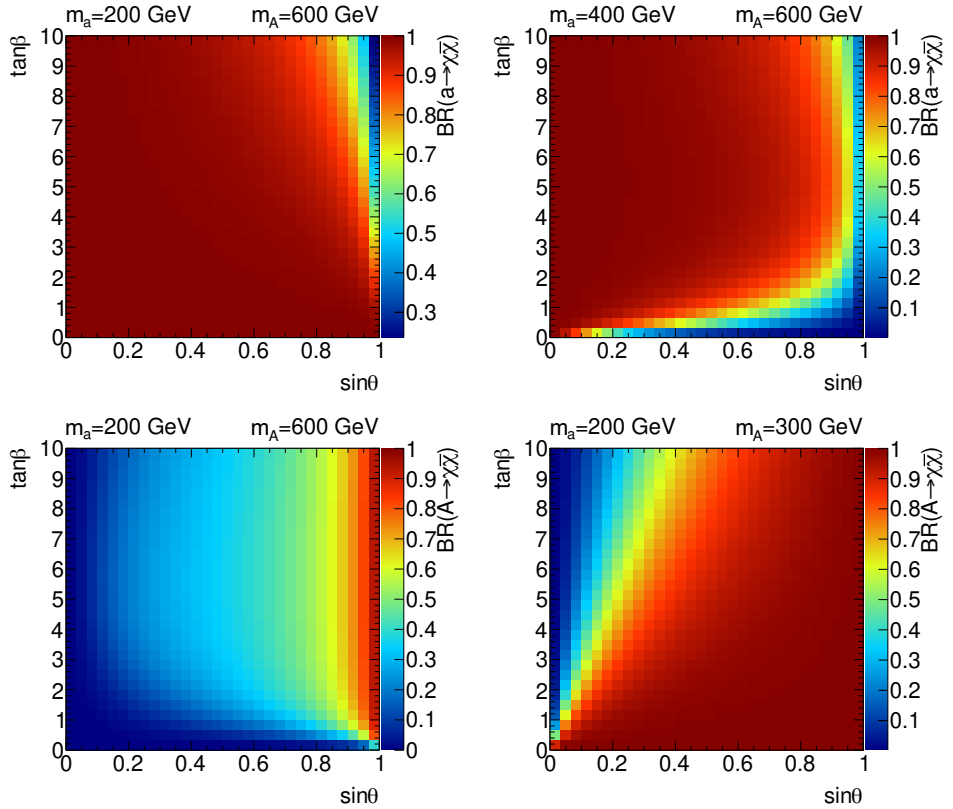


Figure 7.4: Branching fraction into $\chi\bar{\chi}$ for a (top row) and A (bottom row) for different values of m_a and m_A.

$\tan\beta$	$\sin\theta$	$\sigma_{\text{generated}}$ (pb)	σ_{rescaled} (pb)
1	0.5	0.021496	0.022334
5	0.5	0.000870	0.000924
10	0.5	0.000217	0.000224
15	0.5	0.000107	0.000094
1	0.35	0.010759	0.010987
1	0.7	0.041942	0.043662

Table 7.1: Cross-sections for $pp \to t\bar{t}\chi\bar{\chi}$ for different values of $\sin\theta$ and $\tan\beta$ generated directly with MadGraph5_aMC@NLO and rescaled as described in the text.

7.3 ATLAS Searches for DM Produced in Association with a Top Pair in the DMSIMP Model

As discussed in the introduction to this Chapter, the aim of the present study is to reinterpret in the framework of the 2HDMa model the results of existing ATLAS searches for $DMt\bar{t}$. It is therefore necessary to give some details about the model used in the ATLAS analyses and to provide a brief overview of their strategy.

The ATLAS Collaboration has searched for the production of dark matter in association with a top pair in the framework of simplified models, in which the production of DM particles happens through the exchange of a massive mediator ϕ_p, which couples both to SM and DM particles (see Sec. 2.1) [129, 130, 131]. The model has four parameters: the mass of the mediator m_{ϕ_p}, the DM mass m_χ, the DM-mediator coupling g_χ, and the flavor-universal SM-mediator coupling g_f. In the ATLAS searches, only decays of the mediator into a pair of DM particles are considered. A further assumption is that $g_f = g_\chi \equiv g = 1.0$ and that that $m_\chi = 1$ GeV. The model thus defined will be referred to as DMSIMP in the following.

The production of $DMt\bar{t}$ in the context of DMSIMP happens through the same diagrams of Fig. 7.1, with the pseudoscalar mediator ϕ_p replacing a/A. Searches can focus on three different final states, depending on the decays of the two top quarks in the events. If both top quarks decay semileptonically, the final state includes two b-tagged jets, two leptons (e, μ) and $E_{\text{T}}^{\text{miss}}$. If only one top decays semileptonically, the final state only has one lepton, and if both top quarks decay fully hadronically no lepton is present on the event, but only a high multiplicity of jets and $E_{\text{T}}^{\text{miss}}$. ATLAS has searched for $DMt\bar{t}$ production in all

of these channels using 36.1 fb^{-1} proton-proton collision data collected by the ATLAS experiment in 2015 and 2016, and the results are documented in [42, 41].

No excess of events was observed over the predicted SM background, resulting for each of the analyses in a limit at 95% C.L. on the possible number of events from BSM physics. This limit was interpreted in the framework of DMSIMP for each of the analyses as upper limits on the excluded cross-section for $DMt\bar{t}$ production as a function of m_{ϕ_p}. These limits are presented as upper limits on the multiplicative factor with respect to the nominal cross-section for $g = 1.0$, denoted as $\sigma/\sigma(g = 1.0)$. If the multiplicative factor is smaller than 1, the corresponding value of mediator mass is excluded.

Given the fact that the ATLAS Collaboration has three analyses with non-overlapping final state, the optimal coverage for the 2HDMa model results from a reinterpretation of the statistical combination of all three ATLAS analyses. However, since the zero-lepton analysis is somewhat less sensitive than the others, in the following only the one-lepton and two-lepton analyses will be presented.

7.3.1 One-Lepton Analysis

Event Selection

The experimental signature of the one-lepton $DMt\bar{t}$ analysis is composed of one lepton and at least four jets from the top quarks decays, in addition to $E_{\text{T}}^{\text{miss}}$ from DM production. Interesting events are then selected [42] requiring exactly one lepton with $p_T > 25$ GeV, at least four energetic jets ($p_T > (60, 50, 40, 40)$ GeV respectively), of which at least one has to be b-tagged, and $E_{\text{T}}^{\text{miss}} > 250$ GeV. The $E_{\text{T}}^{\text{miss}}$ trigger is required to be fired.

The dominant background for one-lepton $DMt\bar{t}$ final state are:

- Reducible backgrounds, like Z+jets and dibosons
- $t\bar{t}$ and Wt production, with a final state identical to the signal, but with $E_{\text{T}}^{\text{miss}}$ kinematically bounded from W mass
- Irreducible backgrounds, like $t\bar{t}Z$ events with Z decaying invisibly, $t\bar{t}W$ events with a lost lepton from the decay of the W, and W+jets production.

Four different signal regions are defined [42], making use of the variables presented in Secs. 6.3.1-6.3.2 to reduce the background contribution. They are called tN_med, DM_high, DM_low and DM_low_loose. The selection cuts of each signal region are defined in Tab. 7.2.

Variable [Unit]	tN_med	DM_high	DM_low	DM_low_loose
$\mathcal{N}_\ell$	1	1	1	1
$p_T(\ell)$ [GeV]	> 25	> 25	> 25	> 25
$\mathcal{N}_{\text{jets}}$	≥ 4	≥ 4	≥ 4	≥ 4
$\mathcal{N}_{b-\text{jets}}$	≥ 1	≥ 1	≥ 1	≥ 1
$p_T(b_1)$ [GeV]	≥ 25	≥ 25	≥ 25	≥ 25
$p_T(j_1, j_2)$ [GeV]	$\geq (60, 50)$	$\geq (125, 75)$	$\geq (120, 85)$	$\geq (60, 60)$
$p_T(j_3, j_4)$ [GeV]	$\geq (40, 40)$	$\geq (65, 25)$	$\geq (65, 60)$	$\geq (40, 25)$
$E_{\text{T}}^{\text{miss}}$ [GeV]	> 250	> 380	> 320	> 300
$H_{\text{T,sig}}^{\text{miss}}$	> 14	incl.	> 14	> 14
$E_{\text{T},\perp}^{\text{miss}}$ [GeV]	> 230	incl.	incl.	incl.
m_T [GeV]	> 160	> 225	> 170	> 120
$m_{\text{top}}^{\text{reclustered}}$ [GeV]	> 150	> 130	> 130	incl.
am_{T2} [GeV]	> 175	> 190	> 160	> 140
$\Delta R(b, \ell)$	< 2	incl.	incl.	incl.
$\lvert\Delta\phi(j_{1,2}, \vec{p}_{\text{T}}^{\text{miss}})\rvert$	> 0.4	> 0.4	> 0.4	-
$\lvert\Delta\phi(j_i, \vec{p}_{\text{T}}^{\text{miss}})\rvert$	-	> 1.0	> 1.0	> 1.4
$\Delta\phi(\ell, \vec{p}_{\text{T}}^{\text{miss}})$	-	> 1.2	> 1.2	> 0.8
m_{T2}^{τ} [GeV]	> 80	> 80	> 80	> 80

Table 7.2: Signal regions definitions for the one-lepton $DMt\bar{t}$ analysis [42].

Background Estimation

The main backgrounds for the one-lepton $DMt\bar{t}$ analysis come from the production of $t\bar{t}$, single-top, W+jets and $t\bar{t}Z$ with the Z boson decaying into neutrinos. For all these processes control regions are defined, with slightly different cuts based on the corresponding signal region.

The dileptonic $t\bar{t}$ CR (T2LCR) requires m_T above the W-mass endpoint, inverting the SR request on am_{T2} and vetoing hadronic top-quark decays. The CR for semileptonic $t\bar{t}$ (T1LCR) is defined asking for a hadronic top-quark decay and m_T within a window around the mass of the W-boson.

The CRs for W+jets (WCR) and single-top (STCR) production select event with m_T below the W-boson mass, large am_{T2} and a hadronic top-quark veto, in order to suppress the semileptonic $t\bar{t}$ contamination. STCR also requires two b-tagged jets to reduce the W+jets contribution, with a minimum separation $\Delta R(b_1, b_2) > 1.2$ to suppress the semileptonic $t\bar{t}$ contribution. WCR selects instead events with exactly one b-tagged jet or $\Delta R(b_1, b_2) < 1.2$. Moreover, only events with a positively charged lepton are kept in WCR to increase its purity.

Finally, the $t\bar{t}Z$ background is estimated in a three-lepton CR (TZCR), with

Variable [Unit]	tN_med	T1LCR	T2LCR	WCR	STCR	TZCR
$\mathcal{N}_\ell$	1	1	1	1	1	3 (2 SFOS)
$p_T(\ell_1)[GeV]$	> 25	> 25	> 25	> 25	> 25	> 25
Lepton charge	-	-	-	+1	-	-
$\mathcal{N}_{\text{jets}}$	≥ 4	≥ 4	≥ 4	≥ 4	≥ 4	≥ 4
$\mathcal{N}_{b-\text{jets}}$	≥ 1	≥ 1	≥ 1	≥ 1	≥ 2	≥ 1
$p_T(b_1)$ [GeV]	≥ 25	≥ 25	≥ 25	≥ 25	≥ 25	≥ 25
$p_T(j_1, j_2)$ [GeV]	$\geq (60, 50)$	$\geq (60, 50)$	$\geq (60, 50)$	$\geq (60, 50)$	$\geq (60, 50)$	$\geq (60, 50)$
$p_T(j_3, j_4)$ [GeV]	$\geq (40, 40)$	$\geq (40, 40)$	$\geq (40, 40)$	$\geq (40, 40)$	$\geq (40, 40)$	$\geq (40, 40)$
$E_{\text{T}}^{\text{miss}}$ [GeV]	> 250	> 250	> 250	> 250	> 250	incl.
$H_{\text{T,sig}}^{\text{miss}}$	> 14	> 10	> 10	> 10	> 10	incl.
$E_{\text{T},\perp}^{\text{miss}}$ [GeV]	> 230	> 230	> 230	> 230	> 230	incl.
m_T [GeV]	> 160	$[30, 90]$	> 120	$[30, 90]$	$[30, 120]$	incl.
$m_{\text{top}}^{\text{reclustered}}$ [GeV]	> 150	> 150	top veto	top veto	top veto	incl.
am_{T2} [GeV]	> 175	< 200	< 200	> 200	> 200	incl.
$\Delta R(b, \ell)$	< 2	incl.	incl.	incl.	incl.	incl.
$\Delta R(b_1, b_2)$	incl.	incl.	incl.	< 1.2	> 1.2	incl.
$\lvert\Delta\phi(j_{1,2}, \vec{p}_{\text{T}}^{\text{miss}})\rvert$	> 0.4	> 0.4	> 0.4	> 0.4	> 0.4	incl.
m_{T2}^{τ} [GeV]	> 80	> 80	> 80	> 80	> 80	incl.

Table 7.3: Control region definitions corresponding to tN_med SR for the one-lepton $DMt\bar{t}$ analysis [42]. The entry of "incl." represents an inclusive selection with no requirements.

at least one same-flavour-opposite-charge (SFOS) lepton pair compatible with a Z-boson decay.

The selection cuts for all the CRs with their respective SRs are summarised in Tabs. 7.3, 7.4, 7.5 and 7.6.

Results

Background-only fit results of the one-lepton $DMt\bar{t}$ analysis are shown in Tab. 7.7. A good agreement between data and MC expectations is observed in all signal regions, thus exclusion limits were set.

Fig. 7.5 presents the upper limits at 95% C.L. on the ratio of the excluded cross-section to the nominal cross-section for a coupling assumption of $g = 1$. In order to obtain the excluded cross-section σ_{exc} from it, one has to multiply the excluded ratio for $\sigma(g = 1)$, publicly available. The plot shows that the limit given by the one-lepton $DMt\bar{t}$ analysis is stronger for small masses of the mediator. The irregular behaviour of the observed limit is due to the use of different signal regions for each point, selected based on the region with the best

Variable [Unit]	DM_high	T1LCR	T2LCR	WCR	STCR	TZCR
$\mathcal{N}_\ell$	1	1	1	1	1 3 (2 SFOS)	
$p_T(\ell_1)$ [GeV]	> 25	> 25	> 25	> 25	> 25	> 25
Lepton charge	incl.	incl.	incl.	$+1$	incl.	incl.
$\mathcal{N}_{\text{jets}}$	≥ 4	≥ 4	≥ 4	≥ 4	≥ 4	≥ 4
$\mathcal{N}_{b-\text{jets}}$	≥ 1	≥ 1	≥ 1	≥ 1	≥ 2	≥ 1
$p_T(b_1)$ [GeV]	≥ 25	≥ 25	≥ 25	≥ 25	≥ 25	≥ 25
$p_T(j_1, j_2)$ [GeV]	$\geq (125, 75)$	$\geq (125, 75)$	$\geq (125, 75)$	$\geq (125, 75)$	$\geq (125, 75)$	$\geq (125, 75)$
$p_T(j_3, j_4)$ [GeV]	$\geq (65, 25)$	$\geq (65, 25)$	$\geq (65, 25)$	$\geq (65, 25)$	$\geq (65, 25)$	$\geq (65, 25)$
$E_{\text{T}}^{\text{miss}}$ [GeV]	> 380	> 280	> 280	> 280	> 280	incl.
m_T [GeV]	> 225	$[30, 90]$	> 120	$[30, 90]$	$[30, 120]$	incl.
$m_{\text{top}}^{\text{reclustered}}$ [GeV]	> 130	> 130	top veto	top veto	top veto	incl.
am_{T2} [GeV]	> 190	< 200	< 200	> 190	> 200	incl.
$\Delta\phi(\ell, \vec{p}_{\text{T}}^{\text{miss}})$	> 1.2	incl	> 1.2	incl.	incl.	incl.
$\Delta R(b_1, b_2)$	incl.	incl.	incl.	< 1.2	> 1.2	incl.
$\lvert\Delta\phi(j_{1,2}, \vec{p}_{\text{T}}^{\text{miss}})\rvert$	> 0.4	> 0.4	> 0.4	> 0.4	> 0.4	incl.
$\lvert\Delta\phi(j_i, \vec{p}_{\text{T}}^{\text{miss}})\rvert$	> 1.0	> 1.0	incl.	> 1.0	incl.	incl.
m_{T2}^{τ} [GeV]	> 80	> 80	> 80	> 80	> 80	incl.

Table 7.4: Control region definitions corresponding to DM_high SR for the one-lepton $DMt\bar{t}$ analysis [42]. The entry of "incl." represents an inclusive selection with no requirements.

expected limit. The observed limit is better than the expected one for DM_low and DM_high signal regions, where a small deficiency of data with respect to expectations is observed, while it is worse for tN_med and DM_low_loose, where a mild excess is present (Tab. 7.7) [42].

Variable [Unit]	DM_low	T1LCR	T2LCR	WCR	STCR	TZCR
$\mathcal{N}_\ell$	1	1	1	1	13 (2 SFOS)	
$p_T(\ell_1)$ [GeV]	> 25	> 25	> 25	> 25	> 25	> 25
Lepton charge	incl.	incl.	incl.	$+1$	incl.	incl.
$\mathcal{N}_{\text{jets}}$	≥ 4	≥ 4	≥ 4	≥ 4	≥ 4	≥ 4
$\mathcal{N}_{b-\text{jets}}$	≥ 1	≥ 1	≥ 1	≥ 1	≥ 2	≥ 1
$p_T(b_1)$ [GeV]	≥ 25	≥ 25	≥ 25	≥ 25	≥ 25	≥ 25
$p_T(j_1, j_2)$ [GeV]	$\geq (120, 85)$	$\geq (120, 85)$	$\geq (120, 85)$	$\geq (120, 85)$	$\geq (120, 85)$	$\geq (120, 85)$
$p_T(j_3, j_4)$ [GeV]	$\geq (65, 60)$	$\geq (65, 60)$	$\geq (65, 60)$	$\geq (65, 60)$	$\geq (65, 60)$	$\geq (65, 60)$
$E_{\text{T}}^{\text{miss}}$ [GeV]	> 320	> 250	> 230	> 250	> 250	incl.
$H_{\text{T,sig}}^{\text{miss}}$	> 14	> 10	> 10	> 10	> 10	incl.
m_T [GeV]	> 170	$[30, 90]$	> 120	$[30, 90]$	$[30, 120]$	incl.
$m_{\text{top}}^{\text{reclustered}}$ [GeV]	> 130	> 130	top veto	top veto	top veto	incl.
am_{T2} [GeV]	> 160	< 200	< 160	> 160	> 200	incl.
$\Delta\phi(\ell, \vec{p}_{\text{T}}^{\text{miss}})$	> 1.2	-	> 1.2	incl.	incl.	incl.
$\Delta R(b_1, b_2)$	incl.	incl.	incl.	< 1.2	> 1.2	incl.
$\lvert\Delta\phi(j_{1,2}, \vec{p}_{\text{T}}^{\text{miss}})\rvert$	> 0.4	> 0.4	> 0.4	> 0.4	> 0.4	-
$\lvert\Delta\phi(j_i, \vec{p}_{\text{T}}^{\text{miss}})\rvert$	> 1.0	incl.	incl.	incl.	incl.	incl.
m_{T2}^{τ} [GeV]	> 80	> 80	> 80	> 80	> 80	incl.

Table 7.5: Control region definitions corresponding to DM_low SR for the one-lepton $DMt\bar{t}$ analysis [42]. The entry of "incl." represents an inclusive selection with no requirements.

Variable [Unit]	DM_low_loose	TCR	WCR	STCR	TZCR
$\mathcal{N}_\ell$	1	1	1	1	3 (2 SFOS)
$p_T(\ell_1)$ [GeV]	> 25	> 25	> 25	> 25	> 25
Lepton charge	incl.	incl.	$+1$	incl.	incl.
$\mathcal{N}_{\text{jets}}$	≥ 4	≥ 4	≥ 4	≥ 4	≥ 4
$\mathcal{N}_{b-\text{jets}}$	≥ 1	≥ 1	$= 0$	≥ 2	≥ 1
$p_T(b_1)$ [GeV]	≥ 25	≥ 25	≥ 25	≥ 25	≥ 25
$p_T(j_1, j_2)$ [GeV]	$\geq (60, 60)$	$\geq (60, 60)$	$\geq (60, 60)$	$\geq (60, 60)$	$\geq (60, 60)$
$p_T(j_3, j_4)$ [GeV]	$\geq (40, 25)$	$\geq (40, 25)$	$\geq (40, 25)$	$\geq (40, 25)$	$\geq (40, 25)$
$E_{\text{T}}^{\text{miss}}$ [GeV]	> 300	> 230	> 230	> 230	incl.
$H_{\text{T,sig}}^{\text{miss}}$	> 14	> 8	> 8	> 8	incl.
m_T [GeV]	> 120	$[30, 90]$	$[30, 90]$	$[30, 120]$	incl.
am_{T2} [GeV]	> 140	$[100, 200]$	> 100	> 200	incl.
$\Delta\phi(\ell, \vec{p}_{\text{T}}^{\text{miss}})$	> 0.8	> 0.8	> 0.8	incl.	incl.
$\Delta R(b_1, b_2)$	incl.	incl.	incl.	> 1.8	incl.
$\lvert\Delta\phi(j_i, \vec{p}_{\text{T}}^{\text{miss}})\rvert$	> 1.4	> 1.4	> 1.4	> 1.4	incl.
m_{T2}^{τ} [GeV]	> 80	> 80	> 80	> 80	incl.

Table 7.6: Control region definitions corresponding to DM_low_loose SR for the one-lepton $DMt\bar{t}$ analysis [42]. The entry of "incl." represents an inclusive selection with no requirements.

	tN_med	DM_high	DM_low	DM_low_loose
Observed events	50	5	13	65
Fitted bkg events	36.3 ± 6.6	7.4 ± 2.1	13.8 ± 3.6	48.3 ± 8.2
$t\bar{t}2\ell$	12.1 ± 2.9	0.82 ± 0.27	2.21 ± 0.58	16.0 ± 1.3
$t\bar{t}1\ell$	0.19 ± 0.05	-	0.07 ± 0.03	-
Single top	3.5 ± 1.2	0.33 ± 0.16	0.65 ± 0.57	3.4 ± 1.3
$t\bar{t}V$	14.2 ± 5.5	4.0 ± 2.0	6.7 ± 3.2	14.3 ± 5.9
W+jets	4.2 ± 1.1	1.64 ± 0.53	3.2 ± 1.3	11.0 ± 2.8
Z+jets	-	-	-	-
Diboson	1.08 ± 0.7	0.66 ± 0.21	0.98 ± 0.33	3.6 ± 1.3

Table 7.7: Background-only fit results for the one-lepton $DMt\bar{t}$ analysis SRs [42]. The quoted uncertainties on the fitted SM background include both the statistical and systematic uncertainties.

7.3.2 Two-Lepton Analysis

Event Selection

The final state searched for in the two-lepton $DMt\bar{t}$ analysis includes two leptons and two b-jets coming from the decay of two top quarks and $E_{\mathrm{T}}^{\mathrm{miss}}$. Selected events [41] are then required to have at least one b-tagged jet with $p_T > 30$ GeV and exactly two opposite-sign leptons, indifferently same- or different-flavor, with an invariant mass $m_{\ell\ell}$ larger than 20 GeV. In addition, same-flavor lepton pairs events with $m_{\ell\ell}$ less than 20 GeV apart from the Z-boson mass are vetoed. Events are required to pass a two-lepton trigger and the leading and subleading lepton transverse momenta in the event need to be at least 25 and 20 GeV.

The main backgrounds for the two-lepton $DMt\bar{t}$ signature come from:

- Reducible backgrounds, such as Z+jets and dibosons, with moderate hadronic activity
- $t\bar{t}$ and Wt production, with the same final state as the signal, but with $E_{\mathrm{T}}^{\mathrm{miss}}$ deriving only from W decays and therefore kinematically bounded
- Irreducible backgrounds, like ttZ and ttW events with Z decaying to neutrinos or with the loss of a lepton from the decay of the W.

In order to deal with these backgrounds, the kinematic variables introduced in Secs. 6.3.1-6.3.3 are exploited, and the selection cuts summarised in Tab. 7.8 are used to define the final signal region (SRt3) [42].

Background Estimation

The production of dileptonic $t\bar{t}$ and $t\bar{t}Z$, with the Z boson decays into neutrinos, are the most important backgrounds in the SR for the two-lepton analysis. The dileptonic top background is estimated in a dedicated CR, selecting events that fail the c_{EM} requirement (CRTt3). For the $t\bar{t}Z$ background, a three-lepton CR (CR3ℓ) is used, with at least one same-flavour-opposite-charge (SFOS) lepton pair compatible with a Z-boson decay. To better mimic the event topology of the SR, a corrected $E_{\mathrm{T}}^{\mathrm{miss}}$ vector is defined by adding vectorially the transverse momenta of the SFOS pair to the original $E_{\mathrm{T}}^{\mathrm{miss}}$ vector and it is used to recalculate the c_{EM} variable, for which m_{T2} is substituted by the variable m_T (Sec. 6.3.2), calculated with the lepton non-identified as coming from the Z boson. The selection cuts for the two CRs are summarised in Tab. 7.8.

Results

The results of the background-only for the two-lepton $DMt\bar{t}$ analysis are presented in Tab. 7.9. Data and MC expectations show a good agreement and

Variable [Unit]	SRt3	CRt3	CR3ℓ
$\mathcal{N}_\ell$	2 OS	2 OS	3 (1 SFOS)
$p_T(\ell_1, \ell_2)$ [GeV]	$> (25, 20)$	$> (25, 20)$	$> (25, 20)$
$m_{\ell\ell}$ [GeV]	> 20	incl.	incl.
$\mathcal{N}_{\text{jet}}$	≥ 1	≥ 1	≥ 3 or ≥ 4
$\mathcal{N}_{b-\text{jet}}$	≥ 1	≥ 1	≥ 2 or $= 1$
$p_T(b_1)$ [GeV]	≥ 30	≥ 30	≥ 30
$p_T(j_1, j_2)$ [GeV]	≥ 30	≥ 30	≥ 30
$E_{\text{T}}^{\text{miss}}$ [GeV]	incl.	incl.	> 80
$\lvert m_{\ell\ell}^{SF} - m_Z \rvert$ [GeV]	> 20	> 20	< 10
$\Delta\phi_b$	< 0.8	incl.	incl.
$\min m_{b\ell}$ [GeV]	< 170	< 170	< 170
c_{EM} [GeV]	> 170	< 150	> 120
m_{T2} [GeV]	> 100	incl.	incl.

Table 7.8: Signal and control region definitions for the two-lepton $DMt\bar{t}$ analysis [41]. The entry of "incl." represents an inclusive selection with no requirements. In the final states with three leptons, the corrected variables of $E_{\text{T}}^{\text{miss}}$, $\min m_{b\ell}$ and m_{T2} are used instead as described in the main text.

exclusion limits were thus set.

Fig. 7.5 show upper limits at 95% C.L. on the ratio of the excluded cross-section to the nominal cross-section for a coupling assumption of $g = 1$ also for the two-lepton analysis. From the plot, it is possible to observe that the sensitivity for two-lepton $DMt\bar{t}$ production is approximately constant for masses below 100 GeV, and decreases rapidly for increasing masses of the mediator. The observed results in slightly worst than the expected one due to a mild excess of data in the signal region (Tab. 7.9) [41].

These results and the one-lepton ones will be reinterpreted in the next section in the framework of the 2HDMa model, in order to understand which areas of its parameter space are already excluded.

	SRt3
Observed events	18
Fitted bkg events	15.2 ± 4.3
$t\bar{t}$	4.5 ± 2.5
Single top	$0.33^{+0.53}_{-0.33}$
Diboson	0.61 ± 0.25
$t\bar{t}V$	4.4 ± 1.9
$t\bar{t}h$	0.06 ± 0.01
tWZ	0.84 ± 0.15
Fakes	2.7 ± 1.3
Others	2.69 ± 0.93

Table 7.9: Background-only fit results for the two-lepton $DMt\bar{t}$ analysis SR [41]. The quoted uncertainties on the fitted SM background include both the statistical and systematic uncertainties.

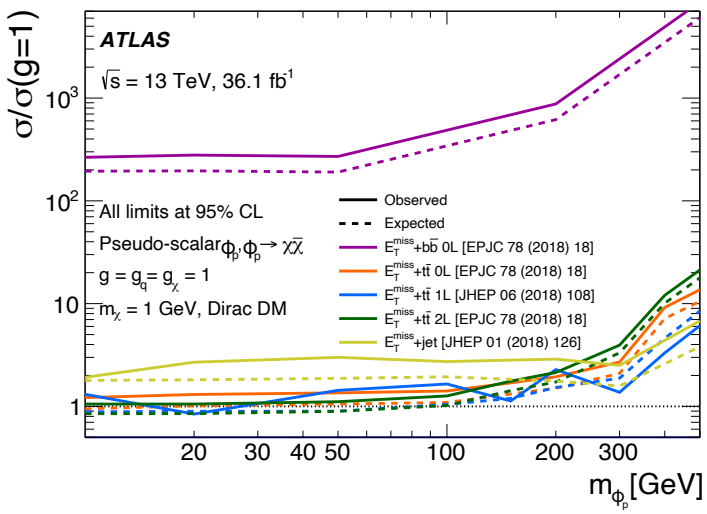


Figure 7.5: Exclusion limits for DMSIMP as a function of the mediator mass for a DM mass of 1 GeV for 36.1 fb^{-1} ATLAS analyses. The limits are calculated at 95% C.L. [42, 41].

7.4 Recasting of ATLAS Results

The results of the analyses performed by the ATLAS Collaboration were obtained in terms of the excluded number of events from BSM physics (N_{exc}), which in turn can be expressed as an excluded production cross-section (σ_{exc}) for the process considered. In formulas, for each value of mediator mass, the exclusion cross-section is given by

$$\sigma_{\text{exc}} = \frac{N_{\text{exc}}}{L \cdot \epsilon \cdot \mathscr{A}}, \tag{7.8}$$

where L is the integrated luminosity, $\mathscr{A}$ is the acceptance at parton level for the requirements defining the analysis, and ϵ is the efficiency of the ATLAS experiment in detecting the final state topology.

In order to map this exclusion in the parameter space of the 2HDMa model, one starts from the value of N_{exc} and has to reconstruct the value of σ_{exc} for the specific model using Eq. 7.8. For this aim, since the ATLAS results are expressed in term of σ_{exc} for DMSIMP, the values of ϵ and $\mathscr{A}$ for both models are needed. The value of ϵ is assumed to be the same for any model once the final state topology has been defined by the cuts, whereas $\mathscr{A}$ depends on the details of the model itself. The excluded cross-section for the 2HDMa model is thus given by:

$$\sigma_{\text{exc}}(\text{2HDMa}) = \sigma_{\text{exc}}(\text{DMSIMP}) \cdot \frac{\mathscr{A}(\text{DMSIMP})}{\mathscr{A}(\text{2HDMa})}. \tag{7.9}$$

In case the acceptance for the 2HDMa model is the same of the DMSIMP model for a given set of analysis cuts, one can simply take the excluded cross-section of DMSIMP without additional steps. It is then necessary to understand whether the analysis acceptance for $DMt\bar{t}$ in 2HDMa and DMSIMP are the same or not. As stated before, there are two different processes contributing to $DMt\bar{t}$ in the 2HDMa model (Eq. 7.2), which, individually taken, are identical to $DMt\bar{t}$ production in DMSIMP if the pseudoscalars a or A (respectively) are identified with the mediator ϕ_p.

In order to verify the effect of the presence of two different processes, the assumption that a corresponds to the DMSIMP mediator ϕ_p was made and it was investigated how the production of A influences the acceptance. To this effect, two benchmark values for m_{ϕ_p} were taken into account, 150 GeV and 300 GeV, and the compatibility of the two models at parton level was studied in detail. The parameters of the 2HDMa model for the benchmark generation were chosen such as the production cross-section for the process $pp \to t\bar{t}a \to t\bar{t}\chi\bar{\chi}$ and the width of the mediator were the same as in DMSIMP. Since a is produced on-shell, the production cross-section in narrow-width approximation is determined by the mass of a and its coupling to the top quark, which scales as the ratio of $\sin\theta$ and

$m_a = 300$ GeV				
Model	Γ_a (GeV)	σ(pb)	$\mathscr{A}$	$\mathscr{A}(m_{\chi\chi} < 400$ GeV)
DMSIMP	11.93	0.0036	0.1061±0.0013	0.1059±0.0013
2HDMa	11.94	0.0037	0.1064±0.0013	0.1048±0.0013

$m_a = 150$ GeV				
Model	Γ_a (GeV)	σ(pb)	$\mathscr{A}$	$\mathscr{A}(m_{\chi\chi} < 200$ GeV)
DMSIMP	5.97	0.0111	0.0749±0.0012	0.0745±0.0011
2HDMa	5.97	0.0112	0.0744±0.0012	0.0733±0.0011

Table 7.10: Mediator width, production cross-section, acceptance for all events and acceptance restricted to the a peak for the two models under consideration, and for the two benchmark points described in the text.

$\tan\beta$ (Eq. 7.7). The width is proportional to $\cos\theta$ and y_χ. A convenient set of parameters is thus $\tan\beta = 1/\sqrt{2}$, $\sin\theta = 1/\sqrt{2}$ - corresponding to the full mixing between a and A - and $y_\chi = \sqrt{2}$. The masses of the Higgs bosons of the model were set to $m_h = 125$ GeV and $m_A = m_{H^\pm} = m_H = 750$ GeV. The cross-sections for the $DMt\bar{t}$ process and the widths of the pseudoscalar a and of the mediator ϕ_p are shown for the parameters given above in Tab. 7.10.

The first parton level comparison performed is shown in Fig. 7.6, which presents the distribution of the invariant mass of the two DM particles. There is a clear difference between the 2HDMa and the DMSIMP model: in 2HDMa there is a contribution from the production of the A pseudoscalar and from the interference between a and A production. The distributions are in logarithmic scale, and the contribution from A production for the chosen sets of parameters is at the level of a few percent, yielding a very small effect on final acceptances. However, the effect may be significantly larger for different parameter choices, calling for a specific strategy to handle this situation when mapping the excluded areas in 2HDMa parameter space.

To further compare the kinematics of the two models, events were selected near the a peak to minimise the A contribution. $m_{\chi\chi}$ was required to be smaller than 200(300) GeV for $m_a = 150(300)$ GeV respectively. The distributions for the invariant mass of the two top quarks (m_{tt}), the difference in azimuthal angle ($\Delta\phi(tt)$), the relative polar angle ($\cos\theta(tt)$) and the transverse momentum of the $\chi\chi$ system are shown in Fig. 7.7 for $m_{\phi_p} = m_a = 150$ GeV. An excellent agreement is observed between DMSIMP and 2HDMa.

A further truth comparison was performed at particle level, on the variables used as main discriminants in the two-lepton analysis described in Sec. 7.3.2.

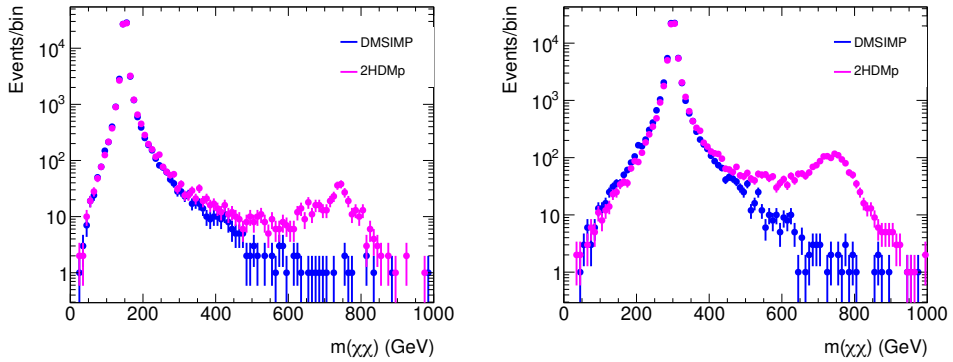


Figure 7.6: Comparison of $m(\chi\chi)$, the invariant mass of the two DM particles for the DMSIMP (blue) and the 2HDMa model (magenta). The plot on the left (right) shows the comparison for $m_a = 150(300)$ GeV respectively.

The distributions of the variables m_{T2} and $E_{\mathrm{T}}^{\mathrm{miss}}$ are shown in Fig. 7.8, always for $m_{\phi_p} = m_a = 150$ GeV. In this case as well the agreement between DMSIMP and 2HDMa is excellent, giving confidence that, once the contribution from A production is separated, the acceptance of the two-lepton $DMt\bar{t}$ analysis is the same for the two models.

For the chosen benchmarks, the A contribution is so small that the 2HDMa acceptances for the two-lepton analysis of Sec. 7.3.2 with and without the $m_{\chi\chi}$ requirements are the same within the statistical error, and perfectly match the ones for the DMSIMP model, as shown in Table 7.10. However, it is worth considering a situation where the ratio between A and a production is larger, to verify if in this case the production of A modifies the acceptance for the analysis. A possible parameters choice is $m_a = 150$ GeV, $m_A = 600$ GeV and $\sin\theta = 0.35$. With respect to the benchmark studied above, the lower value of $\sin\theta$ enhances the production of A. The study was performed for values of $\tan\beta$ varying between 0.5 and 10. The observed acceptance is shown in Fig 7.9 as black circular markers, which show a significant dependence on $\tan\beta$; this dependence is given by the contribution of A production. This is further stressed by the square markers in Fig. 7.9, calculated only for events with $m_{\chi\chi} < 200$ GeV, which show an excellent agreement within statistical errors with the DMSIMP acceptance for $m_{\phi_p} = 150$ GeV, shown as a continuous line, with the dashed lines giving the interval of the statistical error.

It is thus necessary to develop an algorithm for mapping the DMSIMP acceptance into the 2HDMa one, taking into account the production of two separate

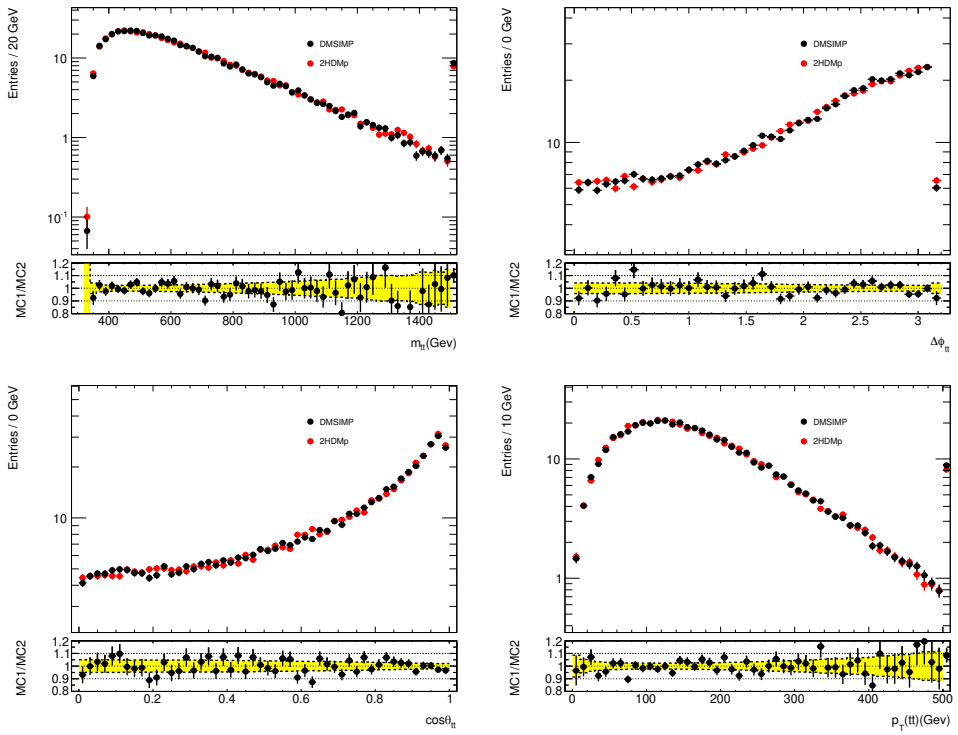


Figure 7.7: Comparison of the distributions of the variables $m(tt)$ (top left), $(\Delta\phi(tt))$ (top right), $(\cos\theta(tt))$ (bottom left) and $p_T(tt)$ (bottom right) as defined in the text for the DMSIMP model (black points) and 2HDMa model (red points). The distributions are provided for $m_a = 150$ GeV, and the 2HDMa distributions are for events passing the $m_{\chi\chi} < 200$ GeV requirement.

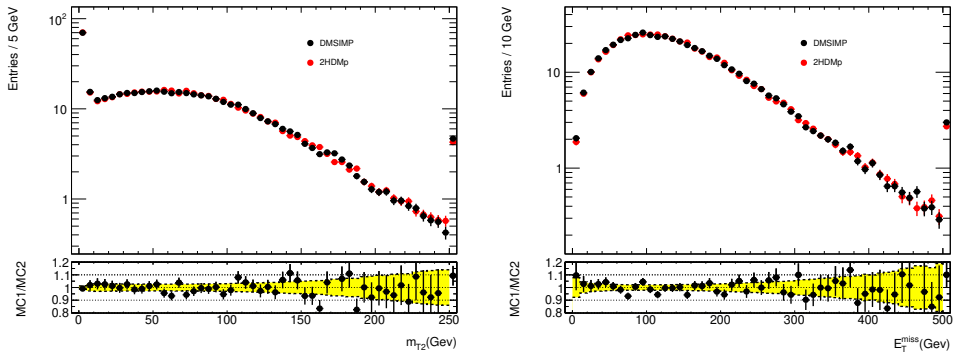


Figure 7.8: Comparison of the distributions of the variables m_{T2} (left) and $E_{\mathrm{T}}^{\mathrm{miss}}$ (right) for the DMSIMP model (black points) and 2HDMa model (red points). The distributions are provided for $m_{\phi_p}/m_a = 150$ GeV, and the 2HDMa distributions are for events passing the $m_{\chi\bar{\chi}} < 200$ GeV requirement.

pseudoscalar bosons. Having verified that the two models yield identical kinematical distributions for the production of a single particle, two input ingredients for the recasting are needed: the dependence of the acceptance on the mass and the width of the mediator. In particular, it is important to verify that the dependence of the acceptance on the width is small, as the acceptance for the DMSIMP analyses was evaluated for a specific value of the coupling, and hence of the widths. In case of a significant acceptance variation with the width, correction factors would have to be applied in the recasting. The width dependence was studied for a 300 GeV mediator and for values of the coupling between 0.1 and 4, yielding widths between ∼100 MeV and ∼170 GeV. The acceptance, illustrated in Fig. 7.10, shows no dependence on the mass as long as the width is smaller than 50% of the mediator mass. The mass dependence was studied for mass values of the DMSIMP mediator between 50 GeV and 1 TeV, and is shown in Fig. 7.11. The acceptance slowly grows with increasing m_{ϕ_p}, as large m_{ϕ_p} give final states with harder $E_{\mathrm{T}}^{\mathrm{miss}}$, favoured by the experimental cuts.

Based on these results, a possible approach for evaluating the acceptance for the 2HDMa model is to take, for each set of parameters, the average of the acceptances for the DMSIMP model for mediator mass equal to m_a and m_A, weighted by the respective cross-sections:

$$\mathscr{A}_{\mathrm{2HDMa}}(m_A, m_a) = \frac{\sigma_a \cdot \mathscr{A}_{\mathrm{DMSIMP}}(m_a) + \sigma_A \cdot \mathscr{A}_{\mathrm{DMSIMP}}(m_A)}{\sigma_a + \sigma_A}, \tag{7.10}$$

where σ_a is the cross-section for the process $pp \to t\bar{t}a \to t\bar{t}\chi\bar{\chi}$ and σ_A the one for

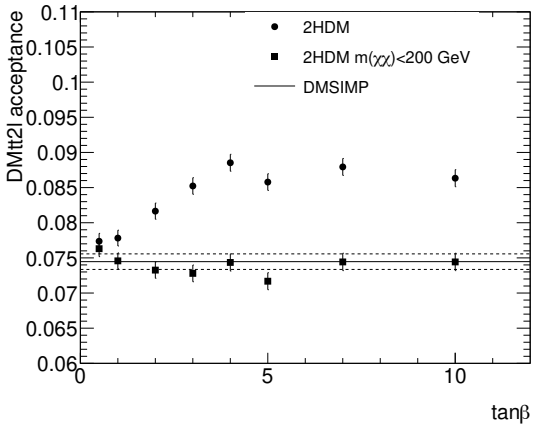


Figure 7.9: Acceptance of the two-lepton analysis as a function of $\tan\beta$ for the 2HDMa model (round markers), for the 2HDMa model considering only events with $m_{\chi\chi} < 200$ GeV (square markers), and for the DMSIMP model (full line) for a mediator mass of 150 GeV. The two dashed lines indicate the statistical error of the DMSIMP. The value of m_A is fixed at 600 GeV, and $\sin\theta = 0.35$.

$pp \to t\bar{t}A \to t\bar{t}\chi\bar{\chi}$.

The acceptance estimated in this way for the parameter choice used above ($m_a = 150$ GeV, $m_A = 600$ GeV and $\sin\theta = 0.35$) is shown as red triangles in Fig. 7.12, and an excellent agreement can be seen with the acceptances evaluated directly on 2HDMa samples.

An additional test of the viability of this approach was performed by addressing the situation in which the masses of a and A are very similar, implying the possibility of a large interference between the production of the two bosons. The analysis acceptance was evaluated both directly and with the procedure defined in Eq. 7.10 for six samples with $m_a = 300$ GeV, $m_A = 250$ or 350 GeV, and different assumptions on the mixing in the pseudoscalar sector, as presented in Tab. 7.11. As it can be seen by comparing the sum of σ_a and σ_A with the cross-section for the full process, labelled as σ_{tot} in the table, the considered benchmarks include a significant level of interference between the two pseudoscalar bosons. The analysis acceptance of the six models is shown in Fig. 7.13, showing that the rescaling of the DMSIMP acceptances with the prescription (7.10) yields the same acceptance as the 2HDMa model despite the interference.

Finally, the last check was to verify at TRUTH level the validity of the approach described above also for the official ATLAS signal samples used for the existing analyses. The acceptances of both one- and two-lepton signal regions

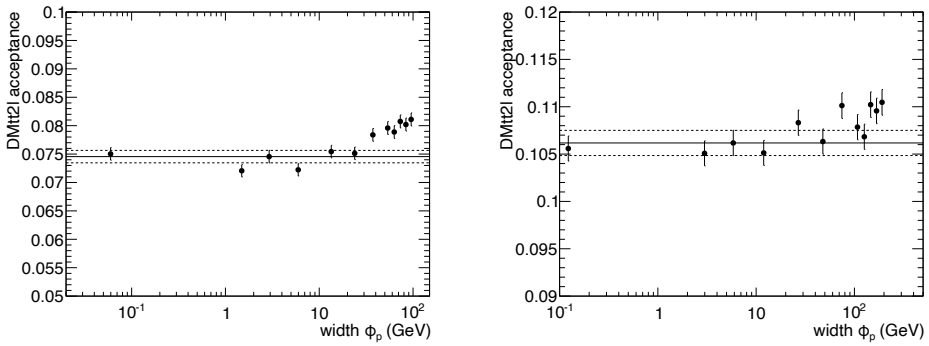


Figure 7.10: Acceptance of the two-lepton analysis as a function of the width of the mediator for a mediator mass $m_{\phi_p} = 150$ GeV (left) and $m_{\phi_p} = 300$ GeV (right).

Param set	m_A (GeV)	$\sin\theta$	σ_a (pb)	σ_A (pb)	σ_{tot} (pb)
1	250	0.7	0.038	0.057	0.078
2	350	0.7	0.040	0.021	0.048
3	250	0.3	0.007	0.100	0.091
4	350	0.3	0.007	0.013	0.016
5	250	0.95	0.061	0.011	0.061
6	350	0.95	0.070	0.0053	0.064

Table 7.11: Definition of the models used for the study addressing the interference between a and A production, and values of the production cross-section.

evaluated on 2HDMa samples was thus compared with the one obtained rescaling DMSIMP samples using Eq. 7.10. The results are shown in Figs. 7.14 and 7.15, and a good agreement is observed in all regions, confirming the validity of the rescaling technique. The values of $\mathscr{A}$ given in Figs. 7.14-7.15 for the two-lepton analysis differ of a factor of about 25 from the acceptances shown in Fig. 7.11 due to the fact that the samples previously used have both top quarks forced to decay semileptonically into electrons or muons, while no specific decay is assumed for the official samples.

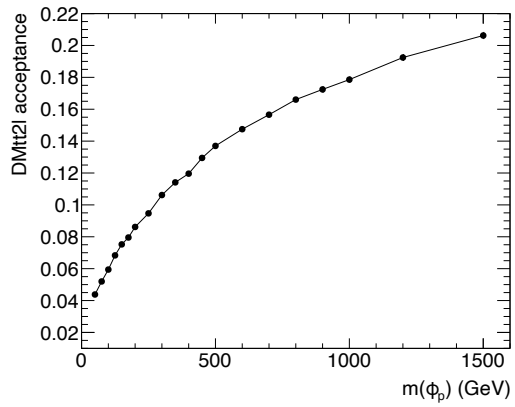


Figure 7.11: Acceptance of the two-lepton analysis as a function of the mediator mass m_{ϕ_p}. The value of the coupling is one for all mass points.

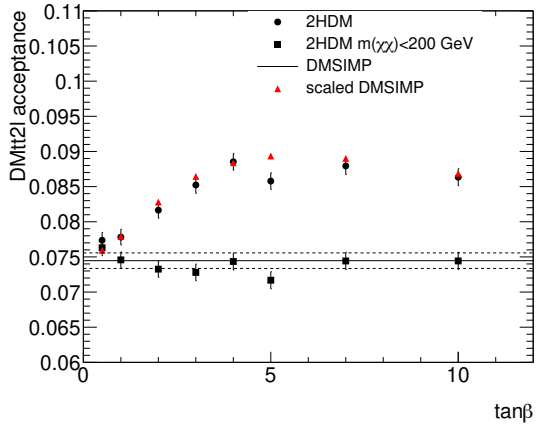


Figure 7.12: Same as Fig. 7.9, with in addition the acceptance calculated following the prescription of Eq. 7.10 (red triangles).

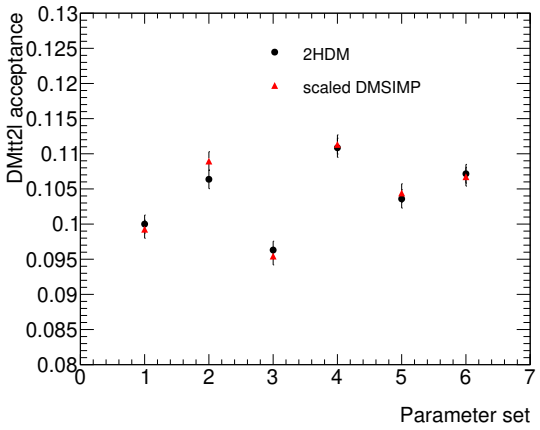


Figure 7.13: Acceptance of the two-lepton analysis for the 2HDMa model (black circles) compared to the DMSIMP acceptance for $m_{\phi_p} = m_a = 300$ GeV rescaled with the prescription (7.10).

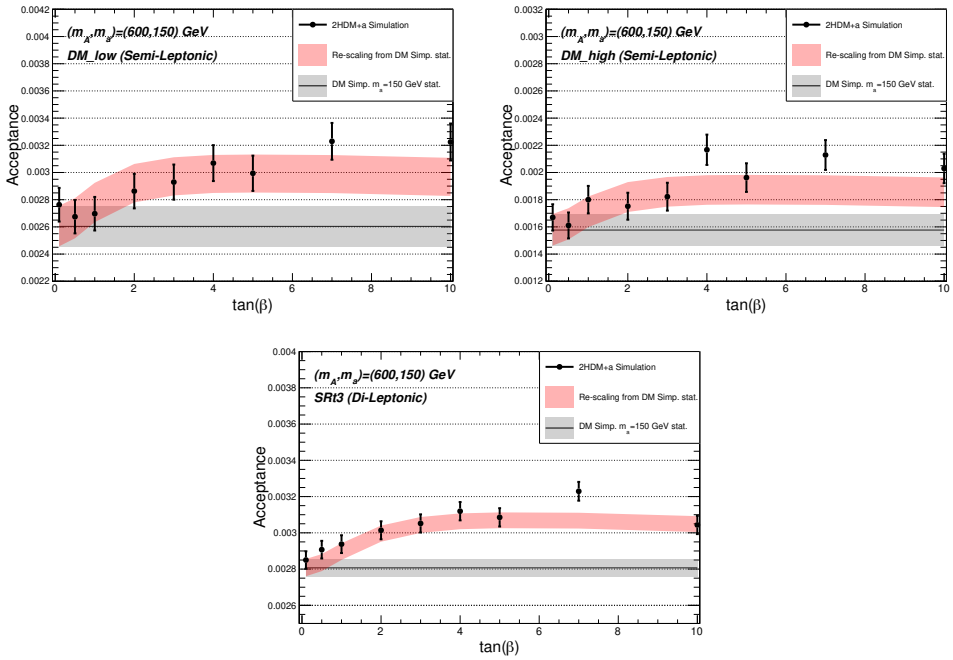


Figure 7.14: Acceptance of the one-lepton (top) and two-lepton (bottom) analyses for the 2HDMa model (black dots) compared to the DMSIMP acceptance rescaled with the prescription (7.10) (red bands) as a function of $\tan\beta$ [44].

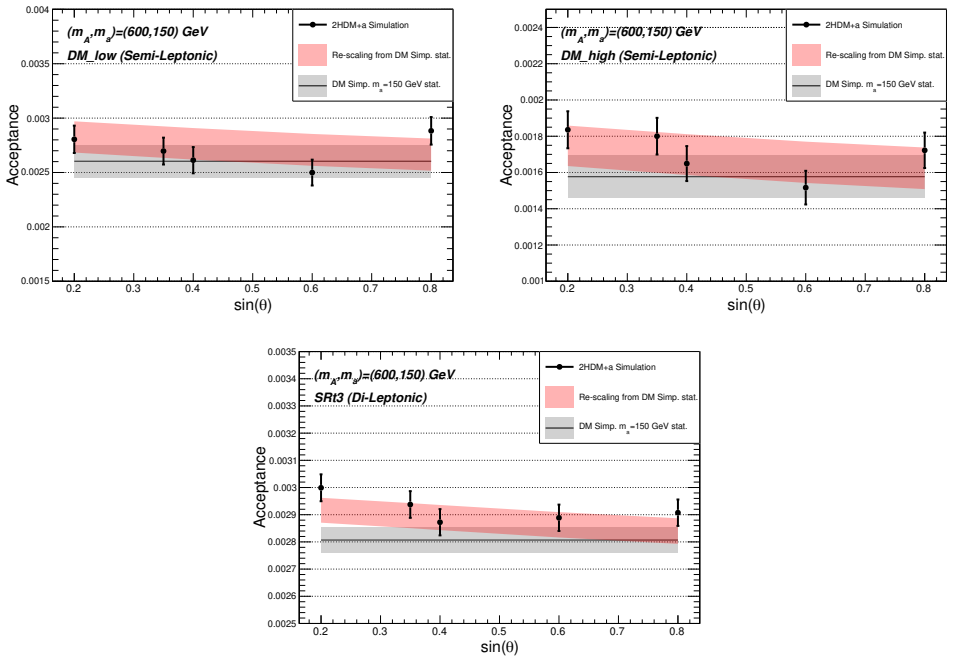


Figure 7.15: Acceptance of the one-lepton (top) and two-lepton (bottom) analyses for the 2HDMa model (black dots) compared to the DMSIMP acceptance rescaled with the prescription (7.10) (red bands) as a function of $\sin\theta$ [44].

7.5 Results of Recasting in the Parameter Space of 2HDMa

Based on the results of the previous section, it is possible to recast the experimental limits on the production of $DMt\bar{t}$ obtained by the ATLAS Collaboration for the DMSIMP model into the parameter space of 2HDMa.

From Eq. 7.9, the main inputs for the calculation of the excluded cross-section in the 2HDMa model are the excluded cross-section of the ATLAS searches for $DMt\bar{t}$ production in the DMSIMP model, and the ATLAS experimental acceptance as a function of m_{ϕ_p}. These values are made available by the ATLAS Collaboration in tabular form for a range of mediator masses [41, 42], and are presented in Table 7.12 and 7.13. When necessary, the data are interpolated to get the acceptance and exclude cross-section also for values of m_{ϕ_p} not included in the tables. For the one-lepton analysis, where more than one signal region is available, the data relative to the SR providing the best expected exclusion are used. During the recasting procedure, the same SR is used when evaluating the contributions from both m_a and m_A, selecting the one with the best expected exclusion for $m_{\phi_p} = m_a$. The acceptances for $m_{\phi_p} = 600$ GeV are reported for all SRs for this reason.

m_{ϕ_p} (GeV)	1L $\mathscr{A}$	2L $\mathscr{A}$
50	$2.26 \cdot 10^{-3}$	$1.68 \cdot 10^{-3}$
100	$3.30 \cdot 10^{-3}$	$2.35 \cdot 10^{-3}$
150	$3.11 \cdot 10^{-3}$	$2.89 \cdot 10^{-3}$
200	$9.29 \cdot 10^{-3}$	$3.18 \cdot 10^{-3}$
300	$3.13 \cdot 10^{-3}$	$4.04 \cdot 10^{-3}$
400	$7.21 \cdot 10^{-3}$	$4.41 \cdot 10^{-3}$
500	$7.66 \cdot 10^{-3}$	$5.04 \cdot 10^{-3}$
600	$1.21 \cdot 10^{-2}$ (tN_med) $1.11 \cdot 10^{-2}$ (DM_high) $1.44 \cdot 10^{-2}$ (DM_low) $2.42 \cdot 10^{-2}$ (DM_low_loose)	$5.54 \cdot 10^{-3}$

Table 7.12: Acceptance of the selections presented in Secs. 7.3.1-7.3.2 for DMSIMP as a function of mediator mass [42, 41].

On the basis of these numbers, for each (m_a, m_A, $\sin\theta$, $\tan\beta$) combination it is possible to calculate the value of σ_{exc}(2HDMa) by using Eqs. 7.9 and 7.10, once the values σ_a and σ_A are available.

Four benchmark scenarios that are consistent with bounds from electroweak precision, flavour and Higgs observables have been recommended to investigate

m_{ϕ_p} (GeV)	1L $\sigma_{\text{exc}}^{\text{Exp}}$ (pb)	1L $\sigma_{\text{exc}}^{\text{Obs}}$ (pb)	2L $\sigma_{\text{exc}}^{\text{Exp}}$ (pb)	2L $\sigma_{\text{exc}}^{\text{Obs}}$ (pb)
50	0.3544	0.5651	0.3555	0.4380
100	0.2582	0.4082	0.2519	0.3125
150	0.1916	0.1809	0.2272	0.2820
200	0.1644	0.2455	0.1868	0.2301
300	0.1004	0.0732	0.1773	0.2107
400	0.0723	0.0521	0.1586	0.1895
500	0.0456	0.0456	0.1335	0.1591
600	0.0456	0.0456	0.1335	0.1591

Table 7.13: ATLAS analyses expected and observed excluded cross-sections for $DMt\bar{t}$ production in the DMSIMP model for different values of m_{ϕ_p}.

the sensitivity to 2HDMa by a dedicated LHC Dark Matter Working Group [44]. The benchmarks are defined as:

1. 2D scan in (m_a, m_A) assuming $\tan\beta = 1$, $\sin\theta = 0.35$.
2. 2D scan in (m_a, $\tan\beta$) assuming $m_A = 600$ GeV, $\sin\theta = 0.35$. In this case the A boson is massive enough to be sub-dominant, but the production cross-section for a is reduced.
3. Two 1D scans in $\sin\theta$ assuming
   1) $m_A = 600$ GeV, $m_a = 200$ GeV and $\tan\beta = 1.0$,
   2) $m_A = 1000$ GeV, $m_a = 350$ GeV and $\tan\beta = 1.0$
4. One scan in m_χ assuming $m_A = 600$ GeV, $m_a = 250$ GeV, $\tan\beta = 1.0$, $\sin\theta = 0.35$

The cross-sections $\sigma(m_a, m_A, \sin\theta, \tan\beta)$ for $pp \to t\bar{t}\chi\chi$ and its component σ_a and σ_A were calculated on these grids and fed in Eqs. 7.9 and 7.10 to obtain limits on these benchmarks.

For the benchmark 1, ATLAS $DMt\bar{t}$ analyses have proven not to be sensitive. The results of the recasting for benchmark 2 are shown in Fig. 7.16. For each value of m_a, the value of $\tan\beta$ such that

$$\sigma(m_a, m_A, \sin\theta, \tan\beta) = \sigma_{\text{exc}}(\text{2HDMa})$$

has been calculated through a quadratic interpolation on the generated grid, thus obtaining the border of the region of parameter space excluded. The result of the combination of one- and two-lepton analyses, obtained using for each mass point the signal region providing the best expected sensitivity, is also presented. For this benchmark, $DMt\bar{t}$ production is dominated by the contribution from a

production due to the high value of m_A. The cross-section drastically decreases for $m_a > 350$ GeV, where the BR of a in DM particles is suppressed by $a \to t\bar{t}$ (see Sec. 7.2.2), and no exclusion is thus possible in that region. For smaller values of m_a, $\tan\beta$ values as high as ~ 0.34 can be excluded.

The interpretation of the results for benchmark 3.1 are presented in Fig. 7.17. In order to be able to show the sensitivity of the analyses, the choice of $\tan\beta$ deviates from the default value proposed in the benchmark and has been set to 0.5. The same choice has been used for benchmark 3.2, shown in Fig. 7.18. The sensitivity of $DMt\bar{t}$ in the latter decreases due to the high value of m_a, as explained above.

Finally, the results of the recasting for benchmark 4 are shown in Fig. 7.19, setting also in this case $\tan\beta = 0.5$. The interpretation of the results stops at $m_\chi = 120$ GeV to avoid the kinematic region where $m_a < 2m_\chi$, which is not correctly handled by the generator and has very low sensitivity.

7.6 Conclusions

The searches for $DMt\bar{t}$ production performed by the ATLAS Collaboration with 36.1 fb^{-1} of pp collisions in the framework of DMSIMP are able to set limits on a part of the parameter space of the 2HDMa model. These results were presented in [1], alongside the results of other analyses reinterpreted in the context of the 2HDMa model (see Sec. 2.3). The final summary plots containing the results from all analyses are shown in Figs. 7.20, 7.21, 7.22 and 7.23 for the different benchmarks. As it can be seen from these plots, the final exclusion is dominated by other channels (in particular the resonant $E_{\mathrm{T}}^{\mathrm{miss}} + Z$ and $E_{\mathrm{T}}^{\mathrm{miss}} + h$) due to the parameter choices of the considered benchmark. The benchmark presented in [44] were in fact optimised based on preliminary results coming from the $E_{\mathrm{T}}^{\mathrm{miss}} + h$ analysis, disadvantaging the $DMt\bar{t}$ analyses, in particular for the choices of $\sin\theta$ and $\tan\beta$.

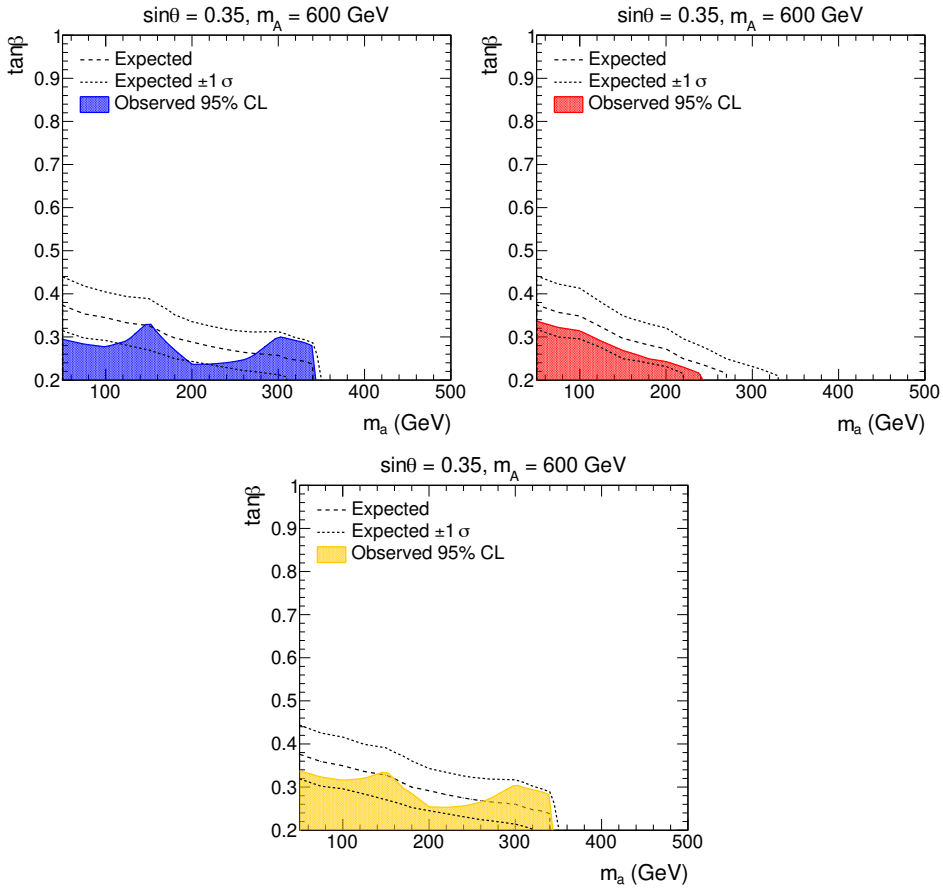


Figure 7.16: Regions in the $m_a - \tan\beta$ plane of the 2HDMa model excluded by ATLAS one-lepton (top left) and two-lepton (top right) $DMt\bar{t}$ analyses and their combination (bottom) for benchmark 2 (see text) [42, 41].

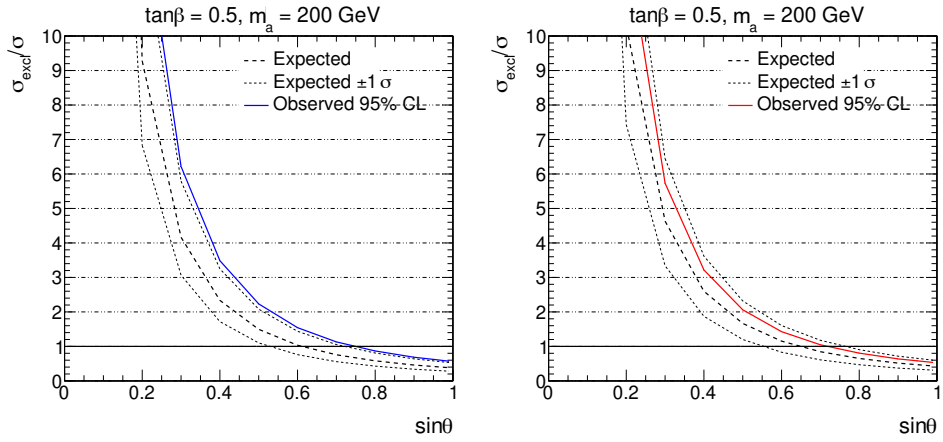


Figure 7.17: Exclusion limits of ATLAS one-lepton (left) and two-lepton (right) $DMt\bar{t}$ analyses as a function of $\sin\theta$ for 2HDMa benchmark 3.1 (see text) [42, 41].

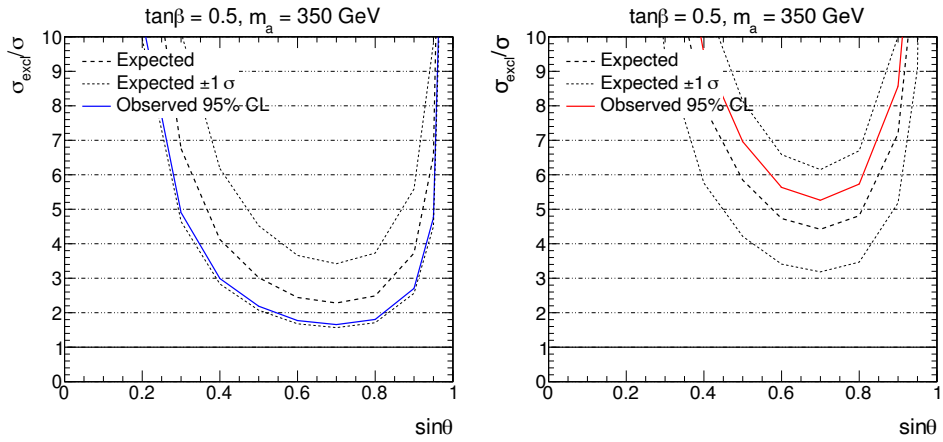


Figure 7.18: Exclusion limits of ATLAS one-lepton (left) and two-lepton (right) $DMt\bar{t}$ analyses as a function of $\sin\theta$ for 2HDMa benchmark 3.2 (see text) [42, 41].

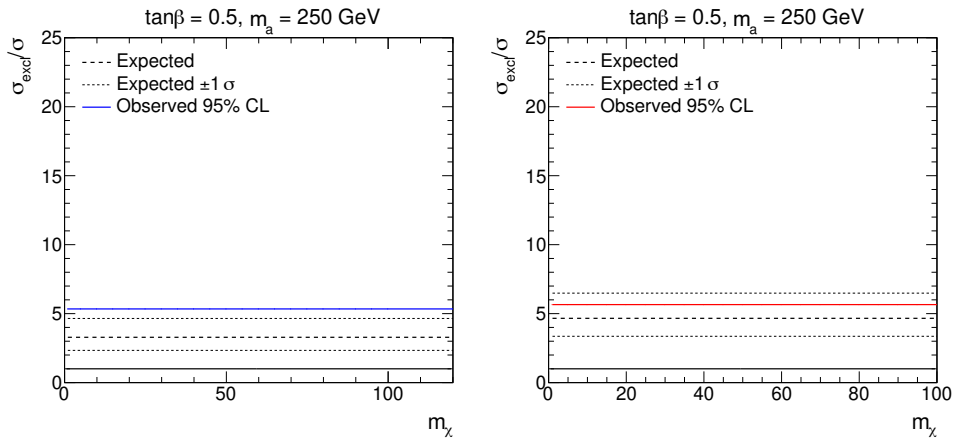


Figure 7.19: Exclusion limits of ATLAS onelepton (left) and two-lepton (right) $DMt\bar{t}$ analyses as a function of m_χ for 2HDMa benchmark 4 (see text) [42, 41].

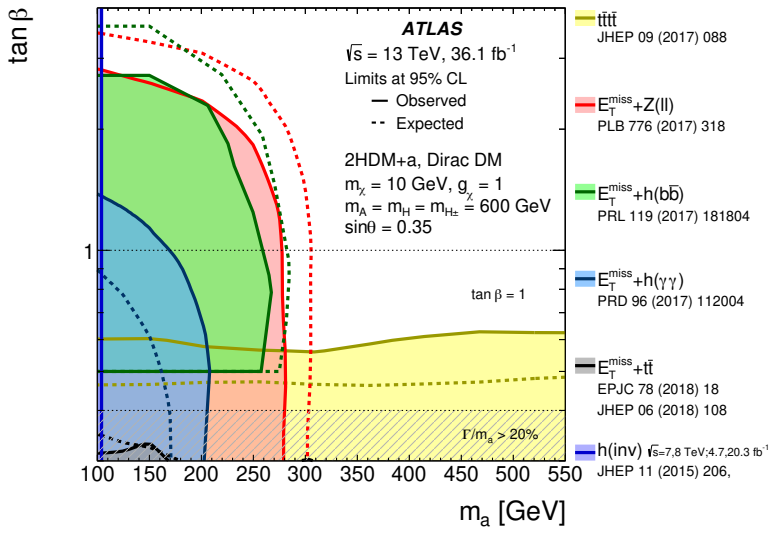


Figure 7.20: Exclusion limits at 95% C.L. of ATLAS analyses targeting the 2HDMa model for benchmark 2 (see text) [1].

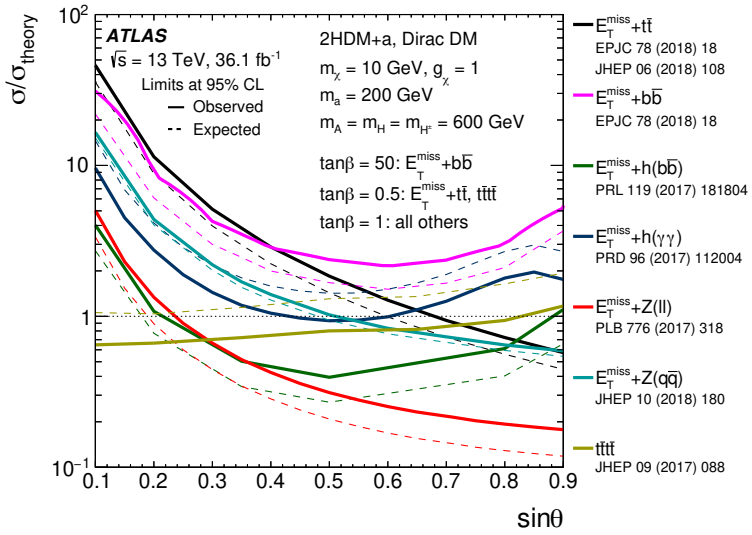


Figure 7.21: Exclusion limits at 95% C.L. of ATLAS analyses targeting the 2HDMa model for benchmark 3.1 (see text) [1].

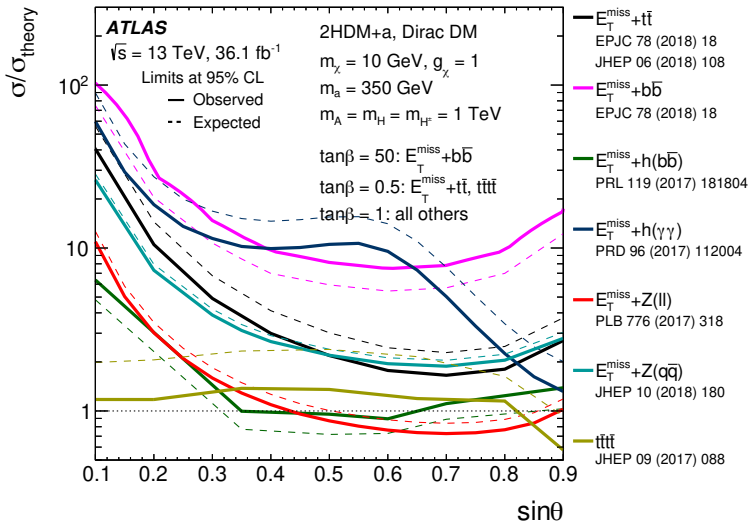


Figure 7.22: Exclusion limits at 95% C.L. of ATLAS analyses targeting the 2HDMa model for benchmark 3.2 (see text) [1].

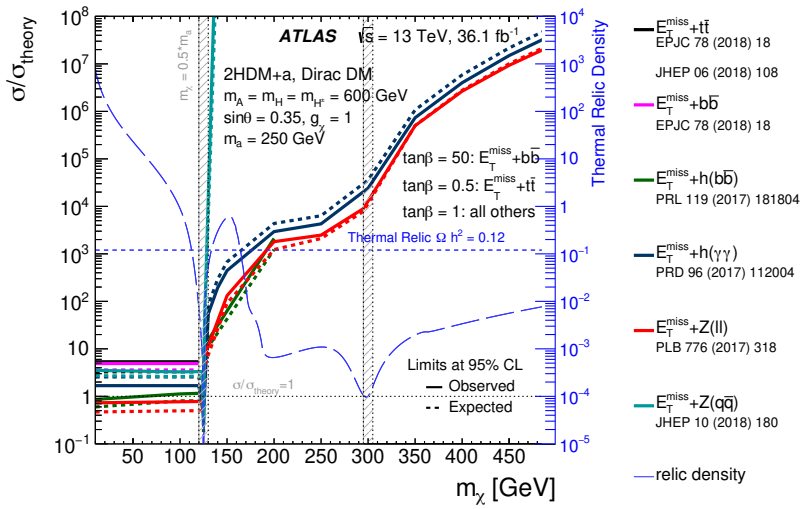


Figure 7.23: Exclusion limits at 95% C.L. of ATLAS analyses targeting the 2HDMa model for benchmark 4 (see text) [1].

Chapter 8

Search for Dark Matter Produced in Association with a Single Top Quark

As anticipated at the beginning of Chapter 7, this Chapter focuses on a new signature for dark matter production in the context of the 2HDMa model, where the dark matter is produced in association with a single top quark (called DMt in the following). Such a signature is the only identified channel sensitive to the production of charged Higgs bosons. This gives the DMt analysis a pivotal role in tackling the 2HDMa model and constraining its parameter space.

Like single top production within the SM, the DMt signature in the 2HDMa model receives three different types of contributions at leading order: t-channel production, s-channel production and associated production together with a W boson (tW). While s-channel production has proven to be negligible, both t-channel and tW productions were studied in detail in [2], using the full Run-2 statistics collected by the ATLAS experiment, corresponding to an integrated luminosity of 139 fb^{-1}. The author of this thesis is in particular the main analyser for the tW analysis in the two-lepton final state (called tW_{2L}), arising when both the associated W-boson and top quark decay leptonically, which will be the main focus on this Chapter. The author took care of all the main steps of the analysis, from the validation and generation of the signal samples to the optimisation of the signal region definition and the development of the strategy for the backgrounds estimation, finishing with the implementation of the statistical analysis. In order to maximise the sensitivity to DMt production, the author implemented also the combination of the two-lepton analysis with the one-lepton one, briefly introduced in this Chapter for this reason.

Given the similarity of the signatures, the contribution from $DMt\bar{t}$ production

in the signal regions of DMt analyses is not negligible. On the basis of the work presented in Chapter 7, the author developed a strategy to include the $DMt\bar{t}$ contribution in both the two- and one-lepton tW-DMt analyses to give a complete picture of the sensitivity of the searches to the 2HDMa model. This is presented at the end of the Chapter.

8.1 Signal Simulation

Signal samples of the process $pp \to tW\chi\bar{\chi}$ in the 2HDMa framework were simulated using the MadGraph5_aMC@NLO [72] v2.6.2 generator at leading order, interfaced to Pythia8 [76] with the A14 tune [70] for the modelling of parton showering, hadronisation and underlying event simulation. Parton distribution functions were provided by the NNPDF3.0 PDF set [60].

The additional samples for the $DMt\bar{t}$ process were generated with MadGraph5_aMC
@NLO v2.6.7 interfaced to Pythia8, using the same PDF set and tune as used for DMt. The samples were simulated at LO with up to one extra parton in the framework of the DMSIMP model, and then reinterpreted in the context of 2HDMa using a similar strategy to the one presented in Chapter 7 that will be explained in the following. $DMt\bar{t}$ cross-sections in the 2HDMa model were calculated at NLO accuracy using the same version of MadGraph5_aMC@NLO. No overlap between DMt and $DMt\bar{t}$ samples is present since the DMt ones are generated at LO.

In the analyses presented in the following, if not stated otherwise, the parameters choice presented in Sec. 2.2.6 are adopted, and the value of $\sin\theta$ is set to $1/\sqrt{2}$.

8.2 Phenomenology of DMt Production in the 2HDMa Model

8.2.1 Process Composition

As explained in the introduction, the production of dark matter in association with a single top quark receives three different contributions at leading order: s-channel production, t-channel production and associated production together with a W boson (tW).

DMt production in the s-channel is characterised by a very small cross-section when compared to the other channels. It is mostly proceed through the diagrams shown in Fig. 8.1, where the pseudoscalar mediator a is radiated from the top

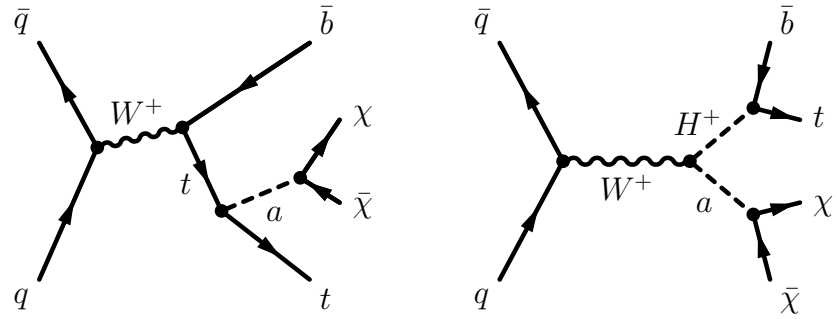


Figure 8.1: Feynman diagram contributing to DMt s-channel production in the 2HDMa model.

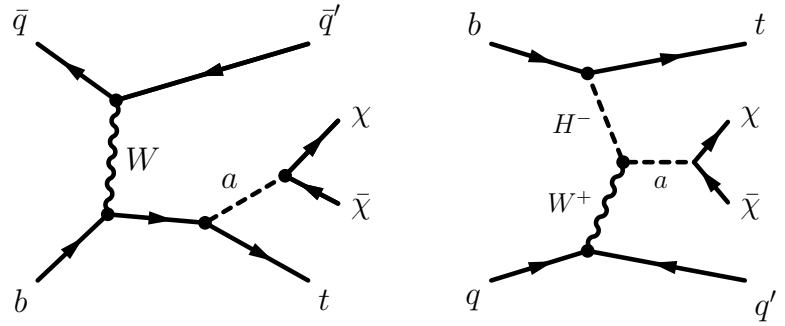


Figure 8.2: Feynman diagrams contributing to DMt t-channel production in the 2HDMa model.

quark in the typical SM single top s-channel diagram (so-called a-strahlung) or is emitted from a W boson together with a charged Higgs.

The t-channel process $pp \to tj\chi\bar{\chi}$ has mainly two different diagrams contributing to it, presented in Fig. 8.2. One is again the SM diagram with a-strahlung, while the other is the t-channel fusion of a charged Higgs and a W boson into the pseudoscalar mediator a. These two diagrams interfere destructively, with interference decreasing at higher $H^\pm$ mass. This destructive interference ensures the perturbative unitarity of the process in the 2HDMa model.

Also the tW production receives the contribution of two different diagrams (Fig. 8.3): the a-strahlung diagram and the associated production of a $H^\pm$ and a top quark. Like in the case of t-channel, the two diagrams interfere destructively to ensure unitarity. When the decay $H^\pm \to W^\pm a$ is possible, the charged Higgs is produced on-shell, and the cross-section of tW production is thus resonantly

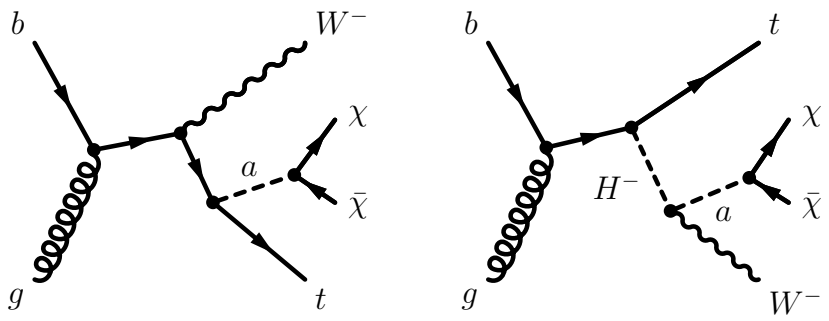


Figure 8.3: Feynman diagrams contributing to DMt tW associated production in the 2HDMa model.

enhanced.

The production cross-section at LO for both the t-channel and tW processes, calculated with MadGraph5_aMC@NLO v2.6.2, is shown in Fig. 8.4 as a function of $\tan\beta$. The cross-section for the charged Higgs on-shell production, contributing to tW final state, is also drawn. Both plots fix $\sin\theta = 1/\sqrt{2}$ and $m_a = 150$ GeV, while $m_A = m_H = m_{H^\pm} = 500$ GeV in the first plot and 1 TeV in the second. As one can see from the figure, the tW contribution to the DMt cross-section is always dominating over the t-channel, especially for lower values of $m_{H^\pm}$, where the on-shell contribution dominates the final cross-section. The tW channel provides thus a better sensitivity over the most part of the parameter space of the 2HDMa model. For this reason, it has been chosen as main focus of this work.

8.2.2 Dependency on $\sin\theta$ and $\tan\beta$

The dependency of tW associated production on $\tan\beta$ shown in Fig. 8.4 is the result of the interplay of different factors. As it can be seen in Sec. 2.2.3, the cross-section for $H^\pm$ production (Fig. 8.3 right) goes as $m_t^2\cot^2\beta + m_b^2\tan^2\beta + \text{const.}$, while the cross-section for a-strahlung (Fig. 8.3 left) is proportional to $\cot^2\beta$. The branching ratio for $H^\pm \to W^\pm a$ depends on $\tan\beta$ due to the competition with the decay $H^\pm \to tb$, which goes as $\cot^2\beta$, and the same thing happens also for the BR of $a \to \chi\bar{\chi}$, which decreases at high $\tan\beta$ since the competing decay $a \to b\bar{b}$ grows as $\tan^2\beta$.

The decay width of $H^\pm \to W^\pm a$ is proportional to $\sin^2\theta$, but the BR has a more complicated dependency on $\sin\theta$ due to the competition with $H^\pm \to tb$, whose decay width is independent from $\sin\theta$. Also the cross-section for a-strahlung depends on $\sin\theta$, going as $\sin^2\theta$ as long as the decay $a \to t\bar{t}$ is kine-

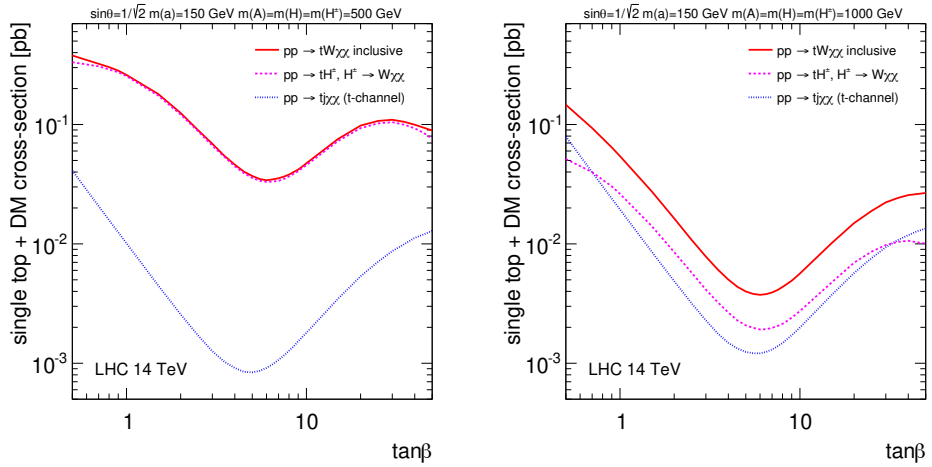


Figure 8.4: Cross-section for DMt production for the t-channel and for tW associated production at 13 TeV as a function of $\tan\beta$ for $m_a = 150$ GeV and $m_{H^\pm} = 500$ GeV (left) and 1000 GeV (right). The full line corresponds to the tW channel, while the dotted line shows the result for t-channel production. The dashed line indicates the contribution to tW production from the on-shell production of a charged Higgs [127].

matically forbidden. For both processes, the total production cross-section of tW production thus increases with $\sin\theta$. For this reason, for the analyses presented in the following a sensible choice is $\sin\theta = 1/\sqrt{2}$, corresponding to the maximum mixing between the pseudoscalar bosons of the 2HDMa model.

8.3 Two-Lepton Search for DM Produced in Association with a Single Top Quark in the 2HDMa Model

8.3.1 Object Selection

All objects were reconstructed as explained in Chapter 5. Two levels of object identification requirements were defined for leptons and jets: "baseline" and "signal". Baseline objects were selected with looser identification criteria, in order to be used in computing the missing transverse momentum $E_{\mathrm{T}}^{\mathrm{miss}}$ as well as in overlap removal (Sec. 5.7). Signal leptons and jets are a subset of the baseline objects, with tighter quality requirements, and were used to define the search

regions.

Baseline electrons need to have $p_T > 10$ GeV and $|\eta| < 2.47$, excluding the transition region between the barrel and end-cap calorimeters ($1.37 < |\eta| < 1.52$). They are requested to pass the Loose likelihood identification operating point [86]. The longitudinal impact parameter z_0 relative to the primary vertex is required to satisfy $|z_0 \sin\theta| < 0.5$ mm. Signal electrons are in addition required to have $p_T > 20$ GeV and to pass the Tight likelihood identification criteria. Moreover, the significance of the transverse impact parameter d_0 must satisfy $|d_0/\sigma(d_0)| < 5$. Signal electrons with $p_T < 200$ GeV are further refined using the FCLoose isolation working point [86], while those with larger p_T are required to pass the FCHighPtCaloOnly isolation working point [86]. Corrections for energy contributions due to pile-up are applied.

Baseline muons are required to have $p_T > 10$ GeV and $|\eta| < 2.5$, and to satisfy the Medium identification criteria [89]. Like electrons, their longitudinal impact parameter z_0 relative to the primary vertex needs to satisfy $|z_0 \sin\theta| < 0.5$ mm. Signal muons are defined with tighter requirements on their transverse momentum and transverse impact parameter significance: $p_T > 20$ GeV and $|d_0/\sigma(d_0)| < 3$. In addition, the FCLoose isolation working point is also required [89].

Jets are reconstructed using the anti-kT algorithm, with a radius parameter $R = 0.4$ [90]. To reduce the effect of pile-up interactions, jets with $|\eta| < 2.4$ and $p_T < 120$ GeV are required to satisfy the Medium working point of the jet vertex tagger (JVT) [94]. Baseline jets are selected in the region $|\eta < 4.5$ and need to have $p_T > 20$ GeV. Signal jets are required to be in the region $|\eta| < 2.5$ and to have $p_T > 30$ GeV. Jets containing b-hadrons are tagged using the MV2c10 algorithm [98]. They are reconstructed in the region $|\eta| < 2.5$ and have $p_T > 20$ GeV. The b-tagging working point chosen is the one providing an efficiency of 77% [99].

8.3.2 Event Selection

The final state searched for by the $\mathrm{tW_{2L}}$ analysis includes two leptons, one b-jet and $E_{\mathrm{T}}^{\mathrm{miss}}$, with one lepton and the b-jet from the decay of the top quark and the other lepton from the decay of the W boson. The main backgrounds for this signature can be divided into four groups:

- backgrounds not including a leptonic top quark decay such as as Z+jets and dibosons
- $t\bar{t}$ and single-top, with a final state similar to the signal but with $E_{\mathrm{T}}^{\mathrm{miss}}$ only produced from W decays and therefore kinematically bounded by the W mass

- $t\bar{t}Z$, $t\bar{t}W$ events, with $E_{\mathrm{T}}^{\mathrm{miss}}$ produced by the Z boson decaying to neutrinos or by the loss of a lepton from the decay of the W. These events contain an additional top-quark decay with respect to the signal

- tWZ events, with the Z boson decaying to neutrinos, with a final state identical to the signal.

The different background contributions require dedicated strategies in order to be reduced. To this end, the discriminant variables introduced in Sec. 6.3.3 were exploited to define an effective SR ($\mathrm{SR_{tW_{2L}}}$) with the best possible statistical significance Z (Eq. 6.1) for the signal in the regions of 2HDMa parameter space investigated in the analysis.

In order to reduce the non-top backgrounds contribution, the minimal requirement of at least one b-tagged jet with $p_T > 50$ GeV is imposed together with the condition $\min m_{b\ell} < 170$ GeV, which forces the compatibility of the event with at least one leptonic top quark decay.

To decrease the background coming from $t\bar{t}$ and $t\bar{t}V$, in events with at least 2 jets $m_{b\ell}^{t}$ is required to be greater than 150 GeV, since for events with two leptonic top-quark decays it has an upper limit at ~ 150 GeV (see Sec. 6.3.3), while it is unbounded for the signal which contains only a single top quark. This is the first time $m_{b\ell}^{t}$ is introduced in a BSM search performed in ATLAS, allowing to significantly reduce the background in the SR.

The two natural variables to exploit the additional $E_{\mathrm{T}}^{\mathrm{miss}}$ coming from DM particles in the signals and further reduce the $t\bar{t}$ and single-top backgrounds are $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_{T2}. Optimal values ($E_{\mathrm{T}}^{\mathrm{miss}} > 200$ GeV and $m_{T2} > 130$ GeV) were found by varying concurrently the two variables after all other cuts were applied, and evaluating the expected signal significance Z for all signal samples.

Finally, a cut on $\Delta\phi_{min}$ is used to suppress events in which the $E_{\mathrm{T}}^{\mathrm{miss}}$ comes from the mismeasurement of a jet, which may help $t\bar{t}$ events to escape the kinematical W mass bound in m_{T2}.

The list of cuts defining the $\mathrm{SR_{tW_{2L}}}$ is presented in Tab. 8.1. The impact of the different selections can be seen in Fig. 8.5, where the kinematic distributions for the main variables are shown for both signal and background when all other cuts except the one on the plotted variable are applied, making clear the effectiveness of the chosen cuts.

8.3.3 Background Estimation

The contributions from the $t\bar{t}$, $t\bar{t}V$ ($V = W, Z$) and diboson background processes are the main one in the $\mathrm{SR_{tW_{2L}}}$, and were thus estimated from MC simulation and dedicated CRs. The remaining sources of background, including the irreducible

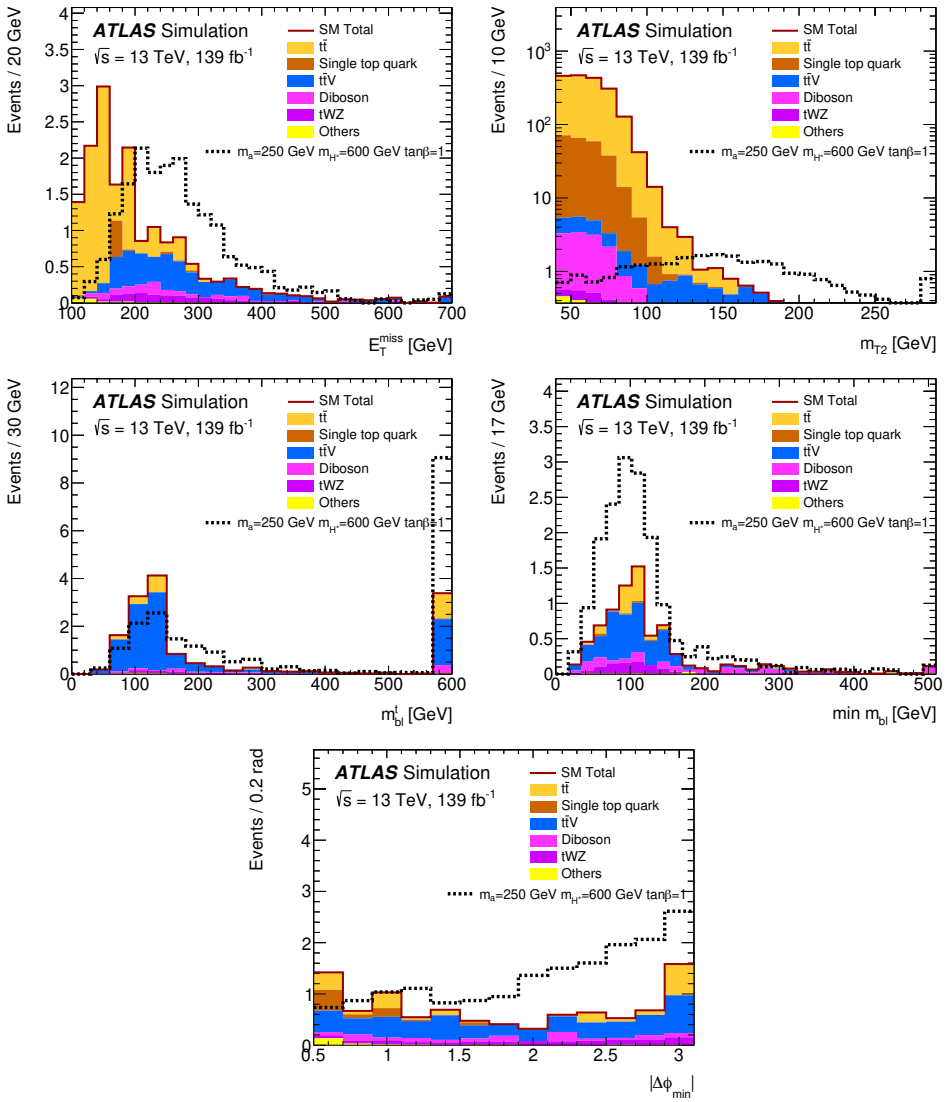


Figure 8.5: Kinematic distributions of main discriminating variables for both background and a signal after all the $\mathrm{SR_{tW_{2L}}}$ cuts except the cut on the variable itself are applied. The signal samples displayed are generated with $\sin\theta = 1/\sqrt{2}$. Please note that m_{bl}^{t} is defined only for events with at least two jets, events with exactly one jet are put in the overflow bin.

Variable [Unit]	SR	CR($t\bar{t}$)	CR($t\bar{t}Z$)	CR(WZ)	VR($t\bar{t}$)	VR(3ℓ)
N_ℓ^{signal}	$= 2$ (OS)	$= 2$ (OS)	$= 3$ (≥ 1 SFOS)	$= 3$ (≥ 1 SFOS)	$= 2$ (OS)	$= 3$ (≥ 1 SFOS)
$p_T(\ell_3)$ [GeV]	-	-	> 20	> 20	-	> 20
$m_{\ell\ell}$ [GeV]	$\notin [71, 111]$	$\notin [71, 111]$	$\in [71, 111]$	$\in [71, 111]$	$\notin [71, 111]$	$\in [71, 111]$
$\mathcal{N}_{\text{jets}}$	≥ 1	≥ 1	≥ 3	$\in [1, 3]$	≥ 1	≥ 1
$\mathcal{N}_{b-\text{jets}}$	≥ 1	≥ 1 (≥ 2 if $\mathcal{N}_{b-\text{jets}} = 3$)	≥ 1	$= 1$	≥ 1	≥ 1
$\min m_{b\ell}$ [GeV]	< 170	< 170	< 170	> 170	< 170	varies
$m_{b\ell}^t$ [GeV]	> 150	< 150	incl.	incl.	> 150	incl.
$E_{\text{T}}^{\text{miss}}$ [GeV]	> 200	> 200	> 200	> 200	> 200	> 200
m_{T2} [GeV]	> 130	$\in [40, 80]$	> 90	> 90	$\in [40, 80]$	> 90
$\Delta\phi_{\min}$ [rad]	> 1.1	> 1.1	incl.	incl.	> 1.1	incl.

Table 8.1: Summary of signal, control and validation region definitions used in the tW_{2L} analysis. The entry of "incl." represents an inclusive selection with no requirements. In the final states with three leptons, the corrected variables of $E_{\text{T}}^{\text{miss}}$, $\min m_{b\ell}$ and m_{T2} are used instead as described in the text. The selection requirement on the corrected $\min m_{b\ell}$ in the VR(3ℓ) region varies according to the jet and b-jet multiplicity as described in the main text. Events with additional baseline leptons are vetoed in addition.

tWZ process, single top quark production and $t\bar{t}h$ production were estimated from simulation.

$t\bar{t}$ Background Estimate

The production of $t\bar{t}$ is the second most important background for the tW_{2L} analysis. It has a final state similar to the signal one, except for the fact that there is an extra b-jet and no additional missing energy coming from DM particles. The two main requests of the SR designed to take advantage of the latter are the cuts on m_{T2} and $m_{b\ell}^t$, both exploiting the bounded kinematic of $t\bar{t}$.

A possible strategy to define a CR for $t\bar{t}$, considered at the beginning of the work, was to reverse the cut on $m_{b\ell}^t$ ($m_{b\ell}^t < 150$ GeV) while loosening the m_{T2} cut. However, the contamination coming from $DMt\bar{t}$ signal proved to be non-negligible in such a region, and a different strategy had to be adopted.

The $\text{CR}_{\text{tW}_{\text{2L}}}(t\bar{t})$ is thus defined inverting both the SR requirements on m_{T2} and $m_{b\ell}^t$, asking for $m_{T2} \in [40, 80]$ GeV and $m_{b\ell}^t < 150$ GeV. In this way the number of $t\bar{t}$ events greatly increases, making the $DMt\bar{t}$ contamination totally negligible. In order to have a CR as close as possible to the SR, all other selection

	$CR_{tW_{2L}}$ ($t\bar{t}$)	$CR_{tW_{2L}}$ ($t\bar{t}Z$)	$CR_{tW_{2L}}$ (WZ)
Observed events	1407	80	48
Fitted bkg events	1406.93 ± 37.51	79.86 ± 8.93	48.09 ± 6.94
$t\bar{t}$	1299.88 ± 39.36	0.00 ± 0.00	0.00 ± 0.00
Single top	93.84 ± 11.59	3.33 ± 0.52	0.78 ± 0.08
W-jets	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
Z-jets	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
Diboson	2.89 ± 1.01	6.60 ± 2.39	28.61 ± 9.64
$t\bar{t}V$	7.64 ± 1.57	56.03 ± 11.03	7.81 ± 2.15
$t\bar{t}h$	2.43 ± 0.21	1.54 ± 0.13	0.19 ± 0.02
tWZ	0.26 ± 0.05	5.82 ± 0.77	3.25 ± 0.35
Fakes	0.00 ± 0.00	6.55 ± 4.69	7.44 ± 5.29

Table 8.2: Background-only fit results for tW_{2L} CRs. The quoted uncertainties on the fitted SM background include both the statistical and systematic uncertainties.

cuts are applied unchanged. The complete list of selection applied is described in Tab. 8.1. The composition of $CR_{tW_{2L}}(t\bar{t})$ after the background-only fit is given in Tab. 8.2, together with the data yield observed. The modelling of E_T^{miss} and $m_{b\ell}^t$ for background events in $CR_{tW_{2L}}(t\bar{t})$ after the fit is also shown in Fig. 8.6.

A validation region closer to the SR is defined to further check the agreement between the fitted background and data ($VR_{tW_{2L}}(t\bar{t})$). All the requirements of $CR_{tW_{2L}}(t\bar{t})$ are applied, except for the one on $m_{b\ell}^t$, which is inverted like in the SR ($m_{b\ell}^t > 150$ GeV). These requests are summarised in Tab. 8.1, while Tab. 8.3 shows the background and data yields in $VR_{tW_{2L}}(t\bar{t})$ after the background-only fit. The distributions of E_T^{miss} and $m_{b\ell}^t$ in $VR_{tW_{2L}}(t\bar{t})$ are presented in Fig. 8.7, which shows a good agreement between the fit results and data.

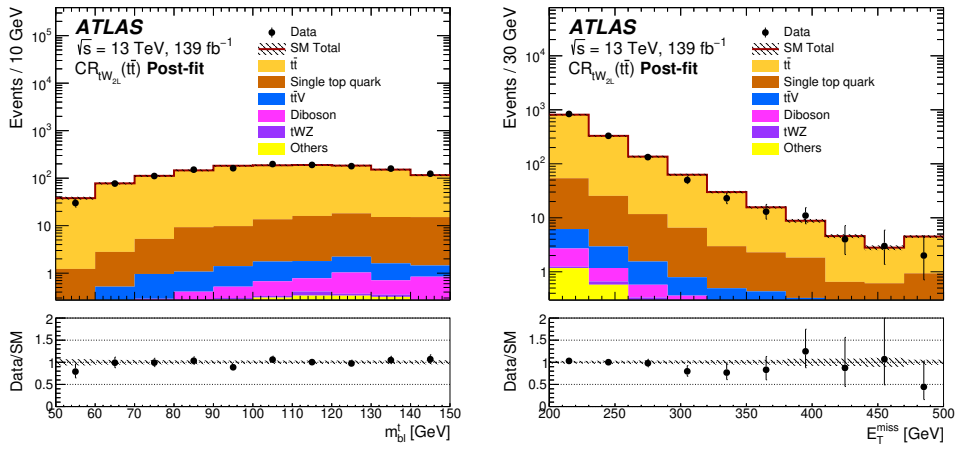


Figure 8.6: Main kinematic distribution in $\mathrm{CR_{tW_{2L}}}(\mathrm{t\bar{t}})$ for data and Monte Carlo samples. The errors includes statistical, experimental and modelling uncertainties. Please note that m_{bl}^{t} is defined only for events with at least two jets, events with exactly one jet are put in the overflow bin.

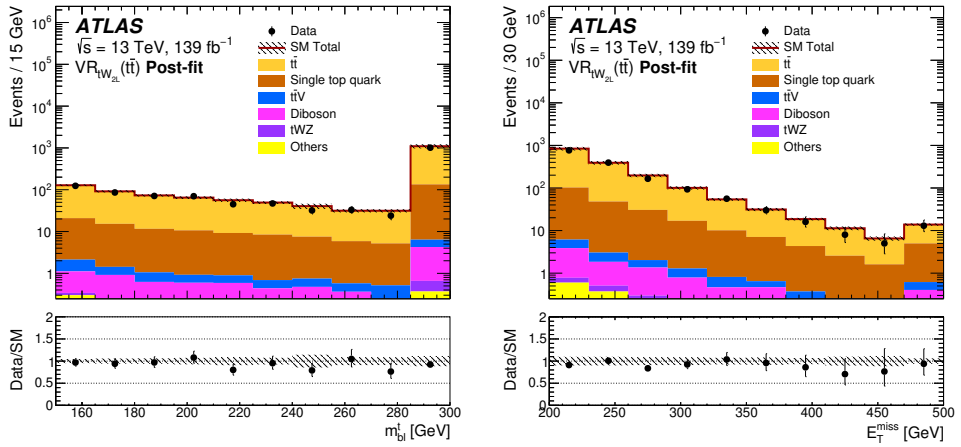


Figure 8.7: Main kinematic distribution in $\mathrm{VR_{tW_{2L}}}(\mathrm{t\bar{t}})$. The errors includes statistical, experimental and modelling uncertainties. Please note that m_{bl}^{t} is defined only for events with at least two jets, events with exactly one jet are put in the overflow bin.

	$VR_{tW_{2L}}$ ($t\bar{t}$)	$VR_{tW_{2L}}(3\ell)$
Observed events	1546	134
Fitted bkg events	1669.05 ± 157.41	121.21 ± 16.06
$t\bar{t}$	1436.35 ± 166.71	0.03 ± 0.00
Single top	213.82 ± 13.72	12.99 ± 0.72
W-jets	0.00 ± 0.00	0.00 ± 0.00
Z-jets	0.00 ± 0.00	0.00 ± 0.00
Diboson	7.58 ± 2.66	27.96 ± 9.74
$t\bar{t}V$	5.75 ± 1.29	46.08 ± 9.72
$t\bar{t}h$	1.50 ± 0.09	1.21 ± 0.11
tWZ	0.55 ± 0.05	11.06 ± 0.75
Fakes	3.50 ± 2.44	21.88 ± 11.63

Table 8.3: Background-only fit results for tW_{2L} VRs. The quoted uncertainties on the fitted SM background include both the statistical and systematic uncertainties.

$t\bar{t}Z$ Background Estimate

The production of $t\bar{t}$ associated with a Z-boson, where the Z decays into neutrinos and the top quarks decay leptonically, is the dominant irreducible background for the tW_{2L} analysis. Its final state consists of two leptons, two b-jets and E_T^{miss}, which is the same final state of the signal with an additional b-jet. The additional E_T^{miss} coming from the invisible Z decay allow this background to evade the kinematic constraints on the W mass to which $t\bar{t}$ production is subjected, making it the most important process in the $SR_{tW_{2L}}$.

The natural control region for the dileptonic $t\bar{t}Z$ with invisible Z decay would be the four-lepton final state, with also the Z-boson decaying into leptons. Indeed, in this case the two leptons from the Z decay could be identified through their invariant mass and added to the E_T^{miss} to mimic the neutrinos, making possible to reproduce all the variables used in the definition of $SR_{tW_{2L}}$. In this way, a control region closely matching the selection cuts of the signal region could be built. However, the observed number of events with four leptons for the available statistics was of the order of only 50 events before any cut, preventing the use of such a region due to its limited statistical power.

For this reason, following the same strategy of Sec. 7.3.2, the three-lepton channel was used instead, with two leptons coming from a leptonic Z decays and treated as invisible, and only one of the two tops required to decay leptonically. The drawback of this choice is that with the three-lepton channel it is not possible to fully reconstruct the discriminant transverse variables employed in the

$\mathrm{SR}_{\mathrm{tW_{2L}}}$ definition which require two leptons, but only an approximation of them, as discussed in [41].

The $\mathrm{CR}_{\mathrm{tW_{2L}}}(\mathrm{t\bar{t}Z})$ is thus defined by selecting three signal leptons out of which one SFOS (opposite-sign same-flavour) pair has to be compatible with a Z-boson decay. The transverse momentum of the leading lepton is required to be above 25 GeV, and above 20 GeV for the other leptons. In order to identify the leptons compatibility with a Z decay, the invariant mass among all SFOS lepton pairs is calculated, and the event is accepted as containing a Z if at least one of the possible combinations has an invariant mass within 20 GeV from the Z-mass peak. In case more than one lepton pair satisfies this condition, the pair yielding the mass nearest to the nominal Z mass is taken. $t\bar{t}Z$ events including one leptonic top decay contain two b-jets and two light jets from the hadronic decay of the W in the final state. Since the jet selection requirement for the $\mathrm{SR}_{\mathrm{tW_{2L}}}$ is of at least one b-tagged jet, the natural extension of that selection would be the requirement of at least four jets of which at least one is b-tagged. This request is loosened by asking for at least one b-jet and at least 3 jets. Among the configurations defined by this requirement, the one with exactly three jets of which exactly one b-tagged is rejected, as it is dominated by diboson processes, in particular WZ production.

The momenta of the leptons compatible with the Z decay are treated as invisible and added vectorially to the $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$, giving rise to the so-called $E_{\mathrm{T,corrected}}^{\mathrm{miss}}$. In order to imitate the signal region selection cuts, since it is not possible to build m_{T2} with the only lepton not treated as invisible, the one-lepton variable m_T (defined in Sec. 6.3.2) calculated using $E_{\mathrm{T,corrected}}^{\mathrm{miss}}$ is employed and indicated as $m_{T2}^{\mathrm{corrected}}$. Moreover, the variable $\min m_{b\ell}^{\mathrm{corrected}}$ is defined as the invariant mass of the leading b-jet with the lepton not compatible with the Z, without actually taking any minimum. A selection on these variables allows the choice of a control region with a good superposition in kinematic space with the signal region.

The $\mathrm{CR}_{\mathrm{tW_{2L}}}(\mathrm{t\bar{t}Z})$ is thus defined requiring $E_{\mathrm{T,corrected}}^{\mathrm{miss}} > 200$ GeV and $m_{T2}^{\mathrm{corrected}} > 90$ GeV, where the cut on $m_{T2}^{\mathrm{corrected}}$ is loosened with respect to the SR to increase the statistics. No selection are imposed on $\Delta\phi_{min}$ and a variation of $m_{b\ell}^{t}$ for the same reason. Finally, $\min m_{b\ell}^{\mathrm{corrected}}$ is asked to be smaller than 170 GeV. These selections are summarised in Tab. 8.1, where it is easily observable the similarity of the definitions of $\mathrm{CR}_{\mathrm{tW_{2L}}}(\mathrm{t\bar{t}Z})$ and $\mathrm{SR}_{\mathrm{tW_{2L}}}$ once the "corrected" variables are employed. The composition of $\mathrm{CR}_{\mathrm{tW_{2L}}}(\mathrm{t\bar{t}Z})$ after the background-only fit is given in Tab. 8.2, alongside the number of observed events in data. The modelling of $E_{\mathrm{T,corrected}}^{\mathrm{miss}}$ and $m_{T2}^{\mathrm{corrected}}$ in the region after the fit is shown in Fig. 8.8.

WZ Background Estimate

The main contamination in the $\mathrm{CR}_{\mathrm{tW_{2L}}}(\mathrm{t\bar{t}Z})$ comes from diboson processes, in particular WZ production. In order to improve the robustness of the $t\bar{t}Z$ estimate,

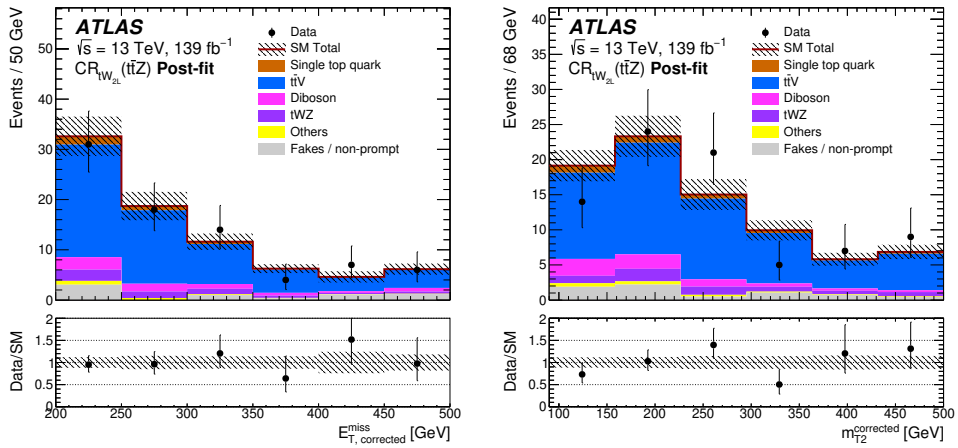


Figure 8.8: Main kinematic distribution in $\mathrm{CR_{tW_{2L}}(t\bar{t}Z)}$ for data and Monte Carlo samples. The errors includes statistical, experimental and modelling uncertainties.

it was thus decided to use a dedicated CR for this background rather than a MC prediction.

In the $\mathrm{CR_{tW_{2L}}(t\bar{t}Z)}$ there are two cuts imposed to reduce the contribution from WZ processes: the request on the number of jets and b-jets and the one on $\min m_{b\ell}^{\text{corrected}}$. The most natural way to define a $\mathrm{CR_{tW_{2L}}(WZ)}$ is then to simply invert these cuts, asking for exactly one b-jet, less than 4 jets and $\min m_{b\ell}^{\text{corrected}} > 170$ GeV. The complete list of requirements for the $\mathrm{CR_{tW_{2L}}(WZ)}$ is presented in Tab. 8.1, while Tab. 8.2 reports the yields of background events in the region after the background-only fit and data. The distributions of $E_{\text{T,corrected}}^{\text{miss}}$ and $m_{T2}^{\text{corrected}}$ in the region after the fit are shown in Fig. 8.9.

$\mathrm{t\bar{t}Z}$/WZ Three-Lepton Validation Region

In order to check the agreement between the predicted $t\bar{t}Z$ and WZ backgrounds after the background-only fit and the observed rates, a common three-lepton validation region is selected with the cuts described in Tab. 8.1. The main difference with the two CRs lies in reverting the $\min m_{b\ell}^{\text{corrected}}$ cut while applying the same jet conditions of the CRs. The number of events in $\mathrm{VR_{tW_{2L}}(3L)}$ after the background-only fit is given in Tab. 8.3 together with the data yield, while the distributions of $E_{\text{T,corrected}}^{\text{miss}}$ and $m_{T2}^{\text{corrected}}$ in the region after the fit are shown in Fig. 8.10. A good agreement between data and simulations is observed.

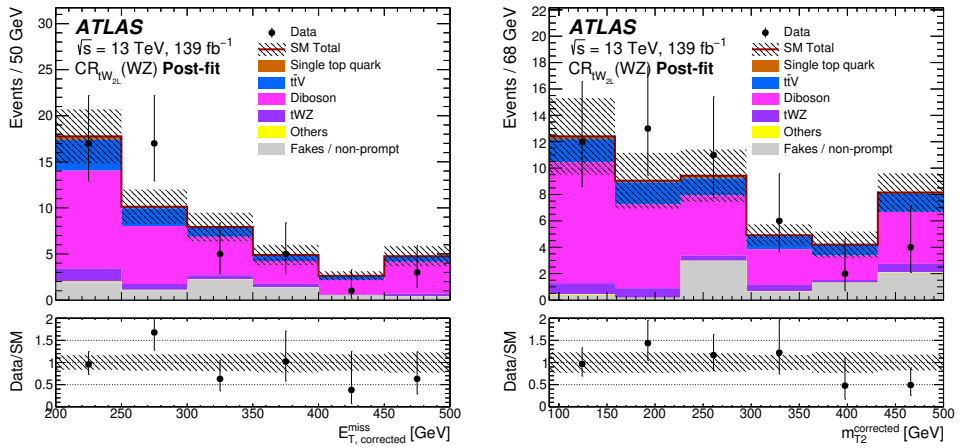


Figure 8.9: Main kinematic distribution in $\mathrm{CR_{tW_{2L}}(WZ)}$ for data and Monte Carlo samples. The errors includes statistical, experimental and modelling uncertainties.

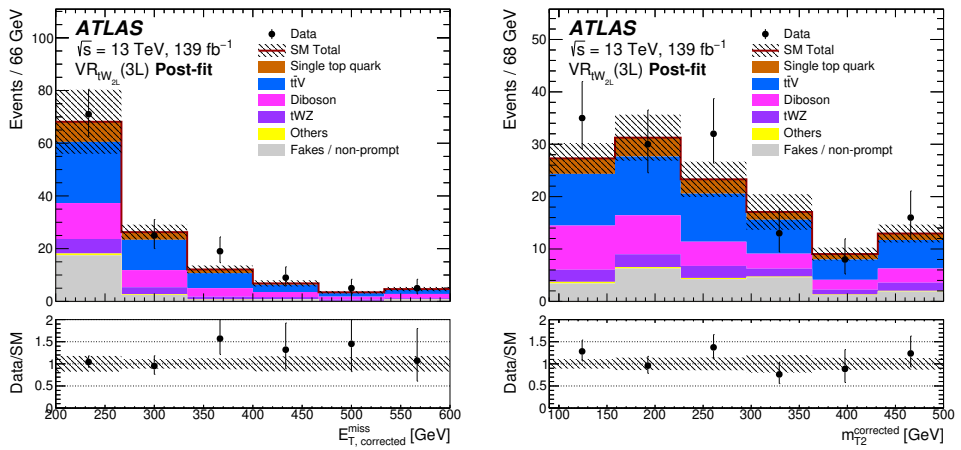


Figure 8.10: Main kinematic distribution in $\mathrm{VR_{tW_{2L}}(3L)}$ for data and Monte Carlo samples. The errors includes statistical, experimental and modelling uncertainties.

Wt Background Estimate

Monte Carlo simulations are used to estimate the background originating from single top processes in the $\mathrm{SR_{tW_{2L}}}$. The only single top contribution in the $\mathrm{SR_{tW_{2L}}}$ comes from Wt associated production, and from initial estimates this background was expected to be only a few percent of the total expected background in the

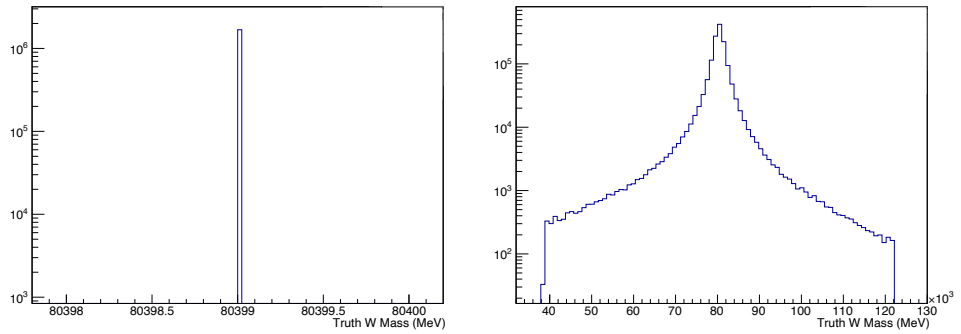


Figure 8.11: W mass lineshapes for nominal POWHEG samples. Left: associated W, Right: W coming from top decay.

signal region.

Additional detailed studies performed at parton level on the recommended Wt POWHEG [73, 74, 75, 113] samples revealed that in POWHEG the W mass line-shapes are not correctly simulated: the associated W-boson is simulated in zero-width approximation, while the Breit-Wigner of the W coming from the top is cut between 40 and 120 GeV (Fig. 8.11). This is a good approximation in the majority of the phase space, but it can have a rather large effect on variables sensitive to the W mass, such as the tail of the m_{T2} distribution. Indeed, the m_{T2} variable has an endpoint at the W mass by construction, by effectively picking the largest value of the two possible reconstructed W candidates: if the W mass is cut off at generator level, this cut will also have an effect on the tail of the the m_{T2} distribution (i.e. through an upper limit), with additional smearing and resolution effects also playing a role. Since $\mathrm{SR_{tW_{2L}}}$ is defined with $m_{T2} > 130$ GeV, the Wt contribution evaluated through the recommended POWHEG samples is then underestimated.

To overcome this issue, the authors of POWHEG were contacted and they provided a patch to the Wt simulation code, introducing the correct line-shape for both W masses (Fig. 8.12). Using this new code, new Wt samples were produced at TRUTH level. Fig. 8.13 shows the comparison of various kinematic variables at TRUTH level between the new "'patched' samples and the nominal ones. All distributions show good agreement, except for m_{T2}, where a discrepancy at high masses is observed as expected.

Given the impossibility to generate FULLSIM patched samples in a short timescale, the most natural way to correct the Wt yield in $\mathrm{SR_{tW_{2L}}}$ would have been to calculate a multiplication factor between the patched and the nominal samples at TRUTH level in the $\mathrm{SR_{tW_{2L}}}$ and apply it to the reconstructed nominal yield. However, this was not possible, since there are no nominal events left

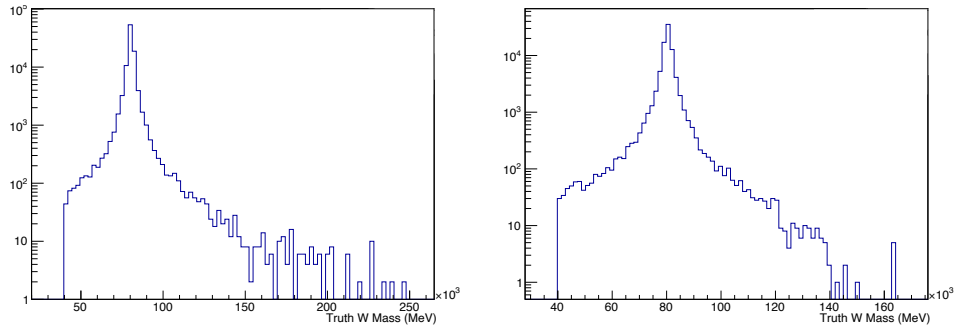


Figure 8.12: W mass lineshapes for patched POWHEG samples. Left: associated W, Right: W coming from top decay.

in the $\mathrm{SR}_{\mathrm{tW_{2L}}}$ at TRUTH level due to the upper limit on m_{T2} introduced by the wrong line-shapes. For this reason, the ratio between patched and nominal samples has been studied in the $m_{T2} - E_{\mathrm{T}}^{\mathrm{miss}}$ plane. As shown in Fig. 8.14, there is a clear trend (within statistical uncertainties) going towards high values of m_{T2} and $E_{\mathrm{T}}^{\mathrm{miss}}$, where the ratio increases. The Wt yield in the $\mathrm{SR}_{\mathrm{tW_{2L}}}$ is then corrected using the ratio of events between the two TRUTH samples for $m_{T2} > 110$ GeV and $E_{\mathrm{T}}^{\mathrm{miss}} > 200$ GeV, near the SR, normalising it to the ratio of events for $m_{T2} < 90$ GeV and $E_{\mathrm{T}}^{\mathrm{miss}} < 90$ GeV.

8.3.4 Fake and Non-Prompt Leptons

The contribution from events with one or more non-prompt or misidentified leptons is expected to be non-negligible for the $\mathrm{tW_{2L}}$ analysis, in particular for the three-lepton control and validation regions. The fake lepton contribution arise when one or more hadronic jets are misidentified as leptons, indicated as the fake leptons, whereas the non-prompt leptons may come from semi-leptonic decays of heavy-flavour jets or leptons produced by photon conversions. Both fake and non-prompt lepton contributions together are referred to as the fakes background in the following.

The rate at which hadronic jets are misidentified as leptons is not accurately described in Monte Carlo simulations, since it depends on the showering development, where non-perturbative QCD is involved. Moreover, the probability to obtain a fake lepton is quite low, thus a large number of events would need to be generated and reconstructed to obtain enough statistics in the phase space targeted by the analyses. An accurate simulation would therefore require a large computational power.

For these reasons, the fake background was determined from data using the data-driven Matrix Method [132, 133]. In this method, the determination of fakes

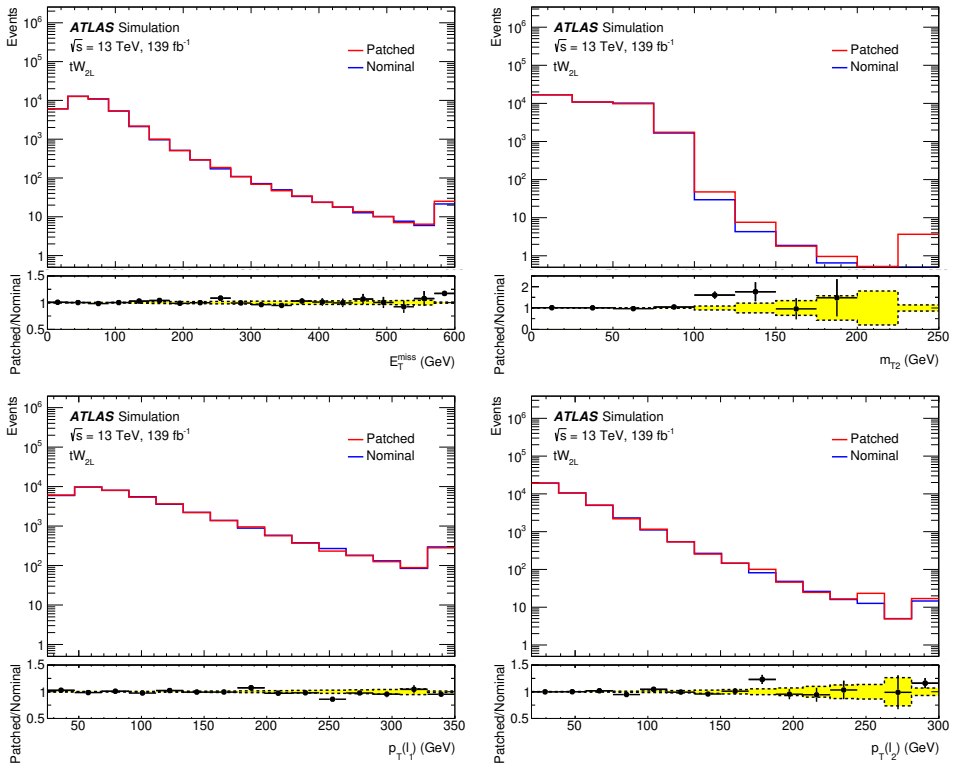


Figure 8.13: TRUTH level comparison of POWHEG Wt nominal and patched samples.

relies on comparing in data the number of events with leptons passing two sets of selection requirements, the so-called "loose" (L) and "tight" (T) selection criteria, where the tight selection is a subset of the loose one. For the tW_{2L} analysis, the tight selection matches the analysis selection for signal leptons, while the loose correspond to the ones used for the definition of baseline lepton candidates (see Sec. 8.3.1).

In case of a selection with two loose leptons, which is presented for clarity and can be extended to any number of leptons, all events can be divided in four categories based on the number of leptons which pass also the tight selection (loose-loose, loose-tight, tight-loose and tight-tight, where the first category refers to the leading lepton). In parallel, the same events can be categorised regarding the true nature of the leptons: real-real events containing two prompt leptons, real-fake and fake-real events with one real and one fake lepton and, finally, fake-fake events in which both leptons are wrongly identified. In the case of two-lepton

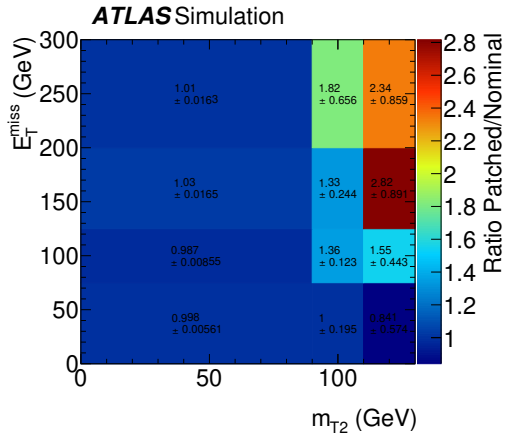


Figure 8.14: Ratio between Powheg Wt nominal and patched samples in the $m_{T2} - E_{\mathrm{T}}^{\mathrm{miss}}$ plane.

events, the number of events in each category can be connected by the equation:

$$\begin{pmatrix} N^{\mathrm{LL}} \\ N^{\mathrm{LT}} \\ N^{\mathrm{TL}} \\ N^{\mathrm{TT}} \end{pmatrix} = M \begin{pmatrix} N^{\mathrm{LL}}_{\mathrm{real,real}} \\ N^{\mathrm{LL}}_{\mathrm{real,fake}} \\ N^{\mathrm{LL}}_{\mathrm{fake,real}} \\ N^{\mathrm{LL}}_{\mathrm{fake,fake}} \end{pmatrix}, \tag{8.1}$$

with

$$M = \begin{pmatrix} (1-r_1)(1-r_2) & (1-r_1)(1-f_2) & (1-f_1)(1-r_2) & (1-f_1)(1-f_2) \\ (1-r_1)r_2 & (1-r_1)f_2 & (1-f_1)r_2 & (1-f_1)f_2 \\ r_1(1-r_2) & r_1(1-f_2) & f_1(1-r_2) & f_1(1-f_2) \\ r_1 r_2 & r_1 f_2 & f_1 r_2 & f_1 f_2 \end{pmatrix}, \tag{8.2}$$

where r_i is the probability for a real lepton which pass the loose selection to pass also the tight one, and f_i is the same probability for a fake lepton. If the number of events passing the loose and tight requirements as well as the efficiencies r_i and f_i are known, the number of events with one real and one fake lepton ($N^{\mathrm{LL}}_{\mathrm{real,fake}}$+$N^{\mathrm{LL}}_{\mathrm{fake,real}}$) or two fake leptons ($N^{\mathrm{LL}}_{\mathrm{fake,fake}}$) passing the loose selection criteria can be obtained by solving the above system of equations. The sum of these numbers, weighted for the corresponding probabilities for loose leptons to pass the tight selection r_i and f_i, gives the fake estimate in each region.

The fake rates f_i and the real efficiencies r_i are calculated in dedicated control regions, selected in order to obtain samples enriched with fake or real leptons, respectively. Their value is expressed as a function of the p_T and η of the lep-

ton, separately for electrons and muons. The measurement of r_i is carried out using the tag-and-probe method on $t\bar{t}Z, Z \to \ell\ell$ Monte Carlo events, while the measurement of f_i is obtained from a sample with at least one b-jet and a pair of same-sign same-flavour leptons [132, 133].

8.3.5 Systematic Uncertainties

In addition to statistical uncertainties in the Monte Carlo simulations and on data, several sources of experimental and theoretical systematic uncertainties were evaluated on all background and signal samples, as explained in Sec. 6.4. Tab. 8.4 shows the contribution of the main sources of systematics uncertainties on the background estimate in $\mathrm{SR_{tW_{2L}}}$ as a relative uncertainty in the total background yield, evaluated after the background-only fit.

	$\mathrm{SR_{tW_{2L}}}$
Total background systematic	20.7%
$t\bar{t}$ modelling uncertainties	3.9%
$t\bar{t}V$ modelling uncertainties	2.2%
Diboson modelling uncertainties	0.4%
tWZ modelling uncertainties	1.8%
Single top modelling uncertainties	4.5%
$t\bar{t}$ normalisation uncertainties	0.6%
$t\bar{t}V$ normalisation uncertainties	10.0%
Diboson normalisation uncertainties	3.0%
MC statistical uncertainties	7.2%
JES uncertainty	9.4%
JER uncertainty	11.8%

Table 8.4: Main sources of systematic uncertainties on the SM background estimates, evaluated after the background-only fit. The values are given as relative uncertainties on the expected background event yields in the $\mathrm{SR_{tW_{2L}}}$.

The dominant experimental systematics in the $\mathrm{SR_{tW_{2L}}}$ are from the jet energy scale and resolution. The uncertainties associated with b-tagging efficiency, trigger requirements, lepton reconstruction and energy measurements have instead a small or negligible impact on the final results.

Uncertainties in the modelling of the SM background processes in simulation and their theoretical cross-section uncertainties were also taken into account. For the $t\bar{t}V$ background, uncertainties due to parton shower and hadronisation modelling are evaluated by comparing the predictions from MadGraph5_aMC@NLO [72]

interfaced to Pythia8 [76] and Herwig7 [78, 79], while the uncertainties related to the choice of renormalisation and factorisation scales (μ_R and μ_F) are obtained by varying them up and down by a factor of two in the nominal generator. A similar approach is used to evaluate the uncertainties in the tWZ process, where an additional 20% uncertainty is added to account for uncertainties in the effects of interference between the ttZ and tWZ.

In the case of $t\bar{t}$ background, the impact of the PS and hadronisation is evaluated by comparing the nominal generator with a Powheg-Box sample interfaced to Herwig++ [78, 79], using the H7UE set of tuned parameters [79]. The uncertainty due to the choice of generator and matching scheme is assessed through an alternative generator setup, using MadGraph5_aMC@NLO interfaced to Pythia8. The renormalisation and factorisation scales uncertainties are evaluated in the same way as for $t\bar{t}V$.

For single-top Wt production, given the possibility of large uncertainties inherent in the TRUTH-level reweighting procedure adopted (see Sec. 8.3.3), a 100% uncertainty is assigned.

For the DM signal processes, both the experimental and theoretical uncertainties in the expected signal yields are considered, including the aforementioned luminosity uncertainty. The dominant experimental uncertainty comes from the jet energy scale and resolution also in this case. In the modelling of the signals, uncertainties due to the variations of the renormalisation and factorisation scales are dominant.

8.3.6 Background-Only Fit Results

The results of the background-only fit for $\mathrm{tW_{2L}}$ in the CRs, VRs and SR are shown respectively in Tab. 8.2, 8.3 and 8.5, alongside with the number of events observed in data in each region. A good agreement between MC and data is observed in all CRs and VRs, while a mild excess of data is observed in $\mathrm{SR_{tW_{2L}}}$, accounting for a discrepancy around 2σ considering statistical and systematic uncertainties. The actual significance of this excess will be calculated with the model independent fit, presented in the next section.

The main kinematic distribution for both data and MC samples in all control and validation regions were presented in Figs. 8.6, 8.8, 8.9, 8.7 and 8.10. Good agreement was found in all distributions, confirming the validity of the approach chosen for the analysis. Fig. 8.15 shows instead comparisons between the observed data and the post-fit SM predictions for some relevant kinematic distributions after applying all $\mathrm{SR_{tW_{2L}}}$ selection requirements except the one on the quantity shown, together with the expected distribution for a couple of representative DMt signal points for illustrative purposes. Reasonable agreement is found between data and MC predictions in all distributions.

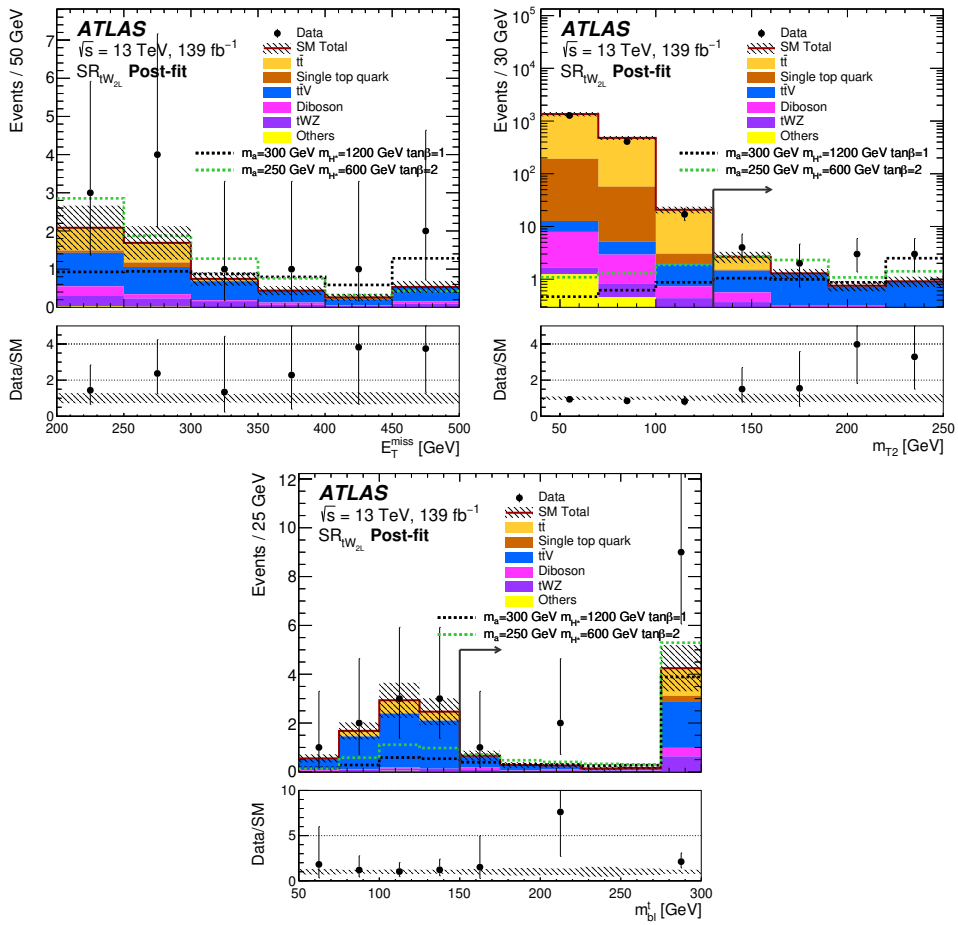


Figure 8.15: Kinematic distributions of main discriminating variables after all the $SR_{tW_{2L}}$ cuts except the cut on the variable itself are applied. The signal samples displayed are generated with $\sin\theta = 1/\sqrt{2}$. Please note that m_{bl}^{t} is defined only for events with at least two jets, events with exactly one jet are put in the overflow bin.

	$SR_{tW_{2L}}$
Observed events	12
Fitted bkg events	5.85 ± 1.21
$t\bar{t}$	1.17 ± 0.90
Single top	$0.26^{+0.27}_{-0.26}$
W-jets	0.00 ± 0.00
Z-jets	0.00 ± 0.00
Diboson	0.52 ± 0.24
$t\bar{t}V$	2.90 ± 0.71
$t\bar{t}h$	0.06 ± 0.01
tWZ	0.84 ± 0.15
Fakes	0.10 ± 0.07

Table 8.5: Background-only fit results for $SR_{tW_{2L}}$. The quoted uncertainties on the fitted SM background include both the statistical and systematic uncertainties.

8.3.7 Model Independent Limits

The results obtained in the $SR_{tW_{2L}}$, where only a small data excess has been observed, can be interpreted as exclusion limits for new physics scenarios. Tab. 8.6 shows upper limits (at the 95% C.L.) on the number of BSM events, calculated given both the observed and expected number of events. In addition, also upper limits on the visible BSM cross-section $\langle \epsilon \mathscr{A} \sigma \rangle^{95}_{\text{obs}}$ are presented, where ϵ is the analysis selection efficiency and $\mathscr{A}$ the acceptance, while σ is the production cross-section for the generic BSM process giving the signature analysed. CL_B values, i.e. the confidence level observed for the background-only hypothesis, are also shown. Finally, the p_0-values, which represent the probability of the SM background to fluctuate to the observed number of events or higher, are also provided and are capped at $p_0 = 0.5$; the associated significance of the excess is provided in parentheses.

Signal channel	$\langle \epsilon \mathscr{A} \sigma \rangle^{95}_{\text{obs}}$[fb]	S^{95}_{obs}	S^{95}_{exp}	CL_B	$p(s=0)$ (Z)
$\text{SR}_{\text{tW}_{2\text{L}}}$	0.10	13.8	$7.3^{+2.9}_{-1.1}$	0.97	0.02 (2.07)

Table 8.6: 95% C.L. upper limits on the visible cross-section ($\langle \epsilon\sigma \rangle^{95}_{\text{obs}}$) and on the number of signal events (S^{95}_{obs}) for $\text{SR}_{\text{tW}_{2\text{L}}}$. The third column ($S^{95}_{\text{exp}}$) shows the 95% C.L. upper limit on the number of signal events, given the expected number (and $\pm 1\sigma$ excursions on the expectation) of background events. The last two columns indicate the CL_B value, i.e. the confidence level observed for the background-only hypothesis, and the discovery p-value ($p(s=0)$). The associated significance is provided in parentheses.

8.4 One-Lepton Search for DM Produced in Association with a Single Top Quark in the 2HDMa Model

The production of dark matter in association with a single top quark has also been investigated in the one-lepton final state [2]. In order to maximise the sensitivity to DMt production, a statistical combination of the one- and two-lepton analyses was performed. A brief overview of the one-lepton analysis ($\text{tW}_{1\text{L}}$ in the following) is thus presented in the next section.

8.4.1 Event Selection

The experimental signature of the $\text{tW}_{1\text{L}}$ analysis includes one lepton, at least three jets from the top quark and W decays and $E^{\text{miss}}_{\text{T}}$ from DM production. Final state events are then selected requiring exactly one lepton with $p_T > 30$ GeV, at least three energetic jets, of which exactly one has to be b-tagged, and $E^{\text{miss}}_{\text{T}} > 250$ GeV. The $E^{\text{miss}}_{\text{T}}$ trigger is required to be fired.

The main background for the one-lepton DMt final state are:

- Reducible backgrounds, like Z+jets and dibosons
- $t\bar{t}$ production, with an additional b-jet and $E^{\text{miss}}_{\text{T}}$ coming only from W decays
- Wt production, with a final state identical to the signal, but with $E^{\text{miss}}_{\text{T}}$ kinematically bounded from W mass
- Irreducible backgrounds, like $t\bar{t}Z$ events with Z decaying invisibly, $t\bar{t}W$ events with a lost lepton from the decay of the W, and W+jets production.

Variable [Unit]	SR	CR($t\bar{t}$)	CR(W)
$\mathcal{N}_\ell$	1	1	1
$p_T(\ell)$ [GeV]	> 30	> 30	> 30
$\mathcal{N}_{\text{jet}}$	≥ 3	≥ 3	≥ 3
$p_T(j_1, j_2, j_3)$ [GeV]	$> (50, 50, 30)$	$> (50, 50, 30)$	$> (50, 50, 30)$
$\mathcal{N}_{b-\text{jet}}$	$= 1$	≥ 2	$= 1$
$p_T(b_2)$ [GeV]	< 50	> 50	< 50
$m_W^{\text{reclustered}}$ [GeV]	> 60	incl.	< 60
m_T [GeV]	> 200	> 200	$\in [40, 100]$
am_{T2} [GeV]	> 220	< 220	> 220

Table 8.7: Summary of signal and control region definitions used in the $\text{tW}_{1\text{L}}$ analysis channel. The W+jets control regions are further split into regions with opposite lepton charges as described in the text.

In order to reduce these backgrounds, the kinematic variables introduced in Secs. 6.3.1 -6.3.2 are used, and the selection cuts summarised in Tab. 8.7 are exploited to define the final signal region $\text{SR}_{\text{tW}_{1\text{L}}}$. To further increase the sensitivity of the analysis, a shape-fit on the $E_{\text{T}}^{\text{miss}}$ variable is performed.

8.4.2 Background Estimation

The dominant background contributions for the $\text{tW}_{1\text{L}}$ analysis come from $t\bar{t}$ and W+jets processes, for which dedicated control regions are defined. The $\text{CR}_{\text{tW}_{1\text{L}}}(\text{t}\bar{\text{t}})$ is obtained requiring at least two b-jets, inverting the SR selection on am_{T2} and removing the requirement on $m_W^{\text{reclustered}}$. The $\text{CR}_{\text{tW}_{1\text{L}}}(\text{W})$ selects events with $40 < m_T < 100$ GeV and $m_W^{\text{reclustered}} < 60$ GeV. In order to exploit the lepton charge asymmetry of the W+jets production with respect to other backgrounds, the CR is split into two regions with opposite lepton charged. The selection cuts for the CRs are summarized in Tab. 8.7.

8.4.3 Background-Only Fit Results

Background-only fit results in the $E_{\text{T}}^{\text{miss}}$-binned SRs of the one-lepton analysis are shown in Tab. 8.8. A good agreement between MC and data is observed, with a $\sim 1\sigma$ excesses in bins 0 and 1 and a $\sim 1\sigma$ underfluctuation of data in bin 4.

	$SR^{250}_{tW_{1L}}$	$SR^{300}_{tW_{1L}}$	$SR^{400}_{tW_{1L}}$	$SR^{500}_{tW_{1L}}$	$SR^{600}_{tW_{1L}}$
Observed events	469	287	96	36	12
Fitted bkg events	431 ± 27	262 ± 20	91 ± 10	36 ± 5	15.5 ± 2.8
$t\bar{t}$	213 ± 25	111 ± 18	28 ± 7	7.5 ± 2.9	2.3 ± 1.5
Single top	43 ± 15	27 ± 12	9 ± 7	3.9 ± 3.8	$1.7^{+2.0}_{-1.7}$
W+jets	91 ± 8	63 ± 5	26 ± 3	12.0 ± 1.7	5.9 ± 1.1
Z+jets	3.8 ± 1.0	1.7 ± 0.7	0.6 ± 0.1	0.30 ± 0.05	0.15 ± 0.02
Diboson	26 ± 5	19.0 ± 3.9	9 ± 2	4.8 ± 1.0	2.6 ± 0.6
$t\bar{t}V$	47 ± 3	34.7 ± 2.4	15.3 ± 1.2	6.6 ± 0.5	2.5 ± 0.3
Others	7.5 ± 0.5	5.3 ± 0.4	2.3 ± 0.1	0.9 ± 0.1	0.40 ± 0.05

Table 8.8: Background-only fit results for tW_{1L} signal region bins. Events with E_T^{miss} above the lowest bin-threshold are retained in each bin. This is indicated by value X in $SR^X_{tW_{1L}}$. The quoted uncertainties on the fitted SM background include both the statistical and systematic uncertainties.

8.4.4 Model Independent Limits

Also for the one-lepton analysis, the results obtained in the SRs can be interpreted as exclusion limits for BSM physics. To retain discovery potential, the E_T^{miss} bins need to be redefined inclusively, i.e. all events above the lowest bin-threshold in E_T^{miss} are taken. The upper limits at 95% C.L. on the number of BSM events and on the BSM visible cross-section for all the SRs are shown in Tab. 8.9.

8.5 Statistical Combination of One- and Two-Lepton Analyses

As explained before, in order to maximise the sensitivity to DMt signatures in the context of the 2HDMa model, a statistical combination of the results from the tW_{1L} and tW_{2L} analyses was performed. The combination was carried out in a more sophisticated way than the $DMt\bar{t}$ one presented in Sec. 7.5, where for each point the analysis providing the best expected exclusion was simply chosen. The DMt combination was indeed performed through a simultaneous fit, which used all tW_{2L} and tW_{1L} regions at the same time, constraining the background normalised in CRs in the same regions as the respective individual analyses. This means that the likelihood of Eq. 6.3 used in the model dependent combined fit (Sec. 6.1.3) contains the product of the Poisson observation of all tW_{2L} and tW_{1L} signal and control regions. The result of using at the same time both tW_{2L} and tW_{1L} signal regions is an increased statistical power, which allows to exclude also

Signal channel	$\langle\epsilon\sigma\rangle^{95}_{\text{obs}}$[fb]	S^{95}_{obs}	S^{95}_{exp}	CL_B	$p(s=0)$ (Z)
$\text{SR}^{250}_{\text{tW}_{1L}}$	0.72	100.6	$66.7^{+32.7}_{-16.4}$	0.85	0.12 (1.16)
$\text{SR}^{300}_{\text{tW}_{1L}}$	0.51	70.8	$54.1^{+16.0}_{-15.9}$	0.85	0.15 (1.02)
$\text{SR}^{400}_{\text{tW}_{1L}}$	0.24	32.9	$29.4^{+10.1}_{-6.4}$	0.64	0.30 (0.52)
$\text{SR}^{500}_{\text{tW}_{1L}}$	0.14	18.9	$18.7^{+7.6}_{-4.9}$	0.52	0.45 (0.13)
$\text{SR}^{600}_{\text{tW}_{1L}}$	0.08	10.6	$12.0^{+2.7}_{-3.6}$	0.24	0.50 (0.00)

Table 8.9: 95% C.L. upper limits on the visible cross section ($\langle\epsilon\sigma\rangle^{95}_{\text{obs}}$) and on the number of signal events (S^{95}_{obs}) for all tW_{1L} SRs. The third column (S^{95}_{exp}) shows the 95% C.L. upper limit on the number of signal events, given the expected number (and $\pm1\sigma$ excursions on the expectation) of background events. The last two columns indicate the CL_B value, i.e. the confidence level observed for the background-only hypothesis, and the discovery p-value ($p(s=0)$). The associated significance is provided in parentheses.

regions of the parameter space not excluded by the individual analysis.

Experimental uncertainties in the background and signal samples are evaluated in the same way described in Sec. 8.3.5 for both analyses and correlated across channels. Modelling uncertainties from the same source for a given process are correlated, like for example the $t\bar{t}$ modelling uncertainties are correlated across tW_{1L} and tW_{2L} analyses. Signal systematic uncertainties are also correlated.

The model dependent limit for the combined analysis were obtained using the sum of tW_{1L} and tW_{2L} signal yield estimates for each generated sample, paying attention to remove possible overlaps between the samples.

The limits in the parameter space of the 2HDMa model obtained by the statistical combination of tW_{1L} and tW_{2L} analyses will be shown in Sec. 8.7, alongside the results coming from the independent analyses.

8.6 Contribution from DM Production in Association with a Top Pair

The signature of dark matter production in association with a single top quark in the Wt channel in the 2HDMa model closely resembles the signature of DM produced together with a top pair, having only one less b-jet in the final state. It is thus interesting to investigate the contribution of $DMt\bar{t}$ production in the SRs of the DMt analyses, in order to asses the complete sensitivity to the 2HDMa model.

Relying on the reweighting studies developed in Chapter 7, $DMt\bar{t}$ samples generated in the context of the DMSIMP model were employed for this study.

m_a	$m_{H^\pm} = m_A$	$\tan\beta$	DMt	$DMt\bar{t}$
100	400	1	109.8	52.3
100	600	1	207.7	48.7
100	800	1	132.5	48.5
100	1200	1	53.7	46.5
200	400	1	43.3	36.3
200	600	1	136.6	31.4
200	800	1	119.6	30.2
200	1200	1	53.3	30.5
250	1250	0.5	109.2	99.6
250	1250	1	41.6	24.8
250	1250	2	11.5	6.3
250	1250	3	6.2	2.8
250	600	0.5	142.8	103.2
250	600	1	108.7	26.7
250	600	2	51.1	6.8
250	600	3	27.4	3.1
300	400	1	3.8	27.5
400	500	1	1.8	11.5

Table 8.10: DMt and $DMt\bar{t}$ production yields in the inclusive $\mathrm{SR}^{250}_{\mathrm{tW_{1L}}}$ for different choice of the parameters of the model. It is assumed $m_{H^\pm} = m_A = m_H$ and $\sin\theta = 1/\sqrt{2}$.

In order to take into account the contribution coming from both the processes $pp \to t\bar{t}a \to t\bar{t}\chi\bar{\chi}$ and $pp \to t\bar{t}A \to t\bar{t}\chi\bar{\chi}$ in 2HDMa (see Sec. 7.2.1), for each parameters choice both the samples having a mediator mass m_{ϕ_p} equal to m_a and m_A are used. Their cross-section is rescaled to the NLO one of $DMt\bar{t}$ production in 2HDMa, and a weight accounting for their relative importance is added, given respectively by

$$w_a = \frac{\sigma_a}{\sigma_a + \sigma_A}, \tag{8.3}$$

$$w_A = \frac{\sigma_A}{\sigma_a + \sigma_A}, \tag{8.4}$$

where σ_a is the cross-section for the process $pp \to t\bar{t}a \to t\bar{t}\chi\bar{\chi}$ and σ_A the one for $pp \to t\bar{t}A \to t\bar{t}\chi\bar{\chi}$. This is the same strategy used in Chapter 7, since in both cases the two contributions are weighted based on their relative contribution to the total cross-section.

m_a	$m_{H^\pm} = m_A$	$\tan\beta$	DMt	$DMt\bar{t}$
100	400	1	22.5	4.3
100	600	1	29.3	4.0
100	800	1	19.6	3.9
100	1200	1	7.0	3.9
200	400	1	7.6	2.9
200	600	1	18.7	2.5
200	800	1	16.5	2.4
200	1200	1	6.8	2.4
250	1250	0.5	14.9	7.9
250	1250	1	5.6	2.0
250	1250	2	1.6	0.5
250	1250	3	0.7	0.2
250	600	0.5	22.2	8.1
250	600	1	14.8	2.1
250	600	2	7.5	0.5
250	600	3	4.3	0.2
300	400	1	0.7	1.8
350	400	1	0.2	1.4
400	500	1	0.3	0.7

Table 8.11: DMt and $DMt\bar{t}$ production yields in the $\mathrm{SR}_{\mathrm{tW_{2L}}}$ for different choice of the parameters of the model. It is assumed $m_{H^\pm} = m_A = m_H$ and $\sin\theta = 1/\sqrt{2}$.

Since the DMt analyses were optimised on single-top signatures, both analyses have a cut reducing the contribution of $DMt\bar{t}$ in the signal regions. For $\mathrm{tW_{1L}}$, this cut is the veto on a second b-jet, while for $\mathrm{tW_{2L}}$ it is the cut on $m^t_{b\ell} > 150$ GeV, which removes the majority of $DMt\bar{t}$ signal, as shown in [134] in the case of simplified models.

The contribution of both DMt and $DMt\bar{t}$ production in the signal regions of $\mathrm{tW_{1L}}$ and $\mathrm{tW_{2L}}$ analyses was investigated for different choice of the parameters of the model, assuming $m_{H^\pm} = m_A = m_H$ and $\sin\theta = 1/\sqrt{2}$. The results are presented in Tabs. 8.10-8.11 for the inclusive $\mathrm{SR}^{250}_{\mathrm{tW_{1L}}}$ and for the $\mathrm{SR}_{\mathrm{tW_{2L}}}$, respectively. From the tables it can be seen that the DMt production generally dominates over $DMt\bar{t}$, especially at low $H^\pm$ masses due to the resonant $H^\pm$ production (Fig. 8.3). However, when the decay $H^\pm \to W^\pm a$ is kinematically forbidden (i.e. $m_{H^\pm} - m_a \lesssim m_W$), the cross-section of DMt production drastically decrease, and the $DMt\bar{t}$ contribution dominates. The $DMt\bar{t}$ yields depend only

slightly on the value of m_A, since for the chosen value of $\sin\theta$ they receive the main contribution from the on-shell production of the a boson. Thus the $DMt\bar{t}$ relative contribution is more important at high $m_{H^\pm}$, where DMt production is dominated by a-strahlung and is not resonantly enhanced.

Tabs. 8.10-8.11 also show that the $DMt\bar{t}$ contribution decreases with $\tan\beta$ since, as explained in Sec. 7.2.2, its cross-section is proportional to $1/\tan\beta^2$. For this reason, $DMt\bar{t}$ production is almost negligible for $\tan\beta > 1$, while at small values of $\tan\beta$ it is enhanced. DMt production has instead a more complex dependence on $\tan\beta$, but it decrease for increasing $\tan\beta$ until $\tan\beta \sim 6$ (Sec. 8.2.2).

8.7 Model Dependent Limits

Since no significant excess of data has been observed in any signal region, the results of the analyses can be interpreted as exclusion limits on the parameter space of the 2HDMa model. Two benchmark has been selected to investigate the sensitivity of DMt analyses. They are defined as follows:

1. 2D scan in (m_a, $m_{H^\pm}$) assuming $\tan\beta = 1$, $\sin\theta = 1/\sqrt{2}$.

2. 2D scan in ($m_{H^\pm}$, $\tan\beta$) assuming $m_a = 250$ GeV, $\sin\theta = 1/\sqrt{2}$.

Fig. 8.16 presents the observed and expected exclusion contours for the two benchmarks for the $\mathrm{tW}_{1\mathrm{L}}$ and $\mathrm{tW}_{2\mathrm{L}}$ analyses, presented both individually and statistically combined. In Fig. 8.16 only the $DM\mathrm{t}$ signal is taken into account to better illustrate the sensitivity of the analyses to single-top signatures. Fig. 8.17 shows instead the observed and expected exclusion contours including also the contribution from the $DMt\bar{t}$ process, as explained in the previous section. For the combined analyses, the 1σ and 2σ uncertainty bands around the observed limit contours are reported, together with the variations obtained by changing the theoretical cross-section predictions for signal to be 15% above or below the nominal value (where 15% is chosen since it is the largest uncertainty in the signal yields across the plane). In both cases, the observed combined exclusion is worst than the one-lepton one at high values of $m_{H^\pm}$ due to the data excess present in the two-lepton signal region, which lowers the constraining power of the combined analyses.

The combination of DMt analysis, taking into account only DMt signal, is able to exclude masses of a below 280 GeV in the range 400 GeV–1.1 TeV, and up to 310 GeV for $m_{H^\pm}$ around 800 GeV. These limits are increased thanks to the contribution of $DMt\bar{t}$, which allows to exclude a masses below 190 GeV for $m_{H^\pm}$ as high as 1.3 TeV, and up to 330 GeV for $m_{H^\pm}$ around 800 GeV. As expected,

the $DMt\bar{t}$ effect in bigger for high values of $m_{H^\pm}$, where its relative contribution is more important. The analyses are not sensitive to higher masses of a because for $m_a > 350$ GeV the channel $a \to t\bar{t}$ opens and the BR in DM drastically lowers.

In the benchmark with $m_a = 250$ GeV, the combination of DMt analysis excludes values of $m_{H^\pm}$ between 550 GeV and 1.1 TeV for $\tan\beta$ around 1 considering only DMt signal, while with $DMt\bar{t}$ all $m_{H^\pm}$ values between 450 GeV and 1.5 TeV are excluded for $\tan\beta$ around and below unity, where the $DMt\bar{t}$ cross-section is bigger. For $H^\pm$ around 800 GeV, the DMt-only results show values of $\tan\beta$ as high as 1.3 excluded, while values below $\tan\beta = 1.5$ are excluded considering also $DMt\bar{t}$.

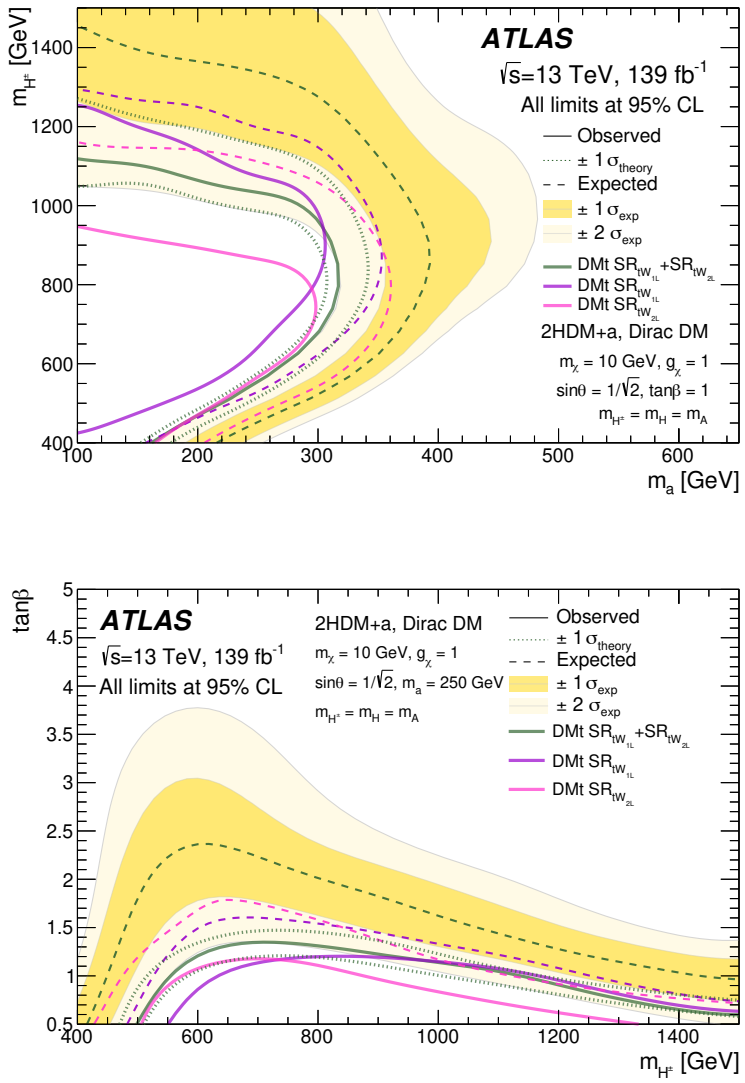


Figure 8.16: The expected and observed exclusion contours for benchmark 1 (top) and benchmark 2 (bottom) assuming only DMt contribution, for the individual tW_{1L} (purple-line) and tW_{2L} (pink-line) analyses and for their statistical combination (green-line).

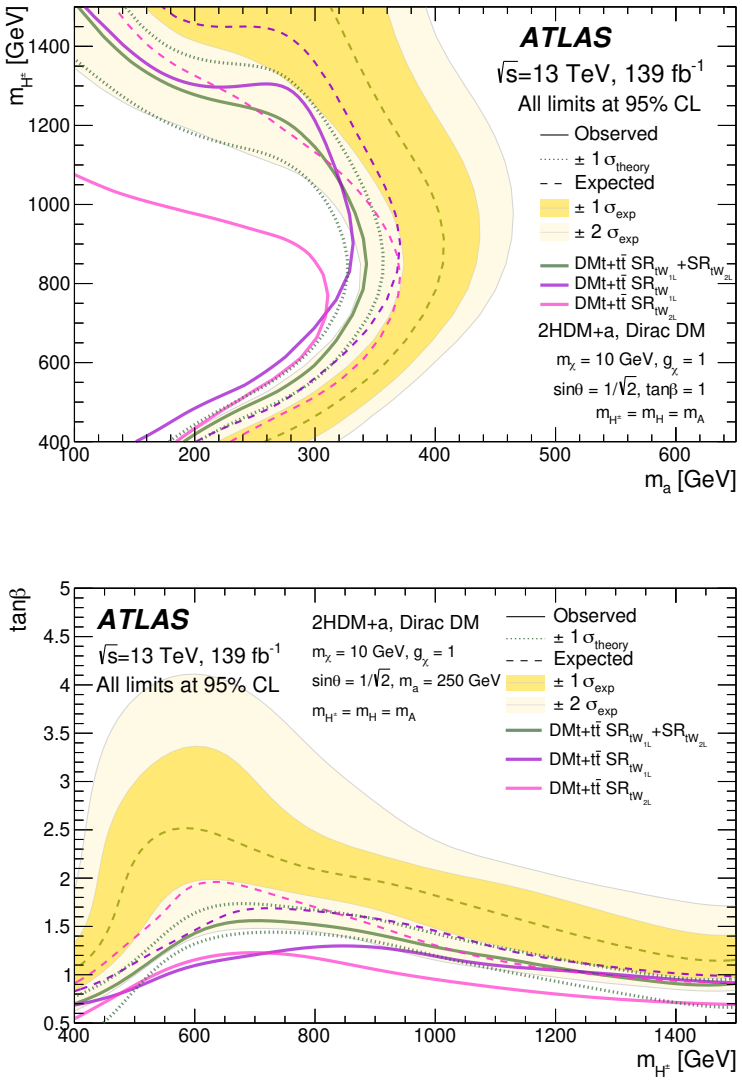


Figure 8.17: The expected and observed exclusion contours for benchmark 1 (top) and benchmark 2 (bottom) assuming $DMt\bar{t}$ and DMt contributions, for the individual tW_{1L} (purple-line) and tW_{2L} (pink-line) analyses and for their statistical combination (green-line).

Conclusions

Many cosmological observations give convincing evidence for the existence of dark matter (DM), estimated to constitute around 26% of the Universe by measurements of the cosmic microwave background. While the existence of DM thus seems well established, very little is known about its nature. Numerous models of DM have been proposed, and a possible strategy to test them is to use particle accelerators.

This thesis focused in particular on the detection of DM produced in the ATLAS experiment at the LHC, studying signatures with top quarks in the final state in the framework of the 2HDMa model. The latter is a recently presented model that introduces in the SM a pseudoscalar particle which mediates the interaction with DM and a second Higgs doublet, giving rise to a rich new physics phenomenology. The searches presented in this work are part of a bigger effort by the ATLAS Collaboration, aiming at exploring the 2HDMa model in its entiriety.

In the first part of this work, the production of DM in association with a pair of top quarks in the 2HDMa model was addressed and the regions of parameter space already excluded by existing analyses of the ATLAS Collaboration were identified. In order to map these results to the 2HDMa model, a dedicated recasting strategy was developed and successfully applied. The results of the analysis were included in the 2019 ATLAS DM summary paper [1].

The second part of the study focused instead on a new signature arising from the 2HDMa model, consisting in the production of DM together with a single top quark. This was the first analysis of such a signature in ATLAS, and new analysis strategies had to be developed for the different final states. The present thesis work concentrated on the two-lepton final state, which was described in detail. Good agreement was found between the observed data and the expected Standard Model yields, thus 95% C.L. limits on the parameters space of the 2HDMa model were set. Due to the similarity of this signature with the production of DM in association with a pair of top quarks, the contribution of the latter in the analyses optimised for single-top final states was also evaluated, allowing to set more stringent limits on the 2HDMa model. These results were published in a dedicated paper by the ATLAS Collaboration [2], and will soon be included in a new ATLAS DM summary paper focusing on 2HDMa searches.

Searching for DM at accelerator experiments is a thriving research field – both in terms of experimental strategies and theoretical developments. A large number of different searches are being performed by the experimental collaborations and a number of different models are available for predicting DM signals and interpreting results. A particularly exciting aspect is the complementarity between accelerator searches and alternative approaches to the search for DM, which promise to cover the full parameter space of many well-motivated models of DM. The discovery of DM is not guaranteed, since many DM models predict production cross-sections that are much too small to be observable. Nevertheless, a large spectrum of well-motivated theories can be tested at particle accelerators, providing precious information about the nature of DM.

Bibliography

[1] ATLAS collaboration, *Constraints on mediator-based dark matter and scalar dark energy models using $\sqrt{s} = 13$ TeV pp collision data collected by the ATLAS detector.*, *JHEP* **05** (2019) 142, [1903.01400].

[2] ATLAS collaboration, *Search for dark matter produced in association with a single top quark in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*, 2011.09308.

[3] K. G. Begeman, *HI rotation curves of spiral galaxies. I. NGC 3198.*, *Astronomy and Astrophysics* **223** (1989) 47–60.

[4] Hupp, E. and Roy, S. and Watzke, M., *NASA Finds Direct Proof of Dark Matter*, *NASA Press Release* **05** (2006) 06–297.

[5] D. Clowe et al., *A Direct Empirical Proof of the Existence of Dark Matter*, *The Astrophysical Journal* **648** (2006) L109–L113, [astro-ph/0608407].

[6] Massey, J. and Kitching, T. and Richard, J., *The dark matter of gravitational lensing*, *Rept.Prog.Phys.* **73** (2010) 086901, [1001.1739].

[7] Lokas, E. L., *Velocity dispersions of dwarf spheroidal galaxies: dark matter versus MOND*, *Mon. Not. Roy. Astron. Soc.* **327** (2001) L21, [astro-ph/0107479].

[8] N. Aghanim et al., *Planck 2018 results. VI. Cosmological parameters*, *Astronomy and Astrophysics* **641** (2020) 67, [1807.06209].

[9] M. Schumann, *Direct detection of WIMP dark matter: concepts and status*, *Journal of Physics G: Nuclear and Particle Physics* **46** (2019) 103003, [1903.03026].

[10] A. Del Popolo and M. Le Delliou, *Small Scale Problems of the λCDM Model: A Short Review*, *Galaxies* **5** (2017) 17.

[11] P. A. Zyla and others (Particle Data Group), *Progress of Theoretical and Experimental Physics*, vol. 8. Oxford University Press, 2020, 10.1093/ptep/ptaa104.

[12] S. Dodelson and L. M. Widrow, *Sterile-neutrinos as dark matter*, *Phys. Rev. Lett.* **72** (1994) 17–20, [hep-ph/9303287].

[13] K. Abazajian, G. Fuller and M. Patel, *Sterile neutrino hot, warm, and cold dark matter*, *Phys. Rev. D* **64** (2001) 22, [astro-ph/0101524].

[14] A. Ringwald, *Alternative dark matter candidates: Axions*, 2016. [1612.08933].

[15] S. Alekhin et al., *A facility to search for hidden particles at the CERN SPS: the SHiP physics case*, *Reports on Progress in Physics* **79** (2016) .

[16] ADMX Collaboration collaboration, *Search for Invisible Axion Dark Matter with the Axion Dark Matter Experiment*, *Phys. Rev. Lett.* **120** (2018) 151301.

[17] L. Zhong et al., *Results from phase 1 of the HAYSTAC microwave cavity axion experiment*, *Phys. Rev. D* **97** (2018) 092001.

[18] J. L. Feng, *Dark Matter Candidates from Particle Physics and Methods of Detection*, *Annual Review of Astronomy and Astrophysics* **48** (2010) 495–545, [1003.0904].

[19] S. J. Witte and G. B. Gelmini, *Updated Constraints on the Dark Matter Interpretation of CDMS-II-Si Data*, *Astropart. Phys.* **64** (2017) 26, [1703.06892].

[20] G. Angloher et al., *Results on low mass WIMPs using an upgraded CRESST-II detector*, *Eur. Phys. J.* **74** (2014) 014, [1407.3146].

[21] J. H. Davis, C. McCabe and C. Boehm, *Quantifying the evidence for Dark Matter in CoGeNT data*, *Astropart. Phys.* **08** (2014) , [1405.0495].

[22] P. Agnes et al., *Low-Mass Dark Matter Search with the DarkSide-50 Experiment*, *Phys. Rev. Lett.* **121** (2018) 081307.

[23] D. Akerib et al., *Dark Matter Searches with LUX*, *Phys. Rev. Lett.* **118** (2017) 021303, [1608.07648].

[24] E. Aprile et al., *XENON100 Dark Matter Results from a Combination of 477 Live Days*, *Phys. Rev. D* **94** (2016) 122001, [1609.06154].

[25] E. Aprile et al., *First Dark Matter Search Results from the XENON1T Experiment*, *Phys. Rev. Lett.* **119** (2017) .

[26] Q. Wang et al., *Results of dark matter search using the full PandaX-II exposure*, *Ch. Phys. C* **44** (2020) 125001.

[27] M. G. Aartsen et al., *Search for neutrinos from dark matter self-annihilations in the center of the Milky Way with 3 years of IceCube/DeepCore*, *Eur. Phys. J.* **77** (2017) 627, [1705.08103].

[28] E. Behnke et al., *Final Results of the PICASSO Dark Matter Search Experiment*, *Astropart. Phys.* **90** (2017) 85–92, [1611.01499].

[29] M. Felizardo et al., *Final Analysis and Results of the Phase II SIMPLE Dark Matter Search*, *Phys. Rev. Lett.* **108** (2012) 201302, [1106.3014].

[30] N. Tomassetti, *AMS-02 in space: physics results, overview, and challenges*, *Nuclear and Particle Physics Proceedings* **265–266** (2015) 245–247, [1511.00052].

[31] W. Atwood et al., *The Large Area Telescope on the Fermi Gamma-ray Space Telescope Mission*, *Astrophys. J* **697** (2009) 1071–1102, [0902.10899].

[32] K. Mora, *Dark Matter Searches with H.E.S.S.*, [1512.00698].

[33] F. Kahlhoefer, *Review of LHC Dark Matter Searches*, *Int. J. Mod. Phys.* **A32** (2017) 1730006, [1702.02430].

[34] M. Bauer, U. Haisch and F. Kahlhoefer, *Simplified dark matter models with two Higgs doublets: I. Pseudoscalar mediators*, *JHEP* **05** (2017) 138, [1701.07427].

[35] 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) 1–29, [1207.7214].

[36] 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) 30, [1207.7235].

[37] 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) 112004, [1706.03948].

[38] 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) 181804, [1707.01302].

[39] 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. B* **776** (2017) 318, [1708.09624].

[40] 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*, *JHEP* **10** (2018) 180, [1807.11471].

[41] ATLAS collaboration, *Search for dark matter produced in association with bottom or top quarks in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*, *Eur. Phys. J. C* **78** (2018) 18, [1710.11412].

[42] ATLAS collaboration, *Search for top-squark pair production in final states with one lepton, jets, and missing transverse momentum using 36 fb^{-1} of $\sqrt{s} = 13$ TeV pp collision data with the ATLAS detector*, *JHEP* **06** (2018) 108, [1711.11520].

[43] ATLAS collaboration, *Search for new phenomena in a lepton plus high jet multiplicity final state with the ATLAS experiment using $\sqrt{s} = 13$ Tev proton-proton collision data*, *JHEP* **09** (2017) 088, [1704.08493].

[44] LHC Dark Matter Working Group collaboration, *LHC Dark Matter Working Group: Next-generation spin-0 dark matter models*, *Phys. Dark Univ.* **27** (2020) 100351, [1810.09420].

[45] L. Evans and P. Bryant, *LHC Machine*, *JINST* **3** (2008) S08001.

[46] F. Marcastel, *CERN's Accelerator Complex*, https://cds.cern.ch/record/1621583, (2013).

[47] J. Stirling, *Parton luminosity and cross section plots*, http://www.hep.ph.ic.ac.uk/~wstirlin/plots/plots.html, (2012).

[48] ATLAS collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003.

[49] ATLAS collaboration, *ATLAS detector and physics performance: Technical Design Report*, https://cds.cern.ch/record/1291633, (1999).

[50] ATLAS collaboration, *ATLAS Insertable B-Layer Technical Design Report*, https://cds.cern.ch/record/391176, (2010).

[51] ATLAS collaboration, *Production and integration of the ATLAS Insertable B-Layer*, *JINST* **13** (2018) T05008.

[52] ATLAS collaboration, *Performance of the ATLAS Trigger System in 2015*, *Eur. Phys. J. C* **77** (2017) 317.

[53] S. Höche, *Introduction to parton-shower event generators*, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders (TASI 2014)*, vol. 108, pp. 235–295, 2015. 1411.4085.

[54] A. Buckley, *General-purpose event generators for LHC physics*, *Phys. Rept.* **504** (2011) 145–233, [1101.2599].

[55] F. Ambroglini et al., *Proceedings of the Workshop on Monte Carlo's, Physics and Simulations at the LHC PART I*, [0902.0293].

[56] J. C. Collins, D. E. Soper and G. F. Sterman, *Factorization of Hard Processes in QCD*, *Adv. Ser. Direct. High Energy Phys.* **5** (1989) 1–91, [hep-ph/0409313].

[57] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, *Nucl. Phys. B* **126** (1977) 298–318.

[58] V. N. Gribov and L. N. Lipatov, *Deep inelastic e p scattering in perturbation theory*, *Phys. Lett. B* **37** (1971) 78–80.

[59] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e^+e^- Annihilation by Perturbation Theory in Quantum Chromodynamics.*, *Sov. Phys. JETP* **46** (1977) 641–653.

[60] NNPDF collaboration, *Parton distributions for the LHC run II*, *JHEP* **2015** (2015) 40.

[61] Hou, T. and others, *New cteq global analysis of quantum chromodynamics with high-precision data from the lhc*, 1912.10053.

[62] A. D. Martin, W. J. Stirling, R. S. Thorne and G. Watt, *Parton distributions for the LHC*, *Eur. Phys. J. C* **63** (2009) 189–285.

[63] S. Catani, F. Krauss, B. R. Webber and R. Kuhn, *QCD Matrix Elements + Parton Showers*, *JHEP* **11** (2001) 063.

[64] L. Lönnblad, *Correcting the Colour-Dipole Cascade Model with Fixed Order Matrix Elements*, *JHEP* **05** (2002) 046.

[65] M. L. Mangano, M. Moretti, F. Piccinini and M. Treccani, *Matching matrix elements and shower evolution for top-pair production in hadronic collisions*, *JHEP* **01** (2007) 013.

[66] B. R. Webber, *A QCD Model for Jet Fragmentation Including Soft Gluon Interference*, *Nucl. Phys. B* **238** (2017) 492–528.

[67] G. Marchesini and B. R. Webber, *Monte Carlo Simulation of General Hard Processes with Coherent QCD Radiation*, *Nucl. Phys. B* **310** (1988) 461–526.

[68] B. Andersson, G. Gustafson, G. Ingelman and T. Sjostrand, *Parton Fragmentation and String Dynamics*, *Phys. Rept.* **97** (1983) 431–145.

[69] T. Sjostrand, *Jet Fragmentation of Nearby Partons*, *Nucl. Phys. B* **248** (1984) 469–502.

[70] ATLAS collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, https://cds.cern.ch/record/1966419, (2014).

[71] ATLAS collaboration, *The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie-Landshoff diffractive model*, https://cds.cern.ch/record/2206965, (2016).

[72] J. Alwall, R. Frederix and S. Frixione, *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **07** (2014) 79, [1405.0301].

[73] P. Nason, *A New Method for Combining NLO QCD with Shower Monte Carlo Algorithms*, *JHEP* **11** (2004) 040, [hep-ph/0409146].

[74] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, *JHEP* **11** (2007) 070, [0709.2092].

[75] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043, [1002.2581].

[76] T. Sjöstrand, S. Mrenna and P. Skands, *A brief introduction to PYTHIA 8.1*, *Comput. Phys. Commun.* **178** (2008) 852–867.

[77] T. Gleisberg, S. Höche, F. Krauss, M. Schönherr, S. Schumann et al., *Event generation with SHERPA 1.1*, *JHEP* **02** (2009) 007, [0811.4622].

[78] Bahr, M.l. and others, *Herwig++ physics and manual*, *Eur. Phys. J. C* **58** (2008) 639–707, [0803.0883].

[79] J. Bellm et al., *Herwig 7.0 / Herwig++ 3.0 Release Note*, *Eur. Phys. J. C* **76** (2016) 196, [1512.01178].

[80] S. Agostinelli et al., GEANT4 – *a simulation toolkit*, *Nucl. Instrum. Meth. A* **506** (2003) 250.

[81] ATLAS collaboration, *The ATLAS Simulation Infrastructure*, *Eur. Phys. J. C* **70** (2010) 823, [1005.4568].

[82] T. Cornelissen et al., *The new ATLAS track reconstruction (NEWT)*, *Journal of Physics: Conference Series* **119** (2008) 032014.

[83] ATLAS collaboration, *Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2*, *Eur. Phys. J. C* **77** (2017) 673, [1704.07983].

[84] ATLAS collaboration, *Reconstruction of primary vertices at the ATLAS experiment in Run 1 protonâĂŞproton collisions at the LHC*, *Eur. Phys. J. C* **77** (2017) 5, [1611.10235].

[85] ATLAS collaboration, *Vertex Reconstruction Performance of the ATLAS Detector at* $\sqrt{s} = 13$ *TeV*, https://cds.cern.ch/record/2037717, (2015).

[86] ATLAS collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at* $\sqrt{s} = 13$ *TeV*, *Eur. Phys. J. C* **79** (2019) 639, [1902.04655].

[87] ATLAS collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data*, *JINST* **14** (2019) P12006, [1908.00005].

[88] ATLAS collaboration, *Calorimeter Clustering Algorithms : Description and Performance*, https://cds.cern.ch/record/1099735, (2008).

[89] ATLAS collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at* $\sqrt{s} = 13$ *TeV*, *Eur. Phys. J. C* **76** (2016) 292, [1603.05598].

[90] M. Cacciari, G. P. Salam and G. Soyez, *The anti-ktjet clustering algorithm*, *JHEP* **04** (2008) 063.

[91] M. Cacciari, G. P. Salam and G. Soyez, *FastJet user manual*, *Eur. Phys. J. C* **72** (2012) 1896, [1111.6097].

[92] ATLAS collaboration, *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, *Phys. Rev. D* **96** (2017) 072002, [1703.09665].

[93] ATLAS collaboration, *Jet energy measurement and its systematic uncertainty in proton-proton collisions at* $\sqrt{s} = 7$ *TeV with the ATLAS detector*, *Eur. Phys. J. C* **75** (2015) 17, [1406.0076].

[94] ATLAS collaboration, *Performance of pile-up mitigation techniques for jets in in pp collisions at* $\sqrt{s} = 8$ *TeV using the ATLAS detector*, *Eur. Phys. J. C* **76** (2016) 581, [1510.03823].

[95] ATLAS collaboration, *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*, http://cdsweb.cern.ch/record/2160731/, (2016).

[96] ATLAS collaboration, *Secondary vertex finding for jet flavour identification with the ATLAS detector*, https://cds.cern.ch/record/2270366, (2017).

[97] ATLAS collaboration, *Topological b-hadron decay reconstruction and identification of b-jets with the JetFitter package in the ATLAS experiment at the LHC*, https://cds.cern.ch/record/2645405, (2018).

[98] ATLAS collaboration, *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*, https://cds.cern.ch/record/2273281, (2017).

[99] ATLAS collaboration, *ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **79** (2019) 970, [1907.05120].

[100] ATLAS collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **78** (2018) 903, [1802.08168].

[101] ATLAS collaboration, *E_T^{miss} performance in the ATLAS detector using 2015-2016 LHC $p-p$ collisions*, http://cds.cern.ch/record/2625233, (2018).

[102] W. Buttinger and M. Lefebvre, *Formulae for Estimating Significance*, https://cds.cern.ch/record/2643488, (2018).

[103] M. Baak, G. J. Besjes, D. Cote, A. Koutsman, J. Lorenz and D. Short, *HistFitter software framework for statistical data analysis*, *Eur. Phys. J.* **C75** (2015) 153, [1410.1280].

[104] A. L. Read, *Presentation of search results: the CL_S technique*, *J. Phys. G* **28** (2002) 2693.

[105] S. Amoroso, F. Siegert, J. Kretzschmar and C. Gutschow, *PMG references document*, https://cds.cern.ch/record/2678867, (2019).

[106] M. Czakon and A. Mitov, *Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders*, *Comput. Phys. Commun.* **185** (2014) 2930, [1112.5675].

[107] S. Frixione, P. Nason and G. Ridolfi, *A Positive-Weight Next-to-Leading-Order Monte Carlo for Heavy Flavour Hadroproduction*, *JHEP* **09** (2007) 126, [0707.3088].

[108] LHC Higgs Cross Section Working Group collaboration, D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, 1610.07922.

[109] M. Aliev, H. Lacker, U. Langenfeld, S. Moch, P. Uwer and M. Wiedermann, *HATHOR – HAdronic Top and Heavy quarks crOss section calculatoR*, *Comp. Phys. Comm.* **182** (2011) 1034–1046, [1007.1327].

[110] P. Kant et al., *HATHOR for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions*, *Comp. Phys. Comm.* (2015) 74–89, [1406.4403].

[111] N. Kidonakis, *Next-to-next-to-leading logarithm resummation for s-channel single top quark production*, *Phys. Rev. D* **81** (2010) 054028, [1001.5034].

[112] N. Kidonakis, *Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production*, *Phys. Rev. D* **83** (2011) 091503, [1103.2792].

[113] E. Re, *Single-top Wt-channel production matched with parton showers using the POWHEG method*, *Eur. Phys. J. C* **71** (2010) 1547, [1009.2450].

[114] S. Frixione, E. Laenen, P. Motylinski, C. White and B. R. Webber, *Single-top hadroproduction in association with a W boson*, *JHEP* **07** (2008) 029.

[115] J. M. Campbell, R. K. Ellis and C. Williams, *Vector boson pair production at the LHC*, *JHEP* **018** (2011) 034, [1105.0020].

[116] M. Grazzini, S. Kallweit, D. Rathlev and M. Wiesemann, *$W^{\pm}Z$ production at hadron colliders in NNLO QCD*, *Phys. Lett. B* **761** (2016) 179–183, [1604.08576].

[117] E. Bothmann et al., *Event Generation with Sherpa 2.2*, *SciPost Phys.* **7** (2010) 034, [1905.09127].

[118] ATLAS collaboration, *ATLAS simulation of boson plus jets processes in Run 2*, https://cds.cern.ch/record/2261937, (2017).

[119] ATLAS collaboration, O. Bessidskaia Bylund, *Modelling Wt and tWZ production at NLO for ATLAS analyses*, in *9th International Workshop on Top Quark Physics*, 11, 2016. 1612.00440.

[120] P. Konar, K. Kong, K. T. Matchev and M. Park, *Dark matter particle spectroscopy at the LHC: generalizing M_{T2} to asymmetric event topologies*, *JHEP* **04** (2010) 086, [0911.4126].

[121] C. G. Lester and B. Nachman, *Bisection-based asymmetric M_{T2} computation: a higher precision calculator than existing symmetric methods*, *JHEP* **03** (2015) 100, [1411.4312].

[122] ATLAS collaboration, *Search for top-squark pair production in final states with one lepton, jets, and missing transverse momentum using* 36.1 fb^{-1} *of* $\sqrt{s} = 13$ *TeV pp collision data with the ATLAS detector*, *JHEP* **06** (2018) 108, [1711.11520].

[123] C. Lester and D. Summers, *Measuring masses of semi-invisibly decaying particles pair produced at hadron colliders*, *Phys. Lett. B* **463** (1999) 99–103, [hep-ph/9906349].

[124] A. Barr, C. Lester and P. Stephens, *A variable for measuring masses at hadron colliders when missing energy is expected; m_{T2}: the truth behind the glamour*, *J. Phys. G* **29** (2003) 2343–2363, [hep-ph/0304226].

[125] U. Haisch and G. Polesello, *Searching for production of dark matter in association with top quarks at the LHC*, *JHEP* **02** (2019) 029, [1812.00694].

[126] U. Haisch, P. Pani and G. Polesello, *Determining the CP nature of spin-0 mediators in associated production of dark matter and $t\bar{t}$ pairs*, *JHEP* **02** (2017) 131, [1611.09841].

[127] P. Pani and G. Polesello, *Dark matter production in association with a single top-quark at the LHC in a two-Higgs-doublet model with a pseudoscalar mediator*, *Phys. Dark Univ.* **21** (2018) 8–15, [1712.03874].

[128] A. Alloul et al., *FeynRules 2.0 – A complete toolbox for tree-level phenomenology*, *Comp. Phys. Comm.* **185** (2014) 2250–2300, [1310.1921].

[129] D. Abercrombie et al., *Dark Matter benchmark models for early LHC Run-2 Searches: Report of the ATLAS/CMS Dark Matter Forum*, *Phys. Dark Univ.* **27** (2020) 100371, [1507.00966].

[130] M. R. Buckley, D. Feld and D. Goncalves, *Scalar simplified models for dark matter*, *Phys. Rev. D* **91** (2015) 015017, [1410.6497].

[131] U. Haisch and E. Re, *Simplified dark matter top-quark interactions at the LHC*, *JHEP* **06** (2015) 078, [1503.00691].

[132] ATLAS collaboration, *Measurement of the top quark-pair production cross section with ATLAS in pp collisions at* $\sqrt{s} = 7$ *TeV*, *Eur. Phys. J. C* **71** (2011) 1571, [1012.1792].

[133] ATLAS collaboration, *Measurement of the top quark pair production cross section in pp collisions at* $\sqrt{s} = 7$ *TeV in dilepton final states with ATLAS*, *Phys. Lett. B* **707** (2012) 459–477, [1108.3699].

[134] U. Haisch and G. Polesello, *Searching for production of dark matter in association with top quarks at the LHC*, *JHEP* **02** (2019) 029, [1812.00694].

Acknowledgements

First of all, I am deeply grateful to my supervisor, Giacomo, for his guidance throughout my entire PhD. Thank you for your patience and support, for your confidence in my abilities and for teaching me how to become an experimental physicist. A big thank you also to the entire group of ATLAS Pavia for having followed me in the last 5 years, since my bachelor thesis. In particular, I would like to thank Gabriella for always making sure everything was fine, and Daniela for introducing me to the group and for the constant support.

I would like to thank the referees of this thesis, Dan Tovey and Filip Moortgat. Thanks for reading and commenting on this work, your input was incredibly valuable and allowed me to improve the final result.

Part of my thesis was funded by the ATLAS PhD grant, a special grant scheme initiated by Fabiola Gianotti and Peter Jenni after the award of the Fundamental Physics Prize and supported by Lombard Odier. I am very grateful to all of them for this extraordinary opportunity.