VNIVERSITAT DE VALÈNCIA

Measurement of top-quark polarisation observables and a search for dark matter in single-top-quark production with the ATLAS detector at the LHC

Tesi doctoral

Galo Gonzalvo Rodríguez

IFIC Universitat de València - CSIC
Departament de Física Atòmica, Molecular i Nuclear
Programa de Doctorat en Física

Directors:

Dra. Mª José Costa Mezquita
Dr. José Enrique García Navarro

València, novembre 2021

La Dra. **María José Costa Mezquita**, investigadora científica del Consejo Superior de Investigaciones Científicas i el Dr. **José Enrique García Navarro**, científic titular del Consejo Superior de Investigaciones Científicas, com a directors de tesi, i el Dr. **Juan Zúñiga Román**, catedràtic del departament de Física Atòmica, Molecular i Nuclear de la Universitat de València, com a tutor de la Universitat

Certifiquen:

Que la present memòria "Measurement of top-quark polarisation observables and a search for dark matter in single-top-quark production with the ATLAS detector at the LHC" ha sigut realitzada sota la seua direcció o supervisió en l'Institut de Física Corpúscular, centre mixte de la Universitat de València i del CSIC, per **Galo Gonzalvo Rodríguez** i constitueix la seua tesi per a optar al títol de Doctor per la Universitat de València una vegada cursats els estudis en el Doctorat en Física.

I per a que així conste, en compliment de la legislació vigent, presenten en el departament de Física Atòmica, Molecular i Nuclear de la Universitat de València la referida tesi doctoral, i firmen el present certificat.

València, a 19 de novembre de 2021

María José Costa Mezquita José Enrique García Navarro Juan Zúñiga Román

Declaration

This dissertation is the result of my own work, except where explicit reference is made to the work of others, and has not been submitted for another qualification to this or any other university.

Galo Gonzalvo Rodríguez

To my family: my guidance, my heart and my spirit.

Acknowledgements

It has been very difficult to write these lines, since there are a lot of people whom I would like to express my gratitude. It has been an absolute pleasure to share this journey with you. Thanks to you all!

First of all, I would like to thank my mentors María José and Jose Enrique. Honestly, I cannot imagine a better combination of mentors. Thanks for giving me the chance to learn from you and develop my skills. From María José I could highlight her critical thinking and her ability to make me pursue the highest standard in every aspect of my work: from writing this thesis to presenting results in weekly meetings. From Jose Enrique, I would like to spotlight his ability to solve any technical problem. I still remember when you taught me how to make a loop over the different backgrounds at the beginning of my Master's project. Now, I feel confident to face any computer programming problem, thanks to your guidance. I am sure these skills will be a key aspect for my near future. However, I want to focus on the human part. Through these years, I had the chance to work with two amazing people, who made me love what I have been doing, love the discussions and made me feel an important part of something bigger, and I cannot express how thankful I feel for that.

I feel very grateful for the wonderful people I met from the ATLAS collaboration. In particular, I would like to thank Joe, Jordy, Jim, Marcel, Francesco, Andrea, Runyu, Nuno, Maura, Matthias and Stefano. Thanks for all those wonderful and inspiring discussions. Certainly, I could not forget about Marc: thanks for helping me during the toughest moments of the unfolding, for the discussions and for your friendship. It has been such a pleasure!

Now, I want to move to the incredible people I had the privilege to work with at IFIC. In particular, I want to thank the seniors Susana, Maria, Carmen, Santi, Luca, Nadia and Urmila and the students Josep, Jesus, Javi, Pablo and Florencia. It has been an absolute pleasure to work with you all, to chat with you and to share some laughs and thoughts. I would like to specially thank Carlos and Paolo for their help during this work. I am sure this thesis will not be the same without your guidance, without your patience and enthusiasm, thank you very much. Of course I could not forget about Salva, an absolute mentor to me. Thanks for your patience during my first steps, when I did not know how to even open a pdf from the terminal. Thanks for teaching me all the mysteries of the alignment and for introducing me to this world during my first year. It was such an incredible experience.

I need to highlight two very special friends: Oscar and Laura. It is hard to write in one paragraph how important you both have been for me during these years. Oscar, thanks for your incredible support during the toughest times. You are one of the most amazing people I met during this period and I wish you all the best in the next steps of the journey. Laura, you know it would not have been the same without you. You are such an incredible person, always willing to give assistance, and I learnt a lot from you. But most importantly, I earned a real friend. You introduced me to the people at IFIC and we had lots of chats about physics and no-physics who made this experience extremely enjoyable. Those chats at the office during the first years will always be remembered with much affection.

I would like to express my gratitude to Juan Zuñiga for introducing me to this unbelievable group of people. In addition, thanks for such a wonderful course on experimental techniques which helped me to find my path, and thank you for being a constant source of support and guidance.

Studying physics is such an interesting experience. Of course it can be very difficult and frustrating sometimes. Nevertheless, if you share this experience with your friends, this will for sure bring you lots of joy and happiness. I would like to thank my friends JJ, Rapha, Ruben and Nayara, whom I owe a lot. We have had such an amazing journey. You are the most beautiful part of studying physics, and I am sure we will keep building memories for many many years to come. I certainly could not forget about you, Hector: how many times studying in the library at

night, how many days booking the "study room" for ourselves. Those moments in the rooftop of the biology building drinking coffee and chatting about everything. We shared the good moments, but you also made the hardest part of the degree extremely pleasant, and I feel very grateful for that.

I want to thank also my co-worker friend Marcos. We worked very hard during the Master's degree and shared lots of resources. I feel we helped each other a lot during the process, and I really appreciate that. But most importantly, I met such an incredible person in you and a good friend. Thanks for all the good memories.

What is life if you cannot share the good and the bad moments with the people you love the most? In my case, these are my family. Thank you, thank you from the bottom of my heart for making me feel such a unique person, for all your love and all your support. Anytime I feel weak, I feel I may give up, I simply look back and see you right there next to me, giving me strength to carry on. You make me work harder and you are my inspiration to become the greatest person I can be. I would not have reached this landmark in my career without your support. Thank you!

I want to express my gratitude to my 'little brother' Sergi. You know to me you are part of my family. We were grown together and you have always been by my side when I needed it. When I look back to all this wonderful journey and see all the memories passing by through my mind, the Toy Story soundtrack comes to my mind: "you've got a friend in me". This friendship will last a lifetime.

I need to dedicate a few words to my grandmothers, Paqui and Isabel. I would not be the person I am today if it was not for you. Thank you!

To my grandmother Paqui... Thank you for taking care of me since I was a child. For all the sacrifices you made. I always love to discuss with you about any aspect of life. I love your advises and your unique point of view. The love between a grandson and a grandmother is something special, and you are one of the most important people in my life. You made me become the adult I am today, and you know you will always have me by your side, taking care of you as you have always done for me.

To my grandmother Isabel... You are my guardian angel. Words cannot describe how much I miss you, how much I wish I could hug you one more time, how much

I wish you were here, sharing this moment with me. Nonetheless, I know you are not far away. Every time I close my eyes I feel you right here, next to me, holding my hand. Tears come to my eyes but they always come with a warm smile in my face. Thank you for all your advises, for teaching me the importance of family, and especially thank you for making me feel so loved. You always have a special place in my heart. I love you, deeply.

Last but no least, I move to the three most important people in my life. My pillars and my inspiration. To my parents, Rafael and Enri. You are the mirror in which I see myself everyday. You taught me everything, you gave me the tools and you trusted in me. Now, I reached this incredible milestone in my life, and this is all thanks to you. You made me the happiest child in the world. Now, I am a grown up, but still there is a little happy child within me who has the need to give you lots of hugs every time I see you. I know I will always be your child, and you know you will always be my biggest source of inspiration, and my motivation to be the best me I can be everyday. I love you so much. To my sister, Paula. We have had such an incredible relationship. We have cried, we have laughed, and at the end of the day there is nothing but love. You are not just my sister, you are a best friend. I am so proud of the person you have become, and I feel extremely happy and grateful for having shared this path with you. Thanks for being always by my side, for listening to me, for all your advises, for all the moments, and for all the love.

Thank you all for this amazing journey!

Contents

Preface

The Standard Model of particle physics is the theoretical framework which provides the best description of the subatomic world. The discovery of the Higgs boson in 2012, the last piece of this theory, is still one of the foremost milestones in the recent history of particle physics. However, the Standard Model cannot be considered the ultimate theory. There are still important mysteries that are not explained within this theory, such as the nature of dark matter, the asymmetry between matter and antimatter observed in the universe and the observed non-zero neutrino masses.

The top quark is the most massive particle in the Standard Model. Due to its large mass, it presents some unique features. First of all, it is the only quark that decays before hadronising, thus enabling the study of its properties through its decay products. It has also the largest Yukawa coupling with the Higgs boson, which indicates that the top quark could play a special role in spontaneous symmetry breaking. Moreover, it is a crucial particle in new physics searches, since many models beyond the Standard Model predict the existence of new particles which would preferentially couple to top quarks.

For all these reasons the top quark constitutes a fundamental part of the LHC physics programme. The LHC is the most powerful particle accelerator in the world. Here, particle beams are accelerated to ultra-relativistic energies and collide at specific points, where the detectors are placed. ATLAS is one of the two general purpose detectors at the LHC. It was designed to discover the Higgs boson, to make precision measurements of Standard Model parameters and to search for new physics, all this in an unprecedented energy regime. In order to exploit maximally

its physics potential, it is essential to reach a deep understanding of the detector performance, including precise calibration and alignment procedures. The work in this thesis has contributed to this task through the alignment of the ATLAS inner detector. The alignment algorithm aims to determine the actual position of each component of the inner detector and to follow its possible changes in time. During Run-2, it was observed that the pixel detector and the IBL innermost layer of the inner detector showed characteristic patterns. These took place during each LHC fill. This work has contributed to the understanding of the pixel detector vertical movement, finding a clear relation between these movements and the luminosity of the fill. This study has been published in Ref. [1]. In addition, baseline alignment constants have been determined for different data-taking periods, on top of which additional refinements on the alignment corrections are provided on the daily basis to account for eventual small variations on the detector positions.

The LHC physics programme encompasses two different approaches in the search for new physics. The first is based on high precision measurements of the Standard Model processes, looking for deviations from the theoretical predictions. The second consists on direct searches of new physics, in particular in the form of new particles. This dissertation presents two different analyses in the top-quark sector each exploring one of these strategies: a high precision measurement of observables sensitive to the top-quark polarisation, and a search for dark matter particles produced in association with a single top quark. They both use pp collision data recorded by the ATLAS detector during Run-2 at $\sqrt{s} = 13$ TeV, which corresponds to a total integrated luminosity of 139 fb^{-1}.

In pp collisions at the LHC top quarks can be produced both in top quark-antiquark pairs in processes mediated by the strong interaction or singly in processes mediated by the electroweak interaction. In the single-top t-channel production at leading order, single top quarks are produced with their spin aligned in a particular direction. This is a consequence of the vector-axial structure of the tWb vertex. Due to its short lifetime, the top-quark spin information is transferred to its decay products. The first analysis described in this work presents the normalised differential cross-section of single-top-quark production in the t-channel, where the top quark decays leptonically, as a function of the three polarisation angles of

the lepton defined in Ref. [2]. These observables are particularly sensitive to new physics effects in the tWb vertex. The results of this analysis have been published in Ref. [3]. In addition, they have been used to set limits in the real and imaginary part of the dipole operator $C_{tW} = 0.4 \pm 1.1$ and $C_{itW} = -0.3 \pm 0.4$ at 95% confidence level as part of the PhD thesis in Ref. [4]. The limits on the imaginary part, which is sensitive to CP violation components, are the best limits from high-energy measurements at the time this work has been presented.

The second analysis described in this dissertation presents a direct search of dark matter particles produced in association with a single top quark. This leads to a characteristic final state in the detector consisting of a single top quark and missing transverse momentum, known as the mono-top signature. The analysis benefits from cutting-edge multivariate-analysis techniques to enlarge the signal discrimination from the Standard Model backgrounds. In the absence of any significant excess of data, these results are used to set upper limits on the corresponding dark matter production cross-section. The observed limits presented in this work constitute the most stringent constraints so far in the dark matter production cross-section for the considered models. This analysis is currently being reviewed within the ATLAS collaboration before being submitted for publication.

1 Introduction

The recent discovery of the Higgs boson [5, 6] at the Large Hadron Collider (LHC) [7, 8] has become a historical landmark in order to confirm the particle content of the Standard Model (SM) theory. This theory was developed in the early 1970s, and it has successfully predicted a wide variety of experimental results.

Within the SM context, the work in this thesis focuses on the top quark physics sector. In addition to the Higgs boson, the top quark is one of the most interesting particles in the SM due to its large mass, which implies that it has the largest coupling to the Higgs boson and makes it the only quark that decays before hadronising. In addition, many new physics theories predict new particles which would couple preferentially to the top quark (see Ref. [9] for a review).

In spite of the wide variety of experimental results in very good agreement with the SM, it cannot be considered the ultimate theory. There are still important mysteries that are not explained within this theory, such as the nature of dark matter (DM), the asymmetry between matter and antimatter observed in the universe and the observed non-zero neutrino masses. The SM does not describe either the gravitational interaction. At the LHC, there is a continuous effort in order to find hints of new physics which could help to explain these unknowns. The strategies can be grouped in two main categories. On the one hand, there are continuous efforts in the field of direct searches. This means to search for new particles not included in the SM that can be directly produced at the LHC. On the other hand, there are also strong efforts which focus on high precision measurements of the SM predictions. Deviations from these predictions could provide hints on where the new physics lies. In particular, if new particles are not directly accessible at the LHC,

they could still have an effect on measurements at a lower energy scale. One can learn about the properties of these new physics phenomena within an Effective Field Theory (EFT) framework. In this thesis two analyses are presented which cover the two possibilities; a direct search of DM particles produced in association with a top quark and a high precision measurement of observables which are sensitive to new physics in the tWb vertex.

This chapter will be organised as follows: Firstly, a description of the SM formalism is given. Secondly, an introduction to the characteristic features of the top quark and to its production mechanisms at the LHC is presented. This chapter concludes with an introduction to the theoretical framework in which the two presented analyses are developed.

1.1 The Standard Model in a nutshell

The SM of particle physics is a relativistic quantum field theory which describes the matter that forms the universe and its interactions. It identifies the matter with a set of fundamental particles, called *fermions*, and describes the interactions of these particles through the exchange of another set of particles, called *gauge bosons*.

Fermions are spin-1/2 particles which obey the Fermi-Dirac statistics and are subject to the Pauli exclusion principle. They can be divided in two main groups: leptons and quarks, depending on whether they carry charge under the strong interaction. They are further arranged in three generations, which share the same quantum numbers[1] but have different masses. The first generation corresponds to the lightest particles and constitutes the ordinary matter. The second and third generations are not stable and eventually will decay to particles from the first generation.

In the case of leptons, each generation is formed by a charged particle (electron, muon or tau) and its associated charge-less neutrino. In contrast, quarks are grouped in generations formed by a couple of charged quarks. Each generation contains an "up"-type quark, with charge $+\frac{2}{3}$, and a "down"-type quark with charge $-\frac{1}{3}$, both in electron charge units. In addition to the electric charge, quarks contain

[1] Quantum numbers describe values of conserved quantities in the dynamics of a quantum system.

a different kind of charge named "colour" charge. By convention, this colour charge can be classified in three categories: red, green and blue. Quarks never appear isolated in common matter. They always appear embedded in hadrons, which can be classified in baryons and mesons. Baryons are made of three different quarks (or antiquarks), whereas mesons are made of a quark-antiquark pair. Protons and neutrons are examples of baryons. Protons are mainly formed by two "up"-type quarks and one "down"-type quark of the first generation, whereas neutrons are formed by two "down"-type quarks and one "up"-type quark also from the first generation. Thus ordinary matter, made of atoms, are all constituted at the most fundamental level by first generation fermions: electrons, up quarks and down quarks.

Gauge bosons are spin-1 particles which obey the Bose-Einstein statistics. The SM describes three of the four interactions present in nature as the exchange of these particles. It includes the electromagnetic interaction, described by Quantum Electro-Dynamics (QED), the strong interaction, described by Quantum Chromodynamics (QCD) and the weak interaction[2]. Each interaction has associated at least one gauge boson. The mediator of the electromagnetic interaction is the photon (γ). Thus, the repulsion between two electrons can be understood as the exchange of photons which pull away one from the other. Similarly, the strong interaction is mediated by a total of 8 gluons (g). Finally, the $W^{\pm}$ and the Z bosons are the mediators of the weak interaction. The former mediates the charged weak currents and the latter the neutral weak currents. This is summarised in Table 1.1. A typical strength value is also given for reference. As one can see, the three interactions described by the SM are indeed very different. At usual energies, the strong interaction is by far the strongest interaction. Indeed, the strength of the strong interaction is such that, in addition to holding the quarks inside the hadrons, a residual force leaks out from it which is also responsible for holding the protons and the neutrons together within the atom nucleus.

Figure 1.1 summarises the components of the SM, including the Higgs boson. The role of this particle in the SM context will be described below. On the whole, 17 particles constitute the SM content. Apart from these particles, each fermion is

[2]Gravity is not included within the SM.

Interaction	Boson	Typical strength
Strong	8 gluons	1
Electromagnetic	Photon	10^{-3}
Weak	$W^{\pm}$, Z bosons	10^{-8}

Table 1.1: The three interactions of the SM and their respective gauge mediators. A typical strength value relative to the strong force is also given to perform a straight comparison [10]. However, the strength of each interaction depends on the distance and the interacting particles, so these numbers should not be used as reference.

accompanied by an antifermion, which has the same mass and opposite quantum numbers.

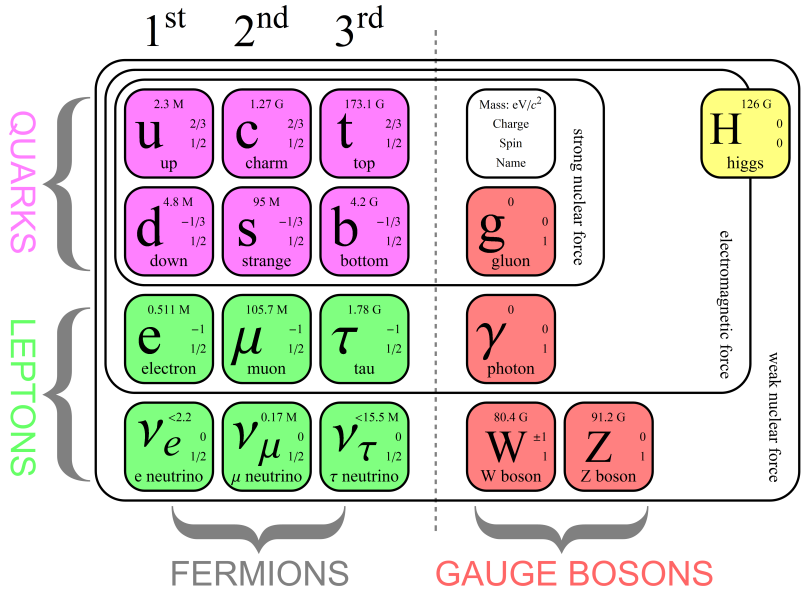


Figure 1.1: Elementary particles that constitute the SM. The first three columns include the three generations of fermions (leptons and quarks), which form the matter of the universe. The fourth and fifth columns include the gauge bosons, which mediate the interactions. Finally, the sixth column includes the Higgs boson. For each particle its mass, spin and electric charge are included [11].

All these particles and their interactions are described in the SM Lagrangian, which is based on the SM symmetry group $SU(3)_C \times SU(2)_L \times U(1)_Y$. The

$SU(3)_C$ is known as the colour group and it is associated with the strong force. The $SU(2)_L \times U(1)_Y$ is known as the electroweak group and it is associated with the electroweak interaction, which unifies both the electromagnetic and weak interactions above the energy scale of the order of 246 GeV. The SM Lagrangian is invariant under local transformations of these groups, called gauge transformations. The local invariance of the SM Lagrangian implies the introduction of bosonic fields associated to the gauge bosons previously described, which are responsible for the interactions. The different components of the SM symmetry group are described in more detail in the following.

1.1.1 The strong interaction

The strong interaction acts on both quarks and gluons, since they are the only colour-charged particles in the SM. Quarks transform as colour triplets under global transformations of $SU(3)_C$: $\boldsymbol{q}^T = (q, q, q)$. This implies that replacing all blue quarks by red quarks and vice versa would leave the Lagrangian unchanged, and the same holds for any orthonormal basis state. This is a global symmetry in the QCD Lagrangian, and it can be extended to a local symmetry by the introduction of the covariant derivatives:

$$D_\mu = \partial_\mu + i g_s T^a G^a_\mu, \tag{1.1}$$

where G^a_μ with $a = 1...8$ are the eight vector boson fields which represent the gluons. These particles are colour charged, and thus gluon self-interactions are allowed in the SM. The generators of the $SU(3)_C$ symmetry group are depicted as T^a, and g_s is the so-called ***bare QCD coupling***. The QCD Lagrangian can be written as:

$$\mathcal{L}_{QCD} = -\frac{1}{4} G^{\mu\nu}_a G^a_{\mu\nu} + \bar{\boldsymbol{q}}[i\gamma^\mu D_\mu - m_q]\boldsymbol{q}, \tag{1.2}$$

where $G^{\mu\nu}_a = \partial^\mu G^\nu_a - \partial^\nu G^\mu_a - g_s f^{abc} G^\mu_b G^\nu_c$, being f^{abc} the structure constants. The first term describes the kinematics of the gluon fields and their self interactions. The second term includes the free Dirac fermion term and the interaction term between the quarks and gluons.

There are some features of the strong interaction which are characteristic of this force. These are related to the coupling constant of the strong interaction $\alpha_s = g_s^2/4\pi$. This parameter increases with the distance between the interacting particles. At high energy scales (i.e. short distances), such as those in proton-proton (pp) collisions at the LHC, the strong interaction can be described using perturbative QCD (pQCD). A one loop calculation in pQCD gives:

$$\alpha_s(Q^2) = \frac{\alpha_s(Q_0^2)}{1 - \beta_1 \frac{\alpha_s(Q_0^2)}{2\pi} \ln Q^2/Q_0^2}, \tag{1.3}$$

where $\alpha_s(Q_0^2)$ is the value of the strong coupling constant at a fixed energy scale Q_0^2, which is determined from experiment. A common choice is the current world average value $\alpha_s(M_Z^2) = 0.1179 \pm 0.0010$ [12]. The term $\beta_1 = (2N_f - 11N_C)/6$ is always negative, being N_f the number of active quark flavours[3] and N_C the number of colours. As one can see, at a very large energy scale Q^2, the coupling constant is actually very weak, and this means that the quarks are not bound. This is the so-called *asymptotic freedom*. In contrast, at low energies (i.e. long distances), the value of the coupling constant is much stronger. Indeed, at some point[4], it is more energetically favorable to produce a new pair of quarks from the vacuum than pulling a couple of quarks away from each other. This is the so-called *confinement*, and implies that quarks cannot be isolated but they are only present within hadrons, which are colour-neutral. The production of hadrons is called *hadronisation*. This is a repetitive process which turns out in the formation of a cascade of hadrons, which is usually named a *jet*.

1.1.2 The electroweak interaction

The electroweak interaction [13–15] unifies both the weak and the electromagnetic interactions. It corresponds to the $SU(2)_L \times U(1)_Y$ component in the SM symmetry group. Electroweak interactions can lead to both charged

[3] If $m_q^2 \gg Q^2$, being m_q the mass of a quark q, then the quark is not active.

[4] This corresponds to the QCD scale $\Lambda_{\text{QCD}} \sim 200$ GeV, which is the energy scale at which the perturbatively defined strong coupling constant diverges. Its value is indicative of the energy range where non-perturbative dynamics dominates.

currents, when the interacting particles couple to a $W^{\pm}$ boson, and neutral currents, when the interacting particles couple to a Z or γ boson.

The weak interaction is left-chiral[5], meaning that the charged W gauge bosons only couple to left-chiral quarks and leptons, and its Lagrangian is invariant under local transformations of the $SU(2)_L$ symmetry group in the massless limit. The weak isospin is always conserved in weak interactions, similarly to the electric and the colour charge in the electromagnetic and the strong interactions, respectively. Fermions transform as weak isospin doublets and singlets under this gauge group according to their chirality (see Table 1.2). The two left-chiral components in each doublet have opposite T_3, which is the third component of the weak isospin. The weak isospin doublets in the lepton sector associate the left-chiral component of the lepton field with the left-chiral component of its associated neutrino. In the quark sector, quarks are arranged in doublets with an up-type quark component and a down-type quark component, both left-chiral. In contrast with leptons, quark mixing between different generations is allowed in the charged-current weak interactions. This is due to the fact that quarks are massive particles in the SM in contraposition with neutrinos[6], and the physical (mass) eigenstates (d, s, b) do not coincide with the flavour eigenstates (d', s', b'). The relation between these eigenstates is given by the Cabibbo-Kobayashi-Maskawa (CKM) matrix:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \tag{1.4}$$

where the different elements of the CKM matrix are determined experimentally and listed in Table 1.3 [12]. The largest values are found in the diagonal elements, implying that flavour-changing charged currents are not preferred but still possible.

[5] Something is chiral if it is not identical to its mirror image. The two possible chiral states are named left- and right-chiral. In the massless limit, chirality coincides with helicity, which is the projection of the spin direction of a particle onto its momentum direction. A chiral theory is asymmetric with respect to chiralities.

[6] In the SM, neutrinos are massless. However, it is known that neutrinos are massive, even if their masses are very low. Due to their tiny masses there is mixing between the flavour and mass eigenstates of neutrinos, which leads to the phenomenon of neutrino oscillations. See Section 1.1.5.

$SU(2)_L$ doublets	$SU(2)_L$ singlets
$\begin{pmatrix} u_L \\ d_L \end{pmatrix}$	u_R, d_R
$\begin{pmatrix} c_L \\ s_L \end{pmatrix}$	c_R, s_R
$\begin{pmatrix} t_L \\ b_L \end{pmatrix}$	t_R, b_R
$\begin{pmatrix} \nu_{e,L} \\ e_L \end{pmatrix}$	e_R
$\begin{pmatrix} \nu_{\mu,L} \\ \mu_L \end{pmatrix}$	μ_R
$\begin{pmatrix} \nu_{\tau,L} \\ \tau_L \end{pmatrix}$	τ_R

Table 1.2: Weak isospin doublets and singlets according to $SU(2)_L$ transformations. Since weak interactions are left-chiral, the left-chiral (L) components of the fermion fields are arranged in doublets, whereas the right-chiral (R) components are singlets under $SU(2)_L$ because they do not interact weakly. Since neutrinos are only left-chiral in the SM, the right-chiral component is not present. The $SU(2)_L$ doublets also share the same weak hypercharge Y.

In particular, the V_{tb} element of the CKM matrix is very close to unity. This implies that the top quark couples almost exclusively to a bottom quark through the weak interactions. This matrix can be parametrised in terms of four parameters, three of which are real and one being a complex phase factor:

$$V = \begin{pmatrix} c_{12}c_{13} & c_{13}s_{12} & s_{13}e^{-i\delta_{13}} \\ -c_{23}s_{12} - c_{12}s_{13}s_{23}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{13}s_{23}e^{i\delta_{13}} & c_{13}s_{23} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - c_{23}s_{12}s_{13}e^{i\delta_{13}} & c_{13}c_{23} \end{pmatrix}, \quad (1.5)$$

where $c_{ij} = \cos\theta_{ij}$, $s_{ij} = \sin\theta_{ij}$ and δ_{13} is the complex phase. The presence of this complex phase implies a violation of the charge-conjugation-parity symmetry (CP violation) in the weak interactions.

CKM element	Global fit value
$\|V_{ud}\|$	0.97401 ± 0.00011
$\|V_{us}\|$	0.22650 ± 0.00048
$\|V_{ub}\|$	$0.00361^{+0.00011}_{-0.00009}$
$\|V_{cd}\|$	0.22636 ± 0.00048
$\|V_{cs}\|$	0.97320 ± 0.00011
$\|V_{cb}\|$	$0.04053^{+0.00083}_{-0.00061}$
$\|V_{td}\|$	$0.00854^{+0.00023}_{-0.00016}$
$\|V_{ts}\|$	$0.03978^{+0.00082}_{-0.00060}$
$\|V_{tb}\|$	$0.999172^{+0.000024}_{-0.000035}$

Table 1.3: Magnitude of the nine CKM matrix elements from a global fit to all available measurements and imposing the SM constraints (i.e. three generation unitarity) [12].

In the case of neutral currents, a fermion and its associated antifermion interact with a Z boson. In contrast with charged currents, flavour-changing neutral currents (FCNC), which would take place when a quark and an antiquark of different flavour exchange a Z boson, are forbidden in the SM at tree level and highly suppressed at higher orders by the GIM mechanism [16].

The electric charge is associated with the $U(1)_{EM}$ symmetry group. In order to incorporate the electromagnetic interaction within the weak interaction formalism, one needs to define the hypercharge. The hypercharge Y is defined from the electric charge Q and the third component of the weak isospin T_3 of a particle, according to the relation $Y = 2(Q - T_3)$. With this definition, all the elements within a weak isospin doublet have the same hypercharge and are invariant under $U(1)_Y$. Thus, incorporating the electromagnetic interaction, one can build a larger symmetry group $SU(2)_L \times U(1)_Y$. This is the so-called *electroweak interaction* gauge group. Similarly to the strong interactions, new covariant derivatives can be introduced to impose local symmetry:

$$D_\mu = \partial_\mu - i\frac{g}{2}T_a W_\mu^a - i\frac{g'}{2}YB_\mu, \tag{1.6}$$

where Y is the weak hypercharge, T_a are the weak isospin components, with $a = 1...3$, and g (g') is the weak (hypercharge) coupling. These covariant derivatives imply the existence of vector boson fields W_μ^a and B_μ, which are intimately related with the $W^\pm$ boson fields $W_\mu^\pm$, the Z boson field Z_μ and the photon field A_μ:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \pm iW_\mu^2), \tag{1.7a}$$

$$\begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix}, \tag{1.7b}$$

where $\cos\theta_W = \frac{g}{\sqrt{g^2+g'^2}}$ and $\sin\theta_W = \frac{g'}{\sqrt{g^2+g'^2}}$, being θ_W the Weinberg angle.

These bosons have a weak isospin component, and thus self-interactions between them take place in the SM. With the introduction of these covariant derivates, the electroweak Lagrangian can be written as:

$$\mathcal{L}_{EW} = -\frac{1}{4}B^{\mu\nu}B_{\mu\nu} - \frac{1}{4}W_a^{\mu\nu}W_{\mu\nu}^a + i\sum_i \bar{\psi}_i\gamma^\mu D_\mu\psi_i, \tag{1.8}$$

where $B^{\mu\nu} = \partial^\mu B^\nu - \partial^\nu B^\mu$ and $W_a^{\mu\nu} = \partial^\mu W_a^\nu - \partial^\nu W_a^\mu - g\epsilon^{abc}W_b^\mu W_c^\nu$, being ϵ^{abc} the Levi-Civita tensor. The index i runs over all weak isospin singlets and doublets listed in Table 1.2.

1.1.3 Spontaneous symmetry breaking

The introduction of mass terms in the electroweak Lagrangian would break the symmetry. However, it is known that fermions and the $W^\pm$ and Z gauge bosons are massive. In order to recover the mass of the particles without losing the properties of this symmetry, the Brout-Englert-Higgs mechanism is introduced [17–20]. This mechanism assumes a global symmetry of the Lagrangian which is not shared by the physical vacuum. A new complex scalar field ϕ is introduced, and the SM Lagrangian is expanded including the Higgs term:

$$\mathcal{L}_{Higgs} = (D_\mu\phi)^\dagger(D_\mu\phi) - V(\phi) \tag{1.9a}$$

$$V(\phi) = \mu^2|\phi|^2 + \lambda|\phi|^4, \tag{1.9b}$$

where ϕ is the Higgs field and $\mu^2 < 0$ and $\lambda > 0$ are the parameters of the Higgs potential $V(\phi)$. A positive value of λ guarantees that the potential is bounded from below, i.e. there is a ground state. On the other hand, if μ^2 were positive this potential would have an unique vacuum state with $\phi = 0$. Since μ^2 is negative, $\phi = 0$ is an unstable local maximum of the effective potential and the minimum has $\phi \neq 0$ with an arbitrary $SU(2)_L \times U(1)_Y$ orientation.

A schematic view of the Higgs potential is displayed in Figure 1.2. As one can see, this potential is symmetric. However, when the field is situated at a minimum, the ground state does no longer share the symmetries of the Lagrangian. This process is the so-called *spontaneous symmetry breaking* (SSB). In the SM, choosing a particular ground state corresponds to breaking the $SU(2)_L \times U(1)_Y$ symmetry group into the $U(1)_{EM}$ symmetry group. Expanding the scalar field around the chosen vacuum state leads to mass terms in the Lagrangian, which is the procedure to recover the masses of the SM particles without breaking gauge invariance. This expansion produces an additional scalar field, which is identified with the Higgs boson. Therefore, the discovery of the Higgs boson in 2012 constitutes one of the most important milestones in the recent history of particle physics because it proved the Brout-Englert-Higgs mechanism and the origin of particle masses.

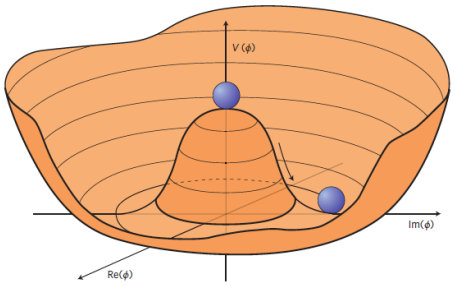


Figure 1.2: Scalar potential with SSB [21].

According to the Goldstone theorem [22–24], if there is an exact continuous symmetry of the Lagrangian which is not shared by the physical vacuum, then there must be at least one massless scalar particle called a Goldstone boson.

However, in the SM context, the degree of freedom associated to this scalar particle is reabsorbed by the vector boson fields, becoming massive in the process. This ensures that the $W^{\pm}$ and the Z bosons recover their masses through SSB. Concretely, the mass of the vector bosons and the Higgs boson are given by:

$$M_W = \frac{1}{2} v g, \tag{1.10a}$$

$$M_W = \frac{M_W}{\cos\theta_W}, \tag{1.10b}$$

$$M_H = \sqrt{2\lambda} v \tag{1.10c}$$

The masses of the SM fermions come from the relation of their quantum fields with the Higgs field. This leads to the so-called Yukawa Lagrangian:

$$\mathcal{L}_{Yukawa} = \sum_{ij} [\bar{l}_{Li} y^e_{ij} e_{R_j} \phi + \bar{Q}_{Li} y^d_{ij} d_{Rj} \phi + i \bar{Q}_{Li} y^u_{ij} u_{Rj} \sigma_2 \phi^*] + h.c., \tag{1.11}$$

where σ_2 is the second Pauli matrix, $y^{e,u,d}_{ij}$ are the matrices of Yukawa couplings, l_{Li} (Q_{Li}) is the left-chiral weak-isospin lepton (quark) doublet field for the i-th generation, e_{Rj} is the right-chiral weak-isospin lepton singlet field, and u_{Rj} (d_{Rj}) is the right-chiral weak-isospin up-(down-)quark singlet field for the j-th generation (see Table 1.2).

The fermion masses are found to be related to the Higgs vacuum expectation value $v = (\sqrt{2} G_F)^{-1/2} \approx 246$ GeV, where G_F is the Fermi constant, according to the relation $m_f = \lambda_f \frac{v}{\sqrt{2}}$, where λ_f is the Yukawa coupling of the given fermion to the Higgs boson. Due to its large mass, the top quark is the SM particle with the highest coupling to the Higgs boson ($\sim$1). This suggests that the top quark may play a special role in SSB.

1.1.4 The SM Lagrangian

The combination of the different elements described in the previous sections forms the SM Lagrangian:

$$\mathcal{L}_{SM} = \mathcal{L}_{QCD} + \mathcal{L}_{EW} + \mathcal{L}_{Higgs} + \mathcal{L}_{Yukawa}, \tag{1.12}$$

where $\mathcal{L}_{QCD}$ and $\mathcal{L}_{EW}$ describe the strong and the electroweak sector, respectively. $\mathcal{L}_{Higgs}$ describes the Higgs potential, which is the basis of SSB and thus provides the mass terms for the $W^{\pm}$ and Z bosons. Finally, $\mathcal{L}_{Yukawa}$ takes into account the interactions between the fermions and the Higgs field, which is the origin of the fermion masses.

1.1.5 Limitations of the Standard Model

As stated at the beginning of this section, the SM is a very successful theory which has been validated through many experimental observations. Indeed, it predicted well in advance the existence of many particles. Not only the Higgs boson, but also the top quark and the W and Z bosons, among others. It is also very predictive in terms of the interactions between the SM particles. However, it cannot be considered the ultimate theory since there are still many unknowns which cannot be explained in the SM framework. The main open questions are listed below:

- **Neutrino masses**: Neutrinos in the SM are massless. However, there are many observations which point that neutrinos are actually massive, even if their masses are very low ($\leq \mathcal{O}(0.1$ eV$)$) compared to the electroweak scale $\mathcal{O}(100$ GeV$)$ or the smallest SM fermion mass $\mathcal{O}($MeV$)$. This statement comes from the observation of neutrino oscillations, which have been observed with great precision [25]. One of the first experimental evidences of neutrino oscillations was found through the observation of solar neutrinos travelling to the Earth. The neutrinos emitted from the solar reactions are expected to be electron-like. However, the amount of electron-like neutrinos which reaches the Earth is smaller by a factor 2 to 3 [26]. The explanation of this phenomenon is that neutrinos are actually massive and their mass eigenstates do not coincide with their flavour eigenstates. This leads to the phenomenon of neutrino oscillations. The Seesaw mechanisms have been proposed beyond the SM theory to accommodate neutrino masses

and explain their smallness in a natural way. These mechanisms can accommodate either Majorana neutrinos [27–33] or Dirac neutrinos [34–37]. However, if they were Dirac particles like the rest of the fermions, this would imply right-chiral neutrinos which have not yet been observed. This is also true for most models which consider neutrinos as Majorana particles[7] [38–40]. One possibility to prove the Majorana nature of neutrinos comes from the Schechter-Valle black-box theorem [41], which states that if neutrinoless double-beta decay is present in nature then neutrinos must be Majorana particles. However, this process would be very rare and has not been observed yet [42–46].

- **Matter-antimatter asymmetry of the universe**: The conservation laws in particle physics imply that matter and antimatter is always created and annihilated in pairs. Additionally, it is generally considered that the Big Bang created the same amount of matter and antimatter. However, everything which is seen in the universe is composed of matter. There are CP violating components in the SM, as in the CKM matrix described before, but they are not sufficient to explain the observed asymmetry. Therefore, the solutions for this problem cannot be addressed within the current description of fundamental physics. The Sakharov conditions [47] describe the circumstances which need to be fulfilled by a theoretical model to explain the matter-antimmater asymmetry of the universe[8]. These include baryon number non-conservation, C and CP violation and departure from thermal equilibrium. The origin of this asymmetry must have occurred at the end or after inflation[9], a stage commonly referred to as baryogenesis [47]. There are many proposed models of baryogenesis which fulfill the Sakharov conditions. These include models which assume that the asymmetry was generated from

[7] A Majorana particle is a fermion that is its own antiparticle. Neutrinos are the only fermions in the SM which can be considered as Majorana particles since they are both charge-less and colour-less. If neutrinos were Majorana particles lepton number would no longer be conserved.

[8] There are examples of models that do not respect these conditions [48], although these are met in the great majority of cases.

[9] Inflation is a theory of exponential expansion of space in the early universe which lasted from 10^{-36} seconds after the Big Bang to some time between 10^{-33} and 10^{-32} seconds after it [49–53].

the decay of heavy particles, such as in GUT baryogenesis [51, 54–58] or leptogenesis [59] models; and models which point to a first order phase transition at the electroweak spontaneous symmetry breaking (electroweak baryogenesis) [60–64]; among others. Extensive reviews on baryogenesis models can be found in Refs. [48, 65–68].

- **Hierarchy problem**: The hierarchy problem appears when the SM is considered as a low energy effective theory that is valid up to a given cutoff. This assumption seems to be the most reasonable given the unsolved issues of the SM. The main example of hierarchy problem in the SM is found in the Higgs mass. The radiative corrections are proportional to the previously mentioned cutoff. If this were to be the Planck scale[10] ($\sim 10^{19}$ GeV), the quantum correction to the Higgs mass would be around 30 orders of magnitude larger than its effective value, which is quite unnatural. To solve this problem within the SM context, the bare parameter of the theory should be extremely fine-tuned to overcome the radiative corrections[11]. It could be that new physics enters in some place between the electroweak scale and the Planck scale, and new particles could cancel these radiative corrections. Examples of these particles arise in theories such as supersymmetry (SUSY) [69, 70], or Composite Higgs Models (CHMs) [71, 72]. However, there is still no evidence of the existence of these particles.

- **Unification of the strong interaction**: The weak and the electromagnetic interactions are unified in the SM context within the electroweak symmetry group. However, this is not the case for the strong interaction. Great Unification Theories (GUT) predict the unification of these three fundamental forces at a larger energy scale [73]. Some consequences which arise from GUT is the instability of the proton or the magnetic monopoles, but these have not yet been observed in nature.

[10] This is the energy scale at which quantum effects of gravity become strong.

[11] The mass of the Higgs boson can be expressed as $m_H^2 \simeq m_0^2 + \delta m_H^2$, where m_0 is the bare parameter of the theory and δm_H^2 is the radiative correction.

- **Strong CP violation**: The SM does not imply that CP symmetry is conserved in strong interactions. However, there is no observation of CP violation in this sector. Indeed, if there is some CP violation, it should be very low given the constraints from the measurement of the neutron dipole moment[12] [81]. Some theories which could explain CP conservation in the strong interaction imply the existence of a new particle, the axion [82], though it has not yet been observed.

- **Flavour puzzle**: The SM has a total of 19 free parameters which are determined from experimental observations. The flavour puzzle refers to the inability of the SM to explain why the free parameters of the theory have the values they have. The fermion masses and their mixing angles appear to be arbitrary: the mass of the electron is found to be $\sim$0.511 MeV, whereas the mass of the top quark is $\sim$173 GeV, a much bigger value. There are very fundamental questions involved in this puzzle, such as why are there three generations of fermions. Experimental observations have confirmed that there are exactly three light neutrino species [83], and so far there is no evidence of additional generations of either charged leptons or quarks. Indeed, a fourth generation of quarks is highly disfavoured from measurements of the Higgs properties [84]. However, there is no strong theoretical argument forbidding additional generations from existing. All this could indicate unknown principles which could lead us to a new and more complete theory.

- **Gravity**: Gravity is the only interaction which is left out of the SM formalism. It is very well understood at large scales in the context of General Relativity. Both General Relativity and the SM have been tested at very high precision and describe data extremely well. However, these two theories cannot be unified in a single theory yet.

[12] QCD includes a term that breaks CP-symmetry, the θ-term [74–78]. CP violation induced by a θ-term leads to a sizeable electric dipole moment for the neutron [79, 80]. The non-observation of such dipole moment imposes a very strict upper bound on the value of θ. This unnaturally small value of θ is known as the *strong CP problem*.

- **Dark matter and dark energy**: Astrophysical and cosmological observations point to the existence of a new kind of matter which constitutes approximately the 25% of the matter-energy content of the universe [85]. This is the so-called dark matter, which is still far from being understood. This kind of matter interacts mainly gravitationally, and thus does neither emit nor absorb light. The SM does not provide any proper candidate which could be identified with DM. More details are given in Section 1.4. In addition to the DM, there is the dark energy which constitutes approximately 70% of the universe matter-energy content. This amount of dark energy is necessary to explain the accelerating expansion of the universe, but there is very little knowledge about it.

1.2 The top quark: an ideal realm for new physics searches

The top quark is a fundamental particle in the SM. It forms the third generation of quarks together with the bottom quark. It is positively charged, with an electric charge of $+\frac{2}{3}$ in electron units, and it is the most massive particle in the SM, with a mass of $m_t = 172.76 \pm 0.30$ GeV[13] [12].

The third generation of quarks was postulated by Kobayashi and Maskawa in 1973 [89]. This prediction was based on the principles of the GIM mechanism [16]. The GIM mechanism predicted the existence of the charm quark and prevented FCNC in the weak interactions by imposing unitarity of the mixing matrix between the first two generations of quarks. If this mixing matrix was generalised to three generations of quarks, the parameters of this matrix would include a complex phase, which could explain the CP violation observed in Kaon decays [90]. The bottom quark was soon discovered by the E288 experiment at Fermilab in 1977 [91]. However, due to its large mass, it was not possible to discover the top quark until much later. Measuring the W and Z boson masses allowed to constrain the

[13]The most precise determination of the top-quark mass as provided by ATLAS (CMS) is $m_t = 172.69 \pm 0.48$ GeV [86] ($m_t = 172.44 \pm 0.48$ GeV [87]). The world average provided by the LHC top working group is $m_t = 173.34 \pm 0.76$ GeV [88].

top-quark mass [92–95], but it took more than 20 years from its prediction to being discovered. It was finally observed in 1995 by the CDF and D0 experiments at the Tevatron in Fermilab [96, 97].

The top quark is a crucial particle in the SM context. Due to its large mass, it is the only quark that decays before hadronising. This is a unique feature among the quarks which allows to study the properties of this particle through its decay products. Furthermore, it decays almost exclusively to a W boson and a bottom quark ($|V_{tb}| \sim 1$, see Table 1.3), allowing precise measurements of the corresponding CKM element and providing characteristic signatures to study the top quark. This gives access to the top quark properties, such as the electric charge and the polarisation, that can both be determined precisely and efficiently from the properties of the charged lepton formed in the W-boson decay. Another consequence of its large mass is found in its Yukawa coupling. Indeed, it is the particle with the largest Yukawa coupling with the Higgs boson, and thus it is expected to play a special role in SSB.

The top quark contributes dominantly in radiative corrections to electroweak processes, allowing for very stringent tests of the SM to be performed. The top-quark mass is also an important input for the studies of the vacuum stability. If the SM was a valid theory up to the Planck scale ($\sim 10^{19}$ GeV), then the SM vacuum should be stable. This is true at tree level since the Higgs potential shows a unique global minimum. However, if one includes radiative corrections, the shape of the Higgs potential is severely affected and this statement is not necessarily valid [98]. The stability of the SM vacuum is approximately equivalent to the question whether λ (see Equation 1.9) stays positive up to the Planck scale under radiative corrections. These have been calculated at next-to-next-to-leading order (NNLO) and depend upon the free parameters of the theory [99]. Figure 1.3 shows the state-of-the-art results from the SM vacuum stability tests in the top mass-Higgs mass plane. The stability region is reached if the SM minimum is the global minimum of the Higgs potential. In case it is a local minimum the SM vacuum is not stable and it can tunnel into an energetically favoured state. As one can see, the stability region is disfavoured by present data by 2.5σ, which could be interpreted as a hint of new physics that would enter below the Planck scale. It

is however not possible to exclude that we live in a metastable universe, i.e. the lifetime of the (local) SM minimum could be longer than the age of the universe. The biggest source of uncertainty in the determination of the radiative corrections which affect the Higgs potential comes from the experimental measurement of the top-quark mass.

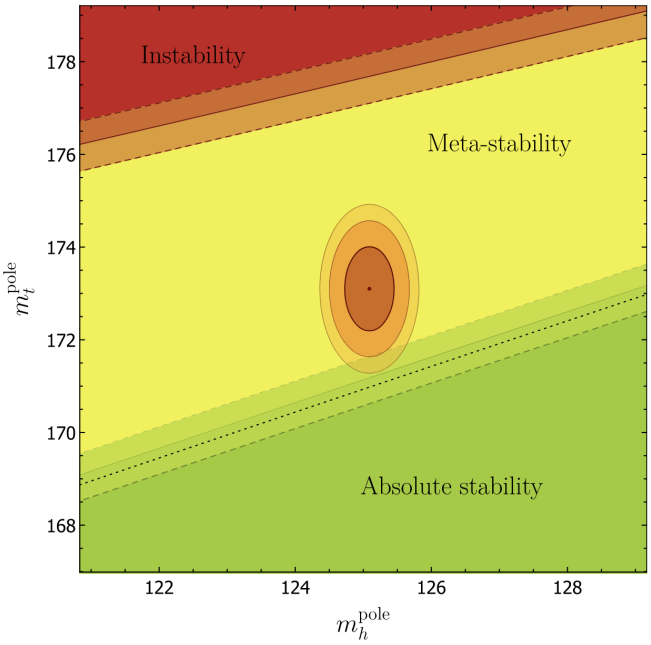


Figure 1.3: Regions of absolute stability, metastability and instability of the SM vacuum in the top-mass–Higgs-mass plane. The ellipses denote the allowed regions at 1, 2 and 3σ from experimental constraints. The shaded bands on the phase boundaries, framed by the dashed lines and centred on the solid lines, are combinations of the α_s experimental uncertainty and the theory uncertainty. The calculations are made assuming a top-quark mass $m_t^{pole} = 173.1 \pm 0.6$ GeV [100].

The top quark is also crucial in new physics searches, since many models beyond the SM predict the existence of new particles which would preferentially couple to top quarks [9]. In addition, it is a very important background in many new physics searches at the LHC.

All these arguments motivate an exhaustive study of the top-quark properties and the comparison of the measurements with the SM predictions. It also motivates the search for new particles which are associated to the top quark. The work done in this thesis covers both aspects, performing a precise measurement of normalised differential cross-sections sensitive to the top-quark polarisation and a direct search for DM particles produced in association with top quarks.

If new physics were associated with the top-quark sector, the LHC would be the ideal place to find it since it is a top factory, producing more than 273 million top quarks (42 million singly-produced) that are recorded by the ATLAS detector during Run-2[14]. This ensures very high statistics and enables high precision measurements of the top-quark properties.

1.2.1 Top-quark production at the LHC

In pp collisions at the LHC top quarks can be produced both in top quark-antiquark pairs in processes mediated by the strong interaction or singly in processes mediated by the electroweak interaction.

Top-quark pair production is the main production mechanism of top quarks at the LHC, with a production cross-section approximately three times larger than single production. There are two processes which contribute to top-quark pair production at leading order (LO) in QCD, $gg \to t\bar{t}$ and $q\bar{q} \to t\bar{t}$. The corresponding Feynman diagrams are depicted in Figure 1.4. At next-to-leading-order (NLO) there are additional processes with gq ($g\bar{q}$) in the initial state which also contribute to top-quark pair production. At the LHC with $\sqrt{s} = 13$ TeV, approximately 90% of the total top quark pair production cross-section is given by the gg initial state, and the rest is mostly given by $q\bar{q}$ in the initial state [12]. The $t\bar{t}$ production cross-section is computed theoretically at NNLO in QCD including resummation of next-to-next-to-leading logarithmic (NNLL) soft-gluon terms with Top++2.0 [101–106] and assuming a top quark mass of $m_t = 172.5$ GeV. The predicted cross-sections are listed in Table 1.4 for different centre-of-mass energies.

[14] This is the period in which the LHC provided pp collisions at a centre-of-mass energy of 13 TeV. See Section 2.1.

$\sqrt{s}$ (TeV)	$\sigma_{t\bar{t}}$ (pb)
7	177^{+5+9}_{-6-9}
8	253^{+6+12}_{-9-12}
13	832^{+20+35}_{-29-35}
14	980^{+20+40}_{-30-40}

Table 1.4: Theoretical predictions for the top-quark pair production cross-section for the different operating centre-of-mass energies at the LHC. The first uncertainty is from scale dependence and the second from parton distribution functions [101–106].

There are other rare top-quark pair production mechanisms which include the production of a $t\bar{t}$ pair in association with a W, Z or Higgs boson, or a photon.

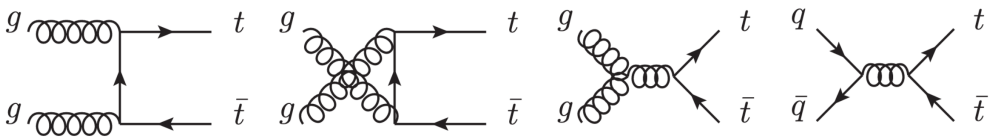


Figure 1.4: Feynman diagrams for $t\bar{t}$ production at LO. The first three diagrams correspond to the gg initial state, whereas the last diagram corresponds to the $q\bar{q}$ initial state.

Single-top-quark production has also a sizeable signal at the LHC. There are three main production mechanisms at LO which are categorised through the characteristics of the exchanged W boson. These include the t-channel, in which the W boson is space-like, the s-channel, in which the W boson is time-like, and the production of a top quark in association with an on-shell[15] W boson (tW-channel). The Feynman diagrams associated with these processes are depicted in Figure 1.5. The predicted single-top-quark production cross-sections for the different contributing processes calculated at NLO in QCD are listed in Table 1.5 for different centre-of-mass energies, assuming a top quark mass of $m_t = 172.5$ GeV.

[15]The term on-shell refers to real particles whose associated fields obey the equations of motion. In contrast, the term off-shell refers to virtual particles whose associated fields do not necessarily satisfy these relations.

$\sqrt{s}$ (TeV)	$\sigma_{t+\bar{t}}$ (pb)		
	t-channel	tW-channel	s-channel
7	$63.9^{+1.9+2.2}_{-1.3-2.2}$	$15.7^{+0.4+1.1}_{-0.4-1.1}$	$4.29^{+0.12+0.14}_{-0.10-0.14}$
8	$84.7^{+2.6+2.8}_{-1.7-2.8}$	$22.4^{+0.6+1.4}_{-0.6-1.4}$	$5.24^{+0.15+0.16}_{-0.12-0.16}$
13	217^{+7+6}_{-5-6}	$71.7^{+1.8+3.4}_{-1.8-3.4}$	$10.32^{+0.29+0.27}_{-0.24-0.27}$
14	248^{+8+7}_{-5-7}	84^{+2+3}_{-2-5}	$11.39^{+0.32+0.29}_{-0.26-0.29}$

Table 1.5: Theoretical predictions for the single-top-quark production cross-section for the different operating centre-of-mass energies at the LHC and including both top quarks and antiquarks. The first uncertainty is from scale dependence and the second from parton distribution functions [107–109].

These are calculated using HATHOR [107, 108] for the t-channel and the s-channel. For the tW-channel, the calculation includes NNLL soft-gluon corrections [109]. Other rare single-top-quark production modes include tZq, $t\gamma q$, tWZ and tHq.

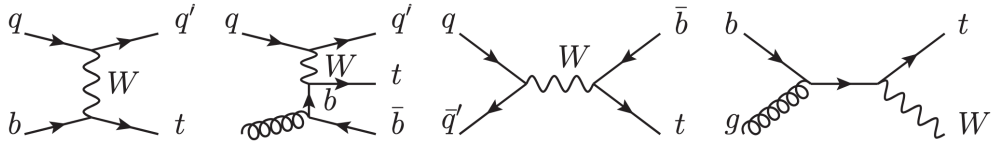
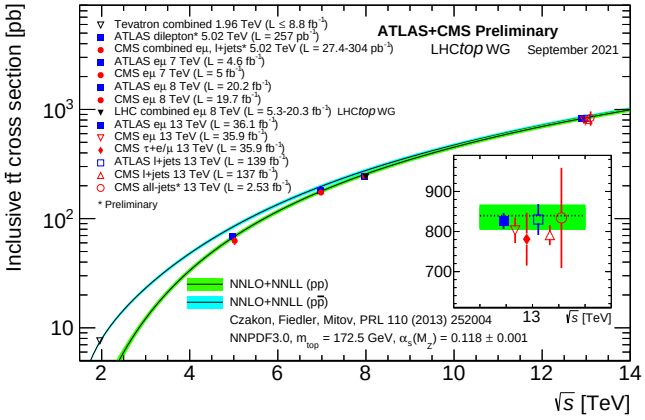


Figure 1.5: Feynman diagrams for single-top-quark production at LO. The first two correspond to the t-channel production, the next diagram corresponds to the s-channel and the last diagram corresponds to the production of a top quark in association with a W boson.

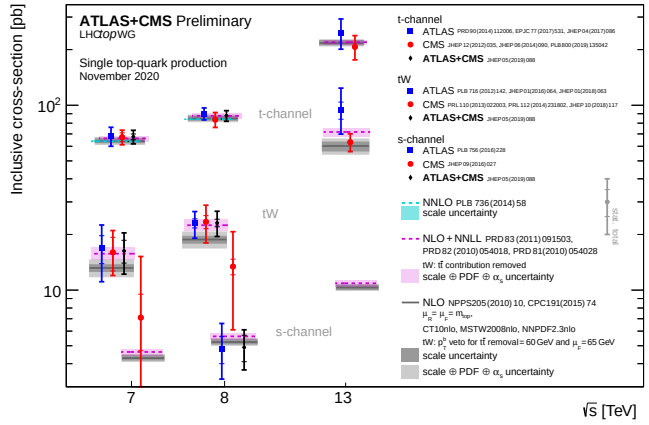
Figure 1.6 summarises the measured cross-sections by both ATLAS and CMS experiments for top-quark pair production and single-top-quark production at different centre-of-mass energies compared with theoretical predictions. There is a good agreement between the expected and the measured values.

1.2.2 Top quark decay

As stated before, the top quark decays almost exclusively to a W boson and a bottom quark. The top quark decay can be further categorised according to the



(a)



(b)

Figure 1.6: Top: Summary of LHC and Tevatron measurements of the top-pair production cross-section as a function of the centre-of-mass energy compared to the NNLO QCD calculation complemented with NNLL resummation as provided by Top++2.0 [102]. The theory band represents uncertainties due to renormalisation and factorisation scale, parton density functions and the strong coupling. The measurements and the theory calculation are quoted at m_t = 172.5 GeV. Measurements made at the same centre-of-mass energy are slightly offset for clarity. Bottom: Summary of ATLAS and CMS measurements of the single-top-quark production cross-sections in various channels as a function of the centre-of-mass energy. The measurements are compared to theoretical calculations based on: NLO QCD, NLO QCD complemented with NNLL resummation and NNLO QCD (t-channel only) [110].

decay of the W boson. This includes two different channels; the leptonic channel, in which the W boson decays to a lepton and a neutrino, and the hadronic channel, in which the W boson decays to a pair of quarks. The branching ratios for these decays are 32.6% and 67.4%, respectively [12]. In the case of $t\bar{t}$ production, this categorisation leads to three different final states: the dilepton final state, in which the two W bosons decay leptonically, the lepton+jets final state, in which one W boson decays hadronically and the other decays leptonically, and the full hadronic final state, in which both W bosons decay hadronically. The branching ratios for these processes are 10.5%, 43.8% and 45.7%, respectively[16].

1.3 Top-quark polarisation in single-top t-channel

The single-top t-channel production is the dominant process that contributes to single-top-quark production at the LHC. It is characterised by the interaction of a light quark q and a b-quark, exchanging a W boson, and producing a top quark and another light quark q' which is usually named the *spectator quark* ($qb \to tq'$). In t-channel production, one can consider that the incoming b-quark arises directly from the sea quarks in the proton in the so-called five flavour scheme (5FS) ($2 \to 2$ process) or that it arises from gluon splitting in the so-called four flavour scheme (4FS)[17] ($2 \to 3$ process). The corresponding Feynman diagrams are depicted in Figure 1.7, including the leptonic decay of the top quark. As it can be seen, the 4FS implies an extra b-quark in the final state. However, this additional b-quark is characterised by its soft transverse momentum[18] spectrum, being most of the times outside the kinematic acceptance.

The production of top quarks and antiquarks is not predicted to be equivalent. The different subprocesses that contribute to top-quark and antiquark productions

[16] In many analyses, leptonic decays of the top quark where the W boson decays to a τ lepton and its associated neutrino are considered within the leptonic channel only when the τ lepton further decays leptonically, which occurs 35.21% of the time.

[17] In the four flavour scheme, only the quarks u, d, s, c are considered inside the proton structure. The initial state b-quark arises from gluon splitting into nearly collinear $b\bar{b}$. In the five flavour scheme, b quarks inside the proton are also considered.

[18] The transverse momentum is defined in the plane perpendicular to the beam direction. See Section 2.3.1.

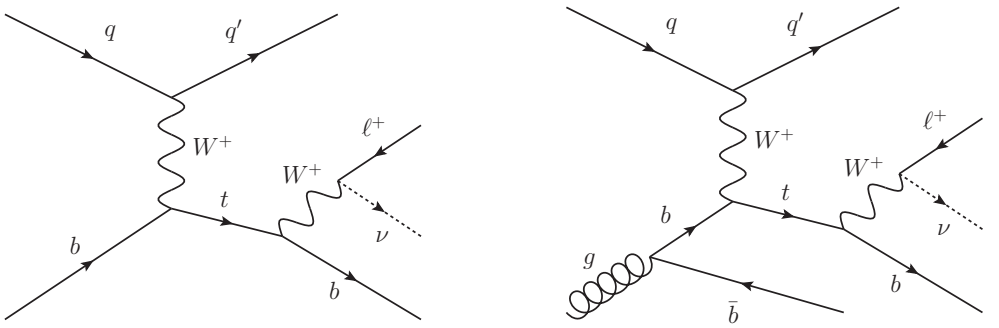


Figure 1.7: LO Feynman diagrams for t-channel single-top-quark production and leptonic decay in (a) the 5FS and (b) the 4FS.

in the t-channel are displayed in Figure 1.8. For the top quark, the dominant process implies a u quark in the initial state, whereas for the top antiquark a d quark is expected in the initial state. Given the predominance of u-type valence quarks in the proton, the cross-section of top-quark production in the t-channel is $\sim$2 times larger than the production of top antiquarks.

As it is shown in the two diagrams of Figure 1.7, the tWb vertex appears both in the production and the decay vertex of the top quark. In the SM, the interaction in the tWb vertex is given by:

$$\mathcal{L}_{tWb} = -\frac{g}{\sqrt{2}}\bar{b}\gamma^{\mu}V_{tb}P_L tW^{-}_{\mu} + h.c., \tag{1.13}$$

where g is the weak coupling constant, $\bar{b}$, t and W are the corresponding fermionic and bosonic fields, γ^{μ} are the gamma matrices, V_{tb} is the corresponding element of the CKM matrix and $P_L = \frac{1}{2}(1-\gamma_5)$ is the left-chiral operator.

The vector-axial (V-A) structure of the tWb vertex implies that the top quarks are produced polarised. This means that their spin is aligned in a particular direction, which in the t-channel corresponds to the direction of the down-type quarks [111, 112]. This is characteristic of single-top-quark production. In top quark pair strong production, due to parity conservation in QCD [113], top quarks are produced unpolarised.

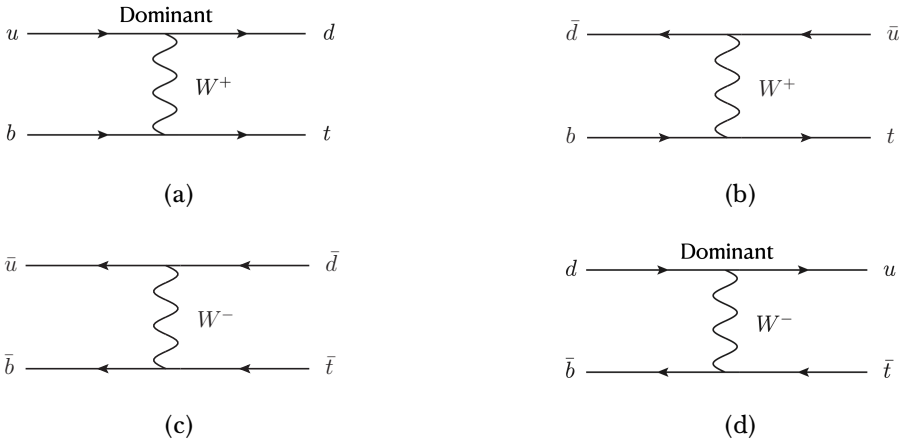


Figure 1.8: Subprocesses contributing to the t-channel production at LO in the 5FS. (a) and (b) correspond to top-quark production, with (a) being the dominant process. (c) and (d) correspond to top-antiquark production, with (d) being the dominant process.

As it can be seen in Figure 1.8, the down-type quarks can be either in the initial or in the final state, depending on the process considered. For the top quarks, in the dominant process the down-type quark is the spectator quark, whereas in the subdominant process this is the initial state down-type quark. Nevertheless, the direction of the spectator quark is expected to be very close to the beam direction. The proximity between these directions produces a very high top-quark (antiquark) polarisation along (opposite to) the direction of the spectator quark.

Angular observables can be defined that are sensitive to the top-quark polarisation. These are defined exploiting the $t \to Wb \to bl^+\nu$ decay mode of the top quark, as well as the charge-conjugate decay mode of the top antiquark. The lepton $l^\pm$ can be either an electron or a muon[19]. In the decay, three orthogonal directions may be defined [2]. As illustrated in Figure 1.9, the $\hat{z}'$ direction is the direction of the momentum of the spectator quark, $\vec{p}_{q'}$, in the top-quark reference frame. The $\hat{y}'$ direction is taken along $\hat{z}' \times \hat{p}_q$, where $\hat{p}_q$ is the direction of the

[19]Events in which the W boson decays into a τ lepton and a neutrino are also considered if the τ lepton subsequently decays to a muon or an electron and an additional neutrino.

incoming light quark[20], in the top-quark reference frame. Finally, the $\hat{x}'$ direction lies in the plane of production, orthogonal to $\hat{y}'$ and $\hat{z}'$, such that (x', y', z') form a right-handed coordinate system:

$$\hat{z}' = \frac{\vec{p}_{q'}}{|\vec{p}_{q'}|}, \qquad \hat{y}' = \frac{\hat{z}' \times \vec{p}_q}{|\hat{z}' \times \vec{p}_q|}, \qquad \hat{x}' = \hat{y}' \times \hat{z}' \tag{1.14}$$

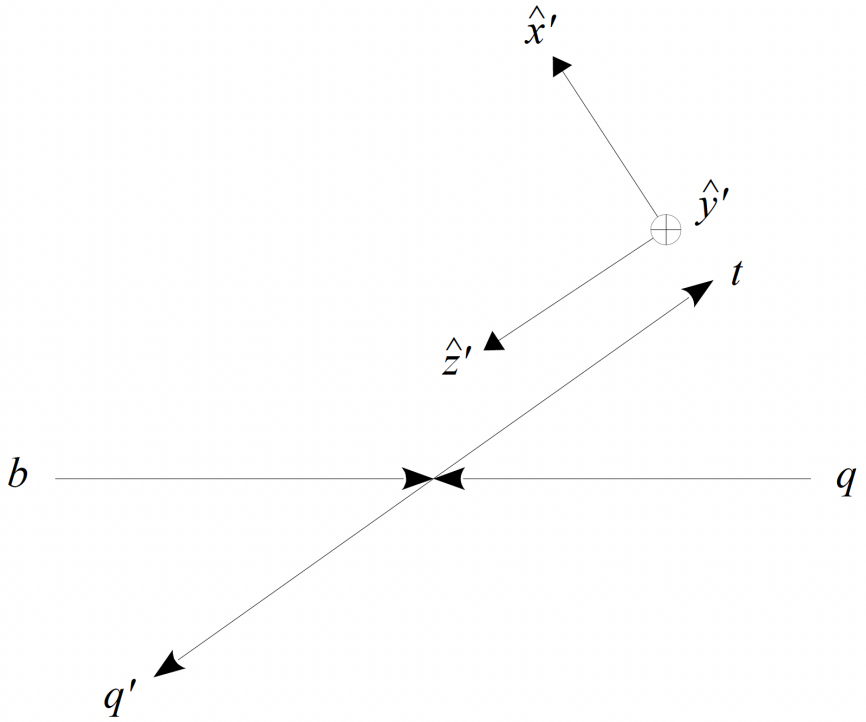


Figure 1.9: Diagram illustrating the three orthogonal directions $\hat{x}'$, $\hat{y}'$ and $\hat{z}'$ used in this analysis, as seen in the zero-momentum frame. The $\hat{z}'$ direction is that of the spectator quark in the top-quark rest frame. The $\hat{x}'$ direction lies in the plane of production, while the $\hat{y}'$ direction is perpendicular to the plane of production.

[20]The direction of the incoming quark cannot be unambiguously determined. To select among the directions of the two beams one can use the momentum of the spectator quark in the laboratory frame, which most of the times follows that of the incoming quark. This choice gives the correct answer 97% (98%) of the time for the dominant process contributing to single-top-quark(antiquark) production in the t-channel [2].

The polarisation vector $\vec{P}$ is defined in this coordinate system; it satisfies $|\vec{P}| \leq 1$, equality holding if the top quarks are produced in a pure quantum mechanical ensemble with respect to spin.

The polarisation vectors are not expected to be the same for top quarks and antiquarks due to the differences in the t-channel production explained before. Concretely, in the SM, the expected values at LO of the polarisation vectors are $\vec{P}^{t} = (0, 0, +0.90)$ and $\vec{P}^{\bar{t}} = (-0.14, 0, -0.86)$ for top quarks and antiquarks, respectively [2]. Acceptance requirements however, which have a large effect on the polarisation, are not accounted for in that calculation. Notice that P_x is not zero for top antiquarks because the spectator quark is not the down-quark in the dominant process. A calculation at NNLO, based on Ref. [114], predicts top-quark polarisation along the direction of the spectator quark of $0.965 \pm 0.003\,(\text{scale})^{+0.003}_{-0.004}\,(\text{PDF} + \alpha_s)$ for top quarks and $-0.957^{-0.003}_{+0.012}\,(\text{scale})^{+0.004}_{-0.002}\,(\text{PDF} + \alpha_s)$ for top antiquarks. This calculation considers stable single top quarks produced in the t-channel from pp collisions at $\sqrt{s} = 13$ TeV, in a region with exactly one light-flavour jet which corresponds to the spectator jet.

Due to its large mass [88], the lifetime of the top quark $\mathcal{O}(10^{-25}\text{s})$ is much shorter than the hadronisation time-scale $\mathcal{O}(10^{-24}\text{s})$. Furthermore, the lifetime of the top quark is much shorter than the depolarisation time scale $\mathcal{O}(10^{-21}\text{s})$[21] [115]. Since the W boson from the top-quark decay is produced on-shell, this implies that the top-quark spin information is transferred to its decay products.

The differential decay rates in the top-quark reference system are given by:

$$\frac{1}{\Gamma}\frac{d\Gamma}{d\cos\theta_{Xi}} = \frac{1}{2}(1 + \alpha_X P_i \cos\theta_{Xi}), \tag{1.15}$$

where α_X is the spin analysing power of the chosen decay product from the top quark ($X = W, l, \nu, b$ in the leptonic channel), θ_{Xi} is the angle between the direction of motion of the top-quark decaying particle X and the axis $i = (\hat{x}', \hat{y}', \hat{z}')$ and P_i is the top-quark polarisation component in a given axis i.

[21] This is the time necessary so that QCD radiation effects affect the spin information.

Spin analyser X	α_X (LO)	α_X (NLO)
l^+	1	0.998
W^+	0.41	0.39
b-quark	-0.41	-0.39
ν	-0.32	-0.33

Table 1.6: Spin analysing power of the top-quark decaying products in the lepton channel both at LO and NLO [116]. In the top-antiquark case, the spin analysing power of the corresponding antiparticles only needs to flip the sign.

The spin analysing power of the top-quark decaying particles in the leptonic channel are listed in Table 1.6, both at LO and NLO. As it can be seen, the charged lepton is the most powerful spin analyser of the top-quark spin, with an analysing power very close to unity at NLO. For this reason, the normalised differential cross-sections measured in this thesis are provided as a function of the three angular distributions $\cos\theta_{li}$ of the charged lepton l associated to the three different polarisation components. In Figure 1.10 these polarisation angles are depicted with respect to a particular axis $i = (\hat{x'}, \hat{y'}, \hat{z'})$, using the reference frame previously introduced.

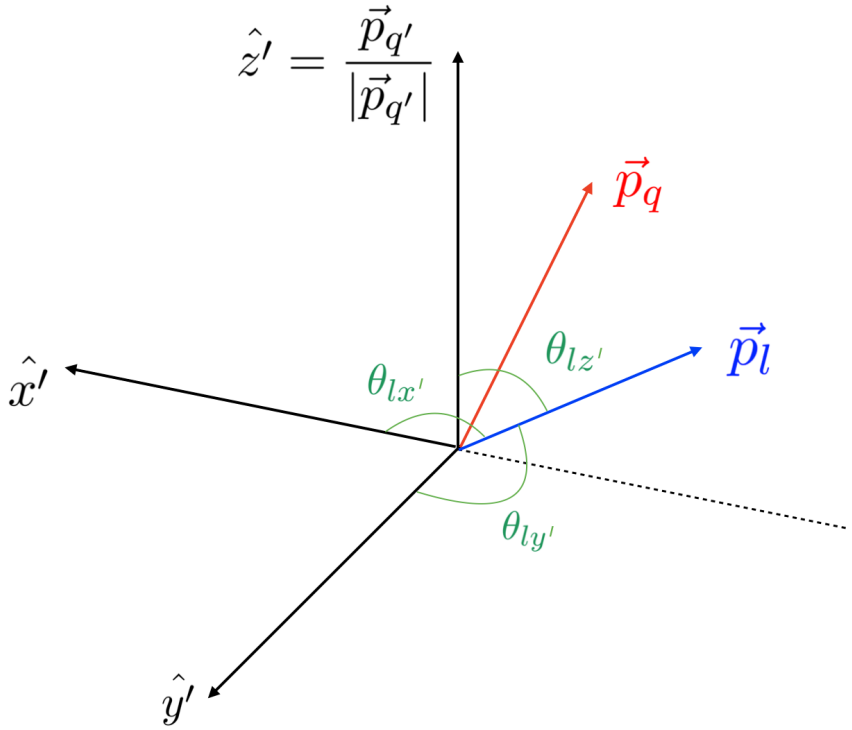


Figure 1.10: Right-handed coordinate system illustrating the three angles of the lepton from the top-quark decay in the top-quark reference frame. The $\hat{z}'$ axis points in the direction of motion of the spectator quark which most of the times corresponds to the top-quark spin direction. The $\hat{y}'$ axis is taken perpendicular to the $\hat{z}'$ axis and the momentum of the initial quark $\vec{p}_q$ (red), and $\hat{x}'$ is perpendicular to the other two axis. The momentum of the charged lepton is represented in blue. The three polarisation angles between the momentum of the charged lepton and the three axes, which are intimately related with the three top-quark polarisation components, are also displayed (green).

1.3.1 Effective field theory in the tWb vertex

Measurements of polarisation observables in the t-channel single-top-quark production are sensitive to new physics phenomena affecting the tWb vertex. EFT is a model independent way to describe possible deviations from the SM. New physics can enter at the Planck scale ($\sim 10^{19}$ GeV), but they could also appear at

lower energies. In that sense, Λ is the energy scale at which new physics enters to alter the high-energy behavior of the SM theory. If Λ is beyond the LHC reach, then new physics effects can be parametrised in terms of an effective Lagrangian [117]. In this context, the SM Lagrangian ($\mathcal{L}_{\text{SM}}$) is augmented with higher-dimensional operators invariant under the SM gauge symmetry. The sum can be truncated at dimension six ($\mathcal{O}_k^{[6]}$), with corresponding Wilson coefficients (C_k) scaled with $1/\Lambda^2$, which is the first term that gives non-zero contributions under the assumption of lepton and baryon number conservation [117]:

$$\mathcal{L}_{\text{EFT}} = \mathcal{L}_{\text{SM}} + \sum_k \frac{C_k}{\Lambda^2}\mathcal{O}_k^{[6]} + ..., \tag{1.16}$$

where $\mathcal{L}_{\text{EFT}}$ is the effective Lagrangian, Λ is the scale of new physics chosen such that higher-dimension operators are sufficiently suppressed by higher powers of Λ, and k runs over all dimension-six operators.

Top-quark physics at the LHC has entered a phase of precision predictions and measurements. A key motivation for a precise scrutiny of the top quark is that its production and decay mechanisms should be especially sensitive to effects of physics beyond the SM. In an EFT context, these measurements can be interpreted to set strong bounds on the phase space in which new physics may hide.

Table 1.7 shows the sensitivity of each of the LHC top processes to the 34 Wilson coefficients that explicitly modify the couplings of the top quark with the other SM fields[22] [118]. Here, the Minimal Flavour Violation (MFV) hypothesis[23] [119] in the quark sector is adopted. In addition, dimension seven or higher operators are neglected, and only CP-conserving extensions of the SM are considered. Since top-quark observables at the LHC are largely blind to the flavour of light quarks with the same quantum numbers, $U(2)_q \times U(2)_u \times U(2)_d$ flavour symmetry is imposed among quarks of the first and second generation [120–122].

There are only three dimension-six effective operators that contribute at order $1/\Lambda^2$ at tree level in the t-channel production of the single top quark as presented

[22] The relations between these operators and those in the Warsaw basis can be found in Ref. [118] (Table 2.1).

[23] Minimal flavour violation requires that all flavour and CP-violating interactions are linked to the known structure of the SM Yukawa couplings.

Parameter	$t\bar{t}$	Single t	tW	tZ	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}t\bar{t}$	$t\bar{t}b\bar{b}$
C_{QQ}^{1}	-	-	-	-	-	-	-	✓	✓
C_{QQ}^{8}	-	-	-	-	-	-	-	✓	✓
C_{Qt}^{1}	-	-	-	-	-	-	-	✓	✓
C_{Qt}^{8}	-	-	-	-	-	-	-	✓	✓
C_{Qb}^{1}	-	-	-	-	-	-	-	-	✓
C_{Qb}^{8}	-	-	-	-	-	-	-	-	✓
C_{tt}^{1}	-	-	-	-	-	-	-	✓	-
C_{tb}^{1}	-	-	-	-	-	-	-	-	✓
C_{tb}^{8}	-	-	-	-	-	-	-	-	✓
C_{QtQb}^{1}	-	-	-	-	-	-	-	-	(✓)
C_{QtQb}^{8}	-	-	-	-	-	-	-	-	(✓)
$C_{Qq}^{1,8}$	✓	-	-	-	✓	✓	✓	✓	✓
$C_{Qq}^{1,1}$	[✓]	-	-	-	[✓]	[✓]	[✓]	✓	✓
$C_{Qq}^{3,8}$	✓	[✓]	-	[✓]	✓	✓	✓	✓	✓
$C_{Qq}^{3,1}$	[✓]	✓	-	✓	[✓]	[✓]	[✓]	✓	✓
C_{tq}^{8}	✓	-	-	-	✓	✓	✓	✓	✓
C_{tq}^{1}	[✓]	-	-	-	[✓]	[✓]	[✓]	✓	✓
C_{tu}^{8}	✓	-	-	-	-	✓	✓	✓	✓
C_{tu}^{1}	[✓]	-	-	-	-	[✓]	[✓]	✓	✓
C_{Qu}^{8}	✓	-	-	-	-	✓	✓	✓	✓
C_{Qu}^{1}	[✓]	-	-	-	-	[✓]	[✓]	✓	✓
C_{td}^{8}	✓	-	-	-	-	✓	✓	✓	✓
C_{td}^{1}	[✓]	-	-	-	-	[✓]	[✓]	✓	✓
C_{Qd}^{8}	✓	-	-	-	-	✓	✓	✓	✓
C_{Qd}^{1}	[✓]	-	-	-	-	[✓]	[✓]	✓	✓
C_{tG}	✓	-	✓	-	✓	✓	✓	✓	✓
C_{tW}	-	✓	✓	✓	-	-	-	-	-
C_{bW}	-	(✓)	(✓)	(✓)	-	-	-	-	-
C_{tZ}	-	-	-	✓	-	✓	-	-	-
$C_{\varphi tb}$	-	(✓)	(✓)	(✓)	-	-	-	-	-
$C_{\varphi Q}^{3}$	-	✓	✓	✓	-	-	-	-	-
$C_{\varphi Q}^{-}$	-	-	-	✓	-	✓	-	-	-
$C_{\varphi t}$	-	-	-	✓	-	✓	-	-	-
$C_{t\varphi}$	-	-	-	-	-	-	✓	-	-

Table 1.7: Wilson coefficients and their contributions to top-quark observables. A check mark outside (inside) brackets indicates that a given process constrains the corresponding operator at $\mathcal{O}(\Lambda^{-2})$ ($\mathcal{O}(\Lambda^{-4})$), and operators in square brackets indicate that an operator contributes at $\mathcal{O}(\Lambda^{-2})$ but only at NLO in QCD. The MFV hypothesis is adopted. Dimension seven and higher operators are neglected, and only CP-conserving extensions of the SM are considered. $U(2)_q \times U(2)_u \times U(2)_d$ flavour symmetry is imposed among quarks of the first and second generation. Single t stands for t and s-channel electroweak production [118].

in Ref. [123]. In the so-called "Warsaw" basis, the three operators are $\mathcal{O}_{\varphi q}$, $\mathcal{O}_{qq}$ and the dipole operator $\mathcal{O}_{tW}$ as described in Ref. [124], where φ stands for the Higgs field. Other effective operators such as $\mathcal{O}_{bW}$ and $\mathcal{O}_{\varphi tb}$ contribute but are suppressed by a factor $1/\Lambda^4$ or by the small value of the bottom-quark mass relative to that of the top quark.

The operator $\mathcal{O}_{\varphi q}$ affects the signal production cross-section and has no effect on normalised distributions. The four-fermion operator, $\mathcal{O}_{qq}$, has a negligible effect on angular distributions [123] and can be ignored. Only the $\mathcal{O}_{tW}$ operator with its complex coefficient has an effect on the polarisation of the top quark. The real and imaginary part of its coefficient are indicated by C_{tW} and C_{itW}, respectively. The coefficient C_{tW} mostly affects $P_{x'}$, whereas C_{itW} affects $P_{y'}$. Since the SM prediction for $P_{z'}$ is already near the maximum allowed value, it is less sensitive to changes in C_{tW} or C_{itW}. A non-zero value for C_{itW} could be a hint of CP violation in the tWb vertex, making it specifically interesting with respect to the matter-antimatter asymmetry present in nature.

Alternatively, the dimension-six operators can also be parametrised in the tWb vertex in terms of the so-called *anomalous couplings* [125]:

$$\mathcal{L}_{tWb} = \frac{g}{\sqrt{2}}\bar{b}\gamma^{\mu}(V_L P_L + V_R P_R)tW^-_{\mu} + \frac{g}{\sqrt{2}}\bar{b}\frac{i\sigma^{\mu\nu}q_{\nu}}{M_W}(g_L P_L + g_R P_R)tW^-_{\mu} + h.c., \tag{1.17}$$

where g and P_L are the weak coupling constant and the left-chiral operator already introduced, $P_R = \frac{1}{2}(1+\gamma_5)$ is the right-chiral operator, m_W and q_{ν} are the mass and the four-momentum of the W boson, and $\sigma^{\mu\nu} = \frac{[\gamma^{\mu},\gamma^{\nu}]}{2}$. $V_{L,R}$ and $g_{L,R}$ are the left- and right-handed vector and tensor couplings, respectively. In the SM at tree level, V_L corresponds to the CKM matrix element V_{tb}, which is very close to one (see Table 1.3). The other three couplings are zero in the SM, and they are referred to as *anomalous couplings*.

These anomalous couplings are related one-to-one with the previously introduced Wilson coefficients [125]:

$$V_L = V_{tb} + \frac{v^2}{\Lambda^2} C_{\varphi q} \qquad V_R = \frac{1}{2}\frac{v^2}{\Lambda^2} C_{\varphi tb} \qquad g_L = \sqrt{2}\frac{v^2}{\Lambda^2} C_{bW} \qquad g_R = \sqrt{2}\frac{v^2}{\Lambda^2} C_{tW}, \tag{1.18}$$

where v = 246 GeV is the Higgs vacuum expectation value. From this notation it can be seen that new physics effects in the V_L coupling would affect just the normalisation (since it is related with $\mathcal{O}_{\varphi q}$). Additionally, V_R and g_L are highly suppressed since $\mathcal{O}_{\varphi tb}$ and $\mathcal{O}_{bW}$ involve right-handed b-quarks. Consequently, the normalised differential cross-sections measured in this thesis would be mostly affected by non-zero values of the g_R anomalous coupling which is related to the $\mathcal{O}_{tW}$ dipole operator.

1.3.2 Scope of this thesis: Measurement of normalised differential cross-sections as a function of polarisation sensitive angular variables

In this thesis the measurement of the normalised differential cross-section of t-channel single-top-quark production, where the W boson from the top quark decays leptonically, is presented as a function of three angular distributions of the charged lepton ($\cos\theta_{lx'}, \cos\theta_{ly'}, \cos\theta_{lz'}$) (see Figure 1.10) which are associated with the three top-quark polarisation components ($P_{x'}$, $P_{y'}$, $P_{z'}$). The measured angular distributions are distorted by detector inefficiencies and acceptance, smeared by reconstruction effects such as detector resolution and sculpted by the event selection effects. To correct for these detector effects, the provided angular distributions are unfolded to the particle level[24], so that the measurements can be compared directly with theoretical predictions within the fiducial reconstructed region[25] described in Section 4.6.2. These results have been published in Ref. [3].

[24] The unfolding technique can be applied at two different levels: particle level and parton level. Particle-level objects are constructed from stable particles with a lifetime > 300 ps within the observable pseudorapidity range (see Section 4.6). In the parton level, quarks and gluons are considered as individual components before they interact through QCD interactions and become embedded within the particles.

[25] A fiducial region is a phase space at particle level (or parton level) that is designed to correspond reasonably closely to what is experimentally accessible by the detector and reconstruction/identification algorithms.

Previous analyses [11, 126–130] have measured spin observables in single-top-quark and/or $t\bar{t}$ events using LHC data. In particular, a previous measurement of t-channel single-top-quark polarisation at $\sqrt{s} = 8$ TeV from ATLAS [129] set a limit $|P_{z'}| > 0.72$ (at 95% confidence level (C.L.)). The spin asymmetry $A_l = (P_{z'}\alpha_l)/2$ was determined to be 0.49 ± 0.06 by ATLAS [128] at $\sqrt{s} = 8$ TeV and measured to be 0.44 ± 0.07 by CMS at $\sqrt{s} = 13$ TeV [130]. Also, the three polarisation components were extracted from the measurement of angular asymmetries at $\sqrt{s} = 8$ TeV as a part of the PhD thesis in Ref. [11]. The measurement was in all cases provided inclusively for top quarks and antiquarks.

An important difference with respect to this analysis, is that previously the unfolding was performed to the parton level in the full phase space, i.e. without imposing additional requirements on the parton-level objects. This allows a direct determination of the top-quark polarisation components. Nonetheless, it was proven that new physics effects would alter the unfolding corrections [131]. This implies that the measurements can depend on the model assumed to calculate the unfolding corrections (in general, the SM). A possibility to overcome this problem is to unfold the angular distributions to the particle level in the reconstructed fiducial region instead. This approach has already been used by the ATLAS collaboration in the measurement of the single-top t-channel cross-section [132] and was also explored in Ref. [133] for polarisation related observables. The main advantage of unfolding to the particle level is that no extrapolations for the transition from partons to particles are needed. Furthermore, if one uses a fiducial region similar to the reconstructed level region as it is done in this analysis, detector acceptance corrections are minimum. Thus, the provided measurements are much more independent of any assumption on the modelling and the underlying physics.

This analysis is based on data from pp collisions at a centre-of-mass energy of $\sqrt{s} = 13$ TeV using the full Run-2 dataset, corresponding to data recorded between 2015 and 2018 and a total integrated luminosity of 139 fb^{-1}. These measurements are provided for top-quark and top-antiquark events, both inclusively and separately. The results are compared with the SM predictions as provided by various Monte Carlo (MC) generators.

These normalised differential cross-sections can be directly compared with the direct measurement of the top-quark and antiquark polarisation vectors [3], which are obtained from a template fit to the joint distributions of the direction cosines of the charged lepton momentum in the top-quark rest frame using the same dataset as in this work. Additionally, the normalised differential cross-sections as a function of $\cos\theta_{lx'}$ and $\cos\theta_{ly'}$ are sensitive to the real and imaginary part of the dipole operator $\mathcal{O}_{tW}$, respectively. Using the formalism previously described, these measurements have been used to perform an interpretation in an EFT context [3], providing limits for this operator. This is part of the PhD thesis in Ref. [4].

1.4 Mono-top search and its interpretation as dark matter

Astronomical observations based on gravitational effects indicate that approximately 25% of the matter-energy content of the universe is composed by a new kind of matter. However, there is still very little knowledge about the properties of this matter. It is known that it does not interact electromagnetically, and thus does neither emit nor absorb visible light. Therefore, it is usually referred to as *dark matter*. It could still interact weakly with the baryonic matter, even though there is no evidence so far of this type of interactions. The SM does not provide any viable candidate for DM, and thus the origin of this matter must arise from physics beyond the SM. The scientific community is making a strong effort to understand the nature of DM, which would dramatically increase our understanding of the universe. This section makes a brief introduction to the state-of-the-art of the DM research, from the evidences of DM to its characteristics and the strategies followed to identify its nature. Finally, this section concludes with an introduction to the mono-top model which is the channel targeted in this thesis.

1.4.1 Observational evidences for dark matter

The first evidence for DM comes from the observation of the rotation curves of disc galaxies. According to classical Newtonian gravitation, the velocity of the

stars orbiting with respect to the centre of a disc galaxy is given by the following expression:

$$v = \sqrt{\frac{GM(r)}{r}}, \tag{1.19}$$

where G is the gravitational constant ($G \sim 6.674 \cdot 10^{-11}\mathrm{m}^3\mathrm{kg}^{-1}\mathrm{s}^{-2}$), r is the distance between the considered star and the centre of the galaxy and $M(r)$ is the mass within the considered orbit. According to this expression, if there is just luminous matter the stars which are found further away from the centre should be moving at slower velocities. Indeed, in the solar system, this is observed for the different planets orbiting around the Sun, with Mercury being the fastest planet and Neptune the slowest. However, this does not match the observations for the stars orbiting the galactic centre. The observations point to an increasing or constant velocity the farther the star from the galactic centre. An example of this behaviour is found in the measured rotation curve of the nearby galaxy M33, depicted in Figure 1.11. According to Equation 1.19, the velocity of the stars should match the short dashed line if one considers just luminous matter. One possible explanation for this observations is given by DM. This relies on the existence of a DM halo surrounding the centre of the galaxy. Thus, as one considers farther distances, the amount of matter within an orbit is increased, and this can overcome the velocity reduction. In particular, $M(r) \propto r$ is enough to make the velocity of the stars constant. This condition would be fulfilled by a DM halo.

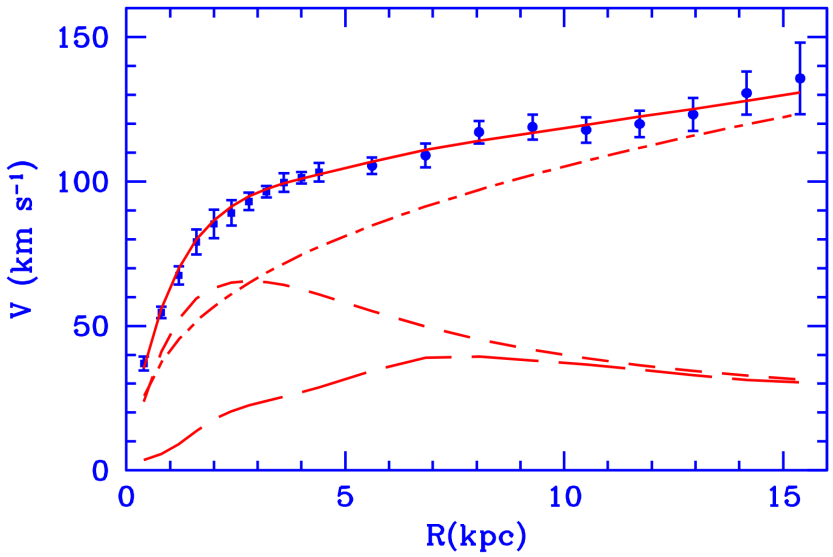


Figure 1.11: M33 rotation curve (points) compared with the best fit model (continuous line). The dashed-dotted line represents the halo contribution, the short dashed line represents the stellar disk and the long dashed line represents the gas contribution [134].

The first observation of this kind of behavior dates back to 1933. Fritz Zwicky studied the motion of the galaxies in the Coma cluster [135] and used the virial theorem [136] to conclude that most of the mass was missing. Concretely, he obtained a value for the mass approximately 400 times greater than the expected mass from the luminous disk, and he referred to this mass as dark matter. Since the postulation from Fritz Zwicky, there have been many observations of different clusters of galaxies with similar results. The existence of a DM halo is now supported by most of the scientific community, and the current efforts are focused on studying its profile density and distributions and understanding the matter which makes it.

Another evidence of DM comes from the so-called *gravitational lensing* [137–139]. According to Einstein's General Relativity theory, the light path is modified due to the space-time curvature induced by massive objects. If

the gravitational lensing is very strong, the different paths of light are diverted, and the light reaching the Earth seems to come from different places. When the source, the massive object and the Earth are in perfect alignment, this leads to a ring structure which is called the Einstein ring. If the gravitational lensing is weak, this results in a distortion of the observed image of the galaxy. An example of weak gravitation lensing is found in the Bullet cluster [140]. The Bullet Cluster includes two colliding clusters. The hot gas of the two colliding components, seen in the form of visible light and X-rays, represents most of the baryonic (luminous) matter in the cluster pair. In theories without DM, the lensing would be expected to follow the baryonic matter. However, the observation of distant quasars beyond the cluster indicates that gravitational lensing is much stronger in two separated regions which do not coincide with the position of baryonic matter. This indicates that most of the matter does not coincide with the luminous matter and therefore constitutes one of the most important arguments in favour of DM.

Nowadays, the model that best describes the evolution and composition of the universe is the Lambda cold dark matter (ΛCDM) model. This model considers three main components in the universe: baryonic matter, dark energy and dark matter. It provides an accurate description of some properties of the cosmos, such as the large scale structure in the distribution of galaxies [141–146] and the existence of the Cosmic Microwave Background (CMB) [147]. Large scale structure is defined as the structure or inhomogeneity of the universe on scales larger than that of a galaxy. The observed clustering of galaxies is expected to evolve with time, as structure continues to grow due to gravity. The comparison of the observed galaxy clustering signatures with N-body simulations of the ΛCDM model is compelling evidence for the existence of DM. Another important observation comes from the CMB. Approximately one second after the Big Bang, the universe was a hot plasma (electrons and protons were separated) and the photons could not travel very far. After $\sim$390 000 years, the temperature had dropped enough so that atoms could be formed (recombination), and the photons could decouple and travel freely (decoupling). These photons have been propagating ever since loosing energy due to the expansion of the universe, and they reach the Earth with microwave wavelengths. These are remnants of the Big Bang and thus contain

information about the early stages of the universe and the formation of matter. The CMB is homogeneous (with equal intensity) and mostly isotropic (uniform in all directions), corresponding to a black-body radiation of 2.7 K. However, there are small temperature fluctuations in the CMB spectrum. These anisotropies are used in simulation to predict the matter content in the universe. The Planck collaboration [85] has used this information to predict the amount of baryonic matter, DM and dark energy in the universe, and the results lead to a 26.8% of DM and 68.3% dark energy, with just 4.9% of the universe being composed by the baryonic matter that is known.

The previous examples can be considered as an evidence for DM, but some of them can be explained also with alternative theories. For instance, the rotation curves of the galaxies can be well accommodated within Modified Newtonian Dynamic (MOND) theories [148] or Tensor-vector-scalar gravity (TeVeS) [149] in the context of General Relativity. However, there are several arguments against this kind of theories. Among them, the most important comes from the observation of the Bullet Cluster, which is incompatible with these theories. Other evidences include the discovery of gravitational waves [150] in 2016. The speed of gravitational waves matched the speed of light [151], and this ruled out many hypothesis which used modified gravity [152–155].

1.4.2 Dark matter candidates

There are several models in the literature which propose different DM candidates. A good DM candidate must fulfill the following conditions (adapted from Ref. [156]):

1. Match the appropriate relic density. The analysis of the CMB yields an estimate of the DM relic density, i.e. the present quantity of DM remaining from the Big Bang [85]. A good DM candidate must have a production mechanism which is able to reproduce this value. In addition, it must be stable, with a lifetime which exceeds the present age of the universe [157].

2. DM candidates are expected to be non-relativistic, which is commonly known as "cold". N-body simulations of the Lambda cold dark matter (ΛCDM)

cosmological model are in agreement with a wide range of observations, such as the abundance of clusters in particular regions of the space and the galaxy-galaxy correlation functions [158].

3. DM candidates should be neutral, with null or very low electric or colour charges, since DM does not interact electromagnetically or through the strong interaction.

4. Be consistent with Big Bang Nucleosynthesis (BBN). BBN predicts the abundances of light elements produced in the first three minutes after the Big Bang with high level of precision. A DM candidate should not induce discrepancies with these observations. Reviews of BBN may be found in Refs. [159, 160].

5. Leave the stellar evolution unchanged. DM candidates may be produced through the weak interaction in the interior of stars. If they escape without further interactions, this implies an energy loss for the star which may modify the stellar evolution. The observation of globular-cluster stars provides upper limits on any additional energy loss rate [161] which can be applied to constrain DM models.

6. Be compatible with the constraints on its self-interactions. Observations of merging clusters [162] and of the ellipticity of certain galaxies as inferred from X-rays [163] constrain the level of DM-DM self-interactions to $\sigma_{DM-DM}/m_{DM} < 0.47$ cm^2/g at 95% C.L. [164].

7. Be compatible with experimental constraints. These include all direct and indirect detection, and accelerator results. Further details will be given later in this section.

8. It must be possible to probe DM experimentally. If there is no possibility to find evidence of a DM candidate in present or future experiments, then it is not considered as a good DM candidate.

There are several candidates which could fulfill these conditions. Some of these are described below. For a full description, see [12]:

- **Sterile neutrinos**: Sterile neutrinos are similar to the SM neutrinos. The main difference is that the sterile neutrinos do not interact weakly, leaving just the gravitational interaction to detect these particles. As explained before, neutrinos are massless within the SM, but it has already been proven that neutrinos have mass, even if it is very low. Thus, a theory beyond the SM is needed to recover the mass of the neutrinos. Staying with the Dirac particle formalism, a right-chiral neutrino which couples to the left-chiral component could be a possibility to recover the neutrino masses. This right-chiral component could be a good DM candidate. However, its mass needs to be equal to its left-chiral counterpart, which in turn is not enough to hold galaxies together [165]. Another possibility is to recover neutrino masses from the Seesaw mechanism (see Section 1.1.5). This formalism predicts a very low Dirac mass term, which would match the SM neutrinos, and also very high Majorana mass terms, which could be a DM candidate. Nowadays there are no evidences of sterile neutrinos [166, 167].

- **Axions**: Axions are attractive DM candidates because they can also explain the strong CP problem described in Section 1.1.5. These particles are pseudo-Nambu-Goldstone bosons which are expected to be light and very weakly coupled. This implies it is very challenging to detect them. A review of the current status of axions searches can be found in Ref. [168].

- **Dark photons**: Dark photons are vector bosons which would act as a portal between the dark and ordinary sectors because of their kinetic mixing with the visible photon. These particles could be either massless or massive. In the case of massless dark photons, the interplay between these particles and the DM could open new windows to study the properties of DM. However, a very light massive dark photon could be a viable DM candidate by itself. These could have been produced in the early universe and would be cosmologically stable. Constraints on the parameter space stem from a combination of direct detection experiments, from stellar cooling constraints and from the CMB. For a full review on the state-of-the-art of dark photon searches, see Ref. [169].

- **Astronomical objects**: These include primordial black holes, failed stars and massive compact halo objects (MACHOs), among others. These candidates could constitute a good DM candidate. However, they seem to be insufficient to explain the total DM content of the universe [170–172].

- **Weakly interacting massive particles (WIMPs)**: These are one of the preferred DM candidates. They often arise from theories beyond the SM which could also solve the hierarchy problem (see Section 1.1.5). One example of these theories is SUSY. This is an extension of the SM which adds another fundamental symmetry, implying that each fermion (boson) has a boson (fermion) super-partner with the same quantum numbers. In many Minimal Supersymmetric Standard Models (MSSM), due to the conservation of R-parity [173, 174], superpartners can only be created or destroyed in pairs. This implies that the lightest supersymmetry particle (LSP) does not decay. Examples of LSP include the sneutrino, the gravitino and the neutralino[26]. All them are characterised by production cross-sections of approximately the weak strength. In the early universe, particles were constantly being created and self-annhilated in the so-called "thermal soup". At some point, due to the expansion of the universe, the universe was sufficiently large and these interactions ceased. This is the so-called "freeze-out". If WIMPs were to exist in the early universe, after the "freeze-out" their relic abundance should stay approximately constant. It is expected that the relic density of WIMPs with masses on the weak scale match naturally this relic density of DM [175], which is usually knows in the literature as the "WIMP miracle". These considerations make WIMPs highly motivated DM candidates.

The range of masses of potential DM candidates could cover 90 orders of magnitude in mass. It could be a particle as light as 10^{-22} eV or as massive as primordial black holes of tens of solar masses. Typical mass ranges are depicted in Figure 1.12. In order to cover all possibilities, an enormous variety of experiments and techniques is needed.

[26] Only one of these can be the LSP in a given SUSY scenario, but different scenarios predict different mass hierarchies.

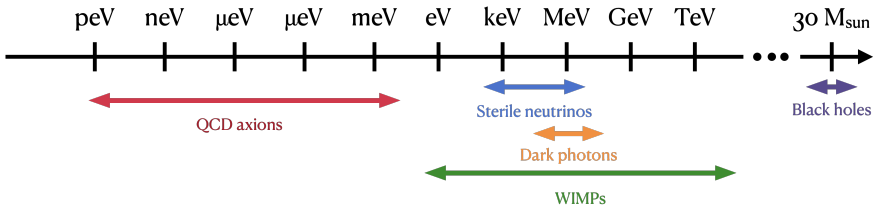


Figure 1.12: Typical mass ranges of several DM candidates. Adapted from Ref. [176].

1.4.3 Dark matter searches

In spite of the strong evidence that indicates the existence of DM, the properties of these particles are still far from being understood. In order to shed some light on its true nature, there are different searches which aim to detect DM particles. Most of these searches assume that DM particles are WIMPs, and would interact weakly with the SM particles. They can be classified in three groups: direct detection, indirect detection, and searches in particle colliders, depending on how they interact with the SM particles. A schematic description of these three categories can be found in Figure 1.13. These searches are complementary, since they are sensitive to different WIMP mass ranges, different couplings of the DM and different systematic uncertainties. A description of the state-of-the-art for the three types of searches is given in the following.

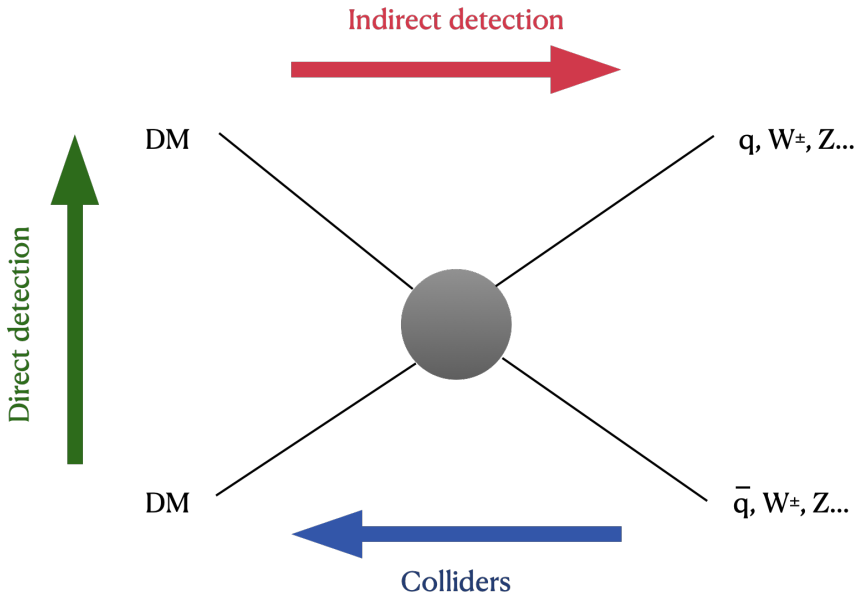


Figure 1.13: Schematic representation of the different strategies in the search for DM. In direct detection, DM particles scatters SM particles. In indirect detection, DM particles annihilate or decay producing SM particles. In DM searches at colliders, SM particles interact to produce DM particles.

1.4.3.1 Direct detection

Direct detection aims to find evidence of DM particles through the elastic or inelastic scattering of WIMP particles with the atomic nucleus. Thus, the aimed signal is a nuclear recoil in the target mass. This can be measured through either charged particles (ionisation), photons (scintillation), an increase of temperature or a bubble nucleation. This searches can be further classified in two groups: spin-dependent (SD), in which the WIMPs spin couples with the spin content of a nucleon, and spin-independent (SI) interactions. If nuclear recoils are SI, then all the nucleons contribute coherently to the cross-section and the amplitude is proportional to the number of nucleons. In contrast, in SD scattering, the spin direction of the different nucleons needs to be considered, which leads to many cancellations. Therefore, direct searches are in general much less sensitive

Experiment	Target	Fiducial mass [kg]	Cross-section [cm^2]	DM mass [GeV]	Ref.
Spin independent high mass (> 5 GeV)					
XENON1T	Xe	1042	4.1×10^{-47}	30	[180]
PandaX-II	Xe	364	8.6×10^{-47}	40	[181]
LUX	Xe	118	1.1×10^{-46}	50	[182]
SuperCDMS	Ge	12	1.0×10^{-44}	46	[183]
DarkSide-50	Ar	46	1.14×10^{-44}	100	[184]
DEAP-3600	Ar	2000	3.9×10^{-45}	100	[185]
Spin independent low mass (< 5 GeV)					
LUX (Migdal)	Xe	118	6.9×10^{-38}	2	[186]
XENON1T (Migdal)	Xe	1042	3×10^{-40}	2	[187]
XENON1T (ionisation only)	Xe	1042	3.6×10^{-41}	3	[188]
DarkSide-50	Ar	20	1×10^{-41}	2	[189]
SuperCDMS (CDMSlite)	Ge	0.6	2×10^{-40}	2	[190]
CRESST	$CaWO_4$ - O	0.024	1×10^{-39}	2	[191]
NEWS-G	Ne	0.3	1×10^{-38}	2	[192]
Spin dependent proton					
Pico60	C_3F_8 - F	49	3.2×10^{-41}	25	[193]
Spin dependent neutron					
XENON1T	Xe	1042	6.3×10^{-42}	30	[194]
PandaX-II	Xe	364	1.6×10^{-41}	40	[195]
LUX	Xe	118	1.6×10^{-41}	35	[196]

Table 1.8: Upper limits at 90% C.L. from direct detection experiments on the DM-nucleon SI (at high > 5 GeV and low < 5 GeV masses) and SD cross-sections. The fiducial mass of the target material for each experiment is given. The DM mass corresponds to the most stringent constraint on the DM-nucleon scattering cross-section [12].

to SD interactions. The target material is chosen accordingly, depending on its sensitiveness to SD and SI couplings. The tightest constraints from direct searches are summarised in Table 1.8. These limits are obtained assuming the local DM density [177–179]. All results are compatible with the background-only hypothesis.

1.4.3.2 Indirect detection

Indirect detection aims to detect the SM particles that are produced in the annihilation or the decay of DM. The setup of the different experiments depends on the target particle. This includes neutrinos, cosmic rays (electrons, positrons, antiprotons and antideuterons) and photons (radio, X-rays and gamma rays).

Neutrino telescopes aim to detect high-energy neutrinos coming from the annihilation of DM particles close to massive objects such as the Sun or the

Galactic Centre. Given that neutrinos interact weakly, they could travel to the Earth without further interactions and could be detected. In order to detect these neutrinos, the target material needs to be very large. Examples of the target material include large pools, the ocean or the Antarctic ice. The neutrino telescope is placed underground, such as Super-Kamiokande (Super-K) [197], under-ice, such as IceCube [198], or underwater, such as ANTARES [199]. The goal is to reduce as much as possible the background from muons produced by cosmic rays in the atmosphere. The neutrinos coming from the muon decays are further rejected selecting just the up-going events or applying requirements on the energy of the neutrinos. All IceCube, Super-K and ANTARES have placed strong constraints on the SD DM annihilation cross-section into W^+W^-, $b\bar{b}$ and $\tau\bar{\tau}$ [200–202].

In the field of cosmic rays, it is of particular interest a positron excess found by PAMELA [203] and later confirmed by AMS [204]. Both detectors are placed in space. PAMELA was attached to an Earth orbiting satellite[27], whereas AMS is installed at the International Space Station and still operating. These detectors target rare particle species, such as positrons, antiprotons or antinuclei, to reduce drastically the background contamination. The excess found by these detectors could be interpreted as DM [206, 207], but it is still compatible with positron production by pulsars or supernova remnants [208, 209].

Finally, regarding photon searches, it is worth to mention the Fermi Large Area Telescope (Fermi-LAT). This is a space observatory which performs gamma ray measurements. So far, it has provided very strong constraints on the DM annihilation cross-section from the observation of dwarf galaxies [210]. In addition, some excesses ($\sim 4.5\sigma$) have been found [211–213] which could be a sign of WIMP particles [214]. However, the possibility of gamma ray emission by pulsars cannot be discarded [215].

1.4.3.3 Detection at colliders

In DM searches at colliders, it is expected that collisions of SM particles could produce a WIMP DM candidate. This DM candidate would not interact with

[27] PAMELA operations finished in 2016 [205].

the material of the detector, and thus would escape undetected as neutrinos do. Therefore, DM searches at the LHC are focused on finding significant amounts of missing transverse momentum (see Section 2.5.5) recoiling against a high transverse momentum SM particle[28]. A particular strength of collider searches lies in the fact that the high-energy collisions of SM particles could not only produce DM directly under controlled experimental conditions, but also provide access to particles mediating the interactions between DM and the SM sector.

Several models have been proposed for the production of DM particles in high energy collisions [216, 217]. These models can be grouped in mainly three categories. The first is based on EFT. In EFT minimal assumptions on the new particle spectrum are made: the DM particle is the only new state beyond the SM kinematically accessible at the LHC. In this context, EFT describes a general contact-like interaction between SM and DM mediated by particles with masses above the kinematical reach of the LHC with a very reduced number of parameters. The simplicity and model independence of the EFT approach clashes with its somewhat limited applicability. In contrast, complete models such as SUSY or Universal Extra Dimensions [218] are the most complete alternative from the theoretical point of view. However, these models imply a large number of parameters, and the realization of DM depends crucially on the fine details of the model. Therefore, complete models lack generality. In between these two possibilities one can find the so-called simplified models. In these simplified models, it is assumed that WIMP particles are single Dirac fermions, usually labelled χ, which would interact with the SM particles through the exchange of a DM-SM mediator. In these models the number of free parameters is limited to the DM and mediator particle masses, spin/parity and couplings. Common benchmark scenarios for simplified models have been agreed by the LHC DM Working Group [219]. To preserve MFV content of the model, it is also assumed that the flavour structure of the couplings between the SM particles and the DM mediators should follow the same structure as in the SM.

The simplified models lead to different types of DM signatures. This can be classified mainly within two different groups:

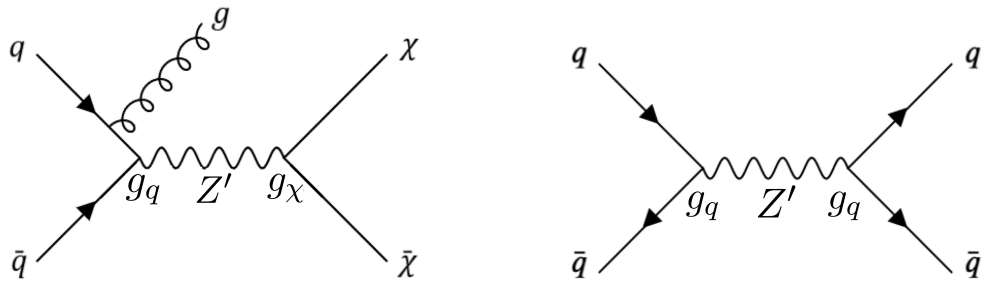
[28] This is necessary to trigger the acquisition of the full event data.

- **Mono-X**: These simplified models assume that DM particles are produced along with SM particles X, which can be either photons, jets, vector or Higgs bosons, or bottom or top quarks. This kind of processes are characterised by a large amount of missing transverse momentum recoiling against the SM particles. These models assume that the coupling between the SM and the DM particles is mediated by another particle, called the *mediator*. This can be either a scalar, a pseudoscalar, a vector or an axial-vector. The free parameters include the couplings between the SM particles and the mediator g_q, the coupling of the DM particles to this mediator g_χ, and the mass of the mediator m_{med} and the DM particles m_χ.

- **Resonances**: These models assume that the DM mediator decays to a couple of SM particles, mainly a couple of jets or leptons, from which its properties could be reconstructed. This would lead to a peak in the dijet or dilepton invariant mass distributions, or an excess of events in specifically chosen angular distributions. The free parameters in these models are reduced to the coupling g_q and the mass of the mediator m_{med}.

Schematic Feynman diagrams of the two categories can be found in Figure 1.14. Both the ATLAS and CMS collaborations have searched for DM particles considering different simplified models during Run-1 and Run-2. In absence of a signal, these measurements have been used to set constraints on the production cross-section of the DM particles or to exclude parameters of the model beyond a particular threshold.

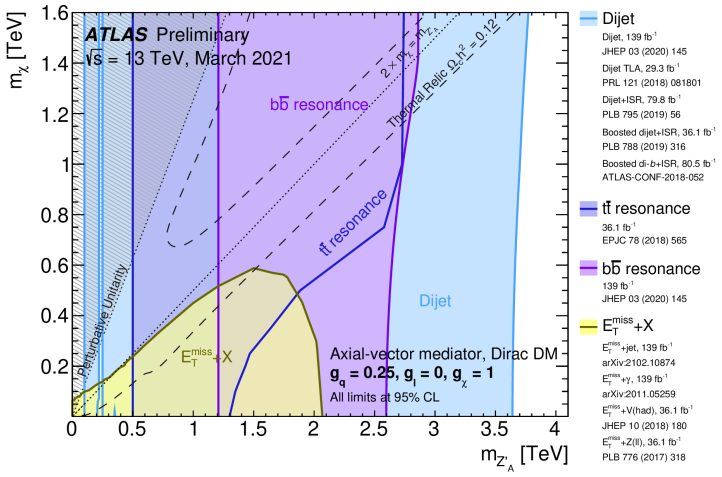
The most recent results in the field of DM searches as provided by ATLAS are displayed in Figures 1.15-1.17 from the combination of different searches. Figure 1.15 shows the excluded regions in the mediator-mass–DM-mass plane for axial-vector mediator simplified models. Both leptophobic and leptophilic scenarios are considered. In the former the mediator couples preferentially to quarks and in the latter lepton couplings are also considered. Figure 1.16 shows the same results for vector mediator simplified models. Finally, Figure 1.17 shows the exclusion limits on the production cross-section for colour-neutral scalar and pseudoscalar mediator dark models as a function of the mediator mass.



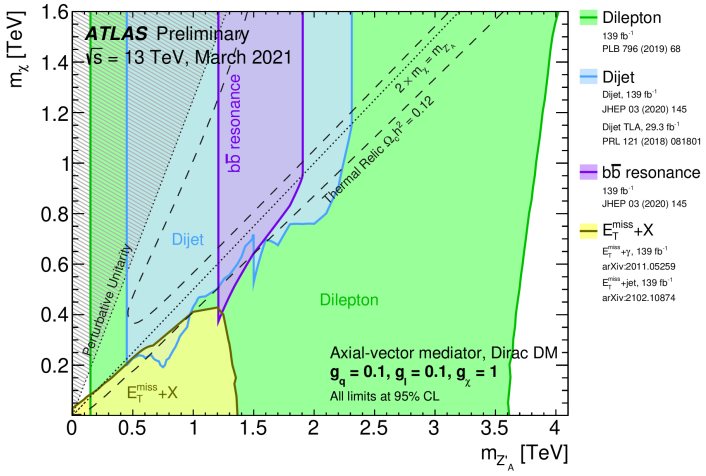
(a) (b)

Figure 1.14: Schematic representation of the two types of simplified models considered in the DM searches at colliders with a vector boson mediator Z'. (a) Mono-X: A pair of DM particles is produced from the decay of the DM mediator along with an additional particle. In this case this additional particle is a gluon, leading to a mono-jet search. (b) Resonant: A pair of SM particles is produced from the decay of the DM mediator.

The latest results from CMS are summarised in Figure 1.18 from the combination of different analyses. The exclusion regions in the mediator-mass–DM-mass plane are computed for both the axial-vector and the vector models.

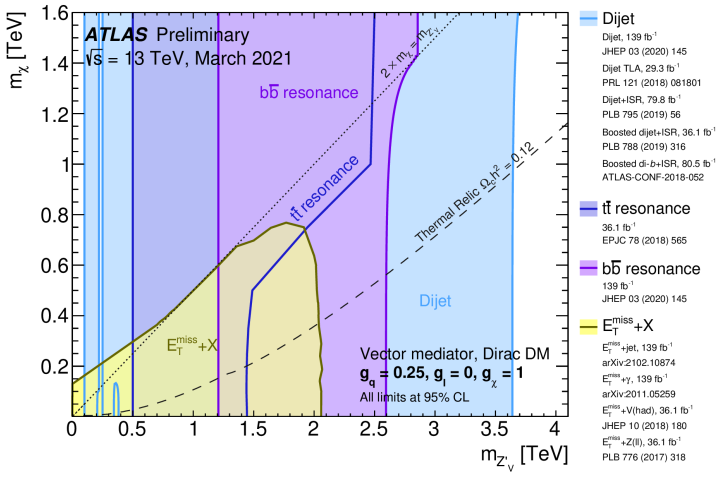


(a)

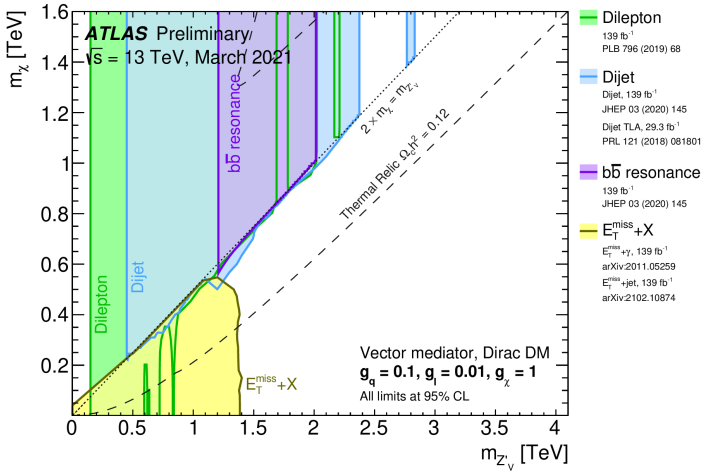


(b)

Figure 1.15: Regions in the mediator-mass–DM-mass plane excluded at 95% C.L. by visible and invisible searches from ATLAS for (a) leptophobic and (b) leptophilic axial-vector mediator simplified models. Each shaded region represents the union of the exclusion contours of the individual analyses listed in the legend, where more than one result contributes. The exclusions are computed for a DM coupling g_χ = 1, quark coupling (a) $g_q = 0.25$ and (b) $g_q = 0.1$, and lepton coupling (a) $g_l = 0$ and (b) $g_l = 0.1$, in both cases universal to all flavours. Dashed curves labelled "thermal relic" correspond to combinations of DM and mediator mass values that are consistent with the DM density $\Omega h^2 = 0.12$. Between the two curves, annihilation processes described by the simplified model deplete Ωh^2 to below 0.12. A dotted line indicates the kinematic threshold where the mediator can decay on-shell into DM. Excluded regions that are in tension with the perturbative unitary considerations of [220] are indicated by shading in the upper left corner [221].

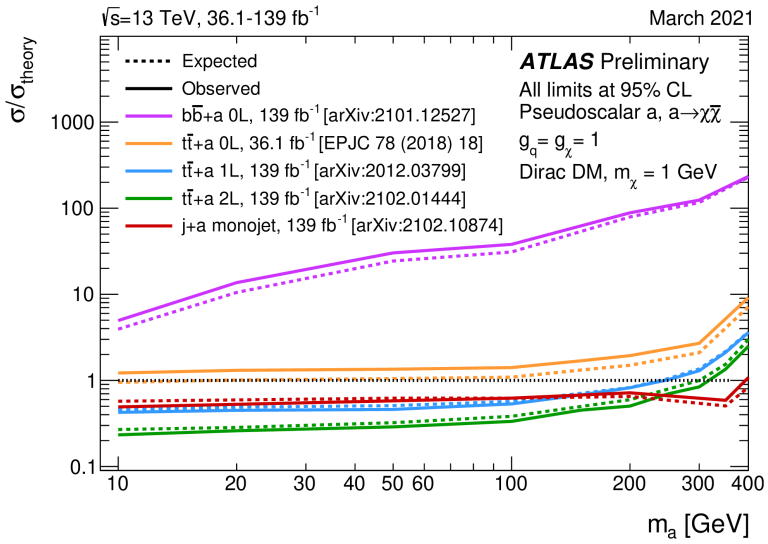


(a)

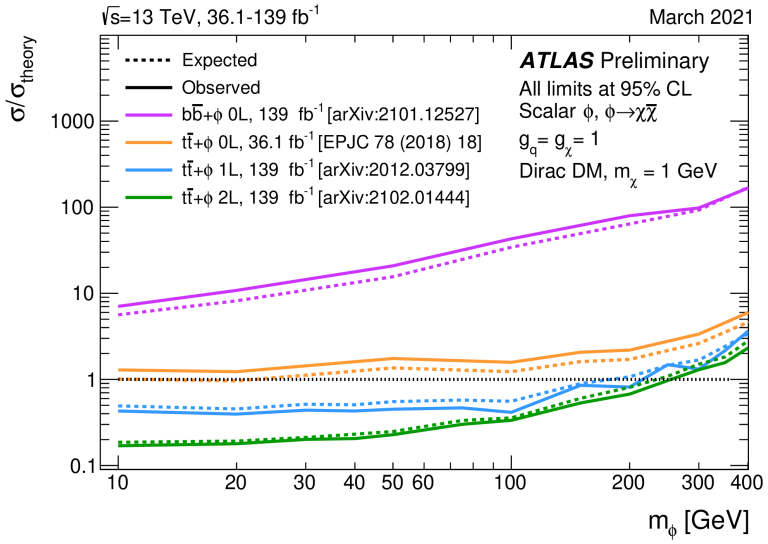


(b)

Figure 1.16: Regions in the mediator-mass–DM-mass plane excluded at 95% C.L. by dijet, dilepton and $E_{\mathrm{T}}^{\mathrm{miss}}$+X searches from ATLAS, for leptophobic (a) and leptophilic (b) vector mediator simplified models. Each shaded region represents the union of the exclusion contours of the individual analyses listed in the legend, where more than one result contributes. The exclusions are computed for a DM coupling $g_\chi = 1$, quark coupling (a) $g_q = 0.25$ and (b) $g_q = 0.1$, and lepton coupling (a) $g_l = 0$ and (b) $g_l = 0.01$, in both cases universal to all flavours. Dashed curves labelled "thermal relic" correspond to combinations of DM and mediator mass values that are consistent with the DM density $\Omega h^2 = 0.12$. (a) Above the curve and (b) between the two dashed curves, annihilation processes described by the simplified model deplete Ωh^2 to below 0.12. The dotted line indicates the kinematic threshold where the mediator can decay on-shell into DM [221].

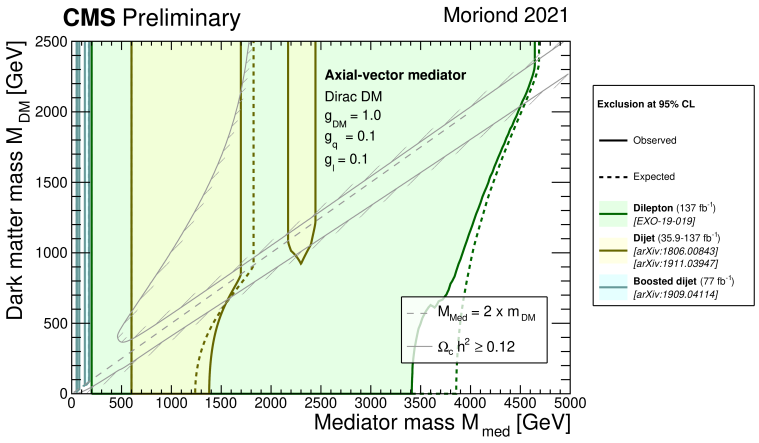


(a)

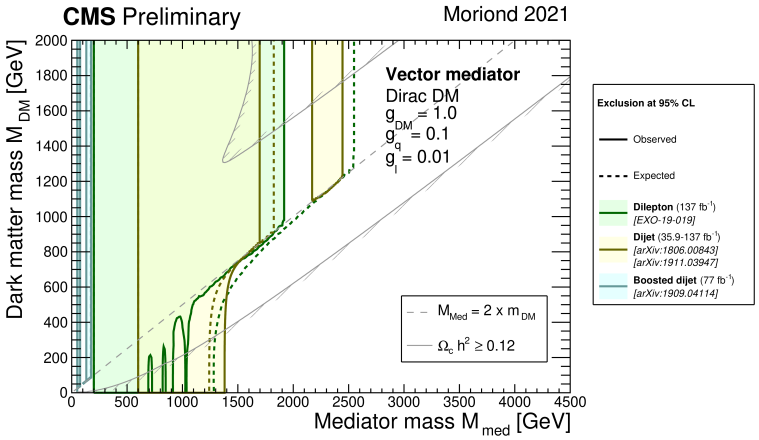


(b)

Figure 1.17: Exclusion limits for colour-neutral (a) pseudoscalar and (b) scalar mediator DM models from ATLAS as a function of the mediator mass (a) m_a and (b) m_φ for a DM mass m_χ of 1 GeV. The limits are calculated at 95% C.L. and are expressed in terms of the ratio of the excluded cross-section to the nominal cross-section for a coupling assumption of $g = g_q = g_\chi = 1$. The solid (dashed) lines show the observed (expected) exclusion limits for different analyses [221].



(a)



(b)

Figure 1.18: 95% C.L. observed and expected exclusion regions in the mediator-mass–DM-mass plane for dijet and dilepton searches from CMS in (a) the axial-vector and (b) the vector model. The exclusion regions are computed for a universal quark coupling of $g_q = 0.1$, lepton coupling $g_l = 0.1$, and for a DM coupling of $g_\chi = 1.0$. The solid grey line corresponds to combinations of DM and mediator mass values that are consistent with a DM density $\Omega h^2 \geq 0.12$. The dotted grey line indicates the kinematic threshold where the mediator can decay on-shell into DM [222].

1.4.4 Mono-top searches at the LHC

Mono-top signatures are characterised by a single top quark produced in association with a large amount of missing transverse momentum, which can be interpreted in several DM scenarios. The simplified model described in Refs. [223–225] covers the different processes leading to this final state, and it is frequently adopted as a benchmark for LHC searches.

Two mono-top production mechanisms arise from imposing that the model Lagrangian respects the electroweak $SU(2)_L \times U(1)_Y$ gauge symmetry and by requiring minimality in terms of new states which supplement the SM fields: the *resonant* and the *non-resonant* models.

The resonant model predicts the existence of a coloured 2/3-charged scalar (ϕ) decaying to a top quark and a spin-1/2 DM particle (χ), as depicted in the left Feynman diagram on Figure 1.19. The Lagrangian of this process can be written as [223, 226]:

$$\mathcal{L}_{\text{res}} = \lambda\phi\bar{d}^c P_R s + y\phi\bar{\chi}P_R t + h.c., \tag{1.20}$$

where d, s and t represent the d-, s- and top-quark fields, respectively. ϕ and χ represent the charged scalar and the DM particle fields. P_R stands for the right-chiral operator. Finally, λ and y are the couplings of the charged scalar to the d- and s-quarks and to the top quark and the DM particle, respectively. This kind of interaction arises in MSSM scenarios with an R-parity violating interaction where a top squark decays to a top quark and a neutralino [227–229] and R-parity conserving interactions of a top quark and a top squark to a neutralino. Another possible interpretation arises from leptoquarks decaying to a top quark and an antineutrino [230], or hylogenesis models of baryon number asymmetry where the top quark is produced in association with several DM candidates [231, 232].

The middle and right Feynman diagrams in Figure 1.19 refer to the non-resonant model. This model is characterised by a FCNC interaction between an up-type quark, a top quark, and a vector mediator particle V. This vector boson particle decays to a couple of DM Dirac fermions $\chi\bar{\chi}$. The Lagrangian that describes this interaction can be written as [226, 233]:

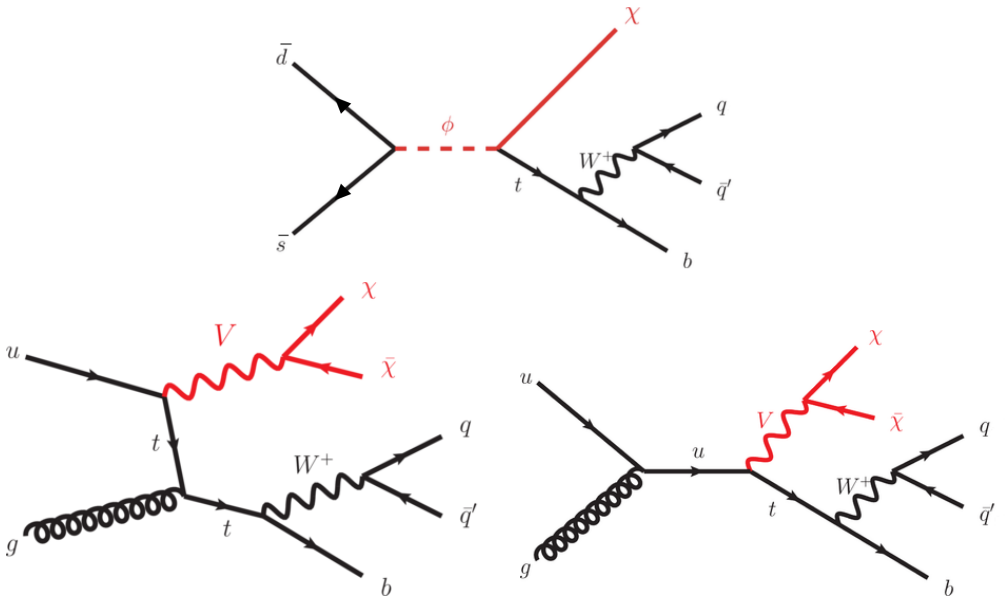


Figure 1.19: Feynman diagrams for the mono-top production models in the DM context. These include the resonant (top) and the non-resonant (bottom) scenarios. The hadronic decay of the top quark is included in the diagram.

$$\mathcal{L}_{\text{non-res}} = aV_\mu \bar{u}\gamma^\mu P_R t + g_\chi V_\mu \bar{\chi}\gamma^\mu \chi + h.c., \tag{1.21}$$

where u and t are the up-type and top-quark fields, and V_μ and χ are the invisible vector boson and Dirac fermion fields. a represents the coupling strength between the massive vector boson V and the up-type and the top quark, g_χ represents the coupling strength between the massive vector boson and the DM particle χ and γ^μ are the gamma matrices. There are several new physics models which predict this kind of interaction [234–236], most notably in models linking cosmological DM with flavour dynamics [235]. Since in the SM FCNCs are disfavored, these vertices Vtu are an obvious place to look for new physics. Indeed, due to loop and renormalization group effects involving SM Yukawa couplings, some amount of flavour violation in the interactions between DM and the SM sector is unavoidable [237]. DM could then be discovered through flavour violating

interactions. The mono-top final state is particularly interesting since it dominates over other DM collider signatures in many models. For instance, its cross-section can be orders of magnitude larger than the mono-jet cross-section even in the case of MFV, if the interactions are chirality flipping [235].

ATLAS and CMS have both searched for such topology at $\sqrt{s} = 8$ TeV using a dataset of 20.3 fb^{-1} [238, 239]. The CDF collaboration also presented a result of such topologies using 7.7 fb^{-1} of Tevatron $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV [240]. ATLAS and CMS released the results of Run-2 searches in this event topology from pp collisions at $\sqrt{s} = 13$ TeV collected during the 2015 and 2016 data-taking [241–243].

For DM production in the non-resonant scenario, masses of a new vector particle coupling to the DM candidate up to 2 TeV are excluded at 95% C.L. for $m_\chi = 1$ GeV, $g_\chi = 1.0$ and $a = 0.5$, while in the resonant case, masses of a new scalar particle coupling to DM up to 3.4 TeV are excluded at 95% C.L. for $m_\chi = 10$ GeV, $y = 0.4$ and $\lambda = 0.2$.

The analysis presented in this thesis describes a search for a top quark produced in association with missing transverse momentum. Previous searches [241] have shown that the channel with a top quark decaying hadronically is much more sensitive to this kind of signature than the leptonic decay. Therefore, the work of this thesis focuses on the hadronic decay of the top quark. The analysis is based on data from pp collisions at a centre-of-mass energy of $\sqrt{s} = 13$ TeV using the full Run-2 dataset. The results are interpreted in the context of DM, considering both the mono-top resonant and non-resonant production modes.

2 Experimental setup: The LHC accelerator and the ATLAS experiment

The two analyses described in this thesis are based on data from pp collisions at a centre-of-mass energy of $\sqrt{s}$ = 13 TeV collected by the ATLAS detector at the LHC during the Run-2 (2015-2018). The LHC is a particle collider in which particle beams are accelerated to ultra-relativistic energies and collide at specific points. The outcome of the collisions is recorded by different detectors, placed at the interaction points. The ATLAS detector is one of the two multi-purpose detectors, aiming to perform high-precision tests of the SM and searches for new phenomena beyond the SM.

This chapter introduces the different components that form the LHC, focusing on the experimental setup needed to produce high energy pp collisions. A brief introduction to the phenomenology of pp collisions is also given. Afterwards, the ATLAS detector and its different constituents are described, with special emphasis on its innermost component, the Inner Detector (ID). The reason is that this work also contributes to the alignment of the ID, as it will be explained in Chapter 3. Then, the different triggers employed to reduce the immense amount of collision data produced at the ATLAS detector are outlined. The techniques that are used in the ATLAS reconstruction of the physics objects utilised in the analyses are also reported, with special emphasis on those used in the presented analyses. Next, an introduction to the computing infrastructure used to process and deliver the

recorded data for physics analyses is given. The chapter concludes with the trigger and physics object selection used in the two analyses presented in this work.

2.1 The Large Hadron Collider

The LHC [7, 8] is placed at the European Council for Nuclear Research (CERN)[1] [244], Geneva, across the border between France and Switzerland, and it is the most powerful particle collider in the world. It is placed ∼100 m underground, and it consists of a 27-kilometre ring. Within the LHC, the particles are accelerated in two separate beam pipes and they collide at specific points, where the detectors are placed. The accelerating particles are mostly protons, even though heavy ions collisions are also recorded (mainly ionised lead, but also xenon). The particles are accelerated using electric fields in radio-frequency cavities. They are guided through the LHC ring using superconducting dipole magnets, with a design magnetic field of 8.33 T. The superconducting state is reached using a cryostat filled with liquid helium, at a temperature of 1.9 K. Additional quadrupole magnets are also used to focus the beams. The particles are accelerated in a ultra-high vacuum environment (10^{-13} atm) to prevent collisions with gas molecules in the beam pipe. The design accelerator parameters are summarised in Table 2.1.

The colliding protons are obtained from ionised hydrogen. Protons are then accelerated in different sequential stages using different accelerators, as depicted in Figure 2.1. Firstly, the protons go through the Linac 2, which is a linear accelerator, and they reach an energy of 50 MeV. Afterwards, the protons are injected in the Proton Synchroton Booster (PSB), where they are accelerated to energies of 1.4 GeV. The next accelerator is the Proton Synchroton (PS), in which the protons reach an energy of 25 GeV. The proton beams are split at this stage in different bunches, each with approximately 10^{11} protons and with a temporal separation of 25 ns between them. Then, the protons are accelerated in the Super Proton Synchroton (SPS) to 450 GeV. Finally, the protons are transferred to the two beam pipes of the LHC. The two beams are accelerated in opposite directions, and each beam carries one half of the desired centre-of-mass energy.

[1] The acronym is adopted from the French name, *Conseil Européen pour la Recherche Nucléaire*.

LHC parameter	Design value	Run-1	Run-2
Beam energy [TeV]	7	3.5/4	6.5
Centre-of-mass energy ($\sqrt{s}$) [TeV]	14	7/8	13
Peak Luminosity ($\mathcal{L}$) [$\cdot 10^{34}$ cm^2 s^{-1}]	1.0	0.8	2.0
Frequency (f) [MHz]	40	20	40
Bunch separation [ns]	25	50	25
Number of particles per bunch (N_b) [$\cdot 10^{11}$]	1.15	1.6	1.2
Transverse emittance (ϵ) [μm]	3.5	2.2	2.2
Amplitude function at the interaction point (β^*) [cm]	55	80	$30 \rightarrow 25$
Nominal magnetic field (B) [T]	8.33	4.16/4.76	7.73
Injection energy [GeV]	450		
Circumference length [km]	26.7		
Radius [km]	4.24		
Number of dipole magnets	1232		
Length of the dipole magnets [m]	14.3		
Number of quadripole magnets	392		
Total mass [tons]	27.5		

Table 2.1: Design value, Run-1 and Run-2 accelerator parameters of the LHC. The tranvserse emittance ϵ and the amplitude function at the interaction point β^* determine the beam size at the intersection point [245].

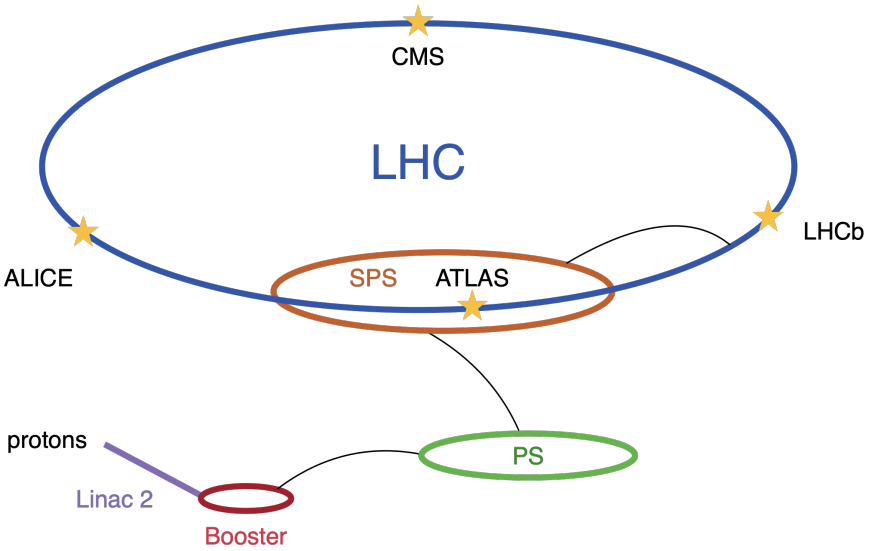


Figure 2.1: Schematic view of the CERN accelerator complex. The four main experiments, placed at the LHC ring collision points, are also displayed.

So far, the LHC has recorded different types of collisions and at different centre-of-mass energies. The pp collisions have been recorded at $\sqrt{s}$ = 5, 7, 8 and 13 TeV. The Pb-Pb collisions have been recorded both at $\sqrt{s}$ = 2.76 and 5.02 TeV, whereas p-Pb collisions have been recorded at $\sqrt{s}$ = 5.02 TeV. Finally, Xe-Xe collisions have been recorded at $\sqrt{s}$ = 5.44 TeV. The data has been collected in two well differentiated stages of consecutive data-taking. The first one, named Run-1, includes the collection of data from 2009 until 2013. The second one, the so-called Run-2, started in 2015 and finished in 2018. The pp collisions were recorded at $\sqrt{s}$ = 7, 8 TeV during Run-1, whereas during Run-2 they were recorded at $\sqrt{s}$ = 13 TeV. The data analysed in this thesis corresponds to the full Run-2 dataset. Between data-taking periods, the detectors and the accelerator are upgraded or refurbished in the so-called Long Shutdowns. The period of time between Run-1 and Run-2 is frequently labelled Long Shutdown 1 (LS1). Currently, the LHC is in its second shutdown (LS2), which will be followed by the Run-3 between 2022 and 2024, in which the LHC will reach an operation energy of $\sqrt{s}$ = 13.6 TeV. Afterwards, in 2025, the LHC is expected to enter the Long Shutdown 3 (LS3), in which the machine will be significantly upgraded to record a much larger amount of data, in the so-called High-Luminosity LHC (HL-LHC)[2] [248].

2.1.1 LHC detectors

The LHC ring can be divided in eight arcs with eight associated intersection points. Four of these intersection points are used for the beam injection, the beam dump system, the beam collimator and to host the radio-frequency cavities, respectively. The other four intersection points host the four main experiments at the LHC. One of these experiments is the ATLAS detector, which will be explained in detail in Section 2.3. The other three experiments include:

- **CMS**: The Compact Muon Solenoid [249] (CMS) is one of the two multi-purpose detectors at the LHC, together with ATLAS. The main

[2]Luminosity is an indicator of the number of collisions that occur in a given amount of time, as described in Section 2.2.3. During Run-2 ATLAS recorded 147 fb^{-1} of data. After detector and data quality requirements, the total integrated luminosity used for Run-2 analyses is 139 fb^{-1} [246, 247]. For the HL-LHC, it is expected to provide up to 3000 fb^{-1}.

motivation to build these two detectors was to discover and study the Higgs boson. In addition, these detectors aim to provide high precision measurements of the SM and search for new physics.

- **ALICE**: A Large Ion Collider Experiment (ALICE) [250] focuses on the collisions of heavy ions at ultra-high energies to produce a new state of matter named quark-gluon plasma, in which strong interactions dominate. It is expected that this state of matter predominated in the early Universe, and so it allows to study the first moments after the Big Bang under laboratory conditions.

- **LHCb**: The Large Hadron Collider beauty (LHCb) [251] is a one-sided detector which focuses on flavour physics of the b-quark. It has a wide coverage in the forward direction, close to the beam line, since B-hadrons are expected to be produced mainly in this direction. It aims at studying the subtle differences between matter and antimatter.

In addition to the four main experiments at the LHC, there are other experiments focusing on specific tasks. These include the total, elastic and diffractive cross-section measurement (TOTEM) [252], which aims to measure the total pp collision cross-section as well as diffractive pp physics; the LHC forward (LHCf) experiment [253], which performs measurements in the beam direction and provides an estimate of the energy of ultra-high energy cosmic rays; the Monopole and Exotics Detector at the LHC (MoEDAL) [254], which aims to detect magnetic monopoles and highly ionising stable massive particles; and the Forward Search Experiment (FASER) [255], which looks for light and extremely weakly interacting new particles and studies neutrinos.

2.2 Phenomenology of proton-proton collisions

The data analysed in this thesis is based on pp collisions recorded at the LHC with the ATLAS detector. In order to analyse this data, one first needs to understand the underlying phenomenology in a pp collision. Protons are not

fundamental particles, since they are made of quarks and gluons. Concretely, they are made by three *valence* quarks, two being up-type quarks and one down-type quark. However, these are not the only constituents in the proton. The valence quarks are continuously interacting within the proton through the strong interaction, and thus they exchange gluons. The gluons can split into a quark-antiquark pair, electrically neutral (the so-called *gluon splitting*). This second set of quarks form the so-called *sea* quarks. These three sets of particles constitute the proton: the valence and sea quarks, and the gluons. This description of the constituents of the proton is known as the parton model.

2.2.1 The parton model

pp interactions can be described in two main regimes depending on the energy scale: the *elastic* and the *inelastic* interactions. The elastic interactions occur at low energies, whereas the inelastic interactions occur at higher energies. At very high energies, such as the ones at the LHC, the deep inelastic regime is reached. In this regime, the quark content of the proton becomes totally accessible.

When two protons collide at the LHC, it is a parton from each proton that interacts in the hard scattering process, and the interaction can be considered as independent of the other partons. The remaining partons also contribute to the final state of the event as low-energy hadronic activity. This is included in the underlying event (see Section 2.2.2). To describe the interaction between the two protons, the different combinations of interacting partons need to be considered. For this purpose, the parton distribution functions (PDFs) $f(x_i)$ provide the probability to find a particular parton within the proton carrying a fraction of its total momentum $x_i P = p_i$, where P is the total momentum of the proton and p_i is the momentum of a parton i. The PDFs depend both on the energy scale (Q^2) and on the flavour of the considered parton. If the protons were formed just by three non-interacting valence quarks, the PDFs would be reduced to a delta-function at $x_i = 1/3$. However, the quarks inside the proton are continuously interacting through the exchange of gluons, which in turn results in a constant exchange of momenta. This smears out the delta-functions. In the end, the PDFs cannot be calculated due to the effects

of non-perburbative QCD. Instead, the PDFs are extracted from global fits to data which are then extrapolated to new energy scales using perturbative QCD. The most common PDF sets are provided by the MSTW [256], the NNPDF [257] and the CTEQ [258, 259] collaborations. One example of PDFs is given in Figure 2.2. The left figure corresponds to a low energy regime, and it can be seen that most of the momenta is carried by the valence quarks. In contrast, at higher energies (right figure), the gluons and the sea quarks also carry an important fraction of the total momenta of the proton.

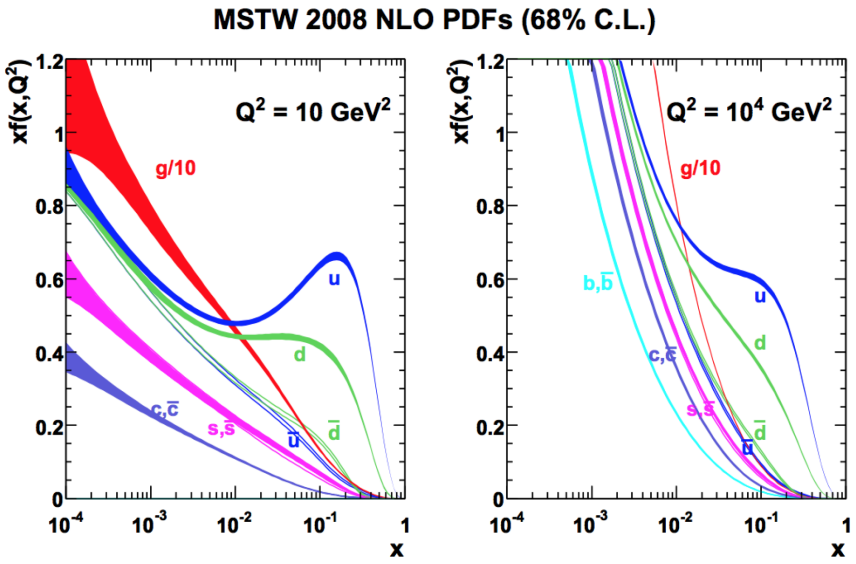


Figure 2.2: Parton distribution functions of a proton at two different energy scales Q^2 = 10 GeV^2 (left) and Q^2 = 10^4 GeV^2 (right) using the MSTW 2008 NLO PDF set [256].

The cross-section of a particular process is calculated taken the input from the PDFs. The interactions can also be categorised in two main groups: *soft* and *hard*. Hard interactions involve parton-parton interactions with large momentum transfer. The cross-section of these processes can be calculated within perturbative theory, and in general are well-understood. In contrast, soft interactions imply a

low momentum transfer and non-perturbative QCD has a much more sizable effect, making much more difficult to calculate the cross-section of these processes.

Hard scattering cross-sections can be calculated using the factorisation theorem. This postulates that the cross-section can be separated in two parts. One part takes care of the parton cross-sections, which are process dependent and calculable within perturbative QCD. The second part, or universal part, takes into account the probability distributions of the partons inside the proton and is modeled with the PDFs. Using the factorisation theorem, the interaction cross-section in hadronic collisions is given by (see Figure 2.3):

$$\sigma(p_1, p_2) = \sum_{i,j} \int dx_1 dx_2 f_i(x_1, \mu_F^2) f_j(x_2, \mu_F^2) \hat{\sigma}_{ij}, \tag{2.1}$$

where $x_1(x_2)$ is the fraction of momenta of the considered parton within p_1 (p_2), and $f_i(x_1, \mu_F^2)$ ($f_j(x_2, \mu_F^2)$) is its associated PDF. The index i (j) runs over all the partons in p_1 (p_2). μ_F is the so-called factorisation scale, which determines the boundary between low and high energy and hence determines the scale at which perturbative calculation becomes valid. Finally, $\hat{\sigma}_{ij}$ is the interaction cross section between the two considered partons. Within perturbative QCD, this can be calculated as:

$$\hat{\sigma}_{ij} = \hat{\sigma}_{ij,LO} + \alpha_s(\mu_R)\hat{\sigma}_{ij,NLO} + \alpha_s^2(\mu_R)\hat{\sigma}_{ij,NNLO} + ..., \tag{2.2}$$

where α_s is the strong coupling constant and μ_R is the renormalisation scale, which cures the ultraviolate divergences in QCD as a result of large momentum in the loops. If one could calculate the entire perturbation series, the μ_F and μ_R parameters would no longer be needed. However, this is unfeasible and the serie is truncated at a particular order. At this point it is important to fix μ_F and μ_R to a particular value. This leads to uncertainties in the calculations which are frequently taken care of by varying these parameters.

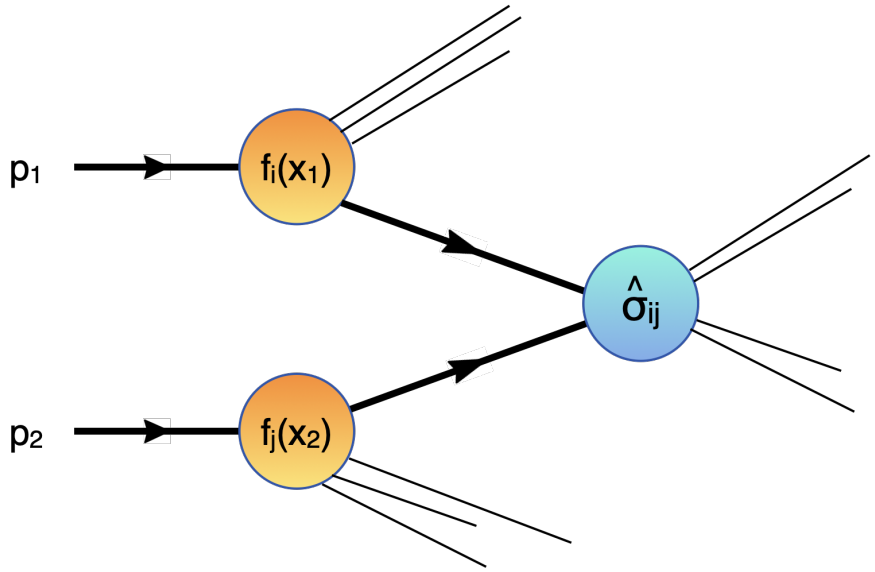


Figure 2.3: Schematic description of a hard process in hadronic collisions using the factorisation theorem.

The hard subprocess, by definition, involves large momentum transfers and therefore the partons involved in it are extremely accelerated. The accelerated coloured partons will emit QCD radiation in the form of gluons. The gluons themselves carry colour charges and can therefore emit further radiation, leading to the so-called *parton showers*. At some point in the evolution of these parton showers perturbation theory becomes invalid and the dynamics enter a non-perturbative phase, which leads to the formation of hadrons (this is the so-called confinement, see Section 1.1.1). These hadrons may be stable and constitute part of the final-state objects, or they may be unstable and decay to stable particles. The parton shower MC event generator algorithms are designed to simulate the final states of high-energy collisions in full detail down to the level of individual stable particles, which are recorded in the detector. The event generation includes all the parton showers, the hadronisation and the decay of unstable hadrons. In addition, these MC event generators are frequently used to simulate the underlying event, as described in the following subsection.

2.2.2 Underlying event

The underlying event was first described by the CDF collaboration [260], and refers to all the activity accompanying the hard scattered processes. This includes all beam-beam remnants, multiple parton interactions (MPI) within one collision and initial and final state radiation (ISR/FSR). It is necessary to provide a good modelling of the underlying event to enable experimental studies and to perform model comparisons, since one needs to distinguish the direct products from the hard scattering from the remaining part of the event.

In general it is not easy to provide a good modelling of the underlying event due to the several soft interactions that occur in which non-perturbative QCD takes place. There are several underlying event simulators, among which one can find PYTHIA [261], HERWIG [262], or SHERPA [263], each with several tunes of model parameters. As an example, PYTHIA simulates the underlying event assuming that the multiple interactions that take place in association with the hard scattered event are independent and can be treated within Poisson statistics. It also assumes that perturbative QCD can be used above a given threshold $p_{\mathrm{T}}^{\mathrm{min}}$, which becomes a key parameter of the model.

A schematic description of the underlying event in a pp collision can be found in Figure 2.4, including the ISR and FSR, the beam-beam remnants and a possible MPI.

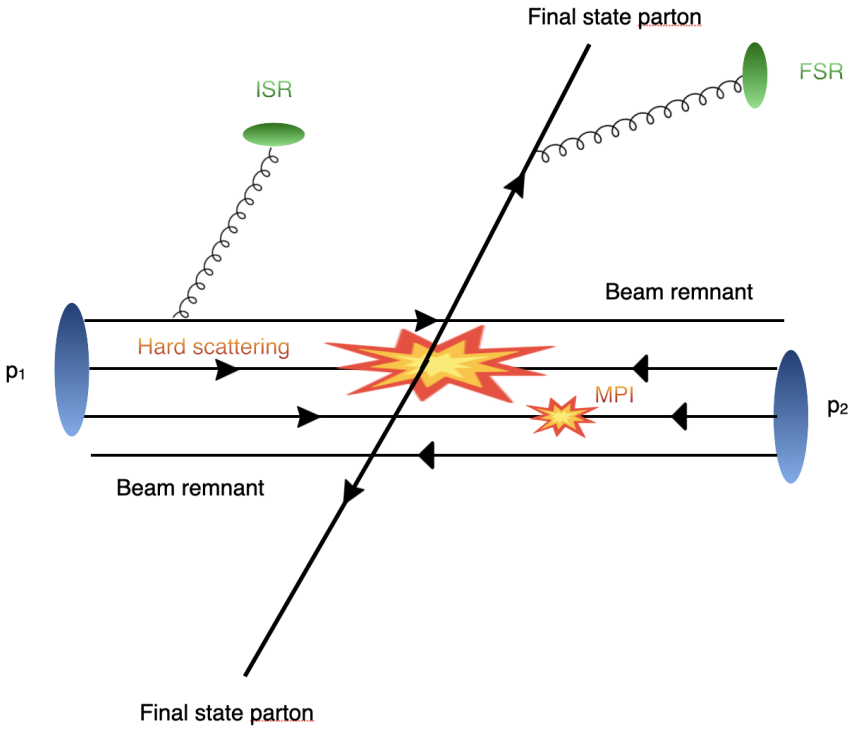


Figure 2.4: Schematic representation of a pp interaction with special emphasise on the underlying event. The collision includes the hard scattering interaction between two partons and the two final state partons resulting from this interaction. The underlying event includes an additional multiple particle interaction, the initial and the final state radiation and the beam remnants.

2.2.3 Luminosity

One important concept in the phenomenology of pp collisions is the luminosity $\mathcal{L}$, which is a measure of the number of collisions that can be produced in a detector per unit area and per unit time:

$$\frac{dN}{dt} = \mathcal{L}\sigma, \tag{2.3}$$

where σ is the production cross-section of a particular process. Since σ is fixed for each process, increasing $\mathcal{L}$ means increasing the associated number of events. This is particularly important for those processes in which the cross-section is very small.

In pp collisions, assuming a Gaussian profile for the colliding beam bunches, the luminosity can be rewritten in terms of the beam parameters as:

$$\mathcal{L} = f_c \frac{N_1 N_2 N_b}{4\pi\sigma_x\sigma_y}, \tag{2.4}$$

where f_c is the crossing frequency of the colliding bunches, N_1 (N_2) is the number of protons in each bunch, N_b is the number of bunches, and σ_x (σ_y) is the Gaussian width of the beam in the horizontal (vertical) plane.

The integrated luminosity $L_{int} = \int \mathcal{L} dt$ corresponds to the integral of the instantaneous luminosity during time. Thus, the total number of collisions of a particular process is given by:

$$N = L_{int}\sigma \tag{2.5}$$

The integrated luminosity during the whole Run-2 yields to a total of 139 fb^{-1} good quality[3] data. The cumulative distributions of the integrated luminosity as a function of time for the 2011-2012 Run-1 and the whole Run-2 data-taking periods are displayed in Figure 2.5.

[3]During data-taking it is important that data are free from integrity issues. This could happen, for example, as a result of a surge in noise from detector electronics or from a high-voltage trip in a detector module. The analysable (or good) data excludes the data collected during these periods of transient detector issues.

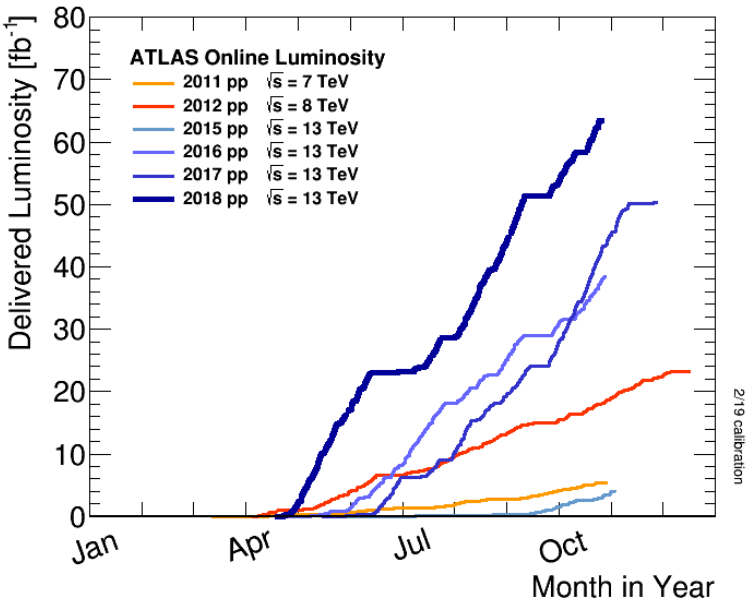


Figure 2.5: Cumulative luminosity versus day delivered to ATLAS during stable beams for high energy pp collisions. Data from 2011-2018 is included in the figure [264].

2.2.4 Pile-up

According to Equation 2.4, there are three possibilities to increase the luminosity in pp collisions:

- Increase the number of protons in each beam.
- Squeeze the width of the beams.
- Increase the number of bunches.

However, this would also imply a larger number of particle interactions occurring per second. The average number of these interactions is referred to as "pile-up". Most of these interactions are uninteresting, secondary or soft interactions, which contaminate the main hard scattered event. Thus, it is important to keep them to the minimum. The first two options from the previous list would

increase the number of pile-up interactions in each bunch crossing (*in-time pile-up*), leading to several interactions at the same time. The latter option would increase the number of pile-up interactions from different bunches (*out-of-time pile-up*). This happens because the spacing of the bunches is lower than the response of the detectors and the different bunches are recorded at the same time.

All things considered, increasing the luminosity comes at the price of increasing the undesired pile-up conditions, which ultimately rely on the beam parameters. The average number of interactions per bunch crossing during the Run-2 can be found in Figure 2.6 for each year of data-taking.

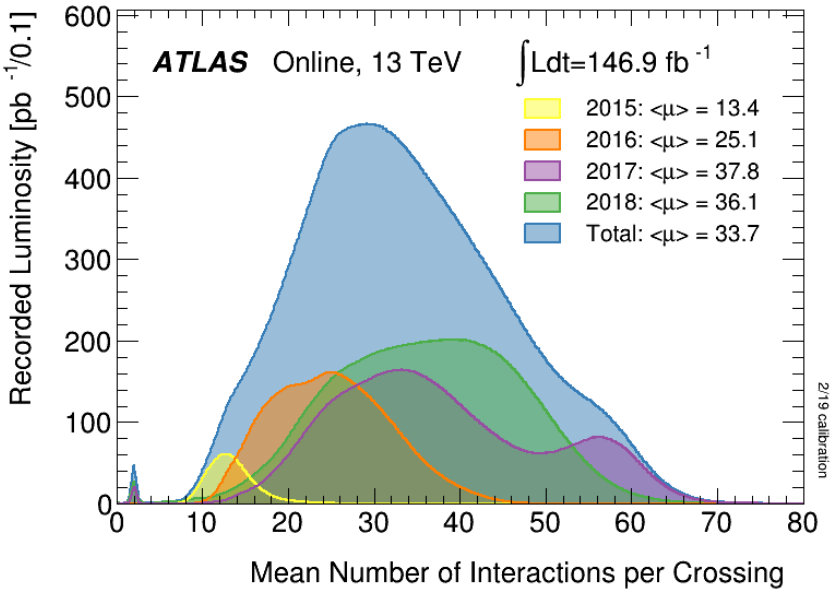


Figure 2.6: Luminosity-weighted distribution of the mean number of interactions per bunch crossing for the 2015-2018 pp collision data. All data recorded by ATLAS during stable beams is shown, together with the total integrated luminosity and the mean $<\mu>$ value for each year [264].

2.3 The ATLAS detector

The ATLAS detector [265] is a multi-purpose detector designed to cover a wide range of physics phenomena: from high precision SM measurements to searches for new physics. The good performance of the ATLAS detector is possible thanks to the ATLAS collaboration, which includes physicists, engineers, technicians and support staff from all around the world. This is one of the largest collaborative efforts ever attempted in science. It involves more than 5000 members, with 3000 scientific authors from 180 institutions in 40 different countries.

This detector is the biggest collider detector ever made, with 25 m of height and 44 m of length, and it is characterised by its cylindrical shape. A schematic view of the ATLAS detector can be seen in Figure 2.7. It is organised in different layers: the innermost layer is the ID, or tracking system, which aims at the reconstruction of the trajectories of the charged particles traversing it. The next layers correspond to the calorimeters. Their purpose is to measure the energy of the particles. There are two types of calorimeters: the electromagnetic calorimeter, which measures mainly the energy of electrons and photons, and the hadronic calorimeter, which measures the energy of strongly interacting particles. Finally, the last layer corresponds to the muon spectrometer (MS). This system aims at the reconstruction of the trajectory of the muons, which are the only particles that escape without stopping in the calorimeters (with the exception of neutrinos). A simplified view of the signal that different types of particles would leave in the different components of the ATLAS detector can be found in Figure 2.8.

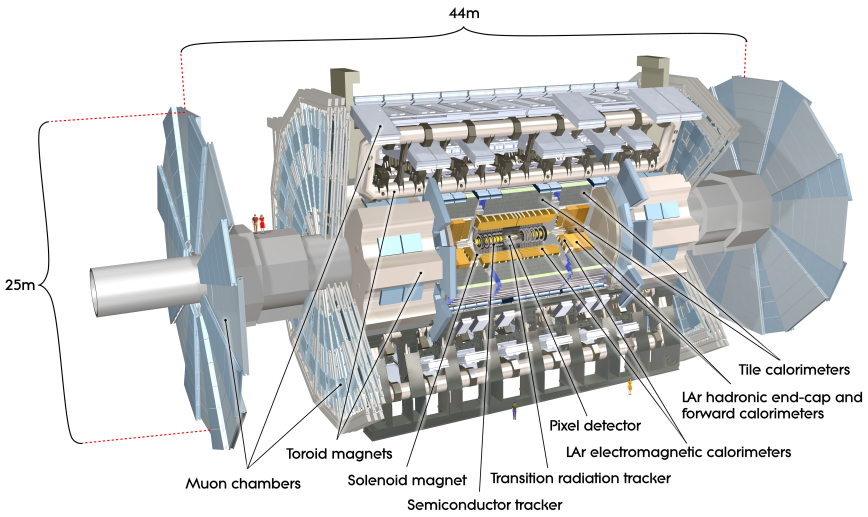


Figure 2.7: Schematic view of the ATLAS detector [266].

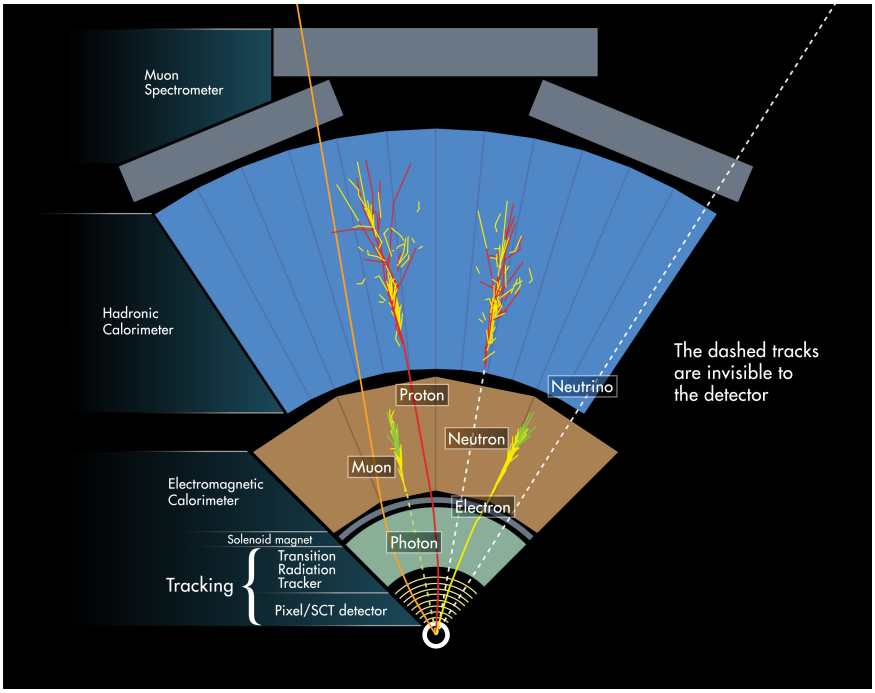


Figure 2.8: Description of the signatures provided by the different types of particles while traversing the different components of the ATLAS detector [267].

2.3.1 ATLAS coordinate system

The position of the different components of the ATLAS detector is provided in the ATLAS coordinate system (x, y, z). The origin of this system is placed in the interaction point. The three axis form a right-handed coordinate system: the z-axis points in the beam direction, the y-axis points upwards, towards the sky, and the x-axis points to the centre of the LHC ring.

However, it is more frequent to express the position of a detector or a particle in terms of polar coordinates (r, ϕ, θ). The r coordinate measures the distance with respect to the origin of the system. The ϕ coordinate determines the angular distance in the plane perpendicular to the beam direction (the x-y transverse plane). Finally, the polar angle θ determines the angular distance with respect to the z-axis.

Instead of the polar angle, the position of a particle in this direction can be expressed in terms of the rapidity y, which is invariant under Lorentz boosts in the z-axis:

$$y = \frac{1}{2}\ln\frac{E + p_z}{E - p_z}, \tag{2.6}$$

where E is the energy of the particle and p_z the longitudinal component of its momentum. At the LHC energies the ultra-relativistic regime is a valid approximation (the masses of the particles can be neglected). Hence, the rapidity can be further approximated to the pseudorapidity η:

$$\eta = -\ln\tan(\theta/2) \tag{2.7}$$

This definition implies that $\eta = 0$ when $\theta = \pi/2$, and $\eta \to \infty$ when $\theta = 0$. However, most of the particles are always produced within $|\eta| < 5$, which corresponds to the detector acceptance. The angular distance between two objects is frequently expressed in terms of $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

Besides, the transverse momentum p_{T} is also preferred. This is calculated in the $x - y$ plane from the transverse energy deposited in the calorimeters. All things considered, in particle physics the linear momentum is represented as $\boldsymbol{p} = (p_{\mathrm{T}}, \eta, \phi)$. The sum of the transverse momenta of all the particles leads to the so-called *missing transverse momentum* $E_{\mathrm{T}}^{\mathrm{miss}}$, as explained in Section 2.5.5.

2.3.2 The Inner Detector

The ID is a spectrometer that reconstructs the curved trajectory of charged particles in a magnetic field. It is the closest component to the interaction point and it is placed surrounding the beam pipe. The main goal of the ID is to reconstruct the momentum of the charged particles traversing it (tracks). As a result, it also provides information on the impact parameters[4] and reconstructs primary and secondary vertices (see Section 2.5.1), which are a key ingredient in b-tagging

[4]The impact parameters determine the distance of a reconstructed track from a charged particle to the perigee, which is the closest point of the track to the global z-axis. These are described in Section 3.1.2.

algorithms (see Section 2.5.4.1). It is characterised by an excellent momentum resolution, and it has a coverage up to $|\eta| < 2.5$. In total, the ID measures 7 m of length and has a diameter of 2.3 m.

The ID is placed within a 2 T solenoid magnetic field in order to bend the trajectory of the particles. From the curvature of the tracks, one can measure the charge and the momentum of the associated particle. There are several subdetectors which constitute the ID. Figure 2.9 provides a schematic view of the different components in the barrel region. The first three subdetectors correspond to the silicon trackers: the Insertable B-Layer (IBL), the pixel detector, and the SemiConductor Tracker (SCT). The last subdetector corresponds to the Transition Radiation Tracker (TRT) straw tubes, filled with gas. Table 2.2 summarises the main features of each subdetector.

Subdetector	Element size [μm]	Intrinsic resolution [μm]
IBL	50×250	8×40
Pixel	50×400	10×115
SCT	80*	17×580
TRT	4000*	130

Table 2.2: Main characteristics of the different components of the ID. The intrinsic resolution is reported along r-ϕ and z except for the TRT, which is insensitive along the z direction. The element sizes in the SCT refers to the spacing of the readout strips, whereas in the TRT they refer to the diameter of the tubes.

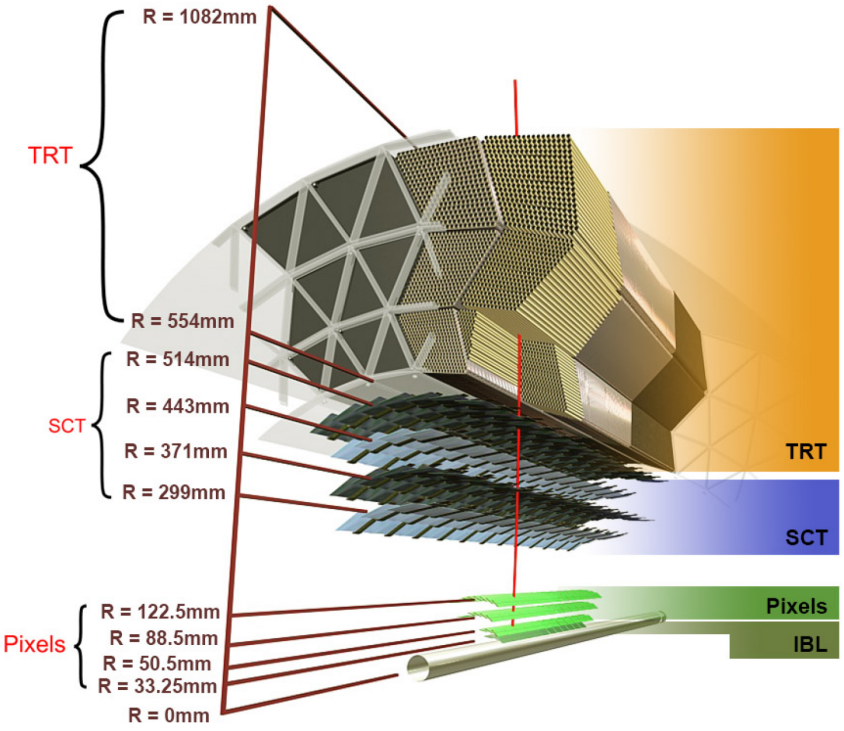


Figure 2.9: 3D visualisation of the barrel of the ID with its different subcomponents and the beam pipe. From innermost to outermost: the IBL, the pixel detector, the SCT and the TRT [268].

2.3.2.1 Insertable B-Layer

The IBL is the innermost component of the ID and it is attached to the beam pipe. It was installed during LS1 and started operations in Run-2, providing one additional measured hit for each track. It contains a total of 280 pixel modules, arranged along 14 azimuthal staves and providing coverage in the $|\eta| < 3$ region. The different staves overlap 1.82° between each other, ensuring a full coverage in the azimuthal direction. A schematic view of the IBL sensors can be found in Figure 2.10.

The insertion of the IBL has increased significantly the tracking performance during Run-2 (in spite of the larger amount of radiation), providing a better impact parameter resolution.

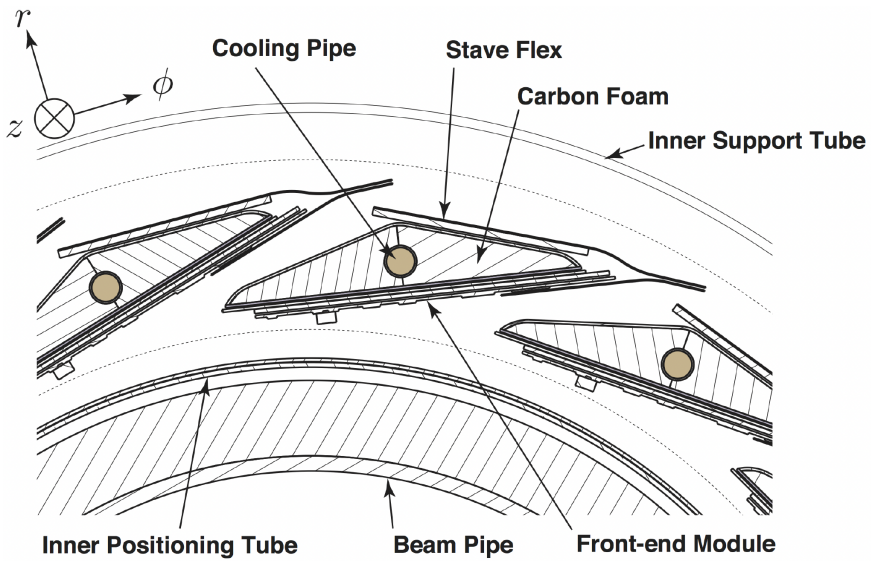


Figure 2.10: Schematic view in the transverse plane of the IBL staves [269].

2.3.2.2 Pixel detector

The pixel detector is made of 1774 identical modules. These are arranged along three barrel layers which surround the beam pipe, and also along two end-caps with three disks each, placed at each end of the barrel. The end-cap disks are set

perpendicularly to the z-axis. A schematic view of the pixel detector can be found in Figure 2.11.

The pixel detector provides full coverage along the azimuthal direction, and covers the region of $|\eta| < 2.5$. It is designed to provide three measured hits for each track.

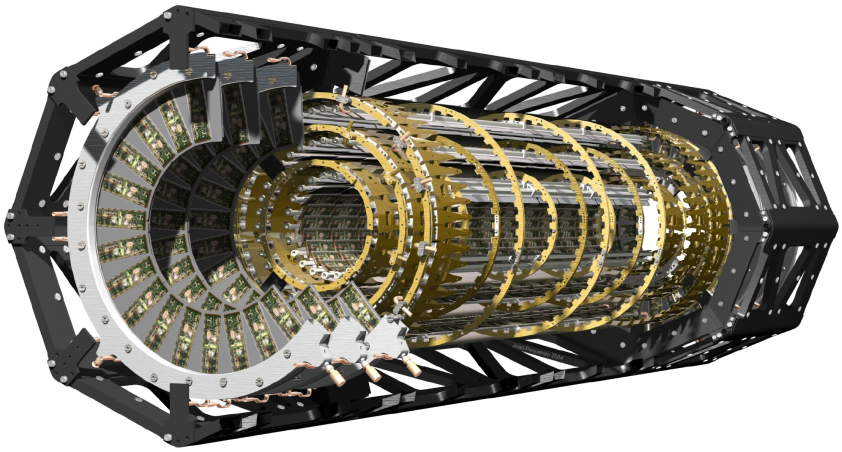


Figure 2.11: Schematic view of the pixel detector [270].

2.3.2.3 SemiConductor Tracker

The SCT has a similar geometry to the pixel detector (see Figure 2.11). It is formed by 4088 modules, arranged in four barrel layers and two end-caps with nine disks each. Each SCT module is composed by two silicon microstrips sensors arranged back-to-back with a deviation of 40 mrad between the two sides to provide a two dimensional measurement ($r - \phi$ and z).

Similarly to the pixel detector, the SCT provides full coverage in the azimuthal direction, and covers the region of $|\eta| < 2.5$. It provides an average of eight hits (four space-points) for each measured track.

2.3.2.4 Transition Radiation Tracker

The TRT is made by a total of 350848 straw tubes, divided in the barrel and the two end-caps. These are placed parallel to the beam pipe for the barrel and radially in the end-caps. Each tube measures 4 mm of diameter and is filled with gas[5], in contrast with the silicon modules. The hits in the TRT are measured through the ionisation of this gas. In this context, the tube works as a cathode, and it has a anode wire in the centre. When a particle traverses a tube, it produces electrons which drift towards the anode generating an electric signal.

The spaces between the straws are filled with polymer fibres in the barrel and foils in the end-caps. When a particle traverses the transition region, it produces a transition radiation which is different depending on the particle type. This radiation may be absorbed by the TRT gas leaving an increased read-out signal. In particular, the emission of transition radiation is much more likely for electrons than for pions with the same momentum, and thus this information can be used in order to discriminate between these particles.

In total, the TRT provides approximately 30 hits for each measured track. The barrel covers the $|\eta| < 1$ region, and the combined barrel and end-caps cover up to $|\eta| < 2$. The TRT provides a measurement in the r-ϕ plane, but it is insensitive to the z direction.

2.3.3 Calorimeters

The next layer of the ATLAS detector is formed by the calorimeters. The calorimeters are responsible for stopping the particles and measure their energy. They are composed by two different materials arranged in alternating layers:

- **Passive material**: The passive material, also named absorber, is made of a dense material to fully stop the traversing particles. When a particle goes through the passive material, it interacts with it, producing a shower of additional particles.

[5] During Run-1, this was a gas mixture of 70% Xe, 27% CO_2 and 3% O_2. In order to mitigate leaks of this Xe-based active gas mixture, the gas system was modified to use a cheaper Ar-based gas mixture in some channels for Run-2 [271].

- **Active material**: The active material detects the particles from the shower originated in the absorber, and transforms them into signal.

The calorimeters are divided in several components. The electromagnetic calorimeter (ECAL) measures the energy from electrons and photons. The hadronic calorimeter (HCAL) measures the energy of the jets. Finally, the forward calorimeter (fCAL) measures the energy of particles produced in the forward direction. The substructure of the calorimeters is depicted in Figure 2.12. The only SM particles that escape from the calorimeter system are muons and neutrinos. Muons are detected in the MS (see Section 2.3.4). Concerning neutrinos, their energies are measured from the reconstruction of the missing transverse momentum $E_{\mathrm{T}}^{\mathrm{miss}}$, computed as the negative sum of all the energy deposits in the transverse plane (see Section 2.5.5).

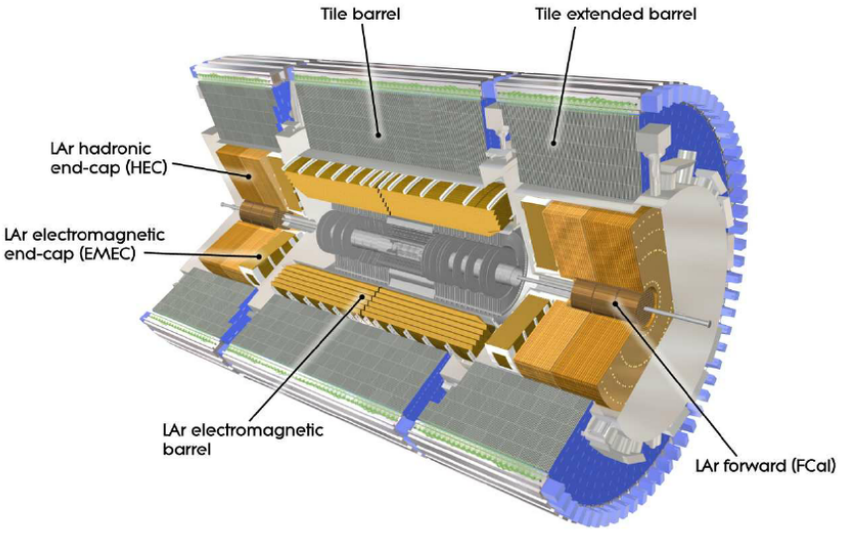


Figure 2.12: Illustration of the different components of the ATLAS electromagnetic and hadronic calorimeters [272].

2.3.3.1 Electromagnetic calorimeters

The ECAL covers the region of space $|\eta| < 3.2$ with a design energy resolution $\frac{\sigma(E)}{E} = \frac{10\%}{\sqrt{E}} \bigoplus \frac{170\text{MeV}}{\sqrt{E}} \bigoplus 0.7\%$ [273]. It is divided in a barrel ($|\eta| < 1.475$) and two end-caps ($1.375 < |\eta| < 3.2$). The barrel part is split in two identical halves separated by a small gap at $z = 0$, and the end-cap is composed by two coaxial wheels. The passive material in the ECAL is made by lead, arranged in an accordion shape to ensure a complete azimuthal coverage. The active material is liquid argon. The calorimeters which use this kind of active material are named LAr calorimeters. The ECAL is covered by a cryostat to ensure the necessary low temperatures to operate.

2.3.3.2 Hadronic calorimeters

The HCAL also covers the region of space $|\eta| < 3.2$ with a design energy resolution $\frac{\sigma(E)}{E} = \frac{52.9\%}{\sqrt{E}} \bigoplus 5.7\%$ [273], and it is also divided into a barrel ($|\eta| < 1.7$) and two end-caps ($1.5 < |\eta| < 3.2$). In the barrel, the absorber is made of steel, whereas the active material are tiles of plastic scintillator. Thus, this part is called the Tile Calorimeter (TileCal). As seen in Figure 2.12, it is formed by a long barrel and two extended barrels. The end-caps are LAr calorimeters and are split in 2 wheels each, sharing the same cryostat as for the ECAL end-caps.

The HCAL is very thick to reduce the punch-through effect, which takes place when the hadronic showers propagate to the MS. This also guarantees a high precision measurement of the $E_{\text{T}}^{\text{miss}}$.

2.3.3.3 Forward calorimeters

The fCALs cover the forward region ($3.1 < |\eta| < 4.9$). They are placed inside the LAr calorimeter end-caps, surrounding the beam pipe. Each fCAL is composed by three modules. One is made of copper, and it aims to measure the energy of electrons and photons. The other two modules are made of tungsten, and aim to measure the energy of the hadrons.

The forward coverage provided by the fCALs is required for many physics tasks, such as the reconstruction of the $E_{\text{T}}^{\text{miss}}$ or the forward-jet tagging.

2.3.4 Muon spectrometer

The MS is the outermost component of the ATLAS detector, and it is characterised by its high-precision tracking and triggering. It aims to identify and reconstruct the trajectories of the muons, which leave a very small signal in the calorimeters.

A schematic view of the MS can be found in Figure 2.13. This detector covers the region $|\eta| < 2.7$, and provides a full azimuthal coverage. It is formed by three cylindrical barrel layers and six end-cap disks. In order to reconstruct the trajectories of the muons, the MS relies on the large superconducting magnet system, which bends the trajectories of the particles in a similar way to the ID.

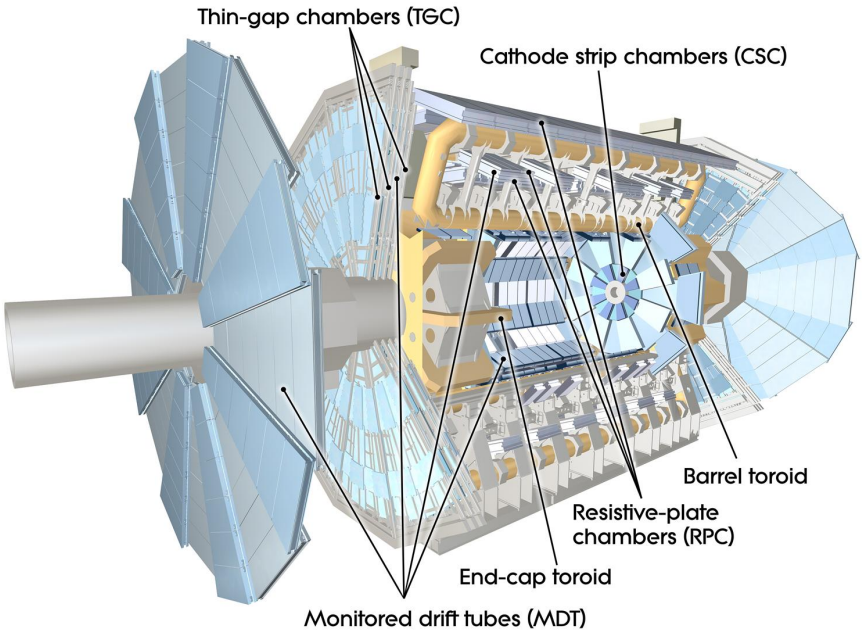


Figure 2.13: Cut away view of the muon spectrometer, including the different types of detectors. The toroid magnets are also displayed [272].

There are different technologies in the subcomponents of the MS. Two of them are used for the reconstruction of the tracks, whereas the other two provide a trigger signal in case a muon is detected.

- **Monitored Drift Tubes**: These detectors provide high precision measurements of the muon tracks in the bending direction. They are placed in the three barrel layers and in two of the end-cap disks.
- **Cathode Strip Chambers**: These components are placed in the innermost barrel layer, in front of the monitored drift tubes. They have a higher granularity to withstand the demanding rate and the higher background conditions in the $2 < |\eta| < 2.7$ region.
- **Resistive Plate Chambers** and **Thin Gap Chambers**: These detectors provide triggers in the barrel and end-cap regions, respectively. In addition, they provide a measurement of the muon tracks in the orthogonal direction to the bending direction.

2.3.5 The magnet system

The ATLAS superconducting magnet is one of its most important technological milestones. It is a unique hybrid system which provides a magnetic field both in the ID and the MS.

The magnetic field in the ID is provided by a superconducting solenoid magnet which is aligned with the beam axis, reaching a central field of 2 T.

In the MS, the magnetic field is provided by a barrel toroid and two end-cap toroids, each made by eight superconducting coils. The magnetic field peaks in the barrel (end-cap) at 3.9 (4.1) T.

2.4 The ATLAS trigger system

Since the amount of data generated by the ATLAS detector is huge, it is unfeasible to store all the collisions taking place in the detector. The ATLAS trigger system [274–276] performs an online selection of events and decides which events contain interesting physics phenomena. The role of the trigger system is thus crucial, since at the LHC the bunch crossing rate was 40 MHz during Run-2, with an average of 25 interactions per bunch crossing. The data acquisition system

(DAQ) channels the data from the detectors through the trigger system and to a data storage system for offline analysis.

To reduce the flow of data to manageable levels, ATLAS used a specialised two-level online event trigger system during Run-2: the Level 1 (L1) trigger and the High Level Triggers (HLT).

The L1 is a hardware-based trigger, which is integrated in the detectors. It uses the information from the different components of the ATLAS detector to decide if an event is stored or rejected, with a processing time of 2.5 μs. The output of the L1 trigger includes Regions of Interest (RoI), which include the position and the transverse momentum of candidate objects. The L1 trigger can save at most 100 000 events each second, which are then transferred to the HLT.

The HLT is a software-based trigger which refines the analysis of the L1 trigger. It conducts a very detailed analysis either by performing overall examination of the whole event for selected layers of the detector, or by utilising the data in the selected RoI. About 1000 events per second are selected by the HLT and are fully assembled into an event record.

2.5 Object reconstruction at the ATLAS detector during Run-2

The different subdetectors that form the ATLAS experiment have been described in Section 2.3. After the trigger selection, the raw data is analysed to build the physics objects which are used in the different analyses: leptons, photons, jets and missing transverse momentum. This process is called reconstruction.

An accurate object reconstruction is essential for any physics analysis. This section begins with a detail description of the ID tracks reconstruction procedure, since these are a key component for the alignment procedure described in Chapter 3. Afterwards, the different reconstruction techniques associated with the physics objects used in the analyses presented in this work are reported.

2.5.1 Tracks and vertices

As explained in Section 2.3.2, the trajectories of charged particles (tracks) are reconstructed in the ID. The ID tracking is performed in two sequences, the main *inside-out* track reconstruction and the consecutive *outside-in* tracking [277, 278].

The inside-out tracking can be divided in different steps. Firstly, a connected component analysis (CCA) [279] groups pixels and SCT strips in a given sensor, where the deposited energy yields a charge above threshold, with a common edge or corner into clusters. In dense environments such as those at the LHC Run-2, where the spatial separation between charged particles traversing the sensor is only a few pixels, it is frequent that one cluster includes energy deposits from multiple particles. These are the so-called *merged* clusters. However, this definition relies on the "truth information", only available in simulation. The track reconstruction algorithm, explained in the following lines, classifies those clusters used in the reconstruction of multiple tracks in two main categories: *identified as merged* and *shared*. Ideally, all clusters identified as merged are merged clusters, and vice versa. Shared clusters are those which are used in multiple reconstructed tracks but are not sufficiently compatible with the properties of a merged cluster to be identified as merged.

From these clusters, three-dimensional measurements referred to as *space-points* are created. In the next step, these space-points are combined in groups of three to form the track seeds. This approach maximizes the possible number of combinations while still allowing a first crude momentum estimate. Different criteria are placed on these seeds to maximize purity, i.e. the fraction of seeds that result in good-quality tracks. The most stringent requirements come from the impact parameters of the track seeds and from the seed-type: SCT-only, pixel-only and mixed-detector seeds are labelled in increasing order of purity. A combinatorial Kalman filter [280] is then applied for the track-fit. This provides track candidates which are built from the chosen seeds by incorporating additional space-points from the remaining layers of the pixel and SCT detectors which are compatible with the preliminary trajectory.

Afterwards, an ambiguity-solver is applied to perform a stringent selection on the track candidates. This compares and rates the individual tracks by assigning a relative *track score* to each track. This is based on simple measures of the track quality. Clusters assigned to a track increase the track score according to the intrinsic resolution and the expected cluster multiplicities in the different subdetectors. Holes, which are defined as intersections of the reconstructed track trajectory with a sensitive detector element that does not contain a matching cluster, reduce the score. The χ^2 of the track-fit is also considered to penalize candidates with a poor fit. Finally, the track momentum is also taken into account to promote energetic tracks.

Track candidates are then processed in descending order of track score to give preference to high-scoring tracks. At this stage, the ambiguity solver deals with those clusters which are used for multiple track candidates. For this purpose, a neural network (NN) is used to identify merged clusters. These are natural in dense environments and can be associated to multiple tracks. However, shared clusters are a strong indicator of incorrect assignments and they must be limited. A track candidate is thus compared to those tracks previously accepted and in case there are more than two tracks in a shared cluster or a track contains more than two shared clusters, this cluster is removed and the track is re-scored.

The ambiguity solver rejects those tracks which fail to fulfill minimum quality requirements. Then, the selected track candidates go through a high-resolution fit which uses all the available information. The fitted tracks which pass the ambiguity solver without modifications are added to the final track collection. This procedure is schematically depicted in Figure 2.14.

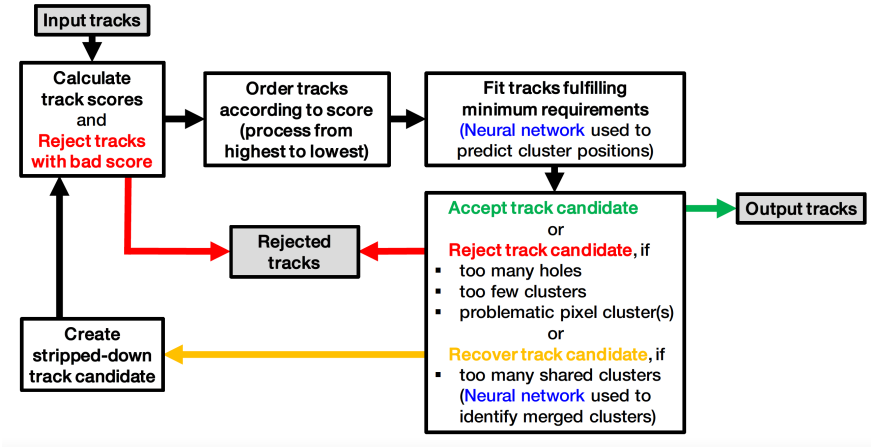


Figure 2.14: Sketch of the flow of tracks through the ambiguity solver [277].

Lastly, the output tracks are extended if there is any compatible measurement in the TRT associated to it. This concludes the inside-out tracking.

A second outside-in tracking is then performed to find small track segments in the silicon detectors which were missed. It starts with a global pattern recognition in the TRT. This produces track segments, preventing hits assigned to tracks used in the inside-out procedure. These segments are followed back into the silicon detectors and matched with the missed track segments. The resulting tracks are added to the track collection resulting from the inside-out procedure.

Another important concept used in object reconstruction are the vertices. The primary vertex determines the interaction point in which the hard interaction takes place. They are reconstructed using the information from the different tracks. For a track to be considered in the reconstruction of a vertex, they must pass quality criteria [281]. For instance, the transverse momentum p_{T} of the different tracks needs to be larger than 400 GeV, they must have a hit in the IBL and they must contain no pixel holes. The reconstruction of primary vertices is produced in two different steps. The initial is called vertex finding, and consists on the association of tracks to vertex candidates. The second step consists on an iterative fit to determine the final position of the vertex. The output of these algorithms is a

three-dimensional point, the reconstructed vertex, and its associated covariance matrix.

2.5.2 Electrons

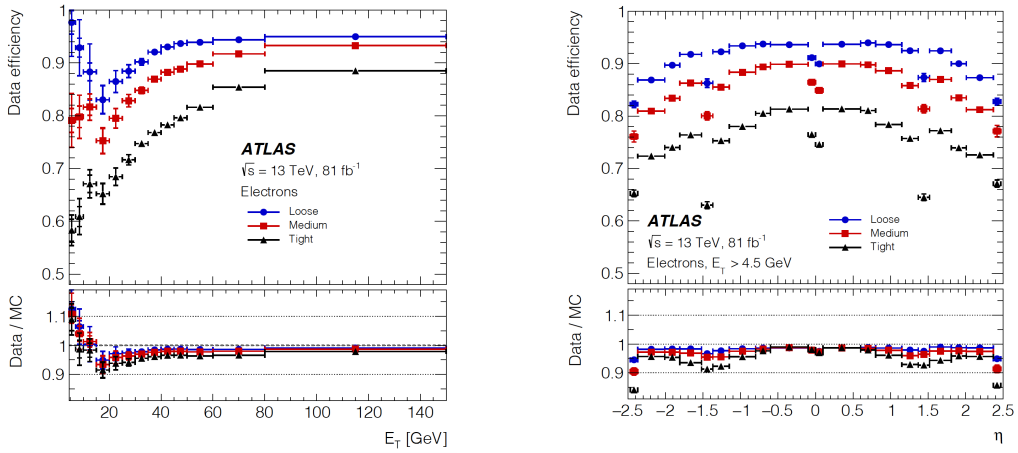
Electron candidates are reconstructed from energy deposits in the ECAL matched to ID tracks, with $|\eta| < 2.5$ [282]. These energy deposits, named superclusters, are dynamic clusters whose size can be adapted as needed to recover energy from bremsstrahlung photons. A topological cell clustering algorithm [283] is used for building these superclusters. Firstly, this algorithm selects clusters of energy deposits measured in topologically connected ECAL cells, denoted topo-clusters, which are matched to ID tracks. These topo-clusters are tested for use as seed cluster candidates, which form the basis of superclusters. For a topo-cluster to become an electron supercluster seed, it is required to have a minimum transverse energy E_{T} of 1 GeV and must be matched to a track with at least four hits in the silicon tracking detectors. Then, topo-clusters near the seed candidates are identified as satellite cluster candidates, which may emerge from bremsstrahlung radiation. These satellite clusters are added to the seed candidates to form the final superclusters if they satisfy the necessary selection criteria: a cluster is considered a satellite either if it falls within a window of $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ around the seed cluster barycentre, or if it is within a window of $\Delta\eta \times \Delta\phi = 0.125 \times 0.300$ around the seed cluster barycentre and its best-matched track is also the best-matched track for the seed cluster. Next, after applying initial position corrections and energy calibrations, the supercluster-building algorithm matches tracks to the electron superclusters.

Electrons can originate either from the hard-scattering event, such as decay products of the W, Z and Higgs bosons (prompt electrons), or as the weak decay products of secondary particles whose lifetimes are relatively long, such as b-hadrons (non-prompt electrons). The identification of prompt electrons relies on a likelihood discriminant constructed from quantities measured in the ID, the calorimeter and the combined ID and calorimeter. These quantities are chosen according to their ability to discriminate prompt isolated electrons from

energy deposits from hadronic jets, from converted photons and from non-prompt electrons. They exploit the properties of the primary electron track, the lateral and longitudinal development of the electromagnetic shower in the ECAL, and the spatial compatibility of the primary electron track with the reconstructed supercluster. Fixed values of the likelihood discriminant correspond to different operating points: *tight*, *medium* and *loose* (at increasing signal efficiency). Tight is the most stringent category, whereas loose is much more relaxed [284].

The energy of the electrons is estimated from the energy deposits in the ECAL. This energy is further calibrated to account for residual differences between data and simulation [282, 285]. These include the intercalibration of the different calorimeter layers, corrections for energy shifts induced by pile-up, corrections to improve the uniformity of the energy response and an adjustment of the overall energy scale, which are all applied to data. In addition, a correction to account for the difference in energy resolution between data and simulation is applied to the simulation. The associated uncertainties on this calibration step are propagated to the different analyses.

The experimentally determined electron spectra must be corrected for the selection efficiencies, which include trigger selection and particle isolation, identification and reconstruction [282, 284]. These efficiencies are determined for both data and simulation using the tag-and-probe method [286] in $Z \to e^+e^-$ and $J/\psi \to e^+e^-$ events. This method selects unbiased samples of electrons (*probes*) by using strict selection requirements on the second object (*tags*) produced from the decay of the resonant particle. From these two objects one can measure the selection efficiency as the ratio between the number of probe electrons passing the associated selection to the number of probe electrons. In order to achieve reliable results, the simulated events need to be corrected to reproduce as much as possible the efficiencies measured in data. This is achieved by applying a multiplicative correction factor, which is defined as the ratio of the efficiency measured in data to that determined from simulation. Figure 2.15 shows the electron identification efficiencies as a function of the E_{T} and η. The lower panel contains the data-to-MC ratio, which are the scale factors applied to the simulated events. The uncertainties on these scale factors are coherently propagated for each analysis.



(a) (b)

Figure 2.15: The electron identification efficiency in $Z \rightarrow e^+e^-$ events in data as a function of (a) E_{T} and (b) η for the loose, medium and tight operating points. The efficiencies are obtained by applying data-to-simulation efficiency ratios measured in $J/\psi \rightarrow e^+e^-$ and $Z \rightarrow e^+e^-$ events to $Z \rightarrow e^+e^-$ simulation. The inner uncertainties are statistical and the total uncertainties are the statistical and systematic uncertainties in the data-to-simulation efficiency ratio added in quadrature. For both plots, the bottom panel shows the data-to-simulation ratios [282].

2.5.3 Muons

Muon candidates with $|\eta| < 2.5$ are reconstructed by combining the information from the ID and the MS as well as the calorimeters [287]. In the MS, tracks are reconstructed by combining hits into local track *segments* by means of a Hough transform [288]. These segments are subsequently combined into track candidates and a fit is performed to obtain the muon trajectory through the magnetic field. Depending on which subdetectors are used in the muon reconstruction, this leads to different types of muons:

- Combined muons: These muons are identified by matching MS tracks to ID tracks and performing a combined track fit based on the ID and MS hits, taking into account the energy loss in the calorimeters.
- Inside-out muons: An *inside-out* algorithm is used to reconstruct this kind of muons. This algorithm extrapolates ID tracks to the MS and searches for at least three aligned MS hits to be used in a combined track fit.
- Muon-spectrometer extrapolated: This type of muons appear when a MS track cannot be matched to an ID track. In these cases, the parameters of the MS track are extrapolated to the beamline and used to define the reconstructed muon.
- Segment-tagged muons: These muons are identified by extrapolating the ID tracks to the MS to search for matching segments. A segment-tagged muon is identified in case an ID track is successfully matched to at least one MS segment, and the muon parameters are taken directly from the ID track fit.
- Calorimeter-tagged muons: In this case muons are identified extrapolating ID tracks through the calorimeters to search for energy deposits consistent with a minimum-ionising particle[6]. In case there is a match, a calorimeter-tagged muon is identified and the muon parameters are again taken from the ID track fit.

[6] A minimum ionising particle is a particle whose mean energy loss rate through matter is close to the minimum.

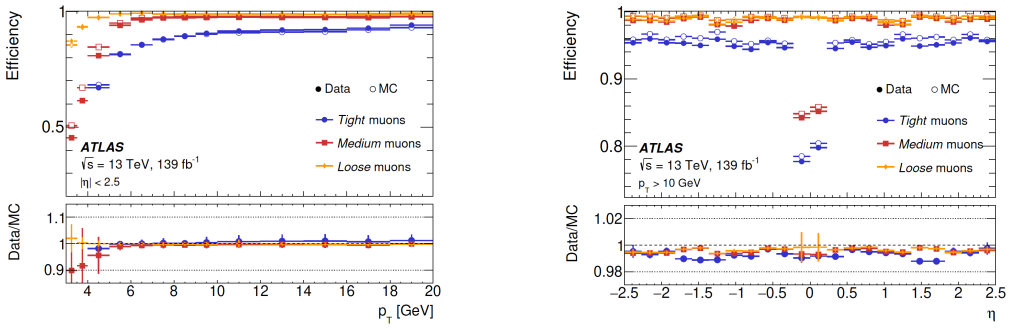
Prompt-muons are identified applying a set of requirements on the number of hits in the ID and the MS, on the track fit properties and on variables that test the compatibility of the individual measurements in the two systems. The stringency of these requirements leads to three main working points: *tight*, *medium* and *loose*. Similarly to electrons, tight muons satisfy more stringent requirements, whereas the requirements for loose muons are much more relaxed. Two additional working points are designed for extreme phase space regions: the *high-p_T* working point, which ensures an optimal momentum measurement for muons with $p_{\mathrm{T}} > 100$ GeV, and the *low-p_T* working point, which targets muons that are less likely to be reconstructed as full tracks in the MS due to their low momentum.

A set of corrections are applied to the simulated muon momentum scale and resolution to account for the differences between data and simulation, studied using $Z \to \mu^+\mu^-$ and $J/\psi \to \mu^+\mu^-$ decays [289]. After applying these corrections, data and simulation agree to the per mille level for the muon momentum scale and to the percent level for the muon momentum resolution. The uncertainties on these corrections are propagated to the different analyses.

The tag-and-probe is used in $Z \to \mu^+\mu^-$ and $J/\psi \to \mu^+\mu^-$ events to measure the reconstruction, identification, isolation and vertex association efficiencies. Probe muons with different definitions and requirements are used to measure the various efficiencies, as described in Ref. [287]. The level of agreement between the efficiency measured in data and simulation for a given algorithm is assessed via the ratio of these two numbers. These ratios are applied to the simulation as correction factors to match data. Figure 2.16 shows the muon reconstruction and identification efficiency for the three main working points. The uncertainties on the correction factors are coherently propagated for each analysis.

2.5.4 Jets

Jets are originated from the hadronization of quarks or gluons produced in the hard scattered event. These produce a spray of stable or nearly-stable particles that is known as a jet. Final-state particles can be clustered by a jet reconstruction algorithm and the jet properties can be approximately mapped onto the partons.



(a) (b)

Figure 2.16: Muon reconstruction and identification efficiencies for the loose, medium, and tight criteria. (a) shows the efficiencies measured in $J/\psi \to \mu^+\mu^-$ events as function of p_{T}, whereas (b) shows the efficiencies measured in $Z \to \mu^+\mu^-$ events as a function of η for muons with $p_{\mathrm{T}} > 10$ GeV. The predicted efficiencies are depicted as open markers, while filled markers illustrate the result of the measurement in collision data. When not negligible, the statistical uncertainty in the efficiency measurement is indicated by the error bars. The panel at the bottom shows the ratio of the measured to predicted efficiencies, with statistical and systematic uncertainties [287].

The primary jet definition used in the majority of physics analysis presented by ATLAS is the anti-k_t algorithm [290]. The inputs to the algorithm include calorimeter energy deposits, charged-particle tracks or algorithmic combinations of these two. The anti-k_t algorithm is based on the concept of distances d_{ij} between two objects, and distances d_{iB} between an object and the beam:

$$d_{ij} = \min(k_{ti}^{-2}, k_{tj}^{-2}) \frac{\Delta_{ij}^2}{R^2} \tag{2.8a}$$

$$d_{iB} = k_{ti}^{-2}, \tag{2.8b}$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ and k_{ti}, y_i and ϕ_i correspond to the transverse momentum, the rapidity and the azimuthal angle of the object i, respectively. Finally, the radius parameter R is the radius of the cone and determines the size of the jets.

The algorithm proceeds as follows: firstly it identifies the smaller distance from all the possible combinations of d_{ij} and d_{iB}. If it is a d_{ij} distance, then the two objects i and j are merged together. In this way, soft components are merged to the hard ones. If it is a d_{iB}, then the entity i is called a jet and it is removed from the list of entities. The algorithm stops when all entities are assigned to a particular jet.

Depending on the radius parameter R, the jets can be classified in two groups: small-R jets with $R = 0.4$ and large-R jets with $R = 1.0$ (see Figure 2.17). Small-R jets are frequently found in the final state topology of many signal processes targeted in different ATLAS analyses, as well as in the various background processes which mimic these signals. In contrast, large-R jets are used more specifically to reconstruct the hadronic decay products of massive particles that are sufficiently boosted, so that their decay products are close together. A description of the two different jet categories and their particularities is given in the following.

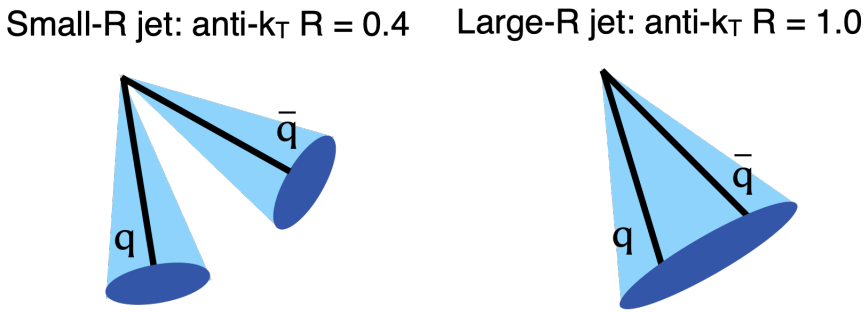


Figure 2.17: Illustration of the two different jet topologies: small-R jets (left) and large-R jets (right).

2.5.4.1 Small-R jets

The presence of small-R jets in the targeted final state of an ATLAS analysis is quite frequent. Therefore, they will be referred to simply as jets in the following. Jets are reconstructed from localised energy deposits in the ECAL and HCAL, from charged-particle tracks or from algorithmic combinations of the latter two [291]. The calorimeter cells are first clustered into three-dimensional, massless, topological clusters (topo-clusters) using a nearest-neighbour algorithm [283]. Cells are added to a topo-cluster if their energy is larger than a given threshold which takes into account the expected noise and controls the growth of the topo-cluster. The energy of the resulting topo-clusters is defined at the electromagnetic scale using the local cell weighting (LCW) scheme [283]. This weighting scheme accounts for the non-compensating response of the calorimeter[7], out-of-cluster energy, and for energy deposited in the dead material of the detector. Only positive-energy topo-clusters are used as inputs to the jet reconstruction. Finally, a correction is applied to each topo-cluster to account for the position of the primary vertex, which is expected to be the origin of hard scattering jets.

These topo-clusters are used as input to the anti-k_t algorithm. Jets reconstructed using only these topo-clusters are referred to as *topo-jets*.

[7]The signal from a hadron in the calorimeters is smaller than the signal of an electron or a photon with the same energy.

Hadronic final-state measurements can be improved by making use of the particle-flow (PF) algorithm [292], which combines the measurements of the calorimeters with the reconstructed tracks in the ID. This exploits the better momentum resolution of the tracker, especially in the low-energy regime. The energy deposited in the calorimeter by charged particles is subtracted from the topo-clusters and replaced by the momenta of the tracks that are matched to those topo-clusters. The resulting jets exhibit improved energy and angular resolution, reconstruction efficiency and pile-up stability compared to calorimeter jets [292].

A schematic view of the PF algorithm can be seen in Figure 2.18. Firstly, well-measured tracks are selected and the algorithm tries to match each track to a single topo-cluster in the calorimeter. The expected energy in the calorimeter, deposited by the particle that also created the track, is computed based on the topo-cluster position and the track momentum. The algorithm also takes into account the probability for a particle to deposit its energy in several topo-clusters. Thus, it decides if it is necessary to include additional topo-clusters to recover the full energy measured from the ID track. Next, the expected energy deposited in the calorimeter is subtracted from the selected topo-clusters. Finally, if the remaining energy of the system is consistent with the expected shower fluctuations of single particles signal, the topo-cluster remnants are removed.

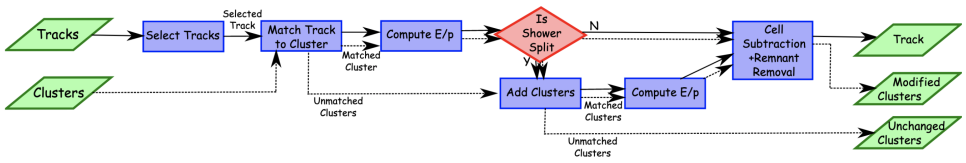


Figure 2.18: Flow chart of how the particle flow algorithm proceeds, starting with track selection and cluster identification and continuing until the energy associated with the selected tracks has been removed from the calorimeter. At the end, charged particles, topo-clusters which have not been modified by the algorithm, and remnants of topo-clusters which have had part of their energy removed remain [292].

The positive energy topo-clusters surviving the energy subtraction step and the selected tracks, which must satisfy $|z_0 \sin\theta| < 2$ mm to suppress the effects of

pile-up, are used as inputs to the jet reconstruction using the anti-k_t algorithm. Jets reconstructed using the PF algorithm are referred to as *PF jets*.

The jet enery scale (JES) is then calibrated in a sequential procedure, described in Figure 2.19. Firstly, a set of corrections are applied to remove the excess energy from pile-up interactions. This is done in two steps: a correction based on the jet area and transverse momentum density, and a residual correction derived from MC simulation. Then, the *absolute* JES and η calibration uses the information from truth jets obtained from dijet MC simulated events to correct the reconstructed jet four-momentum to the particle-level energy scale. This takes into account the non-compensating calorimeter response, possible energy losses in dead material, out-of-cone effects and biases in the jet η reconstruction. Next, the *global sequential calibration*, derived from dijet MC events, is applied to improve the jet p_{T} resolution and the associated uncertainties. This takes into account the variations in the response which may arise from jet-to-jet depending on the flavour and energy distribution of the constituent particles, their transverse distribution, and the fluctuations of the jet development in the calorimeter. Finally, a residual *in situ* calibration is applied to correct for remaining differences between data and MC simulation. These differences are caused by imperfect simulation of both the detector materials and the physics processes involved, which include the hard scattering and underlying event, jet formation, pile-up, and particle interactions with the detector. The jet response is measured in both data and simulation using well-measured reference objects, and the ratio between data and simulation is applied as an additional correction in data. This *in situ* calibration is performed in three sequential steps: Firstly, an η intercalibration corrects the energy scale of forward jets to match those of central jets using the p_{T} balance in dijet events. Secondly, Z+jet and γ+jet analyses use the transverse momentum of a calibrated Z boson or photon to balance the hadronic recoil in an event [293]. Finally, high-p_{T} jets are calibrated using well-calibrated low-p_{T} jets in the multijet balance analysis [294].

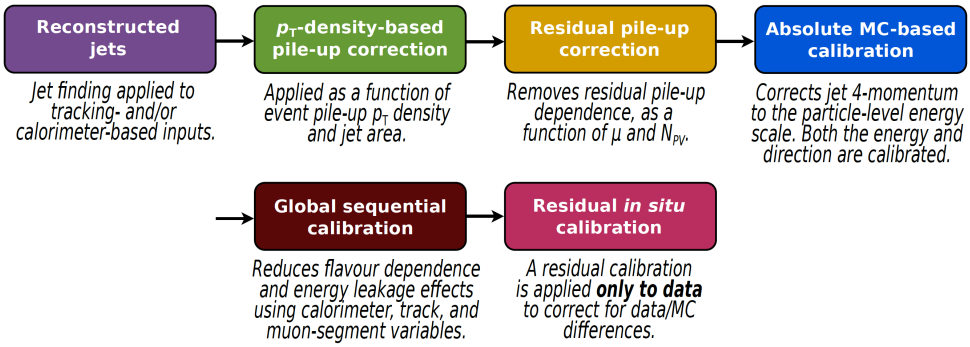
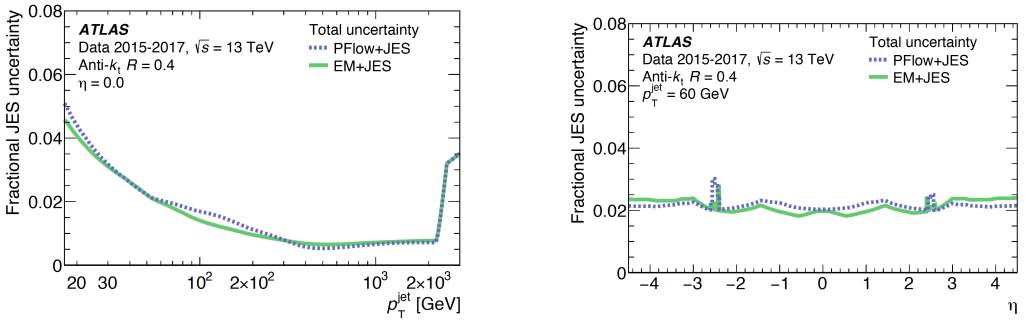


Figure 2.19: Stages of jet energy scale calibrations. Each one is applied to the four-momentum of the jet [291].

The JES calibration is affected by several uncertainties, which need to be taken into account in the different analyses. Figure 2.20 shows the total JES uncertainty. This includes terms from the *in situ* measurements, pile-up effects, flavour dependence, and estimates of additional effects. These uncertainties are reduced to a set of components that preserve as much as possible the correlations between the different p_{T} and η regions. In a given analysis, to estimate the uncertainty due to the JES, the energy is re-scaled up and down by 1σ for each component before re-applying the corresponding object and event selections.

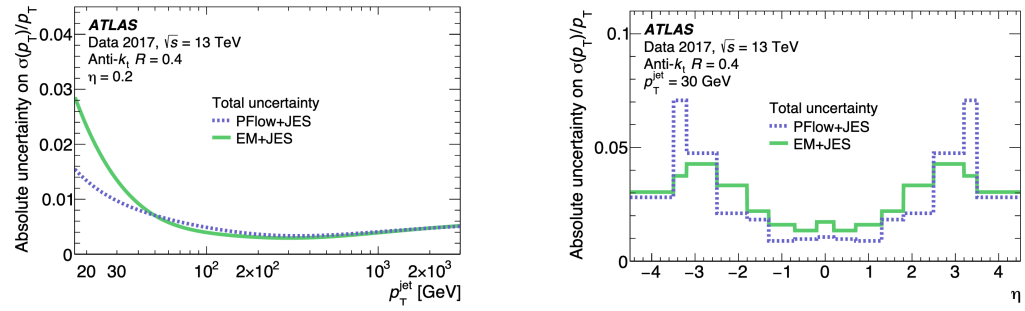
The jet energy resolution (JER) is determined from events in which a jet recoils against a reference object whose momentum can be measured precisely, or dijet events in which the momentum of the two jets is balanced against one another. A smearing procedure is also applied to simulated events to ensure that the JER in simulation matches that in data. This is applied in cases where the jet p_{T} resolution in data is larger than in MC simulation. Otherwise, no smearing is performed since the data should remain unaltered.

JER systematic uncertainties are propagated through physics analyses by smearing jets according to a Gaussian function with width $\sigma^2_{smear} = (\sigma_{nom} - |\sigma_{unc}|)^2 - \sigma^2_{nom}$, where σ_{nom} is the nominal JER of the simulated sample and σ_{unc} is the 1σ variation in the evaluated uncertainty component. The total JER uncertainty is depicted in Figure 2.21.



(a) (b)

Figure 2.20: Fractional jet energy scale systematic uncertainty summed across all components for anti-k_t $R = 0.4$ jets (a) as a function of jet p_{T} at η = 0 and (b) as a function of η at p_{T} = 60 GeV for both topo-jets (solid green line) and PF jets (dashed blue line) [291].



(a) (b)

Figure 2.21: Fractional jet energy resolution systematic uncertainty summed across all components for anti-k_t $R = 0.4$ jets (a) as a function of jet p_{T} at η = 0.2 and (b) as a function of η at p_{T} = 30 GeV for both topo-jets (solid green line) and PF jets (dashed blue line) [291].

In order to further suppress events from in-time pile-up the jet-vertex-tagger (JVT) [295] is used. This algorithm is based on a two-dimensional likelihood discriminant which exploits the information from the tracks matched to each jet to further reject any jets not originating from the hard scattering. The JVT efficiencies on jets coming from the hard scattering are measured in data using the tag-and-probe method in $Z(\rightarrow \mu^+\mu^-)$+jets events. The differences between data and simulation are used to compute simulation-to-data scale factors, which are applied as an additional correction factor to simulation.

The uncertainties associated with the JVT are evaluated in each analysis by varying the correction factors up and down within its uncertainties. Figure 2.22 shows the efficiencies for both data and MC simulation.

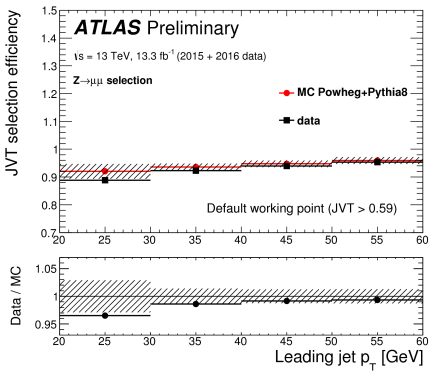


Figure 2.22: The hard-scatter jet selection efficiency measured in $Z(\rightarrow \mu^+\mu^-)$+jets events as a function of the leading jet p_{T} in data and MC simulation for JVT > 0.59. The uncertainties shown are the statistical uncertainty summed in quadrature with the systematic uncertainty. The bottom panels of each figure show the ratio of efficiencies measured in data and MC simulation [296].

b-tagging algorithms: Jets originating from the fragmentation of b-quarks (b-jets) are identified using different algorithms which exploit the main features of this kind of jets [297, 298]. These include the long lifetime of b-hadrons, which generally lead to at least one vertex displaced from the hard scattering or *secondary vertex*; their high mass and high decay multiplicity; and the properties of the b-quark fragmentation. The identification of b-jets is performed in two steps. Firstly,

low-level algorithms reconstruct the main features of the b-jets. Secondly, the results of these algorithms are combined in high-level algorithms consisting of multivariate classifiers. The different low-level algorithms can be grouped in three categories:

- **Impact-parameter-based algorithms**: These algorithms exploit the individual properties of the different tracks associated with a jet. In particular, tracks originated from a b-hadron decay are characterised by their large impact parameters (d_0, z_0). Therefore, the impact parameter significances of these tracks, defined as the impact parameter divided by its uncertainty, differ significantly from zero in contrast with those from tracks in light-jets[8]. The IP2D and IP3D [298] are two examples of these algorithms. The former uses as input the transversal impact parameter significances of all the associated tracks within a jet. The latter includes also the significance of the longitudinal impact parameter, building a two-dimensional template to account for possible correlations. These algorithms use a likelihood ratio technique in which the input variables are compared to predefined distributions provided by MC simulation for each jet-flavour hypothesis (b-, c- and light-jets). Another algorithm is the RNN1P [299]. This algorithm uses a recurrent neural network[9] to exploit the spatial and kinematic correlations between the tracks which are initiated from the same b-hadron.

- **Secondary vertex-based algorithms**: These algorithms use the information of the secondary vertices to construct discriminating variables for b-tagging jets. One example of these algorithms is SV1 [300]. This is a likelihood-based tagger which uses: the information from the invariant mass of all particles in the reconstructed secondary vertex, the ratio of the sum of the energies of the tracks in the vertex to the sum of the energies of all tracks in the jet, the number of two-track vertices and the ΔR between the primary-secondary vertex direction and the jet direction.

[8] These include any non-flavour tagged jet, i.e. jets initiated from u, d or s quarks or gluons.

[9] This is a class of artificial neural network in which the output depends on the prior elements within the sequence. When it makes a decision, it considers the current input and also what is has learned from the inputs it received previously.

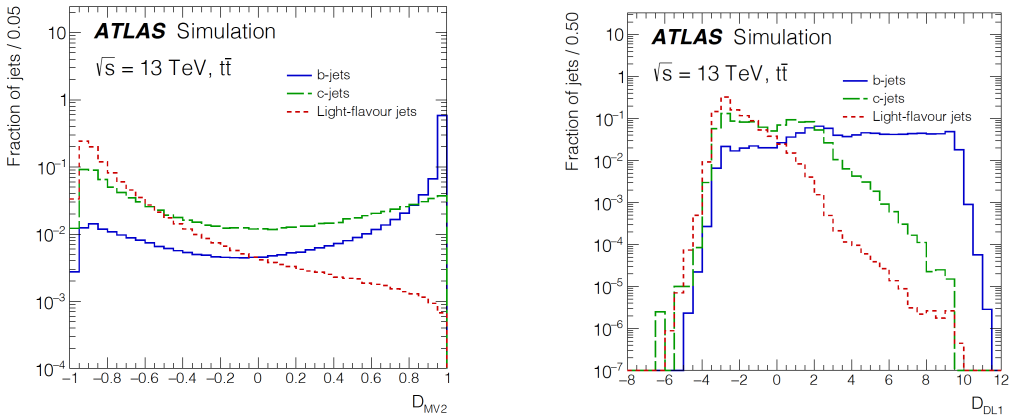
- **Decay chain reconstruction**: These algorithms aim to reconstruct the full b-hadron decay chain. One example is the JetFitter algorithm [301], which exploits the topology of weak b- and c-hadron decays inside the jet. This is done using a Kalman filter [280] to find a common line between the primary, bottom and charm vertices. From this, the b-hadron flight path is reconstructed and the vertex positions are determined.

High-level taggers use the outputs from these algorithms to calculate a final decision on the likelihood that a given jet is a b-, c- or light-jet. Examples of these high-level algorithms are MV2 [298] and DL1 [298]. These differ mainly on the architecture: MV2 is based on a boosted decision tree (BDT) while DL1 is based on a deep neural network (DNN). They both use the input from the IP3D, the SV1 and the JetFitter algorithms as well as the kinematic properties of the jets (namely p_{T} and η).

The BDT employed in the MV2 algorithm is trained on simulated $t\bar{t}$ and Z'[10] events and optimised to discriminate b-jets (signal) with respect to c- and light-jets (backgrounds). Figure 2.23 (a) shows the BDT score, which ranks from -1 to 1. A particular working point is chosen according to the desired b-tagging efficiency, which defines light- and c-jet rejection, after applying a requirement on the minimum threshold of the BDT distribution. There are additional versions of the MV2 algorithm which are used to provide further discrimination between b- and c-jets. An example of these versions is found in the MV2c10 variant. In this case, b-jets are used as signal and the background training sample is composed of 7% (93%) c- (light-) jets.

The DL1 algorithm uses a DNN trained on simulated $t\bar{t}$ and Z' events and treats all flavours equally during training. Thus, the algorithm can be used for both b- and c-tagging. The DL1 algorithm provides a multidimensional output which includes three different probabilities p_b, p_c and p_u for b-, c- and light-flavour tagging, respectively. The final output discriminant combines all b, c and light-flavour probabilities in the form:

[10] Z' refers to a hypothetical gauge boson that arises from extensions of the electroweak symmetry of the SM [302–304].



(a) (b)

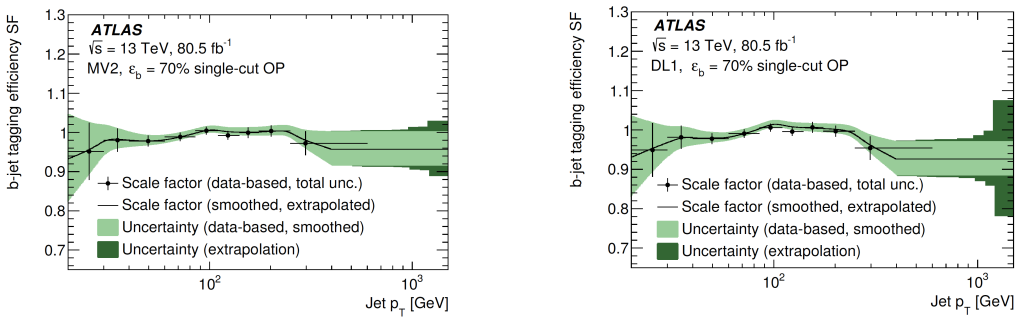
Figure 2.23: Distribution of the output discriminant of the (a) MV2 and (b) DL1 b-tagging algorithms for b-, c- and light-jets in the baseline $t\bar{t}$ simulated events [297].

$$D_{DL1} = \log \frac{p_b}{f_c \times p_c + (1 - f_c) \times p_u}, \tag{2.9}$$

where f_c represents the effective c-jet fraction in the background training sample. The output of the algorithm can be seen in Figure 2.23 (b). Similarly to the BDT case, a working point is chosen according to the desired b-tagging efficiency. One extension of the DL1 algorithm is found in the DL1r variant, which includes the output from the RNN1P algorithm.

The efficiency of the b-tagging algorithm is evaluated on both data and simulation. A set of scale factors, defined as the ratio between the measured b-jet tagging efficiency to the b-jet tagging efficiency derived from the simulation, are calculated to account for the observed differences for each considered working point. The scale factors are smoothed, extrapolated to the high-p_{T} regime and corrected taking into account the generator dependence in the simulation. A weight is then applied in simulation for each jet where b-tagging is applied to match the tagging rate as measured in data.

Figure 2.24 shows the b-tagging scale factors for a 70% efficiency working point for both the MV2 and DL1 taggers with their associated uncertainties, including



(a) (b)

Figure 2.24: b-jet tagging efficiency simulation-to-data scale factors for the 70% working point of the (a) MV2 and (b) DL1 taggers, including the smoothed and extrapolated results. The bin centres are used for the smoothing whereas the dots are located at the mean of the b-jet p_{T} distribution in each p_{T} bin [297].

the smoothed and extrapolated results. These uncertainties are propagated and evaluated in the different analyses. The associated uncertainty components can be separated in three groups: those associated with the b-jet identification efficiency and those related to both c- and light-jets mistagging uncertainties.

2.5.4.2 Large-R jets

At the LHC it is quite frequent to produce very massive particles such as Z, $W^{\pm}$ or Higgs bosons or a top quark with large Lorentz boosts. When these particles decay, their decay products become collimated (boosted) in the direction of the progenitor particle. Therefore, the hadronic decay products can be reconstructed as a single large-R jet [305, 306]. These large-R jets have a characteristic substructure which is distinct from the radiation pattern of jets initiated by either gluons or quarks of all flavours rather than top quarks.

Large-R jets are reconstructed using the anti-k_t algorithm with a radius parameter of 1.0. The input objects can be either topological clusters, particle-flow objects, Track-CaloClusters [307, 308] or Unified Flow Objects (UFOs) [305].

In this work large-R jets are reconstructed from topo-clusters of calorimeter cells. A schematic description of the large-R jet reconstruction and calibration scheme

can be found in Figure 2.25. A *grooming* technique is applied after the large-R jet reconstruction to remove the contamination from other sources such as pile-up or ISR, increasing notably the large-R jet energy and mass resolutions. Different algorithms can be used for this purpose such as pruning [309] or mass drop filtering algorithms [310]. In this case, a trimming technique [311] is used. To trim a large-R jet, the jet constituents are reclustered into sub-jets using the anti-k_t algorithm with a radius parameter of 0.2. Then, sub-jets with a transverse momentum lower than a 5% of the total momentum from the original ungroomed jet are removed from the large-R jet.

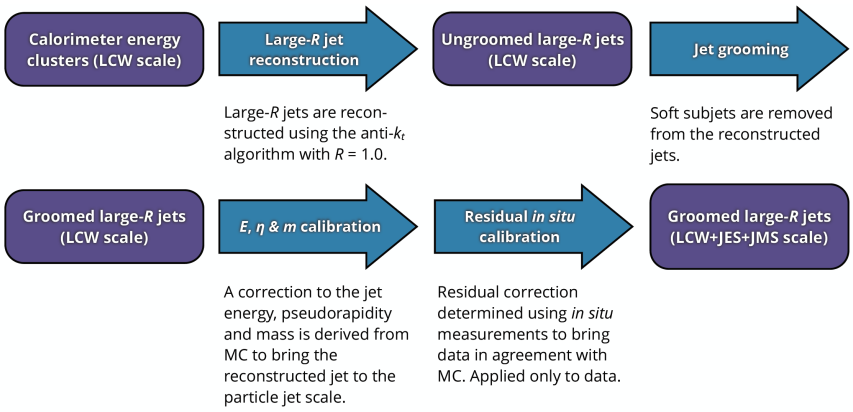


Figure 2.25: Overview of the large-R jet reconstruction and calibration procedure [306].

Next, the energy, pseudorapidity and mass of the large-R jets are corrected to the particle level scale, using MC simulated dijet events. At this stage, the jet mass resolution is improved by combining the jet mass measurement in the calorimeter with the measurement of the charged component of the jet within the ID (the so-called track-assisted mass) [312–320].

A second *in situ* calibration is applied directly to data on two sequential steps: The JES calibration, which is performed using the same procedures as for small-R

jets (see Section 2.5.4.1), and the jet mass scale (JMS) calibration[11], which is obtained from the combination of two different methods:

- R_{trk} **method**: This method relies on the tracker information to provide an independent measurement of the JMS and its associated uncertainty. An average calorimeter-to-track jet response is obtained for both data and simulation. The deviations from unity in the ratio between data and simulation provide an ***in situ*** validation of the modelling of the large-R jet. These ratios are obtained for both the transverse momentum and the mass of the large-R jet.

- **Forward folding**: The forward folding method is applied on $t\bar{t}$ events in the lepton+jets final state. This provides an estimate of the jet mass response using fits to the W boson and top quark mass peaks in the large-R jet invariant mass distribution of the hadronically decaying top quark candidate [321]. The position and shape of these peaks are used to measure the relative JMS and also the relative jet mass resolution (JMR) between data and simulation. The associated data-to-simulation scale factors are shown in Figure 2.26.

[11] This is applied to correct for possible residual problems after the JES calibration.

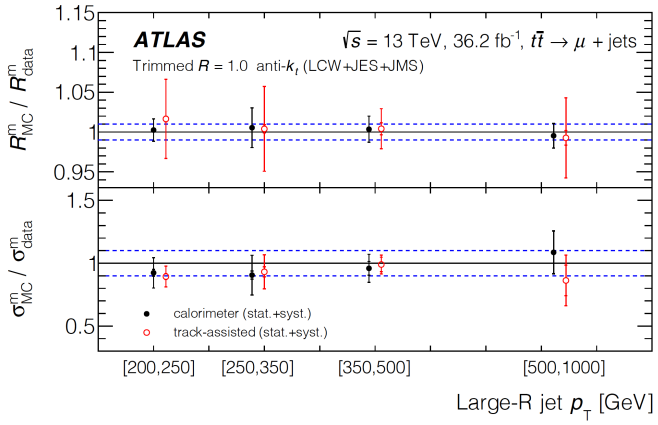


Figure 2.26: Summary of the *in situ* measurements of the large-R jet mass response in $t\bar{t}$ events with a lepton+jets final state as a function of the large-R jet transverse momentum p_{T}. The large-R jet p_{T} is corrected using the simulation calibration, η-intercalibration, and a combination of *in situ* direct balance techniques. The closed circles correspond to the JMS and JMR of trimmed large-R jets reconstructed from calorimeter clusters. The open circles represent the equivalent result for the track-assisted mass. The dashed lines, corresponding to $\pm 1\%$ for the JMS and $\pm 10\%$ for the JMR, are drawn for reference. The results in the first two p_{T} bins (200 GeV $< p_{\mathrm{T}} <$ 250 GeV and 250 GeV $< p_{\mathrm{T}} <$ 350 GeV) correspond to a sample of high-p_{T} W bosons, and the highest two bins (350 GeV $< p_{\mathrm{T}} <$ 500 GeV and 0.5 TeV $< p_{\mathrm{T}} <$ 1 TeV) correspond to high-p_{T} top quarks. In each subsample, the JMS and JMR are extracted simultaneously in a two-parameter fit to the mass distribution. The statistical and total uncertainties are indicated with the small and large error bars on the data points, respectively. [306].

The final estimation of the JES and JMS results from the combination of the different *in situ* techniques. Figures 2.27-2.28 show the data-to-simulation ratio of the average jet p_{T} response and the JMS for two mass intervals as a function of the large-R jet transverse momentum. The results for the JMS are compatible with unity, which indicates that the JES *in situ* calibration is sufficient to correct for the JMS differences between data and simulation.

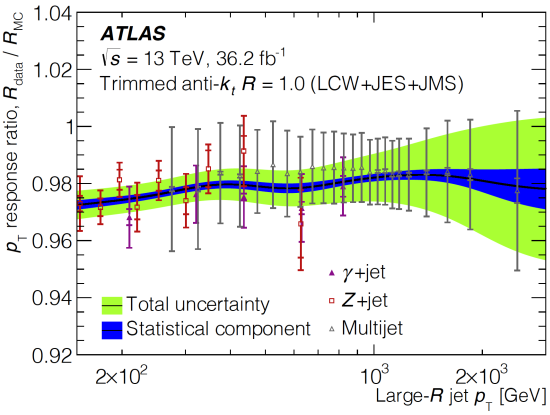
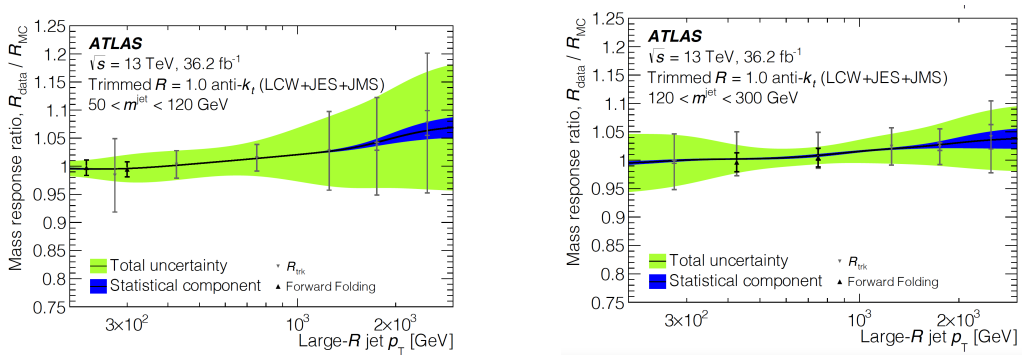


Figure 2.27: Data-to-simulation ratio of the average jet p_{T} response as a function of large-R jet p_{T}. The combined result (band) is based on three *in situ* techniques: the Z+jet balance method (open squares), γ+jet balance method (closed triangles), and the multijet balance (open triangles). The errors represent the statistical (inner error bars) and the total uncertainty (statistical and systematic uncertainties added in quadrature, outer error bars). The results apply to trimmed anti-k_t jets with R = 1.0. The lines shown are smoothed using a sliding Gaussian kernel [306].



(a) (b)

Figure 2.28: Data-to-simulation ratio of the average jet mass response as a function of the large-R jet p_{T}. Corrections using a combination of two *in situ* methods, the R_{trk} and forward-folding approaches, are applied. The fit is performed for large-R jet mass in the W mass range 50–120 GeV (a), and the top mass range 120–300 GeV (b). The error bars represent the statistical and systematic uncertainties added in quadrature. The results apply to anti-k_t jets with R = 1.0 calibrated with the LCW+JES+JMS scheme. The lines shown are smoothed using a sliding Gaussian kernel [306].

The JER of large-R jets is estimated exploiting the momentum balance between leading and subleading large-R jets in dijet events. The momenta of the two jets is used to construct an asymmetry distribution, and the width of this distribution is used to estimate the JER. Figure 2.29 shows the measured relative resolution for both data and simulation as a function of the transverse momentum of the dijet system $p_{\mathrm{T}}^{\mathrm{avg}}$. Given that the observed difference between data and simulation in the JER is consistent with one within uncertainties, no nominal smearing is applied. Instead, these differences are taken as an additional uncertainty.

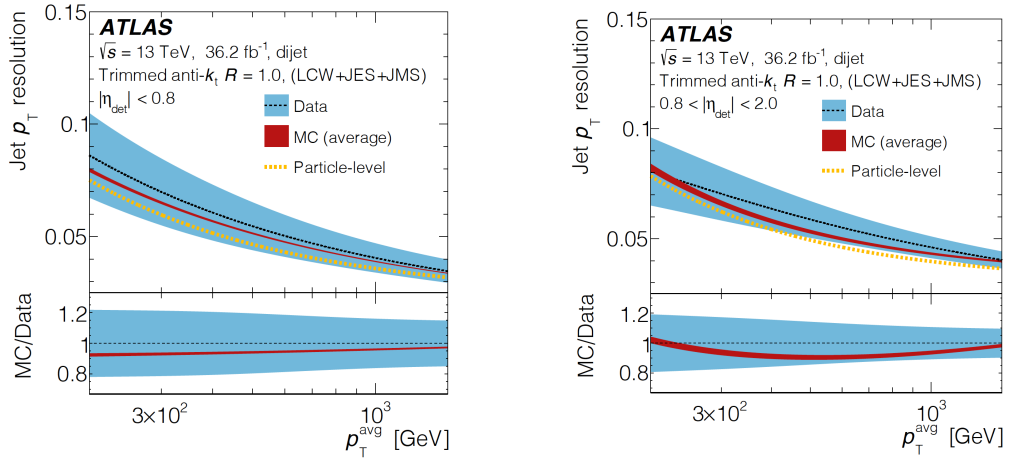
The evaluation of the associated uncertainties in the JES, JMS, JER and JMR is performed in a very similar was as for small-R jets: the uncertainties in the JES, JMS and JMR scale factors are propagated to the analyses in different components. For the JMR, a flat 20% uncertainty is taken where no forward folding measurements are available. For the JER uncertainties, large-R jets are smeared taking into account the 1σ variation of the associated uncertainty component. More details can be found in Section 2.5.4.1.

top-tagging algorithms: Different algorithms can be used to classify a jet as originating from a top quark decay as opposed to a jet originating from gluons and quarks of all flavours other than top quarks [322, 323]. These methods exploit the features of large-R jets originating from a top quark. These include the high reconstructed jet mass, which is expected to have a scale associated with the mass of the progenitor particle, or observables which quantify the extent to which the jet constituents are clustered or uniformly dispersed.

These variables provide complementary information. In order to obtain a higher discrimination power for top-tagged large-R jets, these observables are combined by creating a multivariate top-quark classifier in the form of a BDT or a DNN[12]. The performance of these two algorithms is very similar [322]. Thus, the results shown below are focused on the DNN-based taggers.

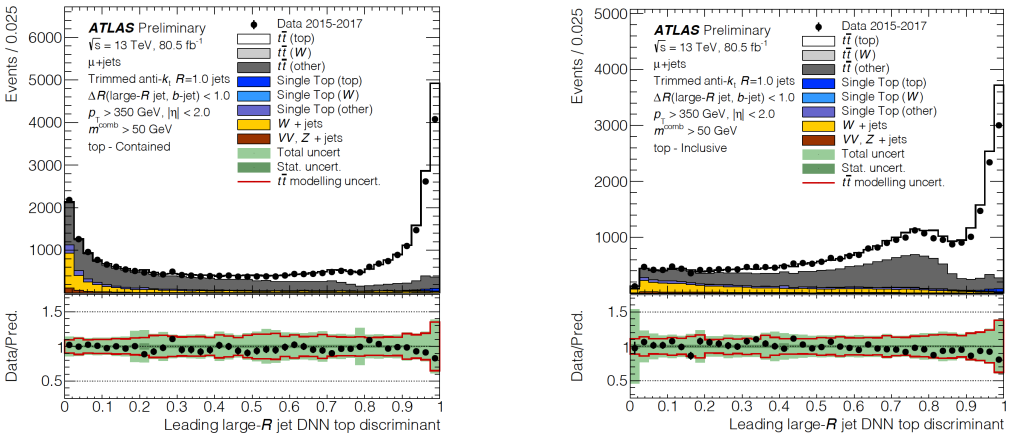
The performance of the top-tagging algorithms is evaluated in $t\bar{t}$ events in the lepton+jets final state. Two flavours of this tagger are available, based on the

[12]There are other possibilities to identify top-tagged jets. Among them, one can find a "TopoDNN", which uses low-level input observables based directly on the jet constituents, shower descontruction [324], or the HEPTopTagger algorithm [316, 317]. For more details, see Ref. [322].



(a) (b)

Figure 2.29: Comparison of the measured jet p_{T} resolution with the resolution determined in simulation, averaged between different generators as a function of the average jet p_{T} and in two bins of detector pseudorapidity η_{det} from (a) $|\eta_{det}| < 0.8$ and from (b) $0.8 < |\eta_{det}| < 2.0$. The large-R jet p_{T} is corrected using the simulation calibration, η-intercalibration, and a combination of *in situ* direct balance techniques. The error band, drawn as a light band, represents the statistical and systematic uncertainties added in quadrature. The determination of the relative resolution using the *in situ* technique for an average of three simulations and their envelope is also shown as a dark band. Inconsistencies between the resolution determined using the *in situ* technique and of the resolution determined from the response in simulation by matching particle-level jets to reco-level jets (light dotted line) are taken as an additional uncertainty in the measurement. The lines shown are obtained by smoothing a binned representation of these uncertainties using a sliding Gaussian kernel [306].

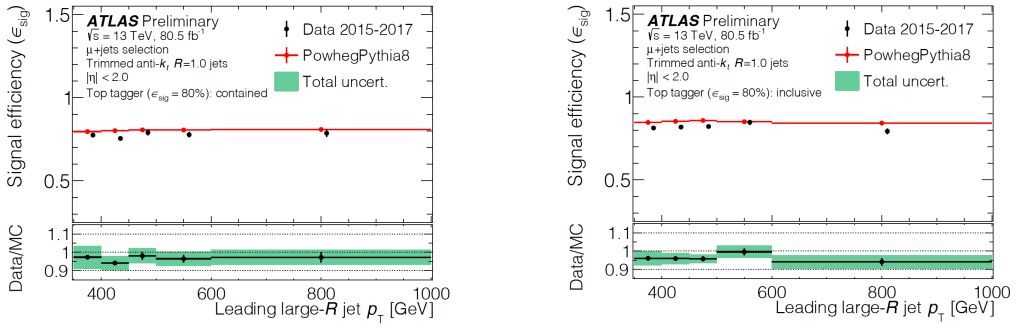


(a) (b)

Figure 2.30: Distribution of the DNN score for the contained top (a) and inclusive top (b) tagger in data and MC simulation for the leading large-R jets in top-enhanced single-muon $t\bar{t}$ events [323].

samples that are used for the training. The *contained* tagger is trained using jets that contain all of the energy from a hadronically decaying top quark. In contrast, the *inclusive* tagger is trained on any jet that contains part of a hadronic top quark decay, regardless of full containment. Figure 2.30 shows the response of the DNN classifier for the two types of algorithms and for both data and simulation. The $t\bar{t}$ sample is divided in three groups according to the progenitor particle that initiated the large-R jet (top, W or other). As it can be seen, the algorithm discriminates quite well those events with a large-R jet originating from a top quark hadronic decay from the different backgrounds.

These results can be used to estimate the signal efficiency (i.e. the number of events that are identified as top-tagged jets over the total number of events) in both data and simulation. The chosen working points are defined so that they yield constant signal efficiencies versus the reconstructed jet p_{T}. The number of signal events on data are estimated from a χ^2 template fit of signal and background distributions predicted by MC simulations. Then, the number of signal-like events in data that pass and fail the top-tagging algorithm as extracted from the fit are used

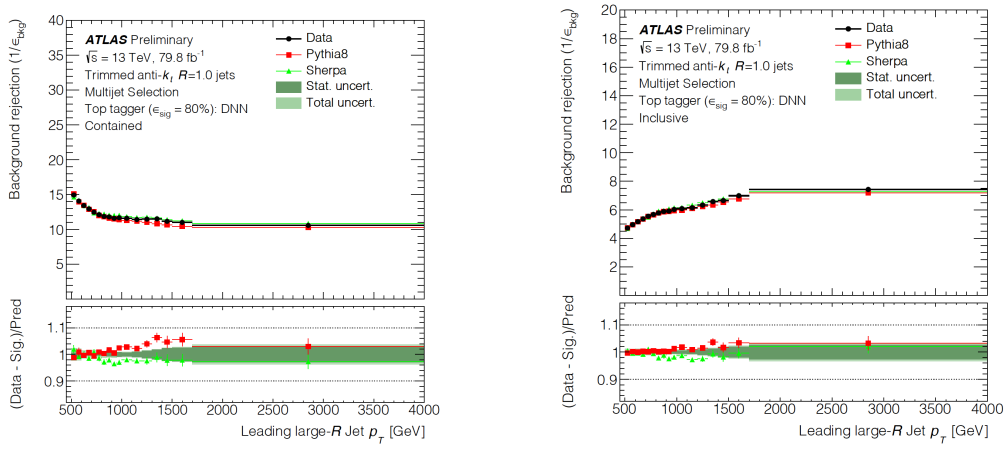


(a) (b)

Figure 2.31: The measured efficiencies for the 80% contained top (a) and 80% inclusive top (b) are shown on the top panel. The bottom panel shows the tagger efficiency scale factors with the total uncertainty bands [323].

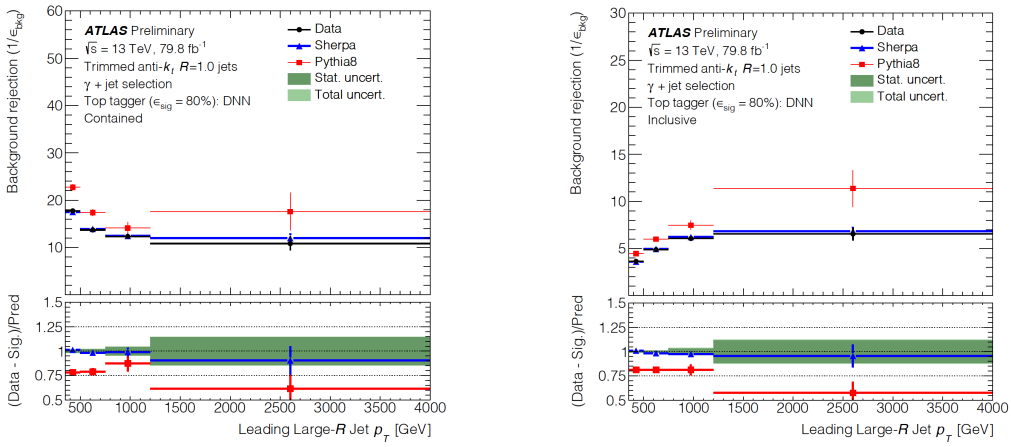
to compute these efficiencies. The simulation efficiencies are extracted directly from the number of simulated signal events. Figure 2.31 shows the signal efficiency on top-quark jets for both data and simulation. This provides an ***in situ*** correction in the form of a jet-by-jet efficiency scale factor [325, 326] which is applied to simulation. These scale factors have a quite limited p_{T} validity range ($p_{\mathrm{T}} \lesssim 1$ TeV) due to the insufficient statistics in data at high-p_{T}. Nevertheless, a method of extrapolating these scale factors and their associated uncertainties is employed to extend the validity range up to $p_{\mathrm{T}} \leq 4$ TeV using only MC simulation [323].

The behavior of background jets originating from gluons or other quarks rather than a top quark is studied in events enriched in multijet and γ+jet processes. In this case, the procedure is much simpler due to the high purity of these samples. The normalisation of these events is derived from the data by subtracting the contributions from hadronically decaying W and Z bosons and $t\bar{t}$ events. The background efficiency is calculated directly as the fraction of events that satisfy the full set of tagging criteria, both for data and simulation. Then, the background scale factors are defined in a similar manner as the signal scale factors, and are applied to simulation. The results are shown in Figures 2.32-2.33, for multijet and γ+jets events, respectively.



(a) (b)

Figure 2.32: Measured background rejection in multijet events for the 80% contained top tagger (a) and the 80% inclusive top tagger [323].



(a) (b)

Figure 2.33: Measured background rejection in γ+jets events for the 80% contained top tagger (a) and the 80% inclusive top tagger [323].

These scale factors are affected by different systematic uncertainties. The associated uncertainties are propagated to the different analyses in different components, which include terms associated to the efficiency on background jets, signal efficiency, extrapolation at high-p_{T} regime and inefficiencies of the taggers.

2.5.5 Missing transverse momentum

Due to the conservation of momentum, the sum of the transverse momenta of all the collision products should be zero in the transverse plane. This is computed from all the detected particles in the detector. However, the presence of undetected particles may lead to an imbalance in this calculation, producing the so-called missing transverse momentum $E_{\mathrm{T}}^{\mathrm{miss}}$. This is associated to the SM neutrinos, but it could also come from other weakly-interacting particles which escape undetected such as possible DM candidates. Thus, the $E_{\mathrm{T}}^{\mathrm{miss}}$ is a required ingredient in the top-quark polarisation analysis, which expects one neutrino in the final state, and also in the DM search presented in this work. Fake $E_{\mathrm{T}}^{\mathrm{miss}}$ can also arise from particles which escape the detector acceptance or particles which are badly reconstructed as a consequence of the limited acceptance of the detector, the finite detector resolution, the possible dead regions or any source of noise.

The $E_{\mathrm{T}}^{\mathrm{miss}}$ is reconstructed as the magnitude of the negative vector sum of transverse momenta of all detected particles [327, 328]. This includes leptons, photons, jets and soft-event signals, which are signals consisting on reconstructed charged-particle tracks associated with the hard scattering vertex but not with any hard object. Since soft-event signals include only tracks associated to the primary vertex, this term is relatively insensitive to pile-up effects.

The magnitude and azimuthal direction of the $E_{\mathrm{T}}^{\mathrm{miss}}$ can be obtained from its x and y components:

$$E_{\mathrm{T}}^{\mathrm{miss}} = \sqrt{(E_{\mathrm{x}}^{\mathrm{miss}})^2 + (E_{\mathrm{y}}^{\mathrm{miss}})^2} \tag{2.10a}$$

$$\phi_{E_{\mathrm{T}}^{\mathrm{miss}}} = \arctan \frac{E_{\mathrm{x}}^{\mathrm{miss}}}{E_{\mathrm{y}}^{\mathrm{miss}}} \tag{2.10b}$$

The quality of the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction is evaluated on data and simulated events with different event topologies. $Z \rightarrow l^{+}l^{-}$ events are used to study the performance of the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction due to its clean detector signature. These events are characterised by low values of the $E_{\mathrm{T}}^{\mathrm{miss}}$, and thus one can obtain information on the intrinsic detector resolutions, the performance of the algorithms involved and the efficiency on the object reconstruction. $W \rightarrow l\nu$ events produce high-p_{T} neutrinos in the final state, and are used to extract information on the scale of the $E_{\mathrm{T}}^{\mathrm{miss}}$. Finally, $t\bar{t}$ events are used to test the robustness of the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation in multijet environments, due to the high-jet multiplicity expected in the final state.

The systematic uncertainties related to the modelling of the $E_{\mathrm{T}}^{\mathrm{miss}}$ in the simulation are estimated by propagating the uncertainties in the energy and momentum scales of the reconstructed hard objects, as well as the uncertainties in the resolution and scale of the soft term [327]. The uncertainties associated to the soft term include the modelling of the underlying event and the impact of the soft terms on the p_{T} scale and resolution of unclustered energy. These are estimated from comparisons between data and MC simulations of observables measuring the contribution of the soft term to the overall $E_{\mathrm{T}}^{\mathrm{miss}}$. The impact of these uncertainties is evaluated in the analyses by including 1σ scale variations and a smearing according to the resolution uncertainty.

2.6 Computing system

In spite of the trigger systems, the total amount of data that is finally stored at the LHC is still huge: between ∼50-70 Petabytes of data are recorded every year. The CERN computing services cannot deal by themselves with such a huge amount of data, which needs first to be stored, then retrieved and finally analysed. This is only possible thanks to the Grid computing technology [329], which shares resources from different computing services around the world. These include storage facilities, a global network and an immense computing power.

This was the origin of the Worldwide LHC Computing Grid (WLCG) [330], with a distributed computing infrastructure arranged in four different levels to provide

almost real-time access to LHC data to physicists all around the world. A diagram showing the different tiers can be found in Figure 2.34.

The Tier-0 is located both at the CERN data centre and the Wigner Research Centre for Physics in Budapest. All the data collected at the LHC passes through the CERN hub, and is then distributed to the different centres. Indeed, CERN provides less than 20% of the total computing capacity. The two Tier-0 centres are responsible not only of keeping the raw data (the first copy, without further reprocessing), but also of performing the first reconstruction of the data and its distribution to the Tier-1 centres. The thirteen Tier-1 centres are placed all around the world and are responsible for storing a proportional share of raw and reconstructed data. In addition, they take care of performing large-scale reprocessing and storing the corresponding output, and also distribute the data to the Tier-2 facilities. The Tier-2 facilities are constituted by different universities and scientific institutes ($\sim$160) which form part of the CERN community. They store sufficient data and provide computing power for specific analysis. In addition, the simulated events are shared proportionally between the different institutes, and they also handle the diverse analysis requirements. Finally, the Tier-3 is formed by the local computing resources which can be accessed by individual scientists to perform specific tasks within each analysis.

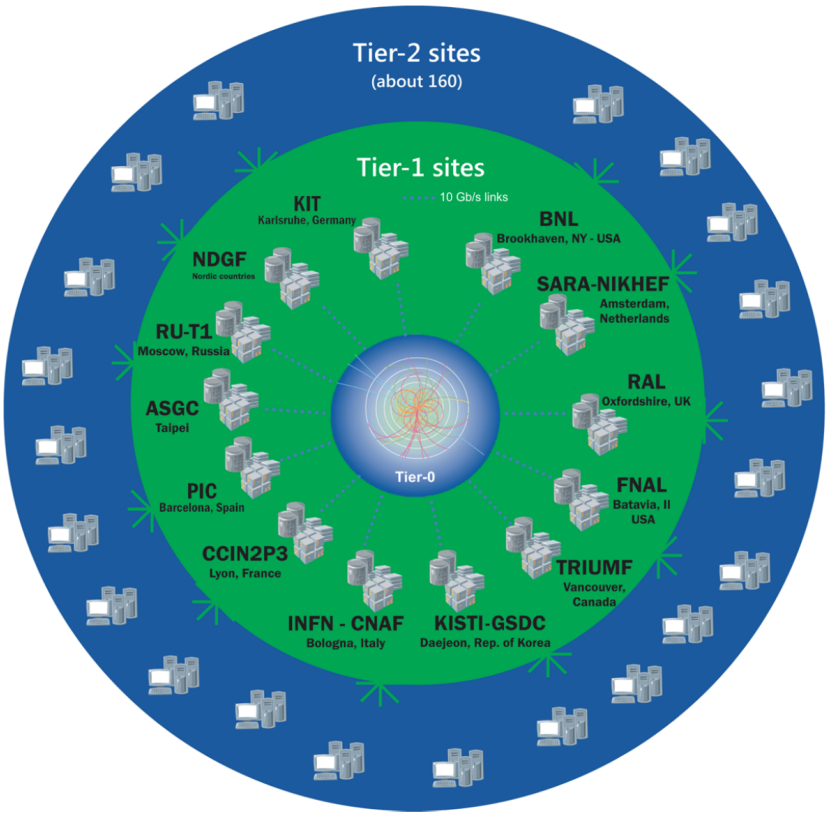


Figure 2.34: Schematic view of the different tiers that form the WLCG, with the exemption of the Tier-3 [330].

2.7 Trigger and object selection in the analyses

This section describes the trigger selection and the candidate objects used for the two analyses presented in this work. This is based on the trigger system defined in Section 2.4 and on the reconstruction techniques defined in Section 2.5. The final selection depends on the topology of the final state, which is very different for each analysis. The analysis described in Chapter 4 is characterised by a top quark in the final state, which decays leptonically, produced along with a light-jet.

In contrast, the analysis described in Chapter 5 features a top quark, which decays hadronically, recoiling against a large amount of missing transverse momentum.

2.7.1 Trigger selection

The analysis described in Chapter 4 expects leptons in the final state. Thus, events were selected by single-lepton triggers [331, 332], imposing low thresholds on the E_{T} of electrons and on the p_{T} of the muons, in addition to isolation requirements, or imposing a looser identification criterion and higher thresholds with no isolation requirement. The lowest thresholds varied from 20 to 26 GeV depending on the lepton flavour and the data-taking period.

The analysis described in Chapter 5 aims to search for DM particles produced in the final-state, which are associated with a large amount of missing transverse momentum. Therefore, events were selected by $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers imposing lower thresholds which varied from 70 to 110 GeV depending on the data-taking period. In addition, high-p_{T} jet triggers ($p_{\mathrm{T}} > 400$ GeV) were also applied to recover data efficiency in the low $E_{\mathrm{T}}^{\mathrm{miss}}$ region.

2.7.2 Object selection

The basic requirements applied on the different physics candidate objects used for the polarisation analysis (Chapter 4) and the mono-top search (Chapter 5) are summarised in Tables 2.3-2.4, respectively. Additional details are given in the following.

Vertex selection: The two analyses presented in this work require at least one primary vertex, built from at least two tracks. In case there is more than one primary vertex candidate, the one with the largest sum of p_{T}^2 of all its associated tracks is used. Indirectly, the analyses also use the information from secondary and tertiary vertices, which are used in the b-tagging algorithms. This is further explained in Section 2.5.4.1.

Electron selection: In the polarisation analysis, electron candidates must satisfy $p_{\mathrm{T}} > 10$ GeV, $|\eta_{\mathrm{cluster}}| < 2.47$ and the tight identification requirement. They are excluded if their calorimeter clusters lie within the transition

Electrons
$p_{\mathrm{T}} > 10$ GeV
$\lvert\eta\rvert < 2.47$, $\notin$ (1.37,1.52)
Working point: Tight
$\lvert d_0/\sigma(d_0)\rvert < 5$
$\lvert z_0 \sin\theta\rvert$ (mm) < 0.5
Muons
$p_{\mathrm{T}} > 10$ GeV
$\lvert\eta\rvert < 2.5$
Working point: Medium
$\lvert d_0/\sigma(d_0)\rvert < 3$
$\lvert z_0 \sin\theta\rvert$ (mm) < 0.5
Jets
$p_{\mathrm{T}} > 30$ GeV, 35 GeV in $\lvert\eta\rvert \in$ (2.75,3.5)
$\lvert\eta\rvert < 4.5$
Tagging algorithm: MV2c10
Working point: 60%
JVT > 0.59 for $\lvert\eta\rvert < 2.4$ and $p_{\mathrm{T}} < 60$ GeV
Missing transverse momentum
$E_{\mathrm{T}}^{\mathrm{miss}} > 35$ GeV

Table 2.3: Basic requirements on the physics candidate objects used in the polarisation analysis (Chapter 4). Additional details are given in the text.

Electrons
$p_{\mathrm{T}} > 30$ GeV
$\lvert\eta\rvert < 2.47$, $\notin$ (1.37,1.52)
Working point: Medium
$\lvert d_0/\sigma(d_0)\rvert < 5$
$\lvert z_0 \sin\theta\rvert$ (mm) < 0.5
Muons
$p_{\mathrm{T}} > 30$ GeV
$\lvert\eta\rvert < 2.5$
Working point: Medium
$\lvert d_0/\sigma(d_0)\rvert < 3$
$\lvert z_0 \sin\theta\rvert$ (mm) < 0.5
Jets
$p_{\mathrm{T}} > 25$ GeV
$\lvert\eta\rvert < 2.5$
Tagging algorithm: DL1r
Working point: 77%
JVT > 0.5 for $\lvert\eta\rvert < 2.4$ and $p_{\mathrm{T}} \in [20, 60]$ GeV
Large-R jets
$p_{\mathrm{T}} > 250$ GeV
$\lvert\eta\rvert < 2.0$
Tagging algorithm: DNN contained
Working point: 50%
Missing transverse momentum
$E_{\mathrm{T}}^{\mathrm{miss}} > 250$ GeV

Table 2.4: Basic requirements on the physics candidate objects used in the mono-top search (Chapter 5). Additional details are given in the text.

region between the barrel and the end-cap of the electromagnetic calorimeter, $1.37 < |\eta_{\text{cluster}}| < 1.52$. The track associated with the electron must pass the requirements $|z_0 \sin\theta| < 0.5$ mm and $|d_0/\sigma(d_0)| < 5$, where z_0 describes the longitudinal impact parameter with respect to the reconstructed primary vertex, d_0 denotes the transverse impact parameter relative to the beam-line axis and $\sigma(d_0)$ is the uncertainty on d_0. Additional isolation criteria are used to identify prompt tight electrons. The scalar sum of the p_{T} of tracks within a variable-size cone around the electron, excluding tracks originating from the electron itself, must be less than 6% of the electron p_{T}. The track isolation cone radius $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ is given by the smaller of $\Delta R = 10$ GeV$/p_{\text{T}}$ and $\Delta R = 0.2$. In addition, the sum of the transverse energy of the calorimeter topo-clusters in a cone of $\Delta R = 0.2$ around the electron is required to be less than 6% of the electron p_{T}, excluding clusters originating from the electron itself. Loose electrons with $p_{\text{T}} > 10$ GeV are also used for the multijet background estimation (see Section 4.4.1) and vetoed in the signal definition.

The mono-top search rejects events with electrons in the final state, using the medium identification. No isolation criteria are applied on the identified electrons. Electron candidates must satisfy $p_{\text{T}} > 30$ GeV and $|\eta_{\text{cluster}}| < 2.47$ (excluding the $1.37 < |\eta_{\text{cluster}}| < 1.52$ region). The track associated with the electron must pass the requirements $|z_0 \sin\theta| < 0.5$ mm and $|d_0/\sigma(d_0)| < 5$.

Muon selection: Both the polarisation and the mono-top analyses consider combined muons. In the polarisation analysis muon candidates must satisfy $p_{\text{T}} > 10$ GeV and $|\eta| < 2.5$ along with the medium identification requirement. This criterion defines requirements on the number of hits in the different ID and MS subsystems and on the significance of the charge-to-momentum ratio q/p. In addition, the track associated with the muon candidate must have $|z_0 \sin\theta| < 0.5$ mm and $|d_0/\sigma(d_0)| < 3$. Isolation criteria are applied to the selected muons. The scalar sum of the p_{T} of tracks within a variable-size cone around the muon (excluding its own track) must be less than 6% of the muon p_{T}, with the track isolation cone radius being given by the minimum of $\Delta R = 10\text{GeV}/p_{\text{T}}$ and $\Delta R = 0.3$. Similarly to electrons, loose muons with

$p_{\mathrm{T}} > 10$ GeV are used for the multijet background estimation (see Section 4.4.1) and vetoed in the signal definition.

The mono-top analysis rejects events with medium muons in the final state. No isolation criteria are required on the identified muons. Muon candidates must satisfy $p_{\mathrm{T}} > 30$ GeV and $|\eta| < 2.5$. In addition, the track associated with the muon candidate must have $|z_0 \sin\theta| < 0.5$ mm and $|d_0/\sigma(d_0)| < 3$.

Jet selection: The polarisation analysis uses topo-jets with $p_{\mathrm{T}} > 30$ GeV and within $|\eta| < 4.5$. The requirement on the p_{T} is increased to 35 GeV in the calorimeter end-cap-forward transition region ($2.75 < |\eta| < 3.5$). To suppress jets from additional pp interactions within a bunch crossing, a JVT larger than 0.59 is required for all jets with $p_{\mathrm{T}} < 60$ GeV and $|\eta| < 2.4$. b-jets are identified (tagged) by the MV2c10 b-tagging algorithm. The chosen working point for this analysis corresponds to a b-tagging efficiency of 60% [325]. The corresponding mistagging rates for the c and light-jets are approximately 2.9% and 0.065%, respectively, as predicted in simulated $t\bar{t}$ events and calibrated in data [333, 334].

The mono-top search uses PF jets with $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$. In addition, jets with $20 < p_{\mathrm{T}} < 60$ GeV and $|\eta| < 2.4$ are rejected if their JVT score is lower than 0.5. The DL1r algorithm is used for b-tagging identification, with the 77% b-tagging efficiency working point. The corresponding mistagging rates for the c and light-jets are approximately 20% and 0.59%.

Large-R jet selection: Large-R jets are used in the mono-top analysis. Large-R jets are considered if they fulfill $|\eta| < 2.0$[13] and $p_{\mathrm{T}} > 250$ GeV. The top-tagging algorithm used in this analysis is the DNN contained, with a working point associated to a top-tagging efficiency of 50%. This algorithm and working point have been chosen taking into account the sensitivity to the mono-top signal and the data/MC agreement.

Misssing transverse momentum selection: The polarisation analysis selects events with a $E_{\mathrm{T}}^{\mathrm{miss}}$ larger than 35 GeV, due to the final state neutrino.

[13] The final state is characterised by a significant amount of $E_{\mathrm{T}}^{\mathrm{miss}}$ recoiling against a boosted large-R jet. Thus, this large-R jet is expected to be produced in the central region of the detector.

In the mono-top search, a large amount of missing amount of missing transverse momentum is expected from the final state DM particles. Selected events must satisfy the requirement $E_{\mathrm{T}}^{\mathrm{miss}} > 250$ GeV.

Overlap removal: The reconstructed objects can satisfy both the jets and leptons selection criteria and as such a procedure called "overlap removal" is applied in order to associate objects to a unique hypothesis. The procedure works as follows: if any electron shares a track with a muon, the electron is removed since it is very likely to correspond to the reconstructed muon. Similarly, if any jet is close to an electron (within $\Delta R < 0.2$) the closest jet is removed. Electrons close to jets (within $\Delta R < 0.4$) are also removed to reduce the impact of non-prompt leptons. To reduce contributions from muons which stem from heavy-flavour decays inside a jet, muons are removed if they are separated from the nearest jet by $\Delta R < 0.4$. Additionally, jets with fewer than three tracks and separated from a muon by $\Delta R < 0.4$ are removed to reduce fake jets from muons depositing a large fraction of their energy in the calorimeters.

3 Alignment of the ATLAS Inner Detector

In order to perform high precision physics measurements, it is crucial to provide data of the highest possible quality. Among the different components of the ATLAS detector, the Inner Detector (ID) is responsible for measuring the trajectory of the charged particles (track reconstruction) and identifying primary and secondary vertices. These functionalities make it a fundamental component for some physics tasks, such as the identification of b-jets or lepton reconstruction.

The ATLAS ID is characterised by its high resolution. To take advantage of this resolution power, one needs to know the exact position and the orientation of the different elements that constitute the ID. These components have a nominal position which is determined at the time of its assembly. However, due to changes in the detector conditions (e.g. temperature, magnetic field) the position and geometry of the ID varies with time. These movements of the detector may lead to biased physics measurements, and they need to be corrected.

Unfortunately, the ID is not accessible during data-taking, and hence the position of the detectors needs to be corrected with an alternative online method. This is the main purpose of the track-based alignment algorithm. This has been developed to determine the actual position of each component of the ID and follow its possible changes in time. The accuracy of the alignment algorithm is such that the position of the subsystems is determined with a precision of the order of a few microns [1], higher than the detector resolution. This level of precision is much better than the accuracy which is obtained from a direct measurement of the module

positions. Therefore, the alignment algorithm provides the most precise estimation of the position and orientation of the different modules.

In this chapter the basics of the ATLAS ID alignment will be introduced. The work done in this thesis, which is centred on the alignment of the ATLAS ID during the Run-2, is then presented. The increasing luminosity has introduced some challenges with respect to Run-1. In particular, time-dependent movements which take place during each LHC fill[1] are characteristic of the Run-2. These movements are described focusing on the pixel vertical displacement, which has been studied in detail during this thesis. Lastly, the high precision alignment corresponding to both heavy ions and pp collisions data-taking periods will be reported.

3.1 Alignment basics

As the ID detector is not built (or even stable) to the tolerances that would be required to ignore the uncertainty in the position, the alignment procedure determines the position of all sensitive elements using tracks in collision data. The track-based alignment algorithm aims at the minimisation of the distances between the reconstructed tracks and the measured hits[2] in the detector. The reconstruction of these tracks is based on two ATLAS reference frames, on the track and the alignment parameters and on the track-to-hit residuals. These concepts are introduced in this section.

3.1.1 Alignment coordinate systems

Two different frames are used to reconstruct the trajectories of the charged particles: the global coordinate system and the local coordinate system. A schematic representation of these frames is shown in Figure 3.1.

[1] During data-taking, the colliding protons at the LHC are not always the same: they are constantly replaced after some time. An LHC fill corresponds to each time the ring is filled with new protons.

[2] A hit is produced when a charged particle crosses a particular module. This produces a signal, which corresponds to a cluster of pixels (strips) in the pixel (SCT) detector, and a drift circle in the TRT.

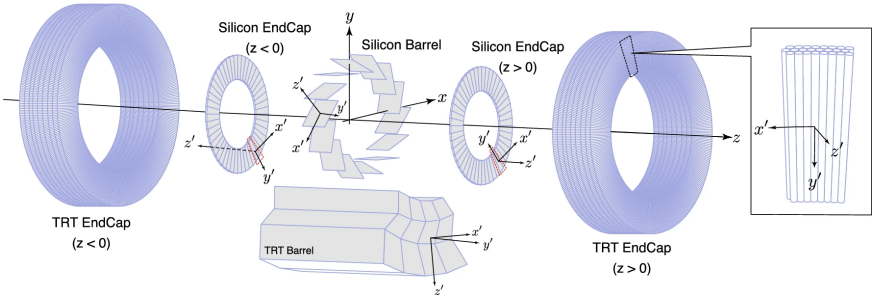


Figure 3.1: Schematic representation of the ATLAS global reference frame (x, y, z). The local reference frame (x', y', z') of the different subdetectors is also displayed [1].

3.1.1.1 Global coordinate system

The global coordinate system was previously introduced in Section 2.3.1. The origin of the global coordinate system (x, y, z) is located at the nominal interaction point of the pp collisions. The z-axis lies along the beam direction and points in the direction of the solenoid magnetic field, the x-axis points to the centre of the LHC ring, and the y-axis points upwards, towards the sky. The position of the different subdetectors that form the ATLAS ID is defined with respect to the global coordinate system.

3.1.1.2 Local coordinate system

The local coordinate system (x', y', z') is a right-handed coordinate system defined for each component of the ATLAS ID, with its centre placed at the geometrical centre of the considered module.

- For the pixel detector, the x'- and y'-axes are defined in the plane of the corresponding component, with the z'-axis pointing outside of the detector plane. The x'-axis points in the most sensitive direction of the module, which is the shorter direction, and the y' axis points along the long side of the detector.

- For the SCT modules, the local coordinate system is defined in a similar way. The x'-axis points in the most sensitive direction, which is perpendicular to the strip-orientation, and the y'-axis is oriented along the long side of the sensor. The z'-axis is perpendicular to the $x' - y'$ plane. However, it is important to mention that for the SCT there are two wafers superposed with slightly different orientations. Hence, two local coordinate systems are defined (one per wafer). This different orientation of the two wafers improves the sensitivity of the SCT along the y'-axis.

- For the TRT straws, the x'-axis points across the straws, which is the most precise measurement direction. The y'-axis points along the wire. In the barrel, the z'-axis points radially outwards (from the origin of the global frame to the straw centre). In the end-caps, the local z'-axis points outwards (parallel to the beam line).

The reconstruction of the hits in the different modules of the detector takes place in the local reference frame.

3.1.2 Track parameters

The trajectory of charged particles in a magnetic field is helicoidal, and can be fully parametrised by five *track parameters*:

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

All these parameters are defined with respect to the perigee, which is the closest point of the track to the global z-axis. d_0 and z_0 are the impact parameters defined in the transverse and longitudinal planes, respectively. ϕ_0 and θ are the azimuthal and polar angles. Lastly, q/p is the charge of the particle over its momentum and it measures the curvature of the tracks. A schematic view of these track parameters is shown in Figure 3.2.

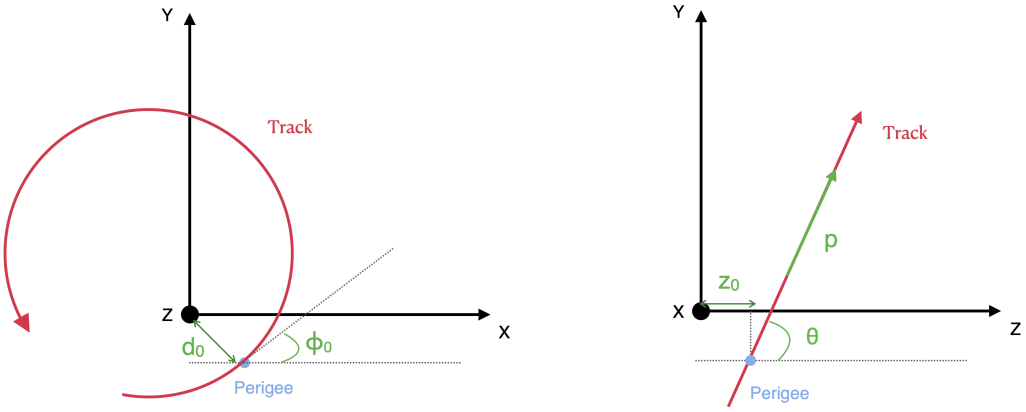


Figure 3.2: Track parameters representation at the perigee in the global reference frame.

3.1.3 Alignment parameters

There are a total of six degrees of freedom (DoF) which fully characterise the position and orientation of a rigid body. In the alignment context, these are the so-called *alignment parameters*:

$$\boldsymbol{a} = (T_x, T_y, T_z, R_x, R_y, R_z), \tag{3.2}$$

where $T_{x,y,z}$ accounts for translations with respect to the origin of the local reference frame in a particular direction and $R_{x,y,z}$ accounts for the rotations around a particular axis. Thus, each module of the ID can be aligned with these six DoF[3]. A schematic representation of these parameters can be found in Figure 3.3.

[3]Additional DoF can be introduced to parametrise concrete deformations of the different modules or substructures. In particular, for the IBL there is an additional DoF, the IBL bowing B_x, which will be introduced in Section 3.4.2.1.

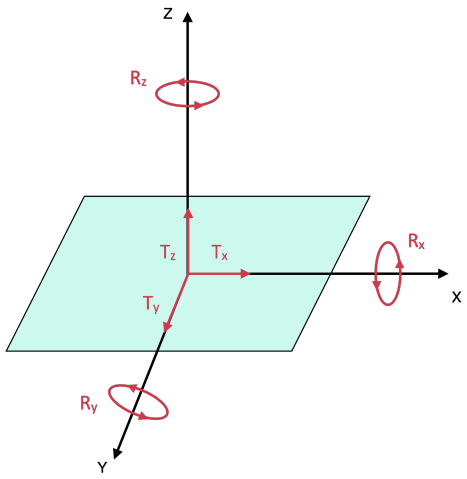


Figure 3.3: Definition of the six alignment parameters which fully characterise the position and orientation for a silicon module.

3.1.4 Residuals

A track-to-hit residual is defined as the distance between a hit in a particular detector element and the extrapolated intersection point of the reconstructed track in the same detector. Mathematically, this can be expressed as:

$$\boldsymbol{r} = \boldsymbol{m} - \boldsymbol{e}(\boldsymbol{\tau}, \boldsymbol{a}), \tag{3.3}$$

where $\boldsymbol{m}$ is the position of the measured hit in a particular detector and $\boldsymbol{e}$ is the intersection point in this detector of the reconstructed track, which depends both on the track $\boldsymbol{\tau}$ and the alignment $\boldsymbol{a}$ parameters. The residuals are defined in the local coordinate system of each subdetector. For each silicon module, a residual in the x and y coordinates is calculated. For the TRT straws, the residual is calculated in the radial direction, taking into account the time extrapolated from the trajectory of the ionised electrons to the centre of the straw.

3.2 Track-fit alignment

In a perfectly aligned detector, the residual distributions should be centred at zero, with a width corresponding to the resolution of the particular module. Any possible deviation in the residual distributions points to a misalignment of the detector. This statement constitutes the basis of the ID alignment.

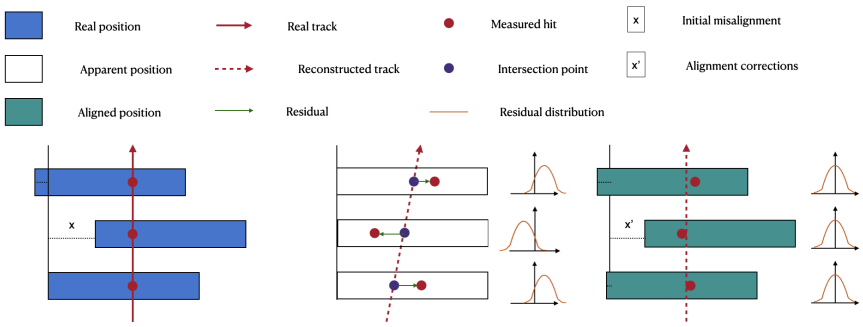


Figure 3.4: Schematic view of the alignment chain. On the left, the true position of the different modules and the true trajectory of the particle are displayed. In the middle, the initial reconstructed track from the apparent position of the detectors is shown (considering there is no misalignment). This leads to residual distributions displaced from zero. On the right, one can see the updated position of the detectors after the alignment, together with an updated reconstructed track and updated residual distributions, much centred at zero.

A schematic description of the alignment procedure can be seen in Figure 3.4. In the left panel, the real position of three different detectors can be seen (blue rectangles). The red dots correspond to the hits that the charged particle leaves in the different modules, and the red line corresponds to the true track of the particle. x stands for the deviation of the modules from their apparent position. In the middle panel, the problem can be seen: the real position of the detector modules is different from the apparent one (white rectangles). The deviation with respect to the true position of the detectors x is unknown. This deviation leads to a discrepancy between the reconstructed and the true track parameters. The green lines correspond to the residuals. These are the differences between the hits (red

dots) and the intersection points from the reconstructed track (blue dots). As it can be seen, the misalignment of the different modules leads to residual distributions displaced from zero. Then, the alignment algorithm corrects the position of the detector modules to minimize these residuals. This can be seen in the right panel: the corrections x' have been applied to the different modules (blue-green rectangles) and they are now much closer to the real position, leading to residual distributions much more centred at zero. However, the position of the detectors in the right panel still does not match perfectly with the real position (left panel). This shows that the alignment is an iterative procedure which needs some iterations to reach its maximum precision.

3.2.1 The global χ^2 algorithm

The ID alignment consists of the minimisation of a χ^2 function defined by the track-to-hit residuals. The global χ^2 algorithm [335] takes care of this procedure and obtains corrections to the position of the different modules of the ID. The χ^2 function is given by:

$$\chi^2 = \sum_t \sum_h \left(\frac{r_{t,h}(\boldsymbol{\tau}, \boldsymbol{a})}{\sigma_h} \right)^2, \tag{3.4}$$

where the t index loops over the reconstructed tracks and the h index loops over the hits associated to each track. $r_{t,h}$ is the residual associated to a particular hit from a reconstructed track and σ_h is the hit uncertainty (see Section 2.3.2).

This equation can be rewritten introducing a vector notation:

$$\chi^2 = \sum_t \boldsymbol{r}(\boldsymbol{\tau}, \boldsymbol{a})^T V^{-1} \boldsymbol{r}(\boldsymbol{\tau}, \boldsymbol{a}), \tag{3.5}$$

where $\boldsymbol{r}$ is the vector of the residuals associated to the track t and V is the covariance matrix. This covariance matrix is in general not diagonal since the multiple Coulomb scattering (MCS)[4] introduces correlations between the measured

[4]This is the elastic scattering of charged particles in a particular material due to the Coulomb interaction. In the multiple Coulomb scattering, the number of interactions is such that the angular dispersion can be modelled as Gaussian.

hits in the different modules. If these correlations were not taken into account, the covariance matrix would be diagonal, being σ_h^2 its diagonal elements.

The minimum of the χ^2 function corresponds to the true position of the detectors. Hence, this function is minimised with respect to the alignment parameters to find corrections for the different detectors:

$$\frac{d\chi^2}{d\boldsymbol{a}} = 0 \rightarrow \sum_t \left[\left(\frac{d\boldsymbol{r}}{d\boldsymbol{a}} \right)^T V^{-1} \boldsymbol{r} \right]^T + \sum_t \left[\boldsymbol{r}^T V^{-1} \left(\frac{d\boldsymbol{r}}{d\boldsymbol{a}} \right) \right] = 0, \tag{3.6}$$

where $\boldsymbol{r}$ stands for $\boldsymbol{r}(\boldsymbol{\tau}, \boldsymbol{a})$. This expression can be further simplified, taking into account that V^{-1} is symmetric:

$$2 \sum_t \left(\frac{d\boldsymbol{r}}{d\boldsymbol{a}} \right)^T V^{-1} \boldsymbol{r} = 0 \tag{3.7}$$

Now, in order to incorporate the full derivatives $(d\boldsymbol{r}/d\boldsymbol{a})$, the differential $d\boldsymbol{r}$ needs to be considered:

$$d\boldsymbol{r} = \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\tau}} d\boldsymbol{\tau} + \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{a}} d\boldsymbol{a} \tag{3.8}$$

The derivative of the residuals with respect to the alignment parameters $d\boldsymbol{r}/d\boldsymbol{a}$ is then:

$$\frac{d\boldsymbol{r}}{d\boldsymbol{a}} = \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\tau}} \frac{d\boldsymbol{\tau}}{d\boldsymbol{a}} + \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{a}} \tag{3.9}$$

This result can then be inserted in Equation 3.7 and the minimum condition turns as follows:

$$\sum_t \left(\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\tau}} \frac{d\boldsymbol{\tau}}{d\boldsymbol{a}} + \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{a}} \right)^T V^{-1} \boldsymbol{r} = 0 \tag{3.10}$$

The derivative $d\boldsymbol{\tau}/d\boldsymbol{a}$ is particularly important as it will be explained below, since it incorporates the relations between the track and the alignment parameters. This, in the end, introduces the correlations between the different modules of the detector.

As it can be seen in the last expression, the minimisation condition implies to calculate the dependence between the residuals and the track parameters $\partial \boldsymbol{r}/\partial \boldsymbol{\tau}$. For this purpose, the χ^2 can be minimised with respect to the track parameters assuming a geometry of the detector $\boldsymbol{a}$. Following a similar procedure as it was done to derive Equation 3.7, one can obtain for a particular track:

$$\frac{d\chi^2}{d\boldsymbol{\tau}} = 0 \rightarrow \left(\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\tau}}\right)^T V^{-1}\boldsymbol{r} = 0 \tag{3.11}$$

Note that here, in contrast to expression 3.7, the full derivative has been replaced by the partial derivative: Equation 3.8 is used to calculate $d\boldsymbol{r}/d\boldsymbol{\tau}$. Since the detector position should not depend on the tracks traversing it, the derivative $d\boldsymbol{a}/d\boldsymbol{\tau}$ is null. Hence, it can be proven that the full derivative $d\boldsymbol{r}/d\boldsymbol{\tau}$ is just the partial derivative with respect to the track parameters.

The initial tracks are constructed from an initial set of track parameters $\boldsymbol{\tau_0}$. Then, one needs to find corrections $\delta\tau$ for these initial track parameters to get closer to the solution $\boldsymbol{\tau} = \boldsymbol{\tau_0} + \boldsymbol{\delta\tau}$. Let us consider that the initial set of track parameters is close to the solution $\boldsymbol{\tau}$. Then, a Taylor expansion of the track parameters can be performed, up to first order, to introduce a small change towards the solution. This will of course affect the residuals:

$$\boldsymbol{r}(\boldsymbol{\tau}, \boldsymbol{a}) = \boldsymbol{r}(\boldsymbol{\tau_0}, \boldsymbol{a}) + \left(\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\tau}}\right)\Bigg|_{\boldsymbol{\tau}=\boldsymbol{\tau_0}} \delta\boldsymbol{\tau} \tag{3.12}$$

This Taylor expansion of the residuals can be introduced in Equation 3.11, and defining $E = \left(\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\tau}}\right)\big|_{\tau=\tau_0}$ this is:

$$E^T V^{-1}\boldsymbol{r}(\boldsymbol{\tau_0}, \boldsymbol{a}) + E^T V^{-1} E\delta\boldsymbol{\tau} = 0 \tag{3.13}$$

From this, the corrections for the track parameters $\delta\boldsymbol{\tau}$ can be obtained. Let us define the matrix $\mathcal{X} = E^T V^{-1} E$ and the vector $\boldsymbol{t} = E^T V^{-1}\boldsymbol{r}(\boldsymbol{\tau_0}, \boldsymbol{a})$. Then, the corrections for the track parameters can be expressed in a compact notation:

$$\delta\boldsymbol{\tau} = -\mathcal{X}^{-1}\boldsymbol{t} \tag{3.14}$$

From this notation it can be proven that $\mathcal{X}^{-1}$ is indeed the covariance matrix of the track parameters [336].

Now the derivative $d\boldsymbol{\tau}/d\boldsymbol{a}$ can be calculated. Recovering $\boldsymbol{\tau} = \boldsymbol{\tau_0} - X^{-1}\boldsymbol{t}$, this is just:

$$\frac{d\boldsymbol{\tau}}{d\boldsymbol{a}} = -\mathcal{X}^{-1}E^T V^{-1}\frac{\partial \boldsymbol{r}(\boldsymbol{\tau_0}, \boldsymbol{a})}{\partial \boldsymbol{a}}, \tag{3.15}$$

where the vector $\boldsymbol{t}$ has been expanded.

Looking back at expression 3.10, the derivative $d\boldsymbol{\tau}/d\boldsymbol{a}$ can now be introduced considering that the alignment corrections are calculated from the initial set of track parameters. Then, the minimum condition is found to be:

$$\sum_t \left(\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{a}}\right)^T W\boldsymbol{r} = 0, \tag{3.16}$$

where the so-called *weight matrix* $W = (I - E(E^T V^{-1}E)^{-1}E^T V^{-1})^T V^{-1}$ has been introduced to make the notation more compact. At this stage, an analogue process can be applied but to the alignment parameters, considering a small change to reach the solution $\boldsymbol{a} = \boldsymbol{a_0} + \delta\boldsymbol{a}$. Afterwards, the residuals can be expanded in a similar way:

$$\boldsymbol{r}(\boldsymbol{\tau_0}, \boldsymbol{a}) = \boldsymbol{r}(\boldsymbol{\tau_0}, \boldsymbol{a_0}) + \left(\frac{\partial \boldsymbol{r}(\boldsymbol{\tau_0}, \boldsymbol{a})}{\partial \boldsymbol{a}}\right)\Bigg|_{\boldsymbol{a}=\boldsymbol{a_0}} \delta\boldsymbol{a} \tag{3.17}$$

Note that here $\boldsymbol{\tau_0}$ is used. This is justified since these are the initial track parameters, which are derived with the initial geometry $\boldsymbol{a_0}$.

Finally, the corrections $\delta\boldsymbol{a}$ can be obtained evaluating $\partial\boldsymbol{r}/\partial\boldsymbol{a}$ from Equation 3.17 and inserting the result in Equation 3.16:

$$\delta\boldsymbol{a} = -\mathcal{M}^{-1}\boldsymbol{v}, \tag{3.18}$$

with $\mathcal{M} = \sum_t (\partial \boldsymbol{r_0}/\partial \boldsymbol{a})^T W (\partial \boldsymbol{r_0}/\partial \boldsymbol{a})$ and $\boldsymbol{v} = \sum_t (\partial \boldsymbol{r_0}/\partial \boldsymbol{a})^T W \boldsymbol{r_0}$ (being $\boldsymbol{r_0} \equiv \boldsymbol{r}(\boldsymbol{\tau_0}, \boldsymbol{a_0})$ and $(\partial \boldsymbol{r_0}/\partial \boldsymbol{a}) \equiv (\partial \boldsymbol{r}/\partial \boldsymbol{a})\big|_{\boldsymbol{\tau}=\boldsymbol{\tau_0}, \boldsymbol{a}=\boldsymbol{a_0}}$). These are the so-called *big matrix* and *big vector*, respectively. In a very similar way to $\mathcal{X}^{-1}$, it can be proven that $\mathcal{M}^{-1}$ is the covariance matrix of the alignment parameters [336].

This formalism explains how to obtain corrections for the alignment parameters by minimising the χ^2 function, which means minimising the residuals. However, it has been assumed that the initial position of the detector is not far from the solution. This is not necessarily the case, and the alignment algorithm would need some iterations to find the real position of the detector. The corrections for a given iteration $\delta \boldsymbol{a}_n$ are computed and the new set of alignment parameters is obtained as:

$$\boldsymbol{a_n} = \boldsymbol{a_{n-1}} + \delta \boldsymbol{a}_n \tag{3.19}$$

In a practical way, the alignment algorithm is divided in two differentiated steps:

- **Accumulation**: Given an initial geometry, the set of initial tracks is reconstructed. This is the so-called track reconstruction. Then, the partial derivatives of the residuals with respect to the track parameters and with respect to the alignment parameters are evaluated. With this information, the big matrix and the big vector are constructed.

- **Solving**: This step takes care of the inversion of the alignment matrix $\mathcal{M}$. In case the number of aligned degrees of freedom is low[5] this step can be very quick. Otherwise, the inversion of the big matrix can be computationally very heavy if the chosen number of alignable structures is large. After this step, the alignment corrections are provided.

As mentioned above, the $d\boldsymbol{\tau}/d\boldsymbol{a}$ is a very important term because it includes the correlations among the different modules of the detector. In the global χ^2 this term is evaluated as it has been described above. Alternatively, the local χ^2 algorithm [337, 338] assumes this term is zero, neglecting the possible correlations between the modules. With this assumption, the big matrix becomes diagonal, making the inversion of the matrix (i.e. the solving) much less CPU demanding. In general, the global χ^2 is preferred since it takes into account these correlations and thus requires less iterations to reach convergence. However, in a practical way the

[5] A description of the different alignment configurations will be given in Section 3.4.

local χ^2 can be used instead when the number of degrees of freedom is very large (e.g. when all the TRT modules are aligned).

3.2.2 Alignment with constraints

As it will be described in Section 3.3, sometimes it is useful to include prior knowledge in the alignment algorithm as extra terms in the χ^2 function. This implies the introduction of constraints in the alignment chain. These extra terms can be grouped in two categories: track parameter constraints (e.g. beam spot position[6]) and alignment corrections constraints (e.g. "soft mode" requirements[7]). An expanded χ^2 function would be given by:

$$\chi^2 = \sum_t [\boldsymbol{r}(\boldsymbol{\tau}, \boldsymbol{a})^T V^{-1} \boldsymbol{r}(\boldsymbol{\tau}, \boldsymbol{a}) + R^T(t) V_t^{-1} R(t)] + R'^T(a) V_a^{-1} R(a), \qquad (3.20)$$

where the second term contains the track parameter constraints and the third term contains the alignment corrections constraints. Here $R(t)$ ($R(a)$) depends only on the track (alignment) parameters and determines their associated constraint, and V_t (V_a) determines the tolerance of this constraint (how restrictive it is).

The use of alignment parameter constraints is of particular interest for two reasons: Firstly, one can make use of the detector assembly survey data and the mechanical tolerances of the detector structure into the alignment, and second these constraints enter in the diagonal terms of the big matrix $\mathcal{M}$ and help to remove the potential singular modes.

[6] This is the position where the collisions of the ATLAS detector take place. Is is expected that the particles are produced close to the beam spot position. This information can be used to set important constraints on the transverse impact parameter d_0.

[7] Sometimes the amount of hits in a particular module is very low and the amount of statistics is insufficient to derive an alignment correction. This can lead to very large and unrealistic alignment corrections which are suppressed including penalty terms, also known as "soft mode" requirements.

3.3 Weak modes

The track-based alignment described in Section 3.2.1 is not sensitive to certain kind of movements or deformations of the ID known as *weak modes*. The weak modes alter the track parameter reconstruction, predominantly d_0, z_0, and the charge over momentum q/p_T. These weak modes affect the alignment procedure since they are near-to-singular solutions of the χ^2 minimisation. They leave the χ^2 unaltered, and thus the alignment algorithm cannot correct for this kind of "invisible" distortions.

The weak modes can be classified mainly in two categories: charge-symmetric distortions and charge-asymmetric distortions. This is illustrated in Figure 3.5, where the reconstructed tracks of two opposite-sign particles is displayed. These are distorted in a different way for a charge-asymmetric distortion (left figure) and in a similar way for a charge-symmetric distortion (right figure).

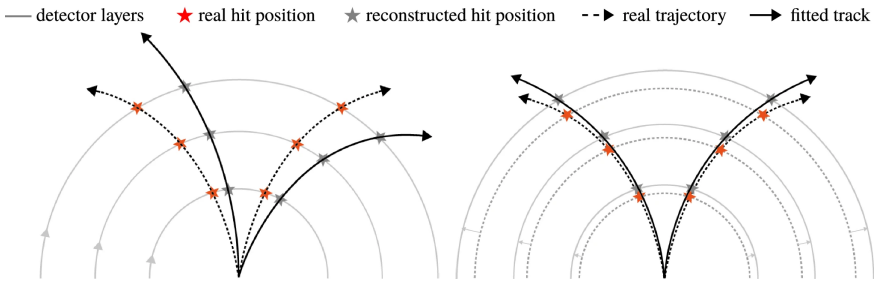


Figure 3.5: Schematic representation of two different types of weak modes. The left picture describes a curl of the detector (charge-asymmmetric), whereas the right picture describes an expansion of the detector (charge-symmetric). The solid gray lines represent the position of the detector, and the arrows indicate the effect of the weak modes. In the right picture, the dashed gray line corresponds to the expected position of the detector, which does not correspond to the true position. The red stars represent the real hit positions in the detector, whereas the gray stars represent the reconstructed hit positions. The black lines represent the trajectories of two opposite-sign particles, being the dashed lines the real trajectories and the solid lines the reconstructed trajectories [1].

Amending these weak modes is a fundamental task since they would introduce biases in the momentum scale of the ID tracks, which would lead to biased physics measurements. Fortunately, as it was described Section 3.2.2, the χ^2 function can be expanded with extra constraints which allow the correction of these weak modes.

First of all, one needs to describe how the weak modes affect the trajectory of the charged particles. The trajectory of a charged particle in a magnetic field is helicoidal, and its transverse momentum in a cylindrical detector of radius R is:

$$p_{\mathrm{T}}[\mathrm{GeV}] = 0.3q[\mathrm{e}]B[\mathrm{T}]\rho[\mathrm{m}] = 0.3qB\left(\frac{R^2}{8s} + \frac{s}{2}\right), \tag{3.21}$$

where q is the charge of the particle in electron units, B is the magnetic field in Teslas, ρ is the radius of the trajectory in metres and s is the sagitta of the trajectory. The sagitta determines how much the track is deviated from a straight trajectory: the larger the sagitta, the larger the bowing of the track. For high-p_T particles, s is much lower than R, and the previous expression can be simplified to:

$$p_{\mathrm{T}} = 0.3qB\frac{R^2}{8s} \tag{3.22}$$

From Equation 3.22 it can be seen that the detector distortions can enter in the reconstructed momentum of the particles in three different manners: sagitta, radial and magnetic field distortions. Figure 3.5 (right) represents a radial distortion: a radial expansion of the detector layers depending linearly on the radius. Detailed studies on radial distortions in Run-2 can be found in Ref. [1]. Magnetic field distortions refer to the relative misalignment between the position of the tracker and the solenoid magnetic field. These were analysed in Run-1 [339]. Among these three, this thesis focuses on the study of sagitta distortions (see Section 3.5.2.2).

In order to correct for this kind of distortion, one can exploit some prior knowledge from particular known events. For instance, the decay of well-known resonances (so-called *standard candles*), can be used to introduce constraints on the alignment algorithm[8]. Concretely, dimuon resonances such as $Z \to \mu^+\mu^-$ are frequently used since the trajectories of the two decaying muons are clearly

[8]There are other possibilities, such as the E/p method [339].

related. There are other dimuon resonances (such as J/ψ and Υ), but the Z boson is preferred due to the high momentum of the decaying muons.

3.3.1 Sagitta bias

As stated above, the sagitta distortions can affect the reconstructed transverse momentum of charged particles. This distortion is charge-asymmetric, and would introduce a bias $\delta_{sagitta}$ in the track parameter q/p_T ($s \propto q/p_T$, see Equation 3.22) of the form:

$$\frac{q}{p_T^{reco}} = \frac{q}{p_T^{true}} + \delta_{sagitta} \tag{3.23}$$

For convenience, the previous expression can be rewritten as:

$$p_T^{reco} = p_T^{true}[1 + qp_T^{true}\delta_{sagitta}]^{-1} \tag{3.24}$$

It is important to note that this kind of distortion will not affect the polar angle. Thus, the previous expression can be extended to the total momentum of the particle as:

$$p^{reco} = p^{true}[1 + qp_T^{true}\delta_{sagitta}]^{-1} \tag{3.25}$$

The invariant mass of the Z boson is reconstructed from the two decaying muons, which are affected by the sagitta bias. The invariant mass of two highly relativistic opposite-charge particles is given by:

$$m_{\mu\mu}^2 = 2|p^+||p^-|[1 - \cos\Delta\phi], \tag{3.26}$$

where p^+ (p^-) is the momentum of the positive (negative) muon, and ϕ is the opening angle between them. If one considers a sagitta distortion, the momentum of the muons would be modified according to Equation 3.25, and then the invariant mass is approximately given by:

$$m^2_{\mu\mu} \simeq m^2_{\mu\mu,reco}[1 + p_T^{+,reco}\delta_{sagitta}(\eta^+, \phi^+) - p_T^{reco}\delta_{sagitta}(\eta^-, \phi^-)] \quad (3.27)$$

Here, the reconstructed invariant mass:

$$m^2_{\mu\mu,reco} = 2|p^{+,reco}||p^{-,reco}|[1 - \cos\Delta\phi^{reco}], \quad (3.28)$$

has been introduced, and the sagitta distortion has been assumed to be small to perform the approximation. $\delta_{sagitta}(\eta^+, \phi^+)$ ($\delta_{sagitta}(\eta^-, \phi^-)$) corresponds to the sagitta distortion measured for the positive (negative) muon.

This kind of sagitta distortion can be measured in specific regions of the detector. The whole space of the detector can be parametrised in terms of η and ϕ, and hence the sagitta biases are calculated as 2D maps in the form $\delta_{sagitta}(\eta, \phi)$[9].

From Equation 3.27, the sagitta corrections can be calculated by comparing the reconstructed mass with the true mass of the Z boson (taken from [12]). The obtained corrections could be introduced either in the positive or the negative muon, and there is no way to know which track is biased in advance. An option to overcome this problem is to initially consider that the sagitta distortion comes from the two muons in a symmetric way, and thus $1/2$ of the computed corrections are assigned for each muon track.

In order to determine the sagitta maps an iterative procedure is followed. In the first iteration, the $Z \to \mu^+\mu^-$ process is reconstructed from the momenta of the two muons. The reconstructed invariant mass is compared with the true mass of the Z boson and the information is used to derive correction maps for $\delta_{sagitta}(\eta, \phi)$. These values are stored and used to recompute the momenta of the two muons. Then, the reconstructed invariant mass is recomputed and compared with the true mass of the Z boson to obtain new sagitta corrections, which are added to the

[9]This method, as described by Equation 3.27, is only sensitive to relative sagitta biases in different sectors of the detector. The presence of a global sagitta bias can be corrected using an alternative method, comparing the p_{T} spectrum of the positive and negative muons [340]. The E/p method is also sensitive to a global sagitta bias. Nevertheless, the presence of a global sagitta bias has been checked to be very small during Run-2 [1].

previous correction maps. The process is repeated until convergence is reached, i.e. the modifications in the sagitta correction maps become negligible.

3.3.2 Track parameter bias

Additionally, the same information from the Z boson standard candle can be used to correct possible biases in the track parameters d_0 and z_0. Since two muons coming from the same particle should originate in the same primary vertex, their impact parameters should be equal. The 2D maps associated to track parameter biases can be computed as:

$$\delta_{d_0} = d_0^+ - d_0^- \quad , \quad \delta_{z_0} = z_0^+ - z_0^- \tag{3.29}$$

If there is any deviation from zero in these maps, it can be inferred there is a systematic distortion in the ID. The impact parameter biases are extracted by iteratively[10] fitting the distribution of impact parameters relative to the primary vertex with a Gaussian function within a $\pm 2\sigma$ range until the fitted μ and σ are stable within 1%. The resulting value of the Gaussian mean (μ) represents the estimate of the impact parameter bias.

3.4 Alignment procedure

This section describes the alignment procedure focusing on the challenges during Run-2. Firstly, the sequential steps taken through the alignment chain are introduced. Afterwards, the operations performed in the ID alignment context after data-taking are described.

3.4.1 Alignment levels

The alignment of the ID is always performed in a hierarchical structure: from large structures to silicon modules and TRT straws. In that sense, the alignment can be divided in three different levels:

[10] This procedure is adopted as the impact parameter distributions have long tails.

- **Level 11**: The largest structures are aligned in this level[11]. The SCT and TRT barrel and end-caps are considered as individual components, whereas the pixel detector on the whole (all barrel and end-caps) is treated as an individual structure since its components are assembled together. Finally, the IBL is treated as a single structure since it is attached to the beam pipe, not to the pixel detector. This is the simplest alignment level, and due to the large amount of statistics in the different aligned detectors and the low number of aligned DoF the number of iterations to reach convergence is, in general, low.

- **Level 2**: The different barrel layers and end-cap disks are treated independently. Since the IBL is a single layer, it is treated equivalently to Level 11.

- **Level 3**: The different silicon modules and TRT wires are aligned independently. This is the most detailed and most complex alignment level. The tracks traversing the different modules are much lower than those traversing the whole structures, and hence a larger number of iterations is required to reach convergence.

It is crucial to perform the alignment of the ID in this increasing level of complexity. Otherwise, the misalignment of a large structure, for instance a vertical displacement of a pixel end-cap disk, could be corrected as unrealistic movements of some of its modules (the ones whose residual distributions are affected by this displacement), which will distort severely the shape of the end-cap disk.

The number of aligned structures and the number of DoF for the three previous alignment levels are displayed in Table 3.1. These three are the most basic alignment levels, but any other possibility can be considered. Among these, the so-called *Level 16* was found to be essential during Run-2. This is very similar to Level 11, but the IBL staves are aligned individually. The importance of this additional alignment level will be described in Section 3.4.2.

[11] During Run-1, this alignment level was named Level 1. However, after the introduction of the IBL in Run-2, this was renamed to Level 11 to distinguish the new alignment configuration (including the IBL) from the previous one.

	Alignable structures				DoF
	IBL	Pixel	SCT	TRT	Number
Level 11	1	1	3	3	48+1
Level 2	1	9	22	176	1248
Level 3	280	1744	4088	350848	738368

Table 3.1: Alignable structures and DoF for each subdetector in the three main alignment levels. The extra DoF considered in the Level 11 stands for the IBL bowing (see Section 3.4.2.1).

3.4.2 Calibration Loop alignment in Run-2

In order to ensure data of the highest possible quality, the alignment algorithm must correct for the different movements of the detector whenever these occur. This can happen from year-to-year, run-to-run or even within the same run[12]. The *Calibration Loop* (CL) [341] is responsible for correcting the most important movements after data collection. It corrects the position of the detector before bulk data-reprocessing[13]. This is done using a subset of the data, called the *express stream*, which is processed at the Tier-0 [342] at CERN, and the alignment corrections are provided within 48 hours.

Due to the movements within a particular run, the CL needs to provide a dynamic alignment. There are mainly two kind of movements which take place during an LHC fill: the pixel vertical displacement and the bowing of the IBL. One of the contributions of this work to the alignment of the ID has been a detailed study on the pixel vertical movements, and will be further described in Section 3.5.1.

3.4.2.1 Temperature-dependent IBL bowing

It was observed that during each LHC fill some misalignments were introduced in the detector. Concretely, the movements related with the IBL were especially

[12]Data at the LHC is recorded in bunches named runs, which can take from some minutes to hours of data-taking.

[13]The bulk includes the full dataset. The computing system used in the reprocessing of this data is explained in Section 2.6.

important. The IBL staves are not stable within a run as they bend with respect to their flat designed shape. These distortions are temperature-dependent, as it can be seen in Figure 3.6.

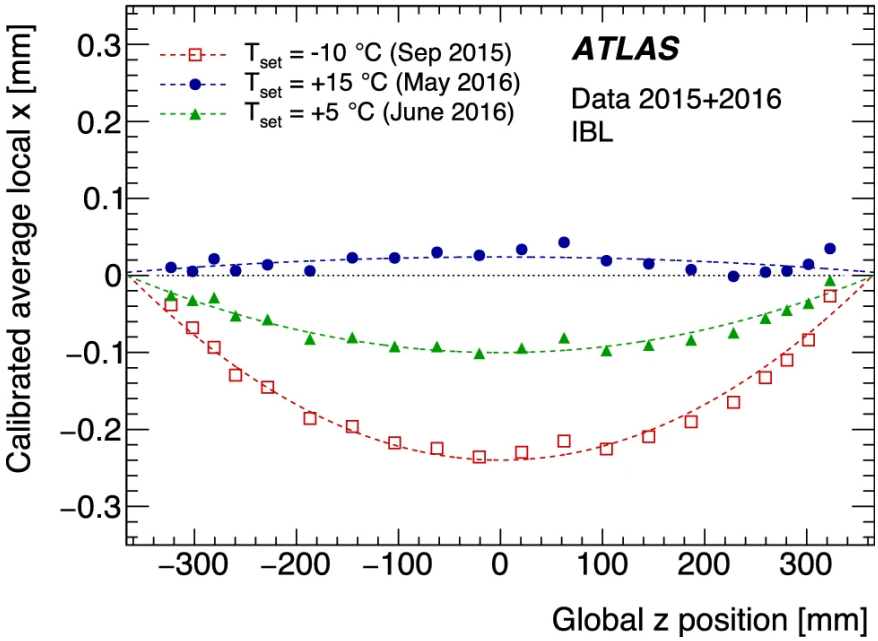


Figure 3.6: IBL local-x position averaged over all the IBL staves for different data-taking periods: September 2015 (red open squares), May 2016 (solid blue circles) and June 2016 (solid green triangles). As it can be seen, each period is associated to a different operating temperature and the bowing of the IBL staves shows a clear dependence with this temperature. The dashed coloured lines correspond to the computed bowing correction from the polynomial function described in the text, whereas the dashed black line corresponds to the nominal position [1].

The underlying reason for this kind of motion is a mismatch between the coefficient of thermal expansion of the different materials which form the IBL staves. Additionally, the resulting bending is not constant within each Run-2 LHC fill. The high luminosity induces an increase in the low-voltage (LV) currents in the IBL module front-end electronics. This increase was traced back to radiation-induced leakage current in transistors [343] and depends on the total ionisation dose. These

variations in the LV currents caused an increase in the IBL module temperatures, which in turn resulted in changes in the IBL distortions on short timescales that need to be corrected.

The staves are fixed at each end, and thus this kind of distortion (*bowing*) alters the local-x residuals in the middle of each stave. This motion could be fixed with a L3 alignment and the usual six alignment parameters. However, updating the L3 corrections is computationally very heavy due to the large number of DoF involved. To solve this problem one can use a model [344] which describes this kind of distortions and relates them with the temperature:

$$\delta x(z) = B - M\left(\frac{z^2 - z_0^2}{z_0^2}\right), \tag{3.30}$$

where z is the position of the module along the global z-axis, z_0 = 366.5 mm corresponds to the fixing points at each end of the stave, B is a global translation in the z-axis which affects the whole stave and M is the bowing magnitude, which is the magnitude of the distortion at the stave centre. The relation of this distortion with the temperature has been studied in detail [344] and a linear dependence has been found. Concretely, the temperature gradient is found to be $dM/dT = (-10.6 \pm 0.7)$ μm/K.

This second order polynomial function parametrises correctly the IBL distortion, and hence it is included as a new DoF associated to the IBL, the bowing B_x. It allows to correct for this motion at Level 11, making the task much simpler. It has to be stated that B_x is quite correlated with R_z, and thus in order to correct properly for the IBL bowing the rotations around the z-axis are usually not aligned at the same time.

3.4.2.2 Calibration Loop alignment iterations

The CL needs to correct for the two kind of distortions which take place within each LHC fill. The importance of the Level 16 alignment is now clear: it allows to correct the distortions associated to the IBL bowing. In the CL, two iterations at Level 11 are performed first to correct for possible global movements of the detector and particularly for the pixel vertical displacement (see Section 3.5.1). Secondly,

Level	IBL	Pixel	SCT barrel	SCT end-caps	TRT barrel & end-caps
Level 11	All except R_z	All	fixed	All except T_z	All except T_z
Level 16	Only B_x	fixed	fixed	fixed	fixed

Table 3.2: Alignment parameters corrected in the CL. The SCT barrel is kept fixed at Level 11 and used as a reference. T_z is not aligned for the SCT end-caps and the TRT due to the lack of sensibility. R_z is not aligned for the IBL since it is correlated with the bowing. At Level 16, only the IBL bowing of the different staves is aligned.

two iterations at Level 16 are performed to correct for the IBL bowing of the different staves. Each run is divided in intervals which are short enough to capture the motion of the subdetectors and also long enough to include enough statistics to compute the corrections. These intervals correspond to groups of lumiblocks[14] (LBs). The first 60 LBs are divided in three groups of 20, which include the biggest movements of the pixel detector due to the increasing temperature. Then, the rest of the run is divided in groups of 100 LBs[15]. Details on the alignment levels which take place in the CL can be found in Table 3.2. As it can be seen, the SCT is kept fixed during the CL due to its large stability during long data-taking periods, which makes it a good reference for the alignment procedure.

3.5 Contributions to the ID alignment during Run-2

The contributions of this work to the ID alignment can be divided in two different parts. The first part constitutes a detailed study on the pixel vertical movements which take place during each LHC fill and their relation with the luminosity. The second part is based on providing the most precise knowledge of the detector position for different data-taking periods, which involve both heavy ions and pp collisions.

[14] Each lumiblock corresponds to $\sim$1 minute of data-taking.

[15] In case the last LB group has less than 100 LB, this group is absorbed in the previous one to ensure there is no group with very low statistics.

3.5.1 Pixel vertical displacement

In addition to the IBL bowing, the vertical displacement of the pixel detector also played a special role in the ID alignment during each LHC fill in Run-2. This kind of motion was already seen in Run-1 to a lesser extent, but it only affected a small fraction of the data and hence no correction was applied. However, during Run-2, due to the increased luminosity in each LHC fill, this motion became one of the most important deformations of the detector. Therefore, it had to be corrected run-by-run to ensure high data-quality.

In this thesis, a detailed study of these movements was performed. It has been observed that the pixel detector rose at the beginning of each LHC fill, and then slowly went down until the end of the fill. This effect is displayed in Figure 3.7. A strong relation between this motion and the luminosity is clearly visible. For this study, a run with high statistics and with stable luminosity (i.e. no abrupt changes) was chosen. The luminosity of this run presented the usual pattern in each Run-2 LHC fill: the peak luminosity is reached at the beginning and then it slowly goes down until the end of the run. This is due to the gradual loss of beam particles after the fill. This effect is mitigated using *luminosity levelling* techniques[16].

[16] In luminosity levelling, the natural decay of the luminosity is pre-empted and the luminosity is spoilt initially with respect to the nominal by adjusting beam parameters. Then, as the luminosity decays, it is spoilt less and less in order to maintain the luminosity of the run as long as possible [345].

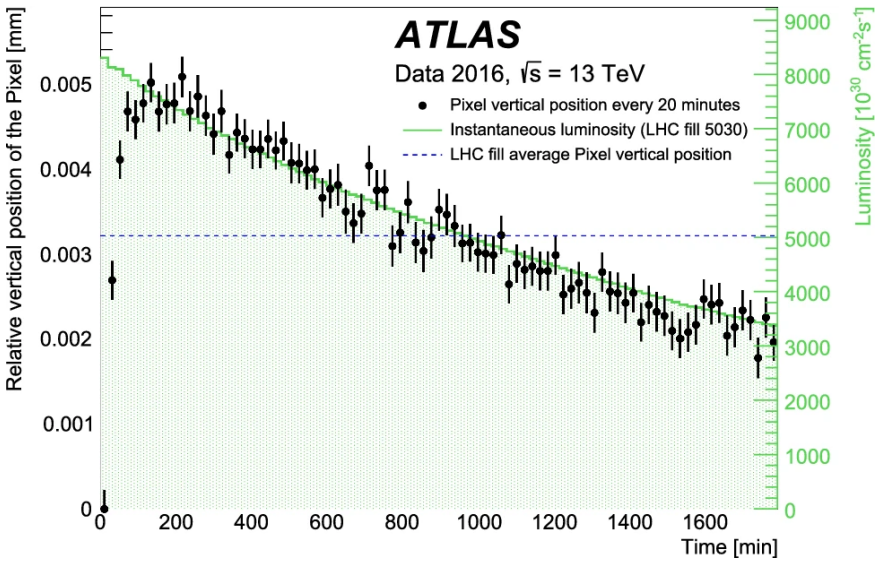


Figure 3.7: Short time scale movements of the pixel detector during a fill. Every time the pixel detector is switched on it rises fast during a transitory regime until it reaches a maximum. Afterwards (in a high luminosity LHC fill) the pixel detector goes down following closely the behaviour of the luminosity of the beams. The black points correspond to the vertical position of the pixel detector as a function of time (every 20 min) and relative to the initial measurement, whereas the instantaneous luminosity is plotted in the green area histogram. The relative normalisation of both histograms allows a straight comparison. The dashed blue line corresponds to the average vertical position that would be computed from the entire range. The uncertainty bars are purely statistical.

In the figure the relative vertical position of the pixel detector in comparison with its baseline position is seen. An updated correction for the pixel vertical position was computed every 20 LBs (∼20 minutes). It can be seen how the detector rises rapidly approximately 5 μm at the beginning of the run, and then there is a slow drift down up to 3 μm as the luminosity goes down. In this plot the average correction using the information from the whole run is also displayed (blue dashed line). This explains why a global correction to the whole run is not enough to describe this motion, and hence it needs to be corrected dynamically within the LHC fill.

The underlying reason for this pixel vertical displacement has been studied in detail [11]. The most probable explanation is related with the bi-phase cooling liquid of the pixel detector[17]. At the beginning of each LHC fill, the pixel detector is switched on and the temperature is increased due to the strong thermal coupling between the modules and the evaporative cooling system and the high data transfer rate, which is linked to the beam luminosity. The temperature in the pixel detector volume keeps rising gradually during the first 60 minutes. Part of the cooling liquid is evaporated, and the mass load in the pixel detector is decreased, which causes the pixel detector to rise. After this initial rise, the instantaneous luminosity decreases over the LHC fill, and thus the occupancy. This decreases the module temperature, which implies a phase transition of the cooling system and then a slow drift in the opposite direction to the initial movement. The speed of the slow drift at the end of the run depends on the peak luminosity, as it can be seen in Figure 3.8.

[17]The evaporative cooling system of the pixel detector uses a gas heater in the exhaust line that heats non-evaporated C_3F_8 cooling liquid plus gas to a constant temperature of 20°C.

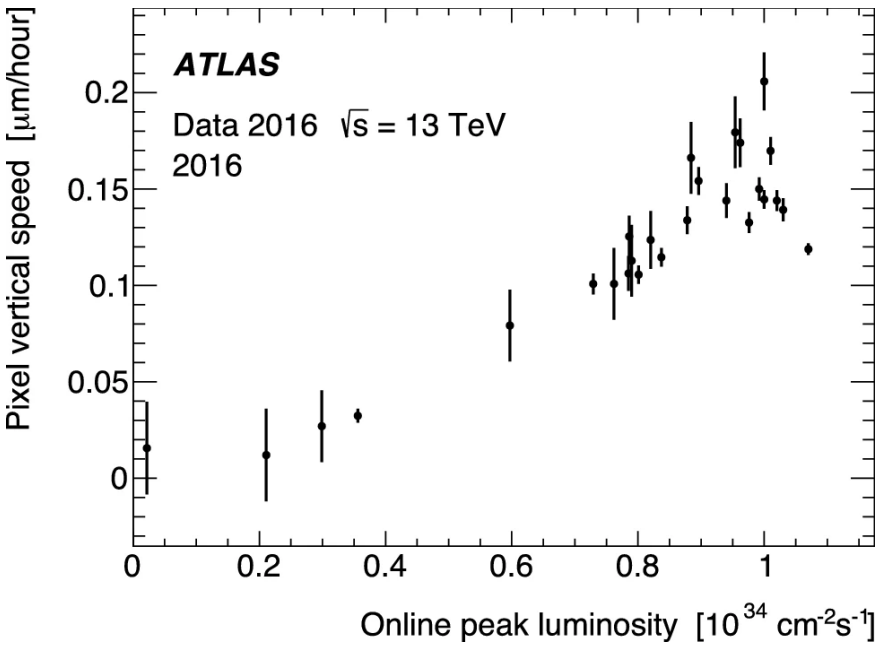


Figure 3.8: Vertical velocity of the pixel drift as a function of the peak luminosity of the considered LHC fill. This vertical velocity is obtained as the average speed during the whole fill, excluding the first 60 minutes which correspond to an increasing temperature. Only fills longer than 8 hours were used [1].

3.5.2 Determination of a baseline alignment

The CL is useful to correct the different movements between subdetectors arising during data-taking. These corrections are applied always on top of a set of baseline alignment constants, which need to be derived whenever the conditions of the detector have changed: after a large shutdown, a change in the machinery, or any other drastic change. It is crucial that these constants describe as accurately as possible the real geometry of the detector because the CL cannot correct by itself very drastic changes of the detector geometry to the desired level of precision, nor track parameter biases which leave the χ^2 unaltered. Indeed, if the baseline alignment constants are biased, this will be propagated to the whole data-taking period, reducing drastically the quality of the recorded data.

The derivation of these baseline alignment constants is usually done by selecting a long run, with high statistics, which takes place at the beginning of the data-taking period under consideration. A detailed alignment is performed and a new set of baseline alignment constants is derived. These constants are valid during long periods of time, as long as the detector is stable, and they constitute the baseline for the daily alignment.

In this section, the work in the derivation of baseline alignment constants for three different scenarios will be discussed: the Heavy Ions (HI) collision periods corresponding to the years 2015 and 2016, and the pp collisions corresponding to 2018.

3.5.2.1 Alignment of 2015 and 2016 heavy ions collisions

Two different HI data-taking periods have been analysed in this thesis:

- HI 2015: $Pb - Pb$ collisions which took place between November and December of 2015. This period began right after the end of pp collisions taken during 2015.
- HI 2016: $p-Pb$ collisions which took place between November and December of 2016. This period began approximately 15 days after the end of the 2016 pp collisions period.

Baseline alignment in 2015

The alignment of the ID for the HI 2015 data-taking period was particularly challenging because of an offset in the detector description which affected mainly the IBL[18]. The rotation of the different staves of the IBL was off by 30 mrad from the y-axis. A description of the problem can be found in Figure 3.9. Additionally, the global rotation of the whole IBL was off 40 mrad from the z-axis[19]. This needs to be

[18]The description of the detector geometry (the so-called Geo-2015 in this case, see Figure 3.9) is used as a starting point for the alignment chain.

[19]A global rotation on the IBL can be easily introduced in the baseline constants. A 40 mrad global R_z rotation was introduced manually so that the starting point for the alignment chain was much closer to the real position. For the R_y rotation of the different staves, the rotation needs to be

corrected prior to the study of any other possible misalignment. For that purpose, a pp 5 TeV run was chosen as a pilot run for the alignment with an integrated luminosity of 13.3 pb^{-1}. Since the HI collisions in 2015 began right after the end of the pp campaign, the description of the detector at the end of the pp campaign is valid for this purpose.

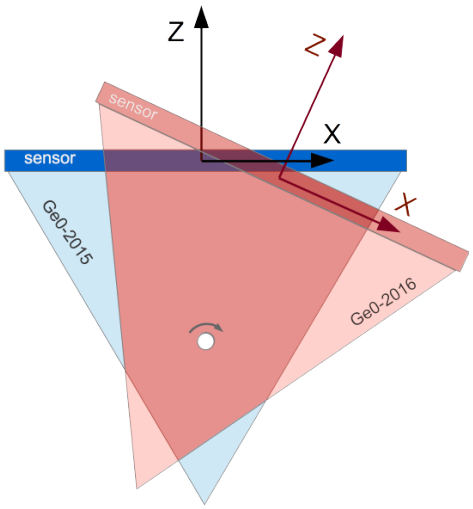


Figure 3.9: Schematic representation of an IBL stave. The predicted position from the initial geometry description is depicted in blue, whereas the actual position is depicted in red. A clear rotation in the actual position can be seen around the y-axis with respect to the predicted position.

Firstly, the global R_z rotation was corrected performing a total of nine iterations at Level 11. Figure 3.10 (a) shows the evolution of the R_z corrections after different iterations. It can be seen that the corrections are quite large but the convergence is quite fast. Details on the configuration can be found in Table 3.3. The IBL bowing is kept fixed in this step due to its large correlation with the rotations around the z-axis. Nonetheless, after reaching convergence for the IBL R_z, four extra iterations were performed at Level 11 correcting for possible bowing misalignments. The

included module by module, which is technically more difficult. Therefore, it was decided to correct the staves rotation directly in the alignment chain.

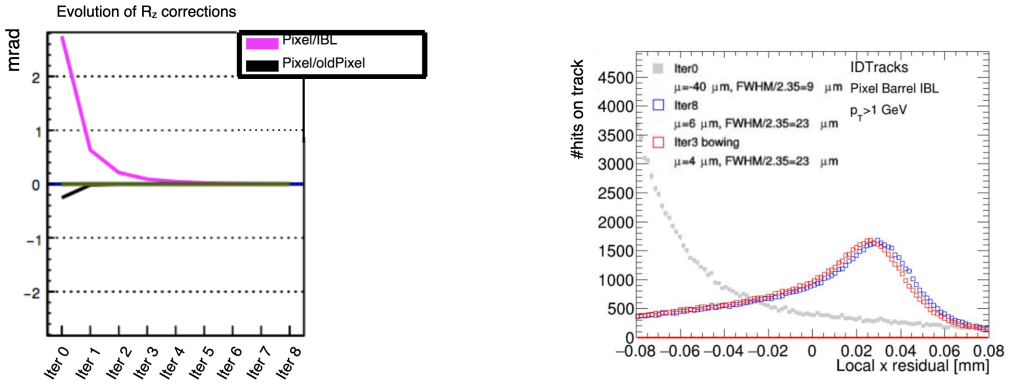
Level	Iterations	IBL	Pixel	SCT	TRT
11	9	All except B_x	All	Barrel: fixed End-caps: T_x, T_y	T_x, T_y
11	4	All except R_z	All	Barrel: fixed End-caps: T_x, T_y, R_z	T_x, T_y, R_z
16	4	T_x, T_y, R_y, R_z	fixed	fixed	fixed
16	4	B_x	fixed	fixed	fixed
2	3	All except B_x	Barrel: All End-caps: T_x, T_y, R_z	fixed	fixed
16	2	T_x, T_y, B_x	fixed	fixed	fixed
3	3	T_x, T_y, R_z	fixed	fixed	fixed

Table 3.3: Steps taken in the derivation of the baseline constants for the HI 2015 campaign.

residual distribution of the IBL in the x-direction after this Level 11 alignment can be seen in Figure 3.10 (b). The residuals clearly improved with respect to the initial position, but still there is a clear bias of $\sim$3 μm.

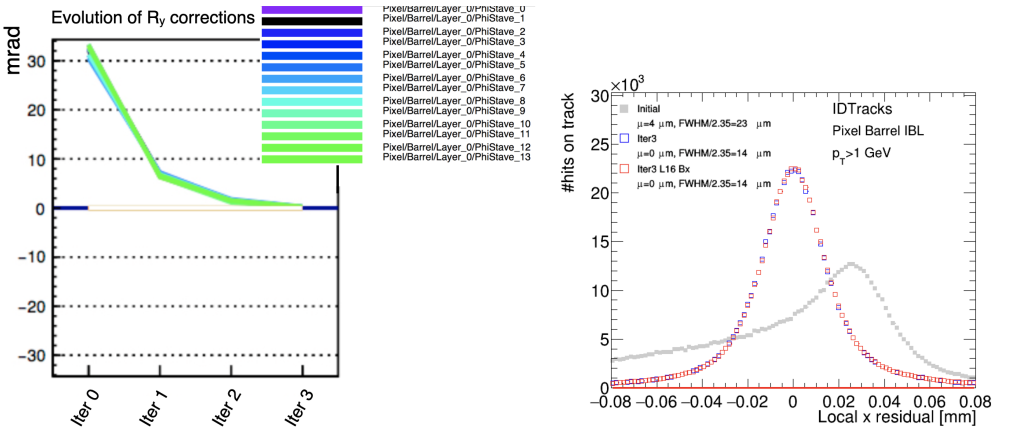
The second step is to correct the rotations of the different IBL staves. For that purpose, four iterations at Level 16 were performed, fixing the other modules. Figure 3.11 (a) shows the R_y corrections for the different alignment iterations. As it can be seen, the corrections are the same for the different staves and the rotation is corrected. The bowing is not aligned in these iterations. Then, four extra iterations are performed aligning the IBL bowing and keeping the rotations fixed around the z-axis. The residual distributions after the IBL stave alignment can be seen in Figure 3.11 (b). One can see the residual distributions are clearly centred at zero, meaning that the biases have been corrected.

After recovering these rotations in the IBL, a more detailed alignment of the whole ID structure was performed. Figure 3.12 shows the residual maps for the x direction in the $\eta - \phi$ plane for the three pixel barrel layers and the IBL. In a perfectly aligned the detector, these maps should be homogeneous and close to zero for the different regions. However, the residual maps shown that the pixel detector required a Level 2 alignment (see Figure 3.12 (a)). After the iterations at Level 2, some extra iterations were performed at Level 16 to ensure the highest precision in the alignment of the IBL staves. The corrected residual maps after these iterations



(a) (b)

Figure 3.10: (a): Evolution of the IBL R_z corrections after different iterations at Level 11. The misalignment in the R_z is recovered after the alignment iterations. The evolution of the pixel detector R_z correction is also displayed. (b): Local x residual distributions in the IBL using the HI 2015 dataset before the alignment (gray), after Level 11 correcting R_z (blue) and after Level 11 correcting B_x (blue).

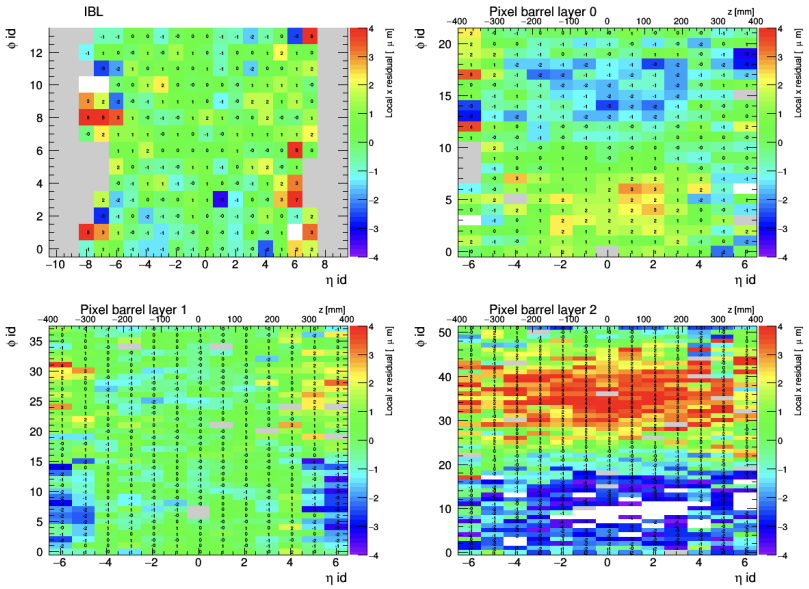


(a) (b)

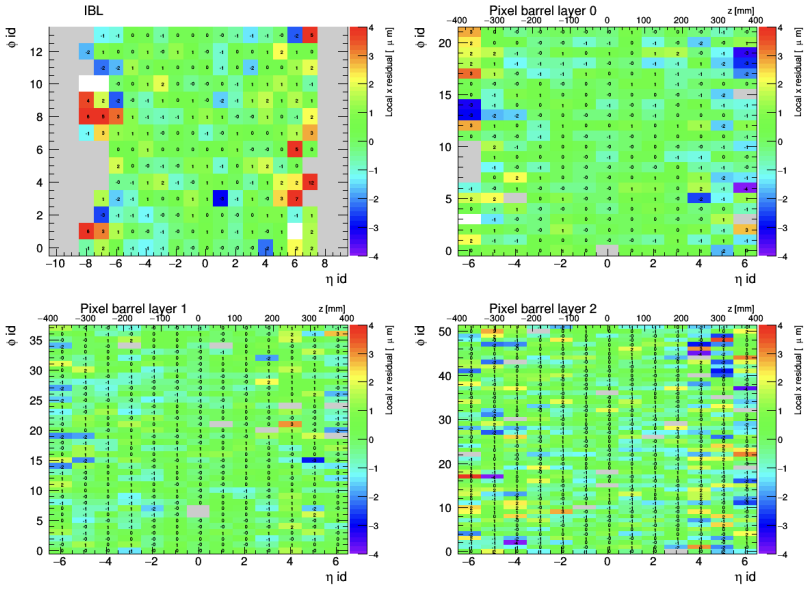
Figure 3.11: (a): Evolution of the IBL R_y stave corrections after different iterations at Level 16. The corrections are very similar for the different staves, and the 30 mrad misalignement is recovered. (b): Local x residual distributions in the IBL using the HI 2015 dataset after the Level 11 alignment (gray), after Level 16 correcting R_y (blue) and after Level 16 correcting B_x (blue). The mean of the distributions is centred at zero after the alignment chain and the biases are corrected.

are displayed in Figure 3.12 (b). As it can be seen, the misalignments of the outer layers are corrected and the reached precision is high ($\sim$4 μm).

From the obtained constants, a new beam spot was determined. To conclude this alignment campaign, a Level 3 alignment of the IBL was performed including the position of the beam spot as a constraint in the alignment algorithm to use it as a reference. This allows to correct misalignments in the y-axis, which has a lower precision as explained in Section 3.1.1.2. Figure 3.13 shows the IBL residual maps in the y direction before (a) and after (b) this Level 3 alignment, with a clear improvement in the results. The reached precision for the IBL in this direction is also very high ($\sim$5 μm) except for the high η indexes, which have much less statistics.

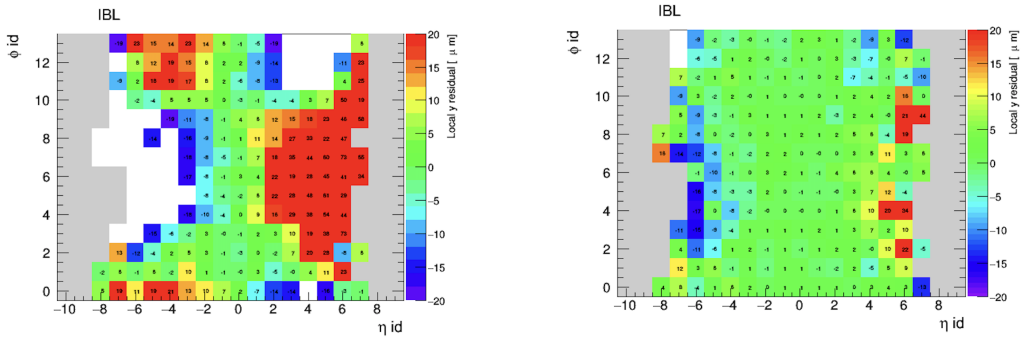


(a)



(b)

Figure 3.12: Residual maps in the x direction as a function of the ϕ and η indexes for the different pixel barrel layers and the IBL before (a) and after (b) correcting pixel movements at Level 2. In (b), additional iterations at Level 16 of the IBL were performed on top of the Level 2 alignment. The empty cells in the IBL residual maps correspond to high η indexes which do not have enough statistics. For the pixel barrel layer 2 in (a) the empty cells correspond to a strong misalignment, which leads to residuals which are beyond the scale of the plot.



(a) (b)

Figure 3.13: Residual maps in the y direction as a function of the ϕ and η indexes for the IBL before (a) and after (b) a Level 3 alignment of this structure. The empty cells correspond to high η indexes which do not have enough statistics.

Baseline alignment in 2016

The procedure followed to derive the baseline alignment constants for 2016 is very similar to the one described above, with the advantage that there were not any important prior misalignments which needed to be corrected.

In order to derive the first set of constants, the latest pp run prior to the HI data-taking period was used, with a total integrated luminosity of 365.5 pb^{-1}. The baseline alignment constants for the 2016 pp dataset were used as a reference. Firstly, an alignment at Level 11 was performed to correct for the movements of the biggest structures. Just four iterations were needed to reach convergence, since the corrections were quite small. In these iterations the beam spot constraint was used to correct for any possible bias in the transverse impact parameter distribution with respect to the beam spot. Figure 3.14 shows the impact parameter distribution. It can be seen that a bias of 1 μm is corrected using this constraint in the alignment algorithm. Finally, an alignment at Level 16 was performed to correct for possible misalignments of the IBL staves. Three iterations were enough to reach convergence. The level of precision reached in the residual maps in the x direction after this alignment is very similar to those presented for 2015.

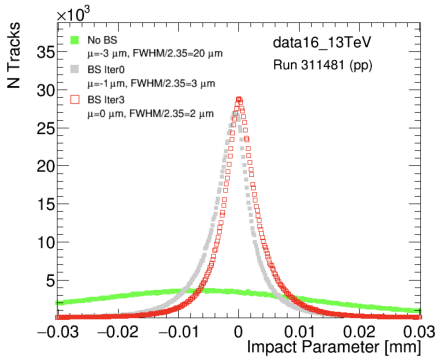
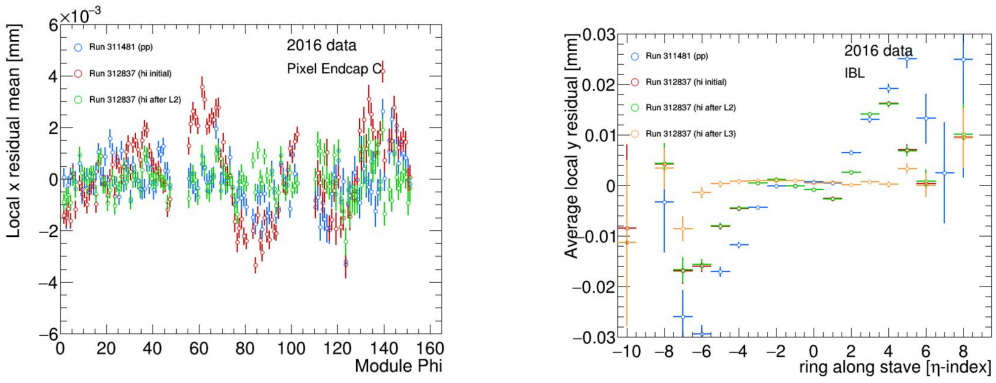


Figure 3.14: Impact parameter distribution with respect to the beam spot position when no beam spot constraint is applied (green), before (gray) and after (red) Level 11 alignment using the beam spot constraint. A bias of 1 μm is corrected after the alignment.

As stated before, HI collisions started two weeks after the end of pp collisions. To make sure the conditions of the detector did not change during this time some extra iterations were performed using one of the HI runs, with an integrated luminosity of 0.109 nb^{-1}. The starting point were the baseline constants derived from the previous pp run. Firstly, three iterations were performed at Level 11. Since the baseline alignment constants were obtained very close in time, the corrections in this case were minor. Figure 3.15 (a) shows the residual distributions for the three disks of one of the pixel end-cap modules as a function of the ϕ coordinate. As it can be seen, a Level 2 alignment was required to correct the observed misalignments. After correcting for these patterns, the last step was an alignment of the IBL at Level 3 as it was done in 2015 to increase the precision along the y-axis. Figure 3.15 (b) shows the average local y residual along the stave as a function of the η coordinate. The improvement after the alignment chain is clearly appreciated. Full details on the different steps taken in the alignment can be found in Table 3.4.

pp run					
Level	Iterations	IBL	Pixel	SCT	TRT
11	4	All except R_z	All	Barrel: fixed End-caps: T_x, T_y, R_z	Barrel: All End-caps: T_x, T_y, R_z
16	4	B_x	fixed	fixed	fixed
HI run					
11	4	All except R_z	All	Barrel: fixed End-caps: T_x, T_y, R_z	Barrel: All End-caps: T_x, T_y, R_z
2	3	All except B_x	Barrel: All End-caps: T_x, T_y, R_z	fixed	fixed
3	3	T_x, T_y, R_z	fixed	fixed	fixed

Table 3.4: Steps taken in the derivation of the baseline alignment constants for the HI 2016 campaign. Two different runs were used: firstly a pp run prior to the beginning of the HI campaign and secondly one of the first HI runs corresponding to the 2016 dataset. The baseline alignment constants derived from the alignment of the pp run were used as a starting point for the alignment of the HI run.



(a) (b)

Figure 3.15: (a): Local x residuals as a function of the ϕ index for one of the pixel end-caps for the pp run (blue), the HI run before the alignment (red) and the HI run after the pixel Level 2 alignment (green). (b): Average local y residual for the IBL staves as a function of the η index for the pp run (blue), the HI run before the alignment (red), the HI run after the pixel Level 2 alignment (green) and the HI run after the IBL Level 3 alignment (yellow).

3.5.2.2 Alignment of 2018 pp collisions

At the end of each data-taking campaign, the reprocessing of the data takes place to provide the highest possible quality. This section focuses on the realignment of the 2018 pp collision dataset to include the best possible knowledge of the detector.

In this realignment the main focus was on the correction of the weak modes described in Section 3.3. As explained, these weak modes leave the χ^2 unaltered, but the information from very well understood processes can be used to correct them. The correction of the weak modes in this realignment was performed using the $Z \rightarrow \mu^+\mu^-$ method. The following selection is applied to select muon candidates produced from a Z boson decay:

- The transverse momentum of the selected muons must satisfy $p_T > 20$ GeV.
- The reconstructed mass of the two muons has to be $70 < m_{\mu\mu} < 110$ GeV, matching the mass of the Z boson.
- Their impact parameters must satisfy $\Delta z_0 < 0.6$ mm and $\Delta d_0 < 0.1$ mm, since they come from the same vertex.
- The angular distance between the two muons $\alpha_{\mu\mu}$ must satisfy $\cos(\alpha_{\mu\mu}) > 0.673$.

From these candidate muons, the information on the δd_0, δz_0 and $\delta_{sagitta}$ can be extracted. This information is used to build 2D maps, which are introduced as additional constraints in the alignment algorithm.

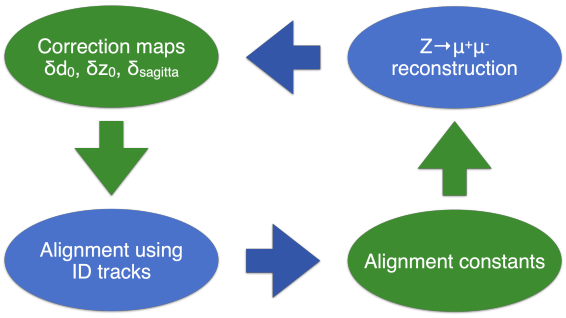


Figure 3.16: Schematic representation of the alignment chain to correct for the weak modes using the $Z \to \mu^+\mu^-$ method. The blue boxes include the two different steps ($Z \to \mu^+\mu^-$ reconstruction and alignment using ID tracks) and the green boxes include the output from each of them (correction maps and alignment constants, respectively), which are used as input for the other step.

This is an iterative procedure, depicted schematically in Figure 3.16. The information from the $Z \to \mu^+\mu^-$ reconstruction is used to build these correction maps, which are used as input for the alignment. The alignment provides a new set of alignment constants, which would alter the impact parameter biases. Therefore, new maps are derived from these constants, and so on.

The 2018 dataset corresponds to a total of 60.6 fb^{-1}. However, at some point during the period, there was a change in the pixel position of approximately 100 μm. Hence, the realignment of the 2018 dataset was divided in two periods, and each of them was performed separately. The realignment of the second period is described below in detail. A similar approach was carried for the first period, leading to very similar results.

The first step in this realignment was to select a set of consecutive runs with high statistics. Concretely, four runs were chosen giving in total 2.2 fb^{-1} of integrated luminosity. Then, a CL-like alignment was performed, including two iterations at Level 11 and two iterations at Level 16 with LB splitting to correct for the detector movements within the selected runs. From this, a first set of constants were derived, which were used as input for the $Z \to \mu^+\mu^-$ reconstruction and the iterative process described above began. Details on the different steps performed to minimise the

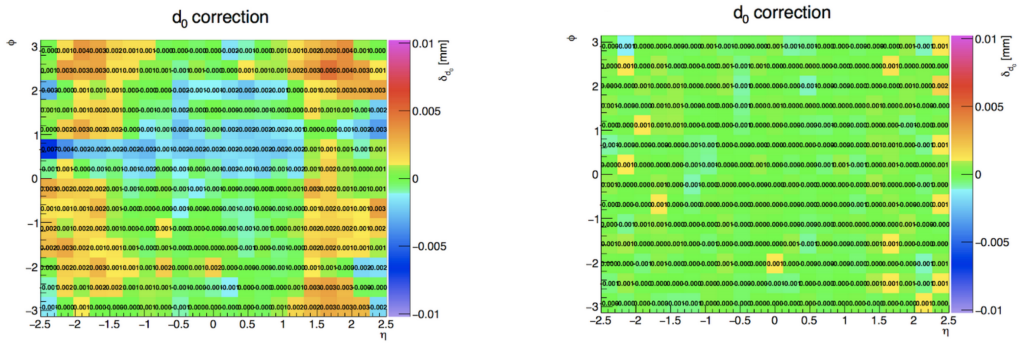
Level	Iterations	Aligned modules
11	2	All
16	2	IBL (bowing)
11	4	All
2	2	IBL and pixel detector (barrel and end-caps)
3	1	IBL and pixel detector (barrel and end-caps)
2	1	IBL and pixel detector (barrel and end-caps)
2	2	SCT barrel and end-caps
2 and 3	1	SCT barrel (Level 3); SCT end-caps (Level 2)
2	1	SCT and end-caps
2	3	TRT barrel and end-caps
11	1	All
3	2	SCT barrel and end-caps
2	1	SCT barrel and end-caps
2 and 3	2	SCT barrel and end-caps (Level 3); TRT barrel and end-caps (Level 2)
2 and 3	1	IBL, pixel detector and SCT barrel and end-caps (Level 3); TRT barrel and end-caps (Level 2)

Table 3.5: Steps taken in the derivation of weak modes free baseline alignment constants for the pp 2018 campaign (second period). A total of 28 iterations were used. The first two steps correspond to a CL alignment. From there, the δd_0, δz_0 and $\delta_{sagitta}$ maps are used as a constraint in the different alignment iterations. As it can be seen, at the beginning the different components of the ID are aligned independently, keeping the other components fixed. Finally, the different subcomponents were aligned together, up to Level 3 for the pixel detector and the SCT and up to Level 2 for the TRT.

impact parameter biases are listed in Table 3.5. In the end, a total of 28 iterations were used to ensure minimum biases.

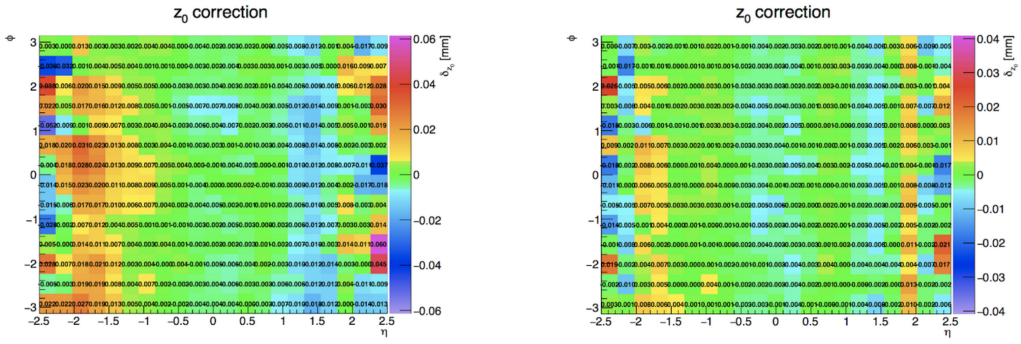
A comparison between the correction maps before and after the alignment chain are shown in Figures 3.17-3.19. The reached precision is very high, reducing biases in the transverse (longitudinal) impact parameter to $\sim$1 μm ($\sim$5 μm, except for the high η indexes, which correspond to the far end-cap regions and have much less statistics), and $\sim$0.15 TeV^{-1} for the sagitta bias.

This set of baseline alignment constants, for both the HI and the pp campaigns, were used as a starting point for a first alignment of the HI campaigns and for a complete realigment of the pp campaign in the CL. This procedure ensures there are no important biases after the CL alignment, which will take care of the possible movements between the large structures and the movements which take place within each run (see Section 3.4.2). After the CL, the alignment provides a very accurate



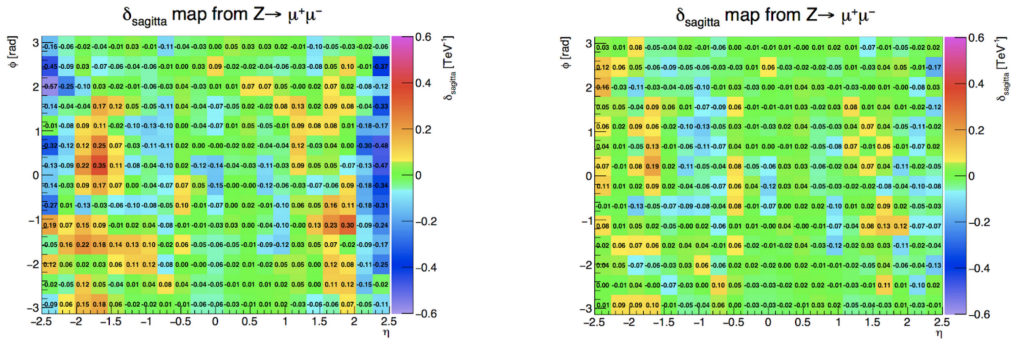
(a) (b)

Figure 3.17: Correction maps as a function of the ϕ and η indexes for the transverse impact parameter d_0 bias before (a) and after (b) the derivation of weak mode free baseline constants for the 2018 reprocessing campaign (second period).



(a) (b)

Figure 3.18: Correction maps as a function of the ϕ and η indexes for the longitudinal impact parameter z_0 bias before (a) and after (b) the derivation of weak mode free baseline constants for the 2018 reprocessing campaign (second period).



(a) (b)

Figure 3.19: Correction maps as a function of the ϕ and η indexes for the sagitta bias before (a) and after (b) the derivation of weak mode free baseline constants for the 2018 reprocessing campaign (second period).

description of the ID for the different LHC fills, with a remarkable precision for the different considered campaigns.

4 Measurement of differential angular distributions sensitive to the top-quark polarisation

The analysis presented in this chapter is developed in the context of t-channel single-top-quark production, where the top quark decays leptonically. The normalised differential cross-section of this process is presented as a function of three angular distributions sensitive to the top-quark polarisation. Due to the short lifetime of the top quark, it decays before depolarisation can take place, and its spin information is transmitted to its decay products. In particular, the lepton is the most powerful spin analyser among the different particles from the top-quark decay. The angular distributions of the lepton from the top-quark decay are measured in the top-quark reference frame with respect to three orthogonal axes. These measurements are unfolded to the particle level and are compared with theoretical predictions within the reconstructed fiducial region. The results are presented for top quarks and antiquarks, both inclusively and separately. They have been published in Ref. [3].

These results are compared with the SM predictions provided by various MC generators and with the direct measurement of the polarisation vector using a template fit [3]. Additionally, these measurements have been used to obtain limits for the real and the imaginary part of the dipole operator $\mathcal{O}_{tW}$ as part of the PhD thesis in Ref. [4]. The bounds for the imaginary part are the most stringent limits from high-energy physics experiments at the time this work is presented.

The data analysed corresponds to pp collisions at a centre-of-mass energy of $\sqrt{s}$ = 13 TeV collected by the ATLAS detector during 2015, 2016, 2017 and 2018, which corresponds to the full Run-2 dataset with a total integrated luminosity of 139 fb^{-1}.

This chapter includes the signal region definition used to enhance the t-channel signal events over the different contaminating backgrounds. A fiducial region at particle level, very close to the reconstruction level region, is also defined. Control regions are used to check the modelling of the most important background sources. The different backgrounds and the t-channel signal are estimated with different MC generators or with data-driven techniques. In order to constrain the normalisation of the most important backgrounds, a fit to the total number of data events is performed in the control regions and in the signal region simultaneously. Afterwards, the different backgrounds are subtracted from the data in the signal region. The resulting distributions, which are expected to be t-channel signal events, are unfolded to the stable particle level within a fiducial region to correct for detector effects and are compared with theoretical predictions. Each bin of the unfolded distributions is divided by the integral over all bins and the results are presented as normalised differential cross-sections as a function of three angular distributions sensitive to the top-quark polarisation.

4.1 t-channel signal and background contamination

This analysis focuses on the leptonic decay channel of the top quark ($t \to Wb, W \to l\nu$, with $l = e, \mu$). The main reason is the spin analysing power of the lepton, very close to unity, which makes it the best candidate to study the top-quark polarisation. Another reason is related to the background contamination. The hadronic channel has a ~ 2 times larger branching ratio [12], but it is expected to have a much larger contamination from multijet production. Events in which the W boson decays into a τ lepton and a neutrino are also considered if the τ lepton subsequently decays to a muon or an electron and an additional neutrino[1].

[1]The τ decays leptonically approximately 35% of the times [12].

The final state considered in this analysis is characterised then by one lepton, which can be either an electron or a muon; two jets, one of them being b-tagged and the other a forward jet originated from a light quark; and a large amount of missing transverse momentum corresponding to the undetected neutrino.

There are several processes which can present the same final state in terms of reconstructed objects, passing the selection requirements. Thus, they contribute as background processes in the signal region. These are depicted in Figure 4.1 and listed below:

- **Top-quark processes**: these include mainly top quarks produced in pairs via the strong interaction due to its very large cross-section. A lower contribution is also expected from the other single-top-quark production channels, the tW- and the s-channel. These processes are one of the most important background sources in this analysis. They are hard to distinguish from the t-channel signal since they contain at least one real top quark in the final state. The corresponding Feynman diagrams can be found in Figures 1.4.

- **W+jets production**: these include the production of a W boson in association with jets, which can either be heavy flavour jets ($W + b\bar{b}$ and $W + c\bar{c}$) or light flavour jets. These processes are also expected to have an important impact in the analysis, in particular if the W-boson decays leptonically.

- **Z+jets production**: this includes the production of a Z boson in association with jets, especially when the Z boson decays leptonically. The associated jets can be either heavy flavour jets ($Z + b\bar{b}$ and $Z + c\bar{c}$) or light flavour jets.

- **Diboson production**: this includes the production of two vector bosons VV (ZZ, WW, ZW), in particular when one boson decays leptonically and the other decays hadronically.

- **Multijet production**: this background includes events in which a jet is misidentified as a lepton, or where a non-prompt lepton is selected (for example, from the decay of a b-hadron).

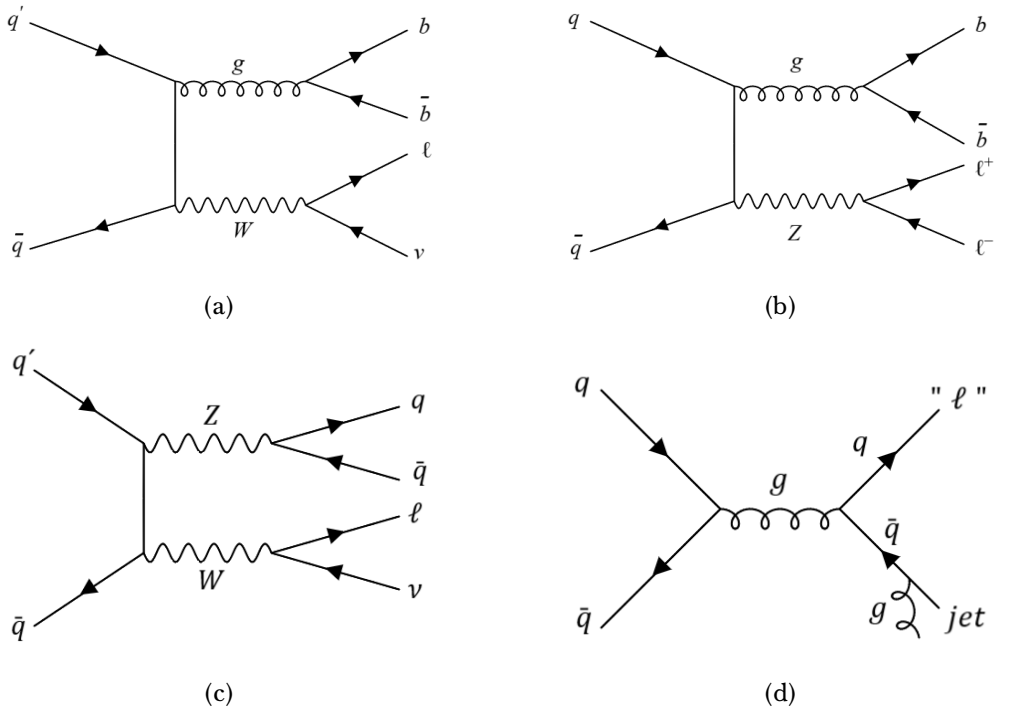


Figure 4.1: Example Feynman diagrams of the main contributing backgrounds (not including top-quark processes) to the t-channel signal final state considered in this analysis for (a) W+jets, (b) Z+jets, (c) Diboson and (d) Multijet.

Other minor backgrounds include the production of a vector boson or a Higgs boson in association with top quarks ($t\bar{t}Z$, $t\bar{t}W$, $t\bar{t}H$, tZq, tHq and tWZ). These backgrounds are also taken into account, even though their contribution to the expected total number of selected backgrounds events is negligible.

4.2 Data and simulated event samples

This section describes the data used in this analysis and the different MC generators employed to simulate the t-channel signal and the background events.

4.2.1 Data event samples

The analysis is performed using pp collision data collected at a centre-of-mass energy of 13 TeV from 2015 to 2018. Stringent detector and data quality requirements were applied [346], resulting in a data sample corresponding to a total integrated luminosity of 139 fb^{-1} [246, 247]. The events were selected by single-lepton[2] triggers [331, 332], as described in Section 2.7.1.

4.2.2 Simulated event samples

Samples of simulated events were produced using different MC event generators including parton shower (PS) and hadronisation models. All the signal or background processes involving one or more top quarks were simulated assuming a top-quark mass of 172.5 GeV, and the top quark was assumed to decay only to a W boson and a b quark. In these samples, top quarks, W and Z bosons were decayed at LO using MadSpin [347, 348] to preserve all spin correlations. Moreover, the PS, hadronisation and underlying-event (UE) modelling was simulated with Pythia8 (v8.230 or v8.212), using the A14 set of tuned parameters (A14 tune) [349] and the NNPDF2.3 LO PDF set. The decays of bottom and charm hadrons were simulated using the EvtGen (v1.6.0 or v1.2.0) program [350].

The effect of multiple interactions in the same and neighbouring bunch crossings (pile-up) is modelled by overlaying simulated minimum-bias events generated with Pythia8 (v8.186) [261, 351] using the NNPDF2.3 LO [352] set of parton distribution functions (PDFs) and the A3 set of tuned parameters [353] over the original hard-scattering event. Events were reweighted such that the distribution of the average number of interactions per bunch crossing matches that observed in data.

***t*-channel signal**: Single-top-quark t-channel events were generated with the NLO Powheg-Box [354–357] (v2) generator which provides matrix elements (MEs) at NLO in the strong coupling constant α_s using the 4FS with the NNPDF3.0 NLO nf4 [358] PDF set. The renormalisation (μ_R) and factorisation (μ_F) scales were set

[2]Henceforth, "lepton" indicates electron or muon.

to $\mu_R^2 = \mu_F^2 = 16(m_b^2 + p_{T,b}^2)$, where m_b and $p_{T,b}$ are the mass and p_{T} of the b quark from the initial gluon splitting.

Additional signal samples of simulated t-channel events were produced within the 4FS with the LO Protos generator (v2.3) [359] using the CTEQ6L1 PDF set [259]. The scales were set to $\mu_F^2 = -p_W^2$ for the light-flavour quark distribution function and $\mu_F^2 = p_{T,\bar{b}}^2 + m_b^2$ for the gluon distribution function, where p_W and $p_{T,\bar{b}}$ are the momentum of the exchanged W boson and the p_{T} of the b antiquark originating from the gluon splitting, respectively [359, 360].

In addition to a SM signal sample, samples with tWb anomalous couplings [361, 362], $V_{L,R}$ and $g_{L,R}$ (see Equation 1.17), enabled in both the production and decay vertices were produced with the Protos event generator. The ranges of the anomalous coupling event samples were selected based on previous established limits [129]: $\mathrm{Re}[g_R]/V_L = \pm 0.18$, $\mathrm{Im}[g_R]/V_L = \pm 0.07$, and $V_R/V_L = \pm 0.4$ and $g_L/V_L = \mp 0.32$. For $V_L = 1$ and $\Lambda = 1$ TeV, following Ref. [363], these correspond to $C_{tW} = \pm 2.0$, $C_{itW} = \pm 0.8$, and $C_{\varphi tb} = \pm 13$ and $C_{bW} = \mp 3.7$, respectively, which are the Wilson coefficients associated to the corresponding operators.

Furthermore, signal samples of simulated events with non-zero Wilson coefficients related to dimension-six effective operators were also produced. These simulation samples were generated at NLO, using the same setup as in Ref. [123], with the MG5_aMC@NLO [364] generator (v2.6.2) using the NNPDF3.0 NLO PDF set. The values of the non-zero Wilson coefficients in the different simulated event samples were set to: $C_{tW} = 2.0$, $C_{itW} = 1.75$, $C_{qq} = -0.4$, $C_{tW} = -2.0$ and $C_{qq} = -0.4$, and $C_{tW} = 2.0$ and $C_{itW} = -1.75$. The EFT operators were allowed to enter in the production vertex as well as the decay vertex, consistently including possible effects on the width of the top quark. The implementation of the NLO effective operators [365] makes use of the 5FS, indicating that the b quark is massless and thus part of the proton. Both μ_F and μ_R were set to the top-quark mass (m_t) in the MC generation.

The t-channel signal production cross-section is calculated at NLO [366] in perturbative QCD using HATHOR (v2.1) [107, 108]. For pp collisions at 13 TeV, this cross-section corresponds to 136 ± 5 pb and 81 ± 4 pb for top-quark and top-antiquark production, respectively. The uncertainties connected with PDFs

and α_s are calculated using the PDF4LHC prescription [367] with the MSTW2008 NLO [256, 368], CT10 NLO [369] and NNPDF2.3 NLO PDF sets, and are added in quadrature to the scale uncertainty.

Top-quark processes: The production of $t\bar{t}$ events, as well as single-top-quark events in the tW process and in the s-channel, were modelled using the POWHEG-BOX (v2) generator which provides MEs at NLO with the NNPDF3.0 NLO PDF set. In $t\bar{t}$ events, the POWHEG-BOX h_{damp} parameter[3] was set to 1.5 m_t [370]. For the $t\bar{t}$ process, the μ_R^2 and μ_F^2 scales were set to $m_t^2 + p_{\text{T,t}}^2$, where $p_{\text{T,t}}$ is the top quark transverse momentum, while for the tW- and s-channel processes these scales were set to m_t. In the case of tW associated production, the diagram removal (DR) scheme [371] was employed to handle the interference with $t\bar{t}$ production [370].

The $t\bar{t}$ cross-section is calculated at NNLO in QCD including resummation of next-to-next-to-leading logarithmic (NNLL) soft-gluon terms with Top++2.0 [101–106]. This cross-section corresponds to 832 ± 35 (PDF + α_s) $^{+20}_{-29}$ (scale) pb. The PDF and α_s uncertainties are calculated using the PDF4LHC prescription with the MSTW2008 68% C.L. NNLO, CT10 NNLO [258] and NNPDF2.3 NNLO PDF sets, and are added in quadrature to the scale uncertainty. The tW events are normalised to the predicted production cross-section of 72 ± 4 pb calculated at NLO in QCD including NNLL soft-gluon corrections [109]. The uncertainty on the cross-section corresponds to the sum in quadrature of the uncertainty derived from the MSTW2008 90% C.L. NNLO PDF set and the scale uncertainties. For the s-channel process, the inclusive cross-section is corrected to the theory prediction calculated at NLO in QCD with HATHOR. This cross-section corresponds to 10.3 ± 0.4 pb. The uncertainties on the cross-section due to PDF and α_s are calculated using the PDF4LHC prescription, similarly as for the t-channel.

***V*+jets processes**: Vector-boson production in association with jets (generally named V+jets, or W/Z+jets) was simulated with the multi-leg SHERPA (v2.2.1) [263] generator. The NNPDF3.0 NNLO set [358] of PDFs as well as the dedicated set

[3]The POWHEG-BOX h_{damp} parameter controls the p_{T} of the first additional emission beyond the LO Feynman diagram in the PS and therefore regulates the high-p_{T} emission against which the $t\bar{t}$ system recoils.

of tuned PS parameters developed by the SHERPA authors for this version were used. The events were filtered according to their b-hadron and c-hadron content at the particle level. The ME+PS matching [372] is employed for different jet multiplicities which were then merged into an inclusive sample using an improved CKKW matching procedure [373, 374] which is extended to NLO accuracy using the MEPS@NLO prescription [375]. These particular simulations are NLO accurate for up to two additional partons and LO accurate for up to four additional partons. The virtual QCD correction for MEs at NLO accuracy are provided by the OPENLOOPS library [376, 377].

The inclusive cross-sections of V+jets production are calculated to NNLO prediction [378] with the FEWZ program [379]. For W+jets, the overall theoretical normalisation uncertainty is 34%. This is the result of adding in quadrature the overall cross-section normalisation uncertainty and 24% per additional jet, according to the Berends–Giele scaling [380].

Diboson production: Fully leptonically and semileptonically decaying diboson (VV) samples were simulated with the SHERPA (v2.2.1) ME+PS generator. The ME+PS matching is the same as for the single-boson processes. These particular simulations are NLO accurate for up to one additional parton and LO accurate for up to three additional parton emissions using factorised on-shell decays.

Minor backgrounds: The production of $t\bar{t}Z$, $t\bar{t}W$, tZq, tHq and tWZ events were modelled using the MG5_AMC@NLO (v2.3.3) generator which provides NLO MEs in α_s with the NNPDF3.0 NLO PDF set. The production of $t\bar{t}H$ events was modelled using the POWHEG-BOX generator at NLO with the NNPDF3.0 NLO PDF set.

Multijet production: For the multijet background estimation in the electron channel[4], dijet events were simulated using PYTHIA8 (v8.186) with the A14 tune, using the NNPDF2.3 NLO PDF set, and where the decays of bottom and charm hadrons were simulated using the EVTGEN (v1.2.0) program. Here, $2 \to 2$ QCD processes were generated, including multijet, $qg \to q\gamma$, $q\bar{q} \to g\gamma$, electroweak (W/Z) and $t\bar{t}$ production processes.

[4]In the muon channel a method based on collision data is used instead. This is described in Section 4.4.1.

Alternative samples of simulated events were also produced for single-top-quark and $t\bar{t}$ events either using different generators or different variations of parameters in the POWHEG-BOX+PYTHIA8 (further details are given in Section 4.8) to estimate the generator modelling uncertainties. For studies of the NLO matching method, MG5_aMC@NLO generator (v2.6.2 or v2.6.0) using either the NNPDF3.0 NLO nf4 PDF set for the t-channel process or the NNPDF3.0 NLO PDF set for the $t\bar{t}$, tW and s-channel processes, is used. In these cases, the ME generator was interfaced to PYTHIA8 (v8.230 or v8.212). To study the PS, the hadronisation and the UE modelling, the POWHEG-BOX (v2) generator interfaced to HERWIG7 (v7.04) [262, 381] using the H7UE set of tuned parameters [381]) and the MMHT2014 LO PDF set [382] was used. In the case of tW associated production, a simulated event sample using the diagram subtraction (DS) scheme [371] is employed to estimate the uncertainty associated to the scheme used to handle the interference with $t\bar{t}$ production.

All baseline simulated event samples were passed through the full simulation (FS) of the ATLAS detector [383] based on the GEANT4 [384] framework. The simulated samples used to estimate the impact of anomalous couplings or EFT coefficients, and samples used to evaluate most of the systematic effects were processed with a fast simulation (AFII) [383] which relies on a parameterisation of the calorimeter response [385].

An extensive software suite [386] is used for real and simulated data reconstruction and analysis, for operation and in the trigger and data acquisition systems of the experiment.

4.3 Event selection in the signal and control regions

The event selection is performed in two steps in this analysis. Firstly, a set of basic kinematic requirements based on the topology of the t-channel signal events are applied on the physics objects described in Section 2.7.2. This is the so-called *pre-selection* level. After reconstructing the W-boson and top-quark kinematics, a set of additional requirements are applied to enhance the signal events over the background contamination. This defines the *signal region*.

Two additional regions are defined in this analysis to monitor the modelling of the main contributing backgrounds ($t\bar{t}$ and W+jets). These are the so-called *control regions*. These control regions are also used to constrain the normalisation of these backgrounds, as it will be explained in Section 4.4.2.

4.3.1 Pre-selection region

At pre-selection level events are required to have at least one vertex reconstructed from at least two inner-detector tracks with transverse momenta of $p_{\mathrm{T}} > 0.5$ GeV. The primary vertex for each event is defined as the vertex with the highest sum of p_{T}^2 over all associated inner-detector tracks [281]. Additional requirements are applied based on the final state topology of the event, depicted in Figure 1.7. The pre-selected events satisfy:

- Exactly one isolated charged lepton (electron or muon) with $p_{\mathrm{T}} > 30$ GeV and $|\eta| < 2.5$. To remove background events from $t\bar{t}$, Z+jets and diboson productions, the event is vetoed if an additional loose lepton candidate is found with $p_{\mathrm{T}} > 10$ GeV, but without isolation requirements and satisfying less stringent lepton identification criteria [282, 285, 387].
- Exactly two jets, one of them being b-tagged. The b-jet is required to have $p_{\mathrm{T}} > 30$ GeV and $|\eta| < 2.5$. As explained before, the second b-jet (see Figure 1.7 (b)) is expected to be very soft and forward, thus falling out of the detector kinematic acceptance. Since it is not expected to be detected, no extra requirement for this b-jet is applied. The non-tagged jet is considered to be the spectator jet, which is required to have $p_{\mathrm{T}} > 30$ GeV and $|\eta| < 4.5$ as it is expected to be produced in the forward direction. The p_{T} requirement is increased for both jets to $p_{\mathrm{T}} > 35$ GeV in the transition region between the central and forward calorimeters ($2.75 < |\eta| < 3.5$).
- The $E_{\mathrm{T}}^{\mathrm{miss}}$ is required to be larger than 35 GeV, to account for the final state neutrino of the signal events.

Two additional multijet background rejection criteria are applied. The transverse mass of the lepton–$E_{\mathrm{T}}^{\mathrm{miss}}$ system,

$$m_T(l, E_{\mathrm{T}}^{\mathrm{miss}}) = \sqrt{2p_{\mathrm{T}}(l)E_{\mathrm{T}}^{\mathrm{miss}}(1 - \cos\Delta\phi(p_{\mathrm{T}}(l), E_{\mathrm{T}}^{\mathrm{miss}}))}, \tag{4.1}$$

where $\Delta\phi(p_{\mathrm{T}}(l), E_{\mathrm{T}}^{\mathrm{miss}})$ is the difference in azimuthal angle between the lepton momentum and the $E_{\mathrm{T}}^{\mathrm{miss}}$ direction, is required to be larger than 60 GeV. This requirement also contributes to the selection of events where the lepton and the neutrino come from the decay of a W boson.

Further reduction of the multijet background is achieved by imposing an additional requirement on events where the lepton and the leading jet in p_{T} have opposite directions in the transverse plane:

$$p_{\mathrm{T}}(l) \; > \; 50 \left(1 - \frac{\pi - |\Delta\phi(p_{\mathrm{T}}(j_1), p_{\mathrm{T}}(l))|}{\pi - 1}\right) \mathrm{GeV}, \tag{4.2}$$

where $\Delta\phi(p_{\mathrm{T}}(j_1), p_{\mathrm{T}}(l))$ is the difference in azimuthal angle between the lepton p_{T} and the leading jet in p_{T}. This requirement allows to significantly reject the background originating from multijet events, where the two jets are generated back-to-back in the azimuthal plane and one of those is misreconstructed as a lepton.

4.3.2 Reconstruction of the W boson and the top quark

The on-shell W boson originating from the decay of the top quark is reconstructed from the momenta of the lepton and the neutrino by imposing four-momentum conservation:

$$(p_W)^2 = (p_l + p_\nu)^2 \rightarrow m_W^2 = m_l^2 + 2(E_l E_\nu - \vec{p}_l \vec{p}_\nu), \tag{4.3}$$

where p is the four-momentum, m is the mass, E is the energy and $\vec{p}$ is the momentum of the corresponding particle.

Given that the neutrino escapes undetected at reconstruction level, its transverse momentum is assumed to correspond to the $E_{\mathrm{T}}^{\mathrm{miss}}$. Then, its energy and its momentum can be expressed as:

$$E_\nu = \sqrt{(E_{\mathrm{T}}^{\mathrm{miss}})^2 + (p_z(\nu))^2}, \tag{4.4a}$$

$$\vec{p_\nu} = (E_{\mathrm{T}}^{\mathrm{miss}} \cos\phi_{E_{\mathrm{T}}^{\mathrm{miss}}}, E_{\mathrm{T}}^{\mathrm{miss}} \sin\phi_{E_{\mathrm{T}}^{\mathrm{miss}}}, p_z(\nu)), \tag{4.4b}$$

where $\phi_{E_{\mathrm{T}}^{\mathrm{miss}}}$ is the azimuthal angle associated with the missing transverse momentum.

With these definitions, Equation 4.3 can be rewritten as:

$$m_W^2 = m_l^2 + 2E_l\sqrt{(E_{\mathrm{T}}^{\mathrm{miss}})^2 + (p_z(\nu))^2} - 2E_{\mathrm{T}}^{\mathrm{miss}}(p_x(l)\cos\phi_{E_{\mathrm{T}}^{\mathrm{miss}}} + p_y(l)\sin\phi_{E_{\mathrm{T}}^{\mathrm{miss}}}) - 2p_x(l)p_z(\nu), \tag{4.5}$$

where all the terms are known except the longitudinal component of the neutrino momentum. To calculate this unknown, the mass of the W boson (m_W = 80.399 $\pm$ 0.023 GeV [12]) is included as a constraint. Equation 4.5 can be expressed as a quadratic equation:

$$a(p_z(\nu))^2 + bp_z(\nu) + c = 0, \tag{4.6}$$

where:

$$a = (E_l)^2 - (p_z(l))^2$$

$$b = p_z(l)(m_l^2 - m_W^2 - 2E_{\mathrm{T}}^{\mathrm{miss}}(p_x(l)\cos\phi_{E_{\mathrm{T}}^{\mathrm{miss}}} + p_y(l)\sin\phi_{E_{\mathrm{T}}^{\mathrm{miss}}})$$

$$c = (E_l)^2(E_{\mathrm{T}}^{\mathrm{miss}})^2 - \frac{1}{4}(m_W^2 - m_l^2 + 2(p_x(l)\cos\phi_{E_{\mathrm{T}}^{\mathrm{miss}}} + p_y(l)\sin\phi_{E_{\mathrm{T}}^{\mathrm{miss}}}))^2$$

Thus, the longitudinal momentum has two possible solutions $p_z(\nu) = \frac{-b\pm\sqrt{\Delta}}{2a}$, with $\Delta = b^2 - 4ac$. From this, three possible scenarios can be considered:

- Δ = 0: There is only one possible solution, which is taken as $p_z(\nu)$.
- Δ > 0: There are two possible solutions. The one with the smaller absolute $p_z(\nu)$ is selected.
- Δ < 0: There is no-real solution for $p_z(\nu)$. This occurs due to the initial assumption $E_{\mathrm{T}}^{\mathrm{miss}} = p_{\mathrm{T}}(\nu)$. This is not always valid, since there can be other contributions to the $E_{\mathrm{T}}^{\mathrm{miss}}$ (like extra neutrinos from B-hadrons or τ-decays). Additionally, non-perfect resolution and calibration of the $E_{\mathrm{T}}^{\mathrm{miss}}$ can also

lead to these complex solutions. This is solved by re-scaling the $E_{\mathrm{T}}^{\mathrm{miss}}$ while preserving the azimuthal angle $\phi_{E_{\mathrm{T}}^{\mathrm{miss}}}$ until a positive solution is found. In other words, one can solve the equation $\Delta = 0$ in terms of the $E_{\mathrm{T}}^{\mathrm{miss}}$ to obtain a re-scaled value labelled $E_{\mathrm{T}}^{\mathrm{miss'}}$. Again, this is a second order equation of the form:

$$A(E_{\mathrm{T}}^{\mathrm{miss'}})^2 + BE_{\mathrm{T}}^{\mathrm{miss'}} + C = 0, \tag{4.8}$$

where:

$$A = (E_l)^2 - (p_z(l))^2 + (p_x(l)\cos\phi_{E_{\mathrm{T}}^{\mathrm{miss}}} + p_y(l)\sin\phi_{E_{\mathrm{T}}^{\mathrm{miss}}})^2$$
$$B = (m_W^2 - m_l^2)(p_x(l)\cos\phi_{E_{\mathrm{T}}^{\mathrm{miss}}} + p_y(l)\sin\phi_{E_{\mathrm{T}}^{\mathrm{miss}}})$$
$$C = (m_W^2 - m_l^2)^2/4$$

From the two possible solutions, if only one is positive then this is chosen as the new $E_{\mathrm{T}}^{\mathrm{miss}}$. In case the two solutions are positive, the one closer to the original $E_{\mathrm{T}}^{\mathrm{miss}}$ is chosen. Once this new $E_{\mathrm{T}}^{\mathrm{miss}}$ is obtained, it is increased by few eV to ensure $\Delta > 0$. Finally, Equation 4.6 is solved again with this new value of the missing transverse momentum to obtain the longitudinal momentum of the neutrino.

After obtaining the four-momentum of the neutrino, the kinematics of the W boson can be fully determined. The top quark can then be reconstructed combining the four-momentum of the W boson with the four-momentum of the b-tagged jet ($t \to Wb$):

$$p_t = p_W + p_b \tag{4.10}$$

4.3.3 Signal region

Further discrimination between t-channel signal and background events is achieved by applying additional criteria that optimise the signal-to-background

ratio (S/B) while maintaining the expected significance[5] always above a fixed threshold, which thereby determine the signal region. The following criteria are applied:

- The invariant mass of the lepton-b-jet system, m_{lb}, is required to be lower than 153 GeV.
- The mass of the reconstructed top quark, $m_{l\nu b}$, is required to be within 120.6 - 234.6 GeV.
- A trapezoidal requirement is applied in order to reject background events which have light-flavour jets in the forward region in events with centrally reconstructed top quarks. This requirement is:

$$\begin{aligned} \eta_j &< (4\ \eta_{l\nu b} + a) \quad \& \\ \eta_j &> (4\ \eta_{l\nu b} - a) \quad \& \\ (\eta_j &> (0.44\ \eta_{l\nu b} + b) \quad \text{OR} \\ \eta_j &< (0.44\ \eta_{l\nu b} - b))\,, \end{aligned} \tag{4.11}$$

  where η_j and $\eta_{l\nu b}$ represent the pseudorapidity of the spectator jet and the reconstructed top quark, respectively, and the intercept parameters a and b of the lines defining the trapezoid were optimised to be 10 and 2, respectively. This requirement is depicted in Figure 4.2.
- The mass of the spectator-jet–top-quark system, $m_{jl\nu b}$, is required to be larger than 320 GeV.
- The scalar sum of the p_{T} of all final-state objects, H_{T}, must be larger than 190 GeV.

[5] This is given by $s/\sqrt{s+b}$.

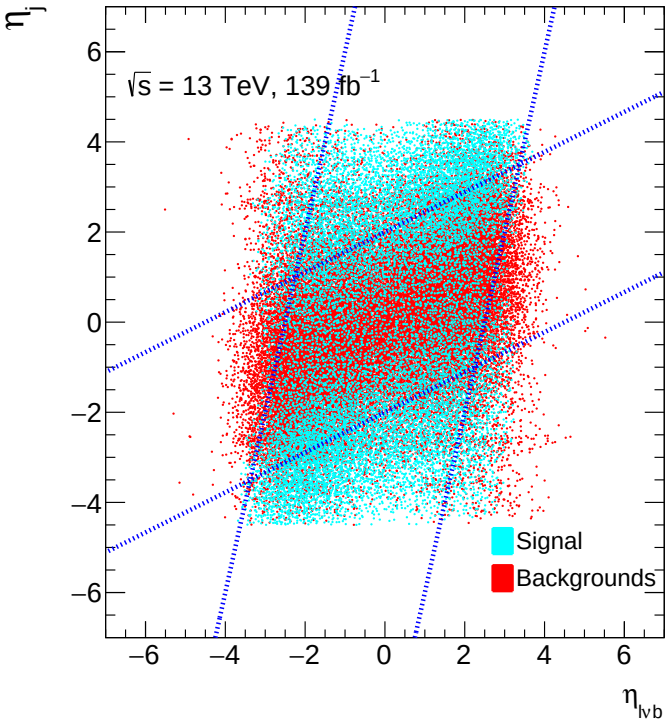


Figure 4.2: Two dimensional correlations among the pseudorapidity of the spectator jet and the reconstructed top quark. The expected signal (background) is represented by the red (blue) dots. The requirement is applied in the region defined by the intersection of the four green lines.

These requirements have been optimised using the full statistics of the Run-2. The optimisation is performed sequentially in the same order as they have been explained before. The m_{lb}, $m_{l\nu b}$, $m_{jl\nu b}$ and H_{T} distributions at pre-selection level can be found in Figure 4.3. The requirements applied in the signal region are also displayed.

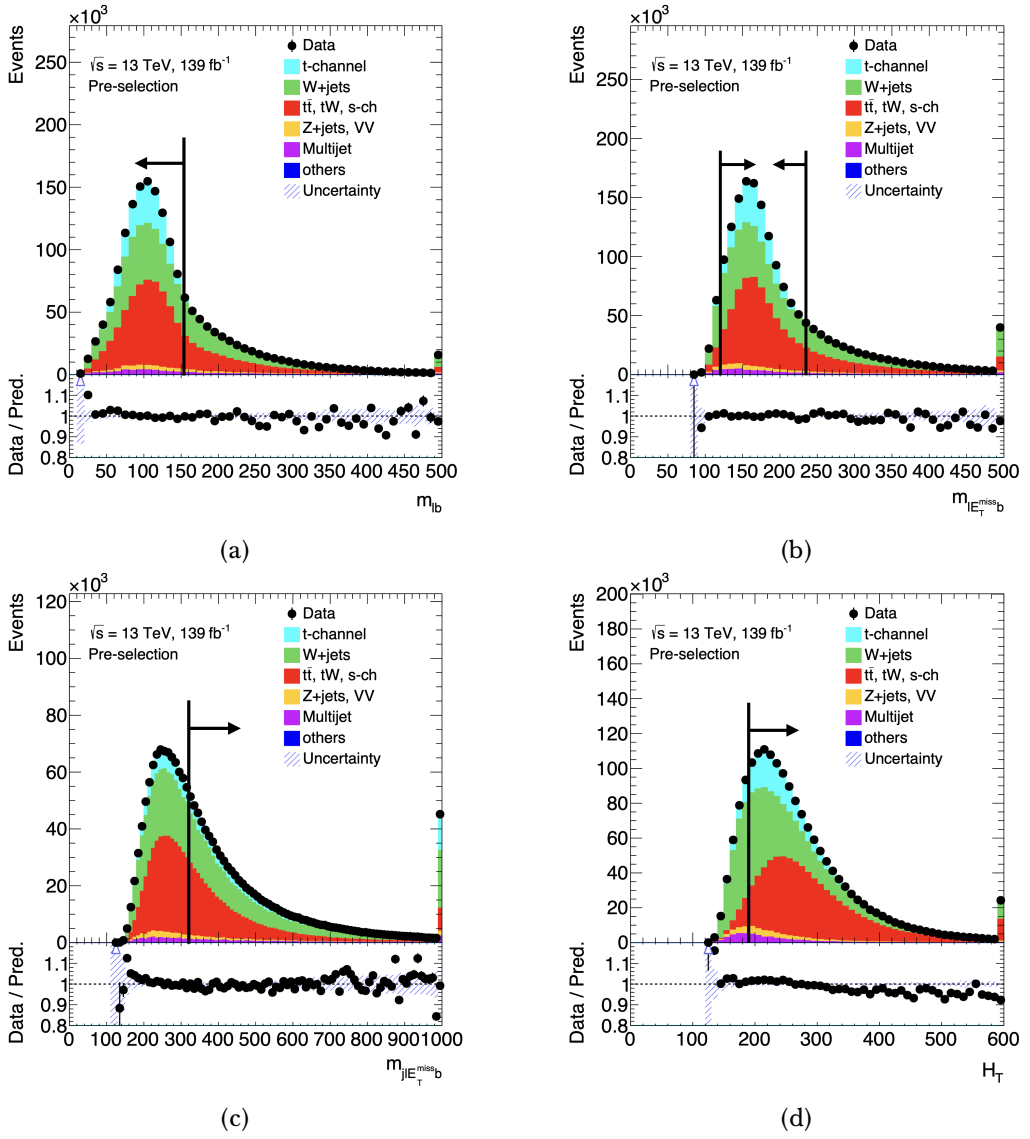


Figure 4.3: Kinematic distributions for data and signal and background expectations of the variables used to define the signal region at pre-selection level. (a) The invariant mass of the lepton-b-jet system, (b) the invariant mass of the reconstructed top quark, (c) the invariant mass of the spectator-jet–top-quark system, (d) the scalar sum of the transverse momentum of the final-state objects. The solid black line represents the requirements applied in the signal region. VV stands for the diboson background. The label "others" represents the $t\bar{t}Z$, $t\bar{t}W$, $t\bar{t}H$, tZq, tHq and tWZ productions. The uncertainty bands correspond to the errors due to the limited size of the simulation samples added in quadrature with the multijet normalisation uncertainty. The lower panel shows the ratio of data to prediction. The last bin of the distributions includes the overflow.

4.3.4 Control regions

Two specific background-enriched control regions are defined to monitor the modelling of the main contributing background processes in the t-channel signal region. Additionally, these control regions (and the signal region) are used to estimate scale factors for the overall normalisations of these backgrounds and the t-channel signal. The definition of these regions ensures orthogonality with the signal region:

- $t\bar{t}$ control region: A control region enriched in $t\bar{t}$ events is defined by applying all the pre-selection requirements, except for the requirement of exactly one b-tagged jet; instead exactly two b-tagged jets are required. This definition ensures a composition of the $t\bar{t}$ background very similar to the composition in the signal region[6].

- W+jets control region: A control region enriched in W+jets events is defined by selecting events satisfying the pre-selection requirements and failing at least one of the signal region selection criteria. This control region has a similar W+jets flavour composition as the signal region. Both regions are dominated by W+heavy jets, which constitutes 95% of the total W+jets background.

Table 4.1 summarises the selection criteria for defining the pre-selection, the signal and the two control regions used in this analysis. For the separate measurements of top-quark and top-antiquark events, the selected events in each region are further divided into two different regions according to the lepton charge.

Figure 4.4 shows the contribution of the different backgrounds and the t-channel signal in the pre-selection, the signal and the control regions. In the signal region, the t-channel signal represents a 47% of the total number of expected events, whereas the two main backgrounds $t\bar{t}$ and W+jets contribute a 26% and a 18%,

[6]The composition of the $t\bar{t}$ background can be split in dileptonic and semileptonic (all hadronic decays are not considered in this analysis). Here, leptonic includes also τ leptons. The composition in the signal region is found to be 68% dileptonic and 32% semileptonic. The numbers in the $t\bar{t}$ control region are found to be 79% dileptonic and 22% semileptonic, which are close to the ones observed in the signal region.

Pre-selection region	Signal region	$t\bar{t}$ control region	W+jets control region	
=1 electron ($p_{\mathrm{T}} > 30$ GeV and $\lvert\eta\rvert < 2.47$, excluding $1.37 < \lvert\eta\rvert < 1.52$) or 1 muon ($p_{\mathrm{T}} > 30$ GeV and $\lvert\eta\rvert < 2.5$)				
Veto secondary low-p_{T} charged loose leptons ($p_{\mathrm{T}} > 10$ GeV and $\lvert\eta\rvert < 2.5$)				
=2 jets ($p_{\mathrm{T}} > 30$ GeV and $\lvert\eta\rvert < 4.5$; $p_{\mathrm{T}} > 35$ GeV within $2.7 < \lvert\eta\rvert < 3.5$)				
$E_{\mathrm{T}}^{\mathrm{miss}} > 35$ GeV				
$m_{\mathrm{T}}(l, E_{\mathrm{T}}^{\mathrm{miss}}) > 60$ GeV				
$p_{\mathrm{T}}(l) > 50 \left(1 - \frac{\pi - \lvert\Delta\phi(p_{\mathrm{T}}(j_1), p_{\mathrm{T}}(l))\rvert}{\pi - 1}\right)$ GeV				
=1 b-jet ($\lvert\eta\rvert < 2.5$; 60% eff.)		=2 b-jets ($\lvert\eta\rvert < 2.5$; 60% eff.)	=1 b-jet ($\lvert\eta\rvert < 2.5$; 60% eff.)	
	$m_{lb} < 153$ GeV		$m_{lb} > 153$ GeV	OR
	$m_{l\nu b} \in [120.6, 234.6]$ GeV		$m_{l\nu b} \notin [120.6, 234.6]$ GeV	OR
	$m_{jl\nu b} > 320$ GeV		$m_{jl\nu b} < 320$ GeV	OR
	trapez. requirement		veto trapez. requirement	OR
	$H_{\mathrm{T}} > 190$ GeV		$H_{\mathrm{T}} < 190$ GeV	

Table 4.1: Summary of the selection criteria for defining the pre-selection, the signal and the two control regions.

respectively. The $t\bar{t}$ control region is very pure in $t\bar{t}$, which represents 74% of the expected events. In the W+jets control region, W+jets events represent 37% of the expected events, very similarly to the $t\bar{t}$ events (39%).

4.4 Background estimation

The normalisation of the t-channel signal and the different background processes (except multijet) is obtained from MC simulation scaled with the theoretical cross-section predictions listed in Section 4.2.2 and the luminosity used in this analysis. The shape modelling is taken from the simulations (for the multijet this is obtained from both data and simulation as it will be explained below).

4.4.1 Estimation of the multijet background

The multijet background is a non-negligible background originating from fake and non-prompt leptons. It is derived in two successive steps: Firstly, the shape for each distribution is obtained separately for the electron and the muon channels. Secondly, the normalisation is obtained from a fit to the data in specifically defined multijet-enriched regions.

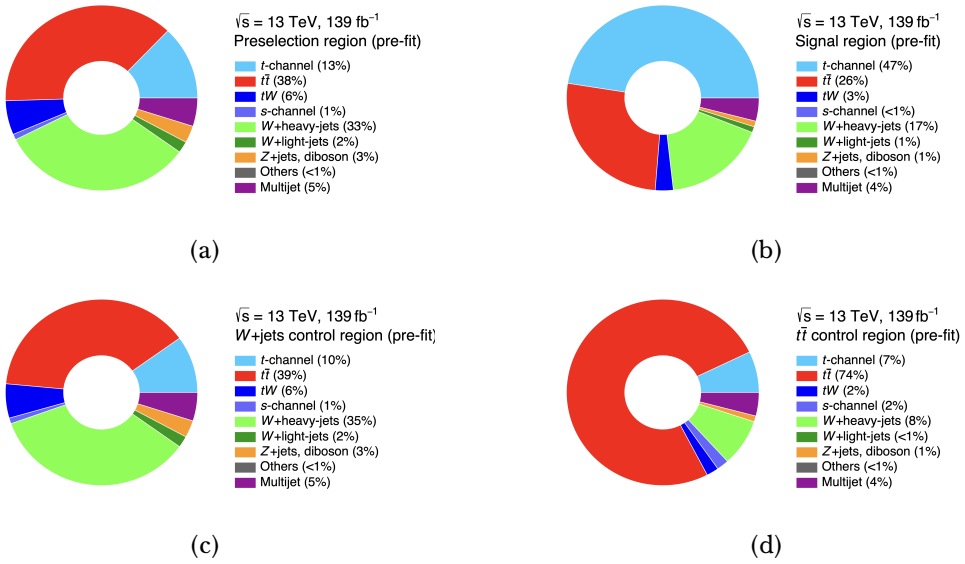


Figure 4.4: Expected contribution of the different backgrounds and the t-channel signal in the different regions used in this analysis: (a) pre-selection region, (b) signal region, (c) W+jets control region and (d) $t\bar{t}$ control region. The label "others" represents the $t\bar{t}Z$, $t\bar{t}W$, $t\bar{t}H$, tZq, tHq and tWZ productions. These results are previous to the normalisation fit described in Section 4.4.2.

The jet-electron model [388] is used to estimate the shape of the multijet background in the electron channel. The idea is to select jets that fake a signal (real) electron. The selected events must have exactly one of these fake electrons, without any additional requirement on true electrons. These fake electrons must satisfy the same p_{T} and η requirements as a signal candidate electron. In addition, they must deposit between 80% and 100% of its energy in the EM calorimeter (as a real electron would do). In order to reduce the contribution from converted photons, the fake electrons are required to have at least four tracks. Finally, a positive or negative charge is randomly assigned to the jet-electron candidate (there is no expected difference in charge for fake electrons). No special trigger requirement is applied. This jet-electron selection is applied to simulated dijet events to build the multijet model in the electron channel.

In contrast with the jet-electron model, the anti-muon model [388] provides an estimate of the multijet shape in the muon channel from collision data. Anti-muon candidates are selected inverting some identification requirements applied to real muon candidates, thus providing muons very likely to originate from multijet production. The requirement on the longitudinal impact parameter z_0 is removed, and the isolation criteria is relaxed. The energy loss in the calorimeters is required to be non-isolated[7], with an upper limit of 6 GeV. Events with one anti-muon candidate and no real muons are selected to build the multijet model in the muon channel.

The multijet normalisation is extracted separately for electrons and muons using a binned likelihood fit to the data. In the case of electrons, the jet-electron model cannot properly fake real electrons with a pseudorapidity falling in the EM crack region. Therefore, the multijet normalisation estimates in the electron channel are derived separately for events with a central jet-electron ($|\eta| < 1.5$) and for events with a forward jet-electron ($|\eta| > 1.5$). The fitted parameters are the normalisation factors associated with the multijet (i.e. central electrons, forward electrons and muons) and the normalisation of the W+jets and the merged top-quark processes

[7] The calorimeter transverse energy within a cone of size $R = 0.2$ around the muon is required to be larger than 3% the transverse momentum of the muon.

(i.e. t-channel signal, $t\bar{t}$ and single-top-quark backgrounds)[8]. The Z+jets, diboson and other minor backgrounds are fixed to their predicted values, which means they are not allowed to vary in the fit.

The used likelihood function is given by the product of the Poisson distributions in each bin of the fitted distribution. It also includes Gaussian priors to constrain the normalisation of the W+jets and the merged top-quark processes:

$$L(\beta_k) = \prod_{i=1}^{N_{\text{selections}}} \prod_{j=1}^{N_{\text{bins}}} \frac{e^{-\mu_{ij}} \cdot \mu_{ij}^{n_{ij}}}{n_{ij}!} \cdot \prod_{k=1}^{N_{\text{process}}} G\left(\beta_k; 1, \Delta_k\right)$$

$$\text{with} \qquad \mu_{ij} = \sum_{k=1}^{N_{\text{process}}} \mu_{ijk} \qquad \mu_{ijk} = \beta_k \cdot \tilde{\nu}_{ijk}, \tag{4.12}$$

where the index i runs over the three possible selections (central and forward electrons, and muons) and the index j runs over all the bins of each histogram. It includes a Poisson term in the observed number of events (n_{ij}) with the expectation value (μ_{ij}) defined as the sum of the expected contributions from the signal and the backgrounds (μ_{ijk} - the index k runs over the considered processes). For a given process, the expectation value in each selection is given by the product combining the predicted number of events ($\tilde{\nu}_{ijk}$) in the considered sample and a scale factor (β_k). The Δ_k term includes the Gaussian constraints, which are applied to the normalisation of the W+jets and the merged top-quark processes. These constraints come from the theoretical uncertainties on the normalisation of these processes[9].

The free parameters of the fit are the multijet normalisations. Although the multijet estimates are extracted separately for the electrons and muon channels, the W+jets and the top-quark scale factors are constrained to be the same for all channels.

The fit is performed simultaneously on the measured distribution of the $E_{\text{T}}^{\text{miss}}$ in the electron channel (central and forward) and the $m_T(l, E_{\text{T}}^{\text{miss}})$ in the muon channel using the previously discussed multijet templates. The variables have been

[8] These processes are merged to ensure the stability of the fit.

[9] The stability of the fit has been checked relaxing the W+jets normalisation constraint. This had no effect in the final multijet yields.

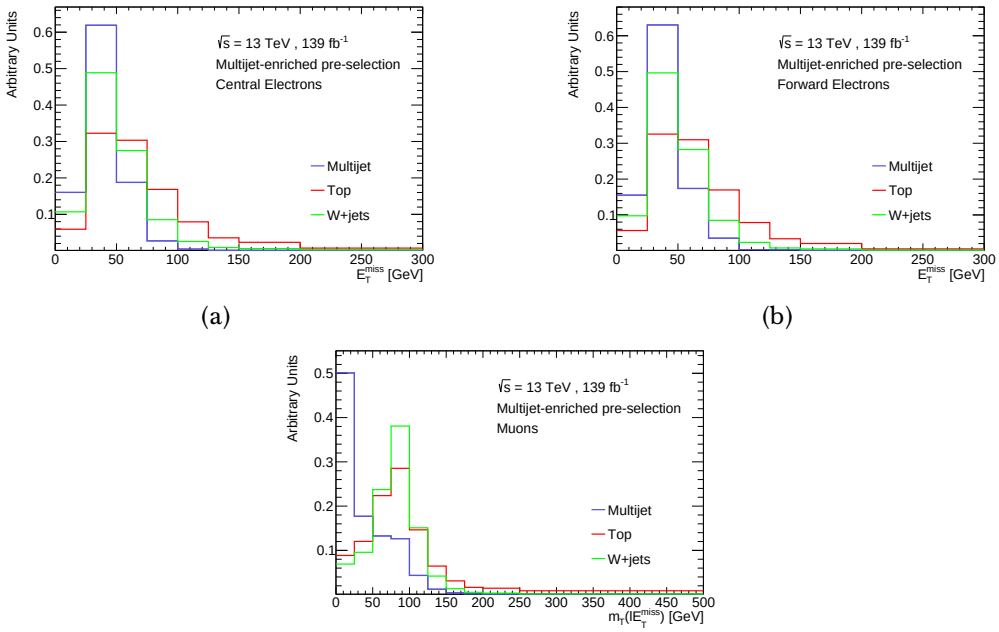


Figure 4.5: Distributions of the $E_{\mathrm{T}}^{\mathrm{miss}}$ (electron channel) and the $m_T(l, E_{\mathrm{T}}^{\mathrm{miss}})$ (muon channel) in the multijet-enriched pre-selection for (a) the central electron channel, (b) the forward electron channel and (c) the muon channel. The distributions are normalised to the total number of events in order to compare the shapes of the MC templates.

chosen carefully to ensure a good discrimination between the multijet shapes and those from the other fitted processes. The shapes of the fitted processes are depicted in Figures 4.5-4.6, and the correlations between the fitted parameters can be found in Figure 4.7. No significant correlations are found between the multijet normalisations and the W+jets or the merged top-quark processes, which is a robustness test of the procedure.

The multijet normalisation is extracted separately for the signal region and the W+jets control region, and for the $t\bar{t}$ control region. The reason is that the signal region and the W+jets control region are subsets of the pre-selection region, in contrast with the $t\bar{t}$ control region. Hence, in total two fits are

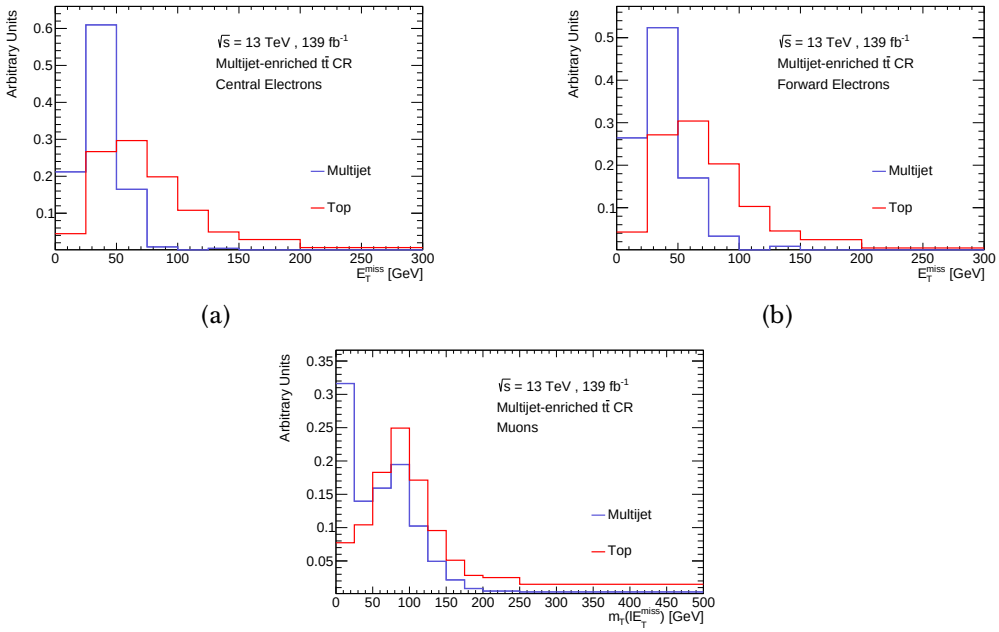
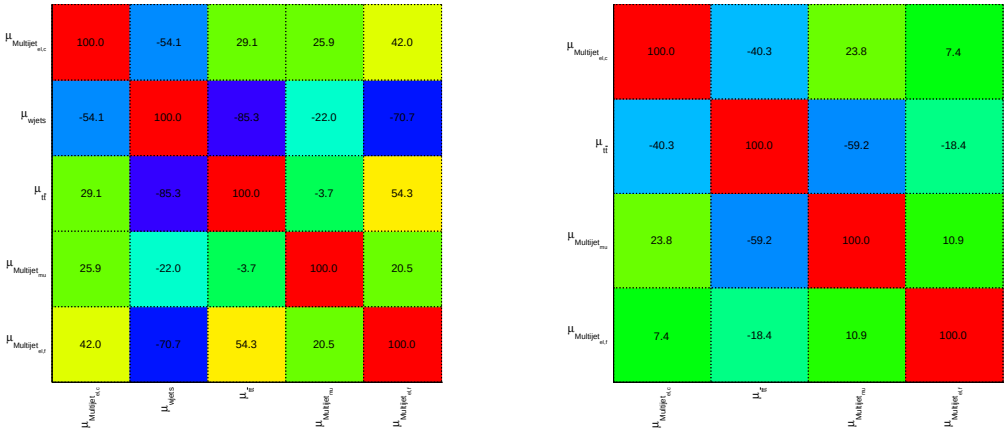


Figure 4.6: Distributions of the $E_{\mathrm{T}}^{\mathrm{miss}}$ (electron channel) and the $m_T(l, E_{\mathrm{T}}^{\mathrm{miss}})$ (muon channel) in the multijet-enriched $t\bar{t}$ control region for (a) the central electron channel, (b) the forward electron channel and (c) the muon channel. The distributions are normalised to the total number of events in order to compare the shapes of the MC templates.



(a) (b)

Figure 4.7: Correlation factors of the different fitted parameters in the multijet normalisation estimate for (a) the multijet-enriched pre-selection and (b) the multijet-enriched $t\bar{t}$ control region.

performed. As mentioned above, the fits are performed in specifically defined multijet-enriched regions. These are defined as the pre-selection region and the $t\bar{t}$ control region but leaving out the $E_{\mathrm{T}}^{\mathrm{miss}} > 35$ GeV requirement (electron channel) or the $m_T(l, E_{\mathrm{T}}^{\mathrm{miss}}) > 60$ GeV requirement (muon channel), and the requirement in Equation 4.2. Due to the low W+jets contribution in the $t\bar{t}$ control region, this background is also kept fixed in this fit.

The fitted distributions of the $E_{\mathrm{T}}^{\mathrm{miss}}$ in the electron channel and the $m_T(l, E_{\mathrm{T}}^{\mathrm{miss}})$ in the muon channel are shown in Figures 4.8-4.9 for the multijet-enriched pre-selection region and for the multijet-enriched $t\bar{t}$ control region, respectively. The number of multijet events containing exactly one jet-electron (either central or forward) or exactly one anti-muon is extracted from the normalisation factors derived from the fit. Table 4.2 shows these numbers for the different regions used in the analysis. The normalisation factors obtained in each of the two fits associated to the W+jets and the merged top-quark contributions are reported in Table 4.3 with their statistical uncertainties. These scale factors are used to estimate the multijet background normalisation but are not propagated to the next steps of the analysis.

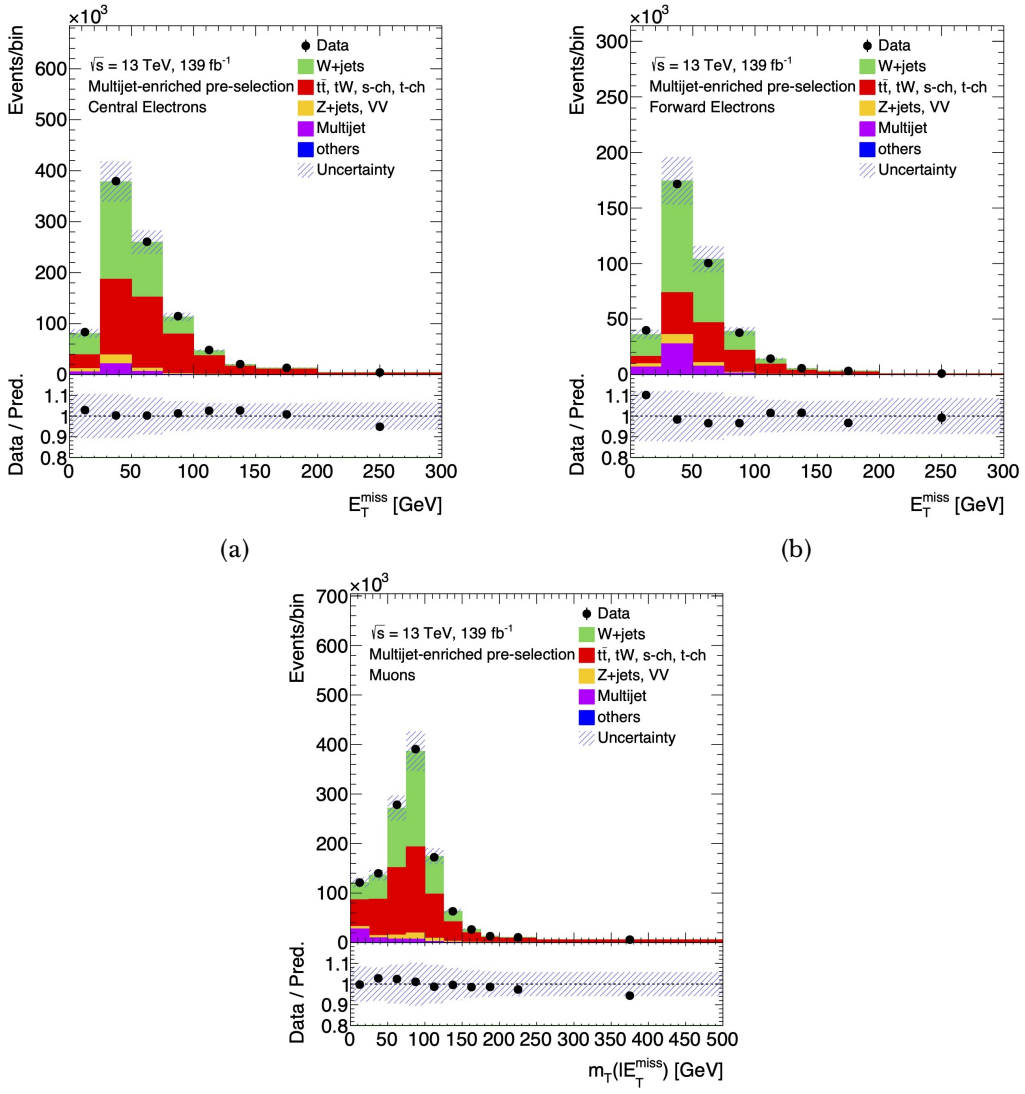
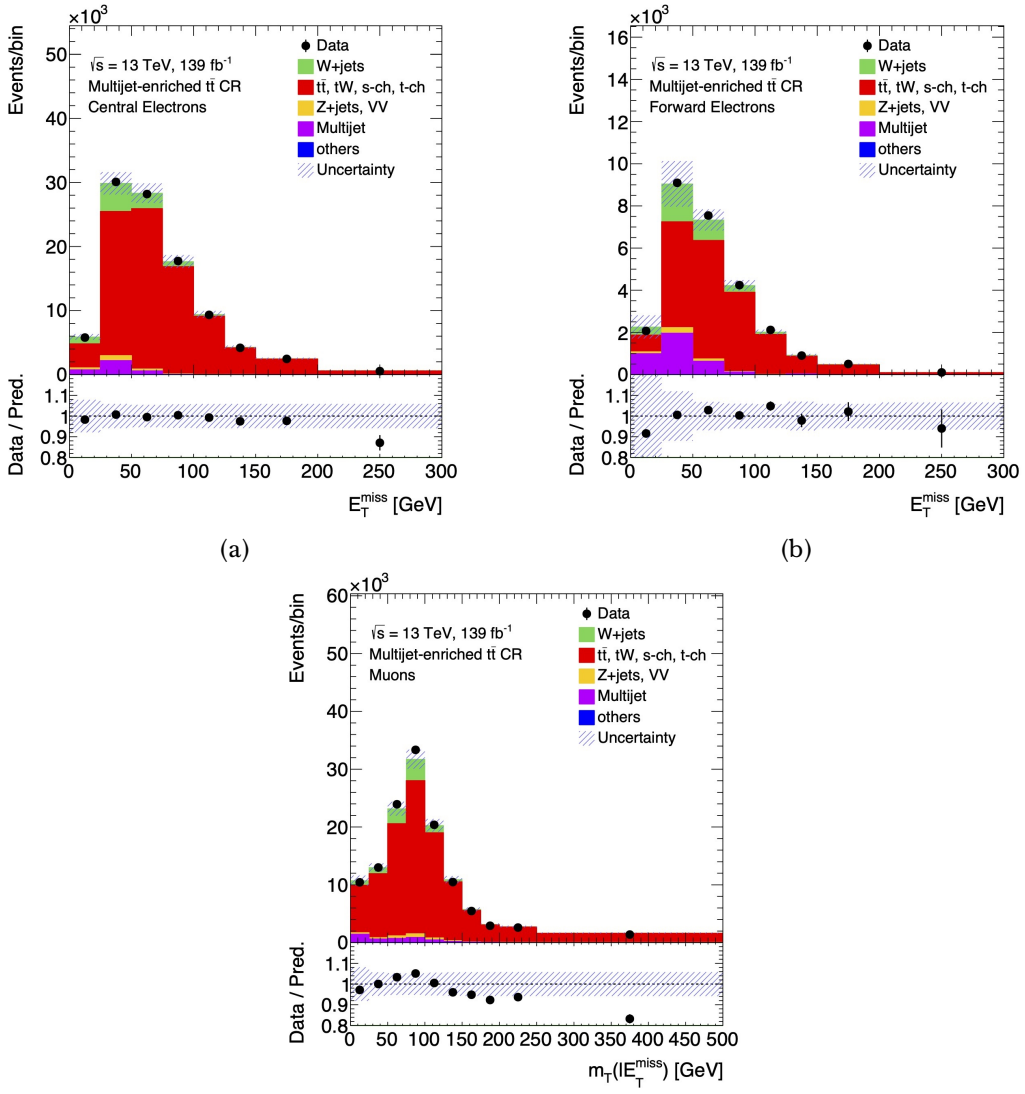


Figure 4.8: Distributions of the post-fit $E_{\mathrm{T}}^{\mathrm{miss}}$ in the multijet-enriched pre-selection region for (a) the central electron channel and (b) the forward electron channel, and distribution of the post-fit $m_T(l, E_{\mathrm{T}}^{\mathrm{miss}})$ in the multijet-enriched pre-selection region for (c) the muon channel. The predicted distributions are rescaled using the fitted normalisation factors. VV stands for the diboson background. The label "others" represents the $t\bar{t}Z$, $t\bar{t}W$, $t\bar{t}H$, tZq, tHq and tWZ productions. The uncertainty bands correspond to the statistical uncertainties together with the merged top-quark processes, W+jets and multijet normalisation uncertainties. The lower panels show the ratio of data to prediction in each bin. The last bin of the distributions includes the overflow.



(a) (b)

(c)

Figure 4.9: Distributions of the post-fit $E_{\mathrm{T}}^{\mathrm{miss}}$ in the multijet-enriched $t\bar{t}$ control region for (a) the central electron channel and (b) the forward electron channel, and distribution of the post-fit $m_T(l, E_{\mathrm{T}}^{\mathrm{miss}})$ in the multijet-enriched $t\bar{t}$ control region for (c) the muon channel. The predicted distributions are rescaled using the fitted normalisation factors. VV stands for the diboson background. The label "others" represents the $t\bar{t}Z$, $t\bar{t}W$, $t\bar{t}H$, tZq, tHq and tWZ productions. The uncertainty bands correspond to the statistical uncertainties together with the merged top-quark processes and multijet normalisation uncertainties. The lower panels show the ratio of data to prediction in each bin. The last bin of the distributions includes the overflow.

Channel	W+jets CR	$t\bar{t}$ CR	Pre-selection	Signal region
Electron	36000 $\pm$ 9000	3400 $\pm$ 1400	39000 $\pm$ 10000	2300 $\pm$ 600
Muon	12200 $\pm$ 2400	2100 $\pm$ 1100	13400 $\pm$ 2700	1230 $\pm$ 250

Table 4.2: Multijet event yields estimated in the electron and muon channels with the jet-electron and anti-muon models. They are given for the W+jets and $t\bar{t}$ control regions as well as for the pre-selection and the signal region. The uncertainties come from the comparison of different fit setups, as explained in Appendix A.

Process	Multijet-enriched $t\bar{t}$ CR	Multijet-enriched pre-selection
W+jets	fixed	1.320 $\pm$ 0.009
$t\bar{t}$,tW,s,t-channel	1.014 $\pm$ 0.004	0.919 $\pm$ 0.005

Table 4.3: Normalisation factors extracted for the W+jets and merged top-quark contributions from the maximum-likelihood fit of the distributions of the $E_{\mathrm{T}}^{\mathrm{miss}}$ (electron channel) and $m_T(l, E_{\mathrm{T}}^{\mathrm{miss}})$ (muon channel) observed in the multijet-enriched $t\bar{t}$ control region and in the multijet-enriched pre-selection region. The uncertainties correspond to the statistical uncertainties provided from the likelihood fit.

The uncertainties on the multijet normalisation estimation are obtained comparing the fit results using alternative variables or different MC generators for the top-quark processes. The overall impact of these sources leads to a 20% normalisation variation in the signal and W+jets control region, and a 40% variation in the $t\bar{t}$ control region. The uncertainties on the shapes provided by the jet-electron and the anti-muon models are estimated by comparing them with modified templates. These are obtained by modifying the parameters used in the jet-electron and anti-muon selections. Further details on the estimation of these uncertainties can be found in Appendix A.

4.4.2 Estimation of the main backgrounds normalisation

The normalisations of the main contributing backgrounds (top-quark processes and W+jets) are constrained from data. A data-driven method is used, based on a maximum-likelihood fit to the number of data events obtained in the control

regions defined in Section 4.3.4. The number of events in the signal region is also included in the fit to constrain the low but non-negligible contribution of t-channel signal events in the control regions. The fit is separately performed for the regions defined with positive (negative) charged leptons for the measurement of top-quark (top-antiquark) events, as well as for the inclusive regions containing both lepton charges.

In this fit, the top-quark backgrounds are merged ($t\bar{t}$, tW and s-channel) and also the W+heavy- and W+light-flavour jets. This is is done to reduce the statistical fluctuations, thus increasing the stability of the fit.

The fitted parameters are the scale factors associated with the W+jets and the top-quark backgrounds, and also the scale factor associated with the t-channel signal events. The Z+jets, diboson, multijet and other minor backgrounds are fixed to their predictions.

The used likelihood function is very similar to Equation 4.4.1. In this case, the fit is performed on the total number of events. Additionally, the signal events are not merged with the other top-quark process but are an independent fitted parameter. Therefore, the expected number of events can be rewritten as $\mu_i = \mu_i^s + \sum_{k=1}^{N_{backgrounds}} \mu_{ik}^b$, where $\mu_i^s = \beta^s \cdot \tilde{\nu}_i^s$ is the expectation value of signal events for a given selection i, and $\mu_{ik}^b = \beta_k^b \cdot \tilde{\nu}_{ik}^b$ is the expectation value of background events for a process k and selection i. Consequently, β^s (β_k^b) and $\tilde{\nu}_i^s$ ($\tilde{\nu}_{ik}^b$) are the scale factor and the predicted number of events of the signal (the considered background process), respectively.

The scale factors extracted for the W+jets and top-quark backgrounds, and also for the t-channel signal are reported in Table 4.4.

Values compatible with unity are found for the merged top-quark background processes, which are dominated by $t\bar{t}$ production. Therefore, the evaluated data-driven normalisations are consistent with the NNLO cross-section calculations, whose central values are used to normalise the corresponding MC simulated event samples previous to the fit.

The scale factor extracted for the t-channel signal contribution is compatible with 1.14 ± 0.19, which is the scale factor obtained from the ATLAS measurement of the t-channel production cross-section at $\sqrt{s}$ = 13 TeV [389].

Process	top quark	top antiquark	top quark and antiquark
t-channel	1.04 ± 0.31	1.13 ± 0.29	1.07 ± 0.29
W+jets	1.2 ± 0.4	1.15 ± 0.35	1.2 ± 0.4
$t\bar{t}$,Wt,s-channel	1.00 ± 0.31	0.99 ± 0.31	1.00 ± 0.31

Table 4.4: Scale factors extracted for the t-channel signal contribution, and for the W+jets and top-quark background processes from the simultaneous maximum-likelihood fit to the numbers of data events observed in the signal region and the two control regions for top quarks and top antiquarks, both inclusively and separately. The uncertainties include both statistical and all systematic effects.

The same set of scale factors is also found when the multijet contribution is not kept fixed but allowed to float within a 50% normalisation constraint, showing that there is no significant sensitivity of the likelihood fit to the multijet contribution.

In the next steps of the analysis, the W+jets and the top-quark backgrounds are renormalised using the scale factors reported in Table 4.4. The scale factor obtained for the t-channel signal events is only used in the plots showing the data/expectation comparisons at reconstruction level. This will be explained in more detail in Section 4.7.

4.5 Event yields and distributions in the signal and control regions

Table 4.5 provides the predicted signal and background event yields for the combined electron and muon channels in the pre-selection, signal and control regions. The multijet background is normalised as explained in Section 4.4.1, whereas the other backgrounds and the t-channel signal are normalised to the theoretical cross-sections. These numbers are compared to the events observed in data. Additionally, the S/B, the signal significance and the ratio of observed to expected number of events (Data/Prediction) are also shown for each region.

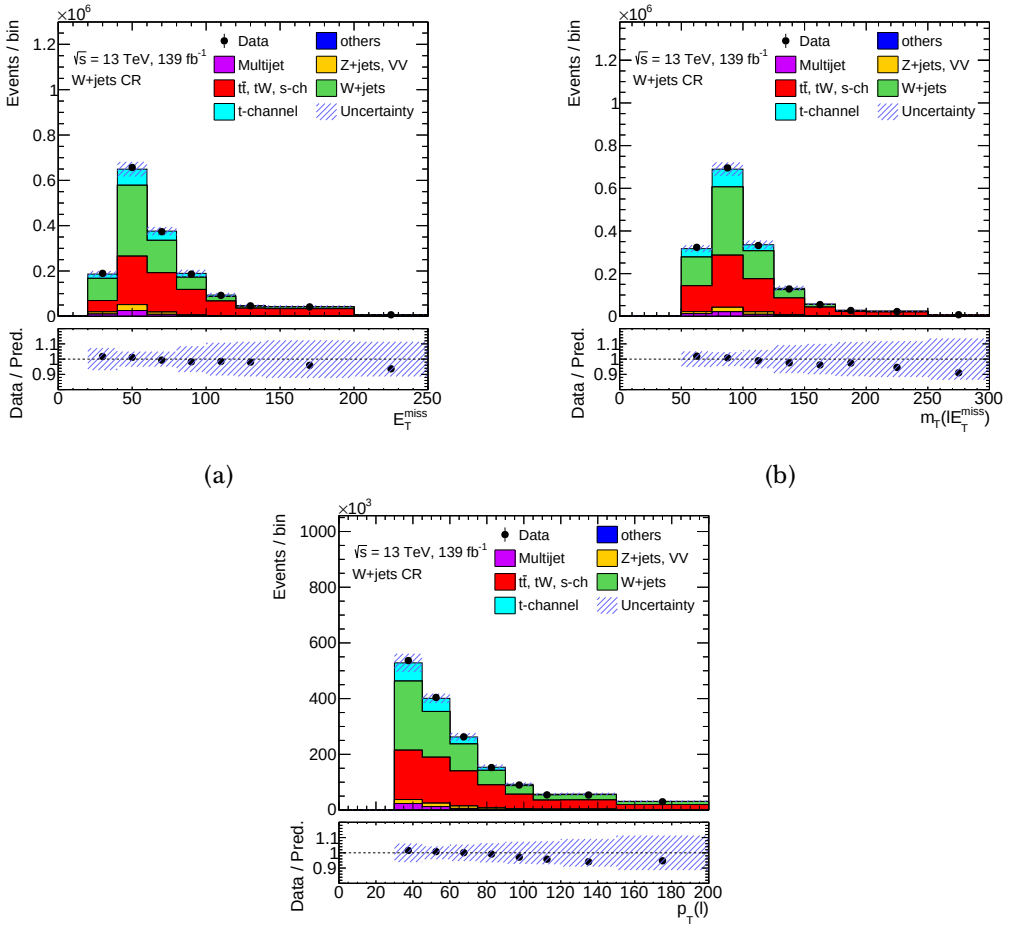
Good overall Data/Prediction agreement is found in the signal and the control regions for relevant kinematic variables, such as the $E_{\mathrm{T}}^{\mathrm{miss}}$, the transverse mass of

Process	Pre-selection	Signal region	$t\bar{t}$ CR	W+jets CR
t-channel	219 000 ± 11 000	70 600 ± 3 500	13 480 ± 680	148 200 ± 7 400
$t\bar{t}$, tW, s-channel	736 000 ± 39 000	43 200 ± 2 400	147 800 ± 8 400	693 000 ± 37 000
W+jets	590 000 ± 200 000	26 200 ± 8 900	16 100 ± 5 500	560 000 ± 190 000
Z+jets, diboson	52 900 ± 5 100	2 120 ± 350	2 620 ± 360	50 800 ± 4 900
Others	494 ± 38	30 ± 4	79 ± 6	464 ± 36
Multijet	52000 ± 10 000	3 500 ± 640	5 500 ± 1 800	48 500 ± 9 400
Total expected	1 650 000 ± 210 000	145 600 ± 9 900	186 000 ± 10 000	1 510 000 ± 2 000 000
Data	1750918	154361	188326	1596557
S/B	0.15 ± 0.02	0.94 ± 0.13	0.08 ± 0.01	0.11 ± 0.02
Significance	170 ± 14	185 ± 11	31.3 ± 1.8	121 ± 10
Data/Prediction	1.06 ± 0.13	1.06 ± 0.07	1.02 ± 0.06	1.06 ± 0.14

Table 4.5: Pre-fit event yields in the pre-selection and signal regions and in the $t\bar{t}$ and W+jets control regions. The predictions are derived from simulated event samples normalised to the theoretical cross-sections. For multijet production the normalisation is estimated using a data-driven likelihood fit. The label "Others" represents the $t\bar{t}Z$, $t\bar{t}W$, $t\bar{t}H$, tZq, tHq, and tWZ productions. The data-driven scale factors obtained for the top-quark and W+jets background processes are not considered to compute these event yields. The uncertainties shown account for all systematic effects and the uncertainty due to limited MC statistics. The statistical significance and the ratios of the expected S/B and of the observed to expected number of events are also given.

the W boson and the lepton p_{T}. Figures 4.10-4.12 show the combined electron and muon events in terms of these three kinematic variables in these regions. The distributions in all figures are normalised to the results of the maximum likelihood fit explained in Section 4.4.2.

From the reconstructed top quark, the lepton momentum is boosted to the top-quark rest frame. The angles $\theta_{lx'}$, $\theta_{ly'}$, $\theta_{lz'}$ are then derived by its projection along the axes defined in Section 1.3. Figures 4.13-4.14 show the angular distributions measured at reconstruction level in the control regions for the combined channel. There is a good level of agreement between the data and the predictions. For the definition of the z'-axis in the $t\bar{t}$ control region, the leading b-jet is taken as the spectator quark.



(a) (b)

(c)

Figure 4.10: Post-fit distributions of (a) the $E_{\mathrm{T}}^{\mathrm{miss}}$, (b) the $m_T(l, E_{\mathrm{T}}^{\mathrm{miss}})$ and (c) the $p_{\mathrm{T}}(l)$ in the W+jets control region for the electron and muon channels merged, comparing observed data, shown as the black points with statistical uncertainties, to SM signal and background predictions. The multijet background is estimated using data-driven techniques, while contributions from simulated W+jets, top-quark backgrounds and t-channel event samples are normalised to the results of a maximum-likelihood fit to event yields in the signal and control regions. VV stands for the diboson background. The label "others" represents the $t\bar{t}Z$, $t\bar{t}W$, $t\bar{t}H$, tZq, tHq and tWZ productions. The uncertainty bands include both statistical and systematic uncertainties (described in Section 4.8.2). The lower panels show the ratio of data to prediction in each bin. The last bin of the distributions includes the overflow.

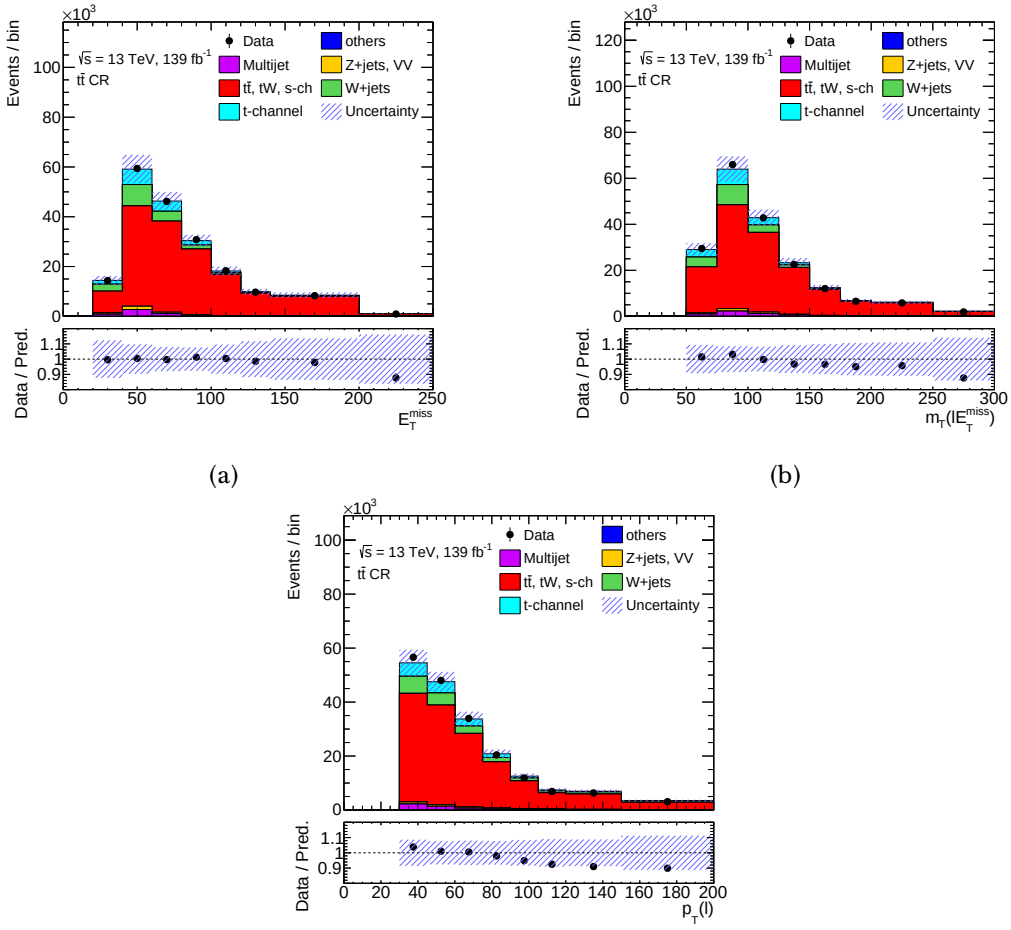
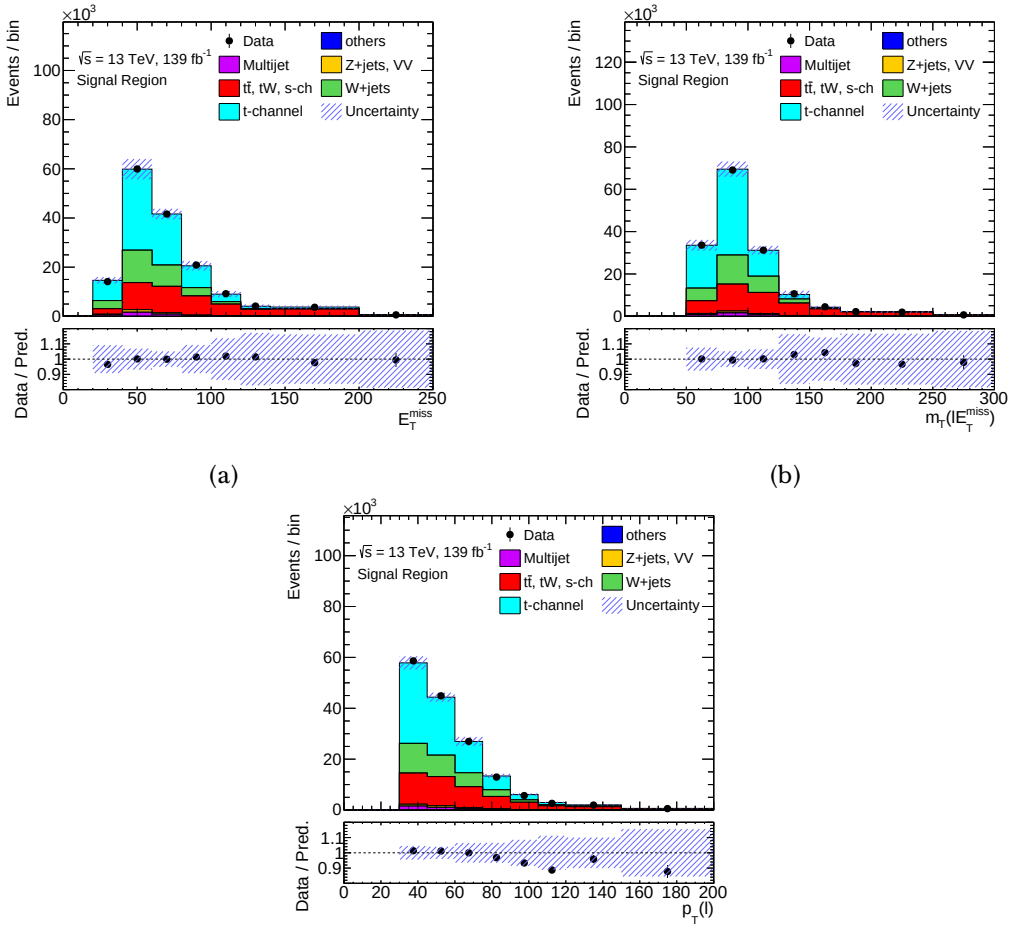


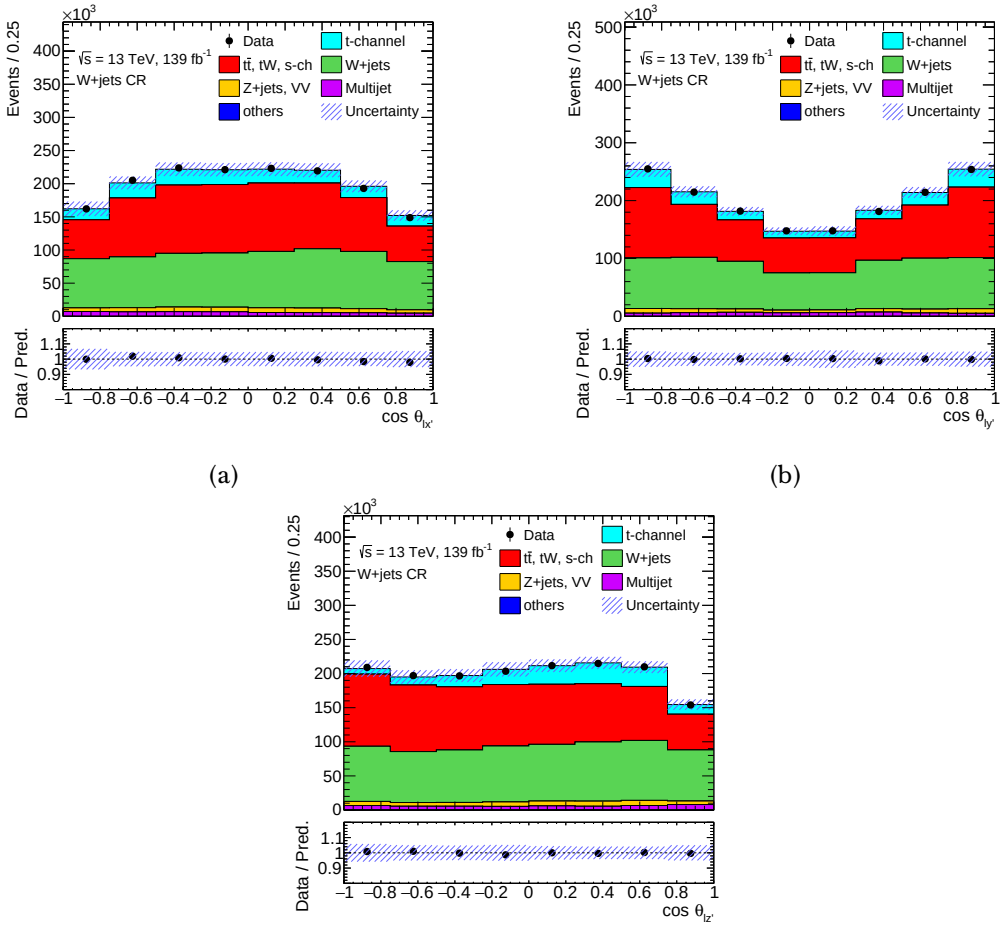
Figure 4.11: Post-fit distributions of (a) the $E_{\mathrm{T}}^{\mathrm{miss}}$, (b) the $m_T(l, E_{\mathrm{T}}^{\mathrm{miss}})$ and (c) the $p_{\mathrm{T}}(l)$ in the $t\bar{t}$ control region for the electron and muon channels merged, comparing observed data, shown as the black points with statistical uncertainties, to SM signal and background predictions. The multijet background is estimated using data-driven techniques, while contributions from simulated W+jets, top-quark backgrounds and t-channel event samples are normalised to the results of a maximum-likelihood fit to event yields in the signal and control regions. VV stands for the diboson background. The label "others" represents the $t\bar{t}Z$, $t\bar{t}W$, $t\bar{t}H$, tZq, tHq and tWZ productions. The uncertainty bands include both statistical and systematic uncertainties (described in Section 4.8.2). The lower panels show the ratio of data to prediction in each bin. The last bin of the distributions includes the overflow.



(a) (b)

(c)

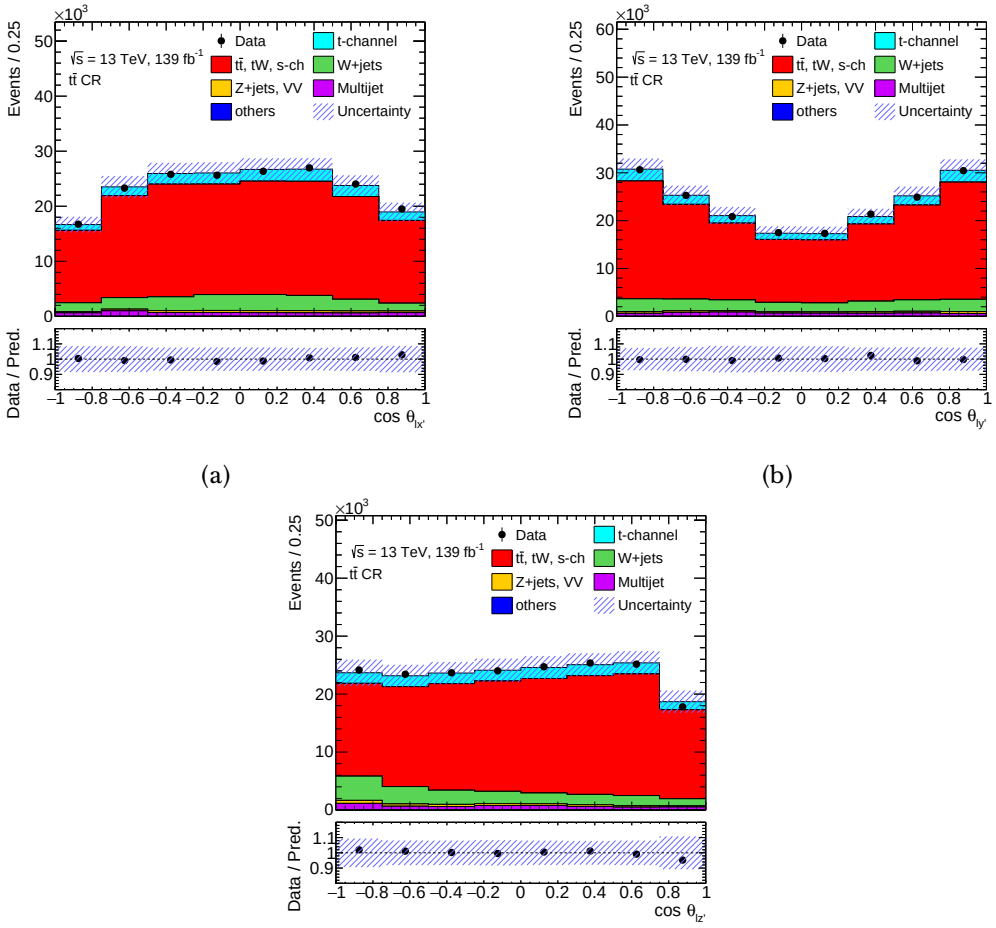
Figure 4.12: Post-fit distributions of (a) the $E_{\mathrm{T}}^{\mathrm{miss}}$, (b) the $m_T(l, E_{\mathrm{T}}^{\mathrm{miss}})$ and (c) the $p_{\mathrm{T}}(l)$ in the signal region for the electron and muon channels merged, comparing observed data, shown as the black points with statistical uncertainties, to SM signal and background predictions. The multijet background is estimated using data-driven techniques, while contributions from simulated W+jets, top-quark backgrounds and t-channel event samples are normalised to the results of a maximum-likelihood fit to event yields in the signal and control regions. VV stands for the diboson background. The label "others" represents the $t\bar{t}Z$, $t\bar{t}W$, $t\bar{t}H$, tZq, tHq and tWZ productions. The uncertainty bands include both statistical and systematic uncertainties (described in Section 4.8.2). The lower panels show the ratio of data to prediction in each bin. The last bin of the distributions includes the overflow.



(a) (b)

(c)

Figure 4.13: Post-fit distributions of (a) $\cos\theta_{lx'}$, (b) $\cos\theta_{ly'}$ and (c) $\cos\theta_{lz'}$ in the W+jets control region for the electron and muon channels merged, comparing observed data, shown as the black points with statistical uncertainties, to SM signal and background predictions. The multijet background is estimated using data-driven techniques, while contributions from simulated W+jets, top-quark backgrounds and t-channel event samples are normalised to the results of a maximum-likelihood fit to event yields in the signal and control regions. VV stands for the diboson background. The label "others" represents the $t\bar{t}Z$, $t\bar{t}W$, $t\bar{t}H$, tZq, tHq and tWZ productions. The uncertainty bands include both statistical and systematic uncertainties (described in Section 4.8.2). The lower panels show the ratio of data to prediction in each bin.



(a) (b)

(c)

Figure 4.14: Post-fit distributions of (a) $\cos\theta_{lx'}$, (b) $\cos\theta_{ly'}$ and (c) $\cos\theta_{lz'}$ in the $t\bar{t}$ control region for the electron and muon channels merged, comparing observed data, shown as the black points with statistical uncertainties, to SM signal and background predictions. The multijet background is estimated using data-driven techniques, while contributions from simulated W+jets, top-quark backgrounds and t-channel event samples are normalised to the results of a maximum-likelihood fit to event yields in the signal and control regions. VV stands for the diboson background. The label "others" represents the $t\bar{t}Z$, $t\bar{t}W$, $t\bar{t}H$, tZq, tHq and tWZ productions. The uncertainty bands include both statistical and systematic uncertainties (described in Section 4.8.2). The lower panels show the ratio of data to prediction in each bin.

4.6 Definition of particle level objects and fiducial region

In order to reduce the dependency on phenomenological models which describe colour reconnection, initial- and final-state radiation, and fragmentation, the measured differential angular distributions are unfolded to particle level.

4.6.1 Particle-level objects

Particle-level objects are constructed from stable particles (lifetime > $0.3 \cdot 10^{10}$ s) of the MC event record (ME+PS) within the observable pseudorapidity range[10]. These objects do not take into account any detector effect or any additional pp interactions (pile-up).

- **Leptons**: Particle-level leptons are defined as electrons, muons or neutrinos that originate from either a W- or Z-boson decay, including those emerging from a subsequent τ-lepton decay. Leptons from hadronic decays are excluded, either directly or via a τ decay. The remaining leptons are assumed to come from a W-boson decay. No isolation requirement is imposed to the selected charged lepton and its four-momentum is calculated including photons within a cone of size $\Delta R = 0.1$ around the charged lepton.

- **Missing transverse momentum**: The $E_{\mathrm{T}}^{\mathrm{miss}}$ is calculated as the vector sum of all the selected neutrinos.

- **Jets**: Particle-level jets are reconstructed using the anti-k_t algorithm with a radius parameter of 0.4. All stable particles are used to reconstruct the jets, excluding electrons, muons, neutrinos, and photons used in the definition of the selected charged leptons. A particle-level jet is identified as b-jet if the jet is within $|\eta| < 2.5$ and a b-hadron is associated with a ghost-matching technique [390]. The hadron must have $p_{\mathrm{T}} > 5\,\mathrm{GeV}$.

All particle-level charged leptons identified within a cone of size $\Delta R = 0.4$ around a selected particle-level jet are removed.

[10] This corresponds to the detector acceptance $|\eta| < 5$.

The W boson and the top quark at particle level are defined applying the same procedure as described in Section 4.3.2, but to particle-level objects.

4.6.2 Fiducial region at particle level

The differential cross-section measurements provided in this analysis are unfolded to the particle level in a fiducial region. This fiducial region is defined to be close to the measured phase-space using the particle-level objects so that acceptance effects are minimal. Exactly one particle-level electron or muon with $p_{\mathrm{T}} > 30$ GeV and $|\eta| < 2.5$ (excluding the region $1.37 < |\eta| < 1.52$ for the case of electrons) is required. There must be two particle-level jets with $p_{\mathrm{T}} > 30$ GeV; exactly one of these jets must be identified as a b-jet with $|\eta| < 2.5$ while the other jet must satisfy $|\eta| < 4.5$. The particle-level $E_{\mathrm{T}}^{\mathrm{miss}}$ is required to be larger than 35 GeV. The two additional multijet background rejection criteria, i.e. $m_T(l, E_{\mathrm{T}}^{\mathrm{miss}})$ larger than 60 GeV and the requirement in Equation 4.2, as well as the remaining signal region requirements are also applied to the particle-level objects.

Table 4.6 summarises the signal selection criteria, performed using particle-level objects, for defining the fiducial region used in this analysis.

Fiducial region
Exactly one electron or muon
Exactly two jets
Exactly one b-tagged jet
$E_{\mathrm{T}}^{\mathrm{miss}} > 35$ GeV
$m_{\mathrm{T}}(l, E_{\mathrm{T}}^{\mathrm{miss}}) > 60$ GeV
$p_{\mathrm{T}}(\ell) > 50 \left(1 - \frac{\pi - \lvert\Delta\phi(p_{\mathrm{T}}(j_1), p_{\mathrm{T}}(l))\rvert}{\pi - 1}\right)$ GeV
$m_{lb} < 153$ GeV
$m_{l\nu b} \in [120.6, 234.6]$ GeV
$m_{jl\nu b} > 320$ GeV
trapez. requirement
$H_{\mathrm{T}} > 190$ GeV

Table 4.6: Summary of the signal selection criteria, performed using particle-level objects, for defining the fiducial region.

4.7 Measurement of the normalised differential cross-sections as a function of the polarisation angles

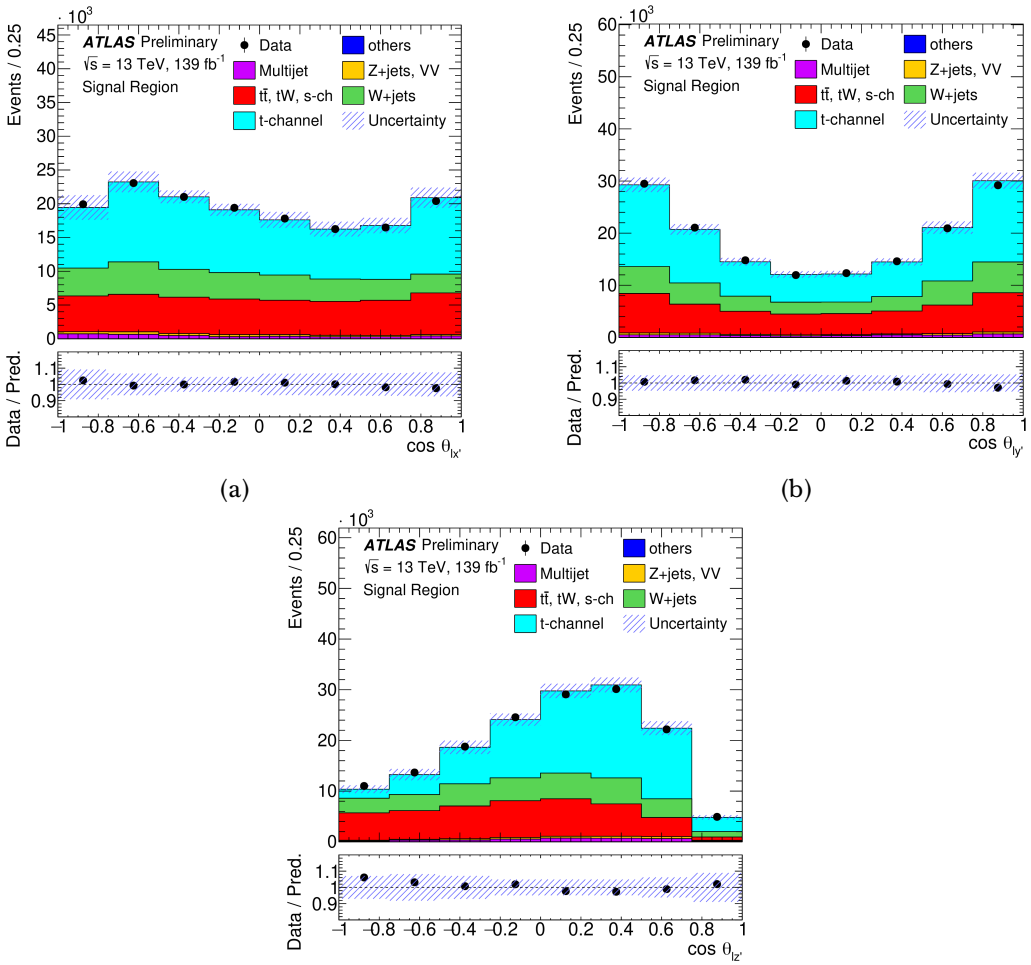
The distributions observed at reconstruction level for the three measured angular distributions $(\cos\theta_{lx'}, \cos\theta_{ly'}, \cos\theta_{lz'})$ are shown in Figure 4.15 for the inclusive signal region. They are compared to the predicted signal and background distributions, normalised to the results of the maximum-likelihood fit described in Section 4.4.2.

In order to compare with theoretical predictions, the different backgrounds are subtracted and the angular distributions are unfolded to the particle level in a fiducial region. The unfolding technique corrects for the detector effects, which include the finite resolution of the detector and the trigger, reconstruction and selection efficiencies. The fiducial region, which has already been defined in Section 4.6.2, is defined as close as possible to the reconstruction level region to reduce the acceptance correction to the minimum.

The procedure can be divided in three different steps. Firstly, the different contributing backgrounds are subtracted from the data; secondly, the background-subtracted data is unfolded to the particle level to correct for the detector effects. Finally, each bin of the unfolded distributions is divided by the integral over all bins and the result is presented as a normalised differential cross-section measurement. A schematic view of the whole process is displayed in Figure 4.16. The normalised differential cross-section measurements of the three angular distributions, which are sensitive to the top-quark polarisation, can be directly compared to theoretical predictions to check for any possible deviation from the SM predictions due to the presence of new physics.

4.7.1 Unfolding technique

The starting point to perform the unfolding are the background subtracted angular distributions at reconstruction level. In order to unfold these distributions, the method pictures the problem with an 'effect' and a 'cause'. The number of



(a) (b)

(c)

Figure 4.15: Post-fit distributions of (a) $\cos\theta_{lx'}$, (b) $\cos\theta_{ly'}$ and (c) $\cos\theta_{lz'}$ in the signal region for the electron and muon channels merged, comparing observed data, shown as the black points with statistical uncertainties, to SM signal and background predictions. The multijet background is estimated using data-driven techniques, while contributions from simulated W+jets, top-quark backgrounds and t-channel event samples are normalised to the results of a maximum-likelihood fit to event yields in the signal and control regions. VV stands for the diboson background. The label "others" represents the $t\bar{t}Z$, $t\bar{t}W$, $t\bar{t}H$, tZq, tHq and tWZ productions. The uncertainty bands include both statistical and systematic uncertainties (described in Section 4.8.2). The lower panels show the ratio of data to prediction in each bin.

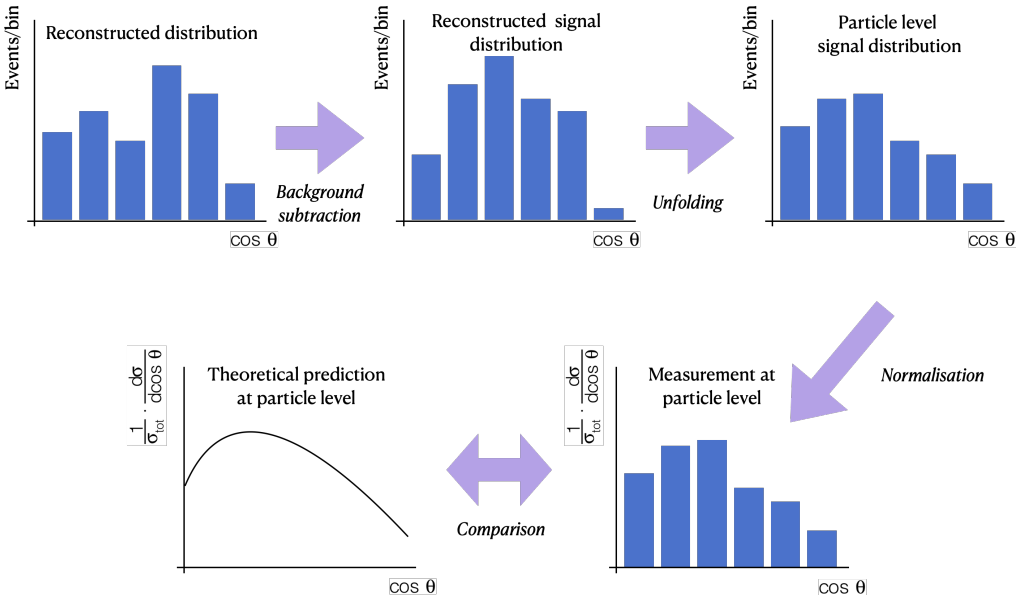


Figure 4.16: Schematic view summarising the different steps to provide the normalised differential cross-section measurements: The estimated background is subtracted from the measured distribution at reconstruction level. The unfolding to particle level technique is applied to the background-subtracted data to obtain the unfolded distributions. These distributions are provided as normalised differential cross-section measurements, which can finally be compared with theoretical predictions.

measured t-channel single-top-(anti)quark events at reconstruction level in a bin j is the effect, whereas the number of produced t-channel events in a pp collisions in a bin k corresponds to the cause. The observed number of data events in a bin N_j^{data} of the measured distribution at reconstruction level can be described by:

$$N_j^{data} = B_j + \sum_k M_{jk}\epsilon_k L_{\text{int}} d\sigma_k, \tag{4.13}$$

where B_j is the sum of all background contributions in bin j, $d\sigma_k$ is the estimated t-channel cross-section in bin k, ϵ_k is the efficiency for an event to be selected in bin k, L_{int} is the integrated luminosity, and M_{jk} is the so-called ***migration matrix***. This describes the probability of migration of particle level events in bin k to bin j after detector reconstruction of the event. The efficiency ϵ_k includes signal acceptance and detector efficiencies:

$$\epsilon_k = S_k^{sel,MC} / S_k^{tot,MC}, \tag{4.14}$$

where $S_k^{tot,MC}$ is the number of generated MC events in bin k and $S_k^{sel,MC}$ is the number of selected MC events in bin k after all requirements are applied.

In order to recover the particle level distribution associated to a reconstruction level measurement, the migration matrix M needs to be inverted. In general, this matrix does not have an exact inverse such that $M \cdot M^{-1} = I$, where I is the identity matrix. Hence, approximations are needed to perform the matrix inversion. In this analysis the iterative bayesian approach [391] implemented in the RooUnfold package [392] is used for this purpose. In this algorithm, the reconstruction level t-channel is used as an initial prior. This distribution is unfolded to the particle level, and the resulting unfolded distribution is used as the new prior for the next iteration. The number of iterations is the regularisation parameter and it is chosen such that spurious fluctuations are suppressed in the inversion result. This parameter is independent for each angular distribution. The determination of the number of iterations for each angular distribution will be explained in Section 4.6.1.

The measured expectation value for the number of signal events at particle level in each bin k of the fiducial volume ν_k^{particle} is obtained from the observed number

of events in each bin j of the reconstructed distribution N_j^{data}, after subtracting the sum of all background contributions B_j, according to:

$$\nu_k^{\text{particle}} = C_k^{\text{particle!reco}} \sum_j M_{jk}^{-1} C_j^{\text{reco!particle}} (N_j^{\text{data}} - B_j), \tag{4.15}$$

where $C_k^{\text{reco!particle}}$ is a correction factor that accounts for events that pass reconstruction selection but not particle level selection. It is defined as:

$$C_j^{\text{reco!particle}} = \frac{S_j^{\text{reco}} - S_j^{\text{reco!particle}}}{S_j^{\text{reco}}}, \tag{4.16}$$

where S_j^{reco} is the number of reconstructed signal events in bin j and $S_j^{\text{reco!particle}}$ is the number of events that pass the reconstruction-level selection but not the particle-level selection. $C_j^{\text{particle!reco}}$ is a correction factor that accounts for signal events that pass the particle-level selection but not the reconstruction-level selection:

$$C_k^{\text{particle!reco}} = \frac{1}{\epsilon_k} = \frac{S_k^{\text{particle}}}{S_k^{\text{particle}} - S_k^{\text{particle!reco}}}, \tag{4.17}$$

where S_k^{particle} is the number of signal events at particle level and $S_k^{\text{particle!reco}}$ is the number of events that pass the particle-level selection but not the reconstruction-level selection. The factor $C_k^{\text{particle!reco}}$ is the inverse of the efficiency ϵ_k.

Each normalised differential cross-section is determined by dividing the obtained ν_k^{particle} values in each bin by the integral over all bins.

The migration matrix and the selection efficiency are computed using samples of t-channel events simulated with the POWHEG-BOX+PYTHIA8 generator. They are calculated for the signal region defined with positive (negative) charged leptons for the measurement of top-quark (top-antiquark) events, as well as for the inclusive signal region containing both lepton charges.

Figure 4.17 shows the migration matrices obtained for each angular distribution for the inclusive measurement. As it can be seen, the number of events in the off-diagonal elements is below 16% for the $\cos\theta_{lx'}$ and $\cos\theta_{ly'}$ distributions, and

below 26% for the $\cos\theta_{lz'}$ distribution. The correction factors for the inclusive measurement are displayed in Figures 4.18-4.19. They are given for the baseline POWHEG-BOX+PYTHIA8 FS sample and also for POWHEG-BOX+PYTHIA8 simulated with AFII and for the SM predictions from PROTOS+PYTHIA8 (LO generator) and MG5_AMC@NLO+PYTHIA8 (alternative NLO generator). The discrepancies observed between the different NLO generators are considered as part of the systematic uncertainty, as described in Section 4.8. Similar migration matrices and correction factors are obtained for both top-quark and top-antiquark events.

The number of bins has been chosen in order to have a stable unfolding response with at least 70% of the events in the diagonal elements of the migration matrix. The number of iterations obtained from the different binning configurations has also been taken into account in this decision[11] (see Section 4.7.2). These stability criteria results in four bins for $\cos\theta_{lz'}$ and eight bins for both $\cos\theta_{lx'}$ and $\cos\theta_{ly'}$ distributions.

The stability of the unfolding procedure has been validated through convergence and closure tests performed by using template distributions constructed from the t-channel POWHEG-BOX+PYTHIA8, PROTOS+PYTHIA8 and MG5_AMC@NLO+PYTHIA8 samples. By using template distributions given by the PROTOS+PYTHIA8 and MG5_AMC@NLO+PYTHIA8 samples, generated including the effects on the tWb vertex of anomalous couplings or additional EFT operators, it is shown that the unfolding method recovers the generated distributions at particle level within the fiducial region. These tests are described in the following subsections for the inclusive measurements. Similar results were obtained for the separate top-quark and top-antiquark measurements.

4.7.2 Convergence test to fix the number of iterations

The regularisation parameter in the iterative bayesian unfolding is the number of iterations which is determined from convergence tests. These are carried out separately for each angular distribution. The tests are performed applying the unfolding corrections to the angular distributions provided at reconstruction level

[11] The larger the number of iterations, the larger the uncertainties induced by the unfolding procedure.

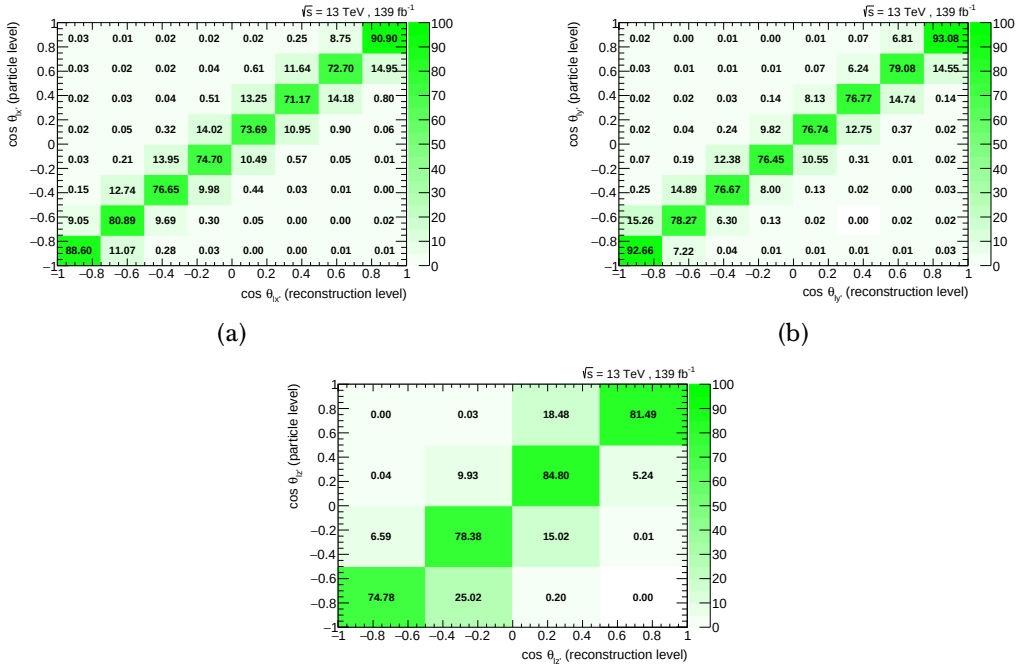
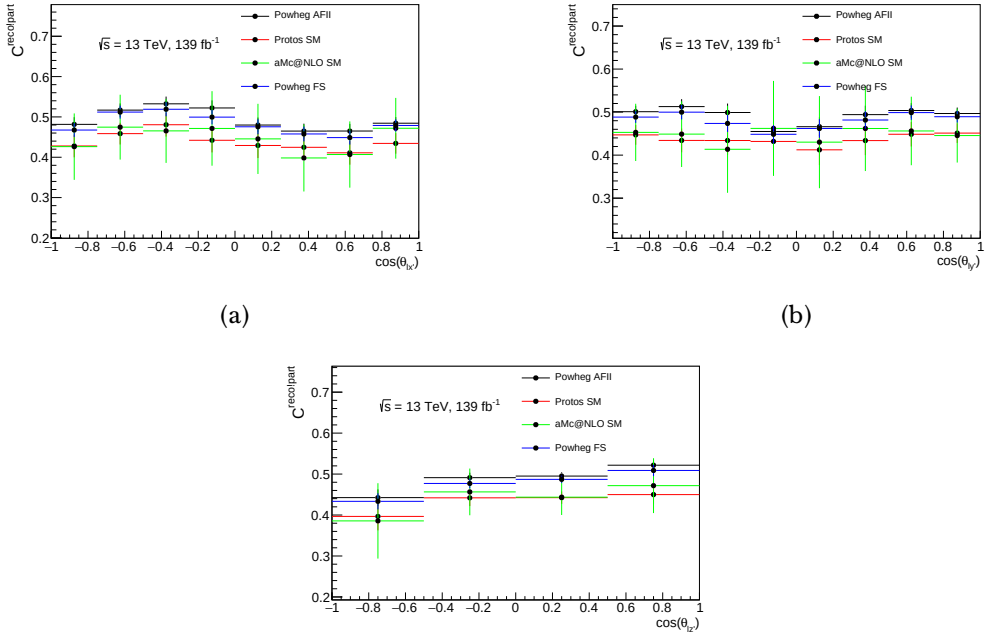


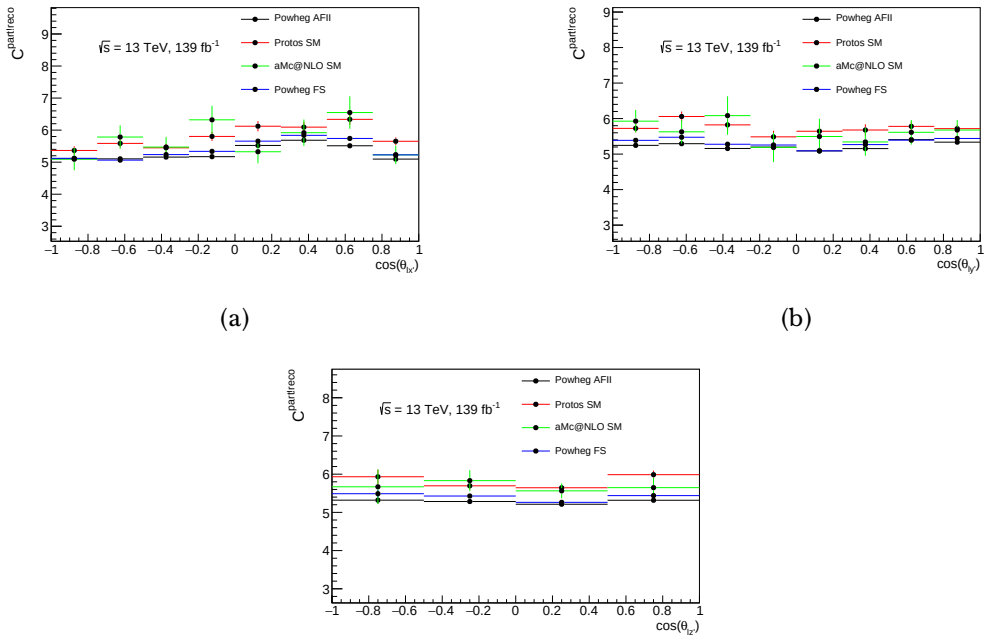
Figure 4.17: Migration matrices for (a) $\cos\theta_{lx'}$, (b) $\cos\theta_{ly'}$ and (c) $\cos\theta_{lz'}$. The angular variable at particle level is shown on the y'-axis while the reconstructed angular variable is shown on the x'-axis. The numbers are normalised per row and given in percentage. They have been computed from the baseline POWHEG-BOX+PYTHIA8 FS sample.



(a) (b)

(c)

Figure 4.18: Correction factors accounting for signal events that pass the reconstruction level selection but not the particle level selection for each bin of (a) $\cos\theta_{lx'}$, (b) $\cos\theta_{ly'}$ and (c) $\cos\theta_{lz'}$. They are given for the baseline Powheg-Box+Pythia8 FS sample and compared with the alternative samples Powheg-Box+Pythia8 AFII, Protos+Pythia8 and MG5_aMC@NLO+Pythia8.



(a) (b)

(c)

Figure 4.19: Correction factors accounting for signal events that pass the particle level selection but not the reconstruction level selection for each bin of (a) $\cos\theta_{lx'}$, (b) $\cos\theta_{ly'}$ and (c) $\cos\theta_{lz'}$. They are given for the baseline POWHEG-BOX+PYTHIA8 FS sample and compared with the alternative samples POWHEG-BOX+PYTHIA8 AFII, PROTOS+PYTHIA8 and MG5_aMC@NLO+PYTHIA8.

Angular distribution	Number of iterations
$\cos\theta_{lx'}$	5
$\cos\theta_{ly'}$	3
$\cos\theta_{lz'}$	3

Table 4.7: Number of bayesian iterations chosen for each measured angular distribution.

by different t-channel MC generators: the baseline POWHEG-BOX+PYTHIA8 FS sample, as well as POWHEG-BOX+PYTHIA8 simulated with AFII, PROTOS+PYTHIA8 and MG5_AMC@NLO+PYTHIA8.

The number of iterations is chosen such that the absolute change between two successive steps becomes negligible (the difference must be smaller than 0.1% with respect to the integral of the associated distribution) and when a convergent state is reached (the difference in the results obtained with the chosen number of iterations and with 15 more iterations should not exceed 0.15% with respect to the integral of the associated distribution) for all the considered samples. Figure 4.20 shows the evolution of the unfolded events to the particle level for the first bin of each angular distribution. Similar results are obtained for the other bins. As it can be seen, the convergence is very fast for the baseline POWHEG-BOX+PYTHIA8 FS sample due to the fact that the same sample is used to build the unfolded angular distributions and to compute the resolution and efficiencies corrections. However, for the other samples this is not the case and thus the convergence is slower. From these studies the number of iterations is fixed for each angular distribution. These numbers are displayed in Table 4.7.

4.7.3 Closure test

The closure test is performed to check the accuracy (i.e. the intrinsic bias) of the unfolding algorithm. For this purpose, the baseline POWHEG-BOX+PYTHIA8 FS sample is divided in two statistically independent subsets of the same size. The former subset is used to determine the migration matrix and the correction factors. This migration matrix and the correction factors are applied to unfold the angular

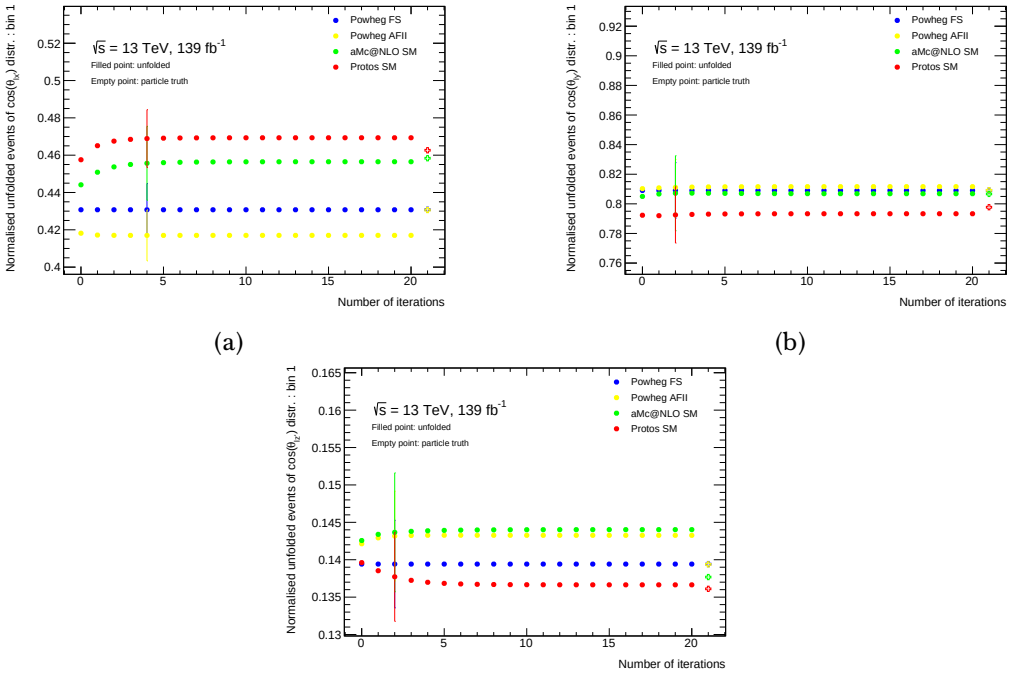


Figure 4.20: Unfolded values for the first bin of the (a) $\cos\theta_{lx'}$, (b) $\cos\theta_{ly'}$ and (c) $\cos\theta_{lz'}$ angular distributions as a function of the number of iterations using different t-channel generators: nominal POWHEG-BOX+PYTHIA8 FS (blue points), POWHEG-BOX+PYTHIA8 AFII (yellow points), PROTOS+PYTHIA8 (red points) and MG5_AMC@NLO+PYTHIA8 (green points). The unfolding is performed in all cases using the baseline POWHEG-BOX+PYTHIA8 FS unfolding matrix and efficiency corrections. The error bars represent the statistical uncertainties for the chosen number of iterations. The empty crosses correspond to the particle-level truth bin content for each sample.

PROTOS anomalous couplings samples
Re$[g_R]/V_L = \pm 0.18$
Im$[g_R]/V_L = \pm 0.07$
$V_R/V_L = \pm 0.4$; $g_L/V_L = \mp 0.32$

Table 4.8: Anomalous couplings PROTOS samples used in the linearity tests.

distributions from the second subset. Lastly, the unfolded angular distributions are compared to the corresponding predictions at particle level from the second subset.

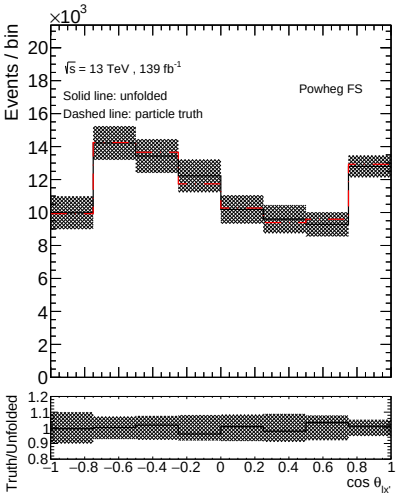
The results of the closure test are displayed in Figure 4.21 for the three angular distributions. The unfolded angular distributions (solid line) obtained from this procedure and the truth distributions (dashed line) are compatible within statistical uncertainties. Thus, the closure test is satisfied and no bias needs to be considered in the measurements of the angular distributions.

4.7.4 Linearity test

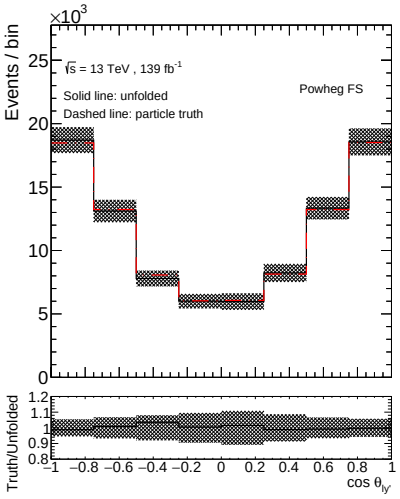
The linearity test is a fundamental part of the unfolding validation to look for possible deviations from the SM predictions. A successful linearity test guarantees that the unfolding is independent on any assumption on the underlying physics.

The procedure consists on comparing the unfolded distributions from different beyond the SM samples with their respective particle-level truth information. The unfolding is performed using the SM migration matrices and correction factors. If the unfolding does not depend on the physics, then the two compared distributions should be compatible within uncertainties.

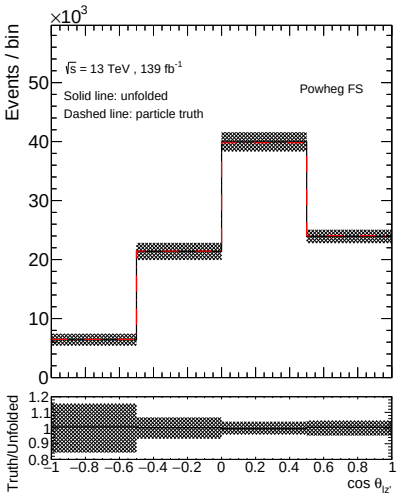
The PROTOS generator is used to compute the non-SM effects in terms of anomalous couplings. The values for the anomalous couplings were discussed in Section 4.2.2 and are summarized in Table 4.8. They are chosen to cover the possible variations with respect to the SM, given the current limits [129]. The migration matrix and correction factors are computed using the PROTOS SM sample so that the results are not affected by the difference between LO and NLO generation.



(a)



(b)



(c)

Figure 4.21: Relative differences between the unfolded (solid) and particle level truth (dashed) angular distributions for (a) $\cos\theta_{lx'}$, (b) $\cos\theta_{ly'}$ and (c) $\cos\theta_{lz'}$. The POWHEG-BOX+PYTHIA8 FS baseline sample is divided in two subsets of the same size. The former is used to compute the unfolding corrections and the latter is used to provide the unfolded distributions. The error bars represent the statistical uncertainties due to the limited size of the two subsets. The lower panels show the ratio between the truth and the unfolded distributions.

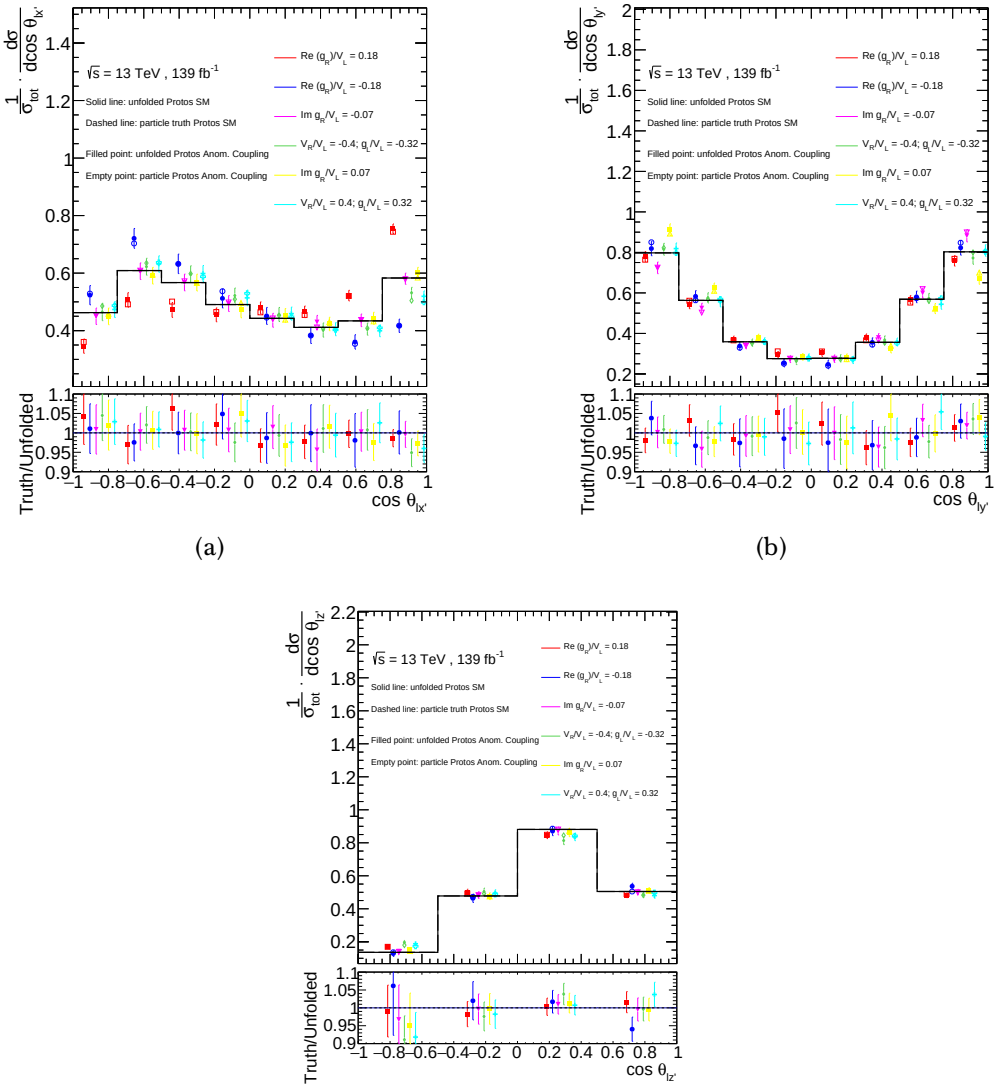
Figure 4.22 shows the results of the linearity test. The solid and dashed black lines represent the unfolded and the truth distributions from SM predictions, respectively. As it can be seen, they overlap. The filled dots represent the unfolded distributions from the samples with anomalous couplings. The empty points represent the truth information from this samples. In the ratio it can be seen the compatibility between the unfolded and truth distributions for all the considered beyond the SM effects and for the three angular distributions. Similar results were obtained using template distributions given by the MG5_aMC@NLO samples including the associated EFT operators. From this, it can be inferred that the unfolding technique has a good linearity and thus it is independent of any tWb model assumption within current constraints.

4.8 Statistical and systematic uncertainties

The systematic uncertainties considered in this analysis can be classified in two different groups: experimental (detector modelling) uncertainties and theoretical (signal and background modelling) uncertainties. The impact of each systematic uncertainty is evaluated from varied templates. These templates are propagated in a correlated way to the signal and the control regions. The impact on the measured differential cross-sections of each systematic source are added in quadrature to obtain the final systematic uncertainty. The statistical uncertainty associated to the limited size of the simulated and the data samples is also taken into account.

4.8.1 Statistical uncertainties

The statistical uncertainty of the data unfolded result is determined by running over an ensemble of 100 000 pseudo-experiments, varying the content of each bin according to its expected statistical uncertainty through Poisson fluctuations. For each pseudo-experiment a new background normalisation is extracted using the procedure described in Section 4.4.2. After background subtraction, each pseudo-experiment is unfolded and normalised. The spread (RMS) of the result in each bin is taken as the measure of the statistical uncertainty.



(a) (b)

(c)

Figure 4.22: Normalised differential cross-sections as a function of (a) $\cos\theta_{lx'}$, (b) $\cos\theta_{ly'}$ and (c) $\cos\theta_{lz'}$ using different new physics samples. The dashed line represents the truth SM values while the solid line represents the unfolded values. The empty dots represent the truth values from different beyond the SM effects in terms of anomalous couplings, while the filled dots represent the unfolded values of the beyond the SM sample using the migration matrix and corrections factors from the SM sample. The uncertainty bands include the statistical error of the unfolded histograms. The lower panels show the ratio between the truth and the unfolded distributions for the SM and the beyond the SM samples considered.

The uncertainty associated to the finite size of the simulated samples is evaluated using 100 000 pseudo-experiments, varying for each process (excluding the data-driven multijet background) the total prediction in each bin according to its statistical uncertainty through Gaussian fluctuations. The same procedure as for the data statistical uncertainty is then used.

4.8.2 Systematic uncertainties

In the following, a description of the procedures to evaluate the systematic uncertainties is discussed. They are grouped into two main categories: experimental and theoretical modelling uncertainties.

4.8.2.1 Experimental uncertainties

Luminosity and pile-up: The uncertainty for the combined 2015–2018 integrated luminosity is 1.7% [246], obtained using the LUCID-2 detector [247] for the primary luminosity measurements. To account for the difference between the pile-up distributions in data and MC simulations, an uncertainty related to the scale factors used to adjust the MC pile-up to the data pile-up profile is applied.

Leptons: For electrons and muons, the reconstruction, identification, isolation, trigger performance differ between data and MC simulation. Scale factors are applied to simulated events to describe the measured efficiencies in data. These scale factors, estimated using the tag-and-probe method [282, 393], as well as the lepton momentum scale and resolution, are assessed using leptonic decays of Z bosons and J/ψ mesons. Corrections to the lepton momentum scale and resolution are also applied to data. The associated systematic uncertainties are then propagated to the distributions used in this analysis.

Jets: To determine the JES uncertainty, information from test-beam data, LHC collision data, and simulation was used, as described in Ref. [291]. The JES uncertainty is decomposed into a set of 29 uncorrelated components, with contributions from pile-up, jet flavour composition, single-particle response, and effects of jets not contained within the calorimeter. The jet-energy resolution (JER) predicted by the MC simulation is verified (and where needed corrected)

using in-situ techniques [291]. Its uncertainty is represented by eight components accounting for jet-p_{T} and η-dependent differences between simulation and data. The systematic uncertainty associated with the JVT is obtained by varying the scale factor used to correct the JVT efficiency in simulation up and down within its uncertainties [394].

***b*-tagging**: The b-tagging efficiency and mistagging rates are measured in data using the methods as described in Refs. [333, 334, 395], with the systematic uncertainties due to b-tagging efficiency and the mistagging rates calculated separately. The impact of the uncertainties on the b-tagging calibration is evaluated separately for b-jets, c-jets and light-flavour jets.

Missing transverse momentum: The systematic uncertainties related to the modelling of the $E_{\mathrm{T}}^{\mathrm{miss}}$ in the simulation are estimated by propagating the uncertainties in the energy and momentum scales of electrons, muons and jets, as well as the uncertainties in the resolution and scale of the soft term [396].

4.8.2.2 Theoretical modelling uncertainties

Systematic uncertainties associated with the signal and background MC modelling are estimated by comparing event samples from different generators and by varying parameters in the event generation.

Matrix element generator: To assess the uncertainty due to the choice of the matching scheme in the t-channel and single-top-quark background ME generation, the nominal MC generator Powheg-Box is compared to MG5_aMC@NLO, both interfaced to Pythia8.

Parton shower generator: The uncertainty in the PS is estimated for all top-quark processes by comparing Powheg-Box interfaced with Pythia8 and Herwig7.

***$t\bar{t}$* and *tW* interference**: An additional uncertainty arising from the method used to handle the interference between tW and $t\bar{t}$ productions is determined by comparing the tW simulated sample that uses the DS method [371] with the nominal one based on the DR technique.

ISR/FSR: The uncertainty due to missing higher-order QCD corrections in the ME computation is estimated for all top-quark processes by independently varying up and down the renormalisation and factorisation scales by factors of 0.5 and 2.0 with respect to the central value. Additionally, uncertainties due to initial-state radiation from the PS are assessed by varying the corresponding parameter of the A14 parton shower tune `Var3c` [349]. For the $t\bar{t}$ process and just in the case of the up variation, the h_{damp} parameter is also changed and set to $3m_t$ [397]. The uncertainties due to final-state radiation from the PS are assessed by varying up and down the renormalisation scale for QCD emission by factors of 0.5 and 2.0.

Parton distribution functions: PDF uncertainties are evaluated using the PDF4LHC15 uncertainty set which consists of 30 eigenvector variations from multiple NLO PDF sets [367].

Cross-sections: The event yields associated with the simulated signal and background processes are estimated using the selection acceptances and the theoretically predicted cross-sections as reported in Section 4.2.2. The uncertainties on these cross-sections are taken into account. However, for those processes whose normalisation is extracted from the data-driven fit the effect is negligible.

Multijet normalisation: The multijet background is normalised through a data-driven analysis based on the techniques described in 4.4.1. A relative systematic uncertainty of 20% (40%) is assigned to this data-driven overall normalisation in the signal and W+jets ($t\bar{t}$) dominated regions. It is estimated from fits, as described in Appendix A.1, using alternative kinematic variables and including modelling uncertainties in the main background processes.

Multijet shape: To evaluate the systematic uncertainty on the shape of the multijet templates, additional MC simulation and data-driven samples were produced varying the fraction of the energy of the jet deposited in the EM calorimeter (jet-electron model) and the ratio of the sum of the transverse momenta of tracks with maximum cone of size 0.4 to the p_{T} of the muon (anti-muon model). These alternative multijet templates are normalised to the nominal yields and compared to the nominal multijet templates. More details can be found in Appendix A.2.

***W*+jets shape**: To evaluate the impact of the systematic uncertainty on the shape of the W+jets simulated templates, an up and down variation of 30% is applied independently to the relative fraction of W+b-jets and W+c-jets in the W+jets distributions [398]. Note that no variation is considered for W+light-jets production since its contribution is negligible.

4.8.3 Treatment of uncertainties

The relative impact of each source of uncertainty is calculated with respect to the Asimov dataset[12]. This is evaluated by unfolding template distributions resulting from simulated pseudo-data modified to reflect the effect of a given systematic source. In each case, a new background normalisation estimation is performed before its subtraction from the simulated pseudo-data, using the fitting procedure described in Section 4.4.2. For all sources of systematic uncertainty the nominal unfolding corrections are considered. Except for the signal modelling uncertainties affecting the unfolding corrections, the systematic uncertainty is evaluated as the difference between the unfolded nominal angular distribution values and the ones measured using the varied normalisations and shapes. For the signal modelling uncertainties, the unfolded varied distribution is compared to the corresponding particle-level spectra. The absolute uncertainty is computed by re-scaling the relative uncertainties with the values from the observed data.

For most of the systematic sources, up and down variations are computed. The impact of these systematic effects is evaluated according to:

$$\text{uncertainty} = \frac{|\text{up}| + |\text{down}|}{2}$$

Then, the resulting uncertainty is symmetrised. In case of one-sided systematics, the absolute value of the considered uncertainty is taken, and it is symmetrised as well.

The procedure described above is valid for most of the uncertainties. However, some considerations need to be taken into account:

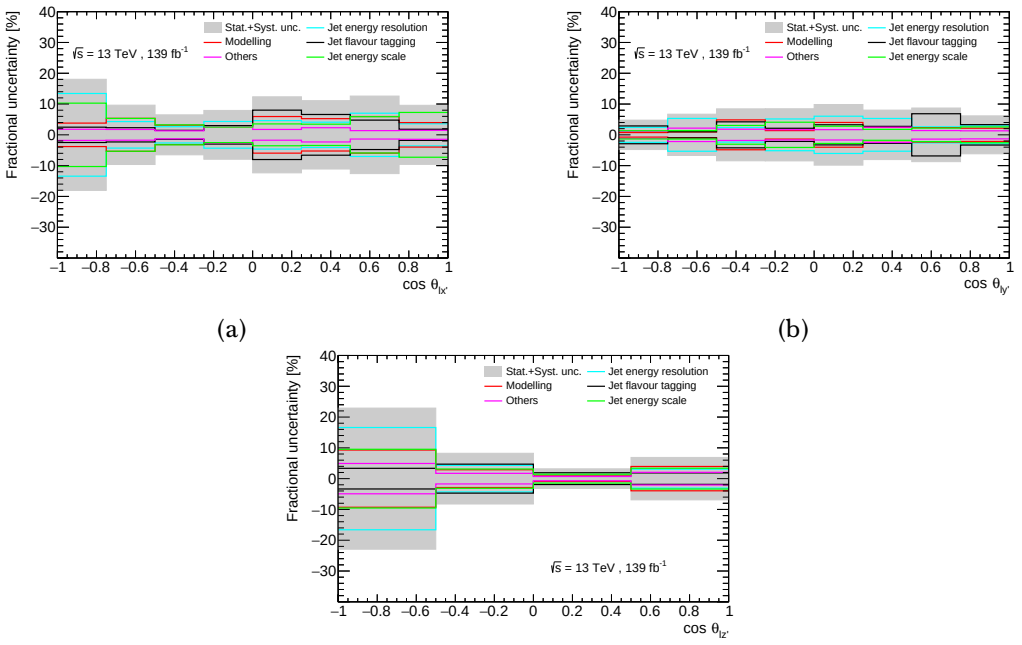
[12] This is the dataset in which all observed quantities are set equal to their expected values.

- For the different background modelling uncertainties, which include $t\bar{t}$, tW- and s-channel modelling (except tW DS samples), and for the $t\bar{t}$ radiation uncertainty a shift is computed from the relative difference between the varied distribution associated to a given systematic source and the reference distribution, both simulated using AFII. This shift is applied to the nominal spectra built from FS samples to compute the systematic template distributions.

- For the signal modelling uncertainties estimated from AFII samples, the shift previously described is applied both at reconstruction and at particle level. The unfolded template distribution for a given signal modelling uncertainty is then compared with the corresponding shifted particle-level distribution.

4.8.4 Uncertainties breakdown

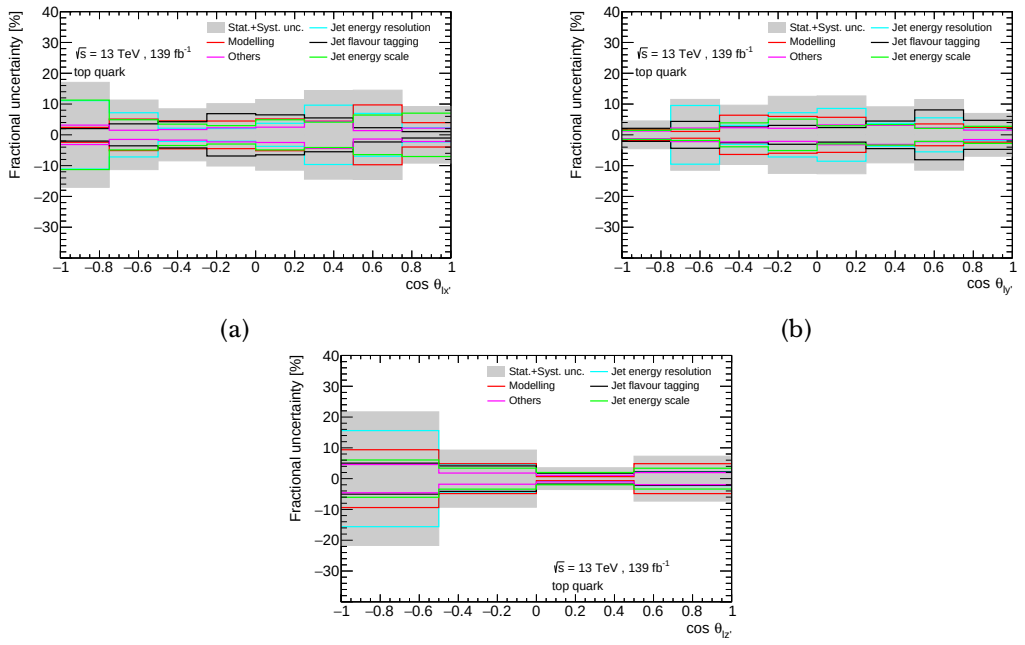
Figures 4.23-4.25 show the breakdown of the different systematic uncertainties in the three measured normalised differential cross-sections for the inclusive, the top-quark and the top-antiquark measurements. As it can be seen, the main contributions are associated with jets, in particular the JER. These are followed by the JES, the modelling and the b-tagging uncertainties. The sensitivity of these measurements to the JER uncertainties has been studied in detail. The reason is that the kinematic angles are determined in the top-quark reference frame, and since jet energy and $E_{\mathrm{T}}^{\mathrm{miss}}$ are used to reconstruct that frame, uncertainties in the measurement of the jet energy are expected to contribute significantly to the total systematic uncertainty.



(a) (b)

(c)

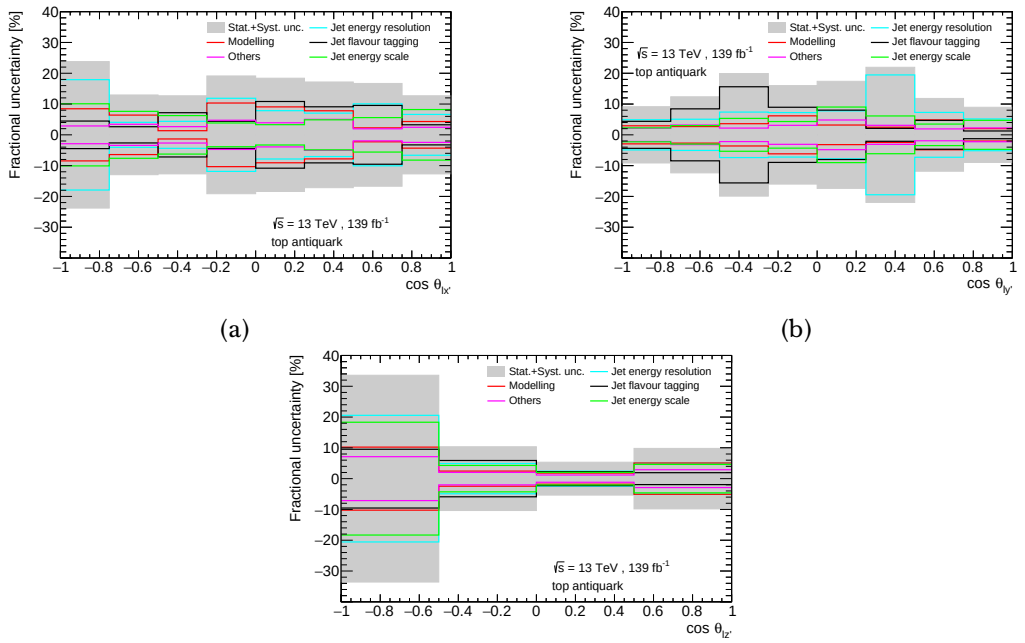
Figure 4.23: Breakdown of the different systematic contributions for the three angular distributions including both top quarks and antiquarks for (a) $\cos\theta_{lx'}$, (b) $\cos\theta_{ly'}$ and (c) $\cos\theta_{lz'}$. The cateogry 'Others' includes all the uncertainties which are not included in any other group.



(a) (b)

(c)

Figure 4.24: Breakdown of the different systematic contributions for the three angular distributions for top quarks: (a) $\cos\theta_{lx'}$, (b) $\cos\theta_{ly'}$ and (c) $\cos\theta_{lz'}$. The cateogry 'Others' includes all the uncertainties which are not included in any other group.



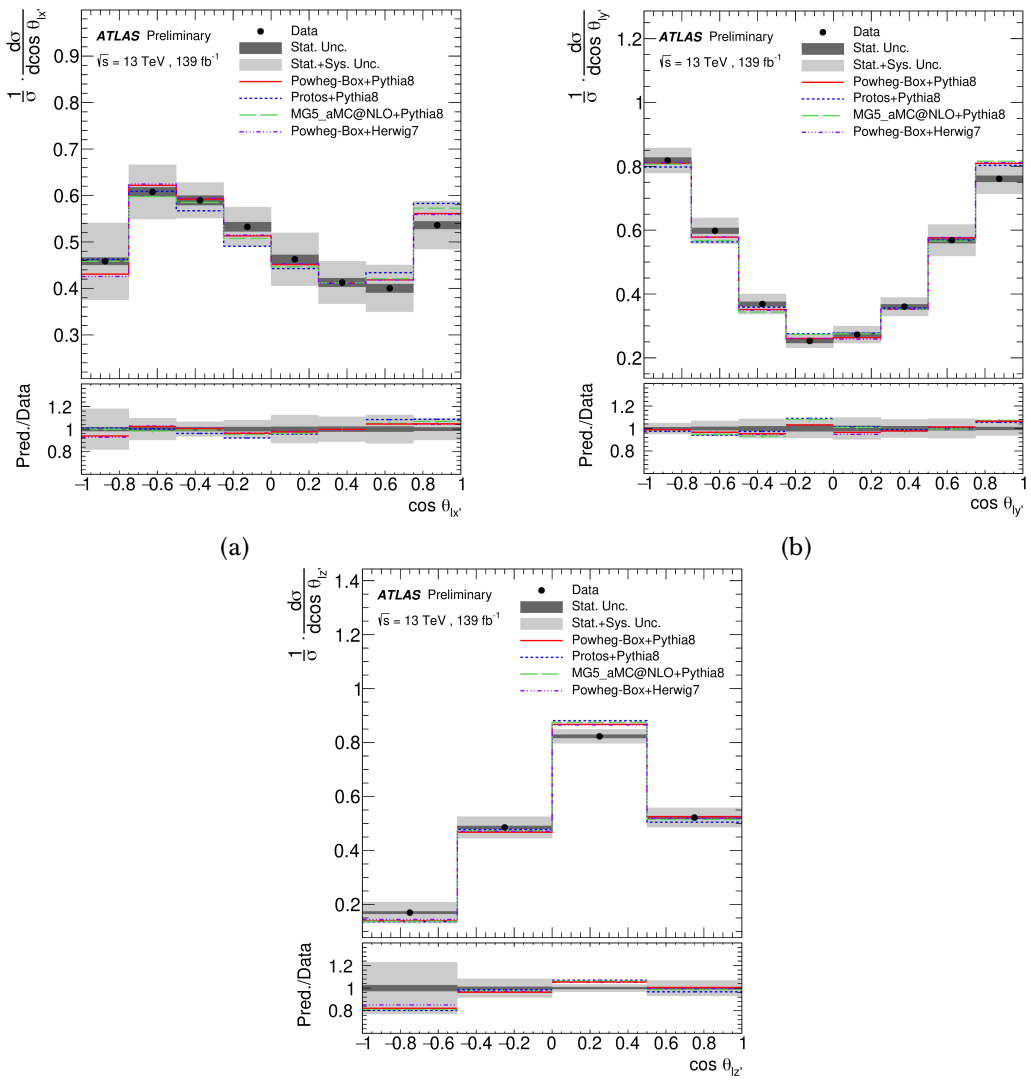
(a) (b)

(c)

Figure 4.25: Breakdown of the different systematic contributions for the three angular distributions for top antiquarks: (a) $\cos\theta_{lx'}$, (b) $\cos\theta_{ly'}$ and (c) $\cos\theta_{lz'}$. The cateogry 'Others' includes all the uncertainties which are not included in any other group.

4.9 Results

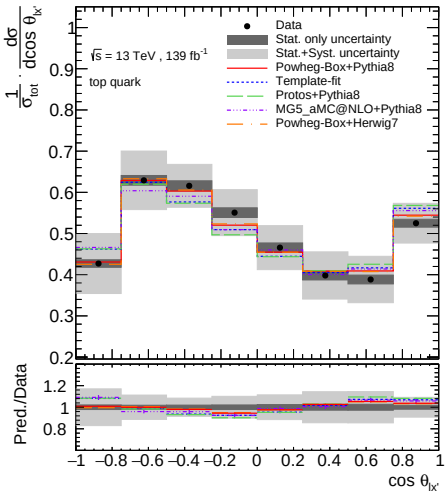
The resulting angular normalised differential cross-sections at particle level within the fiducial region are displayed in Figures 4.26-4.28 for the inclusive, top-quark and top-antiquark measurements, respectively. The measured cross-sections include both statistical and systematic uncertainties, and are compared to the predictions given by the different generators. The results are in agreement with the SM predictions. The top-quark and top-antiquark distributions are also compared to a direct measurement of the top-quark and antiquark polarisation vectors, as described in Section 4.9.2.



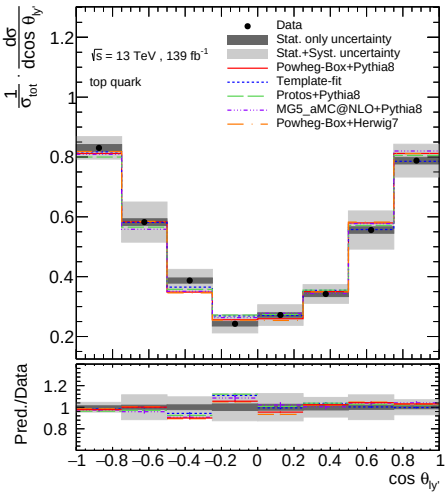
(a) (b)

(c)

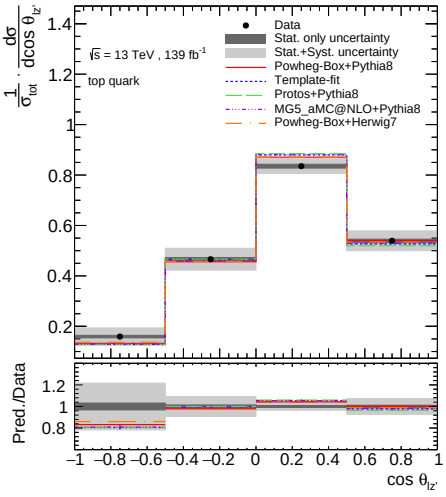
Figure 4.26: Normalised differential cross-sections as a function of (a) $\cos\theta_{lx'}$, (b) $\cos\theta_{ly'}$ and (c) $\cos\theta_{lz'}$, along with various MC predictions of the t-channel signal for both top quarks and top antiquarks. The data, shown as the black points with statistical uncertainties, are compared to predictions using the PowHeg-Box+Pythia8, Protos+Pythia8, MG5_aMC@NLO+Pythia8, and PowHeg-Box+Herwig7 setups. The uncertainty bands include both the statistical and systematic uncertainties. The lower panels show the ratio of data to prediction in each bin.



(a)

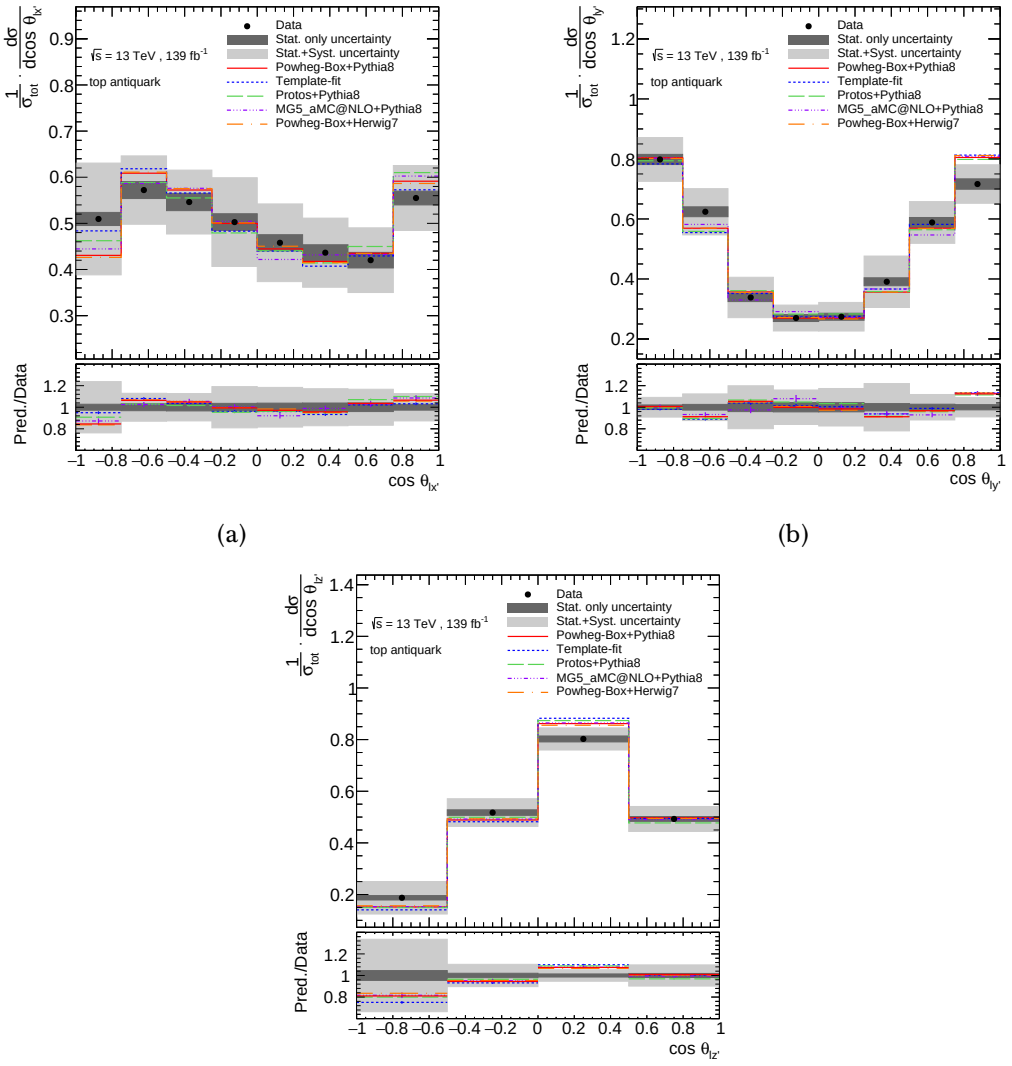


(b)



(c)

Figure 4.27: Normalised differential cross-sections as a function of (a) $\cos\theta_{lx'}$, (b) $\cos\theta_{ly'}$ and (c) $\cos\theta_{lz'}$, along with various MC predictions of the t-channel signal for top quarks. The data, shown as the black points with statistical uncertainties, are compared to predictions using the POWHEG-BOX+PYTHIA8, PROTOS+PYTHIA8, MG5_AMC@NLO+PYTHIA8, and POWHEG-BOX+HERWIG7 setups. The results are also compared with the direct measurement of the top-quark polarisation vector, labelled template-fit. The uncertainty bands include both the statistical and systematic uncertainties. The lower panels show the ratio of data to prediction in each bin.



(a) (b)

(c)

Figure 4.28: Normalised differential cross-sections as a function of (a) $\cos\theta_{lx'}$, (b) $\cos\theta_{ly'}$ and (c) $\cos\theta_{lz'}$, along with various MC predictions of the t-channel signal for top antiquarks. The data, shown as the black points with statistical uncertainties, are compared to predictions using the POWHEG-BOX+PYTHIA8, PROTOS+PYTHIA8, MG5_AMC@NLO+PYTHIA8, and POWHEG-BOX+HERWIG7 setups. The results are also compared with the direct measurement of the top-antiquark polarisation vector, labelled template-fit. The uncertainty bands include both the statistical and systematic uncertainties. The lower panels show the ratio of data to prediction in each bin.

4.9.1 Compatibility of the measurements with the SM predictions

A global covariance matrix of size 20×20 including the effects of all uncertainties and the correlations among the three angular distributions is computed in order to quantitatively compare with theoretical predictions, 20 being the total number of measured bins[13]. The computation of this covariance matrix is the result of three successive steps:

- Computation of the unnormalised covariance matrix which takes into account all systematic uncertainties and the effect from the limited size of the MC samples.
- Computation of the unnormalised covariance matrix which takes into account the statistical uncertainties.
- Addition of the two previous matrices to build the full unnormalised covariance matrix and normalisation.

A first ensemble of pseudo-experiments is used to compute the covariance matrix with the systematic uncertainties and the contribution coming from the limited size of the MC samples. In each pseudo-experiment, Gaussian-distributed shifts are coherently added for each detector-modelling uncertainty by scaling each bin of the data distribution with the expected relative variation from the associated systematic uncertainty effect. The relative variation in each bin j accounting for the i-th systematic type $\Delta_j^{\text{syst},i}$ is defined as the difference between the varied and the nominal pseudo-data events in each bin, normalised to the nominal yields:

$$\Delta_j^{\text{syst},i} \equiv \frac{B_j^{\text{syst},i} + S_j^{\text{syst},i} - B_j^{\text{nominal}} - S_j^{\text{nominal}}}{B_j^{\text{nominal}} + S_j^{\text{nominal}}}, \tag{4.18}$$

where B_j^{nominal} (S_j^{nominal}) is the nominal expected number of background (signal) events in bin j, and $B_j^{\text{syst,i}}$ ($S_j^{\text{syst,i}}$) is the expected number of background (signal) events in bin j for a systematic effect i.

[13] Eight bins are used for $\cos\theta_{lx'}$ and $\cos\theta_{ly'}$, whereas four bins are used for $\cos\theta_{lz'}$.

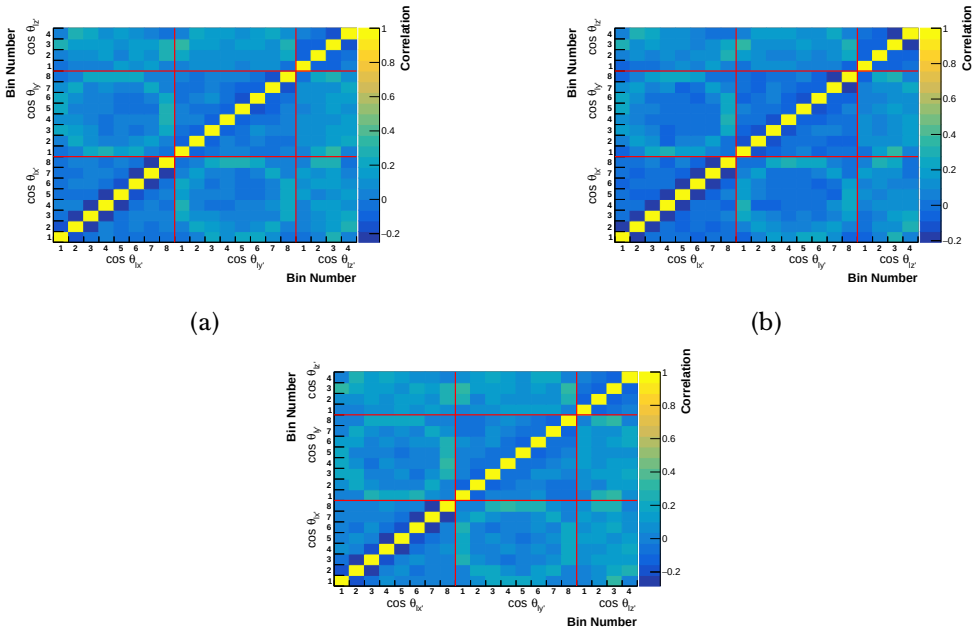
For most systematic sources, up and down variations are provided. In each pseudo-experiment, a random number $\lambda = \mathcal{G}(0,1)$ is drawn for each systematic source, common to all bins. When positive (negative), the up (down) variation is taken. In the case of a one-sided systematic uncertainty, the template distribution is symmetrised and the procedure is applied in the same way. The k-th pseudo-experiment in each bin j is thus constructed as:

$$\text{toy}_j^k \equiv N_j^{\text{obs}} \cdot [1 + \sum_{\text{syst},i} |\lambda_{\text{syst},i}| \Delta_j^{\text{syst},i}] \tag{4.19}$$

An additional shift is added to account for the limited size of the MC samples. In this case, the varied pseudo-data is constructed following the procedure described in Section 4.8.1. The shift is then computed in a similar way to each detector-modelling systematic effect. The only difference is that there is a random number for each bin $\lambda_{\text{syst},i,j}$ rather than for all bins. This term contributes to the systematic shift and is added in the summation defined in Equation 4.19 like any other systematic effect.

The varied distribution is unfolded with the nominal corrections after subtracting the backgrounds with the scale factors derived following the procedure described in 4.4.2. Additional Gaussian-distributed shifts are then coherently added for each signal- and background-modelling systematic uncertainty. The modelling shifts are derived by using the expected relative variations from the associated systematic uncertainty to scale each bin of the data distribution unfolded with nominal corrections. The resulting modified unfolded distributions are used to compute a first matrix.

The determination of statistical correlations between the three different angular distributions are evaluated using the bootstrap method [399], using a set of 1000 bootstrap samples. Each sample is built from events which are fluctuated using a different random seed based on unique identifiers for each event. Thus, the random seed is coherently applied to the three measured quantities. The global statistical matrix is calculated using the statistical error of each bin of the angular distributions and the correlation coefficients, depicted schematically in Figure 4.29. This matrix is then added to the previous one to compute an unnormalised global covariance matrix.



(a)

(b)

(c)

Figure 4.29: Bootstrap correlation matrices for the three unfolded angular measurements for (a) both top quarks and antiquarks, (b) top quarks and (c) top antiquarks.

Finally, a set of 50 000 additional pseudo-experiments is used to normalise the covariance matrix. For each additional pseudo-experiment, each element of the three angular distributions is fluctuated using a multivariate gaussian distribution with the unnormalised global covariance matrix. The resulting pseudo-experiments are normalised and are used to calculate the normalised global covariance matrix. These are reported in Appendix B.

To quantify the level of agreement between each of the measured normalised differential cross-sections and the theoretical predictions, χ^2 values are calculated according to the relation $\chi^2 = V^T_{\mathrm{N}_b-1} \cdot \mathrm{Cov}^{-1}_{\mathrm{N}_b-1} \cdot V_{\mathrm{N}_b-1}$. Here N_b is the number of bins of the spectrum under consideration, V_{N_b-1} is a vector with the differences between the measured and the predicted cross-sections obtained by discarding one of the N_b elements (since these are normalised cross-sections), and $\mathrm{Cov}_{\mathrm{N}_b-1}$ is the $(\mathrm{N}_b-1)\times(\mathrm{N}_b-1)$ sub-matrix derived from the relevant $\mathrm{N}_b \times \mathrm{N}_b$ part of the global covariance matrix by discarding the corresponding row and column. It has been checked that the result is independent of the discarded element. A global χ^2 value is also obtained using the full vector of differences for the three angular distributions and the full global covariance matrix. In this case three elements of the vector and their corresponding rows and columns in the covariance matrix are discarded.

The p-value is a test statistic which quantifies the probability, under assumption of the null hypothesis, of finding data of equal or greater incompatibility with the predictions from the null hypothesis. In this context, the null hypothesis corresponds to the SM predictions. This implies that the smaller the p-value, the highest the incompatibility of the results with the SM predictions.

The p-values are evaluated from the χ^2 values and the number of degrees of freedom (NDF), as shown in Table 4.9. Since these are normalised cross-sections, the NDF of each angular distribution corresponds to the number of bins of each angular distribution minus one. In the case of the global χ^2 value associated with the three angular distributions, the NDF corresponds to the total number of bins minus three. The obtained values show a good agreement of the SM prediction with the measured data.

	top quark		top antiquark		top quark and antiquark	
Angular variable	χ^2/NDF	p-value	χ^2/NDF	p-value	χ^2/NDF	p-value
$\cos\theta_{lx'}$	1.35/7	0.99	0.94/7	1.00	1.32/7	0.99
$\cos\theta_{ly'}$	4.57/7	0.71	2.92/7	0.89	3.78/7	0.81
$\cos\theta_{lz'}$	1.55/3	0.67	2.04/3	0.56	2.26/3	0.52
Global	13.55/17	0.70	6.86/17	0.99	9.25/17	0.93

Table 4.9: The χ^2 and p-value of the three unfolded angular distributions for the top-quark, for the top-antiquark and for both the top-quark and top-antiquark measurements. The numbers are computed comparing the observed data with the POWHEG-BOX+PYTHIA8 SM predictions. The NDF corresponds to the number of bins of each angular distribution minus one. A global χ^2 and p-value for the three angular distributions is also included. In this case, the NDF corresponds to the total number of bins minus three.

4.9.2 Comparison with the direct measurement of the top-quark and antiquark polarisation vectors

The results obtained for top quarks and antiquarks separately can be compared with the direct measurement of the top-quark and antiquark polarisation vectors [3], reported in Table 4.10.

Parameter	Extracted value	(stat.)
$P^t_{x'}$	$+0.01 \pm 0.18$	(± 0.02)
$P^t_{x'}$	-0.02 ± 0.20	(± 0.03)
$P^t_{y'}$	-0.029 ± 0.027	(± 0.011)
$P^t_{y'}$	-0.007 ± 0.051	(± 0.017)
$P^t_{z'}$	$+0.91 \pm 0.10$	(± 0.02)
$P^t_{z'}$	-0.79 ± 0.16	(± 0.03)

Table 4.10: Polarisation values as extracted from a template-fit to the data, including total and statistical-only uncertainties [3].

This comparison can be performed using fully polarised PROTOS samples, interfaced with PYTHIA8. Each sample represents a fully polarised state $P_i = \pm 1$, being $i = x', y', z'$. These are combined taking into account the measured

polarisation components. The sign of the measured polarisation component is used to select either the positive or the negative state sample, and each selected sample is added coherently using the absolute value of the associated polarisation component as a weight:

$$\text{template} = \sum_i |P_i| \times \begin{cases} \text{template}_{i,+} & \text{if} \quad P_i > 0 \\ \text{template}_{i,-} & \text{if} \quad P_i < 0 \end{cases},$$

where $\text{template}_{i,+}$ ($\text{template}_{i,-}$) corresponds to the PROTOS sample associated with the fully polarised positive (negative) state in the i direction.

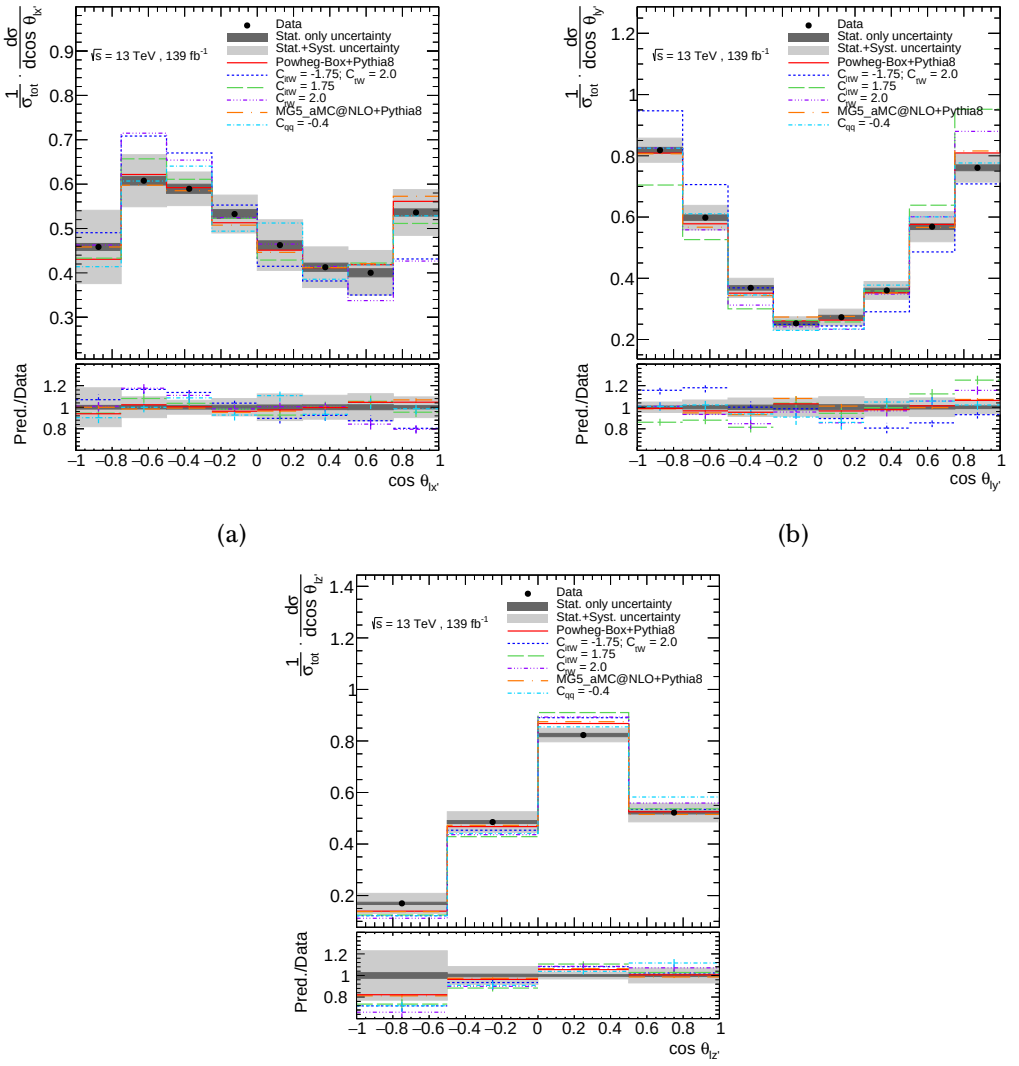
The resulting templates are provided at particle level, which allows a straight comparison. This is a conservative comparison, since just the uncertainties in the unfolded measurements are taken into account. As it can be seen in Figures 4.27-4.28, the results are compatible for the three angular distributions for both top quarks and antiquarks.

4.9.3 Limits on Wilson coefficients

The global covariance matrix introduced in Section 4.9.1, together with the three differential cross-section measurements, have been used to set constraints on the real and imaginary parts of the dipole operator (C_{tW} and C_{itW}, respectively) as part of the PhD thesis in Ref. [4]. One can see in Figure 4.30 that $\cos\theta_{lx'}$ is particularly sensitive to the real part of this operator, whereas $\cos\theta_{ly'}$ is sensitive to the imaginary part. The measured values of C_{tW} = 0.4 $\pm$ 1.1 and C_{itW} = -0.3 $\pm$ 0.4 at 95% C.L. are compatible with the SM predictions.

The best constraints on C_{tW} come from measurements on the values of the W-boson helicity fractions in top-quark pair decays from ATLAS [400] and CMS [401], with the combined result [402] providing bounds of $[-0.48, 0.29]$ at 95% C.L. when only C_{tW} is allowed to vary. Comparable limits are also derived from EFT fits of the top-quark sector [118, 403–405] that include those measurements. However, the constraints obtained in those analyses assume $C_{itW} = 0$, which is not assumed in the result presented here.

Very stringent individual limits on C_{itW} exist from electric dipole moments analyses [406] when only C_{itW} is allowed to be non-zero, but these become much weaker when multiple EFT coefficients are allowed to vary simultaneously. In that case, the limits are dominated by the input from ATLAS [127], which provides bounds on C_{itW} of $[-2.3, 3.0]$ at 95% C.L.. A global fit including real and imaginary parts of all tWb operators [407] gives $C_{tW} \in [-0.8, 0.7]$ and $C_{itW} \in [-2.3, 1.6]$, where the bound on C_{itW} is dominated by the ATLAS result [131]. A later ATLAS result [129] improved the limit to $C_{itW} \in [-0.8, 0.7]$. The bounds presented in this work on C_{itW}, which is sensitive to CP violation components, improve on all of these results.



(a) (b)

(c)

Figure 4.30: Normalised differential cross-sections as a function of (a) $\cos\theta_{lx'}$, (b) $\cos\theta_{ly'}$ and (c) $\cos\theta_{lz'}$, for both top quarks and top antiquarks. The data, shown as the black points with statistical uncertainties, are compared to predictions using different MG5_aMC@NLO+Pythia8 samples, with different values for the real and the imaginary part of the dipole operator. A non-SM value for the four-fermion operator is also considered. The uncertainty bands include both the statistical and systematic uncertainties. The lower panels show the ratio of data to prediction in each bin.

5 Search for dark matter production in association with a single top quark

Among the different limitations of the SM (see Section 1.1.5), understanding the nature of dark matter, the dominant matter component of the universe, is of paramount importance. This kind of matter is expected to interact mainly through the gravitational force, and would couple very weakly to the SM particles. There is very strong evidence for the existence of this kind of non-ordinary matter, but still very little is known about it. The main reason comes from the fact that all DM observations are astrophysical, but it has not yet been observed in the laboratory. WIMPs are natural DM candidates which arise in many theories beyond the SM, and constitute one of the most compelling scenarios for particle DM. One of the main goals at the LHC is to produce WIMP DM particles and to study their properties, which would be a major success in our comprehension of the universe. Among the strategies to search for DM at the LHC one can find signature-based searches, focusing on the final state particles and thus allowing a wide range of interpretability.

This chapter presents a search for single-top-quark production in association with DM particles, i.e. a large amount of missing transverse momentum. This kind of final state is known as mono-top. The two DM production mechanisms that would lead to the mono-top signature are identified as the resonant and the non-resonant DM production (see Section 1.4.4). Both of them are considered in this analysis.

Previous mono-top searches have been carried out by the ATLAS and CMS collaborations, both for Run-1 at 8 TeV [238, 239] and Run-2 at 13 TeV [241–243], and by the CDF collaboration with the 1.96 TeV data collected at Tevatron [240]. The analysis strategy depends on the decay products of the top quark. The work on this thesis focuses on the hadronic channel, since it was proven to be more sensitive to this kind of signature [241].

The analysed data corresponds to pp collisions recorded by the ATLAS detector at the LHC during Run-2 (2015-2018), with a total integrated luminosity of 139 fb^{-1}.

This chapter is organised as follows. Firstly, the DM signal and the SM backgrounds are described in Section 5.1. Section 5.2 introduces the different data and Monte Carlo samples used in the analysis. Then, a summary of the analysis strategy is given in Section 5.3. The event selection in the signal, control and validation regions is discussed in Section 5.4, and the associated event yields and distributions are shown in Section 5.5. Next, the different systematic uncertainties affecting the measurement are described in Section 5.6. Finally, the results of this analysis are reported in Section 5.7.

5.1 Dark matter mono-top signal and background contamination

As mentioned in Section 1.4.4, there are two DM models which can lead to the mono-top signature. Their associated Feynman diagrams were depicted in Figure 1.19. The resonant model is characterised by the interaction of a d-quark and an s-quark from the colliding protons, producing a new scalar particle ϕ, which then decays to an invisible DM particle χ and a top quark. The non-resonant model is characterised by a FCNC between an up-type quark, a top quark and a vector mediator particle V. This vector mediator particle is produced on-shell and then decays to a pair of DM particles $\chi\bar{\chi}$. This analysis focuses on the hadronic decay of the top quark, in which it decays to a W boson and a b-quark and the W boson further decays to a pair of quarks.

The produced top quarks are expected to have a very large Lorentz boost, and thus their decay products are very collimated. The signal final state is then characterised by a top quark, reconstructed as a single large-R jet, and a large amount of missing transverse momentum (DM particles).

There are several SM processes which can mimic this kind of signature. The main background contamination comes from top-quark processes (since these processes have top quarks in the final state), Z+jets processes (especially when the Z boson decays to a couple of neutrinos, thus producing a big amount of missing transverse momentum) and W+jets processes (especially when the W-boson decays to a τ lepton and a neutrino). Other minor backgrounds include diboson processes and the production of a vector boson or a Higgs boson in association with top quarks. The multijet background is expected to have a negligible impact in this analysis, as it will be explained in Section 5.4.4.2. A more detailed description of these background processes can be found in Section 4.1.

5.2 Data and simulated event samples

This section describes the dataset as well as the simulated samples used in this analysis.

5.2.1 Data event samples

The data used in this analysis corresponds to the full LHC Run-2 dataset collected by the ATLAS detector at $\sqrt{s} = 13$ TeV between 2015 and 2018. Only the recorded events in data-taking runs passing high quality criteria are considered [346]. This corresponds to an integrated luminosity of 139 fb^{-1} [246, 247].

The data events are required to pass a set of trigger conditions, selecting events with large missing transverse momentum as described in Section 2.7.1.

5.2.2 Simulated samples

In this analysis, simulated events are used to evaluate background contamination and modelling, signal acceptance, optimization of the event selection, and evaluation of systematic uncertainties. Several MC ME generators are used and interfaced to various parton shower/hadronisation generators. For most of the samples, the full ATLAS detector simulation [383] is performed using GEANT4 [384]. For some modelling uncertainties, AFII is used instead [383].

All the signal or background processes involving top quarks were simulated assuming a top-quark mass of 172.5 GeV, and assuming the top quark decays exclusively to a W boson and a b quark. The decays of bottom and charm hadrons were simulated using the EvtGen program [350]. The top quarks, the W and the Z bosons are decayed using MadSpin [347, 348].

5.2.2.1 Signal samples

Signal events for both the resonant and non-resonant DM scenarios were generated according to the simplified model in Refs. [223–225], described in Section 1.4.4 and implemented in MG5_aMC@NLO [408] (v2.8.1). The generation is done at LO using the NNPDF3.0NNLO PDF set. Parton showering, hadronisation and underlying event modelling were handled using Pythia 8.244 event generator with the A14 set of tuned parameters [349] and using the NNPDF2.3LO PDF set [352].

The generation of these samples for the resonant and non-resonant models depends on different sets of parameters: the resonant model depends on the masses of the scalar mediator m_ϕ and the DM particle m_χ, the coupling constant λ between the scalar mediator and the initial state d- and s-quarks, and the mixing parameter y between the scalar mediator, the top quark and the DM particle. The non-resonant model depends on the masses of the vector mediator m_V and the DM particle m_χ, the coupling a between the vector mediator, the initial state up-type quark and the top quark, and the coupling g_χ between the vector mediator and the DM particles.

For the resonant case, the shape of the kinematic distributions depends slightly on λ, whereas a strong dependence is observed as a function of the scalar mediator

mass m_ϕ. Different values of the mixing parameter y only affect the cross-section, whilst different values of m_χ do not have a sizeable effect either in the cross-section or the shape of the distributions. Consequently, samples with different values of m_ϕ were generated, fixing the values of λ, y and m_χ to 0.6, 0.4 and 1 GeV, respectively.

For the non-resonant case, the shape of the kinematic distributions is barely affected by different values of the parameters m_V, m_χ, a and g_χ. Modifications on these parameters affect mostly the cross-section of the process. Therefore, all samples were generated using coupling constants $a = 0.5$ and $g_\chi = 1$, and $m_\chi = 1$ GeV for the mass of the DM particle.

Table 5.1 summarises the parameters used in the generation of the different signal samples for both cases. The samples were normalised to the theoretical LO cross-section, computed with MG5_aMC@NLO. For the non-resonant case, the remaining kinematic dependence on the mass of the vector mediator particle was accounted for by means of a reweighting procedure based on the generator level $E_{\mathrm{T}}^{\mathrm{miss}}$. The weights were extracted in different bins of the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution as the ratio between the bin content of the additional sample with a different vector mediator mass and of the nominal $m_V = 1.75$ TeV sample, as listed in Table 5.1. Those weights were then applied on an event-by-event basis according to the $E_{\mathrm{T}}^{\mathrm{miss}}$ value of the processed event from the nominal $m_V = 1.75$ TeV sample.

5.2.2.2 Background samples

Table 5.2 shows a summary of the configuration of the nominal background samples used in the analysis. A more detailed description of all the SM background samples can be found in Section 4.2.2.

5.3 Analysis strategy

The DM mono-top search is conducted in two signal regions which are defined to maximise the sensitivity to the two considered DM signals: the resonant and the non-resonant models. The observation of a DM signature is difficult due to the low cross-section of the signal process, which makes its discrimination and

DM mono-top resonant model			
m_ϕ [TeV]	m_χ [GeV]	λ	y
2.5	1	0.6	0.4
3	1	0.6	0.4
4	1	0.6	0.4
5	1	0.6	0.4
6	1	0.6	0.4
DM mono-top non-resonant model			
m_V [TeV]	m_χ [GeV]	a	g_χ
1.75	1	0.5	1
DM mono-top non-resonant model (reweighted samples)			
m_V [TeV]	m_χ [GeV]	a	g_χ
1.25	1	0.5	1
1.5	1	0.5	1
2	1	0.5	1
2.5	1	0.5	1
3.0	1	0.5	1
3.5	1	0.5	1

Table 5.1: Different configurations employed in the simulation of the DM mono-top signal samples for both the resonant and the non-resonant production.

Process	ME event generator	ME PDF	PS and hadronisation	UE tune	Cross-section calculation
$t\bar{t}$	Powheg-Box	NNPDF3.0 NLO	Pythia 8	A14	NNLO+NNLL [101–106]
Single top	Powheg-Box	NNPDF3.0 NLO	Pythia 8	A14	NNLO+NNLL [109, 409, 410]
V+jets ($V = W/Z$)	Sherpa 2.2.1	NNPDF3.0 NNLO	Sherpa	Default	NNLO [378]
Diboson	Sherpa 2.2.1	NNPDF3.0 NNLO	Sherpa	Default	NLO (Sherpa)
$t\bar{t}+V$ ($V = W/Z$); tZq	MG5_aMC@NLO 2.3.3	NNPDF3.0 NLO	Pythia 8	A14	NLO [364]

Table 5.2: Overview of the nominal simulated background samples. Single-top production includes the t-channel, the tW-channel and the s-channel.

observation particularly challenging. In order to enhance the signal-to-background separation, this analysis exploits cutting-edge multivariate analysis techniques (MVA) techniques.

Dedicated control regions are designed to constrain the normalisation of the main contributing background processes in the signal regions: $t\bar{t}$ and V+jets ($t\bar{t}$ CR and V+jets CR). In addition to these control regions, validation regions are also defined to validate the normalisation results and to check the modelling of the backgrounds in regions kinematically closer to the signal region ($t\bar{t}$ VR and V+jets VR). These control and validations regions are defined to be signal-depleted (with a signal contamination lower than 5%) and orthogonal to each other and to the signal region.

The expected SM backgrounds are first determined with a profile likelihood fit in a background-only fit, as explained in Section 5.7.2. Here, the normalisation factors of the $t\bar{t}$ and V+jets backgrounds are extracted by adjusting them simultaneously to match the data yields in the associated control regions. The systematic uncertainties, described in Section 5.6, are included in the fit as nuisance parameters (NPs). Then, a simultaneous fit to the data yields in the control regions and to the MVA output discriminant distribution in the associated signal region of each DM signal model is performed under the background-only hypothesis to test the compatibility of the observed data with the SM predictions. In the absence of a significant event excess in the signal regions, the results of the fit are interpreted to set upper limits on the production cross-section for the considered signal models under the signal-plus-background hypothesis.

5.4 Event selection and background estimation

In this analysis the events are firstly selected if they fulfil a basic set of requirements based on the topology of the DM mono-top signal events. This defines the so-called *pre-selection region*. The requirements are applied on top of the physics objects described in Section 2.7.2.

The signal, control and validation regions are defined on top of this pre-selection region. Table 5.3 provides a summary of the selection requirements applied in the different analysis regions. More details are given in the following.

	Pre-selection	$t\bar{t}$ CR	$t\bar{t}$ VR	V+jets CR	V+jets VR	Signal region
N_{leptons}	0	0	0	0	0	0
Loose jets	0	0	0	0	0	0
$E_{\text{T}}^{\text{miss}}$	> 250 GeV	> 250 GeV	> 250 GeV	> 250 GeV	> 250 GeV	> 250 GeV
$N_{\text{large-}R\text{ jets}}$	≥ 1	≥ 1	≥ 1	≥ 1	≥ 1	≥ 1
N_{top}(cont. 50%)	≥ 1	≥ 1	≥ 1	≥ 1	≥ 1	≥ 1
$N_{b\text{-jets}}$	-	≥ 2	$= 1$	$= 0$	≤ 1	≤ 1
$\Delta\phi_{\text{min}}$	> 0.2	$\in [0.2, 1]$	$\in [0.2, 1]$	$\in [0.2, 1]$	≥ 1	≥ 1
XGBoost score	-	-	-	-	< 0.5	> 0.5

Table 5.3: Selection requirements defining the regions used in this analysis. N_{leptons}, $N_{\text{large-}R\text{ jets}}$, N_{top} and $N_{b\text{-jets}}$ indicate the number of leptons, large-R jets, top-tagged large-R jets and b-jets, respectively. $\Delta\phi_{\text{min}}$ is the minimum angular distance in the transverse plane between the $E_{\text{T}}^{\text{miss}}$ and any small-R jet. XGBoost score represents the score of the MVA discriminant, which is used to define the signal region and the V+jets validation region. Additional details are given in the text. The symbol - indicates that the region is inclusive in that variable.

5.4.1 Pre-selection level

This analysis focuses on the search for DM mono-top events in the hadronic channel. The topology of the signal events in this channel is featured by a large missing transverse momentum due to the final state DM particles. In addition, the produced top quark is expected to be boosted, and its decay products can be collimated into a large-R jet ($R = 1.0$). Therefore, events at pre-selection level are required to satisfy the following criteria:

- The event contains exactly zero leptons.
- The event must fulfil the jet-cleaning criteria with exactly zero *loose* bad jets[1].
- The $E_{\mathrm{T}}^{\mathrm{miss}}$ is required to be larger than 250 GeV. This requirement reduces the multijet background drastically.
- At least one top-tagged large-R jet (J) is required in the final state. The following requirements are applied to top-tagged large-R jets, according to the kinematical range in which the classifier algorithm is calibrated [322, 323]:
  - The p_{T} of the top-tagged large-R jet lies within [350, 2500] GeV.
  - The reconstructed transverse mass of the top-tagged large-R jet lies within [40, 600] GeV.
- The minimum angular distance in the transverse plane between the $E_{\mathrm{T}}^{\mathrm{miss}}$ and any small-R jet $\Delta\phi_{\mathrm{min}}$ is required to be larger than 0.2. This reduces drastically the beam-induced background (BIB)[2] [411, 412].

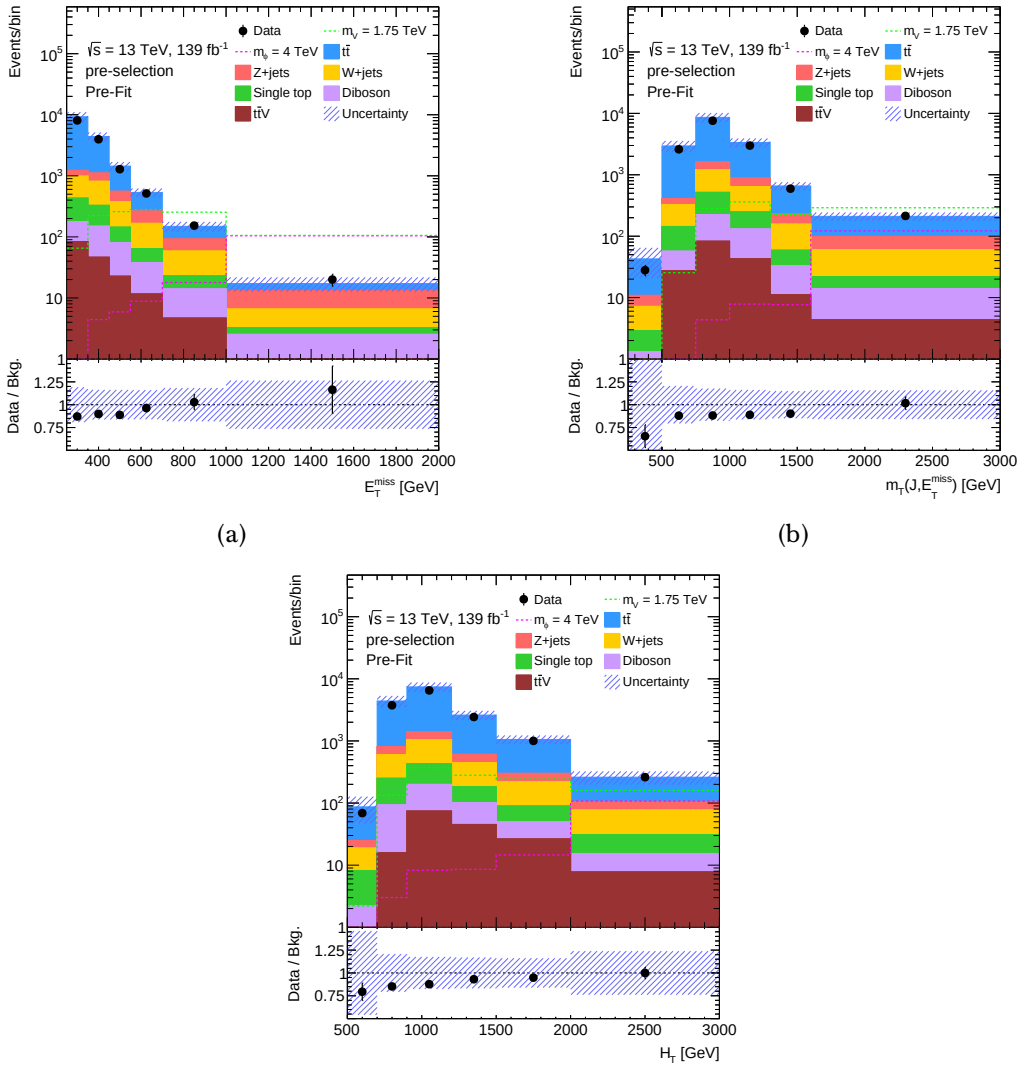
Figure 5.1 shows the distribution of some basic kinematic variables, namely the $E_{\mathrm{T}}^{\mathrm{miss}}$, the transverse mass of the top-tagged large-R jet and the $E_{\mathrm{T}}^{\mathrm{miss}}$ $m_{\mathrm{T}}(E_{\mathrm{T}}^{\mathrm{miss}}, J)$[3] and the sum of the transverse momentum of all the small-R jets in the event and the missing transverse momentum H_{T}, for both data and SM simulations at pre-selection level. All backgrounds are normalised to the theoretical cross-sections.

[1] Jet cleaning implies that the event contains exactly zero *loose* bad jets. Loose jets are likely fake jets, which would affect severely the reconstruction of the $E_{\mathrm{T}}^{\mathrm{miss}}$. The maximum energy fraction in any single calorimeter layer (f_{max}) and the jet charged fraction (f_{ch}), which is defined as the ratio of the scalar sum of the p_{T} of the tracks coming from the primary vertex associated to the jet divided by the jet p_{T}, are very discriminant between fake and true jets. Jets with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.4$ are identified as loose if they satisfy $f_{\mathrm{ch}}/f_{\mathrm{max}} < 0.1$.

[2] This encompasses the upstream proton losses generated by interactions with residual gas in the beam pipe or by beam cleaning losses.

[3] The transverse mass of the top-tagged large-R jet and the $E_{\mathrm{T}}^{\mathrm{miss}}$ is defined as $m_{\mathrm{T}}(E_{\mathrm{T}}^{\mathrm{miss}}, J) = \sqrt{m(J)^2 + 2E_{\mathrm{T}}^{\mathrm{miss}}(E_{\mathrm{T}}(J) - p_{\mathrm{T}}(J)\cos(\phi(J) - \phi(E_{\mathrm{T}}^{\mathrm{miss}})))}$, where $m(J)$, $E_{\mathrm{T}}(J)$, $p_{\mathrm{T}}(J)$, $\phi(J)$ and $\phi(E_{\mathrm{T}}^{\mathrm{miss}})$ are the reconstructed invariant mass, the transverse energy, the transverse momentum, the azimuthal angle of the top-tagged large-R jet and the azimuthal angle of the $E_{\mathrm{T}}^{\mathrm{miss}}$, respectively. In case there is more than one top-tagged large-R jet, the one with highest transverse momentum is taken for the computation of this variable.

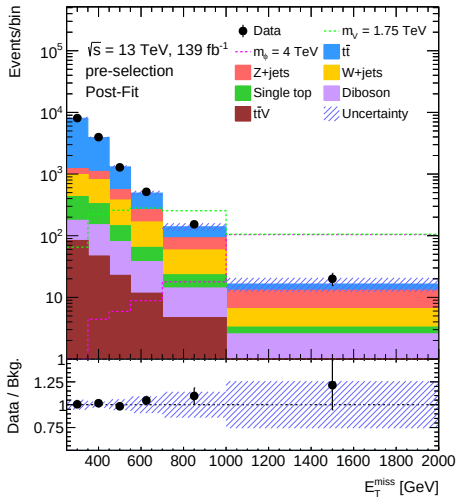
Figure 5.2 shows the same distributions after applying additional normalisation factors to the $t\bar{t}$ and V+jets backgrounds, as obtained from the background-only fit to the yields in the CRs (see Section 5.7.2). There is good agreement between the data and the SM expectations at pre-selection level. The DM mono-top signal distributions are also displayed, both for the DM resonant model with the parameters $\lambda = 0.6$, $y = 0.4$, $m_\chi = 1$ GeV and $m_\phi = 4$ TeV, and for the non-resonant model with the parameters $a = 0.5$, $g_\chi = 1$, $m_\chi = 1$ GeV and $m_V = 1.75$ TeV. The uncertainty bands include all systematic sources as described in Section 5.6.



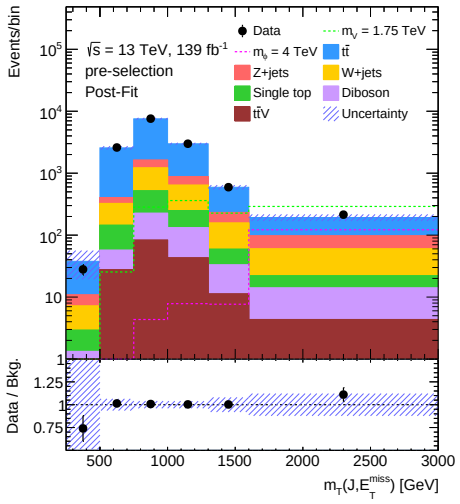
(a) (b)

(c)

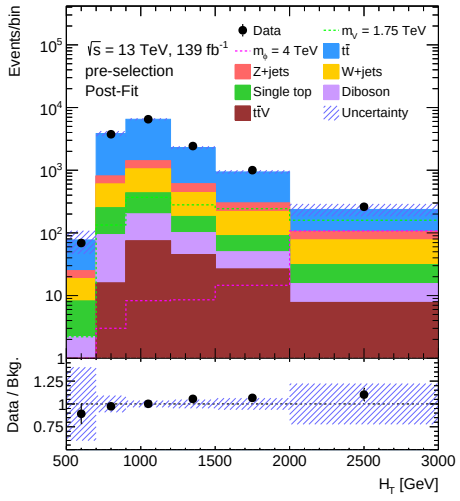
Figure 5.1: Comparison of data and SM predictions (pre-fit) for (a) the $E_{\mathrm{T}}^{\mathrm{miss}}$, (b) the $m_{\mathrm{T}}(E_{\mathrm{T}}^{\mathrm{miss}}, J)$ and (c) the H_{T} distributions at pre-selection level. The expected distributions for the DM mono-top signal models are shown for the DM resonant model, assuming a mass of the mediation of m_{ϕ} = 4 TeV and also for the DM non-resonant model, assuming a mass of the mediator of m_V = 1.75 TeV. The last bin of the distributions includes the overflow. The uncertainty bands include all systematic sources. The lower panel shows the ratio of data to SM prediction.



(a)



(b)



(c)

Figure 5.2: Comparison of data and SM predictions (post-fit) for (a) the $E_{\mathrm{T}}^{\mathrm{miss}}$, (b) the $m_{\mathrm{T}}(E_{\mathrm{T}}^{\mathrm{miss}}, J)$ and (c) the H_{T} distributions at pre-selection level. The $t\bar{t}$ and V+jets backgrounds are normalised to the results of a maximum-likelihood fit to data yields in the control regions. The expected distributions for the DM mono-top signal models are shown for the DM resonant model, assuming a mass of the mediation of m_{ϕ} = 4 TeV and also for the DM non-resonant model, assuming a mass of the mediator of m_{V} = 1.75 TeV. The last bin of the distributions includes the overflow. The uncertainty bands include all systematic sources. The lower panel shows the ratio of data to SM prediction.

5.4.2 Multivariate analysis: The XGBoost classifier

In this analysis, the Extreme Gradient Boosting (XGBoost) classifier [413] has been used as the method to enhance the signal-to-background separation due to its high performance. XGBoost is a library based on the principles of decision tree ensembles [414] which implements the machine learning algorithms under the gradient boosting framework. More details can be found in Appendix C.

The XGBoost classifier is trained using the events which pass the pre-selection level criteria. The classifiers used in this analysis are trained to distinguish the signal from all the sources of background. Two different classifiers are used addressing the two different signal hypotheses. The signal samples used in the training are selected according to the current exclusion limits[4] [415]. Consequently, the signal sample with parameters $\lambda = 0.6$, $y = 0.4$, $m_\chi = 1$ GeV and $m_\phi = 4$ TeV ($a = 0.5$, $g_\chi = 1$, $m_\chi = 1$ GeV and $m_V = 1.75$ TeV) is used for the resonant (non-resonant) model. The number of signal entries used in the training for the resonant (non-resonant) model are approximately 210 000 (57 000).

The events are divided in a training and a testing set, and the k-folding cross-validation method is utilised [416]. This method consists on randomly splitting the dataset in K subsets of identical size. The first set ($k = 1$) is kept for the testing and the remaining $K - 1$ parts are used for the training. After the training, the performance of the classifier is evaluated on the first set. The result is repeated K times, changing the test dataset from $k = 1$ to $k = K$ and using the remaining sets for the training in each case. Then, the global performance of the classifier is obtained averaging over the considered K-folds. In this analysis, K is set equal to five. This implies that 80% of the dataset is used for the training, whereas 20% is used for the testing. This is a common value which ensures a good balance between having a sizeable dataset for the training and having stable results in the testing stage. In the end, the five trained classifiers are employed to calculate the XGBoost score of their associated testing subsets. The data is treated in a similar

[4] These samples were produced with higher statistics, thus enhancing the performance of the classifier.

way: it is divided into five sets of equal size, and the XGBoost score is calculated for each subset using one of the five trained classifiers.

An illustration of the k-folding cross-validation method can be found in Figure 5.3. Notice that the testing subsets are statistically independent between them. This procedure ensures that the performance of the classifier is independent of the considered dataset.

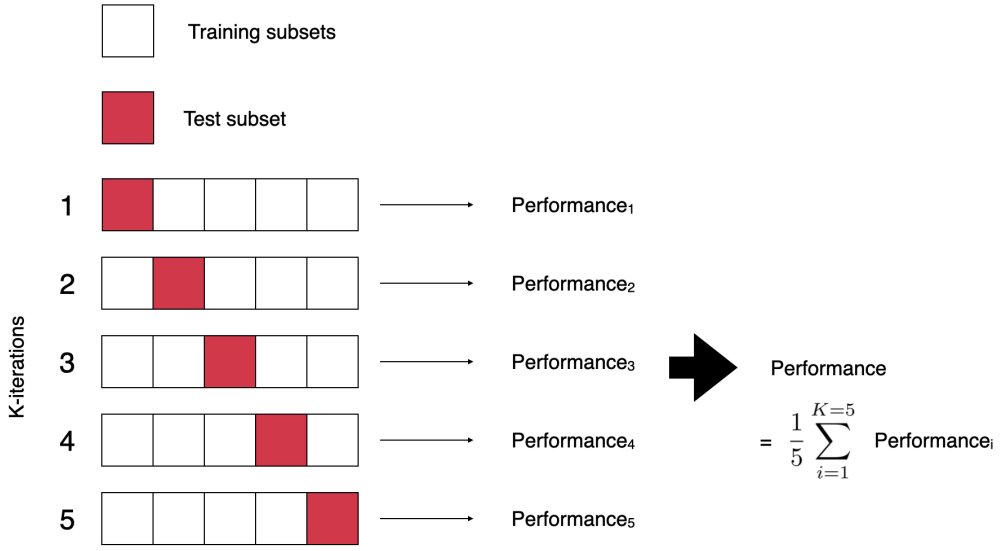


Figure 5.3: Schematic representation of the k-folding cross-validation method, with $K = 5$. For each fold, the dataset is divided in a training and testing subset, including 80% and 20% of the dataset, respectively. The performance of the classifier is evaluated for each of the folds, and the final performance is obtained as the average of all them. Note that the testing set is different in each of the folds, ensuring the independence of the performance of the classifier from the chosen test dataset.

The variables used in the training of the classifiers are optimised, as described in Appendix C[5]. The input variables for the two classifiers are summarised in Table 5.4 and defined in the following:

[5]The number of b-jets, $N_{b\text{-jets}}$, and the $\Delta\phi_{\text{min}}$ are not included in the training since these are used to define the signal region, as described in Section 5.4.3.

- $E_{\mathrm{T}}^{\mathrm{miss}}$: the missing transverse momentum of the event. The DM mono-top signal is expected to have a larger $E_{\mathrm{T}}^{\mathrm{miss}}$ with respect to background events.

- $\Delta p_{\mathrm{T}}(J, \text{jets})$: Difference of the p_{T} of the leading large-R jet with respect to the sum of p_{T} of all the small-R jets in the event. In the signal events, all high-p_{T} jets are expected to come from the top-quark. Therefore, the sum of their p_{T} will be closer to the large-R jet p_{T}, so that this variable peaks around 0.

- Ω: this is the ratio of the difference between $E_{\mathrm{T}}^{\mathrm{miss}}$ and the p_{T} of the leading large-R jet with respect to their sum, $\Omega = \frac{E_{\mathrm{T}}^{\mathrm{miss}} - p_{\mathrm{T}}(J)}{E_{\mathrm{T}}^{\mathrm{miss}} + p_{\mathrm{T}}(J)}$. The large-$R$ jet p_{T} in signal events is expected to be back-to-back with the $E_{\mathrm{T}}^{\mathrm{miss}}$ (especially in the resonant mono-top model). Therefore, this distribution is expected to be centred at zero and symmetric for signal events.

- N_{jets}: this is the small-R jet multiplicity. Signal events are expected to have less jets in the final state with respect to $t\bar{t}$.

- ΔR_{max}: this is the maximum ΔR between small-R jets in the event. Signal events are featured by a boosted decay of the top-quark. Therefore, jets are expected to be closer with respect to $t\bar{t}$.

- $m_{\mathrm{T,min}}(E_{\mathrm{T}}^{\mathrm{miss}}, b\text{-jets})$: this is the reconstructed transverse mass of the $E_{\mathrm{T}}^{\mathrm{miss}}$ with the closest b-jet[6]. This distribution shows a clear peak around the top mass in case of semileptonic $t\bar{t}$ events with a missing lepton.

- m_{top}: mass of the large-R jet that is top-tagged. It is expected to peak around the top-quark mass for the mono-top while a less pronounced peak is expected in both $t\bar{t}$ (as semileptonic events are dominant) and V+jets events.

- H_{T}: sum of the $E_{\mathrm{T}}^{\mathrm{miss}}$ and the p_{T} of all the small-R jets in the event. Different hadronic activity is expected from the signal and backgrounds.

[6]The reconstructed transverse mass of the $E_{\mathrm{T}}^{\mathrm{miss}}$ with the closest b-jet is defined as $m_{\mathrm{T,min}}(E_{\mathrm{T}}^{\mathrm{miss}}, b\text{-jets}) = \sqrt{m(b\text{-jet})^2 + 2E_{\mathrm{T}}^{\mathrm{miss}}(E_{\mathrm{T}}(b\text{-jet}) - p_{\mathrm{T}}(b\text{-jet}) \cos\Delta\phi_{\mathrm{min}}(b\text{-jets}, E_{\mathrm{T}}^{\mathrm{miss}})}$, where $m(b\text{-jet})$, $E_{\mathrm{T}}(b\text{-jet})$ and $p_{\mathrm{T}}(b\text{-jet})$ are the reconstructed invariant mass, the transverse energy and the transverse momentum of the closest b-jet to the $E_{\mathrm{T}}^{\mathrm{miss}}$ in the azimuthal plane, respectively, and $\Delta\phi_{\mathrm{min}}(b\text{-jets}, E_{\mathrm{T}}^{\mathrm{miss}})$ is the minimum azimuthal distance between the $E_{\mathrm{T}}^{\mathrm{miss}}$ and the b-jets.

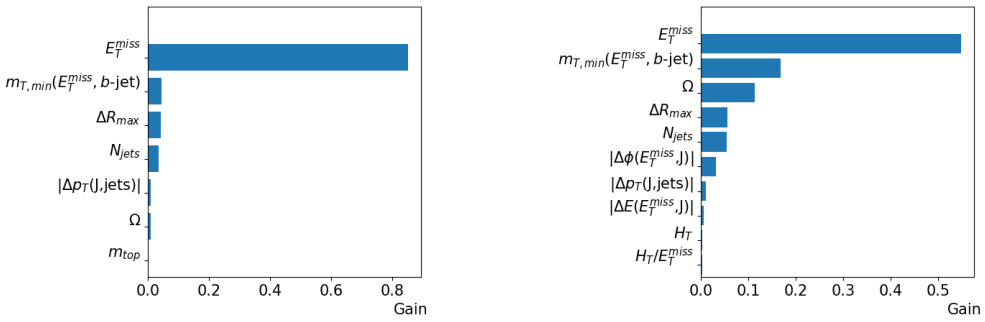
Variable	Resonant model	Non-resonant model
$E_{\mathrm{T}}^{\mathrm{miss}}$	✓	✓
Δp_{T} (J,jets)	✓	✓
Ω	✓	✓
$\mathrm{N}_{\mathrm{jets}}$	✓	✓
ΔR_{max}	✓	✓
$m_{\mathrm{T,min}}(E_{\mathrm{T}}^{\mathrm{miss}}, b\text{-jet})$	✓	✓
m_{top}	✓	
$H_{\mathrm{T}}/E_{\mathrm{T}}^{\mathrm{miss}}$		✓
$\Delta E(E_{\mathrm{T}}^{\mathrm{miss}}, \mathrm{J})$		✓
$\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, \mathrm{J})$		✓
H_{T}		✓

Table 5.4: List of variables used in the training of the XGBoost classifier for the resonant and non-resonant models.

- $H_{\mathrm{T}}/E_{\mathrm{T}}^{\mathrm{miss}}$: signal events expect H_{T} and $E_{\mathrm{T}}^{\mathrm{miss}}$ to be balanced, while it is expected to be asymmetric in $t\bar{t}$ events.
- $\Delta E(E_{\mathrm{T}}^{\mathrm{miss}}, J)$: difference of transverse energy of the $E_{\mathrm{T}}^{\mathrm{miss}}$ and the leading large-R jet. This is expected to be close to zero for signal, while highly unbalanced in the $t\bar{t}$ background.
- $\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, J)$: angular distance in the transverse plane of the $E_{\mathrm{T}}^{\mathrm{miss}}$ and the leading large-R jet. Signal events are expected to have a back-to-back large-R jet and $E_{\mathrm{T}}^{\mathrm{miss}}$.

The modelling of these variables in the training region has been checked carefully. The data and the SM expectations agree within uncertainties. More details on the modelling of the variables can be found in Appendix C. Figure 5.4 shows the ranking of these variables in terms of the XGBoost gain[7]. As one can see, the $E_{\mathrm{T}}^{\mathrm{miss}}$ is clearly the most discriminant variable between the signal and the backgrounds for the two models considered.

[7]Gain is a metric defined by XGBoost which involves evaluation of the structure of the tree. It determines the score it gains after splitting one leaf into two leaves [417].



(a) (b)

Figure 5.4: Ranking of the different input variables as a function of their discrimination power between the signal and the backgrounds for (a) the resonant and (b) the non-resonant models.

There are several parameters which determine the response of the XGBoost classifier. These parameters are also optimised in order to maximise performance avoiding overfitting[8] (see Appendix C). The optimised parameters include:

- *max_depth*: maximum depth of a tree.
- *n_estimators*: the total number of trees.
- *learning_rate*: the step size shrinkage used in each update to prevent overfitting.
- *min_child_weights*: the minimum number of weighted events in each leaf.
- *scale_pos_weight*: the balance between positive (signal) and negative (background) weights.

The results of this optimisation procedure lead to the values of the parameters listed in Table 5.5, for both the resonant and the non-resonant models.

Figure 5.5 shows the response of the classifier. As one can see, the classifier distinguishes quite well the signal from the considered backgrounds and no

[8]Overfitting is an undesirable state when the classifier becomes too specialised in the trained data and thus cannot be extrapolated to other datasets.

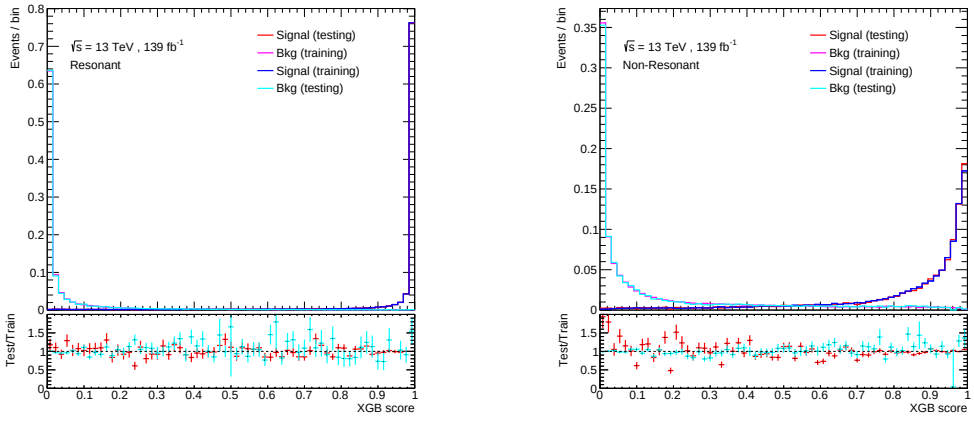
Parameter	Resonant model	Non-resonant model
n_estimators	2000	4000
max_depth	3	3
learning_rate	0.12	0.01
min_child_weights	2	10
scale_pos_weight	8	78

Table 5.5: Optimised parameters in the training of the XGBoost classifier for both the resonant and the non-resonant models.

significant difference is observed in the performance on the training and the testing samples. Figure 5.6 shows the receiver operating characteristic (ROC) curve[9] for both the resonant and the non-resonant models. The area under the ROC curve is one of the main quality metrics of the classifier. The obtained values are of 0.986 for the DM resonant model and 0.918 for the DM non-resonant model. These values are very close to unity, which is the ideal classification scenario. The XGBoost classifier is thus able to distinguish the signal from the different backgrounds with high accuracy.

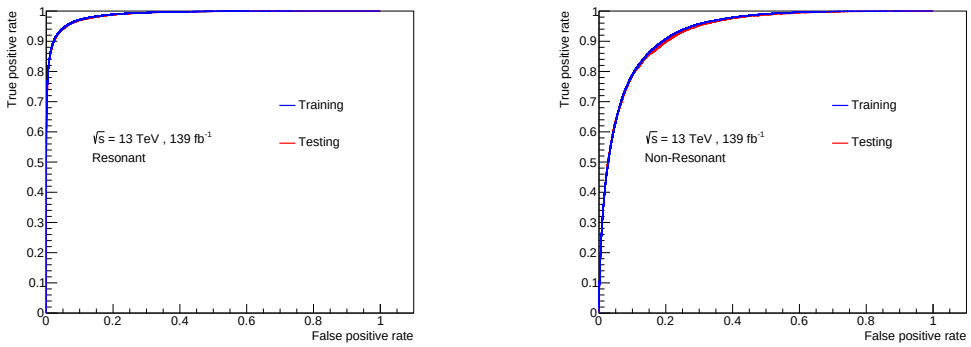
Figure 5.7 shows the XGBoost score at pre-selection level for the two classifiers considered. The $t\bar{t}$ and V+jets backgrounds are normalised using the scale factors obtained from the background-only fit to the yields in the CRs (see Section 5.7.2). As one can see, there is a very good agreement between the data and the expectation, implying there is no mismodelling introduced by the XGBoost classifier. In these distributions, background events accumulate on the left-side of the distribution (corresponding to scores close to zero value), while signal events populate the right-side (with values close to one). Therefore, the score of the XGBoost classifier is an optimal variable to identify the signal region, as discussed in the following section.

[9]The ROC curve illustrates the diagnostic ability of a binary classifier system as its discrimination threshold is varied. It represents the true positive rate against the false positive rate. The true positive rate is defined as the ratio between the number of true positive instances and the number of real positive instances, whereas the false positive rate is defined as the ratio between the number of false positive instances and the real negative instances.



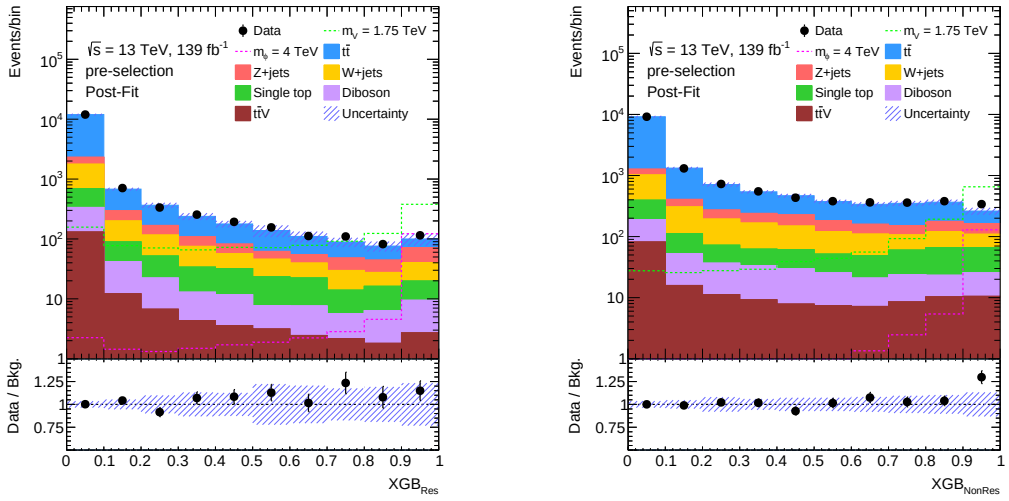
(a) (b)

Figure 5.5: Score of the XGBOOST classifier for the training and the testing samples and for the signal and backgrounds events. In (a), the resonant model is taken as signal, whereas in (b) the non-resonant model is used as signal. The small differences between the training and the testing samples indicate no signs of overtraining. The error bands include the statistical uncertainties.



(a) (b)

Figure 5.6: ROC curve for both the training and the testing dataset in (a) the resonant and (b) the non-resonant models.



(a) (b)

Figure 5.7: Comparison of data and SM predictions (post-fit) for the XGBoost score trained with (a) the resonant model and (b) the non-resonant model at pre-selection level. The $t\bar{t}$ and V+jets backgrounds are normalised to the results of a maximum-likelihood fit to data yields in the control regions. The expected distributions for the DM mono-top signal models which are used in each case for the training of the classifier are also shown. For the DM resonant model, this corresponds to a mass of the mediation of m_ϕ = 4 TeV, whereas for the DM non-resonant model this corresponds to a mass of the mediator of m_V = 1.75 TeV. The uncertainty bands include all systematic sources. The lower panel shows the ratio of data to SM prediction.

5.4.3 Signal region

The signal region is defined by applying additional requirements to make the region signal-enhanced. Two variables are used to define this region: the b-jet multiplicity, $N_{b\text{-jets}}$, and $\Delta\phi_{\text{min}}$. A comparison of their observed and expected distributions at pre-selection level is shown in Figure 5.8.

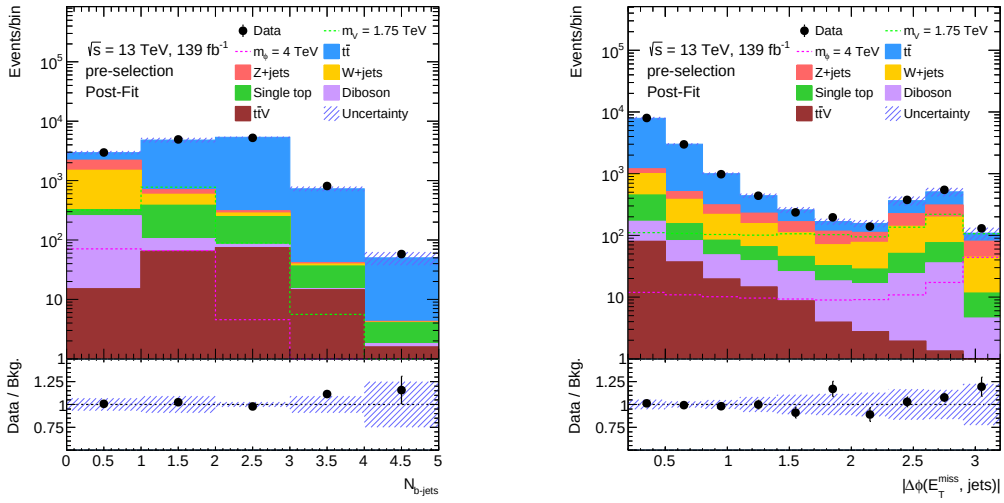
DM events are expected to have a single b-jet from the top-quark decay in the final state in case it is reconstructed. Events from $t\bar{t}$ production are expected to lead to two b-jets, while V+jets are expected to lead to zero or two b-jets (in case of gluon splitting), assuming a perfect reconstruction efficiency. Therefore, the signal region is defined by applying $N_{b\text{-jets}} \leq 1$ to account for reconstruction inefficiencies.

In addition, the DM signal is characterised by a large amount of missing transverse momentum which is more likely to be produced back-to-back with the boosted top-quark. Therefore, a large angular separation between the $E_{\text{T}}^{\text{miss}}$ and the jets (which are expected to come from the top-quark decay) is required. Events from $t\bar{t}$, especially in the semileptonic channel (when the final state lepton is misidentified with a jet), do not have this kind of topology. Events from V+jets have also a different topology, except for the $Z(\rightarrow \nu\bar{\nu})$+jets. Thus, a requirement of $\Delta\phi_{\text{min}} > 1.0$ is applied in the signal region.

The response of the XGBoost classifier is used as the discriminant variable for the profile-likelihood fit, as explained in Section 5.7.1. However, the left side of the distribution is highly signal-depleted and therefore can be used as an additional validation region to check the background modelling, as described in Section 5.4.4. On the other hand, the right side of the XGBoost distribution is highly populated by signal events. Therefore, the XGBoost score is required to be larger than 0.5 in the signal region.

5.4.4 Background estimation

The main contributing backgrounds in the signal regions are the $t\bar{t}$ and the V+jets. The former represents 46% (48%) of the total background in the resonant (non-resonant) signal region, whilst the latter represents 38% (37%) in the resonant (non-resonant) signal region. Single-top-quark production, diboson processes and



(a) (b)

Figure 5.8: Comparison of data and SM predictions (post-fit) for (a) the $N_{b\text{-jets}}$ and (b) the $\Delta\phi_{\text{min}}$. The $t\bar{t}$ and V+jets backgrounds are normalised to the results of a maximum-likelihood fit to data yields in the control regions. The expected distributions for the DM mono-top signal models are shown for the DM resonant model, assuming a mass of the mediation of m_ϕ = 4 TeV and also for the DM non-resonant model, assuming a mass of the mediator of m_V = 1.75 TeV. In (a), the last bin of the distribution includes the overflow. The uncertainty bands correspond to the errors due to the limited size of the simulation samples. The lower panel shows the ratio of data to SM prediction.

the production of a vector boson in association with top quarks have a minor contribution ($\sim$16%) in both cases.

5.4.4.1 Control and validation regions

The control and validation regions employed in this analysis are defined on top of the pre-selection region. As shown in Figure 5.8, the distribution of $N_{b\text{-jets}}$ discriminates $t\bar{t}$ events from V+jets events. Consequently, the $t\bar{t}$ (V+jets) control region is defined by requiring two or more (exactly zero) b-jets. In addition, a $\Delta\phi_{\text{min}} \in [0.2, 1]$ requirement is applied to reduce the signal contribution in both cases.

The $t\bar{t}$ validation region is obtained by changing the $t\bar{t}$ control region requirement on the number of b-jets to exactly one. The validation region for V+jets is defined applying the same requirements on $N_{b\text{-jets}}$ and $\Delta\phi_{\text{min}}$ as in the signal region. Nevertheless, the V+jets validation region corresponds to the lower tail of the XGBoost distribution by requiring the score to be below 0.5. Consequently, two V+jets VRs are defined, one for the resonant and one for the non-resonant case. These regions are still signal-depleted, thanks to the high discrimination power of the classifier, and contain a significant fraction of V+jets events. A summary of the different requirements used to define the signal, control and validation regions can be found in Figure 5.9.

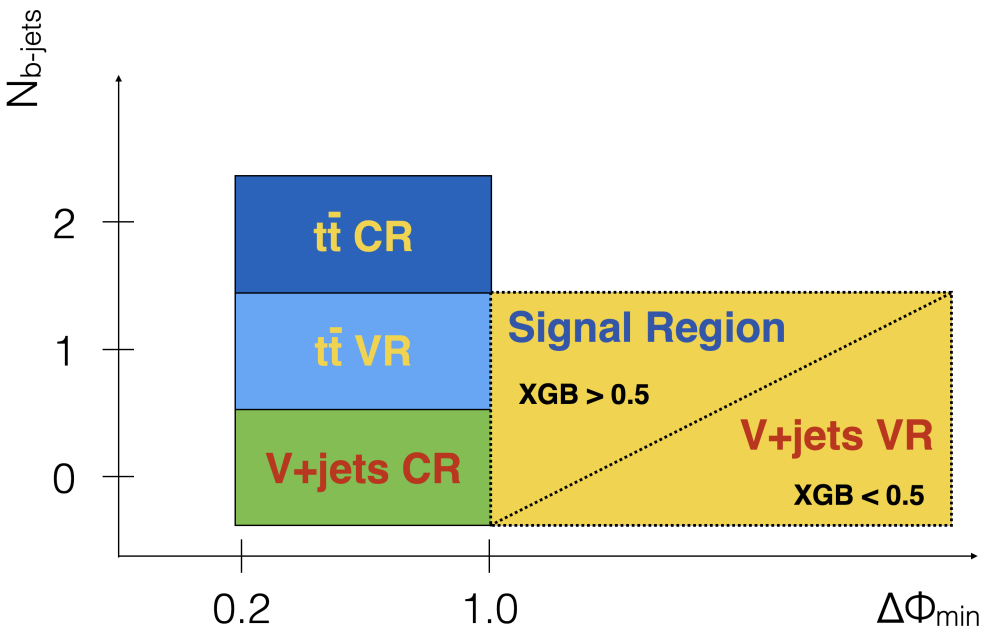


Figure 5.9: Schematic representation of the control, validation and signal regions in the $N_{\text{b-jets}}$ and $\Delta\phi_{\text{min}}$ plane. XGB represents the XGBoost score.

These definitions ensure that the composition of the $t\bar{t}$ and the V+jets backgrounds in the control and validation regions is very similar to the composition in the signal region. Concretely, the $t\bar{t}$ is dominated by its semileptonic component in all cases. Within the semileptonic $t\bar{t}$, $t\bar{t}$+light is the dominant component in the considered analysis regions: this constitutes 82% and 85% of the total $t\bar{t}$ background in the $t\bar{t}$ CR and the $t\bar{t}$ VR, respectively, and in the signal regions it constitutes 97% (92%) for the resonant (non-resonant) case.

The V+jets backgrounds is split in W+jets and Z+jets. The former constitutes 71% of the total V+jets contribution in the V+jets CR, $\sim$55% in the V+jets VR for both the resonant and the non-resonant case, and $\sim$47% in the signal region for both the resonant and the non-resonant case.

For Z+jets events, those where the Z-boson decays to invisible ($Z(\nu\nu)$+jets) have the largest contribution. The relative fraction of $Z(\nu\nu)$+jets with respect to the total

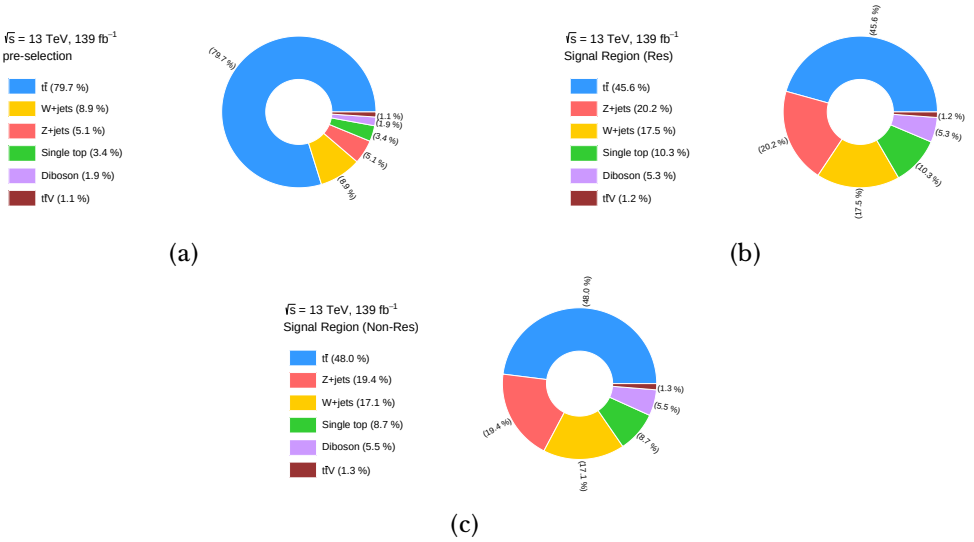


Figure 5.10: Expected contribution of the different backgrounds in (a) the pre-selection region, (b) the DM resonant model signal region and (c) the DM non-resonant model signal region. These results are previous to the normalisation fit described in Section 5.7.2.

Z+jets background is 83% in the V+jets CR, ∼90% in the V+jets VR for both the resonant and the non-resonant case, and ∼94% in the signal regions for both the resonant and the non-resonant case.

For the W+jets background, those were the W-boson decays to a τ lepton and a neutrino ($W(\tau\nu)$+jets) constitute the dominant component. Its relative fraction with respect to the total W+jets background is 70% in the V+jets CR, ∼67% in the V+jets VR for both the resonant and the non-resonant case, and 57% (67%) in the signal region for the resonant (non-resonant) case.

Figures 5.10-5.11 show the background composition in the pre-selection, signal, control and validation regions. As it can be seen, the $t\bar{t}$ control region is very pure in $t\bar{t}$ background, with a relative composition of 96%. In the $t\bar{t}$ validation region, the relative fraction of $t\bar{t}$ is almost 90%. For the V+jets control and validation regions, the relative fraction of V+jets events is approximately 50%.

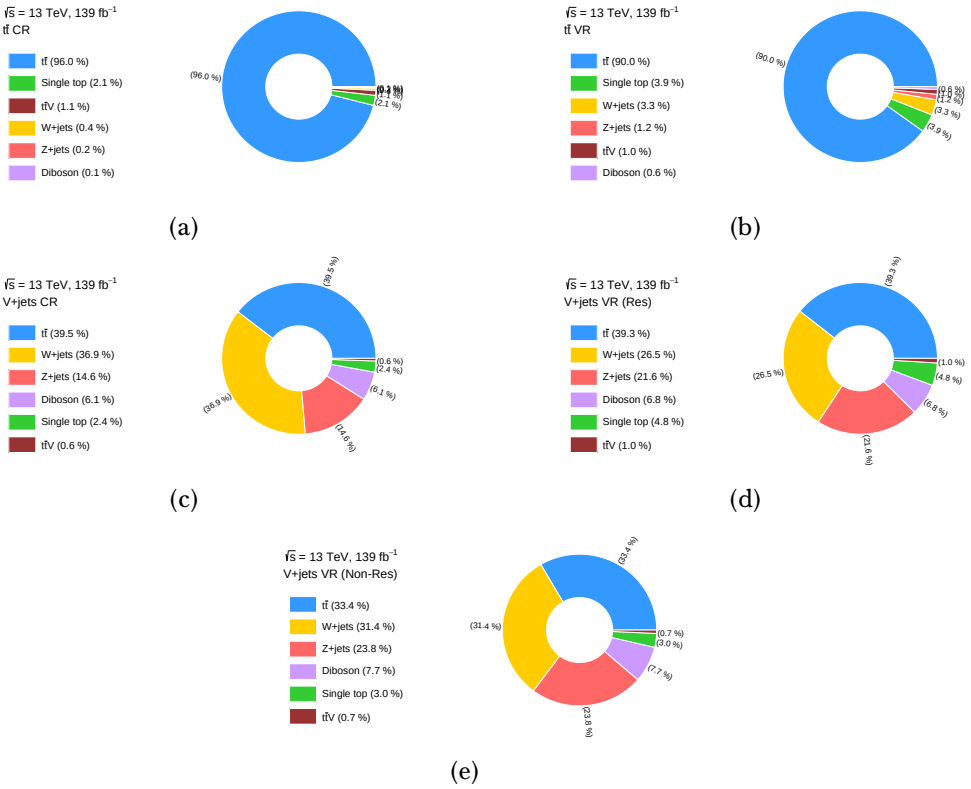


Figure 5.11: Expected contribution of the different backgrounds in (a) the $t\bar{t}$ CR, (b) the $t\bar{t}$ VR, (c) the V+jets CR, (d) the V+jets VR for the DM resonant model and (e) the V+jets VR for the DM non-resonant model. These results are previous to the normalisation fit described in Section 5.7.2.

5.4.4.2 Multijet estimation

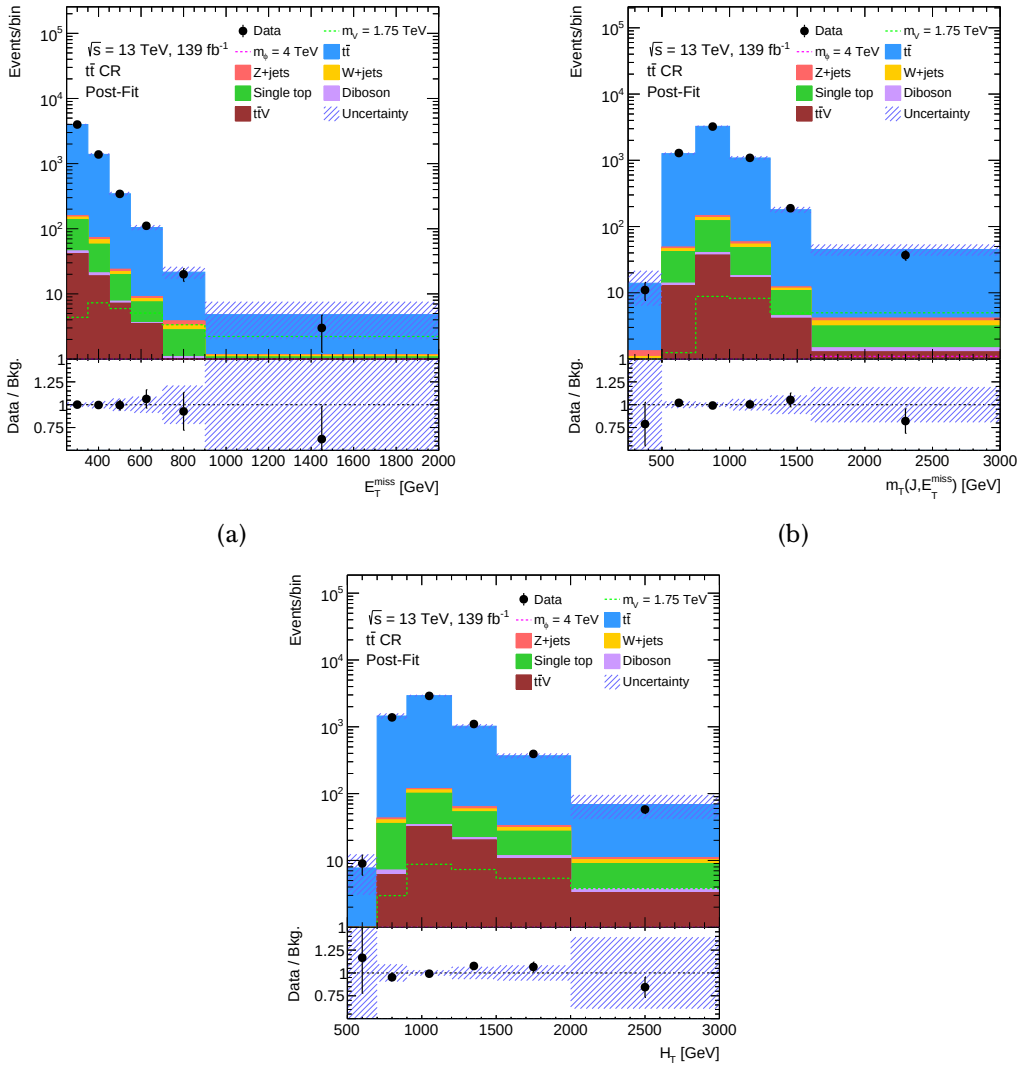
The impact of the multijet background has been studied using MC simulation of dijet events generated with PYTHIA8 and scaled up to the full Run-2 luminosity. The estimation is performed by applying the conditions of each analysis region to the events from this MC sample, except for the requirement on the $E_{\mathrm{T}}^{\mathrm{miss}}$, that is loosened to 50 GeV. The trend of the multijet yields with the $E_{\mathrm{T}}^{\mathrm{miss}}$ is fitted in each region with an exponential function to extrapolate the yields to $E_{\mathrm{T}}^{\mathrm{miss}} > 250$ GeV, which corresponds to the analysis regions. The extrapolated yields result to be systematically lower than the yields obtained directly from the MC in all regions. To correct for this effect, an additional correction factor is applied to the extrapolated yields. This is extracted from a χ^2 minimisation, taking into account the differences between the MC and the extrapolated yields in all regions and assuming that this correction factor is the same in all regions. The estimated multijet yields, compared to the total background yields, are shown in Table 5.6. As one can see, this source of background is estimated to be at maximum at percent level, well within the statistical uncertainty of the expected data yields. Also, the obtained yields are compatible with zero within 1.5 standard deviations.

Region	Total background	Multijet	Multijet / Total Background
Pre-selection	15800 ± 2600	110 ± 90	0.70%
$t\bar{t}$ CR	6800 ± 1500	16 ± 12	0.24%
$t\bar{t}$ VR	4600 ± 900	160 ± 130	3.48%
V+jets CR	1900 ± 400	31 ± 24	1.63%
SR & V+jets VR	2080 ± 290	0.3 ± 0.2	0.01%

Table 5.6: Multijet event yields in the pre-selection, the $t\bar{t}$ control and validation region, the V+jets control region and the inclusive signal region and V+jets validation region (covering the full XGBoost distribution). The numbers are compared with the total number of expected background yields. The multijet yields are estimated from a fit to an exponential function in low $E_{\mathrm{T}}^{\mathrm{miss}}$ regions. The exponential function is then extrapolated to the analysis regions. An additional correction factor is applied to correct for differences between the extrapolated yields and the MC yields.

5.5 Event yields and distributions

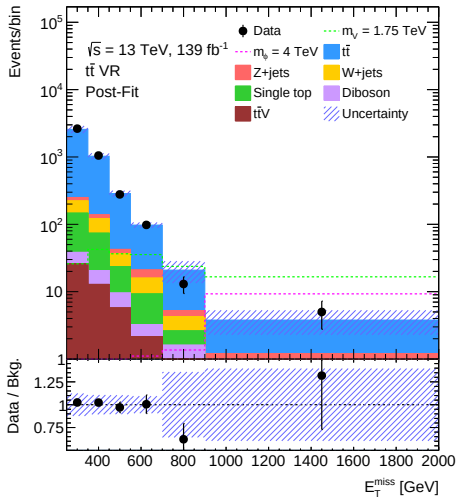
Table 5.7 provides the predicted signal and background event yields in the pre-selection, signal, control and validation regions. The distributions of the $E_{\mathrm{T}}^{\mathrm{miss}}$, the $m_{\mathrm{T}}(E_{\mathrm{T}}^{\mathrm{miss}}, J)$ and the H_{T} are depicted in Figures 5.12-5.16 for the $t\bar{t}$ control region, the $t\bar{t}$ validation region, the V+jets control region and the V+jets validation regions, respectively. Figures 5.17-5.18 show the XGBoost score for the resonant and the non-resonant model, respectively, in all the control and validation regions. In all cases, the $t\bar{t}$ and V+jets backgrounds are normalised using the scale factors obtained from the background-only fit to the yields in the CRs (see Section 5.7.2). Good overall agreement between the data and the SM predictions is observed.



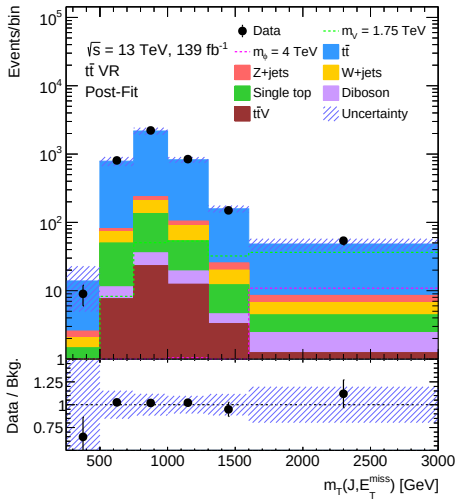
(a) (b)

(c)

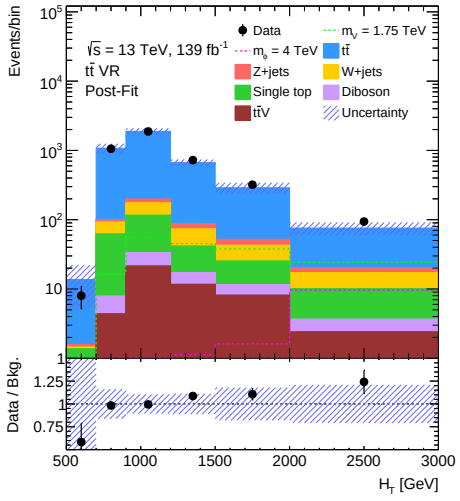
Figure 5.12: Comparison of data and SM predictions (post-fit) for (a) the $E_{\mathrm{T}}^{\mathrm{miss}}$, (b) the $m_{\mathrm{T}}(E_{\mathrm{T}}^{\mathrm{miss}}, J)$ and (c) the H_{T} distributions in the $t\bar{t}$ CR. The $t\bar{t}$ and V+jets backgrounds are normalised to the results of a maximum-likelihood fit to data yields in the control regions. The expected distributions for the DM mono-top signal models are shown for the DM resonant model, assuming a mass of the mediation of m_{ϕ} = 4 TeV, and also for the DM non-resonant model, assuming a mass of the mediator of m_{V} = 1.75 TeV. The last bin of the distributions includes the overflow. The uncertainty bands include all the systematic sources. The lower panel shows the ratio of data to SM prediction.



(a)

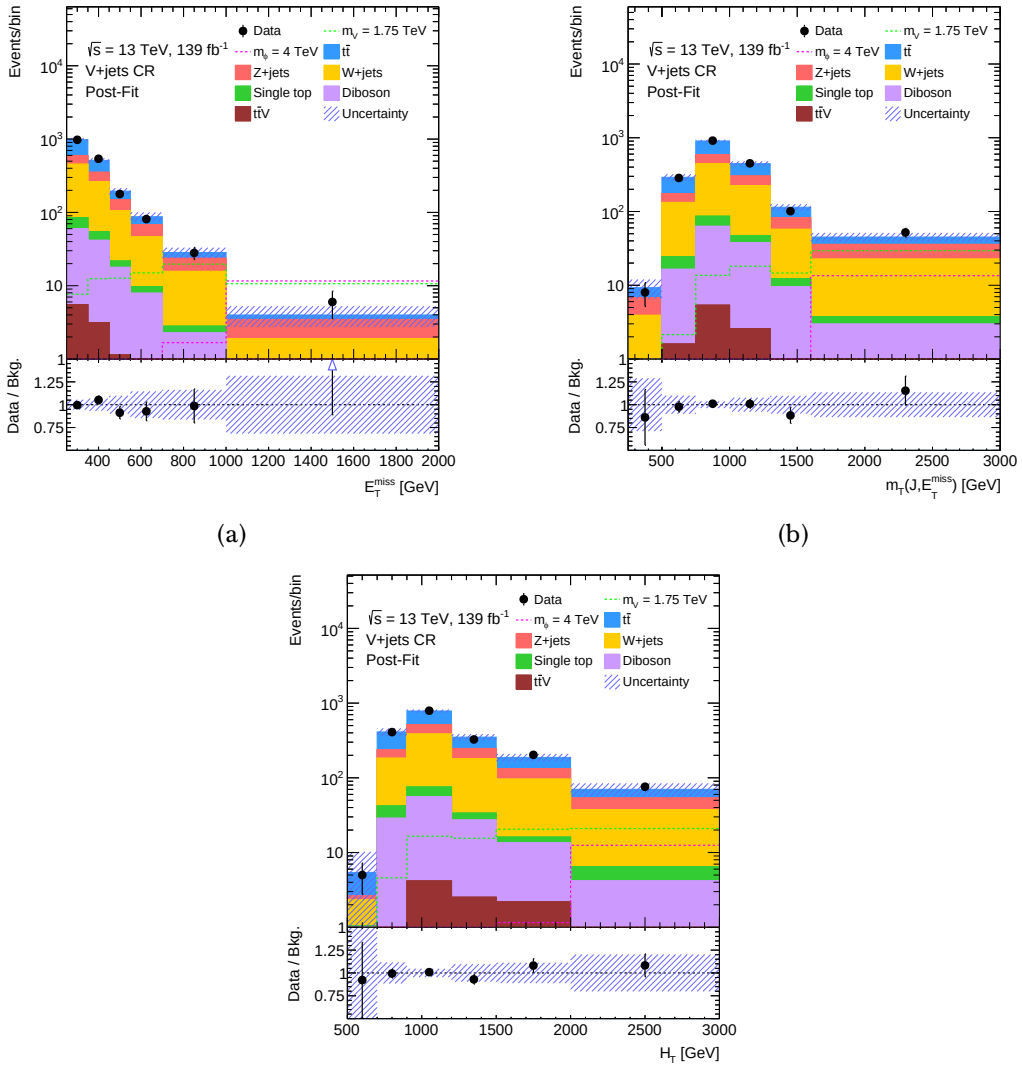


(b)



(c)

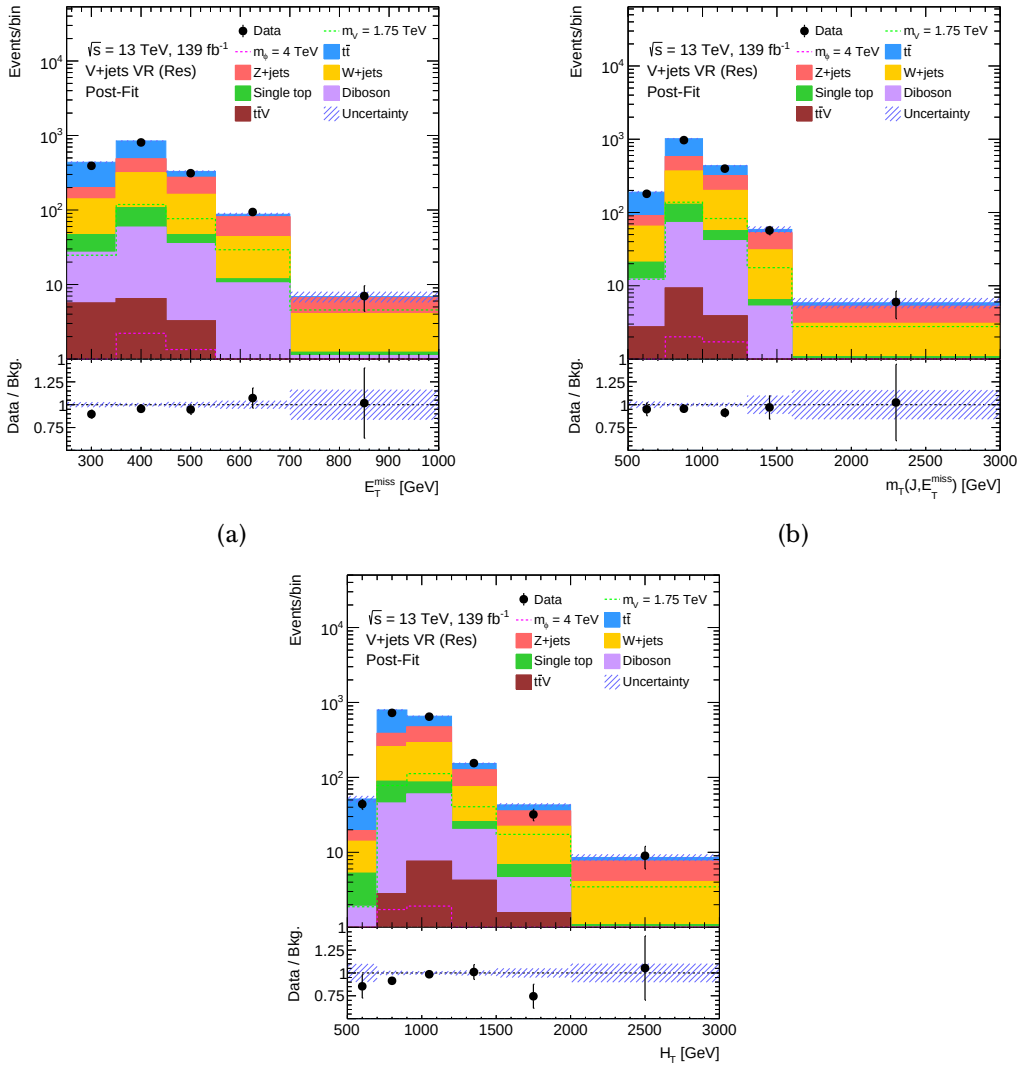
Figure 5.13: Comparison of data and SM predictions (post-fit) for (a) the $E_{\mathrm{T}}^{\mathrm{miss}}$, (b) the $m_{\mathrm{T}}(E_{\mathrm{T}}^{\mathrm{miss}}, J)$ and (c) the H_{T} distributions in the $t\bar{t}$ VR. The $t\bar{t}$ and V+jets backgrounds are normalised to the results of a maximum-likelihood fit to data yields in the control regions. The expected distributions for the DM mono-top signal models are shown for the DM resonant model, assuming a mass of the mediation of m_{ϕ} = 4 TeV, and also for the DM non-resonant model, assuming a mass of the mediator of m_{V} = 1.75 TeV. The last bin of the distributions includes the overflow. The uncertainty bands include all the systematic sources. The lower panel shows the ratio of data to SM prediction.



(a) (b)

(c)

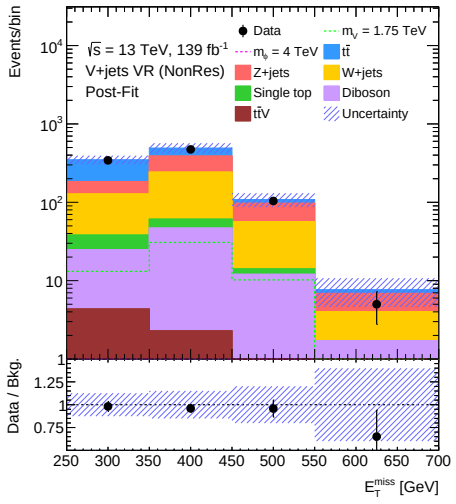
Figure 5.14: Comparison of data and SM predictions (post-fit) for (a) the $E_{\mathrm{T}}^{\mathrm{miss}}$, (b) the $m_{\mathrm{T}}(E_{\mathrm{T}}^{\mathrm{miss}}, J)$ and (c) the H_{T} distributions in the V+jets CR. The $t\bar{t}$ and V+jets backgrounds are normalised to the results of a maximum-likelihood fit to data yields in the control regions. The expected distributions for the DM mono-top signal models are shown for the DM resonant model, assuming a mass of the mediation of m_{ϕ} = 4 TeV, and also for the DM non-resonant model, assuming a mass of the mediator of m_{V} = 1.75 TeV. The last bin of the distributions includes the overflow. The uncertainty bands include all the systematic sources. The lower panel shows the ratio of data to SM prediction.



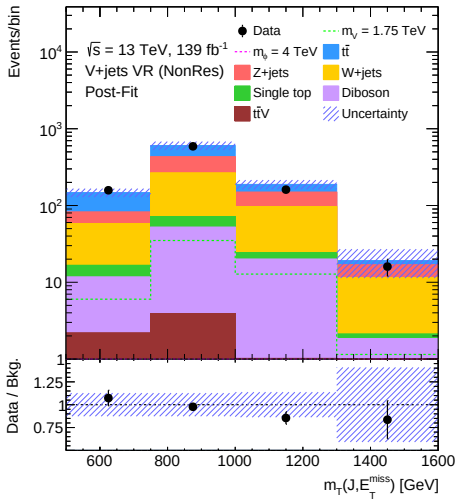
(a) (b)

(c)

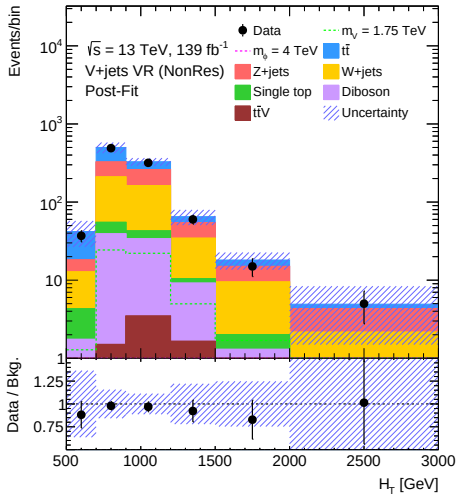
Figure 5.15: Comparison of data and SM predictions (post-fit) for (a) the $E_{\mathrm{T}}^{\mathrm{miss}}$, (b) the $m_{\mathrm{T}}(E_{\mathrm{T}}^{\mathrm{miss}}, J)$ and (c) the H_{T} distributions in the V+jets VR for the DM resonant model. The $t\bar{t}$ and V+jets backgrounds are normalised to the results of a maximum-likelihood fit to data yields in the control regions. The expected distributions for the DM mono-top signal models are shown for the DM resonant model, assuming a mass of the mediation of m_{ϕ} = 4 TeV, and also for the DM non-resonant model, assuming a mass of the mediator of m_V = 1.75 TeV. The last bin of the distributions includes the overflow. The uncertainty bands include all the systematic sources. The lower panel shows the ratio of data to SM prediction.



(a)



(b)



(c)

Figure 5.16: Comparison of data and SM predictions (post-fit) for (a) the $E_{\mathrm{T}}^{\mathrm{miss}}$, (b) the $m_{\mathrm{T}}(E_{\mathrm{T}}^{\mathrm{miss}}, J)$ and (c) the H_{T} distributions in the V+jets VR for the DM non-resonant model. The $t\bar{t}$ and V+jets backgrounds are normalised to the results of a maximum-likelihood fit to data yields in the control regions. The expected distributions for the DM mono-top signal models are shown for the DM resonant model, assuming a mass of the mediation of m_{ϕ} = 4 TeV, and also for the DM non-resonant model, assuming a mass of the mediator of m_{V} = 1.75 TeV. The last bin of the distributions includes the overflow. The uncertainty bands include all the systematic sources. The lower panel shows the ratio of data to SM prediction.

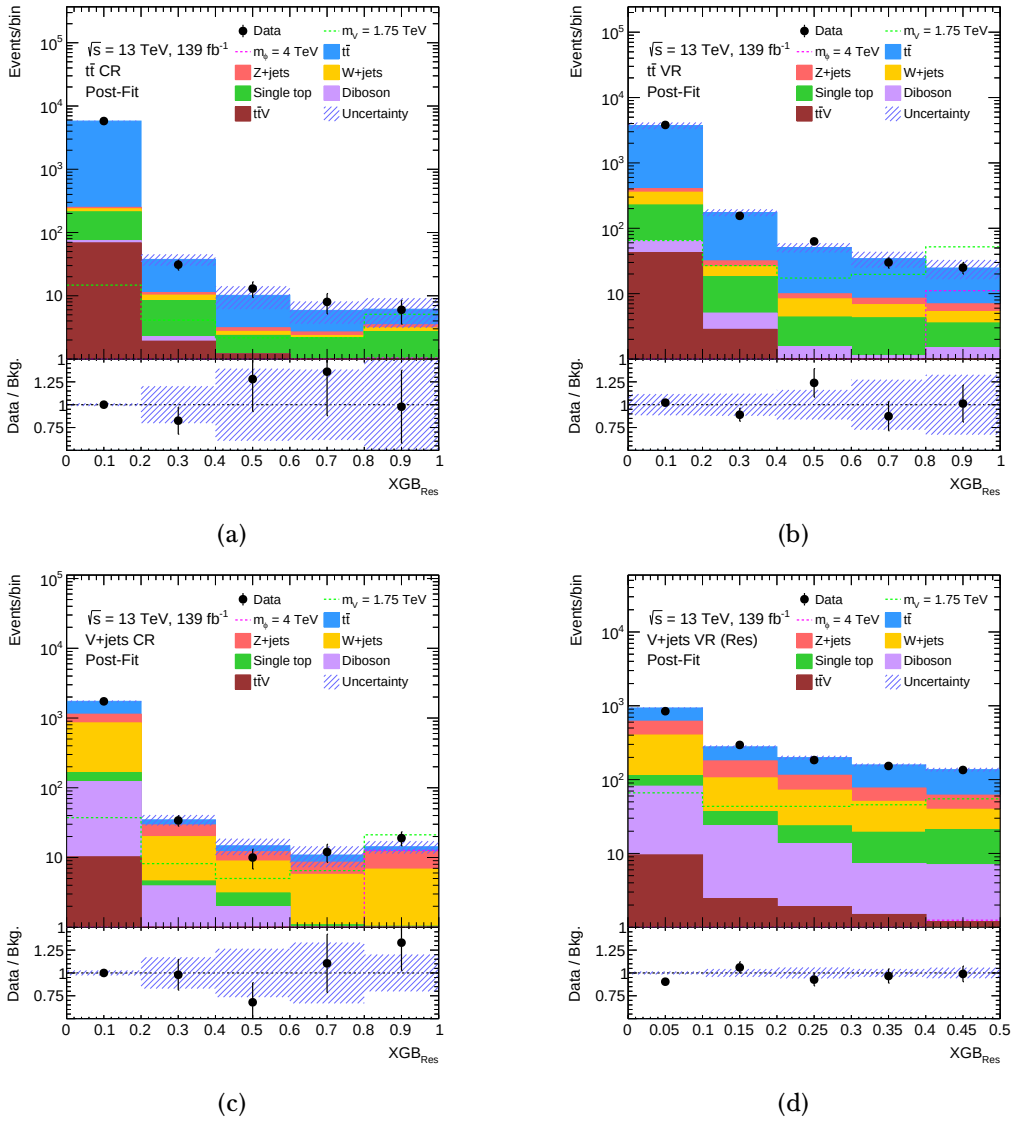


Figure 5.17: Comparison of data and SM predictions (post-fit) for the XGBoost score trained with the resonant model in (a) the $t\bar{t}$ CR, (b) the $t\bar{t}$ VR, (c) the V+jets CR and (d) the V+jets VR for the DM resonant model. The $t\bar{t}$ and V+jets backgrounds are normalised to the results of a maximum-likelihood fit to data yields in the control regions. The expected distributions for the DM mono-top signal models are shown for the DM resonant model, assuming a mass of the mediation of m_{ϕ} = 4 TeV and also for the DM non-resonant model, assuming a mass of the mediator of m_V = 1.75 TeV. The uncertainty bands include all the systematic sources. The lower panel shows the ratio of data to SM prediction.

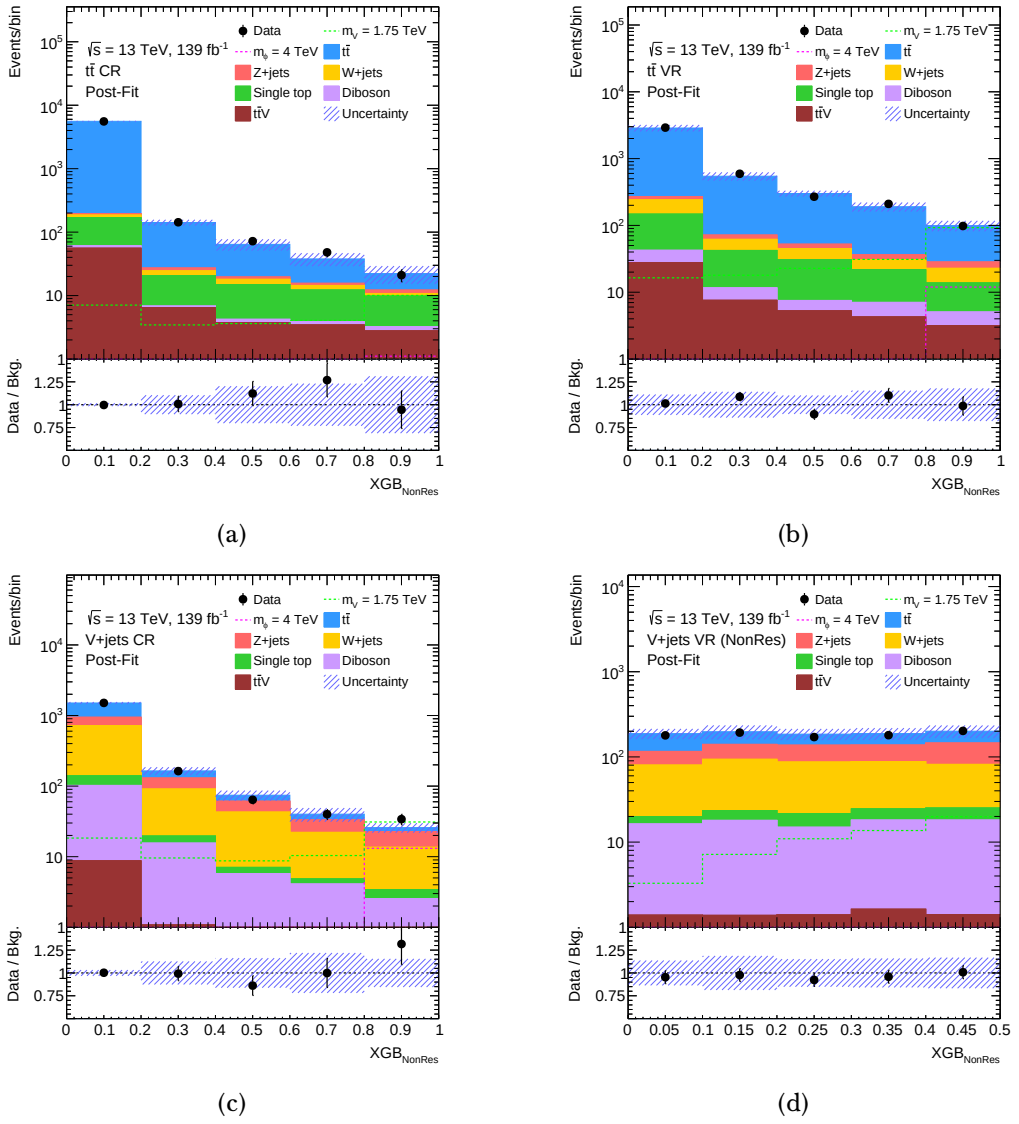


Figure 5.18: Comparison of data and SM predictions (post-fit) for the XGBoost score trained with the non-resonant model in (a) the $t\bar{t}$ CR, (b) the $t\bar{t}$ VR, (c) the V+jets CR and (d) the V+jets VR for the DM non-resonant model. The $t\bar{t}$ and V+jets backgrounds are normalised to the results of a maximum-likelihood fit to data yields in the control regions. The expected distributions for the DM mono-top signal models are shown for the DM resonant model, assuming a mass of the mediation of m_ϕ = 4 TeV and also for the DM non-resonant model, assuming a mass of the mediator of m_V = 1.75 TeV. The uncertainty bands include all the systematic sources. The lower panel shows the ratio of data to SM prediction.

Process	pre-selection	$t\bar{t}$ CR	V+jets CR	$t\bar{t}$ VR	V+jets VR (Res)	Signal Region (Res)	V+jets VR (Non-Res)	Signal Region (Non-Res)
$t\bar{t}$	12600 ± 2500	6600 ± 1500	730 ± 230	4200 ± 800	660 ± 150	170 ± 50	330 ± 80	500 ± 110
Z+jets	850 ± 330	17 ± 7	310 ± 120	58 ± 25	380 ± 150	79 ± 30	250 ± 100	210 ± 80
W+jets	1400 ± 400	27 ± 9	720 ± 230	150 ± 50	450 ± 140	64 ± 19	320 ± 100	200 ± 60
tZq	1.58 ± 0.27	0.49 ± 0.13	0.15 ± 0.07	0.65 ± 0.17	0.18 ± 0.080	0.030 ± 0.023	0.12 ± 0.07	0.10 ± 0.04
tW-channel	510 ± 140	139 ± 29	43 ± 12	170 ± 40	80 ± 20	38 ± 18	29 ± 7	90 ± 33
t-channel (t)	11.2 ± 1.3	4.5 ± 0.7	1.0 ± 0.4	5.2 ± 0.8	0.4 ± 0.5	$(4 \pm 6) \cdot 10^{-6}$	0.4 ± 0.4	0.07 ± 0.19
s-channel	2.8 ± 0.5	1.4 ± 0.4	0.21 ± 0.16	1.13 ± 0.34	0.06 ± 0.07	0.03 ± 0.06	0.03 ± 0.07	0.05 ± 0.07
t-channel ($\bar{t}$)	4.4 ± 0.7	1.5 ± 0.4	0.52 ± 0.25	2.1 ± 0.5	0.19 ± 0.17	0.03 ± 0.07	0.11 ± 0.10	0.11 ± 0.17
Diboson	290 ± 40	6.7 ± 1.2	119 ± 15	26 ± 4	116 ± 18	19.4 ± 3.4	78 ± 12	57 ± 9
$t\bar{t}V$	170 ± 24	72 ± 11	10.7 ± 1.8	48 ± 7	16.5 ± 2.6	4.4 ± 0.7	7.2 ± 1.3	13.7 ± 2.1
Resonant	142 ± 22	1.53 ± 0.24	14 ± 5	13.7 ± 2.2	4.6 ± 0.7	104 ± 16	-	-
Non-resonant	1190 ± 160	28 ± 4	78 ± 12	181 ± 25	-	-	55 ± 9	780 ± 100
Total background	15800 ± 2600	6800 ± 1500	1900 ± 400	4600 ± 900	1710 ± 340	380 ± 70	1010 ± 220	1070 ± 190
Data	14019	5843	1810	4081	1612	403	925	1090

Table 5.7: Predicted (pre-fit) event yields for the pre-selection, the signal, the control and the validation regions. The signal and backgrounds are normalised to their theoretical predictions. The uncertainties shown include all systematic uncertainties. The (resonant) non-resonant yields correspond to the model parameters $\lambda = 0.6$, $y = 0.4$, $m_\chi = 1$ GeV and $m_\phi = 4$ TeV ($a = 0.5$, $g_\chi = 1$, $m_\chi = 1$ GeV and $m_V = 1.75$ TeV).

5.6 Systematic uncertainties

Different sources of systematic uncertainty are considered in this analysis as they can significantly affect both the yields and the shape of the distributions of the expected background and signal processes in the analysis regions. These are classified in two main categories: experimental and theoretical modelling uncertainties. All sources of uncertainty are considered in the measurement by including them as NPs in the profile likelihood fit (see Section 5.7.1). Table 5.8 summarises the number of NPs associated to each source of uncertainty.

5.6.1 Experimental uncertainties

The experimental sources of uncertainty include the uncertainty in the combined integrated luminosity [246], the pile-up reweighting, the uncertainties on the JES [291], JER [291] and the JVT [394], the b-tagging efficiency [333, 334, 395] and the modelling of the missing transverse momentum [396]. More details can be found in Section 2.5 and Section 4.8.2. This analysis also includes uncertainties on the JMS of small-R jets and uncertainties associated to the large-R jets and the top-tagging, as described in the following.

Uncertainty source	Number of nuisance parameters
Luminosity	1
Pile-up	1
Small-R jets	
JES	30
JER	13
JMS	18
JVT	1
b-tagging	13
Large-R jets	
LJES	24
LJER	12
LJMS	18
LJMR	10
top-tagging	23
$E_{\mathrm{T}}^{\mathrm{miss}}$	3
Background modelling	
$t\bar{t}$	6
t-, tW-, s-channel	4
V+jets	6
Normalisation	9
Total	192

Table 5.8: Number of nuisance parameters associated to each systematic uncertainty source included in the profile likelihood fit (described in Section 5.7.1).

Small-R JMS: The JMS uncertainties for small-R jets are derived using the R_{trk} method [306] by comparing the ratio of the jet mass for topo-jets to the jet mass of track-based jets in both data and MC simulation. Since PF jets use both calorimeter and tracking information, two additional uncertainties are derived by comparing the jet mass of topo-jets and PF jets in data and MC. These uncertainties are derived after the application of the JES and JER smearing, which have an impact on the jet mass. The JMS correction and uncertainties are then a residual correction accounting for the distribution of energy within the jet. These uncertainties encompass a total of 18 NPs.

Large-R jets: The large-R jet calibration is obtained from tag-and-probe analyses [306], using similar techniques as for small-R jets. The large-R JES (LJES) uncertainties [306] consist of 24 NPs addressing different sources of uncertainty, including detector effects, selections applied, modelling, statistics, the calibration of jets in the forward region and the differences in the jet response modelling of jets originating from top quark decays, other quarks, or gluons. The large-R JMS (LJMS) uncertainties [306, 418] encompass 18 NPs coming from the uncertainties on the forward folding and R_{trk} techniques, the uncertainties on the interpolation between different mass bins and the uncertainties on the difference in the jet mass response between jets originating from a top quark and jets originating from other quarks or gluons. The large-R JER (LJER) uncertainties [306] come from the data/MC difference, which is taken as an additional uncertainty on top of the uncertainties arising from detector effects, selections applied, statistics and modelling. This results into 12 NPs. The large-R JMR (LJMR) uncertainties [306, 418] include a total of 10 NPs, including the uncertainties on the forward folding technique, a flat 20% uncertainty outside the regime of validity of this technique and the uncertainties on the interpolation between different mass bins.

Large-R jet top-tagging: The uncertainties associated with the top-tagging [322, 323] consist of 23 NPs, including signal and background efficiency components, inefficiencies of the taggers and a component addressing the extrapolation at high-p_{T} regime.

5.6.2 Modelling uncertainties

Modelling uncertainties are associated to the choice of the baseline setup for the simulations of the signal and background events. These uncertainties are considered for the predictions of the three most important backgrounds in the analysis region: $t\bar{t}$, single-top processes and V+jets production. In addition, cross-section uncertainties are associated to all the considered backgrounds.

Modelling of top-quark pairs background: Uncertainties due to missing QCD higher order corrections are estimated by varying the renormalisation and factorisation scale by a factor 0.5 and 2 with respect to the value used in the baseline simulation. Additionally, uncertainties due to initial-state radiation are assessed by varying the corresponding parameters of the A14 parton shower tune `Var3c` [349]. The uncertainties due to final-state radiation are evaluated by varying up and down the renormalisation scale for QCD emission by factors of 0.5 and 2.0. The predictions from these varied settings are compared to the nominal simulation.

The impact of the PDF uncertanties is evaluated varying the event weights and comparing the alternative templates with the nominal sample. A hundred variations of the nominal NNPDF3.0 PDF set are considered. The envelope is then used to estimate the impact of this uncertainty on the considered observables. PDF uncertainties have been evaluated to have an impact of $\approx 1\%$, and thus are pruned away to increase the fit stability.

Uncertainties coming from the choice of the matrix element generator and the parton shower model are evaluated by comparing the baseline setup Powheg-Box+Pythia8 with alternative samples using MG5_aMC@NLO+Pythia8 and Powheg-Box+Herwig7, respectively. The estimated uncertainties are then symmetrised.

Modelling of single-top backgrounds: Uncertainties due to missing QCD higher order corrections, initial-state and final-state radiation and those originating from the PDF choice are evaluated using the same procedure as described for the $t\bar{t}$ background. For the tW-channel, an additional uncertainty is evaluated by using a simulation with the same setup but using the DS scheme. The difference between

the alternative prediction and the baseline prediction, which uses the DR scheme, is used as estimation of this uncertainty.

Modelling of weak boson production associated with jets: Uncertainties originating from missing QCD higher orders, resummation and parton shower matching are evaluated following the procedure described in Ref. [419]. Here, a parametrisation at truth-level is developed to evaluate the relative uncertainties due to the variations of the above-mentioned scales. Through this method, a set of weights based on a minimal set of variables are defined and associated to each event to account for the associated systematic effect.

As the analysis uses also a region with more than one b-jet in the final state, a dedicated uncertainty is associated to the production of a weak boson associated with heavy-flavour jets. A normalisation uncertainty of 30% is associated to the rates of events of V+b-jets and V+c-jets [398]. V+b-jets are identified by events with at least one jet in the final state that is truth-matched with at least one B-hadron[10]. In case no B-hadrons are found in the final state and a D-hadron is truth-matched with a jet, the event is classified as V+c-jets.

Background normalisation: All the backgrounds are subjected to theoretical uncertainties on their respective cross-sections, and therefore on their normalisation. A summary of the values for the background normalisation uncertainties considered for each background process is given in Table 5.9.

5.7 Results

The possible presence of signal is tested employing a profile likelihood fit to the observed number of data events in the control regions and to the XGBoost distribution in the signal region. This fit is performed taking into account all the systematic uncertainties described in Section 5.6 as additional NPs in the likelihood expression. Both the resonant and the non-resonant signal model scenarios are considered. In case there is no significant event excess in the signal region, the

[10]The matching is performed using an improved CKKW procedure [373, 374].

Background	Uncertainty		Ref.
$t\bar{t}$	$+5.6\%$	-6.1%	[105, 420, 421]
t-channel (t)	$+4.0\%$	-3.4%	[409]
t-channel ($\bar{t}$)	$+5.0\%$	-4.5%	[409]
tW-channel	$+5.3\%$	-5.3%	[109]
s-channel	$+3.6\%$	-3.1%	[410]
W+jets	$+5\%$	-5%	[422]
Z+jets	$+5\%$	-5%	[422]
Diboson	$+6\%$	-6%	[422]
$t\bar{t}V$	1.4%	-1.4%	[423]

Table 5.9: Summary of the normalisation uncertainties on the different background processes.

results of the fit are interpreted to set upper limits on the production cross-section for the considered signal models.

5.7.1 Profile likelihood fit

The fitted distribution in the signal region is the score of the XGBoost classifier. The expected number of events in the i-th bin of this distribution is given by:

$$\mathbb{E}[n_i](\mu, \boldsymbol{\theta}) = \mu s_i(\boldsymbol{\theta}) + b_i(\boldsymbol{\theta}), \tag{5.1}$$

where n_i is the observed number of events in the i-th bin, s_i and b_i are the expected number of signal and background events, $\boldsymbol{\theta}$ are the NPs, including the systematic uncertainties which may affect the total number of expected yields, and μ is the signal strength. The compatibility of the observed number of data events with the expected data can be evaluated from two different hypotheses: the signal-plus-background hypothesis ($\mu = 1$) and the background-only hypothesis ($\mu = 0$).

In this analysis, hypothesis testing is performed using a frequentist approach within the RooStats framework [424]. This uses the asymptotic approximation described in Ref. [425]. The binning of the XGBoost distribution in the signal region is chosen to enhance as much as possible the signal sensitivity while keeping

enough statistics in the most significant bins. The chosen binning configuration is $\{0.5, 0.7, 0.85, 0.95, 1\}$ for the XGBoost classifier score.

In addition to the bins in the signal region, the control regions are included in the fit to constrain the normalisation of the main contributing backgrounds ($t\bar{t}$ and V+jets). These are included as two additional bins (one per region), with the total number of observed events in each region. The resulting likelihood function is the product of Poisson probabilities in all the fitted bins:

$$L(\mu, \boldsymbol{\theta}) = \prod_{j=1}^{N} \frac{(\mu s_j + b_j)^{n_j}}{n_j!} e^{-(\mu s_j + b_j)} \prod_{k=1}^{M} \frac{u_k^{m_k}}{m_k!} e^{-u_k}, \tag{5.2}$$

where u_k and m_k are the expected and observed number of events in the additional bins k, respectively. In this case, $N = 4$ for the two considered models, and $M = 2$ since two control regions are considered.

The NPs $\boldsymbol{\theta}$ are included in the fit using gaussian terms with $\sigma = 1$ and floating mean ν, with expected value $\nu = 0$. A value of a particular NP $\theta_i = \pm 1$ corresponds to a $\pm 1\sigma$ variation. Within this formalism, if the fitted systematic has a central value different from zero, it is said to be ***pulled***. Similarly, if the obtained uncertainty on the fitted systematic is lower than one, the systematic is said to be ***constrained***.

In order to test a particular value of an hypothesis μ, the profile likelihood ratio is built:

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}})}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})}, \tag{5.3}$$

where $\hat{\hat{\boldsymbol{\theta}}}$ determines the value of θ that maximises the likelihood for the specified μ. Thus, $L(\mu, \hat{\hat{\boldsymbol{\theta}}})$ is usually known as the conditional maximum-likelihood (ML) estimator of μ (and depends on the particular tested value of μ). The denominator is the maximised (unconditional) likelihood function, and $\hat{\mu}$ and $\hat{\theta}$ are their ML estimators. From this expression it can be seen that $0 \leq \lambda \leq 1$, with λ near 1 implying good agreement between data and the hypothesized value of μ.

The following test statistic can be built from the profile likelihood ratio:

$$q_\mu = \begin{cases} -2\ln\lambda(\mu) & \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases} \tag{5.4}$$

This is the test statistic used in this analysis to set upper limits on the signal strength. The reason is that $q_\mu = 0$ when $\hat{\mu} > \mu$ since data with $\hat{\mu} > \mu$ represents less compatibility with μ than the data obtained, and thus it is not taken as part of the rejection region of the test [425].

In order to quantify the level of agreement of the observed data with a particular hypothesis (a particular value of μ) the p-value is used. This quantity refers to the probability, under a particular assumption on μ, of finding data of equal or greater incompatibility with the predictions. This p-value can be computed as:

$$p_\mu = \int_{q_\mu}^{\infty} f(q_\mu|\mu) dq_\mu, \tag{5.5}$$

where $f(q_\mu|\mu)$ denotes the probability distribution function of the statistic q_μ under the assumption of an hypothesis on μ. This can be approximated by [425]:

$$p_\mu = 1 - \Phi(\sqrt{q_\mu}), \tag{5.6}$$

where Φ is the cumulative distribution of $f(q_\mu|\mu)$. If the p-value is below a particular threshold α the obtained value of μ is said to be excluded at a C.L. $(1-\alpha)$. Then, at a 95% C.L., the upper limit on μ is given by:

$$\mu_{up} = \hat{\mu} + \sigma\Phi^{-1}(1 - 0.05), \tag{5.7}$$

where σ is the standard deviation of $\hat{\mu}$. The upper limit on the signal strength can be transformed into an upper limit on the production cross-section of the considered signal process:

$$\sigma^{obs}_{95\%C.L.} = \mu_{up} \times \sigma_{theory}, \tag{5.8}$$

where σ_{theory} is the predicted cross-section from theory. In the absence of signal, the results of this analysis are interpreted to provide exclusion limits on the cross-section for the two signal models considered, the resonant and the non-resonant. The exclusion limits are provided as a function of different masses of the mediator for each model.

5.7.1.1 Fit setup

The focus of this section is to describe the treatment of the systematic model via the inclusion of NPs in the fit. For each source of systematic uncertainty "up" and "down" variations are usually available, evaluated by varying the specific source by one standard deviation in opposite directions. The corresponding effect on the distributions can then be evaluated. These uncertainties are symmetrised before entering the fit to avoid bad behavior of the fit associated to asymmetric errors. In case of a two-sided variation, the "up" and "down" variations are computed as:

$$\text{up(down)} = \text{nominal} + (-)\frac{\text{up} - \text{down}}{2}$$

In case of one-sided systematics, the variation is defined to be the "up" variation. Then, the "down" variation is simply taken as:

$$\text{down} = \text{nominal} - (\text{up} - \text{nominal})$$

In this analysis, the only systematic uncertainties that are treated as one-sided are the uncertainties on the $E_{\text{T}}^{\text{miss}}$ resolution in the perpendicular and parallel plane to the beam axis, and the modelling uncertainties associated with the choice of parton shower model, the matrix element generator, and the tW overlap removal scheme. All the remaining NPs are treated as two-sided uncertainties.

A pruning procedure is applied to reject those systematic uncertainties which have a negligible impact in the final prediction. This improves the stability of the fit. The impact of each uncertainty on the total background distribution is evaluated in terms of normalisation and shape effect. The normalisation is evaluated by integrating over the nominal distribution and the varied ones. The shape impact

consist of the maximum difference of bin entries between the nominal and the varied distributions due to a given systematic effect, once the normalisation is set to be the same. In case the shape or normalisation impacts are below 1%, the corresponding component of the uncertainty is removed from the systematics model.

5.7.2 Fit of the background model to data in control-regions

Two scale factors are associated to the two main sources of background in this analysis: μ_{top}, associated to the $t\bar{t}$ normalisation, and $\mu_{V\text{+jets}}$, associated to the normalisation of vector boson production in association with jets. These are estimated from a fit to the total number of observed data events in the corresponding control regions under the background-only hypothesis. In this fit, the two scale factors are included as free-floating parameters and the systematic uncertainties are included as NPs. The resulting normalisation factors are propagated to the control and validation regions to evaluate the modelling of the SM backgrounds.

The measured scale factors are shown in Table 5.10. Both values are compatible with unity within uncertainties. The results of the fit are evaluated in the associated control and validation regions, as shown in Section 5.5. Figures 5.12-5.16 showed that there is a good modelling of the backgrounds and no significant shape discrepancy between the data and the SM predictions is observed.

Parameter	Value
μ_{top}	0.85 ± 0.19
$\mu_{V\text{+jets}}$	1.0 ± 0.4

Table 5.10: Measured scale factors for $t\bar{t}$ (μ_{top}) and V+jets ($\mu_{\text{V+jets}}$) from the background-only fit to the data yields in the control regions. The uncertainties include both statistical and systematic effects.

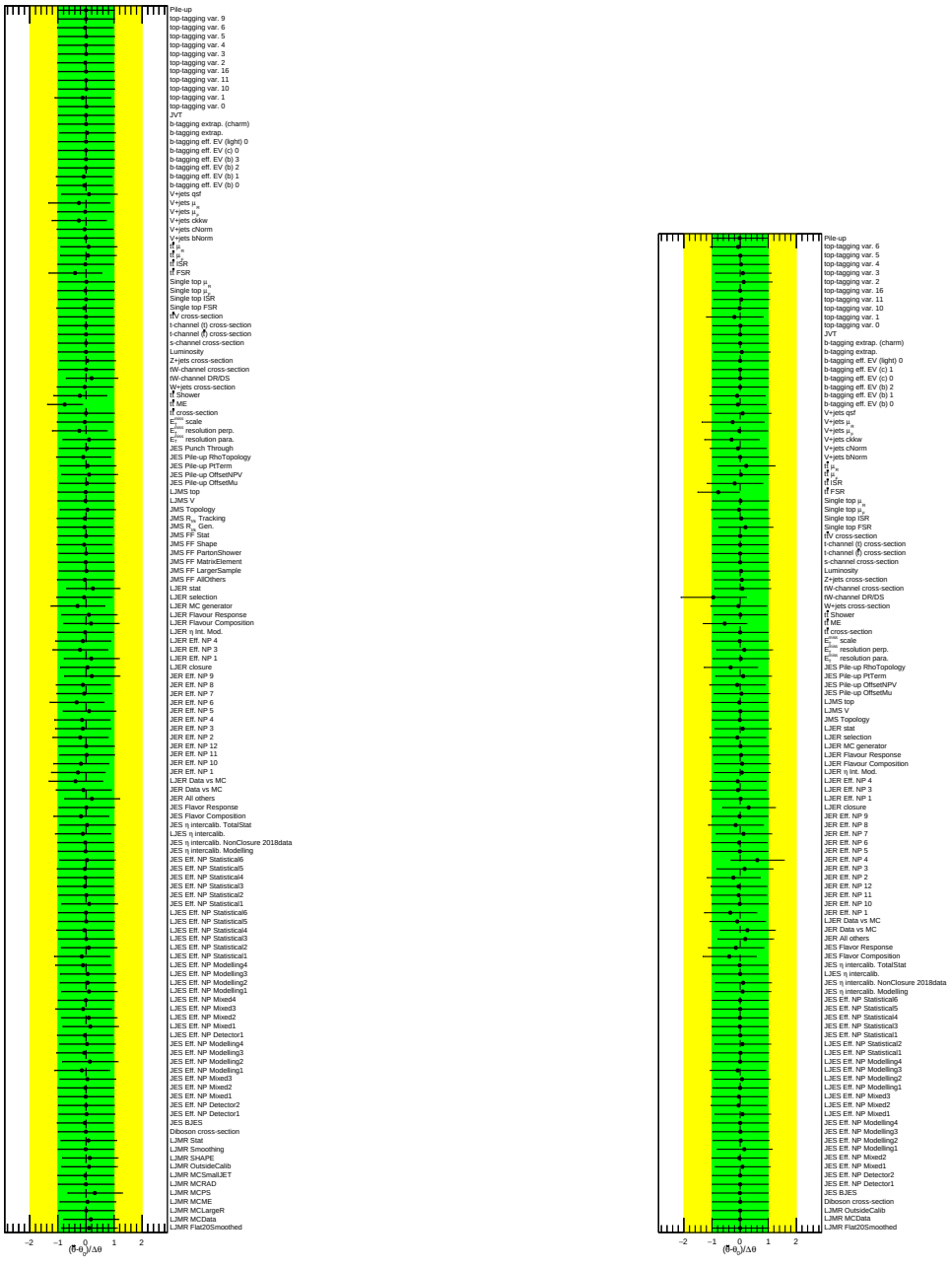
5.7.3 Results of the maximum likelihood fit to data in the signal and control regions under the background-only hypothesis

The compatibility of the observed data with the SM predictions is first evaluated through a simultaneous fit to the yields in the control regions and to the XGBoost distribution in the signal region associated to either the resonant or the non-resonant case under the background-only hypothesis. This fit includes all systematic uncertainties as NPs and has two free parameters: the $t\bar{t}$ and V+jets normalisation factors.

The pulls and constraints of the NPs are displayed in Figure 5.19. The most significant constraints and pulls are observed in the $t\bar{t}$ and single-top modelling uncertainties. The choice of $t\bar{t}$ matrix element generator and the final-state radiation NPs are slightly pulled ($\sim$0.5σ), and the choice of $t\bar{t}$ matrix element generator shows also a sizeable constraint for the resonant case. The sensitivity of the fit to these NPs comes from the fact that their variations have a significant effect on the shape of the XGBoost distribution. In addition, the NP associated to the choice of the tW-channel DR/DS scheme shows a 1σ pull in the non-resonant case. This uncertainty has a big impact in the number of tW-channel yields in the last two bins of the signal region distributions. This trend is more important in the non-resonant case since the shape of this uncertainty follows closely the shape of the observed data, leading to the observed pull. All other uncertainties over the pruning threshold do not show significant constraints nor pulls on the signal strength.

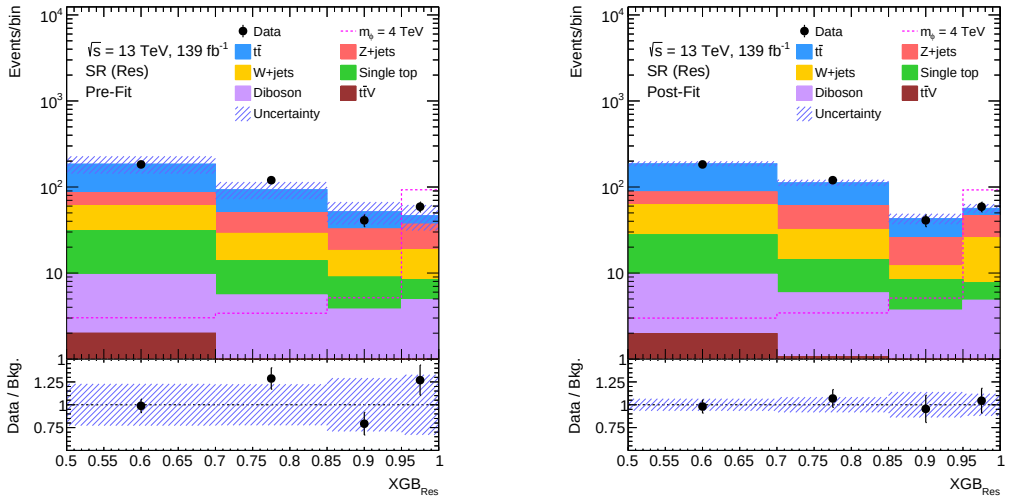
The measured scale factors for the $t\bar{t}$ and V+jets under the background-only hypothesis are shown in Table 5.11 for the signal regions targeting the resonant and non-resonant models. The observed scale factors are compatible with those from the fit in only the control regions within uncertainties.

The pre-fit and post-fit distributions under the background-only hypothesis in the signal region for the XGBoost discriminant are shown in Figures 5.20-5.21 for the resonant and the non-resonant model, respectively. The post-fit p-value obtained for the resonant (non-resonant) case is 0.98 (0.91). No significant excess above the SM is found in the signal region.



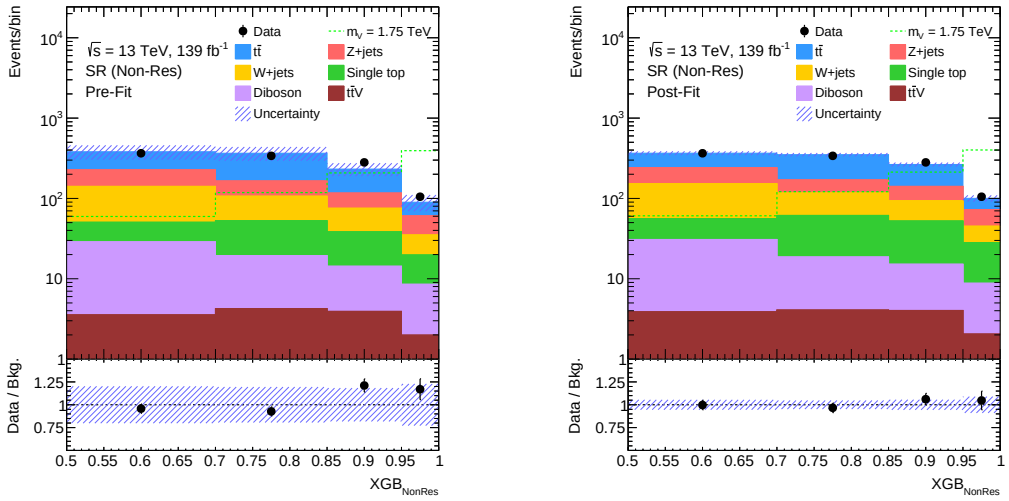
(a) (b)

Figure 5.19: Post-fit systematic uncertainties obtained from the fit under the background-only hypothesis in the control regions and the signal region for (a) the resonant and (b) the non-resonant model. The green (yellow) area represent the $\pm 1(2)\sigma$ band on the pre-fit systematic uncertainty. The black points show the best-fit values of the nuisance parameters, with the error bars representing the post-fit uncertainties. Each nuisance parameter is shown relative to its nominal value, θ_0, and in units of its pre-fit uncertainty.



(a) (b)

Figure 5.20: Distribution of the XGBoost score, trained on the resonant model, in the DM resonant signal region (a) pre- and (b) post-fit. The fit is performed on the observed data under the background-only hypothesis in the control regions and the signal region. The uncertainty bands include all systematic sources. The lower panel shows the ratio of data to SM prediction.



(a) (b)

Figure 5.21: Distribution of the XGBoost score, trained on the non-resonant model, in the DM non-resonant signal region (a) pre- and (b) post-fit. The fit is performed on the observed data under the background-only hypothesis in the control regions and the signal region. The uncertainty bands include all systematic sources. The lower panel shows the ratio of data to SM prediction.

Parameter	Resonant	Non-resonant
μ_{top}	0.88 ± 0.20	0.90 ± 0.19
$\mu_{V\text{+jets}}$	1.2 ± 0.5	1.2 ± 0.5

Table 5.11: Measured scale factors for the $t\bar{t}$ (μ_{top}) and V+jets ($\mu_{V\text{+jets}}$) backgrounds obtained from the data fit of the background-only model in the control regions and the signal region for the resonant and the non-resonant models. The uncertainties include both statistical and systematic effects.

5.7.4 Limits on dark matter production cross-section

In the absence of a significant event excess, a simultaneous fit is performed to the yields of the control regions and to the XGBoost distribution in the signal region under the signal-plus-background hypothesis to extract upper limits on the production cross-section for both the DM resonant and the DM non-resonant models.

The XGBoost classifier was trained using two particular signal models close to the current exclusion limits, specifically with the parameters $\lambda = 0.6$, $y = 0.4$, $m_\chi = 1$ GeV and $m_\phi = 4$ TeV ($a = 0.5$, $g_\chi = 1$, $m_\chi = 1$ GeV and $m_V = 1.75$ TeV) for the resonant (non-resonant) model. Therefore, this set of parameters are used here as reference for the fit performance in each case. No significant difference is observed in the fits with other model parameters.

The correlation matrices between the NPs, the $t\bar{t}$ and V+jets scale factors and the signal strength are given in Figures 5.22-5.23 for the resonant and the non-resonant model, respectively. In both cases, the V+jets normalisation factor is strongly correlated ($\sim$90%) with the uncertainty on the V+jets renormalisation scale μ_R. This is expected since the uncertainty on μ_R has a significantly large normalisation component (30-40% for this background process).

For the non-resonant case, a $\sim$35% correlation is observed between the signal strength and the uncertainty on the choice of tW-channel DR/DS scheme. This is understood since this systematic has a big impact in the number of tW-channel events in the last two bins of the XGBoost distribution, which contain almost all signal events, and its shape follows closely the shape of the observed data. A $\sim$35%

correlation is also observed between the signal strength and one LJER NP, as it implies a $\sim$30% variation in the number of W+jets events in the last bin of the signal region.

The observed correlations between the signal strength and the NPs are lower for the resonant case. A $\sim$20% correlation is observed between the signal strength and one JER NP since this uncertainty implies a $\sim$20% variation in the number of W+jets events in the last bin of the signal region. A $\sim$20% correlation is also observed between the signal strength and the $t\bar{t}$ PS NP, due to the large normalisation and shape effect of this uncertainty in the $t\bar{t}$ background.

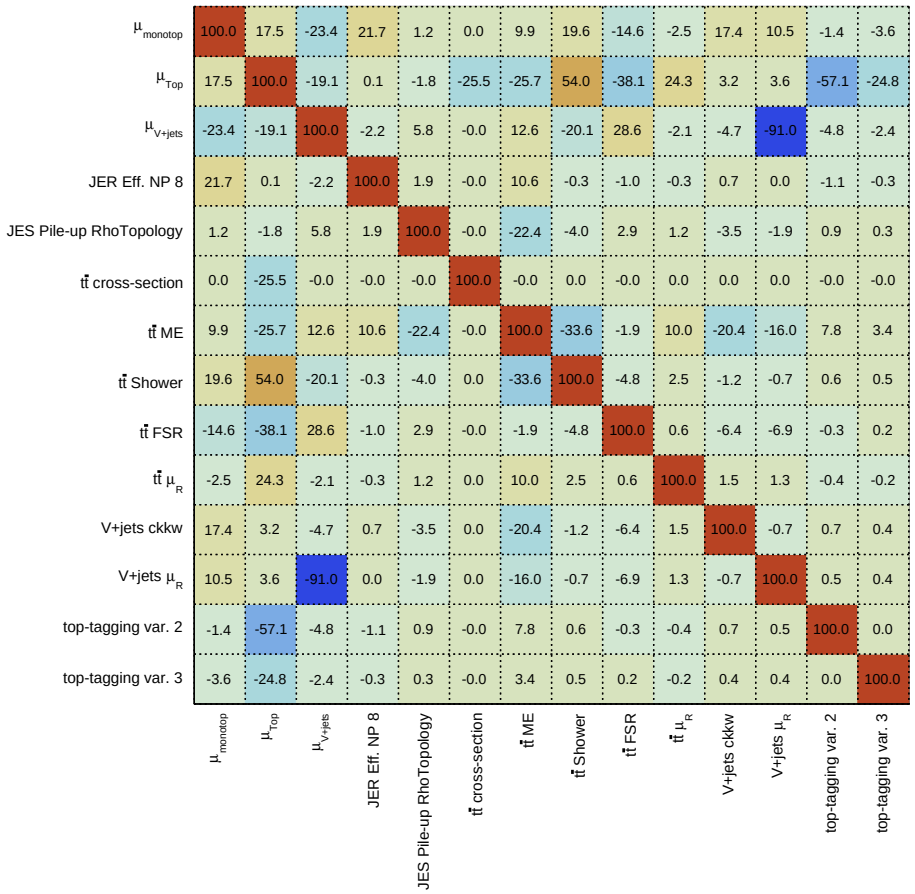


Figure 5.22: Correlation matrix of the free parameters and the NPs included in the data fit in the control regions and the DM resonant signal region under the signal-plus-background hypothesis. Only NPs with a correlation above 20% are displayed. The fit is performed using $\lambda = 0.6$, $y = 0.4$, $m_\chi = 1$ GeV and $m_\phi = 4$ TeV as parameters of the model.

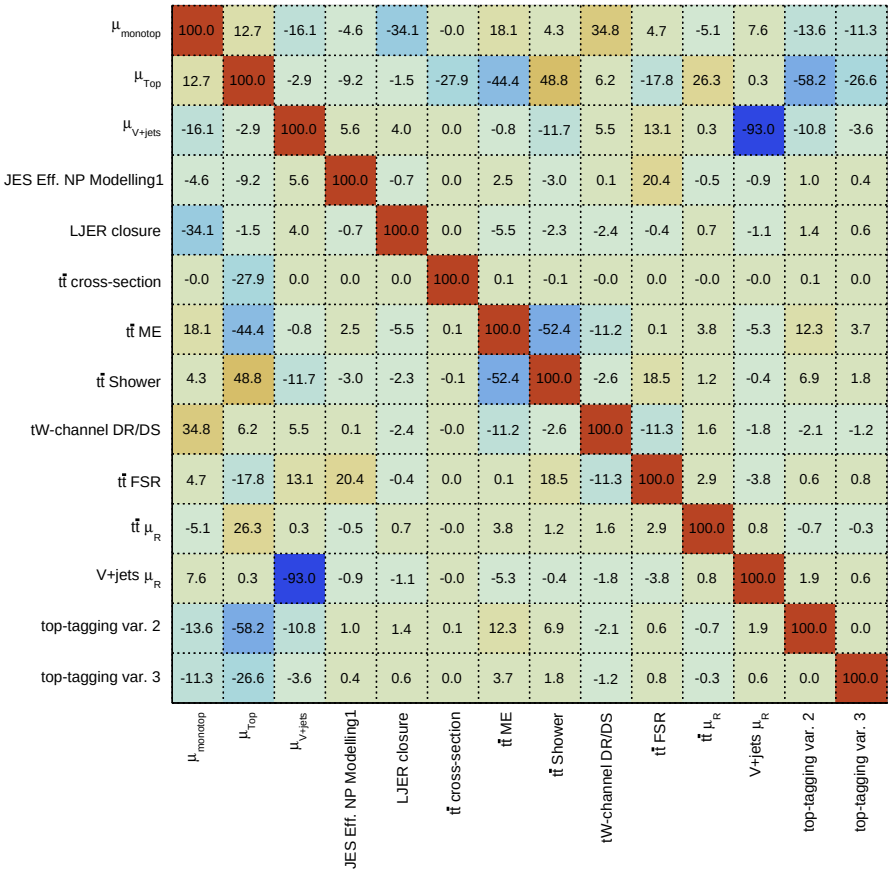


Figure 5.23: Correlation matrix of the free parameters and the NPs included in the data fit in the control regions and the DM non-resonant signal region under the signal-plus-background hypothesis. Only NPs with a correlation above 20% are displayed. The fit is performed using $a = 0.5$, $g_\chi = 1$, $m_\chi = 1$ GeV and $m_V = 1.75$ TeV as parameters of the model.

Figure 5.24 shows the impact of the most important systematic uncertainties on the fitted value of the signal strength, before and after the associated NPs are fitted to the data (empty and full rectangles, respectively). For each NP, the shift in the signal

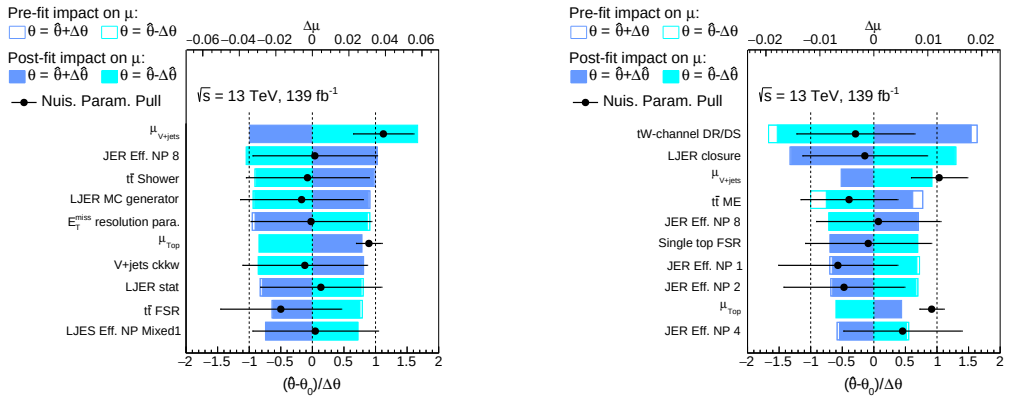


Figure 5.24: The nuisance parameters ranked according to their post-fit impacts on the best-fit value of μ. Only the 10 nuisance parameters with the largest impacts are shown. The empty (solid) blue rectangles illustrate the pre-fit (post-fit) impacts on the parameter of interest μ, corresponding to the top axis. The pre-fit (post-fit) impact of each nuisance parameter, $\Delta\mu$, is calculated as the difference in the fitted value of μ between the nominal fit and the fit when fixing the corresponding nuisance parameter to $\hat{\theta} \pm \Delta\theta$ ($\hat{\theta} \pm \Delta\hat{\theta}$), where $\hat{\theta}$ is the best-fit value of the nuisance parameter and $\Delta\theta$ ($\Delta\hat{\theta}$) is its pre-fit (post-fit) uncertainty. The black points show the best-fit values of the nuisance parameters, with the error bars representing the post-fit uncertainties. Each nuisance parameter is shown relative to its nominal value, θ_0, and in units of its pre-fit uncertainty.

strength is obtained for a fit with the NP fixed to its pre-fit or post-fit $\pm 1\sigma$ values. The uncertainty from the fitted value of the $t\bar{t}$ and V+jets normalisation factors is also displayed. The leading sources of uncertainty come from those parameters which show the largest correlations with the signal strength in each case. For the DM resonant model, the biggest impact on the signal strength comes from the uncertainty on the fitted value of the V+jets normalisation factor, which showed a $\sim$20% correlation with the signal strength. This is followed by the systematic uncertainties on the JER and the choice of $t\bar{t}$ parton shower model. For the DM non-resonant model, the leading sources of uncertainty come from the diagram scheme employed in the tW-channel simulation and the LJER uncertainties.

The fit is performed testing different signal samples, corresponding to different values of the mediator masses m_ϕ (m_V) for the resonant (non-resonant) model. The expected and observed limits on the signal cross-section are computed and displayed in Figures 5.25-5.26. The dashed (solid) black line depicts the expected (observed) exclusion limit, whereas the green (yellow) band indicates the 1σ (2σ) standard deviations. The red line is the cross-section obtained from theoretical predictions. From the intersection of the exclusion band with the theoretical cross-section, the resonant (non-resonant) model is excluded for masses of the mediator below 4.6 (2.6) TeV. The upper limits on the cross-section set for the different masses of the mediator for each model are summarised in Tables 5.12-5.13. These are dominated by the systematic uncertainties. The limits on the mediator mass of the resonant (non-resonant) model are improved by $\sim$1.2 ($\sim$0.6) TeV with respect to the previous analyses [241–243].

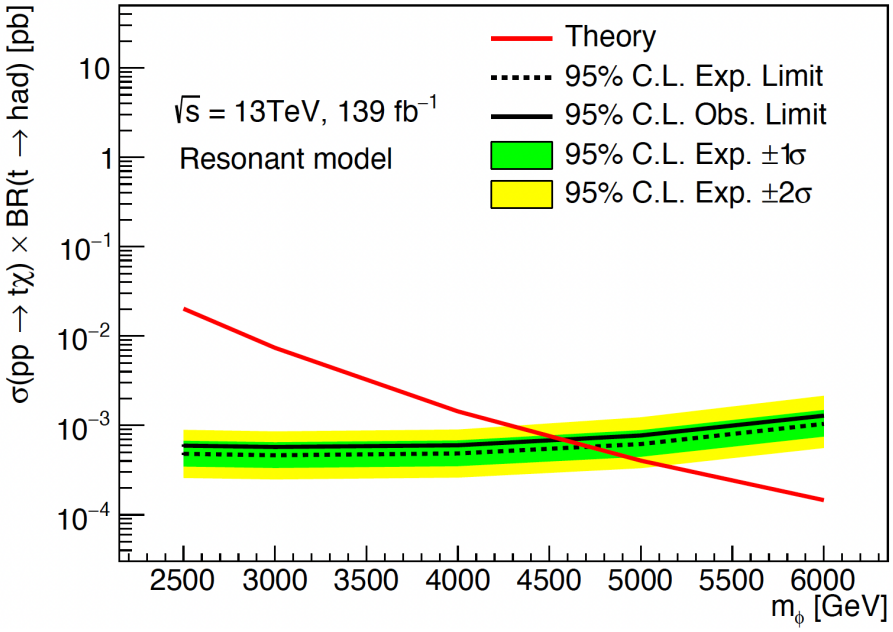


Figure 5.25: Observed and expected limit on the DM resonant cross-section as a function of the mediator mass m_ϕ. The red line in the plot indicates the theoretical cross-section. The mass of the DM particle is set to $m_\chi = 1$ GeV, the coupling constant between the scalar mediator, the top quark and the initial state d- and s-quarks is set to $\lambda = 0.6$ and the mixing parameter between the scalar mediator, the top quark and the DM particle is set to $y = 0.4$.

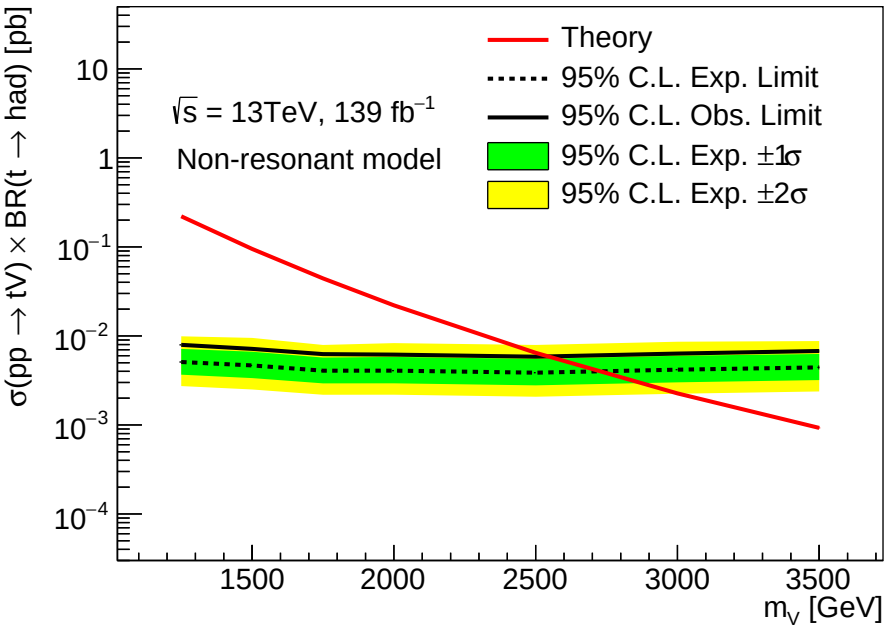


Figure 5.26: Observed and expected limit on the DM non-resonant cross-section as a function of the mediator mass m_V. The red line in the plot indicates the theoretical cross-section. The mass of the DM particle is set to $m_\chi = 1$ GeV, the coupling between the vector mediator, the initial state up-type quark and the top quark is set to $a = 0.5$ and the coupling between the vector mediator and the DM particle is set to $g_\chi = 1$.

m_ϕ [GeV]	Theory, $\sigma(pp \to t\chi)$ $\times$ BR$(t \to qqb)$ [pb]	Expected 95% C.L. limit [pb]	Observed 95% C.L. limit [pb]
2500	0.020218	0.000480	0.000593
3000	0.007374	0.000463	0.000573
4000	0.001433	0.000485	0.000599
5000	0.000403	0.000618	0.000771
6000	0.000146	0.001040	0.001284

Table 5.12: Expected and observed 95% C.L. upper limits on the production cross-section of the resonant model as a function of the mediator mass m_ϕ, compared to LO theoretical predictions in the hadronic channel. The mass of the DM particle is set to $m_\chi = 1$ GeV, the coupling constant between the scalar mediator, the top quark and the initial state d- and s-quarks is set to λ = 0.4 and the mixing parameter between the scalar mediator, the top quark and the DM particle is set to y = 0.4.

m_V [GeV]	Theory, $\sigma(pp \to t\chi)$ $\times$ BR$(t \to qqb)$ [pb]	Expected 95% C.L. limit [pb]	Observed 95% C.L. limit [pb]
1250	0.220800	0.005093	0.007939
1500	0.094930	0.004664	0.007155
1750	0.044280	0.004068	0.006246
2000	0.022070	0.004074	0.006167
2500	0.006463	0.003865	0.005844
3000	0.002262	0.004178	0.006336
3500	0.000926	0.004431	0.006772

Table 5.13: Expected and observed 95% C.L. upper limits on the production cross-section of the non-resonant model as a function of the mediator mass m_V, compared to LO theoretical predictions in the hadronic channel. The mass of the DM particle is set to $m_\chi = 1$ GeV, the coupling between the vector mediator, the initial state up-type quark and the top quark is set to a = 0.5 and the coupling between the vector mediator and the DM particle is set to g_χ = 1.

6 | Conclusions

The SM of particle physics is a very successful theory. It predicted well in advance the existence of many particles, such as the Higgs boson or the top quark. It is also very predictive in terms of the interactions between the elementary particles. However, it cannot be considered the ultimate theory. There are still important mysteries that are not explained by the SM, such as the nature of DM, the asymmetry between matter and antimatter observed in the universe and the observed non-zero neutrino masses. The SM does not describe either the gravitational interaction.

The top quark is a crucial particle in the SM. Due to its large mass, it is the only quark that decays before hadronising. Thus, one can study the properties of the top quark through its decay products. In addition, it has the highest Yukawa coupling with the Higgs boson. This suggests that the top quark may play a special role in spontaneous symmetry breaking. Furthermore, many beyond the SM models predict the existence of new particles which would couple preferentially to top quarks.

At the LHC there is a strong effort towards understanding the open questions unexplained by the SM. The LHC physics programme aims to find hints of new physics using two well differentiated strategies: high precision measurements of the SM processes, looking for deviations from the SM predictions which could indicate signs of new physics; and direct searches, in particular in the form of new particles such as DM candidates. The work in this thesis contributes to these two strategies through two analysis in the top-quark sector: a high precision measurement of observables sensitive to the top-quark polarisation, and a search for DM particles

produced in association with a single top quark, leading to a signature in the detector consisting on a top quark and missing transverse momentum.

The data analysed in this thesis corresponds to pp collisions recorded by the ATLAS detector during the Run-2 (2015-2018) at $\sqrt{s} = 13$ TeV. This yields a total integrated luminosity of 139.0 fb^{-1}. However, in order to perform high precision physics measurements, it is crucial to reach a deep understanding of the detector performance. The work on this thesis also contributed to this effort through the alignment of the ATLAS inner detector. The alignment aims to provide a precise estimation of the position and orientation of the different elements that constitute the inner detector, and to follow its possible changes in time. These movements can occur from year-to-year, run-to-run, or even within the same run. There are two characteristic movements of the inner detector within each run: the bowing of the IBL innermost layer of the inner detector and the pixel vertical displacement. The movements of the pixel detector have been studied in detail in this thesis, and a clear relation has been found between the vertical drift and the luminosity.

These two movements are corrected within the calibration loop. The calibration loop corrections are always applied on top of a set of baseline alignment constants. It is crucial that these constants describe as accurately as possible the real geometry of the detector, because the calibration loop alignment cannot correct by itself very drastic changes of the detector geometry to the desired level of precision, nor track parameter biases which leave the χ^2 function of the alignment algorithm unaltered. The derivation of baseline alignment constants for all the heavy ions 2015, heavy ions 2016 and pp 2018 data-taking periods has been another important contribution of this thesis to the ID alignment.

In the t-channel electroweak production at LO single top quarks are produced with their spin aligned along the direction of the down-type quarks. This is a consequence of the vector-axial structure of the tWb vertex, which is involved in the production and the decay of single top quarks. Due to its short lifetime, the top-quark spin information is transferred to its decay products, in particular to the lepton in the top quark leptonic decay. Measurements of lepton angular observables are therefore sensitive to the top-quark polarisation and allow to probe new physics affecting the tWb vertex, including the presence of new sources of CP violation.

The first analysis performed in this thesis aims to measure the normalised differential cross-section of single-top-quark production in the t-channel, where the top quark decays leptonically, as a function of the three polarisation angles of the lepton. A cut-based event selection is performed in order to build a signal region enriched in t-channel signal events. The modelling of the main contributing backgrounds ($t\bar{t}$ and W+jets) is checked in control regions. All SM backgrounds have been modelled using MC simulation, except multijet. The multijet estimation relies on data-driven techniques. A simultaneous fit to the data yields is performed in the signal and the control regions to constrain the normalisation of the main contributing backgrounds. Then, the expected t-channel angular distributions at reconstruction level are obtained by subtracting the different backgrounds to the data.

These distributions are distorted by inefficiencies and acceptance of the detector, smeared by reconstruction procedures and sculpted by the event selection criteria. Thus, the angular distributions are unfolded to the particle level in a fiducial region to compare with theoretical predictions. This fiducial region is chosen as close as possible to the reconstruction level region, so that acceptance effects are minimum. The robustness of the unfolding technique has been validated through convergence, closure and linearity tests.

The unfolded distributions are divided by the integral over all bins and the results are presented as normalised differential cross-section measurements as a function of three angular distributions sensitive to the top-quark polarisation. The most important uncertainties in this analysis come from the jet energy resolution, since it affects strongly the reconstruction of the top-quark rest frame in which the angular distributions are measured. These measurements are in agreement with the SM predictions as provided by various Monte Carlo generators. The results of this study have been published in Ref. [3].

These results have been interpreted in an EFT context to provide strong constraints on the real part C_{tW} and the imaginary part C_{itW} of the dipole operator Wilson coefficient. This is part of the PhD thesis in Ref. [4]. The measured values of C_{tW} = 0.4 $\pm$ 1.1 and C_{itW} = -0.3 $\pm$ 0.4 at 95% C.L. are compatible with the SM predictions. These limits are comparable with the current limits [118,

129, 402–404] for the real part of the dipole operator, and improve slightly the marginalised limits [406] for the imaginary part, which is sensitive to CP violation components. Therefore, the limits on the imaginary part are the best limits so far from high-energy measurements.

The discovery of DM particles at the LHC laboratory would be a historic milestone in the world of particle physics. This would allow us to study the properties of one of the key constituents of the universe. Focusing on WIMP DM candidates that could be produced at the LHC pp collisions, they could be detected by ATLAS if they were produced in association with a SM object. The second analysis presented in this work was focused on the search for events with large missing transverse momentum (originating from the DM particles) in association with one top quark: a search for the so-called mono-top signature, which is predicted in many beyond the SM scenarios.

This analysis benefits from cutting-edge MVA techniques to enlarge the signal discrimination over the different backgrounds. Control regions are defined in a kinematic region as close as possible to the signal region to constrain the normalisation of the main contributing backgrounds ($t\bar{t}$ and V+jets).

The statistical interpretation of the results is based on a simultaneous fit to the signal and control regions. In this fit, the signal strength and normalisation factors for the main contributing backgrounds are the free parameters. The different systematic uncertainties are included as nuisance parameters in the fit. The results are interpreted in two contexts: the production of a DM signal from the resonant and the non-resonant models.

In the absence of any significant excess of data, the results are used to provide 95% C.L. upper limits on the production cross-section for both models. The observed (expected) mass limits for the scalar mediator mass m_ϕ are 4.6 (4.7) TeV for the resonant model, and 2.6 (2.7) TeV for the vector mediator mass m_V for the non-resonant model, with the chosen model parameters $m_\chi = 1$ GeV, λ = 0.6 and y = 0.4 ($m_\chi = 1$ GeV, a = 0.5 and g_χ = 1) for the resonant (non-resonant) case. The largest uncertainties in this analysis arise from the modelling of the most important background $t\bar{t}$, the method used to treat the overlap between the tW-channel and $t\bar{t}$ generated samples, and the uncertainties on the small-R and

large-R jet energy resolution. These results improve the existing limits [241–243] from previous searches and so far constitute the most stringent constraints on this DM production cross-section.

In conclusion, the work in this thesis has contributed to the LHC physics programme during Run-2 in the search for physics beyond the SM in the top-quark sector: both via a precision measurement of observables which are sensitive to the top-quark polarisation, and via a search for a single top quark in association with missing transverse momentum, in which DM particles are produced along with a top quark. Both measurements are compatible with the SM predictions and no evidence of new physics has been found. These results are used to constrain the parameter space in which new physics may hide. The results of the first analysis have been used to provide the most stringent limits in the imaginary part of the dipole operator Wilson coefficient from high-energy experiments, and the results of the second analysis constitute the most stringent limits in the DM production cross-section for the two models considered: the resonant the non-resonant models. Consequently, these measurements have contributed to the effort of finding an explanation for the matter-antimatter asymmetry of the universe in the form of CP violation in the top-quark sector, and of detecting a viable DM candidate in the laboratory.

Appendices

A Estimation of the uncertainties of the multijet background in the normalised differential cross-section measurements

A.1 Estimation of the normalisation uncertainty of the multijet background

As described in Section 4.4.1, the multijet background estimation is derived in two steps. Firstly, the shape is provided from both the jet-electron model (electron channel) and the anti-muon model (muon channel). Secondly, the multijet normalisation is estimated from a fit to the data. In this analysis a binned likelihood fit is performed on the measured distribution of the $E_{\mathrm{T}}^{\mathrm{miss}}$ (electron channel) and the $m_T(l, E_{\mathrm{T}}^{\mathrm{miss}})$ (muon channel) in specifically defined multijet-enriched regions as defined in Section 4.4.1. However, fitting a different variable in the electron or the muon channels can give different values for the multijet normalisations. In this appendix, the results from alternative fits in the multijet-enriched pre-selection region by using different variables are shown. The $\Delta\phi(l, E_{\mathrm{T}}^{\mathrm{miss}})$ (i.e. the azimuthal angle between the lepton and the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$) distribution is used in the muon channel for both the multijet-enriched pre-selection region and the multijet-enriched $t\bar{t}$ control region. The H_{T} distribution is used in the electron channel for the multijet-enriched pre-selection region whereas the $\Delta\phi(l, j_1)$ (i.e.

the azimuthal angle between the lepton and the leading jet) distribution is used for the multijet-enriched $t\bar{t}$ control region.

Also, the choice of MC generator for the main contributing background process (top-quark processes in this analysis) can lead to different results in the multijet normalisations. Here, the results from alternative fits in the multijet-enriched pre-selection region and $t\bar{t}$ control region using different MC generators for top-quark processes are also presented.

The comparison of these alternative fits with respect to the baseline fits shown in Section 4.4.1 provides an estimate of the multijet background normalisation uncertainty.

In these alternative fits, the multijet-enriched regions defined in Section 4.4.1 are also used. The full shapes of the new fitted variables in the multijet-enriched pre-selection region are shown in Figure A.1, whereas the full shapes of the new fitted variables in the multijet-enriched $t\bar{t}$ control region are shown in Figure A.2.

The $\Delta\phi(l, E_{\mathrm{T}}^{\mathrm{miss}})$ shape also showed a good discrimination power in the electron channel in the multijet-enriched $t\bar{t}$ control region (see Figure A.3). Thus, an alternative fit using the $\Delta\phi(l, E_{\mathrm{T}}^{\mathrm{miss}})$ in the electron channel and the $m_T(l, E_{\mathrm{T}}^{\mathrm{miss}})$ in the muon channel was also considered in the estimation of the multijet normalisation uncertainty. The obtained results were very similar to those obtained using the $\Delta\phi(l, j_1)$ in the electron channel instead, which implies a consistency check of this approach.

The choice of MC generator for the main contributing background (top-quark processes, including t-channel signal for these fits) has also an impact in the normalisation of the multijet background, particularly in the $t\bar{t}$ control region. Two alternatives are considered to compare with respect to the baseline Powheg-Box+Pythia8. Firstly, the baseline event generator is changed and MG5_aMC@NLO+Pythia8 is used. Secondly, the PS generator is modified and Powheg-Box+Herwig7 is employed. These comparisons are performed both in the multijet-enriched pre-selection region and in the $t\bar{t}$ control region. In these alternatives, the baseline distributions are used in the fit (i.e. the $E_{\mathrm{T}}^{\mathrm{miss}}$ in the electron channel and the $m_T(l, E_{\mathrm{T}}^{\mathrm{miss}})$ in the muon channel).

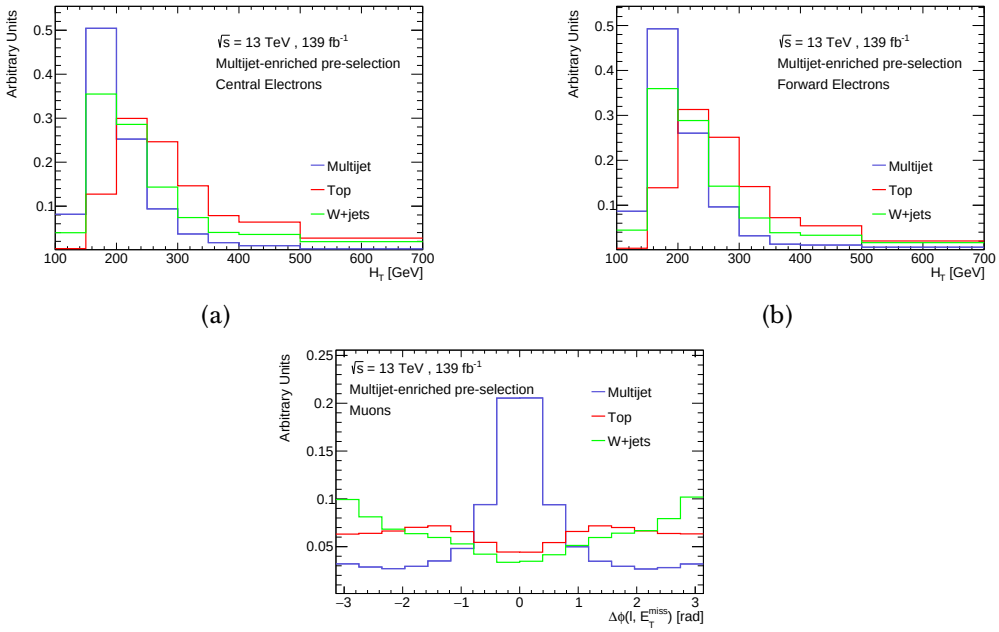


Figure A.1: Distributions of the H_{T} (electron channel) and the $\Delta\phi(l, E_{\mathrm{T}}^{\mathrm{miss}})$ (muon channel) in the multijet-enriched pre-selection for (a) the central electron channel, (b) the forward electron channel and (c) the muon channel. The distributions are normalised to the total number of events in order to compare the shapes of the MC templates. The three distributions show a good discrimination power between the multijet and the other fitted processes.

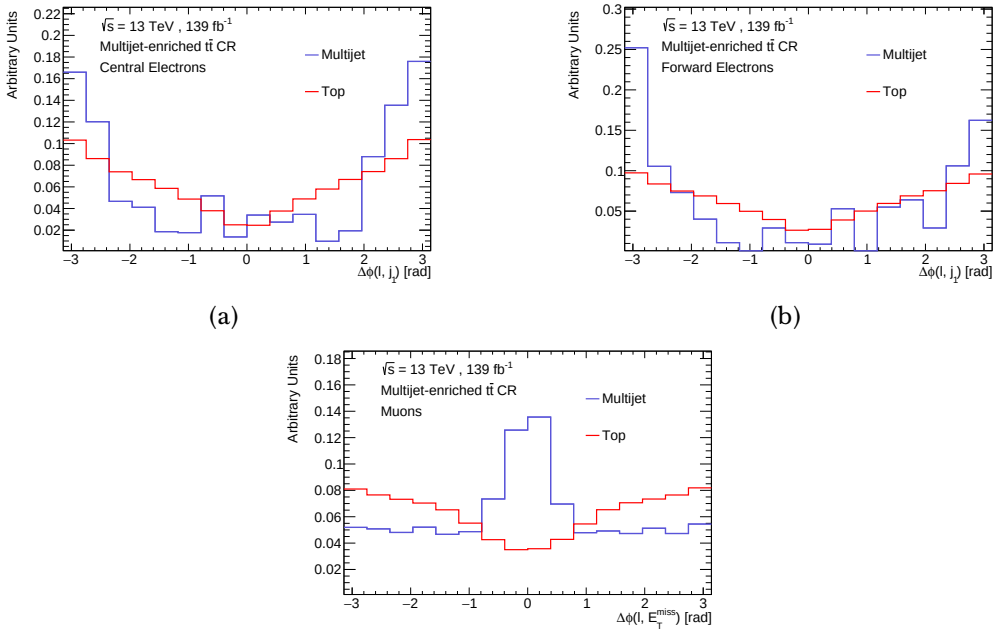
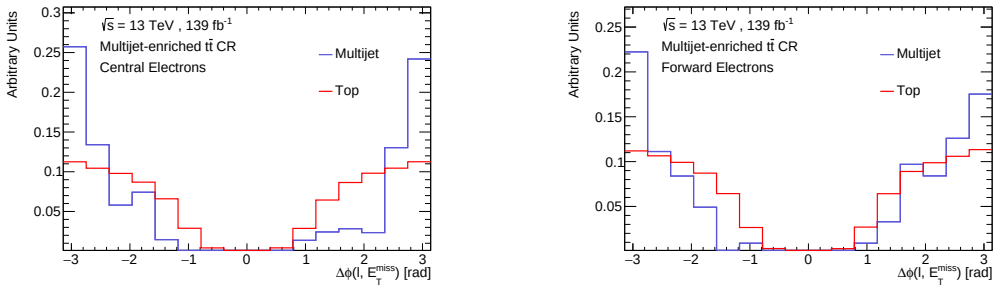


Figure A.2: Distributions of the $\Delta\phi(l, j_1)$ (electron channel) and the $\Delta\phi(l, E_{\mathrm{T}}^{\mathrm{miss}})$ (muon channel) in the multijet-enriched $t\bar{t}$ control region for (a) the central electron channel, (b) the forward electron channel and (c) the muon channel. The distributions are normalised to the total number of events in order to compare the shapes of the MC templates. The three distributions show a good discrimination power between the multijet and the other fitted processes.



(a) (b)

Figure A.3: Distributions of the $\Delta\phi(l, E_{\mathrm{T}}^{\mathrm{miss}})$ in the electron channel in the multijet-enriched $t\bar{t}$ control region for (a) the central electron channel and (b) the forward electron channel. The distributions are normalised to the total number of events in order to compare the shapes of the MC templates. The two distributions show a good discrimination power between the multijet and the other fitted processes.

The number of multijet events using the different approaches is extracted from the normalisation factors derived from the likelihood fit to the data distribution. These are reported in Tables A.1-A.2 for the different control regions and for the signal region. The choice of the fitted variable implies larger deviations in the multijet-enriched pre-selection fit, whereas the choice of MC generator implies larger deviations in the multijet-enriched $t\bar{t}$ control region. As it can be seen, a 20% (40%) multijet normalisation uncertainty covers the differences between the different alternatives in the signal and W+jets control region ($t\bar{t}$ control region).

Region	Preselection	W+jets CR	Signal region	$t\bar{t}$ CR
$E_{\mathrm{T}}^{\mathrm{miss}}$ (e), $m_T(l, E_{\mathrm{T}}^{\mathrm{miss}})$ (μ)	52000	48498	3503	5485
H_T (e), $m_T(l, E_{\mathrm{T}}^{\mathrm{miss}})$ (μ)	42790	39880	2908	-
$E_{\mathrm{T}}^{\mathrm{miss}}$ (e), $\Delta\phi(l, E_{\mathrm{T}}^{\mathrm{miss}})$ (μ)	49360	46170	3190	4660
$\Delta\Phi(l, j_1)$ (e), $m_T(l, E_{\mathrm{T}}^{\mathrm{miss}})$ (μ)	-	-	-	3577

Table A.1: Multijet event yields estimated with the jet-electron and anti-muon models using different variables in the fit to the data shape. These are given for the pre-selection, the $t\bar{t}$ control region, the W+jets control region and the signal region.

Region	Preselection	W+jets CR	Signal region	$t\bar{t}$ CR
POWHEG-BOX+PYTHIA8	52000	48498	3503	5485
POWHEG-BOX+HERWIG7	58586	54673	3913	3173
AMC@NLO+PYTHIA8	58563	54492	4070	4402

Table A.2: Multijet event yields estimated with the jet-electron and anti-muon models using different MC generators for top-quark processes in the fit to the data shape. These are given for the pre-selection, the $t\bar{t}$ control region, the W+jets control and the signal region.

Tables A.3-A.6 include the W+jets and $t\bar{t}$ normalisation factors extracted from the fit using the different methods.

	Multijet-enriched pre-selection	
Fitted variable	$H_{\mathrm{T}}(e)$, $m_T(l, E_{\mathrm{T}}^{\mathrm{miss}})(\mu)$	$E_{\mathrm{T}}^{\mathrm{miss}}(e)$, $\Delta\phi(l, E_{\mathrm{T}}^{\mathrm{miss}})(\mu)$
W+jets	1.328 ± 0.010	1.294 ± 0.011
$t\bar{t}$, tW, s, t-channel	0.933 ± 0.005	0.946 ± 0.006

Table A.3: Normalisation factors extracted for the W+jets and merged top-quark contributions from the maximum-likelihood fit of the distributions of the different variables in the multijet-enriched pre-selection. The uncertainties correspond to the statistical uncertainties provided from the likelihood fit.

	Multijet-enriched $t\bar{t}$ control region	
Fitted variable	$\Delta\phi(l, j_1)(e)$, $m_T(l, E_{\mathrm{T}}^{\mathrm{miss}})(\mu)$	$E_{\mathrm{T}}^{\mathrm{miss}}(e)$, $\Delta\phi(l, E_{\mathrm{T}}^{\mathrm{miss}})(\mu)$
$t\bar{t}$, tW, s, t-channel	1.035 ± 0.005	1.026 ± 0.004

Table A.4: Normalisation factors extracted for the merged top-quark contributions from the maximum-likelihood fit of the distributions of the different variables in the multijet-enriched $t\bar{t}$ control region. The uncertainties correspond to the statistical uncertainties provided from the likelihood fit.

	Multijet-enriched pre-selection	
MC generator	POWHEG-BOX+HERWIG7	aMC@NLO+PYTHIA8
W+jets	1.233 ± 0.011	1.292 ± 0.010
$t\bar{t}$, tW, s, t-channel	1.018 ± 0.006	0.955 ± 0.006

Table A.5: Normalisation factors extracted for the W+jets and merged top-quark contributions from the maximum-likelihood fit using different MC generators for the top-quark processes in the multijet-enriched pre-selection. The uncertainties correspond to the statistical uncertainties provided from the likelihood fit.

	Multijet-enriched $t\bar{t}$ control region	
MC generator	POWHEG-BOX+HERWIG7	aMC@NLO+PYTHIA8
$t\bar{t}$, tW, s, t-channel	1.121 ± 0.004	1.058 ± 0.005

Table A.6: Normalisation factors extracted for the merged top-quark contributions from the maximum-likelihood fit using different MC generators for the top-quark processes in the multijet-enriched $t\bar{t}$ control region. The uncertainties correspond to the statistical uncertainties provided from the likelihood fit.

A.2 Estimation of the shape uncertainty of the multijet background

A multijet shape uncertainty is also considered in this analysis. This is estimated by comparing modified templates to the baseline templates provided by the jet-electron and anti-muon models. The parameters that were changed for these alternative multijet samples are the fraction of the energy of the jet deposited in the EM calorimeter f_{EM} (jet-electron model) and the ratio of the sum of the transverse momenta of tracks within a maximum cone size of 0.4 around the muon to the p_{T} of the muon $\sum_i p_{\text{T},i}^{track}(\Delta R < 0.4)/p_{\text{T}}$ (anti-muon model). The variations of these parameters lead to subsets of the baseline configuration as summarised in Table A.7. Figure A.4 shows the modified multijet shapes for the three angular observables considered in this analysis.

Channel	Variable	Baseline	Subset
Electron	f_{EM}	[0.8, 1.0]	[0.9, 1.0]
Muon	$\sum_i p_{\mathrm{T},i}^{track}(\Delta R < 0.4)/p_{\mathrm{T}}$	[0.0, 0.1]	[0.0, 0.05]

Table A.7: Summary of the configurations used in each subset of the baseline setup to estimate the multijet shape uncertainty templates. The varied parameters are the fraction of the energy of the jet deposited in the EM calorimeter f_{EM} for the jet-electron model and the fraction between the sum of the transverse momenta of tracks within a maximum cone size of 0.4 around the muon and the p_{T} of the muon $\sum_i p_{\mathrm{T},i}^{track}(\Delta R < 0.4)/p_{\mathrm{T}}$ for the anti-muon model.

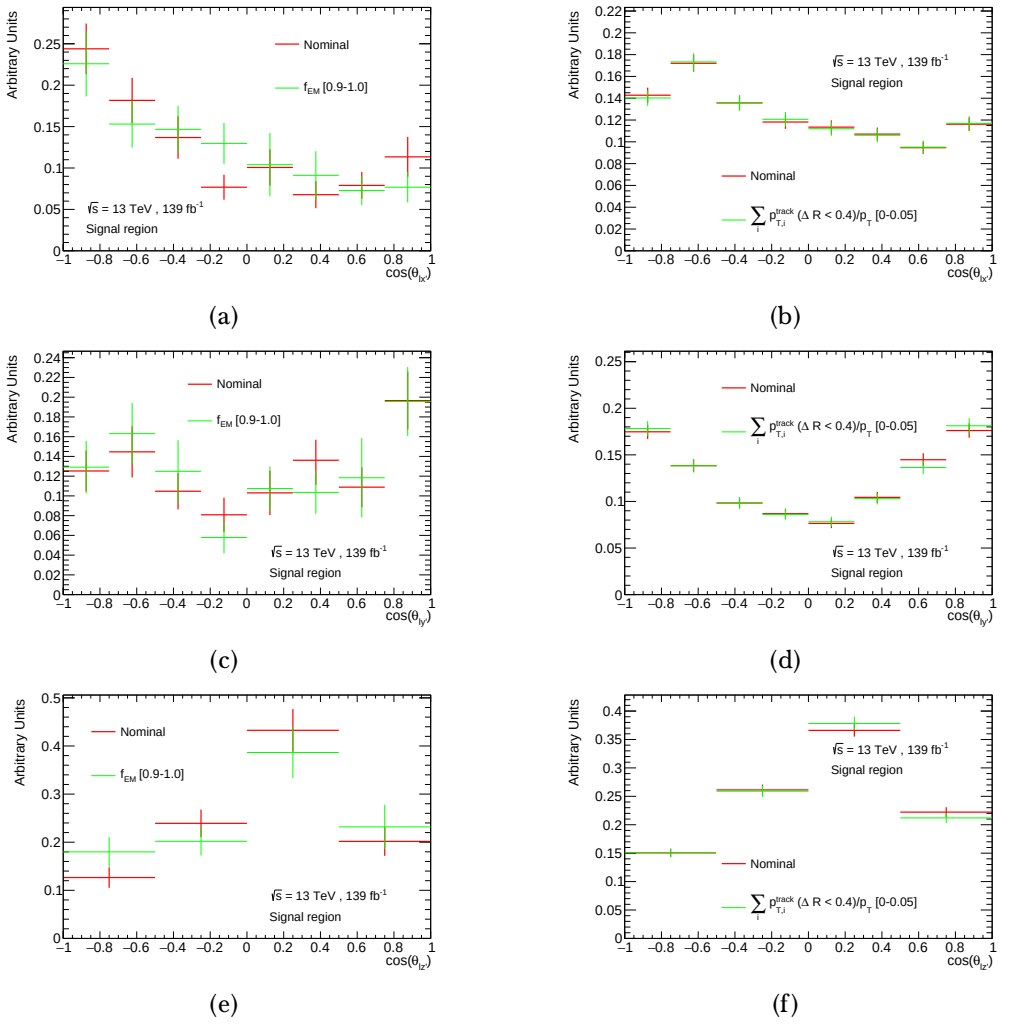


Figure A.4: Baseline and alternative shapes of the multijet background for the three angular distributions considered in this analysis (a) $\cos\theta_{lx'}$, (c) $\cos\theta_{ly'}$ and (e) $\cos\theta_{lz'}$ for the jet-electron model in the electron channel and (b) $\cos\theta_{lx'}$, (d) $\cos\theta_{ly'}$ and (f) $\cos\theta_{lz'}$ for the anti-muon model in the muon channel. The uncertainty bands include the statistical uncertainty due to the limited size of the simulated samples.

B Covariance matrices obtained for the normalised differential cross-section measurements

The global 20×20 covariance matrices obtained for the normalised differential cross-section measurements described in Chapter 4 are listed in Tables B.1-B.3 for the inclusive, the top-quark and the top-antiquark measurements, respectively.

	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8	1	2	3	4
1	9178.00	3522.00	-208.00	516.00	-2459.00	-2598.00	-3194.00	-4757.00	-2.98	-1657.00	871.40	1388.00	814.30	903.00	508.40	-2825.00	1219.00	-556.10	-518.70	-144.50
2	3522.00	3645.00	-83.52	659.10	-2040.00	-1477.00	-1674.00	-2552.00	242.10	-1174.00	485.60	586.10	157.50	371.80	1035.00	-1704.00	456.50	-514.60	-312.30	370.40
3	-208.00	-83.52	1760.00	-594.30	237.00	-296.90	-81.14	-732.70	-114.60	290.20	63.08	-151.00	-244.50	-373.30	44.32	485.70	-716.80	-94.22	81.95	729.10
4	516.00	659.10	-594.30	1797.00	-1436.00	-865.20	-120.40	43.00	-47.67	-577.40	136.00	-10.42	235.10	78.78	1122.00	-936.30	397.20	-781.50	530.60	-146.30
5	-2459.00	-2040.00	237.00	-1436.00	3336.00	1390.00	180.30	791.40	498.90	692.00	-693.50	-234.40	-401.50	84.09	-2120.00	2175.00	-355.30	914.00	-457.90	-100.80
6	-2598.00	-1477.00	-296.90	-865.20	1390.00	2451.00	267.40	1129.00	739.20	479.90	-738.70	-225.00	-387.10	39.15	-1352.00	1444.00	-53.97	1032.00	-543.80	-433.80
7	-3194.00	-1674.00	-81.14	-120.40	180.30	267.40	2599.00	2022.00	-734.00	1035.00	116.20	-596.70	-297.10	-743.30	705.40	514.60	-676.50	61.13	553.30	62.06
8	-4757.00	-2552.00	-732.70	43.00	791.40	1129.00	2022.00	4056.00	-580.90	911.20	-240.00	-756.20	123.50	-360.30	56.75	846.00	-270.40	-60.24	666.80	-336.10
9	-2.98	242.10	-114.60	-47.67	498.90	739.20	-734.00	-580.90	1761.00	-852.50	-596.20	157.90	-310.00	349.80	-832.70	322.50	199.90	444.30	-294.30	-350.00
10	-1657.00	-1174.00	290.20	-577.40	692.00	479.90	1035.00	911.20	-852.50	1788.00	-2.14	-350.40	-242.10	-606.70	-407.30	673.00	-793.50	445.50	-80.63	428.60
11	871.40	485.60	63.08	136.00	-693.50	-738.70	116.20	-240.00	-596.20	-2.14	816.80	-41.25	132.30	-218.30	658.60	-749.80	13.90	-335.70	57.67	264.10
12	1388.00	586.10	-151.00	-10.42	-234.40	-225.00	-596.70	-756.20	157.90	-350.40	-41.25	569.40	37.83	257.90	-230.60	-400.80	273.40	123.10	-227.80	-168.60
13	814.30	157.50	-244.50	235.10	-401.50	-387.10	-297.10	123.50	-310.00	-242.10	132.30	37.83	668.10	133.20	128.10	-547.40	367.10	-364.60	131.40	-133.80
14	903.00	371.80	-373.30	78.78	84.09	39.15	-743.30	-360.30	349.80	-606.70	-218.30	257.90	133.20	896.40	-484.70	-327.60	655.80	-216.60	-198.80	-240.40
15	508.40	1035.00	44.32	1122.00	-2120.00	-1352.00	705.40	56.75	-832.70	-407.30	658.60	-230.60	128.10	-484.70	2792.00	-1623.00	32.14	-1053.00	762.70	258.00
16	-2825.00	-1704.00	485.70	-936.30	2175.00	1444.00	514.60	846.00	322.50	673.00	-749.80	-400.80	-547.40	-327.60	-1623.00	2653.00	-748.70	956.80	-150.30	-57.86
17	1219.00	456.50	-716.80	397.20	-355.30	-53.97	-676.50	-270.40	199.90	-793.50	13.90	273.40	367.10	655.80	32.14	-748.70	1418.00	-396.60	-297.30	-724.00
18	-556.10	-514.60	-94.22	-781.50	914.00	1032.00	61.13	-60.24	444.30	445.50	-335.70	123.10	-364.60	-216.60	-1053.00	956.80	-396.60	1566.00	-752.50	-416.50
19	-518.70	-312.30	81.95	530.60	-457.90	-543.80	553.30	666.80	-294.30	-80.63	57.67	-227.80	131.40	-198.80	762.70	-150.30	-297.30	-752.50	1020.00	29.99
20	-144.50	370.40	729.10	-146.30	-100.80	-433.80	62.06	-336.10	-350.00	428.60	264.10	-168.60	-133.80	-240.40	258.00	-57.86	-724.00	-416.50	29.99	1111.00

Table B.1: Full bin-wise covariance matrix ($\cdot 10^{-6}$) evaluated at particle level for both top quarks and antiquarks. The first eight bins correspond to $\cos\theta_{lx'}$, the next eight bins correspond to $\cos\theta_{ly'}$ and the last four bins correspond to $\cos\theta_{lz'}$.

	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8	1	2	3	4
1	6407.00	3564.00	-80.13	-213.40	-1759.00	-2752.00	-2434.00	-2731.00	-176.10	-1664.00	761.40	1301.00	1351.00	-828.20	1325.00	-2070.00	201.00	153.90	-881.10	526.20
2	3564.00	5006.00	-594.60	130.00	-2025.00	-2469.00	-1814.00	-1797.00	-94.51	-2002.00	324.60	1059.00	998.30	-1014.00	2935.00	-2206.00	707.60	-627.80	-407.70	327.90
3	-80.13	-594.60	2666.00	-728.70	431.80	133.00	-635.20	-1192.00	597.50	-463.30	245.30	-162.10	-196.60	395.00	-1338.00	922.40	-288.40	66.54	-433.90	655.80
4	-213.40	130.00	-728.70	2575.00	-971.20	-1064.00	239.70	32.45	-503.80	778.00	-307.60	219.70	-81.57	-573.20	1371.00	-902.30	-311.40	-1112.00	946.80	476.10
5	-1759.00	-2025.00	431.80	-971.20	2280.00	833.90	536.90	673.00	443.80	-232.80	71.44	-498.50	-276.10	740.20	-2009.00	1760.00	109.00	475.30	-185.90	-398.40
6	-2752.00	-2469.00	133.00	-1064.00	833.90	3408.00	766.20	1144.00	93.19	1697.00	-420.50	-1074.00	-876.00	788.90	-1775.00	1566.00	-356.90	999.10	-109.20	-533.00
7	-2434.00	-1814.00	-635.20	239.70	536.90	766.20	2324.00	1017.00	-112.20	1131.00	-328.20	-317.80	-741.50	270.40	-365.90	463.90	-227.10	18.25	402.60	-193.80
8	-2731.00	-1797.00	-1192.00	32.45	673.00	1144.00	1017.00	2853.00	-247.80	755.80	-346.40	-527.10	-177.60	220.90	-143.90	466.10	166.20	26.24	668.30	-860.80
9	-176.10	-94.51	597.50	-503.80	443.80	93.19	-112.20	-247.80	1138.00	-1224.00	65.98	-44.53	65.06	129.30	-693.20	563.20	250.30	160.60	-214.90	-196.10
10	-1664.00	-2002.00	-463.30	778.00	-232.80	1697.00	1131.00	755.80	-1224.00	4540.00	-705.30	-765.50	-1098.00	125.10	-523.10	-349.50	-1306.00	313.40	598.10	394.80
11	761.40	324.60	245.30	-307.60	71.44	-420.50	-328.20	-346.40	65.98	-705.30	630.60	40.03	211.70	7.21	-127.60	-122.70	138.40	-53.24	-224.80	139.70
12	1301.00	1059.00	-162.10	219.70	-498.50	-1074.00	-317.80	-527.10	-44.53	-765.50	40.03	1006.00	298.90	-340.80	452.80	-647.30	246.90	-140.00	-220.30	113.40
13	1351.00	998.30	-196.60	-81.57	-276.10	-876.00	-741.50	-177.60	65.06	-1098.00	211.70	298.90	904.10	-255.70	325.40	-451.60	374.10	-99.35	-176.10	-98.62
14	-828.20	-1014.00	395.00	-573.20	740.20	788.90	270.40	220.90	129.30	125.10	7.21	-340.80	-255.70	879.10	-1272.00	727.80	-30.37	315.50	-250.40	-34.74
15	1325.00	2935.00	-1338.00	1371.00	-2009.00	-1775.00	-365.90	-143.90	-693.20	-523.10	-127.60	452.80	325.40	-1272.00	4604.00	-2766.00	325.80	-1270.00	697.30	246.90
16	-2070.00	-2206.00	922.40	-902.30	1760.00	1566.00	463.90	466.10	563.20	-349.50	-122.70	-647.30	-451.60	727.80	-2766.00	3046.00	1.18	773.00	-208.90	-565.30
17	201.00	707.60	-288.40	-311.40	109.00	-356.90	-227.10	166.20	250.30	-1306.00	138.40	246.90	374.10	-30.37	325.80	1.18	832.50	-129.00	-282.00	-421.40
18	153.90	-627.80	66.54	-1112.00	475.30	999.10	18.25	26.24	160.60	313.40	-53.24	-140.00	-99.35	315.50	-1270.00	773.00	-129.00	1469.00	-724.60	-615.10
19	-881.10	-407.70	-433.90	946.80	-185.90	-109.20	402.60	668.30	-214.90	598.10	-224.80	-220.30	-176.10	-250.40	697.30	-208.90	-282.00	-724.60	1153.00	-146.80
20	526.20	327.90	655.80	476.10	-398.40	-533.00	-193.80	-860.80	-196.10	394.80	139.70	113.40	-98.62	-34.74	246.90	-565.30	-421.40	-615.10	-146.80	1183.00

Table B.2: Full bin-wise covariance matrix ($\cdot 10^{-6}$) evaluated at particle level for top quarks. The first eight bins correspond to $\cos\theta_{lx'}$, the next eight bins correspond to $\cos\theta_{ly'}$ and the last four bins correspond to $\cos\theta_{lz'}$.

	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8	1	2	3	4
1	14600.00	-66.78	-1893.00	3388.00	-3475.00	485.60	-4737.00	-8302.00	-104.80	833.50	107.70	558.30	-883.00	6282.00	-2130.00	-4664.00	3321.00	-1778.00	741.40	-2284.00
2	-66.78	6634.00	-136.00	-977.50	-414.70	-1186.00	-1205.00	-2647.00	619.40	-952.50	805.90	508.60	-972.60	-387.80	111.90	267.10	-398.70	425.50	-595.80	569.00
3	-1893.00	-136.00	5854.00	-3645.00	-1894.00	-2628.00	2779.00	1563.00	-1589.00	-2506.00	3031.00	-1237.00	1239.00	-1819.00	1838.00	1043.00	-494.90	-846.90	494.20	847.60
4	3388.00	-977.50	-3645.00	8048.00	-1238.00	1403.00	-3841.00	-3137.00	2777.00	1655.00	-3206.00	425.80	-1864.00	3555.00	-1062.00	-2280.00	833.80	-357.50	736.60	-1213.00
5	-3475.00	-414.70	-1894.00	-1238.00	6588.00	214.80	-762.80	982.30	622.30	2143.00	-2241.00	1477.00	-424.00	-2995.00	-68.77	1487.00	-1198.00	2046.00	-1816.00	967.20
6	485.60	-1186.00	-2628.00	1403.00	214.80	5479.00	-2758.00	-1010.00	2359.00	2146.00	-2853.00	455.90	-1026.00	2187.00	-2703.00	-566.10	873.70	550.80	-343.40	-1081.00
7	-4737.00	-1205.00	2779.00	-3841.00	-762.80	-2758.00	6222.00	4304.00	-2761.00	-2501.00	3277.00	-1231.00	1761.00	-3476.00	2675.00	2256.00	-994.70	-193.20	369.40	818.40
8	-8302.00	-2647.00	1563.00	-3137.00	982.30	-1010.00	4304.00	8247.00	-1923.00	-818.60	1079.00	-958.20	2169.00	-3346.00	1340.00	2457.00	-1943.00	153.30	413.60	1376.00
9	-104.80	619.40	-1589.00	2777.00	622.30	2359.00	-2761.00	-1923.00	5258.00	467.70	-2896.00	428.90	-1967.00	802.40	-1121.00	-973.90	-414.70	1031.00	-107.70	-509.00
10	833.50	-952.50	-2506.00	1655.00	2143.00	2146.00	-2501.00	-818.60	467.70	5116.00	-3406.00	1042.00	-758.30	1310.00	-2466.00	-1304.00	-61.40	603.50	-635.20	93.03
11	107.70	805.90	3031.00	-3206.00	-2241.00	-2853.00	3277.00	1079.00	-2896.00	-3406.00	5428.00	-1410.00	1300.00	-1034.00	1876.00	141.80	830.70	-1409.00	296.30	281.80
12	558.30	508.60	-1237.00	425.80	1477.00	455.90	-1231.00	-958.20	428.90	1042.00	-1410.00	1835.00	-739.00	-367.00	-530.50	-259.40	-336.50	851.20	-749.20	234.50
13	-883.00	-972.60	1239.00	-1864.00	-424.00	-1026.00	1761.00	2169.00	-1967.00	-758.30	1300.00	-739.00	2402.00	-731.90	27.88	465.30	19.87	-652.60	307.50	325.30
14	6282.00	-387.80	-1819.00	3555.00	-2995.00	2187.00	-3476.00	-3346.00	802.40	1310.00	-1034.00	-367.00	-731.90	7020.00	-3131.00	-3869.00	2738.00	-1997.00	982.50	-1723.00
15	-2130.00	111.90	1838.00	-1062.00	-68.77	-2703.00	2675.00	1340.00	-1121.00	-2466.00	1876.00	-530.50	27.88	-3131.00	4401.00	943.80	-982.60	252.70	142.30	587.60
16	-4664.00	267.10	1043.00	-2280.00	1487.00	-566.10	2256.00	2457.00	-973.90	-1304.00	141.80	-259.40	465.30	-3869.00	943.80	4855.00	-1793.00	1320.00	-236.50	709.80
17	3321.00	-398.70	-494.90	833.80	-1198.00	873.70	-994.70	-1943.00	-414.70	-61.40	830.70	-336.50	19.87	2738.00	-982.60	-1793.00	3134.00	-1286.00	-391.40	-1456.00
18	-1778.00	425.50	-846.90	-357.50	2046.00	550.80	-193.20	153.30	1031.00	603.50	-1409.00	851.20	-652.60	-1997.00	252.70	1320.00	-1286.00	2279.00	-1059.00	66.13
19	741.40	-595.80	494.20	736.60	-1816.00	-343.40	369.40	413.60	-107.70	-635.20	296.30	-749.20	307.50	982.50	142.30	-236.50	-391.40	-1059.00	2004.00	-553.00
20	-2284.00	569.00	847.60	-1213.00	967.20	-1081.00	818.40	1376.00	-509.00	93.03	281.80	234.50	325.30	-1723.00	587.60	709.80	-1456.00	66.13	-553.00	1943.00

Table B.3: Full bin-wise covariance matrix ($\cdot 10^{-6}$) evaluated at particle level for top antiquarks. The first eight bins correspond to $\cos\theta_{lx'}$, the next eight bins correspond to $\cos\theta_{ly'}$ and the last four bins correspond to $\cos\theta_{lz'}$.

C | Multivariate analysis technique in the dark matter search

This appendix is dedicated to the description, the development and the optimisation of the MVA technique used in the DM search, including the details that are missing in Section 5.4.2.

C.1 Introduction to the XGBoost classifier

Extreme Gradient Boosting (XGBoost) [413] is an optimised distributed gradient boosting library used for supervised learning problems, in which training data with n examples and m features $\mathcal{D} = \{(\boldsymbol{x_i}, y_i)\}(|\mathcal{D}| = n, \boldsymbol{x_i} \in \mathbb{R}^m, y_i \in \mathbb{R})$ is used to predict a target variable y_i. This encompasses a wide variety of tasks, such as regression, classification and ranking.

The mathematical structure which best relates the input data with the predicted output $\hat{y}_i$ is named the *model*. The main goal in supervised learning is to learn the model *parameters* from data. The model choice of XGBoost are decision tree ensembles. The *boosting* [426] ensemble method is employed to combine multiple weak learners sequentially. In XGBoost, the weak learners are decision trees which can approximate discrete functions. The resulting tree structure is flow-chart-like and splits at each of its internal nodes. Each split is a decision. A branch of the tree denotes the classification rules, and each leaf of the decision tree represents a class. In tree ensembles, an additive strategy is employed in which the prediction

of each tree is added together step by step, fixing what was learnt at each stage. Concretely, for a given step t, the prediction is given by:

$$\hat{y}_i^{(t)} = \sum_{k=1}^{t} f_k(\boldsymbol{x_i}) = \hat{y}^{(t-1)} + f_t(\boldsymbol{x_i}), \qquad f \in \mathcal{F}, \tag{C.1}$$

where k runs over the different trees from previous steps, f is a function in the functional space $\mathcal{F}$ which contains the tree structure and the leaf scores, and $\mathcal{F} = \{f(\boldsymbol{x}) = w_{q(\boldsymbol{x})}\}(q : \mathbb{R}^m \rightarrow T, w \in \mathbb{R}^T)$ is the space of all possible trees. Here, w is the vector of scores on leaves (leaf weights), T is the number of leaves in the tree and q is a function assigning each data point to the corresponding leaf.

The training procedure consists on finding the best parameters to fit data and labels. The ***objective function*** is defined to quantify the level of agreement between the fitted model and the data. In XGBoost this function consists of two parts: the training loss $\sum_{i=1}^{n} l(y_i, \hat{y}_i^{(t)})$, where l measures the difference between the prediction $\hat{y}_i^{(t)}$ and the target y_i, and the regularisation term $\sum_{i=1}^{t} \Omega(f_i)$. In order to find the best tree f_t in the t-th iteration, it is necessary to minimise the objective function:

$$\mathcal{L}^t = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{(t-1)} + f_t(\boldsymbol{x_i})) + \Omega(f_t), \tag{C.2a}$$

$$\Omega(f_t) = \gamma T + \frac{1}{2}\lambda ||w||^2, \tag{C.2b}$$

where γ and λ are the regularisation parameters. This additional regularisation term helps to reduce the complexity of the model and to avoid overfitting. All things considered, the training loss determines how predictive our model becomes with respect to the training data and the regularisation term determines the complexity of the model. The trade-off between these two terms is usually referred as the ***bias-variance trade-off*** in machine learning.

A Taylor expansion of the loss function up to the second order can be used to quickly optimise the objective function in the general setting [427]:

$$\mathcal{L}^t \simeq \sum_{i=1}^{n} [l(y_i, \hat{y}^{(t-1)}) + g_i f_t(\boldsymbol{x_i}) + \frac{1}{2} h_i f_t^2(\boldsymbol{x_i})] + \Omega(f_t), \tag{C.3}$$

where $g_i = = \partial_{\hat{y}^{(t-1)}} l(y_i, \hat{y}^{(t-1)})$ and $h_i = \partial^2_{\hat{y}^{(t-1)}} l(y_i, \hat{y}^{(t-1)})$ are first and second order gradient statistics on the loss function. The utilisation of second order derivatives is a characteristic feature of XGBoost and constitutes the origin of the *Extreme Gradient Boosting* name. This technique increases the computation speed and also prevents overfitting. After the constant terms are removed, the simplified objective function at step t becomes:

$$\tilde{\mathcal{L}}^t = \sum_{i=1}^{n} [g_i f_t(\boldsymbol{x_i}) + \frac{1}{2} h_i f_t^2(\boldsymbol{x_i})] + \Omega(f_t) \tag{C.4}$$

From this expression one can derive the optimal weight w_j^* of leaf j for a given structure $q(\boldsymbol{x})$. After some manipulation, this becomes:

$$w_j^* = -\frac{G_j}{H_j + \lambda}, \tag{C.5}$$

where $G_j = \sum_{i \in I_j} g_i$, $H_j = \sum_{i \in I_j} h_i$, and $I_j = \{i | q(\boldsymbol{x}_i) = j\}$ is the instance set of leaf j. The corresponding optimal value is given by:

$$\tilde{\mathcal{L}}^t(q) = -\frac{1}{2} \sum_{j=1}^{T} \frac{G_j^2}{H_j + \lambda} + \gamma T \tag{C.6}$$

This last equation measures the quality of a tree structure q. However, it is not possible to perform a scan over all possible trees and decide which is the best one. In order to overcome this problem, the optimisation is performed level by level. In particular, one can calculate the gain resulting from splitting a leaf into two leaves:

$$\mathcal{L}_{split} = \frac{1}{2} \left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda} \right] - \gamma, \tag{C.7}$$

where the first and second terms determine the score on the new left and right leaves, respectively. The third term is the score on the original leaf, and the last term is the regularisation on the additional leaf. As one can see, if the term between brackets is lower than γ the gain becomes negative, and thus it would be better not to add the branch.

The XGBoost software has become a popular choice for machine learning tasks due to its high performance. There are some characteristic features of this algorithm which make it a good choice to produce competitive results in a wide range of problems:

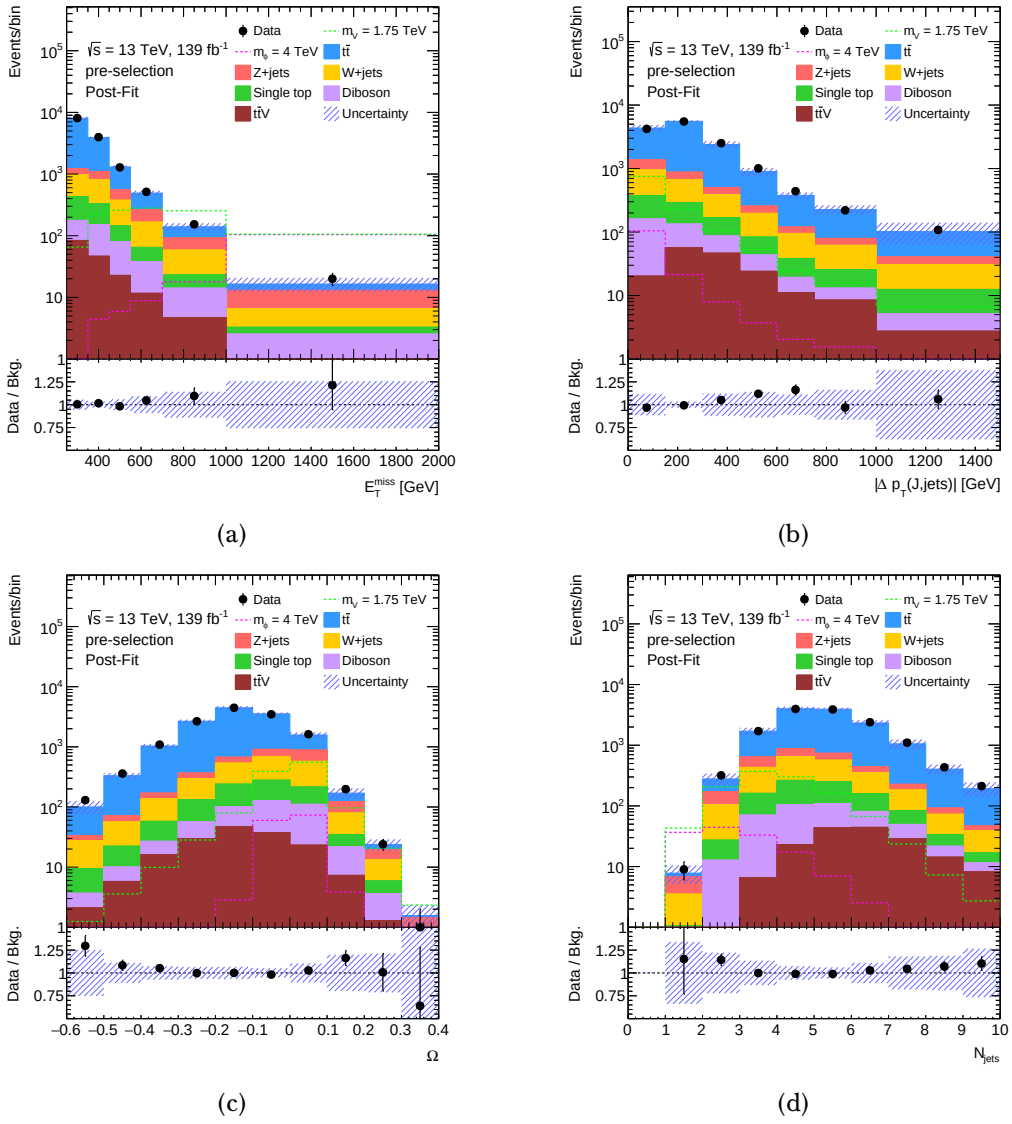
- The introduction of the regularisation term in the loss function reduces the tree complexity. This is an important benefit of the algorithm since the models are trained faster and require less storage space.
- Randomisation techniques are implemented in XGBoost to reduce overfitting and to increment training speed. These include random subsamples to train individual trees and column (feature) subsampling both at tree and tree node levels.
- XGBoost reduces drastically the computational complexity for finding the best split, which is the most time consuming part of tree learning, using different techniques:
  - Specific compressed column based structure, in which the data is stored pre-sorted. This input data layout only needs to be computed once before training, and can be reused in later iterations.
  - The weighted quantile sketch algorithm (see appendix in Ref. [413]) is one of the most innovative techniques introduced in XGBoost. This method is used to propose candidate split points and it is based on percentiles of the data. It is specifically designed to handle weighted datasets.
  - The sparsity-aware split finding algorithm is used to handle missing values in the computation of the loss gain of split candidates. The key improvement in this algorithm is that it only visits non-missing entries.

C.2 Variable and hyperparameter optimisation for the dark matter search

The variables used to train the XGBoost classifier are optimised in an initial stage. Two different criteria are considered in this optimisation. Initially, the variables which do not improve the performance[1] of the classifier by more than 0.001 are pruned. Secondly, if two variables have a correlation coefficient larger than 0.8, one of them is dropped away (keeping the one which has a larger impact on the performance). After these optimisation stage, the resulting variables used to train the classifier are summarised in Table 5.4.

The modelling of the input variables in the pre-selection (in which the classifiers are trained) has been checked to be good within the uncertainties, as shown in Figures C.1-C.3. Normalisation factors are applied to the $t\bar{t}$ and V+jets backgrounds as obtained from the background-only fit to the yields in the CRs (see Section 5.7.2). All other backgrounds are normalised to their theoretical cross-sections. The mono-top signal distributions are also displayed, both for the DM resonant model with the parameters $\lambda = 0.6$, $y = 0.4$, $m_\chi = 1$ GeV and $m_\phi = 4$ TeV, and for the non-resonant model with the parameters $a = 0.5$, $g_\chi = 1$, $m_\chi = 1$ GeV and $m_V = 1.75$ TeV. The uncertainty bands include all systematic sources as described in Section 5.6.

[1] The performance of the classifier can be evaluated according to different criteria. In this case, the area under the ROC curve is used.



(a) (b)

(c) (d)

Figure C.1: Comparison of data and SM predictions (post-fit) for (a) the $E_{\mathrm{T}}^{\mathrm{miss}}$, (b) $|\Delta p_{\mathrm{T}}(J,\mathrm{jets})|$, (c) Ω and (d) N_{jets} at pre-selection level. The $t\bar{t}$ and V+jets backgrounds are normalised to the results of a maximum-likelihood fit to data yields in the control regions. The expected distributions for the mono-top signal models are shown for the resonant model, assuming a mass of the mediation of $m_{\phi} = 4$ TeV and also for the non-resonant model, assuming a mass of the mediator of $m_V = 1.75$ TeV. The last bin of the distributions includes the overflow. The uncertainty bands include all systematic sources. The lower panel shows the ratio of data to prediction.

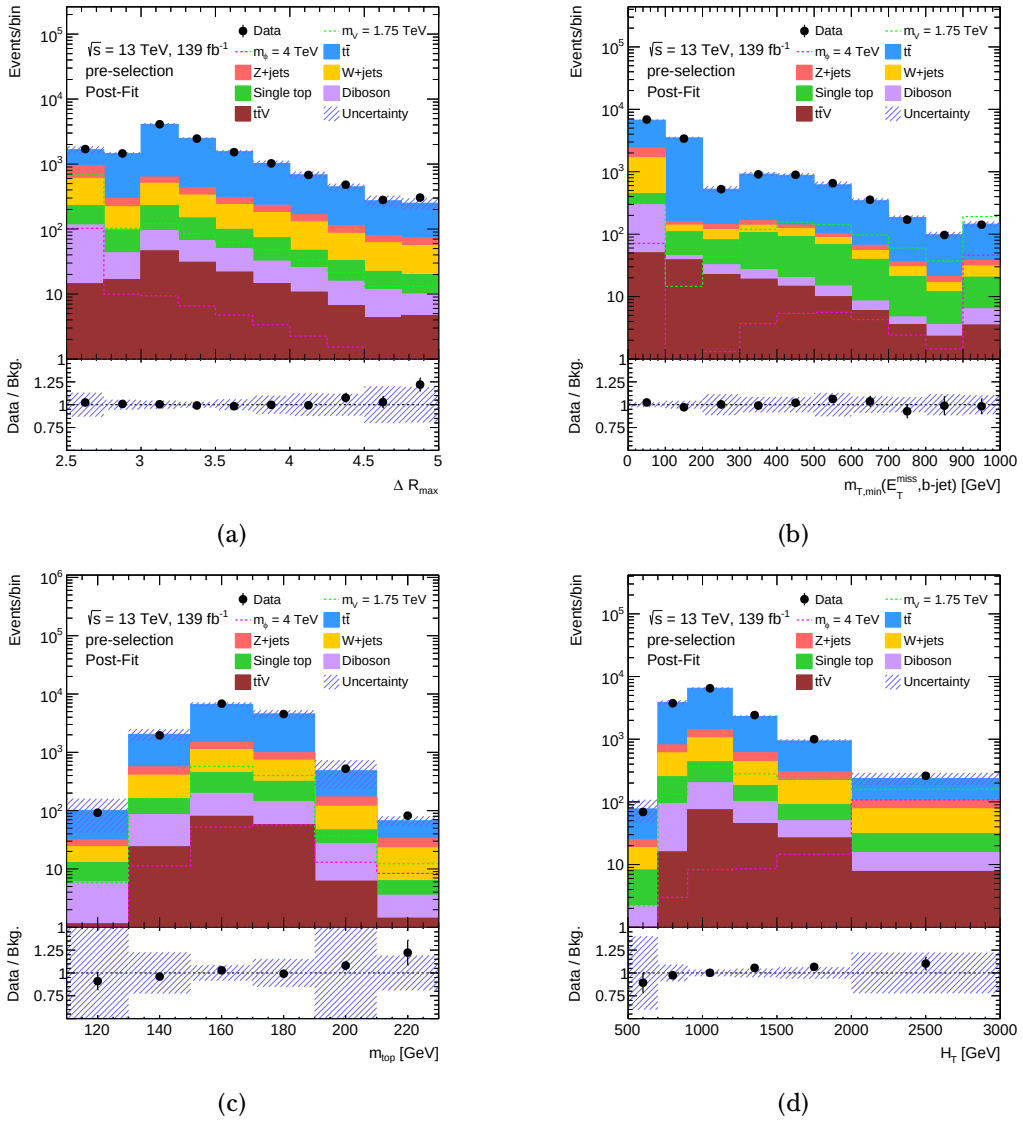
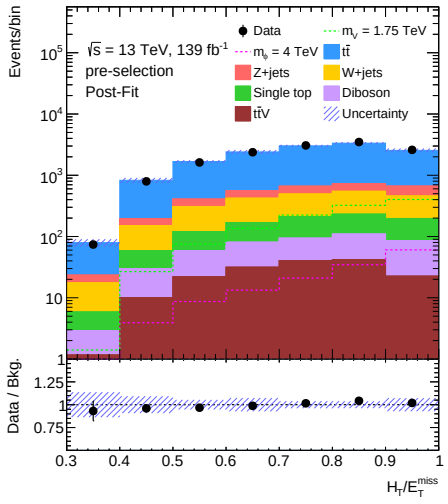
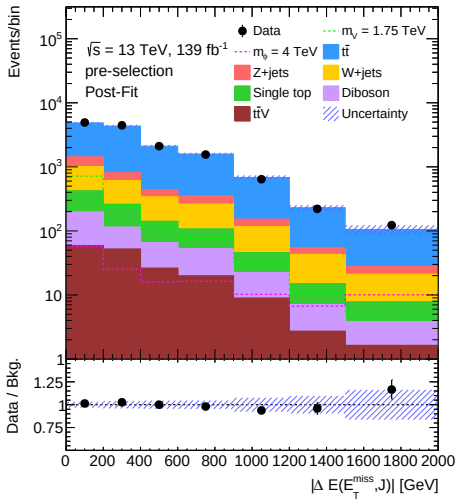


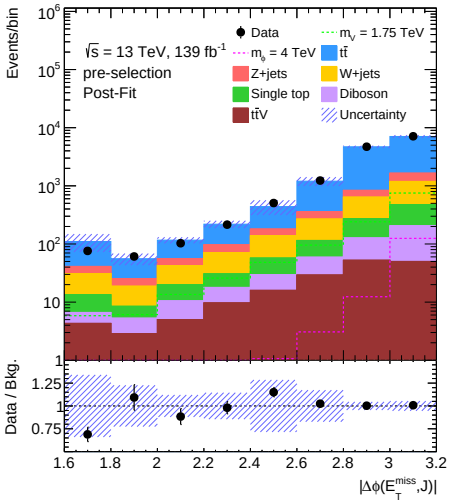
Figure C.2: Comparison of data and SM predictions (post-fit) for (a) ΔR_{max}, (b) $m_{\text{T,min}}(E_{\text{T}}^{\text{miss}}, b\text{-jets})$, (c) m_{top} and (d) H_{T} at pre-selection level. The $t\bar{t}$ and V+jets backgrounds are normalised to the results of a maximum-likelihood fit to data yields in the control regions. The expected distributions for the mono-top signal models are shown for the resonant model, assuming a mass of the mediation of $m_{\phi} = 4$ TeV and also for the non-resonant model, assuming a mass of the mediator of $m_V = 1.75$ TeV. The last bin of the distributions includes the overflow. The uncertainty bands include all systematic sources. The lower panel shows the ratio of data to prediction.



(a)



(b)

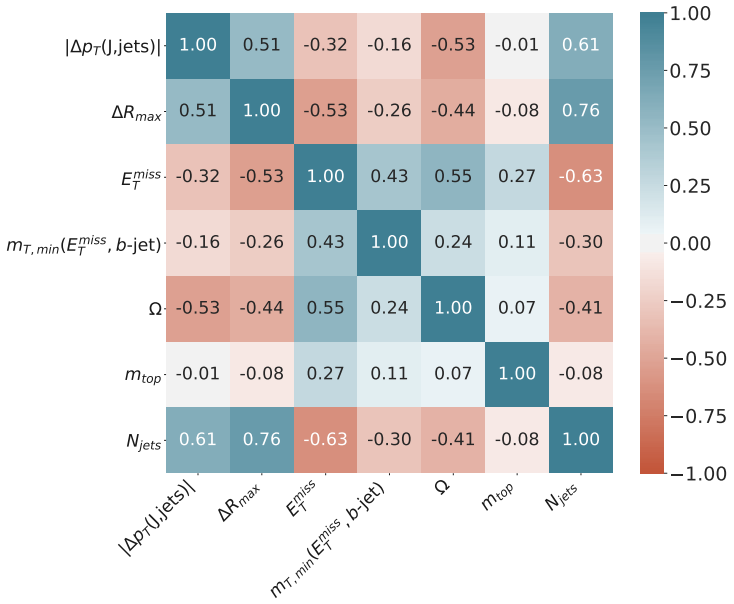


(c)

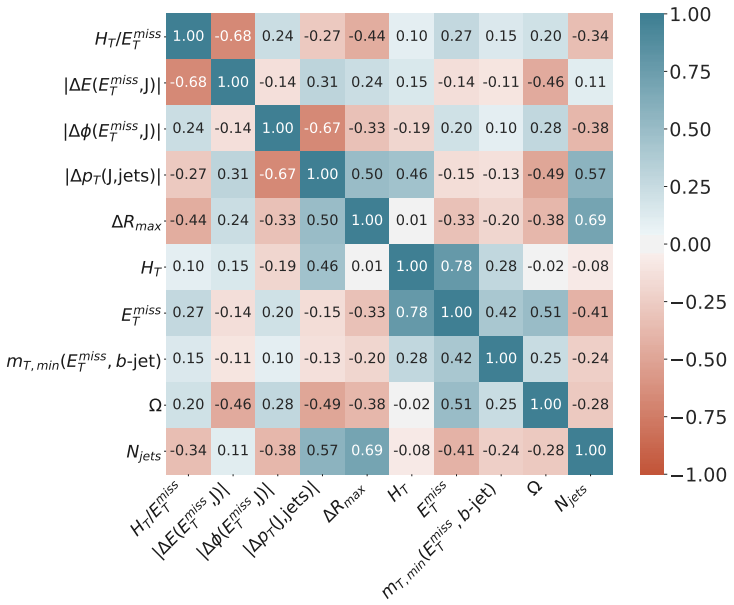
Figure C.3: Comparison of data and SM predictions (post-fit) for (a) $H_{\mathrm{T}}/E_{\mathrm{T}}^{\mathrm{miss}}$, (b) $|\Delta E(E_{\mathrm{T}}^{\mathrm{miss}}, J)|$ and (c) $|\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, J)|$ at pre-selection level. The $t\bar{t}$ and V+jets backgrounds are normalised to the results of a maximum-likelihood fit to data yields in the control regions. The expected distributions for the mono-top signal models are shown for the resonant model, assuming a mass of the mediation of m_{ϕ} = 4 TeV and also for the non-resonant model, assuming a mass of the mediator of m_V = 1.75 TeV. The last bin of the distributions includes the overflow. The uncertainty bands include all systematic sources. The lower panel shows the ratio of data to prediction.

The correlation matrix among the input variables is displayed in Figure C.4 for both the resonant and the non-resonant models. As one can see, the correlation between the selected input variables is always below 0.8.

The XGBoost parameters which affect most the classifier output are optimised to obtain the maximum performance, whilst the other parameters are kept to the default values. The performance of the training with different settings is compared in terms of the area under ROC curve. In addition, every setting has been checked to not significantly degrade the agreement of the classifier performance in the training and the testing subsets. To prevent overfitting and significantly reduce the training time, early stopping is also set to stop the training of the classifier if the score has not improved after 100 rounds. The optimised parameters are listed in Table 5.5.



(a)



(b)

Figure C.4: Correlation matrices between the different variables used in the training of the XGBoost classifier for (a) the resonant and (b) the non-resonant models.

Resum en valencià

El Model Estàndard (SM) de física de partícules és el marc teòric que millor descriu el món subatòmic en l'actualitat. Es tracta d'una teoria extremadament predictiva que ha anticipat l'existència de moltes partícules[1] i que també descriu les interaccions entre aquestes amb molta precisió. El descobriment del bosó de Higgs [5, 6] al Gran Col·lisionador d'Hadrons (LHC) [7, 8] l'any 2012 continua essent un dels esdeveniments més importants en la història recent de la física de partícules. No obstant això, el SM no es pot considerar com la teoria definitiva, ja que hi ha certs fenòmens que no pot explicar.

L'LHC es va construir no solament per a trobar el bosó de Higgs, sinó també per a buscar indicis de nova física que ens ajuden a entendre els misteris que romanen sense resposta. Hi ha dues estratègies ben diferenciades per a buscar nova física a l'LHC. La primera consisteix en mesures precises dels processos predits pel SM, buscant desviacions que ens indiquen on s'amaga la nova física. La segona és la cerca directa de nova física en la forma de noves partícules. El treball d'aquesta tesi analitza les dades recollides pel detector ATLAS (A Toroidal LHC Apparatus) [265] entre els anys 2015 i 2018, i presenta els resultats de dues anàlisis, cadascuna seguint una de les dues estratègies prèviament esmentades: una mesura d'observables sensibles a la polarització del quark cima[2] i una cerca directa de partícules de matèria fosca (DM).

Per poder realitzar mesures d'alta precisió, és indispensable que les dades recollides pels detectors siguen de màxima qualitat. Part del treball d'aquesta tesi

[1] Cal destacar els bosons W i Z, el quark cima (*top*) o el bosó de Higgs, entre altres.

[2] D'ara endavant s'utilitza el nom en anglés d'aquesta partícula, quark *top*.

també contribueix a aquest objectiu, concretament a través de l'alineament del detector intern d'ATLAS.

R.1 Marc teòric

El Model Estàndard de física de partícules és una teoria quàntica de camps que descriu els components fonamentals de la matèria i les seues interaccions (excepte la gravitatòria). La Figura R.1 mostra totes les partícules que conformen el SM. Així, els fermions constitueixen la matèria, i interactuen entre ells mitjançant l'intercanvi d'altres partícules, anomenades bosons. Cada fermió té associat un antifermió que comparteix la mateixa massa, però té els nombres quàntics[3] oposats.

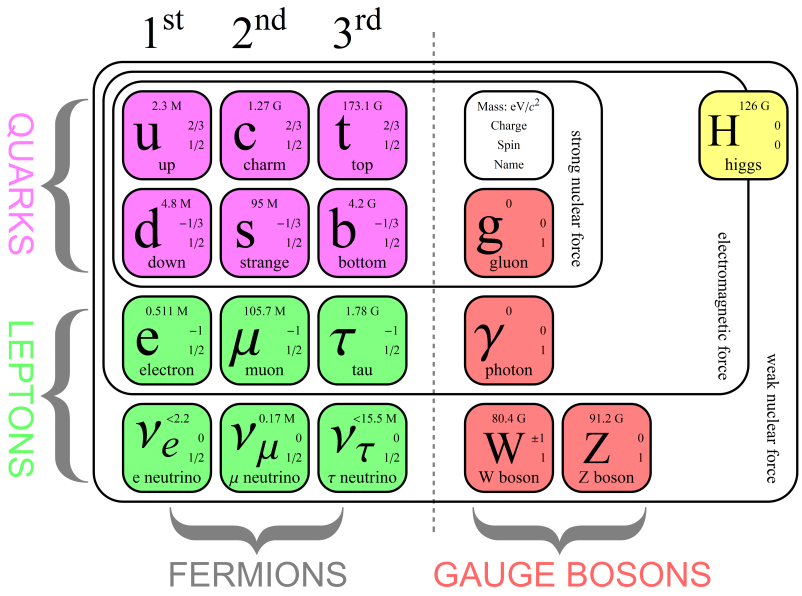


Figure R.1: Partícules fonamentals que constitueixen el SM. Les tres primeres columnes inclouen les tres generacions de fermions (leptons i quarks), que formen la matèria. La quarta i quinta columna inclouen els bosons, que controlen les interaccions. La sexta columna inclou el bosó de Higgs. Cadascuna de les partícules inclou la seua massa, espín i càrrega elèctrica.

[3] Els nombres quàntics descriuen els valors de quantitats conservades en la dinàmica d'un sistema quàntic.

Els fermions són partícules d'espín 1/2 que obeeixen l'estadística de Fermi-Dirac. Es poden dividir en dos grups: leptons i quarks. A més a més, es poden dividir en tres generacions, cadascuna amb els mateixos nombres quàntics però amb massa diferent. La primera generació conté les partícules més lleugeres, les quals formen la matèria ordinària. La segona i tercera generació són inestables i eventualment es desintegren a partícules de la primera generació.

En el cas dels leptons, cada generació està formada per una partícula amb càrrega elèctrica (electró, muó o tau) i el seu neutrí sense càrrega. D'altra banda, cada generació de quarks està formada per dos quarks carregats. Els quarks a més posseeixen un altre tipus de càrrega, la càrrega de color. Una peculiaritat dels quarks és que mai apareixen aïllats a la natura, sinó agrupats en unes partícules de càrrega de color neutra, els hadrons. Aquests hadrons es poden dividir en dos grups: els barions, formats per tres quarks; i els mesons, formats per una parella quark-antiquark. El protó i el neutró són exemples d'aquests hadrons. Així doncs, la matèria al seu nivell més fonamental, els àtoms, està formada per fermions de la primera generació: un nucli format per protons i neutrons, i electrons orbitant al seu voltant.

Pel que fa als bosons, aquests són partícules d'espín 1 que obeeixen l'estadística de Bose-Einstein. Hi ha diferents tipus de bosons, cadascun associat a una de les interaccions descrites pel SM. Els huit gluons (g) són els mediadors de la interacció forta, descrita per la cromodinàmica quàntica (QCD). El fotó (γ) és el mediador de la interacció electromagnètica, descrita per l'electrodinàmica quàntica (QED). Finalment, els bosons $W^{\pm}$ i Z són els mediadors de la interacció feble. Tant els gluons com el fotó són partícules sense massa, mentre que els bosons $W^{\pm}$ i Z són partícules amb massa.

Totes aquestes partícules i les seues interaccions estan descrites pel lagrangià del SM, que es fonamenta en el grup de simetria $SU(3)_C \times SU(2)_L \times U(1)_Y$. Per un costat, el terme $SU(3)_C$ és el grup de color i està associat a la interacció forta. Per l'altre, el grup $SU(2)_L \times U(1)_Y$ és el grup electrofeble i està associat a la interacció electrofeble, que unifica tant la interacció electromagnètica com la interacció feble. El lagrangià del SM és invariant sota transformacions de *gauge* locals d'aquest grup. No obstant això, per mantenir aquesta simetria, és necessari que les partícules

que formen el SM siguen partícules sense massa. Això no es correspon amb la realitat, i per tant és necessari un mecanisme que permeta dotar de massa a les partícules sense trencar la simetria. Amb aquest objectiu, s'introdueix el mecanisme de Brout-Englert-Higgs [17–19]. Aquest mecanisme permet recuperar les masses de les partícules sense perdre les propietats de simetria a través d'un procés anomenat trencament espontani de simetria (SSB). Per poder utilitzar aquest mecanisme, és menester introduir un nou camp escalar. Aquest nou camp es correspon amb una nova partícula, que s'anomena el bosó de Higgs. Per tant, el descobriment d'aquesta partícula a l'LHC és una de les proves més fortes que confirmen la validesa del SM.

Malgrat l'enorme èxit del SM, encara hi queden molts misteris que aquesta teoria no pot explicar. Alguns exemples inclouen les oscil·lacions dels neutrins[4], la inclusió de la interacció gravitatòria, l'asimetria entre matèria i antimatèria, o entendre la natura de la matèria i l'energia fosques (que constitueixen el 26.8% i el 68.3% de l'univers, respectivament [85]). Totes aquestes evidències suggereixen que el SM no és la teoria definitiva, sinó una aproximació d'una teoria més general que donaria resposta a aquestes qüestions.

R.1.1 La física del quark top

El quark top constitueix la tercera generació de quarks junt amb el quark b. Va ser descobert l'any 1995 pels experiments CDF i D0 al Tevatron [96, 97]. Aquesta partícula es caracteritza per una càrrega elèctrica de valor $+\frac{2}{3}$ (en unitats de càrrega de l'electró) i, sobretot, per la seua massa. Amb un valor de $m_t = 172.76 \pm 0.30$ GeV [12], es tracta de la partícula més massiva del SM. Aquesta elevada massa li confereix unes propietats molt característiques. En primer lloc, fa que la seua vida mitjana siga molt curta, i per tant es tracta de l'únic quark que es desintegra abans d'hadronitzar-se. Per aquesta raó, podem estudiar les propietats del quark top a través dels seus productes de desintegració. A més a més, el quark top es desintegra quasi exclusivament a un quark b i un bosó W ($\sim$99.8% de les vegades [12]), proporcionant senyals característics al detector. En segon lloc, l'alta

[4] Aquest fenomen només és possible si aquestes partícules tenen massa. Tanmateix, aquestes partícules no tenen massa dins del formalisme del SM.

massa del quark top li proporciona l'acoblament més gran amb el bosó de Higgs de totes les partícules del SM. Això suggereix que el quark top podria tindre un rol especial en el SSB. Finalment, molts models de nova física prediuen l'existència de noves partícules que s'acoblarien preferiblement al quark top. Tot plegat, la física del quark top constitueix una de les àrees de recerca més importants del programa de física de l'LHC.

A les col·lisions protó-protó de l'LHC els quarks top es poden produir bé en parelles, a través de la interacció forta, o bé en solitari a través de la interacció electrofeble. La producció per parelles és la dominant[5]; nogensmenys, la producció en solitari també resulta interessant d'estudiar, ja que ens permet accedir al vèrtex tWb que intervé tant en la producció de quarks top en solitari com en la seua desintegració. Dins de la producció en solitari trobem tres canals de producció diferents: el canal t i el canal s, on el bosó W es produeix de forma virtual; i el canal tW, on el bosó W es produeix de forma real amb el quark top.

Com s'ha dit abans, el quark top es desintegra fonamentalment en un quark b i un bosó W. Considerant els modes de desintegració del bosó W, la desintegració del quark top es pot classificar en dos canals: el canal hadrònic, quan el bosó W es desintegra a dos quarks; i el canal leptònic, quan es desintegra a un leptó i un neutrí.

R.2 El detector ATLAS a l'LHC

Les dues anàlisis presentades en aquesta tesi estan basades en dades recollides pel detector ATLAS de l'LHC durant el Run-2 (2015-2018). Aquestes dades es corresponen a col·lisions protó-protó a una energia en el centre de masses de $\sqrt{s} = 13$ TeV. En aquest capítol s'expliquen els detalls del dispositiu experimental emprat per a fer aquest treball, així com les tècniques de reconstrucció dels objectes físics que s'utilitzen en les dues anàlisis.

[5] Concretament, la secció eficaç (la probabilitat d'interacció per unitat de superfície) de la producció per parelles és aproximadament 2.75 vegades més gran que la producció en solitari per a les energies de col·lisió amb què s'ha treballat en aquest treball [101–109].

R.2.1 El Gran Col·lisionador d'Hadrons

L'LHC [7, 8] és un anell de 27 km de circumferència que es troba sota terra en el laboratori del CERN [244]. Es tracta de l'accelerador de partícules de més energia del món. Les partícules són accelerades en dos feixos al llarg d'aquesta estructura i es fan interactuar en quatre punts específics, on es col·loquen els detectors. ATLAS i CMS [249] són dos detectors de propòsit general, dissenyats amb l'objectiu de trobar el bosó de Higgs, realitzar mesures de precisió del SM i cercar signes de nova física. ALICE [250] busca reproduir els primers instants de l'univers, just després del Big Bang, mentre que LHCb [251] es dedica a l'estudi de la física del quark b.

R.2.2 El detector ATLAS

El detector ATLAS [265] és el major detector que mai s'ha construït en el món de la física de partícules. Es caracteritza per la seua estructura cilíndrica, com es pot veure a la Figura R.2. Aquest detector s'organitza en diferents capes, cadascuna amb un objectiu determinat:

- **Detector intern**: Aquest és el component més pròxim al punt d'interacció. Té la tasca de reconstruir les trajectòries de les partícules carregades que el travessen (traces), i amb això també reconstruir els vèrtexs primaris i secundaris així com mesurar el moment de les partícules. Això és possible gràcies a un camp magnètic de 2 T proporcionat per un solenoide superconductor, que corba les trajectòries de les partícules carregades. Aquest detector es pot dividir en quatre components diferents. De dins cap a fora tenim els tres detectors de silici: IBL, detector de píxels i SCT; i el TRT, format per tubs plens de gas.

- **Calorímetres**: Els calorímetres són responsables d'aturar les partícules i mesurar la seua energia. Hi ha principalment dos tipus de calorímetres: el calorímetre electromagnètic, que s'encarrega de mesurar l'energia dels

electrons i els fotons; i el calorímetre hadrònic, que mesura l'energia dels *jets*[6].

- **Cambra de muons**: La cambra de muons és el component més extern del detector ATLAS. S'encarrega de reconstruir les trajectòries dels muons, que són les úniques partícules que escapen dels calorímetres sense gaire interactuar (amb l'excepció dels neutrins).

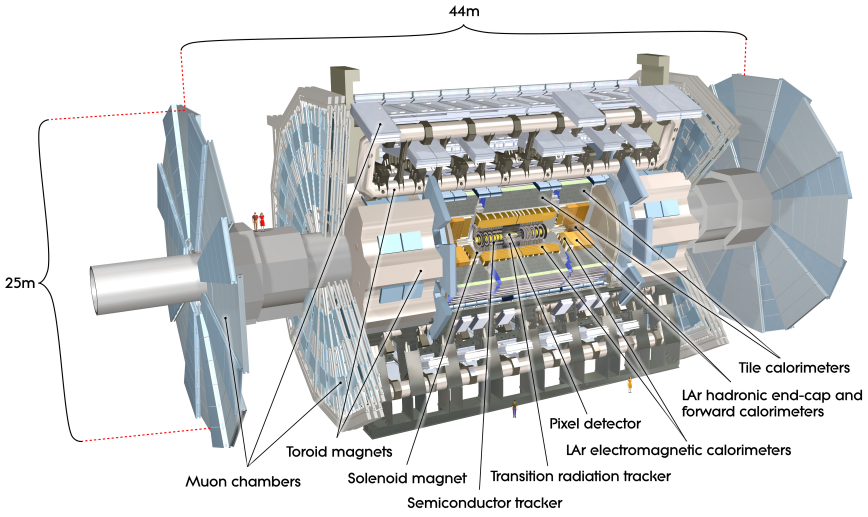


Figure R.2: Visió esquemàtica del detector ATLAS [266].

A causa de la immensa quantitat d'esdeveniments[7] que tenen lloc al detector ATLAS, és necessari aplicar una selecció per a emmagatzemar només aquells que siguen interessants. El sistema de *trigger* [274–276] s'encarrega d'aquest treball, reduint els ∼40 milions d'esdeveniments que tenen lloc a l'LHC per segon a només ∼1000, que es desen per a la seua posterior anàlisi.

[6] Quan es produeix un quark o un gluó en les col·lisions a l'LHC, es forma una cascada d'hadrons que es propaga en la direcció del quark o el gluó original, formant un doll d'hadrons conegut com a *jet*.

[7] S'anomena esdeveniment al resultat d'una col·lisió.

Els senyals proporcionats pels diferents components del detector ATLAS s'han de transformar en objectes físics reconstruïts per poder emprar-los en les diferents anàlisis. A continuació es dóna una descripció dels algoritmes de reconstrucció dels objectes que s'han utilitzat en la realització d'aquest treball:

- **Electrons**: Es reconstrueixen associant un dipòsit d'energia al calorímetre electromagnètic amb una traça al detector intern [282].
- **Muons**: Es reconstrueixen associant traces en la cambra de muons amb traces en el detector intern, complementant aquesta informació amb el senyal dels calorímetres [287].
- **Jets**: Els jets es reconstrueixen a partir dels dipòsits d'energia en els calorímetres i les traces del detector intern. Aquesta informació és combinada utilitzant l'algoritme anti-k_t [290]. Un paràmetre clau en aquest algoritme és el radi R del con d'hadrons. Això permet distingir entre dos tipus de jets: grans ($R = 1.0$) [305, 306] i menuts ($R = 0.4$) [291]. A partir d'ara, ens referirem als jets menuts simplement com a jets.
- ***b*-jets**: Els jets s'identifiquen com a provinents d'un quark b emprant diferents algoritmes que exploten els seus trets característics, com els seus paràmetres d'impacte[8], la presència de vèrtexs secundaris o la seua topologia de desintegració [297, 298].
- **top-jets**: De manera similar, els jets grans es poden identificar com a provinents d'un quark top a partir de determinades característiques, com l'enorme massa del jet reconstruït o que els seus constituents estiguen agrupats [322, 323].
- **Energia transversa mancant**: A causa de la conservació del moment, la suma del moment transvers de tots els productes de col·lisió ha de ser zero en el pla transversal. Això vol dir que la presència de partícules no detectades crearia un desequilibri en aquest càlcul, produint l'anomenada

[8] Els paràmetres d'impacte determinen la distància d'una traça reconstruïda al punt més proper a la direcció dels feixos.

energia transversa mancant $E_{\mathrm{T}}^{\mathrm{miss}}$. En el marc del SM, l'única partícula que produeix $E_{\mathrm{T}}^{\mathrm{miss}}$ és el neutrí, que només interacciona feblement i escapa sense ser detectat. L'$E_{\mathrm{T}}^{\mathrm{miss}}$ es reconstrueix com la suma vectorial negativa del moment transvers de totes les partícules detectades, així com de les traces que no estan associades a cap objecte físic [327, 328].

R.3 Alineament del detector intern d'ATLAS

Per a poder realitzar mesures d'alta precisió és crucial que les dades recollides pel detector ATLAS siguen de màxima qualitat. Els canvis en les condicions del detector intern (temperatura, canvis en el camp magnètic...) poden produir modificacions en la posició de les subestructures que el componen. Si aquestes alteracions en la geometria del detector no es corregeixen, aleshores podrien donar lloc a mesures esbiaixades. Com el detector és inaccessible mentre s'estan prenent dades, cal trobar una manera de corregir aquestes desviacions sense accedir directament al detector intern. De fet, la funció de l'algoritme d'alineament és la de determinar la posició dels diferents elements que conformen el detector intern i seguir la seua evolució temporal.

R.3.1 Algoritme d'alineament

L'alineament del detector intern està basat en les traces reconstruïdes. Es fonamenta en la minimització dels residus, definits com les distàncies entre els punts predits per les traces reconstruïdes i els senyals recollits en el detector. La Figura R.3 mostra el funcionament del procés d'alineament. A la part de l'esquerra es veu la posició real dels detectors. El senyal a cadascun dels tres mòduls representats s'utilitza per a reconstruir la traça. Però com es veu a la figura d'enmig, la posició real dels detectors no és la mateixa que la seua posició aparent. Per aquesta raó, la traça reconstruïda no és la mateixa, i apareixen unes desviacions entre els senyals mesurats i els punts d'intersecció de la traça, donant una distribució de residus desplaçada de zero. L'algoritme tracta de minimitzar aquests residus corregint la posició del detector, com es veu a la figura de la

dreta. Aquest procés és iteratiu i necessita diverses iteracions per assolir la màxima precisió.

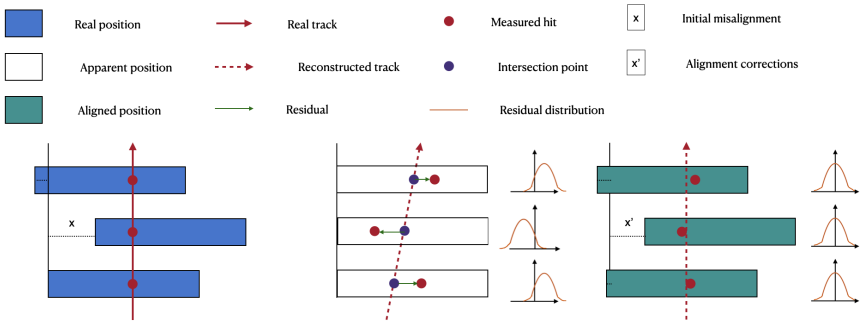


Figure R.3: Visió esquemàtica de la cadena d'alineament.

L'algoritme d'alineament consisteix en la minimització d'una funció χ^2 construïda a partir dels residus en els diferents mòduls. Aquesta minimització la du a terme l'algoritme del χ^2 global [335]. La precisió de l'alineament és molt elevada, de tal forma que es coneix la posició dels diferents mòduls amb una precisió de l'ordre del micròmetre. No obstant això, l'algoritme d'alineament es veu afectat per una sèrie de distorsions sistemàtiques que no pot corregir. Aquestes distorsions reben el nom de *modes febles*. Per tal de corregir els modes febles, s'inclouen unes restriccions addicionals a l'algoritme. Concretament, s'empra la informació de processos ben coneguts, com la desintegració del bosó Z a dos muons. Les traces d'aquests dos muons estan íntimament relacionades. Per tant, la informació de les dues traces reconstruïdes (com, per exemple, les diferències entre els paràmetres d'impacte) pot ser utilitzada per a corregir aquestes desviacions sistemàtiques.

R.3.2 Deformacions dependents del temps

Al Run-2 es va observar que hi havia uns moviments prou significatius que tenien lloc al llarg de cada període de presa de dades. Aquestes deformacions afectaven en particular al detector de píxels i a l'IBL, i era menester entendre-les i corregir-les. Ambdues deformacions estan relacionades amb les altes condicions

de lluminositat[9] característiques del Run-2. Per corregir aquests moviments s'introdueixen unes iteracions d'alineament específiques després de cada presa de dades, conegudes com a *cicle de calibratge* [341].

R.3.2.1 Deformació de les dogues de l'IBL

En el cas de l'IBL, es va observar que les dogues que el conformen s'arquejaven amb els canvis de temperatura. Per a corregir aquesta deformació es va introduir un grau de llibertat addicional en el procediment d'alineament. Concretament, es va veure que aquesta deformació es podia parametritzar amb una paràbola [344]. Això es veu representat a la Figura R.4, on s'observa la deformació de les dogues de l'IBL a diferents temperatures i el seu modelatge amb la introducció del grau de llibertat addicional.

[9] La lluminositat és una mesura del nombre de col·lisions que tenen lloc al detector per unitat d'àrea i de temps.

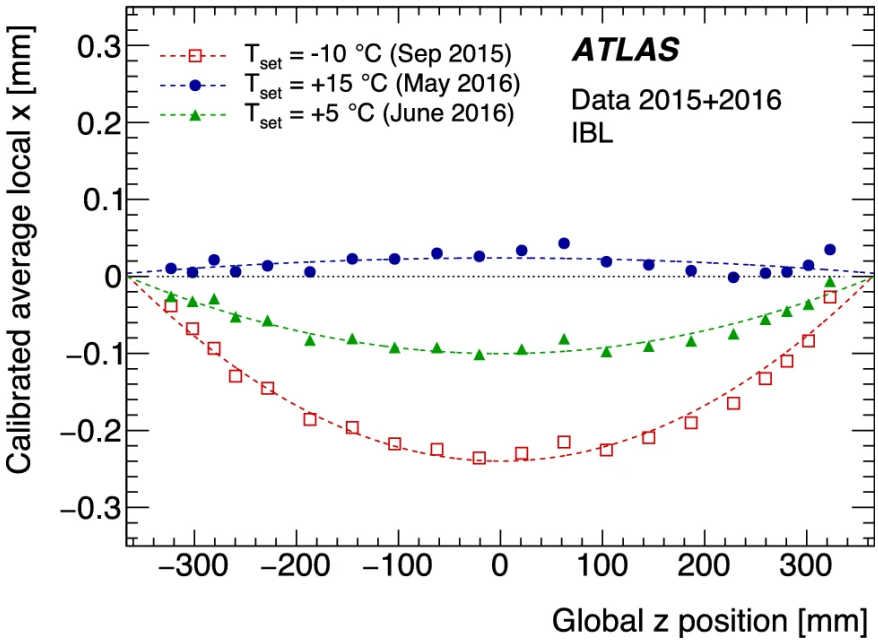


Figure R.4: Arquejat de l'IBL fent una mitjana sobre les diferents dogues per a diferents períodes de presa de dades: setembre de 2015 (quadrats rojos), maig de 2016 (cercles blaus) i juny de 2016 (triangles verds). Cada període té associat una temperatura característica. Es pot percebre una clara dependència entre l'arquejat de les dogues i la temperatura. Les línies de colors representen la correcció proporcionada per l'alineament, mentre que la línia negra representa la posició de referència [1].

R.3.2.2 Desplaçament vertical del detector de píxels

L'altre moviment característic del detector intern durant la presa de dades és un desplaçament vertical del detector de píxels. La primera contribució d'aquest treball a l'alineament consisteix en un estudi detallat d'aquest tipus de deformació del detector. Concretament, es va observar que el detector de píxels s'alçava cada vegada que començava la presa de dades, i després anava caient a poc a poc fins que aquesta finalitzava. Aquest tipus de patró es veu representat a la Figura R.5. El detector de píxels s'eleva al principi, aproximadament uns 5 μm, per a després

anar baixant lentament uns 3 μm. Com es pot apreciar, hi existeix una relació directa entre aquests moviments i la lluminositat.

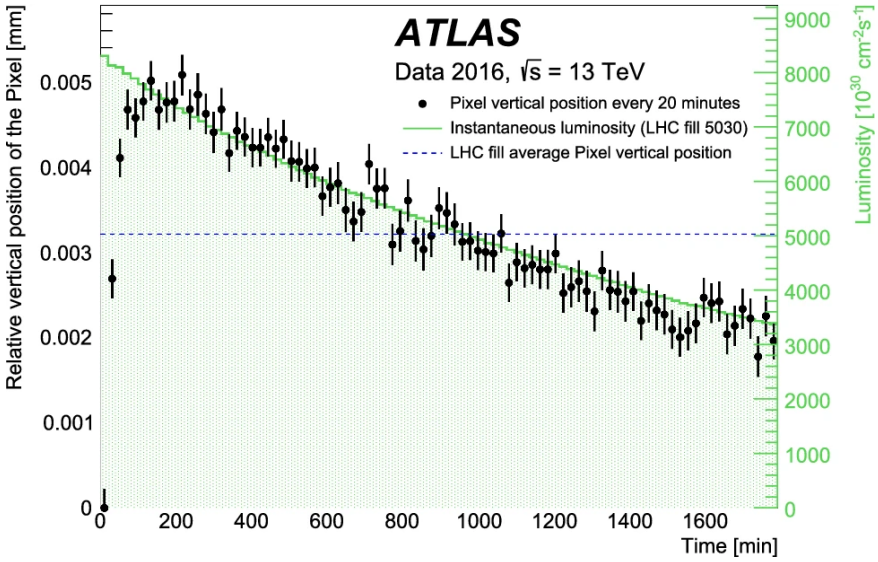


Figure R.5: Moviments del detector de píxels durant la presa de dades. Cada vegada que el detector s'encén, aquest puja en un règim transitori fins que arriba a un màxim. Després el detector de píxels va baixant seguint el comportament de la lluminositat dels feixos. Els punts negres corresponen a la posició vertical del detector de píxels com a funció del temps (cada 20 minuts) i relativa a la posició inicial, mentre que la lluminositat es representa com una àrea de color verd. La normalització relativa d'ambdues quantitats permet una comparació directa. La línia blava correspon a la posició mitjana que s'obtindria considerant la presa de dades sencera. Les incerteses representades són purament estadístiques.

Aquests moviments característics del detector de píxels es deuen al líquid de refrigeració utilitzat per mantenir la temperatura del detector [11]. Al començament de la presa de dades, el detector de píxels s'encén i augmenta la temperatura. Per aquesta raó es produeix una evaporació parcial del líquid de refrigeració, reduint-se la massa i produint-se un ascens del detector. A mesura que la lluminositat decreix, la temperatura baixa i el gas es condensa, produint-se un descens de la posició vertical del detector.

R.3.3 Determinació d'un alineament de referència

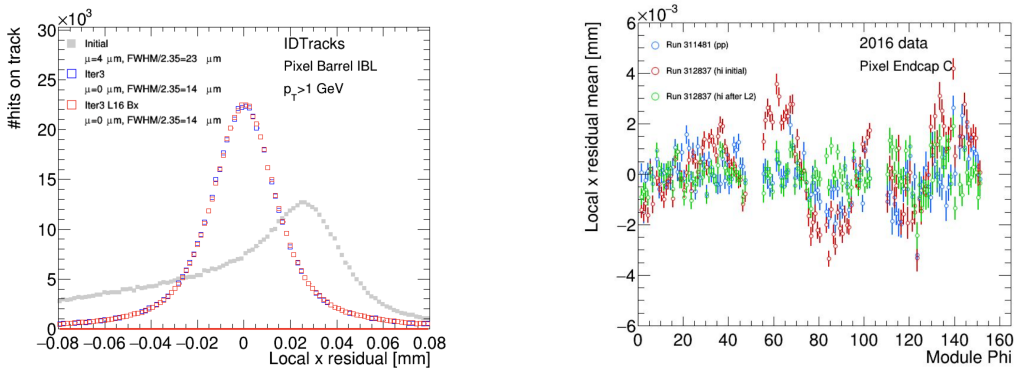
El cicle de calibratge és una ferramenta molt útil per corregir els diferents moviments que esdevenen durant la presa de dades. Aquestes correccions s'apliquen sempre damunt unes constants de referència, que determinen la posició de partida del detector. Atés que el cicle de calibratge no pot corregir per si mateix deformacions molt dràstiques en la geometria del detector, ni tampoc els modes febles, és necessari que aquestes constants de referència descriguen amb suficient precisió la posició del detector. L'altra contribució d'aquest treball a l'alineament del detector intern consisteix en la determinació de constants de referència per a tres períodes de col·lisió diferents: les col·lisions d'ions pesats en 2015 i 2016, i les col·lisions protó-protó corresponents a 2018.

R.3.3.1 Constants de referència per a les col·lisions d'ions pesats

Les constants de referència obtingudes per als dos períodes de col·lisions d'ions pesats van ser el punt de partida per a un primer alineament del detector intern durant les dues campanyes considerades. Per a obtenir aquestes constants s'utilitza només una part mínima de les dades, i es realitza un alineament d'alta precisió. Després, les constants obtingudes s'utilitzen per a un alineament de cadascun dels dos períodes dins del cicle de calibratge. Aquest s'encarrega de corregir els moviments característics dins de cada presa de dades (Secció R.3.2), i qualsevol altre moviment residual que es puga presentar.

La Figura R.6 (a) mostra la distribució de residus a l'IBL per al període de 2015, abans i després de l'alineament de precisió. Com es pot apreciar, la posició del detector al començament estava clarament esbiaixada. Tanmateix, l'alineament és capaç de corregir les deformacions, i a la fi les distribucions de residus estan clarament centrades en zero.

La Figura R.6 (b) mostra la mitjana de residus per a cada mòdul d'un component del detector de píxels en el període de 2016. Es poden apreciar uns patrons molt clars abans de l'alineament que cal corregir. En completar el procediment d'alineament es veu que aquests patrons desapareixen i la mitjana de residus apareix centrada en zero per als diferents mòduls.



(a) (b)

Figure R.6: (a): Distribució de residus a l'IBL per a la campanya d'alineament d'ions pesats de 2015 abans de l'alineament (gris), en un pas intermedi (blau) i després de la cadena d'alineament (blau). (b): Distribució mitjana de residus per a la campanya d'alineament d'ions pesats de 2016 en els diferents mòduls d'un component del detector de píxels abans de l'alineament (roig) i després de l'alineament (verd). Els punts blaus representen un període de presa de dades de col·lisions protó-protó previ a les col·lisions d'ions pesats que es va utilitzar com a punt de partida.

Finalment la precisió que es va assolir en les constants de referència era de l'ordre d'uns pocs micròmetres per ambdós períodes, garantint un alineament d'alta precisió dins del cicle de calibratge per als dos períodes considerats.

R.3.3.2 Constants de referència per a les col·lisions protó-protó

Després de cada campanya de presa de dades es produeix un reprocessat per a garantir la màxima qualitat possible de les dades. Aquesta secció explica el realineament del període corresponent a les col·lisions protó-protó en 2018. De nou, se selecciona una part de les dades per a realitzar aquest estudi, i les constants de referència obtingudes s'empren com a punt de partida en el realineament sencer del període dins del cicle de calibratge.

Aquest treball se centra en la correcció dels modes febles, descrits prèviament en la Secció R.3.1. Per a corregir-los s'empra la informació d'esdeveniments $Z \rightarrow \mu^+\mu^-$. A partir dels dos muons reconstruïts s'extrauen uns mapes de correccions bidimensionals que inclouen tota la regió de l'espai abastada pel detector. Cadascun dels mapes porta una informació: δd_0 (δz_0), que és la diferència en el paràmetre d'impacte transversal (longitudinal)[10] entre els dos muons; i $\delta_{sagitta}$, que és la diferència entre la *sagitta*[11] dels dos muons. Aquests mapes s'utilitzen com una restricció addicional en el procediment d'alineament. Es tracta d'un procés iteratiu, representat a la Figura R.7.

[10] Els paràmetres d'impacte d_0 i z_0 determinen la distància entre una traça i el perigeu, que és la distància més pròxima entre la traça i la direcció dels feixos.

[11] La sagitta determina quant es desvia una trajectòria respecte a una línia recta. Com més gran és la sagitta, més gran és la curvatura de la traça.

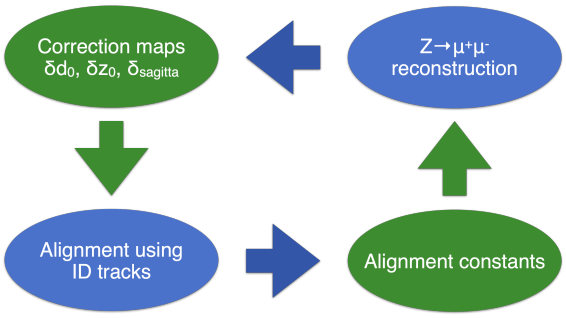


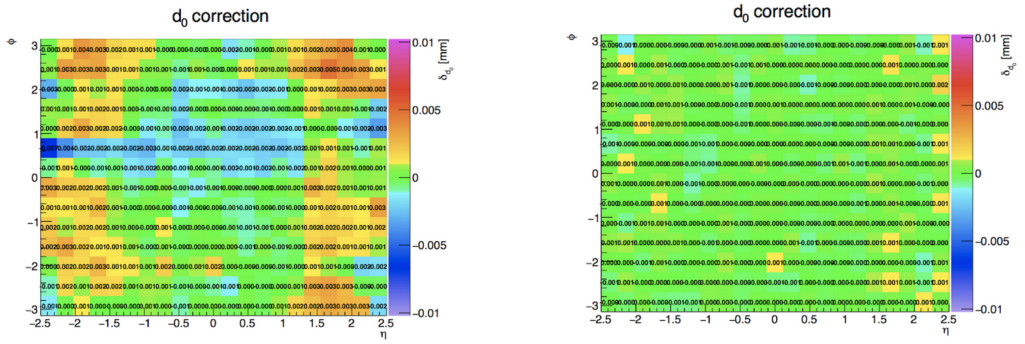
Figure R.7: Representació esquemàtica del procés iteratiu per a la derivació d'unes constants de referència lliures de modes febles emprant la informació dels esdeveniments $Z \rightarrow \mu^+\mu^-$. Les bafarades blaves contenen els dos passos (reconstrucció dels esdeveniments $Z \rightarrow \mu^+\mu^-$ i alineament emprant les traces), mentre que les bafarades verdes inclouen el resultat de cada procés (mapes de correccions i constants d'alineament actualitzades, respectivament) que s'introdueixen com a informació addicional per al pas següent.

Els mapes de correccions abans i després de l'alineament es troben a les Figures R.8-R.10. El nivell de precisió assolit és molt elevat, reduint els biaixos en el paràmetre d'impacte transversal (longitudinal) a $\sim$1 μm ($\sim$5 μm excepte en les regions menys centrals del detector, que tenen menys estadística), i $\sim$0.15 TeV^{-1} per a la sagitta.

R.4 Mesura de seccions eficaces diferencials com a funció de variables angulars sensibles a la polarització del quark top

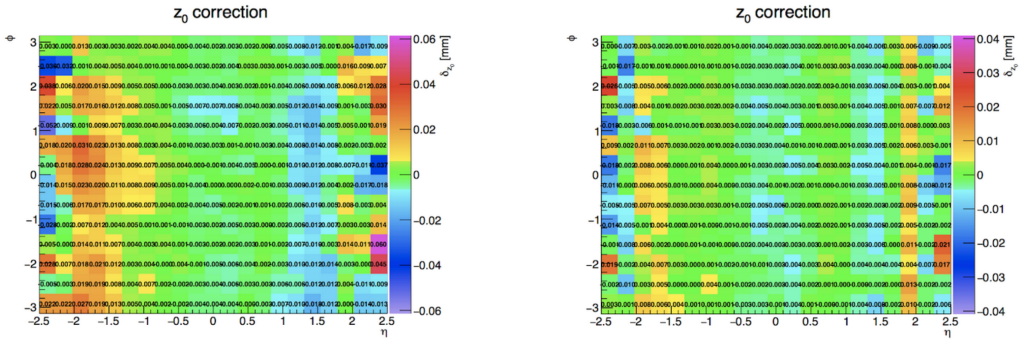
L'anomenat canal t és el mètode dominant de producció de quarks top en solitari. Aquest procés es caracteritza per la interacció d'un quark lleuger[12] amb un quark b, que intercanvien un bosó W produint un altre quark lleuger, anomenat *quark espectador*, i un quark top. La Figura R.11 mostra el diagrama de Feynman associat a aquest procés, on s'inclou la desintegració leptònica del bosó W. Cal dir que el quark b provinent del gluó inicial donaria lloc a un jet que cau fora

[12]Quark lleuger es refereix a qualsevol tipus de quark excepte el quark b i el quark top.



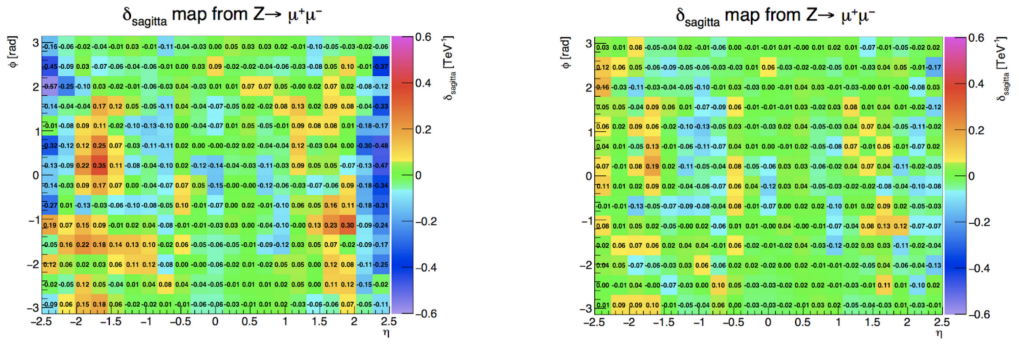
(a) (b)

Figure R.8: Mapes de correccions del paràmetre d'impacte transversal abans (a) i després (b) del realineament de 2018.



(a) (b)

Figure R.9: Mapes de correccions del paràmetre d'impacte longitudinal abans (a) i després (b) del realineament de 2018.



(a) (b)

Figure R.10: Mapes de correccions de la sagitta abans (a) i després (b) del realineament de 2018.

de la regió d'acceptància del detector, i que per tant no es detecta. Com es pot veure en el diagrama, el vèrtex tWb està involucrat tant en la producció com en la desintegració dels quarks top. En el marc del SM, la interacció al vèrtex tWb està donada pel lagrangià:

$$\mathcal{L}_{tWb} = -\frac{g}{\sqrt{2}}\bar{b}\gamma^{\mu}V_{tb}P_L tW^{-}_{\mu} + h.c., \tag{R.1}$$

on g és la constant de l'acoblament dèbil, $\bar{b}$, t i W són els corresponents camps fermiònics i bosònics, γ^{μ} són les matrius gamma, V_{tb} és l'element corresponent a la matriu Cabibbo-Kobayashi Maskawa (CKM), i $P_L = \frac{1}{2}(1-\gamma_5)$ és l'operador de projecció levògira.

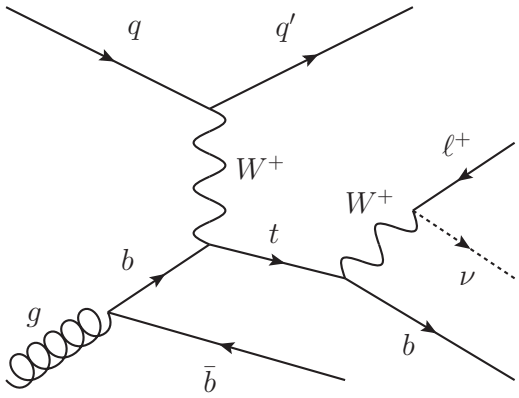


Figure R.11: Diagrama de Feynman a primer ordre que representa la producció d'un quark top solitari en el canal t amb la seua posterior desintegració leptònica.

A conseqüència de l'estructura vector-axial (V-A) del vèrtex tWb, els quarks top produïts en el canal t estan polaritzats. Això significa que el seu espín s'orienta en una direcció en particular, que en aquest cas coincideix amb la direcció dels quarks baix [111, 112]. A més a més, per la proximitat entre la direcció del feix i el quark espectador es produeix un alt grau de polarització dels quarks (antiquarks) top al llarg de (en contra de) la direcció del quark espectador. Aquest fenomen és característic de la producció de quarks top en solitari, degut a que la producció en parelles es produeix a través de la interacció forta, que conserva la paritat [113].

Aquesta anàlisi mesura observables angulars que són sensibles a la polarització del quark top. Per poder realitzar aquesta mesura, el primer de tot és establir un sistema de coordenades [2]. En el sistema de referència del quark top, s'escull l'eix $\hat{z}'$ en la direcció del quark espectador. L'eix $\hat{y}'$ es tria perpendicularment a l'eix $\hat{z}'$ i a la direcció del quark lleuger inicial. Finalment, l'eix $\hat{x}'$ roman al pla de producció, ortogonal als eixos $\hat{z}'$ i $\hat{y}'$, formant un sistema de coordenades dextrògir.

Gràcies a la seua gran massa [88], el quark top es desintegra molt ràpidament, transferint la informació del seu espín als seus productes de desintegració. Les tases de desintegració en el sistema de referència del quark top es poden expressar com:

$$\frac{1}{\Gamma}\frac{d\Gamma}{d\cos\theta_{Xi}} = \frac{1}{2}(1 + \alpha_X P_i \cos\theta_{Xi}), \tag{R.2}$$

on α_X és el poder d'anàlisi de l'espín del producte de desintegració escollit ($X = W, l, \nu, b$ en el canal de desintegració leptònica), θ_{Xi} és l'angle entre la direcció de moviment del producte de desintegració X i l'axis $i = (\hat{x}', \hat{y}', \hat{z}')$, i P_i és la polarització del quark top en un eix determinat i.

D'entre els possibles productes de desintegració del quark top, el leptó és amb molta diferència el que té major poder d'anàlisi de l'espín [116]. Per aquesta raó, aquesta tesi mesura la secció eficaç diferencial del canal t com a funció de les tres distribucions angulars del leptó en el sistema de coordenades referit prèviament i representat a la Figura R.12.

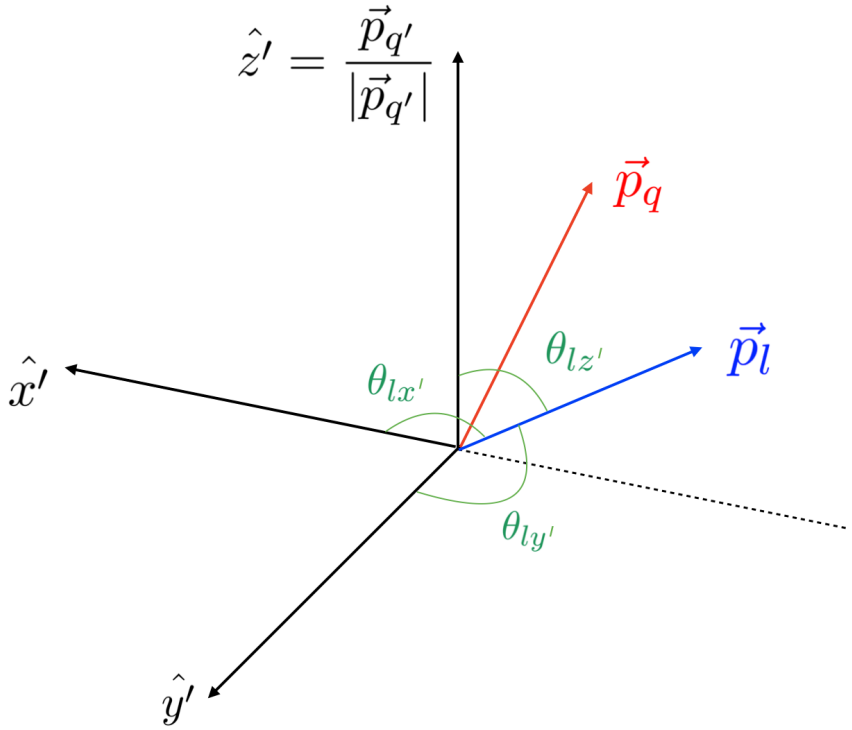


Figure R.12: Sistema de coordenades dextròg ir escollit per mesurar els tres angles del leptó ($\cos\theta_{lx'}, \cos\theta_{ly'}, \cos\theta_{lz'}$) provinent de la desintegració del quark top en el sistema de referència d'aquest últim. $\vec{p}_{q'}$ és el vector moment del quark espectador. La seua direcció es correspon amb l'eix z'. $\vec{p}_q$ és el vector moment del quark lleuger inicial, mentre que $\vec{p}_l$ és el vector moment del leptó.

Aquestes mesures són particularment sensibles a efectes de nova física en el vèrtex tWb. Una manera de descriure possibles desviacions del SM independent de qualsevol model és a partir de les teories de camps efectius (EFT). Si l'escala d'energia Λ a la que apareix la nova física està fora de l'abast de l'LHC, aleshores els efectes de nova física es poden parametritzar en termes d'un lagrangià efectiu [117]:

$$\mathcal{L}_{\text{EFT}} = \mathcal{L}_{\text{SM}} + \sum_k \frac{C_k}{\Lambda^2}\mathcal{O}_k^{[6]} + ..., \tag{R.3}$$

on $\mathcal{L}_{\text{SM}}$ és el lagrangià del SM, Λ és l'escala de nova física, $\mathcal{O}_k^{[6]}$ són els operadors de dimensió sis i C_k són els seus coeficients de Wilson associats. D'aquesta manera, el lagrangià del SM s'expandeix introduint operadors de dimensió major. En aquesta aproximació, l'escala d'energia Λ ha estat escollida de forma que els operadors de dimensions majors que sis estan suprimits. A més, s'han omés els operadors de dimensió cinc, ja que no contribueixen ni a la producció ni a la desintegració del quark top [428].

D'entre tots els possibles operadors, només n'hi ha un que tindria un efecte directe en aquestes mesures [123]. En concret, es tracta de l'operador dipolar $\mathcal{O}_{tW}$. La part real i imaginària del seu coeficient de Wilson es representen com C_{tW} i C_{itW}, respectivament. El coeficient C_{tW} (C_{itW}) afecta principalment a $P_{x'}$ ($P_{y'}$). D'altra banda, un valor diferent de zero en la part imaginària d'aquest coeficient implicaria una violació de la simetria CP[13] en el vèrtex tWb i podria contribuir a explicar l'asimetria entre matèria i antimatèria de l'univers.

Prèviament s'han mesurat alguns observables de l'espín del quark top tant en ATLAS com en CMS [11, 126–130]. En concret, s'ha mesurat la polarització del quark top al llarg de l'eix z', de manera directa i indirecta [128–130], i també el vector de polarització com a part de la tesi en la referència [11]. Aquestes mesures es van realitzar de manera conjunta tant per a quarks com per a antiquarks top. Altrament, aquestes mesures es realitzen a nivell de partons corregint per efectes d'acceptància, d'hadronització, així com d'eficiència i resolució del detector, de tal forma que es poden comparar amb les prediccions teòriques. Tanmateix, es va provar que els efectes de nova física afecten les correccions esmentades [131]. Una possibilitat per a sobreposar-se a aquest problema és realitzar la mesura a nivell de partícules[14] en una regió *fiducial*[15]. D'aquesta manera, les mesures proporcionades són independents de cap assumpció sobre la física subjacent.

Aquesta anàlisi fa ús de 139 fb^{-1} de dades recollides pel detector ATLAS durant el Run-2 per a col·lisions protó-protó a una energia en el centre de masses de $\sqrt{s} =$

[13] C representa la conjugació de càrrega i P la paritat d'una partícula.

[14] A nivell de partícules els objectes es construeixen a partir de partícules estables sense tindre en compte els efectes del detector.

[15] Aquesta és una regió de l'espai de fases que es construeix el més proper possible a la regió que és accessible experimentalment.

13 TeV. Les mesures es proporcionen a nivell de partícules de forma conjunta i separada per a quarks i antiquarks top. Per altre costat, es presenta una comparació d'aquestes mesures amb la mesura directa dels vectors de polarització de quarks i antiquarks top [3], així com els límits extrets a partir d'aquestes mesures per als coeficients C_{tW} i C_{itW} com a part del treball de tesi en la referència [4].

R.4.1 Selecció d'esdeveniments

L'anàlisi descrita en aquesta secció se centra en el canal de desintegració leptònic del quark top ($t \to Wb, W \to l\nu$, amb $l = e, \mu$). Esdeveniments en els quals el bosó W es desintegra a un leptó τ i un neutrí es consideren només si el leptó τ es desintegra a un electró o un muó més un neutrí addicional. En definitiva, l'estat final considerat en aquesta anàlisi està caracteritzat per la presència d'un leptó, dos jets (un d'ells identificat com a b-jet) i una gran quantitat d'energia transversa mancant a causa del neutrí.

Hi ha una sèrie de processos al SM que poden representar el mateix estat final i d'aquesta manera contaminarien els nostres esdeveniments de senyal. Aquest tipus de processos s'anomenen fons. Els principals fons en aquesta anàlisi inclouen: la producció de parelles de quarks top ($t\bar{t}$), els quarks top produïts en solitari en el canal s i el canal tW, la producció de bosons W o Z amb jets (W/Z+jets), la producció de processos dibosònics (WW, ZZ i ZW), i la producció d'esdeveniments de jets múltiples (***multijet***), on un dels jets s'identifica erròniament amb un leptó.

Per tal de desfer-se del fons, es defineix una regió de senyal aplicant una sèrie de requeriments que fomenten la presència d'esdeveniments del canal t sobre el fons. En primer lloc, es requereix la presència d'exactament un leptó (electró o muó), així com de dos jets (essent un d'ells un b-jet) i una alta energia transversa mancant ($E_{\mathrm{T}}^{\mathrm{miss}} > 35$ GeV). S'apliquen dos requeriments addicionals per reduir el fons de multijet. En primer lloc, la massa transversa del sistema leptó–$E_{\mathrm{T}}^{\mathrm{miss}}$:

$$m_T(l, E_{\mathrm{T}}^{\mathrm{miss}}) = \sqrt{2p_{\mathrm{T}}(l)E_{\mathrm{T}}^{\mathrm{miss}}(1 - \cos\Delta\phi(p_{\mathrm{T}}(l), E_{\mathrm{T}}^{\mathrm{miss}}))}, \tag{R.4}$$

on $p_{\mathrm{T}}(l)$ és el moment transvers del leptó i $\Delta\phi(p_{\mathrm{T}}(l), E_{\mathrm{T}}^{\mathrm{miss}})$ és la diferència angular entre el moment del leptó i la direcció de la $E_{\mathrm{T}}^{\mathrm{miss}}$, ha de ser major que 60 GeV. En segon lloc, s'aplica el següent requeriment:

$$p_{\mathrm{T}}(l) \; > \; 50 \left(1 - \frac{\pi - |\Delta\phi(p_{\mathrm{T}}(j_1), p_{\mathrm{T}}(l))|}{\pi - 1}\right) \mathrm{GeV}, \tag{R.5}$$

on $p_{\mathrm{T}}(j_1)$ és el moment transvers del jet de major moment i $\Delta\phi(p_{\mathrm{T}}(j_1), p_{\mathrm{T}}(l))$ és la diferència angular entre el leptó i el jet de major moment. Aquest requeriment elimina esdeveniments on els dos jets es generen en direccions oposades, el qual és un patró característic del fons de multijet.

La selecció d'esdeveniments explicada fins ara defineix la regió de *preselecció*. Per a incrementar encara més el senyal sobre el fons s'apliquen els següents requeriments, que defineixen l'anomenada *regió de senyal*:

- La massa invariant del sistema leptó–b-jet, m_{lb}, ha de ser major que 153 GeV.
- La massa del quark top reconstruït, $m_{l\nu b}$, ha d'estar entre 120.6 i 234.6 GeV.
- S'aplica un requeriment trapezoidal per a eliminar esdeveniments de fons que tenen jets lleugers en una direcció propera als feixos amb quarks top reconstruïts en la regió central del detector:

$$\begin{aligned} \eta_j &< (4\,\eta_{l\nu b} + a) \quad \& \\ \eta_j &> (4\,\eta_{l\nu b} - a) \quad \& \\ (\eta_j &> (0.44\,\eta_{l\nu b} + b) \quad \mathsf{OR} \\ \eta_j &< (0.44\,\eta_{l\nu b} - b))\,, \end{aligned} \tag{R.6}$$

  on η_j i $\eta_{l\nu b}$ representen la pseudorapidesa[16] del quark espectador i del quark top reconstruït, respectivament. Els paràmetres a i b defineixen el trapezoide, i prenen els valor 10 i 2, respectivament.

[16] La pseudorapidesa es defineix com $\eta = -\ln\tan(\theta/2)$, on θ es l'angle polar que representa la distància angular entre la direcció d'una partícula i la direcció dels feixos. Grans valors de η indiquen que la partícula s'ha produït en una direcció propera als feixos, mentre que valors petits de η indiquen que s'ha produït en la regió central del detector.

- La massa del sistema quark-espectador–top-quark $m_{jl\nu b}$ ha de ser major que 320 GeV.

- La suma escalar del moment transvers de tots els objectes a l'estat final, H_{T}, ha de ser major que 190 GeV.

Les seccions eficaces diferencials que es presenten a aquesta anàlisi es proporcionen a nivell de partícules estables en una regió fiducial. Aquesta regió es defineix el més a prop possible a la regió de senyal, aplicant requeriments molt similars però sobre objectes a nivell de partícules.

R.4.2 Estimació del fons

La normalització del senyal i dels diferents fons (excepte el fons de multijet) s'obté a partir de la secció eficaç absoluta de cada procés així com de la lluminositat emprada en aquesta anàlisi. Per al fons de multijet, s'empren les tècniques del model jet-electró [388] i el model anti-muó [388], i s'obté la seua normalització a partir d'un ajust a les dades. Per a constrényer la normalització dels fons principals ($t\bar{t}$ i W+jets) s'aplica un procediment addicional. Concretament, es realitza un ajust al nombre d'esdeveniments mesurats en les dades en dues regions específiques dominades per aquests fons, així com en la regió de senyal. Aquestes regions específiques s'anomenen *regions de control*, i serveixen a més per a verificar el correcte modelatge dels fons que representen. Es defineixen de la següent manera:

- La regió de control de $t\bar{t}$ es defineix aplicant els mateixos requeriments que a la regió de preselecció, però es requereix que els dos jets siguen ambdós b-jets.

- En la regió de control de W+jets se seleccionen tots els esdeveniments que han passat els requeriments de la regió de preselecció, però no han passat algun dels requeriments de la regió de selecció.

Les Figures R.13-R.14 mostren les distribucions d'algunes variables cinemàtiques en les regions de control, després d'aplicar els factors de normalització addicionals provinents de l'ajust a les dades. Com es pot

apreciar, hi ha un bon acord entre la simulació i les dades, la qual cosa confirma un modelatge correcte dels processos del SM.

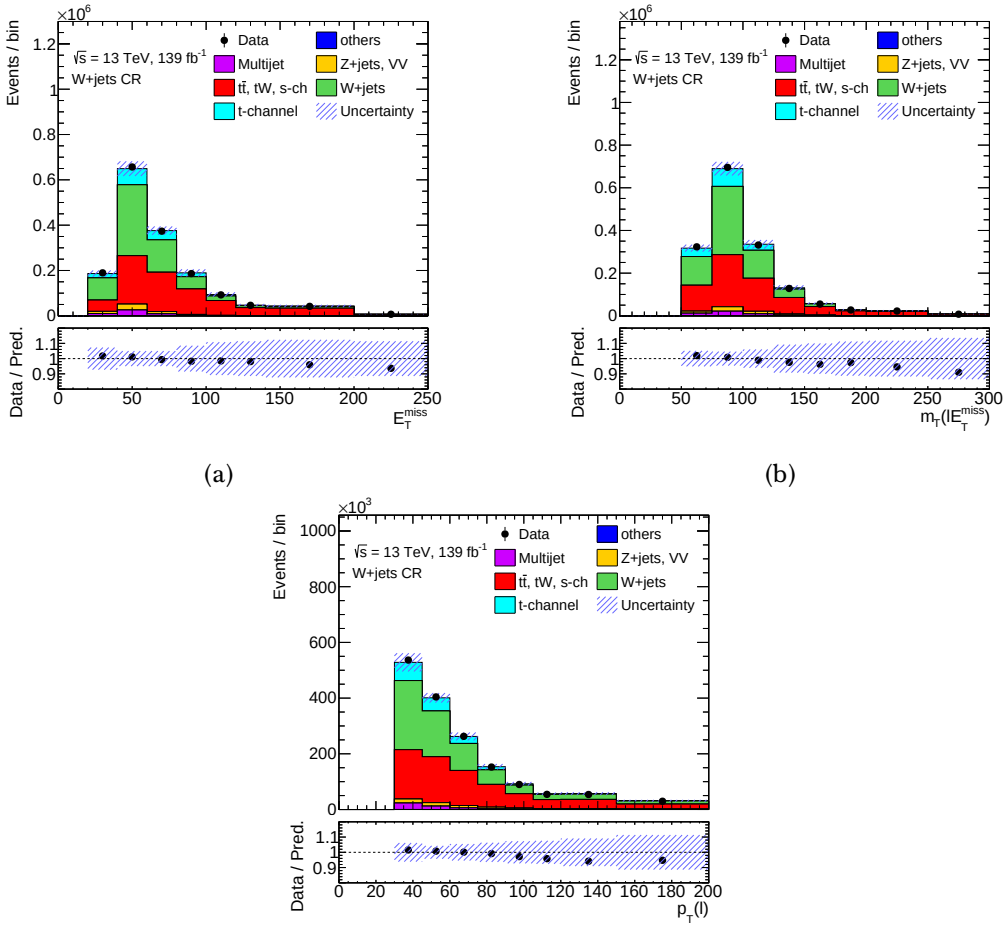


Figure R.13: Distribucions de (a) $E_{\mathrm{T}}^{\mathrm{miss}}$, (b) $m_T(l, E_{\mathrm{T}}^{\mathrm{miss}})$ i (c) $p_{\mathrm{T}}(l)$ en la regió de control de W+jets, on es compara l'acord entre dades i simulacions de Monte Carlo dels processos del SM. Els esdeveniments del canal t, així com els dels fons de $t\bar{t}$ i W+jets estan escalats utilitzant els resultats d'un ajust al nombre d'esdeveniments de dades en les regions de control i la regió de senyal.

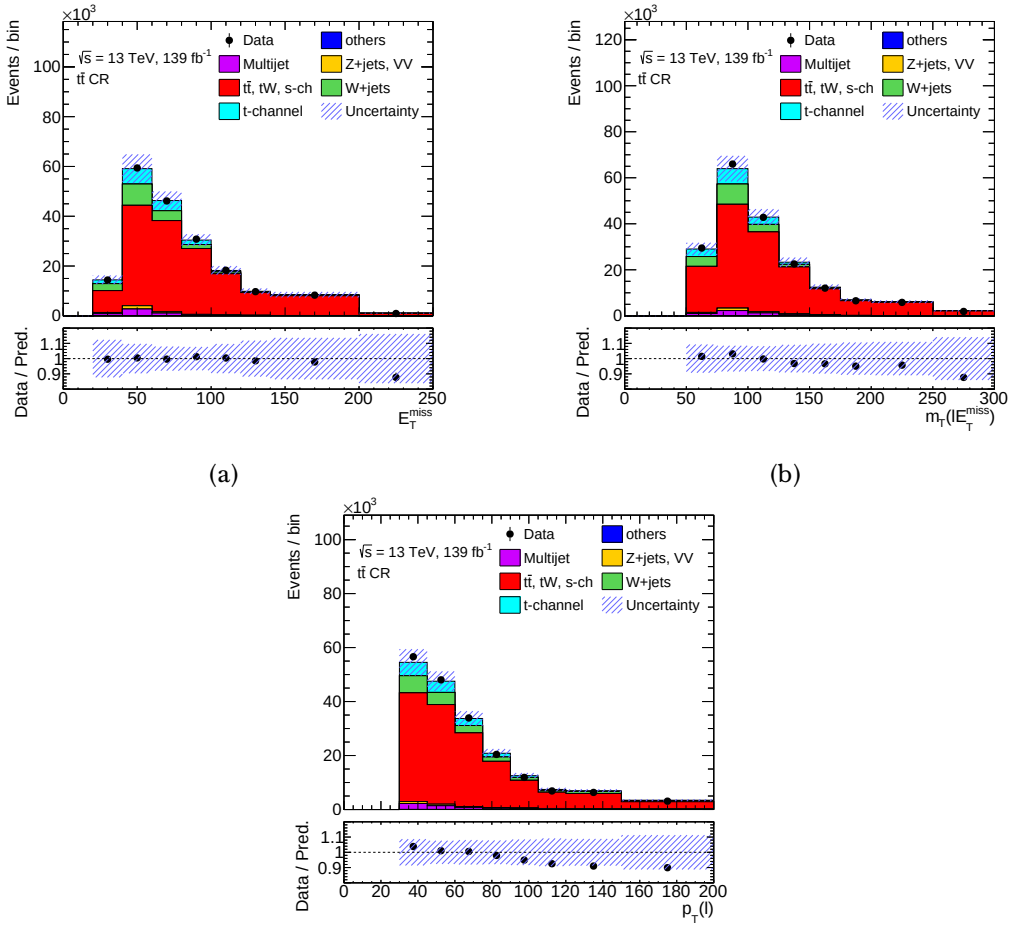


Figure R.14: Distribucions de (a) $E_{\mathrm{T}}^{\mathrm{miss}}$, (b) $m_T(l, E_{\mathrm{T}}^{\mathrm{miss}})$ i (c) $p_{\mathrm{T}}(l)$ en la regió de control de $t\bar{t}$, on es compara l'acord entre dades i simulacions de Monte Carlo dels processos del SM. Els esdeveniments del canal t, així com els dels fons de $t\bar{t}$ i W+jets estan escalats utilitzant els resultats d'un ajust al nombre d'esdeveniments de dades en les regions de control i la regió de senyal.

R.4.3 Correcció a nivell de partícules

Les distribucions de les tres variables angulars ($\cos\theta_{lx'}, \cos\theta_{ly'}, \cos\theta_{lz'}$) a nivell de reconstrucció es mostren a la Figura R.15. Com es pot observar, hi ha un bon acord entre la simulació i les dades.

Per poder comparar aquests resultats amb les prediccions teòriques en la regió fiducial, cal sostraure els diferents fons i realitzar una correcció de les distribucions a nivell de partícules. D'aquesta manera, es corregeixen els efectes del detector. Finalment, les distribucions obtingudes són dividides per la integral de la distribució, de tal forma que els resultats es presenten com seccions eficaces diferencials normalitzades.

Per a corregir els esdeveniments a nivell de partícules es planteja el problema com un "efecte" i una "causa". El nombre d'esdeveniments de senyal a nivell de reconstrucció en un bin j és l'efecte, mentre que el nombre d'esdeveniments del canal t produïts com a resultat de les col·lisions és la causa. Emprant aquesta tècnica, podem recuperar les distribucions a nivell de partícules a partir de les distribucions de senyal reconstruïdes:

$$\nu_k^{\text{particle}} = C_k^{\text{particle!reco}} \sum_j M_{jk}^{-1} C_j^{\text{reco!particle}} (N_j^{\text{data}} - B_j), \tag{R.7}$$

on ν_k^{particle} és el nombre d'esdeveniments de senyal a nivell de partícules en un bin k, $N_j^{\text{data}} - B_j$ és el nombre esperat d'esdeveniments de senyal reconstruïts (com a resultat de llevar tot el fons B_j a les dades mesurades N_j^{data}) en un bin j, i M_{jk} és l'anomenada ***matriu de migració***. Finalment, $C_k^{\text{particle!reco}}$ i $C_j^{\text{reco!particle}}$ són uns factors de correcció que tenen en compte els esdeveniments que han passat els requeriments a nivell de partícules però no a nivell de reconstrucció i viceversa, respectivament.

Com es pot apreciar, per a passar del nivell de reconstrucció al nivell de partícules cal invertir la matriu de migració. En general, aquesta matriu no té una inversa exacta, i per tal d'invertir-la cal fer aproximacions. En aquesta anàlisi s'utilitza una tècnica iterativa bayesiana [391] per a realitzar la susdita inversió.

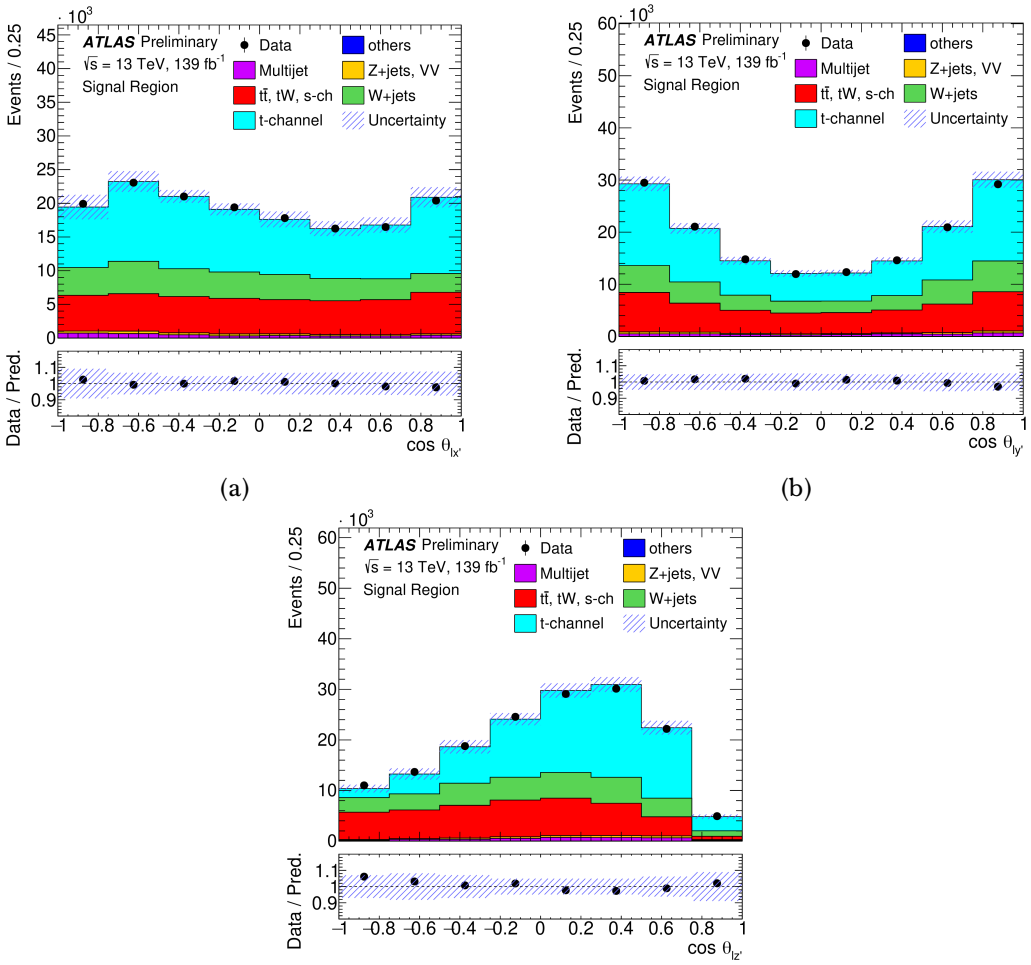


Figure R.15: Distribucions de (a) $\cos\theta_{lx'}$, (b) $\cos\theta_{ly'}$ i (c) $\cos\theta_{lz'}$ en la regió de senyal, on es compara l'acord entre dades i simulacions de Monte Carlo dels processos del SM. Els esdeveniments del canal t, així com els dels fons de $t\bar{t}$ i W+jets estan escalats utilitzant els resultats d'un ajust al nombre d'esdeveniments de dades en les regions de control i la regió de senyal.

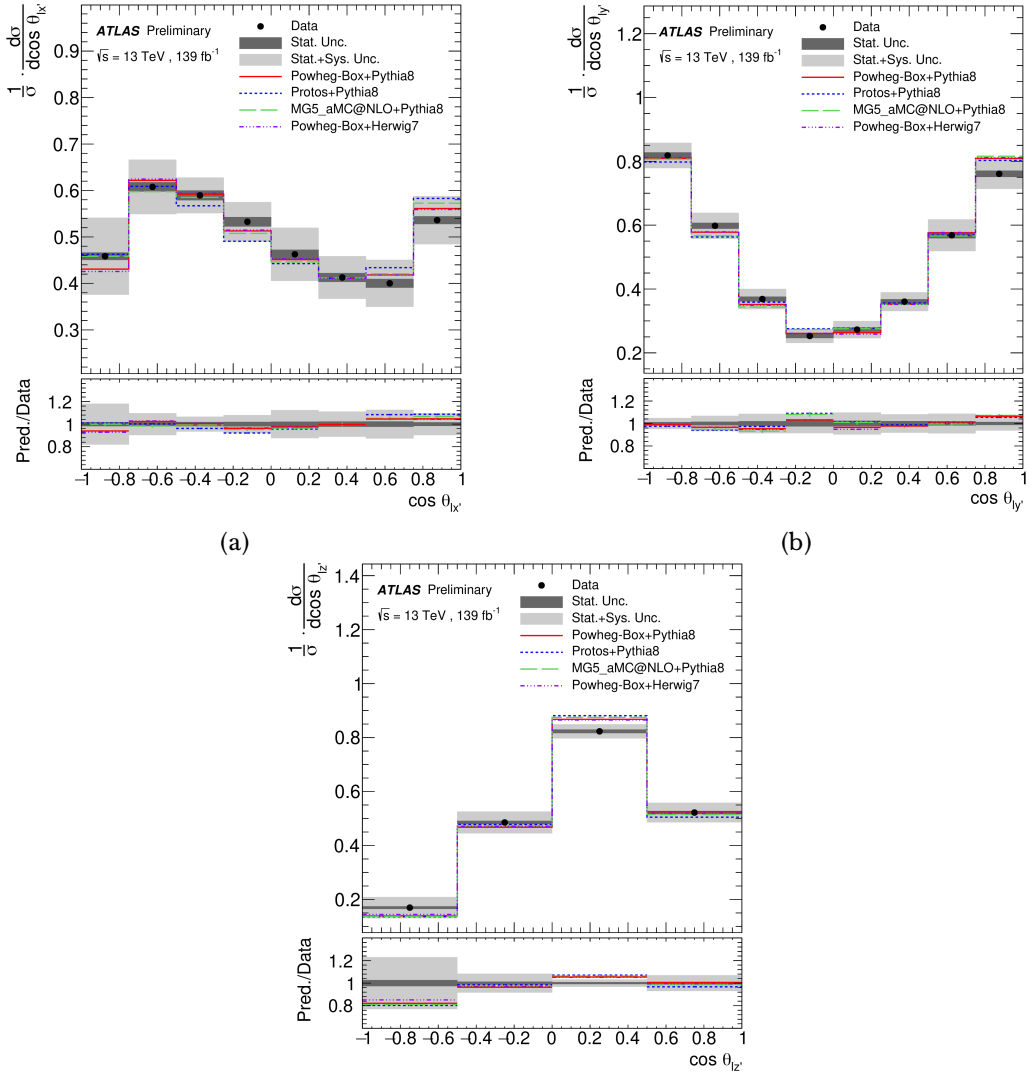
La matriu de migració i els factors de correcció s'obtenen utilitzant mostres simulades a segon ordre (NLO) d'esdeveniments del canal t, emprant el generador POWHEG-BOX+PYTHIA8 [261, 351, 354–357]. El mètode ha estat validat realitzant diferents tests. En primer lloc, es realitza un test de convergència, que determina el nombre d'iteracions emprades en la inversió de la matriu i que mostra que el procés iteratiu convergeix. En segon lloc es realitza un test d'estabilitat estadística, que mostra que la resposta de les correccions és precisa i lliure de biaixos. Finalment, es realitza un test de linealitat que confirma que les correccions són estables independentment de la física subjacent. Aquest test és cabdal per provar que la presència de nova física no esbiaixaria la mesura.

R.4.4 Fonts d'incertesa

Les mesures proporcionades en aquesta anàlisi es veuen afectades per diferents fonts d'incertesa. Aquestes es poden dividir en dos grups: estadístiques (tant de dades com d'esdeveniments simulats) i sistemàtiques. Altrament, les incerteses sistemàtiques es poden dividir en experimentals (modelatge del detector) i teòriques (modelatge del senyal i del fons). L'impacte de cadascuna de les fonts sistemàtiques s'obté a partir de mostres simulades per reflectir cadascuna de les fonts d'incertesa. Aquestes mostres es propaguen tant a la regió de senyal com a les regions de control. Després d'avaluar l'impacte de cada font d'incertesa, aquestes s'afegeixen quadràticament per obtenir l'error final de la mesura. Les fonts principals d'incertesa en aquesta anàlisi venen de les incerteses experimentals relacionades amb els jets, seguides per les incerteses en el modelatge del senyal i del fons de $t\bar{t}$.

R.4.5 Resultats

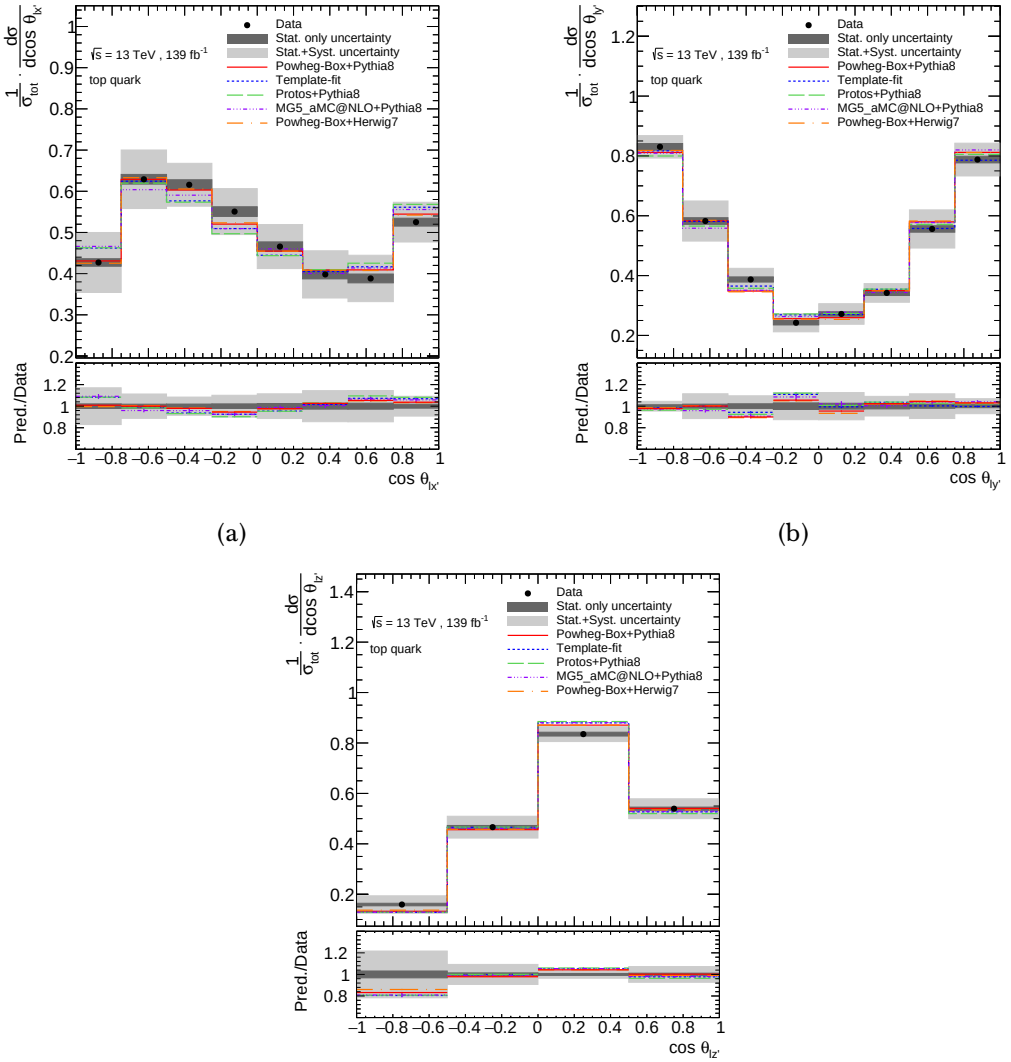
Les seccions eficaces diferencials a nivell de partícules com a funció de les tres variables angulars proposades es presenten a les Figures R.16-R.18 per a quarks i antiquarks top, tant de forma conjunta com separadament. Aquestes es comparen amb prediccions donades pels diferents generadors. Els resultats són compatibles amb les prediccions del SM.



(a) (b)

(c)

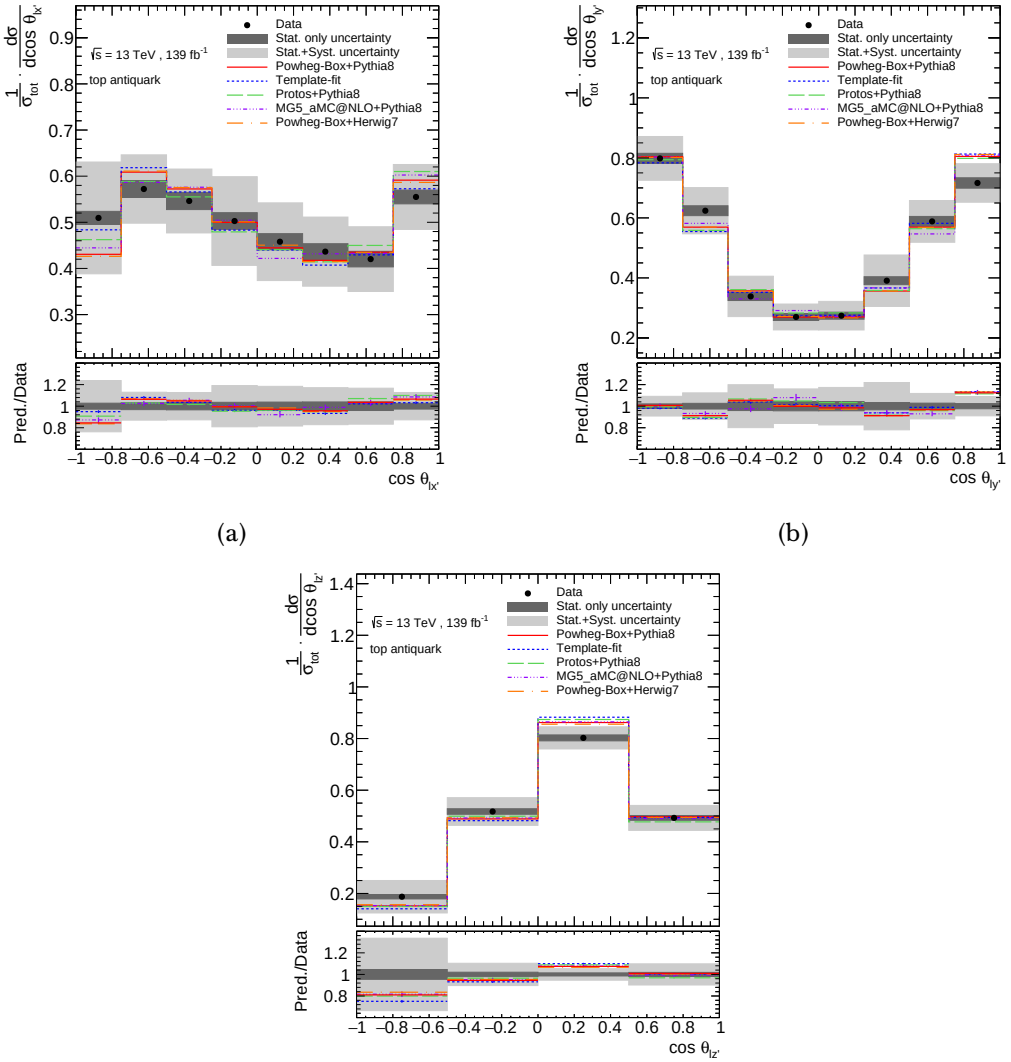
Figure R.16: Seccions eficaces diferencials normalitzades en funció de (a) $\cos\theta_{lx'}$, (b) $\cos\theta_{ly'}$ i (c) $\cos\theta_{lz'}$, comparades amb diverses prediccions del SM obtingudes per simulació Monte Carlo, per a la mesura conjunta de quarks i antiquarks top.



(a) (b)

(c)

Figure R.17: Seccions eficaces diferencials normalitzades en funció de (a) $\cos\theta_{lx'}$, (b) $\cos\theta_{ly'}$ i (c) $\cos\theta_{lz'}$, comparades amb diverses prediccions del SM obtingudes per simulació Monte Carlo, per a la mesura de quarks top. Els resultats també es comparen amb la mesura directa del vector de polarització del quark top, referenciada com a *template-fit*.



(a) (b)

(c)

Figure R.18: Seccions eficaces diferencials normalitzades en funció de (a) $\cos\theta_{lx'}$, (b) $\cos\theta_{ly'}$ i (c) $\cos\theta_{lz'}$, comparades amb diverses prediccions del SM obtingudes per simulació Monte Carlo, per a la mesura d'antiquarks top. Els resultats també es comparen amb la mesura directa del vector de polarització del quark top, referenciada com a *template-fit*.

Els resultats obtinguts per als quarks i antiquarks top per separat es comparen també amb la mesura directa dels vectors de polarització [3], resumida en la Taula R.1 (referenciada com a *template-fit* a la figura). Com es pot veure, hi ha un bon acord entre ambdues mesures.

Paràmetre	Valor mesurat	(estad.)
$P^t_{x'}$	$+0.01 \pm 0.18$	(± 0.02)
$P^{\bar{t}}_{x'}$	-0.02 ± 0.20	(± 0.03)
$P^t_{y'}$	-0.029 ± 0.027	(± 0.011)
$P^{\bar{t}}_{y'}$	-0.007 ± 0.051	(± 0.017)
$P^t_{z'}$	$+0.91 \pm 0.10$	(± 0.02)
$P^{\bar{t}}_{z'}$	-0.79 ± 0.16	(± 0.03)

Table R.1: Vectors de polarització trets a prtir d'un ajust a les dades, incloent les incerteses totals i la incertesa purament estadística [3].

A més, aquests resultats s'han emprat per a obtenir límits d'exclusió per als coeficients C_{tW} i C_{itW} com a part del treball de tesi en la referència [4]. Els valors mesurats C_{tW} = 0.4 $\pm$ 1.1 i C_{itW} = -0.3 $\pm$ 0.4 a un 95% de nivell de confiança (C.L.) són compatibles amb les prediccions del SM. Aquests límits són compatibles amb els límits actuals per a la part real de l'operador [118, 129, 402–404], i milloren els límits actuals en el camp de la física d'altes energies per a la part imaginària [406], que a més és sensible a efectes de violació CP en el vèrtex tWb.

R.5 Cerca de matèria fosca produïda amb un quark top solitari

Les observacions astronòmiques i cosmològiques han determinat que un 26.8% de l'univers està format per un tipus de matèria que no emet ni absorbeix llum. Aquesta és l'anomenada matèria fosca. Se sap que aquesta matèria interactua principalment a través de la interacció gravitatòria, però poc més se sap sobre les seues propietats. La comunitat científica està realitzant un gran esforç per entendre la natura de la DM, la qual cosa augmentaria exponencialment el nostre coneixement de l'univers. El treball d'aquesta tesi contribueix a aquest esforç a

través d'una cerca de DM en el sector de la física dels col·lisionadors, on s'espera que les partícules del SM produïsquen partícules de DM a través de la interacció feble.

Malgrat que encara no s'ha pogut detectar la DM, hi ha moltes evidències de la seua existència. D'entre elles, destaquen les corbes gravitacionals de les galàxies. Aquestes corbes representen la velocitat de les diferents estreles respecte al centre de la galàxia. Assumint que la major part de la matèria es troba al centre lluminós de la galàxia, la velocitat de les estreles s'hauria de reduir amb la distància respecte al centre, seguint la llei de gravitació clàssica de Newton. No obstant això, el que s'observa és que la velocitat de les estreles augmenta o roman constant a mesura que s'allunyen del centre. Aquestes observacions són consistents amb l'existència d'un halo de DM que envolta a la galàxia. Altres exemples inclouen les observacions de lents gravitacionals [137–139] o les anisotropies observades en el fons còsmic de microones [147].

Atés que la DM no s'ha detectat, existeixen diversos models a la literatura que proposen diferents candidats per explicar-la. Tanmateix, un bon candidat a DM ha de complir les següents condicions [156]:

- Tindre un mecanisme de producció que prediga la quantitat de DM observada en l'actualitat a partir de les anisotropies en el fons còsmic de microones.
- Ser no-relativista, el que es coneix com a "freda".
- Ser neutre.
- Ser consistent amb la nucleosíntesi primordial, que prediu l'abundància d'elements lleugers produïts en els primers tres minuts després del big bang.
- No alterar l'evolució de les estreles.
- Ser compatible amb els límits establerts per a les interaccions entre partícules de DM [164].
- Ser compatible amb els límits experimentals provinents de la detecció directa, indirecta, i de les cerques als col·lisionadors.

- Ha de ser possible que un bon candidat de DM puga ser investigat experimentalment. Si no es pot trobar evidència de la seua existència, aleshores no es considera un bon candidat.

Hi ha diversos candidats que poden complir aquestes condicions. Cal destacar els neutrins estèrils, els axions, els fotons foscos, els objectes astronòmics i les partícules massives d'interacció feble (WIMPs). A les cerques dels col·lisionadors, es busca la producció de partícules WIMP junt amb algun objecte energètic del SM, que pot ser un jet, un fotó, un bosó o un quark top. El WIMP produït no interaccionaria amb el material del detector, i per tant escaparia sense ser detectat, com fan els neutrins. Aquestes cerques es caracteritzen doncs per un objecte energètic i una gran quantitat d'energia transversa mancant.

La segona anàlisi descrita en aquest treball presenta una cerca de DM produïda amb un quark top. Aquest estat final s'anomena *mono-top*. El model simplificat descrit en les referències [223–225] comprén els diferents processos que donen lloc a aquest tipus d'estat final. En aquesta tesi, ens centrem en dos mecanismes de producció en particular: el model *ressonant* i el model *no-ressonant*.

El model ressonant prediu l'existència d'una partícula escalar (ϕ) amb càrrega de color que es desintegra a un quark top i una partícula de DM d'espín 1/2 (χ), com es veu al diagrama de Feynman de l'esquerra a la Figura R.19. El lagrangià d'aquest procés es defineix de la següent forma [223, 226]:

$$\mathcal{L}_{\text{res}} = \lambda\phi\bar{d}^{c}P_{R}s + y\phi\bar{\chi}P_{R}t + h.c., \tag{R.8}$$

on d, s i t representen els camps del quark d, del quark s i del quark top, respectivament. ϕ i χ representen el camp de l'escalar carregat i de la partícula de DM, respectivament. $P_R = \frac{1}{2}(1+\gamma_5)$ és l'operador de projecció dextrògira. Finalment, λ i y representen els acoblaments entre l'escalar carregat i els quarks d i s, i entre el quark top i la partícula de DM, respectivament.

Els diagrames de Feynman de la dreta i d'enmig a la Figura R.19 representen el model no-ressonant. Com es pot veure, aquest model es caracteritza per un corrent neutre de canvi de sabor entre un quark u, un quark top, i una partícula vectorial V. Aquest bosó vectorial es desintegra a un parell de partícules de DM ($\chi\bar{\chi}$) que

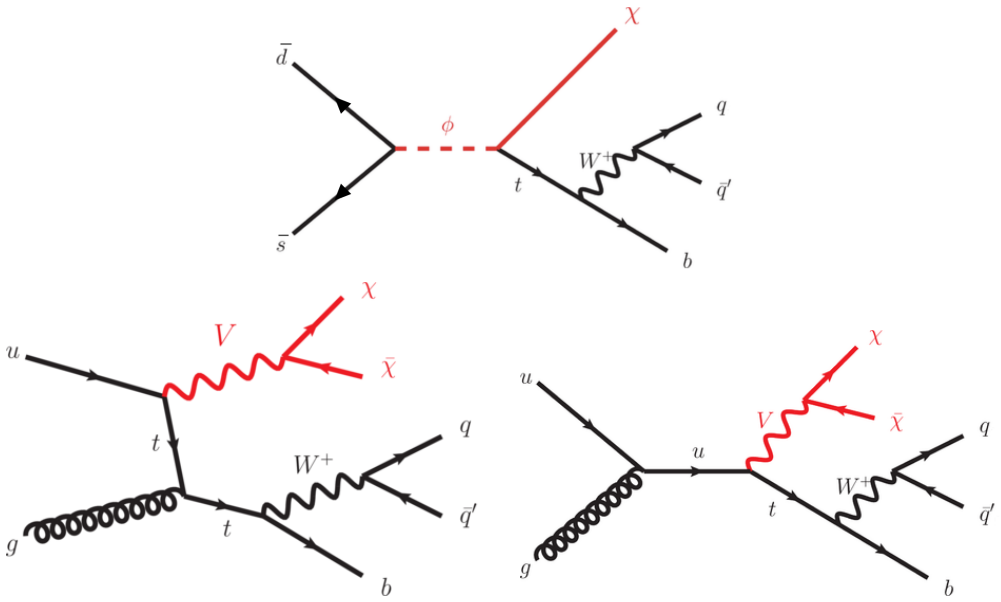


Figure R.19: Diagrames de Feynman associats al model ressonant (esquerra) i al model no-ressonant (mig i dreta). La desintegració hadrònica del quark top s'inclou en els diagrames.

són fermions de Dirac. El lagrangià que descriu aquesta interacció es defineix de la següent manera [226, 233]:

$$\mathcal{L}_{\text{non-res}} = aV_\mu \bar{u}\gamma^\mu P_R t + g_\chi V_\mu \bar{\chi}\gamma^\mu \chi + h.c., \tag{R.9}$$

on u i t són el camp del quark u i del quark top, respectivament, i V_μ i χ són el camp del bosó vectorial i el fermió de Dirac, respectivament. a representa l'acoblament entre el bosó vectorial V, el quark u i el quark top, mentre que g_χ representa l'acoblament entre el bosó vectorial i la partícula de DM χ. γ^μ són les matrius gamma.

Aquest tipus d'estat final ha estat investigat prèviament per ATLAS i CMS [238, 239, 241–243]. Masses de l'escalar carregat (el bosó vectorial) han estat excloses fins a 3.4 TeV (2 TeV) per a $m_\chi = 10$ GeV, $y = 0.4$ i $\lambda = 0.2$ ($m_\chi = 1$ GeV, $g_\chi = 1.0$ i $a = 0.5$) en el model ressonant (no-ressonant) a un 95% C.L..

A l'hora de caracteritzar l'estat final d'aquests processos cal considerar la desintegració del bosó W. Cerques anteriors [241] han demostrat que el canal hadrònic és molt més sensible a aquest tipus de senyal. Així doncs, l'anàlisi presentada en aquesta tesi es centra en la desintegració hadrònica del quark top.

Aquesta anàlisi fa ús de 139 fb^{-1} de dades recollides pel detector ATLAS durant el Run-2 per a col·lisions protó-protó a una energia en el centre de masses de $\sqrt{s} = 13$ TeV.

R.5.1 Selecció d'esdeveniments

En el canal hadrònic, s'espera que els quarks top es produïsquen amb un gran *boost* de Lorentz, i per tant els seus productes de desintegració estan molt col·limats i es poden reconstruir com un jet gran. D'aquesta manera, l'estat final estudiat es caracteritza per un jet gran originat a partir d'un quark top i una gran quantitat d'energia transversa mancant. Hi ha diversos processos al SM que poden replicar aquest estat final i constitueixen el fons en aquesta cerca: processos que inclouen un quark top ($t\bar{t}$ i producció de quarks top en solitari), producció d'un bosó vectorial amb jets (V+jets) i producció de processos dibosònics (WW, ZZ, ZW). El fons de multijet té un efecte insignificant en aquesta anàlisi.

La selecció d'esdeveniments en aquesta anàlisi es realitza en dues parts. En primer lloc, se seleccionen esdeveniments a nivell de preselecció atenent al següent criteri:

- No contenen cap leptó.
- Només contenen jets de bona qualitat[17]
- La $E_{\mathrm{T}}^{\mathrm{miss}}$ ha de ser major que 250 GeV.
- Ha d'haver-hi almenys un top-jet (jet etiquetat com a provinent d'un quark top).

[17] És a dir, no poden tindre cap *loose* jet. Aquests són jets que molt probablement són falsos. La seua presència afectaria dràsticament a la reconstrucció de la $E_{\mathrm{T}}^{\mathrm{miss}}$.

- La distància angular mínima entre la $E_{\mathrm{T}}^{\mathrm{miss}}$ i qualsevol jet $\Delta\phi_{\mathrm{min}}$ ha de ser major que 0.2.

En aquest punt, es realitza l'entrenament d'un classificador XGBoost [413] per assolir una major discriminació entre el senyal i el fons. XGBoost és una llibreria basada en els principis dels arbres de decisió [414] que implementa algoritmes d'aprenentatge automàtic utilitzant la tècnica de potenciació del gradient. S'entrenen dos classificadors diferents: un per al model ressonant i altre per al no-ressonant.

Per a l'entrenament del classificador s'empren una sèrie de variables que posseeixen un elevat potencial de discriminació. Destaquen la $E_{\mathrm{T}}^{\mathrm{miss}}$, la multiplicitat de jets a l'estat final, la massa reconstruïda del quark top o la suma de p_{T} de tots els jets de l'esdeveniment, entre altres.

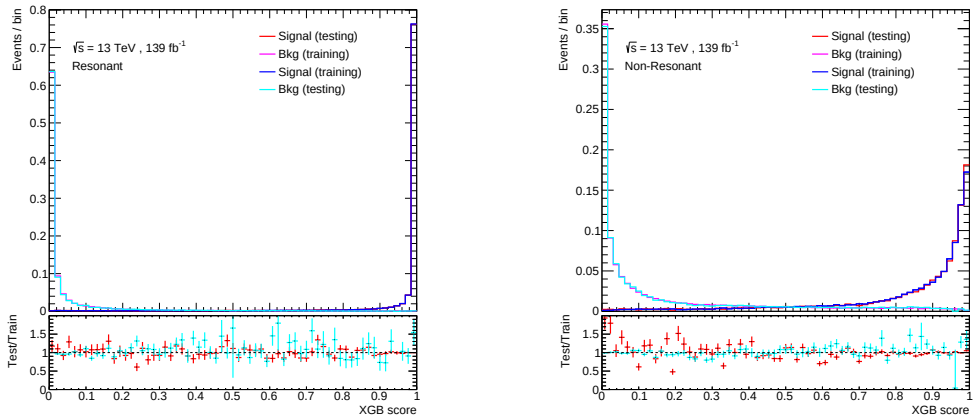
D'altra banda, s'utilitza una tècnica de k-folding [416] per assegurar-se de que el rendiment del classificador no depén de les dades considerades. Aquesta tècnica consisteix en dividir les dades en k parts (en aquest cas $k = 5$), utilitzant una d'elles per al test i les quatre restants per a l'entrenament. El procediment es repeteix cinc vegades, canviant la part utilitzada per al test en cada cas, i després es calcula la mitjana per a obtenir el rendiment del classificador.

La Figura R.20 mostra la resposta del classificador. Com es pot apreciar, aquest és capaç de distingir prou bé el senyal del fons i no s'observen diferències significatives de rendiment entre les parts utilitzades per a l'entrenament i per al test. Per a mesurar la qualitat del classificador, s'empra l'àrea sota la corba ROC[18]. Per al model ressonant s'obté un valor de 0.986, mentre que per al model no-ressonant s'obté un valor de 0.918, indicant que s'ha assolit un molt elevat potencial de discriminació entre senyal i fons.

Amb tot, la regió de senyal es defineix aplicant aquests requeriments addicionals:

- El nombre de b-jets $N_{b\text{-jets}}$ ha de ser menor o igual que 1.
- $\Delta\phi_{\mathrm{min}}$ ha de ser major que 1.0.

[18] La corba ROC representa la raó de vertaders positius front a la raó de vertaders negatius. Valors elevats de l'àrea sota la corba ROC indiquen un alt potencial de discriminació.



(a) (b)

Figure R.20: Resposta del classificador XGBoost per al fons i el senyal. En (a) s'utilitza el model ressonant, mentre que en (b) s'utilitza el model no-ressonant. En ambdós casos es mostra la resposta del classificador tant per a les dades de l'entrenament com per a les dades del test.

- El resultat de la puntuació del XGBoost ha de ser major que 0.5.

R.5.2 Estimació de fons

Els principals fons d'aquesta anàlisi són els fons de $t\bar{t}$ i $V(W/Z)$+jets. La normalització d'aquests fons està restringida a partir d'un ajust al nombre d'esdeveniments de dades en regions de controls específicament definides. El modelatge d'aquests fons i els resultats d'aquest ajust es ratifiquen en regions de validació. Aquestes regions estan definides a partir dels esdeveniments que passen la preselecció. Els requeriments emprats per a definir aquestes regions, junt amb la regió de senyal, es troben resumits a la Figura R.21.

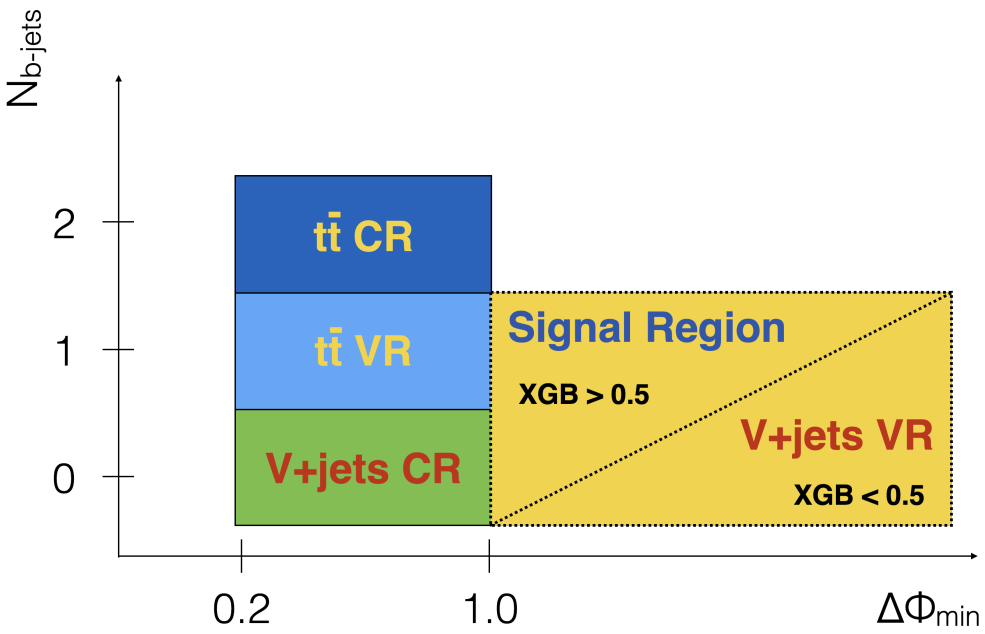
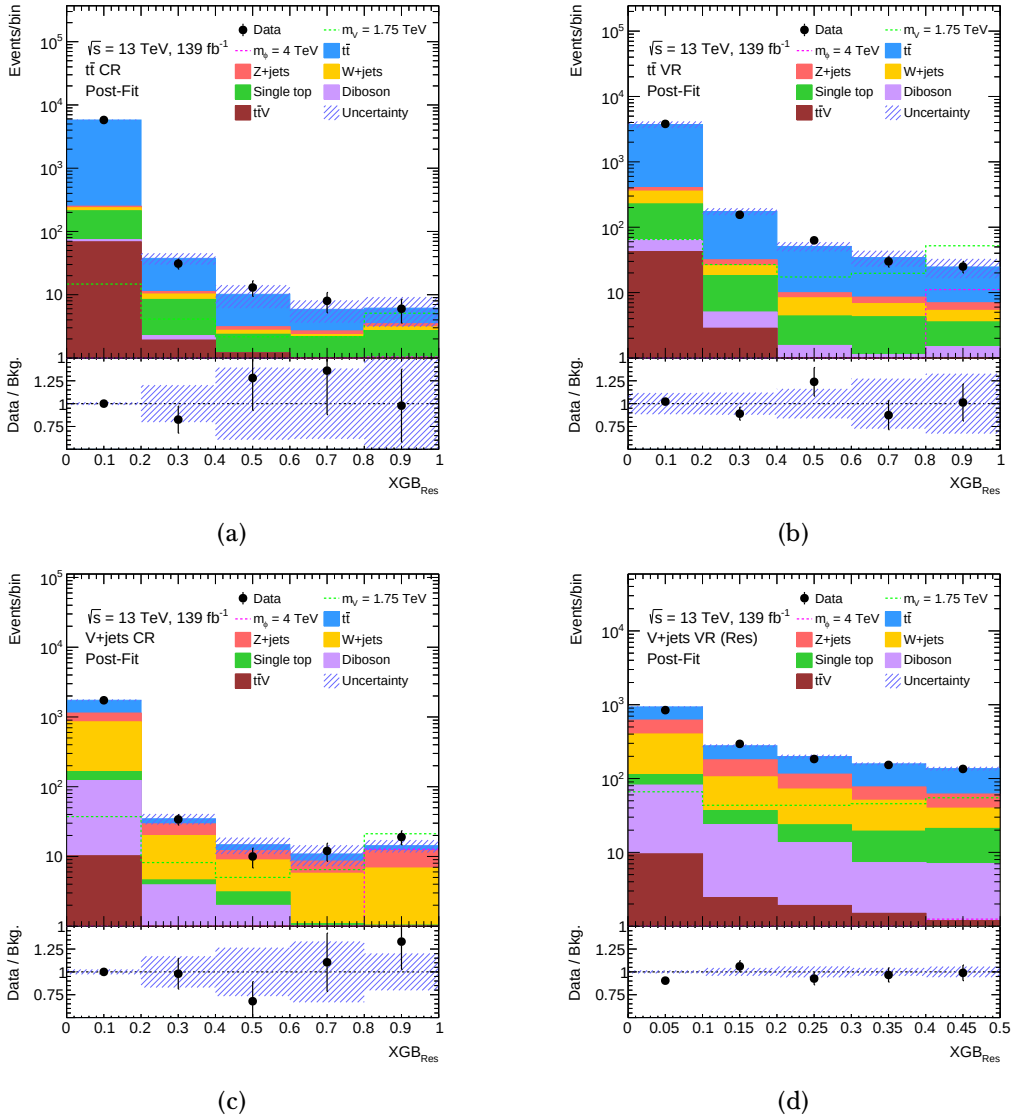


Figure R.21: Representació esquemàtica dels requeriments emprats per a definir les regions de control (CR) i de validació (VR), així com per a la regió de senyal. XGB representa la puntuació del classificador XGBoost.

Les Figures R.22-R.23 mostren la distribució del classificador XGBoost tant per al model ressonant com per al model no-ressonant en les regions de control i validació. Com es pot observar, hi ha un bon acord entre les dades observades i les prediccions del SM.



(a) (b)

(c) (d)

Figure R.22: Distribucions del XGBOOST entrenat amb el model ressonant per a les dades i per als fons del SM, així com per als esdeveniments de senyal del model ressonant amb m_ϕ = 4 TeV i del model no-ressonant amb m_V = 1.75 TeV a (a) la regió de control de $t\bar{t}$, (b) la regió de validació de $t\bar{t}$, (c) la regió de control de V+jets i (d) la regió de validació de V+jets del model ressonant.

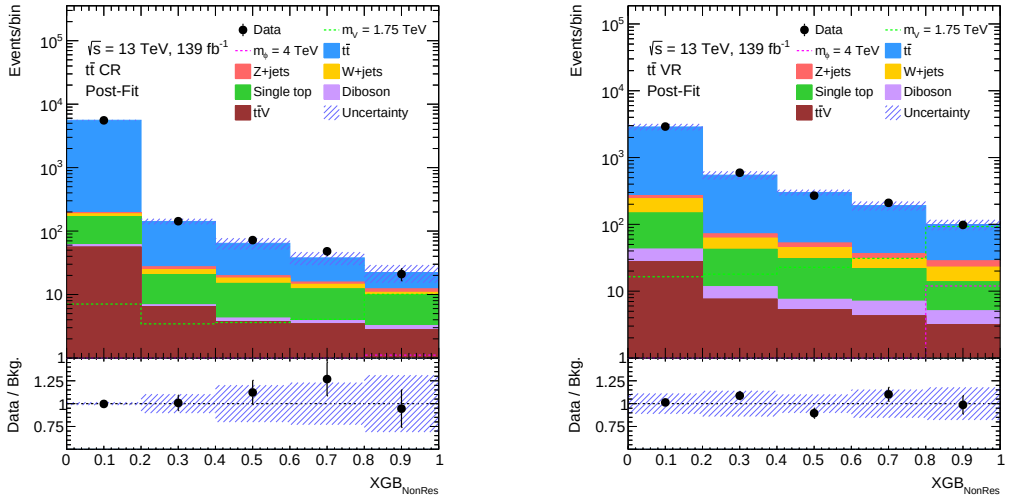


Figure R.23: Distribucions del XGBOOST entrenat amb el model no-ressonant per a les dades i per als fons del SM, així com per als esdeveniments de senyal del model ressonant amb m_ϕ = 4 TeV i del model no-ressonant amb m_V = 1.75 TeV a (a) la regió de control de $t\bar{t}$, (b) la regió de validació de $t\bar{t}$, (c) la regió de control de V+jets i (d) la regió de validació de V+jets del model no-ressonant.

R.5.3 Fonts d'incertesa

En aquesta anàlisi s'han tingut en compte les fonts d'incertesa sistemàtiques provinents de les incerteses experimentals i del modelatge teòric dels diferents processos. Dins de les incerteses experimentals, cal destacar que aquesta anàlisi també té en compte les incerteses associades a la reconstrucció dels jets grans així com a la identificació dels top-jets. Totes les incerteses considerades s'inclouen com paràmetres de soroll en l'ajust final, descrit en la següent secció.

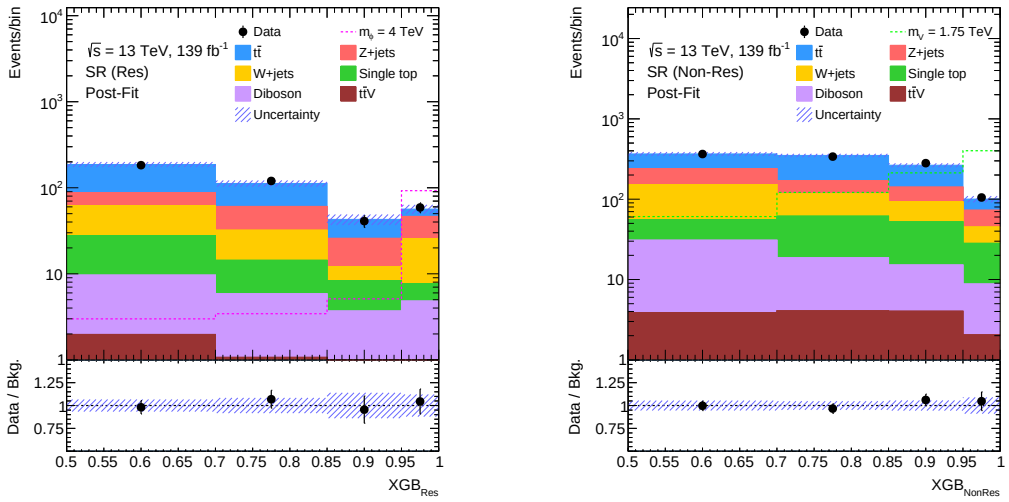
R.5.4 Resultats

Per tal d'estimar la possible presència de senyal de DM es fa un ajust de màxima versemblança a les dades de la distribució del XGBoost a la regió de senyal sota la hipòtesi de presència única de fons, tenint en compte el nombre d'esdeveniments en les regions de control per restringir la normalització dels fons principals. La Figura R.24 mostra la distribució del XGBoost a la regió de senyal per al model ressonant i no-ressonant després de l'ajust. Com es pot veure, les dades són compatibles amb les prediccions del SM.

En absència de senyal, aquests resultat s'empren per a establir un límit inferior en la massa de l'escalar carregat ϕ i el bosó vectorial V. La Figura R.25 mostra els límits en la secció eficaç obtinguts per a diferents valors de la massa d'aquestes partícules. A un 95% C.L. s'han exclòs masses $m_\phi < 4.6$ GeV ($m_V < 2.6$ GeV) per al model ressonant (no-ressonat), assumint uns valors dels paràmetres del model $\lambda = 0.6$, $y = 0.4$ i $m_\chi = 1$ GeV ($a = 0.5$, g_χ=1 i $m_\chi = 1$ GeV).

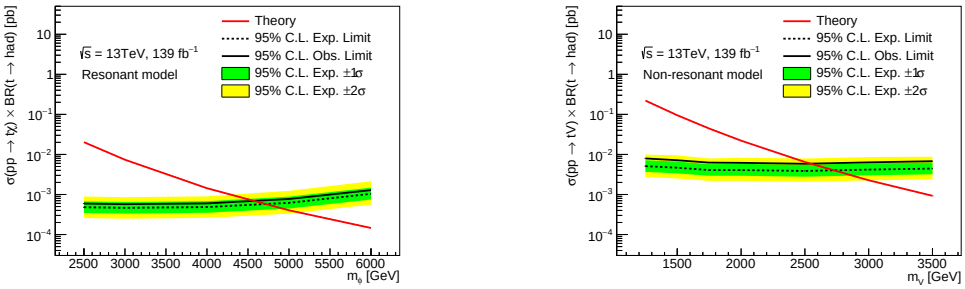
R.6 Conclusions

El treball d'aquesta tesi contribueix al programa de cerca de nova física a l'LHC a través de dues anàlisis amb estratègies ben diferenciades: la mesura de precisió d'observables sensibles a la polarització del quark top, buscant desviacions en les prediccions del SM; i la cerca directa de partícules de DM produïdes amb un quark top en solitari.



(a) (b)

Figure R.24: Distribucions del XGBoost en la regió de senyal dels models (a) ressonant i (b) no-ressonant després de l'ajust a les dades.



(a) (b)

Figure R.25: Límit al 95% C.L. en la secció eficaç de (a) el model ressonant i (b) el model no-ressonant per a diferents valors de la massa de l'escalar carregat i el bosó vectorial, respectivament. La línia roja indica la secció eficaç teòrica.

Les dades analitzades corresponen a les col·lisions protó-protó mesurades al detector ATLAS durant el Run-2 (2015-2018) a una energia en el centre de masses $\sqrt{s} = 13$ TeV, corresponents a una lluminositat integrada de 139 fb^{-1}.

La primera anàlisi presenta una mesura de la secció eficaç diferencial normalitzada del canal de producció t de quarks top produits en solitari, on el quark top es desintegra leptònicament, com a funció de tres variables angulars sensibles a la polarització del quark top. Aquestes mesures són particularment sensibles a efectes de nova física en el vèrtex tWb. Els resultats obtinguts són compatibles amb les prediccions del SM i s'han utilitzat per posar nous límits en la part real i imaginària de l'operador dipolar C_{tW} = 0.4 $\pm$ 1.1 i C_{itW} = -0.3 $\pm$ 0.4 a un 95% C.L. com a part del treball de tesi en la referència [4]. Els límits en la part imaginària són els límits directes més estrictes en el moment en què es presenta aquest treball.

La segona anàlisi consisteix en una cerca de DM a l'estat final mono-top. En aquest estudi s'utilitzen tècniques innovadores d'anàlisi multivariada per a discriminar el senyal del fons. Els resultats obtinguts no mostren cap indici de senyal. Així, aquests resultats són emprats per a posar límits a les seccions eficaces de producció del model ressonant i el model no-ressonant. Tot plegat, masses de l'escalar carregat ϕ inferiors a 4.6 TeV i masses del bosó vectorial V inferiors a 2.6 TeV han estat excloses. Aquests límits són els més restrictius en el moment en què es presenta aquest treball.

Finalment, en aquest treball també s'ha contribuït a l'alineament del detector intern d'ATLAS. Concretament, s'ha analitzat un moviment característic del detector de píxels que té lloc cada vegada que es prenen dades, i s'ha trobat una relació clara amb la lluminositat. D'altra banda, s'hi han obtingut unes constants de referència per als períodes de presa de dades protó-protó en 2018, així com per a col·lisions de ions pesats en 2015 i 2016.

Bibliography

[1] The ATLAS collaboration. "Alignment of the ATLAS Inner Detector in Run-2". In: *Eur. Phys. J. C* 80.arXiv:2007.07624. 12 (July 2020), 1194. 41 p. DOI: 10.1140/epjc/s10052-020-08700-6.

[2] J. A. Aguilar-Saavedra and S. Amor Dos Santos. "New directions for top quark polarization in the t-channel process". In: *Phys. Rev. D* 89.11 (2014), p. 114009. DOI: 10.1103/PhysRevD.89.114009. arXiv: 1404.1585 [hep-ph].

[3] The ATLAS collaboration. *Measurement of the polarisation of top quarks and antiquarks produced in the t-channel collected by the ATLAS detector at $\sqrt{s} = 13$ TeV and bounds on the tWb dipole operator*. Tech. rep. ATL-COM-PHYS-2021-062. Single Top Polarisation Unfolding Anomalous Couplings EFT. Geneva: CERN, Feb. 2021. URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2021-027/.

[4] M. de Beurs. "A top portal to new physics". June 2021. URL: https://www.nikhef.nl/~mdebeurs/thesis.pdf.

[5] The ATLAS Collaboration. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Phys. Lett. B* 716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[6] The CMS Collaboration. "Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC". In: *Phys. Lett. B* 716 (2012),

pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].

[7] L. Evans and P. Bryant. "LHC Machine". In: *JINST* 3 (2008), S08001. DOI: 10.1088/1748-0221/3/08/S08001.

[8] M. Benedikt et al. *LHC Design Report*. CERN Yellow Reports: Monographs. Geneva: CERN, 2004. DOI: 10.5170/CERN-2004-003-V-3. URL: https://cds.cern.ch/record/823808.

[9] J.R. Incandela et al. "Status and Prospects of Top-Quark Physics". In: *Prog. Part. Nucl. Phys.* 63 (2009), pp. 239–292. DOI: 10.1016/j.ppnp.2009.08.001. arXiv: 0904.2499 [hep-ex].

[10] D. Griffiths. *Introduction to elementary particles*. 2008. ISBN: 978-3-527-40601-2.

[11] J. Jimenez Pena. "ATLAS Inner Detector alignment and analysis of the Wtb vertex structure with single top quarks". Presented 05 Oct 2018. July 2018. URL: https://cds.cern.ch/record/2644636.

[12] P.A. Zyla et al. "Review of Particle Physics". In: *PTEP* 2020.8 (2020), p. 083C01. DOI: 10.1093/ptep/ptaa104.

[13] S. Weinberg. "A Model of Leptons". In: *Phys. Rev. Lett.* 19 (21 Nov. 1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264. URL: https://link.aps.org/doi/10.1103/PhysRevLett.19.1264.

[14] S.L. Glashow. "The renormalizability of vector meson interactions". In: *Nuclear Physics* 10 (1959), pp. 107–117. ISSN: 0029-5582. DOI: https://doi.org/10.1016/0029-5582(59)90196-8. URL: https://www.sciencedirect.com/science/article/pii/0029558259901968.

[15] A. Salam and J.C. Ward. "Electromagnetic and weak interactions". In: *Physics Letters* 13.2 (1964), pp. 168–171. ISSN: 0031-9163. DOI: https://doi.org/10.1016/0031-9163(64)90711-5. URL: https://www.sciencedirect.com/science/article/pii/0031916364907115.

[16] S. L. Glashow, J. Iliopoulos, and L. Maiani. “Weak Interactions with Lepton-Hadron Symmetry”. In: *Phys. Rev. D* 2 (7 Oct. 1970), pp. 1285–1292. DOI: 10.1103/PhysRevD.2.1285. URL: https://link.aps.org/doi/10.1103/PhysRevD.2.1285.

[17] P. W. Higgs. “Broken Symmetries and the Masses of Gauge Bosons”. In: *Phys. Rev. Lett.* 13 (16 Oct. 1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.508.

[18] F. Englert and R. Brout. “Broken Symmetry and the Mass of Gauge Vector Mesons”. In: *Phys. Rev. Lett.* 13 (1964). Ed. by J. C. Taylor, pp. 321–323. DOI: 10.1103/PhysRevLett.13.321.

[19] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. “Global Conservation Laws and Massless Particles”. In: *Phys. Rev. Lett.* 13 (20 Nov. 1964), pp. 585–587. DOI: 10.1103/PhysRevLett.13.585. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.585.

[20] P. W. Higgs. “Broken symmetries, massless particles and gauge fields”. In: *Phys. Lett.* 12 (1964), pp. 132–133. DOI: 10.1016/0031-9163(64)91136-9.

[21] J. Ellis. “Higgs Physics”. In: (Dec. 2013). 52 pages, 45 figures, Lectures presented at the ESHEP 2013 School of High-Energy Physics, to appear as part of the proceedings in a CERN Yellow Report, 117–168. 52 p. DOI: 10.5170/CERN-2015-004.117. arXiv: 1312.5672. URL: https://cds.cern.ch/record/1638469.

[22] Y. Nambu. “Quasiparticles and Gauge Invariance in the Theory of Superconductivity”. In: *Phys. Rev.* 117 (1960). Ed. by J. C. Taylor, pp. 648–663. DOI: 10.1103/PhysRev.117.648.

[23] J. Goldstone. “Field Theories with Superconductor Solutions”. In: *Nuovo Cim.* 19 (1961), pp. 154–164. DOI: 10.1007/BF02812722.

[24] J. Goldstone, A. Salam, and S. Weinberg. “Broken Symmetries”. In: *Phys. Rev.* 127 (1962), pp. 965–970. DOI: 10.1103/PhysRev.127.965.

[25] P. F. de Salas et al. “Status of neutrino oscillations 2018: 3σ hint for normal mass ordering and improved CP sensitivity”. In: *Phys. Lett. B* 782 (2018), pp. 633–640. DOI: 10.1016/j.physletb.2018.06.019. arXiv: 1708.01186 [hep-ph].

[26] R. Davis, D. S. Harmer, and K. C. Hoffman. “Search for Neutrinos from the Sun”. In: *Phys. Rev. Lett.* 20 (21 May 1968), pp. 1205–1209. DOI: 10.1103/PhysRevLett.20.1205. URL: https://link.aps.org/doi/10.1103/PhysRevLett.20.1205.

[27] P. Minkowski. “$\mu \to e\gamma$ at a Rate of One Out of 10^9 Muon Decays?” In: *Phys. Lett. B* 67 (1977), pp. 421–428. DOI: 10.1016/0370-2693(77)90435-X.

[28] M. Gell-Mann, P. Ramond, and R. Slansky. “Complex Spinors and Unified Theories”. In: *Conf. Proc. C* 790927 (1979), pp. 315–321. arXiv: 1306.4669 [hep-th].

[29] T. Yanagida. “Horizontal gauge symmetry and masses of neutrinos”. In: *Conf. Proc. C* 7902131 (1979). Ed. by Osamu Sawada and Akio Sugamoto, pp. 95–99.

[30] R. N. Mohapatra and G. Senjanovic. “Neutrino Mass and Spontaneous Parity Nonconservation”. In: *Phys. Rev. Lett.* 44 (1980), p. 912. DOI: 10.1103/PhysRevLett.44.912.

[31] J. Schechter and J. W. F. Valle. “Neutrino Masses in SU(2) x U(1) Theories”. In: *Phys. Rev. D* 22 (1980), p. 2227. DOI: 10.1103/PhysRevD.22.2227.

[32] J. Schechter and J. W. F. Valle. “Neutrino Decay and Spontaneous Violation of Lepton Number”. In: *Phys. Rev. D* 25 (1982), p. 774. DOI: 10.1103/PhysRevD.25.774.

[33] E. Ma. “Pathways to naturally small neutrino masses”. In: *Phys. Rev. Lett.* 81 (1998), pp. 1171–1174. DOI: 10.1103/PhysRevLett.81.1171. arXiv: hep-ph/9805219.

[34] E. Ma and O. Popov. “Pathways to Naturally Small Dirac Neutrino Masses”. In: *Phys. Lett. B* 764 (2017), pp. 142–144. DOI: 10.1016/j.physletb.2016.11.027. arXiv: 1609.02538 [hep-ph].

[35] S. Centelles Chuliá et al. “Dirac Neutrinos and Dark Matter Stability from Lepton Quarticity”. In: *Phys. Lett. B* 767 (2017), pp. 209–213. DOI: 10.1016/j.physletb.2017.01.070. arXiv: 1606.04543 [hep-ph].

[36] S. Centelles Chuliá, R. Srivastava, and J. W. F. Valle. “Seesaw roadmap to neutrino mass and dark matter”. In: *Phys. Lett. B* 781 (2018), pp. 122–128. DOI: 10.1016/j.physletb.2018.03.046. arXiv: 1802.05722 [hep-ph].

[37] C. Yao and G. Ding. “Systematic analysis of Dirac neutrino masses from a dimension five operator”. In: *Phys. Rev. D* 97.9 (2018), p. 095042. DOI: 10.1103/PhysRevD.97.095042. arXiv: 1802.05231 [hep-ph].

[38] E. Majorana. “Teoria simmetrica dell’elettrone e del positrone”. In: *Nuovo Cim.* 14 (1937), pp. 171–184. DOI: 10.1007/BF02961314.

[39] G. Racah. “On the symmetry of particle and antiparticle”. In: *Nuovo Cim.* 14 (1937), pp. 322–328. DOI: 10.1007/BF02961321.

[40] W. H. Furry. “Note on the Theory of the Neutral Particle”. In: *Phys. Rev.* 54 (1938), pp. 56–67. DOI: 10.1103/PhysRev.54.56.

[41] J. Schechter and J. W. F. Valle. “Neutrinoless Double beta Decay in SU(2) x U(1) Theories”. In: *Phys. Rev. D* 25 (1982), p. 2951. DOI: 10.1103/PhysRevD.25.2951.

[42] The KamLAND collaboration. “Search for Majorana Neutrinos near the Inverted Mass Hierarchy Region with KamLAND-Zen”. In: *Phys. Rev. Lett.* 117.8 (2016). [Addendum: Phys.Rev.Lett. 117, 109903 (2016)], p. 082503. DOI: 10.1103/PhysRevLett.117.082503. arXiv: 1605.02889 [hep-ex].

[43] The GERDA Collaboration. “Background-free search for neutrinoless double-β decay of ^{76}Ge with GERDA”. In: *Nature* 544 (2017), p. 47. DOI: 10.1038/nature21717. arXiv: 1703.00570 [nucl-ex].

[44] The CUORE collaboration. “First Results from CUORE: A Search for Lepton Number Violation via $0\nu\beta\beta$ Decay of ^{130}Te”. In: *Phys. Rev. Lett.* 120.13 (2018), p. 132501. DOI: 10.1103/PhysRevLett.120.132501. arXiv: 1710.07988 [nucl-ex].

[45] The NEMO collaboration. "Measurement of the $2\nu\beta\beta$ decay half-life of ^{150}Nd and a search for $0\nu\beta\beta$ decay processes with the full exposure from the NEMO-3 detector". In: *Phys. Rev. D* 94.7 (2016), p. 072003. DOI: 10.1103/PhysRevD.94.072003. arXiv: 1606.08494 [hep-ex].

[46] The EXO collaboration. "Search for Majorana neutrinos with the first two years of EXO-200 data". In: *Nature* 510 (2014), pp. 229–234. DOI: 10.1038/nature13432. arXiv: 1402.6956 [nucl-ex].

[47] A. D. Sakharov. "Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe". In: *Pisma Zh. Eksp. Teor. Fiz.* 5 (1967), pp. 32–35. DOI: 10.1070/PU1991v034n05ABEH002497.

[48] A.D. Dolgov. "Non-GUT baryogenesis". In: *Physics Reports* 222.6 (1992), pp. 309–386. ISSN: 0370-1573. DOI: https://doi.org/10.1016/0370-1573(92)90107-B. URL: https://www.sciencedirect.com/science/article/pii/037015739290107B.

[49] S. Dodelson. *Modern Cosmology*. Academic Press, Elsevier Science, 2003.

[50] W. H. Kinney. "Cosmology, inflation, and the physics of nothing". In: *NATO Sci. Ser. II* 123 (2003). Ed. by Harrison B. Prosper and Michael Danilov, pp. 189–243. DOI: 10.1007/978-94-010-0076-5_5. arXiv: astro-ph/0301448.

[51] E. W. Kolb and M. S. Turner. "Grand Unified Theories and the Origin of the Baryon Asymmetry". In: *Annual Review of Nuclear and Particle Science* 33.1 (1983), pp. 645–696. DOI: 10.1146/annurev.ns.33.120183.003241. URL: https://doi.org/10.1146/annurev.ns.33.120183.003241.

[52] A. R. Liddle and D. H. Lyth. *Cosmological Inflation and Large-Scale Structure*. 2000.

[53] A. D. Linde. *Particle physics and inflationary cosmology*. Vol. 5. 1990. arXiv: hep-th/0503203.

[54] M. Yoshimura. “Unified Gauge Theories and the Baryon Number of the Universe”. In: *Phys. Rev. Lett.* 41 (5 July 1978), pp. 281–284. DOI: 10.1103/PhysRevLett.41.281. URL: https://link.aps.org/doi/10.1103/PhysRevLett.41.281.

[55] D. Toussaint et al. “Matter-antimatter accounting, thermodynamics, and black-hole radiation”. In: *Phys. Rev. D* 19 (4 Feb. 1979), pp. 1036–1045. DOI: 10.1103/PhysRevD.19.1036. URL: https://link.aps.org/doi/10.1103/PhysRevD.19.1036.

[56] J. R. Ellis, M. K. Gaillard, and D. V. Nanopoulos. “Baryon Number Generation in Grand Unified Theories”. In: *Phys. Lett. B* 80 (1979). [Erratum: Phys.Lett.B 82, 464 (1979)], p. 360. DOI: 10.1016/0370-2693(79)91190-0.

[57] S. Weinberg. “Cosmological Production of Baryons”. In: *Phys. Rev. Lett.* 42 (13 Mar. 1979), pp. 850–853. DOI: 10.1103/PhysRevLett.42.850. URL: https://link.aps.org/doi/10.1103/PhysRevLett.42.850.

[58] J. A. Harvey et al. “Calculation of cosmological baryon asymmetry in grand unified gauge models”. In: *Nuclear Physics B* 201.1 (1982), pp. 16–100. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(82)90375-3. URL: https://www.sciencedirect.com/science/article/pii/0550321382903753.

[59] M. Fukugita and T. Yanagida. “Baryogenesis Without Grand Unification”. In: *Phys. Lett. B* 174 (1986), pp. 45–47. DOI: 10.1016/0370-2693(86)91126-3.

[60] V.A. Kuzmin, V.A. Rubakov, and M.E. Shaposhnikov. “On anomalous electroweak baryon-number non-conservation in the early universe”. In: *Physics Letters B* 155.1 (1985), pp. 36–42. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(85)91028-7. URL: https://www.sciencedirect.com/science/article/pii/0370269385910287.

[61] N. Turok and J. Zadrozny. “Phase transitions in the two-doublet model”. In: *Nuclear Physics B* 369.3 (1992), pp. 729–742. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(92)90284-I. URL: https://www.sciencedirect.com/science/article/pii/055032139290284I.

[62] A.E. Nelson, D.B. Kaplan, and A.G. Cohen. "Why there is something rather than nothing: Matter from weak interactions". In: *Nuclear Physics B* 373.2 (1992), pp. 453–478. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(92)90440-M. URL: https://www.sciencedirect.com/science/article/pii/055032139290440M.

[63] M. Pietroni. "The Electroweak phase transition in a nonminimal supersymmetric model". In: *Nucl. Phys. B* 402 (1993), pp. 27–45. DOI: 10.1016/0550-3213(93)90635-3. arXiv: hep-ph/9207227.

[64] S. J. Huber et al. "Electroweak Phase Transition and Baryogenesis in the nMSSM". In: *Nucl. Phys. B* 757 (2006), pp. 172–196. DOI: 10.1016/j.nuclphysb.2006.09.003. arXiv: hep-ph/0606298.

[65] A. Riotto and M. Trodden. "Recent progress in baryogenesis". In: *Ann. Rev. Nucl. Part. Sci.* 49 (1999), pp. 35–75. DOI: 10.1146/annurev.nucl.49.1.35. arXiv: hep-ph/9901362.

[66] P. Di Bari, A. Masiero, and R. Mohapatra. "Focus on the origin of matter". In: *New J. Phys.* 15 (2013), p. 035030. DOI: 10.1088/1367-2630/15/3/035030.

[67] B. Garbrecht. "Why is there more matter than antimatter? Calculational methods for leptogenesis and electroweak baryogenesis". In: *Prog. Part. Nucl. Phys.* 110 (2020), p. 103727. DOI: 10.1016/j.ppnp.2019.103727. arXiv: 1812.02651 [hep-ph].

[68] D. Bodeker and W. Buchmuller. "Baryogenesis from the weak scale to the grand unification scale". In: *Rev. Mod. Phys.* 93.3 (2021), p. 035004. DOI: 10.1103/RevModPhys.93.035004. arXiv: 2009.07294 [hep-ph].

[69] M. Drees. "An Introduction to supersymmetry". In: *Inauguration Conference of the Asia Pacific Center for Theoretical Physics (APCTP)*. Nov. 1996. arXiv: hep-ph/9611409.

[70] S. P. Martin. "A Supersymmetry primer". In: *Adv. Ser. Direct. High Energy Phys.* 18 (1998). Ed. by Gordon L. Kane, pp. 1–98. DOI: 10.1142/9789812839657_0001. arXiv: hep-ph/9709356.

[71] D. B. Kaplan and H. Georgi. "SU(2) x U(1) Breaking by Vacuum Misalignment". In: *Phys. Lett. B* 136 (1984), pp. 183–186. DOI: 10.1016/0370-2693(84)91177-8.

[72] D. B. Kaplan, H. Georgi, and S. Dimopoulos. "Composite Higgs Scalars". In: *Phys. Lett. B* 136 (1984), pp. 187–190. DOI: 10.1016/0370-2693(84)91178-X.

[73] H. Georgi and S. L. Glashow. "Unity of All Elementary Particle Forces". In: *Phys. Rev. Lett.* 32 (1974), pp. 438–441. DOI: 10.1103/PhysRevLett.32.438.

[74] A.A. Belavin et al. "Pseudoparticle solutions of the Yang-Mills equations". In: *Physics Letters B* 59.1 (1975), pp. 85–87. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(75)90163-X. URL: https://www.sciencedirect.com/science/article/pii/037026937590163X.

[75] G. 't Hooft. "Symmetry Breaking through Bell-Jackiw Anomalies". In: *Phys. Rev. Lett.* 37 (1 July 1976), pp. 8–11. DOI: 10.1103/PhysRevLett.37.8. URL: https://link.aps.org/doi/10.1103/PhysRevLett.37.8.

[76] J. Callan et al. "The Structure of the Gauge Theory Vacuum". In: *Phys. Lett. B* 63 (1976). Ed. by J. C. Taylor, pp. 334–340. DOI: 10.1016/0370-2693(76)90277-X.

[77] R. Jackiw and C. Rebbi. "Vacuum Periodicity in a Yang-Mills Quantum Theory". In: *Phys. Rev. Lett.* 37 (3 July 1976), pp. 172–175. DOI: 10.1103/PhysRevLett.37.172. URL: https://link.aps.org/doi/10.1103/PhysRevLett.37.172.

[78] R. J. Crewther. "Effects of Topological Charge in Gauge Theories". In: *Acta Phys. Austriaca Suppl.* 19 (1978), pp. 47–153. DOI: 10.1007/978-3-7091-8538-4_3.

[79] V. Baluni. "CP-nonconserving effects in quantum chromodynamics". In: *Phys. Rev. D* 19 (7 Apr. 1979), pp. 2227–2230. DOI: 10.1103/PhysRevD.19.2227. URL: https://link.aps.org/doi/10.1103/PhysRevD.19.2227.

[80] R.J. Crewther et al. “Chiral estimate of the electric dipole moment of the neutron in quantum chromodynamics”. In: *Physics Letters B* 88.1 (1979), pp. 123–127. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(79)90128-X. URL: https://www.sciencedirect.com/science/article/pii/037026937990128X.

[81] P. G. Harris et al. “New Experimental Limit on the Electric Dipole Moment of the Neutron”. In: *Phys. Rev. Lett.* 82 (5 Feb. 1999), pp. 904–907. DOI: 10.1103/PhysRevLett.82.904. URL: https://link.aps.org/doi/10.1103/PhysRevLett.82.904.

[82] R. D. Peccei and H. R. Quinn. “CP Conservation in the Presence of Pseudoparticles”. In: *Phys. Rev. Lett.* 38 (25 June 1977), pp. 1440–1443. DOI: 10.1103/PhysRevLett.38.1440. URL: https://link.aps.org/doi/10.1103/PhysRevLett.38.1440.

[83] The ALEPH collaboration. “Determination of the Number of Light Neutrino Species”. In: *Phys. Lett. B* 231 (1989), pp. 519–529. DOI: 10.1016/0370-2693(89)90704-1.

[84] G. D. Kribs et al. “Four generations and Higgs physics”. In: *Phys. Rev. D* 76 (7 Oct. 2007), p. 075016. DOI: 10.1103/PhysRevD.76.075016. URL: https://link.aps.org/doi/10.1103/PhysRevD.76.075016.

[85] The Planck collaboration. “Planck 2015 results. XIII. Cosmological parameters”. In: *Astron. Astrophys.* 594 (2016), A13. DOI: 10.1051/0004-6361/201525830. arXiv: 1502.01589 [astro-ph.CO].

[86] The ATLAS Collaboration. “Measurement of the top quark mass in the $t\bar{t} \to$ lepton+jets channel from $\sqrt{s} = 8$ TeV ATLAS data and combination with previous results”. In: *Eur. Phys. J. C* 79.4 (2019), p. 290. DOI: 10.1140/epjc/s10052-019-6757-9. arXiv: 1810.01772 [hep-ex].

[87] The CMS Collaboration. “Measurement of the top quark mass using proton-proton data at $\sqrt{(s)} = 7$ and 8 TeV”. In: *Phys. Rev. D* 93 (7 Apr. 2016), p. 072004. DOI: 10.1103/PhysRevD.93.072004. URL: https://link.aps.org/doi/10.1103/PhysRevD.93.072004.

[88] The ATLAS, CDF, CMS and D0 Collaborations. “First combination of Tevatron and LHC measurements of the top-quark mass”. In: (2014). arXiv: 1403.4427 [hep-ex].

[89] M. Kobayashi and T. Maskawa. “CP-Violation in the Renormalizable Theory of Weak Interaction”. In: *Progress of Theoretical Physics* 49.2 (Feb. 1973), pp. 652–657. ISSN: 0033-068X. DOI: 10.1143/PTP.49.652. eprint: https://academic.oup.com/ptp/article-pdf/49/2/652/5257692/49-2-652.pdf. URL: https://doi.org/10.1143/PTP.49.652.

[90] J. H. Christenson et al. “Evidence for the 2π Decay of the K_2^0 Meson”. In: *Phys. Rev. Lett.* 13 (4 July 1964), pp. 138–140. DOI: 10.1103/PhysRevLett.13.138. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.138.

[91] S. W. Herb et al. “Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions”. In: *Phys. Rev. Lett.* 39 (5 Aug. 1977), pp. 252–255. DOI: 10.1103/PhysRevLett.39.252. URL: https://link.aps.org/doi/10.1103/PhysRevLett.39.252.

[92] G. Arnison et al. “Experimental observation of isolated large transverse energy electrons with associated missing energy at s=540 GeV”. In: *Physics Letters B* 122.1 (1983), pp. 103–116. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(83)91177-2. URL: https://www.sciencedirect.com/science/article/pii/0370269383911772.

[93] M. Banner et al. “Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN pp collider”. In: *Physics Letters B* 122.5 (1983), pp. 476–485. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(83)91605-2. URL: https://www.sciencedirect.com/science/article/pii/0370269383916052.

[94] G. Arnison et al. “Experimental observation of lepton pairs of invariant mass around 95 GeV/c2 at the CERN SPS collider”. In: *Physics Letters B* 126.5 (1983), pp. 398–410. ISSN: 0370-2693. DOI: https://doi.org/10.

1016/0370-2693(83)90188-0. URL: https://www.sciencedirect.com/science/article/pii/0370269383901880.

[95] The UA2 collaboration. "Evidence for $Z^0 \to e^+e^-$ at the CERN $\bar{p}p$ collider". In: *Physics Letters B* 129.1 (1983), pp. 130–140. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(83)90744-X. URL: https://www.sciencedirect.com/science/article/pii/037026938390744X.

[96] The D0 collaboration. "Observation of the top quark". In: *Phys. Rev. Lett.* 74 (1995), pp. 2632–2637. DOI: 10.1103/PhysRevLett.74.2632. arXiv: hep-ex/9503003.

[97] The CDF collaboration. "Observation of top quark production in $\bar{p}p$ collisions". In: *Phys. Rev. Lett.* 74 (1995), pp. 2626–2631. DOI: 10.1103/PhysRevLett.74.2626. arXiv: hep-ex/9503002.

[98] M. F. Zoller. "Vacuum stability in the SM and the three-loop β-function for the Higgs self-interaction". In: *Subnucl. Ser.* 50 (2014). Ed. by A. Zichichi, pp. 557–566. DOI: 10.1142/9789814603904_0034. arXiv: 1209.5609 [hep-ph].

[99] G. Degrassi et al. "Higgs mass and vacuum stability in the Standard Model at NNLO". In: *JHEP* 08 (2012), p. 098. DOI: 10.1007/JHEP08(2012)098. arXiv: 1205.6497 [hep-ph].

[100] A. Andreassen, W. Frost, and M. D. Schwartz. "Scale-invariant instantons and the complete lifetime of the standard model". In: *Phys. Rev. D* 97 (5 Mar. 2018), p. 056006. DOI: 10.1103/PhysRevD.97.056006. URL: https://link.aps.org/doi/10.1103/PhysRevD.97.056006.

[101] P. Bärnreuther, M. Czakon, and A. Mitov. "Percent Level Precision Physics at the Tevatron: First Genuine NNLO QCD Corrections to $q\bar{q} \to t\bar{t} + X$". In: *Phys. Rev. Lett.* 109 (2012), p. 132001. DOI: 10.1103/PhysRevLett.109.132001. arXiv: 1204.5201 [hep-ph].

[102] M. Czakon and A. Mitov. "Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders". In: *Comput. Phys. Commun.* 185

(2014), p. 2930. DOI: 10.1016/j.cpc.2014.06.021. arXiv: 1112.5675 [hep-ph].

[103] M. Czakon and A. Mitov. "NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels". In: *JHEP* 12 (2012), p. 054. DOI: 10.1007/JHEP12(2012)054. arXiv: 1207.0236 [hep-ph].

[104] M. Czakon and A. Mitov. "NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction". In: *JHEP* 01 (2013), p. 080. DOI: 10.1007/JHEP01(2013)080. arXiv: 1210.6832 [hep-ph].

[105] M. Czakon, P. Fiedler, and A. Mitov. "Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$". In: *Phys. Rev. Lett.* 110 (2013), p. 252004. DOI: 10.1103/PhysRevLett.110.252004. arXiv: 1303.6254 [hep-ph].

[106] M. Cacciari et al. "Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation". In: *Phys. Lett. B* 710 (2012), pp. 612–622. DOI: 10.1016/j.physletb.2012.03.013. arXiv: 1111.5869 [hep-ph].

[107] M. Aliev et al. "HATHOR: HAdronic Top and Heavy quarks crOss section calculatoR". In: *Comput. Phys. Commun.* 182 (2011), pp. 1034–1046. DOI: 10.1016/j.cpc.2010.12.040. arXiv: 1007.1327 [hep-ph].

[108] P. Kant et al. "HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions". In: *Comput. Phys. Commun.* 191 (2015), pp. 74–89. DOI: 10.1016/j.cpc.2015.02.001. arXiv: 1406.4403 [hep-ph].

[109] N. Kidonakis. "Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-". In: *Phys. Rev. D* 82 (2010), p. 054018. DOI: 10.1103/PhysRevD.82.054018. arXiv: 1005.4451 [hep-ph].

[110] The ATLAS Collaboration. *Top working group summary plots*. URL: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCTopWGSummaryPlots.

[111] G. Mahlon and S. J. Parke. "Improved spin basis for angular correlation studies in single top quark production at the Tevatron". In: *Phys. Rev. D* 55 (1997), pp. 7249–7254. DOI: 10.1103/PhysRevD.55.7249. arXiv: hep-ph/9611367 [hep-ph].

[112] G. Mahlon and S. J. Parke. "Single top quark production at the LHC: Understanding spin". In: *Phys. Lett. B* 476 (2000), pp. 323–330. DOI: 10.1016/S0370-2693(00)00149-0. arXiv: hep-ph/9912458 [hep-ph].

[113] G. Mahlon and S. J. Parke. "Angular correlations in top quark pair production and decay at hadron colliders". In: *Phys. Rev. D* 53 (1996), pp. 4886–4896. DOI: 10.1103/PhysRevD.53.4886. arXiv: hep-ph/9512264 [hep-ph].

[114] E. L. Berger, J. Gao, and H. X. Zhu. "Differential Distributions for t-channel Single Top-Quark Production and Decay at Next-to-Next-to-Leading Order in QCD". In: *JHEP* 11 (2017), p. 158. DOI: 10.1007/JHEP11(2017)158. arXiv: 1708.09405 [hep-ph].

[115] I. I. Y. Bigi et al. "Production and Decay Properties of Ultraheavy Quarks". In: *Phys. Lett.* B181 (1986), pp. 157–163. DOI: 10.1016/0370-2693(86)91275-X.

[116] J. A. Aguilar-Saavedra and J. Bernabeu. "W polarisation beyond helicity fractions in top quark decays". In: *Nucl. Phys. B* 840 (2010), pp. 349–378. DOI: 10.1016/j.nuclphysb.2010.07.012. arXiv: 1005.5382 [hep-ph].

[117] W. Buchmuller and D. Wyler. "Effective Lagrangian Analysis of New Interactions and Flavor Conservation". In: *Nucl. Phys. B* 268 (1986), pp. 621–653. DOI: 10.1016/0550-3213(86)90262-2.

[118] N. P. Hartland et al. "A Monte Carlo global analysis of the Standard Model Effective Field Theory: the top quark sector". In: *JHEP* 04 (2019), p. 100. DOI: 10.1007/JHEP04(2019)100. arXiv: 1901.05965 [hep-ph].

[119] G. D'Ambrosio et al. "Minimal flavor violation: An Effective field theory approach". In: *Nucl. Phys. B* 645 (2002), pp. 155–187. DOI: 10.1016/S0550-3213(02)00836-2. arXiv: hep-ph/0207036.

[120] D. Barducci et al. “Interpreting top-quark LHC measurements in the standard-model effective field theory”. In: (Feb. 2018). Ed. by J.A. Aguilar-Saavedra et al. arXiv: 1802.07237 [hep-ph].

[121] G. Durieux, F. Maltoni, and C. Zhang. “Global approach to top-quark flavor-changing interactions”. In: *Phys. Rev. D* 91.7 (2015), p. 074017. DOI: 10.1103/PhysRevD.91.074017. arXiv: 1412.7166 [hep-ph].

[122] C. Degrande et al. “Automatic computations at next-to-leading order in QCD for top-quark flavor-changing neutral processes”. In: *Phys. Rev. D* 91 (2015), p. 034024. DOI: 10.1103/PhysRevD.91.034024. arXiv: 1412.5594 [hep-ph].

[123] M. de Beurs et al. “Effective operators in t-channel single top production and decay”. In: *Eur. Phys. J. C* 78.11 (2018), p. 919. DOI: 10.1140/epjc/s10052-018-6399-3. arXiv: 1807.03576 [hep-ph].

[124] B. Grzadkowski et al. “Dimension-Six Terms in the Standard Model Lagrangian”. In: *JHEP* 10 (2010), p. 085. DOI: 10.1007/JHEP10(2010)085. arXiv: 1008.4884 [hep-ph].

[125] J. A. Aguilar-Saavedra. “Effective operators in top physics”. In: *PoS* ICHEP2010 (2010). Ed. by B. Pire et al., p. 378. DOI: 10.22323/1.120.0378. arXiv: 1008.3225 [hep-ph].

[126] The ATLAS Collaboration. “Measurements of top quark spin observables in $t\bar{t}$ events using dilepton final states in $\sqrt{s} = 8$ TeV pp collisions with the ATLAS detector”. In: *JHEP* 03 (2017), p. 113. DOI: 10.1007/JHEP03(2017)113. arXiv: 1612.07004 [hep-ex].

[127] The ATLAS Collaboration. “Search for anomalous couplings in the Wtb vertex from the measurement of double differential angular decay rates of single top quarks produced in the t-channel with the ATLAS detector”. In: *JHEP* 04 (2016), p. 023. DOI: 10.1007/JHEP04(2016)023. arXiv: 1510.03764 [hep-ex].

[128] The ATLAS Collaboration. "Probing the W tb vertex structure in t-channel single-top-quark production and decay in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *JHEP* 04 (2017), p. 124. DOI: 10.1007/JHEP04(2017)124. arXiv: 1702.08309 [hep-ex].

[129] The ATLAS Collaboration. "Analysis of the Wtb vertex from the measurement of triple-differential angular decay rates of single top quarks produced in the t-channel at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *JHEP* 12 (2017), p. 017. DOI: 10.1007/JHEP12(2017)017. arXiv: 1707.05393 [hep-ex].

[130] The CMS Collaboration. "Measurement of differential cross sections and charge ratios for t-channel single top quark production in proton–proton collisions at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J. C* 80.5 (2020), p. 370. DOI: 10.1140/epjc/s10052-020-7858-1. arXiv: 1907.08330 [hep-ex].

[131] The ATLAS Collaboration. "Probing the W tb vertex structure in t-channel single-top-quark production and decay in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *JHEP* 04 (2017), p. 124. DOI: 10.1007/JHEP04(2017)124. arXiv: 1702.08309 [hep-ex].

[132] The ATLAS Collaboration. "Fiducial, total and differential cross-section measurements of t-channel single top-quark production in pp collisions at 8 TeV using data collected by the ATLAS detector". In: *Eur. Phys. J. C* 77.8 (2017), p. 531. DOI: 10.1140/epjc/s10052-017-5061-9. arXiv: 1702.02859 [hep-ex].

[133] O. Estrada Pastor. "Alignment of the ATLAS Inner Detector and probing the top quark polarization in single top t-channel production." Presented 30 Jun 2020. 2020. URL: https://cds.cern.ch/record/2728844.

[134] E. Corbelli and P. Salucci. "The Extended Rotation Curve and the Dark Matter Halo of M33". In: *Mon. Not. Roy. Astron. Soc.* 311 (2000), pp. 441–447. DOI: 10.1046/j.1365-8711.2000.03075.x. arXiv: astro-ph/9909252.

[135] F. Zwicky. "Die Rotverschiebung von extragalaktischen Nebeln". In: *Helv. Phys. Acta* 6 (1933), pp. 110–127. DOI: 10.1007/s10714-008-0707-4.

[136] R. Clausius. "XVI. On a mechanical theorem applicable to heat". In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 40.265 (1870), pp. 122–127. DOI: 10.1080/14786447008640370. URL: https://doi.org/10.1080/14786447008640370.

[137] A. Refregier. "Weak gravitational lensing by large scale structure". In: *Ann. Rev. Astron. Astrophys.* 41 (2003), pp. 645–668. DOI: 10.1146/annurev.astro.41.111302.102207. arXiv: astro-ph/0307212.

[138] M. Bartelmann and M. Maturi. "Weak gravitational lensing". In: Dec. 2016. arXiv: 1612.06535 [astro-ph.CO].

[139] A. N. Taylor et al. "Gravitational Lens Magnification and the Mass of Abell 1689". In: *The Astrophysical Journal* 501.2 (July 1998), pp. 539–553. DOI: 10.1086/305827. URL: https://doi.org/10.1086/305827.

[140] D. Clowe et al. "A direct empirical proof of the existence of dark matter". In: *Astrophys. J. Lett.* 648 (2006), pp. L109–L113. DOI: 10.1086/508162. arXiv: astro-ph/0608407.

[141] P. McDonald and A. Roy. "Clustering of dark matter tracers: generalizing bias for the coming era of precision LSS". In: *Journal of Cosmology and Astroparticle Physics* 2009.08 (Aug. 2009), pp. 020–020. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2009/08/020. URL: http://dx.doi.org/10.1088/1475-7516/2009/08/020.

[142] D. Baumann et al. "Cosmological Non-Linearities as an Effective Fluid". In: *JCAP* 07 (2012), p. 051. DOI: 10.1088/1475-7516/2012/07/051. arXiv: 1004.2488 [astro-ph.CO].

[143] Z. Vlah, M. White, and A. Aviles. "A Lagrangian effective field theory". In: *JCAP* 09 (2015), p. 014. DOI: 10.1088/1475-7516/2015/09/014. arXiv: 1506.05264 [astro-ph.CO].

[144] Z. Vlah, E. Castorina, and M. White. "The Gaussian streaming model and convolution Lagrangian effective field theory". In: *JCAP* 12 (2016), p. 007. DOI: 10.1088/1475-7516/2016/12/007. arXiv: 1609.02908 [astro-ph.CO].

[145] M. Lewandowski et al. “EFT of large scale structures in redshift space”. In: *Phys. Rev. D* 97.6 (2018), p. 063526. DOI: 10.1103/PhysRevD.97.063526. arXiv: 1512.06831 [astro-ph.CO].

[146] S. Chen et al. “Redshift-Space Distortions in Lagrangian Perturbation Theory”. In: *JCAP* 03 (2021), p. 100. DOI: 10.1088/1475-7516/2021/03/100. arXiv: 2012.04636 [astro-ph.CO].

[147] A. A. Penzias and R. W. Wilson. “A Measurement of excess antenna temperature at 4080-Mc/s”. In: *Astrophys. J.* 142 (1965), pp. 419–421. DOI: 10.1086/148307.

[148] M. Milgrom. “MOND theory”. In: *Can. J. Phys.* 93.2 (2015), pp. 107–118. DOI: 10.1139/cjp-2014-0211. arXiv: 1404.7661 [astro-ph.CO].

[149] J. W. Moffat. “Scalar-tensor-vector gravity theory”. In: *JCAP* 03 (2006), p. 004. DOI: 10.1088/1475-7516/2006/03/004. arXiv: gr-qc/0506021.

[150] The LIGO Scientific Collaboration. “Observation of Gravitational Waves from a Binary Black Hole Merger”. In: *Phys. Rev. Lett.* 116.6 (2016), p. 061102. DOI: 10.1103/PhysRevLett.116.061102. arXiv: 1602.03837 [gr-qc].

[151] The LIGO Scientific Collaboration. “GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral”. In: *Phys. Rev. Lett.* 119 (16 Oct. 2017), p. 161101. DOI: 10.1103/PhysRevLett.119.161101. URL: https://link.aps.org/doi/10.1103/PhysRevLett.119.161101.

[152] S. Boran et al. “GW170817 Falsifies Dark Matter Emulators”. In: *Phys. Rev. D* 97.4 (2018), p. 041501. DOI: 10.1103/PhysRevD.97.041501. arXiv: 1710.06168 [astro-ph.HE].

[153] J. Sakstein and B. Jain. “Implications of the Neutron Star Merger GW170817 for Cosmological Scalar-Tensor Theories”. In: *Phys. Rev. Lett.* 119 (25 Dec. 2017), p. 251303. DOI: 10.1103/PhysRevLett.119.251303. URL: https://link.aps.org/doi/10.1103/PhysRevLett.119.251303.

[154] H. Wang and others". "The GW170817/GRB 170817A/AT 2017gfo Association: Some Implications for Physics and Astrophysics". In: *The Astrophysical Journal* 851.1 (Dec. 2017), p. L18. DOI: 10.3847/2041-8213/aa9e08. URL: https://doi.org/10.3847/2041-8213/aa9e08.

[155] J. D. Bekenstein. "Relativistic gravitation theory for the modified Newtonian dynamics paradigm". In: *Phys. Rev. D* 70 (8 Oct. 2004), p. 083509. DOI: 10.1103/PhysRevD.70.083509. URL: https://link.aps.org/doi/10.1103/PhysRevD.70.083509.

[156] M. Taoso, G. Bertone, and A. Masiero. "Dark Matter Candidates: A Ten-Point Test". In: *JCAP* 03 (2008), p. 022. DOI: 10.1088/1475-7516/2008/03/022. arXiv: 0711.4996 [astro-ph].

[157] B. Audren et al. "Strongest model-independent bound on the lifetime of Dark Matter". In: *JCAP* 12 (2014), p. 028. DOI: 10.1088/1475-7516/2014/12/028. arXiv: 1407.2418 [astro-ph.CO].

[158] J. R. Primack. "The Nature of dark matter". In: *Frascati Phys. Ser.* 24 (2002). Ed. by A. Morselli and P. Picozza, pp. 449–474. arXiv: astro-ph/0112255.

[159] B. Fields and S. Sarkar. "Big-Bang nucleosynthesis (2006 Particle Data Group mini-review)". In: (Jan. 2006). arXiv: astro-ph/0601514.

[160] K. A. Olive, G. Steigman, and T. P. Walker. "Primordial nucleosynthesis: Theory and observations". In: *Phys. Rept.* 333 (2000), pp. 389–407. DOI: 10.1016/S0370-1573(00)00031-4. arXiv: astro-ph/9905320.

[161] G. G. Raffelt. "Particle physics from stars". In: *Ann. Rev. Nucl. Part. Sci.* 49 (1999), pp. 163–216. DOI: 10.1146/annurev.nucl.49.1.163. arXiv: hep-ph/9903472.

[162] S. W. Randall et al. "Constraints on the Self-Interaction Cross-Section of Dark Matter from Numerical Simulations of the Merging Galaxy Cluster 1E 0657-56". In: *Astrophys. J.* 679 (2008), pp. 1173–1180. DOI: 10.1086/587859. arXiv: 0704.0261 [astro-ph].

[163] D. A. Buote et al. “Chandra evidence for a flattened, triaxial dark matter halo in the elliptical galaxy ngc 720”. In: *Astrophys. J.* 577 (2002), pp. 183–196. DOI: 10.1086/342158. arXiv: astro-ph/0205469.

[164] D. Harvey et al. “The non-gravitational interactions of dark matter in colliding galaxy clusters”. In: *Science* 347 (2015), pp. 1462–1465. DOI: 10.1126/science.1261381. arXiv: 1503.07675 [astro-ph.CO].

[165] A. Ibarra. “Neutrinos and dark matter”. In: *AIP Conf. Proc.* 1666.1 (2015). Ed. by E. Kearns, p. 140004. DOI: 10.1063/1.4915588.

[166] M. Drewes et al. “A White Paper on keV Sterile Neutrino Dark Matter”. In: *JCAP* 01 (2017), p. 025. DOI: 10.1088/1475-7516/2017/01/025. arXiv: 1602.04816 [hep-ph].

[167] P. B. Denton. “Sterile Neutrino Searches with MicroBooNE: Electron Neutrino Disappearance”. In: (Nov. 2021). arXiv: 2111.05793 [hep-ph].

[168] I. G. Irastorza. “An introduction to axions and their detection”. In: Sept. 2021. arXiv: 2109.07376 [hep-ph].

[169] M. Fabbrichesi, E. Gabrielli, and G. Lanfranchi. “The Dark Photon”. In: (May 2020). DOI: 10.1007/978-3-030-62519-1. arXiv: 2005.01515 [hep-ph].

[170] B. Carr et al. “Primordial black hole constraints for extended mass functions”. In: *Phys. Rev. D* 96 (2 July 2017), p. 023514. DOI: 10.1103/PhysRevD.96.023514. URL: https://link.aps.org/doi/10.1103/PhysRevD.96.023514.

[171] D. S. Graff and K. Freese. “Analysis of a hubble space telescope search for red dwarfs: limits on baryonic matter in the galactic halo”. In: *Astrophys. J. Lett.* 456 (1996), p. L49. DOI: 10.1086/309850. arXiv: astro-ph/9507097.

[172] J. Najita, G. Tiede, and J. Carr. “From stars to super-planets: the low-mass imf in the young cluster ic348”. In: *Astrophys. J.* 541 (2000), pp. 977–1003. DOI: 10.1086/309477. arXiv: astro-ph/0005290.

[173] H. Goldberg. "Constraint on the Photino Mass from Cosmology". In: *Phys. Rev. Lett.* 50 (19 May 1983), pp. 1419–1422. DOI: 10.1103/PhysRevLett.50.1419. URL: https://link.aps.org/doi/10.1103/PhysRevLett.50.1419.

[174] J. R. Ellis et al. "Supersymmetric Relics from the Big Bang". In: *Nucl. Phys. B* 238 (1984). Ed. by M. A. Srednicki, pp. 453–476. DOI: 10.1016/0550-3213(84)90461-9.

[175] J. L. Feng. "Dark Matter Candidates from Particle Physics and Methods of Detection". In: *Ann. Rev. Astron. Astrophys.* 48 (2010), pp. 495–545. DOI: 10.1146/annurev-astro-082708-101659. arXiv: 1003.0904 [astro-ph.CO].

[176] J. Billard et al. *Direct Detection of Dark Matter – APPEC Committee Report.* Tech. rep. 106 pages. Apr. 2021. arXiv: 2104.07634. URL: https://cds.cern.ch/record/2764484.

[177] M. Pato et al. "Systematic uncertainties in the determination of the local dark matter density". In: *Phys. Rev. D* 82 (2 July 2010), p. 023531. DOI: 10.1103/PhysRevD.82.023531. URL: https://link.aps.org/doi/10.1103/PhysRevD.82.023531.

[178] C. Moni Bidin et al. "On the local dark matter density". In: *Astronomy and Astrophysics* 573 (Dec. 2014), A91. ISSN: 1432-0746. DOI: 10.1051/0004-6361/201424675. URL: http://dx.doi.org/10.1051/0004-6361/201424675.

[179] H. Silverwood et al. "A non-parametric method for measuring the local dark matter density". In: *Mon. Not. Roy. Astron. Soc.* 459.4 (2016), pp. 4191–4208. DOI: 10.1093/mnras/stw917. arXiv: 1507.08581 [astro-ph.GA].

[180] The XENON Collaboration. "Dark Matter Search Results from a One Ton-Year Exposure of XENON1T". In: *Phys. Rev. Lett.* 121.11 (2018), p. 111302. DOI: 10.1103/PhysRevLett.121.111302. arXiv: 1805.12562 [astro-ph.CO].

[181] The PandaX-II Collaboration. "Dark Matter Results From 54-Ton-Day Exposure of PandaX-II Experiment". In: *Phys. Rev. Lett.* 119.18 (2017),

p. 181302. DOI: 10.1103/PhysRevLett.119.181302. arXiv: 1708.06917 [astro-ph.CO].

[182] The LUX Collaboration. "Results from a search for dark matter in the complete LUX exposure". In: *Phys. Rev. Lett.* 118.2 (2017), p. 021303. DOI: 10.1103/PhysRevLett.118.021303. arXiv: 1608.07648 [astro-ph.CO].

[183] The SuperCDMS Collaboration. "Results from the Super Cryogenic Dark Matter Search Experiment at Soudan". In: *Phys. Rev. Lett.* 120.6 (2018), p. 061802. DOI: 10.1103/PhysRevLett.120.061802. arXiv: 1708.08869 [hep-ex].

[184] The DarkSide Collaboration. "DarkSide-50 532-day Dark Matter Search with Low-Radioactivity Argon". In: *Phys. Rev. D* 98.10 (2018), p. 102006. DOI: 10.1103/PhysRevD.98.102006. arXiv: 1802.07198 [astro-ph.CO].

[185] The DEAP Collaboration. "Search for dark matter with a 231-day exposure of liquid argon using DEAP-3600 at SNOLAB". In: *Phys. Rev. D* 100.2 (2019), p. 022004. DOI: 10.1103/PhysRevD.100.022004. arXiv: 1902.04048 [astro-ph.CO].

[186] The LUX Collaboration. "Results of a Search for Sub-GeV Dark Matter Using 2013 LUX Data". In: *Phys. Rev. Lett.* 122.13 (2019), p. 131301. DOI: 10.1103/PhysRevLett.122.131301. arXiv: 1811.11241 [astro-ph.CO].

[187] The XENON Collaboration. "Search for Light Dark Matter Interactions Enhanced by the Migdal Effect or Bremsstrahlung in XENON1T". In: *Phys. Rev. Lett.* 123.24 (2019), p. 241803. DOI: 10.1103/PhysRevLett.123.241803. arXiv: 1907.12771 [hep-ex].

[188] The XENON collaboration. "Light Dark Matter Search with Ionization Signals in XENON1T". In: *Phys. Rev. Lett.* 123.25 (2019), p. 251801. DOI: 10.1103/PhysRevLett.123.251801. arXiv: 1907.11485 [hep-ex].

[189] The DarkSide Collaboration. "Low-Mass Dark Matter Search with the DarkSide-50 Experiment". In: *Phys. Rev. Lett.* 121.8 (2018), p. 081307. DOI: 10.1103/PhysRevLett.121.081307. arXiv: 1802.06994 [astro-ph.HE].

[190] The SuperCDMS Collaboration. "Search for Low-Mass Dark Matter with CDMSlite Using a Profile Likelihood Fit". In: *Phys. Rev. D* 99.6 (2019), p. 062001. DOI: 10.1103/PhysRevD.99.062001. arXiv: 1808.09098 [astro-ph.CO].

[191] The CRESST Collaboration. "First results from the CRESST-III low-mass dark matter program". In: *Phys. Rev. D* 100.10 (2019), p. 102002. DOI: 10.1103/PhysRevD.100.102002. arXiv: 1904.00498 [astro-ph.CO].

[192] The NEWS-G Collaboration. "First results from the NEWS-G direct dark matter search experiment at the LSM". In: *Astropart. Phys.* 97 (2018), pp. 54–62. DOI: 10.1016/j.astropartphys.2017.10.009. arXiv: 1706.04934 [astro-ph.IM].

[193] The PICO Collaboration. "Dark Matter Search Results from the Complete Exposure of the PICO-60 C_3F_8 Bubble Chamber". In: *Phys. Rev. D* 100.2 (2019), p. 022001. DOI: 10.1103/PhysRevD.100.022001. arXiv: 1902.04031 [astro-ph.CO].

[194] The XENON Collaboration. "Constraining the spin-dependent WIMP-nucleon cross sections with XENON1T". In: *Phys. Rev. Lett.* 122.14 (2019), p. 141301. DOI: 10.1103/PhysRevLett.122.141301. arXiv: 1902.03234 [astro-ph.CO].

[195] The PandaX-II Collaboration. "PandaX-II Constraints on Spin-Dependent WIMP-Nucleon Effective Interactions". In: *Phys. Lett. B* 792 (2019), pp. 193–198. DOI: 10.1016/j.physletb.2019.02.043. arXiv: 1807.01936 [hep-ex].

[196] The LUX Collaboration. "Limits on spin-dependent WIMP-nucleon cross section obtained from the complete LUX exposure". In: *Phys. Rev. Lett.* 118.25 (2017), p. 251302. DOI: 10.1103/PhysRevLett.118.251302. arXiv: 1705.03380 [astro-ph.CO].

[197] The Super-Kamiokande Collaboration. "The Super-Kamiokande detector". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 501.2 (2003), pp. 418–462.

ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(03)00425-X. URL: https://www.sciencedirect.com/science/article/pii/S016890020300425X.

[198] The IceCube Collaboration. "The IceCube Neutrino Observatory: Instrumentation and Online Systems". In: *JINST* 12.03 (2017), P03012. DOI: 10.1088/1748-0221/12/03/P03012. arXiv: 1612.05093 [astro-ph.IM].

[199] The ANTARES Collaboration. "ANTARES: The first undersea neutrino telescope". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 656.1 (2011), pp. 11–38. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2011.06.103. URL: https://www.sciencedirect.com/science/article/pii/S0168900211013994.

[200] The IceCube Collaboration. "Search for annihilating dark matter in the Sun with 3 years of IceCube data". In: *Eur. Phys. J. C* 77.3 (2017). [Erratum: Eur.Phys.J.C 79, 214 (2019)], p. 146. DOI: 10.1140/epjc/s10052-017-4689-9. arXiv: 1612.05949 [astro-ph.HE].

[201] The Super-Kamiokande Collaboration. "Search for Neutrinos from Annihilation of Captured Low-Mass Dark Matter Particles in the Sun by Super-Kamiokande". In: *Phys. Rev. Lett.* 114 (14 Apr. 2015), p. 141301. DOI: 10.1103/PhysRevLett.114.141301. URL: https://link.aps.org/doi/10.1103/PhysRevLett.114.141301.

[202] The ANTARES Collaboration. "Limits on Dark Matter Annihilation in the Sun using the ANTARES Neutrino Telescope". In: *Phys. Lett. B* 759 (2016), pp. 69–74. DOI: 10.1016/j.physletb.2016.05.019. arXiv: 1603.02228 [astro-ph.HE].

[203] The PAMELA Collaboration. "An anomalous positron abundance in cosmic rays with energies 1.5-100 GeV". In: *Nature* 458 (2009), pp. 607–609. DOI: 10.1038/nature07942. arXiv: 0810.4995 [astro-ph].

[204] The AMS Collaboration. "Towards Understanding the Origin of Cosmic-Ray Positrons". In: *Phys. Rev. Lett.* 122 (4 Jan. 2019), p. 041102. DOI: 10.1103/PhysRevLett.122.041102. URL: https://link.aps.org/doi/10.1103/PhysRevLett.122.041102.

[205] The PAMELA Collaboration. "Ten years of PAMELA in space". In: *Riv. Nuovo Cim.* 40.10 (2017), pp. 473–522. DOI: 10.1393/ncr/i2017-10140-x. arXiv: 1801.10310 [astro-ph.HE].

[206] L. Bergström, T. Bringmann, and J. Edsjö. "New positron spectral features from supersymmetric dark matter: A way to explain the PAMELA data?" In: *Phys. Rev. D* 78 (10 Nov. 2008), p. 103520. DOI: 10.1103/PhysRevD.78.103520. URL: https://link.aps.org/doi/10.1103/PhysRevD.78.103520.

[207] J. Kopp. "Constraints on dark matter annihilation from AMS-02 results". In: *Phys. Rev. D* 88 (7 Oct. 2013), p. 076013. DOI: 10.1103/PhysRevD.88.076013. URL: https://link.aps.org/doi/10.1103/PhysRevD.88.076013.

[208] Q. Yuan et al. "Implications of the AMS-02 positron fraction in cosmic rays". In: *Astropart. Phys.* 60 (2015), pp. 1–12. DOI: 10.1016/j.astropartphys.2014.05.005. arXiv: 1304.1482 [astro-ph.HE].

[209] K. Kohri et al. "Can we explain AMS-02 antiproton and positron excesses simultaneously by nearby supernovae without pulsars or dark matter?" In: *Progress of Theoretical and Experimental Physics* 2016.2 (Feb. 2016). 021E01. ISSN: 2050-3911. DOI: 10.1093/ptep/ptv193. eprint: https://academic.oup.com/ptep/article-pdf/2016/2/021E01/19301380/ptv193.pdf. URL: https://doi.org/10.1093/ptep/ptv193.

[210] The Fermi-LAT Collaboration. "Searching for Dark Matter Annihilation from Milky Way Dwarf Spheroidal Galaxies with Six Years of Fermi Large Area Telescope Data". In: *Phys. Rev. Lett.* 115.23 (2015), p. 231301. DOI: 10.1103/PhysRevLett.115.231301. arXiv: 1503.02641 [astro-ph.HE].

[211] D. Hooper and L. Goodenough. "Dark Matter Annihilation in The Galactic Center As Seen by the Fermi Gamma Ray Space Telescope". In: *Phys. Lett. B* 697 (2011), pp. 412–428. DOI: 10.1016/j.physletb.2011.02.029. arXiv: 1010.2752 [hep-ph].

[212] The Fermi-LAT Collaboration. "The Fermi Galactic Center GeV Excess and Implications for Dark Matter". In: *Astrophys. J.* 840.1 (2017), p. 43. DOI: 10.3847/1538-4357/aa6cab. arXiv: 1704.03910 [astro-ph.HE].

[213] E. Tempel, A. Hektor, and M. Raidal. "Fermi 130 GeV gamma-ray excess and dark matter annihilation in sub-haloes and in the Galactic centre". In: *JCAP* 09 (2012). [Addendum: JCAP 11, A01 (2012)], p. 032. DOI: 10.1088/1475-7516/2012/09/032. arXiv: 1205.1045 [hep-ph].

[214] F. Calore, I. Cholis, and C. Weniger. "Background Model Systematics for the Fermi GeV Excess". In: *JCAP* 03 (2015), p. 038. DOI: 10.1088/1475-7516/2015/03/038. arXiv: 1409.0042 [astro-ph.CO].

[215] K. N. Abazajian and M. Kaplinghat. "Detection of a Gamma-Ray Source in the Galactic Center Consistent with Extended Emission from Dark Matter Annihilation and Concentrated Astrophysical Emission". In: *Phys. Rev. D* 86 (2012). [Erratum: Phys.Rev.D 87, 129902 (2013)], p. 083511. DOI: 10.1103/PhysRevD.86.083511. arXiv: 1207.6047 [astro-ph.HE].

[216] The ATLAS and CMS Collaborations. "Searches for Dark Matter Particles at the LHC". In: *53rd Rencontres de Moriond on Cosmology*. 2018, pp. 327–336. arXiv: 1809.06341 [hep-ex].

[217] Stefano Giagu. "WIMP Dark Matter Searches With the ATLAS Detector at the LHC". In: *Frontiers in Physics* 7 (2019), p. 75. ISSN: 2296-424X. DOI: 10.3389/fphy.2019.00075. URL: https://www.frontiersin.org/article/10.3389/fphy.2019.00075.

[218] T. Appelquist, H. Cheng, and B. A. Dobrescu. "Bounds on universal extra dimensions". In: *Phys. Rev. D* 64 (2001), p. 035002. DOI: 10.1103/PhysRevD.64.035002. arXiv: hep-ph/0012100.

[219] A. Albert et al. “Recommendations of the LHC Dark Matter Working Group: Comparing LHC searches for dark matter mediators in visible and invisible decay channels and calculations of the thermal relic density”. In: *Phys. Dark Univ.* 26 (2019), p. 100377. DOI: 10.1016/j.dark.2019.100377. arXiv: 1703.05703 [hep-ex].

[220] F. Kahlhoefer et al. “Implications of unitarity and gauge invariance for simplified dark matter models”. In: *JHEP* 02 (2016), p. 016. DOI: 10.1007/JHEP02(2016)016. arXiv: 1510.02110 [hep-ph].

[221] The ATLAS Collaboration. *Dark matter summary plots for s-channel mediators*. Tech. rep. Geneva: CERN, Mar. 2021. URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2021-006.

[222] The CMS Collaboration. *DM summary plots*. URL: https://twiki.cern.ch/twiki/bin/view/CMSPublic/SummaryPlotsEXO13TeV#DM_summary_plots.

[223] J. Andrea, B. Fuks, and F. Maltoni. “Monotops at the LHC”. In: *Phys. Rev. D* 84 (2011), p. 074025. DOI: 10.1103/PhysRevD.84.074025. arXiv: 1106.6199 [hep-ph].

[224] J.L. Agram et al. “Monotop phenomenology at the Large Hadron Collider”. In: *Phys. Rev. D* 89.1 (2014), p. 014028. DOI: 10.1103/PhysRevD.89.014028. arXiv: 1311.6478 [hep-ph].

[225] I. Boucheneb et al. “Revisiting monotop production at the LHC”. In: *JHEP* 01 (2015), p. 017. DOI: 10.1007/JHEP01(2015)017. arXiv: 1407.7529 [hep-ph].

[226] D. Abercrombie et al. “Dark Matter Benchmark Models for Early LHC Run-2 Searches: Report of the ATLAS/CMS Dark Matter Forum”. In: *Phys. Dark Univ.* 27 (2020). Ed. by Antonio Boveia et al., p. 100371. DOI: 10.1016/j.dark.2019.100371. arXiv: 1507.00966 [hep-ex].

[227] E. L. Berger, B. W. Harris, and Z. Sullivan. “Single top squark production via R-parity violating supersymmetric couplings in hadron collisions”. In: *Phys.*

Rev. Lett. 83 (1999), pp. 4472–4475. DOI: 10.1103/PhysRevLett.83.4472. arXiv: hep-ph/9903549.

[228] E. L. Berger, B. W. Harris, and Z. Sullivan. "Direct probes of R-parity violating supersymmetric couplings via single top squark production". In: *Phys. Rev. D* 63 (2001), p. 115001. DOI: 10.1103/PhysRevD.63.115001. arXiv: hep-ph/0012184.

[229] N. Desai and B. Mukhopadhyaya. "R-parity violating resonant stop production at the Large Hadron Collider". In: *JHEP* 10 (2010), p. 060. DOI: 10.1007/JHEP10(2010)060. arXiv: 1002.2339 [hep-ph].

[230] S. M. Barr. "A New Symmetry Breaking Pattern for SO(10) and Proton Decay". In: *Phys. Lett. B* 112 (1982), pp. 219–222. DOI: 10.1016/0370-2693(82)90966-2.

[231] H. Davoudiasl et al. "Hylogenesis: A Unified Origin for Baryonic Visible Matter and Antibaryonic Dark Matter". In: *Phys. Rev. Lett.* 105 (2010), p. 211304. DOI: 10.1103/PhysRevLett.105.211304. arXiv: 1008.2399 [hep-ph].

[232] H. Davoudiasl et al. "Baryon Destruction by Asymmetric Dark Matter". In: *Phys. Rev. D* 84 (2011), p. 096008. DOI: 10.1103/PhysRevD.84.096008. arXiv: 1106.4320 [hep-ph].

[233] The ATLAS Collaboration. *Search for invisible particles produced in association with single-top-quarks in proton-proton collisions at* $\sqrt{s}$ *= 8 TeV with the ATLAS detector*. Tech. rep. 2. 2015, p. 79. DOI: 10.1140/epjc/s10052-014-3233-4. arXiv: 1410.5404 [hep-ex].

[234] F. del Aguila, J. A. Aguilar-Saavedra, and L. Ametller. "Z t and gamma t production via top flavor changing neutral couplings at the Fermilab Tevatron". In: *Phys. Lett. B* 462 (1999), pp. 310–318. DOI: 10.1016/S0370-2693(99)00929-6. arXiv: hep-ph/9906462.

[235] J. F. Kamenik and J. Zupan. "Discovering Dark Matter Through Flavor Violation at the LHC". In: *Phys. Rev. D* 84 (2011), p. 111502. DOI: 10.1103/PhysRevD.84.111502. arXiv: 1107.0623 [hep-ph].

[236] E. Alvarez et al. “Leptonic Monotops at LHC”. In: *Phys. Rev. D* 89.1 (2014), p. 014016. DOI: 10.1103/PhysRevD.89.014016. arXiv: 1310.7600 [hep-ph].

[237] J. F. Kamenik and C. Smith. “FCNC portals to the dark sector”. In: *JHEP* 03 (2012), p. 090. DOI: 10.1007/JHEP03(2012)090. arXiv: 1111.6402 [hep-ph].

[238] The CMS Collaboration. “Search for Monotop Signatures in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV”. In: *Phys. Rev. Lett.* 114.10 (2015), p. 101801. DOI: 10.1103/PhysRevLett.114.101801. arXiv: 1410.1149 [hep-ex].

[239] The ATLAS Collaboration. “Search for invisible particles produced in association with single-top-quarks in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector”. In: *Eur. Phys. J. C* 75.2 (2015), p. 79. DOI: 10.1140/epjc/s10052-014-3233-4. arXiv: 1410.5404 [hep-ex].

[240] The CDF Collaboration. “Search for a dark matter candidate produced in association with a single top quark in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV”. In: *Phys. Rev. Lett.* 108 (2012), p. 201802. DOI: 10.1103/PhysRevLett.108.201802. arXiv: 1202.5653 [hep-ex].

[241] The ATLAS Collaboration. “Search for large missing transverse momentum in association with one top-quark in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *JHEP* 05 (2019), p. 041. DOI: 10.1007/JHEP05(2019)041. arXiv: 1812.09743 [hep-ex].

[242] The CMS Collaboration. “Search for dark matter in events with energetic, hadronically decaying top quarks and missing transverse momentum at $\sqrt{s} = 13$ TeV”. In: *JHEP* 06 (2018), p. 027. DOI: 10.1007/JHEP06(2018)027. arXiv: 1801.08427 [hep-ex].

[243] The CMS Collaboration. “Search for dark matter produced in association with a single top quark or a top quark pair in proton-proton collisions at $\sqrt{s} = 13$ TeV”. In: *JHEP* 03 (2019), p. 141. DOI: 10.1007/JHEP03(2019)141. arXiv: 1901.01553 [hep-ex].

[244] CERN: Accelerating science. URL: https://home.cern.

[245] J. T. Boyd. "LHC Run-2 and Future Prospects". In: *2019 European School of High-Energy Physics*. Jan. 2020. arXiv: 2001.04370 [hep-ex].

[246] The ATLAS Collaboration. *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*. Tech. rep. ATLAS-CONF-2019-021. Geneva: CERN, June 2019. URL: https://cds.cern.ch/record/2677054.

[247] G. Avoni et al. "The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS". In: *JINST* 13.07 (2018), P07017. DOI: 10.1088/1748-0221/13/07/P07017.

[248] G. Apollinari et al. *High-Luminosity Large Hadron Collider (HL-LHC): Preliminary Design Report*. CERN Yellow Reports: Monographs. Geneva: CERN, 2015. DOI: 10.5170/CERN-2015-005. URL: https://cds.cern.ch/record/2116337.

[249] The CMS collaboration. "The CMS experiment at the CERN LHC. The Compact Muon Solenoid experiment". In: *JINST* 3 (2008). Also published by CERN Geneva in 2010, S08004. 361 p. DOI: 10.1088/1748-0221/3/08/S08004. URL: https://cds.cern.ch/record/1129810.

[250] The ALICE Collaboration. "The ALICE experiment at the CERN LHC". In: *JINST* 3 (2008), S08002. DOI: 10.1088/1748-0221/3/08/S08002.

[251] The LHCb Collaboration. "The LHCb Detector at the LHC". In: *JINST* 3 (2008), S08005. DOI: 10.1088/1748-0221/3/08/S08005.

[252] The CMS-TOTEM Collaboration et al. *CMS-TOTEM Precision Proton Spectrometer*. Tech. rep. Sept. 2014. URL: https://cds.cern.ch/record/1753795.

[253] The LHCf Collaboration. *LHCf experiment: Technical Design Report*. Technical design report. LHCf. Geneva: CERN, 2006. URL: https://cds.cern.ch/record/926196.

[254] J. Pinfold et al. *Technical Design Report of the MoEDAL Experiment*. Tech. rep. June 2009. URL: https://cds.cern.ch/record/1181486.

[255] The FASER Collaboration. "FASER: ForwArd Search ExpeRiment at the LHC". In: (Jan. 2019). arXiv: 1901.04468 [hep-ex].

[256] A. D. Martin et al. "Parton distributions for the LHC". In: *Eur. Phys. J. C* 63 (2009), pp. 189–285. DOI: 10.1140/epjc/s10052-009-1072-5. arXiv: 0901.0002 [hep-ph].

[257] R. D. Ball et al. "Parton distributions for the LHC Run II". In: *JHEP* 04 (2015), p. 040. DOI: 10.1007/JHEP04(2015)040. arXiv: 1410.8849 [hep-ph].

[258] H.L. Lai et al. "New parton distributions for collider physics". In: *Phys. Rev. D* 82 (2010), p. 074024. DOI: 10.1103/PhysRevD.82.074024. arXiv: 1007.2241 [hep-ph].

[259] J. Pumplin et al. "New generation of parton distributions with uncertainties from global QCD analysis". In: *JHEP* 07 (2002), p. 012. DOI: 10.1088/1126-6708/2002/07/012. arXiv: hep-ph/0201195.

[260] R. D. Field. "The Underlying Event in Hard Scattering Processes". In: *eConf* C010630 (2001). Ed. by N. Graf, P501. arXiv: hep-ph/0201192.

[261] T. Sjöstrand et al. "An introduction to PYTHIA 8.2". In: *Comput. Phys. Commun.* 191 (2015), pp. 159–177. DOI: 10.1016/j.cpc.2015.01.024. arXiv: 1410.3012 [hep-ph].

[262] M. Bahr et al. "Herwig++ Physics and Manual". In: *Eur. Phys. J. C* 58 (2008), pp. 639–707. DOI: 10.1140/epjc/s10052-008-0798-9. arXiv: 0803.0883 [hep-ph].

[263] E. Bothmann et al. "Event Generation with Sherpa 2.2". In: *SciPost Phys.* 7.3 (2019), p. 034. DOI: 10.21468/SciPostPhys.7.3.034. arXiv: 1905.09127 [hep-ph].

[264] The ATLAS Collaboration. *Luminosity Public Results*. URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults.

[265] The ATLAS Collaboration. *ATLAS detector and physics performance: Technical Design Report, 1*. Technical design report. ATLAS. Geneva: CERN, 1999. URL: https://cds.cern.ch/record/391176.

[266] The ATLAS Collaboration. “Computer generated image of the whole ATLAS detector”. Mar. 2008. URL: http://cds.cern.ch/record/1095924.

[267] The ATLAS Collaboration. “How ATLAS detects particles: diagram of particle paths in the detector”. Jan. 2013. URL: http://cds.cern.ch/record/1505342.

[268] The ATLAS Collaboration. *ATLAS Inner Detector Alignment Performance with February 2015 Cosmic Rays Data*. Tech. rep. Geneva: CERN, Apr. 2015. URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2015-009/.

[269] The ATLAS Collaboration. *ATLAS Insertable B-Layer Technical Design Report*. Tech. rep. Sept. 2010. URL: https://cds.cern.ch/record/1291633.

[270] The ATLAS Collaboration. “The ATLAS Experiment at the CERN Large Hadron Collider”. In: *JINST* 3 (2008). Also published by CERN Geneva in 2010, S08003. 437 p. DOI: 10.1088/1748-0221/3/08/S08003. URL: https://cds.cern.ch/record/1129811.

[271] The ATLAS Collaboration. *ATLAS Transition Radiation Tracker (TRT): Straw tubes for tracking and particle identification at the Large Hadron Collider*. Tech. rep. 2017, pp. 257–261. DOI: 10.1016/j.nima.2016.04.026.

[272] The ATLAS Collaboration. “Design and implementation of the Front End Board for the readout of the ATLAS liquid argon calorimeters”. In: *Journal of Instrumentation* 3 (Mar. 2008), P03004. DOI: 10.1088/1748-0221/3/03/P03004.

[273] F. Cavallari. “Performance of calorimeters at the LHC”. In: *Journal of Physics: Conference Series* 293 (Apr. 2011), p. 012001. DOI: 10.1088/1742-6596/293/1/012001. URL: https://doi.org/10.1088/1742-6596/293/1/012001.

[274] The ATLAS Collaboration. “Performance of the ATLAS Trigger System in 2010”. In: *Eur. Phys. J. C* 72 (2012), p. 1849. DOI: 10.1140/epjc/s10052-011-1849-1. arXiv: 1110.1530 [hep-ex].

[275] The ATLAS Collaboration. "Performance of the ATLAS Trigger System in 2015". In: *Eur. Phys. J. C* 77.5 (2017), p. 317. DOI: 10.1140/epjc/s10052-017-4852-3. arXiv: 1611.09661 [hep-ex].

[276] The ATLAS Collaboration. *2015 start-up trigger menu and initial performance assessment of the ATLAS trigger using Run-2 data*. Tech. rep. ATL-DAQ-PUB-2016-001. Geneva: CERN, Mar. 2016. URL: http://cds.cern.ch/record/2136007.

[277] The ATLAS Collaboration. "Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC Run 2". In: *Eur. Phys. J. C* 77.10 (2017), p. 673. DOI: 10.1140/epjc/s10052-017-5225-7. arXiv: 1704.07983 [hep-ex].

[278] T. Cornelissen et al. "The new ATLAS track reconstruction (NEWT)". In: *Journal of Physics: Conference Series* 119.3 (July 2008), p. 032014. DOI: 10.1088/1742-6596/119/3/032014. URL: https://doi.org/10.1088/1742-6596/119/3/032014.

[279] Z. Rosenfeld and J. L. Pfaltz. "Sequential Operations in Digital Picture Processing". In: *J. ACM* 13.4 (Oct. 1966), pp. 471–494. ISSN: 0004-5411. DOI: 10.1145/321356.321357. URL: https://doi.org/10.1145/321356.321357.

[280] R. Fruehwirth. *Application of Kalman filtering to track and vertex fitting*. Dec. 1987. DOI: 10.1016/0168-9002(87)90887-4.

[281] The ATLAS Collaboration. *Vertex Reconstruction Performance of the ATLAS Detector at "$\sqrt{s}$ = 13 TeV"*. Tech. rep. ATL-PHYS-PUB-2015-026. Geneva: CERN, July 2015. URL: https://cds.cern.ch/record/2037717.

[282] The ATLAS Collaboration. "Electron and photon performance measurements with the ATLAS detector using the 2015-2017 LHC proton-proton collision data". In: *JINST* 14.12 (2019), P12006. DOI: 10.1088/1748-0221/14/12/P12006. arXiv: 1908.00005 [hep-ex].

[283] The ATLAS Collaboration. "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1". In: *Eur. Phys. J. C* 77 (2017), p. 490. DOI: 10.1140/epjc/s10052-017-5004-5. arXiv: 1603.02934 [hep-ex].

[284] The 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". In: *Eur. Phys. J. C* 79.8 (2019), p. 639. DOI: 10.1140/epjc/s10052-019-7140-6. arXiv: 1902.04655 [physics.ins-det].

[285] The ATLAS Collaboration. "Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton-proton collision data". In: *JINST* 14.03 (2019), P03017. DOI: 10.1088/1748-0221/14/03/P03017. arXiv: 1812.03848 [hep-ex].

[286] The ATLAS Collaboration. "Electron efficiency measurements with the ATLAS detector using 2012 LHC proton–proton collision data". In: *Eur. Phys. J. C* 77.3 (2017), p. 195. DOI: 10.1140/epjc/s10052-017-4756-2. arXiv: 1612.01456 [hep-ex].

[287] The ATLAS Collaboration. "Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J. C* 81 (2021), p. 578. DOI: 10.1140/epjc/s10052-021-09233-2. arXiv: 2012.00578 [hep-ex].

[288] J. Illingworth and J. Kittler. "A survey of the hough transform". In: *Computer Vision, Graphics, and Image Processing* 44.1 (1988), pp. 87–116. ISSN: 0734-189X. DOI: https://doi.org/10.1016/S0734-189X(88)80033-1. URL: https://www.sciencedirect.com/science/article/pii/S0734189X88800331.

[289] The ATLAS Collaboration. "Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$ =13 TeV". In: *Eur. Phys. J. C* 76.5 (2016), p. 292. DOI: 10.1140/epjc/s10052-016-4120-y. arXiv: 1603.05598 [hep-ex].

[290] M. Cacciari, G. P. Salam, and G. Soyez. "The anti-k_t jet clustering algorithm". In: *JHEP* 04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph].

[291] The ATLAS Collaboration. "Jet energy scale and resolution measured in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: (July 2020). arXiv: 2007.02645 [hep-ex].

[292] The ATLAS Collaboration. "Jet reconstruction and performance using particle flow with the ATLAS Detector". In: *Eur. Phys. J. C* 77.7 (2017), p. 466. DOI: 10.1140/epjc/s10052-017-5031-2. arXiv: 1703.10485 [hep-ex].

[293] The ATLAS Collaboration. *Determination of the jet energy scale and resolution at ATLAS using Z/γ–jet events in data at $\sqrt{s} = 8$ TeV*. Tech. rep. Geneva: CERN, Oct. 2015. URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2015-057.

[294] The ATLAS Collaboration. *TeV-scale jet energy calibration using multijet events including close-by jet effects at the ATLAS experiment*. Tech. rep. Geneva: CERN, Jan. 2013. URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2013-003.

[295] The ATLAS Collaboration. "Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector". In: *Eur. Phys. J. C* 76.11 (2016), p. 581. DOI: 10.1140/epjc/s10052-016-4395-z. arXiv: 1510.03823 [hep-ex].

[296] The ATLAS Collaboration. *JVT Public Plots for ICHEP 2016*. URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/JETM-2016-011/.

[297] The ATLAS Collaboration. "ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J. C* 79.11 (2019), p. 970. DOI: 10.1140/epjc/s10052-019-7450-8. arXiv: 1907.05120 [hep-ex].

[298] The ATLAS Collaboration. *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*. Tech. rep. Geneva: CERN, July 2017. URL: https://cds.cern.ch/record/2273281.

[299] The ATLAS Collaboration. *Identification of Jets Containing b-Hadrons with Recurrent Neural Networks at the ATLAS Experiment*. Tech. rep. Geneva: CERN, Mar. 2017. URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2017-003.

[300] The ATLAS Collaboration. *Secondary vertex finding for jet flavour identification with the ATLAS detector*. Tech. rep. Geneva: CERN, June 2017. URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2017-011.

[301] The ATLAS Collaboration. *Topological b-hadron decay reconstruction and identification of b-jets with the JetFitter package in the ATLAS experiment at the LHC*. Tech. rep. Geneva: CERN, Oct. 2018. URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2018-025.

[302] P. Langacker. "The Physics of Heavy Z' Gauge Bosons". In: *Rev. Mod. Phys.* 81 (2009), pp. 1199–1228. DOI: 10.1103/RevModPhys.81.1199. arXiv: 0801.1345 [hep-ph].

[303] A. Leike. "The Phenomenology of extra neutral gauge bosons". In: *Phys. Rept.* 317 (1999), pp. 143–250. DOI: 10.1016/S0370-1573(98)00133-1. arXiv: hep-ph/9805494.

[304] J. L. Hewett and T. G. Rizzo. "Low-energy phenomenology of superstring-inspired E6 models". In: *Physics Reports* 183.5 (1989), pp. 193–381. ISSN: 0370-1573. DOI: https://doi.org/10.1016/0370-1573(89)90071-9. URL: https://www.sciencedirect.com/science/article/pii/0370157389900719.

[305] The ATLAS Collaboration. "Optimisation of large-radius jet reconstruction for the ATLAS detector in 13 TeV proton–proton collisions". In: *Eur. Phys. J. C* 81.4 (2021), p. 334. DOI: 10.1140/epjc/s10052-021-09054-3. arXiv: 2009.04986 [hep-ex].

[306] The ATLAS Collaboration. "In situ calibration of large-radius jet energy and mass in 13 TeV proton–proton collisions with the ATLAS detector". In: *Eur.*

Phys. J. C 79.2 (2019), p. 135. DOI: 10.1140/epjc/s10052-019-6632-8. arXiv: 1807.09477 [hep-ex].

[307] The ATLAS Collaboration. "Search for diboson resonances in hadronic final states in 139 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *JHEP* 09 (2019). [Erratum: JHEP 06, 042 (2020)], p. 091. DOI: 10.1007/JHEP09(2019)091. arXiv: 1906.08589 [hep-ex].

[308] The ATLAS Collaboration. *Improving jet substructure performance in ATLAS using Track-CaloClusters*. Tech. rep. Geneva: CERN, July 2017. URL: https://cds.cern.ch/record/2275636.

[309] S. D. Ellis, C. K. Vermilion, and J. R. Walsh. "Recombination Algorithms and Jet Substructure: Pruning as a Tool for Heavy Particle Searches". In: *Phys. Rev. D* 81 (2010), p. 094023. DOI: 10.1103/PhysRevD.81.094023. arXiv: 0912.0033 [hep-ph].

[310] J. M. Butterworth et al. "Jet substructure as a new Higgs search channel at the LHC". In: *Phys. Rev. Lett.* 100 (2008), p. 242001. DOI: 10.1103/PhysRevLett.100.242001. arXiv: 0802.2470 [hep-ph].

[311] D. Krohn, J. Thaler, and L. T. Wang. "Jet Trimming". In: *JHEP* 02 (2010), p. 084. DOI: 10.1007/JHEP02(2010)084. arXiv: 0912.1342 [hep-ph].

[312] The ATLAS Collaboration. *Jet mass reconstruction with the ATLAS Detector in early Run 2 data*. Tech. rep. Geneva: CERN, July 2016. URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2016-035.

[313] M. Son, C. Spethmann, and B. Tweedie. "Diboson-Jets and the Search for Resonant Zh Production". In: *JHEP* 08 (2012), p. 160. DOI: 10.1007/JHEP08(2012)160. arXiv: 1204.0525 [hep-ph].

[314] A. Katz, M. Son, and B. Tweedie. "Jet Substructure and the Search for Neutral Spin-One Resonances in Electroweak Boson Channels". In: *JHEP* 03 (2011), p. 011. DOI: 10.1007/JHEP03(2011)011. arXiv: 1010.5253 [hep-ph].

[315] S. Schaetzel and M. Spannowsky. "Tagging highly boosted top quarks". In: *Phys. Rev. D* 89.1 (2014), p. 014007. DOI: 10.1103/PhysRevD.89.014007. arXiv: 1308.0540 [hep-ph].

[316] T. Plehn et al. "Stop Reconstruction with Tagged Tops". In: *JHEP* 10 (2010), p. 078. DOI: 10.1007/JHEP10(2010)078. arXiv: 1006.2833 [hep-ph].

[317] T. Plehn, G. P. Salam, and M. Spannowsky. "Fat Jets for a Light Higgs". In: *Phys. Rev. Lett.* 104 (2010), p. 111801. DOI: 10.1103/PhysRevLett.104.111801. arXiv: 0910.5472 [hep-ph].

[318] A. J. Larkoski, F. Maltoni, and M. Selvaggi. "Tracking down hyper-boosted top quarks". In: *JHEP* 06 (2015), p. 032. DOI: 10.1007/JHEP06(2015)032. arXiv: 1503.03347 [hep-ph].

[319] S. Bressler et al. "Hadronic Calorimeter Shower Size: Challenges and Opportunities for Jet Substructure in the Superboosted Regime". In: *Phys. Lett. B* 756 (2016), pp. 137–141. DOI: 10.1016/j.physletb.2016.02.068. arXiv: 1506.02656 [hep-ph].

[320] B. T. Elder and J. Thaler. "Aspects of Track-Assisted Mass". In: *JHEP* 03 (2019), p. 104. DOI: 10.1007/JHEP03(2019)104. arXiv: 1805.11109 [hep-ph].

[321] The ATLAS Collaboration. *Measurement of large radius jet mass reconstruction performance at $\sqrt{s} = 8$ TeV using the ATLAS detector.* Tech. rep. Geneva: CERN, Mar. 2016. URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2016-008.

[322] The ATLAS Collaboration. "Performance of top-quark and W-boson tagging with ATLAS in Run 2 of the LHC". In: *Eur. Phys. J. C* 79.5 (2019), p. 375. DOI: 10.1140/epjc/s10052-019-6847-8. arXiv: 1808.07858 [hep-ex].

[323] The ATLAS Collaboration. *Boosted hadronic vector boson and top quark tagging with ATLAS using Run 2 data.* Tech. rep. Geneva: CERN, July 2020. URL: https://cds.cern.ch/record/2724149.

[324] D. E. Soper and M. Spannowsky. "Finding top quarks with shower deconstruction". In: *Phys. Rev. D* 87 (2013), p. 054012. DOI: 10.1103/PhysRevD.87.054012. arXiv: 1211.3140 [hep-ph].

[325] The ATLAS Collaboration. "Measurements of b-jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s} = 13$ TeV". In: *JHEP* 08 (2018), p. 089. DOI: 10.1007/JHEP08(2018)089. arXiv: 1805.01845 [hep-ex].

[326] The ATLAS Collaboration. "Performance of b-Jet Identification in the ATLAS Experiment". In: *JINST* 11.04 (2016), P04008. DOI: 10.1088/1748-0221/11/04/P04008. arXiv: 1512.01094 [hep-ex].

[327] The ATLAS Collaboration. "Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s}$ = 13 TeV". In: *Eur. Phys. J. C* 78.11 (2018), p. 903. DOI: 10.1140/epjc/s10052-018-6288-9. arXiv: 1802.08168 [hep-ex].

[328] The ATLAS Collaboration. *E_T^{miss} performance in the ATLAS detector using 2015-2016 LHC p-p collisions*. Tech. rep. ATLAS-CONF-2018-023. Geneva: CERN, June 2018. URL: https://cds.cern.ch/record/2625233.

[329] K. Bos et al. *LHC computing Grid: Technical Design Report. Version 1.06 (20 Jun 2005)*. Technical design report. LCG. Geneva: CERN, 2005. URL: https://cds.cern.ch/record/840543.

[330] Worldwide LHC Computing Grid. URL: https://wlcg.web.cern.ch.

[331] The ATLAS Collaboration. "Performance of electron and photon triggers in ATLAS during LHC Run 2". In: *Eur. Phys. J. C* 80.1 (2020), p. 47. DOI: 10.1140/epjc/s10052-019-7500-2. arXiv: 1909.00761 [hep-ex].

[332] The ATLAS Collaboration. "Performance of the ATLAS muon triggers in Run 2". In: *JINST* 15.09 (2020), P09015. DOI: 10.1088/1748-0221/15/09/p09015. arXiv: 2004.13447 [physics.ins-det].

[333] The ATLAS Collaboration. *Measurement of b-tagging Efficiency of c-jets in $t\bar{t}$ Events Using a Likelihood Approach with the ATLAS Detector*. Tech. rep. ATLAS-CONF-2018-001. Geneva: CERN, Mar. 2018. URL: https://cds.cern.ch/record/2306649.

[334] The ATLAS Collaboration. *Calibration of light-flavour b-jet mistagging rates using ATLAS proton-proton collision data at* $\sqrt{s} = 13$ *TeV*. Tech. rep. ATLAS-CONF-2018-006. Geneva: CERN, Apr. 2018. URL: https://cds.cern.ch/record/2314418.

[335] P. Brückman, A. Hicheur, and S. J. Haywood. *Global chi2 approach to the Alignment of the ATLAS Silicon Tracking Detectors*. Tech. rep. ATL-INDET-PUB-2005-002. ATL-COM-INDET-2005-004. CERN-ATL-INDET-PUB-2005-002. Geneva: CERN, 2005. URL: https://cds.cern.ch/record/835270.

[336] C. Escobar. "Alignment of the ATLAS Silicon Tracker and measurement of the top quark mass". Presented on 09 Jul 2010. 2010. URL: https://cds.cern.ch/record/1277903.

[337] R. Härtel. "Iterative local Chi2 alignment approach for the ATLAS SCT detector". Presented 28 Nov 2005. Nov. 2005. URL: http://cds.cern.ch/record/1511044.

[338] T. Göttfert. "Iterative local Chi2 alignment algorithm for the ATLAS Pixel detector". Presented 26 May 2006. May 2006. URL: http://cds.cern.ch/record/1511043.

[339] The ATLAS Collaboration. *Study of alignment-related systematic effects on the ATLAS Inner Detector tracking*. Tech. rep. Geneva: CERN, Oct. 2012. URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2012-141.

[340] The CMS Collaboration. "Alignment of the CMS tracker with LHC and cosmic ray data". In: *JINST* 9 (2014), P06009. DOI: 10.1088/1748-0221/9/06/P06009. arXiv: 1403.2286 [physics.ins-det].

[341] The ATLAS Collaboration. "ATLAS data quality operations and performance for 2015-2018 data-taking". In: *JINST* 15.04 (2020), P04003. DOI: 10.1088/1748-0221/15/04/P04003. arXiv: 1911.04632 [physics.ins-det].

[342] M. Elsing et al. “The ATLAS Tier-0: Overview and operational experience”. In: *Journal of Physics: Conference Series* 219.7 (Apr. 2010), p. 072011. DOI: 10.1088/1742-6596/219/7/072011. URL: https://doi.org/10.1088/1742-6596/219/7/072011.

[343] The ATLAS Collaboration. *Radiation induced effects in the ATLAS Insertable B-Layer readout chip*. Tech. rep. Geneva: CERN, Nov. 2017. URL: http://cds.cern.ch/record/2291800.

[344] The ATLAS Collaboration. *Study of the mechanical stability of the ATLAS Insertable B-Layer*. Tech. rep. ATL-INDET-PUB-2015-001. Geneva: CERN, June 2015. URL: https://cds.cern.ch/record/2022587.

[345] B. Muratori and T. Pieloni. “Luminosity levelling techniques for the LHC”. In: (2014), pp. 177–181. DOI: 10.5170/CERN-2014-004.177. arXiv: 1410.5646 [physics.acc-ph].

[346] The ATLAS Collaboration. “ATLAS data quality operations and performance for 2015-2018 data-taking”. In: *JINST* 15.04 (2020), P04003. DOI: 10.1088/1748-0221/15/04/P04003. arXiv: 1911.04632 [physics.ins-det].

[347] S. Frixione et al. “Angular correlations of lepton pairs from vector boson and top quark decays in Monte Carlo simulations”. In: *JHEP* 04 (2007), p. 081. DOI: 10.1088/1126-6708/2007/04/081. arXiv: hep-ph/0702198.

[348] P. Artoisenet et al. “Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations”. In: *JHEP* 03 (2013), p. 015. DOI: 10.1007/JHEP03(2013)015. arXiv: 1212.3460 [hep-ph].

[349] The ATLAS Collaboration. *ATLAS Pythia 8 tunes to 7 TeV data*. Tech. rep. ATL-PHYS-PUB-2014-021. Geneva: CERN, Nov. 2014. URL: https://cds.cern.ch/record/1966419.

[350] D. J. Lange. “The EvtGen particle decay simulation package”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 462.1 (2001). BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines, pp. 152–155. ISSN:

0168-9002. DOI: https://doi.org/10.1016/S0168-9002(01)00089-4. URL: https://www.sciencedirect.com/science/article/pii/S0168900201000894.

[351] T. Sjostrand, S. Mrenna, and P. Z. Skands. "A Brief Introduction to PYTHIA 8.1". In: *Comput. Phys. Commun.* 178 (2008), pp. 852–867. DOI: 10.1016/j.cpc.2008.01.036. arXiv: 0710.3820 [hep-ph].

[352] R. D. Ball et al. "Parton distributions with LHC data". In: *Nucl. Phys. B* 867 (2013), pp. 244–289. DOI: 10.1016/j.nuclphysb.2012.10.003. arXiv: 1207.1303 [hep-ph].

[353] The ATLAS Collaboration. *The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie-Landshoff diffractive model.* Tech. rep. ATL-PHYS-PUB-2016-017. Geneva: CERN, Aug. 2016. URL: https://cds.cern.ch/record/2206965.

[354] R. Frederix, E. Re, and P. Torrielli. "Single-top t-channel hadroproduction in the four-flavour scheme with POWHEG and aMC@NLO". In: *JHEP* 09 (2012), p. 130. DOI: 10.1007/JHEP09(2012)130. arXiv: 1207.5391 [hep-ph].

[355] P. Nason. "A New method for combining NLO QCD with shower Monte Carlo algorithms". In: *JHEP* 11 (2004), p. 040. DOI: 10.1088/1126-6708/2004/11/040. arXiv: hep-ph/0409146.

[356] S. Frixione, P. Nason, and C. Oleari. "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method". In: *JHEP* 11 (2007), p. 070. DOI: 10.1088/1126-6708/2007/11/070. arXiv: 0709.2092 [hep-ph].

[357] S. Alioli et al. "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX". In: *JHEP* 06 (2010), p. 043. DOI: 10.1007/JHEP06(2010)043. arXiv: 1002.2581 [hep-ph].

[358] R. D. Ball et al. "Parton distributions for the LHC Run II". In: *JHEP* 04 (2015), p. 040. DOI: 10.1007/JHEP04(2015)040. arXiv: 1410.8849 [hep-ph].

[359] J. A. Aguilar-Saavedra. "Single top quark production at LHC with anomalous Wtb couplings". In: *Nucl. Phys. B* 804 (2008), pp. 160–192. DOI: 10.1016/j.nuclphysb.2008.06.013. arXiv: 0803.3810 [hep-ph].

[360] T. Stelzer, Z. Sullivan, and S. Willenbrock. "Single top quark production via W - gluon fusion at next-to-leading order". In: *Phys. Rev. D* 56 (1997), pp. 5919–5927. DOI: 10.1103/PhysRevD.56.5919. arXiv: hep-ph/9705398.

[361] E. Boos, L. Dudko, and T. Ohl. "Complete calculations of $Wb\bar{b}$ and $Wb\bar{b}$ + jet production at Tevatron and LHC: Probing anomalous Wtb couplings in single top production". In: *Eur. Phys. J. C* 11 (1999), pp. 473–484. DOI: 10.1007/s100520050647. arXiv: hep-ph/9903215.

[362] C.R. Chen, F. Larios, and C. P. Yuan. "General analysis of single top production and W helicity in top decay". In: *Phys. Lett. B* 631 (2005), pp. 126–132. DOI: 10.1016/j.physletb.2005.10.002. arXiv: hep-ph/0503040.

[363] J. A. Aguilar-Saavedra. "A Minimal set of top anomalous couplings". In: *Nucl. Phys. B* 812 (2009), pp. 181–204. DOI: 10.1016/j.nuclphysb.2008.12.012. arXiv: 0811.3842 [hep-ph].

[364] J. Alwall et al. "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations". In: *JHEP* 07 (2014), p. 079. DOI: 10.1007/JHEP07(2014)079. arXiv: 1405.0301 [hep-ph].

[365] C. Zhang. "Single Top Production at Next-to-Leading Order in the Standard Model Effective Field Theory". In: *Phys. Rev. Lett.* 116.16 (2016), p. 162002. DOI: 10.1103/PhysRevLett.116.162002. arXiv: 1601.06163 [hep-ph].

[366] J. M. Campbell et al. "Next-to-Leading-Order Predictions for t-Channel Single-Top Production at Hadron Colliders". In: *Phys. Rev. Lett.* 102 (18 May 2009), p. 182003. DOI: 10.1103/PhysRevLett.102.182003. URL: https://link.aps.org/doi/10.1103/PhysRevLett.102.182003.

[367] J. Butterworth et al. “PDF4LHC recommendations for LHC Run II”. In: *J. Phys. G* 43 (2016), p. 023001. DOI: 10.1088/0954-3899/43/2/023001. arXiv: 1510.03865 [hep-ph].

[368] A. D. Martin et al. “Uncertainties on alpha(S) in global PDF analyses and implications for predicted hadronic cross sections”. In: *Eur. Phys. J. C* 64 (2009), pp. 653–680. DOI: 10.1140/epjc/s10052-009-1164-2. arXiv: 0905.3531 [hep-ph].

[369] J. Gao et al. “CT10 next-to-next-to-leading order global analysis of QCD”. In: *Phys. Rev. D* 89.3 (2014), p. 033009. DOI: 10.1103/PhysRevD.89.033009. arXiv: 1302.6246 [hep-ph].

[370] The ATLAS Collaboration. *Studies on top-quark Monte Carlo modelling for Top2016*. Tech. rep. ATL-PHYS-PUB-2016-020. Geneva: CERN, Sept. 2016. URL: https://cds.cern.ch/record/2216168.

[371] S. Frixione et al. “Single-top hadroproduction in association with a W boson”. In: *JHEP* 07 (2008), p. 029. DOI: 10.1088/1126-6708/2008/07/029. arXiv: 0805.3067 [hep-ph].

[372] S. Hoeche et al. “A critical appraisal of NLO+PS matching methods”. In: *JHEP* 09 (2012), p. 049. DOI: 10.1007/JHEP09(2012)049. arXiv: 1111.1220 [hep-ph].

[373] S. Catani et al. “QCD matrix elements + parton showers”. In: *JHEP* 11 (2001), p. 063. DOI: 10.1088/1126-6708/2001/11/063. arXiv: hep-ph/0109231.

[374] S. Hoeche et al. “QCD matrix elements and truncated showers”. In: *JHEP* 05 (2009), p. 053. DOI: 10.1088/1126-6708/2009/05/053. arXiv: 0903.1219 [hep-ph].

[375] S. Hoeche et al. “QCD matrix elements + parton showers: The NLO case”. In: *JHEP* 04 (2013), p. 027. DOI: 10.1007/JHEP04(2013)027. arXiv: 1207.5030 [hep-ph].

[376] F. Cascioli, P. Maierhofer, and S. Pozzorini. "Scattering Amplitudes with Open Loops". In: *Phys. Rev. Lett.* 108 (2012), p. 111601. DOI: 10.1103/PhysRevLett.108.111601. arXiv: 1111.5206 [hep-ph].

[377] A. Denner, S. Dittmaier, and L. Hofer. "Collier: a fortran-based Complex One-Loop LIbrary in Extended Regularizations". In: *Comput. Phys. Commun.* 212 (2017), pp. 220–238. DOI: 10.1016/j.cpc.2016.10.013. arXiv: 1604.06792 [hep-ph].

[378] The ATLAS Collaboration. *ATLAS simulation of boson plus jets processes in Run 2*. Tech. rep. Geneva: CERN, May 2017. URL: https://cds.cern.ch/record/2261937.

[379] C. Anastasiou et al. "High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at NNLO". In: *Phys. Rev. D* 69 (2004), p. 094008. DOI: 10.1103/PhysRevD.69.094008. arXiv: hep-ph/0312266.

[380] F. A. Berends et al. "On the production of a W and jets at hadron colliders". In: *Nucl. Phys. B* 357 (1991), pp. 32–64. DOI: 10.1016/0550-3213(91)90458-A.

[381] J. Bellm et al. "Herwig 7.0/Herwig++ 3.0 release note". In: *Eur. Phys. J. C* 76.4 (2016), p. 196. DOI: 10.1140/epjc/s10052-016-4018-8. arXiv: 1512.01178 [hep-ph].

[382] L. A. Harland-Lang et al. "Parton distributions in the LHC era: MMHT 2014 PDFs". In: *Eur. Phys. J. C* 75.5 (2015), p. 204. DOI: 10.1140/epjc/s10052-015-3397-6. arXiv: 1412.3989 [hep-ph].

[383] The ATLAS Collaboration. "The ATLAS Simulation Infrastructure". In: *Eur. Phys. J. C* 70 (2010), pp. 823–874. DOI: 10.1140/epjc/s10052-010-1429-9. arXiv: 1005.4568 [physics.ins-det].

[384] S. Agostinelli et al. "GEANT4: A simulation toolkit". In: *Nucl. Instrum. Meth.* A506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.

[385] The ATLAS Collaboration. *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim*. Tech. rep. Oct. 2010.

[386] The ATLAS Collaboration. *The ATLAS Collaboration Software and Firmware*. Tech. rep. Geneva: CERN, May 2021. URL: https://cds.cern.ch/record/2767187.

[387] The ATLAS Collaboration. *Electron identification measurements in ATLAS using $\sqrt{s}$ = 13 TeV data with 50 ns bunch spacing*. Tech. rep. ATL-PHYS-PUB-2015-041. Geneva: CERN, Sept. 2015. URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2015-041/.

[388] The ATLAS Collaboration. *Estimation of non-prompt and fake lepton backgrounds in final states with top quarks produced in proton-proton collisions at $\sqrt{s}$=8 TeV with the ATLAS detector*. Tech. rep. ATLAS-CONF-2014-058. Geneva: CERN, Oct. 2014. URL: https://cds.cern.ch/record/1951336.

[389] The ATLAS Collaboration. "Measurement of the inclusive cross-sections of single top-quark and top-antiquark t-channel production in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector". In: *JHEP* 04 (2017), p. 086. DOI: 10.1007/JHEP04(2017)086. arXiv: 1609.03920 [hep-ex].

[390] M. Cacciari, G. P. Salam, and G. Soyez. "The Catchment Area of Jets". In: *JHEP* 04 (2008), p. 005. DOI: 10.1088/1126-6708/2008/04/005. arXiv: 0802.1188 [hep-ph].

[391] G. D'Agostini. "A multidimensional unfolding method based on Bayes' theorem". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 362.2 (1995), pp. 487–498. ISSN: 0168-9002. DOI: https://doi.org/10.1016/0168-9002(95)00274-X. URL: https://www.sciencedirect.com/science/article/pii/016890029500274X.

[392] T. Adye. "Unfolding algorithms and tests using RooUnfold". In: *PHYSTAT 2011*. Geneva: CERN, May 2011. DOI: 10.5170/CERN-2011-006.313. arXiv: 1105.1160 [physics.data-an].

[393] The ATLAS Collaboration. "Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s}$ =13 TeV". In: *Eur. Phys. J. C* 76.5 (2016), p. 292. DOI: 10.1140/epjc/s10052-016-4120-y. arXiv: 1603.05598 [hep-ex].

[394] The ATLAS Collaboration. "Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector". In: *Eur. Phys. J. C* 76.11 (2016), p. 581. DOI: 10.1140/epjc/s10052-016-4395-z. arXiv: 1510.03823 [hep-ex].

[395] The ATLAS Collaboration. "ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J. C* 79.11 (2019), p. 970. DOI: 10.1140/epjc/s10052-019-7450-8. arXiv: 1907.05120 [hep-ex].

[396] The ATLAS Collaboration. "Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s}$ = 13 TeV". In: *Eur. Phys. J. C* 78.11 (2018), p. 903. DOI: 10.1140/epjc/s10052-018-6288-9. arXiv: 1802.08168 [hep-ex].

[397] The ATLAS Collaboration. *Study of top-quark pair modelling and uncertainties using ATLAS measurements at $\sqrt{s}$=13 TeV*. Tech. rep. Geneva: CERN, Sept. 2020. URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2020-023.

[398] The ATLAS Collaboration. "Evidence for the $H \to b\bar{b}$ decay with the ATLAS detector". In: *JHEP* 12 (2017), p. 024. DOI: 10.1007/JHEP12(2017)024. arXiv: 1708.03299 [hep-ex].

[399] G. Bohm and G. Zech. *Introduction to Statistics and Data Analysis for Physicists; 3rd revised*. Hamburg: Verlag Deutsches Elektronen-Synchrotron, 2017, p. 488. ISBN: 978-3-945931-13-4. DOI: 10.3204/PUBDB-2017-08987. URL: https://bib-pubdb1.desy.de/record/389738.

[400] The ATLAS Collaboration. "Measurement of the W boson polarisation in $t\bar{t}$ events from pp collisions at $\sqrt{s}$ = 8 TeV in the lepton + jets channel with ATLAS". In: *Eur. Phys. J. C* 77.4 (2017). [Erratum: Eur.Phys.J.C 79, 19

(2019)], p. 264. DOI: 10.1140/epjc/s10052-017-4819-4. arXiv: 1612.02577 [hep-ex].

[401] The CMS Collaboration. "Measurement of the W boson helicity fractions in the decays of top quark pairs to lepton + jets final states produced in pp collisions at $\sqrt{s} = 8$TeV". In: *Phys. Lett. B* 762 (2016), pp. 512–534. DOI: 10.1016/j.physletb.2016.10.007. arXiv: 1605.09047 [hep-ex].

[402] The ATLAS and CMS Collaborations. "Combination of the W boson polarization measurements in top quark decays using ATLAS and CMS data at $\sqrt{s} = 8$ TeV". In: *JHEP* 08.08 (2020), p. 051. DOI: 10.1007/JHEP08(2020)051. arXiv: 2005.03799 [hep-ex].

[403] I. Brivio et al. "O new physics, where art thou? A global search in the top sector". In: *JHEP* 02 (2020), p. 131. DOI: 10.1007/JHEP02(2020)131. arXiv: 1910.03606 [hep-ph].

[404] G. Durieux et al. "The electro-weak couplings of the top and bottom quarks — Global fit and future prospects". In: *JHEP* 12 (2019). [Erratum: JHEP 01, 195 (2021)], p. 98. DOI: 10.1007/JHEP12(2019)098. arXiv: 1907.10619 [hep-ph].

[405] J. J. Ethier et al. "Combined SMEFT interpretation of Higgs, diboson, and top quark data from the LHC". In: *JHEP* 11 (2021), p. 089. DOI: 10.1007/JHEP11(2021)089. arXiv: 2105.00006 [hep-ph].

[406] V. Cirigliano et al. "Constraining the top-Higgs sector of the Standard Model Effective Field Theory". In: *Phys. Rev. D* 94.3 (2016), p. 034031. DOI: 10.1103/PhysRevD.94.034031. arXiv: 1605.04311 [hep-ph].

[407] F. Déliot et al. "Global Constraints on Top Quark Anomalous Couplings". In: *Phys. Rev. D* 97.1 (2018), p. 013007. DOI: 10.1103/PhysRevD.97.013007. arXiv: 1711.04847 [hep-ph].

[408] J. Alwall et al. "MadGraph 5 : Going Beyond". In: *JHEP* 06 (2011), p. 128. DOI: 10.1007/JHEP06(2011)128. arXiv: 1106.0522 [hep-ph].

[409] N. Kidonakis. "Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production". In: *Phys. Rev. D* 83 (2011), p. 091503. DOI: 10.1103/PhysRevD.83.091503. arXiv: 1103.2792 [hep-ph].

[410] N. Kidonakis. "NNLL resummation for s-channel single top quark production". In: *Phys. Rev. D* 81 (2010), p. 054028. DOI: 10.1103/PhysRevD.81.054028. arXiv: 1001.5034 [hep-ph].

[411] The ATLAS Collaboration. *Non-collision backgrounds as measured by the ATLAS detector during the 2010 proton-proton run.* Tech. rep. Geneva: CERN, Sept. 2011. URL: https://cds.cern.ch/record/1383840.

[412] The ATLAS Collaboration. "Characterisation and mitigation of beam-induced backgrounds observed in the ATLAS detector during the 2011 proton-proton run". In: *JINST* 8 (2013), P07004. DOI: 10.1088/1748-0221/8/07/P07004. arXiv: 1303.0223 [hep-ex].

[413] T. Chen and C. Guestrin. "XGBoost: A Scalable Tree Boosting System". In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '16. San Francisco, California, USA: Association for Computing Machinery, 2016, pp. 785–794. ISBN: 9781450342322. DOI: 10.1145/2939672.2939785. URL: https://doi.org/10.1145/2939672.2939785.

[414] B. P. Roe et al. "Boosted decision trees as an alternative to artificial neural networks for particle identification". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 543.2 (2005), pp. 577–584. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2004.12.018. URL: https://www.sciencedirect.com/science/article/pii/S0168900205000355.

[415] The ATLAS Collaboration. "Search for large missing transverse momentum in association with one top-quark in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *JHEP* 05 (2019), p. 041. DOI: 10.1007/JHEP05(2019)041. arXiv: 1812.09743 [hep-ex].

[416] M. Stone. "Cross-Validatory Choice and Assessment of Statistical Predictions". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 36.2 (1974), pp. 111–133. DOI: https://doi.org/10.1111/j.2517-6161.1974.tb00994.x. URL: https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/j.2517-6161.1974.tb00994.x.

[417] XGBoost documentation. URL: https://xgboost.readthedocs.io/en/latest/tutorials/model.html.

[418] The ATLAS Collaboration. *Measurement of the ATLAS Detector Jet Mass Response using Forward Folding with 80 fb^{-1} of $\sqrt{s} = 13$ TeV pp data*. Tech. rep. Geneva: CERN, July 2020. URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2020-022.

[419] J.K. Anders and M. D'Onofrio. *V+Jets theoretical uncertainties estimation via a parameterisation method*. Tech. rep. Geneva: CERN, Jan. 2016. URL: https://cds.cern.ch/record/2125718.

[420] M. Botje et al. "The PDF4LHC Working Group Interim Recommendations". In: (Jan. 2011). arXiv: 1101.0538 [hep-ph].

[421] U. Langenfeld, S. Moch, and P. Uwer. "New results for t anti-t production at hadron colliders". In: *17th International Workshop on Deep-Inelastic Scattering and Related Subjects*. July 2009, p. 131. arXiv: 0907.2527 [hep-ph].

[422] J. Butterworth et al. *Single Boson and Diboson Production Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV*. Tech. rep. Geneva: CERN, Aug. 2010. URL: https://cds.cern.ch/record/1287902.

[423] The ATLAS Collaboration. *Cross-section NLO ttV*. URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/CrossSectionNLOttV.

[424] W. Verkerke and D. P. Kirkby. "The RooFit toolkit for data modeling". In: *eConf* C0303241 (2003). Ed. by L. Lyons and Muge Karagoz, MOLT007. arXiv: physics/0306116.

[425] G. Cowan et al. “Asymptotic formulae for likelihood-based tests of new physics”. In: *Eur. Phys. J. C* 71 (2011). [Erratum: Eur.Phys.J.C 73, 2501 (2013)], p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0. arXiv: 1007.1727 [physics.data-an].

[426] F. H. Friedman. “Greedy function approximation: A gradient boosting machine.” In: *The Annals of Statistics* 29.5 (2001), pp. 1189–1232. DOI: 10.1214/aos/1013203451. URL: https://doi.org/10.1214/aos/1013203451.

[427] J. Friedman, T. Hastie, and R. Tibshirani. “Additive logistic regression: a statistical view of boosting (With discussion and a rejoinder by the authors)”. In: *The Annals of Statistics* 28.2 (2000), pp. 337–407. DOI: 10.1214/aos/1016218223. URL: https://doi.org/10.1214/aos/1016218223.

[428] C. Zhang and S. Willenbrock. “Effective-Field-Theory Approach to Top-Quark Production and Decay”. In: *Phys. Rev. D* 83 (2011), p. 034006. DOI: 10.1103/PhysRevD.83.034006. arXiv: 1008.3869 [hep-ph].