

---

# Search for supersymmetric particles in final states with multiple heavy-flavour jets with the ATLAS detector

---

*Ph.D. Dissertation*

**Carlos Moreno Martínez**

Institut de Física d'Altes Energies  
Universitat Autònoma de Barcelona  
Departament de Física  
Facultat de Ciències  
Edifici Cn E-08193 Bellaterra (Barcelona)

*Thesis directors*

Aurelio Juste Rozas  
ICREA/Institut de Física d'Altes Energies  
Universitat Autònoma de Barcelona  
Edifici Cn E-08193 Bellaterra (Barcelona)

Immaculada Riu Dachs  
Institut de Física d'Altes Energies  
Universitat Autònoma de Barcelona  
Edifici Cn E-08193 Bellaterra (Barcelona)

*Thesis tutor*

Rafel Escribano Carrascosa  
Institut de Física d'Altes Energies  
Universitat Autònoma de Barcelona  
Edifici Cn E-08193 Bellaterra (Barcelona)





# Abstract

This dissertation presents a search for supersymmetry in final states containing multiple  $b$ -quarks and large missing transverse energy using proton-proton collisions at a center-of-mass energy of  $\sqrt{s} = 13$  TeV as provided by the Large Hadron Collider (LHC) and recorded by the ATLAS experiment. The collisions used in the analysis correspond to the data taking period from 2015 to 2018, also known as LHC Run 2, amounting to a total integrated luminosity of  $139 \text{ fb}^{-1}$ .

The search targets a model for higgsino pair production in the context of R-parity conserving supersymmetry coupled with gauge-mediated symmetry-breaking mechanisms, where the higgsino is expected to be light due to naturalness arguments and decay to a Higgs boson and a gravitino, which is the lightest supersymmetric particle (LSP). The analysis is optimized focusing on the dominant decay channel of the Higgs boson into bottom-antibottom quark pairs. Thus, the characteristic final states contain high multiplicity of  $b$ -jets. The gravitinos do not interact with the detector, and therefore the final states are expected to exhibit large values of missing transverse energy. This is the only analysis in ATLAS optimized to probe the high-mass regime for the targeted model with large higgsino decay branching ratio into a Higgs boson.

No significant excess over the Standard Model prediction is observed in the analysis. The results are interpreted in the context of simplified models in order to translate them into upper limits on the higgsino pair production cross section. The search excludes at 95% confidence level (CL) all higgsino mass hypotheses between 200 and 950 GeV for scenarios with 100% decay branching ratio into a Higgs boson and a gravitino. Additionally, 95% CL exclusion limits for mixed branching ratio scenarios are derived, for which all mass hypotheses between 250 and 800 GeV are excluded for models with a branching ratio of 50% into a Higgs boson. The lowest value excluded for the decay branching ratio into a Higgs boson is 15% for a higgsino mass of 400 GeV.



# Resumen

En esta tesis se presenta una búsqueda de supersimetría en estados finales con múltiples quarks  $b$  y alto momento transverso faltante usando colisiones de protón-protón a una energía de centro de masa de  $\sqrt{s} = 13$  TeV proporcionadas por el Gran Colisionador de Hadrones (LHC, por sus siglas en inglés) y detectadas con el experimento ATLAS. Las colisiones usadas en el análisis corresponden al periodo entre 2015 y 2018 conocido como Run 2, con una luminosidad integrada total de  $139 \text{ fb}^{-1}$ .

Esta búsqueda tiene como objetivo la producción de pares de higgsinos en el contexto de supersimetría con paridad-R conservada y con mecanismos de ruptura de simetría mediada por interacciones gauge, en los cuales se espera que el higgsino sea ligero y que se desintegre en un bosón de Higgs y un gravitino, considerado como la partícula supersimétrica más ligera. El análisis está optimizado para el canal de desintegración del bosón de Higgs a parejas de quarks bottom-antibottom, el cual tiene la mayor probabilidad. Por tanto, los estados finales característicos contienen alta multiplicidad de  $b$ -jets. Los gravitinos no interactúan con el detector, por lo que los estados finales contienen también alto momento transverso faltante. Este es el único análisis en ATLAS optimizado para cubrir el espacio de parámetros con alta masa para el higgsino y alta probabilidad de desintegrarse en un bosón de Higgs.

En el análisis no se ha encontrado ningún exceso significativo con respecto a la predicción del Modelo Estándar. Los resultados de la búsqueda se han interpretado en el contexto de modelos simplificados para estimar el límite superior en la sección eficaz de la producción de pares de higgsinos. El análisis excluye a un nivel de confianza del 95% las hipótesis de masa para el higgsino entre 200 y 950 GeV en modelos con 100% de probabilidad de desintegrarse en un bosón de Higgs y un gravitino. Además, los límites para modelos en los que el higgsino se desintegra en un bosón de Higgs con probabilidad variable se calculan en base a las observaciones, para los cuales todas las hipótesis de masa entre 250 y 800 GeV están excluidas a un nivel de confianza del 95% en modelos con probabilidad de desintegrarse a un bosón de Higgs del 50%. El valor más bajo de la probabilidad de desintegrarse en un bosón de Higgs es de 15% para higgsinos con masa de 400 GeV.



# Acknowledgements

This thesis does not only reflect my own work during these four years of PhD, but also that of many other people. Whether it was help in the field of particle physics, or being there to support me in the process, I am very grateful to all those that walked this path with me.

My supervisors at IFAE, Aurelio Juste and Imma Riu, have made an extraordinary display of patience while guiding me throughout these years. Thanks to their expertise I have been able to grow both professionally and personally in a nurturing environment, making the PhD in the group a very enjoyable experience. I have really appreciated all the help received when working on physics analyses and the ATLAS trigger, together with the fact that my input was always taken seriously, even when they were just bad ideas. I am also most grateful for their very much needed guidance at the time of writing this dissertation, in particular to Aurelio, for doing the impossible to help me meet the deadlines I was already late for.

I have really enjoyed working alongside many great people in the ATLAS SUSY group. The good organization there allowed for an innovative analysis while profiting from the experience of other efforts. In particular, big thanks to the SUSY conveners and SUSY Electroweak subconveners for all the help in the review of the analysis. Many thanks to Rosa, who has guided me since my very first days in ATLAS, providing me with useful insight both in SUSY and L1Topo matters. Special thanks to Marco, Matt, Max, Michael and the rest of the multi-b group, with whom I have worked closely throughout the PhD. We may not have found SUSY this time, but we had fun in the process! Also big thanks to Emily and Lorenzo, for their enormous help and huge enthusiasm when starting to build new a new effort. Hopefully it will see the light soon enough!

Working in L1Topo I have had the chance to meet great people within the ATLAS Trigger group. Special mention to Johannes, Murrough and all the L1Topo hardware/firmware experts, who have provided invaluable insight during the commissioning and validation process of the system. Great thanks as well to the Trigger Core Software group for always providing the necessary help when Athena gets too complicated. In particular, I greatly appreciated working alongside with Joerg, whose infinite wisdom has helped avoid breaking the nightlies countless times. Big thanks as well to the Trigger Menu coordinators throughout the years, who have always provided a great amount of help when needed.

I would also like to thank the other fellow scientists in the making from IFAE, with whom I have shared great moments in these past years. Thanks to Adrián, Anil, Paula, Sergio, Shalini and Stergios for keeping me sane for this long. And to all of the already-made scientists, some of which I have had the pleasure to see evolve from students to the other side of the graduation. Thanks to Chiara, Julian, Nazlim, Nicola, Tal and Trisha, for being those “older siblings” that have always been happy to help. Big thanks go as well to Caspard, César, Christoph, Cristina, Danaisis, Darío, Emily, Kevin, Lourdes, Marta, Pablo, Saúl, Stephen and the rest of friends at CERN, for all the fun times spent together throughout these years.

The work carried out during the PhD has been possible thanks to the support of great people who, in spite of my mood swings and questionable ability to handle stress, have always stayed by my side (even in the distance). Huge thanks to Diego, Estela, Marta and Vanessa, for so many things that I would probably need another 200 pages just to mention them all. May we find Móspide soon! Big thanks as well to Antonio Jesús, Cristina, Javier and Javier, José, José Manuel, Raquel, Tapi and all the friends who have been there in the good and the bad times. Without you lot I would not have got where I am now.

Finally, I feel that a special mention needs to go to those who were there from the very beginning, years before I even expressed any curiosity about science. Muchas gracias a mi familia, por acompañarme a cada paso del camino, cerca o lejos, pero siempre a mi lado. A mis padres, por motivarme cada día a ser la mejor versión de mí mismo, y por ayudarme siempre en todo lo que he necesitado. Y a mi hermana, por su energía contagiosa y por ser un faro de positividad en los días oscuros. And last, but most certainly not least, thanks to Nhat, for bringing joy into my life every day. You have been there for me in my ups and downs, and it is thanks to your daily support that I have been able to bring this work to a conclusion. Em yêu anh.

# Contents

<b>Introduction</b>	<b>1</b>
<b>1 The Standard Model of Particle Physics</b>	<b>3</b>
1.1 Introduction . . . . .	3
1.2 Particle content of the Standard Model . . . . .	4
1.3 Interactions in the Standard Model . . . . .	6
1.3.1 Quantum Chromodynamics . . . . .	8
1.3.2 Electroweak interactions . . . . .	10
1.4 Spontaneous symmetry breaking and particle masses . . . . .	14
1.4.1 The Higgs mechanism . . . . .	16
1.4.2 Massive particles in the SM . . . . .	18
1.5 Limitations of the Standard Model . . . . .	19
1.6 Experimental tests of the SM . . . . .	21
<b>2 Supersymmetry</b>	<b>25</b>
2.1 Introduction . . . . .	25
2.2 General principles of Supersymmetry . . . . .	26
2.2.1 Dynamics and interactions of chiral supermultiplets . . . . .	28
2.2.2 Gauge supermultiplets and gauge interactions . . . . .	31
2.2.3 Supersymmetry breaking . . . . .	33
2.3 Minimal Supersymmetric Standard Model . . . . .	36
2.3.1 Soft SUSY breaking in the MSSM . . . . .	38
2.3.2 R-parity . . . . .	39
2.3.3 Mass spectrum of the MSSM . . . . .	40
2.4 Approaches in the search for SUSY . . . . .	42
2.4.1 Simplified models . . . . .	43
2.4.2 Phenomenological MSSM . . . . .	44

<b>3 The Large Hadron Collider and the ATLAS experiment</b>	<b>45</b>
3.1 Introduction . . . . .	45
3.2 The Large Hadron Collider . . . . .	46
3.2.1 Magnet system and resonant cavities . . . . .	47
3.2.2 Accelerator chain . . . . .	48
3.2.3 LHC operations . . . . .	50
3.3 The ATLAS experiment . . . . .	51
3.3.1 Coordinate system . . . . .	53
3.3.2 Inner detector . . . . .	54
3.3.3 Calorimeters . . . . .	56
3.3.4 Muon spectrometer . . . . .	59
<b>4 The ATLAS trigger system</b>	<b>63</b>
4.1 Introduction . . . . .	63
4.2 The Trigger and Data Acquisition system in ATLAS . . . . .	64
4.2.1 Level-1 Trigger . . . . .	64
4.2.2 High-level trigger . . . . .	70
4.2.3 Read-out data path . . . . .	71
4.2.4 Trigger menu . . . . .	72
4.3 Level-1 Topological trigger . . . . .	73
4.3.1 Description of the hardware . . . . .	75
4.3.2 Commissioning and validation . . . . .	76
4.3.3 Physics performance . . . . .	78
4.4 Phase-1 Upgrade . . . . .	82
<b>5 Simulation of proton collisions</b>	<b>87</b>
5.1 Introduction . . . . .	87
5.2 Physics processes inside proton-proton collisions . . . . .	88
5.2.1 Factorization theorem and parton distribution functions . . . . .	88
5.3 Monte Carlo simulation and generators . . . . .	91
5.3.1 Event simulation . . . . .	91
5.3.2 Monte Carlo generators . . . . .	95
5.4 Simulation of ATLAS detector . . . . .	96
<b>6 Object reconstruction in ATLAS</b>	<b>99</b>
6.1 Introduction . . . . .	99
6.2 Tracks and primary vertex . . . . .	99

6.3	Lepton reconstruction . . . . .	101
6.3.1	Electron reconstruction . . . . .	101
6.3.2	Electron identification . . . . .	101
6.3.3	Electron isolation . . . . .	102
6.3.4	Electron energy calibration . . . . .	102
6.3.5	Muon reconstruction . . . . .	102
6.3.6	Muon identification . . . . .	103
6.3.7	Muon isolation . . . . .	103
6.3.8	Muon energy calibration . . . . .	104
6.4	Jet reconstruction . . . . .	104
6.4.1	Jet reconstruction algorithms . . . . .	105
6.4.2	Jet calibration and associated uncertainties . . . . .	106
6.4.3	Jet Vertex Tagger . . . . .	108
6.4.4	Jet cleaning . . . . .	109
6.5	Jet-flavour tagging . . . . .	110
6.5.1	$b$ -tagging algorithms . . . . .	111
6.5.2	Calibration and uncertainties . . . . .	113
6.6	Missing transverse energy . . . . .	117
<b>7</b>	<b>Search for higgsino pair production</b>	<b>119</b>
7.1	Introduction . . . . .	119
7.2	Multi-variate analysis techniques . . . . .	120
7.2.1	Machine learning . . . . .	120
7.2.2	Boosted decision trees . . . . .	121
7.2.3	Neural networks . . . . .	123
7.2.4	ML-assisted reweighting . . . . .	124
7.3	Signal model . . . . .	125
7.4	Event selection . . . . .	127
7.5	Relevant kinematic variables . . . . .	129
7.5.1	Higgs Boson Reconstruction . . . . .	130
7.6	Main background processes . . . . .	134
7.6.1	Pair production of top quarks . . . . .	134
7.6.2	Single-top-quark production . . . . .	136
7.6.3	Vector boson production in association with jets . . . . .	137
7.6.4	Diboson production . . . . .	137
7.6.5	$t\bar{t}+X$ production . . . . .	137

7.6.6	Modelling uncertainties . . . . .	138
7.7	QCD estimation with NN-assisted reweighting . . . . .	138
7.7.1	QCD multijet estimate . . . . .	139
7.8	Data agreement at pre-selection level . . . . .	144
7.9	Signal-to-background discrimination . . . . .	149
7.9.1	BDT discriminant . . . . .	149
7.9.2	Optimization and performance . . . . .	150
7.9.3	Definition of analysis regions . . . . .	152
<b>8</b>	<b>Statistical methods and results</b>	<b>163</b>
8.1	Introduction . . . . .	163
8.2	Parameter estimation . . . . .	164
8.2.1	Parameters of interest and nuisance parameters . . . . .	165
8.2.2	Maximum likelihood method . . . . .	166
8.3	Hypothesis testing . . . . .	169
8.3.1	Test statistic . . . . .	169
8.3.2	Likelihood ratio discriminant . . . . .	169
8.3.3	Accepting the alternative hypothesis . . . . .	171
8.3.4	Rejecting the alternative hypothesis . . . . .	173
8.4	Systematic uncertainties . . . . .	174
8.4.1	Experimental systematic uncertainties . . . . .	175
8.4.2	Modelling systematic uncertainties . . . . .	176
8.5	Background-only fit . . . . .	179
8.6	Interpretations . . . . .	183
8.6.1	Model-independent upper limits . . . . .	183
8.6.2	Model-dependent upper limits . . . . .	184
<b>9</b>	<b>Perspective and prospects of the search</b>	<b>187</b>
9.1	Introduction . . . . .	187
9.2	Comparison of the results to other analyses . . . . .	187
9.2.1	Low- $E_T^{\text{miss}}$ channel . . . . .	187
9.2.2	Comparison to CMS results . . . . .	189
9.2.3	ATLAS exclusion in branching ratio plane . . . . .	191
9.3	Prospects and reach of the search . . . . .	194
9.3.1	Non-simplified models . . . . .	194
9.3.2	Long-lived scenarios . . . . .	196

9.3.3 New dark matter models . . . . .	198
<b>Conclusion</b>	<b>201</b>
<b>A <math>E_T^{\text{miss}}</math> trigger calibration</b>	<b>203</b>
<b>B Definitions of all analysis regions</b>	<b>213</b>
B.1 Region definitions . . . . .	213
B.2 Background composition . . . . .	218
B.3 Post-fit data comparison in all analysis regions . . . . .	221
<b>Bibliography</b>	<b>227</b>

# Introduction

The quest to understand the fundamental structure of matter has pushed the limits of human knowledge throughout the ages. From ancient times until now, steady progress in the field of physics have allowed to unveil many of the mysteries in nature, leading to technological advancements that have shaped the development of society itself. With the discovery of new phenomena, there is always a new window into unexplored regions of the universe that becomes available, and so was the case of the subatomic world. With the discovery of the electron and the nucleons inside the atoms, the indivisible was divided, and everything we knew about the smallest scales of nature was put under the microscope, quite literally. This journey culminated in the Standard Model of particle physics, one of the most complex and successful theories ever weaved, which was completed in 2012 with the discovery of the Higgs boson. Some of its properties, however, are yet to be understood both from theoretical and experimental points of view. For this reason, the large amounts of data collected in the LHC physics program are thoroughly inspected, aiming at unravelling the most fundamental matter constituent and their interactions in nature.

In spite of its accuracy, the Standard Model has intrinsic shortcomings that are reinforced by the experimental confirmation of the Higgs boson, one particular detail being its relatively light mass compared to the Planck scale. The theory is therefore considered to be a low-energy approximation to a more generic model, with several extensions being proposed that can successfully reproduce its behaviour in the infrared limit. One such theory, characterized by its mathematical elegance, is supersymmetry. It predicts the existence of extra superpartner particles for each of the fields that exists in the Standard Model, which can in turn solve some of its inconsistencies. While there are a plethora of supersymmetric models, with very different characteristics and phenomenology, a particular interesting scenario is that of “natural” theories where the fine-tuning problems of the Standard Model are solved without needing to introduce ad hoc constraints. In these models, the phenomenology typically include light superpartners to the Higgs boson, also known as higgsinos, that can be produced in the energy scale explored by the LHC.

The work presented in this dissertation includes a search for scenarios where higgsinos are produced in pairs in the LHC proton-proton collisions as recorded by the ATLAS experiment, where they are assumed to involve Higgs bosons in their decay chain. It exploits the large decay branching fraction of the Higgs boson to  $b$ -quarks in order to design the final states targeted, together with missing energy carried by other supersymmetric particles that leave the experiment undetected. The analysis was largely optimized with respect to a similar search carried out in the past, resulting in excellent discovery potential and sensitivity.

This thesis is organized as follows. In chapters 1 and 2 the theoretical framework for the models targeted in this search are developed, detailing the motivation for supersymmetric models. Chapters 3 and 4 provide an overview of the experimental setup, while chapters 5 and 6

present essential aspects of the data analysis carried out. In chapter 7 the analysis strategy for this search is discussed, with all results collected in chapter 8. Finally, in chapter 9 the results of the search are compared to other analyses targeting similar model, and prospects for future search are briefly discussed.

Most of the work collected in this thesis is presented in the context of the mentioned search. However, I have been involved in other searches for supersymmetric particles involving multiple  $b$ -jets and missing energy in the final state, both in strong-production and electroweak scenarios. Throughout my Ph.D studies I have contributed to the analyses that lead to the following publications:

- ATLAS Collaboration, “Search for pair production of higgsinos in final states with at least three  $b$ -tagged jets in  $\sqrt{s} = 13$  TeV  $pp$  collisions using the ATLAS detector”, Phys. Rev. D 98, 092002 (2018).
- ATLAS Collaboration, “Search for gluino-mediated stop and sbottom pair production in events with  $b$ -jets and large missing transverse momentum”, ATLAS-CONF-2018-041.

The search for higgsino pair production presented in chapters 7 and 8 is currently at the final stages of internal approval within the ATLAS Collaboration, and is expected to be submitted to a scientific journal by summer 2022.

In addition to the contributions to supersymmetry searches, I have also been heavily involved in the commissioning and validation of a part of the trigger system in the ATLAS experiment. As discussed in chapter 4, the Level-1 Topological system provides the experiment with large reduction of trigger rates without biasing complex signatures that are essential for some of the analyses in the collaboration. This system was validated in real time during LHC Run 2 by means of a simulation that reproduces the expected behaviour of the firmware implementation. This simulation is also ported to the software that will be used during LHC Run-3, where I have largely contributed as well, benefiting from the experienced gained during Run 2 data taking periods.

# Chapter 1

# The Standard Model of Particle Physics

## 1.1 Introduction

The journey into the subatomic world began with the discovery of the electron at the end of the 19th century, and throughout the following decades remarkable advances were made to the field, from Gell-Mann’s quark model [1] that unifies atomic nucleons with the plethora of observed mesonic and baryonic resonances, to the development of a theory for particle behaviour at the quantum scale. All of these advancements culminated in the second half of the 20th century with the formulation of the *Standard Model* (SM) [2–4], a theory describing the most fundamental particles in nature and their interactions. The model has been tested experimentally ever since to unprecedented levels of precision, and so far the SM can predict the empirical observations with outstanding accuracy. Many of the predictions of the model, such as the existence of the top quark or the Higgs boson, have been corroborated in the Large Electron-Positron Collider (LEP), the Tevatron Collider and, most recently, the LHC. Nonetheless, this theory is considered to be a low energy approximation to a more complete description of nature, although no clear evidence has been found to this day of what physics beyond the Standard Model (BSM) could be.

The mathematical formalism used in the SM is Quantum Field Theory (QFT), where all particles are assumed to be excitations of fields that permeate the universe. The description of nature in the SM starts with a set of fermionic and scalar fields whose actions are invariant to a set of local gauge transformations, and from here it is possible to extract the complete matter content in the model and the interactions they experience. This paradigm includes different phenomena such as non-perturbative behaviour in the low energy regime of the strong interactions or spontaneous symmetry breaking in the electroweak sector, while the predictive power of the theory is ensured thanks to it being renormalizable.

This chapter is organized as follows: in section 1.2 an overview of the fields described by the SM and their properties is provided, while the general mechanism that give rise to interactions in the SM is discussed in section 1.3, with detailed descriptions for the strong and electroweak interactions in subsections 1.3.1 and 1.3.2, respectively. The spontaneous symmetry breaking mechanism that leads to particle masses is described in section 1.4. Finally, the general limitations of the SM and experimental tests of the theory are discussed in sections 1.5 and 1.6, respectively.

## 1.2 Particle content of the Standard Model

The SM is a quantum field theory, in which the particles can be understood as excitations of quantum fields. These fields have different properties, but they can generally be divided in three sectors: the fermions, which constitute all existent visible matter; the vector bosons, that carry the interactions between particles, and the scalar boson sector, containing only the Higgs boson, responsible for the mass of all fermions and vector bosons.

Firstly, all the fundamental building blocks of matter are represented by spin- $1/2$  fermionic fields. According to the interactions that these particles can undergo, the fermionic sector can be further divided in two groups:

- The leptons can only experience electroweak interactions. They come in three identical copies commonly referred to as generations or families, which only differ in the particle masses, as shown in figure 1.1. The three copies of these fields give raise to the electron, muon and tau particles, and the three of them have negative equal charge (positive in the case of their anti-particles). Each of these three particles is associated to a neutrino which, although assumed to be massless in the SM description, in reality they are electrically neutral, very light massive particles.
- The quarks can experience strong and electroweak interactions. They also come in three generations, each of them containing an “up-type” particle (quarks up, charm and top) and a “down-type” one (quarks down, strange and bottom). All of them have fractional electric charge compared to that of the electron as specified in figure 1.1, with the up-type quarks having charge equal to  $2/3$  and the down-type quarks  $-1/3$ . They also carry a “color” charge associated to the strong interaction. Quarks do not exist in free states in nature since the strong interaction binds them together fast enough after they are produced, a process commonly referred to as “hadronization”. The only exception is the top quark: due to its short lifetime it decays before hadronizing, providing an experimental probe for free quarks to study their properties outside of bound states.

In the SM framework, fermions are described by means of Dirac spinors  $\psi$ . They are 4-component vectors that arise as solutions of the Dirac equation

$$(i\gamma^\mu \partial_\mu - m) \psi = 0, \quad (1.1)$$

where  $\gamma^\mu$  are the Dirac matrices and  $m$  is the mass of the fermion, and where natural units are used ( $\hbar = c = 1$ ). The Dirac equation describes fermions when they are free from any potential, and it has four independent solutions: two with energy  $E > 0$  and two with energy  $E < 0$ , which describe a fermion and its antiparticle, respectively. This twofold degeneracy of the energy eigenstates provides the necessary means to have a mathematical description of the spin components of the solutions. This degeneracy can be broken using the spin component of the (anti-) fermion to label the two solutions with positive (negative) energy. The solutions with positive helicity are generally referred to as “right-handed” while the negative helicity solutions are known as “left-handed”.

The fermion sector of the SM can be mathematically described by three identical copies of the same structure, one for each family. For each of them, two left-handed  $SU(2)$  doublets are included, one for leptons and one for quarks, plus three right-handed  $SU(2)$  singlets that account for the up- and down-type quarks and the massive lepton in the family. The three copies are

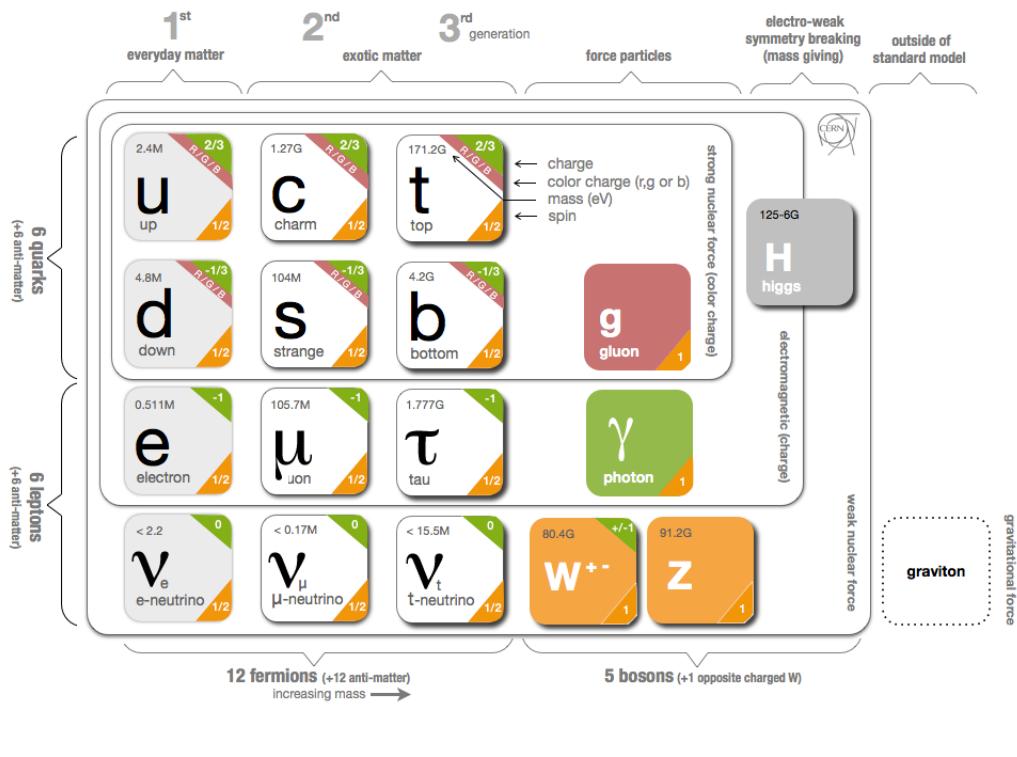


Figure 1.1: Summary table of all fundamental particles considered in the SM together with their physical properties as measured experimentally [5].

written explicitly with the following notation, that will be used throughout this chapter:

$$\begin{aligned}
 \Psi_{L,i}^{\text{lep}} &= \begin{pmatrix} \nu_{e,L} \\ e_L \end{pmatrix}, \begin{pmatrix} \nu_{\mu,L} \\ \mu_L \end{pmatrix}, \begin{pmatrix} \nu_{\tau,L} \\ \tau_L \end{pmatrix}; \\
 \Psi_{R,i}^{\text{lep}} &= e_R, \quad \mu_R, \quad \tau_R; \\
 \Psi_{L,i}^q &= \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \begin{pmatrix} c_L \\ s_L \end{pmatrix}, \begin{pmatrix} t_L \\ b_L \end{pmatrix}; \\
 \Psi_{R,i}^u &= u_R, \quad c_R, \quad t_R; \\
 \Psi_{R,i}^d &= d_R, \quad s_R, \quad b_R;
 \end{aligned} \tag{1.2}$$

where the  $u, d$  superscripts in the last two equations indicate the “up” and “down” quark flavours, and the  $L$  and  $R$  subscripts denote left- and right-hand projections of the fermion field. These projections can be obtained in the relativistic regime, or in the case of massless particles, by means of the chirality operator,  $\gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3$ , as

$$P_R \equiv \frac{1}{2} (1 + \gamma^5), \quad P_L \equiv \frac{1}{2} (1 - \gamma^5), \tag{1.3}$$

which are the right-hand and left-hand projection operators, since in these cases the chirality and helicity operators are approximately equal. The lack of right-handed neutrinos in eq. 1.2 is due to the nature of the weak interactions: it couples charged leptons only to left-handed neutrinos (or right-handed anti-neutrinos). Given that neutrinos can only experience weak interactions, there is no experimental evidence of right-handed neutrinos (or left-handed anti-neutrinos), and

therefore they are excluded from the SM description. Mathematically this is only compatible with the Dirac equation if neutrinos are purely massless, in which case their properties can completely be described with 2-component Weyl spinors.<sup>1</sup>

Secondly, the force carriers of the electroweak and strong interactions are represented by spin-1 bosonic fields. They arise in the SM as a consequence of the local gauge invariance requirement and are therefore denominated gauge bosons. The mediators of the strong interaction, the gluons, are massless and electrically neutral, and carry color charge which results in interactions not only with quarks but also with other gluons. The mediators of the weak interaction are the  $W^\pm$  and  $Z$  bosons, the former being electrically charged with charge  $Q = \pm 1$  respectively, and the latter being electrically neutral, and they are also self-interacting. Their masses are specified in figure 1.1. The carrier of the electromagnetic force, the photon ( $\gamma$ ), is massless and electrically neutral, and cannot interact with other photons.

Finally, the last piece of the particle content in the SM is the Higgs boson. It does not arise in the same way than the gauge bosons, but rather from the Higgs mechanism that allows for mass terms for all the particles in the Lagrangian, as described in section 1.4. It is a scalar (spin-0), electrically neutral particle that was first observed in 2012 by the ATLAS [6] and CMS [7] experiments using data from the LHC. The Higgs boson has tree-level couplings to other particles only if they are massive, and the coupling strength increases with the mass of such particles. In addition, it also has self-interactions, where vertices with three or four Higgs bosons are possible.

### 1.3 Interactions in the Standard Model

In the SM, interactions arise due to the presence of symmetries, which can be described mathematically as transformations that leave the Lagrangian of the system unchanged. These transformations are described in the framework of Group Theory as elements of a group,<sup>2</sup> with some relevant examples given by the set of all phase transformations,  $U(1)$ , or the set of all unitary 2-dimensional matrices with determinant 1,  $SU(2)$ . Symmetries are of great importance in physics since they lead, through Noether’s theorem, to conservation laws. In the case of space-time transformations, the Poincaré group comprises translations, rotations and Lorentz boosts, and it is invariance under such group that gives rise to momentum, energy and angular momentum conservation laws. Other “internal” transformations (i.e. transformations that commute with the space-time components of the fields) are referred to as “gauge” transformations.

A Lagrangian that remains invariant under a gauge transformation is called gauge invariant, and it is the promotion of a global gauge symmetry to a local one that provides the necessary fields that mediate the interactions among particles. A good way to illustrate this could be to examine the action of the group  $U(1)$  on a free fermion  $\psi$  with electric charge  $q$ . The Lagrangian

---

<sup>1</sup>The SM description is also compatible with neutrinos being *Majorana particles*, which follow a real wavefunction equation, the Majorana equation (in contrast to the complex Dirac equation). Phenomenologically, it would imply that neutrinos are their own antiparticles, which has direct consequences that could be observed in particle and nuclear physics experiments. The most typical example is the *neutrinoless double β-decay*, where two neutrons inside an atomic nucleus decay simultaneously into two protons and emitting only two electrons in absence of any neutrino. So far no strong evidence has been found to accept or reject the Majorana neutrino hypothesis.

<sup>2</sup>A *group* is a set of elements equipped with an operation, called group multiplication, that associates any pair of elements with a third element inside the same set with the properties (i) that the group multiplication is associative, (ii) that there exists an identity element such that the multiplication of any element by the identity results in the same element, and (iii) that for each element in the group there exists an inverse element such that the multiplication of both results in the identity.

of such physical system is

$$\mathcal{L} = i\bar{\psi}\gamma_\mu\partial^\mu\psi - m\bar{\psi}\psi, \quad (1.4)$$

which, after the application of Euler-Lagrange equations, results in the Dirac equation 1.1. The group  $U(1)$  is composed of one-dimensional unitary matrices, condition that is only fulfilled by complex numbers with unit modulus. The elements of  $U(1)$  can therefore be expressed as  $e^{i\alpha}$ , where  $\alpha$  is constant. The action of  $U(1)$  on the Lagrangian given by eq. 1.4 translates into a global phase shift of  $\alpha$  that clearly leaves the Lagrangian, and consequently the equations of motion of the system, unchanged, making the global phase  $\alpha$  physically unmeasurable.

This symmetry, however, could be made more general by requiring that the phase by which the system is shifted depends on space-time coordinates,  $\alpha = \alpha(x)$ . The phase introduced in the system by the  $U(1)$  transformation is not constant, or “global”, anymore and rather it changes for each space-time point. On closer inspection, however, it is easy to see that the Lagrangian is no longer invariant under this type of transformation, since the derivative now transforms as

$$\partial^\mu\psi \rightarrow e^{i\alpha(x)}\partial^\mu\psi + ie^{i\alpha(x)}\psi\partial^\mu\alpha(x), \quad (1.5)$$

where the second term breaks the gauge invariance since it cannot be cancelled out with any other term. This can be fixed by implementing a redefinition of the derivative acting on the fermion field, in such way that the derivative term transforms covariantly under the group  $U(1)$ . The new “covariant derivative” must include a new vector field that, after a local phase shift, can cancel out the unwanted symmetry breaking term, which can be achieved by the definition

$$D_\mu \equiv \partial_\mu - iqA_\mu, \quad (1.6)$$

where the new “gauge” field  $A_\mu$  transforms under  $U(1)$  as  $A_\mu \rightarrow A_\mu + \frac{1}{q}\partial_\mu\alpha(x)$ . The derivative term with the new definition thus transforms as

$$\begin{aligned} D^\mu\psi \rightarrow D'^\mu\psi' &= \partial^\mu\left(e^{i\alpha(x)}\psi\right) - iq\left(A^\mu + \frac{1}{q}\partial^\mu\alpha(x)\right)\left(e^{i\alpha(x)}\psi\right) \\ &= e^{i\alpha(x)}(\partial^\mu - iqA^\mu)\psi \\ &= e^{i\alpha(x)}D^\mu\psi, \end{aligned} \quad (1.7)$$

and therefore the Lagrangian is now invariant under local gauge transformations. Writing the new derivative explicitly, the Lagrangian is

$$\begin{aligned} \mathcal{L} &= i\bar{\psi}\gamma^\mu D_\mu\psi - m\bar{\psi}\psi \\ &= \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi + q\bar{\psi}\gamma^\mu\psi A_\mu, \end{aligned} \quad (1.8)$$

where the new term that appears accounts for the coupling between the fermion and the new vector gauge field. To complete the physical description of the new field, a kinetic term for the new particle needs to be added to the Lagrangian. The only possibility is to add a term involving the gauge invariant tensor  $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ , and the complete Lagrangian is

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi + q\bar{\psi}\gamma^\mu A_\mu\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \quad (1.9)$$

A mass term for the new field cannot be added due to gauge invariance. Indeed, a term proportional to  $A_\mu A^\mu$ , after a  $U(1)$  transformation, would produce unwanted terms proportional to the derivatives of the parameter  $\alpha(x)$ . The only way for the gauge field to acquire a mass without breaking the gauge invariance of the Lagrangian is through the Higgs mechanism described in section 1.4.

It is easy to see that the application of the Euler-Lagrange equations for the new gauge field  $A_\mu$  to eq. 1.9 results in the Maxwell equations in presence of a source of electromagnetic field  $\psi$ , while the equations of motion for the fermion  $\psi$  are those of a charged fermion moving in an electromagnetic potential  $A_\mu$ . Thus, this new interaction that appears could be seen as the coupling between the field  $A_\mu$  to an electromagnetic current

$$j_{EM}^\mu = -q\bar{\psi}\gamma^\mu\psi. \quad (1.10)$$

This is no accident, since  $U(1)$  was chosen specifically because it is the symmetry that leads to quantum electrodynamics (QED), and the resulting Lagrangian is the complete description for that theory.

In the SM description, all the interactions arise through the same mechanism: a global gauge symmetry of the Lagrangian is promoted to local symmetry, and the required addition of new gauge fields lead to interactions between the fermions through couplings to such new fields. In the example presented above the symmetry group is abelian, which means that the multiplication of any two elements of the group is commutative. The symmetries that lead to the strong and electroweak interactions do not have this property, and it needs to be accounted for in the corresponding covariant derivative since otherwise there would appear extra terms breaking the gauge invariance. The SM is defined by the symmetry  $SU(3)_C \times SU(2)_L \times U(1)_Y$  that leads to the fundamental interactions between particles. The invariance of the Lagrangian under  $SU(3)$  alone provides the mathematical description for the strong interactions (the subscript “C” stands for “color”), while the symmetry  $SU(2)_L \times U(1)_Y$  leads to the unified description of the weak and electromagnetic interactions (subscripts stand for “left”, referring to the weak interaction coupling only to left-handed fermions, and “hypercharge”, respectively). In the following subsections the different interactions are explained in a detailed manner.

### 1.3.1 Quantum Chromodynamics

The same procedure illustrated above can be applied to other symmetries of the Lagrangian to find the mathematical description for other interactions. In particular, the theory of strong interactions can be obtained by requiring local gauge invariance under the group  $SU(3)$ , and is commonly referred to as *Quantum Chromodynamics* (QCD). The charge associated to this interaction is called “color”, and it can take three different values denoted as red, green or blue.

In contrast with the group  $U(1)$ , any two elements  $g_1, g_2$  from  $SU(3)$  do not generally commute;  $SU(3)$  is a so-called *non-abelian* group. In this case, the *generators*<sup>3</sup>  $T_a$ ,  $a = 1, \dots, 8$ , of the group do not commute, and due to the closure property of the group the product of any two generators results in an element of the group, which can in turn be written as a linear combination of generators. The commutation relations for the generators of any group can therefore be written as

$$[T_a, T_b] = if_{abc}T_c, \quad (1.11)$$

where  $f_{abc}$  are the *structure constants* of the group, and there is an implicit summation over repeated indices. In the case of *abelian* groups, all the generators commute with each other and the structure constants are zero. Any element of the group can be written in terms of the generators as

$$U = e^{i\alpha_a T_a}, \quad (1.12)$$

---

<sup>3</sup>The generators of a Lie group are operators that describe the behaviour of the group elements around the identity. Given the analytical continuity of these groups, the generators are enough to span all the elements of the group.

where  $\alpha_a$  are real parameters that characterize the element of the group. Without loss of generality, the study could be restricted to infinitesimal transformations ( $\alpha \ll 1$ ), for which it is enough to perform a Taylor expansion to first order in  $\alpha$  of eq. 1.12.

The effect of promoting the global  $SU(3)$  symmetry to local gauge invariance can be examined by looking at the Lagrangian for free quarks,

$$\mathcal{L} = \sum_f \bar{\psi}_f (i\gamma_\mu \partial^\mu - m_f) \psi_f, \quad (1.13)$$

where  $f = u, d, s, c, b, t$  denotes the flavour of the quarks. The quark fields  $\psi_f$  and their derivatives transform under an infinitesimal element of  $SU(3)$  as

$$\psi_f \rightarrow [1 + \alpha_a(x)T_a] \psi_f, \quad (1.14a)$$

$$\partial_\mu \psi_f \rightarrow [1 + i\alpha_a(x)T_a] \partial_\mu \psi_f + iT_a \psi_f \partial_\mu \alpha_a(x), \quad (1.14b)$$

where once again the derivative term introduces unwanted terms, breaking the gauge invariance of the Lagrangian: new fields need to be included in order to restore the symmetry. However, it is easy to see that a similar definition of the gauge fields to the one described in section 1.3 is not enough to define a new derivative that result in  $D_\mu \psi_f$  transforming covariantly under  $SU(3)$ . In order to have a derivative term that transforms as the field  $\psi_f$  in eq. 1.14a the new gauge fields  $G_\mu^a$  introduced in the covariant derivative therefore need to transform in such a way that these new terms that appear due to  $SU(3)$  being non-abelian are absorbed as well. The way to achieve that is to define a covariant derivative as

$$D_\mu \equiv \partial_\mu + ig_s T_a G_\mu^a, \quad (1.15)$$

where  $g_s$  is the coupling strength, and the fields  $G_\mu^a$  transform under the group as

$$G_\mu^a \rightarrow G_\mu^a - \frac{1}{g_s} \partial_\mu \alpha_a - f_{abc} \alpha_b G_\mu^c. \quad (1.16)$$

It is common to use the strong fine-structure constant  $\alpha_s = \frac{g_s}{4\pi}$  as well as a replacement of the coupling strength  $g_s$ , in particular to express experimental measurements.

The addition of a kinetic energy term for the gauge fields needs to be made covariant as well, which can only be achieved by a redefinition of the gauge invariant field strength tensor  $F_{\mu\nu}^a$  that accounts for the non-commuting generators:

$$F_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a - g_s f_{abc} G_\mu^b G_\nu^c. \quad (1.17)$$

The new Lagrangian, once all these elements are incorporated, is

$$\begin{aligned} \mathcal{L} &= -\frac{1}{4} F_{\mu\nu}^a F^{\mu\nu,a} + \sum_f \bar{\psi}_f (i\gamma^\mu D_\mu - m_f) \psi_f \\ &= -\frac{1}{4} F_{\mu\nu}^a F^{\mu\nu,a} + \sum_f \bar{\psi}_f (i\gamma^\mu \partial_\mu - m_f) \psi_f - g \sum_f \bar{\psi}_f \gamma^\mu T_a \psi_f G_\mu^a. \end{aligned} \quad (1.18)$$

Once again the new term that appears describes the coupling of the quark fields to the new gauge bosons, the gluons. It is easy to see that the contraction of the strength field tensor  $F_{\mu\nu}^a F^{\mu\nu,a}$  leads to self-coupling terms for the gauge field, responsible for Feynman diagrams with vertices involving three and four gluons. This feature of the theory is a direct consequence of the group  $SU(3)$  being non-abelian and does not appear in abelian theories, like the case of QED

illustrated above. A mass term for the gauge field is also forbidden to preserve gauge invariance in the Lagrangian. It is important to note that in the expression 1.18 both the quark kinetic and mass terms, as well as the quark-gluon coupling term, only couple same flavour quarks together. This is relevant from a phenomenological point of view, as quark flavour mixing is forbidden in strong interactions and can only be realized through electroweak interactions.

Similarly to QED, QCD is also a renormalizable theory: the infinites appearing in the theory can be “absorbed” by redefining some of the physical parameters measured experimentally. This leads to a dependence of the redefined parameters of the theory on the energy scale at which they are measured. In the case of QCD, the fine-structure constant  $\alpha_s$  is modified by the renormalization requirements and, at first order, can be expressed as [8]

$$\alpha_s(Q^2) = \frac{12\pi}{(11n_C - 2n_f) \log\left(\frac{Q^2}{\Lambda^2}\right)}, \quad (1.19)$$

where  $n_C$  is the number of colors,  $n_f$  is the number of active quark flavours (i.e. with mass below the energy scale  $Q^2$ ) and  $\Lambda^2$  is the energy scale below which the coupling constant is large enough to break the perturbative approximation. In a universe with less than 16 quark flavours, the factor  $(11n_C - 2n_f)$  will always remain positive, which in turn results in  $\alpha$  decreasing when the energy scale  $Q^2$  increases. This is the origin of the asymptotic freedom in the theory: at small distances (or high  $Q^2$ ) the coupling constant is small enough to consider the quarks as free from the strong interaction. Inside hadrons, for example protons, the quarks behave like free particles. As distance between bound quarks increases (i.e.  $Q^2$  decreases), on the other hand, the coupling becomes larger and larger, leading to *confinement*. When two quarks are pulled apart, at a certain point the attraction between them due to the strong interaction is so intense that, in order to minimize the energy of the system, a new pair of quark-antiquark is created from the vacuum and combined with each of the original two quarks to form two new hadrons. For this reason, quarks can never be observed isolated, except for the top quark, whose lifetime is so short that there is no time for hadronization before it decays. In particle accelerators and colliders, the typical signature of strong interactions where a quark or a gluon is radiated is a collimated group of hadrons denominated *jet*, which result from the sequential hadronization process of the radiated quark/gluon and its decay products.

In the large  $Q^2$  limit, a perturbation expansion in terms of the coupling constant can be performed, leading to the Feynman diagrams for QCD. This approximation is only valid when the energy scale of the process  $Q^2$  is above the scale  $\Lambda^2$  in eq. 1.19. In this limit, the quarks and gluons inside a hadron can be considered as *partons*: point-like particles that are free inside the hadrons they constitute. Due to the uncertainty principle, the large energy scale  $Q^2$  translates into short interaction times which also justifies considering the partons involved in a given process as independent from the rest of “spectator” partons. This approximation is particularly useful in the study of scattering processes at very high energies, such as the ones produced in colliders like the LHC, as the interactions are not seen to happen between the complex bound systems of quarks that compound the hadrons involved in the collision but rather as a scattering process between the partons themselves.

### 1.3.2 Electroweak interactions

Similar to the case of QCD, the description for the weak and electromagnetic interactions can be achieved by examining another symmetry group of the Lagrangian:  $SU(2)_L \times U(1)_Y$ . In the first effort of unifying the charged and neutral weak currents together a main problem arises:

while the charged currents only couple to left-handed fermions, thus presenting a clear vector-axial vector (V-A) form, from neutrino scattering experiments it is known that the neutral currents have a small but non-zero component that couples to right-handed fermions [8]. It could be tempting to associate the weak interactions to the action of the group  $SU(2)_L$  alone coupling only to left-handed fields. This approach could successfully describe the vector-axial vector structure of the weak charged currents but it fails to provide an accurate depiction of the weak neutral currents. Using the Pauli matrices  $\sigma_i$ , it is possible to generate a set of weak currents that form an “isospin” triplet under  $SU(2)$  and whose corresponding charges generate the  $SU(2)_L$  algebra. However, the phenomenological implications from this approach do not match the observed behaviour of the weak neutral currents. Introducing the operators

$$\sigma^+ = \frac{1}{2}(\sigma_1 + i\sigma_2) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \sigma^- = \frac{1}{2}(\sigma_1 - i\sigma_2) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad (1.20)$$

and using as example the currents involving the first generation of leptons, it is immediate to arrive to the following expressions for the charged and neutral currents, where the notation introduced in eq. 1.2 is used:

While this formulation correctly describes the charged currents that in the SM couple to the  $W^\pm$  bosons, it also implies that there must be a neutral boson  $W^0$  that couples exclusively to left-handed fermions, which does not match the observed behaviour of the  $Z$  boson. Therefore the association of the  $SU(2)_L$  symmetry group to the weak interaction is not satisfactory.

Interestingly, a current originating from the action of a “hypercharge”  $U(1)_Y$  group, similar to eq. 1.10, is also neutral and couples to both left- and right-handed fermions,

$$J_\mu^Y = -q\bar{e}\gamma_\mu e = -q\bar{e}_R\gamma_\mu e_R - q\bar{e}_L\gamma_\mu e_L. \quad (1.22)$$

The solution presented in the SM is to associate the physical weak neutral and electromagnetic currents to orthogonal combinations of the currents  $J_\mu^3$ , coupling only to left-handed fermions,

and  $J_\mu^Y$ , that couples to both left- and right-handed fermions. This way, it is possible to work with currents that have well defined transformation properties under  $SU(2)_L$  and  $U(1)_Y$  and, while the currents  $J_\mu^3$  and  $J_\mu^Y$  lack of physical meaning by themselves, the combination of both can describe the experimental behaviour of electromagnetic and weak neutral currents.

The symmetry group  $SU(2) \times U(1)$  acts on the fermionic fields of the SM as

$$\begin{aligned}\Psi_L &\rightarrow \Psi'_L = U_{SU(2)} U_{U(1)} \Psi_L = e^{i\alpha_a T^a + i\beta Y} \Psi_L, \quad \text{with } a = 1, 2, 3 \\ \Psi_R &\rightarrow \Psi'_R = U_{U(1)} \Psi_R = e^{i\beta Y} \Psi_R,\end{aligned}\tag{1.23}$$

where  $\alpha_a$ ,  $\beta$  are arbitrary parameters and  $T^a$  and  $Y$  are the generators of the groups  $SU(2)_L$  and  $U(1)_Y$ , respectively. Once again, if the symmetry group  $SU(2) \times U(1)$  were to be promoted from global to local symmetry, a new covariant derivative would need to be defined in order to preserve gauge invariance. Similar to the case of QCD, the new derivative can be written as

$$D_\mu = \partial_\mu + igT_a W_\mu^a + i\frac{g'}{2} Y B_\mu, \quad \text{with } a = 1, 2, 3\tag{1.24}$$

In this way, four new fields are introduced: three fields  $W_\mu^a$  for the generators of the group  $SU(2)_L$  with coupling constant  $g$ , and one field  $B_\mu$  for  $U(1)_Y$  with the coupling  $g'$ . In order to preserve gauge invariance, the gauge fields need to transform as

$$\begin{aligned}W_\mu^a &\rightarrow W_\mu^{a'} = W_\mu^a - \frac{1}{g} \partial_\mu \alpha_a - f_{abc} \alpha_b W_\mu^c, \\ B_\mu &\rightarrow B'_\mu = B_\mu - \frac{1}{g'} \partial_\mu \beta.\end{aligned}\tag{1.25}$$

After including the new covariant derivative, the Lagrangian for the electroweak interactions reads

$$\begin{aligned}\mathcal{L} &= -\frac{1}{4} W_{\mu\nu}^a W^{\mu\nu, a} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \\ &+ \sum_{\text{flavor}} \left[ \bar{\Psi}_L \gamma^\mu \partial_\mu \Psi_L + ig W_\mu^a \bar{\Psi}_L T_a \gamma^\mu \Psi_L + i\frac{g'}{2} B_\mu \bar{\Psi}_L Y \gamma^\mu \Psi_L \right] \\ &+ \sum_{\text{flavor}} \left[ \bar{\Psi}_R \gamma^\mu \partial_\mu \Psi_R + i\frac{g'}{2} B_\mu \bar{\Psi}_R Y \gamma^\mu \Psi_R \right].\end{aligned}\tag{1.26}$$

The first line in eq. 1.26 contains the kinetic terms for the gauge fields  $W_\mu^a$  and  $B_\mu$  through the field strength tensors

$$\begin{aligned}W_{\mu\nu}^a &= \partial_\mu W_\nu^a - \partial_\nu W_\mu^a - g f_{abc} W_\mu^b W_\nu^c, \\ B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu,\end{aligned}\tag{1.27}$$

where  $f_{abc}$  are  $SU(2)_L$  structure constants. The second and third rows in eq. 1.26 are the kinetic and coupling terms for left- and right-handed fermions, respectively, where it is easy to identify the interaction terms with the gauge fields. The interaction between the fermions and the gauge bosons could be also be seen as the coupling of the gauge bosons to the  $SU(2)_L$  isospin triplet of currents defined in eqs. 1.21a–1.21c and the  $U(1)_Y$  hypercharge current:

$$\mathcal{L}_{\text{int}} = ig W_\mu^i J_i^\mu + i\frac{g'}{2} B_\mu J_Y^\mu.\tag{1.28}$$

No mass term has been added in the Lagrangian for the gauge bosons since, as mentioned above, it would break gauge invariance. However, mass terms for the fermions cannot be added either.

Using the electron field as an example, it is easy to see that a Dirac mass term like the one in eq. 1.4 leads to mixing between left- and right-handed fermions,

$$\begin{aligned} -m_e \bar{e} e &= -m_e \bar{e} \left[ \frac{1}{2} (1 - \gamma^5) + \frac{1}{2} (1 + \gamma^5) \right] e \\ &= -m_e (\bar{e}_R e_L + \bar{e}_L e_R), \end{aligned} \quad (1.29)$$

where terms like  $\bar{e}_L e_L$  vanish due to properties of  $\gamma^5$ . Given that  $e_L$  is part of a doublet and  $e_R$  is a singlet, a Dirac mass term does not transform well under  $SU(2)_L$  and therefore breaks the gauge invariance of the Lagrangian. This problem did not arise in the discussion of the strong interaction because it is blind to fermion chirality, but since the Lagrangian must be invariant under all  $SU(3)_c \times SU(2)_L \times U(1)_Y$ , no Dirac mass term can be present in the description at all and, just like the gauge bosons, the fermions only gain mass due to their coupling to the Higgs boson.

These new gauge fields cannot readily be identified with the  $W^\pm$ ,  $Z$  and  $\gamma$  bosons from section 1.2. As mentioned above, the physical weak and electromagnetic currents are obtained as linear combinations of the currents  $J_\mu^i$  and  $J_\mu^Y$ . In the same way, the physical bosonic fields can be obtained from linear combinations of the ones introduced by the covariant derivative. To reach the description for the  $W^\pm$  bosons, that couple only to the weak charged currents, it is enough to find the combination that couples to the physical currents  $J_\mu^\pm$  from eqs. 1.21a and 1.21b. Rewriting the interaction terms of the Lagrangian given by eq. 1.28 in the basis introduced in eq. 1.20, it is easy to see that the physical bosons  $W^\pm$  that couple to the currents  $J_\mu^\pm$  are

$$W_\mu^\pm = \frac{1}{\sqrt{2}} (W_\mu^1 \mp i W_\mu^2). \quad (1.30)$$

This description, however, cannot explain some well known processes, such as the weak decay  $K^+(u\bar{s}) \rightarrow \mu^+ \nu_\mu$ , since no mixing of different quark families is allowed in the Lagrangian in eq. 1.26. The solution lies in the distinction between the weak interaction eigenstates and the mass eigenstates: the physically observable quarks do not necessarily have to be the same fields that diagonalize the interaction terms. That way, it is possible to identify the interaction eigenstates with linear combinations of the mass eigenstates in a way such that the  $W^\pm$  bosons couple to physical states involving quarks from two different families. It is enough to introduce this mixing in the down-type quarks only while keeping the physical up-type quarks aligned with their corresponding interaction eigenstates. The interaction eigenstates of the down-type quarks  $d'$ ,  $s'$ ,  $b'$  can then be rewritten in terms of their mass eigenstates  $d$ ,  $s$ ,  $b$  as

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{\text{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{cd} & V_{td} \\ V_{us} & V_{cs} & V_{ts} \\ V_{ub} & V_{cb} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \quad (1.31)$$

where the *Cabibbo-Kobayashi-Maskawa* matrix  $V_{\text{CKM}}$  parametrizes the mixing among the three quark families. Taking into account properties of unitary matrices and symmetries of the quark fields, it is possible to reparameterize the CKM matrix in terms of only three angles and a complex phase.

Similar to the case of the  $W^\pm$  bosons, the  $Z$  boson and the photon  $\gamma$  must be a combination of the fields  $W_\mu^3$  and  $B_\mu$  as well. The relation between the two sets of fields can be obtained as a rotation of angle  $\theta_W$ , called *Weinberg* or *electroweak mixing angle*,

$$\begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix} = \begin{pmatrix} \cos \theta_W & -\sin \theta_W \\ \sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix}, \quad (1.32)$$

Fermion	Electric charge ( $Q$ )	Weak isospin ( $T^3$ )	Hypercharge ( $Y$ )
$\begin{pmatrix} \nu \\ e \end{pmatrix}_L$	0	$1/2$	-1
	-1	$-1/2$	-1
$\begin{pmatrix} u \\ d \end{pmatrix}_L$	$2/3$	$1/2$	$1/3$
	$-1/3$	$-1/2$	$1/3$
$e_R$	-1	0	-2
$u_R$	$2/3$	0	$4/3$
$d_R$	$-1/3$	0	$-2/3$

Table 1.1: Charge, weak isospin and hypercharge for all the fermion fields in the SM.

where  $A_\mu$  and  $Z_\mu$  are the mathematical description of the physical photon and  $Z$  boson, respectively. In terms of these new fields, the interaction term for neutral currents in the Lagrangian can be rewritten as

$$\mathcal{L}_{\text{int}}^{\text{NC}} = (ig \sin \theta_W J_3^\mu + ig' \cos \theta_W J_Y^\mu) A_\mu + (ig \cos \theta_W J_3^\mu - ig' \sin \theta_W J_Y^\mu) Z_\mu. \quad (1.33)$$

The first term, coupling only to the photon  $A_\mu$ , is the electromagnetic current. Comparing it to the current in eq. 1.10, it is easy to see that for this identification to be fulfilled the following relation must hold:

$$g \sin \theta_W = g' \cos \theta_W = e, \quad (1.34)$$

where  $e$  is the charge of the electron. Furthermore, for the identification of this linear combination of  $J_3^\mu$  and  $J_Y^\mu$  with the electromagnetic current, the weak isospin and hypercharge operators need to fulfil the *Gell-Mann - Nishijima* relation

$$J_{\text{EM}}^\mu = J_3^\mu + \frac{J_y^\mu}{2} \rightarrow Q = T^3 + \frac{Y}{2}. \quad (1.35)$$

The weak isospin values for each fermionic field can be obtained by acting on them with the operator  $T^3 = \frac{1}{2}\sigma^3$ , and together with their electric charges, the hypercharge  $Y$  for each fermion can be obtained. The three values are collected in table 1.1.

The second term in eq. 1.33 can be identified as the coupling between the  $Z$  boson and the weak neutral current. From the relations between the coupling constants in eq. 1.34 and the  $J_Y^\mu$  and  $J_{\text{EM}}^\mu$  currents, the physical weak neutral current can be written as

$$J_{\text{NC}}^\mu = J_3^\mu - \sin^2 \theta_W J_{\text{EM}}^\mu. \quad (1.36)$$

It is immediate to see now how the weak neutral current has a right-handed component that comes into play through  $J_{\text{EM}}^\mu$  whose impact on the pure left-handed  $SU(2)$  current  $J_3^\mu$  will be determined by the magnitude of  $\sin^2 \theta_W$ , solving the original problem of finding a unified description for all weak currents.

## 1.4 Spontaneous symmetry breaking and particle masses

As mentioned in section 1.3, neither vector bosons nor fermions are allowed to have a mass term in the SM Lagrangian due to gauge invariance. However, all the fermions in the SM that have been observed so far are massive, and so are the  $W^\pm$  and  $Z$  bosons. A different approach to introduce particle masses in the Lagrangian without explicit Dirac mass terms was proposed

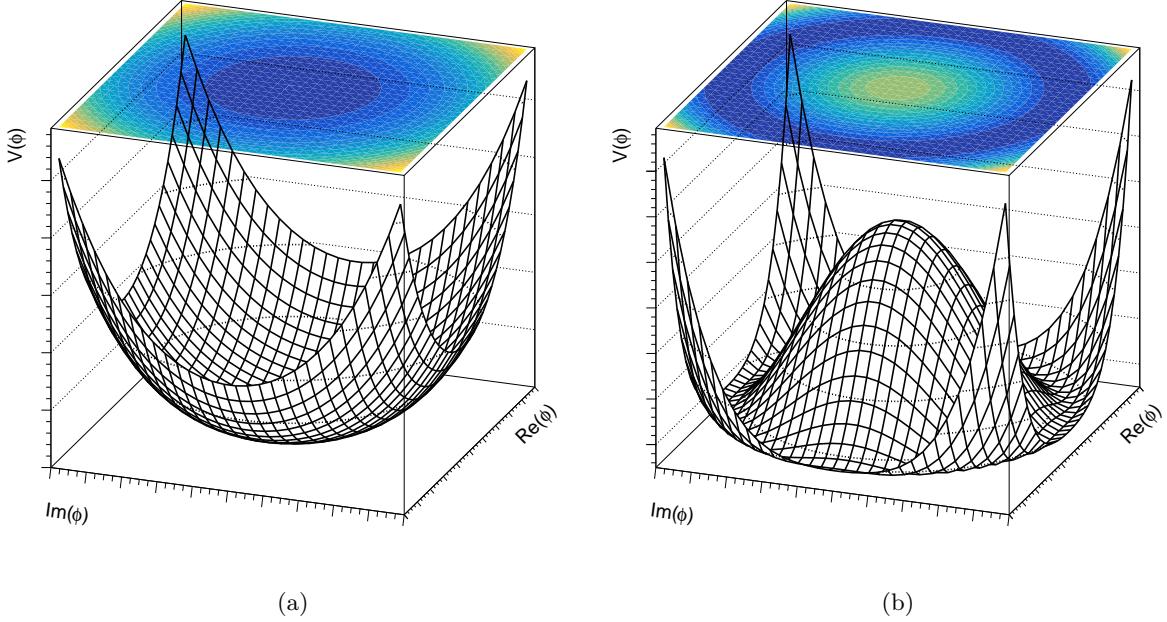


Figure 1.2: Shape of the potential  $V(\phi)$  for the cases (a)  $\lambda > 0, \mu^2 > 0$  and (b)  $\lambda > 0, \mu^2 < 0$ .

in 1964 by Peter Higgs [9] and by François Englert and Robert Brout [10] in what today is known as the *Higgs mechanism*. In essence, a new (Higgs) field that couples to fermions and the  $SU(2)_L \times U(1)_Y$  gauge bosons is introduced into the Lagrangian together with its corresponding potential. Initially the Higgs potential is invariant under the symmetry group that characterizes the gauge interactions. However, after choosing one particular ground state for the Higgs field the symmetry is *spontaneously broken*, or rather it remains hidden, which leads to new mass terms arising in the SM Lagrangian for each of the particles that is coupled to the new field. These new terms that appear after the symmetry is broken actually preserve the renormalizability [11], and hence the predictability, of the theory, in contrast with Dirac mass terms added “by hand” that would lead to a non-renormalizable Lagrangian.

The Higgs mechanism relies on spontaneous symmetry breaking and the Goldstone theorem, which states that for each symmetry generator that is broken when choosing a specific *vacuum expectation value* (VEV) for the Higgs field there will be a new massless boson added to the Lagrangian. These principles can be illustrated with a simple Lagrangian including only a scalar complex field  $\phi = (\phi_1 + i\phi_2)/\sqrt{2}$  and its potential,

$$\mathcal{L} = (\partial_\mu \phi)^* (\partial^\mu \phi) - V(\phi), \quad V(\phi) = \mu^2 \phi^* \phi + \lambda (\phi^* \phi)^2. \quad (1.37)$$

The behaviour of the field will depend on the sign of the constant  $\mu^2$ . If  $\mu^2 > 0$  (figure 1.2a), the potential  $V(\phi)$  just describes a massive field  $\phi$  with mass  $\mu^2$  and with a self-interacting vertex with coupling strength  $\lambda$ . The case where  $\mu^2 < 0$  (figure 1.2b), however, creates a potential with an infinite number of minima that can lead to different ground states for the field  $\phi$ . It is nature that chooses one of those minima as the physical ground state, thus breaking the symmetry in the potential. It is easy to gain some intuition about the behaviour of the scalar field in this case by performing small perturbations around the extremes of the potential,

$$\phi_1 = \phi_2 = 0, \quad (1.38a)$$

$$\phi_1^2 + \phi_2^2 = v^2 = -\frac{\mu^2}{\lambda}. \quad (1.38b)$$

The extreme in eq. 1.38a is a maximum in the potential, as shown in figure 1.2b, and therefore any small perturbation around it would make the field *roll* to smaller values of  $V(\phi)$ , breaking the perturbative approximation. The set of extremes in eq. 1.38b, on the other hand, correspond to a ring of local stable minima. If the field is assumed to take one of them in particular, it is easy to see that the perturbations around that point behave differently depending on the direction. For any movement along the tangent to the circle given by eq. 1.38b the field changes to a configuration with exactly the same potential, so the field would not feel any “resistance” in the motion along this direction as there is no needed energy for the transition to one ground state to another. Movement along the radial direction, however, would require the field to acquire extra energy as every configuration has a larger potential value than the ground state. In fact, if the perturbations are small enough, the local potential in the radial direction seen by the field would approximately be the same as the one shown in figure 1.2a, which describes the equation of motion of a massive particle.

To prove this, it is enough to choose one ground state for the field  $\phi$  and perform small perturbations in the Lagrangian in eq. 1.37 around that point. Without loss of generality the ground state could be chosen to be  $\phi_1 = v$ ,  $\phi_2 = 0$ , which spontaneously breaks the symmetry in the potential. Small perturbations can be studied by rewriting the field as

$$\phi(x) = \frac{1}{\sqrt{2}} [v + \eta(x) + i\xi(x)], \quad (1.39)$$

where  $\eta(x)$  and  $\xi(x)$  represent quantum fluctuations around the minimum in the radial and tangent directions, respectively. The new Lagrangian with the redefinition of the field  $\phi$  is

$$\mathcal{L} = \frac{1}{2} (\partial_\mu \eta)^2 + \frac{1}{2} (\partial_\mu \xi)^2 - \lambda v^2 \eta^2 - \lambda \left[ v\eta^3 + v\eta\xi^2 + \frac{1}{2}\eta^2\xi^2 + \frac{1}{4}\eta^4 + \frac{1}{4}\xi^4 \right], \quad (1.40)$$

where now, for the field  $\eta$  (i.e. fluctuations in the radial direction around the minimum) there is a mass term with the right sign with respect to its kinetic term  $\frac{1}{2} (\partial_\mu \eta)^2$ , where as for the field  $\xi$  there is no mass term at all. The rest of the terms in the Lagrangian are the cubic and quartic couplings between the new fields  $\eta$  and  $\xi$ . The field  $\eta$  has obtained a mass  $m_\eta = \sqrt{2\lambda v^2}$ , as expected from the shape of the potential. At the same time, a new massless field  $\xi$  has been introduced in the Lagrangian, completing the degrees of freedom that were originally introduced with the complex scalar field  $\phi$ .

### 1.4.1 The Higgs mechanism

As shown in the example above, the spontaneous breaking of a symmetry can raise a mass term for the fields that acquire a VEV, but it does not affect other fields that could be present in the Lagrangian as long as they do not interact with the ones obtaining mass. In addition, a number of massless bosons are introduced into the description after the symmetry breaking, which could potentially lead to the prediction of unobserved SM massless particles.

The subtlety of the Higgs mechanism lies on the fact that when the gauge bosons associated to the broken symmetry are coupled to the scalar fields that assist in the symmetry breaking, then the massless Goldstone bosons are “absorbed” by the gauge bosons, turning into their longitudinal degrees of freedom and allowing them to acquire a mass. This is easily accomplished in the framework of local gauge symmetries since the covariant derivatives have an implicit

coupling to the gauge fields. In the case of the SM, given that the gluons are expected to be massless, it is enough to introduce a  $SU(2) \times U(1)$  invariant term in the Lagrangian such as

$$\mathcal{L}_H = (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi), \quad (1.41)$$

where  $\phi$  is a  $SU(2)$  Higgs complex doublet with weak hypercharge  $Y = 1$ ,

$$\phi \equiv \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}, \quad (1.42)$$

where all fields  $\phi_i$  are real, and  $V(\phi)$  is the Higgs potential,

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2. \quad (1.43)$$

In this case,  $D_\mu$  is the covariant derivative introduced in the discussion of the  $SU(2) \times U(1)$  local symmetry. The interesting scenario is where  $\mu^2 < 0$  and  $\lambda > 0$ , so some of the new Higgs fields can acquire mass after breaking the gauge symmetry by choosing a VEV for  $\phi$ . Any choice of the ground state leads to the breaking of the  $SU(2) \times U(1)$  symmetry, generating a mass term for each broken generator. However, if there is a subset of transformation under which the chosen ground state remains invariant, the associated gauge boson(s) to said subset of transformations will not acquire mass. Given that in the SM only three out of the four  $SU(2) \times U(1)$  gauge bosons have been observed to be massive, the chosen vacuum needs to remain invariant under at least one generator. The only choice for such vacuum is  $\phi^+ = 0, \phi^0 = v = \sqrt{-\mu^2/\lambda}$ , which is invariant<sup>4</sup> under the  $U(1)_{\text{em}}$  symmetry group with generator  $Q = T^3 + Y/2$ , resulting in massless photons. All other possible combinations of the generators of  $SU(2) \times U(1)$  will be broken with this choice of vacuum, leading to massive  $W^\pm$  and  $Z$  bosons.

Small perturbations around the chosen vacuum state can be parametrized as

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} \xi_1(x) + i\xi_2(x) \\ v + h(x) + i\xi_3(x) \end{pmatrix}, \quad (1.44)$$

where the new fields  $\xi_i, h$  are introduced to account for quantum fluctuations in each direction of the  $\phi_i$  space. Substituting this approximation of the Higgs doublet about the chosen energy minimum into the Lagrangian in eq. 1.41 the following expression is obtained:

$$\begin{aligned} \mathcal{L}_H = & \sum_i \frac{1}{2} \partial_\mu \xi_i \partial^\mu \xi_i + \frac{1}{2} \partial_\mu \eta \partial^\mu h \\ & + \frac{v}{2} [g \partial_\mu \xi_1 W_2^\mu + g \partial_\mu \xi_2 W_1^\mu - \partial_\mu \xi_3 (g W_3^\mu + g' B^\mu)] \\ & + \sum_i \frac{g^2 v^2}{8} W_{i,\mu} W_i^\mu + \frac{g'^2 v^2}{8} B_\mu B^\mu - \frac{gg' v^2}{4} W_{3,\mu} B^\mu - \lambda v^2 h^2 + \text{interaction terms}, \end{aligned} \quad (1.45)$$

where the gauge bosons  $W_i^\mu, B^\mu$  of section 1.3.2 have been written explicitly. After the spontaneous symmetry breaking the sought mass terms for the gauge bosons appear, but the massless goldstone bosons  $\xi_i$  seem to still be present in the Lagrangian. However, the off-diagonal terms  $\partial_\mu \xi_i W_j^\mu$  are an indication that these fields do not correspond to the real physical degrees of freedom available.

---

<sup>4</sup>Since  $Y(\phi^0) = 1$  and  $T^3(\phi^0) = -\frac{1}{2}$ , the field is electrically neutral ( $Q\phi^0 = 0$ ), and therefore any transformation generated by its charge will leave the field unchanged, i.e.  $\phi^0 \rightarrow \phi^{0'} = e^{i\alpha(x)Q} \phi^0 = \phi^0$ .

It is easy to see that the Higgs doublet can be re-parameterized without loss of generality<sup>5</sup> as

$$\phi(x) = \frac{1}{\sqrt{2}} e^{\frac{1}{v} \sum_i \sigma_i \xi_i} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \quad (1.46)$$

and so now the massless goldstone bosons  $\xi_i$  appear now only in the exponential function. Given that the Lagrangian needs to be invariant under  $SU(2)$ , it is always possible to find a transformation  $e^{-\frac{1}{v} \sum_i \sigma_i \xi_i}$  such that the massless goldstone bosons are *gauged away*, disappearing completely from the description. The only field that remains,  $h(x)$  is the massive Higgs field, while the degrees of freedom corresponding to all  $\xi_i$  are absorbed by the gauge bosons in order to acquire mass.

### 1.4.2 Massive particles in the SM

After choosing the right vacuum for the Higgs fields and removing the non-physical degrees of freedom  $\xi_i$ , the remaining Lagrangian describes one massive scalar (Higgs) boson  $h$ , three massive spin-1 bosons and one massless spin-1 boson. The mass of the Higgs boson can be readily obtained by comparing the quadratic term in  $h$  from eq. 1.45 with a general mass term for a scalar boson,  $\frac{1}{2}m_h^2 h^2$ , which is completely determined by the vacuum expectation value of the Higgs field and the coupling  $\lambda$ , and yields

$$m_h = \sqrt{2\lambda v^2}. \quad (1.47)$$

The description of the gauge bosons in eq. 1.45, however, does not correspond directly with the physical states  $W^\pm, Z$  and  $\gamma$  since there is a mixing term combining the fields  $W_3^\mu$  and  $B^\mu$ . In order to obtain the description for the actual physical bosons, it is enough to substitute the definitions from eqs. 1.30 and 1.32 into eq. 1.45, resulting in the mass terms for the  $W^\pm$  and  $Z$  bosons and where the mass term for the photon vanishes. From here the masses of the gauge bosons can be readily identified as

$$M_W = \frac{1}{2}vg, \quad M_Z = \frac{1}{2}v\sqrt{g^2 + g'^2}. \quad (1.48)$$

Taking into account the relation between the electroweak mixing angle  $\theta_W$  and the electroweak coupling constants  $g, g'$  (see eq. 1.34), the ratio of the masses for the gauge bosons can be expressed as

$$\frac{M_W}{M_Z} = \cos \theta_W, \quad (1.49)$$

which constitutes a prediction of the SM and can be compared to the experimental results. The massless photon however was a requirement introduced in the construction of the theory and therefore constitutes only a consistency check.

The mass terms for the fermions in the model can be obtained by introducing the so called *Yukawa interactions*, that couple the fermions to the Higgs scalar field. Using the notation

---

<sup>5</sup>Performing a series expansion to the first order in all the fields it is possible to see that this parameterization leads to

$$\phi(x) \approx \frac{1}{\sqrt{2}} \begin{pmatrix} 1 + i\xi_3/v & i(\xi_1 - i\xi_2)/v \\ i(\xi_1 + i\xi_2)/v & 1 - i\xi_3/v \end{pmatrix} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \approx \frac{1}{\sqrt{2}} \begin{pmatrix} \xi_2 + i\xi_1 \\ v + h - i\xi_3 \end{pmatrix},$$

which describes four linearly independent fields and is completely equivalent to the parametrization used originally in eq. 1.44.

introduced in eq. 1.2, these new interaction can be added to the Lagrangian with the following  $SU(2) \times U(1)$  invariant terms

$$\mathcal{L}_{\text{Yukawa}} = -\hat{Y}_{ij}^l \bar{\Psi}_{L,i}^l \phi \Psi_{R,j}^l - \hat{Y}_{ij}^d \bar{\Psi}_{L,i}^d \phi \Psi_{R,j}^d - \hat{Y}_{ij}^u \bar{\Psi}_{L,i}^u \phi \Psi_{R,j}^u + \text{h.c..} \quad (1.50)$$

These terms are written in the basis of electroweak interaction eigenstates, and therefore there are mixing terms between different fermion flavours represented by the matrices  $\hat{Y}_{ij}^X$ . Rotating the system to the basis formed by the mass eigenstates from eq. 1.31 also diagonalizes the Yukawa interaction terms, which after the spontaneous symmetry breaking lead to mass terms for each of the fermions introduced in eq. 1.50. In the diagonal base, the Yukawa couplings turn into  $\hat{Y}_{ij}^X \rightarrow y_i^X \delta_{ij}$ , and the resulting mass for each fermion is  $m_i^X = y_i^X v / \sqrt{2}$ . It is important to note that the masses in the fermion sector are determined by these Yukawa couplings, which in principle are completely independent and therefore the fermion masses are free parameters in the SM that need to be determined experimentally.

## 1.5 Limitations of the Standard Model

The SM is a very successful theory that accurately describes nature at its most fundamental level. However, it is widely accepted that the model is incomplete as it has a number of shortcomings, both structural and phenomenological, that point in the direction of it being only an *effective theory* arising from a more universal description of nature. Some of the most relevant issues in the SM are summarized below.

### Disconnection from cosmology

The SM lacks a description for the gravitational interaction that rules the universe at the astronomical and cosmological scales. The irreconcilable differences between quantum field theory and general relativity make it virtually impossible to incorporate a description of gravity in the SM as we know it. In addition, the SM also lacks an explanation for the largest phenomena that drive the dynamics of galaxies and the universe itself, namely dark matter and dark energy. Dark matter, on the one hand, is a non-luminous type of matter that dominantly interacts gravitationally. It is a well accepted explanation for a variety of experimental observations at different scales, such as the unexpected rotational curves of galaxies if only baryonic matter is accounted for [12], the gravitational lensing effect produced by the bullet cluster that does not match the position of luminous matter [13], or the temperature fluctuations in the cosmic microwave background [14]. All the matter particles described by the SM add up to only 5% of the content of the universe, while dark matter constitutes roughly 27%. Several extensions of the SM have been proposed to include a description for dark matter as well, but so far no experimental evidence has been found to confirm any of them. The remaining 68% of the content of the universe is known as dark energy, an energy density with negative pressure that is responsible for the accelerated expansion of the universe. A description for dark energy arises in Einstein's equations in the context of general relativity as a constant term, also known as the *cosmological constant*. In the quantum field theory description, however, no successful approach has yet been found for this energy density and, while several models have been proposed, the predictions from the theory are many orders of magnitude off with respect to the measured values from the redshift in distant galaxies.

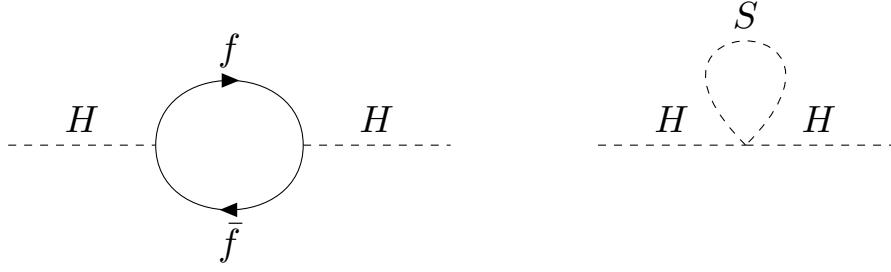


Figure 1.3: Corrections to the Higgs propagator from a new fermion  $f$  and a new scalar particle  $S$  added to the particle content of the SM.

### Hierarchy problem

One of the most outstanding unresolved issues of the SM is the so called *hierarchy problem*, the main idea being that the large differences between the Planck scale, at which quantum effects in the gravitational interaction are expected to dominate, and the electroweak breaking scale are completely unjustified in the theory. It is easy to illustrate the difference in scales by looking at the ratio between the coupling constants that characterize gravity ( $G_N$ ) and the electroweak interaction ( $G_F$ ), which yields a value of  $G_N/G_F \approx 10^{34}$ . It could be argued that this difference in the magnitude of both interactions is just the way nature turns out to be. However, the problem has the additional complexity coming from the mass of the Higgs boson being sensitive to these scale differences. Indeed, the expression for the Higgs mass in eq. 1.47 was computed only at tree-level, while a more correct approach needs to include the most relevant corrections arising from higher order terms in the perturbative expansion. At one-loop all massive particles in the SM contribute to its mass with diagrams similar to the ones shown in figure 1.3, which lead to divergent integrals. It is possible to regularize the infinities arising from these integrals by the introduction of a *cut-off* energy scale  $\Lambda_{\text{SM}}$  at which new physics are expected to appear. This however results in a quadratic dependence on the cut-off  $\Lambda_{\text{SM}}$ , explicitly

$$\Delta m_H^2 = \frac{\Lambda_{\text{SM}}}{32\pi^2} \left[ 6\lambda + \frac{1}{4} (9g^2 + 3g'^2) - y_t^2 \right] + \mathcal{O}(\ln(\Lambda_{\text{SM}})), \quad (1.51)$$

where  $\lambda$  is the self-coupling strength for the Higgs field,  $g, g'$  are the coupling strength for the electroweak gauge bosons and  $y_t$  is the Yukawa coupling for the top quark, which dominates over the rest of fermions. If no new physics appear between the electroweak and Planck scales, the cut-off then must take the value  $\Lambda_{\text{SM}} \approx M_P = (8\pi G_N)^{1/2} = 2.4 \times 10^{18}$  GeV, which leads to corrections to the Higgs mass much larger than the tree-level expected value and that results in a theoretical prediction that is hard to reconcile with the experimentally measured value of  $m(h) \approx 125$  GeV. It is possible that whatever new physics appear in the ultraviolet extreme introduce a new correction to the Higgs mass of the same magnitude but opposite sign than the one introduced by the SM physics, and so the hierarchy problem turns into one of fine tuning to exactly cancel out the two corrections.

It could also be argued that the introduction of the quadratic terms in  $\Lambda_{\text{SM}}$  depends on the method used to regularize the divergent integrals from figure 1.3, and therefore lacks of physical meaning. Even in that case, the addition of any new particle with mass  $m_{\text{BSM}}$  to the matter content of the SM leads to extra terms in  $\Delta m_H^2$  that are proportional to the mass square of said new particle,  $\Delta^{\text{BSM}} m_H^2 \sim m_{\text{BSM}}^2 \ln(\Lambda_{\text{SM}}/m_{\text{BSM}})$ . The mass  $m_{\text{BSM}}$  is ultimately an indication of the energy scale where new physics may manifest, and so the Higgs mass is still sensitive to BSM physics.

## Fermion masses

The Yukawa sector is the least satisfying part of the SM, since it cannot provide a prediction for the mass of any fermion. Each of them is an input parameter of the model that needs to be measured experimentally. Even though the SM successfully predicts that the Higgs boson couples to fermions with strength directly proportional to their mass, as shown in figure 1.6b, it fails to provide an underlying mechanism that can link the observed values without considering them as input parameters of the model. This issue in the SM is even more outstanding when considering the different ranges that the fermion masses span into, from the electron to the top quark, showing 6 orders of magnitude in difference. Furthermore, no mass term is introduced in the SM Lagrangian for neutrinos since they are assumed to be massless. However, the observation of neutrino oscillations is not compatible with massless particles and, in fact, the mass splitting between the different flavour of neutrinos has been determined experimentally. There have been several proposals as to how neutrinos can obtain a mass, such as Majorana mass terms or see-saw mechanisms involving heavier right-handed neutrinos, but so far no clear evidence has been found to favour one scenario over the rest.

## Matter-antimatter asymmetry and CP violation

The amount of matter observed in the universe dominates over its antimatter content, asymmetry for which the SM cannot fully provide an explanation. As discussed in section 1.3.2, the CKM matrix includes a complex phase which leads to CP violation in the quark sector. This complex phase, however, is not enough to explain the observed asymmetry and new mechanisms need to be included. The SM, on the other hand, allows for a CP-violating angle  $\theta_{\text{QCD}}$  in the strong interaction Lagrangian that could contribute to the matter-antimatter asymmetry without breaking the renormalizability of the theory. However, having such a term would also lead to phenomenological predictions that have not been observed empirically, such as an electric dipole moment in neutrons. Experiments have set tight upper limits on the dipole moment in neutrons,  $< 3.6 \times 10^{-26}$  e cm at 95% confidence level [15], which largely constrains the possible values for this CP-violating angle to  $\theta < 10^{10}$  rad. This poses another fine-tuning problem in the SM, since there is no reason for which  $\theta_{\text{QCD}}$  could not be  $\mathcal{O}(1)$ .

## 1.6 Experimental tests of the SM

The SM has undergone a great number of precision measurements throughout the decades, both to determine the experimental values of the free parameters in the model and to test different observables as predicted by the theory. With the arrival of the LHC the particle physics community have gained access to unprecedented levels of accuracy, and so far most of the experimental tests are in good agreement with the SM. In particular, the properties of the strong interaction that drives the proton-proton collisions have been exhaustively studied. As mentioned in section 1.3.1 the value of the QCD coupling constant  $\alpha_s$  changes with the energy scale, a thorough scan of which has been performed with results from different experiments and is summarized in figure 1.4. For the first time, thanks to the addition of data from the LHC and Tevatron, the behaviour of the QCD coupling constant has been studied at energies up to 2 TeV, and is in good agreement with the expected curve from eq. 1.19.

The cross section for different processes happening during proton-proton collisions at the LHC can be predicted from the SM with the usual perturbation approximation in quantum field

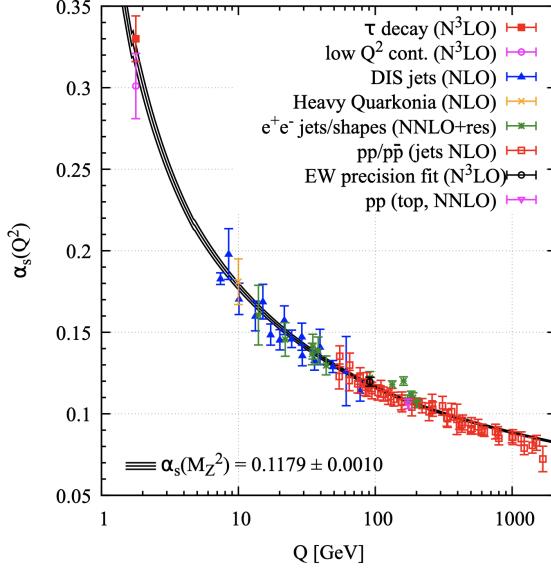


Figure 1.4: Measured QCD coupling constant  $\alpha_s$  at different energy scales by different experiments [8].

theory. These include the production of  $W^\pm$  and  $Z$  bosons, in different channels depending if they are produced in association with other particles, top quarks and Higgs bosons. During LHC Run-1 and Run-2, the ATLAS Collaboration has measured several of these cross sections in a number of analyses targeting different final states. The experimental results are compared to the SM prediction at next-to-leading order or higher in the perturbation expansion, and are found to be in excellent agreement, as seen in figure 1.5.

One of the main proofs of the predictive power of the theory was found in 2012 with the discovery of the Higgs boson. Since then, the ATLAS and CMS Collaborations at CERN have measured the properties of the SM Higgs boson to astounding precision during LHC Run-1 and Run-2 periods. Using a combination of di-photon and four-lepton analyses from both experiments, the mass of the Higgs boson was first determined with Run-1 data with excellent precision. The results of the combination can be observed in figure 1.6a, yielding a final value of

$$m_H = 125.09 \pm 0.24 \text{ GeV.} \quad (1.52)$$

In addition, the coupling of the Higgs boson to all massive particles has been studied extensively by the ATLAS Collaboration. As discussed in the previous section, a tree-level prediction from the Higgs mechanism in the SM is that the coupling strength between the Higgs boson and a massive particle is directly proportional to the particle's mass. In figure 1.6b the coupling between the Higgs boson and different massive particles as measured by the ATLAS experiment can be observed, which are all compatible with the prediction assuming the existence of a SM Higgs boson.

The Higgs sector in the SM provides other observables that can be tested experimentally in order to constrain the validity of the theory. A new observable can be constructed with the masses of the gauge bosons and the electroweak mixing angle as

$$\rho = \frac{M_W^2}{M_Z^2 \cos^2 \theta_W}, \quad (1.53)$$

which, as per eq. 1.49, is predicted at tree-level to equal exactly one in the SM. This observable

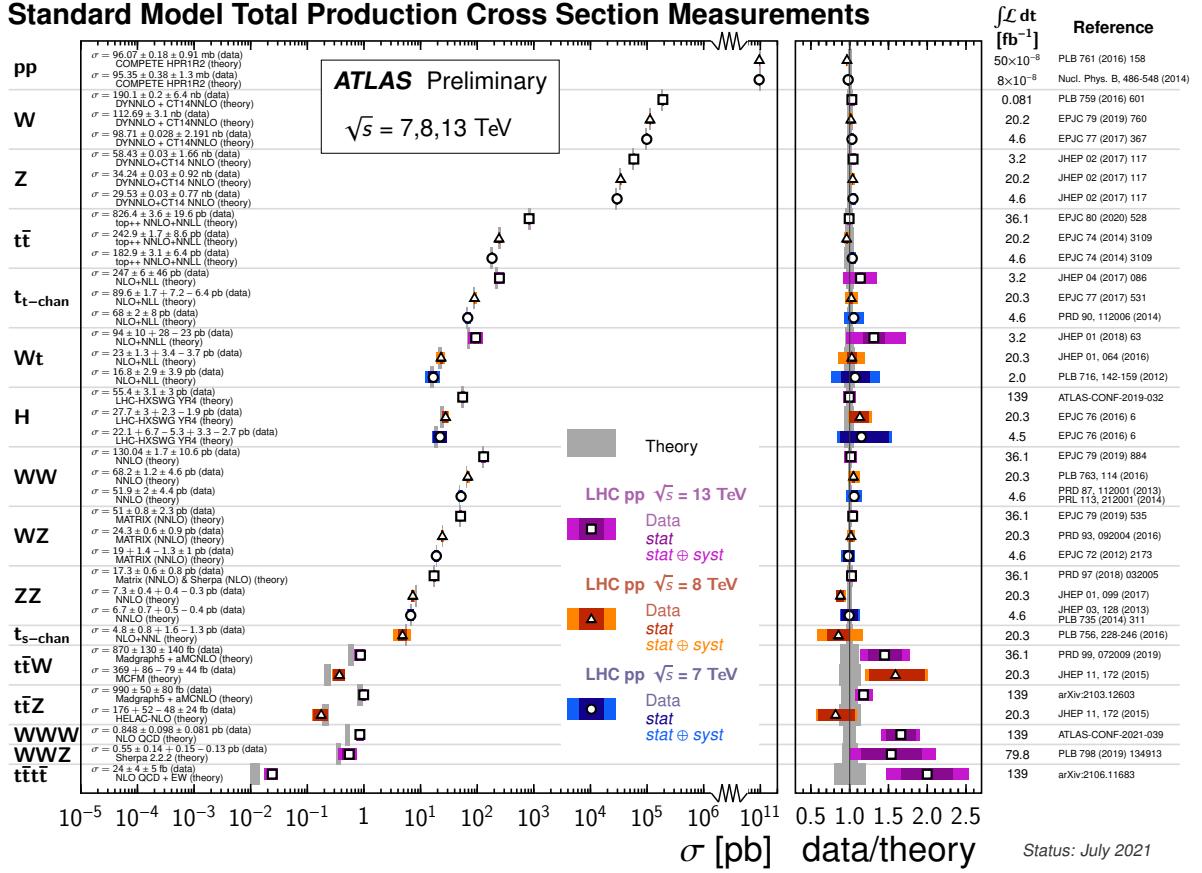


Figure 1.5: Measured production cross section for several particles and processes using LHC Run-1 and Run-2 data with the ATLAS detector [16].

is very sensitive to the structure of the Higgs fields implemented in the model, as it directly determines the relation between the  $W^\pm$  and  $Z$  boson masses, and therefore it can also be used to constrain new physics in the Higgs sector. The most recent combination of the measurements of  $M_W$ ,  $M_Z$  and the electroweak mixing angle from a number of experiments, in the modified minimum subtraction renormalization scheme, yield [8]

$$\rho = 1.00038 \pm 0.00020, \quad (1.54)$$

which is still compatible with the SM prediction.

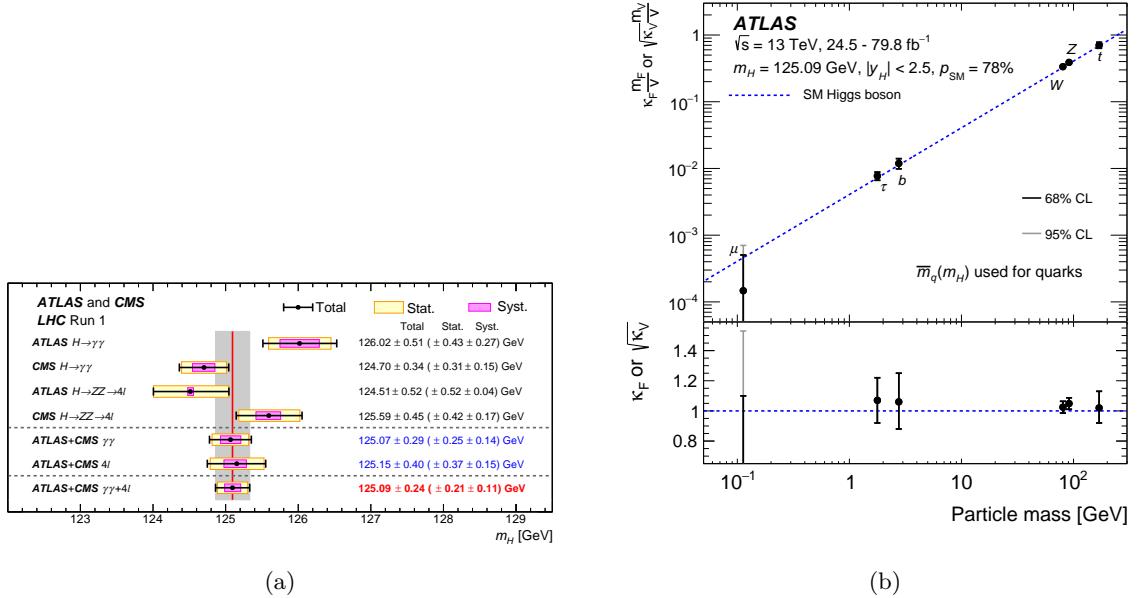


Figure 1.6: Experimental probes of the Higgs sector in the SM. (a) Combination of ATLAS and CMS measurements of the mass of the Higgs boson [17]. (b) Measured coupling of the Higgs boson to massive fermions and gauge bosons with the ATLAS detector [18].

# Chapter 2

## Supersymmetry

### 2.1 Introduction

The Standard Model of particle physics provides an unprecedented level of accuracy in the description of nature at its most fundamental level. As discussed in chapter 1, however, the model has some theoretical and phenomenological shortcomings, which makes the particle physics community consider it as a low-energy approximation of a more complete theory. Currently many different experiments are able to test the SM predictions to extremely high precision and, while the theory successfully passes most of the empirical tests, recent data seems to show some tension with the predictions from the model. For example, while in the SM all leptons couple to the electroweak gauge bosons equally (concept known as *lepton universality*), the LHCb Collaboration has observed that the branching fractions of the decay  $B^+ \rightarrow K^+ l^+ l^-$  for different leptons, generally parametrized with

$$R_K = \frac{\mathcal{B}(B^+ \rightarrow K^+ \mu^+ \mu^-)}{\mathcal{B}(B^+ \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) K^+)} \Big/ \frac{\mathcal{B}(B^+ \rightarrow K^+ e^+ e^-)}{\mathcal{B}(B^+ \rightarrow J/\psi(\rightarrow e^+ e^-) K^+)} , \quad (2.1)$$

seems to favour decays involving electrons over muons [19]. In the lepton sector, other measurements also show a certain level of disagreement with the SM, such as the muon anomalous magnetic moment  $a_\mu = (g_\mu - 2)/2$  that yields a larger value than predicted [20].<sup>1</sup>

Several extensions of the SM have been proposed throughout the decades, some proposed shortly after the SM itself [21–23], to address the deficiencies found in the description. Generally, the approach used to extend the SM could include increased matter content with heavy new particles, new symmetries that could be broken or could generate new interactions, or new degrees of freedom such as extra dimensions. Some well motivated examples are Grand Unified Theories (GUT), where the SM symmetry group  $SU(3)_C \times SU(2)_L \times U(1)_Y$  is assumed to arise after the breaking of a larger symmetry group, composite Higgs models, that state that the observed Higgs boson is a light bound state from a new strong-like interaction, or dark sectors that couple to the SM particles through new interactions. Most BSM models provide rich phenomenology that can be tested at colliders, with predictions of yet to be observed particles

<sup>1</sup>These measurements, however, cannot be claimed as *discoveries* yet since the statistical deviation from the SM, although quite large, are still  $< 5\sigma$ . The most precise measurement of  $R_K$  by the LHCb Collaboration, using  $9 \text{ fb}^{-1}$  data from LHC proton-proton collisions in the dilepton mass-squared region  $1.1 < q^2 < 6.0 \text{ GeV}^2/C^4$ , yields a statistical deviation of  $3.1\sigma$  from the SM prediction, while the  $a_\mu$  measurement by the Muon  $g - 2$  Collaboration stands out with a deviation of  $4.2\sigma$ . In the case of the muon anomalous magnetic momentum, most of the debate is about the predicted values by the SM. More data, or higher precision, is still needed to have enough sensitivity in order to reject the SM prediction.

such as leptoquarks, vector-like fermions or new generations of gauge bosons. A large component of the physics program in the ATLAS experiment consists of searches for new phenomena in LHC proton-proton collisions.

Proposed for the first time in the 1970s [21, 22, 24], one of the most elegant extensions of the SM is supersymmetry (SUSY). In essence, a SUSY transformation  $Q$  changes a fermion into a boson, and vice versa,

$$Q |\text{fermion}\rangle = |\text{boson}\rangle, \quad Q |\text{boson}\rangle = |\text{fermion}\rangle.$$

A Lagrangian that is invariant under this kind of transformation is called *supersymmetric*. For this to happen in the SM, each of the fields described in the theory must have a partner such that their spin differ in  $\frac{1}{2}$ , which leads to a number of new particles with well defined properties. In general, SUSY is not a single model but a whole theoretical framework that shows great flexibility to accommodate a large range of phenomena, and that provides solutions to some of the SM shortcomings highlighted in chapter 1.

In this chapter an overview of the theory and phenomenology of supersymmetric models is provided. It is organized as follows: in section 2.2 the general framework of SUSY is discussed, highlighting the relevant features for model building. Next, the minimal extension of the SM into a SUSY model is developed in section 2.3, and finally a discussion of the most common approaches to search for SUSY in collider experiments is provided in section 2.4.

## 2.2 General principles of Supersymmetry

In the discussion of SUSY it is convenient to use a 2-component Weyl spinors to describe fermions, instead of 4-component Dirac spinors as used in chapter 1. In this notation, any Dirac spinor can be expressed as

$$\Psi_D = \begin{pmatrix} \xi_\alpha \\ \chi^{\dagger\dot{\alpha}} \end{pmatrix}, \quad (2.2)$$

where  $\xi_\alpha, \chi^{\dagger\dot{\alpha}}$  are anti-commuting, complex Weyl spinors, and where both indices  $\alpha, \dot{\alpha}$  are independent from each other. In general, undotted indices run over the first two components of the original Dirac spinor, while dotted indices run over the last two components. In reality, this separation of the Dirac spinor is equivalent to decomposing the spin  $SU(2)$  representation into two irreducible chiral representations, and therefore  $\xi_\alpha$  and  $\chi^{\dagger\dot{\alpha}}$  live in different representations so the dotted and undotted indices are not interchangeable. In this representation, commonly known as Weyl representation<sup>2</sup> or chiral basis, the spinors  $\xi_\alpha$  and  $\chi^{\dagger\dot{\alpha}}$  are the left-handed and right-handed components of original Dirac spinor,

$$P_L \Psi_D = \frac{1}{2} (1 - \gamma_5) \Psi_D = \begin{pmatrix} \xi_\alpha \\ 0 \end{pmatrix}, \quad P_R \Psi_D = \frac{1}{2} (1 + \gamma_5) \Psi_D = \begin{pmatrix} 0 \\ \chi^{\dagger\dot{\alpha}} \end{pmatrix}, \quad (2.3)$$

where  $P_{L,R}$  are the left-hand and right-hand projection operators as defined in eq. 1.3. Hermitian conjugation transforms a left-handed Weyl spinor into a right-handed one, and vice versa, and so all fermionic degrees of freedom can be described by either left-handed spinors or right-handed ones equivalently.

---

<sup>2</sup>In this representation the Dirac  $\gamma$  matrices take the form

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix}, \quad \gamma_5 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix},$$

where  $\sigma^\mu \equiv (1, \sigma^i)$ ,  $\bar{\sigma}^\mu \equiv (1, -\sigma^i)$ , and  $\sigma^i$  are the Pauli matrices.

The group of SUSY transformations can be defined in terms of its generators  $Q$ , therefore it is important to understand their properties and the (anti-)commutation relations they follow. An operator  $Q$  that transforms an integer spin field into a spinor field must be fermionic (i.e. carry spin 1/2) and so it carries a spinorial index itself, which implies they do not commute with Lorentz transformations. For this reason, SUSY cannot be an internal symmetry but rather a space-time symmetry. The Coleman-Mandula no-go theorem [25] highly restricts all possible Lie algebras that can be included consistently in a four-dimensional theory, and so the generators of SUSY transformations, in the Weyl representation, need to satisfy the following anti-commutation relations [26]:

$$\{Q_\alpha, Q_{\dot{\alpha}}^\dagger\} = -2\sigma_{\alpha\dot{\alpha}}^\mu P_\mu, \quad (2.4a)$$

$$\{Q_\alpha, Q_\beta\} = 0, \quad \{Q_{\dot{\alpha}}^\dagger, Q_{\dot{\beta}}^\dagger\} = 0, \quad (2.4b)$$

$$[Q_\alpha, P^\mu] = 0, \quad (2.4c)$$

where  $P_\mu$  is the generator of translations from the Poincaré group. It is possible to define multiple independent copies of the supersymmetry generators  $Q_\alpha, Q_{\dot{\alpha}}^\dagger$ , in which case the commutation relations would become

$$\{Q_\alpha^i, Q_{\dot{\alpha}}^{\dagger j}\} = -2\delta^{ij}\sigma_{\alpha\dot{\alpha}}^\mu P_\mu, \quad (2.5)$$

where the superscripts  $i, j$  label the different copies of the SUSY generator, and  $\delta^{ij}$  takes the value 1 if  $i = j$ , and 0 otherwise. These “extended” SUSY paradigms are interesting in higher dimensional theories. In four dimensions, however, extended SUSY models do not generally reproduce the SM phenomenology, therefore the rest of the discussion will be focused on cases where there is only one SUSY generator.

According to Noether’s theorem, any continuous symmetry on a system implies the existence of a conserved current. In the case of SUSY, the conserved *supercurrents*  $J_\alpha^\mu, J_{\dot{\alpha}}^{\dagger\mu}$  are anti-commuting 4-vectors, carrying spinor indices [27], and that verify separately

$$\partial_\mu J_\alpha^\mu = 0, \quad \partial_\mu J_{\dot{\alpha}}^{\dagger\mu} = 0. \quad (2.6)$$

It is possible to arrive at the conserved charges integrating over spatial coordinates the temporal component of the supercurrent  $J_\alpha^\mu$  and its hermitian conjugate  $J_{\dot{\alpha}}^{\dagger\mu}$ ,

$$Q_\alpha = \sqrt{2} \int d^3 \vec{x} J_\alpha^0, \quad Q_{\dot{\alpha}}^\dagger = \sqrt{2} \int d^3 \vec{x} J_{\dot{\alpha}}^{\dagger 0}, \quad (2.7)$$

which are the generators of the SUSY transformations discussed above and, together with the anti-commutation relations from eqs. 2.4a and 2.4b, define the algebra of the group. Given that the relations between the SUSY generators include generators from the Poincaré group, the SUSY algebra is extended to include also the Poincaré algebra into a *superalgebra* with well defined properties.

In SUSY models all the matter content is mathematically described with irreducible representations of the SUSY algebra, called *supermultiplets*. Each supermultiplet must contain both fermions and bosons that are related by a SUSY transformation and that are commonly referred to as *superpartners*. Supermultiplets have the following properties:

- All the fields inside a given supermultiplet have the same mass, since the operator  $P^\mu P_\mu$  commute with the SUSY generator  $Q_\alpha, Q_{\dot{\alpha}}^\dagger$ , as per eq. 2.4c.

- All the fields share the same gauge quantum numbers, since the SUSY generators  $Q_\alpha, Q_{\dot{\alpha}}^\dagger$  commute with the gauge symmetry generators.
- In each supermultiplet there are the same number of fermionic and bosonic degrees of freedom.

There are several ways to group the matter fields together in supermultiplets respecting the requirement over the degrees of freedom. A supermultiplet including a single Weyl fermion<sup>3</sup> and two real scalars (or one complex scalar) is the simplest possibility, which is commonly known as *chiral supermultiplet*. Another possibility is a supermultiplet formed by a spin-1 boson together with a spin-1/2 Weyl fermion, which is called a *gauge supermultiplet*. For the requirement over the degrees of freedom to be realized the spin-1 boson needs to be massless (which allows only for two transversal polarization modes), which is satisfied by the SM gauge bosons prior to spontaneous symmetry breaking. In order to reproduce the SM interactions that couple differently to left-handed and right-handed fermions, the matter fields of the SM have to be necessarily written in terms of chiral supermultiplets, with the left-handed fields living in a separate supermultiplet to their right-handed equivalents. In models without extended SUSY and renormalizable interactions, any other possible combination of fermionic and bosonic fields into supermultiplets is reducible to a combination of chiral and gauge supermultiplets.

The adopted nomenclature for the particle content in SUSY extensions of the SM is to add “s-” before fermion names to label their superpartners, and so the quarks live in a chiral supermultiplets with *squarks* and leptons live together with *sleptons*. For the superpartners of the gauge bosons it is customary to add “-ino” at the end of their name, leading to *gluino*, *higgsino*, *wino* and *bino* to label the superpartners of the gluon, Higgs and  $W$  and  $B$  bosons, respectively.

### 2.2.1 Dynamics and interactions of chiral supermultiplets

The Wess-Zumino model describes a single free chiral supermultiplet, and it can be used to illustrate some properties of the Lagrangians built with chiral supermultiplets. Considering a supermultiplet with a complex scalar field  $\phi$  and a Weyl fermion  $\psi$ , the simplest possible non-interacting Lagrangian that could be written contains only the kinetic terms for both fields,

$$\mathcal{L}_{\text{chiral}} = \mathcal{L}_{\text{scalar}} + \mathcal{L}_{\text{fermion}} = -\partial^\mu \phi^* \partial_\mu \phi + i\psi^\dagger \bar{\sigma}^\mu \partial_\mu \psi, \quad (2.8)$$

from which the equations of motion for  $\phi, \psi$  are

$$\partial^\mu \partial_\mu \phi = 0, \quad \partial^\mu \partial_\mu \phi^* = 0 \quad (2.9a)$$

$$\bar{\sigma}_{\alpha\dot{\alpha}}^\mu \partial_\mu \psi^\alpha = 0, \quad \bar{\sigma}_{\alpha\dot{\alpha}}^\mu \partial_\mu \psi^{\dagger\dot{\alpha}} = 0. \quad (2.9b)$$

Since  $\psi$  and  $\phi$  live in the same supermultiplet, a SUSY transformation applied on  $\phi$  must result in a function of  $\psi$ , the easiest possibility being

$$\delta\phi = \varepsilon\psi, \quad \delta\phi^* = \varepsilon^\dagger\psi^\dagger, \quad (2.10)$$

where  $\delta$  indicates the SUSY transformation and  $\varepsilon^\alpha$  parametrizes the transformation. Since the variation connects a bosonic field with a fermionic one, clearly the object  $\varepsilon$  must be fermionic

---

<sup>3</sup>Since the fermions in the theory still need to satisfy the Dirac equation, having a single Weyl fermion with two degrees of freedom is equivalent to a Majorana fermion in the Dirac representation, with the phenomenological implication that said fermion is its own anti-particle.

as well, therefore it is described by an anti-commuting, 2-component Weyl spinor. With this definition of the action of SUSY over the scalar field of eq. 2.10, the way that the fermionic field  $\psi$  transforms can be found by imposing that the action of the system  $S = \int d^4x \mathcal{L}_{\text{chiral}}$  remains invariant, i.e.  $\delta S = 0$ , which leads to

$$\delta\psi_\alpha = -i(\sigma^\mu\varepsilon^\dagger)_\alpha\partial_\mu\phi, \quad \delta\psi_{\dot{\alpha}}^\dagger = i(\varepsilon\sigma^\mu)_{\dot{\alpha}}\partial_\mu\phi^*. \quad (2.11)$$

With the previous definitions in eqs. 2.10 and 2.11 the Lagrangian is supersymmetric by construction, however this does not grant that the algebra generated by the conserved charges satisfy the anti-commutation relations from eq. 2.4a. For any field  $X = \phi, \psi$  in the supermultiplet, since the generators  $Q, Q^\dagger$  are quantum mechanical operators they are related to the transformations given by eqs. 2.10 and 2.11 through

$$[\varepsilon Q + \varepsilon^\dagger Q^\dagger, X] = -i\sqrt{2}\delta X, \quad (2.12)$$

and therefore it is possible to see if the SUSY transformations derived above close the SUSY algebra. By examining the commutator of two consecutive SUSY transformations,  $[\delta_1, \delta_2]X \equiv \delta_1(\delta_2 X) - \delta_2(\delta_1 X)$ , it is enough to verify that<sup>4</sup>

$$[\delta_1, \delta_2]X = i(\varepsilon_1\sigma^\mu\varepsilon_2^\dagger - \varepsilon_2\sigma^\mu\varepsilon_1^\dagger)\partial_\mu X. \quad (2.13)$$

For the scalar field  $\phi$  this condition is always fulfilled. For the fermion  $\psi$ , however, the commutator  $[\delta_1, \delta_2]X$  yields

$$[\delta_1, \delta_2]\psi_\alpha = i(\varepsilon_1\sigma^\mu\varepsilon_2^\dagger - \varepsilon_2\sigma^\mu\varepsilon_1^\dagger)\partial_\mu\psi_\alpha + i(\varepsilon_{2\alpha}\varepsilon_{1\dot{\beta}}^\dagger - \varepsilon_{1\alpha}\varepsilon_{2\dot{\beta}}^\dagger)\bar{\sigma}^{\mu,\dot{\beta}\beta}\partial_\mu\psi_\beta. \quad (2.14)$$

The first term in the right-hand side of eq. 2.14 is the sought term to satisfy the condition from eq. 2.13, although unwanted extra terms appear. Observing that the second term in the right-hand side of eq. 2.14 vanishes when the equations of motion of the fermionic field given by eq. 2.9b are verified, one concludes that the SUSY algebra only closes if the fermion field is *on-shell*, which is still unsatisfactory since it would imply that the SUSY transformations do not have well defined properties at a quantum level. This is due to the number of degrees of freedom associated to  $\phi$  and  $\psi$ : while a complex scalar field always has two degrees of freedom, a complex Weyl fermion initially has four degrees of freedom that are reduced to two when it is on-shell.

In order to have the SUSY algebra closing both on-shell and off-shell an auxiliary complex scalar field  $F$  can be introduced in the description, with Lagrangian

$$\mathcal{L}_{\text{aux}} = F^*F, \quad (2.15)$$

that describe no dynamics (eq. of motions  $F = F^* = 0$ ), and that under a SUSY transformation behaves as

$$\delta F = -i\varepsilon^\dagger\bar{\sigma}^\mu\partial_\mu\psi, \quad \delta F^* = i\partial_\mu\psi^\dagger\bar{\sigma}^\mu\varepsilon, \quad (2.16)$$

---

<sup>4</sup>Taking into account the expanded commutator

$$[\varepsilon_1^\alpha Q_\alpha + \varepsilon_1^{\dot{\alpha}} Q_{\dot{\alpha}}^\dagger, \varepsilon_2^\alpha Q_\alpha + \varepsilon_2^{\dot{\alpha}} Q_{\dot{\alpha}}^\dagger] = \varepsilon_1^\alpha\{Q_\alpha, Q_\beta\}\varepsilon_2^\beta + \varepsilon_1^\alpha\{Q_\alpha, Q_{\dot{\alpha}}^\dagger\}\varepsilon_2^{\dot{\alpha}} + \varepsilon_1^{\dot{\alpha}}\{Q_{\dot{\alpha}}^\dagger, Q_\alpha\}\varepsilon_2^\alpha + \varepsilon_1^{\dot{\alpha}}\{Q_{\dot{\alpha}}^\dagger, Q_{\dot{\beta}}^\dagger\}\varepsilon_2^{\dot{\beta}},$$

and applying the anti-commutation relations from eqs. 2.4a and 2.4b, and taking into account that the momentum operator generates the space-time translations,  $[P_\mu, X] = i\partial_\mu X$ , it is immediate to arrive to the condition specified above.

which vanish on-shell. If, in addition, the transformation rule for the fermionic field in eq. 2.11 is modified to become

$$\delta\psi_\alpha = -i(\sigma^\mu\varepsilon^\dagger)_\alpha\partial_\mu\phi + \varepsilon_\alpha F, \quad \delta\psi_{\dot{\alpha}}^\dagger = i(\varepsilon\sigma^\mu)_{\dot{\alpha}}\partial_\mu\phi^* + \varepsilon_{\dot{\alpha}}^\dagger F^*, \quad (2.17)$$

then the unwanted terms in eq. 2.14 are cancelled out with those coming from  $\delta F$  (up to a total derivative, which still leaves the action of the system invariant), and now the SUSY algebra closes both on-shell and off-shell. These auxiliary fields  $F$ , even though they do not have an impact in the dynamics of the system, bring two new bosonic degrees of freedom into the chiral supermultiplet that compensate with the extra two fermionic degrees of freedom when the system is off-shell. When the system becomes on-shell, the auxiliary field vanishes and the remaining two bosonic degrees of freedom from  $\phi$  compensate the two on-shell fermionic degrees of freedom from  $\psi$ . Nevertheless, the auxiliary fields  $F$  have other important implications in a SUSY model as they play a key role in the mechanisms for SUSY breaking, as will be discussed in section 2.2.3.

The framework to include non-gauge interactions for chiral supermultiplets can be studied using as example a collection of  $N$  copies of such supermultiplets. The gauge interactions will be addressed in section 2.2.2. For such system, all possible interacting terms to be added to the Lagrangian are limited by the closure of the SUSY algebra and the renormalizability of the theory. The interaction terms can be written with the help of the *superpotential*, a holomorphic function of the scalar fields in the supermultiplet (but not their complex conjugates). The most general superpotential that can be written while respecting all requirements on Lagrangian invariance is

$$W(\vec{\phi}) = \frac{1}{2}M^{ij}\phi_i\phi_j - \frac{1}{6}y^{ijk}\phi_i\phi_j\phi_k, \quad (2.18)$$

where  $M^{ij}$  is a complex symmetric matrix that parametrizes the mass of the fermions and  $y^{ijk}$  is a Yukawa coupling that is totally symmetric under the exchange of  $i, j, k$ . The complete Lagrangian, including all auxiliary fields  $F^i$  and  $F^{*i}$ , is

$$\mathcal{L}_{\text{chiral}} = -\partial^\mu\phi^{*i}\partial_\mu\phi_i + i\psi^{\dagger i}\bar{\sigma}^\mu\partial_\mu\psi_i + F^{*i}F_i - \frac{1}{2}W^{ij}\psi_i\psi_j + W^iF_i + \text{complex conj.}, \quad (2.19)$$

where the index  $i = 1, \dots, N$  labels the copy of the supermultiplet, and an implicit sum over repeated indices is assumed, and where the objects  $W^i, W^{ij}$  can be written as

$$W^i = \frac{\partial W}{\partial\phi_i} = M^{ij}\phi_j + \frac{1}{2}y^{ijk}\phi_j\phi_k, \quad (2.20a)$$

$$W^{ij} = \frac{\partial W}{\partial\phi_i\partial\phi_j} = M^{ij} + y^{ijk}\phi_k. \quad (2.20b)$$

The equations of motion for the auxiliary field in this new Lagrangian from eq. 2.19 are now  $F^i = -W^{*i}$ , and the dynamics of the system can be written mathematically without explicit dependence on  $F^i, F^{*i}$ . In this form, and making the substitutions from eqs. 2.20a and 2.20b, it is easy to identify mass terms for both the fermion and the scalar fields, together with Yukawa interaction terms connecting one scalar field  $\phi_i$  to two fermions  $\psi_j, \psi_k$  with coupling strength  $y^{ijk}$ , plus cubic and quartic interaction terms in the scalar fields. Deriving the equations of motion for scalars and fermions it is also possible to see that for fields inside the same supermultiplet  $i$ , both fermion and scalar fields have the same mass.<sup>5</sup>

---

<sup>5</sup>The equations of motion for scalars and fermions are

$$\partial^\mu\partial_\mu\phi_i = M_{ik}^*M^{kj}\phi_j + \dots, \quad \partial^\mu\partial_\mu\psi_i = M_{ik}^*M^{kj}\psi_j + \dots$$

The superpotential  $W(\vec{\phi})$  therefore defines all possible non-gauge interactions to be added to the chiral supermultiplet Lagrangian, and it is of great relevance when defining extensions of the SM. As will be mentioned in section 2.3 when discussing the minimal possible supersymmetric extension of the SM, the matter content and the superpotential completely define the model, with the exception of explicit SUSY breaking terms.

## 2.2.2 Gauge supermultiplets and gauge interactions

In the study of the dynamics of the gauge supermultiplets it is easy to arrive to similar conclusions regarding the need for auxiliary fields to ensure the SUSY algebra closes both on-shell and off-shell. A supermultiplet containing a massless gauge boson  $A_\mu$  and the corresponding gaugino  $\lambda$  has two bosonic and two fermionic on-shell degrees of freedom, each one balancing out the other. When the system is off-shell, however, the gaugino corresponds to four degrees of freedom while the gauge boson corresponds to only three degrees of freedom, one of the original four being fixed by gauge invariance. Similar to the procedure developed for chiral supermultiplets, a real scalar auxiliary field  $D$  can be included in the description so that the SUSY anti-commutation relations from eqs. 2.4a and 2.4b are verified. The Lagrangian for a free system containing only a gauge supermultiplet therefore is

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4}F_{\mu\nu}^a F^{\mu\nu a} + i\lambda^{\dagger a}\bar{\sigma}^\mu \nabla_\mu \lambda^a + \frac{1}{2}D^a D^a, \quad (2.21)$$

where the index  $a$  runs over the number of generators in the gauge symmetry group that result in the gauge boson(s)  $A_\mu^a$  (e.g. eight bosons for  $SU(3)$  or three for  $SU(2)$ ), and an implicit sum over repeated indices is assumed. The term  $-\frac{1}{4}F_{\mu\nu}^a F^{\mu\nu a}$  represents the kinetic energy of the vector fields  $A_\mu^a$ , with the field strength tensor  $F_{\mu\nu}^a$  defined as in eq. 1.17. The kinetic term for the gaugino,  $i\lambda^{\dagger a}\bar{\sigma}^\mu \nabla_\mu \lambda^a$ , is identical to the one introduced for the chiral supermultiplet in the previous section, with the difference that the derivative has now been substituted by the covariant derivative  $\nabla_\mu$  defined as

$$\nabla_\mu \lambda^a = \partial_\mu \lambda^a + g f^{abc} A_\mu^b \lambda^c, \quad (2.22)$$

where  $f^{abc}$ , as in chapter 1, are the symmetry group structure constant.

Following a similar procedure discussed in section 2.2.1, it is straightforward to arrive to the transformation laws for the fields in the Lagrangian given by eq. 2.21:

$$\delta A_\mu^a = -\frac{1}{\sqrt{2}} \left( \varepsilon^\dagger \bar{\sigma}_\mu \lambda^a + \lambda^{\dagger a} \bar{\sigma}_\mu \varepsilon \right), \quad (2.23a)$$

$$\delta \lambda_\alpha^a = \frac{i}{2\sqrt{2}} (\sigma^\mu \bar{\sigma}^\nu \varepsilon)_\alpha F_{\mu\nu}^a + \frac{1}{\sqrt{2}} \varepsilon_\alpha D^a, \quad (2.23b)$$

$$\delta D^a = \frac{i}{\sqrt{2}} \left( -\varepsilon^\dagger \bar{\sigma}^\mu \nabla_\mu \lambda^a + \nabla_\mu \lambda^{\dagger a} \bar{\sigma}^\mu \varepsilon \right), \quad (2.23c)$$

which, thanks to the addition of the real scalar field  $D^a$ , verify the conditions to close the SUSY algebra both on-shell and off-shell. Since the covariant derivative introduces a coupling between the gauge fields  $A_\mu^a$  and the gauginos  $\lambda^a$ , and given that the field strength tensor  $F_{\mu\nu}^a$  is quadratic in the gauge fields, the SUSY transformations above are not linear in the fields for non-abelian

---

Both bosons and fermions share the same squared-mass matrix  $(M^2)_i^j = M_{ik}^* M^{kj}$ , which can be diagonalized by redefining the fields  $\phi_i, \psi_i$ , leading to a basis formed by supermultiplets for which the scalar field  $\phi'_i$  and the Weyl fermion  $\psi'_i$  have the same squared-mass parameter  $M'^2_i = M'^*_i M'^{ki}$ .

symmetry groups. It is possible to add new auxiliary fields in addition to  $D^a$  in order to make these SUSY transformation laws linear in the fields, similar to the ones derived for the chiral supermultiplet, although in this discussion the Wess-Zumino gauge [28] will be used, in which those additional auxiliary fields are eliminated and only  $D^a$  remains.

Interactions in the SM arise from the coupling between gauge bosons, that in SUSY models must live inside a gauge supermultiplet, and quarks and leptons, which live in chiral supermultiplets. It is therefore interesting to understand how these two sets of supermultiplets can couple while still leaving the Lagrangian of the system SUSY invariant. Firstly, gauge interactions arise through (local) gauge invariance in the Lagrangian. In order to have a chiral supermultiplet including gauge-mediated interactions, it is necessary to ensure that the Lagrangian  $\mathcal{L}_{\text{chiral}}$  of the system, eq. 2.19, remains invariant under the gauge group, which can be achieved by replacing the partial derivatives  $\partial_\mu$  for the covariant derivative<sup>6</sup>  $\nabla_\mu$  as introduced in section 1.3:

$$\partial_\mu X_i \rightarrow \nabla_\mu X_i = \partial_\mu X_i - i g_a A_\mu^a (T^a)_i^j X_j, \quad X_i = \phi_i, \psi_i, F_i \quad (2.24)$$

where  $g_a$  is the coupling strength for a certain gauge group,  $T^a$  are its generators and a sum over all of the gauge groups (i.e. over repeated index  $a$ ) is implied. Similarly to what happened in the development of QCD and electroweak interactions, through this requirement a new coupling between chiral supermultiplets and the gauge bosons have appeared. In addition, covariant derivatives need to be included in the SUSY transformation laws for the chiral supermultiplet derived in the previous section to ensure simultaneous SUSY and (local) gauge invariance of the Lagrangian.

In replacing the space-time derivatives by covariant derivatives it is straightforward to see that the gauge vector field  $A_\mu^a$  couples to both the scalar  $\phi_i$  and the Weyl fermion  $\psi_i$  in the chiral supermultiplet. However there are still possible interaction terms that can be added to the description without spoiling any of the requirements imposed so far. In particular, terms coupling the gaugino to the fields  $\phi_i$  and  $\psi_i$  from the chiral supermultiplets can be added, which can be interpreted as the SUSY symmetrization of the gauge boson-matter fields couplings. Combining the Lagrangians for the chiral and gauge supermultiplets, from eqs. 2.19 and 2.21 respectively, the most generic Lagrangian that still respects gauge and SUSY invariance is

$$\mathcal{L}_{\text{SUSY}} = \mathcal{L}_{\text{chiral}} + \mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{int.}}$$

$$\begin{aligned}
&= -\nabla_\mu \phi^{*i} \nabla_\mu \phi_i + i \psi^{\dagger i} \bar{\sigma}^\mu \nabla_\mu \psi_i \\
&\quad - \frac{1}{4} F_{\mu\nu}^a F^{\mu\nu a} + i \lambda^a \bar{\sigma}^\mu \nabla_\mu \lambda^a \\
&\quad - \frac{1}{2} W^{ij} \psi_i \psi_j - \sqrt{2} \sum_a g^a (\phi^* T^a \psi) \lambda^a \\
&\quad - V(\vec{\phi}, \vec{\phi}^*),
\end{aligned}
\tag{2.25}$$

<sup>6</sup>This is identical to the covariant derivative introduced in the previous chapter in the discussion of local gauge symmetries as the origin of the SM interactions. The notation has been changed with respect to that discussion ( $D_\mu \rightarrow \nabla_\mu$ ) to avoid confusion with the  $D^a$  fields introduced in the current section.

where the scalar potential  $V(\vec{\phi}, \vec{\phi}^*)$  collects the terms involving the auxiliary  $F$  and  $D$  fields. In each line the interactions generated by the Lagrangian terms are represented by means of Feynman diagrams, where the superpartners particles have been drawn in red to establish a difference from the fields that appear normally in the SM (the scalars  $\phi_i$  as superpartners of leptons/quarks  $\psi_i$  and the gauginos  $\lambda^a$  as superpartners of the gauge bosons). The second line in eq. 2.25 corresponds to the kinetic terms for the fields in the chiral supermultiplet, which leads to three- and four-vertex interactions between the gauge bosons and the scalars  $\phi_i$ , together with the gauge coupling for the fermions  $\psi_i$  that is identical to the gauge interaction terms in the SM. The third line in eq. 2.25, on the other hand, corresponds to the kinetic terms for the fields inside the gauge supermultiplet, which leads to cubic and quartic interactions between the gauge bosons, as seen in the SM, but also to couplings between the gauge bosons and their corresponding superpartners, the gauginos  $\lambda^a$ . Finally, the fourth line in eq. 2.25 collects the other possible interaction terms that could be added to the Lagrangian without spoiling any symmetry, which lead to Yukawa couplings between scalars and fermions from the chiral supermultiplets, in addition to mass terms for the fermions  $\psi_i$ , and couplings between scalars and fermions from the chiral supermultiplets to gauginos.

As mentioned above, the scalar potential  $V(\vec{\phi}, \vec{\phi}^*)$  can be defined by the terms quadratic in the auxiliary fields  $F, D$  that were introduced to ensure the SUSY algebra closing off-shell. It plays an essential role in SUSY spontaneous breaking scenarios, as will be discussed in section 2.2.3. The scalar potential also introduces new interactions in the description, in addition to new mass terms. Explicitly written, it yields

$$\begin{aligned} V(\vec{\phi}, \vec{\phi}^*) &= F^{*i} F_i + \sum_a \frac{1}{2} D^a D^a \\ &= \left( M_{ij}^* \phi^{*j} + \frac{1}{2} y_{ijk}^* \phi^{*j} \phi^{*k} \right) \left( M^{il} \phi_l + \frac{1}{2} y^{ilm} \phi_l \phi_m \right) + \sum_a g_a^2 (\phi^* + T^a \phi)^2, \end{aligned} \quad (2.26)$$

where the equations of motion for both  $F$  and  $D$  have been used to write the scalar potential in terms of the  $\phi_i$  fields.<sup>7</sup> The first term in the right-hand side of the first line in eq. 2.26 is commonly referred to as “F-term”, while the second one is referred to as “D-term”. As can be seen from the extended potential in the second line of eq. 2.26 mass terms (i.e. proportional to  $\phi_i \phi_i^*$ ) for the scalars are raised together with additional cubic and quartic interactions between the scalars, that need to be added to the interactions arising in the Lagrangian from eq. 2.25.

As can be seen from eqs. 2.25 and 2.26, the addition of SUSY to a theory with local gauge invariance as origin for interactions is in fact an extension of such theory, since all of the original interactions can be retrieved while new couplings appear, enriching the matter content of the model and modifying its phenomenological description.

### 2.2.3 Supersymmetry breaking

As discussed in the previous section, any supersymmetric model comes with mass-degenerate sets of bosons and fermions. Given that none of the superpartners of the SM particles have been observed so far, if SUSY were to exist in nature it would necessarily be a broken symmetry of the Lagrangian. The mechanism that breaks SUSY in nature is unknown although it is

<sup>7</sup>The equations of motion for the  $F$  field can be obtained directly from the Lagrangian for the chiral supermultiplets and are not affected by the coupling to the gauge supermultiplets. The equations of motion for the  $D$  field are obtained after the coupling between chiral and gauge supermultiplets, as new interaction terms proportional to  $(\phi^* T^a \phi) D^a$  can be introduced which, together with the rest of the Lagrangian, lead to  $D^a = g_a (\phi^* T^a \phi)$  as equation of motion.

generally accepted that it must happen through spontaneous symmetry breaking, similarly to what happens to the electroweak gauge group. The SUSY breaking energy scale highly depends on the breaking mechanism, ranging from somewhere near the electroweak breaking scale all the way to the Planck scale.

If SUSY is broken, according to the Nambu-Goldstone theorem, there should be a massless particle associated to the broken generator, with its same quantum numbers, that arises in the symmetry breaking. Since the broken generators in this case would be the conserved charges  $Q_\alpha, Q_{\dot{\alpha}}^\dagger$ , which are fermionic, the expected Nambu-Goldstone particle should be a massless Weyl fermion. It is commonly referred to as *goldstino*, labelled by  $\tilde{G}$ , and it can be understood as the superpartner of the field that acquires a VEV. It is expected that  $\tilde{G}$  couples to all the boson-fermion pairs in the model. If SUSY were to be promoted from global to local symmetry,<sup>8</sup> the massless spin-2 graviton that mediates the gravitational interactions should be included in the particle content of the model, and it would necessarily need to live in a gauge supermultiplet together with a massless spin-3/2 *gravitino*. Similarly to what happens in the electroweak sector in the SM, the *super-Higgs mechanism* [29] allows the gravitino to get a mass term after the symmetry breaking by “eating up” the goldstino. In this case, the goldstino becomes the longitudinal degrees of freedom of the gravitino, for which the mass generated depends heavily on the SUSY breaking mechanism and it could range from just a few MeV all the way to the Planck scale. The gravitino inherits the non-gravitational interactions of the goldstino, and therefore the longitudinal components of the gravitino (helicity  $\pm \frac{1}{2}$ ) could play an important role in the particle physics phenomenology explored in colliders, according to the goldstino-gravitino equivalence theorem [30]. The transversal gravitino components (helicity  $\pm \frac{3}{2}$ ) can therefore be ignored, and so generally the gravitino is also labeled as  $\tilde{G}$  in experimental searches.

Spontaneous SUSY breaking happens when the vacuum  $|0\rangle$  is not invariant under the generators  $Q, Q^\dagger, Q|0\rangle \neq 0$ , which implies that the vacuum expectation value for the SUSY Hamiltonian is non-zero, i.e.  $\langle 0 | H | 0 \rangle \neq 0$ . Generally this corresponds to a non-zero vacuum expectation value for the scalar potential  $V(\vec{\phi}, \vec{\phi}^*)$  introduced in eq. 2.26, and so SUSY is spontaneously broken if the VEV for both the  $F_i$  and  $D^a$  auxiliary fields do not vanish simultaneously. If SUSY is a global symmetry, the scalar potential is always non-negative, and while it is possible that its global minimum is non-zero and corresponds to the vacuum in which we live, it is also possible that we live in a local minimum of the scalar potential. This kind of meta-stable scenarios is compatible with the global minimum preserving SUSY (i.e.  $\langle 0_{\text{global}} | V | 0_{\text{global}} \rangle = 0$ ) but then it is required that the SUSY breaking local minimum in which we live has a lifetime at least of the order of the universe, so that in the actual moment SUSY is still broken [31–33].

It is possible to produce spontaneous SUSY breaking separately through the F-term or the D-term in the scalar potential. Models including SUSY breaking due to the F-term use the O’Raifeartaigh mechanism [34]. It relies on choosing the superpotential  $W(\vec{\phi})$  such that it is not possible to satisfy simultaneously the equations  $F_i = -\partial W^*/\partial\phi^{i*} = 0$ , leading to  $V = \sum_i |F_i|^2 > 0$  at the minimum. This mechanism relies on the matter content of the model including a gauge singlet so that linear terms in the superpotential can be included. The linear terms then raise a non-vanishing contribution to the scalar potential once the fields acquire a VEV, thus breaking the SUSY invariance of the Lagrangian. Similarly, spontaneous SUSY breaking can happen through the D-term in the scalar potential, an example of which is the Fayet-Iliopoulos mechanism [35]. It relies on the addition of linear terms in the  $D$  fields, which after them obtaining a VEV, lead to positive values for the scalar potential at its minimum. In both cases of SUSY breaking, for the needed linear terms to be added in the Lagrangian, it

---

<sup>8</sup>When SUSY is made a local transformation (i.e. the transformation parameter  $\varepsilon$  changes from space-time point to point), gravitational interactions arise in a natural way, and the resulting theory is called *supergravity*.

needs to have an exact  $U(1)$  symmetry,<sup>9</sup> otherwise these terms would not be allowed by gauge invariance requirements.

The previous conditions for spontaneous SUSY breaking lead to phenomenological predictions that are hard to reconcile with the minimal SUSY extension of the SM, which is discussed in section 2.3. For this reason it is often considered the existence of extra *hidden or dark sectors*, mostly decoupled from the SM visible matter content and its interactions, where SUSY is spontaneously broken. In this way it is possible to introduce one of the mechanisms mentioned above in the hidden sector without needing to impose any requirements on the SM supermultiplets. The effects of SUSY breaking are then transmitted to the visible sector by some mechanism, usually involving an extra set of *messenger* particles, in a way that they manifest as explicit SUSY breaking terms in the Lagrangian. This type of approach to SUSY breaking can successfully describe models with renormalizable interactions and still reproduce phenomenologically acceptable extensions of the SM.

The way that SUSY breaking is transmitted from the hidden to the visible sector depends on the coupling between the two, which also conditions the phenomenological implications of the model. There is a number of proposed mechanisms through which the SUSY breaking effects are transmitted to the visible sector. Scenarios where the SM gauge interactions mediate between the hidden and visible sectors, usually referred to *gauge-mediated SUSY breaking* (GMSB) models [37–39], are of particular interest. The general approach is to include chiral supermultiplets of messenger particles that couple directly to the matter content in the hidden sector and indirectly to the visible chiral supermultiplets through the SM gauge interactions. SUSY breaking effects therefore arise in the visible sector through radiative corrections to the masses for all its particles, except for the gauge bosons as they are protected by gauge invariance. An attractive feature of this mechanism is that flavour-changing effects introduced by SUSY breaking are automatically suppressed due to gauge interactions being flavour blind, without needing to fine tune specific couplings on account of experimental constraints. In addition, in this kind of models the goldstino/gravitino is typically very light, with its mass ranging from a few eV to the GeV scale depending on the SUSY breaking scale and the mass scale of the messenger particles, and constitutes a feasible dark matter candidate. Besides, the goldstino/gravitino couples to the particles in the visible sector with strength large enough to be phenomenologically relevant in collider physics.

Supergravity models, on the other hand, provide a natural framework for SUSY breaking effects to be transmitted from the hidden to the visible sector, where gravity acts as the messenger of SUSY breaking. This kind of scenarios are commonly known as *gravity-mediated* or *Planck-mediated SUSY breaking* (PMSB) scenarios, and typically the SUSY breaking scale needs to be close to the Planck scale for the effects on the visible sector not to be negligible.<sup>10</sup> In these models the gravitino mass is at least of the order of the electroweak scale, but it can be much larger, in contrast with GMSB models. Its interactions, however, are purely gravitational and so it plays no role in collider physics.

Given that SUSY has not yet been observed, it is difficult to understand, if SUSY were to exist, which symmetry breaking mechanism would be preferred by nature. A way to parametrize SUSY breaking is to add terms in the Lagrangian of the visible sector that break SUSY explicitly.

---

<sup>9</sup>In the case of the Fayet-Iliopoulos mechanism this requirement is by construction. For the O’Raifeartaigh models, the Nelson-Seiberg theorem [36] states that in scalar potentials with SUSY breaking global minimum through F-terms there needs to be an exact  $U(1)$  symmetry in the theory.

<sup>10</sup>If spontaneous SUSY breaking happens in the hidden sector through the O’Raifeartaigh mechanism, the mass scale of SUSY breaking terms in the Lagrangian of the visible sectors is of the order of  $m \sim \langle F \rangle / \Lambda_{\text{Planck}}$ , where  $\langle F \rangle$  is the VEV of the auxiliary F-field.

One of the most attractive features of SUSY models is that they solve the hierarchy problem of the SM by construction, as the quadratic divergences in radiative corrections to the Higgs mass from each particle are always cancelled by their superpartners. When SUSY is broken, however, particles living in the same supermultiplet will not be mass degenerate any longer, and so the Higgs boson is at risk of receiving once again large radiative corrections from heavy particles. It is possible to restrict the SUSY breaking terms added to the Lagrangian in a way that the loop corrections affecting the Higgs mass coming from the SM fermions still cancel out with that of their superpartners. These terms, commonly referred to as *soft* SUSY breaking contributions, need to include only couplings between fields with positive mass dimension and mass terms in order to fulfil the previous requirement. The most general soft breaking Lagrangian that can be written is

$$\mathcal{L}_{\text{soft}} = - \left( \frac{1}{2} M_a \lambda^a \lambda^a + \frac{1}{6} a^{ijk} \phi_i \phi_j \phi_k + \frac{1}{2} b^{ij} \phi_i \phi_j + t^i \phi_i \right) + \text{complex conj.} - (m^2)_j^i \phi^{j*} \phi_i, \quad (2.27)$$

with new couplings  $M_a, a^{ijk}, b^{ij}, t^i$  and  $(m^2)_j^i$ . This is the most general, phenomenologically relevant addition to the Lagrangian so that quadratic divergences in the correction to scalar masses are exactly cancelled at all orders in the perturbative expansion [26]. In addition,  $\mathcal{L}_{\text{soft}}$  completely describes all possible SUSY breaking effects transmitted from the hidden sector where SUSY is spontaneously broken. The mass term  $M_a$  for the gauginos  $\lambda^a$  is always allowed by gauge invariance, and so are the mass terms  $(m^2)_j^i$  for the scalar fields  $\phi_i$ . Therefore the gauginos and scalars in the model can always have a mass term regardless of their superpartners, which in the SM are required to be massless by gauge invariance (i.e. all SM fermions and bosons). The rest of the terms are allowed by gauge invariance depending on the structure of the couplings  $a^{ijk}, b^{ij}$  and  $t^i$ , which are similar to the terms introduced in the superpotential from eq. 2.18, with the exception of the linear terms in the scalar fields  $\phi_i$ . If such a superpotential is allowed by gauge invariance, then the soft breaking terms will be allowed as well. This kind of parameterization allows to write SUSY models regardless of the breaking mechanism that may be acting, and is interesting in order to understand the complete phenomenology of the theory and potential signatures that could arise in experimental searches for SUSY.

## 2.3 Minimal Supersymmetric Standard Model

The minimal supersymmetric Standard Model (MSSM) [40] is an extension of the SM with the minimal matter content so that SUSY is conserved, before the addition of SUSY breaking terms. In order to achieve both requirements, all fermion and gauge fields in the SM need to live in chiral and gauge supermultiplets, respectively. Since the left-handed and right-handed components of the SM fermions couple differently to the gauge bosons they forcefully need to live in separated supermultiplets. As discussed above, in each supermultiplet there need to be also the corresponding superpartners to the original SM particles, all sharing the same quantum numbers. In the case of the quarks and leptons, their superpartners are scalar fields, while for the gauge bosons the superpartners are spin-1/2 gauginos. In total, the MSSM include three chiral supermultiplets to accommodate right-handed leptons ( $e_R, \mu_R, \tau_R$ ) together with the “right-handed” sleptons<sup>11</sup> ( $\tilde{e}_R, \tilde{\mu}_R, \tilde{\tau}_R$ ) and six chiral supermultiplets for the right-handed quarks  $q_R$  and the corresponding squark fields  $\tilde{q}_R$  (three for up-type quarks and three for the down-type ones). In the SM the left-handed fermions live in  $SU(2)$  doublets, and so the corresponding

---

<sup>11</sup>It is common to label the scalars in the MSSM using the chirality eigenstates, often indicated by a subscript “L” or “R”, to indicate that they are the superpartners of left- or right-handed SM fermions. However, since they are scalars, it does not make sense to talk about their own chirality.

Names		Spin 0	Spin 1/2	$SU(3)_C \times SU(2)_L \times U(1)_Y$
squarks, quarks (x3 families)	$Q$	$(\tilde{u}_R \quad \tilde{d}_L)$	$(u_L \quad d_L)$	$(\mathbf{3}, \mathbf{2}, \frac{1}{6})$
	$\bar{u}$	$\tilde{u}_R^*$	$u_R^\dagger$	$(\bar{\mathbf{3}}, \mathbf{1}, -\frac{2}{3})$
	$\bar{d}$	$\tilde{d}_R^*$	$d_R^\dagger$	$(\bar{\mathbf{3}}, \mathbf{1}, \frac{1}{3})$
sleptons, leptons (x3 families)	$L$	$(\tilde{\nu} \quad \tilde{e}_L)$	$(\nu \quad e_L)$	$(\mathbf{1}, \mathbf{2}, -\frac{1}{2})$
	$\bar{e}$	$\tilde{e}_R^*$	$e_R^\dagger$	$(\mathbf{1}, \mathbf{1}, 1)$
Higgs, higgsinos	$H_u$	$(H_u^+ \quad H_u^0)$	$(\tilde{H}_u^+ \quad \tilde{H}_d^0)$	$(\mathbf{1}, \mathbf{2}, +\frac{1}{2})$
	$H_d$	$(H_d^0 \quad H_d^-)$	$(\tilde{H}_d^0 \quad \tilde{H}_d^-)$	$(\mathbf{1}, \mathbf{2}, -\frac{1}{2})$

Table 2.1: Chiral supermultiplets described by the MSSM. The notation used is to have all fields written in terms of left-handed Weyl spinor, hence the conjugate in the right-handed quark and lepton fields. The last column indicates the quantum numbers of each field with respect to the SM gauge symmetry group. Table taken from ref. [26].

chiral supermultiplets are also  $SU(2)$  doublets. There are three of such chiral supermultiplets for the left-handed leptons, which include also the neutrinos and the corresponding sneutrinos, and another three supermultiplets for the left-handed quarks and squarks. All of the chiral supermultiplets, together with the notation commonly used to designate them, is summarized in table 2.1. The interaction and mass eigenstates for the sfermions do not need to be aligned in general, and mixing between superpartners of the the left-handed and right-handed fields of the same SM particle is expected, to form two physical states with different masses.

The color-octet gluon  $g$  lives together with the gluinos  $\tilde{g}$  in the same gauge supermultiplet, sharing a similar  $SU(3)$  structure. In the description of the electroweak interaction as gauge invariance under the group  $SU(2) \times U(1)$  in section 1.3 four gauge fields,  $W_i, B$ , were introduced to account for the 3+1 generators of the group. In the MSSM the three  $W_i$  fields corresponding to the  $SU(2)$  generators live together with the three corresponding winos in a gauge supermultiplet, both vector bosons and fermions keeping a  $SU(2)$  triplet structure. The remaining  $B$  field lives in a separate gauge supermultiplet with the bino. A summary of all gauge supermultiplets is shown in table 2.2.

The Higgs sector in the MSSM is slightly more complex than in the SM. In order to have masses for all SM fermions through Yukawa interactions that are SUSY invariant, there needs to be two complex scalar Higgs  $SU(2)$  doublets, instead of just one as discussed in section 1.4 for the SM. Therefore the Higgs sector in the MSSM description is composed by two chiral supermultiplets, each containing a complex scalar  $SU(2)$  doublet with the corresponding spin-1/2 higgsinos. The two scalar Higgs doublets translate to eight degrees of freedom, out of which only three are needed to provide mass to the gauge bosons  $W^\pm, Z$  and the rest become observable particles. It is customary to label the Higgs doublet that couple to the up-type quarks in the Yukawa interactions as  $H_u$ , and the one that couples to the down-type quarks as  $H_d$ . In order to have a consistent theory, they must have hypercharge  $Y = +1/2$  and  $Y = -1/2$ , respectively, which leads to the phenomenological prediction of the existence of charged Higgs bosons  $H^\pm$ . The remaining fields mix to produce two CP-even neutral Higgs bosons,  $h^0, H^0$ , with  $H^0$  being heavier than  $h^0$ , and a CP-odd neutral Higgs boson  $A^0$ . Similar to the case of the sfermions, the interaction eigenstates stated in tables 2.1 and 2.2 do not necessarily have to match the physical observable states. In particular, it is possible in the model for the higgsinos to mix

Names	Spin 1/2	Spin 1	$SU(3)_C \times SU(2)_L \times U(1)_Y$
gluino, gluon	$\tilde{g}$	$g$	(8, 1, 0)
winos, W bosons	$\tilde{W}^\pm \quad \tilde{W}^0$	$W^\pm \quad W^0$	(1, 3, 0)
bino, B boson	$\tilde{B}^0$	$B^0$	(1, 1, 0)

Table 2.2: Gauge supermultiplets described by the MSSM. The last column indicates the quantum numbers of each field with respect to the SM gauge symmetry group. Table taken from ref. [26].

with the winos and binos to form neutral and charged physical states commonly referred to as neutralinos and charginos, respectively.

The Yukawa interactions from the SM, together with a higgsino mass term, are included into the model through the superpotential, which takes the form

$$W_{\text{MSSM}} = \tilde{u}\hat{Y}^u\tilde{Q}H_u - \tilde{d}\hat{Y}^d\tilde{Q}H_d - \tilde{e}\hat{Y}^e\tilde{L}H_d + \mu H_u H_d, \quad (2.28)$$

where  $\hat{Y}^{u,d,l}$  are 3x3 matrices that encode the Yukawa couplings for the up-type quarks, down-type quarks and leptons, respectively. Following the discussion from the previous section, the superpotential only includes scalar fields, and not their complex conjugates. The SM Yukawa interactions appear in the MSSM through derivatives of the superpotential, together with new vertices of the form squark-higgsino-quark and slepton-higgsino-lepton. These interactions, however, are proportional to the Yukawa couplings, which are known to be very small for the first two fermion generations. From a detection point of view at colliders, the most interesting mechanisms to generate SUSY particles are those mediated by the gauge interactions, since particles in the same supermultiplet couple to the gauge bosons with the same strength.

### 2.3.1 Soft SUSY breaking in the MSSM

As discussed in section 2.2.3, SUSY must necessarily be broken and the physical states in the same supermultiplet in the MSSM cannot be mass degenerate. The most general SUSY breaking terms that can be included in the MSSM Lagrangian, provided that they do not introduce the hierarchy problem back into the model, are similar to those in eq. 2.27, and can be written explicitly for the fields in the MSSM as

$$\begin{aligned} \mathcal{L}_{\text{soft}} = & -\frac{1}{2} \left( M_1 \tilde{B} \tilde{B} + M_2 \tilde{W} \tilde{W} + M_3 \tilde{g} \tilde{g} + \text{compl. conj.} \right) \\ & - \tilde{Q}^\dagger \hat{m}_Q^2 \tilde{Q} - \tilde{L}^\dagger \hat{m}_L^2 \tilde{L} - \tilde{u} \hat{m}_u^2 \tilde{u}^\dagger - \tilde{d} \hat{m}_d^2 \tilde{d}^\dagger - \tilde{e} \hat{m}_e^2 \tilde{e}^\dagger \\ & - \left( \tilde{u} \hat{a}_u \tilde{Q} H_u - \tilde{d} \hat{a}_d \tilde{Q} H_d - \tilde{e} \hat{a}_e \tilde{L} H_d + \text{compl. conj.} \right) \\ & - m_{H_u}^2 H_u^* H_u - m_{H_d}^2 H_d^* H_d - (b H_u H_d + \text{compl. conj.}) \end{aligned} \quad (2.29)$$

The first and second lines of eq. 2.29 consist of mass terms for the superpartners of the SM fields, with  $M_1, M_2$  and  $M_3$  being the mass terms for the bino, wino and gluino, respectively, and  $\hat{m}_X^2$  are hermitian 3x3 matrices that lead to squark/slepton mass terms and, if they are not diagonal, to squark and slepton mixing terms. The terms in the third line of eq. 2.29 are analogous to the Yukawa terms introduced in the superpotential, with  $\hat{a}_X$  being 3x3 matrices, while the last line represent modifications to the Higgs potential. All of these new couplings

have positive mass dimension and of order  $m_{\text{soft}}$  (or  $m_{\text{soft}}^2$  according to the dimension), where  $m_{\text{soft}}$  is the SUSY breaking scale. To ensure that the quadratic divergences in the corrections to scalar masses are exactly cancelled in the theory without large fine tuning,  $m_{\text{soft}}$  should not be larger than  $\sim 10^3$  GeV, argument commonly referred to as *naturalness* of the theory.

The soft SUSY breaking terms introduce a large number of new free parameters in the MSSM, compared to the relatively few introduced by the superpotential from eq. 2.28. In total, the MSSM has [41] 105 parameters that account for mass terms, phases and mixing angles that cannot be removed. Many of these terms introduce direct and indirect flavour-changing effects together with CP-violating enhancement on the gauge interactions that have not been observed experimentally, and so they can be constrained or completely eliminated from the theory. To ensure that the phenomenology of the SM can be reproduced from the MSSM, it is possible to introduce the *soft SUSY-breaking universality* conditions. To avoid flavour-changing effects, it is enough to require the sfermion mass terms to be flavour-blind,

$$(\hat{m}_X^2)_j^i = m_X^2 \delta_j^i, \quad (2.30)$$

and the matrices  $\hat{a}_X$  to be proportional to the corresponding Yukawa couplings,

$$\hat{a}_X = A_{X0} \hat{Y}^X, \quad (2.31)$$

where  $A_{X0}$  could be in principle a complex number that changes for each sfermion  $X$ . In addition, to restrict all CP-violating effects to those introduced by the CKM matrix in the SM, it is possible to require that the gaugino masses  $M_1, M_2$  and  $M_3$ , together with the couplings  $A_{X0}$ , are all real. Considering that the Yukawa couplings are negligible for the first and second fermion generation compared to those of the third generation, these conditions reduce the number of free parameters in the model drastically, which stand now at only 14,<sup>12</sup> without counting those of the SM.

As mentioned in section 2.2.3, the form of the soft breaking terms that arise in the visible sector depends on how the SUSY breaking effects are transmitted from the hidden sector. While it is common for a SUSY breaking mechanism to partially include some of these conditions, such as the suppressed flavour-changing effects in GMSB scenarios, it is difficult in general to find a way to include all of these conditions naturally without needing to impose them in an *ad hoc* manner.

### 2.3.2 R-parity

The superpotential written for the MSSM in eq. 2.28 can successfully reproduce the SM Yukawa interactions and Higgs mass term. However, there are extra terms that could be added to the superpotential without breaking SUSY nor gauge invariance and without posing any renormalization problems, such as

$$W_{\Delta L=1} = \frac{1}{2} \lambda^{ijk} \tilde{L}_i \tilde{L}_j \tilde{e}_k + \lambda'^{ijk} \tilde{L}_i \tilde{Q}_j \tilde{\bar{d}}_k + \mu''^i \tilde{L}_i H_u, \quad W_{\Delta B=1} = \frac{1}{2} \lambda''^{ijk} \tilde{u}_i \tilde{d}_j \tilde{\bar{d}}_k, \quad (2.32)$$

where the constants  $\lambda, \lambda'$  and  $\lambda''$  are the coupling strength for each of the interactions, and the indices  $i, j, k = 1, 2, 3$  run over the generations of chiral supermultiplets. Given that all

---

<sup>12</sup>The left free parameters are 3 gaugino masses, 5 squared masses for sleptons and squarks, 3 cubic couplings for scalars and 3 Higgs mass parameters

particles living in the same supermultiplet share the same quantum numbers,<sup>13</sup> these terms in the superpotential violate lepton and baryon number conservation in precisely one unit.

It is possible to impose a discrete  $\mathbb{Z}_2$  symmetry on the MSSM such that these terms, and the phenomenological implications they bring into the model, are forbidden. For each particle, defining the quantum number

$$P_R = (-1)^{3(B-L)+2s}, \quad (2.33)$$

where  $B, L$  are its baryon and lepton numbers, respectively, and  $s$  its spin, and requiring that each interaction vertex has  $P_R = +1$ , the lepton and baryon number violating terms in the Lagrangian disappear. This symmetry is commonly known as *R-parity* [42], and the MSSM is necessarily a R-parity conserving (RPC) model if it were to reproduce the phenomenology of the SM. In addition, the quantum number  $P_R$  is useful to label the SM particles, as they all have even R-parity ( $P_R = +1$ ), while all their superpartners, or supersymmetric particles, introduced in the MSSM have odd R-parity ( $P_R = -1$ ).

R-parity conservation also leads to some interesting phenomenological consequences for the model, specially from the point of view of particle colliders. In particular, any RPC model predicts that R-parity odd particles can only be produced in pairs, and in the decay of any of them there always needs to be an odd number of supersymmetric particles. That means that the lightest supersymmetric particle (LSP) is stable, and any other R-parity odd particle necessarily needs to decay into the LSP eventually. From the cosmological point of view, RPC models provide a natural candidate for dark matter.

The addition of such symmetry into the model, however, might seem ad hoc just to preserve the conservation of lepton and baryon numbers. Nevertheless, it is still possible that couplings such as the ones proposed in eq. 2.32 exist in nature. From experimental limits on proton decay rate and other well known processes [43] it is possible to greatly constraint the parameter space of the couplings  $\lambda$ , although it is still possible that some of them survive. In such case, this type of extension of the MSSM superpotential can provide rich phenomenology that is currently targeted by multiple searches in the LHC experiments.

### 2.3.3 Mass spectrum of the MSSM

The mass terms for gauge bosons in the SM appear after the electroweak spontaneous symmetry breaking, where the Higgs field obtains a VEV due to non-zero minima in the Higgs potential. The masses of all SM fermions arise as well after the Higgs field gets a VEV through the Yukawa interactions. In order to have a consistent theory with the SM in the infrared limit, the Higgs sector of the MSSM needs to reproduce such behaviour. From the superpotential in eq. 2.28 and the SUSY breaking soft terms it is possible to define a scalar potential  $V(\vec{\phi}, \vec{\phi}^*)$  for the MSSM, whose minimum need to break the gauge symmetry group  $SU(3)_C \times SU(2)_L \times U(1)_Y$  into  $SU(3)_C \times U(1)_{em}$ . The Higgs potential in the model is therefore the part of the scalar potential  $V$  that includes only terms with the fields  $H_u$  and  $H_d$ :

$$\begin{aligned} V_H = & (|\mu|^2 + m_{H_u}^2)|H_u^0|^2 + (|\mu|^2 + m_{H_d}^2)|H_d^0|^2 - (bH_u^0H_d^0 + \text{compl. conj.}) \\ & + \frac{1}{8}(g^2 + g'^2)(|H_u^0|^2 - |H_d^0|^2)^2, \end{aligned} \quad (2.34)$$

---

<sup>13</sup>Baryon and lepton numbers are not strictly quantum numbers as they are not associated to any symmetry group of the SM. However, since both of them are accidentally conserved in the SM, it is easy to refer to them as such. The chiral supermultiplets carry the following baryon numbers:  $Q_i \rightarrow 1/3, \bar{u}_i, \bar{d}_i \rightarrow -1/3$ , and 0 for the rest. The lepton number of the chiral supermultiplets is:  $L_i \rightarrow +1, \bar{e}_i \rightarrow -1$  and 0 for the rest.

where the term in the second line come from the D-terms in the scalar potential, and the  $SU(2)_L$  symmetry has been used to rotate the Higgs doublet  $H_u$  so only the neutral component gets a VEV. By similar arguments to those used in section 1.4.1, if the charged component of  $H_u$  cannot get a VEV, the charged component of  $H_d$  cannot get a VEV either, therefore preserving electromagnetism at the minimum of the potential. To ensure there is electroweak symmetry breaking, the scalar potential introduced in the previous equation needs to be bound from below so that a minimum can exist.

Under the right conditions on the potential  $V_H$ , the neutral components of the Higgs doublets will therefore obtain a VEV that can be labelled as

$$v_u = \langle H_u^0 \rangle, \quad v_d = \langle H_d^0 \rangle. \quad (2.35)$$

Given that these components mix to create different mass eigenstates, out of which the lightest one is assumed to be the SM Higgs boson, it is possible to relate these VEVs to the one measured for the SM Higgs as

$$v_u^2 + v_d^2 = v^2, \quad \tan \beta = \frac{v_u}{v_d}, \quad (2.36)$$

where  $\beta$  is the neutral Higgs mixing angle. The expressions for the gauge boson mass terms in the SM can be therefore retrieved to be identical to those in eq. 1.48. It is also possible to find the mass terms for all the MSSM Higgs physical states by minimizing the potential  $V_H$ . The masses for all of them can be arbitrarily large except for the lightest state,  $h^0$ , that must be bound from above.<sup>14</sup> Including all relevant radiative corrections in the computation of the upper bound on the mass of the lightest neutral Higgs  $h^0$ , and assuming that all relevant particles contributing to such correction have masses below 1 TeV, due to naturalness arguments, the upper limit found is [44]

$$m_{h^0} \lesssim 135 \text{ GeV}. \quad (2.37)$$

This constraint becomes more and more difficult to reconcile, however, with the measured mass of the SM Higgs boson as experiment have excluded large sections of the SUSY parameter space, tightening this upper bound.

Even though the masses for all of the supersymmetric particles ( $P_R = -1$ ) depend highly on the SUSY breaking mechanism, there are some general trends that appear in the MSSM regarding the mass hierarchy. Firstly, due to R-parity conservation in the MSSM the LSP is usually the lightest neutralino  $\tilde{\chi}_1^0$  (except for scenarios with very light gravitinos, e.g. GMSB as supersymmetry breaking mechanism). The main component of the lightest neutralino is generally determined by the mass hierarchy of the higgsino, wino and bino, which usually also determines the mass relation between  $\tilde{\chi}_1^0$  with the next neutralino  $\tilde{\chi}_2^0$  and the first chargino  $\tilde{\chi}_1^\pm$ . The gluino  $\tilde{g}$ , in comparison, is expected to have a heavier mass, given that the larger coupling for QCD than electroweak interactions would make the gluino mass larger when the renormalization group equations are evolved down to the electroweak scale. Finally, the lightest scalars are typically the superpartners of the heaviest quarks and leptons, i.e.  $\tilde{t}_1, \tilde{b}_1$  and  $\tilde{\tau}_1$ , with the rest of the scalars usually much heavier, while the remaining squarks are generally almost degenerate and heavier than the sleptons.

---

<sup>14</sup>The masses for the CP-odd and charged mass eigenstates are fixed by the parameters in the scalar potential  $V$ , namely  $\mu, m_{H_u}^2$  and  $m_{H_d}^2$ . The masses for the CP-even neutral states, however, depend on the mixing angle  $\beta$ , as expected, and can be computed at tree level as

$$m_{h^0, H_0}^2 = \frac{1}{2} \left( 2|\mu|^2 + m_{H_u}^2 + m_{H_d}^2 + m_Z^2 \mp \sqrt{(2|\mu|^2 + m_{H_u}^2 + m_{H_d}^2 - m_Z^2)^2 + 4m_Z^2(2|\mu|^2 + m_{H_u}^2 + m_{H_d}^2) \sin^2(2\beta)} \right),$$

and while the mass for  $H^0$  can increase without limit, at tree level the mass of the lightest CP-even neutral state is bounded from above as  $m_{h^0} < m_Z |\cos(2\beta)|$ .

Naturalness arguments, on the other hand, can also guide the discussion on mass hierarchy in the MSSM. For the quadratic divergences in the loop correction to scalar masses to disappear without the need for large fine tuning, the stop masses need to be bounded from above and so they need to be generally light ( $< 1$  TeV). Similarly, the lightest sbottom should have a similar mass to that of the stops and be mostly composed by the  $\tilde{b}_L$  component. In addition, the two lightest neutralinos  $\tilde{\chi}_{1,2}^0$  and the lightest chargino  $\tilde{\chi}_1^\pm$  are expected to be dominated by the higgsino component, almost degenerate and lighter than the stops. The mass of the gluino is also constrained and, while allowed to be heavier than the stops, it should not be much larger than 2 TeV. The conditions applied to the rest of the sparticles are much looser: they are allowed to get much larger values without bringing back the hierarchy problem into the model. These requirements, however, are becoming more and more difficult to reconcile with current experimental exclusion limits, which seems to lead to the conclusion that SUSY cannot fulfil the purpose that it was originally meant to and so it is less likely to be realized in nature. Nevertheless, this discussion about naturalness only applies to the MSSM, and larger supersymmetric extensions of the SM could still be phenomenologically feasible within current experimental limits. And even if SUSY cannot fix the hierarchy problem in a straightforward manner, there is not a reason why such a symmetry should not exist in nature, and, should it be found, it could provide a deeper understanding of the relations between the particles that constitute the building blocks of matter and the ones that carry their interactions.

## 2.4 Approaches in the search for SUSY

Any SUSY extension of the SM brings rich phenomenology that can be tested experimentally. In the MSSM the particle spectrum is increased by more than a factor two compared to that of the SM. Many constraints can be imposed on SUSY models from SM precision measurements and cosmology. However, the main driver for SUSY discovery and exclusion are direct searches in particle colliders. Already the generations of colliders previous to the LHC provided large sensitivity to multiple SUSY scenarios, mainly on strongly produced particles. At the time of this writing, the LHC has the largest potential in history to access a large portion of the phase space where the MSSM could live. In figures 2.1a and 2.1b the production cross section for several supersymmetric particles at a center-of-mass energy  $\sqrt{s} = 13$  TeV in proton-proton collisions at the LHC is shown in a wide range of possible masses. These large cross sections, combined with the large integrated luminosity of  $3000 \text{ fb}^{-1}$  expected by the end of the LHC program, constitute an incomparable opportunity to explore a number of SUSY scenarios.

In most of the direct searches carried out at the LHC the MSSM phenomenology is used to design the analysis strategies. In spite of its clear predictions in terms of particle content, the large number of free parameters introduced by the SUSY breaking soft terms make it difficult to delimit the region of the phase space that should be scanned, since the different signatures in particle detectors can vary. It is for this reason that several assumptions are often made when modelling the signal processes used to design the searches carried out at the LHC experiments. The most common ones are detailed below. It is clear that the results yielded by the experimental observations always need to be interpreted inside the corresponding framework, and in some cases the extrapolation to more general SUSY scenarios might not be straightforward.

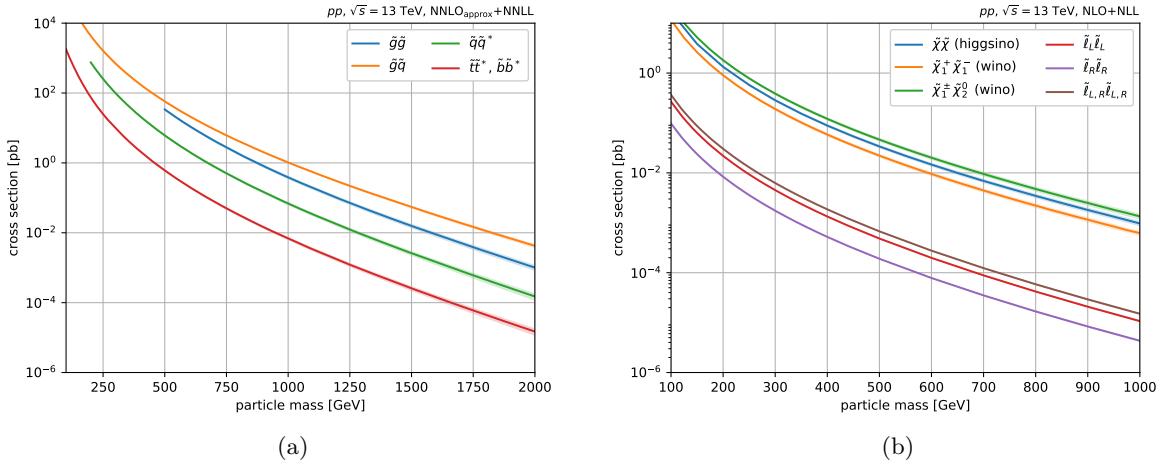


Figure 2.1: Production cross sections for several SUSY (a) strongly-produced particles and (b) electroweak particles in proton-proton collisions at a center-of-mass energy  $\sqrt{s} = 13$  TeV in the LHC at NNLO+NNLL orders in the perturbation expansion.

### 2.4.1 Simplified models

The most common approach in LHC experiments to search for new SUSY phenomena is to interpret the results in terms of *simplified models* [45, 46]. The idea is to divide a full model into reduced particle spectra, in a way that only a subset of all the particles in the model are considered relevant for the search. In the case of searches targeting the MSSM usually this means considering the effect of only two or three particles in the model: the ones produced in the proton-proton collision, and the LSP. A search for gluino pair production, for example, would only consider the gluino  $\tilde{g}$  and the LSP (typically the lightest neutralino  $\tilde{\chi}_1^0$ ) as relevant, with the rest of the particles being considered too heavy to have any effect on the couplings, cross sections and branching ratios. This also implies that often the searches consider only one possible decay mode of the supersymmetric particles in the interpretation of their results.

This approach is very convenient in LHC experiments, as it allows to optimally design data analyses by splitting the immense available phase space into manageable models with well defined signatures. It also allows to interpret the results in a concise way, usually as direct exclusion limits on the target particle mass and the mass splitting with the LSP without needing to worry about all other parameters that may come into play. This is the way in which most experimental results are usually presented.

Simplified models are a useful way to decouple the phenomenology accessible in collider experiments from the unknown SUSY breaking mechanism, providing tools to explore multiple corners of the MSSM parameter space. This approach, however, may fail to capture the large complexity of the full model and the results need to be interpreted carefully. In general, the exclusion limits derived using simplified models tend to loosen up when extrapolated to more complex scenarios, due to the common assumptions on branching ratios and heavy particles mentioned above. Nevertheless, with the appropriate statistical treatment they present an advantage over the full model, and the complexity of the bigger picture can often be restored, or at least partially, through statistical combinations and re-interpretations of the results from such searches.

### 2.4.2 Phenomenological MSSM

The phenomenological MSSM (pMSSM) makes use of the soft SUSY-breaking universality conditions introduced in section 2.3.1 to largely constrain the available phase space of the MSSM. While these conditions are motivated by the needed suppression of CP-violation and flavour-changing effects introduced by some of the SUSY breaking soft terms, it is often hard to find a breaking mechanism that naturally leads to all of them. Nevertheless, it is a useful tool to approach the MSSM in a way that is phenomenologically consistent with the SM, with a reduced parameter space that makes it easier to explore experimentally. In general, the available free parameters in the pMSSM are increased with respect to those from the soft SUSY-breaking universality conditions to allow the first and second generation sfermions to have an independent mass from those of the third generation. In total, there are 19 free parameters, that can be computed as:

- Three gaugino masses:  $M_1, M_2, M_3$ .
- Three Higgs potential parameters:  $\tan \beta, m_A, \mu$ .
- Five squared-mass parameters for the first and second generations sfermions, which are assumed to be degenerate:  $m_{Q_{1,2}}^2, m_{\tilde{u}_{1,2}}^2, m_{\tilde{d}_{1,2}}^2, m_{L_{1,2}}^2, m_{\tilde{e}_{1,2}}^2$ .
- Five squared-mass parameters for the third generation sfermions:  $m_{Q_3}^2, m_{\tilde{u}_3}^2, m_{\tilde{d}_3}^2, m_{L_3}^2, m_{\tilde{e}_3}^2$ .
- Three cubic couplings:  $A_{t0}, A_{b0}, A_{\tau0}$ .

The cubic couplings  $A_{t0}, A_{b0}, A_{\tau0}$  are often not included, given that their impact tends to be negligible. However, in the light of recent muon  $g - 2$  measurements they may play a relevant role in model building as well. Even though these constraints lead to a large reduction in the possible free parameters of the MSSM, it is still computationally expensive to use such a model to provide potential signatures from which design data analyses in the LHC experiments. The usual strategy is to scan separately the different corners of the available phase space with the help of the data analyses designed to target different simplified models.

## Chapter 3

# The Large Hadron Collider and the ATLAS experiment

### 3.1 Introduction

After the scientific breakthroughs achieved in particle physics during the mid 1980s and the later precision measurements performed with the Large Electron-Positron Collider at CERN and Tevatron at Fermilab, the Standard Model had been solidly established as the most accurate description of the elementary particles in nature. However, by the late 1990s the confirmation of the last piece of the puzzle had not yet arrived: the discovery of the Higgs boson, finally providing proof for our best understanding as to how particles obtain mass. Furthermore, even though physics beyond the Standard Model had been deeply explored from a theoretical point of view during the previous decades, by the end of LEP operations there was still no trace as to where new physics could be found. The task of shedding new light upon the unsolved problems was then left to LEP’s successor: the Large Hadron Collider (LHC) [47].

Even though the LHC first operated in September 2008 at CERN, it was conceived around the time when LEP was proposed. The LHC was presented already in 1984 during the International Committee for Future Accelerators workshop, only three years after the  $e^+e^-$  collider was approved. The plan back then was different to the final design eventually carried out for the hadron collider. At first the LHC was conceived to be operated at the same time as LEP, with both colliders hosted in the same 27 kilometre long underground tunnel in the CERN area, near the Franco-Swiss border. As shown in figure 3.1, LHC would be installed right above LEP and would be ready for proton-proton and Pb-Pb collisions, as well as having the potential to combine particle collisions from both accelerators. The LHC was finally approved for construction by the CERN Council in 1994, although the idea of parallel operation with LEP was eventually abandoned due to the high complexity and economic requirements of operating both colliders together. After its last operation in 2000, LEP was removed from the underground tunnel and the works to install the LHC began.

With the arrival of the new colossus, a state-of-the-art generation of particle detectors was needed to take the most from the high energy collisions. The ATLAS (A Toroidal LHC ApparatuS) Collaboration first expressed the interest in building a general purpose detector in 1992, with several proposals included in their Letter of Intent [49]. In the following years, the technical details for the detector were progressively worked out and published in the corresponding technical design reports, culminating in the formal approval for the construction of the experiment

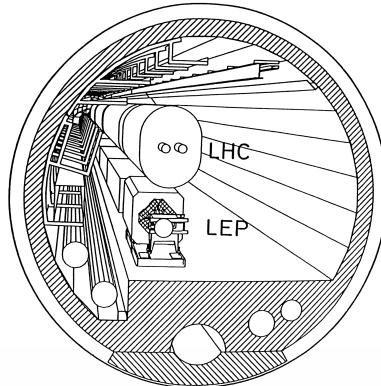


Figure 3.1: Original conception of LHC operating together with LEP [48].

in 1997 by the LHC Experiments Committee and the CERN Council. For the next ten years the excavation of the cavern that would hold the experiment and the process of building the different parts of the detector was carried out. At the end of 2008 the construction and emplacement of the ATLAS detector were finalized, right on time for LHC proton-proton collisions. Such a feat came together with great advancements in the fields of cryogenics and electronics, pushing the frontier of technology to unprecedented levels.

The LHC, together with the particle detectors that record the collisions therein, has open the largest window into the subatomic world. Not only has it provided the means to observe the Higgs boson for the first time, but it also allows for the most accurate precision measurements of the Standard Model up to the date of this writing. In addition, the potential for discovery of physics beyond the Standard Model is much larger than that of the LHC's predecessors, LEP and Tevatron, given the large luminosity delivered combined with the unprecedented energy levels reached in the collisions. The first collisions were awaited with great excitement and expectations, however the total recorded data so far constitute only a small percentage of the total expected by the end of the lifetime of the accelerator, which has large discovery potential.

In this chapter an overview of the LHC and the ATLAS detector are provided. In section 3.2 the LHC system is discussed, with emphasis on the magnets and resonant cavities on subsection 3.2.1. In subsection 3.2.2 the chain of accelerators that provide the LHC with protons is briefly examined, while the operation parameters of the LHC are discussed in 3.2.3. Finally, the section 3.3 provides an overview of the sub-detectors in the ATLAS experiment, focusing on the inner detector, the calorimeters and the muon chambers in subsections 3.3.2, 3.3.3 and 3.3.4, respectively.

## 3.2 The Large Hadron Collider

The LHC is a quasi-circular accelerator with two rings of superconducting magnets that provide the necessary magnetic field to bend the trajectory of two counter-rotating particle beams, placed at the end of a complex accelerator chain at CERN. A circular accelerator provides the means to increase the beam energy greatly through successive rotations, an advantage over linear accelerators. With circular motion of charged particles, however, synchrotron radiation is produced, which can severely decrease the energy levels reached by the beam. Since the energy lost due to synchrotron radiation is inversely proportional to the mass of the charged particles

as

$$\frac{dE}{dt} \propto \frac{E^4}{m^4 r^2}, \quad (3.1)$$

the heavier the particles are, the lower the energy loss is. Thus, the LHC is designed to operate with protons and Pb ions, in order to reduce the loss of energy in the beam acceleration compared to its predecessor, LEP, which used to collide the lighter electrons. This increase does not come for free, since the magnetic field needed to bend the trajectory of the beam increases with the mass and the energy of the charged particles, and so very large magnetic fields are needed in the LHC to keep the beams rotating in the adequate trajectory once the high energy regime is reached. A collider such as the LHC presents a large advantage over fix-target experiments since the energy reached in the collisions is twice the energy of each single beam. The designed center-of-mass energy for proton-proton collisions at the LHC is 14 TeV.

The LHC contains eight straight sections, where the radiofrequency cavities and the physics experiments are located, and eight arcs where the magnets are placed. In total there are four interaction points where the accelerated particles collide. It is where the main LHC experiments, ALICE [50], ATLAS [51], CMS [52] and LHCb [53], are placed to collect data from the collisions. Two of them, ATLAS and CMS, are general purpose experiments, exploring a wide range of physics, from SM precision measurements to BSM searches. LHCb uses data with reduced luminosity to study flavour physics in detail, while ALICE is optimized to study quark-gluon plasma states from Pb-Pb collisions.

Several systems inside the LHC make use of high vacuum levels to ensure an appropriate performance. In the cryogenic system for the magnets and the helium distribution lines, vacuum is used as thermal insulation in order to reduce the energy needed to keep such low temperatures. Keeping the right level of vacuum inside the beampipe is crucial to have a high quality beam with long lifetime, since particle-gas interactions inside can compromise the quality of the beam. The target pressure level is  $10^{-13}$  bar, and corresponds to a beam lifetime of 100 hours. In order to achieve such low pressures a complex system is designed, combining turbomolecular pumps with cryopumping, a technique where the walls of the beam pipe are continuously cooled down with liquid helium in order to condensate the gas inside and get it attached to the wall, thus reducing the number of molecules floating freely in the pipe. However, synchrotron radiation arising from the change in the trajectory of the charge particles in the beam can compromise the vacuum level inside the beam pipe, since it can increase the temperature of the cold bores, releasing the condensed molecules. For this reason, an extra piece of hardware is included, named *beam screen* [54], that runs in parallel and inside the beam pipe, as shown in figure 3.2a. It is designed to shield the cold bores by capturing the heat load from synchrotron radiation, being continuously cooled down to 20 K with helium that runs inside two diametrically opposed tubes. The combination of all these elements allow for the high quality vacuum required for the physics program in all the LHC experiments.

### 3.2.1 Magnet system and resonant cavities

The two key components of the LHC are its magnet system, which allows to bend the trajectory of the particle beam into the circular shape of the tunnel, and the radiofrequency (RF) cavities, that accelerate the particles in order to reach large center-of-mass energy in the collisions. Both of them rely on the cryogenic and vacuum systems mentioned above..

A large number of NbTi superconductive electromagnets are used in the LHC, and can be classified in two categories. On the one hand, the *bending* dipole magnets modify the direction of the beam and ensure a circular trajectory of the particles. In total, there are 1232 dipole magnets

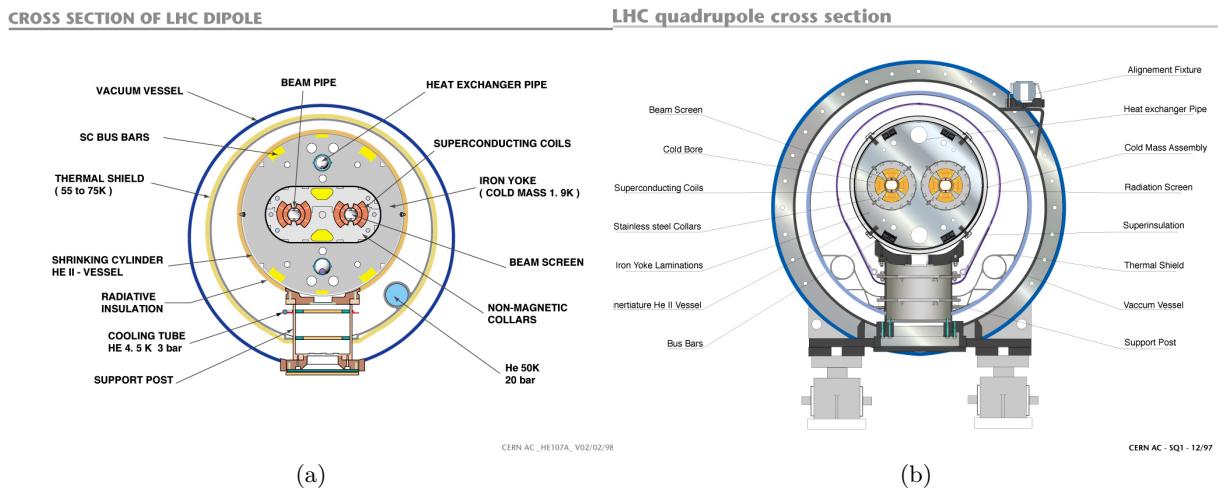


Figure 3.2: The LHC uses twin-bore superconducting magnets due to small space available in LEP tunnel. Cross section of (a) dipole and (b) quadrupole magnets, with different elements highlighted [55].

in the LHC, each of 15 meters in length and reaching a magnetic field of up to 8.33 T through an electric current of 11.8 kA. Figure 3.2a shows the cross section of one dipole. On the other hand, higher-order magnets are used to control the spread of the beam. A total of 858 quadrupole magnets are used to focus the beam, a sketch of which is shown in figure 3.2b. These magnets can focus the beam either on the horizontal or the vertical plane at a time, while defocusing it on the non-focusing plane. For this reason an array of quadrupoles separated by some drift distance is needed to achieve the necessary beam properties. Higher-order magnets are used to control effects raising from the focusing action on the beam, such as betatron oscillations. Large magnetic fields need to be kept at very low temperatures to maintain the superconductivity in the magnets. The LHC uses a complex cryogenics system with about 6000 tons of liquid nitrogen to cool the magnets down to 80 K and 140 tons of liquid helium to further reduce the temperature to the operating value of 1.9 K, in a process that lasts approximately one month.

The particles in the beam are accelerated by means of quickly alternating longitudinal electric fields thanks to superconducting RF cavities. A total of 16 RF cavities, 8 per beam, are located in the straight sections of the LHC in order to accelerate protons from the injection energies of 450 GeV to the nominal 7 TeV. For the acceleration to be efficient, the particles in the beam need to be grouped into *bunches* that travel inside the phase stability regions created by the RF cavities. The RF system, in addition to particle acceleration, also provides *longitudinal focusing* to the beam, since the particle bunches entering a cavity get synchronized with the RF signal. As a result, slow particles in a bunch get accelerated while fast particles get decelerated, keeping the bunches compact and ready for collisions.

### 3.2.2 Accelerator chain

The protons to be used in the collisions undergo a series of steps prior to their injection in the LHC. A scheme of the different accelerators involved in the process is shown in figure 3.3, together with the four main particle detectors that record collisions delivered by the LHC. The stages of proton acceleration at CERN can be summarized as follows:

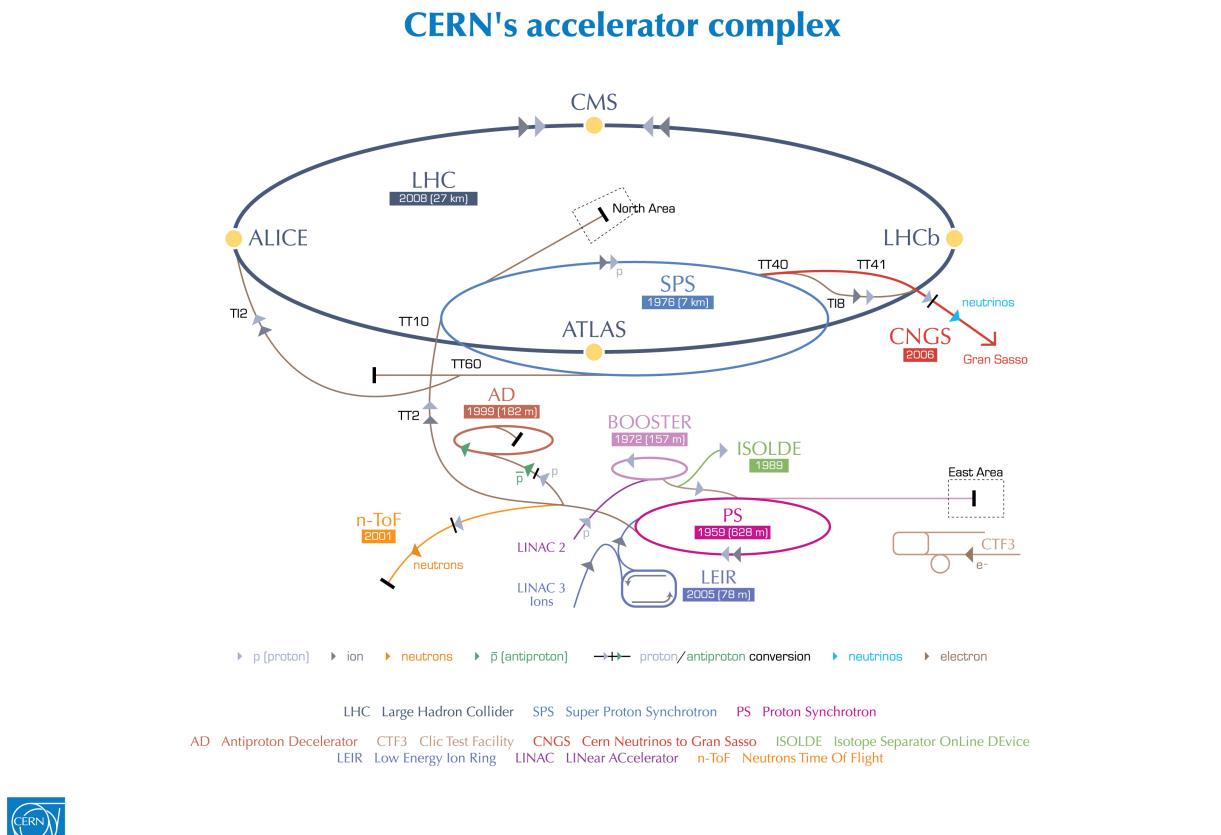


Figure 3.3: Accelerator complex at the CERN laboratory in Geneva, with the different experiments and particle detectors highlighted [56].

- First the protons are extracted by ionizing hydrogen gas thanks to large electric fields and accelerated to an initial energy of 50 MeV. This happens in a 33 meter long linear accelerator named Linac2.
- The protons are injected from Linac2 to the Proton Synchrotron Booster (PSB), a circular accelerator of 157 meter in circumference and made of four superimposed synchrotron rings, where the protons are accelerated from the initial 50 MeV to 1.4 GeV.
- Afterwards, the protons are accelerated by the Proton Synchrotron (PS) [57], the oldest synchrotron experiment at CERN. It is a circular accelerator of 628 meter in circumference and 100 dipole magnets that can accelerate the protons up to energies of 26 GeV.
- The protons are subsequently injected in the Super Proton Synchrotron (SPS), where they are accelerated until they reach energies of 450 GeV, the nominal energy needed to be injected into the LHC. The SPS is underground, unlike the previous accelerators in the chain, and its circumference is 6.9 km. It was used during the late 70s and early 80s to provide proton-antiproton collisions, which revealed for the first time the existence of the  $W$  and  $Z$  bosons [58], confirming the electroweak sector of the SM. Currently it is used to provide the beam not only to the LHC but also to the CERN North Area, where several experiments not related to the LHC are located.
- Finally, the protons arrive from the SPS to the LHC, where they are accelerated to the final energy and prepared for collisions. During LHC Run 2 the protons were accelerated

to 6.5 TeV, resulting in collisions at a center-of-mass energy of 13 TeV. In each turn at the LHC the protons are accelerated of about 0.5 MeV on average, and so the acceleration from 450 GeV to 6.5 TeV takes about 15 minutes.

In the case of operation with heavy ions, the chain is slightly different. They are first extracted and accelerated in Linac3 and then injected to the Low Energy Ion Ring (LEIR), where they are further accelerated to energies compatible with injection into the PS. Afterwards, the acceleration chain is identical to that of protons.

### 3.2.3 LHC operations

The key parameters of the LHC are the center-of-mass energy reached in the collisions and the amount of physics events delivered to the particle detectors. The center-of-mass energy of the proton collisions has evolved throughout the years, from 7 GeV in 2010 to 13 TeV during the whole Run 2. The design energy in the collisions is 14 TeV, while in the next period of data taking, known as Run 3, the energy reached is expected to be 13.6 GeV. Run 3 is planned to span from 2022 to 2026, with annual stops for reparations and small upgrades.

The amount of physics events provided by the LHC can be parametrized in terms of its *instantaneous luminosity*  $\mathcal{L}$ , which relates to the number of events  $N$  corresponding to a certain physics process as

$$\frac{dN}{dt} = \sigma \mathcal{L}, \quad (3.2)$$

where  $\sigma$  is the cross section for such process. The instantaneous luminosity in the LHC can be approximated by

$$\mathcal{L} = \frac{f N_b n_1 n_2}{4\pi \sigma_x \sigma_y}, \quad (3.3)$$

where  $f$  is the revolution frequency of the beam,  $N_b$  the number of bunches,  $n_{1,2}$  the number of protons within each colliding bunch that is colliding, and  $\sigma_x, \sigma_y$  are the transverse beam sizes in the horizontal and vertical directions, respectively. This approximation, however, assumes head-on collisions between bunches, while they actually happen at a certain angle. In addition, the transverse beam size  $\sigma_{x,y}$  is directly determined by the action of the focusing magnets on the beam near the interaction point, which has non-trivial effects on the finite-sized proton bunches. The total amount of delivered collisions is parametrized in terms of the integrated luminosity,  $\mathcal{L}_{\text{int}}$ , by simply integrating the instantaneous luminosity over time,

$$\mathcal{L}_{\text{int}} = \int \mathcal{L} dt. \quad (3.4)$$

Table 3.1 summarizes a list of selected parameters for the LHC throughout the years. In Run 1, the integrated luminosity delivered by the LHC was  $5.46 \text{ fb}^{-1}$ , while for Run 2 it was  $156 \text{ fb}^{-1}$ . The luminosity delivered for each year in Run 1 and Run 2 is shown in figure 3.4a.

In any bunch crossing multiple proton-proton interactions are expected to happen, which are commonly referred to as *pileup*, leading to a cascade of signals in the particle detectors coming from a number of interaction vertices. Pileup can be separated into two categories, depending on the origin of the signals in the detectors. If the multiple interactions arise within one single bunch crossing the collisions are known as in-time pileup, whereas out-of-time pileup involves the bunch crossing immediately before or after the crossing taking place at the interaction point. It is customary to parametrize the pileup in terms of the average number of interactions per bunch crossing,  $\langle \mu \rangle$ . Figure 3.4b shows the luminosity-weighted average number of interactions during

Parameter	2011	2012	2015	2016	2017	2018
center-of-mass energy [TeV]	7	8	13	13	13	13
Av. interactions/crossing	9.1	20.7	13.4	25.1	37.8	36.1
Bunch spacing [ns]	50	50	25	25	25	25
Peak luminosity [ $\text{cm}^{-2}\text{s}^{-1}$ ]	3.65	7.73	5.03	13.8	20.9	21.0
Integrated luminosity	5.5	23.1	4.2	38.5	50.2	63.4

Table 3.1: Summary of the main operational parameters of the LHC during Run 1 and Run 2.

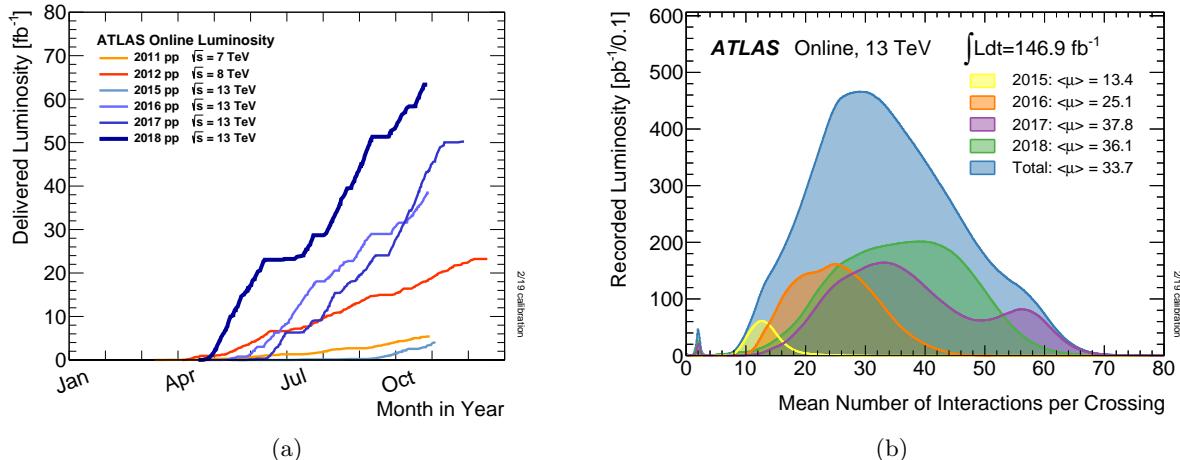


Figure 3.4: (a) Integrated luminosity delivered by the LHC to the ATLAS experiment for each year. (b) Distribution of the luminosity-weighted mean number of interactions per crossing for each year in LHC Run 2. Figures taken from [59].

Run 2. Keeping pileup to manageable levels is crucial for understanding the performance of the physics detectors. The operational parameters of the LHC are key for the pileup during data taking. Pileup reduction is the reason why the bunch spacing was reduced from 50 ns in Run 1 to 25 ns in Run 2, and most likely will be one of the biggest challenges for the High-Luminosity LHC [60] physics program.

### 3.3 The ATLAS experiment

The ATLAS detector is the largest of the four experiments that collect data from the proton-proton and heavy ion LHC collisions, with 44 meter length, 25 meter height, and weighting 7000 tonnes. Due to the large amount of radioactivity from the beams and the collisions, the detector is highly resistant to radiation, with electronics fast enough to handle the interactions happening every 25 ns. The ATLAS Collaboration carries out a wide physics program with data collected by the detector. A number of analyses target SM precision measurements, reaching so-far unexplored kinematic regions thanks to the large energies reached by the LHC. In particular, the large number of top-quarks produced by the LHC brings the possibility to study in detail its production mechanisms and interactions. In addition, with the discovery of the Higgs boson a new field of study has opened up, with several analyses aimed to understand its properties accurately, which is not only relevant to complete the SM description but also to set constraints in new physics models [8]. A large portion of the phase space for direct production of heavy

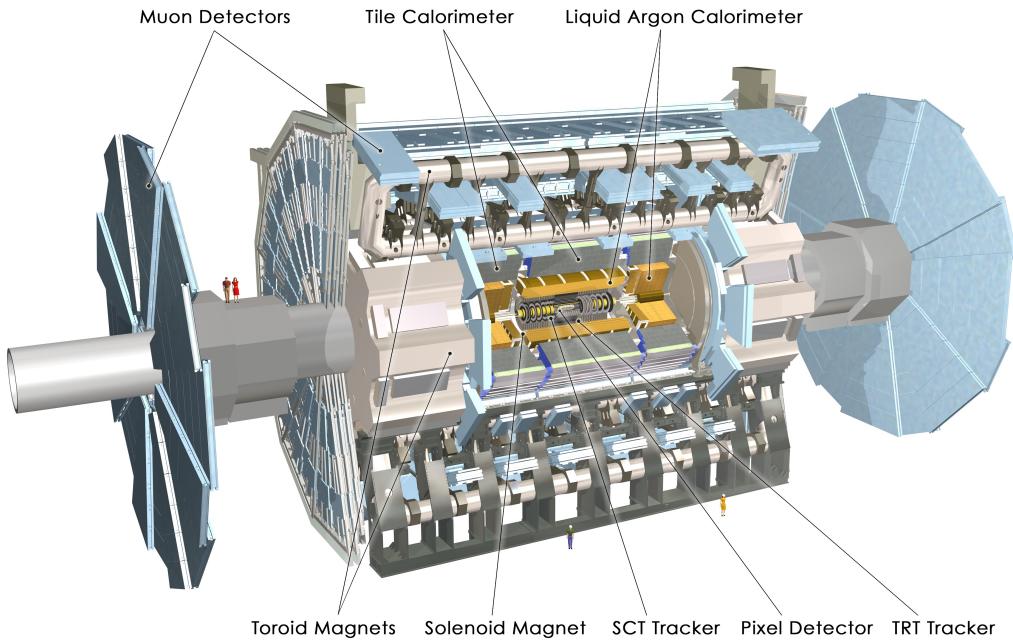


Figure 3.5: Scheme of the ATLAS detector with the different sub-detectors highlighted [51].

resonances is now accessible, thus a strong physics program to search for BSM physics is well established in the ATLAS Collaboration. Finally, flavour physics and heavy ions are also part of the scope of the physics program in ATLAS, and even though it most likely cannot match the sensitivity of dedicated experiments such as LHCb and ALICE, it can provide validation of their results which, at the time of this writing, are of extreme importance given the large tensions with the SM theory that have been measured.

The experiment is divided in different specialized sub-detectors that provide efficient particle identification and determine their position and momentum with very high precision. The ATLAS detector consists of an inner detector (ID) or tracker, that accurately measures the momentum and trajectory of light charged particles; the calorimeters, that collect the energy deposited by particles arising from the LHC collisions; and the muon spectrometer, that determines kinematic information from muons as they are not expected to leave a significant signal in the calorimeters. The calorimeters are further divided in two different systems optimized for different particle detection: the electromagnetic calorimeter (ECal) for electrons and photons, and the hadronic calorimeter (HCal) for strongly interacting particles. The geometry of the detector is divided in the barrel and the end-caps. The barrel is the main cylinder placed around the beampipe, that collects information from particles that are emitted mostly perpendicularly (to a certain extent) to the beampipe. The end-caps close the two sides of the barrel and provide coverage of the forward regions to the collisions. During LHC run periods, two beams of particles travelling in opposite directions are made to collide in the center of the barrel. A scheme of the geometry of the ATLAS detector is shown in figure 3.5.

An important feature of the tracker and the muon spectrometers is the need for a large magnetic field in order to bend the very fast particles produced in the LHC collisions. From the bent trajectories it is then possible to extract information from the original particles such as their charge and their momentum with high precision. These large magnetic fields are provided by a complex magnet system that is part of the ATLAS detector. The ID is embedded in a 2 T axial magnetic field, generated by the central solenoid electromagnet. It is constructed from a

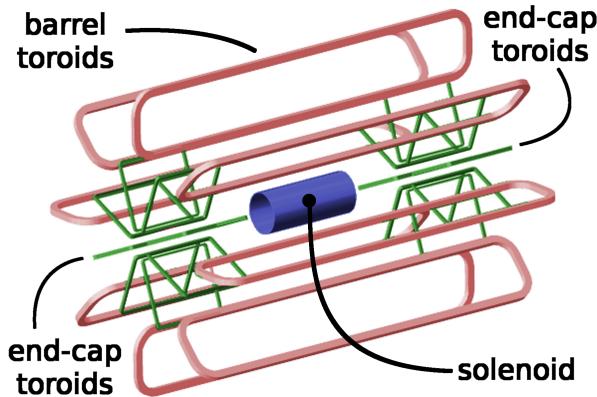


Figure 3.6: Scheme of the magnet system in ATLAS, with the central solenoid providing a 2 T magnetic field for the ATLAS ID, while the toroidal magnets provide 4 T fields to the muon spectrometer in the barrel and the end-caps. Figure taken from ref. [61].

NbTi alloy with Al used for stability, with 5.8 m length and placed between the ID and the ECal. On the other hand, the magnetic field for the muon spectrometers is twice as large as the one generated for the ID. For this, open-air toroidal electromagnets were chosen, since they allow for excellent muon measurements and reconstruction without information from the ID. Additionally, the reduction in material that the muons have to go through also minimizes the energy loss due to scattering in the detector. This choice, however, comes with an increase in the necessary current to reach the desired magnetic fields. The toroidal magnets are divided in two: the magnets in the barrel and those in the end-caps. The magnetic field in the barrel is provided by eight independent toroidal coils, with each of them placed inside individual vacuum vessels, and can reach a peak field of 3.9 T. The fields in the end-caps are provided by two separated systems, one for each end-cap, that are made of eight coils each inside a single cold mass, and that can provide a peak field of 4.1 T. A schematic view of the magnet system in ATLAS is shown in figure 3.6.

### 3.3.1 Coordinate system

One coordinate system is widely used inside the ATLAS Collaboration and is common to most analyses. In this coordinate system, the origin is defined as the interaction point, with the  $z$ -axis running along the beampipe, the Y-axis pointing upwards and the X-axis pointing towards the center of the LHC. The part of the detector that corresponds to positive  $z$  values is referred to as *A-side*, while the part that corresponds to negative values is known as *C-side*. Given that energy is conserved in the XY-plane, transverse to the beampipe, it is typically useful to define the momentum projection over the transverse plane, or *transverse momentum*, as

$$p_T = \sqrt{p_x^2 + p_y^2}, \quad (3.5)$$

where  $p_{x,y}$  are the X,Y components of the momentum  $p$  of a certain particle.

It is common to use spherical coordinates in ATLAS, where the X,Y coordinates are replaced by the angles  $\phi, \theta$ . The azimuthal angle  $\phi$  is defined as the angle between the X-axis and the projection of a point over the XY-plane, and ranges between  $-\pi$  and  $\pi$ , while the polar angle  $\theta$  is the angle between a given point and the transverse plane. It is also common to replace the

polar angle by the *pseudorapidity*,  $\eta$ , defined as

$$\eta = -\ln \tan \left( \frac{\theta}{2} \right), \quad (3.6)$$

which, in the limit of massless particles, coincides with its rapidity. For massive particles it can be obtained without needing to know its total energy as  $\eta$  depends only on its position, which makes it a useful variable with wide applicability. Defined in this way, it is clear that the pseudorapidity for a certain object increases rapidly as it approaches the  $z$ -axis. However, this is not an issue given the limited coverage of the ATLAS detector up to  $|\eta| < 4.9$ , and the even tighter regions to which physics analyses are usually restricted to. It is customary to define angular distances in the  $\eta - \phi$  plane as

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}, \quad (3.7)$$

where  $\Delta\eta$  is the difference in the  $\eta$  coordinate between the two objects, and  $\Delta\phi$  their difference in  $\phi$ .

### 3.3.2 Inner detector

The ID is the closest sub-detector to the beampipe and so it is the first layer of material that the particles produced in the LHC collisions encounter. With a length of 5.4 m, it is placed between the beampipe and the first of the calorimeters, and provides coverage in the region  $|\eta| < 2.5$ . Its purpose is to provide good momentum resolution for charged particles with momentum as low as 100 MeV, and to find secondary vertices in the reconstructed data. It is divided into the barrel, that covers the region  $|\eta| < 1.2$ , and the end-caps, which provide information in the region  $1.2 < |\eta| < 2.5$ . It is composed of three subsystems, the pixel detector, the semiconductor tracker (SCT) and the transition radiation tracker (TRT), each using a different type of technology in order to reach very precise spatial and momentum resolution while reducing the production costs. An overview of each of them is provided below. A scheme of the barrel section of the ID is shown in figure 3.7. The total momentum resolution of the inner detector is

$$\frac{\sigma(p_T)}{p_T} = 0.036\% p_T \oplus 1.3\%, \quad (3.8)$$

where the symbol  $\oplus$  represents sum in quadrature.

#### Pixel detector

The pixel system is the closest to the interaction point, and is composed of radiation-hard silicon detectors called *pixels*. They are p-n junctions that produce a small current when charged particles come through, ionizing electrons from the valence band to the conduction band. The barrel is composed of four concentric cylinders or layers, with the beampipe in the center. The three outermost layers contain 80 pixels of  $50 \times 400 \mu\text{m}$  in each module. The innermost layer, the Insertable B-Layer (IBL) [62], was commissioned during the shutdown period between LHC Run 1 and Run 2, and is placed at only 33.25 mm from the center of the beampipe. It is composed of 14 staves of 20 modules, 12 of which contain planar silicon sensors, while the remaining eight contain 3D pixel sensors, all of them with an area of  $50 \times 250 \mu\text{m}$ . With more than 12 million pixels, the IBL can efficiently reconstruct secondary vertices associated to decays of long-lived particles, providing a great improvement in the identification of  $b$ -hadrons with respect to Run 1. In the end-cap sector the pixel detector has three wheels of 34 cm radius to provide coverage in

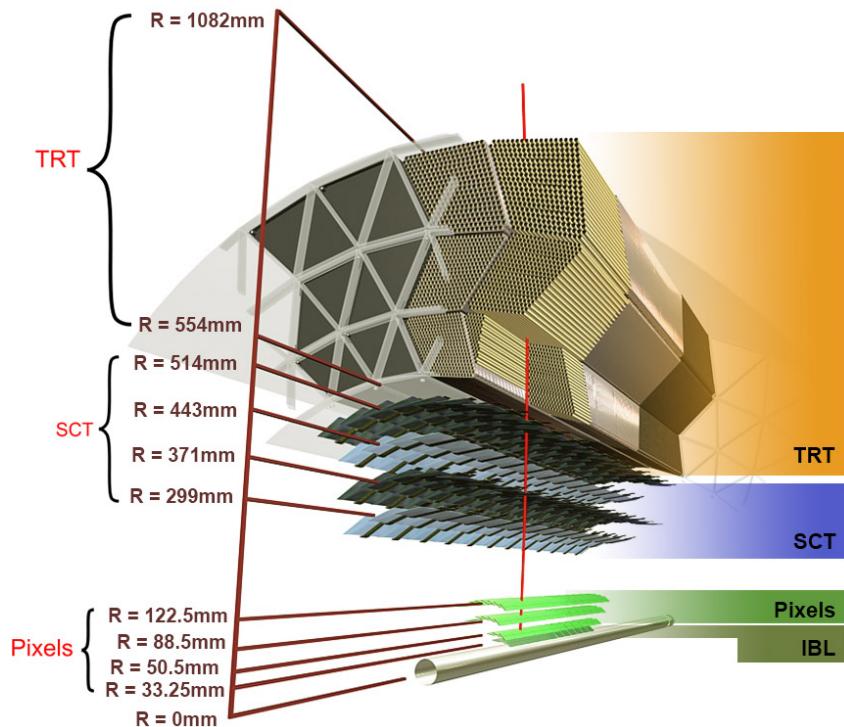


Figure 3.7: Scheme of the transverse view on the barrel sector of the ATLAS ID [51].

the forward region, which are located at 49.5, 58.8 and 65.0 cm from the interaction point. In total, more than 80 million pixels are included in this subsystem, each with a spatial resolution of  $115 \mu\text{m}$  in the  $z$ -direction and  $10 \mu\text{m}$  in the transverse direction.

Since the pixel detector is the closest one to the interaction point, it is the most exposed one to radiation. For this reason it needs to be resistant enough for several years of operations. In order to reduce damage from radiation, the temperature of the pixel detector needs to be kept between  $-10^\circ\text{C}$  and  $-5^\circ\text{C}$  during LHC run periods. In addition, the bias voltage applied to the pixel sensors needs to increase, from the initial 150 V to 600 V in a period of 10 years, modified by several factors such as pixel position or total luminosity accumulated, in order to collect the charge efficiently.

### Semi-conductor tracker

The SCT is placed between the pixel detector and the TRT. It is composed of four concentric cylinders in the barrel, at distances of 30.0, 37.4, 44.7 and 52.0 cm from the interaction point, and provide coverage in the region  $|\eta| < 1.1 - 1.4$ . In each end-cap, there are nine disks that cover the region  $1.1 - 1.4 < |\eta| < 2.5$ . In total, 4088 mono-dimensional p-in-n silicon micro-strip modules to provide binary signals that are used to reconstruct the trajectory of the charged particles. Different to the planar and 3D pixels, the technology used in the SCT is cheaper and has a simpler readout, resulting in lower resolution that is compensated with a larger area compared to the pixel detector. Each of the modules in the barrel is composed of two layers of strips tilted at an angle of 40 mrad and a pitch of  $80 \mu\text{m}$  in order to provide information about the  $z$ -coordinate at the same time than suppressing the rate of fake hits. The spatial resolution of the SCT is  $580 \mu\text{m}$  in the longitudinal direction and  $17 \mu\text{m}$  in the transverse direction.

### Transition radiation tracker

The outermost subsystem in the ATLAS ID is the TRT, which combines the ability that highly energetic particles have to ionize gas molecules with the properties of transition radiation material. It is composed of over 52000 straw tubes in the barrel sector, each with a length of 1.5 m placed in parallel to the beampipe, with multiple tubes combined horizontally in single chains. In the end-cap sectors there are more than 122000 tubes, of 0.4 m length placed perpendicular to the beam. The TRT covers the region  $|\eta| < 2$ .

Each of the tubes has a diameter of 4 mm and is filled with a mixture of gas<sup>1</sup> that is ionized after a charged particle originated in the collision passes through. The electrons produced in the ionization are then collected in a gold plated tungsten wire that serves as anode and piled up among the connected tubes. For this reason, the TRT can only provide tracking information in the  $(r - \phi)$  plane, as two objects hitting the same tube chain in two different points cannot provide a distinguishable signal. The spatial resolution on this plane is  $130 \mu\text{m}$ . In addition, the space between the tubes is filled with a transition radiation material, that emits X-ray when particles travel through. This radiation highly depends on the incoming particle, and therefore provides a useful tool in their identification, in particular for incoming electrons.

#### 3.3.3 Calorimeters

The calorimeters cover the region  $|\eta| < 4.9$  and are placed between the central solenoid and the muon spectrometers. They provide precise measurements of the total energy carried by both charged and neutral particles generated in the interaction point. This task is performed by absorbing the energy of the incoming particle and transforming it into a quantity that can be measured, such as electric charge or light. The approach followed in ATLAS to measure the energy of both electromagnetically interacting particles and hadrons by means of the so called *sampling calorimeters*. In them, layers of an *absorber material* are used to trigger the generation of particle cascades or showers, and are alternated with spaces filled with an active medium, that can create a readable signal from the particles generated in the cascade. The typical physical processes undergone in the shower determine the characteristics needed for efficient calorimetry.

In the case of incoming electrons,<sup>2</sup> the passage through the absorber layers triggers the radiation of a photon through Bremsstrahlung, which in turn generates an electron-positron pair that keeps propagating through the detector. Inversely, when the incoming particle is a photon it will create an electron-positron pair when propagating through the absorber material, which will further radiate photons through Bremsstrahlung. Incoming electrons and photons will therefore have a similar signature in the calorimeter, and it is only through the signals left behind in the tracker that they can be identified. An important quantity that determines the properties of the electromagnetic showers is the *radiation length*,  $X_0$ , which is defined as the mean distance in the calorimeter at which an incoming electron has lost a factor  $1/e$  of its initial energy due to radiation and production of other particles. It depends on the material used to build the detector and is typically used to guide the design of the calorimeter and determine the needed dimensions to be sensitive to a specific range of energies. For example, for incoming electrons of photons with energy between 1 GeV and 1 TeV, 99% of the shower can be contained between  $11 X_0$  and  $22 X_0$ . In ATLAS, the size of the ECal corresponds to between  $22 X_0$  ( $\eta = 0$ )

<sup>1</sup>Originally, they were filled with a mixture of 70% Xe, 27% CO<sub>2</sub> and 3% O<sub>2</sub>. However, due to a gas leakage in 2016, the affected tubes were refilled with a mixture of 80% Ar and 20% CO<sub>2</sub>, which resulted in a reduced cost with respect the original one and similar performance.

<sup>2</sup>The discussion applies as well to its antiparticle, the positron.

and  $33 X_0$  ( $\eta = 1.3$ ) in the barrel sector, and between 24 and  $38 X_0$  in the endcap.

The case of hadrons, such as neutrons or pions, is more complicated due to the large number of particles that are typically involved in the cascades, that can be both strongly and electromagnetically interacting. The principle is similar to the electromagnetic showers: the incoming particle travels through the absorber material in the calorimeter, which triggers collisions that break the nuclear bonds in the particle, leading to new ones produced from its fragments. The new particles further travel inside the detector and experience the same process, leading to a large cascade of particles of many types, potentially including electrons/photons that produce electromagnetic showers as well. The size of a hadronic cascade can be parametrized through the *nuclear interaction length*,  $\lambda_I$ , defined as the mean distance that a hadronic particle travels through the detector before undergoing inelastic scattering. It depends on the material used in the calorimeter, and is typically much larger than the radiation length  $X_0$ . For this reason, ATLAS is designed to include two sets of calorimeters. The ECal is optimized to capture the energy from incoming electrons and photons. Incoming hadrons, however, leave a signature in the ECal but, since  $\lambda_I \gg X_0$ , they would need a much larger calorimeter to be contained. In order to efficiently capture all hadronic cascades, a second calorimeter with much shorter  $\lambda_I$ , the HCal, is placed after the ECal. The total thickness of both calorimeters combined is around 11 times the nuclear interaction length thus a good acceptance of high-energy jets is ensured. An overview of each of them is provided below.

An important feature of the hadronic cascades is that a fraction of the energy of the incoming particle cannot be measured through calorimetry, since it is invested into breaking the nuclear bonds in the particles travelling through the detector. In addition, it is also possible that neutrinos are generated in the decay of some of the particles in the cascade, leaving no signal in the detector. In general, the response of the calorimeter to an electromagnetic shower is larger than that of the hadronic component in a hadronic cascade. The so called *compensating calorimeters* can provide an estimate of the actual energy carried by the hadronic component of the cascade by matching its energy response to that of the associated electromagnetic shower. In the case of ATLAS, the calorimeters are non-compensating, and this issue is dealt with by dedicated calibrations applied downstream, as described in chapter 6. The total energy resolution of the ATLAS calorimeter system is given by

$$\frac{\sigma(E)}{E} = \frac{10\%}{\sqrt{E}} \oplus \frac{17\%}{E} \oplus 0.7\%. \quad (3.9)$$

### Electromagnetic calorimeter

As mentioned above, the ECal is placed outside the central solenoid, enveloping it and the ID. It is divided into the barrel and end-cap sectors to provide optimal coverage in pseudorapidity ( $|\eta| < 1.475$  in the barrel and  $1.375 < |\eta| < 3.2$  in the end-caps). In both cases, they are composed of alternate layers of liquid argon (LAr) as active material and lead plates as absorber. The charged particles that travel through the LAr layers ionize the Ar molecules, and the produced electrons in the ionization are collected in the lead plates, thanks to a drift electric field, and sent to the read-out. The plates are designed in an accordion shape for optimal charge collection with reduced read-out latency and are distributed differently in the different sectors of the calorimeter. In the barrel, they are placed in three layers of decreasing granularity, as can be observed in figure 3.8a. The layer closest to the ID provides very fine  $\eta$  granularity since it is divided in strips of  $\Delta\eta \times \Delta\phi = 0.0031 \times 0.098$ . The intermediate layer is divided in towers of  $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ , while the outermost layer has a coarser granularity, of  $\Delta\eta \times \Delta\phi = 0.05 \times 0.025$ . In the end-caps, the plates are placed in two coaxial wheels, one

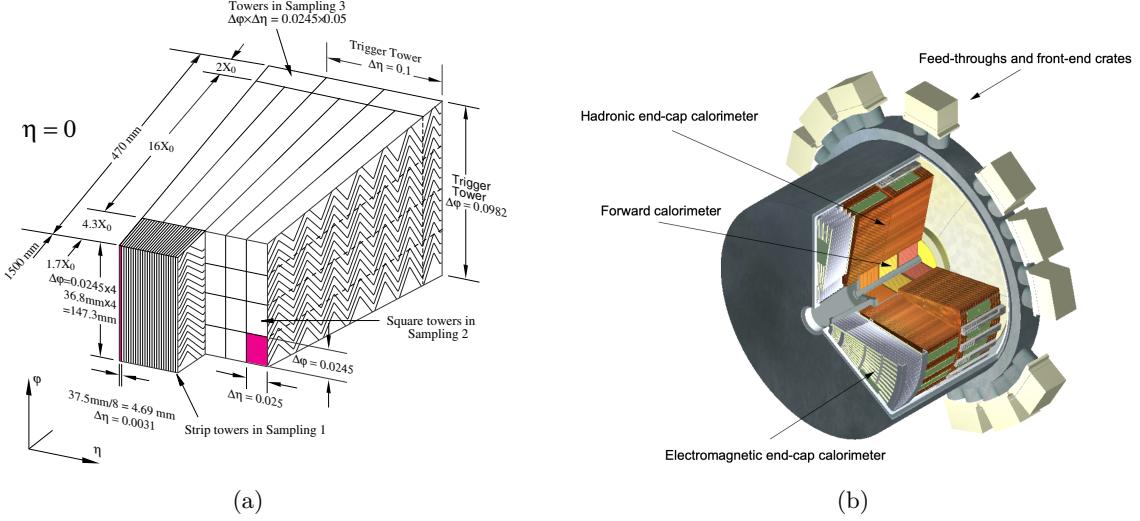


Figure 3.8: (a) Scheme of the granularity and plate disposition in the barrel sector of the ECal. (b) Drawing of the end-cap cryostat containing the end-cap ECal sector, together with the HEC and the FCal. Figures taken from ref. [51].

covering the regions  $1.375 < |\eta| < 2.5$  and the other covering  $2.5 < |\eta| < 3.2$ . The inner wheels are designed with two layers each, while the outer wheels are split into a region for precision physics ( $1.5 < |\eta| < 2.5$ ), where a third layer is included to enhance coverage, and a region with two layers ( $|\eta| < 1.5$ ), that mostly overlaps with the barrel sector. In addition, a presampler is placed before the ECal and provides an estimate of the energy lost when the particles travel through the ID and the central solenoid. The total energy resolution of the ECal is given by

$$\frac{\sigma(E)}{E} = \frac{10\%}{\sqrt{E}} \oplus \frac{17\%}{E} \oplus 0.7\%. \quad (3.10)$$

### Hadronic calorimeter

The HCal is placed outside of the ECal, and is a sampling calorimeter as well. It is divided in three subsystems, each optimized for accurate measurement of the energy carried by hadrons in the region they cover. The first subsystem is composed of an array of scintillator tiles that act as the active medium and steel as absorber material, and is referred to as tile calorimeter (TileCal). The particles from the cascades created by the absorber trigger light emission, that is later collected in photomultiplier tubes. It is divided in two sectors: the central barrel, that covers the region  $|\eta| < 1.0$  and is 5.8 m long, and the extended barrels, that provide sensitivity in the region  $0.8 < |\eta| < 1.7$  and are 2.6 m long each. It has an inner radius of 2.28 m and an outer radius of 4.25 m, which correspond to a thickness of  $9.7 \lambda_I$  at  $\eta = 0$ . The barrel and both extended barrels are all divided into 64 modules of 0.1 rad in size.

The second subsystem is the hadronic end-cap calorimeter (HEC), which provides coverage of the region  $1.5 < |\eta| < 3.2$ . Similarly to the ECal, the HEC uses liquid argon as active medium but it uses copper as absorber, instead of lead. The electrons produced in the ionization of the liquid argon are collected in the copper plates and produce the signal saved in the read-out. In each end-cap, the HEC consists of two wheels of outer radius 2.03 m with two layers each, the

closest to the interaction point having copper plates of 25 mm thickness, while the others have plates 50 mm thick. The plates are separated by gaps of 8.5 mm filled with liquid argon. To avoid extra material in the transition from the ECal to the HCal in the end-caps, the HEC and end-cap sector of the ECal live in the same cryostat, the configuration of which can be seen in figure 3.8b.

The third and final subsystem is the forward calorimeter (FCal), which provides calorimetry information for both electromagnetic and hadronic incoming particles in the region  $3.2 < |\eta| < 4.9$ . Since this subsystem is the most exposed to radiation from the beam and the collisions, the materials are chosen to ensure high resistance to radiation damage. In each end-cap the FCal is divided in three modules. While all modules use liquid argon as active material, the module closest to the interaction point is optimized to capture electromagnetic showers and so it uses copper as absorber, with the other two modules using tungsten for capturing the hadronic cascades. They also share the same cryostat than the end-cap sector of the ECal and the HEC, and the total thickness of the three modules is approximately ten interaction lengths. The total energy resolution of the HCal

$$\frac{\sigma(E)}{E} = \frac{50\%}{\sqrt{E}} \oplus \frac{1\%}{E} \oplus 3\%. \quad (3.11)$$

### 3.3.4 Muon spectrometer

The muon chambers constitute the outermost subdetector in the ATLAS experiment, furthest away from the interaction point than the calorimeters and the ID. Their function is to provide high precision measurement of muon momentum and charge identification by deflecting their tracks through the magnetic field generated by the toroid magnets. The magnetic field created by the barrel toroids bends the tracks for muons in the region  $|\eta| < 1.4$ , while the end-cap toroids provide the necessary field in the region  $1.6 < |\eta| < 2.7$ . The region  $1.4 < |\eta| < 1.6$  is covered by the superposition of the magnetic fields created by the barrel and end-cap toroid magnets. The spatial resolution is about 10% for high energy tracks (1 TeV), and about 3% for intermediate incoming muon energy (a few GeV).

There are different types of muon chambers used. The Monitored Drift Tubes (MDT) and the Cathode Strip Chambers (CSC) are combined for precision tracking. The MDTs cover most of the range  $|\eta| < 2.7$  and amount to 1088 chambers with 339000 channels in total. Each chamber consists of six layers of drift tubes filled with a mixture of Ar and CO<sub>2</sub> gas with a tungsten-rhenium wire in the middle at high electric potential, all mounted on a support structure. When muons travel through the drift tubes, the electrons produced in the ionization of the gas are collected in the wire, raising a current that is collected by the read-out electronics. The CSC are used in the forward region due to the smaller drift time compared to the MDT system, since the expected rate of particles is much larger than at low  $|\eta|$ . They consist of multi-wire proportional chambers that provide detection with large granularity thanks to the segmented cathodes into strips, which result in spatial resolution of 40 μm in the bending plane and 5 mm in the transverse plane. A total of 32 CSC chambers and 31000 channels are used as part of the ATLAS detector.

The layout of the precision track measurement chambers includes three cylindrical layers of MDT chambers in the barrel with the beam in the center. This sector of the detector is further divided in eight equal octants in the  $\phi$  plane. In each of the end-caps, the muon chambers are placed in concentric wheels at distances of 7.4 m, 10.8 m, 14 m, and 21.5 m. MDT chambers are used in all of them except for the region  $2 < |\eta| < 2.7$  of the innermost layer in each end-cap, where CSCs are used. A scheme of the layout of the muon detector system is shown in figure

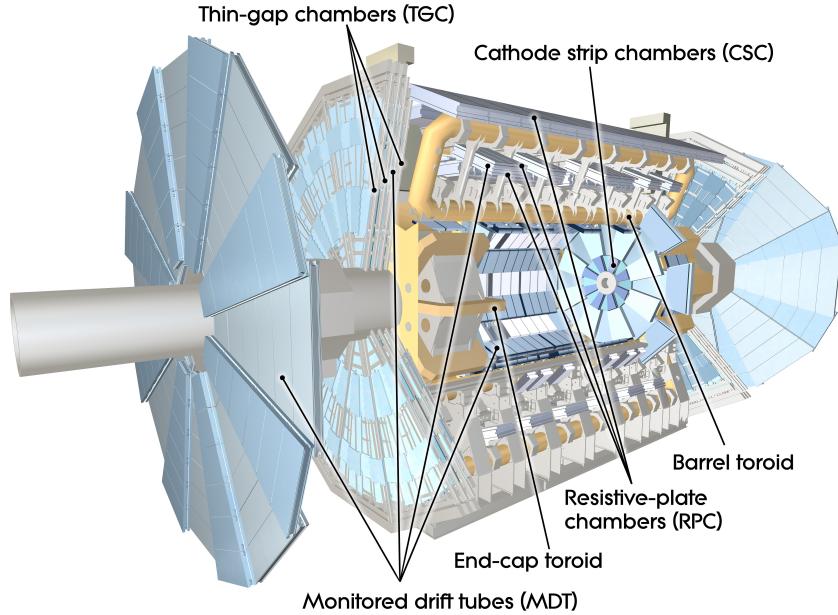


Figure 3.9: Scheme of the ATLAS muon spectrometer [51].

### 3.9.

In addition to the chambers used for precision tracking, the **Resistive Plate Chambers (RPC)** and the **Thin Gap Chambers (TGC)** are used for triggering purposes, bunch crossing identification and for providing the muon location in the direction orthogonal to the plane containing its track. They are conceived to provide much faster signals than the MDTs and CSCs. The RPCs are used in the barrel, and consist of parallel electrode-plate systems filled with gas that, after the passage of muons, produce electrons through ionization that are carried to the anode by a drift electric field. The TGCs are suited to be used in the end-cap region and consist of multi-wire proportional chambers, similarly to the CSC, but with a **wire-to-cathode** distance smaller than the **wire-to-wire** distance, which leads to a faster response. The RPC and TGC systems combined provide coverage of the region  $|\eta| < 2.4$ , as required by the trigger system, and a time resolution of about 1.5 ns.

During the LHC Long-Shutdown 2 (2019-2021), the ATLAS muon system is being upgraded in preparation for the conditions to be expected during the HL-LHC program. The main change is the replacement of the innermost end-cap layer, or Small Wheels, by the updated *New Small Wheels* (NSW) [63]. They are more resistant to degradation due to the larger incoming particle flux expected during HL-LHC runs, and largely reduce the rate of high  $p_T$  muon fakes at the trigger level. The configuration of the NSW includes eight detector planes in two multilayers in each end-cap. Each of the multilayers has four small-strip TGC (sTGC) used for triggering together with four Micromegas (MM) detector planes. A scheme of the design of the MM for the NSW is shown in figure 3.10b. The sTGCs are very similar to the TGCs but with a much smaller pitch in the strips (3.2 mm) that allows for very good spatial resolution (better than 1 mrad) as well as a fast enough response to meet the trigger requirements. A scheme of the internal structure of the sTGC is shown in figure 3.10a. The MM, on the other hand, rely on signal amplification through electron avalanche processes. They consist of a gas-filled gap where electrons are produced through ionization and drifted towards the read-out electrodes. A metallic micro mesh is placed close to the read-out electrodes at an electric potential such that the electric field between the two of them is very large, triggering the avalanche processes

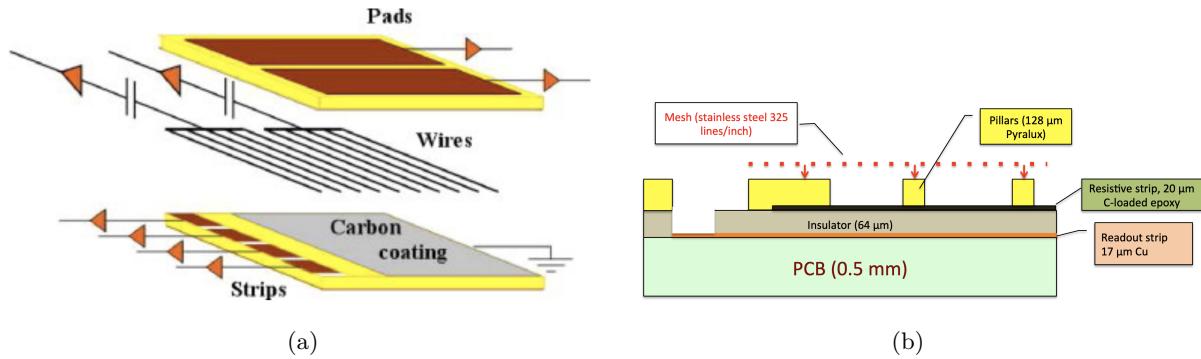


Figure 3.10: Scheme of the elements in the upgraded NSW to be installed during the LHC Long-Shutdown 2. In (a) a scheme of the sTGC is shown, together with the scheme of the MM in (b). Figures taken from ref. [63].

and amplifying the signal. In the NSW design, the read-out electrodes are immersed in an insulator medium and capacitively connected to resistance plates that collect the electrons from the avalanche, reducing the risk of sparks between the mesh and the read-out electrodes.



## Chapter 4

# The ATLAS trigger system

### 4.1 Introduction

The ATLAS experiment collects information from the proton-proton and heavy-ion collisions delivered by the LHC. As developed in the previous chapter, the detector is a complex machine with a very large number of read-out channels (around 100 million). The complete information from the detector read-out for each of the LHC collisions amounts therefore to sizeable disk memory, about 1.6 megabytes per event. Considering the 25 ns bunch spacing in the LHC during Run 2, this translates to filling about 64 TB of memory every minute, which leads to very large resource usage only for storage. In addition, not all collisions are interesting physics events, as many of them are only elastic collisions between the protons, which do not reveal much information about their internal structure.

The solution to both issues, namely recording only the physically interesting events with low memory consumption, is accomplished by a complex trigger system [64, 65]. It has a low enough latency to allow it to deal with the 40 MHz rate of LHC collisions is able to decide which events to record based on the information from the detector. This is achieved with a trigger strategy in two stages, where first the collisions are processed by the *Level-1 trigger* (L1) [66], which uses coarse granularity detector information to provide fast decisions and reduce the event rate from 40 MHz to a maximum of 100 kHz. The selected events are further processed by the *High-Level trigger* (HLT) [67] which uses the complete detector information to apply complex selections to reduce the rate of recorded events to an average of 1 kHz. Due to the low latency required for the L1 system, it is exclusively composed of fast processing hardware electronics that receive reduced information from the different detector sub-systems in ATLAS. The HLT, on the other hand, has less stringent constraints in terms of processing time, which allows for a software-based system that runs in a computing farm located at the LHC Point-1 (P1) and performs complex operations, such as object calibration or jet flavour tagging.

The trigger system in any physics experiment is of crucial importance since it is the first door to the data analysis. Any event that is not selected by the trigger is not recorded, and therefore lost for data analysis. To ensure excellent performance, the ATLAS trigger system is extensively validated and continuously monitored. A large number of selections are applied [68–70] during LHC data taking periods in order to have wide acceptance of interesting physics scenarios. These selections are typically optimized by the physics groups using Monte Carlo (MC) simulation to model the targeted physics processes. It is thanks to a rich variety of trigger selections that the ATLAS Collaboration can carry out a wide physics program, ranging from QCD processes at high energy to low  $p_T$  flavour physics.

This chapter gives an overview of the whole ATLAS trigger system, with special emphasis on the Level-1 Topological (L1Topo) system, in which I have worked developing the simulation and validation tools. The chapter is organized as follows: in section 4.2 the two stages of the ATLAS trigger system are examined, together with the path the data follows from the detectors to storage in disk and with an overview of the trigger selections applied during LHC Run 2. In section 4.3 a detailed description of the L1Topo hardware system is provided, together with the commissioning and validation of that system and its impact on physics analyses. Finally, in section 4.4 the upgrades during the LHC Long-Shutdown 2 (LS2) to the L1 system are reviewed.

## 4.2 The Trigger and Data Acquisition system in ATLAS

As mentioned above, the trigger system in ATLAS is split in two steps to allow for fast event processing together with detailed selections to maximize the suppression of uninteresting data. As shown in the scheme in figure 4.1, the ATLAS trigger and data acquisition (DAQ) systems are connected and rely on one another: the trigger needs the read-out and proper data flow of information from the detector provided by the DAQ system, and the trigger determines which events are sent further down the DAQ system all the way into storage. A period of data acquisition with the ATLAS detector is known as *run*, and each of them is identified with an unique number. Runs used for physics analyses always happen during LHC operation with stable beams, but other types of runs can be used for detector performance studies, such as runs with cosmic particles or with LHC test beams. Inside each run, every event that is accepted by the L1 and the HLT is labelled with a unique number, so each portion of recorded data can always be identified by its run and event numbers. Each run is divided in units of constant integrated luminosity called Luminosity Blocks (LB) that correspond to roughly one minute of data taking. The Trigger and Data Acquisition (TDAQ) system as a whole can be operated from the ATLAS Control Room in P1, where a new run can be started once the detector is ready.

### 4.2.1 Level-1 Trigger

The L1 trigger system is a hardware-based system with four components, making large use of Field Programmable Gate Array (FPGA) technology. The first two subsystems in the data path are the Level-1 Calorimeter (L1Calo) [73] and the Level-1 Muon (L1Muon) [74] triggers. They can access information from the ATLAS calorimeters and muon trigger chambers, respectively, to identify the interesting fragments of detector read-out, called Regions of Interest (RoI), from which object candidates commonly referred to as *trigger objects* (TOB) are formed. Additionally, both systems compute trigger decisions based on the number of TOBs above a certain energy threshold. The L1Calo system can also compute event-level quantities, such as calorimeter-based missing transverse energy ( $E_T^{\text{miss}}$ ) or total energy in the event, and provide trigger decisions based on those as well. The third subsystem is the Level-1 Topological (L1Topo) trigger [75], that receives the TOBs formed by L1Calo and L1Muon as input to compute decisions based on topological selections, such as angular separation between TOBs or invariant masses. A detail description of the L1Topo system can be found in section 4.3. The last subsystem in the L1 trigger is the Central Trigger Processor (CTP) [76], which collects the information from the other three components and issues a L1-Accept (L1A) signal, streaming down the accepted events to the HLT.

The L1 system ensemble is composed of custom high-speed electronics in order to handle the high rate of collisions at the LHC, with a rate of accepted events up to 100 kHz, representing a

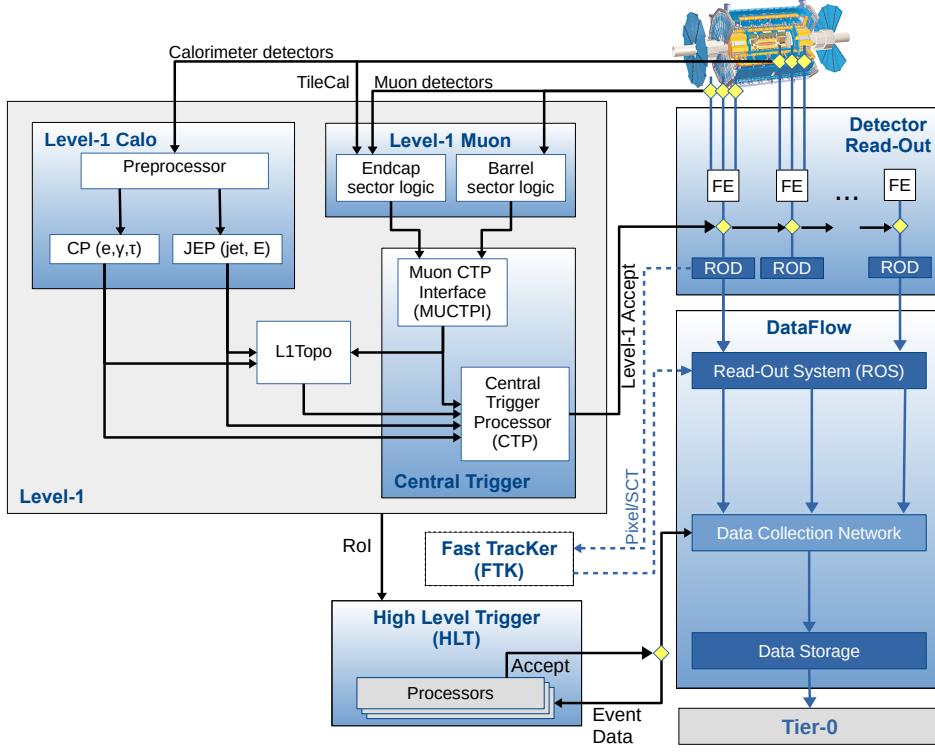


Figure 4.1: Scheme of the ATLAS Trigger and Data Acquisition system, with the data flow represented from detector read-out to storage [71]. The Fast TracKer (FTK) [72] is a hardware-based system for inner-detector track reconstruction designed to provide tracks to the HLT at the L1 accept rate. It was undergoing partial commissioning during Run 2 and was not used by the HLT for trigger decisions.

rate reduction factor of 400 over the LHC collision frequency. The total latency of the whole L1 trigger, from detector read-out to issuing the L1-accept, is  $2.5 \mu\text{s}$ . Events are stored in readout buffers while waiting for the L1 trigger decision.

### Level-1 Calorimeter trigger

The L1Calo system uses reduced granularity information, typically  $\eta \times \phi = 0.1 \times 0.1$ , from around 7200 trigger towers<sup>1</sup> from both the electromagnetic and hadronic calorimeters to identify electron, photon, tau and jet candidates. It forms the TOBs in order to compute trigger decisions based on object multiplicity above energy thresholds, and also sends them to L1Topo for more complex trigger selections. The total processing time of an event in L1Calo is inferior to  $1 \mu\text{s}$ , while the total latency of the system, including cabling, is about  $2.1 \mu\text{s}$ .

A scheme of the architecture of the L1Calo system is shown in figure 4.2a. Its three main components are the PreProcessor (PPr), that handles the analogue output of the detector read-out and digitizes it, and the Cluster Processor (CP) and the Jet/Energy-sum Processor (JEP), that take the digitized outputs of the Preprocessor and execute the object identification and

<sup>1</sup>The trigger towers are computed by summing up the energy from calorimeter cells in the detector. The number of cells used to form each tower depends on the detector granularity and goes from just a few in the end-cap and forward regions, to 60 cells per tower in the barrel sector of the ECal.

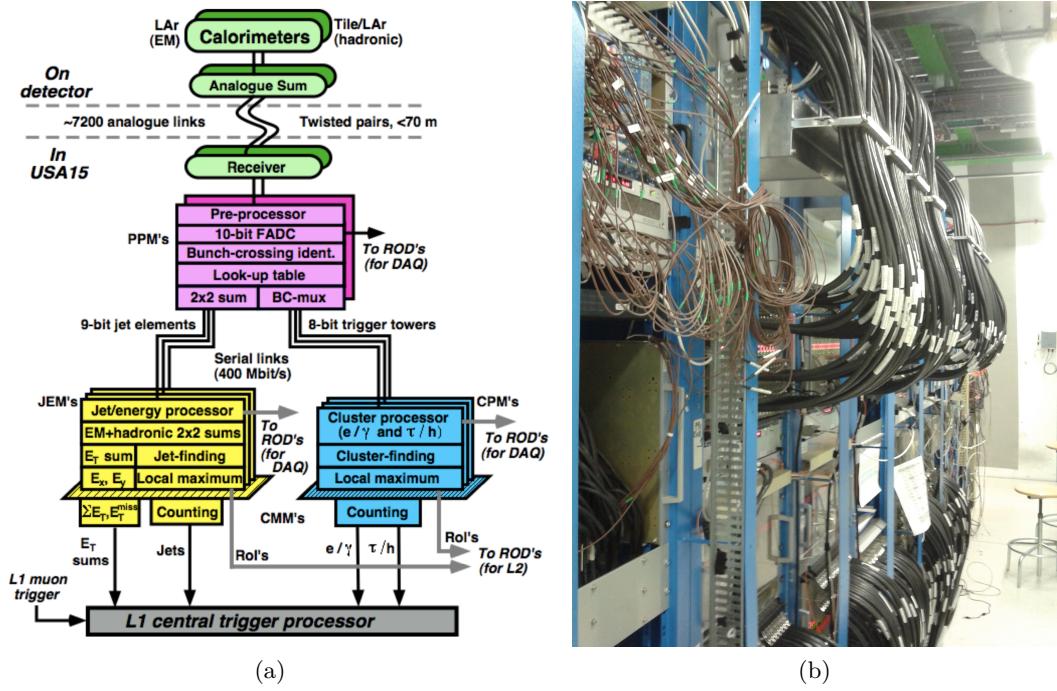


Figure 4.2: (a) Architecture of the L1Calo trigger system, with the real-time data flow represented by the solid black lines [73]. (b) Picture of the trigger electronic crates located in the separate cavern USA15 [77].

counting algorithms. The L1Calo hardware is not located on the detector but rather in a different underground cavern called USA15, that hosts all the trigger electronics in ATLAS (figure 4.2b). All the custom modules are 366 mm high and 400 mm deep, and are designed to be compact in order to reduce the latency of the system. A common module design, compatible with different functions depending on the firmware installed, is used for low production cost and simple maintenance. All the modules have voltage and temperature monitoring systems implemented to avoid extreme conditions that could hurt the hardware.

The PPr is composed of 124 custom PreProcessor Modules (PPM) split in eight crates, four of them dedicated to processing the signals from the electromagnetic calorimeters and four for the hadronic calorimeters. The inputs to the PPMs are provided by the Tower Builder Boards and sent through 616 16-way twisted-pair cables, with propagation speed of 4.76 ns/m and optimised for cross-talk between towers lower than 0.5%. Before arriving to the PPMs, the Receiver Modules apply different corrections, such as  $E_T$  calibration, to signals from both the electromagnetic and hadronic calorimeters. Each of the PPMs receives four analogue cables as input, that amount to a total of 64 trigger towers, and includes 16 Multi-Chip Modules (MCM) that process four trigger towers each. The signals are digitized to a precision of 10 bits and time corrections are applied to account for different in-flight time and cable lengths. A series of operations are performed on the signals by means of look-up tables, such as  $E_T$  tuning through linear transformations, pedestal subtraction and dead channel removal, and after which the energy resolution is fixed to 1 GeV. Look-up tables operations are used for flexibility and possibility for parallelization. Finally, the data in each PPM is serialized at 400 Mbit/s before being sent to the CP and the JEP in order to reduce the number of connections. During the LHC Long-Shutdown 1 (LS1), the MCM were replaced by the new Multi-Chip Modules (nMCM), with similar functionalities but updated components that allows for improved performance. This

allowed to overcome one of the major limitations experienced during Run 1, related to the high missing energy trigger rates due to the liquid argon calorimeter pulse shapes. By allowing for more flexible signal processing in the nMCM, it is possible to implement dynamical pedestal corrections that lead to a large reduction in the rate for such triggers.

Both CP and JEP carry out similar tasks and have similar functionalities, which is reflected in the similar design of their modules. In both cases a sliding-window [78] algorithm on the input trigger towers is used for object identification. For both CP and JEP, the detector is divided in four  $\phi$  quadrants and split among the available crates/modules. In general, to avoid effects coming from the finite coverage in  $\eta, \phi$  of each module, large amounts of data are duplicated and made available where needed. The CP system is composed of 4 crates with 14 Cluster Processor Modules (CPM) each, and offers coverage of the region  $|\eta| < 2.5$ . Each CPM included 8 Xilinx XCV-1000E FPGAs for algorithm execution. They execute two different algorithms to identify electrons/photons and taus, respectively, and build the TOBs that will be used for multiplicity counting and streamed down to the L1Topo system. For electron/photon identification, the algorithm looks for narrow and energetic showers in the electromagnetic calorimeter that do not reach the hadronic calorimeter, with additional transverse isolation requirements for background reduction. In the case of tau identification, the algorithm searches for collimated clusters of hadrons as a result of the hadronic tau lepton decay, and with further isolation requirements if needed. The results from each CPM are the sum of the multiplicity counts from each of its FPGAs. They are limited to 3 bits for each energy threshold configured, which means they are saturated if the counts exceed 7 ( $2^3 - 1$ ).

The JEP system consists of 2 crates with 16 JET/Energy Modules (JEM) each, and covers the region  $|\eta| < 3.2$  for jet identification. Each JEM has a total of 6 FPGAs, 4 for input processing and de-serialization, one for jet identification and TOB formation, and another one for total energy computation. For energy sum and missing transverse energy computation it also accesses information from the FCal system, with coverage of the full  $|\eta| < 4.9$  range for accurate calculation. Neither the jet identification or energy summation algorithms need information with granularity as fine as the CP algorithms, and so the granularity used by the JEP is coarser, of  $\eta \times \phi = 0.2 \times 0.2$ , in order to reduce resource consumption and execution time. The jet identification algorithm is very similar to the one used for tau identification, also based on a sliding-window approach over the available trigger towers for each module/crate. The energy summation algorithm computes the total transverse energy in the event, together with  $E_T^{\text{miss}}$ , by summing independently the components in each direction,  $E_x$  and  $E_y$ , over the region covered by each module.

The outputs from both CP and JEP are sent to the Extended Common Merger Modules<sup>2</sup> (CMX), where the multiplicity of each object and the total energy and  $E_T^{\text{miss}}$  are summed up for the whole  $\eta$  coverage for each system. The outputs from the CMX are sent to the CTP in the form of bits that encode the multiplicity of each object above each of the energy thresholds configured, together with one bit for each total energy/ $E_T^{\text{miss}}$  threshold configured, which is set to 1 if the computed quantity is above such value. In addition, they also stream the information of the reconstructed TOBs to the L1Topo system.

---

<sup>2</sup>During Run 1 the outputs from CP and JEP were merged using Common Merger Modules (CMM), which were upgraded to the CMX in order to allow for larger processing power and bandwidth to stream data to the L1Topo system.

## Level-1 Muon trigger

The L1Muon system aims at providing estimates of muon transverse momentum ( $p_T$ ) by comparing the trajectories of muon candidates with the expected detector hits from a muon with infinite momentum. During LHC Run 2 it was able to handle six configurable momentum thresholds and compute the muon multiplicity for each of them. It provides coverage of the barrel and the end-cap sectors through separated high-speed electronics that process detector information from the RPCs and TGCs, respectively. The geometrical coverage in the end-cap sector is almost total. In the barrel, however, due to ATLAS support structures, there is a gap at  $\eta = 0$  that reduced the acceptance originally down to 80% of the total geometry of the detector. During the LS1 a fourth layer of RPC was included to provide increased sensitivity in that region, that resulted in 20% increased acceptance compared to LHC Run 1. The total latency of the L1Muon system, from detector read-out to CTP, is about 2  $\mu$ s.

The L1Muon trigger has both on-detector and off-detector electronics that carry out different functionalities. The on-detector modules read the information from the RPCs and TGCs to form the muon candidates and compute the thresholds over limited sectors of the detector, while the counting of objects is handled by the off-detector electronics to avoid double counting. In the barrel, the information from the RPCs is processed locally by the so-called Splitter Boards and Processor Box (PAD), with two separated algorithms for low and high momentum muons. The results are sent to the corresponding Sector Logic modules, where they are combined with the trigger information over the whole geometry of the detector. The reconstruction algorithm requires hit coincidence in 3 concentric RPC stations for high transverse momentum muon, and hits in only 2 stations for low energy muons. Both algorithms are executed simultaneously in different PADs. The Barrel Sector Logic (BSL) [79] modules are located outside the ATLAS main cavern in 16 crates. Each BSL module contains two VIRTEX II Xilinx FPGAs, and up to four mezzanine receivers and one serializer for data flow. The BSL electronics receive the information processed by the on-detector electronics and compute the trigger decisions on muon multiplicity above the configured energy thresholds.

Similarly, information from the end-cap sector is first processed by the on-detector electronics and the results are sent to the End-cap Sector Logic (ESL) modules [80]. The muon reconstruction algorithm requires coincidence of two TGC stations for low momentum muons, and an extra hit in a third TGC station compatible with the previous two for high momentum muons. Additionally, overlap checks are performed with the TileCal detector for background suppression. The ESL modules use a look-up table method to compute charge and momentum for its high speed and flexibility. SRAM-embedded type FPGAs are implemented due to their capability to store large look-up tables in the same device without needing extra chips, thus reducing the wiring in the board, the complexity and the latency. The final ESL module design includes two Xilinx XCV405E FPGAs and one Xilinx XCV400E FPGA to carry out all decoding and look-up table operations.

The results from both the BSL and ESL are sent to the CTP for the final L1A. They are collected and interfaced by an additional system named Muon-to-CTP-Interface (MuCTPI) [81]. It consists of 16 MIOCT cards that span over the whole geometry of the detector, with each card having 2 electrical outputs and designed to operate at the same rate than the LHC collisions, 40 MHz. Both the hardware and the firmware of the MuCTPI were upgraded during the LS1 to cope with the larger speed and bandwidth needed to include the L1Topo system. During Run 2 it could send information at an aggregated rate of 10.24 Gb/s. An extra interface, called MuCTPI-to-Topo [81], is included to transform the electrical outputs from the MuCTPI into optical inputs for the L1Topo system. It consists of a Xilinx VC707 FPGA and two mezzanine

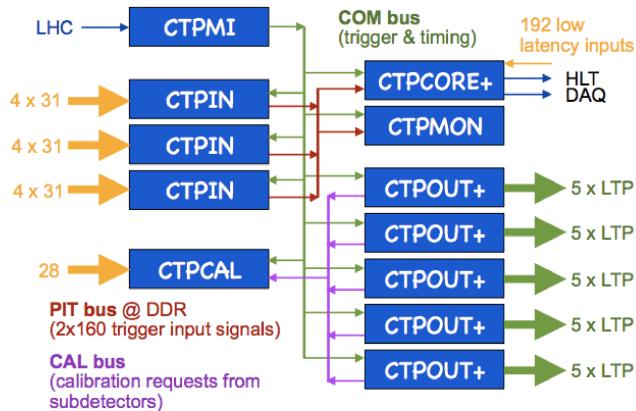


Figure 4.3: Scheme of the connections between CTP modules, with the number of bits sent by each connection highlighted where relevant [82].

cards, and the total latency from input to output during Run 2 was measured to be 85 ns.

## CTP

The outputs from the L1Calo, L1Muon and L1Topo systems are collected by the CTP in order to issue a final L1A, which determines if the detector information related to the collision under study gets streamed down further to the HLT, together with timing signals that identify the piece of sub-detector read-out being analysed. The L1A is based on a list of configurable selections that are defined before the data taking run begins. Additionally, the CTP also provides trigger summary and monitoring information.

The CTP performs these tasks by means of different custom-built modules, which are shown in the scheme of figure 4.3 as installed in Run-2, and are briefly described below:

- CTP Machine Interface (CTPMI): it handles the timing signals from the LHC and shares them with the rest of the modules in the CTP.
- CTP Input Modules (CTPIN): they take the inputs from L1Calo and L1Muon and format them to be processed for the L1A signal. There are three CTPIN modules in total, and each of them can receive up to 124 inputs on a total of four cables.
- Updated CTP Core Module (CTPCORE+): it produces the L1A signal with the inputs from the other L1 systems. It was re-designed during the LS1, and includes two Virtex 7 FPGAs for enhanced processing power, larger number of accepted inputs to accommodate the inclusion of the L1Topo system and double the output trigger bits compared to Run 1 (512 in Run 2 vs 256 in Run 1). The inputs from L1Topo are received directly through low-latency cables, summing up to 192 bits, to account for the extra processing time needed by said systems before sending their outputs to the CTP.
- Upgraded CTP Out Modules (CTPOUT+): they handle the distribution of the L1A and timing signals to the sub-detectors for read-out, and was upgraded during the LS1 for enhanced processing power for improved diagnostics and monitoring features.
- CPT Monitoring module (CTPMON): it monitors per-bunch trigger signals for correct functioning.

- CTP Calibration module (CPTCAL): it handles calibration requests from sub-detectors.

The L1A signal is formed combining logically the 512 available output bits, using look-up tables to select object multiplicities and perform logical OR operations. The latency of the CTP system during Run 2 , from CTPIN to the CTPOUT+ modules, was of 6 bunch-crossings (150 ns), within the latency budget for the CTP of 8 bunch-crossings.

### 4.2.2 High-level trigger

After an event has been accepted by the L1 trigger, it is streamed down to the HLT for further processing. The HLT is a software-based system that performs the final decision for event data storage. The HLT software is based on the ATLAS offline software Athena, which allows to perform analysis-like selections at the trigger level. It is executed on a computing farm located at LHC P1 that includes about 40000 instances of the selection application named Processing Units (PUs), which are optimized to perform decisions with a latency of  $\mathcal{O}(100)$  ms. The events are processed in a sequential manner while the different reconstruction and selection algorithms for a given event are executed in parallel to reduce the total computing time needed for each collision.

In general, after the L1 trigger issues the L1A signal a set of fast trigger algorithms are executed in the HLT for early rejection, after which dedicated CPU-intensive reconstruction algorithms are executed for the full evaluation of the configured trigger selections. The reconstruction algorithms in the HLT use more precise detector information than the L1 trigger, and is provided by specific requests to the sub-detector read-out systems. The reconstruction algorithms for specific objects, such as electrons, photons or jets, typically use RoIs, while event-level objects, such as missing transverse energy, need complete detector information. After the different object reconstruction algorithms have been executed, a decision for each trigger selection is computed based on the features of the event.

The HLT reduces the rate of L1-accepted events from 100 kHz to an average of about 1 kHz. The average rate of events recorded during ATLAS runs in 2018 is shown in figure 4.4a, and amounts to 1.2 kHz in total, with the average data bandwidth stored being 1.2 GB/s. This trigger rate needs to be split among all selections configured, including those used for calibration and debugging purposes. During Run 2, the recorded data was efficiently organized according to its purpose through the *stream model*, that steered the event-data information into different files depending on the selections each event satisfies. Five types of streams were included during the data-taking period 2015-2018, and are described below:

- Physics streams, that collect events for offline data analysis. This type of stream dominates over the rest in terms of stored data, with multiple dedicated instances for the different physics analysis groups in ATLAS. The ‘standard’ physics stream that collects the events recorded by the most general trigger selections is called Physics Main, and is the one used by analysis in the collaboration. Dedicated streams for high-rate or complex topology triggers are included as well, such as streams for B-physics or tau analyses.
- The express stream, that contains a small subset of events from physics streams used for real-time reconstruction in order to have fast data quality information and monitoring.
- Debug streams, that store events that are neither accepted or rejected by the trigger system in order to identify potential problems with either the trigger or the data acquisition systems.

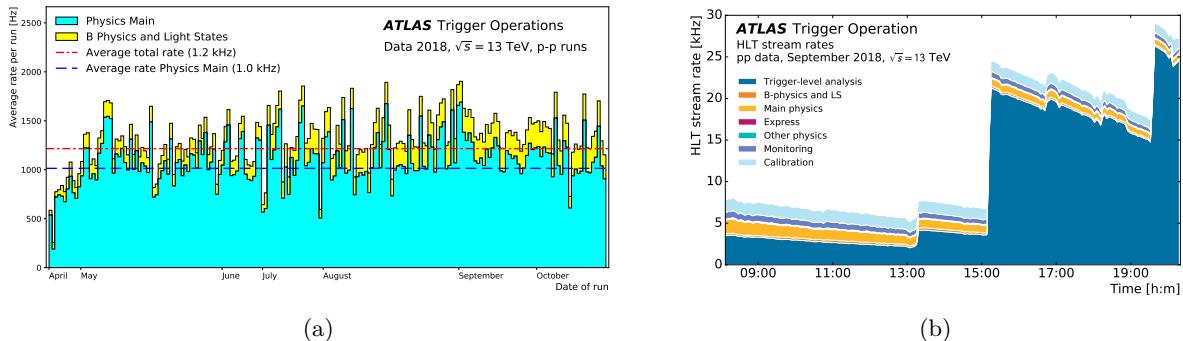


Figure 4.4: ATLAS trigger rates in 2018, with the average over the whole year shown in (a) and the trigger rates shown as a function of time for one run shown in (b), where the rates of different streams are displayed explicitly. Figures taken from ref. [65].

- Calibration streams, used for dedicated sub-detector or HLT calibration. In the case of sub-detector calibration, only a fraction of the detector information is generally recorded to reduce memory consumption, using the so-called Partial Event Building strategy [64].
- Trigger-Level Analysis streams, containing only reduced detector information and reconstructed TOBs for physics analysis at the trigger stage. It is a low-level approach to physics analysis that can reduce bias introduced by the trigger selections and reduce the loss of interesting events, but that generally lacks precise object calibration that cannot be applied online.

Events may end up in different streams, and therefore the sum of rates of two streams is typically higher than the rate of a stream that combines the trigger selections in the previous two. Similarly to the physics streams, there can be multiple streams of each type except for the express stream. Figure 4.4b shows the trigger rates for different streams as a function of time for a particular ATLAS run. An extra type of streams, the monitoring streams, are essential to steer events into the online monitoring tools available at P1. These streams, however, are not recorded.

The HLT computing farm is updated every year in LHC data-taking periods, replacing the older hardware for more recent systems to increase the processing power and speed. The software deployed there is under constant monitoring for excellent performance, with a thorough 2-level revision system in place to avoid the introduction of problems with any modification of the code. The whole framework is built and tested on a daily basis, and only selected stable versions of the software get built into a release that is deployed in P1 after extensive validation.

### 4.2.3 Read-out data path

The ATLAS trigger system determines the events to be sent to storage. The path followed by the data from read-out to storage is therefore interconnected with the real-time data path followed during trigger processing, but it is not identical, as shown in figure 4.1. From the moment a collision happens in the LHC, the detector information is extracted by the corresponding Front-End (FE) electronics implemented for the trackers, calorimeters and muon chambers and stored in data buffers to wait for the L1A signal issued by the CTP. The size of these buffers determines the maximum latency allowed for the L1 trigger, showing a compromise between the amount of

data that can be held and the complexity of the trigger algorithms executed. During Run 2 the buffers could store up to  $2.5\ \mu\text{s}$  of detector read-out, which corresponds to 100 bunch crossings. The timing signals from the LHC play an essential role in bunch crossing identification, and are handled by the Timing, Trigger and Control (TTC) network.

To ensure proper detector read-out functionality and avoid data loss due to partial data buffering, a dead time is implemented at the CTP. It consists of a trigger veto that limits the number of events accepted at the L1. This veto is implemented in two flavours: simple dead time, that automatically rejects a fixed number of bunch crossings after an event has been accepted by the L1 (4 bunch-crossings at the end of Run 2), and complex dead time, where only a determined number of L1A are allowed over a certain period of time. The application of dead time avoids the overlap between the read-out windows of accepted events and prevents overflow in the FE read-out buffers.

Once an event is accepted by the L1 trigger the data collected by the FE electronics and stored in buffers is sent to the Read-Out Drivers (ROD) for formatting and processing. The data is then sent to the Read-Out System (ROS), which collects the data while waiting for a trigger decision from the HLT. These buffers can store data for longer than those used for the L1 trigger as the HLT requires more time to process events. The HLT then requests the detector information from the ROS for processing the final trigger decision, and, if the event is accepted, it sends a signal to the Sub-Farm Output (SFO) to stream the data to the CERN Tier-0 facilities for offline processing and permanent storage.

#### 4.2.4 Trigger menu

As mentioned in the discussion above, the trigger system applies a series of selections configured prior to data taking. These selections, commonly referred to as *trigger chains*, are typically composed of one or more requirements applied at the L1 trigger combined with reconstruction and decision algorithms executed in the HLT. Each of them is related to one type of object (e.g. single lepton triggers) or to an event-level quantity (such as missing momentum triggers), or could be complex combinations involving more than one type of object (e.g. simultaneous selections on  $b$ -jet multiplicity and  $E_{\text{T}}^{\text{miss}}$ ). For a given event, the result of each of the steps executed in a trigger chain is cached, together with the partial detector information used, in order to reduce the time needed to provide a final decision for each event. The set of all trigger chains in an ATLAS run is called *trigger menu*, and it is of great importance as it defines the regions of the phase space that will be available for offline analysis.

Each of the trigger chains configured in the menu include a prescale with value  $n \geq 1$ . The prescales are used to control the trigger rate of a given chain by randomly recording only 1 in  $n$  accepted events. They are a very useful way to record events with high trigger rate signatures, such as low momentum muons, without introducing any bias in the relevant kinematic distributions such as a high- $p_{\text{T}}$  threshold. The trigger prescales can be changed mid-run for luminosity-dependent trigger configurations in order to optimize the resources used. The map linking prescale values for each trigger chain with the corresponding LB are stored into tables and distributed over all the CPUs in the HLT farm.

The trigger menu is carefully designed to fulfil all trigger requirements in terms of latency, resource usage and rate at the same time than accepting large portions of the available phase space as required by the physics analyses. The menu for Run 2 was designed to keep unprescaled the single-lepton triggers with transverse momentum thresholds around 25 GeV, optimized for accepting a large portion of events with leptonically decaying  $W$  or  $Z$  bosons. In addition, it

covers all relevant signatures for physics analysis of Standard Model precision measurements to searches for BSM physics. Different trigger menus are designed for the different types of collisions in the LHC, taking into account the topologies and trigger rates expected from both proton-proton and heavy ion interactions. Additionally, specific menus are designed for special ATLAS runs, such as the ones used for commissioning, validation or calibration of the detector.

The trigger menu also evolves depending on the expected LHC conditions for each year in a data-taking period. In Run 2 the delivered instantaneous luminosity increased each year, which was reflected in the modifications in the menu,<sup>3</sup> typically leading to tighter kinematic selections with each passing year. An exception is the change from 2017 to 2018, where the increase in luminosity took place without an increment in the average number of interactions. This resulted in reduced CPU usage during 2018 that allowed the menu from 2017 to stay within the allocated rates/resources during 2018 as well. For this reason, few modifications were needed to design the menu for the last year of Run 2, with most of them aimed at improving the phase space coverage with respect to the previous year.

### 4.3 Level-1 Topological trigger

The L1Topo system was implemented and commissioned in 2016, after the first year of data taking in Run 2, and allows to perform complex topological trigger selections at the L1 at low latency thanks to custom electronics based on FPGA technology. Its implementation in the ATLAS trigger system required the additions to the other L1 systems mentioned in subsection 4.2.1, namely the use of nMCM in the PreProcessor and the switch to the CMX for output merging in L1Calo, as well as the inclusion of the new MuCTPI-to-Topo interface at the end of the L1Muon system.

The L1Topo board take as input the TOBs formed by L1Calo and L1Muon and combine them in order to compute decisions about configurable selections. This was achieved by means of two different types of algorithms: sort/select algorithms, that handle the input data, and decision algorithms, where a hypothesis is contrasted against data and a trigger decision bit is formed. The sort/select algorithms receive the inputs from the L1Calo and L1Muon systems and perform two tasks over the data. First, every TOB is converted into a common granularity of  $\eta \times \phi = 0.1 \times 0.1$  to allow for easy computation of kinematic quantities that combine objects extracted from sub-detector read-outs of different granularity. Many of the computations carried out by L1Topo are performed by means of look-up tables, and having unified granularity allows for object-independent tables, thus reducing the necessary space to store them. The inputs are then grouped into TOB-specific lists, where they can either be ordered in decreasing transverse momentum, or selected if their momentum is higher than certain configurable energy thresholds. The size of these lists is fixed in the firmware, and each type can contain either 6 or 10 TOBs for sorted or selected lists, respectively. In the case of global event quantities, such as total or missing energy, only one TOB of each type is computed per event.

The decision algorithms use the lists produced by the sort/select algorithms as inputs, apply configurable  $p_T/\eta$  selections, perform a series of bitwise operations with them and compare the results to user configurable parameters in order to provide a decision. Each type of operation is executed through a specific algorithm, and in general each algorithm can take input TOBs from either L1Calo or L1Muon, or combine both into one single selection. Some examples of trigger selections applied during Run 2 include requirements on quantities computed from pairs of TOBs, such as invariant or transverse masses or angular separations, or requirements on event

---

<sup>3</sup>For detailed discussions of the trigger menu per year in Run 2 see refs. [68–70].

Algorithm	Definition
Pseudorapidity distance: $x_1 \leq \Delta\eta =  \eta_1 - \eta_2  \leq x_2$	Pseudorapidity difference between the objects of two given input lists. The lower and upper bounds ( $x_1, x_2$ ) are configurable parameters of the algorithm.
Azimuthal distance: $y_1 \leq \Delta\phi =  \phi_1 - \phi_2  \leq y_2$	Azimuthal distance between the objects of two given input lists. The lower and upper bounds ( $y_1, y_2$ ) are configurable parameters of the algorithm.
Box cuts: $x_1 \leq \Delta\eta \leq x_2$ and $y_1 \leq \Delta\phi \leq y_2$	Combination of $\Delta\eta$ and $\Delta\phi$ requirements with configurable lower and upper bounds ( $x_1, y_1, x_2, y_2$ ). The same pair of objects must satisfy both conditions simultaneously to be selected by the algorithm.
Window cuts: $x_1 < \eta < x_2$ and $y_1 < \phi < y_2$	Requirement on the absolute pseudorapidity and azimuthal positions of the objects given in an input list. The lower and upper bounds ( $x_1, y_1, x_2, y_2$ ) are configurable parameters of the algorithm.
Angular distance: $x_1^2 \leq \Delta R^2 = (\Delta\phi)^2 + (\Delta\eta)^2 \leq x_2^2$	Angular distance between the objects of two given input lists. The lower and upper bounds ( $x_1, x_2$ ) are configurable parameters of the algorithm.
Disambiguation: $\phi_1 \neq \phi_2$ OR $\eta_1 \neq \eta_2$ , $\Delta R > x$	Takes two or three input lists and checks the spatial separation between each pair, accepting an event if none of the input objects in one list overlaps with the objects in the other(s).
Ratio: $O_1 \geq x O_2$	Takes two input lists and for two objects in the same location ( $\phi_1 = \phi_2$ AND $\eta_1 = \eta_2$ ) the ratio of two quantities is required to be above a threshold. The threshold ( $x$ ) and the two quantities ( $O_1, O_2$ ) are configurable parameters of the algorithm.
Invariant mass: $x_1^2 \leq m_{\text{inv}}^2 = 2E_{\text{T},1}E_{\text{T},2}(\cosh \Delta\eta - \cos \Delta\phi) \leq x_2^2$	Invariant mass of each pair of objects in two input lists. If the two input lists are different, the two objects used in the mass computation must satisfy the disambiguation criteria. The lower and upper bounds ( $x_1, x_2$ ) are configurable parameters of the algorithm.
Transverse mass: $x_1^2 \leq m_{\text{T}}^2 = 2E_{\text{T},1}E_{\text{T}}^{\text{miss}}(1 - \cos \Delta\phi) \leq x_2^2$	Transverse mass of the $E_{\text{T}}^{\text{miss}}$ and the objects in one input list. The lower and upper bounds ( $x_1, x_2$ ) are configurable parameters of the algorithm.
Event hardness: $H_{\text{T}} > x$	Scalar sum of the transverse momenta of all or a selected number of jets in the event. Additional requirements on the jet $p_{\text{T}}$ or $\eta$ are possible. The threshold ( $x$ ) is a configurable parameter of the algorithm.
Simple cone: $\sum_{\Delta R < 1.0} E_{\text{T}} > x$	Sums the energies of sliding-window jet TOBs with $E_{\text{T}} \geq 15$ GeV and centre within a cone of radius $\Delta R = 1.0$ around each jet. The threshold ( $x$ ) for the total energy sum is a parameter of each algorithm.
Late muon	Finds the highest- $p_{\text{T}}$ muon in the next bunch crossing and combines it with the input lists associated with the current bunch crossing. This results in a tighter latency budget for this algorithm.

Table 4.1: Description of all available L1Topo decision algorithms during Run 2 together with the different parameters that can be configured through the menu for each of them. Taken from ref. [83].

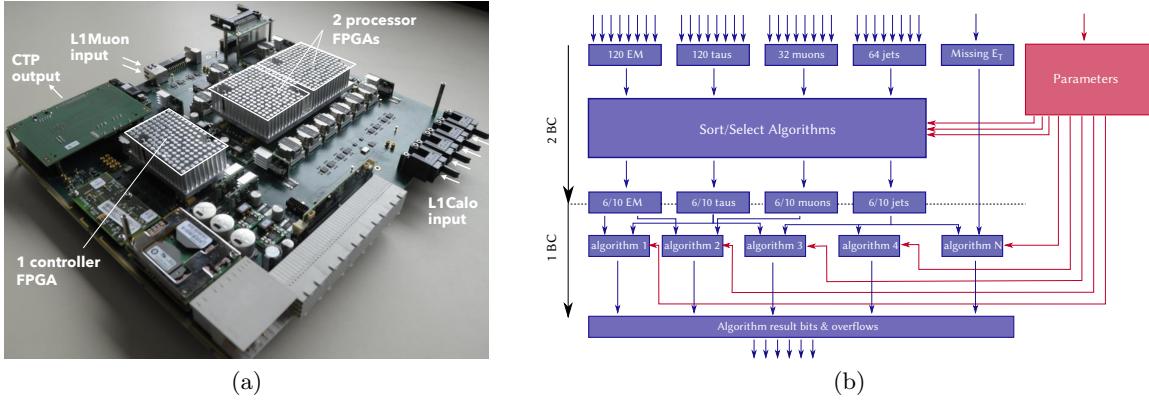


Figure 4.5: (a) Picture of a L1Topo module during Run 2 with its main components and I/O paths highlighted. (b) Scheme of the data-flow and algorithms executed in the processor FPGAs in L1Topo. Figures taken from ref. [83].

level variables such as total hadronic energy. Additionally, L1Topo can access information from neighbouring bunch crossings. This feature was used during Run 2 to trigger on delayed muons originating in the decay from potential long-lived particles where the products would be detected simultaneously to the next collision. A complete description of the L1Topo decision algorithms implemented during Run 2 is shown in table 4.1. Each selection applied in L1Topo is referred to as *trigger item* and is configured through the menu, in terms of TOBs used for the computation and parameters used in the hypothesis formation. The L1Topo system could provide the CTP with up to 128 trigger decision bits during Run 2, out of which 113 were in use.

### 4.3.1 Description of the hardware

During Run 2, the L1Topo system consisted of two AdvancedTCA-compliant hardware boards housed in a single processor crate. Each of the modules contains two Xilinx Virtex 7 FPGAs to process algorithms, and one Kintex 7 FPGA for data formatting and input-output (I/O) control. Figure 4.5a shows a picture of one of the modules with its main components highlighted. The inputs from L1Calo are received through optical Fibre cables in the back-plane of the module, while the inputs from L1Muon arrive through optical fibres connected to the front panel, both at a speed of 6.4 Gb/s. The maximum number of TOBs that can be received are 64 jets, 120  $e/\gamma$  candidates, 120  $\tau$ s, 32 muons and 1  $E_T^{\text{miss}}$  object. Each of the TOBs encode the position in the  $\eta - \phi$  plane and the  $p_{\text{T}}\text{as}$  bit arrays, with a granularity of  $\eta \times \phi = 0.1 \times 0.1$  and  $p_{\text{T}} = 1 \text{ GeV}$ . Upon arrival, the optical signals are converted into electrical ones and sent to the processor FPGAs where they are de-serialized. High-speed inter-FPGA communication is available inside each module, however during Run 2 it was not used as the input data was replicated in each of the two processor FPGAs.

The latency of the L1Topo system is about 200 ns, which during LHC Run 2 corresponded to a spacing of 8 bunch crossings. Out these 200 ns, around 50 ns correspond to data transmission from input sources to L1Topo and from L1Topo to the CTP. The remaining time is split between data reception and de-serialization (50 ns), data synchronization in FPGAs (25 ns) and algorithm processing (75 ns). This last step is divided in sort/select algorithms to format and reduce the size of the inputs, which takes 50 ns, and decision algorithms that produce the bits sent to the CTP in 25 ns. A scheme of the algorithm processing workflow is shown in figure 4.5b, with the time needed to execute each step highlighted on the left.

The outputs from L1Topo consist of two bits per trigger item, one that encodes the decision from the dedicated topological algorithm and another one to store overflow information. The overflow bit for a certain trigger item is set if the corresponding inputs come from a saturated module from either L1Calo or L1Muon.<sup>4</sup> Additionally, if the select algorithms receive more TOBs with identical  $p_T$  than can be stored in a list, an overflow signal is also issued. The overflow bits are propagated into the algorithm processing, and a logical OR of the decision and overflow bits for each trigger item is sent to the CTP for the final L1A. The outputs from L1Topo are transmitted through low-latency electrical cables, and a maximum of 32 output bits can be sent per FPGA, adding up to 128 in total.

### 4.3.2 Commissioning and validation

Since its deployment in P1, the L1Topo hardware has undergone an extensive commissioning and validation process to ensure excellent performance. Before its installation, several checks were performed in laboratory conditions on the communication logic at the hardware level. Once deployed in P1, more tests were performed with the complete system integrated in the real data-path with the rest of the ATLAS trigger. These tests include latency studies and time alignment between the L1Topo outputs and the triggered events.

The topological algorithms were studied in depth using different techniques in order to have a complete understanding of their behaviour prior to data-taking. The firmware installed in the modules was inspected by means of standalone VHDL<sup>5</sup> simulation using well defined inputs to reproduce each of the configured algorithms, through which the decision outputs can be examined in detail. Additionally, the outputs produced by the L1Topo hardware system were studied using *hot tower* patterns as inputs. These hot tower signals are generated in the L1Calo electronics by replacing the usual pedestal correction with alternative calibration for selected channels, that leads to high- $p_T$  TOBs in predefined positions. The decisions produced by the firmware could then be crosschecked with the expected outcome, in particular for algorithms relying on angular selections, since the inputs are well defined.

A very useful tool on the validation process of the topological algorithms was an online simulation of the L1Topo functionalities based on Athena and integrated in the HLT. During Run 2, it would process every L1-accepted event in real time, performing bitwise operations to produce simulated decisions for each of the configured trigger items. The simulated decision bits could then be compared to those produced by the firmware, with the advantage that a software-based simulation provides over the firmware in terms of flexibility and possibility of examining intermediate steps of the computation.<sup>6</sup> This online simulation can be executed offline in standalone mode as well using the HLT-accepted events as input for detailed studies. It played a key role in the continuous validation strategy implemented during the commissioning stage, and was kept running online for the rest of Run 2 for monitoring purposes. Additionally, during data-taking periods, a special debug data stream was included to store events that showed problems related to L1Topo, such as timing issues in the TTC.<sup>7</sup>

---

<sup>4</sup>Overflow on L1Calo inputs are set when there are more than 5 jets in the same JEP CMX or 4 clusters in the same CP CMX, while an overflow on inputs from L1Muon are set if more than 2 muons in the same octant.

<sup>5</sup>Very High Speed Integrated Circuit Hardware Description Language (VHSIC-HDL), that is usually shortened as VHDL.

<sup>6</sup>The only stored information from the L1Topo hardware system are the decision and overflow bits, while all input information is deleted after each computation, making it difficult to understand if a computed decision is correct

<sup>7</sup>These were typically observed during TTC initialization or whenever the TTC system needed to be restarted. The rate of events in this debug stream with stable detector conditions was very low during 2018.

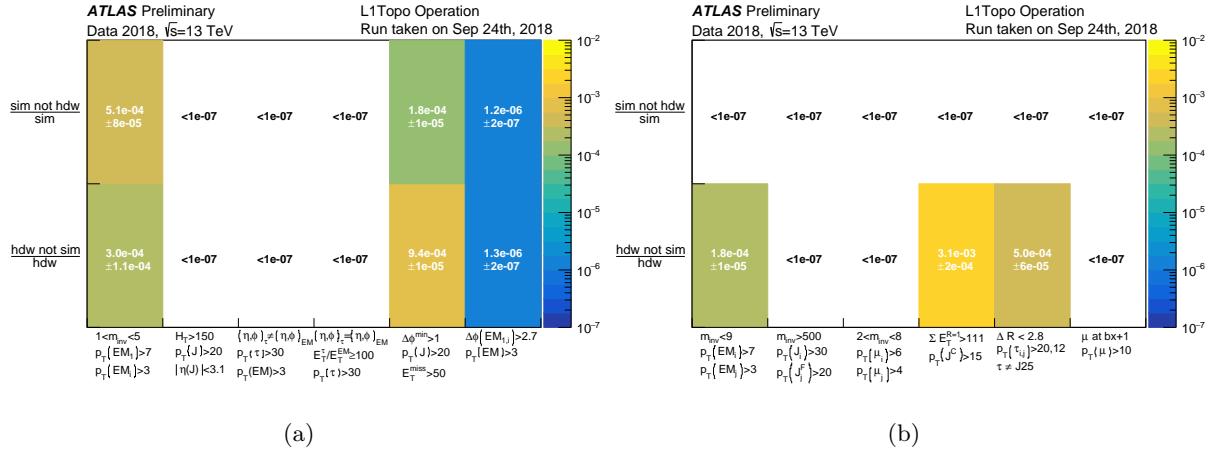


Figure 4.6: Rate of disagreement between simulated and real trigger decisions for few main trigger items used in physics analyses during 2018. The mismatch rates are presented in the form of simulation false positives (top row) and false negatives (bottom row). The data used for these figures corresponds to a proton-proton collision run taken on the 24th of September [84].

The usual comparison between firmware and simulation results is shown in the form of mismatch rates for each configured trigger item and is stored for offline reprocessing. An example of the mismatch rates for selected topological algorithms in a run taken during 2018 is shown in figures 4.6a and 4.6b. These mismatch rates are also shown in P1 during data-taking for a fixed number of LBs in order to have accurate time-dependent monitoring and spot problems that may appear mid-run.

In the last year of Run 2 both the hardware and simulation results were inspected thoroughly with different types of runs and under a number of conditions. An example of the powerful validation strategy followed is one of the issues found and solved in spring 2018, before data-taking started, that was related to a hardware cable swap in one of the L1topo modules. This problem manifested as four malfunctioning invariant mass trigger items combining baseline jets ( $|\eta| < 3.1$ ) with jets in all detector range ( $|\eta| < 4.9$ ),<sup>8</sup> with large mismatch rates between firmware and simulation. During hot tower runs the issue was identified as a problem in the hardware and not the algorithm structure, since similar items using the same topological algorithm with different configured parameters were working as expected. Additionally, for hot towers in the barrel sector, where all input jets are expected to have  $|\eta| < 3.1$ , the trigger rates for the problematic items should have been identical to those of other invariant mass items with the same thresholds but using exclusively jets over all detector range.<sup>9</sup> The hardware was then thoroughly inspected and the cable swap was identified and fixed by re-mapping them inside the firmware.

This approach also helped identify configuration issues in the menu during 2018. An example is one of the angular selections, providing trigger decisions based on  $\Delta\eta$  for the two leading forward jets ( $1.5 < |\eta| < 3.1$ ) in the event. The mismatch rates for an example run in May 2018 with proton-proton collisions is shown in figure 4.7. The differences between simulation

<sup>8</sup>The affected trigger items were 200INV9999-J30s6-AJ20s6, 400INV9999-J30s6-AJ20s6, 500INV9999-J30s6-AJ20s6, 600INV9999-J30s6-AJ20s6, where XINV9999 means applying a lower threshold X on the invariant mass, JXs6 is the first input list made of the 6 leading baseline jets with  $p_T > X$  GeV, and AJXs6 is the second input list made of the 6 leading jets over all detector geometry with  $p_T > X$  GeV.

<sup>9</sup>Namely 400INV9999-AJ30s6-AJ20s6 and 500INV9999-AJ30s6-AJ20s6.

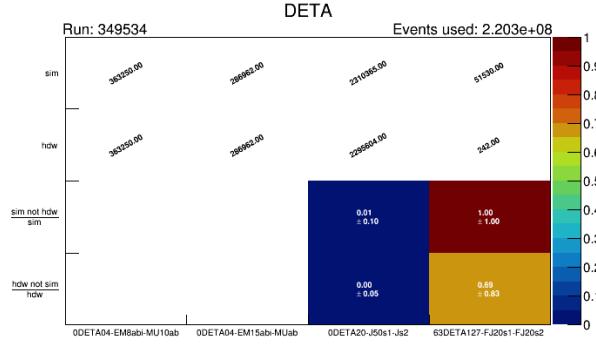


Figure 4.7: Rate of disagreement between simulated and real trigger decisions for triggers selections based on  $\Delta\eta$  of jets, with large disagreement found for one of them due to a misconfiguration in the menu.

and firmware were found to be due to the  $\eta$  threshold for selected jets not being properly configured in the menu, which translated in different selections between the decisions computed in the hardware and the simulated ones. These differences were easily fixed by re-building the corresponding menu files and the firmware to include the new selections.

In addition to issues related with the firmware and the menu, the continuous validation approach was useful to understand the differences between hardware and simulation in detail. Since the L1Topo simulation runs as part of the detector reconstruction in the production of Monte Carlo samples to model different physics processes, it needs to reproduce the behaviour of the firmware ideally with no differences. Some discrepancies between both systems were found during 2018 that led to sub-percent level mismatch rates, most of them related to cases where there is ambiguity in the selection of the leading TOBs due to multiple objects having identical  $p_T$ . In these cases, the firmware typically selects a specific TOB as the leading one and computes the decisions based on it, while the simulation selects a different TOB, resulting in different decisions for the same events. The criteria to handle events with ambiguity used in the hardware system were inspected carefully and the simulation was modified to reduce the differences between both systems. The complete hardware description was implemented in the simulation only for muons, while for other objects approximations were needed due to the limited access to hardware information.<sup>10</sup> By the end of 2018 the agreement between hardware and simulation of many L1Topo items was improved with respect to the conditions at the beginning of the year, and in particular the agreement for all triggers used in physics analysis was higher than 99.9%, showing good understanding of the selections employed.

### 4.3.3 Physics performance

The addition of L1Topo to the ATLAS trigger system opened the possibility to select kinematic regions with challenging signatures that would otherwise lead to very high trigger rates. By applying complex selections, the topological algorithms allow to reject events that would eventually be removed by offline analysis selections while keeping the collisions in the interesting

<sup>10</sup>The TOBs sent to the L1Topo modules are pre-sorted in the L1Calo system, with a complex strategy implemented based on the crate/module number of origin. This information is not easily accessible at the HLT level and, while approximations based on the absolute position of each TOB can be used in the L1Topo simulation, the firmware approach is hard to match perfectly.

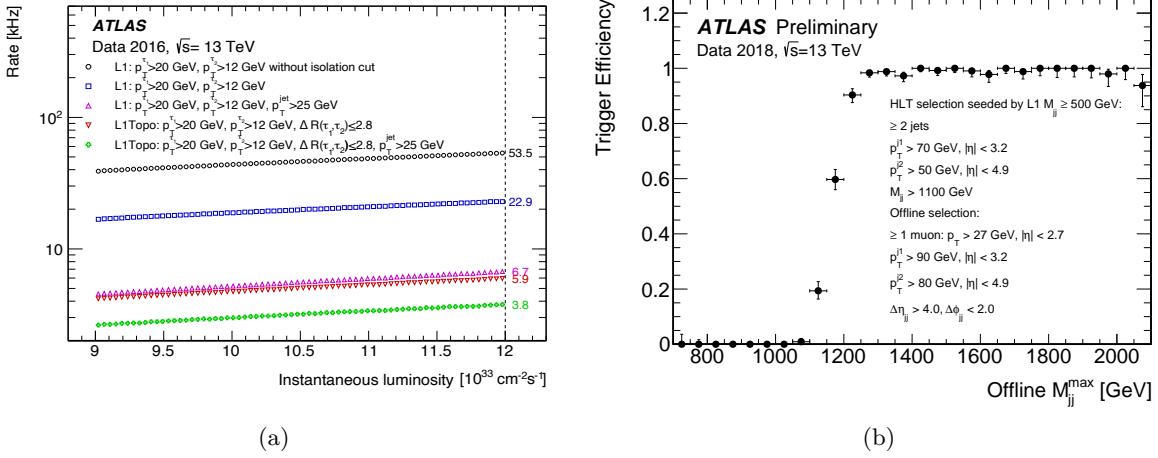


Figure 4.8: (a) Trigger rates for topological and non-topological di-tau triggers as a function of the luminosity, where a large rate reduction is observed for the L1Topo triggers compared to the conventional L1Calo ones. (b) Trigger efficiency curve for one di-jet trigger including an invariant mass requirement at the L1. Figures taken from ref. [83].

kinematic regions. Thanks to this approach, the  $p_T$  thresholds for different triggers configured in the menu could be lowered without a significant increase in the corresponding trigger rates, increasing the sensitivity to softer kinematic regions that are crucial for several physics analyses.

One example of analysis that profited greatly from topological selections at the trigger level during Run 2 is the cross section measurement of the process  $H \rightarrow \tau_{\text{had}}\tau_{\text{had}}$ , involving hadronically decaying  $\tau$ -leptons. The sensitivity of this analysis relies heavily on low- $p_T$  $\tau$ -lepton thresholds in order to have large signal acceptance, and so the trigger strategy adopted needed of the L1Topo selections in order to keep the event rates within the L1 allocated budget. In order to allow for the low  $p_T$  thresholds of 20 GeV and 12 GeV for the leading and sub-leading  $\tau$ -lepton candidates, the selection used includes a cut in their maximum radial separation of  $\Delta R(\tau_1, \tau_2) < 2.9$ . This topological requirement is optimized for large acceptance of  $H \rightarrow \tau_{\text{had}}\tau_{\text{had}}$  events while rejecting di- $\tau$ -lepton events arising through other mechanisms, and translates in a trigger rate reduction of a factor 4 with respect to the non-topological selections, as shown in figure 4.8a. Additionally, to further reduce the trigger rate, an additional requirement was included, targeting extra jets that could arise from QCD initial state radiation or from vector-boson fusion (VBF) Higgs production. This selection is added in the trigger in the form of an extra jet with  $p_T > 25$  GeV that does not overlap with any of the  $\tau$ -lepton candidates, ensured through the disambiguation algorithm described in table 4.1. This extra requirement results in a further trigger rate reduction of a factor 1.5.

Higgs analyses targeting VBF production also made use of L1Topo selections for large background rejection at the trigger level. The typical signature from this physics process includes highly energetic jets coming from the hard scattering, which can be exploited in order to trigger on events without targeting a particular Higgs decay channel. The common topological selections used in such trigger strategy during Run 2 target di-jet signatures and were based on invariant mass or event hardness criteria. The trigger efficiency for one of these selections is shown in figure 4.8b, where the leading jet is required to have  $p_T > 70$  GeV and be in the central region of the detector,  $|\eta| < 3.2$ , while the second one is required to have  $p_T > 50$  GeV without constraints on its position. The trigger rate for a signature with such low- $p_T$  thresholds would

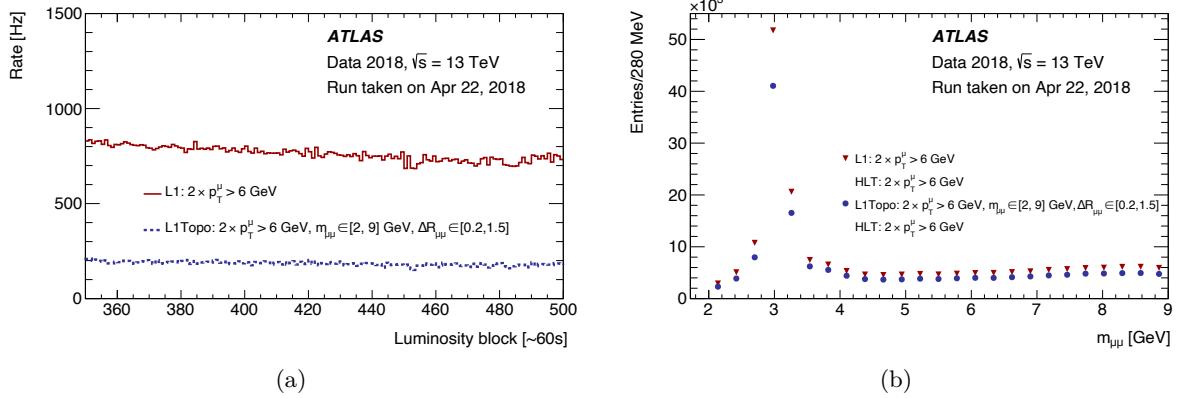


Figure 4.9: Comparison of the (a) trigger rates and (b) invariant mass distribution of the accepted events for two di-muon triggers with and without topological requirements. Figures taken from ref. [83].

be several times larger than the allowed budget at the L1 trigger, and so in this case the extra requirement of the invariant mass of the jets to be higher than 500 GeV is included in L1Topo both to reduce the trigger rate at the same time than ensuring that the jets come from the scattering process and not from the decay of any intermediate particle. Further selections applied on the HLT based on this L1Topo criterion include angular selections ( $\Delta\eta_{jj} > 4$  and  $\Delta\phi_{jj} < 2$ ) in order to further reject uninteresting events for these analyses, such as those produced in QCD multijet processes with similar signatures. Given that this trigger approach does not target any particular Higgs decay channel it could also be used for BSM Higgs decay searches during Run 2.

The L1Topo triggers have been crucial for the ATLAS  $B$ -physics program, as they rely heavily on low momentum leptons to reach the kinematically interesting regions. Di-lepton triggers with low  $p_T$  thresholds have very large trigger rates, and it is thanks to topological selections that they can stay at the few GeV level without exceedingly large trigger rates. One example is the  $J/\psi \rightarrow \mu\mu$  analysis [85], for which two muons with  $p_T > 6$  GeV are required at the L1. In order to keep these di- $\mu$  trigger rates at the same level than those of the corresponding selections during Run 1, additional topological requirements are included: a selection on the angular separation of the muons,  $0.2 < \Delta R_{\mu\mu} < 1.5$ , and an invariant mass cut,  $2 \text{ GeV} < m_{\mu\mu} < 9 \text{ GeV}$ . These selections are optimized for large signal acceptance and background rejection, with a resulting rate reduction of a factor 4, as shown in figure 4.9a. The trigger efficiency as a function of the offline invariant mass of the di- $\mu$  system is shown in figure 4.9b, where can be observed that no bias is introduced with respect to the non-topological di- $\mu$  trigger, with just a small efficiency loss at the relevant  $J/\psi$  mass peak. In final states with electrons, similar topological selections are used in order to keep the lepton  $p_T$  thresholds low. Another example of  $B$ -physics analyses using L1Topo triggers is the test for lepton flavour universality, measured through the decay asymmetry of  $K^{*0}$  to electrons or muons [19]. In this case, di-electron triggers are used with  $p_T$  thresholds of 7 and 5 GeV with a complex strategy that involves a logical OR of two different selections. The first one requires an extra selection on the invariant mass of the two electrons,  $m_{ee} < 9$  GeV, that helps reduce the large trigger rates due to the low momentum thresholds. A second algorithm is then included to target events in which the two electrons are very collimated and overlap in the detector. These cases typically result in the total energy of the electron being mismeasured due to momentum leaking out of the size of the electron TOB. In order to capture all the energy of the di-electron system, the second algorithm requires one electron with  $p_T > 7$  GeV to overlap with a large-R jet ( $\Delta\eta \times \Delta\phi = 0.8 \times 0.8$ ) with  $p_T > 15$  GeV,

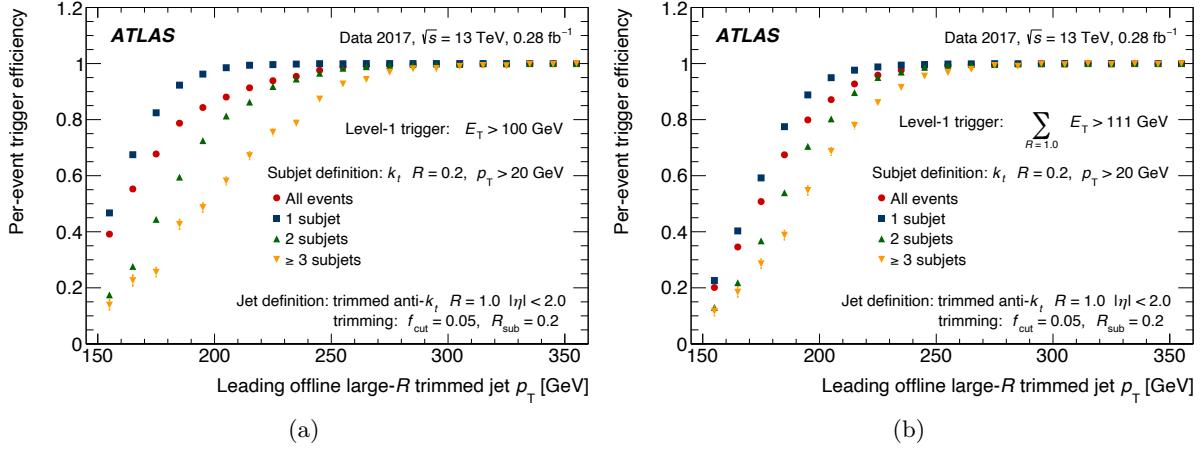


Figure 4.10: Jet energy trigger efficiency for different jet structures for (a) L1Calo trigger and (b) L1Topo trigger. Figures taken from ref. [83].

that is meant to record all the energy leaking out from the electron TOB.

The topological triggers have also been used during Run 2 for accurate measurement of jet properties at the L1. The simple cone algorithm can efficiently capture the total jet energy in cases where there are multiple sub-components inside of it. As shown in figure 4.10a, a L1Calo trigger selection with the usual sliding window algorithm with parameter  $\Delta R = 1.0$  starts to lose efficiency when there are more than two sub-jets inside, while the performance of the simple cone algorithm (figure 4.10b) is much less affected by the jet substructure. The energy threshold in the case of the L1Topo simple cone trigger is chosen so that the trigger rate matches that of the non-topological selection with  $E_T > 100$  GeV.

Additionally, the L1Topo triggers were also used for analyses targeting long-lived particles (LLP). Given the possibility to access information from neighbouring bunch crossings, the L1Topo system has been used to implement trigger selections based on certain TOBs being present in consecutive proton-proton collisions. In particular, a selection requiring one jet with  $p_T > 50$  GeV in one bunch crossing together with a muon with  $p_T > 10$  GeV in the following one was used in the search for heavy, long-lived, highly ionising particles, allowing for an efficiency increase of a factor 2-3 with respect to non-topological triggers. Another example of physics analysis making use of L1Topo selections is the search for neutral scalars decaying to LLPs. The typical expected signature of these processes are displaced jets with little energy deposited in the electromagnetic calorimeter in comparison with the energy collected by the hadronic calorimeter, since the LLP travels some distance inside the detector before decaying. A non-topological selection exists for this type of signatures, called High- $E_T$  CalRatio, based on jets with  $p_T > 60$  GeV that combine information from both calorimeters. This threshold, however, limits the sensitivity to scenarios where the scalars are light, and cannot be lowered due to the high trigger rates during Run 2 LHC conditions. For this reason, a second selection, named Low- $E_T$  CalRatio, is performed by the L1Topo system to target these scenarios. In this trigger, the fact that all jets coming from the LLP decay are expected to be very collimated is exploited, and the  $\tau$  TOB is used to measure all the energy more precisely instead of the jet TOB. The selection requires then the leading  $\tau$  TOB to have  $p_T > 30$  GeV and not to overlap with any  $e/\gamma$  TOB with energy higher than 3 GeV. Since the  $e/\gamma$  TOB are constructed exclusively using information from the electromagnetic calorimeter, this selection effectively translates in having a large energy ratio between hadronic and electromagnetic calorimeters in a certain part of the

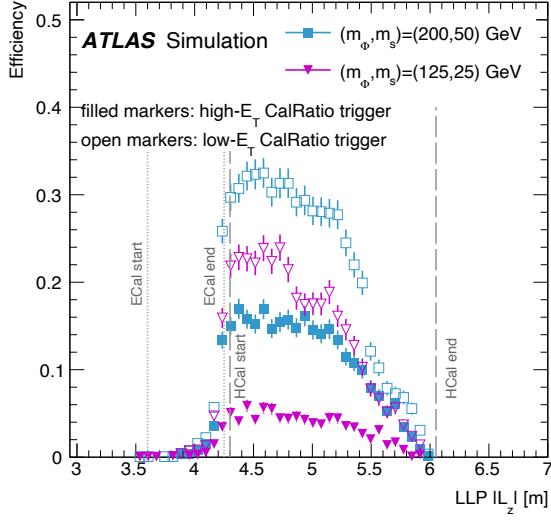


Figure 4.11: Comparison of signal acceptance for light scalars by the high- $E_T$  (non-topological) and low- $E_T$  (topological) CalRatio triggers. Figure taken from ref. [83].

detector. The signal efficiency is shown in figure 4.11 for two low mass hypotheses for both CalRatio triggers, with a factor 4 (2) gain in signal acceptance for a scalar mass of 125 GeV (200 GeV) obtained thanks to the topological selection.

## 4.4 Phase-1 Upgrade

During the LS2, between 2018 and 2022, the ATLAS L1 system is undergoing the Phase-1 upgrade [86] for enhanced processing power that allows for better object reconstruction in the trigger. The L1Calo system is completely replaced by the Feature Extractor (FEX) sub-systems for optimal TOB reconstruction. Three types of FEX systems will be used during Run 3: the electron Feature Extractor (eFEX), for reconstructing  $e/\gamma$  and  $\tau$  candidates, the jet Feature Extractor (jFEX), that provides jet candidates together with total energy and missing energy measurements, and that can also provide  $\tau$  candidates, and the global Feature Extractor (gFEX), mainly for energy measurements but that can also provide jet candidates. The TOB type overlap existing between the different FEX systems allows for flexible configuration and excellent reconstruction by combining object candidates from two FEXes. One example is the  $\tau$  candidates from eFEX ( $e\tau$ ) and jFEX ( $j\tau$ ), that can be combined inside of the L1Topo system to make a new TOB type that profits from the higher resolution in jFEX and more accurate isolation criteria from eFEX. In the new system, the L1Calo FEXes will only reconstruct the object candidates and send them to the L1Topo modules, where the multiplicity triggers will be processed. The L1Muon system will be upgraded as well, with the inclusion of new trigger chambers as part of the New Small Wheels, as discussed in the previous chapter. In addition, the MUCTPI system will also be upgraded for larger bandwidth, which will allow in Run 3 to send twice as many  $p_T$  thresholds to the CTP with respect to Run 2. A new system, the Topo-Fibre Optic eXchange (TopoFOX), is included in the real data path after L1Calo/MUCTPI to remap their outputs before feeding them to the L1Topo modules, removing the need for the CMX/MUCTPI-to-Topo features from Run 2. The CTP, in turn, will not be upgraded during the LS2. A scheme of the changes in the L1 system with respect to Run 2 is shown in figure 4.12.

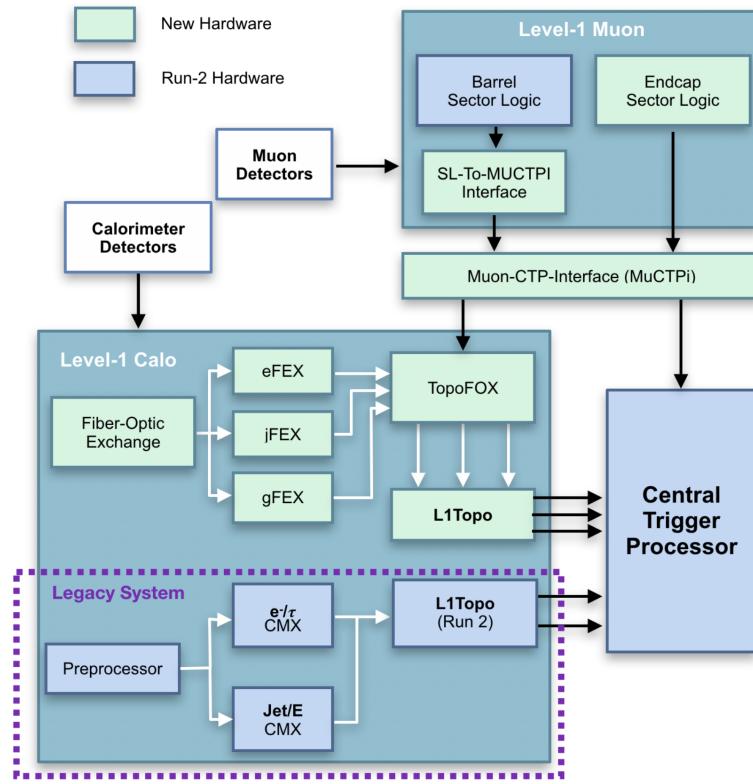


Figure 4.12: Scheme of the Phase-1 upgrade in the L1 system.

The two L1Topo modules from Run 2 will be replaced by three new AdvancedTCA-compliant hardware boards, each with two Xilinx Ultrascale+ FPGAs used for algorithm processing with enhanced capacity in comparison with the previous system and with larger input (118 fibres) and output (24 fibres) capabilities. Two mezzanine cards are included in each module, one for power and control and the other for output handling to the CTP. Out of the three modules, one will be dedicated exclusively to computing multiplicity over energy thresholds for L1Calo triggers (TOPO1) while the other two (TOPO2, TOPO3) will handle the topological selections. The TOPO1 board will be connected to the CTP through optical fibres to allow for a larger amount of output bits as needed by the multiplicity triggers,<sup>11</sup> while the other two will be connected through electrical cables as the requirement on the number of output bits is lower. Similarly to Run 2, the L1Topo system will provide the CTP with the decisions over each topological trigger item configured in the menu. For multiplicity triggers, however, the output from L1Topo will consist of only the counts and not any decisions, which are left to the CTP to form together with the L1A signal.

The increase in processing power that comes with the new L1Topo system allows for more complex topological algorithms that could not be implemented during Run 2 due to hardware and latency limitations. During the LS2 different signatures have been studied through Monte Carlo simulation and new trigger selections involving complex kinematic quantities will be included

<sup>11</sup>The output from a multiplicity trigger consists of the number of TOBs of a certain type above a configured energy threshold, sent to the CTP in the form of bits. In order to allow for a reasonable flexibility while taking into account the hardware limitations, the typical number of output bits per multiplicity trigger is 3, which corresponds to counts of up to 7 TOBs in an event. If more TOBs are present, all bits are set to one and the event is said to be “saturated”, in which case the CTP cannot distinguish between events having exactly or more than the maximum number of TOBs allowed.

in the Run 3 menu. The new algorithms implemented in the L1Topo firmware are collected in table 4.2, and include simultaneous selections on invariant mass and angular separations. These algorithms also come with the possibility of adding muon charge information to form a decision, with applications in  $B$ -physics analyses. All these algorithms have been implemented in the Athena-based simulation and tested offline with Run 2 recorded data. Additionally, the algorithms that will be implemented in the TOPO1 module for L1Calo multiplicity triggers have also been simulated during the LS2 with detailed studies on their expected performance in preparation for the firmware implementation.

An extensive commissioning plan has been laid out for the new L1Topo system, both during the LS2 in laboratory conditions and once installed in P1 with different ATLAS run types. First tests were performed on module prototypes during 2018 at the CERN laboratories with successful results, where communication with other L1 systems was studied. Part of the commissioning plan once the system is deployed include triggering on cosmic rays and hot tower patterns in the calorimeters, together with comparisons of the hardware decisions with simulated ones, similarly to what was done in the validation of the L1Topo modules in Run 2. A new feature of this commissioning plan includes ATLAS runs during LHC proton-proton collisions with both the new L1 system and the former hardware from the L1Calo and L1Topo systems used in Run 2. The old modules, referred to as *Legacy system*, will run in parallel and the trigger menu will include identical selections configured in both Run 2 and Run 3 systems for comparison. Given that the CTP does not change with respect to Run 2, the number of allowed input bits will still be fixed to 512. This translates into the need for a reduced trigger menu during the commissioning period with the Legacy system, that can potentially affect the kinematic reach of different physics analyses. For this reason, the trigger menu for this period has been designed carefully to allow for detailed validation of the new systems together with minimum impact on the physics performance. After the commissioning period has ended, the Legacy system will be removed and the L1 trigger will operate with the full Phase-1 implementation.

Algorithm	Definition
Invariant mass with azimuthal distance: $x_1^2 \leq m_{\text{inv}}^2 \leq x_2^2$ AND $y_1 \leq \Delta\phi \leq x_2$	Simultaneous selection on the invariant mass and the separation in the transverse plane of two given input lists. The parameters $(x_1, x_2, y_1, y_2)$ are configured through the trigger menu.
Invariant mass with angular distance: $x_1^2 \leq m_{\text{inv}}^2 \leq x_2^2$ AND $y_1^2 \leq \Delta R^2 \leq y_2^2$	Simultaneous selection on the invariant mass and the angular separation of two given input lists. The parameters $(x_1, x_2, y_1, y_2)$ are configured through the trigger menu.
Invariant mass with angular distance and muon charge: $x_1^2 \leq m_{\text{inv}}^2 \leq x_2^2$ AND $y_1^2 \leq \Delta R^2 \leq y_2^2$ AND Total charge $ Q_\mu  \neq 2$	Only applicable to muons TOBs. Simultaneous selection on the invariant mass and the angular separation of two given input lists. Additionally, events are accepted only if the muons verifying the selections have opposite charge. The parameters $(x_1, x_2, y_1, y_2)$ are configured through the trigger menu.
Angular distance with muon charge: $x_1^2 \leq \Delta R^2 \leq x_2^2$ AND Total charge $ Q_\mu  \neq 2$	Only applicable to muon TOBs. Angular distance between the objects of two given input lists. Additionally, events are accepted only if the muons verifying the selection have opposite charge. The lower and upper bounds $(x_1, x_2)$ are configurable parameters of the algorithm.
Invariant mass of three objects: $x_1^2 \leq m_{\text{inv},3}^2 = m_{\text{inv},1-2}^2 + m_{\text{inv},2-3}^2 + m_{\text{inv},1-3}^2 \leq x_2^2$	Selection on the invariant mass of three input lists. The computation is done with the approximation of individual TOB masses being negligible for reduced resource usage. The parameters $(x_1, x_2)$ are configured through the trigger menu.
Invariant mass of three objects and muon charge: $x_1^2 \leq m_{\text{inv},3}^2 \leq x_2^2$ AND Total charge $ Q_\mu  \neq 2$	Only applicable to muons. Selection on the invariant mass of three input lists. Additionally, events are accepted only if the muons that satisfy the selections have opposite charge. The parameters $(x_1, x_2)$ are configured through the trigger menu.
Exclusive jets: $\chi_{\min} \leq \chi_1 = E_{\text{T},1} e^{\eta_1} + E_{\text{T},2} e^{\eta_2} \leq \chi_{\max}$ AND $\chi_{\min} \leq \chi_2 = E_{\text{T},1} e^{-\eta_1} + E_{\text{T},2} e^{-\eta_2} \leq \chi_{\max}$	Takes the two leading jets in the event and combines them into the quantities $\chi_1$ and $\chi_2$ and accepts the events with both in a defined range. The parameters $(\chi_{\min}, \chi_{\max})$ are configurable through the trigger menu.

Table 4.2: Description of new algorithms implemented in L1Topo preparation for Run 3.



# Chapter 5

## Simulation of proton collisions

### 5.1 Introduction

The processes undergone in the proton-proton collisions in the LHC are very complex due to the nature of the strong interaction as described by the SM. Understanding in detail the properties of the collisions and the particles produced therein is fundamental for physics analyses in the ATLAS Collaboration, and so a large effort goes into modelling studies for key physics processes through Monte Carlo (MC) simulation. The typical output of these simulated collisions are all of the particles produced in the event together with their 4-momentum vectors. An extra step commonly carried out after the simulation of the physics process is to infer the signal left in the detector by each of the particles produced in the collision, reproducing the behaviour of real recorded data. A number of tools exist in order to provide accurate descriptions of many SM and BSM physics phenomena, with the models they generate being continuously validated by contrasting them with data in well controlled regions of the available phase space.

The MC simulated collisions have many applications. Good predictions from the simulated models are essential for analysis design, to be able to accurately extract the kinematic features of a selected signal process in order to efficiently select the targeted corner of the phase space. In many cases, the background processes are also modelled using MC simulation, and therefore being able to predict their behaviour as close to nature as possible is of extreme relevance, especially in searches for BSM physics where mismodelled backgrounds could be mistaken for signal processes. Additionally, MC simulation is generally used for efficiency studies when designing new trigger selections as well to understand the acceptance window and also estimate the rate of recorded events for specific physics processes. Object identification and calibration studies can also be performed with the help of MC generated samples, in particular those requiring knowledge about the actual particles produced in the collision such as flavour tagging. Even though the differences between MC simulation and real data are minimized by constantly improving the MC tools, small discrepancies between the two usually still survive, which are generally corrected downstream in the physics analyses with dedicated calibrations or statistical methods.

This chapter provides an overview of the simulation process of proton-proton collisions in the ATLAS experiment, and is organized as follows. In section 5.2 the physics processes that take place during a hadron collision are discussed, while the methods and tools to simulate such processes are briefly described in subsections 5.3.1 and 5.3.2, respectively. Finally, in section 5.4 the simulation of the ATLAS detector is briefly discussed.

## 5.2 Physics processes inside proton-proton collisions

The interactions between protons during a bunch crossing can be separated in elastic scattering, when the transferred momentum between the protons is low and they just change their trajectories as a result, and deep-inelastic or hard scattering, when the energy exchanged is large enough to break the protons into their constituent partons. The latter are the most interesting scenarios in the LHC collisions, since new particles that were not in the initial state are generally produced and a wide range of physics processes can be explored. These interactions are driven by QCD, which defines two approaches for their simulation depending on the energy scale of the process. When the energy exchanged between the particles is very high, the value of the strong coupling  $\alpha_s$  is small enough to follow a perturbative approach, as discussed in chapter 1. In this regime it is simpler to consider the collision among the constituent partons in the protons rather than the scattering between protons themselves. As the energies become lower due to it being shared in large decay chains, the magnitude of the strong coupling increases and the perturbative approximation can no longer be used. In this regime, phenomenological models are employed in order to provide predictions in that regime.

A typical proton-proton collision involves high and low energy stages, which are simulated following the corresponding approach. A scheme of a general collision is shown in figure 5.1 with all of the intermediate steps between the hard scattering and the final state particles highlighted. The inelastic scattering is produced between one parton from each proton involved in the collision. The original partons from the colliding protons, however, can also interact, producing extra particles in a process referred to as *underlying event*. The products of the hard scattering that are subject to the strong interaction (quarks and gluons) will lose energy in the form of radiated gluons, that will split into quark-antiquark pairs which can further radiate new gluons. This complex process is referred to as *parton shower*, and is the reason why the final states generally contain very large numbers of particles. Finally, once quarks and gluons in the shower have energies low enough to get bound in colorless states, they combine among themselves into heavier composite particles through the *hadronization* process. Additionally, the initial partons can also radiate gluons before the hard scattering, triggering extra parton showers producing what is called *initial state radiation* (ISR), while radiated gluons after the hard scattering can also occur resulting in *final state radiation* (FSR). These processes are discussed in detail below.

### 5.2.1 Factorization theorem and parton distribution functions

The complexity of a proton-proton collision, as displayed in figure 5.1, makes it hard to perform a direct computation of any related observable. The factorization theorem, however, allows for a division of any calculation in two parts: a short-distance (or high energy) section that can be computed by means of the perturbative approximation for QCD, and a long-distance part, which is assumed universal enough to be extrapolated from fits to data from different experiments. Furthermore, the same theorem allows to split the computation into the contribution from each of the partons involved, instead of considering the complete bound states. The differential cross section for an observable  $\mathcal{O}$  in a hadronic collision can be written as [87]

$$\frac{d\sigma}{d\mathcal{O}} = \sum_{a,b} \int_0^1 dx_a dx_b \sum_F \int d\Phi_F f_a^{h_1}(x_a, \mu_F) f_b^{h_2}(x_b, \mu_F) \frac{d\hat{\sigma}_{ab \rightarrow F}}{d\hat{\mathcal{O}}} D_F(\hat{\mathcal{O}} \rightarrow \mathcal{O}, \mu_F), \quad (5.1)$$

where the indices  $a, b$  indicate all possible partonic constituents in both hadrons  $h_1, h_2$ , respectively,  $x_i$  is the momentum fraction carried by each parton and  $d\Phi_F$  is the phase space of the

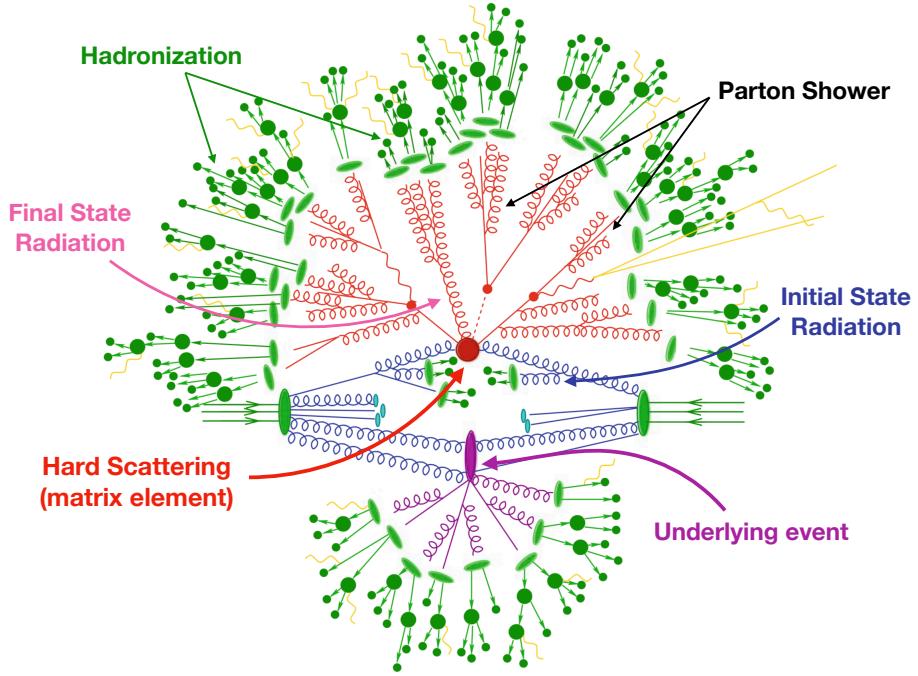


Figure 5.1: Scheme of a proton-proton collision in the LHC.

final state  $F$ , with the sum over  $F$  running all over possible combinations such that  $ab \rightarrow F$ . The variables written with a hat represent the corresponding quantities at a parton level instead of those related to the hadrons. The perturbative part of the computation enters through the partonic cross section  $\hat{\sigma}_{ab \rightarrow F}$ . The non-perturbative part of the calculation enters through the parton distribution functions (PDF),  $f_a^{hi}(x_a, \mu_F)$ , and the fragmentation function (FF),  $D_F(\hat{O} \rightarrow O, \mu_F)$ , which parametrizes the transition from the free quarks to the bound states in the hadronization process. Both the PDF and the FF need to be parametrized with fits to experimental data in well defined processes as the perturbative approximation does no longer apply in those regimes. The turning point between the perturbative and non-perturbative regimes is defined as the *factorization scale*,  $\mu_F$ , which is arbitrary and needs to be defined by the user. It is typically chosen so that the PDFs include a resummation of all emissions efficiently without double counting nor accidentally removing some of the corrections needed.

The partonic cross section is proportional to the matrix element of the process  $ab \rightarrow F$  as

$$d\hat{\sigma}_{ab \rightarrow F} \propto |\mathcal{M}_{ab \rightarrow F}|^2(\Phi_F; \mu_F, \mu_R), \quad (5.2)$$

where the matrix element  $\mathcal{M}_{ab \rightarrow F}$  is computed through the usual procedure involving the Feynman rules for QCD. They include a second arbitrary parameter, the *renormalization scale*,  $\mu_R$ , at which the strong coupling constant  $\alpha_s$  is evaluated. Additionally, the proper computation of the matrix elements need to be done at the partonic center-of-mass energy,  $\sqrt{\hat{s}}$ , as each parton only carries a fraction of the momentum of the original hadrons. The dependence of the matrix elements with the factorization and renormalization scales decreases as the order in the perturbative approximation increases, disappearing completely if all orders were included in the computation. For this reason, the estimate at the highest order in perturbations available should always be used.

The PDF  $f_a^{hi}(x_a, \mu_F)$  is the parametrized distribution of partons of type  $a$  carrying a momentum fraction  $x_a$  inside the hadron  $h_i$  at the energy scale  $\mu_F$ . They need to be estimated

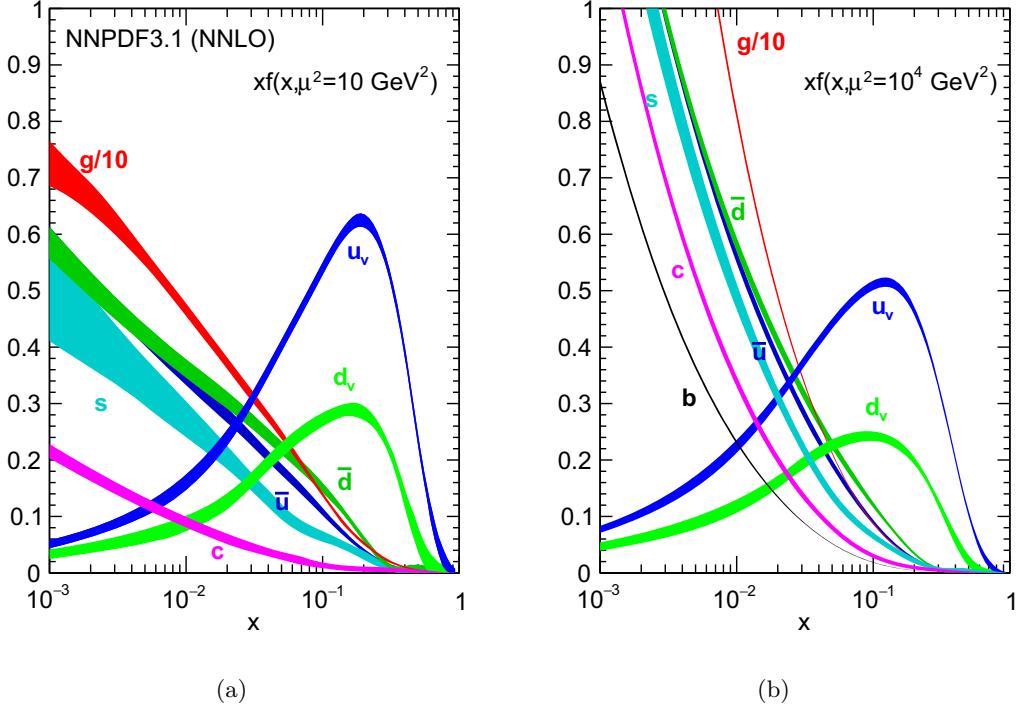


Figure 5.2: PDFs as determined by the NNPDF Collaboration for (a)  $Q^2 = 10 \text{ GeV}$  and (b)  $Q^2 = 10^4 \text{ GeV}$  [88].

from data measured at a certain energy scale, since they cannot be computed using perturbative QCD. Their evolution with energy  $Q^2$  can be obtained by means of the Altarelli-Parisi equations:

$$\frac{\partial q(x, Q^2)}{\partial \log Q^2} = \frac{\alpha_s}{2\pi} (P_{qq} \otimes q + P_{qg} \otimes g), \quad (5.3a)$$

$$\frac{\partial g(x, Q^2)}{\partial \log Q^2} = \frac{\alpha_s}{2\pi} \left( \sum_i P_{gq} \otimes (q_i + \bar{q}_i) P_{gg} \otimes g \right), \quad (5.3b)$$

The symbol  $\otimes$  in the previous equations denotes the convolution integral defined as

$$P_{qq} \otimes q \equiv \int_x^1 \frac{dy}{y} f_q(y) P_{qq} \left( \frac{x}{y} \right), \quad (5.4)$$

where  $P_{ij}$  is the *splitting function*, encoding the probability for the parton  $i$  to emit a parton  $j$  carrying a fraction  $x/y$  of the original parton momentum. The determination of PDFs is carried out by many collaborations performing deep inelastic scattering experiments on hadrons. A commonly used set of PDFs is the one released by the NNPDF Collaboration [88], using data from HERA-II [89], ATLAS, CMS and LHCb. In figures 5.2a and 5.2b are shown the proton PDFs from the latest NNPDF3.1 set determined at an energy scale squared of  $10 \text{ GeV}^2$  and  $10^4 \text{ GeV}^2$ , respectively, where it can be appreciated that the valence quarks are more likely to carry most of the momentum of the proton, with the sea partons carrying a lower fraction.

## 5.3 Monte Carlo simulation and generators

The general approach to simulate proton-proton collisions is by means of MC methods, which rely on the generation of pseudo-random numbers. The computation is split on the different processes mentioned above and represented in figure 5.1, with dedicated tools existing that can either reproduce the complete process chain or target specific steps the simulation of the collisions.

### 5.3.1 Event simulation

In the following paragraphs, the general methodology for each of the steps of a simulated collision is examined, from the matrix element computation all the way to the hadronization.

#### Matrix element

The matrix element for a certain partonic process,  $ab \rightarrow F$ , can be obtained with the usual tools of perturbation theory at different orders and constitute the main contribution to its cross section. The inclusive total cross section<sup>1</sup> for the process  $ab \rightarrow F$ , where inclusive states that the final state should be composed of  $F$  plus any other particles produced in the process, can be expressed as [87]

$$\hat{\sigma}_{ab \rightarrow F} = \sum_{k=0}^{\infty} \int d\Phi_{F+k} \left| \sum_{l=0}^{\infty} \mathcal{M}_{F+k}^{(l)} \right|^2, \quad (5.5)$$

where the sum over  $k$  indicates all possible “real-emission” corrections or legs, while the sum over  $l$  indicates virtual corrections or loops. The order in the perturbation expansion is fixed by the combination of real and virtual corrections included in the computation. If  $k = 0, l = 0$  the computation is performed at Leading Order (LO) for inclusive  $F$  production, without any corrections from loops of external legs included. Increasing the number of real-emission corrections,  $k = n$  with  $l = 0$ , provides the cross section for  $F$  production in association with  $n$  jets at LO, while varying both  $k, l$  such that  $k + l \leq n$  provides the  $N^n$ LO cross section for inclusive  $F$  production, which includes the  $N^{n-1}$ LO computation for  $F$  plus one jet,  $N^{n-2}$ LO for  $F$  plus two jets, etcetera.

The computation at all orders in the perturbative approximation is impossible in most cases, with the cross section for most final states being computed at a fixed order  $n$ . The introduction of virtual corrections in the computation of the matrix element usually leads to divergences that need to be suppressed for the computation to be meaningful. In general, only partons contributing with low momentum compared to the energy scale of the process  $ab \rightarrow F$  are allowed in the matrix element corrections, to avoid the presence of divergent integrals. The soft and collinear regimes, which are typically inaccessible due to this requirement, are generally treated as part of the parton shower process.

As mentioned above, the computation of the matrix element  $\mathcal{M}_{F+k}^{(l)}$  includes a dependency on the arbitrary energy scales  $\mu_F$  and  $\mu_R$ . While these scales could be well defined in some specific processes, in general they need to be defined based on assumptions. For this reason, it is customary to derive systematic variations to the cross section of the process by raising and

---

<sup>1</sup>The differential cross section for a determined observable  $\mathcal{O}$ , as expressed in 5.1, can be obtained by adding a  $\delta(\mathcal{O} - \mathcal{O}(\Phi_{F+k}))$  in eq. 5.5.

lowering the chosen values for both  $\mu_F$  and  $\mu_R$ , providing an uncertainty on the cross section based on the differences with the nominal computation.

### Parton shower

The parton shower process is simulated through dedicated algorithms that aim at handling the emissions from all initial and final partons before and after the hard scattering. They complement the matrix element computation by incorporating the partons removed from the matrix element step as described above. The general workflow of a parton shower includes recursively generating sequences of emissions, ordered in decreasing momentum or angle, based on the probability that each parton involved splits into two lower-energy ones. The splitting processes considered typically in parton shower algorithms are  $g \rightarrow q\bar{q}$ ,  $q \rightarrow gq$  and  $g \rightarrow gg$ . In the regimen where the momentum of the involved partons is similar, the cross section for producing  $n + 1$  particles in the final state,  $\sigma_{n+1}$ , can be obtained from the cross section for  $n$  particles,  $\sigma_n$ , and the splitting functions  $P_{ab}$  described above as

$$d\sigma_{n+1} \approx d\sigma_n \frac{\alpha_s}{2\pi} \frac{dQ}{Q} P_{ab}(x, Q) dx, \quad (5.6)$$

where  $x$  is the momentum fraction carried by the outgoing parton and  $Q$  is the energy scale at which the splitting happens.

The parton shower algorithms produce iteratively all possible emissions coming from the partons involved in the hard scattering process. For outgoing particles, the emissions are expected to have lower energy the farther removed they are from the inelastic scattering, becoming progressively softer as the initial energy is split throughout the shower. The algorithm recursively loops over all outgoing partons in order to perform the splitting until the energy of the shower reaches the hadronization scale at around 1 GeV. Inversely, in the case of ISR the emissions are expected to be more energetic when they are removed far away from the hard scattering, and so the iterative process is typically applied backwards when modelling it.

The parton shower step can provide an estimate for the full perturbation expansion of the cross section, where the matrix element is accurately computed to a certain order and terms above are computed typically with leading-logarithmic (LL) precision. The full computation of a cross section at order  $n = k + l \geq 1$ , however, introduces an overlap between the matrix element at order  $n$  and the parton shower for the order  $n - 1$ , which needs to be treated carefully when combining computations at different orders. This happens since the cross section is always computed for the inclusive final state  $F$ , which includes all emissions  $F+k$  through the showering algorithm. When computing the cross section for the process  $F+1$  at the same order  $n$  the matrix element provides duplicated information with respect to the parton shower for  $F$ , in a certain kinematic region. This leads to event double-counting that needs to be handled properly for the process  $F$  not to be overestimated.

The most common strategy used in current parton shower generators in order to avoid overlaps with the computed matrix element is *slicing*. It consists on splitting the available phase space by defining a matching scale, above which only the matrix element is kept while below only the parton shower is used. This approach leaves the more accurate computation at N<sup>n</sup>LO from the matrix element in the region where it can be computed without divergent integrals appearing, while the more problematic soft and collinear regimes are left for the parton shower algorithms to handle. Other matching approaches are discussed in ref. [87].

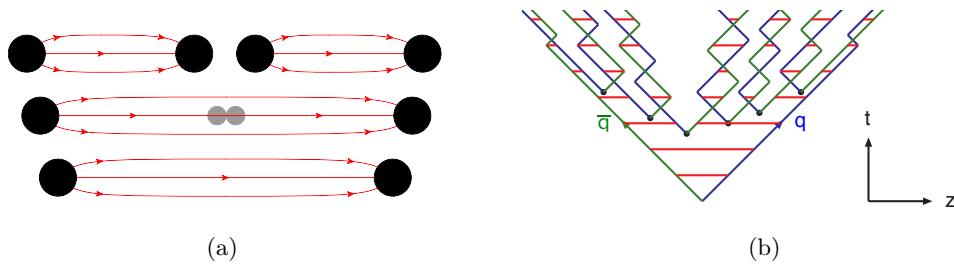


Figure 5.3: Representation of (a) the general principle behind the string models [87], and (b) the hadronization process from a initial state with two quarks in the Lund model [91].

## Hadronization

The hadronization step starts once the energy of the showers is low enough so the perturbative approximation for QCD breaks. It is performed based on phenomenological models, making use of the FFs fitted with experimental data. In this step, a reconstruction algorithm is executed in order to group together the partons after the shower into colorless states, using as inputs the already organized distributions of momentum, color and flavour as produced by the parton shower generator. The hadronization algorithm acts recursively on all available partons, which can be merged first into unstable hadrons that would further split, until all partons are confined into stable bound states.

Two approaches for hadronization are widely used in MC generators to accurately model the final stages of a hadron collision. The first one, the string model [90], is based on linear confinement in QCD, where the potential energy between quarks is stored in a virtual “string” connecting them. The contribution to the potential energy from the strong interaction grows linearly with the distance separating the partons, while a contribution from the electromagnetic interaction is also included when it is relevant.<sup>2</sup> The most commonly used model within this approach is the Lund model. When the distance between two quarks increases, so does the potential energy stored in the string, which can eventually break in two, adding a new quark-antiquark pair to the newly created end points of the string, as shown in figure 5.3a. In this model, the starting point is always the leading hadrons that contain the original end point quarks, and iterates towards the middle of the string, as shown in figure 5.3b. This allows for single on-shell hadron fragmentation, which makes it simple to ensure that only resonance with the appropriate mass are produced.

The cluster model [92], on the other hand, relies on the pre-confinement property of the strong interaction, which states that color-singlet combinations of partons will be formed in the shower following an invariant mass distribution that depends only on the factorization and renormalization scales set in the generator. As shown in figure 5.4, the invariant mass of the color-singlet “clusters” does not depend on the energy scale  $Q$  of the process. In these models, the partons are grouped into color-singlet clusters which are allowed to decay into two lighter clusters or a cluster plus a hadron. This operation is performed iteratively until only stable hadrons remain in the final state.

<sup>2</sup>The Coulomb potential decreases as  $1/r$ , and therefore only at short distances it has a non-negligible impact compared to QCD. At distances below 1 fm the potential due to the strong interaction is expected to be comparable to the electromagnetic one and therefore both are taken into account. At higher distances, QCD dominates and the Coulomb potential becomes negligible in comparison.

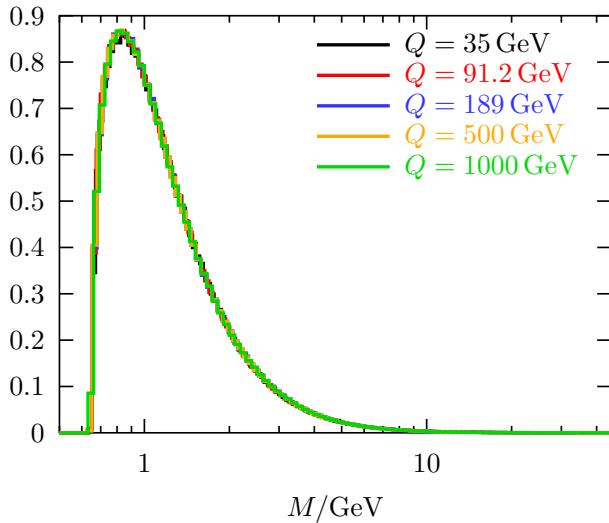


Figure 5.4: Invariant mass distribution for the primary clusters of partons, computed at different energy scales  $Q$  [91].

### Multiple parton interactions, pileup and underlying event

In the hadron interactions at the LHC protons do not interact one-by-one since they travel in the beam inside bunches. Each bunch crossing produces multiple proton-proton interactions, out of which generally only one of them results in large enough momentum transfer for the deep inelastic scattering to happen. It is possible, however, that multiple hard scatterings are produced in one bunch crossing in what is commonly referred to as *multiple parton interactions*. The typical signature of these processes are back-to-back jets with little total transverse momentum. This type of extra interactions can generally be computed in the perturbative regime as well, similar to the main event, where the bunches crossing can be considered as a collection of the hadron constituent partons. The fraction of such processes, however, is typically small, with the majority of events including only one hard scattering process. When the partons interacting are the other constituents of the hadrons involved in the hard scattering the process is referred to as *underlying event*, while if they are produced due to multiple proton interactions it is called *pileup*.

Additional interactions involving low momentum transfer, on the other hand, are much more common, and can have a larger effect on the kinematics of the event, such as the number of particles in the final state or the total scattered energy. In this case, where the energy scale is low enough to break the QCD perturbative regime, phenomenological models need to be used in order to accurately describe the processes involved. The typical signature of the underlying event in this regime are soft jets in the beam direction with little activity in the transverse plane, with the interaction probability depending on the overlap between the hadrons involved in the hard scattering. The soft interactions arising due to in-time pileup are modelled with similar techniques to those used for the underlying event, while for out-of-time pileup additional elements are incorporated in the computation to account for the detector response and signal processing. Both pileup and underlying event processes are present simultaneously in the simulation of any collision.

### 5.3.2 Monte Carlo generators

The simulation of the hadron collisions is performed through dedicated software tools called *Monte Carlo generators*. They are generally classified according to which of the steps described above they can perform, with general purpose generators being capable of simulating the whole event process, while dedicated generators target specific parts of the chain, such as the matrix element computation or the showering process.

#### General purpose generators

A number of independent software packages exist to simulate all steps involved in the collision, from the matrix element computation to the hadronization process, that can be used standalone. Since they provide the means to compute each of the steps involved in the process, they can also be interfaced with other step-specific generators for increased accuracy. This is a common practice in ATLAS analyses, where general purpose generators are often exclusively used for the parton shower and hadronization steps, while the matrix element computation is performed by dedicated generators. The most widely used general purpose generators at the time of this writing are summarized below.

- PYTHIA [93] is used to generate events with LO precision in the matrix element computation for  $2 \rightarrow n$  ( $n \leq 3$ ), with the subsequent parton shower and hadronization steps applied on top. The parton shower modelling can accurately describe both low- and high- $p_T$  radiation regimes, which include ISR and FSR. The hadronization step is carried out following the Lund model described above.
- HERWIG [94] can also simulate events at LO precision in the matrix element computation for  $2 \rightarrow 2$  processes. The parton shower step includes full spin correlation for all partons, which leads to more accurate computation than PYTHIA since it includes gluon splitting asymmetries in the collinear approximation. The hadronization step is performed based on the cluster model.

#### Matrix element generators

Some tools provide highly accurate matrix element computations at the expense of not including the simulation steps from the parton shower onwards. These can be combined with other generators for a complete description, as mentioned above. The main matrix element generators used in ATLAS at the time of this writing are mentioned below:

- POWHEG-BOX [95] can generate events with matrix element computations at NLO, using a scheme that includes a 5-flavour scheme. It makes use of a custom matching method named POWHEG [96], that allows for easy interfacing with other generators to perform the parton showering, hadronization and multiple parton interaction steps.
- MADGRAPH5\_AMC@NLO [97] can generate events at NLO for specific processes, which makes use of the MC@NLO method, and also LO computations for user-defined lagrangians. This is particularly useful in the design and simulation of BSM signal models. The matrix element computation can be interfaced with a second generation for the parton shower and hadronization processes, for which a matching method based on slicing is implemented.

- SHERPA [98] computes matrix elements for the process  $F + k$  at NLO precision for  $k \leq 2$  and at LO precision for  $k \leq 4$ , and includes also its own parton shower and hadronization mode. The parton shower matching to the matrix element computation based on the slicing method. The hadronization process is carried out following a cluster model.

## 5.4 Simulation of ATLAS detector

The output of the different MC generators mentioned above are the 4-momenta of all stable particles in the final state after the hadron scattering, usually encoded in a widely used format named *HepMC*. The simulated events can be readily used to understand the kinematics of specific processes, in what is called as *particle-level* analysis. The samples produced in this way, however, cannot be compared to the real data collected by the ATLAS detector as it does not include any estimate for particle-matter interaction nor detector response. In order to have simulated events that can be compared to real data, an extra step needs to be included in the simulation chain to account for detector effects, producing samples at the so-called *reconstruction level*.

The simulation of the ATLAS detector is carried out by means of the GEANT4 package [99], which can efficiently reproduce effects arising from interactions with matter. A scheme of the general process is shown in figure 5.5, and can be summarized as follows:

1. The stable particles produced in the event generation are interfaced with GEANT4 in order to emulate their interactions with each of the detector sub-systems, reproducing all electromagnetic and hadronic showers in the calorimeters. This process is highly configurable, and all detector conditions during data-taking can be included in the simulation. The output of this step is recorded as detector hits.
2. The hits are digitalized to emulate the behaviour of the detector electronics, which also includes detector noise to match the expected signals from the real read-out.
3. All L1 algorithms configured in the trigger menu are simulated based on the digitalized signals, and a L1A decision is provided for each event. An output raw data object (RDO) is produced, which encodes all detector information in bytestream format.
4. The information stored in RDO format is used as input for the object reconstruction algorithms, as described in chapter 6, and also to simulate the decisions computed at the HLT.

The simulation of the ATLAS detector can be performed in two ways. The full description of all interactions of particles with matter can be included in what is called ATLAS Full Simulation (FullSim) [100]. This operation, however, is computationally expensive and might not be optimal for processing large volumes of simulated samples. A second approach, called ATLFAST-II (AFII) [101] is included, in which the detector response is parameterized in a simplified way in order to reduce the resources needed. It counts with a fast calorimeter simulation (FastCaloSim) that reproduces the behaviour of the ECal and HCal in a simplified way, while the full Geant4 simulation is used for the ID and the muon chambers. Another modality of the fast simulation, called ATLFAST-IIF [102], includes also a simplified emulation of the tracker and muon spectrometer through the fast ATLAS tracking simulation (Fatras). The output of both FullSim and AFII are written out in the necessary format to be comparable to data.

Different corrections are applied to the reconstructed simulated samples to make them as similar to the real data taken with the ATLAS detector as possible. They are typically scaled by

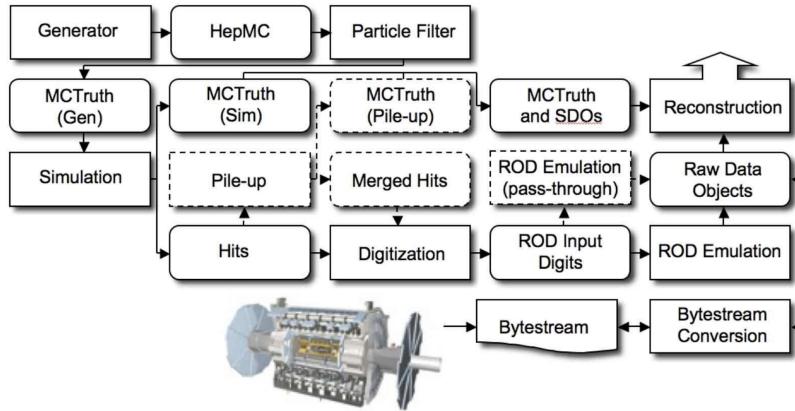


Figure 5.5: Scheme of the ATLAS detector simulation [100].

the cross section for the specific process computed at the highest order available. Additionally, a correction in the form of event weight is applied so that the pileup distribution in the simulated samples matches the LHC conditions, which is done on a yearly basis. Even after these corrections, the behaviour of the simulated samples can still show residual differences with the observed LHC collisions due to approximations performed during the simulation and reconstruction steps. In those cases, scale factors are derived by directly comparing data and simulation in well defined kinematic regions, which is done on a per-object basis to improve the performance of the reconstruction algorithms and selection efficiencies in the simulated samples.



# Chapter 6

# Object reconstruction in ATLAS

## 6.1 Introduction

As described in the previous chapters, the process of detecting the particles produced in the LHC collisions is long and complex, with several stages that involve different sub-detectors. The output of such process is the complex readout of the whole detector, made of signals generated by the readout electronics. In order to relate these signals to the physical particles that travelled through the detector, the events are *reconstructed* following a number of highly complex algorithms that identify leptons, photons and hadronically interacting particles.

The analysis described in this thesis relies on the hadronic activity in the event, and in particular the flavour of the quarks that originate the cascades inside the detector. Additionally, non-interacting BSM particles leave the detector carrying a fraction of the momentum transferred in the collision, resulting in net missing energy in the transverse plane, which is reconstructed with dedicated algorithms. As will be discussed in chapter 7, events with leptons are vetoed in this search, for which they need to be reconstructed prior to the analysis implementation. Information related to photons produced in the hard scattering is not used in the analysis, and therefore it will not be discussed further in this chapter.

This chapter is organized as follows. In section 6.2 the procedures to identify tracks from the ID and the formation of the primary vertices in the event are discussed. In section 6.3 the strategies to reconstruct electrons and muons are briefly discussed. Section 6.4 provides an overview of the reconstruction and calibration techniques for jets, and section 6.5 focuses on the strategy to identify the flavour of the quarks seeding the jets. Finally, in section 6.6 the approach to compute the missing transverse energy in the events is developed.

## 6.2 Tracks and primary vertex

The first step in the process of reconstructing the different objects present in an event is identifying all tracks left by charged particles in the detector. This information is essential for lepton identification, jet reconstruction and jet tagging. As particles travel through the ID, if they are charged, they will produce hits in the different tracking sub-systems that can be combined to reconstruct their trajectory. The path they follow is modified by the magnetic field generated by the central solenoid according to Lorentz's law, resulting in circular trajectories with radius proportional to the ratio between momentum and charge,  $r \propto p/|q|$ . The sign of the charge can be inferred from the direction in the transverse plane in which the trajectories are curved.

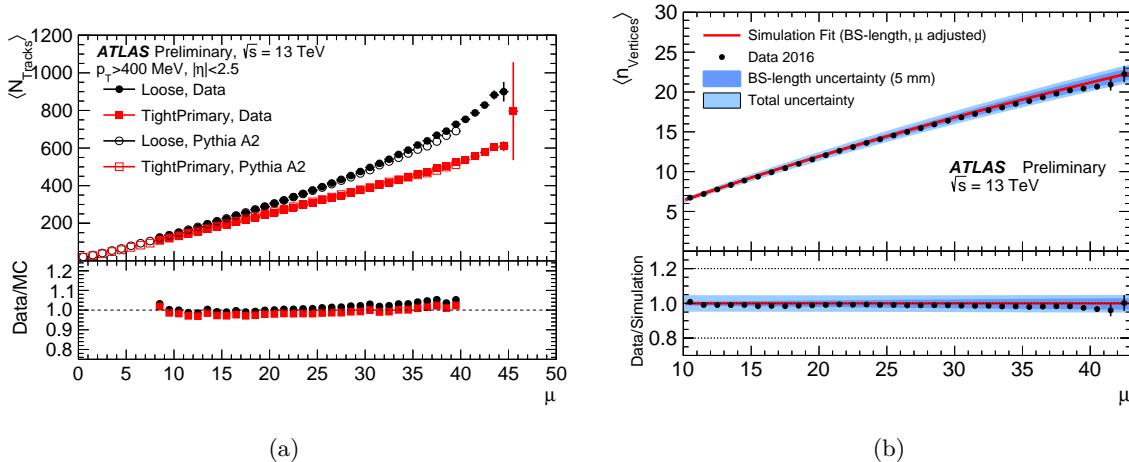


Figure 6.1: Average number of (a) reconstructed tracks and (b) reconstructed vertices per event as a function of the average number of collisions, comparing simulation with data collected during 2016. In the reconstruction of tracks, two different selection criteria are shown, used for large rejection of fake tracks during LHC Run 2. Figures taken from refs. [105, 106].

The tracks are typically reconstructed using an inside-out algorithm, starting in the inner layers of the ID, with combinations of 3 hits in the silicon detector as seed, and adding hits in the outer layers of the tracker that are compatible with the trajectory [103, 104]. New spatial points are added to the tracks by means of a Kalman filter that efficiently creates multiple track candidates per seed while suppressing contribution from random collections of hits (fake tracks). The ambiguity in the final track associated to a given seed is solved with a dedicated algorithm relying on statistical and machine learning methods and involving resolution information. The tracks associated with particles originating from the hard scattering are referred to as *primary tracks*. Tracks originating from long-lived particles, on the other hand, are reconstructed using hits in the TRT that are not compatible with primary tracks as seeds, using a outside-in algorithm that looks for compatible hits in the SCT and pixel detectors. Once reconstructed, tracks are parametrized by their polar angle  $\theta$ , azimuthal angle  $\phi$ , the ratio of charge and momentum  $q/p$ , and the distance in the transverse plane  $d_0$  and in the  $z$ -axis  $z_0$  to the originating point (transverse and longitudinal impact parameters, respectively). The average number of tracks as a function of pileup during 2016, for two different fake track suppression criteria, are shown in figure 6.1a.

Once the tracks for a given event have been reconstructed, they are used to find the points where particle interactions took place, which are referred to as *primary vertices* (PVs). The vertices in the event are reconstructed through a dedicated vertex finding algorithm employing an iterative process in which candidate PVs are formed based on track spatial information. For a given candidate PV, each track is assigned a weight depending on the compatibility with it, with less compatible tracks being associated a smaller weight. The position of the PV is recomputed using the weighted tracks, and tracks are re-weighted based on the new position of the candidate PV, repeating the process until an optimal convergence level is reached. After the last iteration, the tracks that are not compatible with the position of a given candidate PV are used as input in the reconstruction of other PVs, in a way that they are used only once. The reconstructed PV with highest  $p_{\text{tsum}}$ , computed from the associated tracks, is considered as the main PV, where the hard scattering process happened, while the rest are considered to arise from other interactions such as pileup in the collision. All objects in the event are expressed with

coordinates relative to the main PV. Reconstructed displaced vertices that are not compatible with the beam axis are named *secondary vertices*, and their identification plays a key role in flavour tagging and reconstruction of long-lived particles. The number of vertices in the event as a function of pileup is shown in figure 6.1b.

## 6.3 Lepton reconstruction

Leptons travelling through the detector leave characteristic signatures that aid in their reconstruction. In case of electrons, as discussed in chapter 3, they are expected to develop electromagnetic showers that are contained within the ECal, without much energy leaking to the hadronic calorimeters. Additionally, since they are charged, they are expected to leave hits in the ID compatible with the trajectory from the hard scattering to the electromagnetic shower. Muons, on the other hand, are expected to deposit much less energy on the calorimeters, and so their reconstruction relies on the tracks from the ID and hits in the muon chambers. Tau leptons, given their very large mass, have typically short lifetimes so they are expected to either decay leptonically, in which case the electron/muon can be reconstructed as detailed below, or hadronically into jets that are reconstructed as discussed in the following sections.

### 6.3.1 Electron reconstruction

The starting point for electron reconstruction is the identification of clusters in the electromagnetic calorimeter associated to the shower triggered by the incoming particle. This is performed using a clustering algorithm that takes the ECal towers of granularity  $\eta \times \phi = 0.025 \times 0.025$  as input. Prior to 2017, a *sliding-window* [107] algorithm was used in order to build the clusters. An improved approach was implemented in 2017, relying on variable-sized clusters for the reconstruction of photons and electrons. These *superclusters* are optimized to capture the energy from bremsstrahlung photons radiated by incoming electrons, as well as electron-positron pairs created by incoming photons. They are formed using the topo-clusters from jet reconstruction as input, where seeds for electron superclusters are required to have  $E_T > 1$  GeV and be compatible with a track containing at least 4 hits in the silicon tracker. Satellite clusters (topo-clusters within a window of  $\Delta\eta \times \Delta\phi = 0.125 \times 0.300$  centred on the seed) are added to the supercluster to account for secondary showers generated by the original electron.

Tracking information is used at this stage to disentangle electrons originated in the hard scattering from photons converting to an electron-positron pair. Tracks are extrapolated to the ECal, and are matched to the superclusters based on angular requirement. For electron candidates, if more than one track are compatible with the cluster, only one is selected on a cluster-track separation basis. Finally, selections based on the impact parameters of the track associated to the cluster are applied to ensure compatibility with the main PV in the event.

### 6.3.2 Electron identification

After the reconstruction step, the electron candidates are filtered in order to reject objects such as hadronic jets or converted photons since they could be wrongly reconstructed as electrons. An electron identification step [107] is therefore executed after the reconstruction, based on a likelihood discriminant built from information on the track quality, the longitudinal and lateral development of the electromagnetic shower and cluster-track matching, which is used to separate real electrons from fake ones. This discriminant is estimated for low-momentum electrons from

simulated  $J/\Psi \rightarrow ee$ , while for high-momentum electrons simulated samples for the processes  $Z \rightarrow ee$  and  $W \rightarrow e\nu$  are used.

Three working points, Loose, Medium and Tight, are optimized in bins of  $E_T$  and  $\eta$  based on the likelihood discriminant, with average electron efficiencies of 98%, 90% and 80%, respectively. The Medium and Tight working points include additional requirements on electron  $E_T$  and track quality.

### 6.3.3 Electron isolation

Prompt electrons produced in the hard scattering are expected to be more isolated in the detector than those originated from hadronic decays or photon conversions. In order to select electrons from the main event, a track-based isolation criterion is used, where the  $p_T$  sum of all tracks within a cone of radius  $\Delta R < \min(10\text{GeV}/E_T, 0.2)$  centred in the reconstructed electron is compared to the measured energy of said electron. The operating point used in the analysis discussed in this thesis corresponds to the loosest isolation requirement, with an efficiency of about 99% for all  $E_T$  range.

### 6.3.4 Electron energy calibration

The final reconstructed electrons are calibrated to correct different effects that can distort their measured energy. A summary of the steps followed in the energy scale calibration is given below:

- Data-driven corrections are applied to data in order to reduce the impact of non-uniform calorimeter response effects.
- MC-driven corrections are applied to both data and MC to account for energy loss by the electrons before entering the calorimeter. This calibration is performed by means of a BDT based on information about the electromagnetic shower in the calorimeter.
- An in-situ calibration is also included to account for differences between data and MC, derived with a template procedure on  $Z \rightarrow ee$  events. The energy scale calibration is afterwards applied to data, while the energy resolution corrections are applied to MC.

### 6.3.5 Muon reconstruction

Muons are expected to leave hits in both the ID and the MS, and therefore the first step in their identification is to perform separated reconstructions for the tracker, as described in section 6.2, and for the muon chambers [108]. For the MS, each muon chamber is checked for track patterns using the hits left in the read-out, and track segments are formed. Muon candidate tracks are then constructed by fitting the track segments together, starting from the middle layers of the MS and looking for compatible segments in the inner and outer layers. In most of the geometry of the detector, two matching segments are required in order to build a muon track, except for the transition region between the barrel and the end-cap, where a single high-quality segment is enough. Further quality criteria, based on a  $\chi^2$  test after fitting all hits in a candidate track, are performed in order to reject fake muons.

Four types of muon candidates are created, depending on how tracks from the ID and the MS are combined together:

- Combined muons (CB): combined tracks are formed by applying a global fit using hits from both the tracker and muon chambers. In general, an outside-in pattern recognition approach is followed, identifying tracks in the MS and extrapolating to the ID in order to look for matching tracks. An inside-out algorithm, starting in the tracker and extrapolating outwards towards the muon chambers, is in place as a complementary approach.
- Segment-tagged muons (ST): if a muon track candidate from the ID is matched to at least one segment in the muon chambers it is considered a ST muon. They are particularly useful for reconstructing low- $p_T$  muons and muons in detector regions with limited acceptance.
- Calorimeter-tagged muons (CT): similar to ST muons, CT muons are formed from a track in the ID matched to a calorimeter energy deposit compatible with the expected signature of muons, and are helpful in the reconstruction of muons in regions of the detector with reduced MS acceptance due to detector services and cabling. The CT muon identification criteria are optimised in a momentum range of  $15 < p_T < 100$  GeV, precisely in the detector region with lower acceptance by the muon chamber,  $| \eta | < 0.1$ .
- Extrapolated muons (ME): they are constructed solely from MS information in the region without coverage from the ID ( $2.5 < |\eta| < 2.7$ ). A requirement on the track extrapolation to be compatible with the main PV is applied, and the parameters of the muon track are afterwards defined at the interaction point, with corrections for the expected energy lost in the calorimeters.

### 6.3.6 Muon identification

Muons from the hard scattering are separated from those originating in hadron decays, such as pions or kaons, with requirements on the quality of the candidate tracks. Hadronic decays in the ID typically lead to muon tracks with a distinctive “kink”, resulting in lower track fit quality when segments from the MS are combined with hits in the tracker. Simulated  $t\bar{t}$  events are used to study different discriminants between prompt muons and those originated in hadron decays. The final discrimination is based on the quality of the track fit, the number of holes in the candidate muon trajectory (defined as points compatible with the trajectory where no hit was recorded) and momentum compatibility between ID and MS measurements.

Four identification working points, Loose, Medium, Tight and High- $p_T$ , are defined and provided centrally for physics analyses. They are inclusive categories, with each selection defined by tighter requirements applied on top of the previous working point. The prompt muon efficiency, as estimated from MC, ranges at low muon  $p_T$  from 78.1% in the tightest working point, High- $p_T$ , up to 96.7% for Loose. At high muon  $p_T$ , the efficiencies range between 80.4% to 98.1%.

### 6.3.7 Muon isolation

Prompt muons originated in the hard scattering are usually well separated from other objects in the event. Muons that are generated in the hadronic showers in jets, however, are expected to be aligned with said jet. Normally, it is desirable to select only muons that can provide information about the main event, thus isolation requirements are applied to remove muons coming from a jet. ATLAS provides centrally different isolation operating points, targeting different performance as required by the physics analyses. The OPs are defined by means of a selection on the ratio of the  $p_T$  sum of tracks within a cone centered in the reconstructed muon and the measured  $p_T$  of the muon. The one used in this analysis corresponds to the loosest requirement available, with an efficiency of 99% over most of the  $\eta$  and  $p_T$  ranges.

### 6.3.8 Muon energy calibration

After the reconstruction step, corrections are derived and applied to correct the muon candidate energy scale and resolution. They are computed from  $Z \rightarrow \mu\mu$  and  $J/\Psi \rightarrow \mu\mu$  events by performing a maximum-likelihood fit of the distribution of the invariant mass of the di-muon system and comparing data with simulated samples. The corrections are applied to MC events to improve the similarities with data, accounting for effects not taken into account in the simulation such as inhomogeneities of the magnetic field. The target precision for the muon energy scale and resolution after the calibration is at the per-mille and percent levels, respectively.

## 6.4 Jet reconstruction

As developed in chapter 5, strongly interacting particles evolve into complex showers, typically collimated due to the confinement property in QCD, as they travel through the detector. This makes particle identification hard, with the original partons being impossible to detect directly. These collimated sprays of hadrons, however, conserve some of the kinematic properties of the particles originated at the hard scattering, providing the key to understanding the physics processes undergone in the hadron collision. The cascades are typically reconstructed into single objects named *jets*, in a way that each jet usually collects the complete hadronic shower of one of the partons from either the hard scattering or the initial/final state radiation processes.

The hadronic cascades are characterized by large energy deposits left in the calorimeters within a cone of small area, leading to large signals in groups of adjacent cells denominated *topo-clusters*. For each event, clusters are identified by means of a dedicated algorithm [78, 109]. Each topo-cluster candidate starts with a seed formed by calorimeter cells with readout signal four times higher than the electronics noise, to which adjacent cells are added if they have a signal-to-noise ratio equal or higher than 2. Finally, the clusters are completed by adding a ring of guard cells around the candidate, regardless of their measured signal.

Topo-clusters are the first step in the chain of jet identification, since they provide a first measurement of the energy deposited by collimated showers. This approach alone, however, is not enough to reconstruct the total energy of the cascade in cases where the spread of the particles is wide enough to fall within two reconstructed topo-clusters. For this reason, more complex algorithms are employed to reconstruct jets using the previously identified topo-clusters as inputs.

At the time of this writing, two different approaches are followed by most ATLAS analyses in the reconstruction of jets, which rely on the same algorithms them but take different inputs. The so-called *EMTopo* jets, on the one hand, are formed using topo-clusters as inputs. These have been used for a large number of analyses published using LHC Run 2 data, with jets constructed exclusively using information from the calorimeters. On the other hand, a more complex approach has been explored, that involves information from both ID and calorimeters. It benefits from the higher momentum resolution of the ATLAS tracking system (eq. 3.8 in comparison to eq. 3.11) by combining topo-clusters with the compatible reconstructed tracks in the ID from charged particles in the hadronic cascade. In this way, the so-called *particle-flow* (PFlow) jets are constructed using as inputs the 4-vectors obtained from the accurate spatial coordinates of the topo-clusters and the high-resolution momentum measurement from the tracks. This approach leads to improved jet momentum measurements with respect to *EMTopo* jets, and also increases the acceptance to softer kinematic regimes since the ID can measure tracks for particles with energies as low as a few MeV.

### 6.4.1 Jet reconstruction algorithms

The algorithms used to form the jets take the 4-momenta of the corresponding input objects, whether topo-clusters or PFlow objects, with the center of the detector as origin and group them together in an iterative process. The most common approach to jet reconstruction in ATLAS is the family of  $k_t$  algorithms, in which all input objects are eventually sorted into jets based on their relative distances, with the possibility for a jet to contain more than one object. In this class of algorithms, the following operations are performed recursively:

1. For a given input object  $i$ , the distance to all other objects is computed as

$$d_{ij} = \min(k_{T,i}^{2n}, k_{T,j}^{2n}) \frac{\Delta R_{ij}^2}{R^2}, \quad (6.1)$$

where  $j$  labels all other inputs,  $k_{T,i}$  is the transverse momentum of object  $i$ ,  $\Delta R_{ij}$  is their angular separation, and  $R$  is a parameter that is related to the size of the final reconstructed jet. The label  $n$  is arbitrary, and defines the specific algorithm used.

2. For the same input object  $i$ , the distance to the beam axis is computed as

$$d_{i,B} = k_{T,i}^{2n}. \quad (6.2)$$

If there is any object  $j$  closer to input  $i$  than its distance to the beam axis,  $d_{ij} < d_{i,B}$ , then objects  $i, j$  are merged together in a single object.

3. If there is no object  $j$  closer to  $i$  than the beam axis, then  $i$  is defined to be a jet and removed from the pool of inputs.
4. The previous steps are repeated over all available inputs until all of them have been merged into jets.

The value of  $n$  defines the properties of the algorithm, which are reflected in the reconstructed jets as shown in figure 6.2. The case with  $n = 0$  defines the Cambridge-Aachen algorithm, in which the dependency with the input transverse momenta disappears and all jets are formed based exclusively on geometrical requirements. The case with  $n = 1$  is known as  $k_t$  algorithm, in which jets are grouped starting with soft inputs in the event. This can be problematic in the reconstruction of the jets arising in the hard scattering process, since pileup jets typically have lower  $p_T$  and therefore they are given priority in this approach. The most widely used case in ATLAS analyses is the anti- $k_t$  algorithm [110], with  $n = -1$ , since jets are reconstructed starting with the most energetic inputs. Additionally, the reconstructed jets in this case have a cone-like structure, approximating better the expected shapes of the hadronic showers in the detector.

The radius  $R$  of the jet is important in order to efficiently capture all the signals left in the detector by the cascade and provide an accurate measurement of the total energy of the process. For compressed topologies, where the products of the hard scattering are not expected to be extremely energetic, the so-called *small-R jets* are typically used, which are computed with a radius of  $R = 0.4$ . In topologies that lead to very energetic hadronic showers, a single small- $R$  jet might fail to capture all the energy deposited in the detector. Additionally, in these *boosted* regimes, two final state partons can be collimated enough so that their cascades overlap, with a total size larger than the volume covered by small- $R$  jets. In these cases, *large-R jets* can be constructed using a jet radius of  $R = 1.0$ , in order to better capture the whole cascade. These two are the only jet radii that are centrally supported in the ATLAS Collaboration, since each value of  $R$  needs dedicated calibrations.

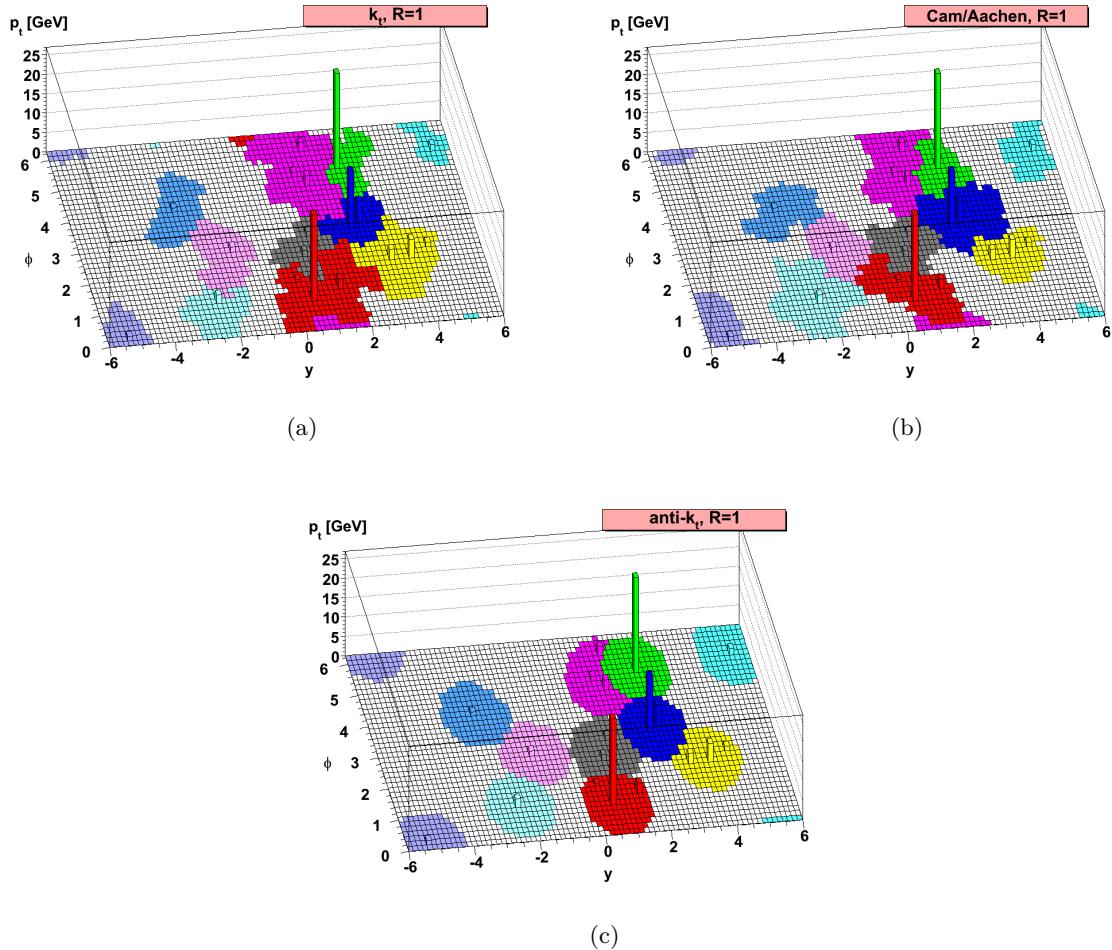


Figure 6.2: Calorimeter cells combined into jets using different algorithms in the  $k_t$  family: (a) the baseline  $k_t$  algorithm, corresponding to  $n = 1$ , (b) the Cambridge-Aachen algorithm, with  $n = 0$ , and (c) the anti- $k_t$  algorithm, with  $n = -1$ . Figures taken from ref. [110].

The inputs to the jet reconstruction algorithms can also be already formed jets. This is particularly useful when a different value for  $R$  is needed, since small- $R$  jets can be combined into *reclustered* jets with larger radii for optimal performance in a given kinematic regime. Differently to running the jet clustering algorithm for a custom radius, the reclustered jets benefit from the already provided calibrations for the input small- $R$  jets, and where the corresponding uncertainties can simply be propagated from those of the input jets.

#### 6.4.2 Jet calibration and associated uncertainties

As discussed in chapter 3, ATLAS uses non-compensating calorimeters, which do not account for the energy used in hadronic showers to break the QCD bound states and is therefore not recorded. The jet energy scale (JES) is corrected downstream after the reconstruction algorithms have been executed, together with corrections applied on the jet position. The sequence of corrections applied [111] is highlighted in figure 6.3, and summarized below:

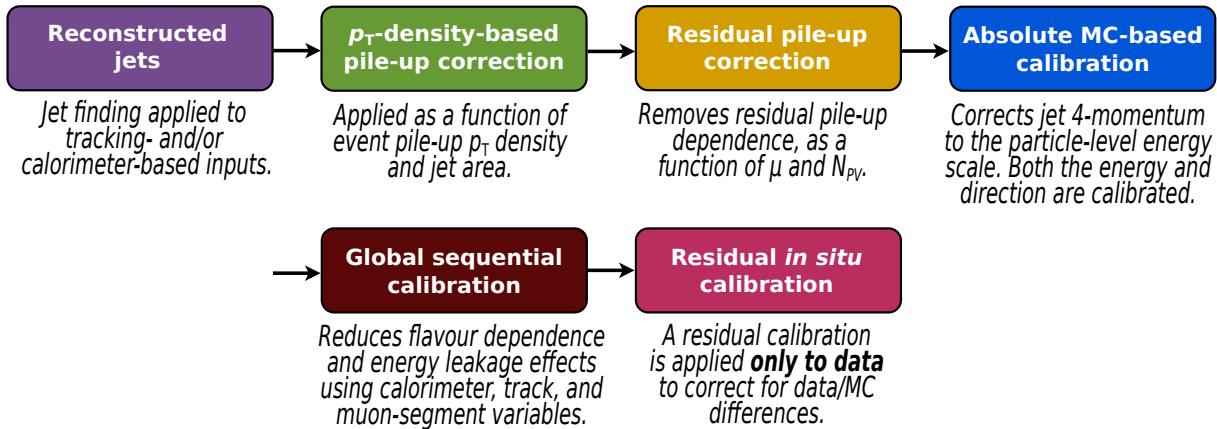


Figure 6.3: Sequence of calibration steps applied to reconstructed jets. Figure taken from reference [111].

- A first correction is applied to re-compute the coordinates of the topo-clusters with respect to the main PV in the event, instead of using the center of the detector as origin, resulting in improved  $\eta$  resolution. Additionally, the energy scale of the topo-clusters is calibrated at the electromagnetic scale.<sup>1</sup> In the case of EMTopo jets, this energy correction affects all jets. In the case of PFlow jets, the track-measured momentum is replaced only in low energy cases ( $< 100$  GeV), since for high energy jets the tracks typically provide a less accurate measurement than the topo-clusters. The energy correction therefore applies as well to high- $p_T$ PFlow jets.
- Soft particles arising in pileup collisions can be wrongly added to jets from the hard scattering, modifying their area and reconstructed energy. For this reason, a second correction is applied in two steps to suppress pileup effects. The first stage consists in estimating the average energy due to pileup collisions in the event from the number of reconstructed PVs, after which it is subtracted from the energy of the jets. At the second stage a residual correction is applied, using information from the number of interactions and the bunch crossing ID of the event to disentangle effects from in-time and out-of-time pileup.
- Next, absolute calibrations are applied to correct the jet  $\eta$  and energy scale. These corrections are derived from simulated MC events, in which information of the truth jet energy and direction are available. Truth jets are reconstructed in MC using the record of simulated particles as inputs, and they are later matched to the calorimeter jets based on an angular separation requirement,  $\Delta R < 0.3$ .
- The Global Sequential Calibration (GSC) [112] procedure is applied on top, which aims at improving the jet energy response using global properties of the jet in MC simulated di-jet events. In these corrections, information from the tracks associated to the jet is used to reduce the difference in the calorimeter response observed in different flavour jets. In addition, calorimeter energy deposit information is used to improve the jet energy resolution (JER), to which information from the muon chambers is added in the case of very energetic jets that leak out of the calorimeter. The GSC improves the performance of the jet energy response without changing the average JES in the di-jet sample used.

<sup>1</sup>Defined as the baseline energy scale at which energy deposited in the calorimeter through electromagnetic showers is known to be measured correctly.

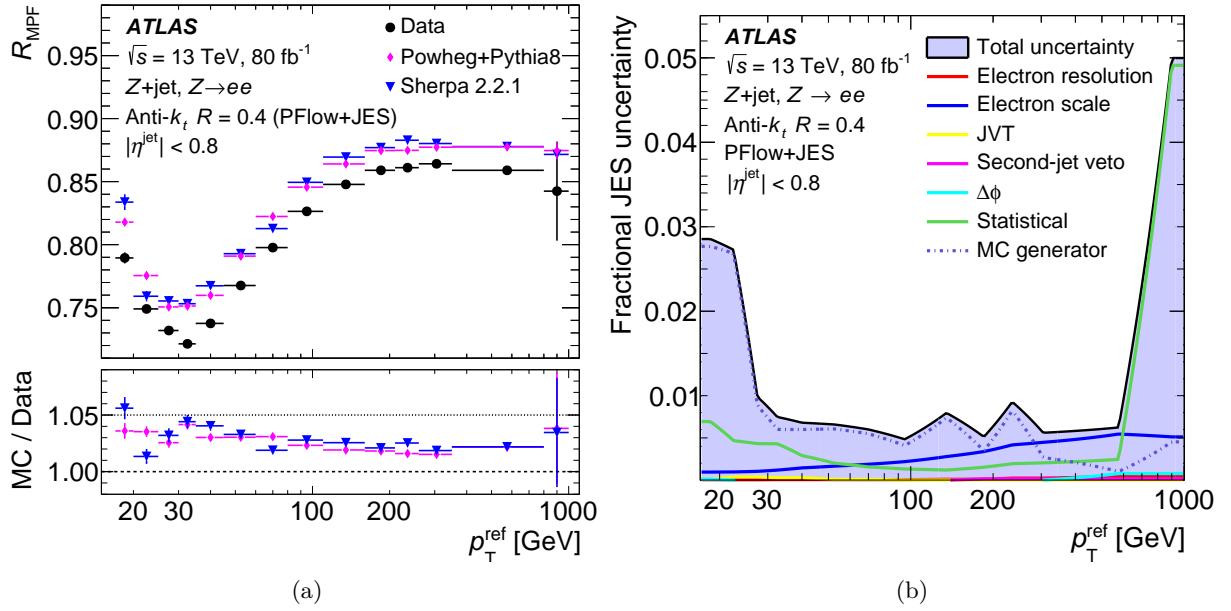


Figure 6.4: (a) Average PFlow jet response for  $Z(\rightarrow ee)$ +jet events, where data is compared to simulated events, with the ratio being the in-situ calibration applied to the simulation. (b) PFlow jets systematic uncertainties for  $Z(\rightarrow ee)$ +jet events, split in the contribution from the different sources and with the total uncertainty shown with the solid purple band. Figures taken from reference [111].

- Finally, an in-situ calibration is derived from data to correct the differences observed between MC simulation and the real detector read-out. These corrections are derived in events where jets are produced back-to-back with well calibrated objects, such as leptons or jets. Jet  $\eta$  is calibrated in the forward region of the detector with well measured central jets from QCD di-jet events. Additionally, the jet momentum in central jets is calibrated against well measured  $Z$  bosons or photons in  $Z/\gamma$ + jets production. For high- $p_T$  central jets, the calibration is performed against low-momentum jets in QCD multijet events. The size of the in-situ correction applied to PFlow jets derived with  $Z$ +jets events is shown in figure 6.4a.

These corrections are centrally applied to all samples processed in the ATLAS Collaboration, and the corresponding uncertainties are provided, which need to be propagated accordingly in each analysis. The ATLAS JES calibration in Run 2 comes with 80 dedicated systematic uncertainties, most of which are related to the in-situ calibration step. These account for jet modelling uncertainties, systematic uncertainties in the calibration of the well-measured objects that balance against the jets in the event or statistical uncertainties in the samples used. Other uncertainties are also included in the description, that account for systematic effects arising in the pileup correction. The size of the uncertainties on the JES in  $Z$ +jets events is shown in figure 6.4b.

#### 6.4.3 Jet Vertex Tagger

Large pileup in the LHC collisions, in addition to distorting the reconstructed jet energy spectrum, can also lead to a large number of reconstructed spurious jets that do not provide infor-

mation about the main event. In order to keep only the jets originated in the hard scattering, a track-based algorithm is used in ATLAS to identify jets arising from pileup effects, called *Jet Vertex Tagger* (JVT) [113, 114]. It relies on multi-variate analysis (MVA) techniques, incorporating a k-nearest neighbour (kNN) algorithm to estimate the probability of a given jet to have originated in the hard scattering process.

The inputs for the JVT algorithm are jet-based quantities computed as follows. After the main PV ( $\text{PV}_0$ ) has been identified, for each jet in the event the “jet vertex fraction” (JVF) is computed as

$$\text{JVF} = \frac{\sum_m p_{\text{T},m}^{\text{trk}}(\text{PV}_0)}{\sum_n \sum_l p_{\text{T},l}^{\text{trk}}(\text{PV}_n)}, \quad (6.3)$$

where the numerator is the scalar sum of the based transverse momentum of all tracks associated the jet and compatible with the main PV, while the denominator is the scalar sum of the transverse momentum of all tracks in the jet associated with other PV in the event. For a jet originating in the hard scattering, the JVF will be much larger than if it had originated in the pileup, since most of its tracks will be pointing to the main PV. However, in events where there are a very high number of pileup tracks, the JVF can become very small for both hard scattering and pileup jets, since the sum of the momenta of all tracks in the denominator would become large. The JVF is therefore further improved by adding a correction in the denominator to reduce the dependence on the number of pileup tracks in the event,  $n_{\text{trk}}$ , as

$$\text{corrJVF} = \frac{\sum_m p_{\text{T},m}^{\text{trk}}(\text{PV}_0)}{\sum_l p_{\text{T},l}^{\text{trk}}(\text{PV}_0) + \frac{\sum_{n \geq 1} \sum_l p_{\text{T},l}^{\text{trk}}(\text{PV}_n)}{10^{-3} n_{\text{trk}}}}. \quad (6.4)$$

Another useful discriminant used as input for the JVT algorithm is the ratio  $R_{p_{\text{T}}}$  of the transverse momentum of the tracks associated to a jet that are compatible with the main PV and its reconstructed transverse momentum,

$$R_{p_{\text{T}}} = \frac{\sum_k p_{\text{T},k}^{\text{trk}}(\text{PV}_0)}{p_{\text{T}}^{\text{jet}}}, \quad (6.5)$$

that is expected to be close to 1 if the jet is originated in the hard scattering, and become small if the jet comes from pileup.

The JVT algorithm computes the probability for a given jet to be originated in the pileup, making use of corrJVF and  $R_{p_{\text{T}}}$ . It is optimized separately for central jets, in the nominal JVT algorithm, and in the forward region ( $|\eta| < 2.5$ ), where a forward-JVT (fJVT) algorithm is used following a similar approach<sup>2</sup> [114]. The score for both of them is shown in figure 6.5 for pileup and hard scattering jets in MC simulation.

#### 6.4.4 Jet cleaning

Apart from the pileup effects that can result in spurious jets, other sources of signals in the detector read-out can lead to fake reconstructed jets. These are typically due to the so-called non-collision backgrounds, such as cosmic rays, in which a muon coming from the atmosphere crosses the detector with the subsequent energy deposited in the calorimeters, or beam-induced effects, such as muons generated from beam losses. In addition, detector effects such as coherent noise in the calorimeter read-out can also contribute to these sources for fake jets.

<sup>2</sup>In this case, due to detector limitations, the algorithm identifies pileup jets based on the fraction of missing transverse momentum in the direction of each jet

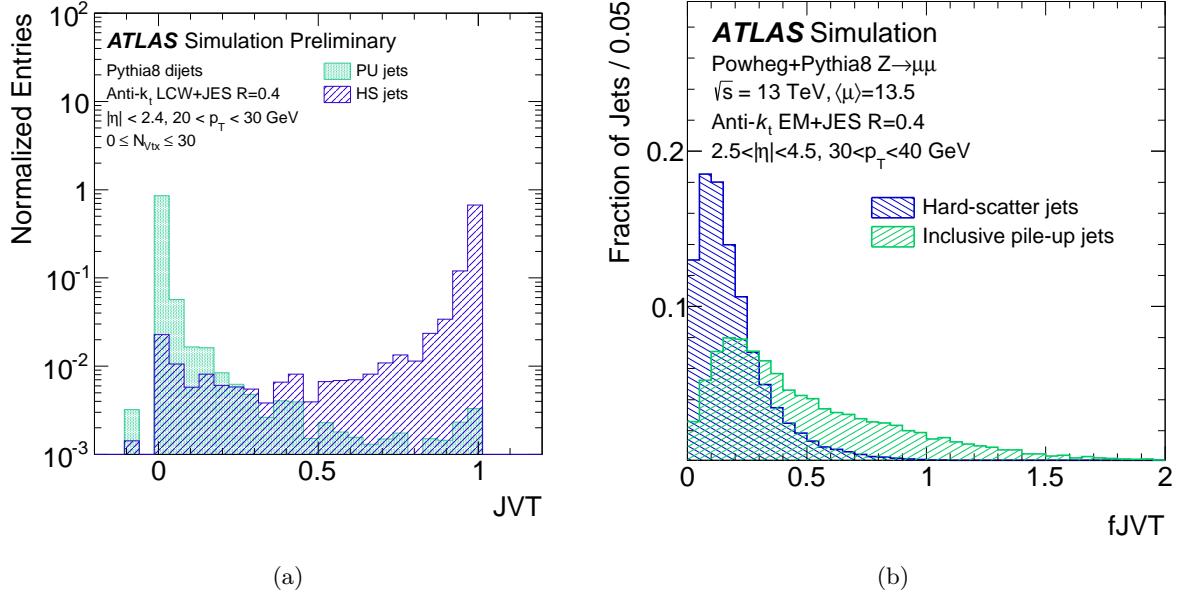


Figure 6.5: Output score of the (a) JVT algorithm for di-jet events, in a selection with  $20 \text{ GeV} < p_T < 30 \text{ GeV}$  [113], and (b) fJVT algorithm for  $Z + \text{jets}$  events in a selection with  $30 \text{ GeV} < p_T < 40 \text{ GeV}$  [114]. In both cases the jets arising from the hard scattering are separated from those arising from pileup, which allows to see the good separation achieved with the algorithms.

A jet cleaning procedure is implemented in physics analyses to reduce the presence of fake jets that are not originated in the LHC collision [115], generally applied as an event veto. In order to reject events where beam-induced background and coherent noise signals are present, certain criteria are implemented based on the energy ratio in different layers of the calorimeter. Furthermore, tracking information is also used to construct similar variables to  $R_{p_T}$  in order to assess the probability of each jet being originated in one of such processes mentioned above. Additional criteria on the ionization signal shape from the liquid argon calorimeters is also applied to reduce the contribution from coherent read-out noise.

These requirements are combined, using different thresholds, in *working points* with a definite efficiency for certain kinematic regimes. During Run 2 two working points were implemented, namely BadTight and BadLoose, with efficiencies of 95% and 99.5% for jets with  $p_T > 20 \text{ GeV}$ , respectively. For jets with  $p_T > 100 \text{ GeV}$  their efficiencies were 99.5% and 99.9%, respectively.

## 6.5 Jet-flavour tagging

As discussed above, the partons generated in the hard scattering evolve into hadronic showers inside the detector that are reconstructed as jets. The structure of these jets shows in general different traits depending on the flavour of the parton they generate from. When the jet is originated by a gluon or quarks  $u$  or  $d$ , they are commonly referred to as light-jets; the cases where the initial parton is a  $c$ -quark are called  $c$ -jets, while jets originated from a  $b$  quark are known as  $b$ -jets. The task of identifying the type of jet according to the original parton is called *flavour tagging*, and is of great relevance in physics analyses targeting final states with heavy flavour quarks.

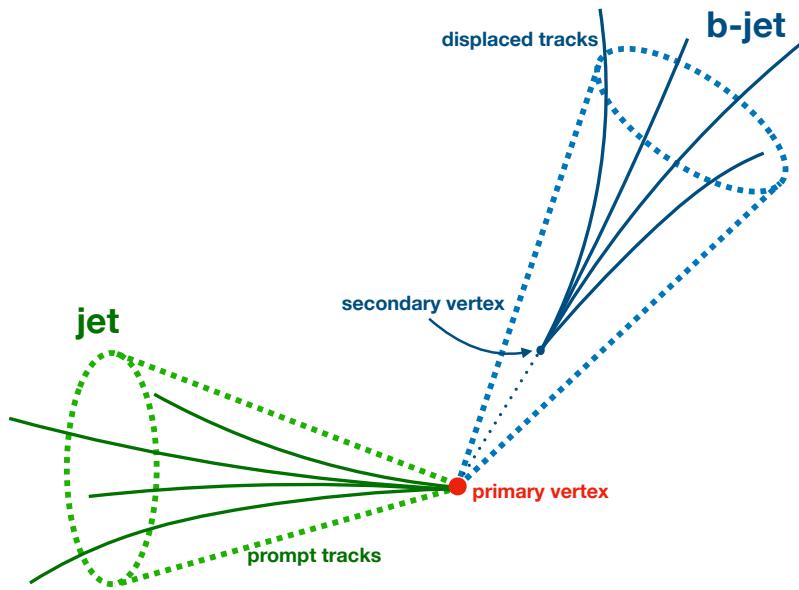


Figure 6.6: Scheme of the typical decay topology of jets originated by a  $b$ -quark from the hard scattering.

In particular, given the large amount of top quarks produced in the LHC and their decay chain involving  $b$  quarks, the correct identification of  $b$ -jets, usually referred to as  $b$ -tagging, is of importance and has been explored in depth in the ATLAS Collaboration. The general approach is to exploit the fact that  $b$ -hadrons have a characteristic lifetime of around 1.5 ps, allowing them to travel a distance inside the detector of about  $c\tau \approx 4.5$  mm. As shown in figure 6.6, this is reflected in a large impact parameter  $d_0$  compared to light-jets, with a secondary vertex close to the main PV. The addition during LHC LS1 of the IBL to the ATLAS pixel detector improved largely the impact parameter resolution, resulting in better performing  $b$ -tagging during Run 2 compared to Run 1.

### 6.5.1 $b$ -tagging algorithms

The identification of  $b$ -jets in ATLAS is done by means of low-level algorithms, that exploit the basic features of  $b$ -hadrons, and high-level algorithms, that combine the outcome of the low-level ones to achieve high  $b$ -tagging efficiency and low rates of jets being wrongly tagged (mistag) [116]. The low-level algorithms take the jet tracks as input, and can be classified in 3 families, depending on which of the following features they exploit:

**Impact parameter:** as mentioned above, the tracks associated to  $b$ -jets have typically a larger impact parameter  $d_0$  than those of light-jets, due to the lifetime of  $b$ -hadrons. This offers a good handle for the correct identification of  $b$ -jets, and is exploited by the complementary algorithms IP2D and IP3D [117]. They both build discriminating variables based on the significance of both longitudinal and transverse impact parameters and their correlations, and include MC-derived templates for the probability distribution functions for  $b$ -,  $c$ - and light-jets in order to compute the contribution for each track associated to a jet.

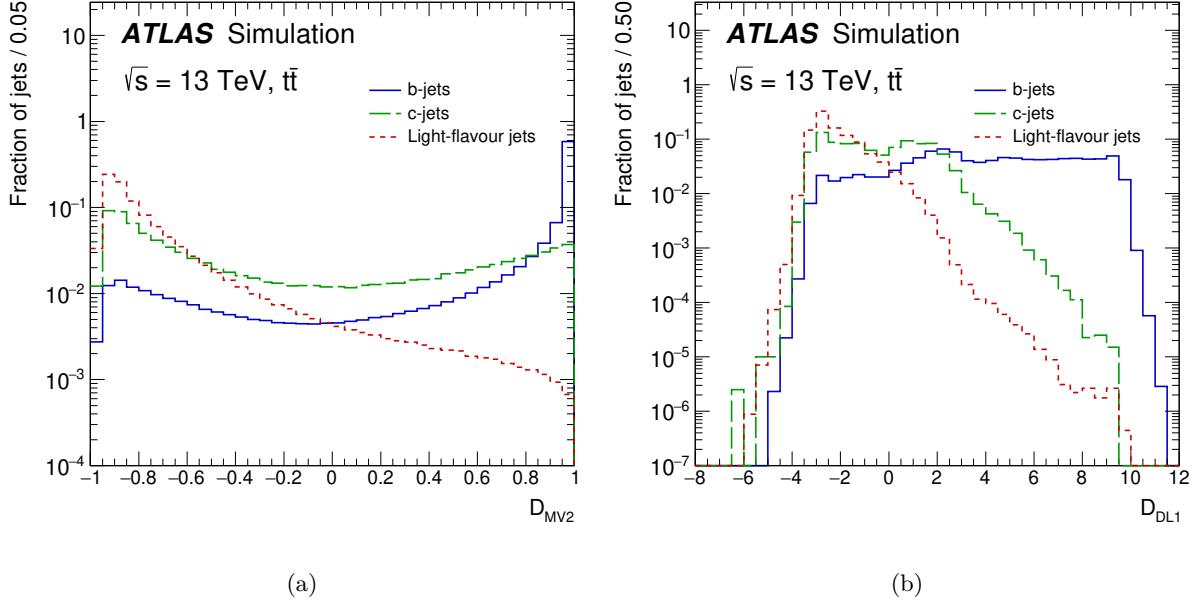


Figure 6.7: Output discriminant computed by the (a) MV2c10 and (b) DL1 *b*-tagging algorithms for  $t\bar{t}$  simulated events, split in the different jet flavours at truth level. Figures taken from ref. [116].

**Secondary vertex:** since jets are always reconstructed with the main PV as origin, the defined *b*-jet cone by the clustering algorithm must include a second vertex where the *b*-hadron decays. This type of topology is exploited by the SV1 algorithm [118], which uses track information to build a likelihood discriminant in order to identify jets with a vertex compatible with the decay of a *b*-hadron. This is done by testing all track pairs in a jet against a two-track vertex hypothesis, rejecting pairs that are compatible with long-lived resonances, such as  $K_0$  or  $\Lambda$  hadrons, photon conversions or hadronic interactions in the detector.

**Multiple vertices topology:** since *b*-quarks have a large branching ratio to decay to *c*-quarks, and D-hadrons also travelling some distance inside the detector before decaying, *b*-jets are expected to contain a third vertex due to the D-hadron decay. This type of topology is exploited by the JetFitter (JF) algorithm [119], which looks for the intersection between tracks and the jet axis in order to compute the probability of there being a third vertex inside the jet cone.

The outputs of these low-level algorithms are combined in high-level ones, employing MVA techniques to compute the likelihood of jets being initiated by *b*-quarks. Two main high-level algorithms have been used in ATLAS during LHC Run 2:

- The MV2c10 algorithm [120], belonging to the MV2 family, relies on boosted decision trees trained with  $t\bar{t}$  simulated events to discriminate between *b*-jets and *c*- and light-jets. The output of the MV2c10 algorithm for each jet is a score between 0 and 1 representing the probability of such jet to be originated from a *b*-quark. The distribution of the MV2c10 output for *b*-, *c*- and light-jets for a simulated  $t\bar{t}$  sample is shown in figure 6.7a. This was the recommended *b*-tagging algorithm in ATLAS for most part of LHC Run 2.
- The DL1 algorithm [117] relies on an artificial deep neural network (DNN) trained on a

hybrid simulated sample where all quark flavours are treated equally. The output is a multi-dimensional array that provides the probability for each jet to be a real  $b$ -,  $c$ - or light-jet. The distribution of the DL1 outputs for the three jet flavours, evaluated on a simulated  $t\bar{t}$  sample, is shown in figure 6.7b. A new version of this tagger, namely DL1r, includes as an additional input a new IP reconstruction algorithm based on recurrent neural networks (RNNIP) [121]. This has become the default  $b$ -tagging algorithm for the late Run 2 physics analyses, partly due to lower mistag rates compared to MV2c10.

The cuts over the output of a  $b$ -tagging algorithm are combined into the so-called *operating points* (OP), that are defined to provide specific  $b$ -jet acceptance efficiency in simulated  $t\bar{t}$  samples. Four OP are provided centrally to be used in physics analyses, with their corresponding calibrations, that translate into  $b$ -tagging efficiencies of 60%, 70%, 77% and 85%. The larger the acceptance of real  $b$ -jets, however, the larger the probability of accepting mistags as the selection becomes looser. The rejection of  $c$ - and light-jets as a function of the  $b$ -jet efficiency are shown in figure 6.8 for both MV2c10 and DL1 algorithms.

### 6.5.2 Calibration and uncertainties

The  $b$ -tagging algorithms need of the truth information about the flavour of the jets used in the training, and therefore they need to be studied first using MC simulation. To address differences with real data, dedicated calibrations are derived that result in SFs with their corresponding uncertainties. These SFs are later on applied to MC samples downstream in the physics analyses for a more accurate description of the recorded data. The uncertainty on the SFs include statistical errors in the data and MC samples used in their estimation, jet reconstruction systematic uncertainties propagated through the  $b$ -tagging algorithm, and modelling systematic uncertainties to account for differences in the simulated physics process compared to the real collisions. The obtained SFs for the DL1 algorithm at the 70% OP, with their uncertainties, for each jet flavour are shown as a function of jet  $p_T$  in figure 6.9.

The efficiency and mistag rate for the different jet flavours are derived in dedicated analyses. The corresponding SFs are estimated in well measured processes comparing data and MC, and are parametrized as a function of the jet flavour and its  $p_T$  and  $\eta$ , in flavour-dependent analyses:

- The  $b$ -tagging efficiency is extracted from  $t\bar{t}$  di-leptonic events with exactly two jets, where a combinatorial likelihood approach is adopted to extract the jet flavour composition in the sample at the same time than the  $b$ -tagging efficiency [123]. The  $b$ -tagging probability is determined for a jet momentum range between 20 GeV and 600 GeV. Flavour tagging in jets with  $p_T > 600$  GeV need to make use of a high- $p_T$  extrapolation in order to determine the probability of being tagged. Scale factors are determined by comparing the efficiency estimated from simulation to the values observed in data, which are measured to be very close to 1 with uncertainties ranging between 1% and 8%.
- The mistag rate for  $c$ -quarks is estimated from  $t\bar{t}$  semi-leptonic events, where one of the  $W$  bosons decays to a  $c$ - and a  $s$ -quark [124]. A likelihood fit is used to sort the jets into the corresponding  $t\bar{t}$  decay products and the  $c$ -jet mistag rate is extracted from those associated to the hadronic  $W$  decay. The SFs are extracted from the comparison between observed data and simulation values, which for the MV2c10 tagger are measured to be between 1.1 at 85% WP and 1.6 for 60% WP. A similar analysis is performed for the DL1r tagger, for which the SFs are close to one, with efficiencies ranging between 3% and 17% [125].

- The mistag rate for light-jets is studied with two methods [126]. On the one hand, the negative-tag calibration relies on the comparison between the mistag rate in a data sample enriched with light-jets to simulation, where jets are selected based on the performance of the IP2D and IP3D algorithms after flipping the sign of their impact parameters. The adjusted-MC method, on the other hand, relies on adjusting for each jet the impact parameter performance in MC simulation to that observed in data, and re-executing the  $b$ -tagging algorithms. The estimated SFs for MV2c10 are around 2 for both methods. A similar analysis is about to be published for DL1r tagger. They have been determined for the 70% WP to be very close to 1, smaller than those for MV2c10.

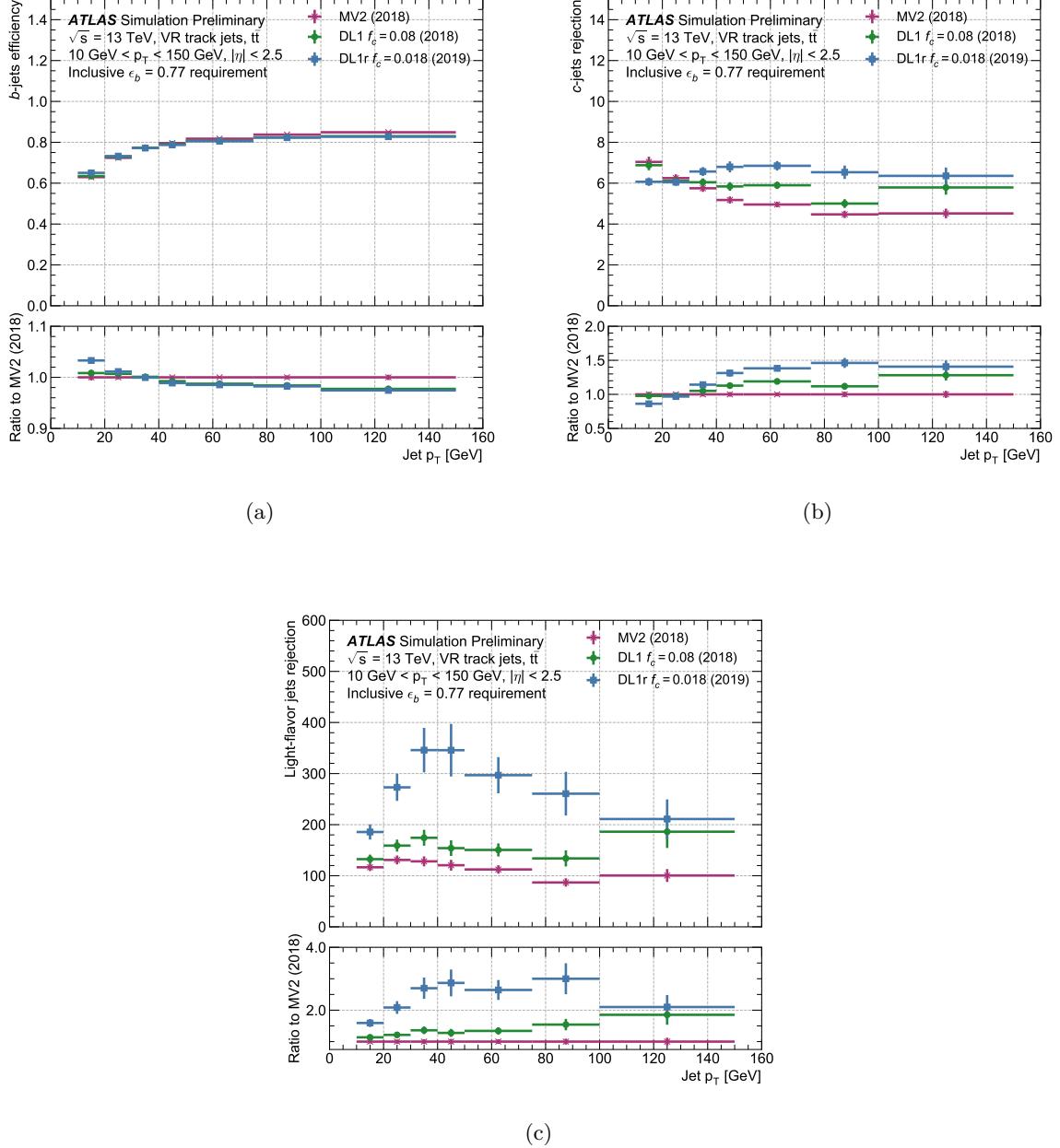


Figure 6.8: (a) Identification efficiency for  $b$ -jets, and rejection rate for (b)  $c$ -jets and (c) light-jets for the MV2c10 and DL1 taggers as evaluated in a simulated  $t\bar{t}$  sample with a requirement over the jet  $p_T$  of  $20 \text{ GeV} < p_T < 150 \text{ GeV}$ . Figures taken from ref. [122].

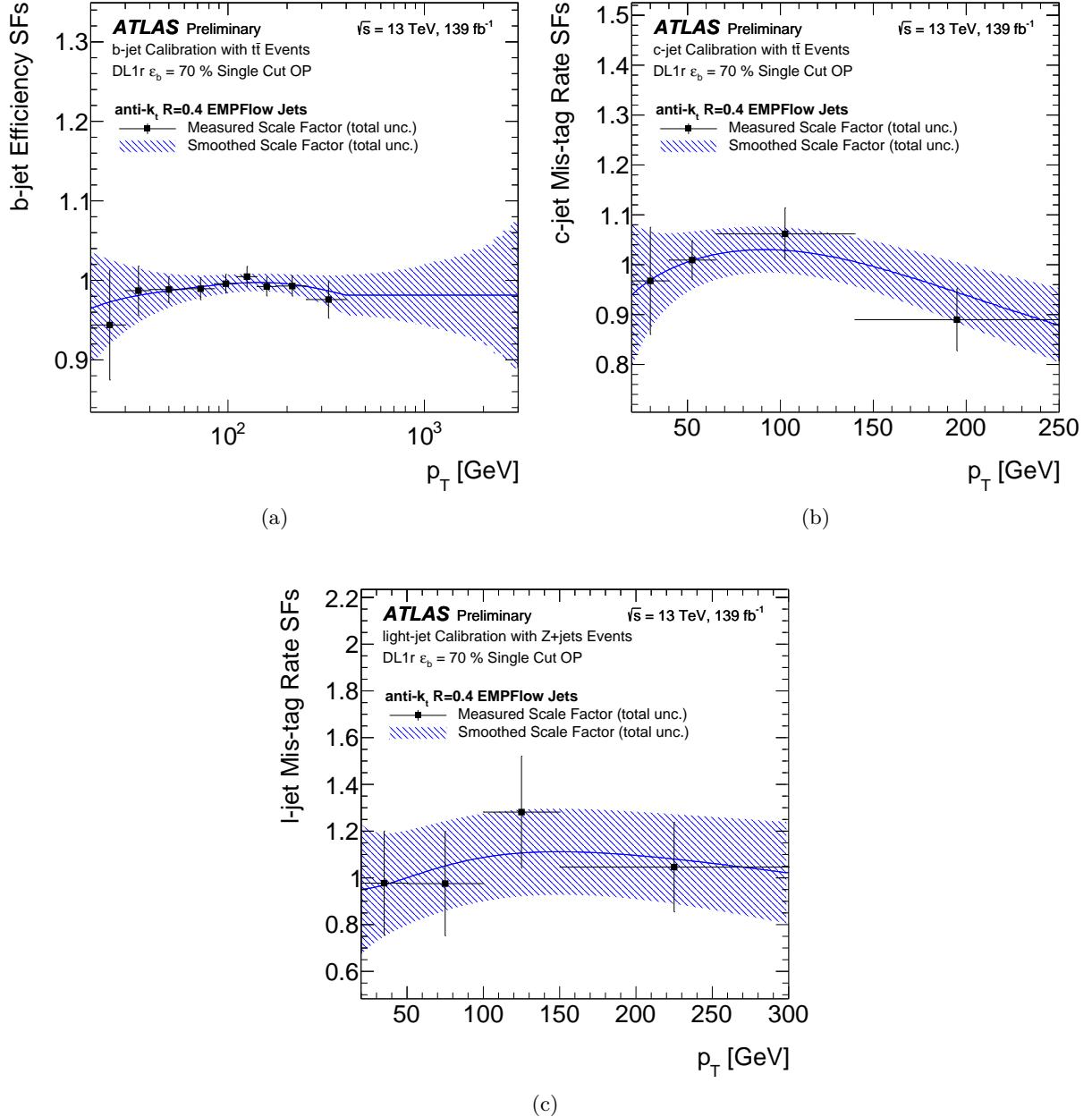


Figure 6.9:  $b$ -tagging scale factors for (a)  $b$ -jet efficiency, (b)  $c$ -jet mistag rate and (c) light-jet mistag rate at 70%OP, as a function of jet  $p_T$ . Figures taken from ref. [127].

## 6.6 Missing transverse energy

Weakly interacting particles, such as neutrinos, or BSM hypothetical particles with no electroweak or strong couplings, do not leave any trace in the detector as they travel through after being produced in the collisions. They cannot be reconstructed explicitly, and therefore the only way to measure the energy they carry is through the energy imbalance in the event. Since the transverse momentum of all particles in the event needs to vanish when added vectorially, the missing energy in an event can be defined as the negative vector sum of the momentum of all reconstructed particles, after the corresponding calibrations are applied:

$$\vec{E}_T^{\text{miss}} = - \sum \vec{p}_T^e - \sum \vec{p}_T^\gamma - \sum \vec{p}_T^\tau - \sum \vec{p}_T^{\text{jets}} - \sum \vec{p}_T^\mu - \sum \vec{p}_T^{\text{soft-terms}}, \quad (6.6)$$

where the term  $\sum \vec{p}_T^{\text{soft-terms}}$  is included to account for all energy measured that is not included in any of the reconstructed objects. The magnitude of the vector  $\vec{E}_T^{\text{miss}}$  is the total energy imbalance in the transverse plane, and is commonly denoted as  $E_T^{\text{miss}}$ .

Different approaches are followed to reconstruct the soft term [128, 129]. An algorithm is in place to compute the *calorimeter soft term* (CST) based on energy deposits in the calorimeters that are not associated to any reconstructed object in the event. The recommended algorithm for physics analysis, on the other hand, computes the *track soft term* (TST) using solely track information for better pileup treatment. Since neutral particles do not leave any trace in the ID, the TST does not include information about the momentum carried by those. The performance of the  $E_T^{\text{miss}}$  reconstruction algorithm during Run 2 is shown in figure 6.10a, where data is compared to MC simulated events in a selection that targets the process  $Z \rightarrow ee$  in association with jets.

It is important to note that mismeasurement in the energy of reconstructed objects has a direct impact in the amount of  $E_T^{\text{miss}}$  observed in a event, either due to detector resolution effects or because of decay chains involving neutrinos. Both real and “fake” sources of  $E_T^{\text{miss}}$  result in indistinguishable contributions to the missing transverse momentum vector. In many physics analyses, specially in searches for BSM scenarios with stable non-interacting particles in the final state, understanding the source of  $E_T^{\text{miss}}$  can be crucial in order to reject events produced in background SM processes. For this reason, the  $E_T^{\text{miss}}$  significance is constructed in order to test the hypothesis that the observed  $\vec{E}_T^{\text{miss}}$  corresponds to fake sources against the hypothesis that it comes from actual invisible particles [131]. It is built based on the log-likelihood ratio for the two hypotheses, where the likelihood functions depend on the types of objects reconstructed in the event, their resolution and their kinematic properties. Due to this dependence on all reconstructed particles, this variable is commonly referred to as *object-based  $E_T^{\text{miss}}$  significance*, and its performance over the full Run 2 dataset in both data and MC simulated events, in a selection targeting the process  $Z \rightarrow ee$ , can be observed in figure 6.10b. Typically, low values of object-based  $E_T^{\text{miss}}$  significance correspond to fake sources of  $E_T^{\text{miss}}$  while high values correspond to actual invisible particles carrying the missing momentum observed.

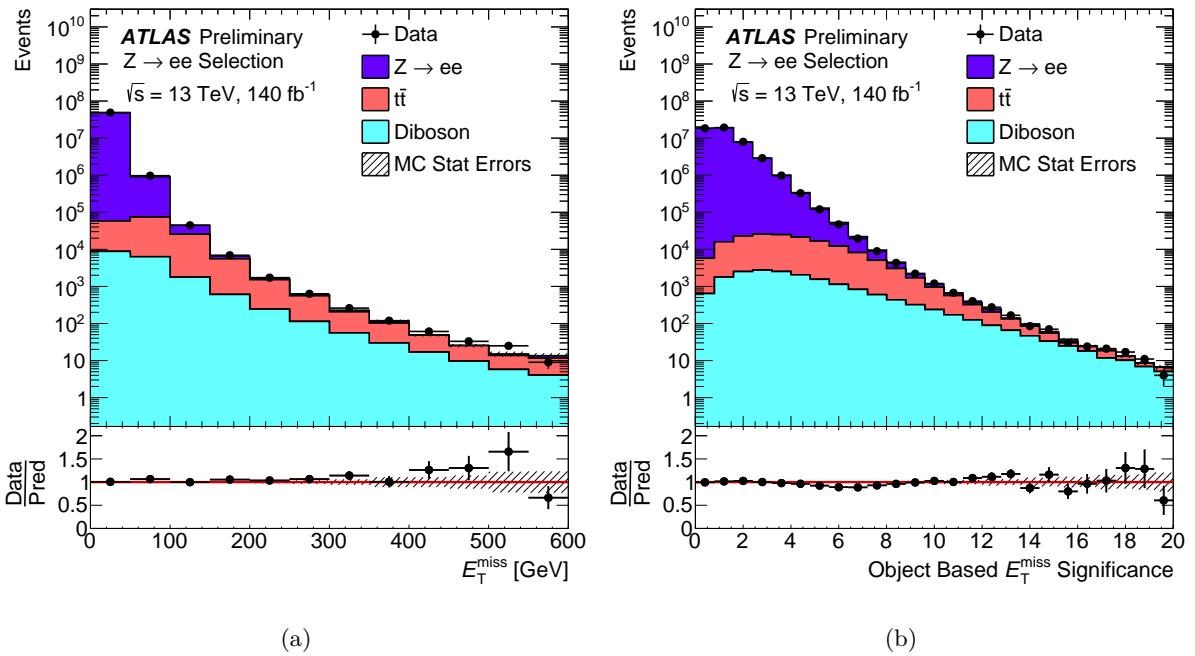


Figure 6.10: Performance of (a) reconstructed  $E_T^{\text{miss}}$  and (b) object-based  $E_T^{\text{miss}}$ -significance in  $Z(\rightarrow ee) + \text{jets}$  events, where recorded data is compared to MC simulated samples.  $E_T^{\text{miss}}$  is computed including the TST and taking into account PFlow jets with  $p_T > 30$  GeV. Figures taken from ref. [130].

# Chapter 7

## Search for higgsino pair production

### 7.1 Introduction

As discussed in chapter 2, naturalness arguments justify a mass hierarchy in the MSSM with light neutralinos ( $\approx 1$  TeV) dominated by the higgsino component, with slightly heavier third-generation squarks and gluinos. This type of scenarios have been explored for a long time at LEP and Tevatron, and now they are under investigation in searches using LHC data, resulting in large excluded regions of the available parameter space for strongly-produced particles. The SUSY electroweak sector, however, due to significantly lower production cross section, has been hard to study so far, and it is only now thanks to the large energies reached in the LHC and the large integrated luminosity recorded so far that it can be explored to the limits of natural SUSY.

Previous ATLAS analyses using partial Run 2 data have excluded a portion of the available phase space in the SUSY electroweak sector [132–134]. In particular, searches targeting scenarios with higgsino-dominated  $\tilde{\chi}_1^0$  in the context of gauge-mediated SUSY breaking have previously excluded neutralino masses of up to 800 GeV [135]. In that search, two approaches were combined for maximal sensitivity to a wide range of mass hypotheses: one channel targeting high- $E_T^{\text{miss}}$  events for sensitivity to high  $\tilde{\chi}_1^0$  mass, and one channel relying on events with large  $b$ -jet multiplicity sensitive to low  $\tilde{\chi}_1^0$  mass. A  $2\sigma$  excess was found for  $m(\tilde{\chi}_1^0) \approx 250$  GeV, in the region with liminal sensitivity from both channels.

The search presented in this work, which uses the full LHC Run 2 dataset recorded with the ATLAS detector amounting to an integrated luminosity of  $139 \text{ fb}^{-1}$ , targets similar topologies, focusing on the high- $E_T^{\text{miss}}$  region that corresponds to high  $\tilde{\chi}_1^0$  mass. The analysis strategy has been optimized from scratch in order to significantly enhance the sensitivity in the high-end of the neutralino mass spectrum. In addition, different aspects of the analysis have been improved to increase the sensitivity at low  $\tilde{\chi}_1^0$  mass as well, targeting the excess of the previous publication, leading to a more complex trigger strategy and robust background estimation in a kinematic region previously inaccessible to the high- $E_T^{\text{miss}}$  channel.

The analysis strategy relies on the definition of dedicated kinematic regions that target different physics processes. In general, a set of *signal regions* (SR) are built for each  $\tilde{\chi}_1^0$  mass hypothesis targeted which include a large percentage of signal events with low contribution from SM background processes. These regions are used to test the hypothesis of only SM processes being present (null hypothesis) against the hypothesis of both SM and BSM physics existing in the recorded data. A set of *control regions* (CR) is constructed with the requirement of enhanced

acceptance of background events with low levels of signal contamination. These regions are used to derive the normalization of the main SM processes present in the signal regions, and are required to be kinematically close to those. A final set of *validation regions* (VR) is used to test the extrapolation of the corrections derived in the control regions to the signal regions. In general, dedicated CR/VR are used to target a specific SM process for better accuracy. This is performed by means of statistical methods, relying on fits to data to derive corrections and on hypothesis test techniques. The statistical treatment of the analysis regions in this search is discussed in chapter 8.

This chapter is organized as follows. In section 7.2 a brief overview of multi-variate analysis techniques is provided, as this analysis heavily relies on some of them. In section 7.3 the signal model is discussed, with the particular topology targeted in this search. The first event selection used for early background rejection and providing a basis for the optimization of the analysis is developed in section 7.4, while the relevant kinematic variables in this search are detailed in section 7.5. The main background processes are discussed in section 7.6, where details about their generation and associated modelling uncertainties are provided. Additionally, a novel strategy used to estimate the contribution from QCD multijet processes is discussed in section 7.7, relying on a data-driven method combined with machine learning techniques for highly accurate predictions. The agreement between data and the background estimate is provided in section 7.8 for loose kinematic cuts where the expected signal contribution is low. Finally, the main discrimination between signal and background processes is discussed in section 7.9, with detailed information about the control, validation and signal region definition procedure.

## 7.2 Multi-variate analysis techniques

The problem of separating signal from background in LHC physics analyses can profit from the ability to identify two different classes  $\mathcal{C}_0, \mathcal{C}_1$  within a given data sample provided by multi-variate analysis (MVA) techniques. They aim at exploiting not only the different input features in a data collection, but also their correlations in order to provide optimal separation between signal and background events. Since the analysis presented in this chapter relies heavily on them, this section provides an overview of key concepts that will be used to develop the strategy followed in this search.

### 7.2.1 Machine learning

MVA techniques generally involve complex algorithms with a large number of parameters that determine the performance of the classification. The task of finding the most optimal set of parameters is usually complex and the most efficient way to carry it out is to allow the algorithm in question “learn” from an input dataset that is representative of the data that it will be evaluated on. This process, referred to as *training*, allows to tune the algorithm parameters in an iterative way, constantly evaluating its performance and modifying the values accordingly. The training can be done providing the algorithm with a label for each element of the input dataset that encodes its true classification, in which case the prediction for each element is compared to the real value to determine the accuracy of the algorithm in what is called *supervised learning*. This is done by minimizing a metric or loss function, which encodes the differences between predicted and true classification for each element in the dataset. *Unsupervised learning*, on the other hand, is an approach where the input datasets are not labelled, typically adopted when

the algorithm is expected to learn some key features of the dataset in order to find, for example, anomalies present in similar input data.

The learning process for an algorithm needs to be carried out in a different dataset to the one on which it is evaluated, in order to avoid introducing any bias in the predictions provided. In the case of physics analysis, MVA algorithms are typically trained using simulated events. This approach allows for supervised learning since each event in the training corresponds to a well defined physics process, and also provides an independent dataset for the algorithm to learn from. The final trained algorithm is evaluated with recorded collisions in order to identify the presence of signal events.

Given the limited number of events in any sample used in the training, it is possible for the algorithm to learn specific features that appear in the dataset due to statistical fluctuations. In these situations, commonly referred to as *overtraining*, the expected performance of the algorithm might be better than the actual separation provided when it is evaluated on a different sample. In order to detect cases of overtraining, the input dataset is generally divided in two samples, one used for the training and the other to test the expected performance of the algorithm.

In addition to the key parameters of the algorithm, which are determined through the learning process, the “structure” of the algorithm can also impact its performance. It is encoded in the so-called *hyperparameters*, which are user-defined and do not change during the training process. Due to the complexity of these algorithms, the usual approach to determine the most optimal combination of hyperparameters is to perform a scan over a wide range of values, where for each possibility the algorithm is fully trained and the expected performance for each combination is compared. The structure of an algorithm is related to the amount of overtraining present after the learning process, where algorithms with higher complexity typically result in higher accuracy but also being more prone to overtraining.

### 7.2.2 Boosted decision trees

A common algorithm used in multi-variate analysis are decision trees. They are simple tools that can be combined with learning techniques to provide good separation between two targeted classes  $\mathcal{C}_0, \mathcal{C}_1$ . They act on the whole dataset to design a series of selection cuts optimized for maximum separation at each step. For each element in the dataset, the algorithm receives a set of  $N$  input features,  $\vec{x} = (x_1, x_2, \dots, x_N)$ , and a label indicating the class to which it belongs,  $y^{\text{true}}$ , where the parametrization typically follows the convention of using  $y^{\text{true}} = 0$  for  $\mathcal{C}_0$ , or “background”, and  $y^{\text{true}} = 1$  for  $\mathcal{C}_1$ , or “signal”. The algorithm then forms a root *node* including the whole dataset, and performs a selection cut based on one of the input features, splitting it into two smaller nodes. The resulting nodes are further split based on other selections on the input features, without constraints on repeating the same ones. The objective is to reach nodes where either one of the classes to separate dominates over the other, in which case the node is called *leaf* and the splitting process stops. The sequence of selections applied to arrive to each leaf is commonly referred to as *branch*. The purity  $P$  of a node is defined as the fraction of elements belonging to the signal class  $\mathcal{C}_i$ , and an optimal decision tree is that for which the leaves have large purity. A scheme of a decision tree is shown in figure 7.1a.

The selection cuts are optimized for each node in order to provide maximal separation between the two classes at each splitting step. The choice on the feature used and the split point is performed by means of a metric that is defined by the user. One example of a widely

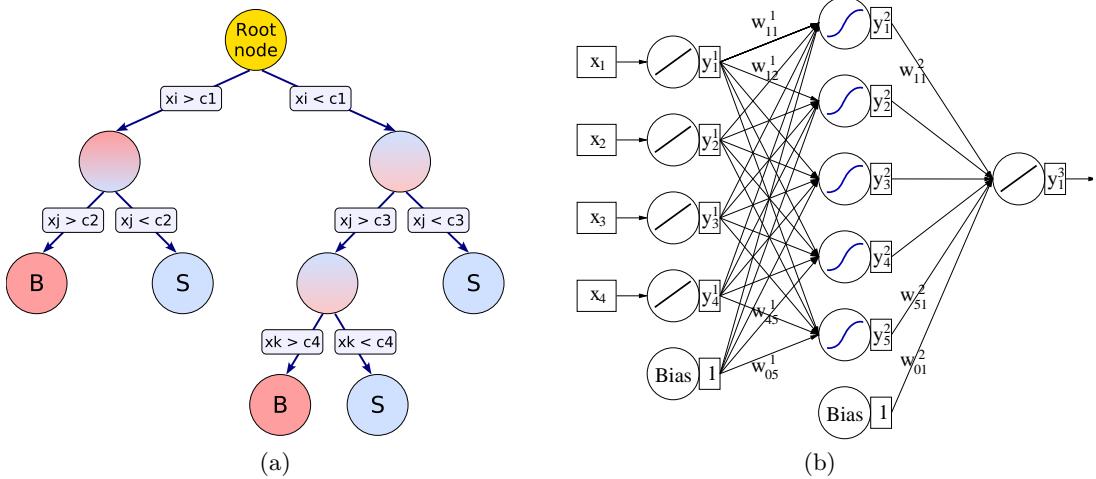


Figure 7.1: Schematic views of (a) a decision tree and (b) a neural network, used for a classification problem [136].

used metric is the Gini index,

$$G = P(1 - P), \quad (7.1)$$

and selection cuts are chosen so that they maximize  $G$  at each node. Another example is the *cross entropy*, defined as

$$E = -(P \log P + (1 - P) \log(1 - P)). \quad (7.2)$$

The gain in each node splitting is defined as the difference in the metric between the initial node and the two resulting ones, and usually a scan in different possible cuts is performed to find the optimal gain. Different hyperparameters of the tree, such as the *depth*, which encodes the maximum number of nodes in a branch, or the number of bins used to scan the optimal splitting point in each node, can be configured by the user and have a direct impact in the performance of the algorithm.

Usually one single decision tree does not provide enough separation power, and so they are commonly referred to as *weak learners*. However, multiple trees can be combined in *decision forests* for enhanced performance that typically improves with the number of trees. One example of this type of algorithms are *random forests*, in which the input data is randomly re-sampled and multiple trees are optimized independently of each other using subsets of the available data. The final output of this type of forests is typically the average of the binary output of each decision tree.

A common technique to “grow” decision forests used in physics analyses is *boosting*, an iterative procedure in which the performance of the whole classifier is evaluated in each step and the input data is reweighted based on its performance. A new tree optimized with the reweighted input data is then added to the forest. The resulting collection after the boosting process is called *boosted decision trees* (BDT). A usual boosting algorithm is called *adaptive boosting* (AdaBoost), in which the elements of the input dataset that are misclassified in the previous step have an applied weight equal to  $w = (1-f)/f$ , where  $f = N_{\text{misclassified}}/N_{\text{total}}$ . This weight is always larger than 1 and therefore the misclassified events in a given step acquire more importance in the optimization process for the next one, which results in the new trees added to the forest becoming more and more accurate in classifying the mislabelled elements. In this

approach, the final output of the ensemble of trees includes the same weight applied to each classifier,

$$y(\vec{x}) = \sum_{i=1}^{N_{\text{trees}}} \log \left( \frac{1 - f^{(i)}}{f^{(i)}} \right) y^{(i)}(\vec{x}), \quad (7.3)$$

where  $y^{(i)}$  is the output of the  $i$ th tree and  $f^{(i)}$  its fraction of misclassified events.

Another common strategy used to build BDTs is *gradient boosting*. This case is more complex than AdaBoost, since at each step the gradient of the defined loss function is found before optimizing the new tree. For a given model  $F$  with  $M$  boosting stages and loss function  $L$ , at each step  $m < M$  the algorithm computes the *pseudo-residuals*  $r_{im}$  for each element  $i$  in the training dataset as

$$r_{im} = -\frac{\partial L(y_{im}, y_i^{\text{true}})}{\partial y_{im}}, \quad (7.4)$$

where  $y_{im}$  is the prediction for the element  $i$  from the algorithm at stage  $m$ . The new tree added in each step is trained using the input features for each element  $\vec{x}$ , and instead of the original labels  $\vec{y}^{\text{true}}$ , the labels used for the new tree are the pseudo-residuals  $\vec{r}_m = (r_{1m}, r_{2m}, \dots, r_{Nm})$ . The model is then updated with the new weak learner, for which a weight  $\gamma_m$  proportional to the gradient of the loss function is applied.

### 7.2.3 Neural networks

Neural networks (NN) are algorithms designed to imitate a simplified model of neuron cells, and are heavily used in the resolution of pattern recognition problems. These algorithms consist of multiple nodes, or neurons, organized in layers, with the first layer handling the inputs and the last one producing the output. The rest of layers between the first and the last ones are called *hidden layers*. They are generally used in classification problems according to the patterns existing in a given input dataset. The information used as input is a set of  $N$  values,  $\vec{x} = (x_1, x_2, \dots, x_N)$  ranging between 0 and 1. The output of each node is produced by evaluating a user-defined *activation function*,  $\phi(\nu)$ , on the received input information,  $\nu$ , with the addition of a bias term. In the case of the first layer, each node receives as input one of the values in  $\vec{x}$ , i.e.  $\nu_i = x_i$ , and therefore there must be as many nodes in the first layer as input features. For the nodes in the hidden and output layers, the input information is a weighted average of the outputs of the nodes in the previous layer. The output  $y_k^{(n)}(\nu)$  of the node  $k$  in the hidden layer  $n$  can be written as

$$y_k^{(n)}(\nu) = \phi \left( w_0^{(n)} + \sum_{j=1}^{p^{(n-1)}} w_{kj}^{(n)} y_j^{(n-1)} \right), \quad (7.5)$$

where  $w_0^{(n)}$  is the bias term, the sum runs over all the nodes in the layer  $(n-1)$ ,  $p^{(n-1)}$  is the number of nodes in layer  $(n-1)$  and  $w_{kj}^{(n)}$  is the weight matrix relating the output of node  $j$  in layer  $(n-1)$  with node  $k$  in layer  $(n)$ . The output layer can contain either one or multiple nodes, depending on the number of classes that need to be separated. In the case of  $m$  available classes,  $\mathcal{C}_i, i = 0, \dots, m$ , the elements of each class  $i$  are typically labelled with  $y^{\text{true}} = i$ . This kind of network, in which the information flows in one direction (from input to output layer) is commonly known as *feedforward multilayer NN*. An schematic view of such algorithm is shown in figure 7.1b.

The activation function for the nodes is chosen arbitrarily, with constraints in its inputs and

outputs to ensure compatibility with the NN algorithm. A common choice is a sigmoid function,

$$\phi(\nu) = \frac{1}{1 + e^{-\lambda\nu}}, \quad (7.6)$$

with  $\lambda$  determined arbitrarily. Other choices, such as a sinusoidal or an arctangent function are common as well, and the final choice is typically made based on the problem at hand.

The weights  $w_{kj}^{(n)}$  are optimized through the NN training procedure, in which an initial value is provided that is modified iteratively in order to minimize the loss function. The simplest definition for the loss function for a training dataset with  $N$  observations is the root-mean-square error (RMS)

$$L(w) = \sum_{j=1}^N (y_j^{\text{true}} - y(\vec{x}_i))^2, \quad (7.7)$$

where  $y_j^{\text{true}}$  is the true classification of observation  $j$  and  $y(\vec{x}_i)$  the NN prediction for said observation. More complex metrics can be defined in order to improve the performance of the NN, which depends on the problem to solve. Its minimization is done iteratively by modifying the weights  $w_{kj}^{(n)}$  and evaluating the loss function at each step on a reduced dataset to estimate the overall performance. A common method to vary the weights in an efficient way is the so-called *stochastic gradient descent*, in which the weights change as

$$w_{kj}^{(n)} \rightarrow w_{kj}^{(n)} - \eta \frac{\partial L(w)}{\partial w_{kj}^{(n)}}, \quad (7.8)$$

with  $\eta$  known as the *learning rate*. From this definition, it is clear that the variation of the weights always happens in the direction of maximum decrease, although evaluating  $\partial L(w)/\partial w_{kj}^{(n)}$  might not be computationally accurate in some cases, depending on the dataset.

#### 7.2.4 ML-assisted reweighting

Often the problem of determining the multidimensional probability distribution function (*pdf*) of a given physics process is hard to solve. In collider experiments it is common to use MC simulated samples in order to provide estimates of the *pdfs* for the relevant processes contributing in a given kinematic region. This procedure, however, comes with limitations associated to the simulation process. In particular, since the estimated *pdf* does not have an analytical expression, the available statistics in the simulated samples can have an impact in the accuracy of the prediction in extreme kinematic regions. Furthermore, if a given process is hard to model due to inaccuracies in any of the simulation steps discussed in chapter 5, as is the case for SM QCD multijet processes, the predicted *pdf* might show large discrepancies with the observations.

MVA techniques can provide a different way to approach the estimation of the *pdf* for a given process. In the case of a classification problem with only two mutually exclusive classes,  $\mathcal{C}_0$  and  $\mathcal{C}_1$ , the output of the optimized MVA algorithm is approximately equal to the probability of an input set of features,  $\vec{x}$ , to belong to class  $\mathcal{C}_1$ , i.e.  $y(\vec{x}) \approx P(\mathcal{C}_1|\vec{x})$  [137]. This relation between the performance of the MVA algorithm and statistics provides the basis for *pdf* estimation for the class  $\mathcal{C}_1$  using that of  $\mathcal{C}_0$  as reference.

The *pdf* for the class  $i$  encodes the probability of observing a set of values  $\vec{x}$  in an element from  $\mathcal{C}_i$ ,  $P(\vec{x}|\mathcal{C}_i)$ . They can be related to the probability of the observed values  $\vec{x}$  to belong to the class  $\mathcal{C}_i$ , and therefore to the output of the neural network for the same input  $y(\vec{x})$ , through

Bayes theorem<sup>1</sup> as

$$y(\vec{x}) \approx P(\mathcal{C}_1|\vec{x}) = \frac{P(\vec{x}|\mathcal{C}_1)P(\mathcal{C}_1)}{P(x)} = \frac{\alpha_{\mathcal{C}_1}P(\vec{x}|\mathcal{C}_1)}{\alpha_{\mathcal{C}_1}P(\vec{x}|\mathcal{C}_1) + \alpha_{\mathcal{C}_0}P(\vec{x}|\mathcal{C}_0)}, \quad (7.9)$$

where  $\alpha_{\mathcal{C}_i} = P(\mathcal{C}_i)$  is the probability of an element to belong to the class  $\mathcal{C}_i$ , with  $\alpha_{\mathcal{C}_0} + \alpha_{\mathcal{C}_1} = 1$ , which also correspond to the relative size of the class with respect to the total dataset. In the second equality of the previous expression the probability  $P(\vec{x})$  to observe the values  $\vec{x}$  has been split into the conditional probabilities of each class, which can be done since their intersection is empty. The MVA-approximated *pdf* for the class  $\mathcal{C}_1$  can be obtained from the previous expression as

$$P(\vec{x}|\mathcal{C}_1) \approx P(\vec{x}|\mathcal{C}_0)\frac{\alpha_{\mathcal{C}_0}}{\alpha_{\mathcal{C}_1}}\frac{y(\vec{x})}{1-y(\vec{x})}. \quad (7.10)$$

This allows to estimate  $P(\vec{x}|\mathcal{C}_1)$  from a reference *pdf*  $P(\vec{x}|\mathcal{C}_0)$ . Such approximation is very useful in cases where, for example, the class  $\mathcal{C}_1$  follows a complex *pdf* which cannot be readily computed analytically. If a known reference *pdf*  $P(\vec{x}|\mathcal{C}_0)$  provides a good approximation, it can be used to estimate that of  $\mathcal{C}_1$  and, in particular, if  $P(\vec{x}|\mathcal{C}_0)$  has an analytical expression, so does the estimate for  $P(\vec{x}|\mathcal{C}_1)$  since the output of the NN can also be described by the analytical expression in eq. 7.5 once all the weights are known. The validity of this approximation, however, depends on the NN output being a good estimate of  $P(\mathcal{C}_1|\vec{x})$ .

This procedure can help improve the estimated *pdfs* from MC simulated samples in the context of collider experiments, for example comparing the prediction to data in a well controlled kinematic region. Using a NN to separate simulated events ( $\mathcal{C}_0$ ) from recorded data events ( $\mathcal{C}_1$ ), it is possible to have a better estimate for the data *pdf* using the simulated *pdf* as reference. For each simulated event, an event weight is defined as

$$W = \frac{y(e)}{1-y(e)} \approx \frac{\alpha_{\text{data}}}{\alpha_{\text{MC}}} \frac{P(e|\text{Data})}{P(e|\text{MC})}, \quad (7.11)$$

where  $e$  represents the event kinematic information. The event weight defined in this way allows to correct efficiently the normalization of the simulated sample, as can be seen in the ratio of  $\alpha$  factors in the previous equation, at the same time as the shape of the *pdf* and the correlations of the inputs. This procedure can also be applied in MC-to-MC reweighting, as developed in section 7.7 to estimate the contribution of QCD multijet processes in the analysis.

### 7.3 Signal model

The analysis presented in this thesis is a search for electroweakino pair production in scenarios where the lightest neutralino,  $\tilde{\chi}_1^0$ , is mostly dominated by the higgsino component. In this case, the terms higgsino and lightest neutralino are interchangeable. As discussed in chapter 2, these scenarios have typically nearly degenerated  $\tilde{\chi}_1^0$ ,  $\tilde{\chi}_2^0$  and  $\tilde{\chi}_1^\pm$ , and so any decay  $\tilde{\chi}_1^\pm \rightarrow \tilde{\chi}_1^0$  or  $\tilde{\chi}_2^0 \rightarrow \tilde{\chi}_1^0$  is assumed to happen in association with other objects too soft to be reconstructed. The effective cross section for the targeted process is therefore computed inclusively with contributions from  $\tilde{\chi}_1^0\tilde{\chi}_2^0$ ,  $\tilde{\chi}_1^0\tilde{\chi}_1^\pm$ ,  $\tilde{\chi}_2^0\tilde{\chi}_1^\pm$ ,  $\tilde{\chi}_1^\pm\tilde{\chi}_1^\pm$  production channels.

In GMSB scenarios the higgsino-dominated  $\tilde{\chi}_1^0$  is expected to favour decays to a gravitino  $\tilde{G}$  (LSP) plus either a Higgs or a  $Z$  boson, depending on the mass of the higgsino. This signal

---

<sup>1</sup>For two subsets  $A, B$ , the probability of their intersection can be computed in function of their conditional probabilities as  $P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$ .

process is modelled by means of MC simulated event samples, in which matrix elements (ME) for higgsino pairs are generated with up to two additional partons using MADGRAPH v2.6.1 at (LO) with the NNPDF 2.3 PDF sets. PYTHIA v8.230 is subsequently used to model parton showering (PS), hadronization and underlying event (UE). The signal samples are normalised using the best cross-sections calculated at NLO in the strong coupling constant, adding the resummation of soft gluon emission at next-to-leading-logarithmic (NLL) accuracy.

The MC simulated signal samples used in this analysis are generated with a truth-level filter that corresponds to a higgsino decay branching ratio of 50% into  $Z + \tilde{G}$  and 50% into  $h + \tilde{G}$ . This analysis is optimized for final states with decays into  $h + \tilde{G}$  where all Higgs bosons decay to  $b\bar{b}$ , which is ensured with a selection on the Higgs decay mode at truth level. The results are then interpreted using simplified models with a single production mode and 100% branching ratio to a specific decay chain. A diagram of the simplified models is shown in Figure 7.2. In addition to the simplified models with 100% branching ratio scenarios described, interpretations for mixed decay scenarios are provided as exclusion limits in the  $\text{BR}(\tilde{\chi}_1^0 \rightarrow h\tilde{G})$ - $m(\tilde{\chi}_1^0)$  space. These are performed by scaling the events corresponding to each neutralino decay mode by the corresponding branching ratio scanned.

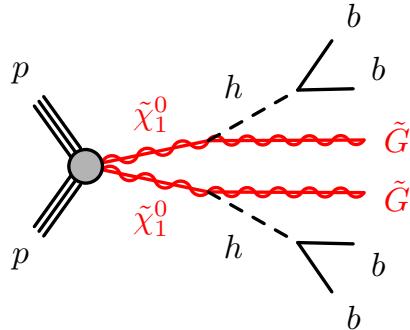


Figure 7.2: The decay topologies of the  $\tilde{\chi}_1^0$  in the simplified model considered in this search. In the model, the Higgs decay fractions are the Standard Model ones; in the figure, only the  $b\bar{b}$  decay mode (target of this search) is represented. The pair production of any of the combinations  $(\tilde{\chi}_1^0, \tilde{\chi}_2^0)$ ,  $(\tilde{\chi}_1^0, \tilde{\chi}_1^\pm)$ ,  $(\tilde{\chi}_2^0, \tilde{\chi}_1^\pm)$ ,  $(\tilde{\chi}_1^\pm, \tilde{\chi}_1^\pm)$  is effectively equivalent to a  $(\tilde{\chi}_1^0, \tilde{\chi}_1^0)$  pair production (plus soft particles) under the assumptions described in the text.

Figure 7.3 shows the distribution of  $E_T^{\text{miss}}$  for some signal masses in the  $hh \rightarrow 4b$  topology. The expected  $E_T^{\text{miss}}$  distribution for intermediate and high masses is generally high, as expected from boosted  $\tilde{G}$  from the higgsino decay. This search is optimized for large acceptance in the high- $E_T^{\text{miss}}$  regime, with best sensitivity in the region  $m(\tilde{\chi}_1^0) \geq 400$  GeV. The reach of this *high-mass channel* presented in this chapter is as low as  $m(\tilde{\chi}_1^0) = 200$  GeV, with a complementary *low-mass channel* optimized for the low- $E_T^{\text{miss}}$  regime. This second channel, targeting low- $E_T^{\text{miss}}$  final states, is a separated ATLAS analysis and will not be the focus of this thesis. While targeting the same signal model, the low- and high-mass analyses differ substantially in strategy and background composition.

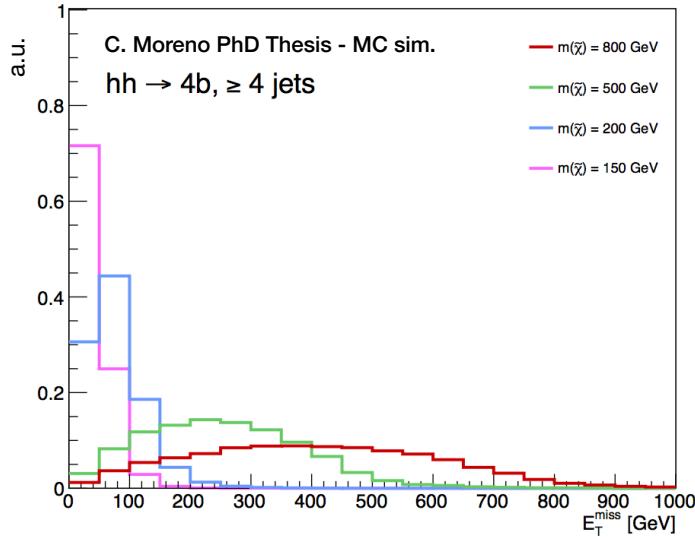


Figure 7.3: Distribution of  $E_T^{\text{miss}}$  in the the  $hh \rightarrow 4b$  topology. To ease the comparison of the shape between the different signals, the distributions are normalised to unit area.

## 7.4 Event selection

The general strategy in the high-mass channel is to target boosted topologies that are typical of the decay of resonances with large mass. In particular, the gravitinos are expected to carry large momentum, resulting in large  $E_T^{\text{miss}}$  that grows with the neutralino mass, as shown in figure 7.3. The Higgs bosons produced in the decay are also expected to become more energetic as the higgsino mass increases, resulting in boosted  $b$ -jets which appear collimated at very high masses.

Events are selected by means of a high- $E_T^{\text{miss}}$  requirement at the trigger level to capture the mentioned topologies. The  $E_T^{\text{miss}}$  cut applied at the L1 and the HLT changed each year of Run 2, which are summarized in table 7.1. While the L1  $E_T^{\text{miss}}$  is computed using exclusively calorimeter information during all Run 2, two different algorithms were used to compute  $E_T^{\text{miss}}$  in the HLT [138]. During 2015 and 2016, the `mht` algorithm computed  $E_T^{\text{miss}}$  based exclusively on calorimeter jets after a calibration procedure similar to the one applied during offline reconstruction. This approach is motivated by topologies where the momentum in the event is dominated by the jets, which already include pileup corrections applied in the calibration. During 2017 and 2018 the `pufit+cell` algorithms are combined to compute  $E_T^{\text{miss}}$  at the HLT. On the one hand, the total  $E_T^{\text{miss}}$  is estimated with the `cell` algorithm, which determines  $E_x$  and  $E_y$  over 188 thousand calorimeter channels without including corrections for the hadronic contributions nor pileup effects, with loose criteria applied to reduce electronics noise. The `pufit` algorithm, on the other hand, computes pileup-corrected  $E_T^{\text{miss}}$  by applying a calibration in high energy calorimeter signals based on a fit applied to low energy signals, following the assumption that the latter arise mainly in pileup collisions. The efficiency of the  $E_T^{\text{miss}}$  trigger selection as a function of the offline  $E_T^{\text{miss}}$  for both data and simulated  $t\bar{t}$  events is shown in figure 7.4.

An additional selection is applied on the offline-computed  $E_T^{\text{miss}}$  (as described in chapter 6, taking into account all calibrated objects in the event), which is tighter than that applied at the trigger level, of  $E_T^{\text{miss}} > 150 \text{ GeV}$ . As shown in figure 7.4, the trigger efficiency below this point is low, generally at the 50% (10%) at 150 (100) GeV. This tighter selection does not result in significant signal loss, especially for high higgsino mass hypothesis, but leads to large

Year	Trigger name	HLT algorithm	L1 threshold [GeV]	HLT threshold [GeV]	$\int \mathcal{L} dt$ [fb $^{-1}$ ]
2015	HLT_xe70_mht_L1XE50	mht	50	70	3.5
2016	HLT_xe90_mht_L1XE50	mht	50	90	12.7
2016	HLT_xe110_mht_L1XE50	mht	50	110	30.0
2017	HLT_xe90_pufit_L1XE50	pufit, cell	50	90,50	21.8
2017	HLT_xe100_pufit_L1XE50	pufit, cell	50	100,50	33.0
2017	HLT_xe110_pufit_L1XE50(55)	pufit, cell	50(55)	110,50	47.7
2018	HLT_xe110_pufit_xe65_L1XE50	pufit, cell	50	110,65	57.0
2018	HLT_xe110_pufit_xe70_L1XE50	pufit, cell	50	110,70	62.6

Table 7.1: Evolution in the trigger menu of the primary  $E_T^{\text{miss}}$  triggers throughout the LHC Run 2. The algorithm used for each year is shown together with the energy thresholds applied at the L1 and the HLT, as well as the integrated luminosity collected by each selection. For 2017 and 2018, two HLT thresholds are provided, which correspond to the missing energy computed by the pufit and cell algorithms, respectively. The integrated luminosity showed are not exclusive, meaning that the same events can be recorded by two trigger selections in a given data taking period, so the total luminosity does not correspond to the sum of the ones showed in this table [138].

background rejection rates of processes with low- $E_T^{\text{miss}}$  such as QCD multijet. In the region selected, however, there are expected discrepancies between the turn-on efficiency curve in data and MC simulation, arising from differences in the detector simulation step and reconstruction procedure. These are corrected with dedicated scale factors applied to MC as event weights on a yearly basis, derived in different energy regimes based on the jet activity in the event. The calibration procedure is detailed in appendix A. The largest differences observed between data and simulation are of the order of 10% in the lowest jet energy regions.

In addition to the  $E_T^{\text{miss}}$  selections, extra requirements referred to as *pre-selection cuts* are included based on kinematic properties of the events in order to largely reduce contributions from SM processes with minimal impact on the signal acceptance. First, a selection on the number of jets in the event is applied. This analysis relies on PFlow jets, reconstructed with the Medium JVT working point, with threshold of  $p_T > 25$  GeV and over the detector region of  $|\eta| < 2.8$ . Events with less than 4 or more than 7 jets are rejected. Of all the jets in the event, at least three of them are required to be tagged as  $b$ -jets with the DL1 algorithm at 77% OP. In addition, all events with any reconstructed lepton are vetoed, since the targeted signal does not include any leptonic decay. For this last requirement, electrons (muons) are only considered if they are reconstructed with the Loose (Medium) ID working point and over a threshold of  $p_T > 20$  GeV. This set of selections will be referred to in the following as Presel-0L.

A second set of selections, also relying on the  $E_T^{\text{miss}}$  requirements at the trigger level mentioned above, is included in order to study the modelling of SM production of  $Z$  bosons in association with jets. These events typically dominate in the kinematic region favoured by high  $\tilde{\chi}_1^0$  mass hypotheses, due to the large  $E_T^{\text{miss}}$  in the final states coming from the  $Z \rightarrow \nu\nu$  decays. In this case, the targeted events are those where the  $Z$  boson decays to a pair of muons which are used to study the production cross section of the  $Z + \text{jets}$  process. Thus, two opposite-sign muons reconstructed with the Medium ID working point and with  $p_T$  threshold of 20 GeV are required. In order to reduce the contribution from other SM processes, an additional requirement on the invariant mass of the di-muon pair is applied to target the  $Z$  mass peak,  $|M_Z - M_{\mu\mu}| < 20$  GeV. In this case, two versions of  $E_T^{\text{miss}}$  are used. On the one hand, the usual  $E_T^{\text{miss}}$  computed as described in chapter 6 is used to apply a tight selection of  $E_T^{\text{miss}} < 75$  GeV to reduce the con-

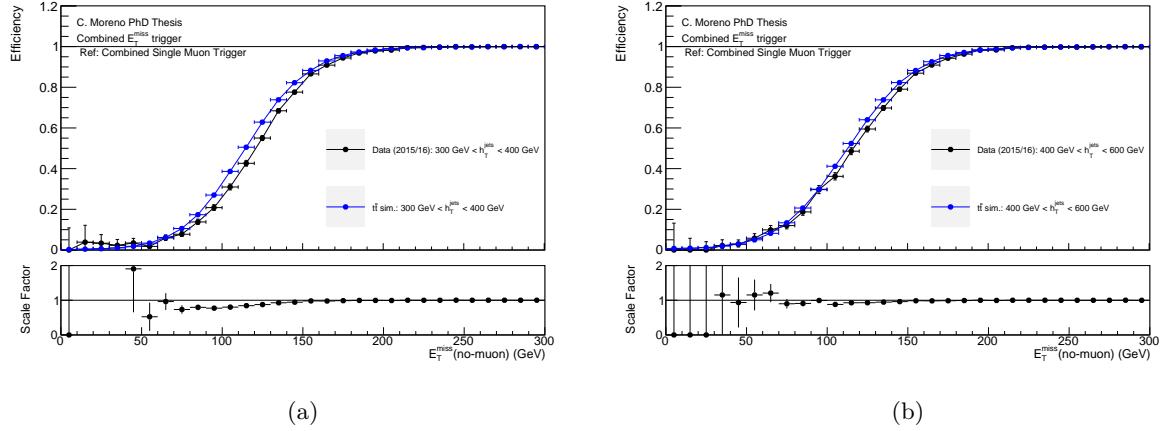


Figure 7.4: Comparison of the efficiency of the  $E_T^{\text{miss}}$  selection at the trigger level in the combined period of 2015 + 2016 for data and simulated  $t\bar{t}$  events in two different energy regimes, (a)  $300 \text{ GeV} < H_T < 400 \text{ GeV}$ , and (b)  $400 \text{ GeV} < H_T < 600 \text{ GeV}$ , measured as a function of the offline- $E_T^{\text{miss}}$ . Single-muon events are used as reference, in a selection with at least four jets and at least 2  $b$ -jets. The calibration procedure followed is detailed in appendix A.

tribution from  $t\bar{t}$  events. A second version is computed, where the momentum of the muons is vectorially added to  $\vec{E}_T^{\text{miss}}$ , which allows for the applied selection of  $E_T^{\text{miss}-\mu} > 175 \text{ GeV}$ .<sup>2</sup> This set of selections is labelled as Presel-ZCR.

## 7.5 Relevant kinematic variables

The following kinematic variables are computed for each event passing the selections mentioned above, and are used in order to study the modelling of SM processes contributing as background and, ultimately, as source of information to discriminate them from the targeted signal.

$N_{\text{jet}}$  : is the number of jets with  $|\eta| < 2.8$  and  $p_T > 25 \text{ GeV}$ .

$N_{b\text{-jet}}$  : is the number of  $b$ -jets with  $|\eta| < 2.5$  with  $p_T > 25 \text{ GeV}$ . By default this is computed with DL1r at the 77% O.P.

$m_{\text{eff}}$  : the effective mass is defined as the sum of the  $p_T$  of all jets with  $p_T > 25 \text{ GeV}$  and  $\eta < 2.8$ , and  $E_T^{\text{miss}}$ :

$$m_{\text{eff}} = \sum_{i \leq n} p_T^{j_i} + E_T^{\text{miss}}. \quad (7.12)$$

this variable reflects the overall energy scale of the event, and in the signal is therefore correlated to the mass of the SUSY particles produced.

$M_J^{\sum,4}$  : the total jet mass variable, defined as:

<sup>2</sup> $E_T^{\text{miss}-\mu}$  is closer to the values computed at the HLT since it relies solely on calorimeter information, hence in this set of selections the requirement on offline  $E_T^{\text{miss}}$  at 75 GeV does not pose a problem in terms of trigger acceptance.

$$M_J^{\sum,4} = \sum_{i \leq 4} m_{J,i}^{R=0.8} \quad (7.13)$$

where  $m_J^{R=0.8}$  refers to the mass of large-radius re-clustered jets in the event; the sum is taken on up to the four leading re-clustered jets in the event.<sup>3</sup> The presence of boosted and semi-boosted top quarks in the  $t\bar{t}$  and single-top-quark events leads to the formation of high- $p_T$ , massive jets at the scale of  $R \approx 0.8$ . The  $M_J^{\sum,4}$  variable is sensitive to this large-angle clustering of hard constituents which distinguishes the signal from background events.

$\Delta\phi_{\min}^{4j}$  : is defined as the minimum  $\Delta\phi$  between any of the four leading jets and  $E_T^{\text{miss}}$ :

$$\Delta\phi_{\min}^{4j} = \min(|\phi_1 - \phi_{E_T^{\text{miss}}}|, \dots, |\phi_4 - \phi_{E_T^{\text{miss}}}|) \quad (7.14)$$

where the index refers to the  $p_T$  ordered list of the four leading jets. This variable is extremely useful in the suppression of the multijet background: all-hadronic backgrounds without neutrinos in the decay chain can have a sizeable amount of  $E_T^{\text{miss}}$  only if the energy deposited by one of the jets in the event is mismeasured. In these cases, it is likely that the measured (fake)  $E_T^{\text{miss}}$  will be aligned to one of the most energetic jets in the event. Requiring high values of  $\Delta\phi_{\min}^{4j}$  helps in suppressing events with no real  $E_T^{\text{miss}}$ . A selection on  $\Delta\phi_{\min}^{4j}$  is therefore present in all the analysis regions. Additionally, it provides a good discriminant at high neutralino mass.

$H_T$  : the sum of the  $p_T$  of all available jets in the event with  $p_T > 25$  GeV and  $\eta < 2.8$

$$H_T = \sum_{i < 8} p_T^{j_i} \quad (7.15)$$

$m_{T,\min}^{b\text{-jets}}$  : is defined as the minimum transverse mass between the  $E_T^{\text{miss}}$  and the three leading  $b$ -jets:

$$m_{T,\min}^{b\text{-jets}} = \min_{i \leq 3} \sqrt{(E_T^{\text{miss}} + p_T^{j_i})^2 - (p_x^{\text{miss}} + p_x^{j_i})^2 - (p_y^{\text{miss}} + p_y^{j_i})^2} \quad (7.16)$$

The variable has a kinematic endpoint near the top mass for semileptonic  $t\bar{t}$  events, if the  $E_T^{\text{miss}}$  from the neutrino is well measured and if the  $b$ -jet from the top that decays semileptonically is among the three leading  $b$ -tagged jets. Since the  $\tilde{G}$  which produces the  $E_T^{\text{miss}}$  in SUSY events is largely independent of the  $b$ -jet kinematics, the value of  $m_{T,\min}^{b\text{-jets}}$  can be much larger in signal processes.

### 7.5.1 Higgs Boson Reconstruction

This analysis targets events where both higgsinos decay into Higgs bosons, with the Higgs further decaying into pairs of  $b\bar{b}$  quarks. Different strategies to reconstruct both Higgs bosons in the event have been tested with the goal of achieving maximum signal acceptance in a certain mass window centred around the Higgs mass. The reconstruction algorithms always take four jets and pairs them into two Higgs candidates. Two different options have been explored to select the four jets in the events:

---

<sup>3</sup>In the case of events with  $< 4$  re-clustered jets, all of them are used.

- The  $p_T$  method aims at targeting boosted Higgs bosons (likely to be produced from higgsinos with high masses) by selecting jets with high transverse momentum. The criteria to select the jets that will be considered as Higgs decay products is:
  - If there are exactly four  $b$ -tagged jets in the event, pick those.
  - If there are more than four  $b$ -tagged jets in the event, pick the ones with highest  $p_T$ .
  - If there are less than four  $b$ -tagged jets in the event, pick all of the  $b$ -tagged and the non  $b$ -tagged jets with highest  $p_T$ .
- The “new” method has been implemented to improve the  $p_T$  method at very high higgsino mass by considering very boosted Higgs bosons whose decay products are merged into a single jet. The criteria to select the Higgs decay products is:
  - If there are exactly four  $b$ -tagged jets in the event, pick those.
  - If there are more than four  $b$ -tagged jets in the event, pick the ones with highest  $p_T$ .
  - If there are only three  $b$ -tagged jets in the event, first look at the mass of all the  $b$ -tagged jets. If the mass of any of the  $b$ -jets is higher than 100 GeV, consider it as a Higgs candidate where the  $b\bar{b}$  quark pair is merged into one single jet. In any other case, select as fourth jet the non-tagged jet that minimizes either the distance  $\Delta R$  between Higgs decay products or the mass difference of Higgs candidates, depending on the criteria used to pair the four selected jets. The fraction of events where the decay products of a Higgs boson are collimated into a single jet increases with the neutralino mass, from about 2% for  $m(\tilde{H}) = 500$  GeV up to almost 40% for  $m(\tilde{H}) = 1.5$  TeV. At  $m(\tilde{H}) = 1$  TeV, the fraction of events with a heavy  $b$ -jet that is considered as a Higgs candidate in this approach is 14.9%.

Once the jets have been selected as Higgs decay products, they need to be paired into two Higgs candidates. Two methods have been explored to form the candidates in the most efficient way:

- The “min-diff” method aims at minimizing the mass difference between the two Higgs candidates,  $m(H_1)$  and  $m(H_2)$ , where  $H_1$  is always picked to be the heaviest of the two candidates.
- The “min-dR” method minimizes  $\Delta R_{\max}^{bb} = \max(\Delta R(H_1), \Delta R(H_2))$ , where  $\Delta R(H_i)$  is the separation of the jets coming from the same Higgs candidate.

The final algorithms are a combination of one of the jet selection methods and one of the pairing methods. The performance of the different reconstruction algorithms is shown in figure 7.5 for three different higgsino mass points,  $m(\tilde{H}) = 250, 500, 1000$  GeV, where the last bin in each distribution is the overflow bin, after a preselection requiring  $E_T^{\text{miss}} > 100$  GeV,  $\geq 4$  jets,  $\geq 3$   $b$ -jets, 0 signal leptons, and a further selection at truth level targeting events in which both Higgs bosons decay to a pair of  $b\bar{b}$  quarks. The background showed in these plots is estimated as described in section 7.6. The jet selection method “new” is more efficient at recovering the events in the tails of the mass distributions compared to the “ $p_T$ ” method, specially at high masses where the decays of at least one Higgs candidate are likely to be merge into a single jet. While the jet pairing method “min-diff” is more efficient at low higgsino mass, “min-dR” is also more efficient at high higgsino mass as it aims at reconstructing boosted Higgs candidates

whose decay products are expected to be collimated. In this analysis, the combination “new” + “min-dR” will be used to reconstruct the Higgs candidates. As described in section 7.9.2, both Higgs candidate masses will be used as training variables for the signal-background discriminating BDT, as well as the  $\Delta R^{bb}$  separation between the decay products of both Higgs candidates.

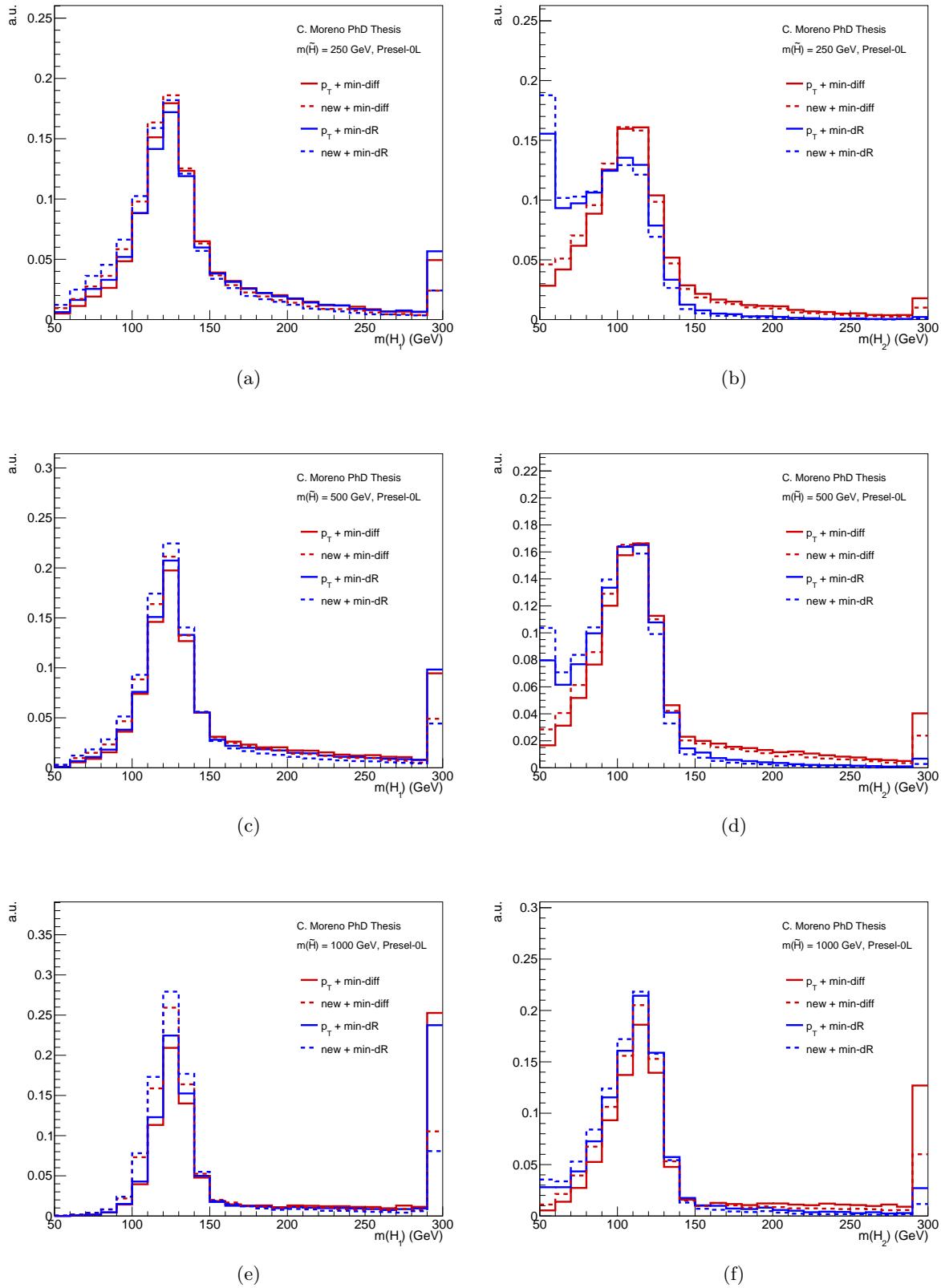


Figure 7.5: Reconstructed mass of the leading (a, c, e) and subleading (b, d, f) Higgs candidates for higgsino masses of 250, 500 and 1000 GeV. The four possible reconstruction algorithms are compared in each plot, with the overflow bin of each distribution in the last bin. The distributions are normalized to unit area.

## 7.6 Main background processes

The event selection specified in section 7.4 result in large acceptance of SM processes that produce high number of  $b$ -jets together with  $E_T^{\text{miss}}$ , such as processes involving leptonic decays of top quarks or  $Z$  boson decays to neutrinos. In the following subsections the main processes contributing to the background in the analysis signal regions are discussed, which are all estimated from MC simulation. The QCD multijet background is an exception, since it is estimated from a data-driven method as detailed in section 7.7.

Process	Generator + fragmentation/hadronization	Tune	PDF set	Cross-section order
$t\bar{t}$	POWHEG-BOX v2 + PYTHIA v8.230	A14	NNPDF3.0 (ME) NNPDF2.3 (UE)	NNLO+NNLL [139]
single-top-quark	POWHEG-BOX v2 + PYTHIA v8.230	A14	NNPDF3.0 (ME) NNPDF2.3 (UE)	NLO [140] ( $t/s$ -channel) NLO+NNLL [141] ( $Wt$ )
$t\bar{t}W/t\bar{t}Z$	MADGRAPH5_aMC@NLO v2.3.3 + PYTHIA-8.210	A14	NNPDF3.0 (ME) NNPDF2.3 (UE)	NLO [97]
4-tops	MADGRAPH5_aMC@NLO v2.2.2 + PYTHIA v8.186	A14	NNPDF3.0 (ME) NNPDF2.3 (UE)	NLO [97]
$t\bar{t}h$	POWHEG-BOX v2 + PYTHIA v8.230	A14	NNPDF3.0 (ME) NNPDF2.3 (UE)	NLO [142]
Dibosons $WW, WZ, ZZ$	SHERPA v2.2.1	Default	NNPDF3.0	NLO [143, 144]
$W/Z+\text{jets}$	SHERPA v2.2.1	Default	NNPDF3.0	NNLO [145]

Table 7.2: List of generators used for the different processes. Information is given about the underlying-event tunes, the PDF sets and the order of the perturbative QCD calculation used for the normalization of the different samples.

### 7.6.1 Pair production of top quarks

Top quark pair production,  $t\bar{t}$ , is one of the largest irreducible backgrounds in this analysis. The main decay mode for top quarks in the SM is to a  $W$  boson plus a  $b$  quark, as ensured by the CKM matrix element  $|V_{tb}| \approx 1$ , which already provides two  $b$ -jets in the final state of  $t\bar{t}$  events. Hadronic decays of one of the produced  $W$  bosons, typically to a  $c$ -quark and a  $s$ -quark, can result in a mistag of the  $c$ -quark as a  $b$ -jet, increasing the probability of these events to survive the selections of Presel-0L. The partons radiated by the particles interacting in the hard scattering can also contribute to the  $b$ -jet multiplicity of the event.

The  $E_T^{\text{miss}}$  requirement is usually not fulfilled by  $t\bar{t}$  events where both  $W$  bosons decay hadronically (hadronic  $t\bar{t}$ ), since the only source of missing energy in that case would come from jet energy mismeasurements. A larger component of events surviving the  $E_T^{\text{miss}}$  selection are due to leptonic decays of the  $W$  bosons, in which cases the process is commonly referred to as semi-leptonic  $t\bar{t}$ , if it is only one  $W$  boson, or di-leptonic  $t\bar{t}$  in case both  $W$  bosons decay to leptons plus their associated neutrino, the latter constituting the main source of  $E_T^{\text{miss}}$ . The electron/muon veto applied means that the only  $t\bar{t}$  events accepted in the analysis are those where the  $W$  boson(s) decay to either a tau lepton, which further decays hadronically and is

therefore reconstructed as a jet, or to electron/muon either with momentum below the required thresholds to be reconstructed or that are wrongly identified as a different particle.

The extra partons radiated in the hard scattering, introduced in cross section computations at NLO and higher, play an important role in the modelling studies for the  $t\bar{t}$  process. Simulated collisions resulting in  $t\bar{t}$  events can be classified into three production modes according to their truth information. In order to identify the hadrons inside a reconstructed jet in the event, *particle jets* are reconstructed using exclusively truth particle information with the same jet reconstruction algorithms discussed in chapter 6, which are matched based on topological requirements to jets reconstructed with detector information [146]. Since the relevant information for this classification is related exclusively to the initial and final state radiation of the event, all jets originating in the top decay chain are disregarded. The remaining jets are labelled according to their particle jet content:

- If there is a radiated jet containing a  $B$ -meson at truth level, the jet is considered a  $b$ -jet and the event is classified as  $t\bar{t}+ \geq 1b$ .
- If there is a radiated jet containing a  $D$ -meson at truth level, such jet is considered to be originated from the hadronization of a  $c$ -quark and therefore the event is classified as  $t\bar{t}+ \geq 1c$ .
- The rest of events are classified as  $t\bar{t}+\text{light-jets}$ .

Events in which  $t\bar{t}$  is produced in association with at least one  $b$ -jet or at least one  $c$ -jet are commonly referred to as  $t\bar{t}+\text{heavy flavour (HF)}$ , and their correct modelling is important in analyses requiring multiple number of  $b$ -jets, since they become dominant over the  $t\bar{t}+\text{light jets}$  component at high  $b$ -jet multiplicity.<sup>4</sup> Figure 7.9b shows the change in composition of the  $t\bar{t}$  background as a function of the number of  $b$ -jets for events selected with Presel-0L, where the agreement to data at high  $b$ -jet multiplicity reveals that the normalization of the  $t\bar{t}+\text{HF}$  component is underestimated by the MC generator used, consistently with what was observed in dedicated  $t\bar{t}+\text{HF}$  cross section measurements [147]. This motivates a complex strategy in terms of dedicated control regions used in the fit to data to correct separately the normalization of each component, as developed in section 7.9.

The estimate for this background used in this search is generated with PowHEG-Box v2 for the matrix element computation, interfaced with PYTHIA v8.230 for the parton shower and hadronization steps, using the NNPDF3.0 set in the matrix element step and the NNPDF2.3 set for the underlying event simulation. The samples are normalized to the cross section computed at NNLO+NNLL. The simulated samples are used to estimate the shape of relevant distributions, while the normalization is corrected by means of a fit to data in dedicated control regions. In order to increase the available MC statistics in kinematic regions with low cross sections, additional samples are generated with filters at truth level on relevant kinematic variables. Two schemes are used in this analysis:

- Filtered samples in truth  $E_T^{\text{miss}}$ , in bins with  $200 \text{ GeV} \leq E_T^{\text{miss}} < 300 \text{ GeV}$ ,  $300 \text{ GeV} \leq E_T^{\text{miss}} < 400 \text{ GeV}$  and  $E_T^{\text{miss}} \geq 400 \text{ GeV}$ .
- Filtered samples in truth  $H_T$ , in bins with  $600 \text{ GeV} \leq H_T < 1 \text{ TeV}$ ,  $1 \text{ TeV} \leq H_T < 1.5 \text{ TeV}$  and  $H_T \geq 1.5 \text{ TeV}$ .

---

<sup>4</sup> $t\bar{t} + \geq 1c$  can also contribute to regions with large number of  $b$ -jets due to mistag of the  $c$ -jet.

These samples are combined with an unfiltered sample using requirements at truth  $E_T^{\text{miss}}$  and  $H_T$  to avoid double counting of simulated events. The filtered samples are then scaled by the corresponding cross section of the given slice.

### 7.6.2 Single-top-quark production

In addition to pair production, top quarks can be produced singly in LHC collisions through the electroweak interaction. The cross section for such process is lower than that of  $t\bar{t}$ ,  $247 \pm 46$  pb ( $t$ -channel) and  $94 \pm 23$  pb ( $Wt$ -channel) compared to  $826.4 \pm 19.6$  pb at  $\sqrt{s} = 13$  TeV, due to the lower interaction strength characteristic to each process. Selections with large number of  $b$ -jets and high  $E_T^{\text{miss}}$  capture a considerable amount of single-top-quark events due to  $W$  decays to hadrons or to tau leptons, and in certain kinematic regions, such as those with high  $m_{T,\text{min}}^{b\text{-jets}}$  values as shown in figure 7.10d, its contribution is comparable to that of  $t\bar{t}$ .

The production of single-top-quark in proton-proton collisions receives contributions from three distinct channels. In the s-channel and t-channel, the top quark is produced together with another quark, which typically do not fulfil the simultaneous selections on  $E_T^{\text{miss}}$  and  $b$ -jet multiplicity used in this analysis. The  $Wt$ -channel, in which the top quark is produced together with an on-shell  $W$  boson, is the dominant component of single-top-quark events in Presel-0L. The modelling of this channel needs careful treatment due to quantum interference with  $t\bar{t}$  process: at NLO the  $W$  boson and top quark in the final state can be produced in association with an extra  $b$ -quark, resulting in a similar topology than LO  $t\bar{t}$ . In both cases, the final state is constituted by two on-shell  $W$  bosons and two  $b$ -quarks, where for  $t\bar{t}$  events they come from a double top-quark resonance, while for single-top-quark events there is exclusively one resonance formed by one of the  $W$  bosons and one of the  $b$ -quarks. The total cross section for the  $WWbb$  final state therefore receives contributions from both topologies, which leads to an interference terms from both matrix elements:

$$|\mathcal{M}_{WWbb}|^2 \propto |\mathcal{M}_{\text{LO}-t\bar{t}}|^2 + |\mathcal{M}_{\text{NLO}-Wt}|^2 + 2 \text{Re} \{ \mathcal{M}_{\text{LO}-t\bar{t}} \times \mathcal{M}_{\text{NLO}-Wt}^* \} .$$

The general approach to handle this interference is to introduce a correction in the single-top-quark simulated samples, while leaving those for  $t\bar{t}$  unaffected. This correction aims at removing the contribution from doubly resonant  $WWbb$  from the NLO prediction for the  $Wt$ -channel, and is implemented according to one of the two following methods:

- In the *diagram removal* (DR) scheme all NLO diagrams contributing to the  $WWbb$  final state are removed from the matrix element computation, and while this is not a gauge-invariant approach, it has been found that there is only weak dependence on the gauge choice.
- In the *diagram subtraction* (DS) scheme an extra subtraction term is introduced in the matrix element computation in order to cancel explicitly the contribution from LO  $t\bar{t}$  to  $WWbb$  final states.

The simulated samples used to estimate the contribution from single-top-quark events in this analysis are generated separately for each channel. They are all generated with POWHEG-BOX for the matrix element computation and interfaced with PYTHIA v8.230 to simulate the showering step, and they are normalized to the NLO+NNLL computation of the cross section. The samples for the  $Wt$ -channel are generated following the DR scheme, and, in order to increase the available MC statistics in this channel, extra samples in different kinematic slices are generated by applying filters at truth level, with a similar scheme to that used for  $t\bar{t}$  simulated samples.

### 7.6.3 Vector boson production in association with jets

Processes in which electroweak bosons are generated in the hard scattering together with jets can have a significant contribution in the targeted kinematic regions. In particular,  $Z$  bosons decaying to neutrinos result in large  $E_T^{\text{miss}}$  in the event, making this process the leading background in the kinematic region preferred by high-mass neutralinos. In order to produce  $E_T^{\text{miss}}$ , events with  $W$  bosons need to decay leptonically, and due to the electron/muon veto applied in this analysis its contribution is suppressed.

These backgrounds are simulated separately for process leading to a  $W$  or a  $Z$  boson. In both cases the samples are generated using SHERPA v2.2.1, which combines COMIX [148] and OPENLOOPS [149] for the matrix element computation and merges the output with the parton shower process using the CKKW prescription [150]. The samples are then normalized to the NNLO computation of the cross section. In the case of  $Z+\text{jets}$ , the simulated samples are used exclusively to compute the shape of kinematic distributions, while the normalization is estimated from a fit to data.

### 7.6.4 Diboson production

Vector boson pair production is also possible in proton-proton collisions at the LHC. The usual kinematic of these events make them likely to pass a selection with high  $b$ -jet multiplicity and large  $E_T^{\text{miss}}$ , since both  $b$ -jets and  $E_T^{\text{miss}}$  can arise independently from the decays of  $Z$  and  $W$  bosons. Events combining  $Z \rightarrow b\bar{b}$  and  $W \rightarrow \tau\nu_\tau$  could be an example, or events with  $Z \rightarrow \nu\nu$  and  $W \rightarrow cs$  where the  $c$ -quark could be mistagged as a  $b$ -jet. The production cross section for these processes, however, is suppressed with respect to single vector boson production, and therefore their contribution in this analysis is expected to be small.

Simulated samples are produced for each possible combination of  $W$  and  $Z$  bosons, with a sample per decay mode. They are generated with SHERPA v2.2.1 using the NNPDF3.0 PDF set, and the cross section is normalized to the NLO prediction.

### 7.6.5 $t\bar{t}+X$ production

Top quark pair production can also happen in association with other particles, such as electroweak bosons, a Higgs boson or another  $t\bar{t}$  pair. All of these processes can contribute to the kinematic region selected in this analysis already through the  $t\bar{t}$  pair, with increased probability when the associated  $W$  bosons decay hadronically or to tau leptons, the  $Z$  boson to either neutrinos or a  $b\bar{b}$  pair and the Higgs boson to a  $b\bar{b}$  pair. Similarly to the case of diboson production, the production cross section for these processes is very small compared to other SM processes contributing to the same kinematic region. Since their effect is expected to be minor in this analysis, all these processes are grouped together under the category  $t\bar{t}+X$ .

All  $t\bar{t}+X$  processes are simulated at NLO precision, with separate samples produced for each of them. Events where  $t\bar{t}$  is produced in association with a Higgs boson are generated with POWHEG-BOX v2 for the matrix element computation interfaced with PYTHIA v8.230 for showering, with the NNPDF3.0 and NNPDF2.3 sets for hard scattering and underlying event, respectively. The processes  $t\bar{t}+W/Z$  and  $t\bar{t}t\bar{t}$  are generated with MADGRAPH5\_aMC@NLO for the matrix element step, with the parton shower and hadronization modelled with PYTHIA as well.

### 7.6.6 Modelling uncertainties

The samples specified in table 7.2 used to model the main background processes are chosen for being the closest to measured data in certain well controlled kinematic regions. The choice of generator, however, generally comes with assumptions over the physics processes that may not reproduce the exact same predictions in all parts of the phase space. For this reason, alternative samples are generated to evaluate the differences to those considered as the nominal prediction, which are later applied as uncertainties in the modelling of the SM processes.

For  $t\bar{t}$  and single-top-quark events, three sources of uncertainty are considered:

- The uncertainty due to the matrix element generator choice is evaluated with an alternative sample generated with a different matrix element MC generator, MADGRAPH5\_AMC@NLO, interfaced with the same parton shower generator, PYTHIA.
- The uncertainty in the parton shower process is estimated with an alternative sample produced with the same matrix element generator as the nominal ones, PowHEG-Box, interfaced with a different parton shower generator, HERWIG.
- The uncertainties associated to emissions of ISR and FSR are evaluated with variations in the renormalization and factorization scales used in the generation of the samples, and are estimated by separately scaling them up and down in the PYTHIA configuration.

The single-top-quark samples include an extra source of uncertainty for the interference of the  $Wt$ -channel at NLO precision with LO  $t\bar{t}$ . While the nominal samples for said channel are generated following the DR scheme, alternative samples following the DS scheme are used to evaluate the differences with the nominal prediction, which are applied as uncertainties to the predicted contribution of  $Wt$ -channel single-top-quark production.

For  $W/Z$  bosons produced in association with jets, an uncertainty associated to ISR/FSR emissions is estimated by varying the factorization and renormalization scales in SHERPA and comparing the results with the nominal samples. For the rest of the backgrounds (top+ $X$ , dibosons) no dedicated uncertainties are derived as their contributions to the kinematic region explored in this analysis are sub-leading.

## 7.7 QCD estimation with NN-assisted reweighting

Collisions in which the final state consist of multiple jets due to QCD processes, such as quark pair production or di-gluon emission, are commonly referred to as *QCD multijet*. They result in topologies with large number of jets, hence the name, and low  $E_T^{\text{miss}}$  since there is no real invisible particle produced. The only source of  $E_T^{\text{miss}}$  in this case is jet energy mismeasurement, either due to detector effects or due to neutrinos produced in the decay chain of any of the hadrons inside the jet.

Events produced by QCD multijet processes make up a small percentage of the SM background present after the preselection cuts, which nonetheless needs to be estimated in an accurate way. This process is particularly challenging to simulate through MC methods, since they typically happen at the limit of the perturbative approximation of QCD, and they include multiple factors proportional to the strong interaction coupling,  $\alpha_s$ , which makes it hard to accurately estimate the cross section. Even though QCD di-jet MC samples were available to be used at

the time of this writing, they do not contain enough statistics for a robust estimate for this analysis after the pre-selection cuts.

Finding an alternative technique to estimate this background was part of the optimization process for this analysis, in which different options were explored. A data-driven strategy was preferred for increased accuracy, combined with multivariate analysis (MVA) techniques to improve its performance. In particular, a reweighting procedure relying on neural networks was implemented, which led to a background prediction with excellent agreement with data both in the analysis pre-selection region and in kinematic regions with enhanced contribution from QCD multijet events.

### 7.7.1 QCD multijet estimate

A data-driven method was chosen for higher accuracy with large available statistics. In this approach, a template for the QCD multijet process is filled with data events with  $\Delta\phi_{\min}^{4j} < 0.2$ , where the non-QCD backgrounds as estimated from MC in this kinematic region are subtracted.<sup>5</sup> In order to have an accurate prediction for this background, this template needs to be corrected for different effects:

- The normalization of the template does not correspond to the actual QCD multijet background at  $\Delta\phi_{\min}^{4j} > 0.4$  due to the exponential shape of the  $\Delta\phi_{\min}^{4j}$  distribution.
- As shown in figure 7.6, the shape of relevant kinematic variables, such as  $E_T^{\text{miss}}$  or leading jet  $p_T$ , is correlated to  $\Delta\phi_{\min}^{4j}$ , and so their shapes change for different  $\Delta\phi_{\min}^{4j}$  slices. Building a template from data at low  $\Delta\phi_{\min}^{4j}$  might not be an accurate representation of the QCD multijet behaviour at high  $\Delta\phi_{\min}^{4j}$ , and these differences need to be corrected as well.
- There is no prediction for the shape of the  $\Delta\phi_{\min}^{4j}$  distribution at  $\Delta\phi_{\min}^{4j} > 0.2$ , which is a powerful discriminant variable against signal.

In order to tackle these problems, a fake  $\hat{\Delta\phi}_{\min}^{4j}$  distribution is created for the template with random numbers that follow the real  $\Delta\phi_{\min}^{4j}$  distribution from dedicated di-jet MC samples. This fake  $\hat{\Delta\phi}_{\min}^{4j}$  distribution, however, does not carry the correct correlations to other kinematic variables since it is generated randomly. In order to correct the shapes and normalization of the template, as well as the correlation between  $\hat{\Delta\phi}_{\min}^{4j}$  and other variables, a kinematic reweighting using multiple variables is performed in order to match the behaviour of QCD multijet events at high  $\Delta\phi_{\min}^{4j}$ . This reweighting is performed by means of a neural network, following the procedure discussed in section 7.2.4. The NN is trained on di-jet MC samples for high purity of QCD multijet events, and the reweighting is later applied to the template derived from data. The event selection used for the training is Presel-0L, with the relaxed cuts  $E_T^{\text{miss}} > 100 \text{ GeV}$  and  $N^{\text{b-jet}} \geq 2$  in order to increase the available statistics. The main goal for the reweighting is to match the behaviour of di-jet MC high  $\hat{\Delta\phi}_{\min}^{4j}$  (which in reality corresponds to low- $\Delta\phi_{\min}^{4j}$  values) into the behaviour of high  $\Delta\phi_{\min}^{4j}$  dijet MC events. The reweighted data-driven template is used in the analysis as the final QCD estimate.

To train the neural network, all di-jet MC events are split into low- and high- $\Delta\phi_{\min}^{4j}$  slices at  $\Delta\phi_{\min}^{4j} = 0.2$ . The neural network is then trained to separate the events from both  $\Delta\phi_{\min}^{4j}$  slices,

---

<sup>5</sup>This is done on a histogram-level by adding the non-QCD background events with negative event weights in the template.

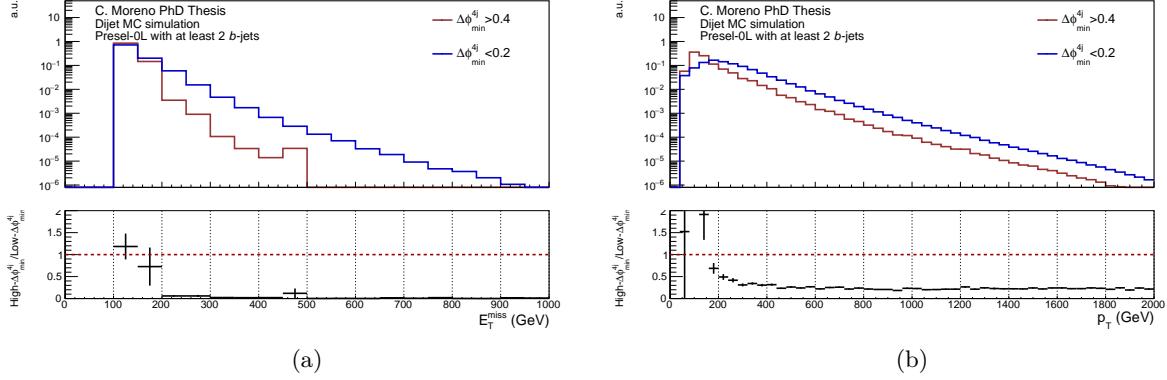


Figure 7.6: Comparison of dijet simulated events with  $\Delta\phi_{\min}^{4j} > 0.4$  and  $\Delta\phi_{\min}^{4j} < 0.2$  for the distributions (a)  $E_T^{\text{miss}}$  and (b) leading jet  $p_T$  in the region defined with Presel-0L with a looser requirement of at least 2  $b$ -jets.

where the  $\hat{\Delta\phi}_{\min}^{4j}$  distribution is used in the training for the low- $\Delta\phi_{\min}^{4j}$  sample. An event weight for the low- $\Delta\phi_{\min}^{4j}$  events is then derived as eq. 7.11. The combination of hyperparameters for the neural network that leads to the best reweighting is chosen from an extensive scan of different values as the one with best quality in the prediction of the *pdf*. The final structure of the neural network is summarized in table 7.3. The variables used to train the neural network, after a scan comparing the performance with several input variable combinations, are  $E_T^{\text{miss}}$ ,  $\Delta\phi_{\min}^{4j}$  and  $H_T$ .

Hyperparameter	value
Epochs	100
Hidden layers	2
Nodes per layer	67
Learning rate $\eta$	0.0002
Batch size	30
Optimizer	Adam
Loss function	Binary Cross entropy

Table 7.3: Selected combination of hyperparameters that optimize the NN training based on the reweighting performance

The performance of the reweighting applied to dijet MC events in a baseline selection ( $E_T^{\text{miss}}$  trigger,  $E_T^{\text{miss}} > 100$  GeV, 0L,  $4 \leq N_{\text{jet}} \leq 7$ ) for different kinematic variables is summarized in figure 7.7, for events with at least 2  $b$ -jets, in the left column, and events with at least 3  $b$ -jets, in the right column.

Once the best model has been chosen for the NN-assisted reweighting, the template filled as discussed above<sup>6</sup> is evaluated with said model and event weights are derived as per eq. 7.11. The reweighted template is the final QCD multijet estimate sample. In order to validate the performance of the complete procedure, a validation region is defined with the Presel-0L set of kinematic cuts, with the replacement of  $\Delta\phi_{\min}^{4j} \geq 0.4$  by  $0.2 \leq \Delta\phi_{\min}^{4j} < 0.4$ , which enhances the contribution of QCD multijet events.

Dedicated uncertainties for this background are included in the fit strategy in order to cover remaining mismodelling that are not corrected with the NN-reweighting procedure. For this, two

<sup>6</sup>Data events with  $\Delta\phi_{\min}^{4j} < 0.2$  after subtracting the non-QCD backgrounds estimated from MC simulation.

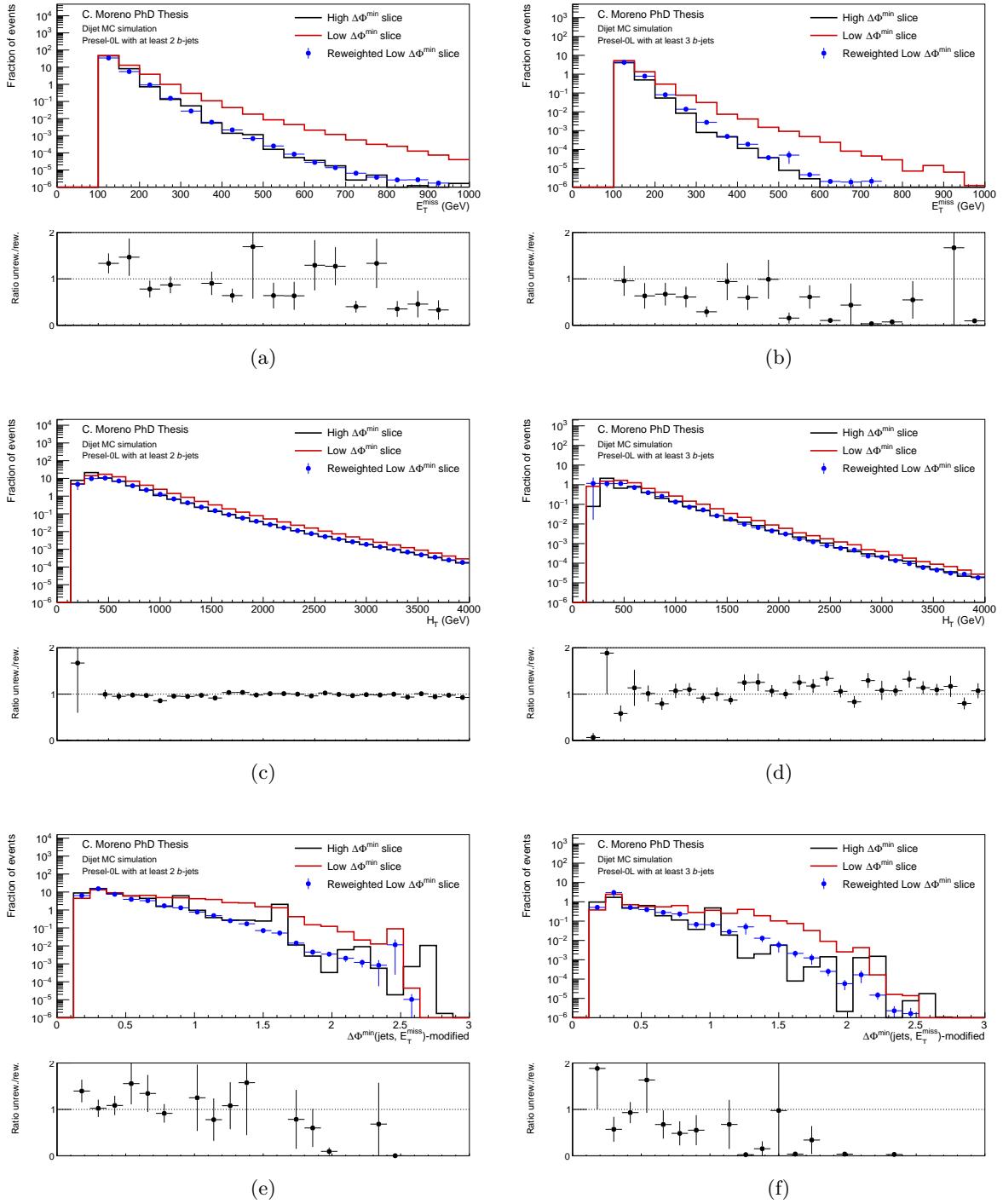


Figure 7.7: Performance of NN-based reweighting in dijet MC, comparing the reweighted and un-reweighted distributions from the low- $\Delta\Phi_{\min}^{4j}$  slice to the target high- $\Delta\Phi_{\min}^{4j}$  slice in a baseline selection with (a,c,e) at least 2  $b$ -jets and (b,d,f) at least 3  $b$ -jets.

sources of contribution to the total uncertainty are studied. On the one hand, the non-closure of the NN-reweighting procedure in the region  $\Delta\phi_{\min}^{4j} > 0.4$  is evaluated in dijet MC events in order to understand potential mismodelling introduced in the final template. The differences between reweighted low- $\Delta\phi_{\min}^{4j}$  and high- $\Delta\phi_{\min}^{4j}$  events is found to be at most 40%. The uncertainty in the

application of the reweighting to the data-driven template is also considered, which is evaluated as the non-closure observed in the validation region defined with  $0.2 \leq \Delta\phi_{\min}^{4j} < 0.4$ . The data agreement with the prediction in this region is displayed in figure 7.8, which shows a level of disagreement typically smaller to 10%.

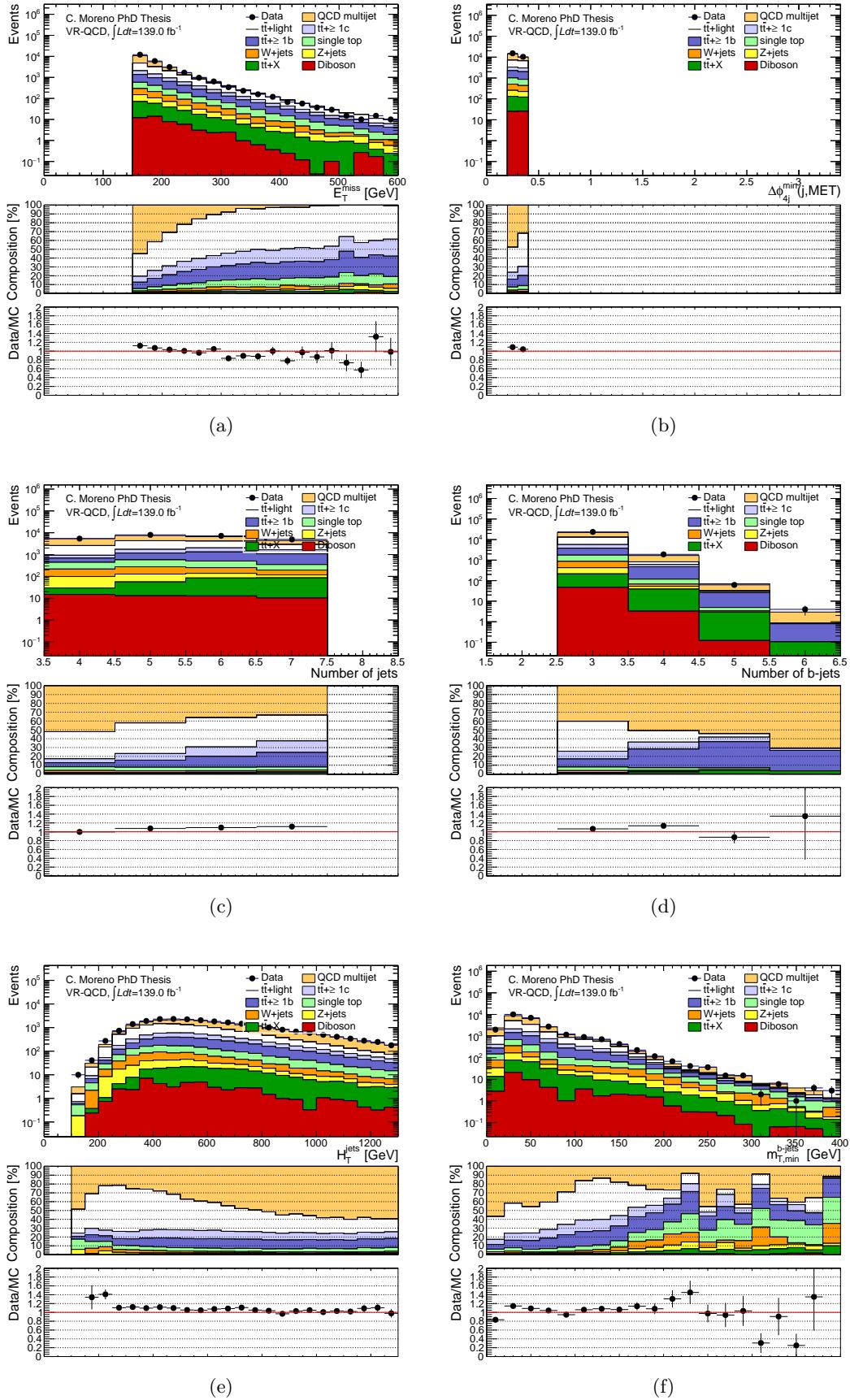


Figure 7.8: Data and background agreement in the region defined by the selections Presel-0L, replacing the requirement on  $\Delta\phi_{4j}^{\min}$  by  $0.2 \leq \Delta\phi_{4j}^{\min} < 0.4$  to enhance the contribution of QCD multijet events, for the kinematic variables (a)  $E_T^{\text{miss}}$ , (b)  $\Delta\phi_{4j}^{\min}$ , (c) number of jets, (d)  $b$ -jet multiplicity, (e)  $H_T$  and (f)  $m_{T,\min}^{b-\text{jets}}$ . Only the statistical uncertainty is included.

## 7.8 Data agreement at pre-selection level

In this section the modelling of the different kinematic observables described in section 7.5 is discussed. All of the figures show the agreement between data and the background estimation prior to the fit. There are two regions of interest: the  $t\bar{t}$  enriched selection in the 0-lepton channel defined with the selection cuts of Presel-0L, and a  $Z+jets$  enriched region constructed with events with exactly two leptons defined with Presel-Z. The data agreement to the background prediction in Presel-0L is shown for some key kinematic distributions in figures 7.9 and 7.10. Figure 7.11 shows the agreement in Presel-Z for relevant kinematic variables, some of which are re-computed with the corrected  $E_T^{\text{miss}}$  accounting for the muons in the event. The  $t\bar{t}$  background process is split into  $t\bar{t} + \text{light}$ ,  $t\bar{t} + \geq 1b$  and  $t\bar{t} + \geq 1c$  to study separately the modelling of  $t\bar{t}$  produced in association with heavy flavour quarks. The reconstructed Higgs candidate mass are shown for both Presel-0L and Presel-Z in figure 7.12.

As expected, different processes show characteristic signatures that can be exploited in the signal-to-background separation stage, which is done by means of a BDT as will be discussed in section 7.9. Top quark pair production events typically populate all the kinematic region explored, with the  $t\bar{t} + \text{HF}$  components contributing more at large  $b$ -jet multiplicity. This background typically shows  $m_{T,\min}^{b\text{-jets}}$  values below the top-quark mass, and its contribution steeply falls as  $m_{T,\min}^{b\text{-jets}}$  increases, where single-top-quark and  $Z+jets$  processes dominate. A similar behaviour can be seen in the  $E_T^{\text{miss}}$  distribution, where the total contribution from  $t\bar{t}$  events falls from about 85% at  $E_T^{\text{miss}} = 200$  GeV to 50% at  $E_T^{\text{miss}} = 600$  GeV, where the contributions from single-top-quark and  $Z+jets$  become more relevant. Events from high higgsino mass hypotheses typically live in the kinematic region with high- $E_T^{\text{miss}}$  and high- $m_{T,\min}^{b\text{-jets}}$  and therefore  $t\bar{t}$  becomes a subleading background in the regions that target high  $\tilde{\chi}_1^0$  masses. The QCD multijet background shows very low  $E_T^{\text{miss}}$ , with relevant contributions showing typically in the region with  $E_T^{\text{miss}} < 200$  GeV. This justifies the robust strategy developed in section 7.7 for its estimation, since the low  $E_T^{\text{miss}}$  requirement applied in the analysis does not totally suppress the contribution from this background and it can populate regions that target lower  $\tilde{\chi}_1^0$  masses. The reconstructed  $E_T^{\text{miss}}$  in these events is purely due to jet energy mismeasurement, which grows with the jet  $p_T$ . This feature is shown in the distribution of  $H_T$ , where QCD multijet events show very large values since highly energetic jets need to be present for a mismeasurement of  $E_T^{\text{miss}} \approx 150$  GeV to happen.

These figures include only statistical uncertainties. No background correction is implemented before the final likelihood fit. Therefore, the agreement in the following plots is expected to be generally good, and any remaining discrepancy will be dealt with by the background normalization scale factors from the final fit as well as with relevant systematic uncertainties. As can be seen in figure 7.10d, a certain mismodelling in the  $m_{T,\min}^{b\text{-jets}}$  distribution is observed after the top quark mass ( $m_{T,\min}^{b\text{-jets}} > 180$  GeV), which motivates the inclusion of high- $m_{T,\min}^{b\text{-jets}}$  control regions to correct these effects in the fit to data.

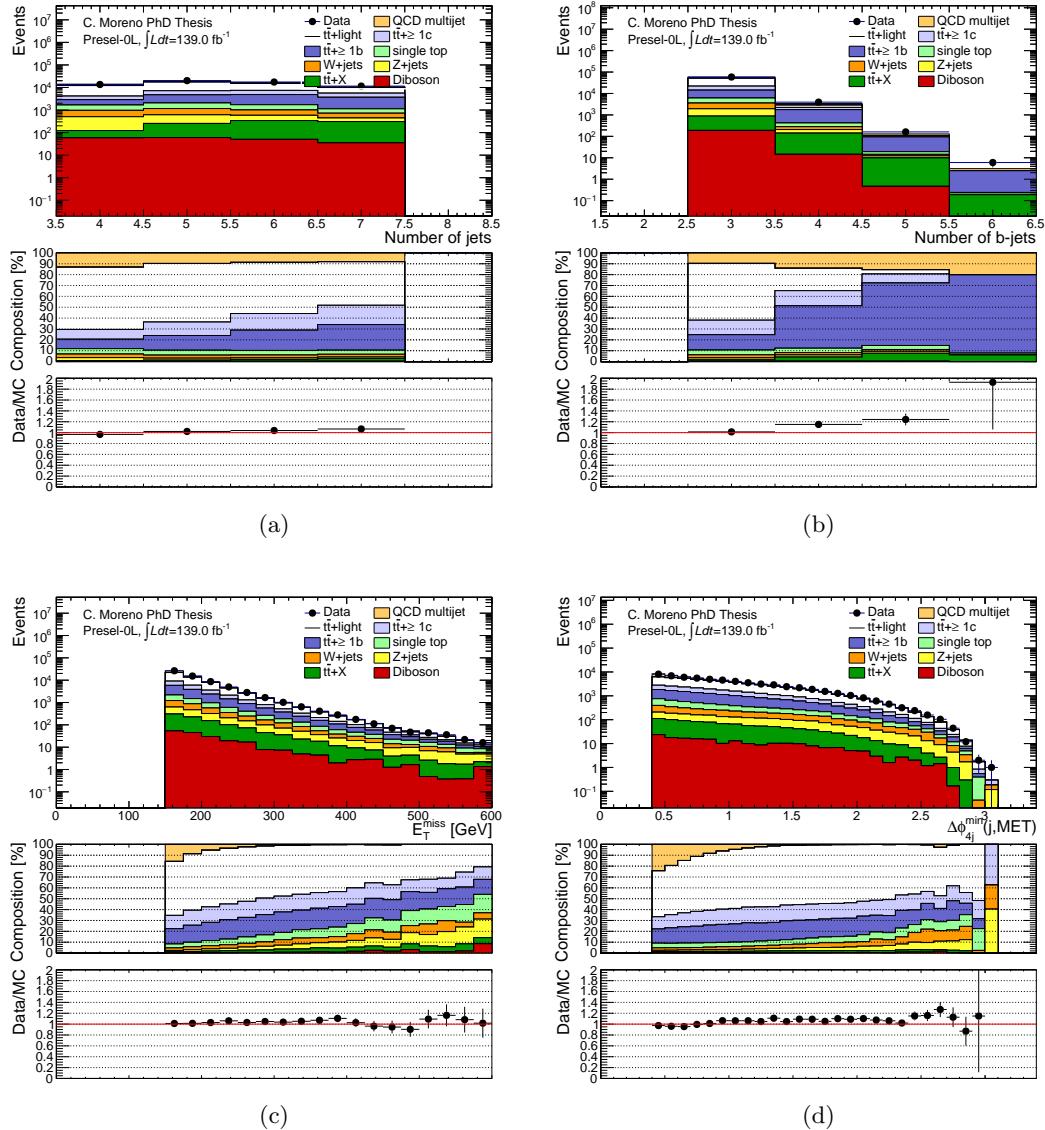


Figure 7.9: Comparison of data and background prediction for the kinematic distributions of (a) number of jets, (b)  $b$ -jet multiplicity (c)  $E_T^{\text{miss}}$  and (d)  $\Delta\phi_{\text{min}}^{4j}$  in region Presel-0L. Only statistical uncertainties are included.

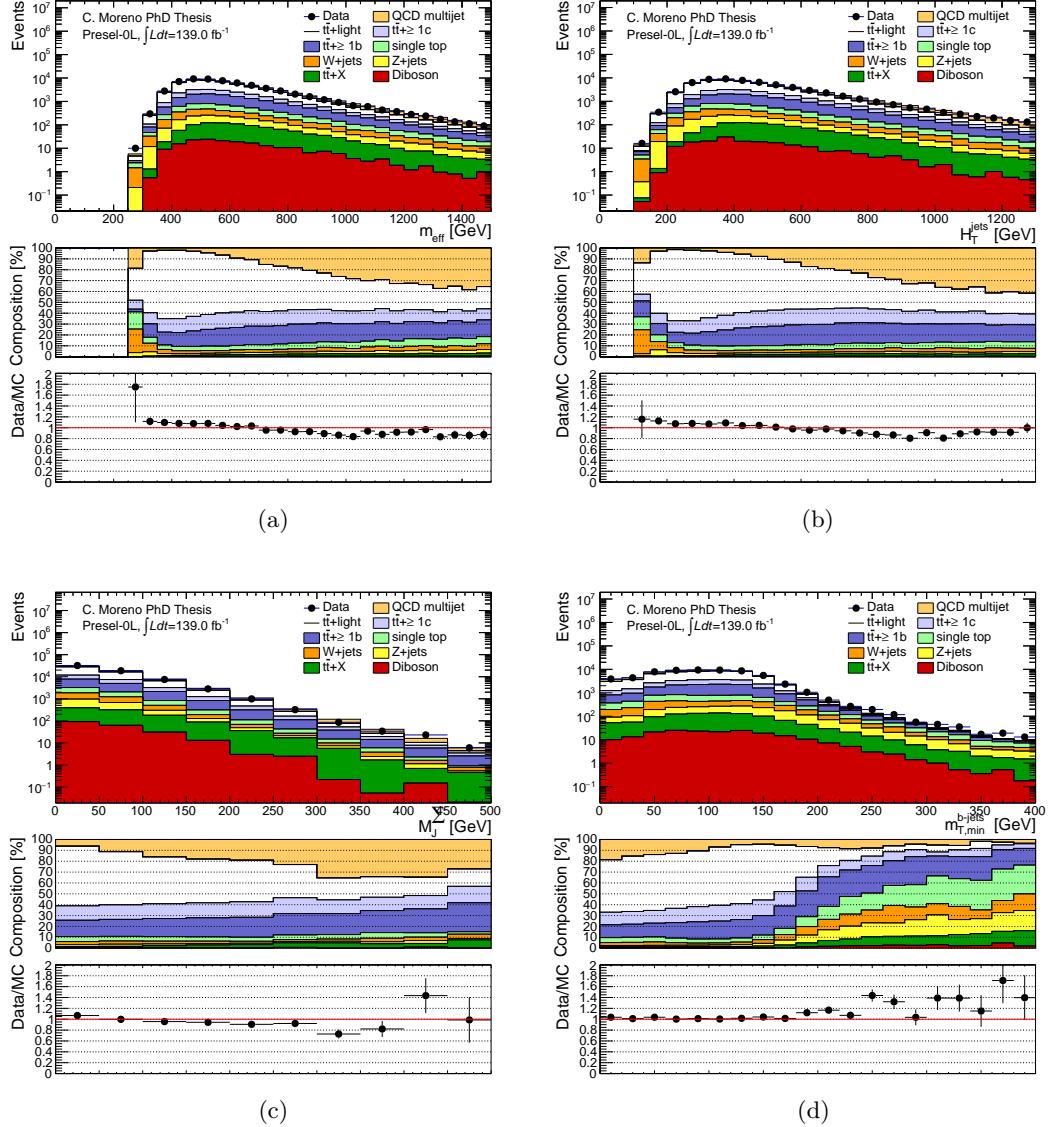


Figure 7.10: Comparison of data and background prediction for the kinematic distributions of (a)  $m_{\text{eff}}$ , (b)  $H_T$ , (c)  $M_J^{\sum,4}$  and (d)  $m_{T,\min}^{b\text{-jets}}$  in region Presel-0L. Only statistical uncertainties are included.

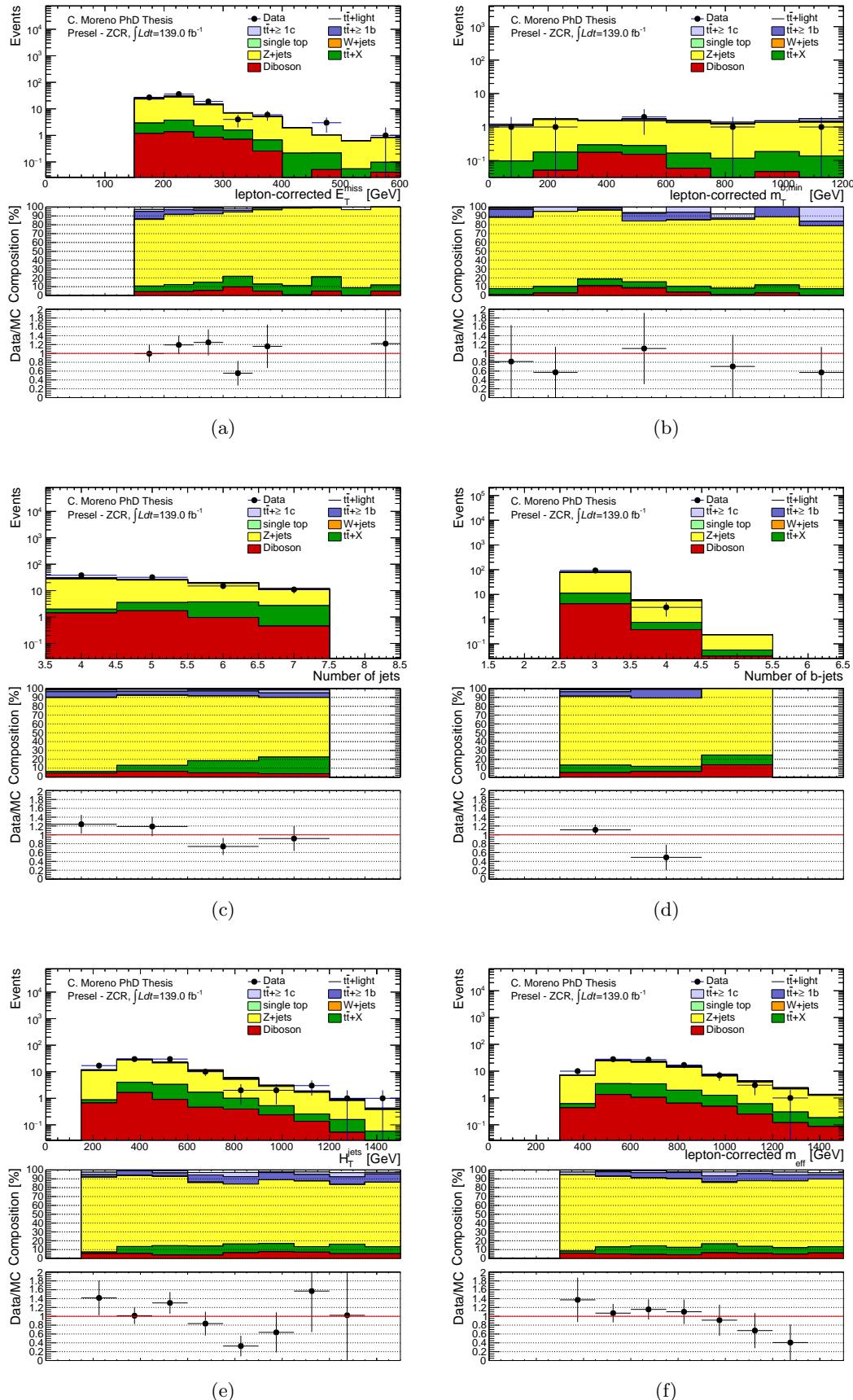


Figure 7.11: Comparison of data and background prediction for the kinematic distributions of (a)  $E_T^{\text{miss}}$ , (b)  $m_{T,\min}^b$ , (c)  $N_{\text{jet}}$ , (d)  $N_{b\text{-jet}}$ , (e)  $H_T$  and (f)  $m_{\text{eff}}$  in region Presel-Z. Only statistical uncertainties are included.

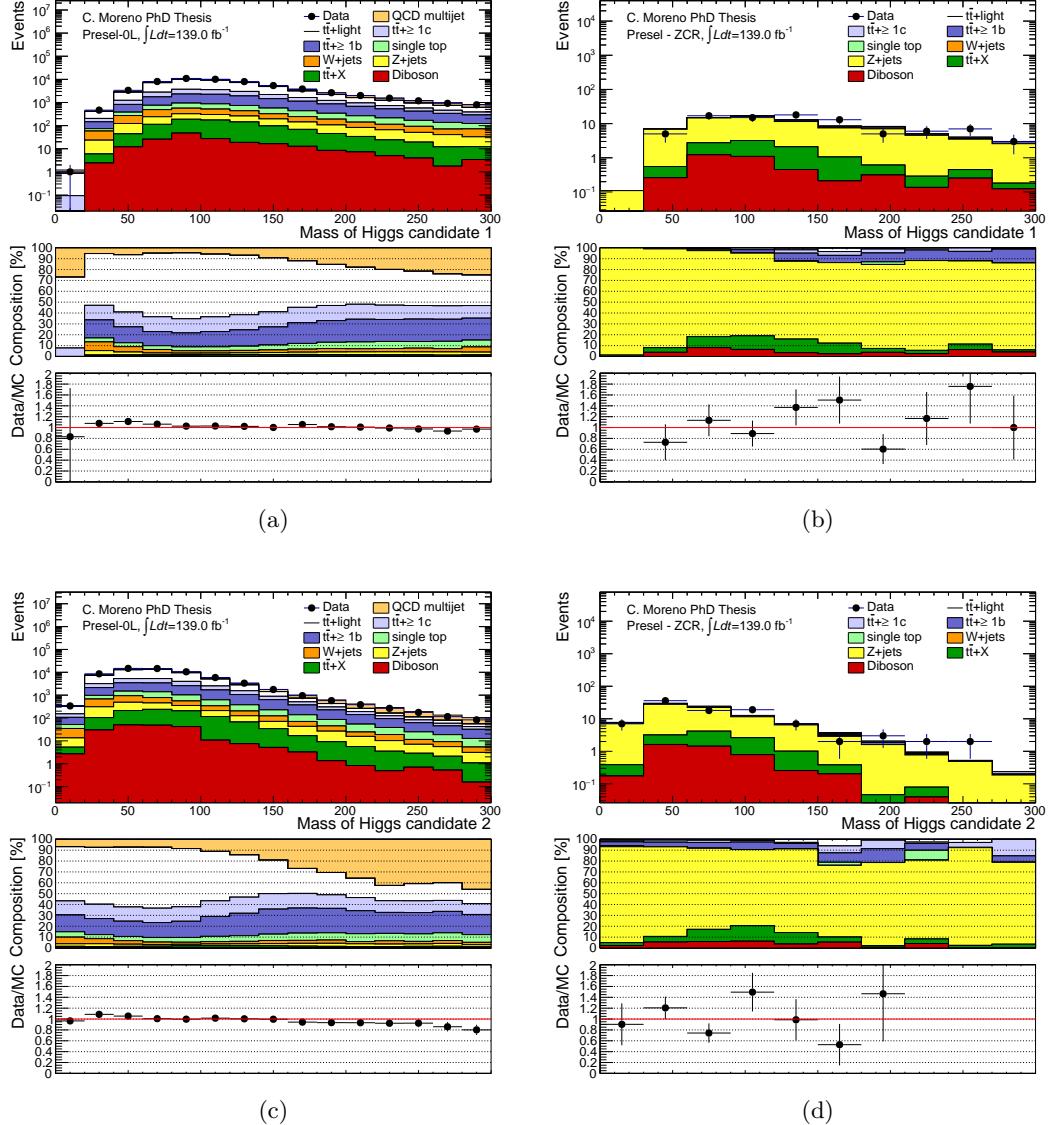


Figure 7.12: Comparison of data and background prediction for the mass of the leading (top row) and subleading (bottom row) reconstructed Higgs boson candidates, in regions given by the requirements defined in (a, c) Presel-0L and (b, d) Presel-Z. Only statistical uncertainties are included.

## 7.9 Signal-to-background discrimination

As specified in section 7.1, this analysis relies on defining different regions with different signal-to-background ratio. In order to do so, it is essential to find a way to separate the processes as efficiently as possible. In the previous analysis targeting the same signal model and using partial Run 2 dataset, the discrimination between signal and background was performed by means of selections on key kinematic variables that showed good separation between both processes. These cuts were optimized for maximal statistical significance in the signal regions. To define the control and validation regions, some of the kinematic cuts were inverted in order to make them orthogonal to the signal regions while still being kinematically similar.

In this case, part of the optimization for this round of the analysis included exploring MVA tools to carry out the signal-to-background separation task and use them to define the analysis regions for the statistical analysis. Both NN and BDT approaches were investigated, and finally a BDT was chosen to provide the discrimination as its performance matched that of the explored NN while being a simpler algorithm overall. In the following subsections the approach used in this analysis to separate signal and background is developed.

### 7.9.1 BDT discriminant

In order to train the BDT, a master dataset is formed from all the background and signal simulated samples for all events surviving the set of selections Presel-0L with a looser cut in offline- $E_T^{\text{miss}}$  of 100 GeV. This not only allowed for the flexibility of choosing at a later moment the final selection on offline- $E_T^{\text{miss}}$ , but also allowed the BDT to train more efficiently for low higgsino masses due to the increase of statistics. To ensure the training is optimal for all of the higgsino mass hypotheses, the BDT is parametrized as a function of the  $\tilde{\chi}_1^0$  mass [151]. All of the available samples with simulated signal events are used together, with the higgsino mass for each sample used as an input variable in the training. In order to prevent the BDT from learning to discriminate the signal and background processes based on the higgsino mass variable, the same distribution of higgsino mass is assigned randomly to background events. The BDT then provides one output score per mass hypothesis, which can be interpreted as the probability of an event to be signal-like with higgsino mass determined by the parameter.

Both signal and background samples are split into two equal sized sets, one used for the BDT training and the other for model selection (test set). A series of scans in the BDT hyperparameters and variables used in the training is performed, which is discussed in section 7.9.2. The background events are scaled down prior to the training to match the number of signal events in each dataset. All BDTs are trained using weighted signal and background events, as event weights modify the shapes of the input variables for the training. Once the selected BDT model is in place after hyperparameter/variable scans, the signal and background samples are evaluated and the BDT score is used to construct the analysis regions. After the optimal combination of hyperparameters and input variables has been found, a second BDT with the same hyperparameters/variables is trained using the test set, allowing to have an unbiased BDT score for the training set as well, and allowing for the usage of all the MC statistics in the analysis. For data samples, all of the events are evaluated using both BDTs by splitting the complete dataset into two samples prior to the evaluation, similarly to what is done in the simulated samples to obtain the training and test sets.

All the considered BDTs are trained using the *Extreme Gradient Boosting* algorithm (XG-Boost) [152] configured with regression trees as weak learners, and using an approximate method

Input variables	$m(\tilde{H}) = 250$ GeV	$m(\tilde{H}) = 500$ GeV	$m(\tilde{H}) = 1000$ GeV
Base variables + $m_{\text{eff}}$	0.7900	0.9549	0.9938
Base variables + $H_T$	0.7901	0.9550	0.9939
Base variables + $H_T + E_T^{\text{miss}}$ sign.	0.7997	0.9581	0.9941
Base variables + $H_T + E_T^{\text{miss}}$ sign. + $\Delta R_{\min}^{bb}$	0.8073	0.9591	0.9944

Table 7.4: Comparison of BDT performance in terms of AUC for three benchmark signal mass points. The "base variables" is a set of input variables formed by number of jets, number of b-jets,  $E_T^{\text{miss}}$ ,  $\Delta\phi_{\min}^{4j}$ ,  $m_{T,\min}^{b\text{-jets}}$ , the masses of both Higgs candidates, the separation  $\Delta R$  of the decay products of both Higgs candidates and  $M_J^{\sum}$ .

to select the split points based on pre-train sketching. Before the training, candidate split points are computed based on percentiles of the input variable weighted distributions, and these candidate split points are used throughout the training process. This method is less accurate than the XGBoost *Exact Greedy Algorithm* where all possible split points are examined, but it is also less resource consuming, especially given the large size of the datasets used in the analysis.

### 7.9.2 Optimization and performance

The BDT is trained using the following "high-level" kinematic variables:  $N_{\text{jet}}$ ,  $N_{b\text{-jet}}$ ,  $H_T$ ,  $E_T^{\text{miss}}$ , object-based  $E_T^{\text{miss}}$  significance,  $\Delta\phi_{\min}^{4j}$ ,  $M_J^{\sum,4}$ ,  $m_{T,\min}^{b\text{-jets}}$ . Additionally, the minimum separation between  $b$ -jets in the event,  $\Delta R_{\min}^{bb}$ , is used in the training as well, together with the mass of the leading and subleading reconstructed Higgs masses,  $m(\tilde{H}_{1,2})$ , and the separation of their decay products,  $\Delta R_{H_{1,2}}^{bb}$ . The list of variables used in the BDT training was selected from a wide range of kinematic distributions, including Higgs-related variables built with different Higgs reconstruction algorithms. First the list of variables is reduced to a nominal set of distributions based on the separation between signal and background events, computed as

$$\text{separation} = \frac{1}{2} \sum_{i=0}^{\text{n. bins}} \frac{(N_i^{\text{sig}} - N_i^{\text{bkg}})^2}{N_i^{\text{sig}} + N_i^{\text{bkg}}}, \quad (7.17)$$

where  $N_i$  is the number of signal/background events in each bin of the distribution. The nominal set of variables was selected maximizing the separation between three higgsino mass hypotheses (250 GeV, 500 GeV and 1000 GeV) and the sum of all background processes, and giving priority to the variables with higher separation for high mass signal points. The separation between the three mass hypotheses and background for the main discriminant variables is shown in figure 7.13.

Then separated trainings were performed using different subsets of kinematic distributions to asses the impact of reducing the number of input variables in terms of the performance of the BDT. The area under the ROC curve (AUC) was computed for all mass points for each training as a measure of the discriminating power of the BDT. Even though all the tests showed similar results, training with all the variables in the nominal set was found to provide the best performance (see table 7.4).

To select the optimal combination of hyperparameters, a wide scan for different values was performed. For each set of hyperparameters, the BDT is evaluated over the test sample and a test signal region is created for three benchmark mass points: 250 GeV, 500 GeV and 1000 GeV. The test SRs are defined to maximize the statistical significance, computed under the assumption

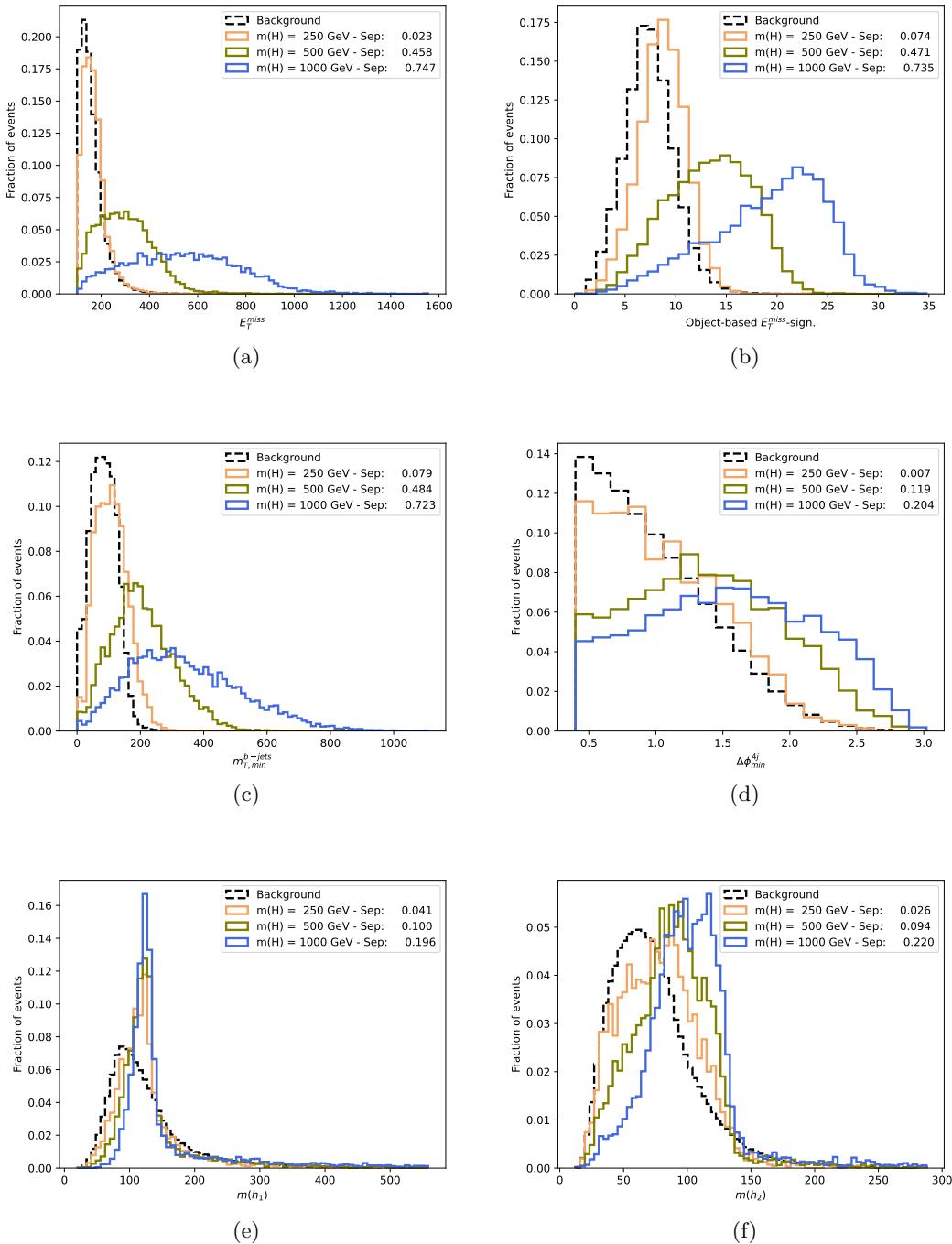


Figure 7.13: Main discriminant kinematic distributions for signal and background events in region Presel-0L, normalized to unit area.

of a normalization uncertainty of 50% for both signal and background processes. The selected hyperparameters are the ones that lead to maximum significance in the signal region for the mass point 1000 GeV, and while they are not the most optimal for low mass higgsinos, the performance of the chosen BDT for  $m(\tilde{H}) = 250, 500$  GeV is still in the top 10% of all models tested. The finally selected hyperparameters are collected in table 7.5.

Hyperparameter	value
Number of trees	500
Learning rate $\eta$	0.5
Subsample rate	0.8
Maximum number of bins	50
Maximum depth	6

Table 7.5: Selected combination of hyperparameters that optimize the BDT performance.

The distribution of the BDT output for these three mass hypothesis, both for signal and background, are shown in figure 7.14, with no signs of overtraining. The approach followed in the hyperparameter optimization prioritizes the background suppression at large BDT score values, rather than overall performance in the whole range between 0 and 1. Given the sharp shape of the BDT output distribution for the signal near 1, the SRs will be defined at very large BDT score values, while the rest of the distribution will not be considered in the statistical fit. In this case, a small percentage of surviving background in the SR translates into large event yields given the high cross section of the dominant processes. Therefore it was found that prioritizing the background reduction in the tail of BDT score distribution leads to enhanced sensitivity than optimization based on ROC curve.

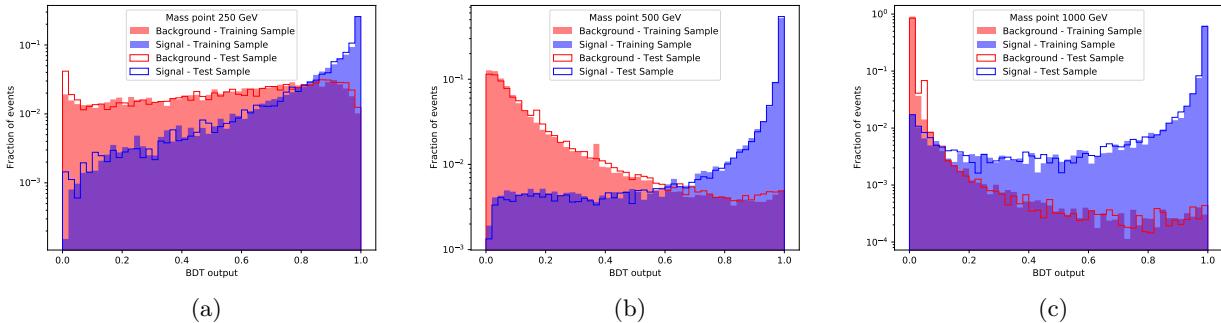


Figure 7.14: BDT score for both signal and background tested with the signal mass hypotheses (a) 250 GeV, (b) 500 GeV and (c) 1000 GeV. Both signal and background processes are normalized to unit.

### 7.9.3 Definition of analysis regions

All of the analysis regions that will be included in the likelihood fit are defined based on the BDT output score on top of the Presel-0L set of selections defined in section 7.4, with the exception of the control and validation regions for  $Z+jets$ , which use the set Presel-ZCR. The data and background prediction agreement for the BDT score distribution is shown in figure 7.15 for four higgsino mass hypotheses, with good agreement observed overall. The small disagreement at BDT output values close to 1 is corrected in the fit to data by means of the control regions. For each mass hypothesis<sup>7</sup> used in the training of the BDT an independent set of SRs, CRs and VRs is built. A fit is performed for each set of regions (i.e. one fit per mass point), to maximize sensitivity in the dedicated kinematic region that each mass hypothesis defines. The multiple

<sup>7</sup>The list of mass hypotheses considered is: 200, 250, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400 and 1500 GeV.

SRs for a given mass point are combined in a multi-bin fit in order to maximize exclusion, hence the regions in a given set are defined to be orthogonal.

The strategy to define all analysis regions, summarized in figure 7.16, is the following. Starting at the maximum BDT score, the SRs are constructed with all the events with BDT output in the range  $(X, 1]$ , where  $X$  is the cut that maximizes the expected significance in the SR. Multiple SRs can be defined as long as the expected significance in each of them is higher than 1. After finding all SRs, the lower threshold in BDT score of the last SR is used as upper cut of the VR, and the procedure is repeated separately in regions of low and high- $m_{T,\min}^{b\text{-jets}}$  to define VRs for  $t\bar{t}$  background and high- $m_{T,\min}^{b\text{-jets}}$  selections. Once the VRs are found, the process is repeated to define CRs for  $t\bar{t}$  and high- $m_{T,\min}^{b\text{-jets}}$  events, with a further splitting of the  $t\bar{t}$  CRs in b-jet multiplicity. To compute the number of signal events in each region, the signals are scaled down to the already excluded cross section by the previous round of this analysis in ATLAS [135]. All CRs and VRs are higgsino-mass dependent, and while they are expected to have a low signal contamination from the mass hypothesis they target, they can still have contributions from events corresponding to other signal mass points. Thus, in order to constrain the contribution from other higgsino mass hypotheses, an additional selection is added to CR and VR targeting the SM Higgs boson mass window in the leading reconstructed Higgs candidate,  $100 \text{ GeV} < m(h_1) < 150 \text{ GeV}$ . For all regions the signal contamination levels from other mass points are kept below 20% in the VRs and 10% in the CRs.

## Signal regions

An independent set of signal regions is defined for each point in the signal grid that is included in the BDT training. For a given mass hypothesis, in order to define the SRs, at each BDT output value  $X_1$  the expected significance is computed by taking the signal and background events above that threshold, BDT score  $> X_1$ . The expected significance is calculated with the function `BinomialExpZ` in the package `RooStats` [153] for a luminosity of  $139 \text{ fb}^{-1}$  and assuming a normalization uncertainty for signal and background of 50%. The signal region is constructed with all of the events with BDT score in the interval  $(X_1, 1.0]$ , choosing the value of  $X_1$  that maximizes the expected significance and provided that the background relative statistical uncertainty is lower than 30%. After finding one signal region the process is repeated, using the lower BDT score threshold  $X_1$  of the first SR as upper cut. If it is possible to find a value  $X_2$  that maximizes the expected significance in the interval  $(X_2, X_1]$  and fulfills: 1) such maximum significance is above 1.0, and 2) the expected significance in the region is at maximum 1.5 times lower than the significance of the previous signal region, then a second signal region is defined with all of the events whose BDT score is inside that range. This process is repeated until it is not possible to find regions with expected significance above 1. If more than 4 signal regions are found, then all of the regions found after the third one are merged into a single signal region, in a way that there are never more than 4 signal regions for any signal mass point.

A set of signal regions was found for all signal mass points in the interval  $[200, 1100]$ . For the points  $m(\tilde{H}) > 1100 \text{ GeV}$  the only signal region found had an expected significance below 1. For this reason, the signal region defined for the mass point  $m(\tilde{H}) = 1100 \text{ GeV}$  is used to assess the sensitivity of the analysis to all the signal mass points above 1100 GeV. The signal regions found for three benchmark mass points, together with the corresponding thresholds in BDT score and the expected signal and background events in each one, are summarized in tables 7.6 - 7.8. Similar tables for all of the mass points can be found in appendix B.

The expected background composition in the leading SR for each mass point is shown in

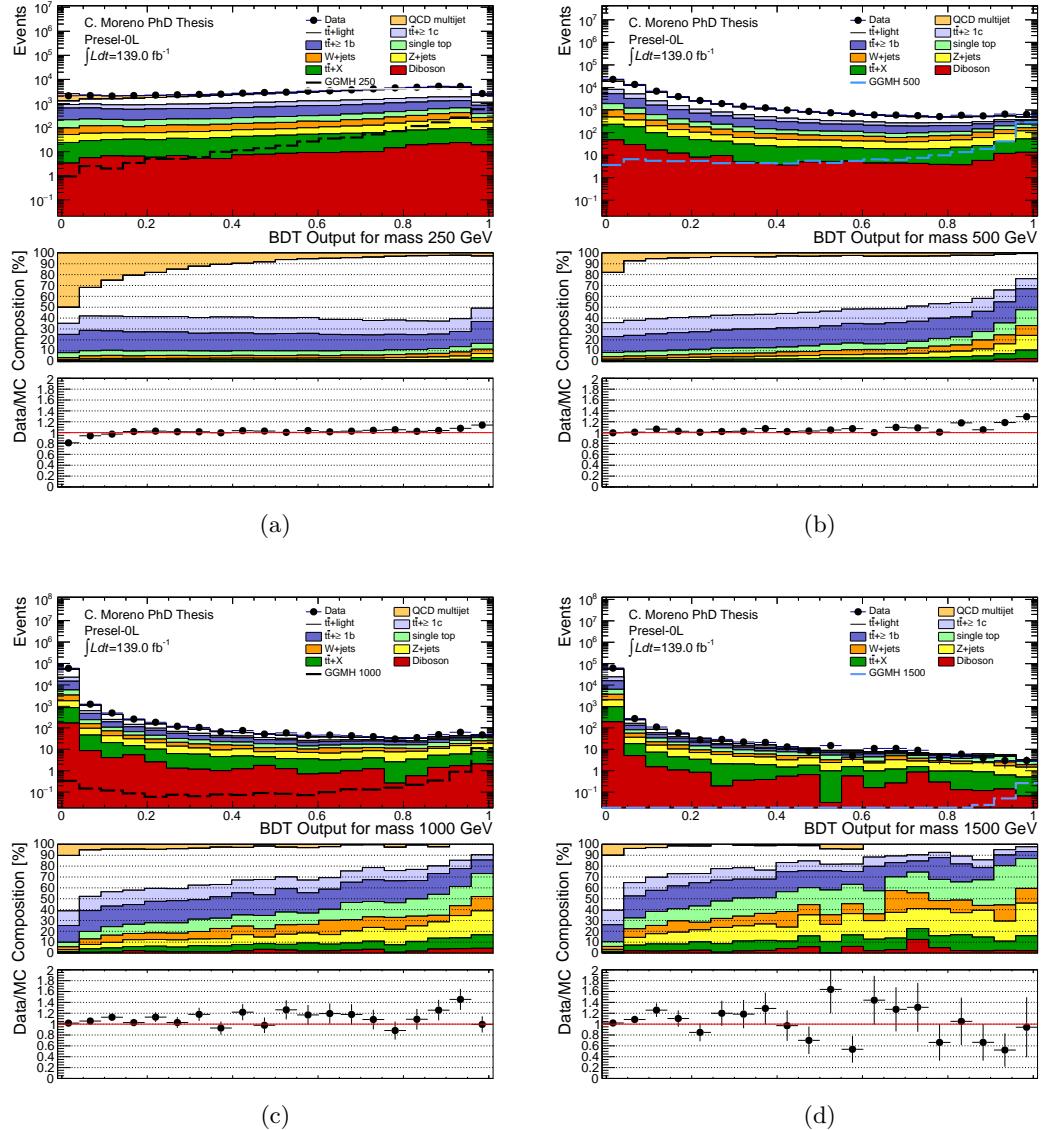


Figure 7.15: Comparison of data and background prediction for the BDT score distribution for the higgsino mass hypotheses (a)  $m(\tilde{H}) = 250 \text{ GeV}$ , (b)  $m(\tilde{H}) = 500 \text{ GeV}$ , (c)  $m(\tilde{H}) = 1000 \text{ GeV}$  and (d)  $m(\tilde{H}) = 1500 \text{ GeV}$ . Only statistical uncertainties are included.

Region	N. b-jets	$m_{T,\min}^{b\text{-jets}}$	BDT score	$N^{\text{sig}}$	$N^{\text{bkg}}$
SR-1-250	$\geq 3$	-	(0.9997, 1]	27.9031	1.56719
SR-2-250	$\geq 3$	-	(0.9994, 0.9997]	23.6513	5.06638
SR-3-250	$\geq 3$	-	(0.99846, 0.9994]	47.8136	21.4034
SR-4-250	$\geq 3$	-	(0.99786, 0.99846]	18.9394	14.4228
VR-tt-250	$\geq 3$	< 200 GeV	(0.969239, 0.99786]	80.3503	401.223
VR-st-250	$\geq 3$	$\geq 200$ GeV	(0.98232, 0.99786]	3.56394	24.9764
CR-tt3b-250	= 3	< 200 GeV	(0.963639, 0.969239]	7.96057	99.8456
CR-tt4b-250	$\geq 4$	< 200 GeV	(0.955179, 0.969239]	9.24107	99.8399
CR-st-250	$\geq 3$	$\geq 200$ GeV	(0.940219, 0.98232]	5.64447	99.8735
VR-Z-250	$\geq 3$	-	(0.538771, 1]	0	24.9861
CR-Z-250	$\geq 3$	-	(0.0961619, 0.538771]	0	99.9791

Table 7.6: Selection requirements for all the analysis regions for  $m(\tilde{H}) = 250$  GeV. The selection Presel-0L is applied to all except for CR-Z and VR-Z, that use the selection Presel-ZCR. The expected numbers of signal and background events in each region are collected in the last two columns.

Region	N. b-jets	$m_{T,\min}^{b\text{-jets}}$	BDT score	$N^{\text{sig}}$	$N^{\text{bkg}}$
SR-1-500	$\geq 3$	-	(0.9997, 1]	8.75667	0.618144
SR-2-500	$\geq 3$	-	(0.9994, 0.9997]	5.14364	1.61243
SR-3-500	$\geq 3$	-	(0.99878, 0.9994]	6.41886	4.84762
VR-tt-500	$\geq 3$	< 200 GeV	(0.974519, 0.99878]	1.54755	24.9946
VR-st-500	$\geq 3$	$\geq 200$ GeV	(0.98888, 0.99878]	4.20567	25.0773
CR-tt3b-500	= 3	< 200 GeV	(0.924498, 0.974519]	2.20954	99.9869
CR-tt4b-500	$\geq 4$	< 200 GeV	(0.787956, 0.974519]	1.15961	99.9717
CR-st-500	$\geq 3$	$\geq 200$ GeV	(0.954699, 0.98888]	5.46219	99.9548
VR-Z-500	$\geq 3$	-	(0.0454809, 1]	0	24.9997
CR-Z-500	$\geq 3$	-	(0.00534011, 0.0454809]	0	99.7545

Table 7.7: Selection requirements for all the analysis regions for  $m(\tilde{H}) = 500$  GeV. The selection Presel-0L is applied to all except for CR-Z and VR-Z, that use the selection Presel-ZCR. The expected numbers of signal and background events in each region are collected in the last two columns.

Region	N. b-jets	$m_{T,\min}^{b\text{-jets}}$	BDT score	$N^{\text{sig}}$	$N^{\text{bkg}}$
SR-1-1000	$\geq 3$	-	(0.9991, 1]	4.75919	0.554438
VR-tt-1000	$\geq 3$	< 200 GeV	(0.721194, 0.9991]	0.236075	24.9972
VR-st-1000	$\geq 3$	$\geq 200$ GeV	(0.922958, 0.9991]	1.77659	24.7843
CR-tt3b-1000	= 3	< 200 GeV	(0.331647, 0.721194]	0.145877	99.9196
CR-tt4b-1000	$\geq 4$	< 200 GeV	(0.0630013, 0.721194]	0.0657878	100.001
CR-st-1000	$\geq 3$	$\geq 200$ GeV	(0.332187, 0.922958]	0.531473	99.9527
VR-Z-1000	$\geq 3$	-	(0.000820016, 1]	0	23.5204
CR-Z-1000	$\geq 3$	-	(0.000100002, 0.000820016]	0	98.7919

Table 7.8: Selection requirements for all the analysis regions for  $m(\tilde{H}) = 1000$  GeV. The selection Presel-0L is applied to all except for CR-Z and VR-Z, that use the selection Presel-ZCR. The expected numbers of signal and background events in each region are collected in the last two columns.

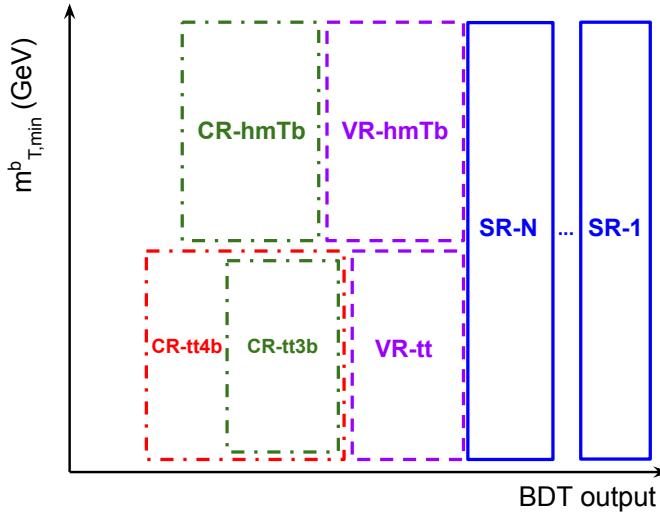


Figure 7.16: General scheme followed in the definition of SR, VR and CR for each mass point.

figure 7.17. For low mass points, the SM process expected to contribute the most in the SRs is  $t\bar{t}$ , with equally relevant contributions from  $t\bar{t}$  produced in association with light jets and in association with  $b$ -jets. As the higgsino mass increases, so does the contribution from single-top-quark production, up to 30% of the total yield in the SRs, due to the BDT selecting events in the tail of the  $m_{T,\min}^{b\text{-jets}}$  distribution as signal-like. At very high masses, the background in the SRs is dominated by processes in which  $Z$  bosons are generated in association with jets, with the  $Z$  boson decaying to a pair of neutrinos that make up most of the missing energy in the event. Dedicated CR and VR for  $t\bar{t}$  and  $Z$ +jets backgrounds are defined below. The case of single-top-quark is treated differently, since it was not possible to find a selection with high enough purity for this process to resolve its normalization in the fit to data. Instead, extra CR/VR are defined using events with high  $m_{T,\min}^{b\text{-jets}}$  values in order to improve the agreement between background estimate and data in this region.

### Control and validation regions for low and high $m_{T,\min}^{b\text{-jets}}$ events

After finding all possible signal regions for a given mass point, a similar procedure is carried out in order to define the validation and control regions. The regions for  $t\bar{t}$  background are kept orthogonal from the regions for high  $m_{T,\min}^{b\text{-jets}}$  events by applying complementary cuts on the variable  $m_{T,\min}^{b\text{-jets}}$  on top of the Presel-0L selections, with the splitting point at  $m_{T,\min}^{b\text{-jets}} = 200$  GeV.

The procedure to define the validation region is identical for low and high  $m_{T,\min}^{b\text{-jets}}$  selections and is run independently of each other. For each of them, the lower threshold of the last signal region found is used as upper cut in the BDT output. The validation regions are constructed with all of the events with BDT output in the range  $(X_{\text{VR}}, X_{\text{SR}_i}]$ , where the value  $X_{\text{VR}}$  is chosen so that there are at least 25 weighted background events and that the signal contamination in the region is lower than 20%.

Once the high- $m_{T,\min}^{b\text{-jets}}$  validation region has been found, the same procedure is followed to define the corresponding control region using the same selection as the VR-hmTb. The constructed control region  $(X_{\text{CR}}, X_{\text{VR}}]$  is required to have at least 100 weighted background

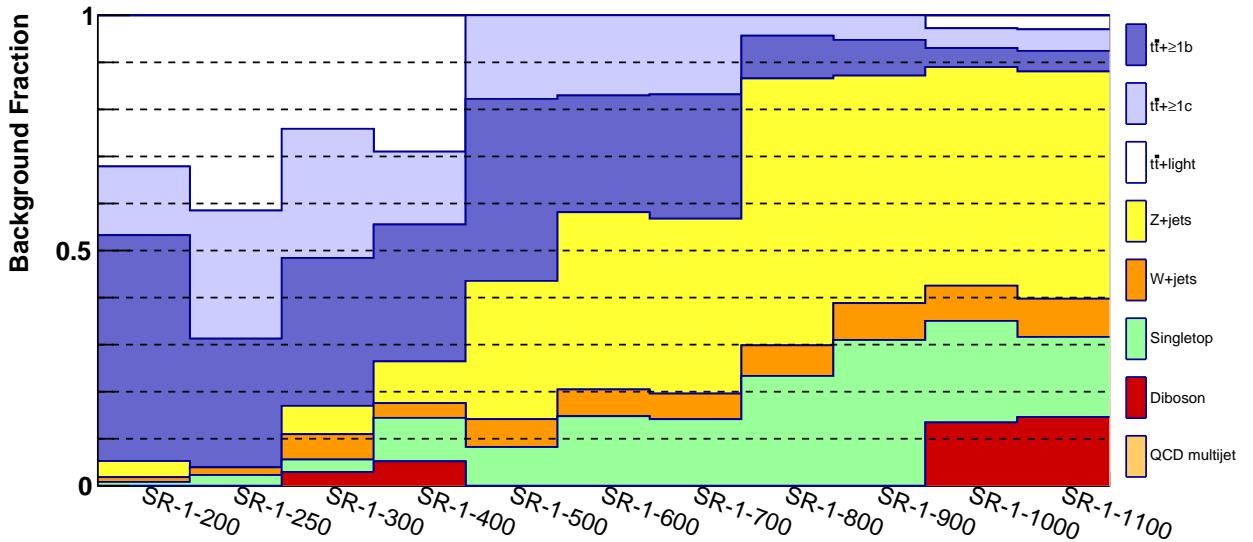


Figure 7.17: Background composition in the leading signal region for each mass point. As the neutralino mass increases, the dominant background shifts from  $t\bar{t}$  events to  $Z + \text{jets}$ .

events and signal contamination lower than 10%.

For  $t\bar{t}$  background, the subset of events used to define the control regions is further split into two subsets formed by events with exactly 3  $b$ -jets and events with at least 4  $b$ -jets. This strategy allows to define two control regions enriched in  $t\bar{t}$  events but with different percentage of  $t\bar{t} + \text{light}$ ,  $t\bar{t} + \geq 1c$  and  $t\bar{t} + \geq 1b$ , which is useful to fit separately the normalization of each component. The procedure to define both control regions is identical to the case described above, and is performed separately for each of the subsets defined in terms of  $b$ -jet multiplicity.

All of the validation and control regions defined for all of the mass points are summarized in tables 7.6–7.8. The background composition in the control and validation regions for  $t\bar{t}$  and high- $m_{T,\min}^{b\text{-jets}}$  events is shown in figures 7.18–7.22. Control and validation regions for  $t\bar{t}$  achieve large purity of such process, with varying composition of the heavy flavour components in the regions with 3 and 4  $b$ -jets, which will allow for good accuracy in the estimation of its normalization when fitted to data. The regions for high- $m_{T,\min}^{b\text{-jets}}$  events show large contribution from single-top-quark events, and even though its purity of about 30% is typically not enough to correct the normalization for such background, the addition of this control region to the fit helps correct the normalization of the other processes and better constrain their uncertainties in this extreme kinematic region.

### Control and validation regions for $Z + \text{jets}$

As discussed above, the process  $Z + \text{jets}$  becomes the dominant background in the signal regions for high mass due to highly energetic neutrinos emitted in  $Z \rightarrow \nu\nu$  decays. In order to define control and validation regions for these processes, events with exactly two leptons are used, where the leptons are added to  $E_T^{\text{miss}}$  in order to simulate the presence of neutrinos. A new sample is therefore constructed with the selections from Presel-ZCR defined in section 7.4. The BDT is then evaluated over the sample, using the lepton-corrected version of the different input variables in the training. With the resulting BDT output, for each signal mass point a similar procedure to the one described above is followed in order to define the control and validation for  $Z + \text{jets}$ , with the difference that no signal region is looked for this time since the two lepton

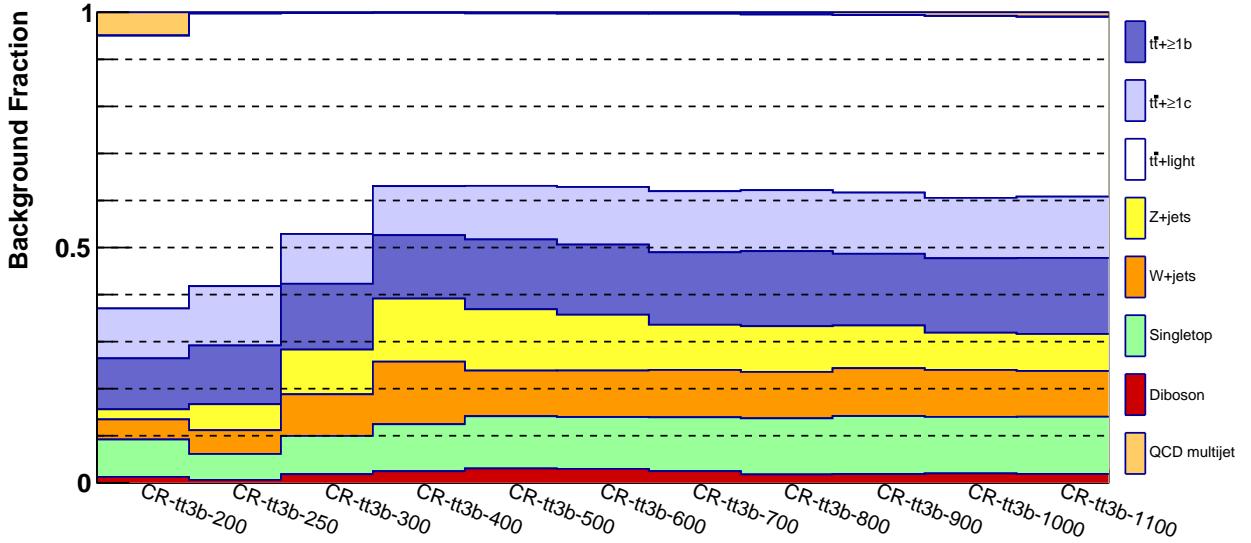


Figure 7.18: Background composition in the 3  $b$ -jet control regions for  $t\bar{t}$  events for each mass point.

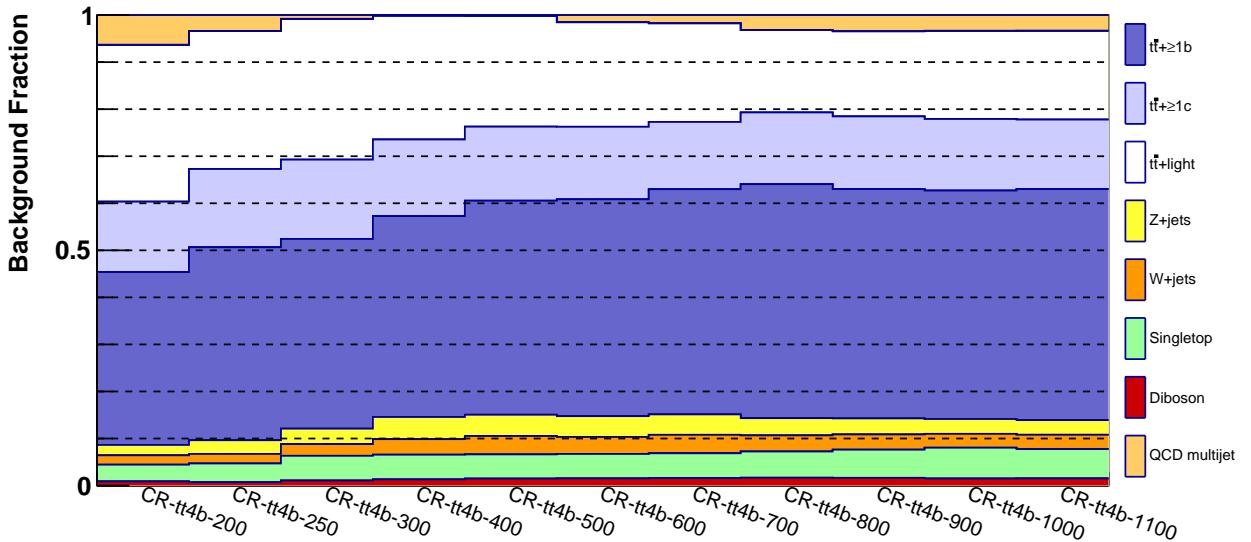


Figure 7.19: Background composition in the 4  $b$ -jet control regions for  $t\bar{t}$  events for each mass point.

requirement effectively vetoes all signal events. The requirements for the validation region are a minimum of 25 weighted background events, while for the control region is a minimum of 100 weighted background events. The background composition of these regions are shown in figures 7.23 and 7.24, where a large purity of  $Z + \text{jets}$  events is observed.

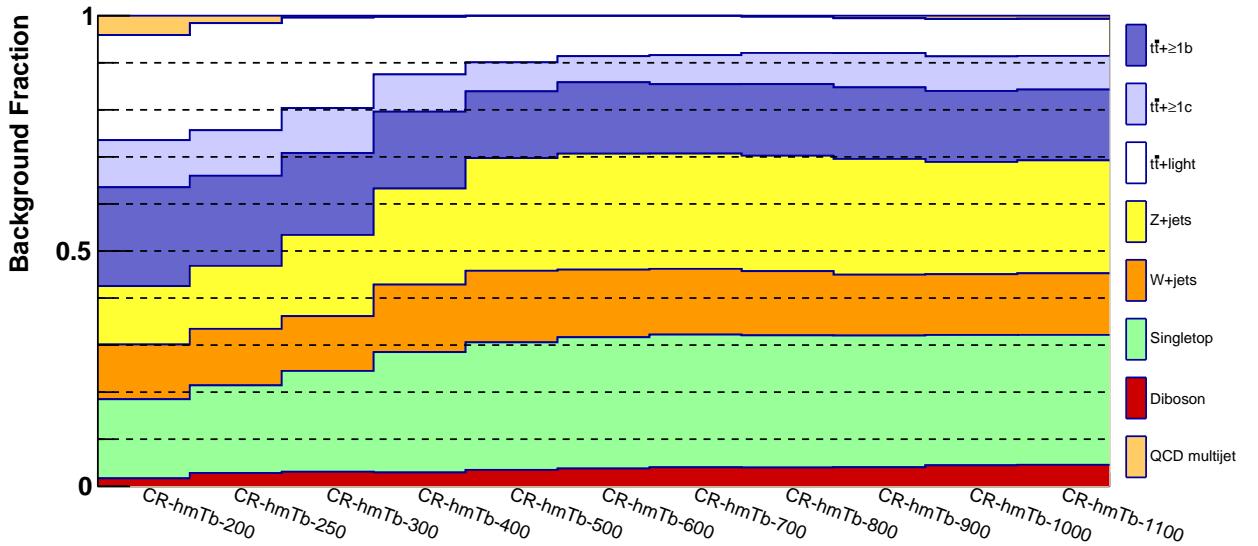


Figure 7.20: Background composition in the control regions for high- $m_{T,\min}^{b\text{-jets}}$  events for each mass point.

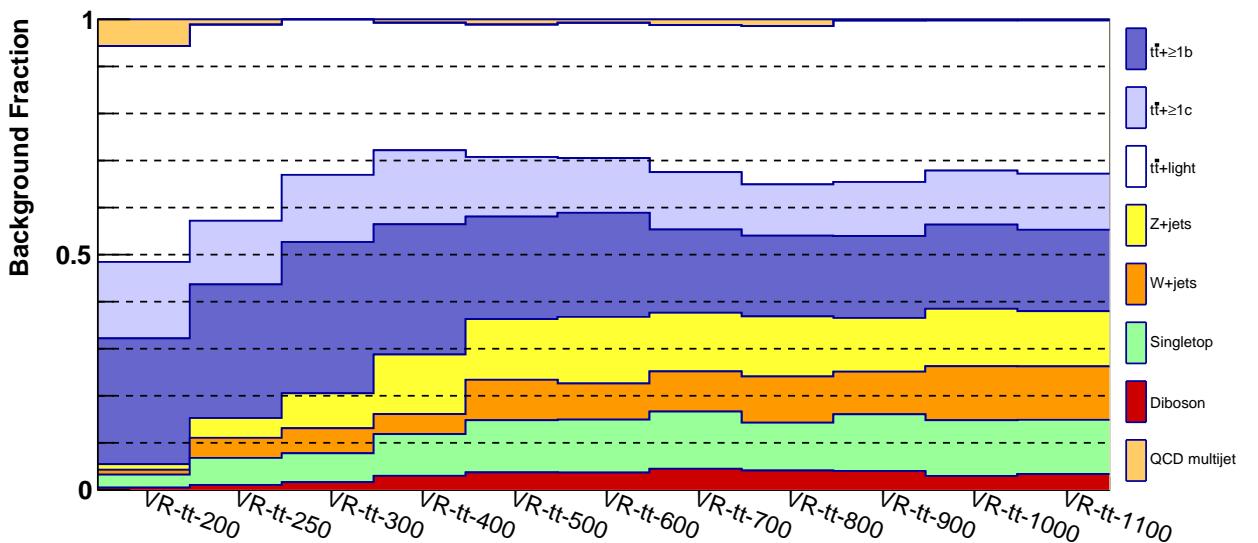


Figure 7.21: Background composition in the validation regions for  $t\bar{t}$  events for each mass point.

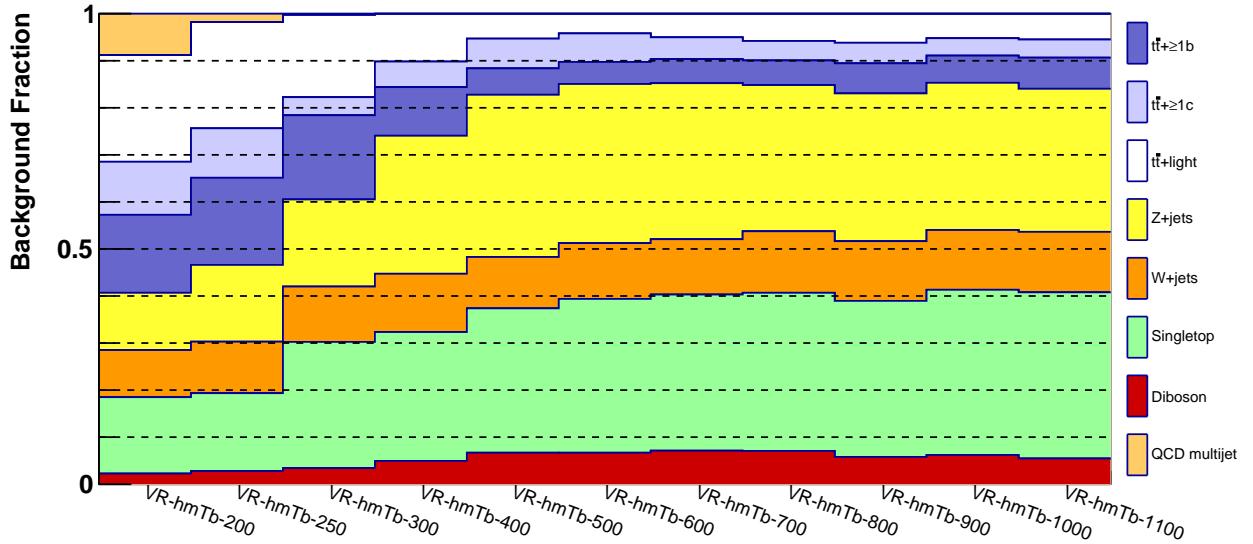


Figure 7.22: Background composition in the validation regions for high- $m_{T,\min}^{b-\text{jets}}$  events for each mass point.

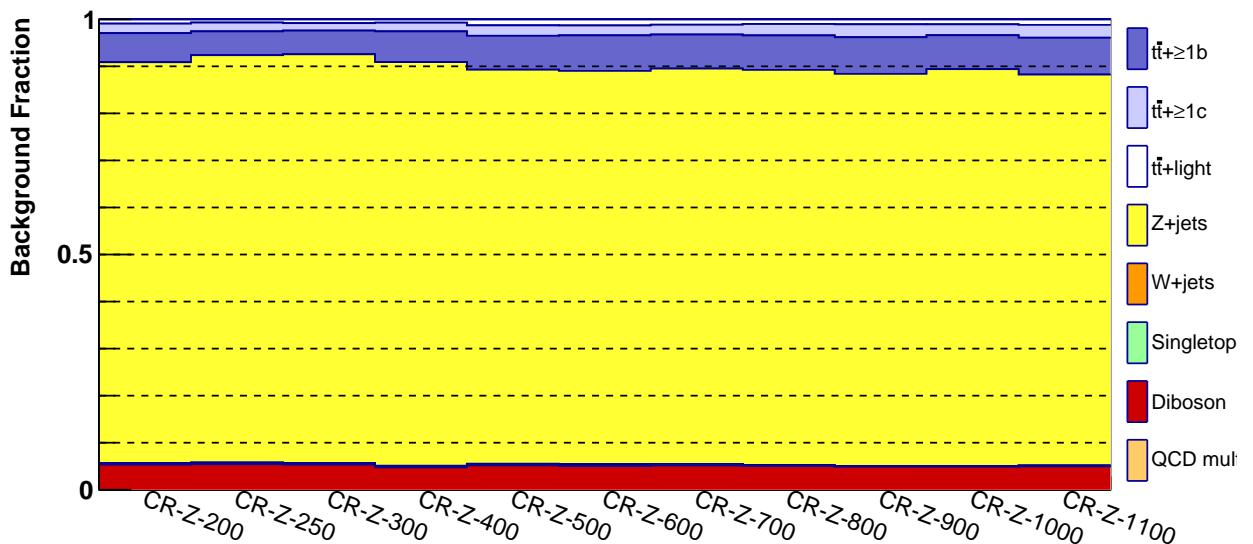


Figure 7.23: Background composition in the control regions for  $Z + \text{jets}$  events for each mass point.

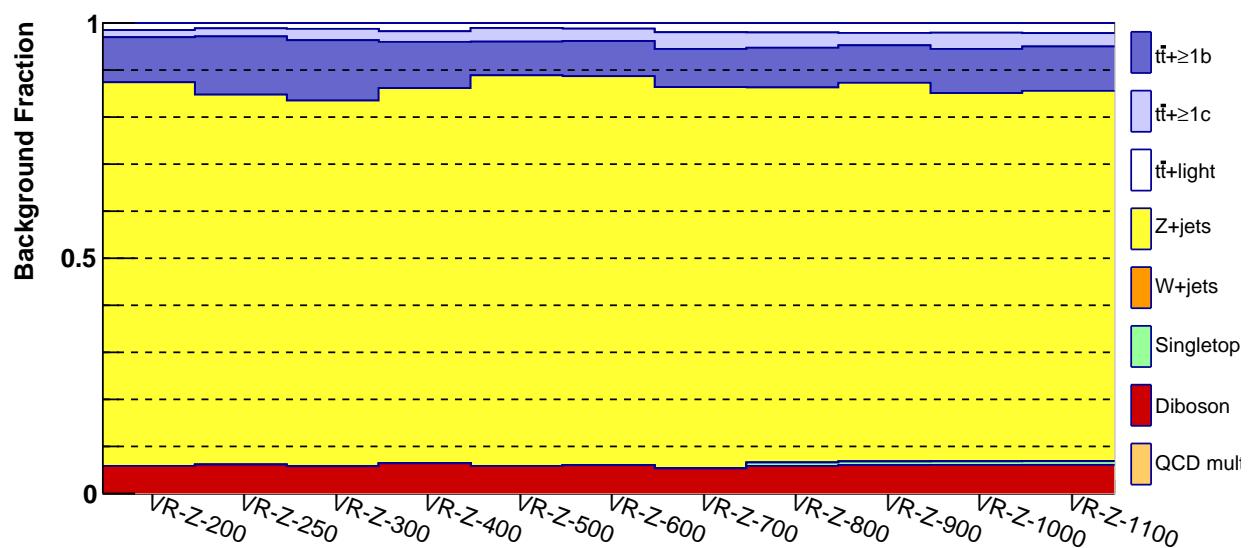


Figure 7.24: Background composition in the validation regions for  $Z + \text{jets}$  events for each mass point.



# Chapter 8

## Statistical methods and results

### 8.1 Introduction

The strategy presented in the previous chapter is a complex approach to search for new resonant phenomena that could be present in the recorded data. The predictions over the signal and the background processes, however, are not expected to match perfectly with the observations due to different effects. On the one hand, the possibility of statistical fluctuations, in particular in tight regions with low number of events, can lead to discrepancies between data and predictions. On the other hand, the modelling of the physics processes happening in the collisions might not be perfect, which also results in differences between data and simulation. In addition, while the understanding of the detector response is generally good, its behaviour as an ensemble is extremely complex and differences between its modelling in simulated samples and the actual recorded data are expected. These differences are typically included in the analysis as uncertainties in the measured properties of the reconstructed objects in the event, as discussed in chapter 6.

The agreement between data and predictions derived with the methods developed in chapter 7 is therefore treated by means of a statistical analysis, which allows for both the extraction of relevant information from the observations and testing the hypothesis of BSM signal present in the recorded dataset. This is performed using all recorded events during LHC Run 2 with the ATLAS detector, which after the trigger requirement and good data quality criteria correspond to an integrated luminosity of  $139 \text{ fb}^{-1}$ . The strategy followed relies on determining the cross section of the relevant backgrounds in the kinematic region explored,  $t\bar{t}$  and  $Z+\text{jets}$ , together with the values of the signal cross section compatible with the observation. The approach discussed in the previous chapter, which defines control and signal regions dominated by background and signal events, respectively, is key in the determination of the normalization for the different processes. Through statistical inference tools, the observed cross section for the main backgrounds and the signal can be extracted in the dedicated control and signal regions. From the estimated values of the signal normalization compatible with the observation, a statement can be made on statistical terms about the presence of the predicted signal.

This chapter is organized as follows. In section 8.2, the main concepts and tools for parameter estimation from data observations are introduced, while in section 8.3 the key ideas in hypothesis testing theory are developed. Section 8.4 provides an overview of the systematic uncertainties included in the analysis of the recorded data, which is split in the contributions coming from detector response and those associated with the modelling of the background processes. The results from the fit to the control regions in order to estimate the normalization of

the backgrounds is provided in section 8.5, while the interpretations of the results, in terms of exclusion limits, are provided in section 8.6.

## 8.2 Parameter estimation

In the statistical interpretation of the LHC physics analyses, the recorded properties of the collisions are considered as random variables and treated by means of probability theory. These random variables follow a given probability distribution function (*pdf*) that connects the measurements with the theoretical models developed in the previous chapters. In this context, if all the information regarding the physics process is encoded in a set of  $n$  random variables,  $\vec{x} = (x_1, \dots, x_n)$ , each event is considered as a measurement where each of those variables take a specific value,  $\vec{\bar{x}} = (\bar{x}_1, \dots, \bar{x}_n)$ . If the physics model that describes the recorded data depends on a series of parameters  $\vec{\theta} = (\theta_1, \dots, \theta_m)$ , so does the *pdf* that describes the behaviour of the random variables  $\vec{x}$ , whose joint *pdf* can therefore be written as  $f(\vec{x}|\vec{\theta})$ .

The first step in the statistical treatment in a search for new physics is to determine the key parameters  $\vec{\theta}$  from the experimental data in a procedure called *statistical inference* or *fit to data*. The determined values from data are later used to test the hypothesis of BSM physics present in the data sample against the hypothesis of only SM events processes taking place. Due to the intrinsic randomness of the observables in the analysis, a perfect knowledge of the true values of the parameters is never possible, and therefore any estimated value always comes with an associated uncertainty. For a given parameter  $\theta$  of the physics model describing the data, the following notation is used,

$$\theta = \hat{\theta} \pm \delta\theta, \quad (8.1)$$

where  $\hat{\theta}$  is the estimated value of the parameter, and  $\delta\theta$  the associated uncertainty. In this notation it is clear that in the process of estimating a parameter from data, an *uncertainty interval* is defined as  $[\hat{\theta} - \delta\theta, \hat{\theta} + \delta\theta]$ , with  $\hat{\theta}$  being the *central value* of the interval. In some cases the uncertainty interval is not symmetric with respect to the central value, for example to avoid getting into regions with unphysical values for  $\hat{\theta}$ . The usual notation in cases where the positive and negative uncertainties are different is

$$\theta = \hat{\theta}_{-\delta\theta_-}^{+\delta\theta_+}. \quad (8.2)$$

The size of the uncertainty interval is determined by a given probability level based on the accuracy on the measurement of the true value. Its interpretation therefore depends on the concept of probability itself, which differs in the two following approaches:

- In the *frequentist approach*, probability is understood as the fraction of favourable outcomes in the total number of repetitions of an experiment. In this sense, if an uncertainty interval is defined to provide a coverage of  $\alpha \leq 1$ , the true value of  $\theta$  is expected to fall inside the uncertainty interval a fraction  $\alpha$  of the repeated experiments. The probability  $\alpha$  is commonly referred to as *confidence level*.
- In the *bayesian approach*, probability is a measure of the degree of belief about a certain outcome to be true, which is modified through repeated experiments. In this context, the uncertainty or *credible* interval is believed to contain the true value of  $\theta$  with a probability of  $\alpha$ .

The two approaches typically lead to different results due to the different treatment of the data. In the following, only the frequentist approach will be discussed, as it is the base for the methods followed in this analysis.

The function that takes recorded data as input and returns the determined value  $\hat{\theta}$  for a given parameter is called *estimator*. They can be simple, such as measured observable, or complex algorithms combining multiple observables across the whole dataset. The statistical properties of a given estimator depend on the way it is constructed and its complexity. Some of the main properties used to evaluate the performance of an estimator are the following:

- Consistency, which encodes information about the accuracy of the estimate in the case of a large number of experiments. An estimator is said to be consistent if the estimated value  $\hat{\theta}$  converges to the true value in the limit of infinite measurements,

$$\lim_{N \rightarrow \infty} (\hat{\theta} - \theta) = 0.$$

- The bias of an estimator is defined as the difference between its expected value and the true value of the given parameter,

$$b(\hat{\theta}) = \langle \hat{\theta} \rangle - \theta.$$

In general, an *unbiased* estimator is preferred, which can be constructed from the estimator of  $\theta$  and subtracting the bias term.

- Efficiency, which indicates the precision of the estimator in terms of its variance  $V[\hat{\theta}]$ . The minimum variance of any consistent estimator has a lower limit given by the Cramér-Rao inequality [154, 155]. The efficiency is defined as the ratio between this theoretical minimum,  $V_{\text{Cramér-Rao}}(\hat{\theta})$ , and the actual variance of the estimator,

$$\epsilon(\hat{\theta}) = \frac{V_{\text{Cramér-Rao}}(\hat{\theta})}{V[\hat{\theta}]}.$$

The efficiency defined in this way is always between 0 and 1, and an estimator is said to be efficient when  $\epsilon(\hat{\theta})$  is close to 1.

- An estimator is said to be *robust* if it has limited sensitivity to fluctuations in the data sample in the tails of the *pdf*, commonly referred to as *outliers*.

### 8.2.1 Parameters of interest and nuisance parameters

A given physics theory modelled by a *pdf* of the form  $f(\vec{x}|\vec{\theta})$  needs to describe every effect that is observed in data. In the case of ATLAS analyses, the *pdf* not only needs to describe the different particle physics processes involved, but also the detector response, which typically result in complex forms of the *pdf* with large number of unknown parameters to be determined. It is typically desired, however, to express the results of the statistical analysis in function of the most relevant parameters, without showing explicit dependence on all of them. In the case of a search for BSM resonant particle production, for example, the *parameters of interest* (POI) usually are the background and signal cross sections and the mass of the BSM particle. The rest of them, known as *nuisance parameters* (NP), such as detector response characterization, are not as relevant in such context and, while they also impact the results of the search, their determination is not the main goal of the analysis. The choice of POI and NP depends on the analysis in

question. For a search of new physics, for example, the main POIs are usually those related to the modelling of the physics processes, whereas in analyses focusing on object calibration the parameters related to detector response become the POIs while parameters related to the physics processes happening in the collisions are typically treated as NPs.

The process of fitting the predictions to experimental data has as goal estimating both POIs and NPs in a given model and, as mentioned above in both cases an associated uncertainty is also computed. The uncertainties on the NPs is reflected in the accuracy with which the POI can be determined. In this case, the uncertainties arising from the propagation of our imperfect knowledge of the true value for the NPs are referred to as *systematic uncertainties*. Uncertainties on the final measurements that arise due to pure fluctuations in the experimental data are called *statistical uncertainties*.

Given the complex nature of the analyses performed in ATLAS, the number of total parameters in a *pdf* is usually very large. In order to improve the performance of the final measurement, an uncertainty interval for each NP is typically provided as a first estimate.<sup>1</sup> This uncertainty interval is usually computed by evaluating alternative samples where the value of a given NP is modified. The final measurement for the NP is provided after the fit to the observed data, where both the central value and the post-fit uncertainty may be different to those provided initially.

### 8.2.2 Maximum likelihood method

The *likelihood function* provides a powerful means for estimating the values of the parameters  $\vec{\theta}$  that determine the *pdf* describing the experimental data. The likelihood function  $L(\vec{x}|\vec{\theta})$  is defined as the compatibility of the measured values of the random variables,  $\vec{x}$ , with given values of the parameters  $\vec{\theta}$ . It is constructed from the *pdf* evaluated at measured data,  $\vec{x}$ ,

$$L(\vec{x}|\vec{\theta}) = f(\vec{x}|\vec{\theta}), \quad (8.3)$$

where the parameters  $\vec{\theta}$  are left to change freely as arguments of the function. The likelihood function defined in this way cannot be interpreted as a probability distribution, even though each point represents the probability of measuring the values  $\vec{x}$ , since the hypothesis under which  $\vec{x}$  is sampled does not remain constant. This is made clear by the fact that the integral of the likelihood function over all parameter space does not necessarily add up to one.

With this definition, it is intuitive to use the likelihood function to find which set of parameters  $\vec{\theta}$  result in highest probability for the measured variables  $\vec{x}$ . This method consists of finding the values  $\hat{\vec{\theta}}$  that maximize the likelihood function  $L(\vec{x}|\vec{\theta})$ , which translates the task of estimating the parameters of the model to that of finding the extremes of a function. Depending on the complexity of the problem, the likelihood function may not have an expression that can be maximized analytically and numerical methods need to be employed. In particular, the negative logarithm of the likelihood function,  $-\log L(\vec{x}|\vec{\theta})$  is often used, since its minima corresponds to the maxima of  $L(\vec{x}|\vec{\theta})$  and is more convenient for the numerical computations, in particular in the case of multiple independent measurements. The estimator obtained from maximizing the likelihood function, or equivalently from minimizing  $-\log L(\vec{x}|\vec{\theta})$ , is known as *maximum likelihood estimator* (MLE), and in the limit of a large number of experiments fulfils the requirements specified above: it is consistent; its bias, if present, tends to zero, and its efficiency tends to one. They are also invariant under reparametrizations of the likelihood function: if  $\hat{\vec{\theta}}$  are

---

<sup>1</sup>The central value for NPs related to detector response is typically provided centrally by the detector performance groups in ATLAS. In cases of specific NPs affecting one analysis exclusively, the uncertainty is determined according to the analysis strategy.

the estimated values of the parameters after maximizing the likelihood function, then, for any transformation  $g$  over  $L(\vec{x}|\vec{\theta})$ , the values that maximize the transformed likelihood function are  $g(\hat{\vec{\theta}})$ .

Multiple measurements can also be combined in the determination of parameters in the model. If the measurements are statistically independent from each other, which is the case of physics analyses in ATLAS since different collisions do not bias each other, the global likelihood of the combination of experiments can be obtained from the product of *pdfs*. For a given number  $N$  of measurements of the set of  $n$  variables,  $\vec{x} = \{(\bar{x}_1^1, \dots, \bar{x}_n^1), \dots, (\bar{x}_1^N, \dots, \bar{x}_n^N)\}$ , the likelihood function can then be written as

$$L(\vec{x}|\vec{\theta}) = \prod_{i=1}^N f(\bar{x}_1^i, \dots, \bar{x}_n^i | \theta_1, \dots, \theta_m). \quad (8.4)$$

The procedure can be illustrated with the following example. Consider a single random variable  $x$  distributed as a Gaussian with unknown average  $\mu$  and known standard deviation  $\sigma$ ,

$$f(x|\mu) = A e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad (8.5)$$

where  $A$  is a normalization constant that does not affect the maximization of the likelihood function. If  $N$  measurements of the variable  $x$  are performed,  $\vec{x} = (\bar{x}^1, \dots, \bar{x}^N)$ , the likelihood function can be written as

$$L(\vec{x}|\mu) = \prod_{i=1}^N \frac{1}{\sigma\sqrt{2}} e^{-\frac{1}{2}\left(\frac{\bar{x}^i-\mu}{\sigma}\right)^2} = A' e^{-\frac{1}{2\sigma^2} \sum_{i=1}^N (\bar{x}^i - \mu)^2}. \quad (8.6)$$

In order to find the value of  $\mu$  that maximizes the likelihood function it is enough to find the zeroes of its derivative,

$$0 = \frac{\partial L(\vec{x}|\mu)}{\partial \mu} \Big|_{\mu=\hat{\mu}} = A' e^{-\frac{1}{2\sigma^2} \sum_{i=1}^N (\bar{x}^i - \mu)^2} \sum_{i=1}^N (\bar{x}^i - \mu) \Big|_{\mu=\hat{\mu}}. \quad (8.7)$$

The zeroes of the derivative only happen when the last factor is equal to zero, since  $A$  is constant and the exponential function is always positive. The MLE for the average  $\mu$  of the Gaussian distribution can readily be obtained from equation 8.7 as

$$\hat{\mu} = \frac{1}{N} \sum_{i=1}^N \bar{x}^i, \quad (8.8)$$

which, for this very simple example, corresponds to the average of the measured values of the random variable  $x$ , as expected. If the variance of the Gaussian  $\sigma$  was also an unknown parameter, the same procedure can be followed to determine its MLE as well, which yields

$$\widehat{\sigma^2} = \frac{1}{N} \sum_{i=1}^N (\bar{x}^i - \mu)^2. \quad (8.9)$$

Differently to the estimator for the average,  $\hat{\mu}$ , the estimator for the variance is biased, since

$$\langle \widehat{\sigma^2} \rangle = \frac{N-1}{N} \sigma^2, \quad (8.10)$$

which differs from the true value  $\sigma$  by a factor  $(N - 1)/N$ . This factor, however, tends to one in the limit of large number of experiments,  $N \rightarrow \infty$ , and therefore the expected value of the estimator matches its true value in said limit, as mentioned above.

The NPs included in the problem are also estimated in the minimization of the likelihood function. They are generally studied in auxiliary measurements where their MLE is determined. Ideally, the likelihood of these auxiliary measurements in which their central values are determined should be included in the likelihood function of the problem under study. In practice, however, this is usually not feasible and they enter the likelihood function through constraining terms. In general, a Gaussian constraint term is used, justified by the central limit theorem, although other functional terms are used when the MLE of the NP in question does not follow a Gaussian distribution.

In cases with large number of measurements  $N$ , the numerical computation of the likelihood function and finding its extremes can become inaccessible. In these cases, the sampled values of the random variables can be stored in histograms and the inputs to the fit would be the number of events in each of their bins, reducing the size of the computation involved in the estimate of the parameters  $\vec{\theta}$ . If the measurements are all independent, the expected number of events in each bin,  $n_i$ , is Poisson-distributed, with the average number of expected events being determined by the original *pdf*  $f(\vec{x}|\vec{\theta})$ . This is the case of LHC data, where events are statistically independent of each other.

In this context, for a given random variable  $x$ , each bin of the histogram storing it is considered as a measurement. The likelihood function can be written, including the NP constraining terms, as

$$L(\vec{x}|\vec{\theta}) = \prod_{i=1}^M \frac{\mu_i(\vec{\theta})^{n_i} e^{-\mu_i(\vec{\theta})}}{n_i!} \times \prod_{k=1}^{\# \text{NP}} \rho(\theta_k^{\text{NP}}), \quad (8.11)$$

where  $M$  is the total number of bins,  $n_i$  the number of observed events in the  $i$ -th bin, and  $\mu_i(\vec{\theta})$  is the expected number of events in the  $i$ -th bin, which depends on the original *pdf*  $f(x|\vec{\theta})$ , as is shown with the dependence of  $\mu$  on the parameters  $\vec{\theta}$ . In this equation, the term  $\rho(\theta_k^{\text{NP}})$  corresponds to the constraint on the  $k$ -th NP. As mentioned above, in these cases it is useful to minimize the negative logarithm of the likelihood function,  $-2 \log L$ , since the product over the number of bins turns into a sum of logarithms, making the problem easier to solve numerically.

Once the values of the parameters have been estimated, an associated uncertainty needs to be computed, as discussed above. In the approximation of large number of measurements, the MLE for a given parameter is expected to follow a Gaussian distribution, and the statistical uncertainty can be obtained from its variance. This variance can be accessed in two ways. On the one hand, the experimental result can be simulated multiple times using MC methods, which allows to get a different estimate of the parameter in each iteration. The variance of the distribution of the MLE after combining the measurement from all pseudo-experiments is already the statistical uncertainty on the estimated value of the parameter. On the other hand, the likelihood function of the problem is expected to tend to a Gaussian when the size of the sample used is large. The  $-\log L$  is then a parabola, and the uncertainty on the parameter can be obtained as the values for which  $-\log L$  increases by a factor of 2 with respect to its minimum.

## 8.3 Hypothesis testing

The final goal of a search for new physics is to provide a statement about the presence of new BSM processes populating the recorded data. This is carried out on a statistical basis by means of hypothesis testing techniques. In this case, the null hypothesis  $H_0$  is usually considered to be the presence of only SM events in the experimental data, and it is commonly referred to as *background-only* hypothesis. The alternative hypothesis,  $H_1$ , where signal processes also contribute to the recorded collisions, is tested against the null hypothesis based on the observations. This process leads to the rejection of one of them, within a certain confidence level based on the propagated uncertainties associated to the measurement of the relevant parameters estimated for each hypothesis.

### 8.3.1 Test statistic

The process of testing two different hypotheses,  $H_0$  and  $H_1$ , can be intuitively understood as finding which one of them is more compatible with the measured values of a collection of random variables,  $\vec{x} = (\bar{x}_1, \dots, \bar{x}_n)$ . Each hypothesis can be modelled with a different *pdf*, which describes the behaviour of the random variables  $\vec{x}$ , and can be written using the notation  $f(\vec{x}|H_0)$ , for the null hypothesis, and  $f(\vec{x}|H_1)$ , for the alternative one.

When the number of random variables associated to a problem is large, it is impractical to examine their individual behaviour under both hypotheses to determine which one is more likely to describe the observation. Instead, a *test statistic*  $t$  is defined that can summarize the information provided by both *pdfs* and that is distributed differently for each hypothesis.<sup>2</sup> It is a function of the random variables,  $t = t(\vec{x})$ , which is evaluated on the observed values,  $\bar{t} = t(\bar{x})$ , in order to test the compatibility of the measurement with the two models under comparison. The less overlap present in the distribution of the test statistic under both hypotheses, the easier it is to provide a statement about which of them should be rejected.

Figure 8.1 shows a practical example of the application of a test statistic to test the hypothesis  $H_1$  of a  $J/\Psi$  meson produced in a proton-proton collision against the null hypothesis  $H_0$ . In the case of  $J/\Psi \rightarrow l^-l^+$  decays being targeted, a good test statistic is the invariant mass of the di-lepton pair, which provides good separation from the background, consisting of all other SM events where no  $J/\Psi$  are produced. In the case of a measurement with low  $\hat{m}_{ll}$ , the most likely hypothesis is  $H_0$ , as seen from the distribution  $f(m_{ll}|H_0)$ . In the case of a second measurement,  $\hat{m}'_{ll}$ , close to the mass of the  $J/\Psi$  meson, the hypothesis  $H_1$  is more likely than  $H_0$ . Since these statements are made in probabilistic terms, as the *pdfs* for each hypothesis are involved, a proper statistical treatment is needed to ensure the correct extrapolation to the physics models under examination.

### 8.3.2 Likelihood ratio discriminant

Intuitively, if the hypothesis testing procedure aims at quantifying the compatibility between the observed data and both  $H_0$  and  $H_1$ , the likelihood functions for each of the hypotheses can

---

<sup>2</sup>This is precisely the purpose of the MVA techniques developed in chapter 7: to combine the information of different input features for two different hypotheses, signal and background, and provide a new random variable that is distributed differently under each of them.

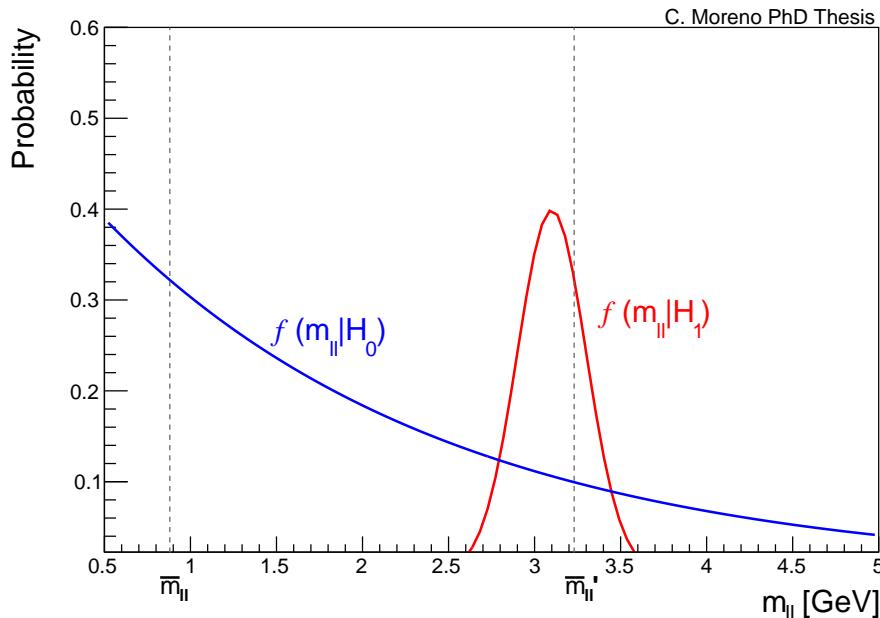


Figure 8.1: Illustration of a practical example of a test statistic built to discriminate signal and background. In this case, the test statistic is the invariant mass of di- $\mu$  pair used to identify  $J/\Psi$  mesons from the rest of processes considered as background. This example is used since it also portrays an actual use case at the trigger level to identify events with  $J/\Psi$  mesons present. For simplicity, the *pdf* of  $m_{ll}$  under the background hypothesis is assumed to follow a pure exponential curve, while for the signal it is assumed to follow a narrow Gaussian distribution with average equals to the  $J/\Psi$  meson mass. Two measurements are shown,  $\bar{m}_{ll}$  and  $\bar{m}'_{ll}$ , with the first one being more compatible with the background-only hypothesis and the second one more compatible with the signal.

be used to build a test statistic. In particular, the discriminant built as

$$\lambda(\vec{x}) = \frac{L(\vec{x}|H_1)}{L(\vec{x}|H_0)}, \quad (8.12)$$

offers in general good discrimination between the two hypothesis. In fact, according to the Neyman-Pearson lemma [156], the test statistic constructed in this way is the most optimal possibility in any problem, since for any given background misidentification rate,  $\alpha$ , it provides the largest possible signal efficiency. This test, however, can only be accurately constructed when the *pdfs* describing both hypotheses are well known, and in most application cases it is not entirely possible.

One particularly interesting case arises when the random variables  $\vec{x} = (x_1, \dots, x_n)$  are independent from each other. In these cases, the *pdfs* for both hypotheses can be factorized as the product of the *pdf* of each of the variable  $x_i$ , and so the likelihood ratio is

$$\lambda(\vec{x}) = \frac{L(\vec{x}|H_1)}{L(\vec{x}|H_0)} = \frac{\prod_{i=1}^n f_i(x_i|H_1)}{\prod_{i=1}^n f_i(x_i|H_0)}, \quad (8.13)$$

where  $f_i(x_i|H_j)$  is the *pdf* of variable  $x_i$  under hypothesis  $j = 0, 1$ . Even when the *pdfs* cannot be factorized exactly, this test statistic, called *projective likelihood ratio*, can be a powerful

discriminant between signal and background, and it is widely used due to simplicity<sup>3</sup> in spite of its worse performance compared to the original likelihood ratio.

If the two hypotheses under study are nested, i.e. the null hypothesis can be considered as a particular case of the alternative hypothesis, a powerful discriminant can be constructed with the likelihood functions. This is the case for this analysis, where the alternative hypothesis can be seen as that where the signal cross section can have any value greater than or equal to zero, while the null hypothesis is that where the signal cross section is exactly zero. It is useful to define the *signal strength* as the ratio between the actual cross section for the signal process and the cross section predicted theoretically,  $\mu = \sigma/\sigma_{\text{th}}$ , in terms of which the null hypothesis can be written as that with  $\mu = \mu_0 = 0$ , while the alternative hypothesis is that in which the signal strength can take any value,  $\mu = \hat{\mu}$ . The likelihood functions can be written then as  $L(\vec{x}|\mu_0, \vec{\theta})$ , for the background-only hypothesis, and  $L(\vec{x}|\mu, \vec{\theta})$ , for the background plus signal hypothesis. A widely used test statistic can be constructed in this case as

$$\lambda_p(\mu_0) = \frac{L(\vec{x}|\mu_0, \hat{\vec{\theta}}(\mu_0))}{L(\vec{x}|\hat{\mu}, \hat{\vec{\theta}})}, \quad (8.14)$$

where, as usual,  $\hat{\mu}$  and  $\hat{\vec{\theta}}$  are the values estimated from data, and therefore the ones that maximize the likelihood function in the case of the alternative hypothesis, while  $\hat{\vec{\theta}}$  are the values that maximize the likelihood function under the assumption that  $\mu = \mu_0$ . These are obtained by performing two separate fits to data, where the values in the numerator are estimated fitting the *pdf* of the background-only hypothesis, while the values in the denominator are obtained by fitting the *pdf* of the background plus signal hypothesis. The test statistic built in this way is known as *profile likelihood ratio* and is frequently used in data analyses at the LHC to determine exclusion limits in case of no discovery.

### 8.3.3 Accepting the alternative hypothesis

The final purpose of the statistical analysis of the collected data in a search for new physics is to determine whether the observation is compatible with the background-only hypothesis or if an alternative theory can be proven to describe data more accurately, the latter case leading to a discovery of the new physics model. In order to claim such thing, the observed data needs to be sufficiently inconsistent with the null hypothesis such that it is highly unlikely that the observation is due to a statistical fluctuation. The compatibility of each hypothesis with the observation can be evaluated through a test statistic as discussed above.

In collider physics a test statistic built from the profile likelihood ratio is widely used to test the compatibility of the measurement with the background hypothesis in the context of new physics discovery. In the case of an excess over the background prediction, the test statistic is defined as

$$q_0 = \begin{cases} -2 \ln \lambda_p(0) & \hat{\mu} \geq 0, \\ 0 & \hat{\mu} < 0. \end{cases} \quad (8.15)$$

Assigning the value 0 to the test statistic prevents the exclusion of negative signal cross sections in the case of an observed deficit over the background prediction.

For the two hypotheses  $H_0$  and  $H_1$ , the test statistic is expected to have different *pdfs*, with some separation. In figure 8.2a a test statistic  $t$  is shown whose *pdf* prioritizes high values of  $t$

<sup>3</sup>In general, the marginal *pdfs* for each of the random variables  $x_i$  can be obtained from MC generated experiments and they are in practice easy to compute.

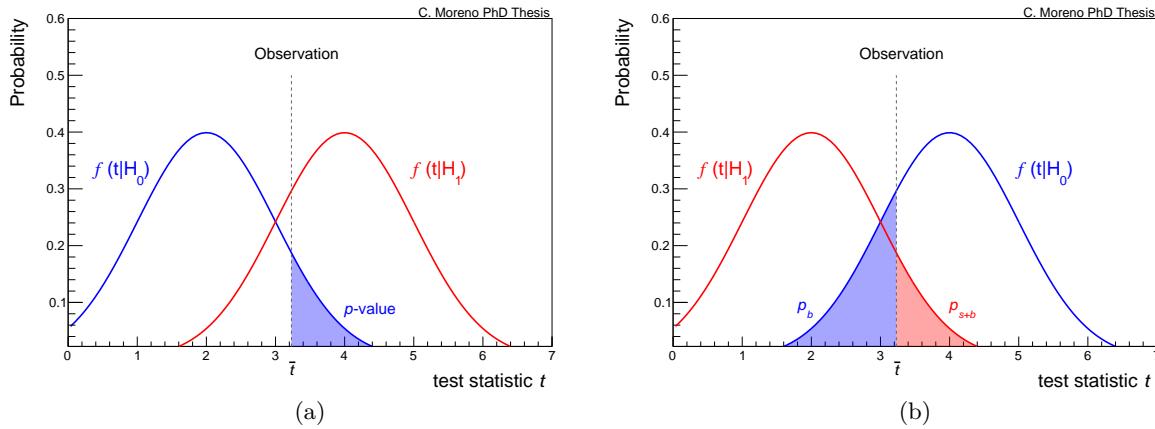


Figure 8.2: (a) Illustration of building a  $p$ -value based on an observation  $\bar{t}$  of the given test statistic. The  $p$ -value corresponds to the probability of measuring any possible value of  $t$  between the measured value  $\bar{t}$  in the direction favoured by the signal hypothesis. (b) Illustration of the process of building the  $p_b$  and  $p_{s+b}$  quantities in order to compute the CL<sub>s</sub> as discussed in the text.

under the alternative hypothesis, and low values of  $t$  for the null hypothesis. In this case, with good separation between both, it is clear that an observation with high  $\bar{t}$  is more compatible with the signal model than the background-only hypothesis. However, given the probabilistic nature of the problem, it is also possible that the observation is just a statistical fluctuation of the background, in which case one needs to be careful about claiming the discovery of new physics. In this sense, the concept of  $p$ -value provides the means to understand the compatibility of the measurement with the background-only hypothesis in probabilistic terms. In the case under discussion, represented in figure 8.2a, where high value of the test statistic  $t$  are more compatible with the signal hypothesis, the  $p$ -value for a given observation  $\bar{t}$  is defined as the probability  $p = P(t \geq \bar{t})$ . Hence, the smaller the  $p$ -value for a given observation, the less compatible it is with the null hypothesis. In the example given in figure 8.2a the  $p$ -value is defined looking at the tail of the  $pdf$  for the test statistic under background-only hypothesis, since the signal favours large values of  $t$ . It is also possible that the test statistic is constructed in a way that the signal is more likely to show low values of  $t$  and the background favours high values of  $t$ , in which case the criterion to define the  $p$ -value for an observation needs to be inverted as well. In general, the idea is to find the probability of the observation to be compatible with the null hypothesis in the signal-like region.

A key concept related to the  $p$ -value that is widely used to present the results of a search for new physics is the *significance level*. It is defined as the distance from the average, measured in standard deviations  $Z\sigma$ , at which a Gaussian random variable needs to be observed to provide a  $p$ -value equal to the original experiment. In case that the test statistic for the background-only hypothesis is Gaussian distributed, the significance level  $Z$  of an observation corresponds to the ratio of the measured value  $\hat{t}$  and the standard deviation  $\sigma$  of the  $pdf$ . However this is typically not the case. The test statistic usually has a complex  $pdf$  under both signal and background hypotheses, in which case the significance level needs to be found as the value  $Z$  that, for a given  $p$ -value, fulfils the following equation

$$p = \int_Z^\infty \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = 1 - \Phi(Z) = \frac{1}{2} \left[ 1 - \text{erf} \left( \frac{Z}{\sqrt{2}} \right) \right], \quad (8.16)$$

where the integrand in the first term is the *pdf* of a normal distribution,  $\Phi(Z)$  is the cumulative of the normal distribution evaluated at the value  $Z$  and  $\text{erf}$  is the error function. Under this definition, the significance level is universal for any experiment since, for a given  $p$ -value, it does not rely on the particular *pdf* of the null hypothesis.

The usual convention in particle physics is to claim there is evidence for the alternative hypothesis when the observations result in a significance level of at least  $3\sigma$ , corresponding to a  $p$ -value of  $1.35 \times 10^{-3}$ , while discovery of new physics is claimed when the significance of the observation is  $5\sigma$ , equivalent to a  $p$ -value of  $2.87 \times 10^{-7}$ . The low probability  $p$  that needs to be reached in order to reject the  $H_0$  hypothesis and accept  $H_1$  reflects the level of incompatibility needed between data and the background only.

### 8.3.4 Rejecting the alternative hypothesis

In case the observation matches, to a certain level, the predictions from the background-only hypothesis,  $H_1$  is examined to understand the compatibility of the measurement with an underfluctuation of the signal process, with the main objective of rejecting it in favour of the null hypothesis. In this sense, the two hypotheses are inverted with respect to the computation of the  $p$ -value, which is now computed for the *pdf* of the test statistic under  $H_1$ . Differently to the case of discovery of new physics, the  $p$ -value required to reject the signal hypothesis is usually set to  $p < 0.05$ , which corresponds to confidence level of 95%.

In a search for new physics, the usual approach is to perform a scan over possible values of the signal strength  $\mu$  and perform the test against the background-only hypothesis based on the observation. If the data matches the prediction of the background, this scan allows to find all values of the signal strength,  $\mu$ , that are not compatible with the observation at a CL of 95%. These values of the signal cross section are said to be *excluded* at the defined CL. The lowest excluded value,  $\mu_{\text{up}}$ , is referred to as *upper limit*, and represents the limit between the parameter region that is not compatible with the observations and the region where the signal hypothesis cannot be rejected against the background-only hypothesis (which does not imply that it can be accepted either). In the region  $\mu < \mu_{\text{up}}$  the observation is compatible with both the null and the alternative hypotheses, and in that case the designed experiment cannot offer enough separation between both theories, or *sensitivity*, and no further statement can be made about the signal process in that region.

A common test statistic for exclusion used in high-energy physics is the so-called *one-sided profile likelihood ratio*, defined as

$$q_\mu = \begin{cases} -2 \ln \lambda_p(\mu) & \hat{\mu} \leq \mu, \\ 0 & \hat{\mu} > \mu. \end{cases} \quad (8.17)$$

The computation of the upper limits differs between the frequentist and Bayesian approaches, and in general each of them will lead to different results. In the case of Bayesian upper limits, they come with inherent subjectivity that is introduced through the prior uncertainties defined for the parameters in the fit. The frequentist limits, on the other hand, can lead to unphysical results, such as excluding negative values of the signal strength  $\mu$  in the case of good agreement between data and background prediction in regions with yields close to zero. The Feldman-Cousins unified approach was introduced to provide a meaningful computation of the upper limits in a search that is mostly compatible with both approaches, although it can also introduce counter-intuitive effects [157]. The topic of upper limit computation still shows discrepancies in the community, without a one-for-all solution useful for all application cases.

In the case of searches for new phenomena in the context of particle physics, the most widely used approach is the so-called  $\text{CL}_s$  method [158]. It aims at solving the problems arising in the frequentist approach by introducing a correction in the  $p$ -value that avoids excluding unphysical regions of the parameter space. This approach is commonly followed using the likelihood ratio of equation 8.12 or the profile likelihood ratio of equation 8.14 as test statistics, although it applies to any test that can be constructed with the observed variables from an experiment. It consists in defining two  $p$ -values, the usual one for the background-only hypothesis,  $p_b$ , as defined above, and a second one for the signal plus background hypothesis,  $p_{s+b}$ . An example is shown in figure 8.2b. If the test statistic favours lower values for the signal than the background, then for a given observation  $\bar{t}$  these  $p$ -values are defined as

$$p_b = P(t \leq \hat{t}|H_0); \quad p_{s+b} = P(t \geq \hat{t}|H_1). \quad (8.18)$$

With these two probabilities, the  $\text{CL}_s$  can be built as

$$\text{CL}_s = \frac{p_{s+b}}{1 - p_b}, \quad (8.19)$$

and is used to define the confidence level when estimating upper limits. A signal hypothesis is excluded at 95% CL, for the measured  $\bar{t}$ , if the corresponding  $\text{CL}_s$  is lower than 0.05. This way, an upper limit corresponding to a 95% CL is obtained by requiring that the measurement of the test statistic  $\hat{t}$  has  $\text{CL}_s \leq 0.05$ . The upper limits obtained in this way are usually more conservative than those obtained using the pure frequentist approach, but it comes free of counter-intuitive effects such as those that appear in the Feldman-Cousins approach, which is the reason why it is used widely in the particle physics community. In the case of the analysis presented in this work, the upper limits are estimated using the  $\text{CL}_s$  method, with the profile likelihood ratio as test statistic.

It is often the case that the *pdf* for the test statistic under both hypotheses is not known, and therefore computing  $p_{s+b}$  and  $p_b$  is not straightforward. Two approximations are commonly used to determine both of them in order to compute the upper limits for a given observation. On the one hand, they can be estimated using pseudo-random experiments by means of MC simulation, where the *pdfs* are sampled by iteratively estimating the  $p$ -values when the expected observation is replaced by Poisson fluctuations. This method can result in highly accurate results at large number of experiments, but it can be computationally expensive. A second approach useful in the case of the profile likelihood ratio as test statistic relies on its asymptotic behaviour for large number of experiments [159]. In this case, the negative logarithm of the test statistic is assumed to be parabolic around its minimum (i.e. the estimated values of the parameters) up to corrections that depend on the inverse of the size of the sample. This allows for an analytical expression for the test statistic, which in turns makes the computation of the  $\text{CL}_s$  straightforward.

## 8.4 Systematic uncertainties

Two sources of systematic uncertainties are considered in this analysis. On the one hand, *experimental uncertainties* associated to the detector response and object reconstruction are introduced to account for the unknown calibration in the extreme kinematic region targeted in this search. Additionally, *theoretical uncertainties* on the modelling of the background and signal processes are also included in the description, which aim at providing a robust post-fit background estimate in the analysis.

The systematic uncertainties are included in the fit to data in the form of NPs, aiming at determining the value of the different unknown parameters related to detector response, object calibration and process modelling. In practice, in order to accurately estimate them, an independent measurement per NP would be needed in the fit. The approach followed is to use MC simulation to define an uncertainty interval around their nominal value for each them, where the real value is expected to be found with a confidence level of 68.27% (i.e.  $1\sigma$ ). The final goal of the fit is to constrain those initial confidence intervals based on the observations. In most cases, the analysis is not very sensitive to the given parameters, meaning that the size of their uncertainties is small compared to the statistical precision of the data sample used. The post-fit uncertainty interval in such cases is very similar to the pre-fit one, and the parameter is said to be *unconstrained*. This is expected for most NPs covering experimental uncertainties, since their pre-fit uncertainty intervals are derived in dedicated analyses with larger sensitivity to the related effect than this search. For example, dedicated analyses exist to measure the jet energy scale and its uncertainties with large coverage of the parameter phase space that are expected to be more sensitive to variations in the targeted quantity.

In general, different prescriptions are centrally provided for the experimental uncertainties to be included in the fit, which determines the number of NPs associated to a given aspect of the calibration of the objects used in the analysis. In the following, a summary of the NPs introduced in the fit is provided for both experimental and theory-related systematic uncertainties.

#### 8.4.1 Experimental systematic uncertainties

**Jet-related uncertainties:** As discussed in chapter 6, the calibration of the jet energy scale and its resolution come with associated uncertainties. In the case of this analysis, the signal regions combined in a multi-bin fit for a given mass hypothesis can lead to shape-dependent effects that need to be accounted for. For this reason, a complex scheme of NPs is implemented. For JES-related uncertainties, a total of 31 NPs are added in the fit, 22 for the global JES calibration, seven for the  $\eta$ -intercalibration non-closure, and two for the jet flavour response. The impact of the JER, on the other hand, is addressed by using  $\pm 1\sigma$  variations derived from data vs MC differences, a noise term for evaluation from random cones in zero bias data, and in-situ measurement of the di-jet  $p_T$  balance asymmetry, for a total of 13 NPs implemented in the fit. Additionally, a set of 18 NPs is introduced for the calibration of  $R = 0.4$  jet mass. And another NP is included to account for JVT-related uncertainties. All uncertainties are propagated to the reclustered jets used in the computation of the variable  $M_J^{\sum,4}$ .

**Flavour tagging uncertainties:** The implementation of uncertainties on the  $b$ -tagging efficiency relies on an eigenvector decomposition approach, where each source of uncertainty is varied by  $\pm 1\sigma$  and the effect on the scale factor for each bin used in the calibration is considered. The total correlation matrix is built from the contribution of each source of uncertainty, and it is used to find the “directions” in which the variations are independent from each other (i.e. the eigenvectors of the covariance matrix). The sizes of the variations corresponding to the  $\pm 1\sigma$  effects are then given by the square root of the corresponding eigenvalues of the correlation matrix. This approach results in nine NPs for the  $b$ -tagging efficiency, together with eight NPs for  $c$ - and light-jet misidentification rate. Two extra NPs are included for the extrapolation of  $b$ -tagging to high- $p_T$  jets.

**$E_T^{\text{miss}}$  uncertainties:** Variations in the  $E_T^{\text{miss}}$  soft-terms are performed according to the measured uncertainty in the energy scale and resolution. A total of 3 NPs, one for the energy scale

and 2 for the energy resolution, are included in the fit. In addition, one extra NP is included to account for the uncertainties measured in the calibration of the  $E_T^{\text{miss}}$ -trigger.

**Luminosity and pileup:** The measured integrated luminosity has an uncertainty of 1.7% [160], which is included in the fit with a single NP. The simulated samples are reweighted to reproduce the pileup behaviour of the different data taking periods considered. An extra NP to account for uncertainties in this reweighting procedure.

**Others:** The uncertainties associated to other reconstructed objects, such as electrons or muons, were found to be negligible and therefore not included in the description of this analysis.

The dominant experimental uncertainties on the background in the signal regions of this analysis are typically the ones related to jet reconstruction. The uncertainties associated to JES and JER are the largest ones, reaching up to 30% on the background prediction in the SRs per source of uncertainty.

#### 8.4.2 Modelling systematic uncertainties

The main sources of modelling uncertainties for all background processes estimated through MC simulation are discussed, together with the procedure followed to estimate them, in section 7.6.6. For  $t\bar{t}$  and single-top backgrounds, a NP is introduced for each to account for the uncertainty related to the choice of MC generator. In addition, an extra NP is introduced for single-top  $Wt$ -channel to account for the uncertainty due to the scheme chosen to handle the interference with  $t\bar{t}$ . In each case, the uncertainty interval is estimated from the difference between the nominal sample with alternative ones generated with the corresponding variations. The difference in the expected yield in each control, validation and signal region between nominal and modified samples is taken as the effect of the  $1\sigma$  variation, and the interval is found by symmetrizing the uncertainty around the nominal value. The size of these uncertainties in all regions for three signal mass hypotheses, representative of the low-, intermediate- and high-mass regimes, are shown in figure 8.3. The NPs introduced in the fit to account for these uncertainties are correlated for all regions, reducing the impact on the sensitivity of the analysis after the fit.

In the case of  $t\bar{t}$ , each component ( $t\bar{t}$  + light-jets,  $t\bar{t} + \geq 1b$  and  $t\bar{t} + \geq 1c$ ) enters the fit as an independent sample. Two NPs are included to account for uncertainties in the normalization of the  $t\bar{t} + \geq 1b$  and  $t\bar{t} + \geq 1c$  components. For each of these NPs, a pre-fit uncertainty of 50% is assumed, as it is consistent with the normalizations observed in other ATLAS analyses [147]. In addition, a NP is introduced with a pre-fit 50% uncertainty on the normalization of the single-top-quark sample.

Uncertainties related to the emissions of ISR/FSR are also included in the description for  $t\bar{t}$ , single-top and  $Z/W+jets$  backgrounds. For  $t\bar{t}$  and single-top, two NPs are implemented: one for variations on the factorization scale, and one for variations on the renormalization scale. Similarly, for  $Z/W+jets$  processes, these two NPs are also implemented, with the addition of an extra NP to account for simultaneous variations on both the factorization and renormalization scales. In all cases, the uncertainty on the yield of each affected background is estimated by increasing and reducing the concerned parameters in the MC generator by a factor of 2. Additionally, for  $t\bar{t}$  and single-top backgrounds, an extra uncertainty on the ISR radiation is derived by changing the tunes in the parton shower generator. The NPs included in the fit for these uncertainties are also correlated across all regions considered.

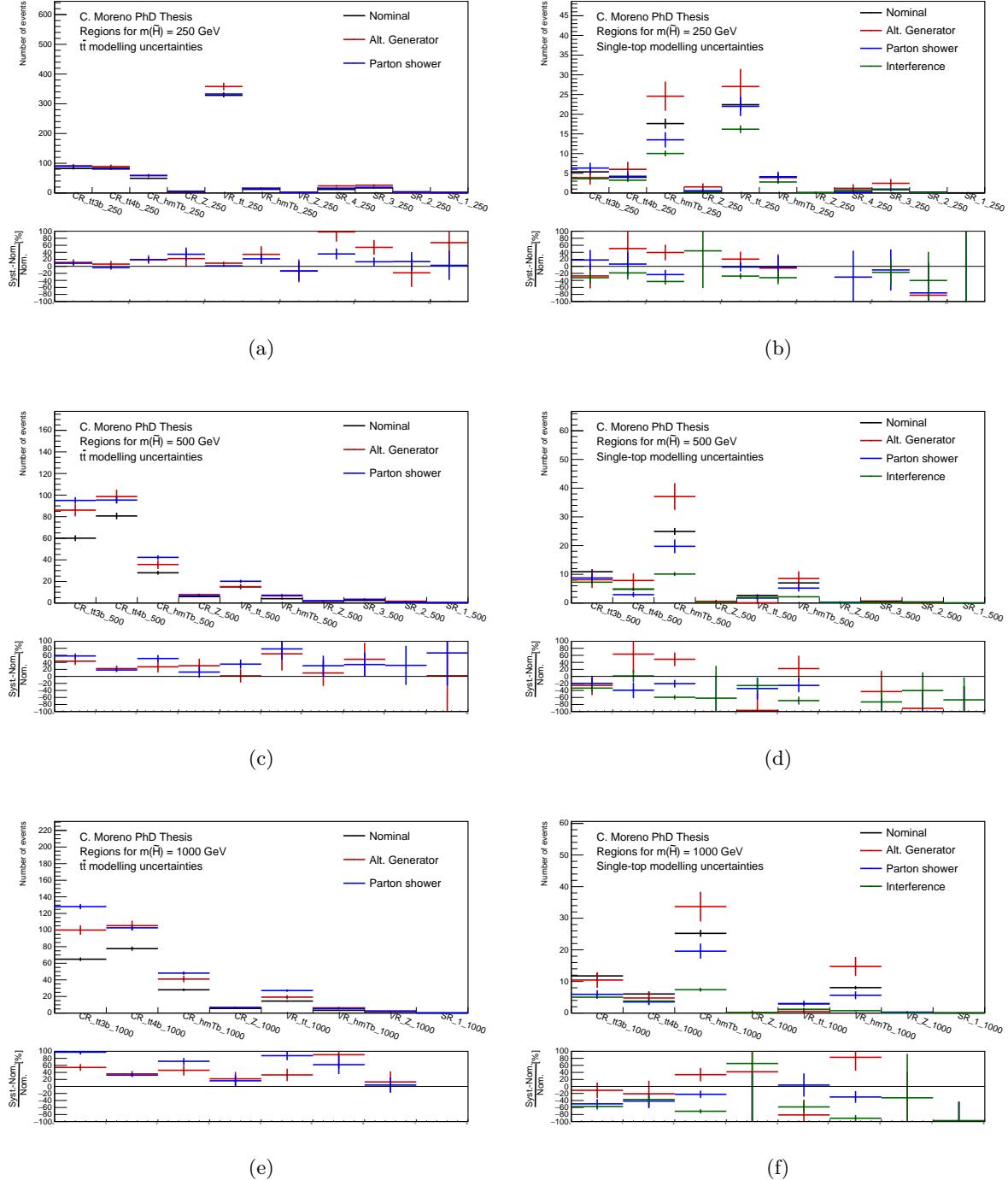


Figure 8.3: Size of the generator, parton shower and interference uncertainties for  $t\bar{t}$  (left) and single-top (right) backgrounds in the regions defined for the mass hypotheses (a,b)  $m(\tilde{H}) = 250 \text{ GeV}$ , (c,d)  $m(\tilde{H}) = 500 \text{ GeV}$  and (e,f)  $m(\tilde{H}) = 1000 \text{ GeV}$ .

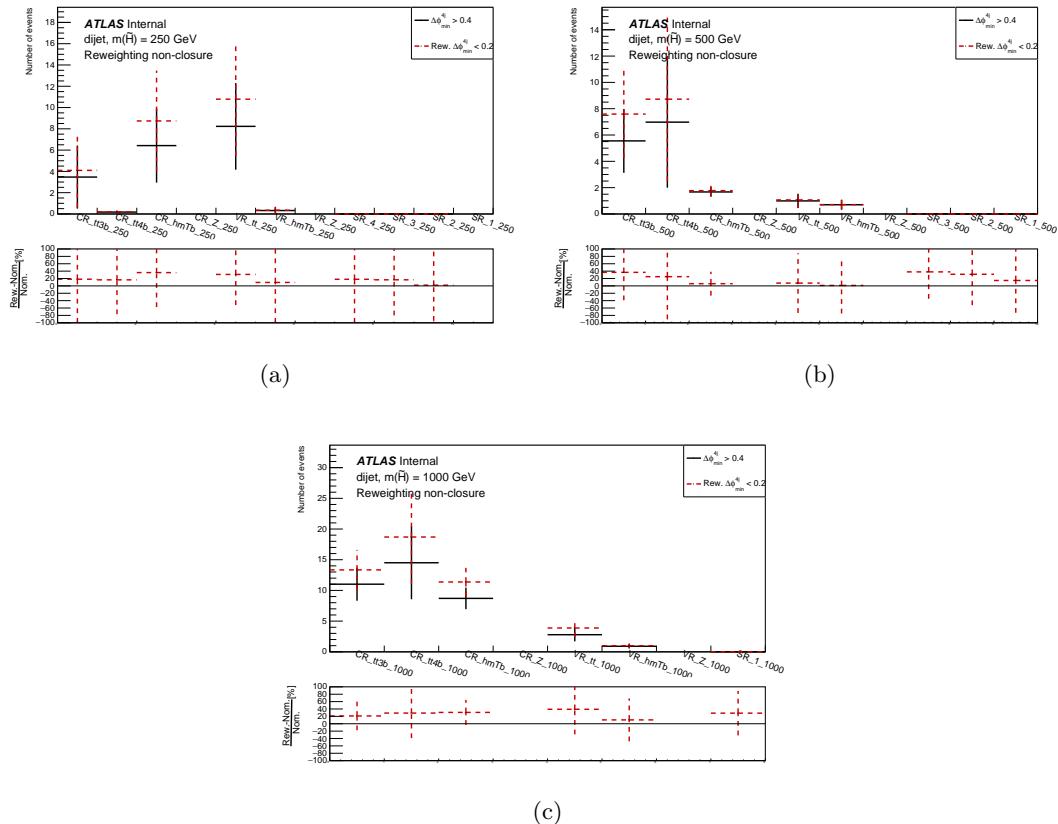


Figure 8.4: NN-based reweighting non-closure in CR/VR/SR for mass points (a)  $m(\tilde{H}) = 250$  GeV, (b)  $m(\tilde{H}) = 500$  GeV and (c)  $m(\tilde{H}) = 1000$  GeV.

As discussed in section 7.6.6,  $t\bar{t} + X$  and diboson processes have minor contributions in the regions of this analysis and therefore no dedicated modelling studies are performed. In order to have robust description of the background in the signal regions, a conservative 50% uncertainty in the normalization of these two backgrounds is included as uncorrelated NPs for each bin in the fit.

The modelling of QCD multijet processes, since it is estimated by means of data-driven methods, is not affected by MC-related uncertainties. In this case, the main source of uncertainty arises from the accuracy of the reweighting strategy, based on a NN. This uncertainty is evaluated directly on di-jet MC samples, by comparing the nominal expected yield in the control, validation and signal regions with the prediction of a low- $\Delta\phi_{\min}^{4j}$  slice after being reweighted with the NN. In order to increase the available statistics for this comparison, an alternative strategy for flavour tagging based on truth information of the event is used, called *tag rate function* or *truth tagging*, instead of direct  $b$ -tagging. The size of the expected uncertainty in the regions for three signal mass points representative of the low, intermediate and high mass regimes is shown in figure 8.4. These uncertainties are typically small, the largest ones are about 40%. In addition, the data agreement in a QCD-enhanced region is examined, which is collected in figure 7.8, and shows slight non-closure effects of about 10% in extreme kinematic regions. In order to cover these two effects together, a 50% uncertainty on the QCD background estimation is applied to each bin included in the fit, which completely covers any mismodelling observed.

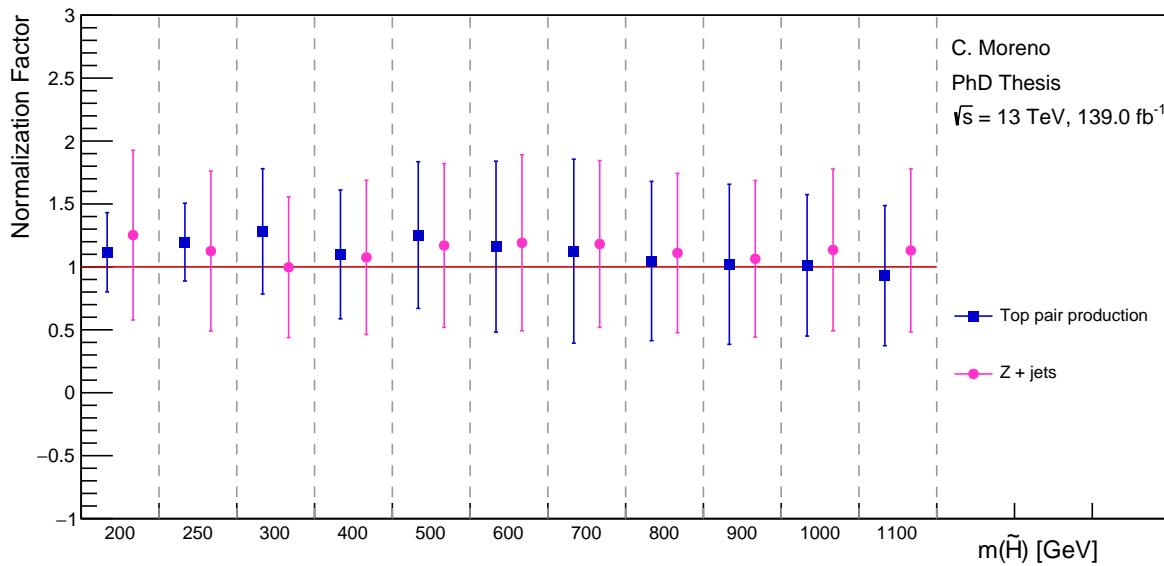


Figure 8.5: Fitted normalization factors for  $t\bar{t}$  and  $Z+\text{jets}$  backgrounds for all mass hypotheses considered in the analysis.

## 8.5 Background-only fit

As discussed in section 7.9.3, a robust background estimate in the signal regions is achieved by defining process-specific control regions to estimate their normalization. The extrapolation from control to signal regions is cross-checked in dedicated validation regions. While the normalization for all sub-leading background processes is estimated directly from the MC samples, for both  $t\bar{t}$  and  $Z+\text{jets}$  events a normalization factor,  $\mu = \hat{\sigma}/\sigma_{\text{theo}}$ , is as free-floating in the fit to be estimated from data. The two NPs included for the normalization of the  $t\bar{t} + \text{HF}$  components further helps improve the agreement of data and the prediction. In addition the NP included for the normalization of single-top events helps correct the mismodelling observed in the tail of the  $m_{T,\min}^{b-\text{jets}}$  distribution, which becomes relevant for high signal masses.

In order to understand the performance of the background estimate, a *background-only fit* is carried out using exclusively the control regions and without accounting for the presence of any signal. In this case, no hypothesis test is performed as the signal hypothesis is disregarded. An independent fit to data in the regions defined for each signal mass point is carried out. The approach to determine the values of all parameters relies on the methods discussed in section 8.2, in particular the maximum likelihood method as implemented in the HistFitter framework [161] as interface for the RooStat package [153]. In this context, the two normalization factors mentioned, one for  $t\bar{t}$  and one for  $Z+\text{jets}$ , are the parameters of interest, and are fitted together with normalization NPs for the relevant backgrounds, as well as all NPs implemented to account for systematic uncertainties. The measured normalization factors are shown in figure 8.5 for the regions defined for each signal mass hypothesis, which shows that all fitted parameters are very closed to one, as expected from the overall pre-fit agreement observed between data and background prediction at pre-selection level. The estimated values of the NPs introduced in the description are shown for the fit corresponding to the mass hypothesis  $m(\tilde{H}) = 1$  TeV in figure 8.6. In general, the analysis does not have enough sensitivity to constrain the  $t\bar{t}$ -related normalization NPs, however the one for single-top is consistently pulled across all mass points, corresponding to a variation on the normalization of about 30%. The experimental uncertainties

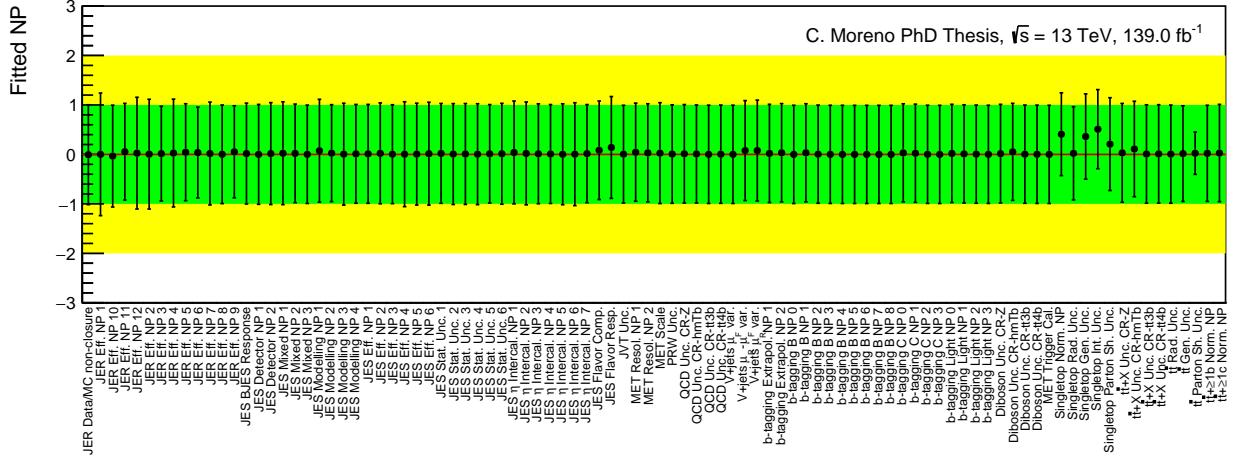


Figure 8.6: Estimated values of the NPs introduced to account for all systematic uncertainties described in section 8.4 for the fit corresponding to higgsino mass hypothesis of 1 TeV.

are in general unconstrained.

The post-fit agreement between data and background estimate in all validation regions for three representative higgsino mass hypotheses,  $m(\tilde{H}) = 250, 500$  and  $1000$  GeV, is shown in figures 8.7–8.9. The significance in each bin is computed as the difference between the observed and expected yields divided by the total uncertainty, and is included in the bottom panel of each figure. As shown, very good agreement is achieved, with “pulls” lower than  $1\sigma$  in all cases, indicating a robust background estimate and a consistent fit approach.

Once a robust background estimate is achieved by determining the normalization of the main processes and cross checking the extrapolation to the validation regions, the signal regions are examined. The corrections derived in the control regions through the background-only fit are also propagated to the signal regions in order to improve the background prediction there. The expected yields in the SRs are compared to the observation in order to asses the presence of excess data over the prediction, which can be observed in figures 8.7–8.9 for three representative mass points. As observed from the significance of the measurement in these signal regions, there are no major deviation from the expected yields.

The numeric values for the background expected yields in every signal region in the analysis, together with the observed data events, are collected in tables 8.1–8.3. In these tables, the pre-fit prediction for each background process is also included in the last rows. The largest differences between data and the background prediction appear in the SR-1 for the very high signal mass hypothesis,  $m(\tilde{H}) \geq 800$  GeV, which target the most extreme kinematic region to which this analysis provides sensitivity. In all of them, around 0.5 background events are expected while 3 events are observed (2 in the case of SR-1-1100). These events are identified to be the same across all regions thanks to the unique run and event numbers assigned during data taking.

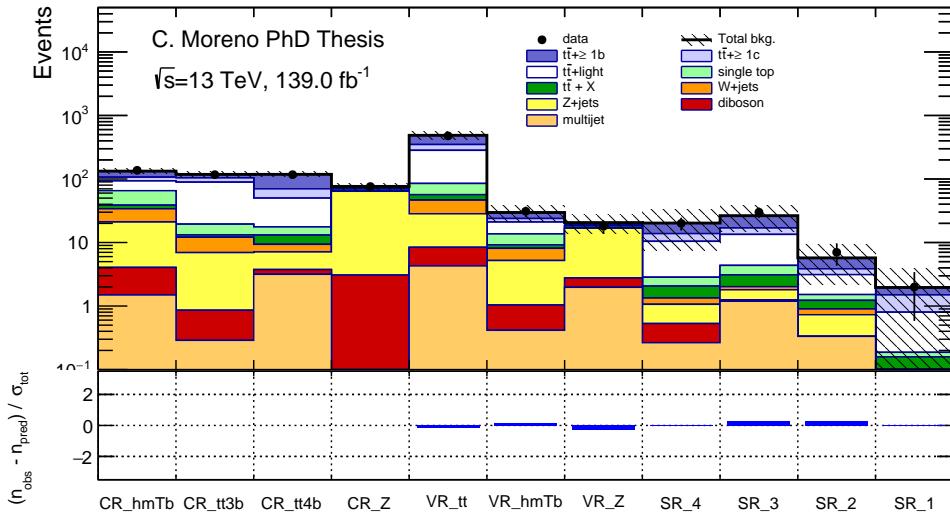


Figure 8.7: Comparison between data and post-fit background prediction in all regions dedicated to the higgsino mass hypothesis  $m(\tilde{H}) = 250$  GeV, with the significance in each bin shown in the bottom panel.

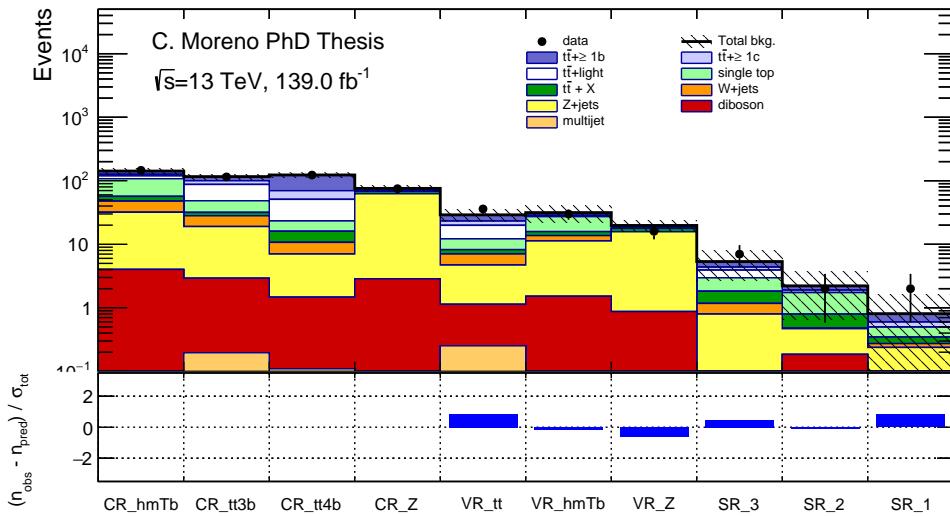


Figure 8.8: Comparison between data and post-fit background prediction in all regions dedicated to the higgsino mass hypothesis  $m(\tilde{H}) = 500$  GeV, with the significance in each bin shown in the bottom pannel.

table.results.yields.channel	SR_1_200	SR_1_250	SR_1_300	SR_1_400	SR_1_500	SR_1_600	SR_1_700	SR_1_800	SR_1_900	SR_1_1000	SR_1_1100
Observed events	2	1	3	3	2	2	2	3	3	3	2
Fitted bkg events	1.42 <sup>+2.73</sup> <sub>-1.42</sub>	1.96 <sup>+2.00</sup> <sub>-1.96</sub>	1.80 <sup>+1.24</sup> <sub>-1.24</sub>	1.08 <sup>+1.03</sup> <sub>-1.03</sub>	0.81 <sup>+0.80</sup> <sub>-0.80</sub>	0.87 <sup>+0.85</sup> <sub>-0.85</sub>	0.78 <sup>+1.03</sup> <sub>-0.78</sub>	0.62 <sup>+0.83</sup> <sub>-0.58</sub>	0.58 <sup>+0.83</sup> <sub>-0.53</sub>	0.61 <sup>+0.53</sup> <sub>-0.53</sub>	0.56 <sup>+0.45</sup> <sub>-0.45</sub>
Fitted ttbb events	0.44 <sup>+0.44</sup> <sub>-0.42</sub>	0.46 <sup>+0.55</sup> <sub>-0.46</sub>	0.55 <sup>+0.65</sup> <sub>-0.55</sub>	0.34 <sup>+0.49</sup> <sub>-0.49</sub>	0.21 <sup>+0.20</sup> <sub>-0.20</sub>	0.12 <sup>+0.21</sup> <sub>-0.12</sub>	0.17 <sup>+0.17</sup> <sub>-0.17</sub>	0.05 <sup>+0.05</sup> <sub>-0.05</sub>	0.04 <sup>+0.04</sup> <sub>-0.04</sub>	0.02 <sup>+0.24</sup> <sub>-0.02</sub>	0.02 <sup>+0.22</sup> <sub>-0.02</sub>
Fitted ttcc events	0.27 <sup>+0.27</sup> <sub>-0.27</sub>	0.70 <sup>+1.19</sup> <sub>-0.70</sub>	0.52 <sup>+0.69</sup> <sub>-0.52</sub>	0.17 <sup>+0.45</sup> <sub>-0.17</sub>	0.10 <sup>+0.64</sup> <sub>-0.10</sub>	0.11 <sup>+0.69</sup> <sub>-0.11</sub>	0.09 <sup>+0.53</sup> <sub>-0.09</sub>	0.03 <sup>+0.61</sup> <sub>-0.03</sub>	0.04 <sup>+0.45</sup> <sub>-0.03</sub>	0.02 <sup>+0.18</sup> <sub>-0.02</sub>	0.02 <sup>+0.05</sup> <sub>-0.02</sub>
Fitted tflight events	0.48 <sup>+1.11</sup> <sub>-0.48</sub>	0.61 <sup>+1.61</sup> <sub>-0.61</sub>	0.43 <sup>+0.43</sup> <sub>-0.43</sub>	0.24 <sup>+0.51</sup> <sub>-0.24</sub>	0.09 <sup>+0.09</sup> <sub>-0.09</sub>	0.00 <sup>+0.00</sup> <sub>-0.00</sub>	0.00 <sup>+0.00</sup> <sub>-0.00</sub>	0.00 <sup>+0.00</sup> <sub>-0.00</sub>	0.00 <sup>+0.00</sup> <sub>-0.00</sub>	0.01 <sup>+0.04</sup> <sub>-0.01</sub>	0.02 <sup>+0.03</sup> <sub>-0.02</sub>
Fitted singletop events	0.06 <sup>+0.06</sup> <sub>-0.06</sub>	0.03 <sup>+0.03</sup> <sub>-0.03</sub>	0.07 <sup>+0.07</sup> <sub>-0.07</sub>	0.01 <sup>+0.06</sup> <sub>-0.01</sub>	0.15 <sup>+0.15</sup> <sub>-0.15</sub>	0.23 <sup>+0.37</sup> <sub>-0.27</sub>	0.11 <sup>+0.11</sup> <sub>-0.11</sub>	0.06 <sup>+0.32</sup> <sub>-0.06</sub>	0.12 <sup>+0.45</sup> <sub>-0.12</sub>	0.12 <sup>+0.26</sup> <sub>-0.12</sub>	0.09 <sup>+0.20</sup> <sub>-0.09</sub>
Fitted topEW events	0.08 <sup>+0.08</sup> <sub>-0.08</sub>	0.12 <sup>+0.26</sup> <sub>-0.12</sub>	0.06 <sup>+0.13</sup> <sub>-0.06</sub>	0.13 <sup>+0.09</sup> <sub>-0.13</sub>	0.07 <sup>+0.07</sup> <sub>-0.07</sub>	0.07 <sup>+0.23</sup> <sub>-0.23</sub>	0.11 <sup>+0.13</sup> <sub>-0.13</sub>	0.15 <sup>+0.15</sup> <sub>-0.15</sub>	0.15 <sup>+0.18</sup> <sub>-0.15</sub>	0.11 <sup>+0.13</sup> <sub>-0.13</sub>	0.09 <sup>+0.09</sup> <sub>-0.09</sub>
Fitted W_jets events	0.03 <sup>+0.09</sup> <sub>-0.03</sub>	0.04 <sup>+0.72</sup> <sub>-0.09</sub>	0.08 <sup>+0.04</sup> <sub>-0.04</sub>	0.03 <sup>+0.09</sup> <sub>-0.03</sub>	0.04 <sup>+0.07</sup> <sub>-0.09</sub>	0.05 <sup>+0.07</sup> <sub>-0.05</sub>	0.04 <sup>+0.05</sup> <sub>-0.05</sub>	0.04 <sup>+0.05</sup> <sub>-0.05</sub>	0.04 <sup>+0.13</sup> <sub>-0.10</sub>	0.04 <sup>+0.10</sup> <sub>-0.10</sub>	0.04 <sup>+0.09</sup> <sub>-0.09</sub>
Fitted Z_jets events	0.06 <sup>+0.04</sup> <sub>-0.04</sub>	0.00 <sup>+0.00</sup> <sub>-0.00</sub>	0.13 <sup>+0.18</sup> <sub>-0.12</sub>	0.24 <sup>+0.16</sup> <sub>-0.13</sub>	0.31 <sup>+0.30</sup> <sub>-0.30</sub>	0.26 <sup>+0.18</sup> <sub>-0.26</sub>	0.29 <sup>+0.18</sup> <sub>-0.26</sub>	0.20 <sup>+0.26</sup> <sub>-0.20</sub>	0.20 <sup>+0.31</sup> <sub>-0.27</sub>	0.23 <sup>+0.23</sup> <sub>-0.23</sub>	0.23 <sup>+0.23</sup> <sub>-0.23</sub>
Fitted diboson events	0.00 <sup>+0.00</sup> <sub>-0.00</sub>	0.00 <sup>+0.00</sup> <sub>-0.00</sub>	0.02 <sup>+0.08</sup> <sub>-0.02</sub>	0.03 <sup>+0.12</sup> <sub>-0.03</sub>	0.00 <sup>+0.00</sup> <sub>-0.00</sub>	0.07 <sup>+0.06</sup> <sub>-0.06</sub>	0.07 <sup>+0.06</sup> <sub>-0.06</sub>				
Fitted QCD events	0.00 <sup>+0.00</sup> <sub>-0.00</sub>										
MC exp. SM events	1.51	1.62	2.05	1.28	0.62	0.62	0.66	0.61	0.56	0.55	0.50

Table 8.1: Observed and expected (post-fit) number of events in all SR-1 regions. The SM prediction is split in the contributions from the main backgrounds, with the pre-fit total prediction shown in the last row.

table.results.yields.channel	SR_2_200	SR_2_250	SR_2_300	SR_2_400	SR_2_500	SR_2_600	SR_2_700	SR_2_800	SR_2_900	SR_2_1000	SR_2_1100
Observed events	9	7	5	9	2	4	4	4	3	3	3
Fitted bkg events	8.48 <sup>+2.9</sup> <sub>-2.9</sub>	5.73 <sup>+3.58</sup> <sub>-3.58</sub>	5.50 <sup>+4.17</sup> <sub>-4.17</sub>	3.96 <sup>+2.21</sup> <sub>-2.21</sub>	2.21 <sup>+1.56</sup> <sub>-1.56</sub>	2.18 <sup>+2.04</sup> <sub>-2.04</sub>	1.96 <sup>+1.6</sup> <sub>-1.6</sub>	2.86 <sup>+1.83</sup> <sub>-1.83</sub>	2.23 <sup>+1.81</sup> <sub>-1.81</sub>	2.23 <sup>+1.81</sup> <sub>-1.81</sub>	2.23 <sup>+1.81</sup> <sub>-1.81</sub>
Fitted ttbb events	2.73 <sup>+2.94</sup> <sub>-2.73</sub>	1.88 <sup>+1.84</sup> <sub>-1.84</sub>	2.11 <sup>+1.85</sup> <sub>-1.85</sub>	1.14 <sup>+1.13</sup> <sub>-1.13</sub>	0.32 <sup>+0.51</sup> <sub>-0.32</sub>	0.30 <sup>+0.53</sup> <sub>-0.30</sub>	0.25 <sup>+0.51</sup> <sub>-0.30</sub>	0.61 <sup>+0.72</sup> <sub>-0.61</sub>	0.61 <sup>+0.72</sup> <sub>-0.61</sub>	0.34 <sup>+0.75</sup> <sub>-0.34</sub>	0.34 <sup>+0.75</sup> <sub>-0.34</sub>
Fitted ttcc events	1.22 <sup>+1.49</sup> <sub>-1.22</sub>	0.73 <sup>+1.31</sup> <sub>-1.31</sub>	0.60 <sup>+1.00</sup> <sub>-1.00</sub>	0.51 <sup>+0.66</sup> <sub>-0.66</sub>	0.16 <sup>+0.39</sup> <sub>-0.16</sub>	0.11 <sup>+0.48</sup> <sub>-0.11</sub>	0.18 <sup>+0.35</sup> <sub>-0.18</sub>	0.22 <sup>+0.42</sup> <sub>-0.22</sub>	0.19 <sup>+0.49</sup> <sub>-0.19</sub>	0.19 <sup>+0.49</sup> <sub>-0.19</sub>	0.19 <sup>+0.49</sup> <sub>-0.19</sub>
Fitted tflight events	3.34 <sup>+3.06</sup> <sub>-3.06</sub>	1.60 <sup>+1.60</sup> <sub>-1.60</sub>	1.24 <sup>+1.24</sup> <sub>-1.24</sub>	0.53 <sup>+0.53</sup> <sub>-0.53</sub>	0.00 <sup>+0.00</sup> <sub>-0.00</sub>	0.13 <sup>+0.13</sup> <sub>-0.13</sub>	0.28 <sup>+0.25</sup> <sub>-0.25</sub>	0.21 <sup>+0.22</sup> <sub>-0.21</sub>	0.05 <sup>+0.05</sup> <sub>-0.05</sub>	0.05 <sup>+0.05</sup> <sub>-0.05</sub>	0.05 <sup>+0.05</sup> <sub>-0.05</sub>
Fitted singletop events	0.07 <sup>+0.40</sup> <sub>-0.07</sub>	0.29 <sup>+0.78</sup> <sub>-0.29</sub>	0.28 <sup>+0.62</sup> <sub>-0.28</sub>	0.31 <sup>+0.97</sup> <sub>-0.31</sub>	0.94 <sup>+1.13</sup> <sub>-0.94</sub>	0.97 <sup>+1.37</sup> <sub>-0.97</sub>	0.51 <sup>+0.86</sup> <sub>-0.51</sub>	0.42 <sup>+0.92</sup> <sub>-0.42</sub>	0.47 <sup>+0.88</sup> <sub>-0.47</sub>	0.47 <sup>+0.88</sup> <sub>-0.47</sub>	0.47 <sup>+0.88</sup> <sub>-0.47</sub>
Fitted topEW events	0.33 <sup>+0.33</sup> <sub>-0.33</sub>	0.33 <sup>+0.33</sup> <sub>-0.33</sub>	0.55 <sup>+0.42</sup> <sub>-0.42</sub>	0.53 <sup>+0.49</sup> <sub>-0.49</sub>	0.32 <sup>+0.30</sup> <sub>-0.30</sub>	0.25 <sup>+0.24</sup> <sub>-0.24</sub>	0.27 <sup>+0.25</sup> <sub>-0.25</sub>	0.38 <sup>+0.34</sup> <sub>-0.33</sub>	0.33 <sup>+0.33</sup> <sub>-0.33</sub>	0.33 <sup>+0.33</sup> <sub>-0.33</sub>	0.33 <sup>+0.33</sup> <sub>-0.33</sub>
Fitted W_jets events	0.09 <sup>+0.47</sup> <sub>-0.09</sub>	0.17 <sup>+0.47</sup> <sub>-0.17</sub>	0.14 <sup>+0.47</sup> <sub>-0.14</sub>	0.24 <sup>+0.66</sup> <sub>-0.24</sub>	0.01 <sup>+0.41</sup> <sub>-0.01</sub>	0.01 <sup>+0.27</sup> <sub>-0.01</sub>	0.03 <sup>+0.13</sup> <sub>-0.03</sub>	0.31 <sup>+0.24</sup> <sub>-0.24</sub>	0.23 <sup>+0.37</sup> <sub>-0.23</sub>	0.23 <sup>+0.37</sup> <sub>-0.23</sub>	0.23 <sup>+0.37</sup> <sub>-0.23</sub>
Fitted Z_jets events	0.08 <sup>+0.36</sup> <sub>-0.08</sub>	0.40 <sup>+0.40</sup> <sub>-0.40</sub>	0.36 <sup>+0.36</sup> <sub>-0.36</sub>	0.45 <sup>+0.45</sup> <sub>-0.45</sub>	0.28 <sup>+0.27</sup> <sub>-0.27</sub>	0.23 <sup>+0.23</sup> <sub>-0.23</sub>	0.33 <sup>+0.24</sup> <sub>-0.24</sub>	0.64 <sup>+0.38</sup> <sub>-0.38</sub>	0.55 <sup>+0.55</sup> <sub>-0.55</sub>	0.55 <sup>+0.55</sup> <sub>-0.55</sub>	0.55 <sup>+0.55</sup> <sub>-0.55</sub>
Fitted diboson events	0.00 <sup>+0.00</sup> <sub>-0.00</sub>	0.00 <sup>+0.00</sup> <sub>-0.00</sub>	0.00 <sup>+0.00</sup> <sub>-0.00</sub>	0.09 <sup>+0.42</sup> <sub>-0.09</sub>	0.18 <sup>+0.28</sup> <sub>-0.18</sub>	0.18 <sup>+0.18</sup> <sub>-0.18</sub>	0.13 <sup>+0.13</sup> <sub>-0.13</sub>	0.08 <sup>+0.49</sup> <sub>-0.08</sub>	0.07 <sup>+0.07</sup> <sub>-0.07</sub>	0.07 <sup>+0.07</sup> <sub>-0.07</sub>	0.07 <sup>+0.07</sup> <sub>-0.07</sub>
Fitted QCD events	0.63 <sup>+0.33</sup> <sub>-0.33</sub>	0.33 <sup>+0.17</sup> <sub>-0.17</sub>	0.20 <sup>+0.11</sup> <sub>-0.11</sub>	0.17 <sup>+0.09</sup> <sub>-0.09</sub>	0.00 <sup>+0.00</sup> <sub>-0.00</sub>						
MC exp. SM events	7.88	4.78	4.82	3.95	1.60	1.83	1.79	2.92	2.16	2.16	2.16

Table 8.2: Observed and expected (post-fit) number of events in all SR-2 regions. The SM prediction is split in the contributions from the main backgrounds, with the pre-fit total prediction shown in the last row.

table.results.yields.channel	SR_3_200	SR_3_250	SR_3_300	SR_3_400	SR_3_500	SR_3_600	SR_3_700	SR_4_250	SR_4_300	SR_4_400	SR_4_500
Observed events	21	30	17	6	7	8	3	20	18	21	21
Fitted bkg events	3.49 <sup>+3.18</sup> <sub>-3.18</sub>	9.52 <sup>+6.34</sup> <sub>-6.34</sub>	3.04 <sup>+2.93</sup> <sub>-2.93</sub>	1.61 <sup>+1.84</sup> <sub>-1.61</sub>	0.98 <sup>+0.82</sup> <sub>-0.82</sub>	1.61 <sup>+1.51</sup> <sub>-1.51</sub>	0.90 <sup>+0.90</sup> <sub>-0.90</sub>	6.44 <sup>+5.68</sup> <sub>-5.68</sub>	6.75 <sup>+5.43</sup> <sub>-5.43</sub>	5.48 <sup>+3.48</sup> <sub>-3.48</sub>	5.48 <sup>+3.48</sup> <sub>-3.48</sub>
Fitted ttbb events	1.27 <sup>+1.59</sup> <sub>-1.27</sub>	3.53 <sup>+3.34</sup> <sub>-3.34</sub>	1.83 <sup>+2.06</sup> <sub>-1.83</sub>	0.52 <sup>+0.52</sup> <sub>-0.52</sub>	0.45 <sup>+0.75</sup> <sub>-0.45</sub>	0.58 <sup>+0.58</sup> <sub>-0.58</sub>	0.47 <sup>+0.69</sup> <sub>-0.47</sub>	3.37 <sup>+3.37</sup> <sub>-3.37</sub>	1.86 <sup>+2.56</sup> <sub>-1.86</sub>	1.88 <sup>+1.71</sup> <sub>-1.71</sub>	1.88 <sup>+1.71</sup> <sub>-1.71</sub>
Fitted tflight events	4.93 <sup>+3.59</sup> <sub>-3.59</sub>	9.04 <sup>+6.06</sup> <sub>-6.06</sub>	2.40 <sup>+2.09</sup> <sub>-2.09</sub>	1.19 <sup>+2.03</sup> <sub>-1.19</sub>	0.95 <sup>+0.95</sup> <sub>-0.95</sub>	1.54 <sup>+1.22</sup> <sub>-1.22</sub>	1.19 <sup>+1.41</sup> <sub>-1.41</sub>	7.59 <sup>+6.19</sup> <sub>-6.19</sub>	4.64 <sup>+4.14</sup> <sub>-4.14</sub>	4.26 <sup>+3.38</sup> <sub>-3.38</sub>	4.26 <sup>+3.38</sup> <sub>-3.38</sub>
Fitted singletop events	0.19 <sup>+0.89</sup> <sub>-0.89</sub>	1.26 <sup>+1.24</sup> <sub>-1.24</sub>	1.11 <sup>+2.02</sup> <sub>-1.11</sub>	1.85 <sup>+1.93</sup> <sub>-1.85</sub>	1.14 <sup>+1.43</sup> <sub>-1.43</sub>	2.28 <sup>+2.13</sup> <sub>-2.13</sub>	1.25 <sup>+0.93</sup> <sub>-0.93</sub>	0.78 <sup>+2.41</sup> <sub>-2.41</sub>	1.89 <sup>+2.88</sup> <sub>-2.88</sub>	6.23 <sup>+5.13</sup> <sub>-5.13</sub>	6.23 <sup>+5.13</sup> <sub>-5.13</sub>
Fitted topEW events	0.52 <sup>+0.52</sup> <sub>-0.52</sub>	1.08 <sup>+0.75</sup> <sub>-0.75</sub>	0.83 <sup>+0.60</sup> <sub>-0.60</sub>	0.66 <sup>+0.60</sup> <sub>-0.60</sub>	0.65 <sup>+0.75</sup> <sub>-0.65</sub>	1.19 <sup>+0.75</sup> <sub>-0.75</sub>	0.74 <sup>+0.50</sup> <sub>-0.50</sub>	0.72 <sup>+0.66</sup> <sub>-0.66</sub>	0.95 <sup>+0.68</sup> <sub>-0.68</sub>	1.59 <sup>+0.94</sup> <sub>-0.94</sub>	1.59 <sup>+0.94</sup> <sub>-0.94</sub>
Fitted W_jets events	0.41 <sup>+1.14</sup> <sub>-0.41</sub>	0.22 <sup>+1.36</sup> <sub>-0.22</sub>	0.20 <sup>+6.39</sup> <sub>-0.20</sub>	0.20 <sup>+0.20</sup> <sub>-0.20</sub>	0.38 <sup>+0.64</sup> <sub>-0.38</sub>	1.00 <sup>+1.07</sup> <sub>-1.00</sub>	0.53 <sup>+0.68</sup> <sub>-0.53</sub>	0.27 <sup>+0.82</sup> <sub>-0.27</sub>	0.41 <sup>+0.77</sup> <sub>-0.41</sub>	0.96 <sup>+1.08</sup> <sub>-1.08</sub>	0.96 <sup>+1.08</sup> <sub>-1.08</sub>
Fitted Z_jets events	0.23 <sup>+0.38</sup> <sub>-0.23</sub>	0.55 <sup>+0.57</sup> <sub>-0.55</sub>	0.48 <sup>+0.52</sup> <sub>-0.48</sub>	0.67 <sup>+0.73</sup> <sub>-0.67</sub>	0.76 <sup>+0.63</sup> <sub>-0.76</sub>	1.69 <sup>+1.05</sup> <sub>-1.69</sub>	0.91 <sup>+0.48</sup> <sub>-0.48</sub>	0.70 <sup>+0.74</sup> <sub>-0.70</sub>	2.10 <sup>+0.99</sup> <sub>-0.99</sub>	2.10 <sup>+0.99</sup> <sub>-0.99</sub>	2.10 <sup>+0.99</sup> <sub>-0.99</sub>
Fitted diboson events	0.12 <sup>+0.32</sup> <sub>-0.12</sub>	0.06 <sup>+0.35</sup> <sub>-0.06</sub>	0.24 <sup>+0.23</sup> <sub>-0.24</sub>	0.22 <sup>+0.51</sup> <sub>-0.22</sub>	0.03 <sup>+0.27</sup> <sub>-0.03</sub>	0.21 <sup>+0.17</sup> <sub>-0.21</sub>	0.16 <sup>+0.16</sup> <sub>-0.16</sub>	0.29 <sup>+0.27</sup> <sub>-0.29</sub>	0.05 <sup>+0.17</sup> <sub>-0.05</sub>	0.53 <sup>+0.83&lt;/sup</sup>	

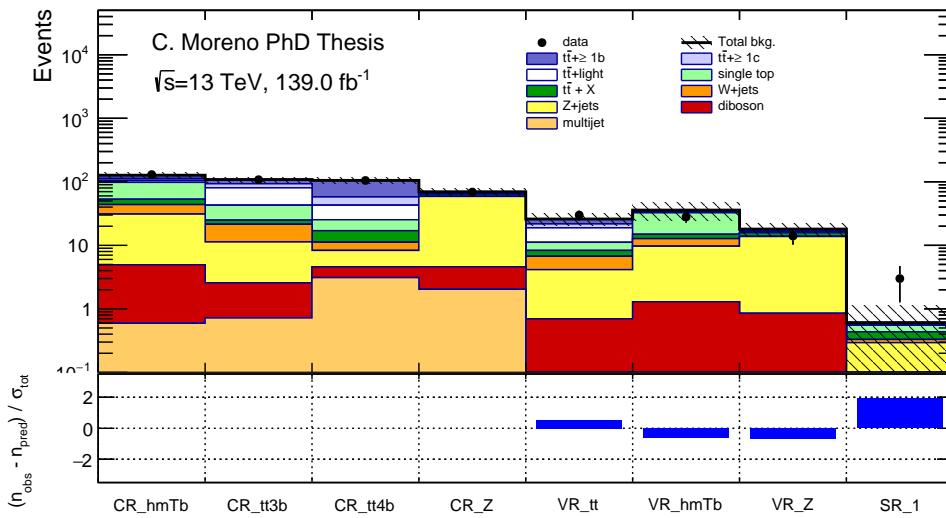


Figure 8.9: Comparison between data and post-fit background prediction in all regions dedicated to the higgsino mass hypothesis  $m(\tilde{H}) = 1000$  GeV, with the significance in each bin shown in the bottom pannel.

## 8.6 Interpretations

Given that no significant excess over the SM prediction is found in the signal regions, the tools introduced in section 8.3 are used to derive statements about the compatibility of these measurements with BSM physics models. This is done using two different approaches. On the one hand, model-independent upper limits are derived by determining the range of observed yields that would be compatible with the background prediction with a significance lower than  $5\sigma$ . Model-dependent upper limits, on the other hand, are determined by performing a scan on the signal strength for the particular signal model under examination to find the lowest value possible that is excluded according to the observation. The results for both approaches are discussed below.

### 8.6.1 Model-independent upper limits

The model-independent limits are derived by means of the so-called *discovery fit*, that aims at determining the largest possible contribution for any signal model in the fitted bins without ruling out the background hypothesis. In this case, both control and signal regions enter the fit to data, where the signal yield  $S$  is included as an additional parameter of interest to be estimated from data. This fit is performed with a description that includes exclusively one signal region, without the possibility of combining multiple signal regions together. The model-independent limits are estimated only for three *discovery regions*, with their corresponding control regions included in the prescription, for a total of three independent fits. They are optimized for sensitivity to the low-, intermediate- and high-neutralino mass regimes, which coincide with the leading signal regions for the signal mass points  $m(\tilde{H}) = 250, 500$  and  $1000$  GeV. Each signal region is fitted together with the control regions for the corresponding mass points, and the limits on the expected signal yield are estimated. For this, instead of the contribution from the signal

Signal channel	$\langle\epsilon\sigma\rangle_{\text{obs}}^{95} [\text{fb}]$	$S_{\text{obs}}^{95}$	$S_{\text{exp}}^{95}$	$p(s = 0)$
SR-1-250	0.04	6.2	$5.9_{-0.9}^{+1.7}$	0.48 (0.05)
SR-1-500	0.04	5.5	$4.0_{-0.6}^{+1.7}$	0.18 (0.92)
SR-1-1000	0.05	6.7	$4.3_{-0.9}^{+0.9}$	0.03 (1.89)

Table 8.4: 95% CL model-independent upper limits as extracted from the three discovery signal regions: SR-1-250, SR-1-500 and SR-1-1000. The limits are estimated with 50000 pseudo experiments, and are consistent with the results obtained under the asymptotic approximation. The table shows the 95% CL observed upper limits on the visible cross section ( $\langle\epsilon\sigma\rangle_{\text{obs}}^{95}$ ), the observed and expected 95% CL upper limits on the number of BSM events ( $S_{\text{obs}}^{95}$  and  $S_{\text{exp}}^{95}$ ), and the  $p$ -value together with the significance  $Z$  of the observation.

model described in section 7.3, a “dummy” signal that only contributes to the signal region with a yield  $S$  is used and no contribution in any control region. The upper limit is then estimated performing a scan of the signal yield  $S$ , excluding those values with  $\text{CL}_s \leq 0.05$ , corresponding to a confidence level of 95%. The results for the three discovery regions considered are shown in table 8.4, with upper limits of about 6 signal events in each, as expected from the similar level of predicted number of background events. In the case of SR-1-1000, the significance of the measurement is estimated to be  $1.89\sigma$  above the background expectation, as already shown in figure 8.9.

### 8.6.2 Model-dependent upper limits

Similarly to the model-independent limits, an *exclusion fit* is performed in order to derive upper limits in the cross section of the signal process under study. For each mass hypothesis, all available signal and control regions are included in the fit, and the signal strength is included in the description as a parameter of interest. In order to determine the upper limit on the allowed signal cross section according to the observation, a scan in a wide range of values of the signal strength is performed, with all values corresponding to a  $\text{CL}_s \leq 0.05$ , computed with the profile likelihood ratio as test statistic, are excluded.

The 95% CL upper limits for the signal model assuming  $\text{BR}(\tilde{H} \rightarrow h\tilde{G}) = 100\%$  are shown in figure 8.10. The expected exclusion limits cover all mass points up to approximately  $m(\tilde{H}) = 1050$  GeV. Signal masses below 200 GeV are not included in the description since the selection on  $E_T^{\text{miss}}$  applied in the analysis reduces the signal acceptance in that kinematic region to negligible levels. The sensitivity of this analysis is complemented by a low- $E_T^{\text{miss}}$  channel, and the mass hypotheses not included in this search are handled by the low-mass analysis. The observed limits in the low-mass region match perfectly the expected ones, while at high neutralino mass they start to diverge due to the small excess found in the leading signal regions for  $m(\tilde{H}) \geq 800$  GeV. According to the observed data, all mass hypotheses are excluded at 95%CL up to  $m(\tilde{H}) = 950$  GeV. The impact of the systematic uncertainties in the sensitivity of the analysis is minor for high neutralino mass hypothesis, with the expected limits computed with the full implementation of all NPs differing by a few percent from the limits computed with only statistical uncertainties. For mass hypotheses below 400 GeV, where the high signal cross section leads to larger SRs, the impact of the systematic uncertainties dominates over the statistical component, and expected limits become up to 25% weaker when the all systematic uncertainties are included.

In addition, the 95% CL exclusion limits for models with different higgsino decay branching

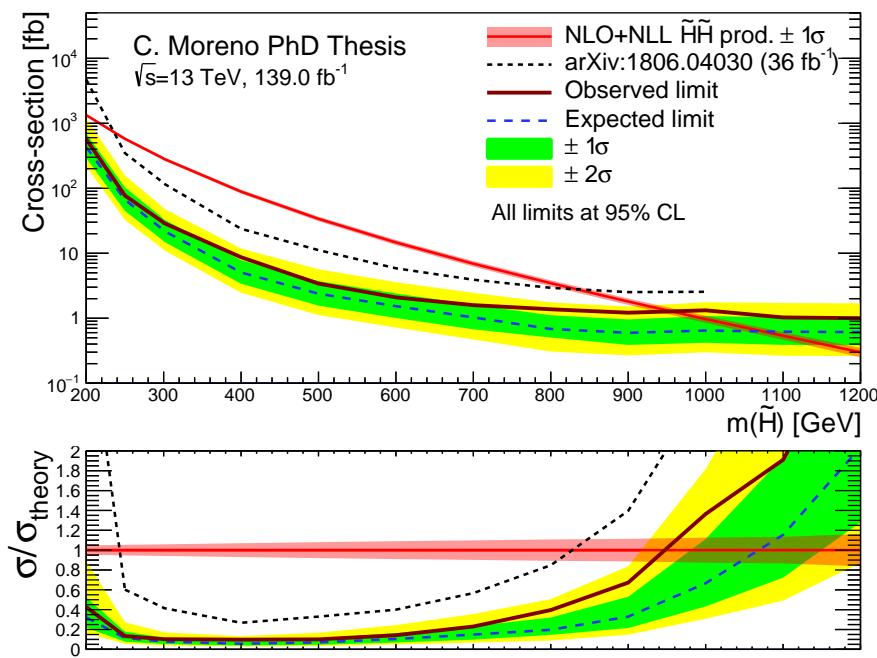


Figure 8.10: Expected and observed 95% CL exclusion limits for all mass hypotheses under consideration in the analysis. For each signal mass point, the limits are estimated using the asymptotic approximation, and the exclusion between mass points is estimated with linear interpolation. The observed 95% CL upper limits from the analysis with integrated luminosity of  $36 \text{ fb}^{-1}$  are included for reference.

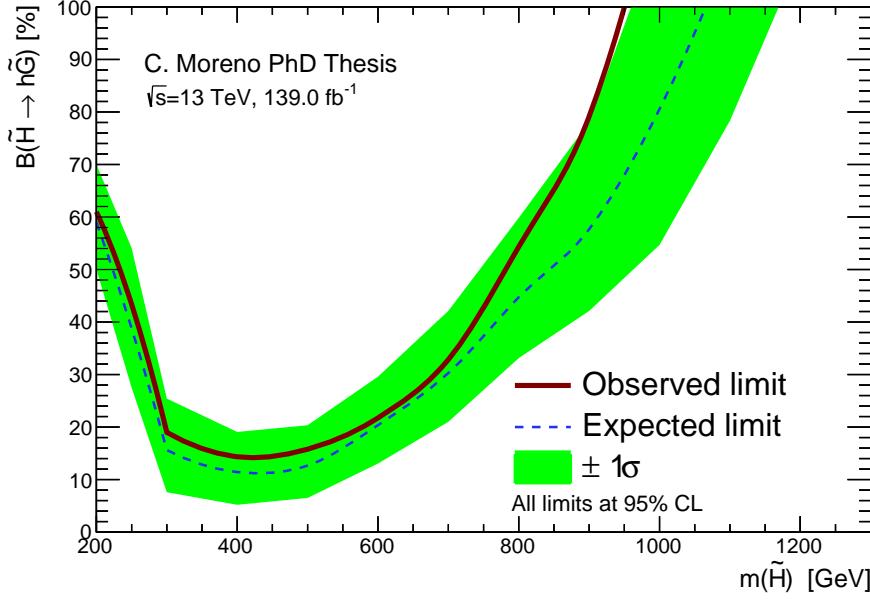


Figure 8.11: 95% CL exclusion limits in the higgsino mass - higgsino decay branching ratio plane, where  $\text{BR}(\tilde{H} \rightarrow h\tilde{G}) + \text{BR}(\tilde{H} \rightarrow Z\tilde{G}) = 1$  is assumed. For each signal mass point, the limits are estimated using the asymptotic approximation, and the exclusion between mass points is estimated with linear interpolation.

ratios to Higgs and gravitino are provided in figure 8.11. They are estimated using MC samples generated with  $\text{BR}(\tilde{H} \rightarrow h\tilde{G}) = \text{BR}(\tilde{H} \rightarrow Z\tilde{G}) = 50\%$ , where events are reweighted by a factor depending on a truth-level classification based on the number of Higgs and  $Z$  bosons in the decay chain. Even though the analysis is optimized exclusively for the scenario where all neutralinos decay to a Higgs boson and a gravitino, thanks to the good signal-to-background discrimination achieved with the BDT and the good sensitivity provided by the signal region strategy adopted, a large portion of the phase space can be excluded. For intermediate signal masses, models with  $\text{BR}(\tilde{H} \rightarrow h\tilde{G})$  as low as 15%, which represents a large improvement with respect to the previous analysis with  $36 \text{ fb}^{-1}$  luminosity [135]. As the mass increases, the sensitivity to lower BR scenarios decreases, as the higgsino pair production cross section rapidly decreases. This large exclusion observed is realized partially due to the large acceptance of events with only Higgs bosons in the decay chain, and also thanks to a non negligible contribution from events with one Higgs and one  $Z$  boson as accepted by the BDT selection.

# Chapter 9

## Perspective and prospects of the search

### 9.1 Introduction

The search presented in the previous chapter is optimized for higgsino pair production in a natural SUSY GMSB scenario. As mentioned briefly in chapters 7 and 8, there is another analysis in ATLAS optimized for the same signal but in a different kinematic region with low  $E_T^{\text{miss}}$ . Other analyses target mixed higgsino decay branching ratio to  $Z$ -bosons as well. In addition, an equivalent search carried out by CMS has been published using the full Run 2 dataset, following a different approach than the one chosen in the ATLAS analyses.

Given the good sensitivity provided by this search to the targeted model, it is expected that it also provides some coverage of other signal scenarios that lead to similar final states. In particular, the optimization of the analysis is performed based on a simplified model that does not account for complex SUSY scenarios that may reproduce the behaviour of nature more accurately.

This chapter provides a short overview to both topics. In section 9.2 the results of this search are compared to other analyses carried out by ATLAS and CMS, while section 9.3 provides an overview of possible signal models to which this analysis could be sensitive.

### 9.2 Comparison of the results to other analyses

The model targeted in this search, focusing mainly on the scenario where  $\text{BR}(\tilde{H} \rightarrow h\tilde{G})=100\%$ , has been explored as well with other approaches. As mentioned in the previous chapters, a low- $E_T^{\text{miss}}$  channel exists as well in ATLAS to target the very compressed scenarios in topologies with little to no  $E_T^{\text{miss}}$  at all. This model has also been explored by a search carried out in the CMS experiment. The results of both analyses are provided below, as a comparison with the ones achieved in this work.

#### 9.2.1 Low- $E_T^{\text{miss}}$ channel

As shown in figure 7.3, signal models with higgsino mass below 300 GeV typically result in topologies with low  $E_T^{\text{miss}}$ , due to large part of its mass needed to produce the intermediate

on-shell Higgs boson in the decay chain. Due to the  $E_T^{\text{miss}}$  requirement at the trigger level, and downstream in the pre-selection, the analysis presented in this work does not provide sensitivity to signals with  $m(\tilde{H}) < 200$  GeV. A complementary channel therefore exists in ATLAS, optimized for signal acceptance in that kinematic regime, where the  $E_T^{\text{miss}}$  requirements of the high-mass analysis make them inaccessible.

A similar approach to the one followed in the publication with integrated luminosity of  $36 \text{ fb}^{-1}$  is adopted [135]. Events are selected by means of a requirement on the number of  $b$ -jets at the trigger level, instead of a  $E_T^{\text{miss}}$ -based selection, which allows for large signal acceptance with negligible bias in  $E_T^{\text{miss}}$ . The selection in the number of tagged jets and their properties, such as the  $p_T$  or the  $b$ -tagging OP used, change for every data taking period. Details about the available triggers that were used in this analysis can be found in ref. [68–70]. The complex trigger strategy adopted required an equally complex calibration procedure, where events are sorted into orthogonal ‘‘buckets’’, depending on certain kinematic properties, in order to avoid applying corrections to a given event twice due to it fulfilling two different trigger requirements simultaneously. Further selections are applied to reduce the contribution from SM processes, such as veto for events with at least one electron or one muon with  $p_T > 20$  GeV, or a veto of events containing top quarks. The definitions of all objects used in this analysis match those used in the high-mass search.

Similarly to the high- $E_T^{\text{miss}}$  analysis, this analysis also relies on the intermediate Higgs bosons in the neutralino decay in order to identify signal events. The  $b$ -jets arising from the decay of the two Higgs bosons are selected as those with largest  $p_T$ , and are paired into Higgs boson candidates such that the separation between two  $b$ -jets coming from the same Higgs boson is minimized. The analysis regions are then defined by requirements on the mass of the reconstructed Higgs boson candidates, which form three concentric and exclusive ellipsoids, meaning that one event cannot be in two regions at the same time. Three variables are defined as follows in order to build the analysis regions:

$$\begin{aligned} X_{hh}^{\text{CR}} &= \sqrt{(m_{h,\text{lead}} - 120 \times 1.05 \text{ GeV})^2 + (m_{h,\text{subl}} - 110 \times 1.05 \text{ GeV})^2}, \\ X_{hh}^{\text{VR}} &= \sqrt{m_{h,\text{lead}} - 120 \times 1.03 \text{ GeV})^2 + (m_{h,\text{subl}} - 110 \times 1.03 \text{ GeV})^2}, \\ X_{hh}^{\text{SR}} &= \sqrt{\left(\frac{m_{h,\text{lead}} - 120 \text{ GeV}}{0.1m_{h,\text{lead}}}\right)^2 + \left(\frac{m_{h,\text{subl}} - 110 \text{ GeV}}{0.1m_{h,\text{subl}}}\right)^2}, \end{aligned} \quad (9.1)$$

where  $m_{h,\text{lead}}$  and  $m_{h,\text{subl}}$  indicate the heavier and lighter reconstructed Higgs boson candidates in the event, respectively. The regions are then defined by applying the following selections on these variables:

- Control region:  $X_{hh}^{\text{CR}} < 55$  GeV,  $X_{hh}^{\text{VR}} \geq 30$  GeV and  $X_{hh}^{\text{SR}} \geq 1.6$ .
- Validation region:  $X_{hh}^{\text{VR}} < 30$  and  $X_{hh}^{\text{SR}} \geq 1.6$ .
- Signal region:  $X_{hh}^{\text{SR}} < 1.6$ .

Two versions of each of these regions is included, one with exactly two  $b$ -tagged jets ( $2b$ ) used exclusively for the background estimate, and another with at least 4  $b$ -jets ( $4b$ ) where the statistical analysis takes place. The final region that is used to estimate the parameters of the theory and perform the hypothesis test is the  $4b$  signal region, which is binned in both  $E_T^{\text{miss}}$  and  $m_{\text{eff}}$  in order to optimize the sensitivity of the analysis.

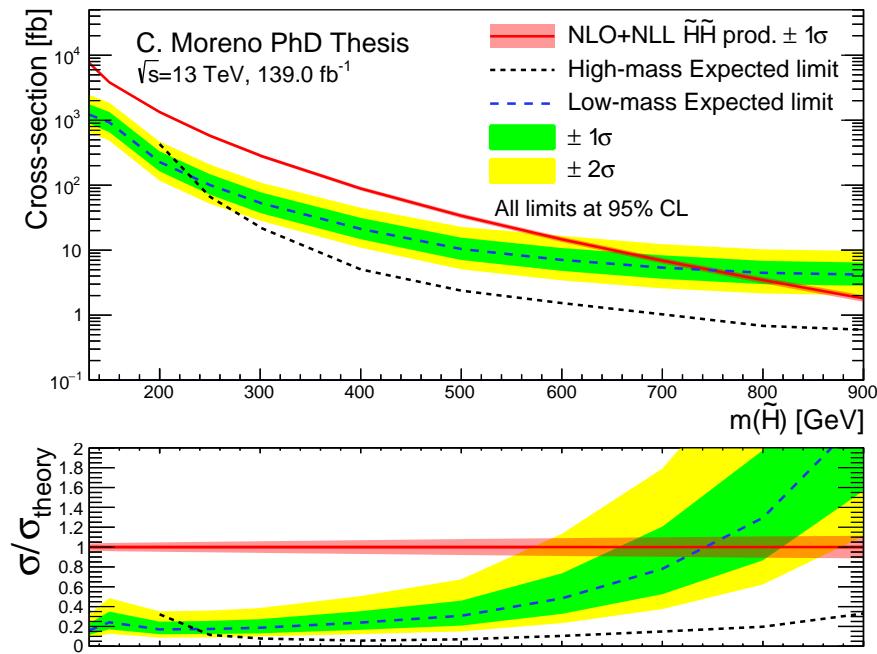


Figure 9.1: 95% CL expected exclusion limits for low-mass channel, where observation is set to match the background prediction in the signal region.

The selection used in this search, without any lower cut on  $E_T^{\text{miss}}$ , enhances the contribution from QCD multijet events in the signal region which largely dominates over other sources of background events. Given the usual difficulties encountered in the simulation of these processes, and the large number of events expected in this region, a data-driven approach for the background estimate is followed. This is carried out by means of a similar reweighting method as developed in section 7.2.4, in this case using a BDT instead of a NN. The BDT is trained to separate the  $2b$  from the  $4b$  events in the control region, and an event weight is derived for the  $2b$  events based on the BDT evaluation process.<sup>1</sup> The final background estimate in any of the  $4b$  region is formed by the observation in the equivalent  $2b$  sample after the reweighting.

The 95% CL exclusion limits provided with this channel are shown in figure 9.1, with large sensitivity to low  $\tilde{\chi}_1^0$  mass scenarios that are out of reach for the high- $E_T^{\text{miss}}$  channel. There is large improvements compared to the published results with integrated luminosity of  $36 \text{ fb}^{-1}$ , with a gain in sensitivity of a factor 3 in the targeted region. In addition, there is good coverage of intermediate mass hypotheses as well, with expected exclusion reaching up to  $m(\tilde{H}) = 750 \text{ GeV}$ . The transition point where low-mass analysis stops leading in sensitivity and the high- $E_T^{\text{miss}}$  channel starts dominating is about  $m(\tilde{H}) = 225 \text{ GeV}$ .

### 9.2.2 Comparison to CMS results

A search for the same SUSY model has been carried out by the CMS Collaboration using the dataset recorded from the full Run 2 LHC period [162]. Given the same final state, where the

<sup>1</sup>The procedure followed is more complex than that used for the estimate of QCD multijet events in the high-mass analysis. In this case, the weight is computed as the ratio of accumulated weights from the training in each leaf for each weak learner. The final event weight is computed as an average of the weights from each of the trees in the BDT, weighted by the learning rate.

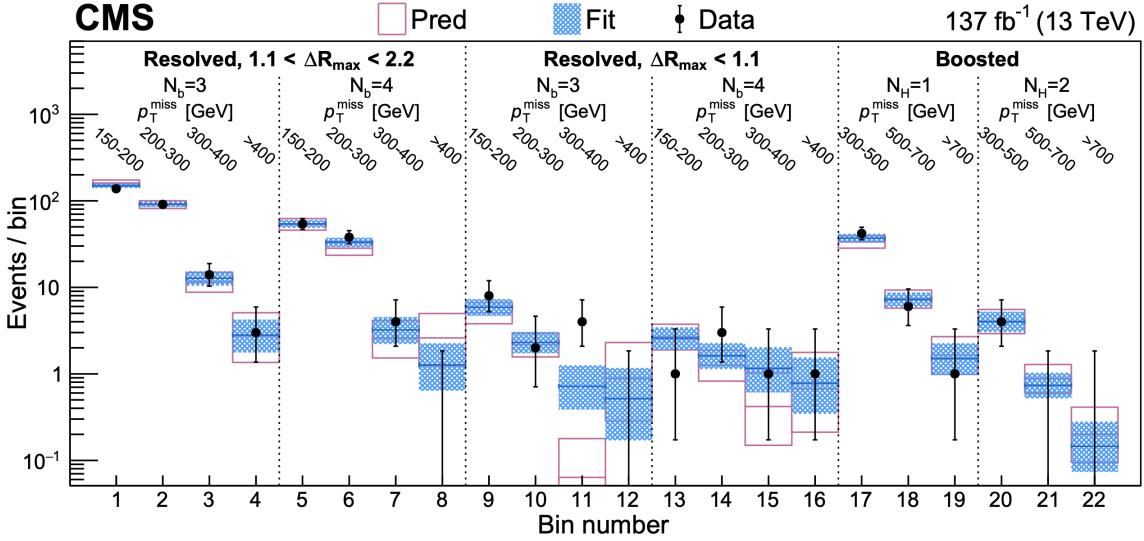


Figure 9.2: Observed and predicted yields in the search regions split in resolved and boosted topologies. The resolved channel is further split in two slices of  $\Delta R_{\text{max}}$ , which is defined as the maximum separation between the  $b$ -jets originating from the same Higgs candidate, and each slice is divided in bins  $b$ -jet multiplicity and  $E_T^{\text{miss}}$ . The boosted channel is divided in bins of number of reconstructed Higgs and  $E_T^{\text{miss}}$ . The number of Higgs is defined as the number of  $R = 0.8$  jets that contain two  $b$ -jets inside, corresponding to collimated decays of boosted Higgs bosons. The points with error bars represent the observed yields, with the pre-fit background prediction shown in magenta and the post-fit estimate in blue [162].

neutralino decays to a Higgs boson and a gravitino, the strategy followed has some common points with the approach adopted in ATLAS. In the analysis, two topologies are targeted: one channel is optimized for boosted Higgs bosons and another one targets resolved scenarios where the Higgs boson decays can be efficiently reconstructed. In both cases the strategy is to identify the Higgs bosons in order to extract properties from the parent neutralinos produced in the collision.

Events in this search are selected by means of a  $E_T^{\text{miss}}$  selection at the trigger level, together with an offline- $E_T^{\text{miss}}$  selection, which is higher for the boosted channel ( $E_T^{\text{miss}} > 300$  GeV) than the resolved one ( $E_T^{\text{miss}} > 150$  GeV). Instead of simulating the trigger selection on the MC samples for the signal and background processes, its efficiency is estimated from data in different bins in the variable  $H_T$  and applied as event weight to the simulated samples, removing the need for a dedicated calibration procedure. Similarly to the ATLAS analyses, a lepton veto is applied in the definition of the signal regions, although  $t\bar{t}$  leptonic events are used to derive the normalization of this background in the relevant kinematic regions. The intermediate Higgs bosons are then reconstructed<sup>2</sup> and selections are applied on the mass of the Higgs candidates, which differ depending on the topology targeted. The resolved signal region is then divided in bins of  $b$ -jet multiplicity,  $E_T^{\text{miss}}$  and the maximum separation between Higgs decay products. The boosted signal region is split in the number of reconstructed Higgs bosons and in bins of  $E_T^{\text{miss}}$ . In each case, the bins available in the SRs are combined in a fit for enhanced statistical power.

<sup>2</sup>Two approaches are followed. For resolved topologies, the four jets in the event with higher  $b$ -tagging score are paired in such a way that the mass difference between the reconstructed Higgs bosons is minimal. For boosted topologies,  $R=0.8$  jets are used to reconstruct the cases where both  $b$  quarks are collimated, and the mass of the two leading  $R=0.8$  jets is considered to be equivalent to the mass of the intermediate Higgs bosons.

The comparison of the observed data and the expected number of events from SM processes is shown in figure 9.2, with no significant discrepancies. The obtained exclusion limits are displayed in figure 9.3, where mass hypotheses up to  $m(\tilde{H}) = 1000$  GeV are excluded in scenarios with  $\text{BR}(\tilde{H} \rightarrow h\tilde{G}) = 100\%$  and nearly-degenerate  $\tilde{\chi}_1^0$ ,  $\tilde{\chi}_1^\pm$  and  $\tilde{\chi}_2^0$ . Even though very different approaches are followed between the ATLAS and CMS searches, similar results are achieved, with sensitivity to a wide range of mass hypotheses in both cases and the same phase space excluded by both experiments. The CMS analysis, however, does not provide sensitivity to signal masses below 200 GeV, which is inaccessible due to the  $E_T^{\text{miss}}$  requirements applied both at the trigger level and downstream in the analysis event selection. In the case of ATLAS, on the other hand, large sensitivity is provided to all kinematically allowed signal mass hypotheses below  $m(\tilde{H}) = 1050$  GeV thanks to the low-mass analysis being free from  $E_T^{\text{miss}}$  requirements and relying on  $b$ -jet identification at the trigger level. One of the main features of the CMS analysis is the use of large- $R$  jets to efficiently reconstruct collimated Higgs decays arising in scenarios with high  $\tilde{\chi}_1^0$  mass, which allows for sensitivity to signal cross sections as low as  $1.2 \text{ fb}^{-1}$  at  $m(\tilde{H}) = 950$  GeV. While this approach provides excellent reach at high masses, the chosen method in the ATLAS high- $E_T^{\text{miss}}$  analysis, with optimization performed for each mass point by means of the BDT and dedicated signal regions, results in a large increase in expected sensitivity in comparison, up to a factor 3 for the mass range  $500 \leq m(\tilde{H}) \leq 900$ , and probing cross sections as low as  $0.5 \text{ fb}^{-1}$  at  $m(\tilde{H}) = 950$  GeV. Due to the  $2\sigma$  excess in the signal regions for  $m(\tilde{H}) \geq 800$  GeV, the largest excluded signal mass hypothesis is similar in both ATLAS and CMS analyses, although the final exclusion in the intermediate mass range, between 300 and 700 GeV, is deeper in the case of the ATLAS high- $E_T^{\text{miss}}$  channel, with excluded cross sections a factor 3 lower than the CMS analysis.

### 9.2.3 ATLAS exclusion in branching ratio plane

Two other ATLAS searches can provide sensitivity to similar signal scenarios than the ones targeted in the analysis presented in this work. They are optimized to exploit topologies involving  $Z$  boson that arise in the neutralino decay chain. One of them, which is referred to as 4-lepton analysis [163], targets final states where both  $Z$  bosons decay leptonically, either to electron or muon pairs but without vetoing decays to  $\tau$ -leptons. This strategy includes selections with at least 4 reconstructed leptons, either electrons, muons or a combination of both, large  $E_T^{\text{miss}}$  to account for the  $\tilde{G}$  escaping the detector, and a veto on events with  $b$ -tagged jets. This search is particularly sensitive to low-mass signal hypotheses in the region with large decay branching ratio to  $Z$  bosons, with the signal acceptance decreasing rapidly as the higgsino mass increases due to topologies becoming collimated, where leptons start to overlap and events are lost due to electron/muon isolation criteria and the requirement for at least 4 leptons in the final state.

Another search, referred to as 0-lepton or full-hadronic analysis [164], targets final states where the  $Z$  boson decay hadronically, with looser selections to allow for acceptance of scenarios with mixed higgsino decay branching fraction. The signal regions therefore include events with large jet multiplicity plus  $E_T^{\text{miss}}$ . It can provide sensitivity to boosted scenarios thanks to the use of large- $R$  jets that capture collimated decays. Additionally, events are selected as signal-like by means of vector and Higgs boson tagging techniques to identify the intermediate resonances in the higgsino decay chain. The signal regions that are most sensitive to the SUSY scenarios under examination require two tagged  $Z$  bosons, in two channels with and without  $b$ -tagging requirements,<sup>3</sup> and one  $Z$  boson together with a Higgs boson, with additional selections on the

---

<sup>3</sup>The track-jets inside the large- $R$  jet are evaluated with  $b$ -tagging algorithms in order to identify cases where two  $b$ -quarks are merged into a single jet.

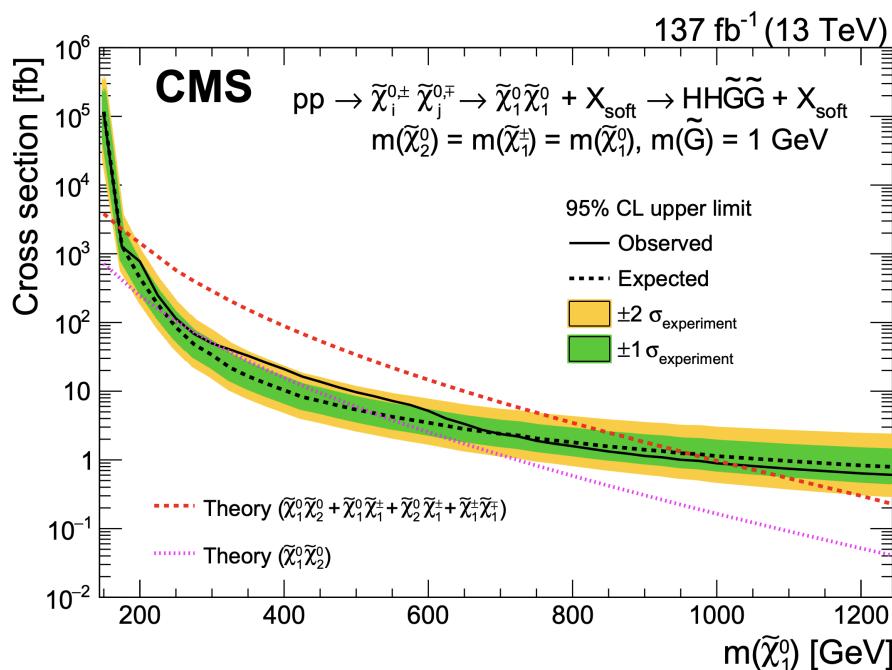


Figure 9.3: Expected and observed exclusion limits at 95% CL in the search carried out by the CMS Collaboration. Two theoretical cross sections are provided in the figure. The red dashed line represents the production cross section under the assumption of nearly degenerate  $\tilde{\chi}_1^0$ ,  $\tilde{\chi}_1^\pm$  and  $\tilde{\chi}_2^0$ , where all possible production of  $\tilde{\chi}_2^0$  and  $\tilde{\chi}_1^\pm$  effectively results in a pair of  $\tilde{\chi}_1^0$ , since the intermediate decay products are too soft to be reconstructed. Additionally, the cross section for producing exclusively a pair of  $\tilde{\chi}_1^0$  is provided in pink [162].

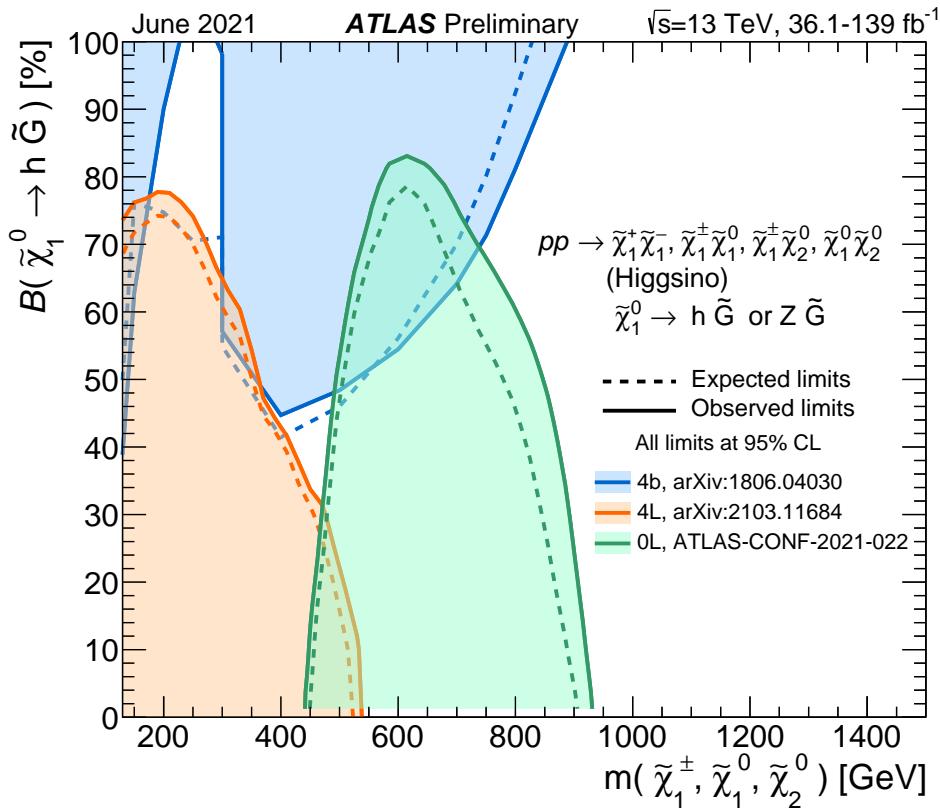


Figure 9.4: Expected and observed exclusion limits at 95% CL in the higgsino mass - decay branching ratio plane with the contribution from three different ATLAS analyses overlaid. The results from the 4-lepton and full-hadronic analyses using the full Run 2 dataset are shown in orange and green, respectively, while the limits obtained with the previous iteration of the analysis presented in this work is shown in blue [164].

number of  $b$ -jets.

The 95% CL upper limits in the higgsino decay branching ratio plane for both analyses using the full Run 2 dataset are shown in figure 9.4, together with the exclusion limits from the previous round of the analysis discussed in this thesis. Both the 4-lepton and full-hadronic analyses provide good sensitivity in their respective targeted mass regimes, reaching exclusion of branching ratio values of up to 80% for  $m(\tilde{H}) = 200$  GeV and  $m(\tilde{H}) = 600$  GeV, respectively. The previous round of the effort presented in this thesis shows a gap at low masses consistent with the  $2\sigma$  excess observed in the dataset corresponding to the LHC data taking years 2015 and 2016. As shown, a small gap in sensitivity at masses around 500 GeV and  $\text{BR}(\tilde{H} \rightarrow h\tilde{G})=40\%$  was left uncovered by those three analyses. The current version of the high-mass channel alone is able to exclude a large portion of this branching ratio plane, covering the two gaps that appear in figure 9.4, plus a large portion of previously unexplored phase space. Thanks to the optimization process carried out for each mass hypothesis in this analysis, the gain in sensitivity is very large in comparison to the previous iteration.

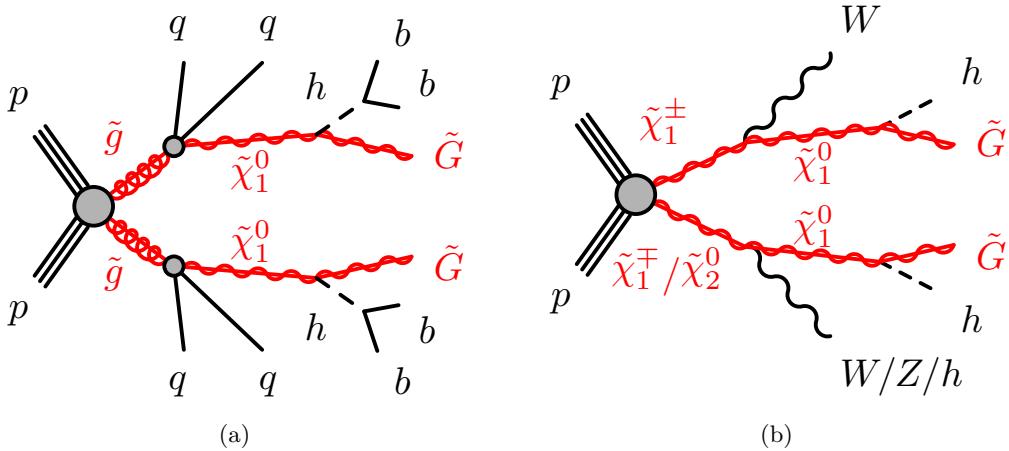


Figure 9.5: Feynman diagrams for non-simplified SUSY scenarios, where (a) shows a gluino pair production model coupled with the GMSB mechanism with a long decay chain, and (b) shows a case with non-degenerate electroweakinos.

### 9.3 Prospects and reach of the search

The analysis presented in this thesis has been optimized for the specific scenario where SUSY breaking is mediated by gauge interactions and the next-to-lightest supersymmetric particle (NLSP) is dominated by the higgsino component. While the followed approach with dedicated optimization performed for each mass hypothesis considered leads to excellent sensitivity to the model under investigation, it can result in an analysis that is less optimal with similar signatures but slightly different topologies.

#### 9.3.1 Non-simplified models

The search for higgsino pair production presented relies on the assumptions made in the simplified models for SUSY, which largely condition the topologies targeted in the analysis strategy. In this approximation, only the main particle(s) searched for is assumed to have a mass in the energy range scanned by the LHC, while the rest are assumed to be completely decoupled. The way in which SUSY may manifest in nature can be more complex than these models, in which case the predictions over the expected topologies and production cross sections may vary from the ones assumed in this analysis. Furthermore, larger supersymmetric extensions of the SM than the MSSM can also reproduce viably the physics in the infrared limit while resulting in more complex topologies.

More complex models, such as the one shown in figure 9.5a, result in final states with higher jet/lepton activity than the one considered in this search. In the case of the model shown, gluino pair production events can decay to a pair of  $\tilde{\chi}_1^0$ , which further decay into a pair of Higgs boson plus  $\tilde{G}$ , resulting in a similar topology than the one targeted but with a larger number of expected jets. The veto applied in this search to events with more than 7 jets already poses a problem in terms of signal acceptance of this type of non-simplified models, and while the veto is very helpful in reducing the contribution of background events in the analysis, it was applied based solely on studies on the relevant signal model. In addition, the large jet activity arising in the longer decay chains in the non-simplified models leads to different correlations between  $E_T^{\text{miss}}$  and jet-related kinematic variables, which the BDT used in the analysis exploits based on

the simplified model explored, and can also limit the sensitivity of the search.

Besides the assumptions made in the simplified model used to design the search, naturalness arguments in the MSSM introduces further constraints in the signal under study. As discussed in section 2.3.3, it is generally expected to have nearly degenerate  $\tilde{\chi}_1^0$ ,  $\tilde{\chi}_1^\pm$  and  $\tilde{\chi}_2^0$  with the higgsino component being dominant over the wino and bino, which are expected to be much heavier. However, the case where  $\tilde{\chi}_1^0$ ,  $\tilde{\chi}_1^\pm$  and  $\tilde{\chi}_2^0$  are not so close in mass is also a possibility to consider, which leads to more complex topologies in general. One example of such cases, coupled with GMSB mechanisms, is shown in figure 9.5b, with chargino pair production or a  $\tilde{\chi}^\pm$  produced together with a  $\tilde{\chi}_2^0$ , which all later decay to a  $\tilde{\chi}_1^0$  and the familiar decay to Higgs boson plus  $\tilde{G}$  is assumed. Due to the mass difference between the lightest neutralino and the rest of electroweakinos, the intermediate products in the decay chain ( $W$ ,  $Z$  and extra Higgs bosons) might decay into hadrons or leptons with  $p_T$  large enough to be reconstructed. In these cases, the various lepton vetoes applied in the analysis, together with the restriction on large jet multiplicity, the signal acceptance would be reduced significantly. In SUSY scenarios with such phenomenology, the only production mechanism for which this analysis would be sensitive is pure  $\tilde{\chi}_1^0\tilde{\chi}_2^0$  production, and given the significantly lower expected cross section for such process the only excluded mass hypotheses based on observations are in the range  $250 \text{ GeV} \leq m(\tilde{H}) \leq 600 \text{ GeV}$ , which is significantly lower than the one reported under the assumption of nearly degenerate  $\tilde{\chi}_1^0$ ,  $\tilde{\chi}_2^0$  and  $\tilde{\chi}_1^\pm$ .

Additionally, other SUSY breaking scenarios could be considered in order to propose signal models with similar topologies to the ones explored in the analysis presented in this work. In generic soft-breaking scenarios where the LSP is a higgsino-dominated  $\tilde{\chi}_1^0$ , decays like the one shown in figure 9.6a are possible, where  $\tilde{\chi}_1^0$  now turns into the source of  $E_T^{\text{miss}}$  in the final states. The topology of this kind of models in the regions of the phase space with large mass splitting between the  $\tilde{\chi}_1^0$  and the  $\tilde{\chi}_2^0$ ,  $\tilde{\chi}_3^0$  is very similar to the one targeted in this search, since Higgs bosons are allowed to be produced on-shell, with the subsequent decay to  $b\bar{b}$  pairs. The theoretical  $\tilde{\chi}_2^0\tilde{\chi}_3^0$  production cross section, however, is much lower than that of  $\tilde{\chi}_1^0\tilde{\chi}_2^0$  and therefore the expected sensitivity is reduced. In other SUSY models coupled with the Peccei-Quinn mechanism to solve the strong CP problem, similar topologies can arise where a massive *axino*  $\tilde{a}$ , the superpartner of the axion field, is the LSP in RPC scenarios, which also provides a strong connection with the dark matter problem [165–170]. In natural scenarios with light higgsino triplet, neutralinos can be produced in the same way as in the model targeted by the search discussed in this work, which decay to the axino LSP in association with a Higgs boson. The Feynman diagram for such process is shown in figure 9.6b.

The search discussed in this work can provide sensitivity to the scenarios shown in figure 9.6 as well, which result in final states with large  $b$ -jet multiplicity from the decay of the intermediate Higgs bosons together with  $E_T^{\text{miss}}$  arising from the LSP. The phase space covered by the search is determined by the kinematic properties of the signal model, which in turn depend strongly on the mass splitting between the produced SUSY particle and the LSP. In cases with small mass splitting between the two, the decay products are expected to be soft and therefore the final states are expected to have low- $p_T$   $b$ -jets and low values of  $E_T^{\text{miss}}$ . This compressed kinematic regime is identical to the one targeted by the low-mass signal regions in the high- $E_T^{\text{miss}}$  channel of the search for GMSB higgsino models, and therefore those are expected to provide the largest sensitivity. Cases with large mass splitting between the heavy SUSY particle and the LSP, on the other hand, are expected to lead to final states with boosted, and even collimated,  $b$ -jets and large values of  $E_T^{\text{miss}}$  due to highly energetic LSP leaving the experiment undetected. This kinematic regime is the same as the one targeted by the high-mass signal regions in the analysis presented in this work, and therefore they can provide the most sensitivity to those corners of

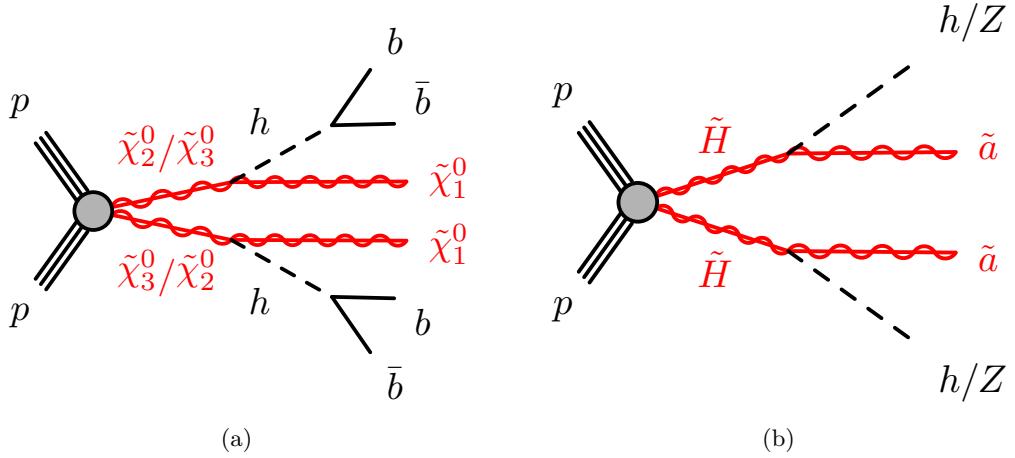


Figure 9.6: Feynman diagrams for non-simplified SUSY scenarios, where (a) shows a model for  $\tilde{\chi}_2^0\tilde{\chi}_3^0$  production in a scenario where they are heavier than  $\tilde{\chi}_1^0$ , and (b) shows a scenario with SUSY coupled with the Peccei-Quinn mechanism, where the axino  $\tilde{a}$  is the LSP.

the parameter space.

### 9.3.2 Long-lived scenarios

The search presented in this work is optimized for scenarios where all SUSY particles decay promptly, effectively with negligible lifetime. One of the features of GMSB scenarios is that the NLSP lifetime can vary largely depending on its mass and increases with the SUSY breaking scale  $\sqrt{F}$  [171]. In cases with low values of  $\sqrt{F}$ , the distance inside the detector travelled by the NLSP before decaying is typically very small and the decay can be considered prompt. For larger values of  $\sqrt{F}$ , however, the decay distance can be large enough for the NLSP to decay deeper inside the detector, in the middle of the tracking system or in the calorimeters, in which case the object reconstruction algorithms can have worse performance over the decay products. When the SUSY breaking scale is very large, the NLSP decay lifetime is long enough for it to travel the whole detector without interacting, and all of its decay products are lost to the analysis, contributing to the total  $E_T^{\text{miss}}$  in the events. In the case of the model targeted by this search, very long higgsino lifetime would lead to topologies with only  $E_T^{\text{miss}}$  in the final state, as none of the Higgs bosons would be reconstructed if the neutralinos decay outside of the detector.

The case with intermediate NLSP decay lifetimes, enough for the detector to measure most of its energy but still well differentiated from the prompt decay case, can still lead to large differences in the sensitivity of the analysis. In particular, this search relies heavily on the correct identification of  $b$ -jets both to reconstruct the intermediate Higgs bosons and to construct variables that are good at rejecting  $t\bar{t}$  background, such as  $m_{T,\min}^{b\text{-jets}}$ . As discussed in chapter 6, the  $b$ -tagging algorithms exploit the typical displacement of the  $B$ -hadrons to tag jets as originated from  $b$ -quarks, and therefore having long-lived neutralinos will result in displaced Higgs bosons whose decay products will be further away from the interaction point than in the prompt case. Scenarios where the NLSP does not decay promptly are susceptible to have distorted performance of the  $b$ -tagging algorithms and therefore the sensitivity of this analysis in particular can be severely affected.

Figure 9.7 shows the efficiency of the DL1r  $b$ -tagging algorithms as a function of the neu-

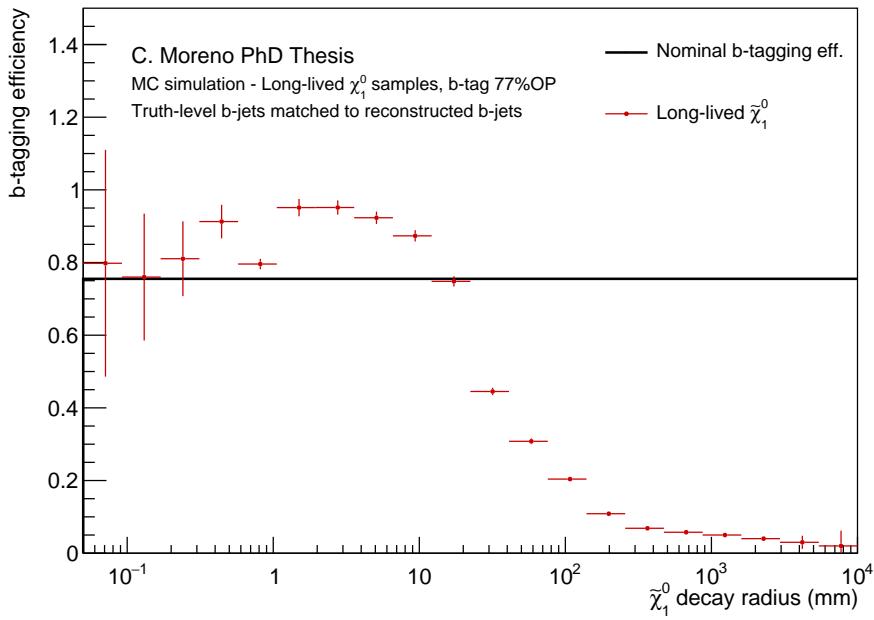


Figure 9.7: Efficiency of the  $b$ -tagging algorithm DL1 evaluated on MC simulated events with long-lived neutralinos as a function of its decay radius. In order to correctly identify the jets arising in the neutralino decay chain, a matching procedure using truth information is followed. The efficiency is estimated using multiple samples where the neutralino lifetime is a configurable input in the simulation, and multiple samples are combined for increased statistics.

tralino decay radius for  $B$ -hadrons, which is computed by matching reconstructed jets to truth jets that originate in the  $\tilde{\chi}_1^0$  decay. As the lifetime of the neutralino increases, so does the  $b$ -tagging efficiency, due to the jets arising in the Higgs boson decay becoming more likely to have optimal displacement to be tagged as  $b$ -jets. However, as the lifetime keeps increasing, the efficiency drops rapidly, specially when the  $\tilde{\chi}_1^0$  travels deeper in the detector without leaving any trace in the tracker. The sensitivity of the analysis is expected to follow this trend as well, with a slight increase at low lifetimes due to the higher  $b$ -tagging efficiency, and little to no sensitivity at all to scenarios with large  $\tilde{\chi}_1^0$  lifetimes.<sup>4</sup> The calibration of the  $b$ -tagging algorithm is usually carried out in ATLAS using simulated  $t\bar{t}$  events as a reference. Due to the nature of the particles involved in these events, decay chains with long-lived particles other than the  $B$ -hadrons are not accounted for in the calibration. In case this search were to be used to provide sensitivity to the long-lived GMSB scenarios, an extra uncertainty on the  $b$ -tagging uncertainty for the signal would need to be included in the fit to account for possible differences with the calibration on prompt decays. This uncertainty was derived in previous interpretations of prompt analyses with long-lived scenarios in ATLAS, and was found to be between 10% and 20%, depending on the kinematics of the events [172].

The loss of sensitivity due to the reduction in  $b$ -tagging efficiency at large  $\tilde{\chi}_1^0$  lifetime could be partially circumvented by replacing the Higgs reconstruction techniques to rely on other algorithms optimized for finding displaced vertices. In ATLAS such algorithms have been developed,

<sup>4</sup>The  $b$ -tagging mistag rate increases drastically as well with the displacement of the jets, in particular light-jets, for which it increases from about 3% for prompt decays up to 80% in long-lived scenarios. This search however is not affected by mistag rate increase at large lifetimes, since the background is composed by processes where all particles decay promptly and the signal model only has decay chains involving  $b$ -quarks.

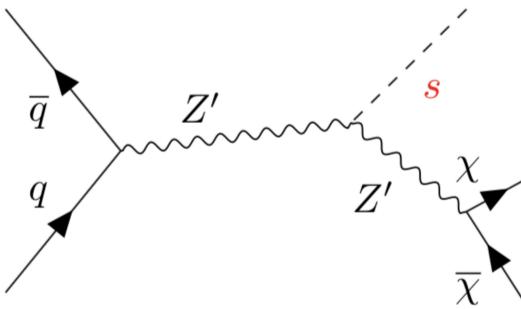


Figure 9.8: Feynman diagram for a simplified DM model, where the mediator  $Z'$  is produced in the proton collisions at the LHC and radiates a dark Higgs  $s$  before decaying to a pair of dark fermions  $\chi$ . In the scenario discussed in the test, the  $s$  is assumed to decay to a pair of Higgs bosons.

such as the VrtSecInclusive (VSI) [173], that have been used in other searches to reconstruct similar topologies than the one discussed above. All other variables using  $b$ -jet information would also need to be recomputed to include the correct information about the displaced jets.

### 9.3.3 New dark matter models

The analysis proposed in this work is in essence a search for di-Higgs plus  $E_T^{\text{miss}}$  final states, and can be used to target other BSM physics models that result in similar topologies. One of such cases arises in the context of direct dark matter detection at the LHC, where new exotic particles typically result in large  $E_T^{\text{miss}}$  in the final state, if they do not interact with SM particles, or can be identified by their decay products if they do. DM models often consider the existence of a dark or hidden matter sector that couples to the SM by means of new mediators, typically a new neutral, spin-1  $Z'$  boson that arises from a new  $U'(1)$  symmetry. In some scenarios, these mediators can establish thermal equilibrium between DM and SM sectors in the early universe, providing the framework for dark fermion annihilation processes,  $\chi\chi \rightarrow \text{SM}$ , that determines the DM relic abundance [174–176].

These models usually suffer from strong constraints based on the Higgs-boson production and decay modes, with the observed relic abundance being hard to reproduce [177]. Scenarios where the particles in the hidden sector obtain their mass through a Higgs mechanism are particularly appealing in this context, since they typically allow for a lighter dark scalar, or dark Higgs,  $s$ , such that the dark fermions do not need to annihilate to the SM particles directly, and processes of the type  $\chi\chi \rightarrow ss$  are allowed. The relic abundance is determined by the coupling strength between  $\chi$  and  $s$  [178, 179], while small mixing between the dark Higgs and the SM Higgs boson ensures decays of the  $s$  to observable particles.

In these models, the coupling of the  $Z'$  to the SM matter content allows for the production of such processes in the collisions at the LHC, while the coupling between the  $s$  and the SM ensures that it can be detected through its decay products [180]. A particularly interesting topology is shown in figure 9.8, where a  $Z'$  is produced and radiates a dark Higgs, just to decay into a pair of dark fermions,  $\chi\chi$ . In this case, the  $s$  is assumed to have the same decay branching ratio as a function of  $m(s)$  than the SM Higgs. For  $m(s) \approx m(h)$ , the expected dominant decay channel is  $s \rightarrow b\bar{b}$ , but as the mass increases new interesting channels open up. When the  $m(s)$  is large enough for it to decay to a pair of SM vector bosons, this channel dominates the decays

of the dark Higgs particle for most of the parameter space. A search for this topology has been carried out by the ATLAS experiment targeting final states with  $WW$  or  $ZZ$  pairs that decay hadronically, without significant excess observed [181].

An interesting topology is kinematically allowed when  $m(s) \geq 2m(h)$ , since it is possible to have  $s \rightarrow hh$  decays to which the analysis presented in this work could provide sensitivity. In this case, Higgs boson decays to  $b\bar{b}$  pairs would result in exactly the same final states as the ones targeted in the SUSY high-mass channel: multiple  $b$ -jets together with large  $E_T^{\text{miss}}$  from the dark fermions  $\chi$ , which do not leave any trace in the detector. The expected behaviour is also similar to the GMSB higgsino model, with low mass hypotheses resulting in soft  $b$ -jets and low values of  $E_T^{\text{miss}}$ , while high  $m(s)$  leads to boosted Higgs bosons and large  $E_T^{\text{miss}}$ . The same analysis as is designed could provide good sensitivity to this model, although one key feature of this DM model is not exploited in the current version of the search. While in the SUSY scenario each Higgs boson is produced in separated decay chains, in the DM signal both Higgs bosons are produced together and therefore it is possible to reconstruct the kinematic properties of dark Higgs. This can also lead to different correlations between the key kinematic variables used in the training of the BDT in the analysis, which are not exploited in the current version of the search but that could be included by simply retraining the BDT with this model to improve the sensitivity. The phase space region where the dark Higgs mass is so large that the decay products start to become very collimated could result in significant loss of the sensitivity provided with this search. In this case, not only the decay products of each Higgs boson could merge into a single jet, but also the two Higgs bosons could become very close in the detector and very hard to disentangle, breaking the Higgs reconstruction techniques used currently. This kinematic regime, however, seems hard to achieve since the mass of the mediator  $Z'$  would also need to be very large, and if the production cross section is large enough to achieve significant sensitivity with the LHC a dedicated analysis might be the best way to explore this region of the parameter space.



# Conclusion

This dissertation presents a search for higgsino pair production in final states with multiple  $b$ -jets and large missing transverse energy, making use of recorded proton-proton collisions at a center-of-mass energy of  $\sqrt{s} = 13$  TeV with the ATLAS detector. The data used in the analysis corresponds to the LHC data taking period of 2015-2018, which amounts to a total integrated luminosity of  $139 \text{ fb}^{-1}$ .

The analysis strategy was completely optimized from scratch compared the previous ATLAS analysis targeting the same signal model, SUSY GMSB natural scenarios with light higgsinos, focusing on high-mass signal hypotheses. It relies on the reconstruction of the intermediate Higgs bosons expected to be originated in the higgsino decay chain, for which a new approach was developed to improve the Higgs-boson identification in the boosted regimes. Even though it is optimized for scenarios with heavy higgsinos, the search also provides good sensitivity to low-mass hypotheses thanks to an optimized trigger strategy with a dedicated calibration. A robust estimate for the main background process in the kinematic regime favoured by these signal mass points, generally regions with low- $E_T^{\text{miss}}$  values where QCD multijet events dominate, is provided by means of a data-driven approach improved with machine learning techniques. The final discrimination between signal and background events is done by means of a BDT optimized for each mass hypothesis considered in the range  $200 \text{ GeV} < m(\tilde{H}) < 1500 \text{ GeV}$ . A set of analysis regions is optimized for maximum exclusion at each mass point, and an independent fit is performed with each set, in some cases combining multiple signal regions in a multi-bin approach.

No significant excess is observed in the analysis over the Standard Model prediction. The largest disagreement between data and background estimate is found in the leading signal regions for high mass hypotheses, SR-1-800, SR-1-900 and SR-1-1000, where three events are observed in contrast with the  $\approx 0.5$  events expected, corresponding to a local significance of about  $1.9\sigma$ . The 95% CL expected exclusion limits for the model with  $\text{BR}(\tilde{H} \rightarrow h\tilde{G})=100\%$  run up to higgsino mass of  $m(\tilde{H}) = 1050 \text{ GeV}$  while, due to the small excess, the actual excluded mass range is  $200 \text{ GeV} \leq m(\tilde{H}) \leq 950 \text{ GeV}$ . A complementary ATLAS analysis targets low- $E_T^{\text{miss}}$  final states and provides better sensitivity than the search discussed in this work in the range  $m(\tilde{H}) < 250 \text{ GeV}$ . Additionally, a CMS search targeting the same SUSY model has been published, excluding signal scenarios with higgsino mass below 950 GeV at 95% CL. The excluded values for the higgsino pair production cross section by the analysis presented in this thesis significantly lower than those obtained by the CMS analysis, of up to a factor 3 in the mass range  $500 \text{ GeV} < m(\tilde{H}) < 900 \text{ GeV}$ .

Besides the exclusion limits for the model with  $\text{BR}(\tilde{H} \rightarrow h\tilde{G})=100\%$ , the exclusion depth in the  $\text{BR}(\tilde{H} \rightarrow h\tilde{G}) - m(\tilde{H})$  plane obtained with this analysis is also provided. This search shows large sensitivity to scenarios with mixed decay branching ratio, being able to exclude at 95% CL models with  $\text{BR}(\tilde{H} \rightarrow h\tilde{G})$  as low as 15% for  $m(\tilde{H})=400 \text{ GeV}$ . These exclusion limits

are complemented with those of other ATLAS searches to set stringent limits on SUSY GMSB natural scenarios with light higgsinos.

The results of this search have a large impact in the SUSY electroweak program in ATLAS. This is the only search able to probe the high mass regime in GMSB SUSY scenarios with light higgsinos and large decay branching ratio into Higgs bosons, setting stringent exclusion limits that put naturalness assumptions in the MSSM in a difficult position. However, this search is interpreted with simplified models, and the sensitivity becomes worse when more complex SUSY scenarios are explored. Even though no BSM physics have been found yet, the LHC program is very long, which is expected to collect an integrated luminosity of about  $3000 \text{ fb}^{-1}$  at the end of its operation. In spite of their typically low production cross section, the large amount of data expected will allow to scan thoroughly the available phase space of SUSY electroweak models, with the searches performed so far paving the way for efficient and innovative future analyses.

## Appendix A

# $E_T^{\text{miss}}$ trigger calibration

The recorded events used in the analysis presented in this work are selected using the lowest unprescaled  $E_T^{\text{miss}}$  trigger, i.e. the selection on  $E_T^{\text{miss}}$  at the trigger level with the lowest threshold that is not prescaled to reduce the rate of accepted events. The algorithms used to compute  $E_T^{\text{miss}}$  at the trigger level and the thresholds applied differ from year to year during Run 2, as described in chapter 7. For all of them, in regions with  $E_T^{\text{miss}} > 230$  GeV the efficiency of these selections reaches a plateau at 100%. To increase sensitivity to lower mass points in this analysis, however, different  $E_T^{\text{miss}}$  cuts are explored, and the final preselection level used with  $E_T^{\text{miss}} > 150$  GeV means that some of the selected events are in the turn-on region of the trigger efficiency. To assess differences in the trigger performance between data and the MC samples used to model the background processes and the signal, trigger efficiency curves are computed using events selected with single-lepton selections at the trigger level as reference, which allows to study the behaviour of the trigger selection on  $E_T^{\text{miss}}$  without bias. The selection used for these studies is:

- at least 4 PFlow jets with  $p_T > 25$  GeV,
- at least 2  $b$ -jets with  $p_T > 25$  GeV,
- exactly 1 muon with  $p_T > 20$  GeV,
- events selected with single muon triggers as reference.

Events selected with a single-electron trigger requirement were also studied, but the final calibration was performed using exclusively the muon channel. The main reason is that muons leave little energy in the calorimeter, which is the only information used to compute  $E_T^{\text{miss}}$  at the trigger level, and therefore they provide a probe for missing energy that can be contrasted with muon reconstruction information in the event. In events that include jets and electrons exclusively  $E_T^{\text{miss}}$  is more likely to arise from energy mismeasurement and detector resolution effects, which led to a much less stable calibration than the muon channel.

The single muon triggers used as reference are the lowest unprescaled ones for each data taking period and are collected below:

- 2015:
  - HLT\_mu20\_iloose\_L1MU15
  - HLT\_mu40

- HLT\_mu60\_0eta105\_msonly
- 2016:
  - A - B:
    - \* HLT\_mu24\_iloose
    - \* HLT\_mu24\_ivarloose
    - \* HLT\_mu40
    - \* HLT\_mu50
  - B - D3:
    - \* HLT\_mu24\_ivarmedium
    - \* HLT\_mu24\_imedium
    - \* HLT\_mu50
  - D4 - E:
    - \* HLT\_mu24\_ivarmedium
    - \* HLT\_mu26\_ivarmedium
    - \* HLT\_mu50
  - F -:
    - \* HLT\_mu26\_ivarmedium
    - \* HLT\_mu50
- 2017:
  - HLT\_mu26\_ivarmedium
  - HLT\_mu50
  - HLT\_mu60\_0eta105\_msonly
- 2018:
  - HLT\_mu26\_ivarmedium
  - HLT\_mu50
  - HLT\_mu60\_0eta105\_msonly

In order to correct any differences in the trigger performance between data and MC simulated samples, dedicated scale factors (SF) are derived and applied to the simulated events. This  $E_T^{\text{miss}}$  trigger calibration is derived by comparing the trigger efficiency turn-on curve between data and  $t\bar{t}$  samples. Due to the different pileup conditions for each year during Run 2, the simulated samples are generated in three *campaigns*: mc16a, which reproduces the pileup conditions of the data taking periods 2015 and 2016, mc16d for the year 2017 and mc16e for 2018. To account for differences in the trigger performance in each period, the calibration is studied separately in each of these periods. For each MC simulation campaign the scale factors are derived inclusively for all  $E_T^{\text{miss}}$  trigger selections used (one SF for mc16a, one for mc16d and one for m16e) instead of being derived independently for each of the lowest unprescaled  $E_T^{\text{miss}}$  triggers used. This approach is followed to have access to enough statistics by using a wider data taking period and to avoid fluctuations in the SF.

For each campaign, in order to correctly model the trigger calibration in different energy ranges, the events are divided in six  $H_T$  bins and the scale factors are derived independently in

each of them. In figure A.1 the trigger efficiency turn-on curves are presented for each  $H_T$  bin and data taking period/campaign for data and  $t\bar{t}$  simulated events, respectively, as a function of the offline  $E_T^{\text{miss}}$  after being corrected by the  $p_T$  of the muon in the event, such that it resembles the values computed at the trigger level.

In figures A.2 - A.4 the trigger turn-on curves from data are compared to  $t\bar{t}$  MC individually for each  $H_T$  bin, where the bottom panel in each plot is the actual scale factor obtained as the ratio between the trigger efficiency in data and the trigger efficiency in simulation. For the preselection cut of  $E_T^{\text{miss}} > 150$  GeV the scale factor applied for most  $H_T$  bins and MC campaigns range between 0.9 and 1.

To smooth any fluctuations in the scale factors shown in figures A.2 - A.4 caused by low statistics, the turn-on curves in data and  $t\bar{t}$  simulated events are fitted independently to a generalized sigmoid function:

$$f(x) = \frac{p_2}{[1 + (2^{p_3} - 1) e^{-p_0(x-p_1)}]^{\frac{1}{p_3}}}, \quad (\text{A.1})$$

where  $p_0$ ,  $p_1$ ,  $p_2$  and  $p_3$  are the free parameters to be fitted. The final parametrization of the scale factors used as an event weight in the analysis is the ratio between the fitted curve in data and the fitted curve in  $t\bar{t}$  simulated events. The comparison of the bin-to-bin computation of the scale factor and its parametrization can be found in figures A.5 - A.7.

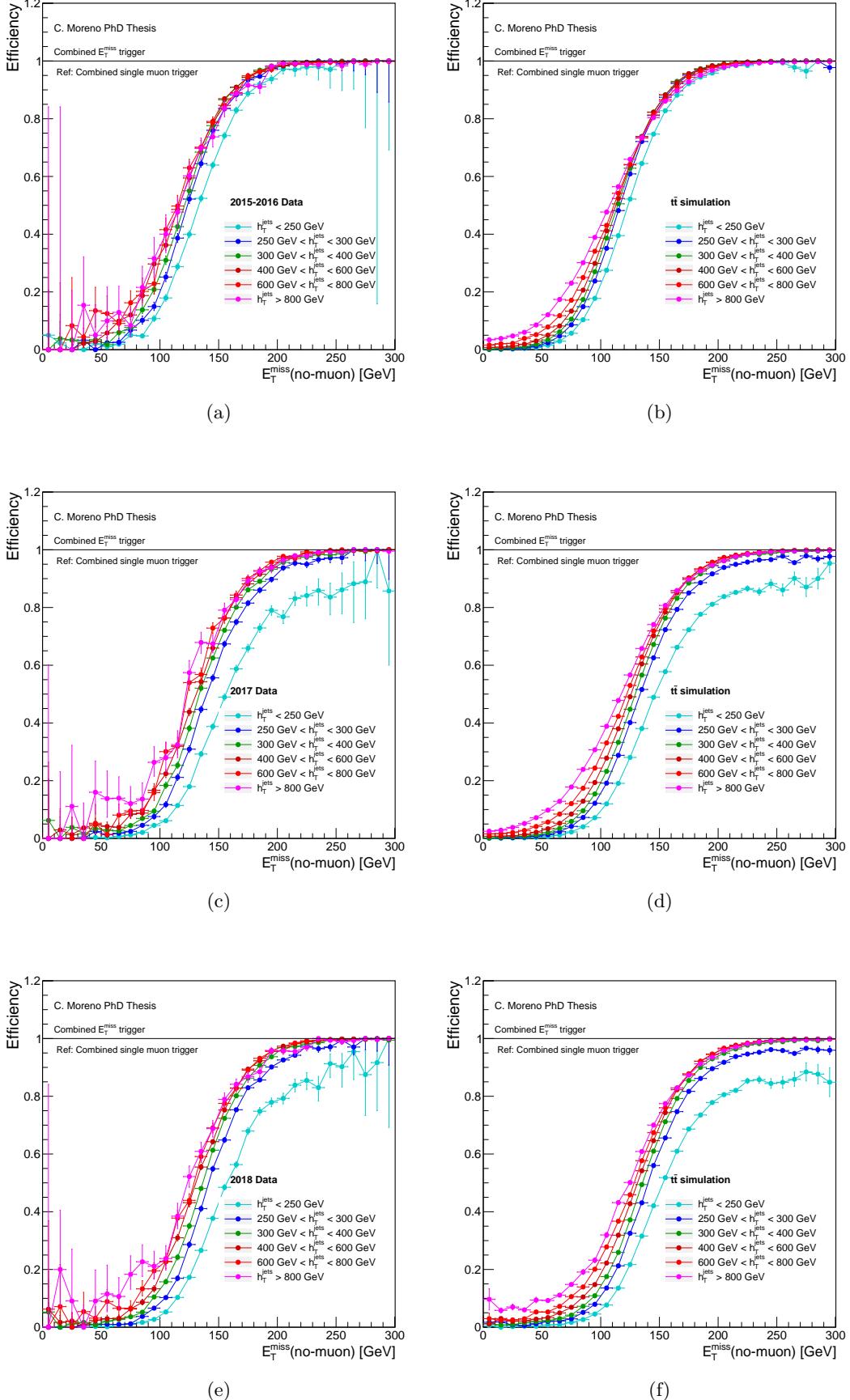


Figure A.1: Inclusive efficiency of the  $E_T^{\text{miss}}$  selection at the trigger level in the  $1\mu$  channel, using single-muon trigger selections as references. The efficiency is shown in six  $H_T$  bins in data (left column) for the data taking periods (a) 2015 and 2016, (c) 2017 and (e) 2018, and  $t\bar{t}$  simulated events (right column) for the campaigns (b) mc16a, (d) m16d and (f) mc16e.

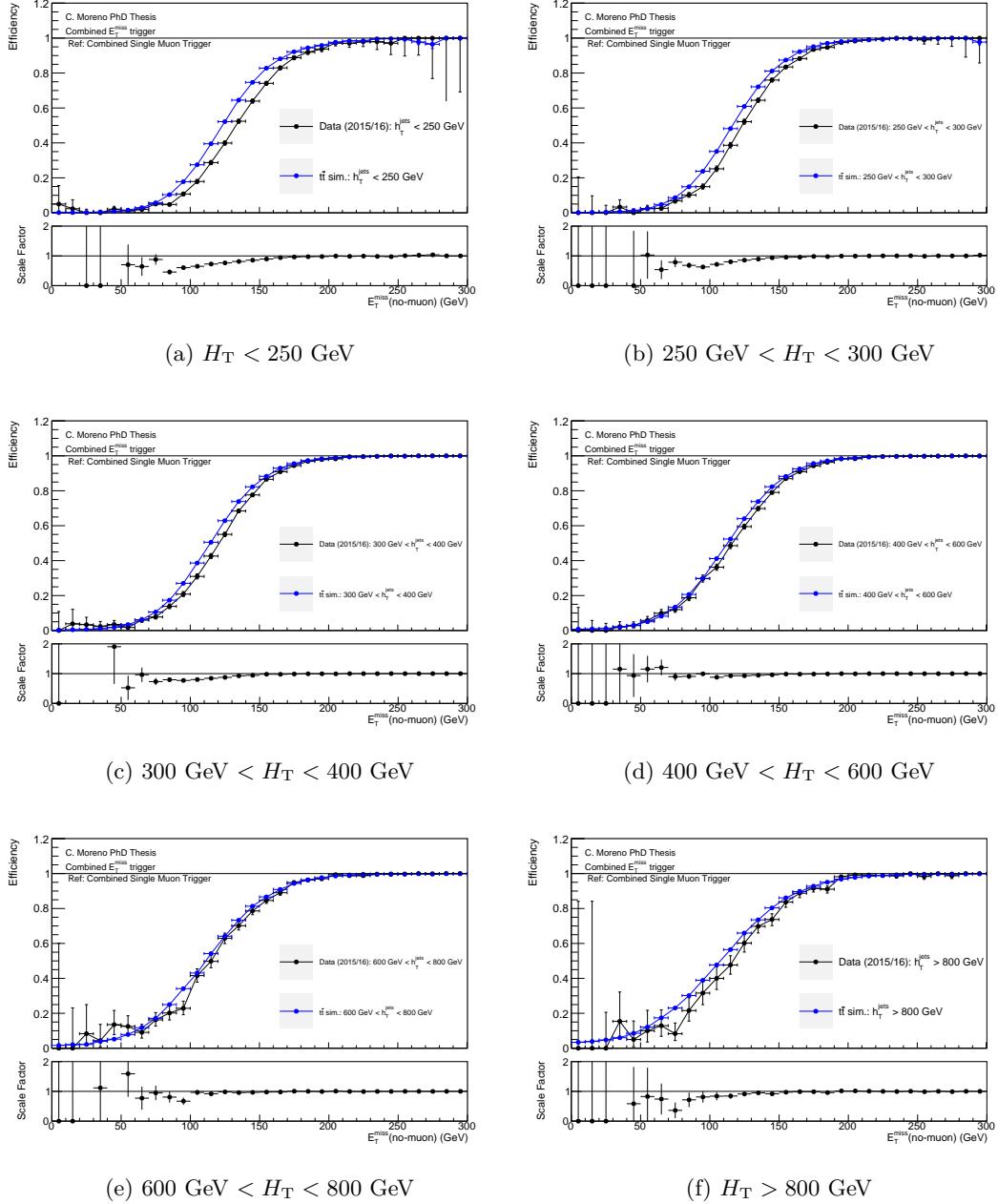


Figure A.2: Scale factor in each  $H_T$  bin for the data taking period of 2015 and 2016, corresponding to MC campaign mc16a.

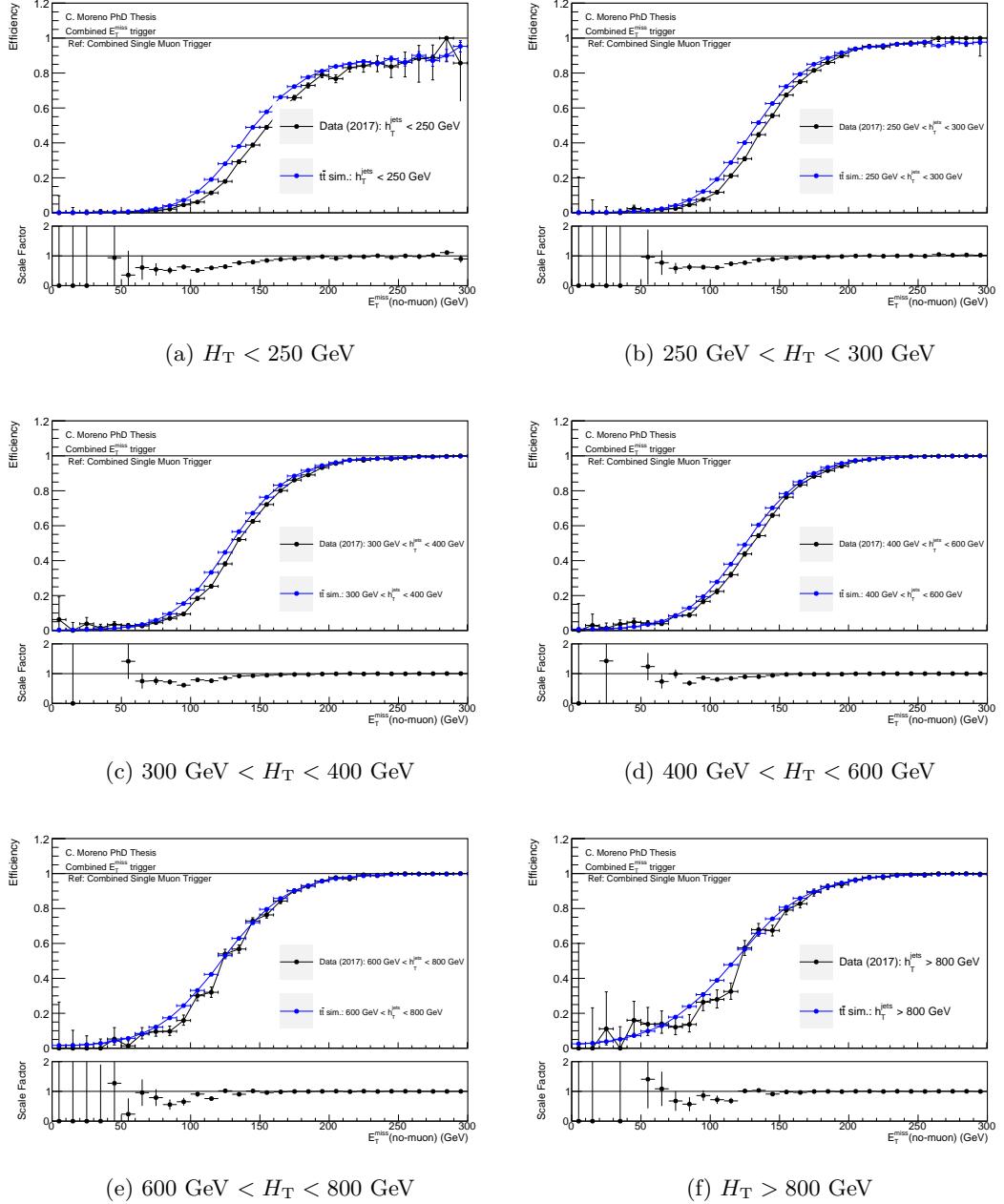


Figure A.3: Scale factor in each  $H_T$  bin for the data taking period of 2017, corresponding to MC campaign mc16d.

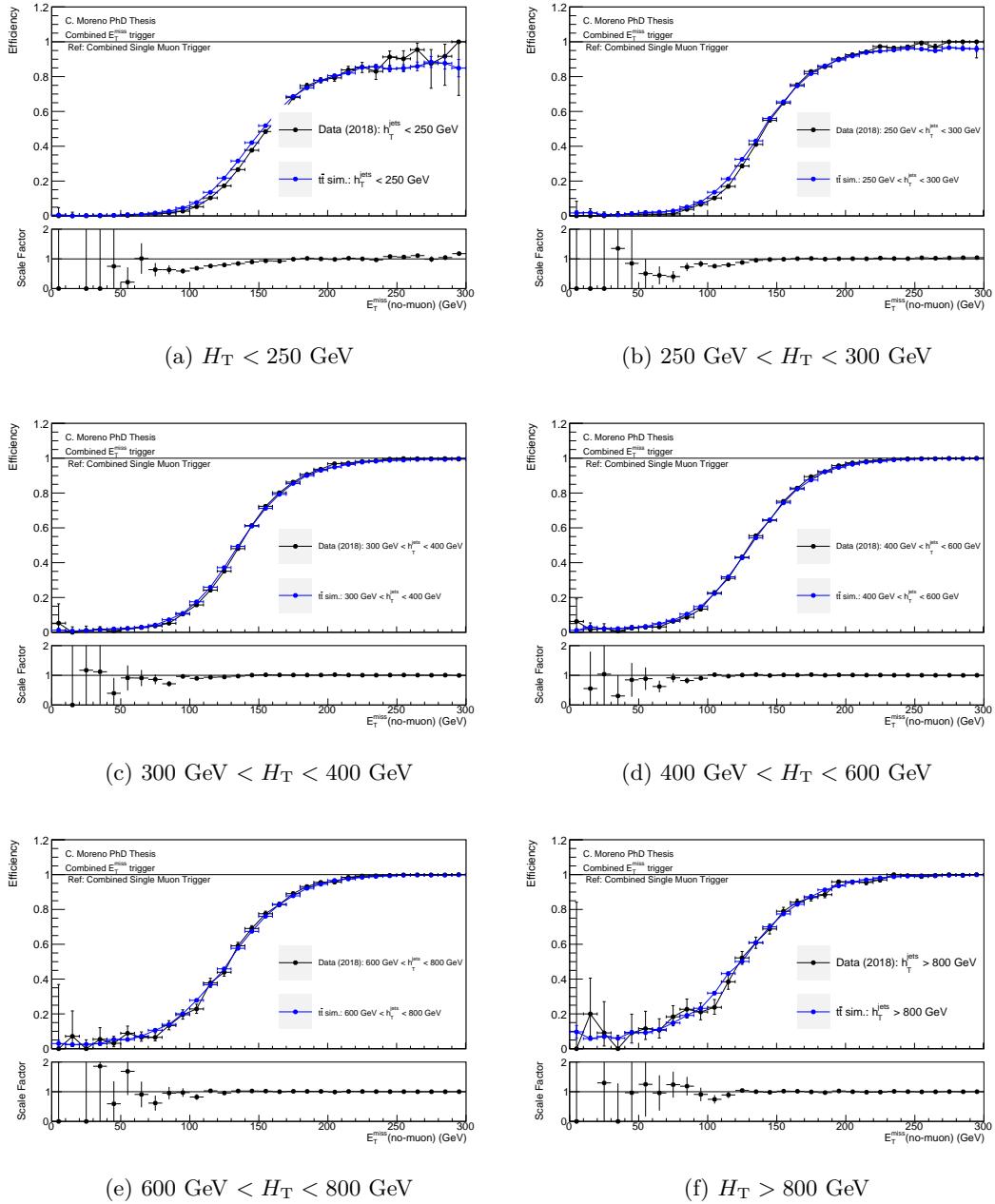


Figure A.4: Scale factor in each  $H_T$  bin for the data taking period of 2018, corresponding to MC campaign mc16e.

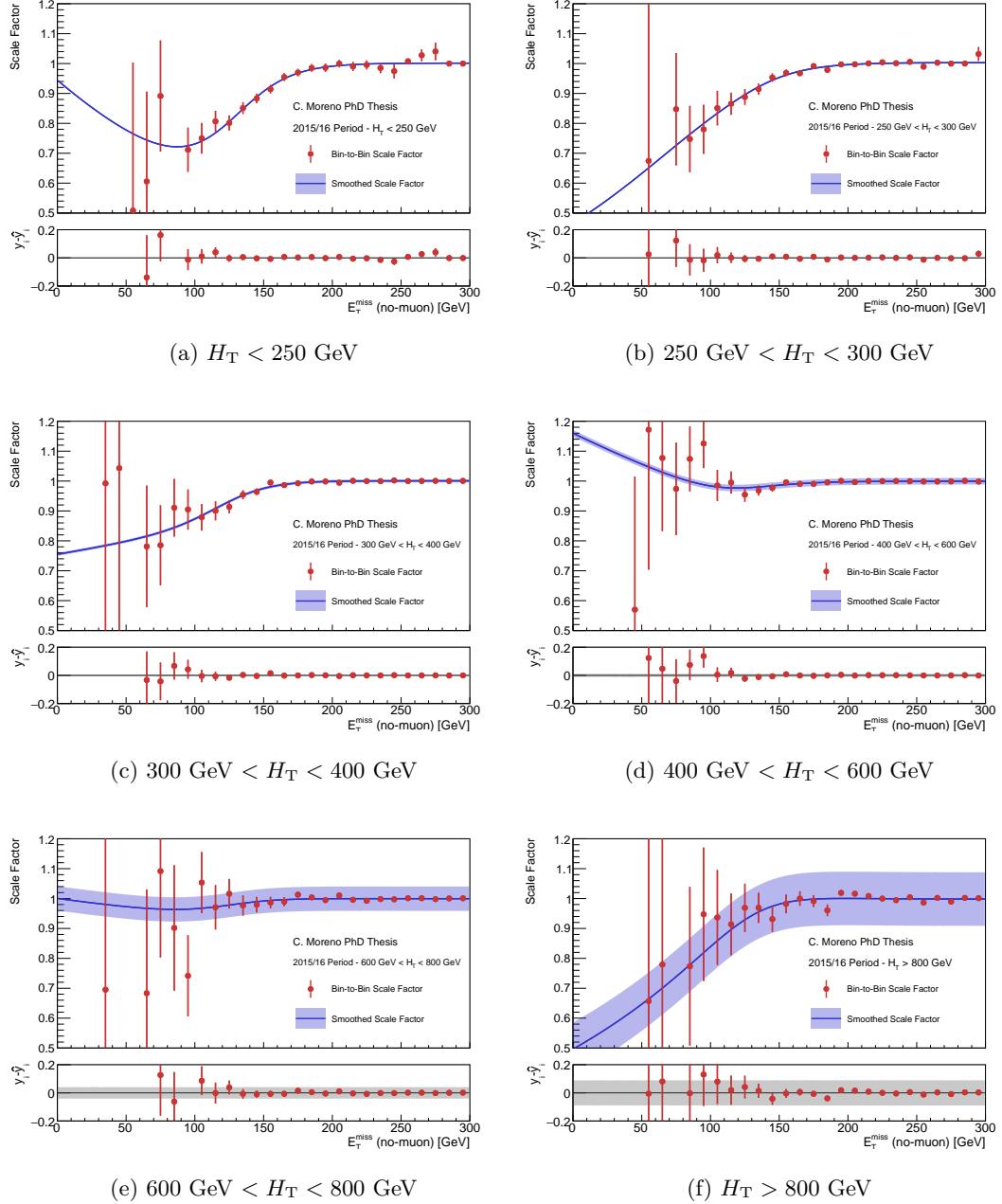


Figure A.5: Scale factor and its parametrization in each  $H_T$  bin for the data taking period of 2015 and 2016, corresponding to MC campaign mc16a.

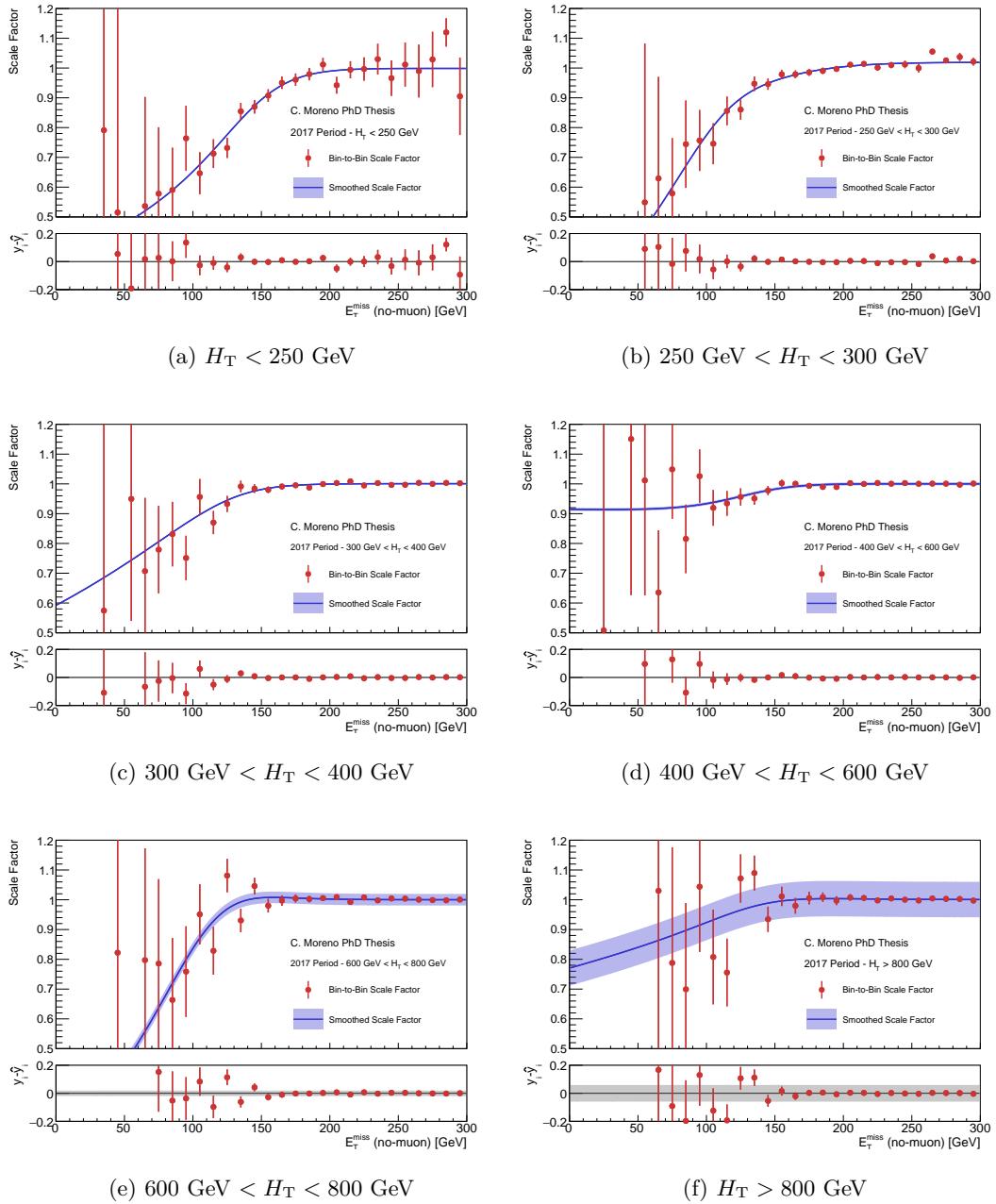


Figure A.6: Scale factor and its parametrization in each  $H_T$  bin for the data taking period of 2017, corresponding to MC campaign mc16d.

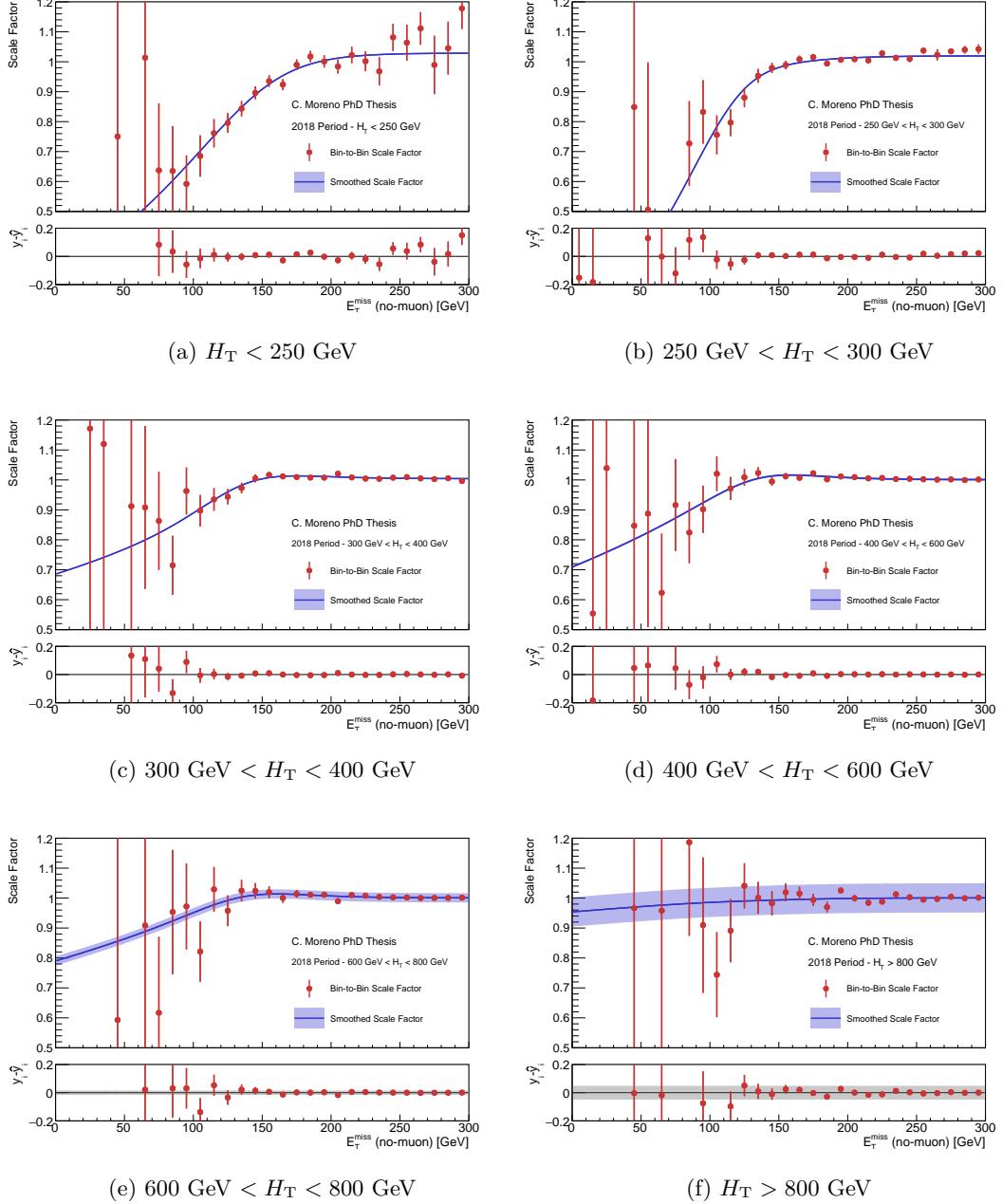


Figure A.7: Scale factor and its parametrization in each  $H_T$  bin for the data taking period of 2018, corresponding to MC campaign mc16e.

## Appendix B

# Definitions of all analysis regions

In this appendix the definitions and background composition of all the regions defined in the analysis discussed in this work are collected.

### B.1 Region definitions

The definitions of all the signal, validation and control regions for all of the signal mass hypotheses explored in the analysis are collected in tables B.1–B.10. The regions are defined according to the methodology discussed in section 7, and all extra cuts shown in the tables are applied on top of the selection Presel-0L, except for the validation and control regions for  $Z+jets$  processes, that are defined from the region Presel-ZCR. There are no regions defined for the signal mass hypotheses above 1300 GeV as it was not possible to find any signal regions with significance above 1. For these points the regions defined for higgsinos with mass 1100 GeV are used. For all control and validation regions, in addition to the selections specified in the tables there is also a requirement on the mass of the leading reconstructed Higgs boson,  $100 \text{ GeV} < m(\tilde{H}) < 150 \text{ GeV}$ , to reduce signal contamination from other higgsino mass hypotheses.

<b>Region</b>	N. b-jets	$m_{T,\min}^{b\text{-jets}}$	BDT score	$N^{\text{sig}}$	$N^{\text{bkg}}$
SR-1-250	$\geq 3$	-	(0.9997, 1]	27.9031	1.56719
SR-2-250	$\geq 3$	-	(0.9994, 0.9997]	23.6513	5.06638
SR-3-250	$\geq 3$	-	(0.99846, 0.9994]	47.8136	21.4034
SR-4-250	$\geq 3$	-	(0.99786, 0.99846]	18.9394	14.4228
VR-tt-250	$\geq 3$	< 200 GeV	(0.969239, 0.99786]	80.3503	401.223
VR-st-250	$\geq 3$	$\geq 200$ GeV	(0.98232, 0.99786]	3.56394	24.9764
CR-tt3b-250	= 3	< 200 GeV	(0.963639, 0.969239]	7.96057	99.8456
CR-tt4b-250	$\geq 4$	< 200 GeV	(0.955179, 0.969239]	9.24107	99.8399
CR-st-250	$\geq 3$	$\geq 200$ GeV	(0.940219, 0.98232]	5.64447	99.8735
VR-Z-250	$\geq 3$	-	(0.538771, 1]	0	24.9861
CR-Z-250	$\geq 3$	-	(0.0961619, 0.538771]	0	99.9791

Table B.1: Definition cuts for all the analysis regions for  $m(\tilde{H}) = 250$  GeV. The selection Presel-0L is applied to all except for CR-Z and VR-Z, that use the selection Presel-ZCR. The expected numbers of signal and background events in each region are collected in the last two columns.

<b>Region</b>	N. b-jets	$m_{T,\min}^{b\text{-jets}}$	BDT score	$N^{\text{sig}}$	$N^{\text{bkg}}$
SR-1-300	$\geq 3$	-	(0.99956, 1]	16.3282	1.99846
SR-2-300	$\geq 3$	-	(0.99904, 0.99956]	16.9777	4.83311
SR-3-300	$\geq 3$	-	(0.99826, 0.99904]	18.2452	9.80585
SR-4-300	$\geq 3$	-	(0.99738, 0.99826]	15.7593	14.7624
VR-tt-300	$\geq 3$	< 200 GeV	(0.97996, 0.99738]	18.0802	90.3731
VR-st-300	$\geq 3$	$\geq 200$ GeV	(0.98548, 0.99738]	4.76103	24.7499
CR-tt3b-300	= 3	< 200 GeV	(0.965719, 0.97996]	5.97029	100.09
CR-tt4b-300	$\geq 4$	< 200 GeV	(0.952459, 0.97996]	8.7066	100.005
CR-st-300	$\geq 3$	$\geq 200$ GeV	(0.959699, 0.98548]	7.09248	99.7868
VR-Z-300	$\geq 3$	-	(0.282506, 1]	0	24.9993
CR-Z-300	$\geq 3$	-	(0.0400408, 0.282506]	0	99.9609

Table B.2: Definition cuts for all the analysis regions for  $m(\tilde{H}) = 300$  GeV. The selection Presel-0L is applied to all except for CR-Z and VR-Z, that use the selection Presel-ZCR. The expected numbers of signal and background events in each region are collected in the last two columns.

Region	N. b-jets	$m_{T,\min}^{b\text{-jets}}$	BDT score	$N^{\text{sig}}$	$N^{\text{bkg}}$
SR-1-400	$\geq 3$	-	(0.99974, 1]	16.2405	1.27451
SR-2-400	$\geq 3$	-	(0.99936, 0.99974]	14.6679	4.33183
SR-3-400	$\geq 3$	-	(0.999, 0.99936]	9.47731	6.16063
SR-4-400	$\geq 3$	-	(0.99798, 0.999]	15.407	16.9254
VR-tt-400	$\geq 3$	< 200 GeV	(0.98858, 0.99798]	3.91262	25.0634
VR-st-400	$\geq 3$	$\geq 200$ GeV	(0.98814, 0.99798]	10.98	54.6342
CR-tt3b-400	= 3	< 200 GeV	(0.968759, 0.98858]	4.9225	99.8348
CR-tt4b-400	$\geq 4$	< 200 GeV	(0.935299, 0.98858]	3.65794	99.9092
CR-st-400	$\geq 3$	$\geq 200$ GeV	(0.969879, 0.98814]	6.58229	99.9995
VR-Z-400	$\geq 3$	-	(0.139823, 1]	0	24.9958
CR-Z-400	$\geq 3$	-	(0.0177804, 0.139823]	0	100.014

Table B.3: Definition cuts for all the analysis regions for  $m(\tilde{H}) = 400$  GeV. The selection Presel-OL is applied to all except for CR-Z and VR-Z, that use the selection Presel-ZCR. The expected numbers of signal and background events in each region are collected in the last two columns.

Region	N. b-jets	$m_{T,\min}^{b\text{-jets}}$	BDT score	$N^{\text{sig}}$	$N^{\text{bkg}}$
SR-1-500	$\geq 3$	-	(0.9997, 1]	8.75667	0.618144
SR-2-500	$\geq 3$	-	(0.9994, 0.9997]	5.14364	1.61243
SR-3-500	$\geq 3$	-	(0.99878, 0.9994]	6.41886	4.84762
VR-tt-500	$\geq 3$	< 200 GeV	(0.974519, 0.99878]	1.54755	24.9946
VR-st-500	$\geq 3$	$\geq 200$ GeV	(0.98888, 0.99878]	4.20567	25.0773
CR-tt3b-500	= 3	< 200 GeV	(0.924498, 0.974519]	2.20954	99.9869
CR-tt4b-500	$\geq 4$	< 200 GeV	(0.787956, 0.974519]	1.15961	99.9717
CR-st-500	$\geq 3$	$\geq 200$ GeV	(0.954699, 0.98888]	5.46219	99.9548
VR-Z-500	$\geq 3$	-	(0.0454809, 1]	0	24.9997
CR-Z-500	$\geq 3$	-	(0.00534011, 0.0454809]	0	99.7545

Table B.4: Definition cuts for all the analysis regions for  $m(\tilde{H}) = 500$  GeV. The selection Presel-OL is applied to all except for CR-Z and VR-Z, that use the selection Presel-ZCR. The expected numbers of signal and background events in each region are collected in the last two columns.

Region	N. b-jets	$m_{T,\min}^{b\text{-jets}}$	BDT score	$N^{\text{sig}}$	$N^{\text{bkg}}$
SR-1-600	$\geq 3$	-	(0.99974, 1]	9.65155	0.621246
SR-2-600	$\geq 3$	-	(0.99944, 0.99974]	6.20185	1.94919
SR-3-600	$\geq 3$	-	(0.99842, 0.99944]	9.25954	8.57574
VR-tt-600	$\geq 3$	< 200 GeV	(0.973659, 0.99842]	1.0921	24.8371
VR-st-600	$\geq 3$	$\geq 200$ GeV	(0.98982, 0.99842]	4.21624	24.9691
CR-tt3b-600	= 3	< 200 GeV	(0.917818, 0.973659]	1.1727	99.9743
CR-tt4b-600	$\geq 4$	< 200 GeV	(0.739655, 0.973659]	0.647523	99.9917
CR-st-600	$\geq 3$	$\geq 200$ GeV	(0.949279, 0.98982]	4.38007	99.9308
VR-Z-600	$\geq 3$	-	(0.0252005, 1]	0	25.0057
CR-Z-600	$\geq 3$	-	(0.00302006, 0.0252005]	0	99.8073

Table B.5: Definition cuts for all the analysis regions for  $m(\tilde{H}) = 600$  GeV. The selection Presel-OL is applied to all except for CR-Z and VR-Z, that use the selection Presel-ZCR. The expected numbers of signal and background events in each region are collected in the last two columns.

Region	N. b-jets	$m_{T,\min}^{b\text{-jets}}$	BDT score	$N^{\text{sig}}$	$N^{\text{bkg}}$
SR-1-700	$\geq 3$	-	(0.99966, 1]	9.85744	0.661642
SR-2-700	$\geq 3$	-	(0.99928, 0.99966]	4.8243	1.98883
SR-3-700	$\geq 3$	-	(0.99832, 0.99928]	6.02274	5.04849
VR-tt-700	$\geq 3$	< 200 GeV	(0.953299, 0.99832]	0.704959	25.0545
VR-st-700	$\geq 3$	$\geq 200$ GeV	(0.98378, 0.99832]	3.95301	24.9855
CR-tt3b-700	= 3	< 200 GeV	(0.846257, 0.953299]	0.759155	99.9418
CR-tt4b-700	$\geq 4$	< 200 GeV	(0.49627, 0.953299]	0.39706	99.9297
CR-st-700	$\geq 3$	$\geq 200$ GeV	(0.899218, 0.98378]	3.0526	99.9981
VR-Z-700	$\geq 3$	-	(0.0110402, 1]	0	24.8456
CR-Z-700	$\geq 3$	-	(0.00112002, 0.0110402]	0	99.5064

Table B.6: Definition cuts for all the analysis regions for  $m(\tilde{H}) = 700$  GeV. The selection Presel-0L is applied to all except for CR-Z and VR-Z, that use the selection Presel-ZCR. The expected numbers of signal and background events in each region are collected in the last two columns.

Region	N. b-jets	$m_{T,\min}^{b\text{-jets}}$	BDT score	$N^{\text{sig}}$	$N^{\text{bkg}}$
SR-1-800	$\geq 3$	-	(0.99948, 1]	9.71687	0.606678
SR-2-800	$\geq 3$	-	(0.99852, 0.99948]	5.48814	2.90483
VR-tt-800	$\geq 3$	< 200 GeV	(0.905638, 0.99852]	0.519291	24.9774
VR-st-800	$\geq 3$	$\geq 200$ GeV	(0.9761, 0.99852]	3.89144	24.9665
CR-tt3b-800	= 3	< 200 GeV	(0.694074, 0.905638]	0.488781	100.028
CR-tt4b-800	$\geq 4$	< 200 GeV	(0.250605, 0.905638]	0.307182	99.9036
CR-st-800	$\geq 3$	$\geq 200$ GeV	(0.806536, 0.9761]	1.95232	99.9915
VR-Z-800	$\geq 3$	-	(0.00418008, 1]	0	24.9708
CR-Z-800	$\geq 3$	-	(0.00050001, 0.00418008]	0	98.243

Table B.7: Definition cuts for all the analysis regions for  $m(\tilde{H}) = 800$  GeV. The selection Presel-0L is applied to all except for CR-Z and VR-Z, that use the selection Presel-ZCR. The expected numbers of signal and background events in each region are collected in the last two columns.

Region	N. b-jets	$m_{T,\min}^{b\text{-jets}}$	BDT score	$N^{\text{sig}}$	$N^{\text{bkg}}$
SR-1-900	$\geq 3$	-	(0.99924, 1]	8.96145	0.559449
SR-2-900	$\geq 3$	-	(0.99804, 0.99924]	3.96884	2.15716
VR-tt-900	$\geq 3$	< 200 GeV	(0.827217, 0.99804]	0.433946	24.5465
VR-st-900	$\geq 3$	$\geq 200$ GeV	(0.957619, 0.99804]	2.93679	24.883
CR-tt3b-900	= 3	< 200 GeV	(0.50943, 0.827217]	0.299463	99.9899
CR-tt4b-900	$\geq 4$	< 200 GeV	(0.120882, 0.827217]	0.160532	99.9908
CR-st-900	$\geq 3$	$\geq 200$ GeV	(0.624412, 0.957619]	1.25459	99.992
VR-Z-900	$\geq 3$	-	(0.00174003, 1]	0	24.943
CR-Z-900	$\geq 3$	-	(0.000220004, 0.00174003]	0	100.414

Table B.8: Definition cuts for all the analysis regions for  $m(\tilde{H}) = 900$  GeV. The selection Presel-0L is applied to all except for CR-Z and VR-Z, that use the selection Presel-ZCR. The expected numbers of signal and background events in each region are collected in the last two columns.

Region	N. b-jets	$m_{T,\min}^{b\text{-jets}}$	BDT score	$N^{\text{sig}}$	$N^{\text{bkg}}$
SR-1-1000	$\geq 3$	-	(0.9991, 1]	4.75919	0.554438
VR-tt-1000	$\geq 3$	< 200 GeV	(0.721194, 0.9991]	0.236075	24.9972
VR-st-1000	$\geq 3$	$\geq 200$ GeV	(0.922958, 0.9991]	1.77659	24.7843
CR-tt3b-1000	= 3	< 200 GeV	(0.331647, 0.721194]	0.145877	99.9196
CR-tt4b-1000	$\geq 4$	< 200 GeV	(0.0630013, 0.721194]	0.0657878	100.001
CR-st-1000	$\geq 3$	$\geq 200$ GeV	(0.332187, 0.922958]	0.531473	99.9527
VR-Z-1000	$\geq 3$	-	(0.000820016, 1]	0	23.5204
CR-Z-1000	$\geq 3$	-	(0.000100002, 0.000820016]	0	98.7919

Table B.9: Definition cuts for all the analysis regions for  $m(\tilde{H}) = 1000$  GeV. The selection Presel-0L is applied to all except for CR-Z and VR-Z, that use the selection Presel-ZCR. The expected numbers of signal and background events in each region are collected in the last two columns.

Region	N. b-jets	$m_{T,\min}^{b\text{-jets}}$	BDT score	$N^{\text{sig}}$	$N^{\text{bkg}}$
SR-1-1100	$\geq 3$	-	(0.99912, 1]	2.68503	0.503031
VR-tt-1100	$\geq 3$	< 200 GeV	(0.702774, 0.99912]	0.125126	24.9547
VR-st-1100	$\geq 3$	$\geq 200$ GeV	(0.911338, 0.99912]	1.01447	24.8116
CR-tt3b-1100	= 3	< 200 GeV	(0.306246, 0.702774]	0.0763907	99.9529
CR-tt4b-1100	$\geq 4$	< 200 GeV	(0.0536611, 0.702774]	0.0194561	99.9503
CR-st-1100	$\geq 3$	$\geq 200$ GeV	(0.295986, 0.911338]	0.257944	99.9052
VR-Z-1100	$\geq 3$	-	(0.000720014, 1]	0	22.7746
CR-Z-1100	$\geq 3$	-	(8.00016e-05, 0.000720014]	0	95.8699

Table B.10: Definition cuts for all the analysis regions for  $m(\tilde{H}) = 1100$  GeV. The selection Presel-0L is applied to all except for CR-Z and VR-Z, that use the selection Presel-ZCR. The expected numbers of signal and background events in each region are collected in the last two columns.

## B.2 Background composition

In figures B.1–B.6 the background composition for all analysis regions is shown, together with the expected number of signal events in each of them.

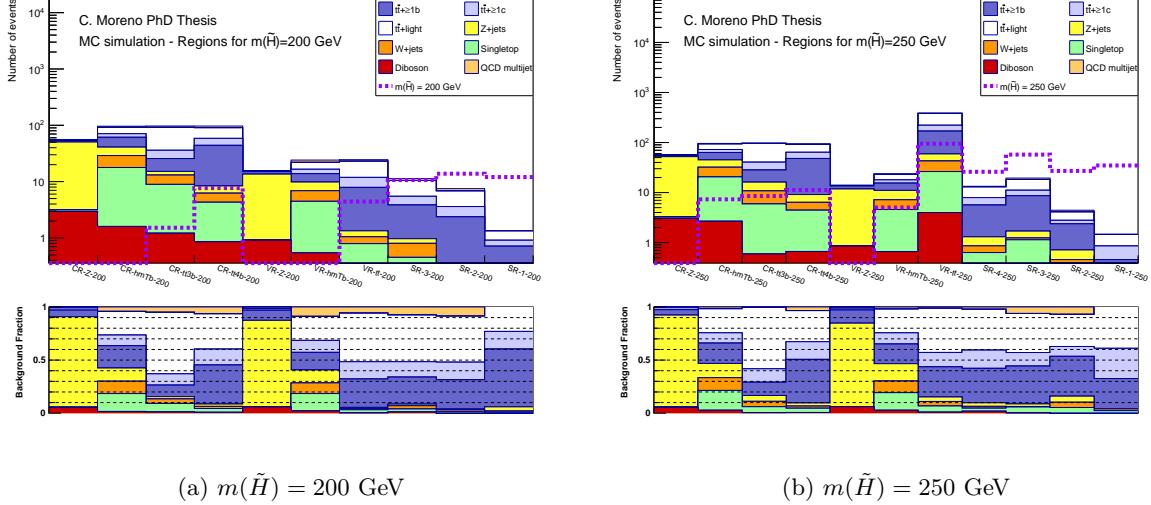
(a)  $m(\tilde{H}) = 200 \text{ GeV}$ (b)  $m(\tilde{H}) = 250 \text{ GeV}$ 

Figure B.1: Background composition in all regions defined for the higgsino mass hypotheses (a) 200 GeV and (b) 250 GeV. The purple dashed line represents the contribution from signal events in each bin corresponding to the targeted higgsino mass hypothesis, which is normalized to the previously excluded cross section as described in section 7.9.3.

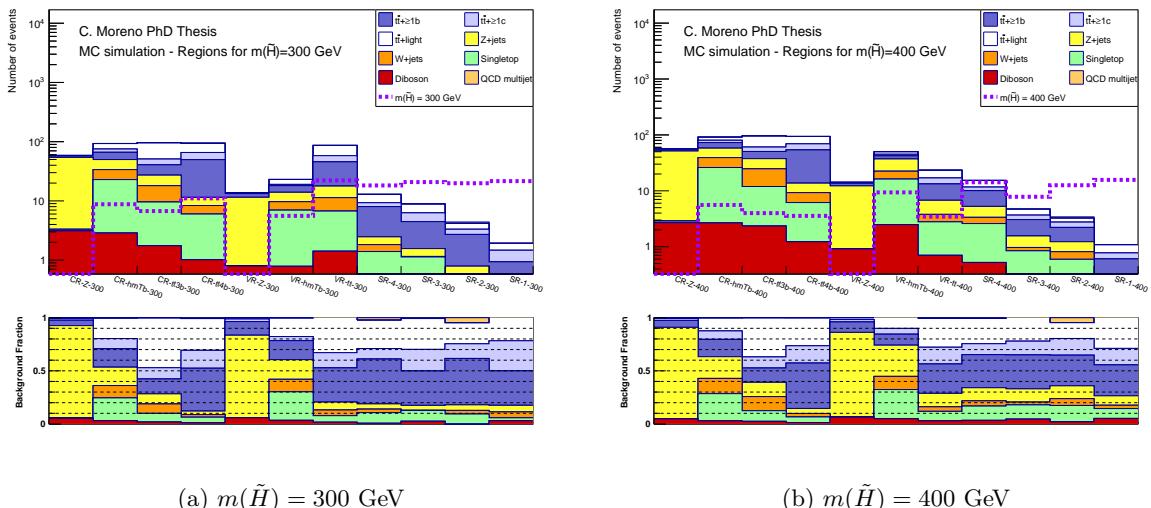
(a)  $m(\tilde{H}) = 300 \text{ GeV}$ (b)  $m(\tilde{H}) = 400 \text{ GeV}$ 

Figure B.2: Background composition in all regions defined for the higgsino mass hypotheses (a) 300 GeV and (b) 400 GeV. The purple dashed line represents the contribution from signal events in each bin corresponding to the targeted higgsino mass hypothesis, which is normalized to the previously excluded cross section as described in section 7.9.3.

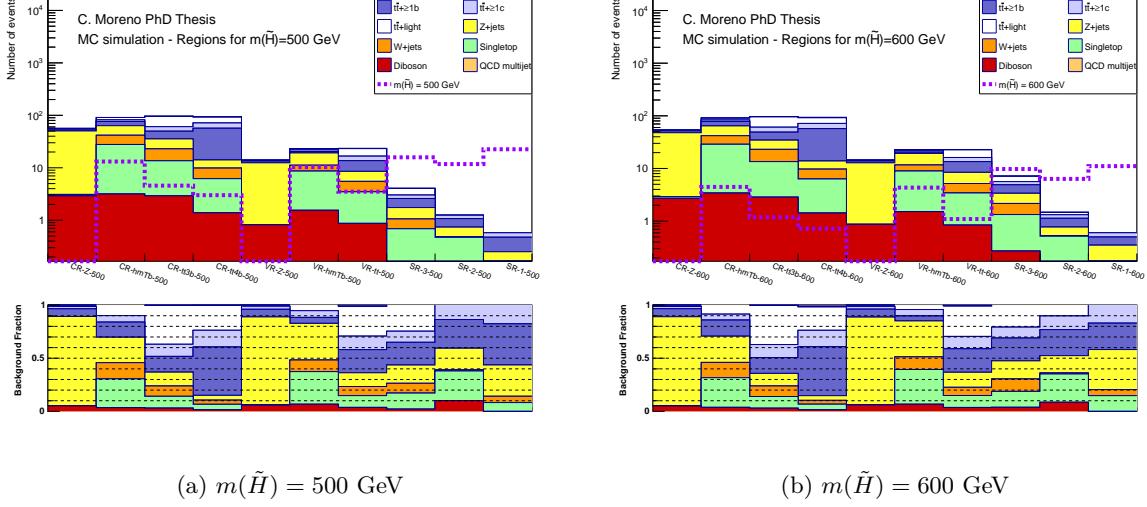


Figure B.3: Background composition in all regions defined for the higgsino mass hypotheses (a) 500 GeV and (b) 600 GeV. The purple dashed line represents the contribution from signal events in each bin corresponding to the targeted higgsino mass hypothesis, which is normalized to the previously excluded cross section as described in section 7.9.3.

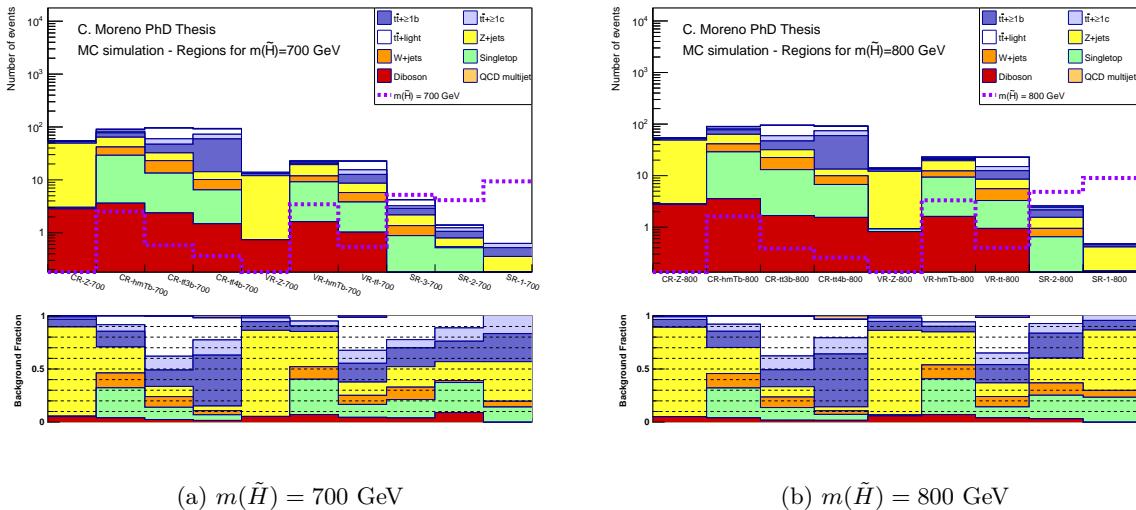


Figure B.4: Background composition in all regions defined for the higgsino mass hypotheses (a) 700 GeV and (b) 800 GeV. The purple dashed line represents the contribution from signal events in each bin corresponding to the targeted higgsino mass hypothesis, which is normalized to the previously excluded cross section as described in section 7.9.3.

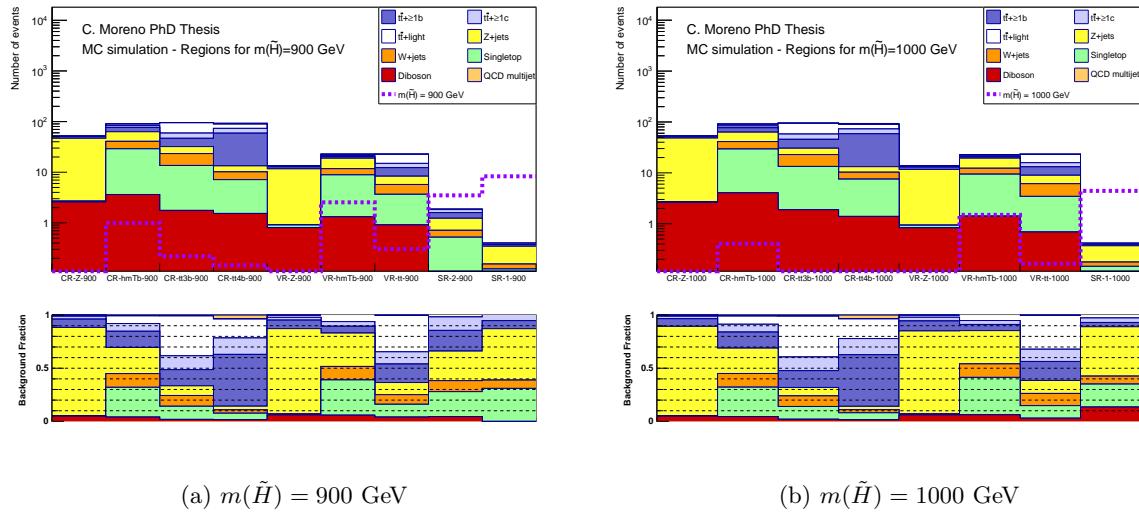


Figure B.5: Background composition in all regions defined for the higgsino mass hypotheses (a) 900 GeV and (b) 1000 GeV. The purple dashed line represents the contribution from signal events in each bin corresponding to the targeted higgsino mass hypothesis, which is normalized to the previously excluded cross section as described in section 7.9.3.

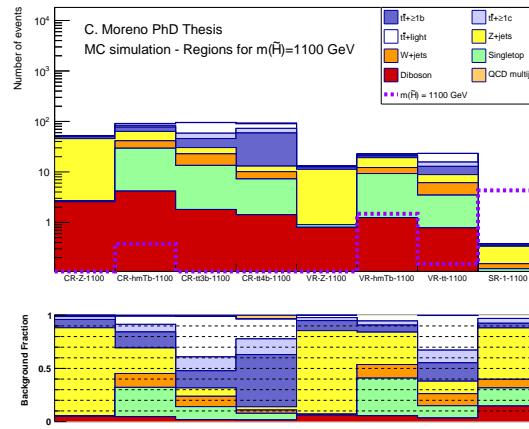


Figure B.6: Background composition in all regions defined for the higgsino mass hypothesis 1100 GeV. The purple dashed line represents the contribution from signal events in each bin corresponding to the targeted higgsino mass hypothesis, which is normalized to the previously excluded cross section as described in section 7.9.3.

### B.3 Post-fit data comparison in all analysis regions

In the following, figures where the data is compared to the post-fit background prediction in all control, validation and signal regions in the analysis is provided. Figures B.7–B.17 show the agreement for the regions corresponding to all higgsino mass hypotheses considered in the analysis, where the significance estimated as the difference between observed and predicted events divided by the total uncertainty is shown in the bottom panel. As discussed in chapter 8, the post-fit agreement is very good in general, with a small excess in the SRs corresponding to mass hypotheses above 800 GeV.

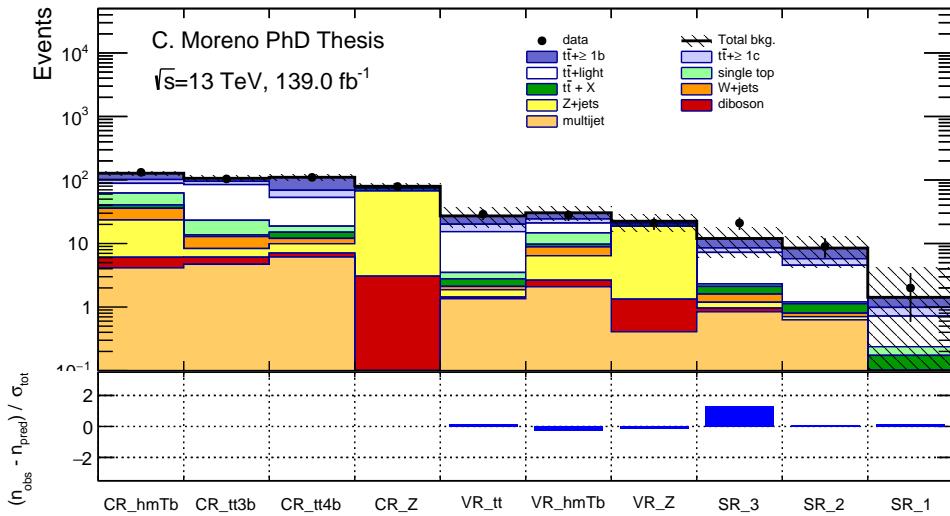


Figure B.7: Comparison between data and post-fit background prediction in all regions dedicated to the higgsino mass hypothesis  $m(\tilde{H}) = 200$  GeV, with the significance in each bin shown in the bottom panel.

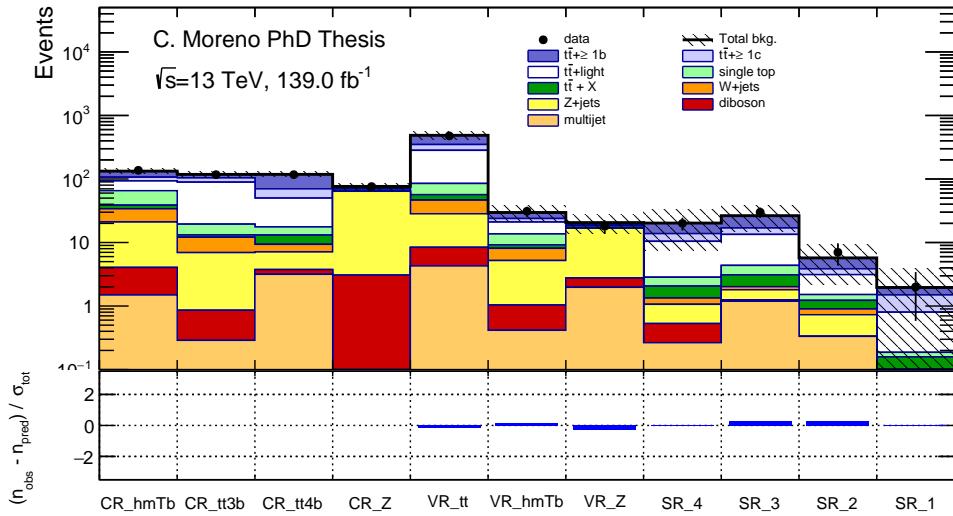


Figure B.8: Comparison between data and post-fit background prediction in all regions dedicated to the higgsino mass hypothesis  $m(\tilde{H}) = 250$  GeV, with the significance in each bin shown in the bottom panel.

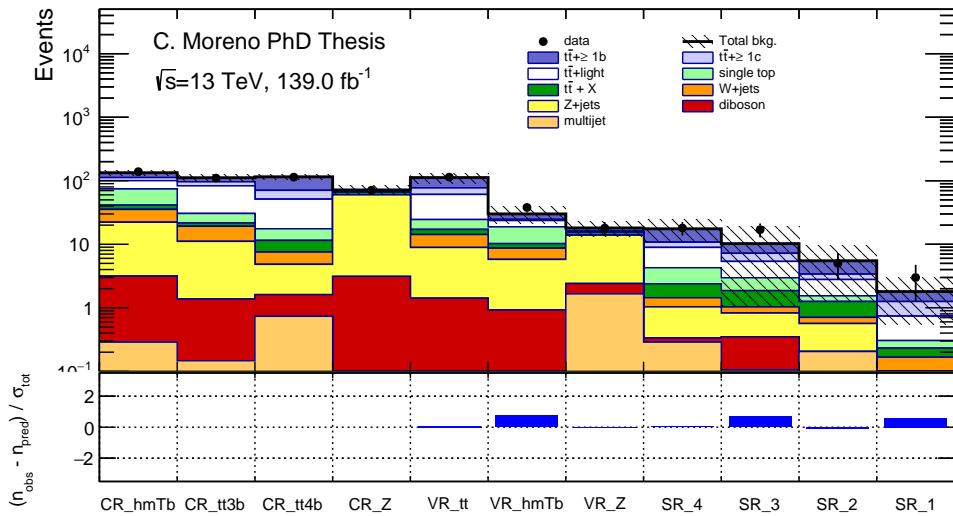


Figure B.9: Comparison between data and post-fit background prediction in all regions dedicated to the higgsino mass hypothesis  $m(\tilde{H}) = 300$  GeV, with the significance in each bin shown in the bottom panel.

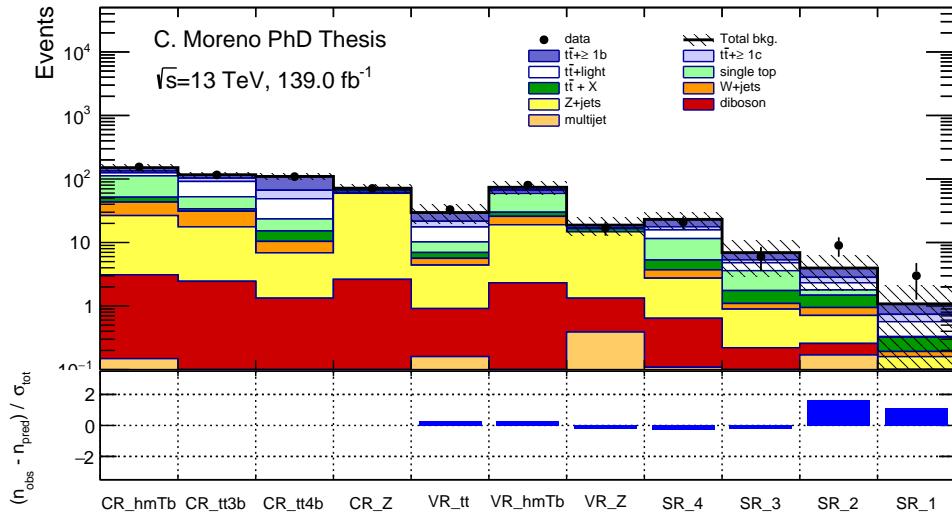


Figure B.10: Comparison between data and post-fit background prediction in all regions dedicated to the higgsino mass hypothesis  $m(\tilde{H}) = 400 \text{ GeV}$ , with the significance in each bin shown in the bottom panel.

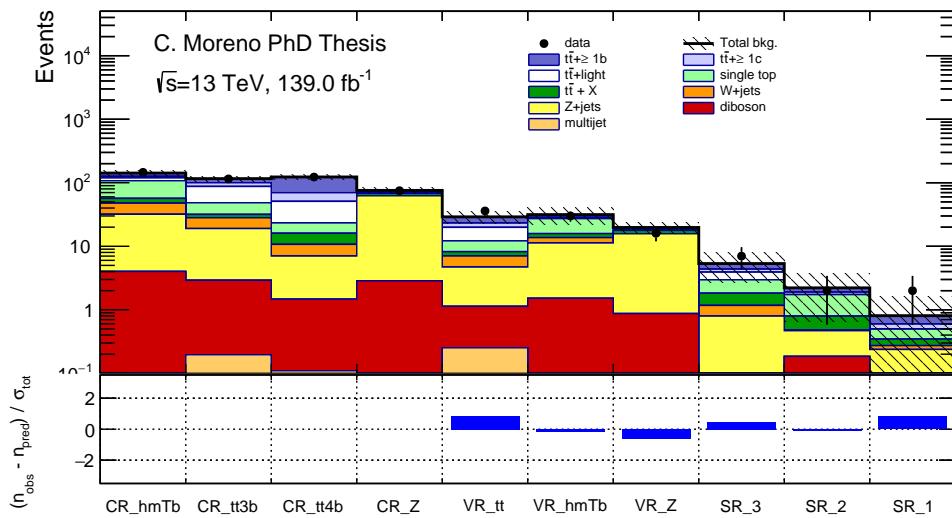


Figure B.11: Comparison between data and post-fit background prediction in all regions dedicated to the higgsino mass hypothesis  $m(\tilde{H}) = 500 \text{ GeV}$ , with the significance in each bin shown in the bottom panel.

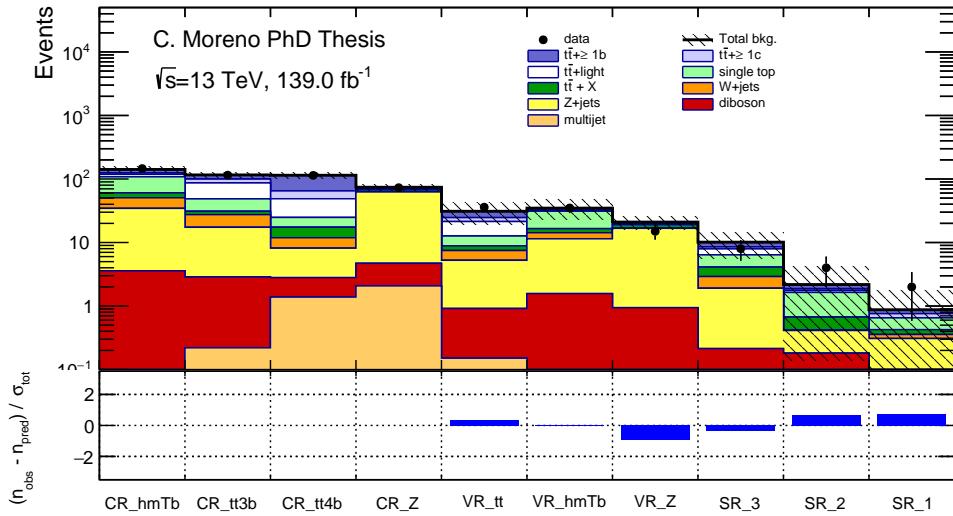


Figure B.12: Comparison between data and post-fit background prediction in all regions dedicated to the higgsino mass hypothesis  $m(\tilde{H}) = 600 \text{ GeV}$ , with the significance in each bin shown in the bottom panel.

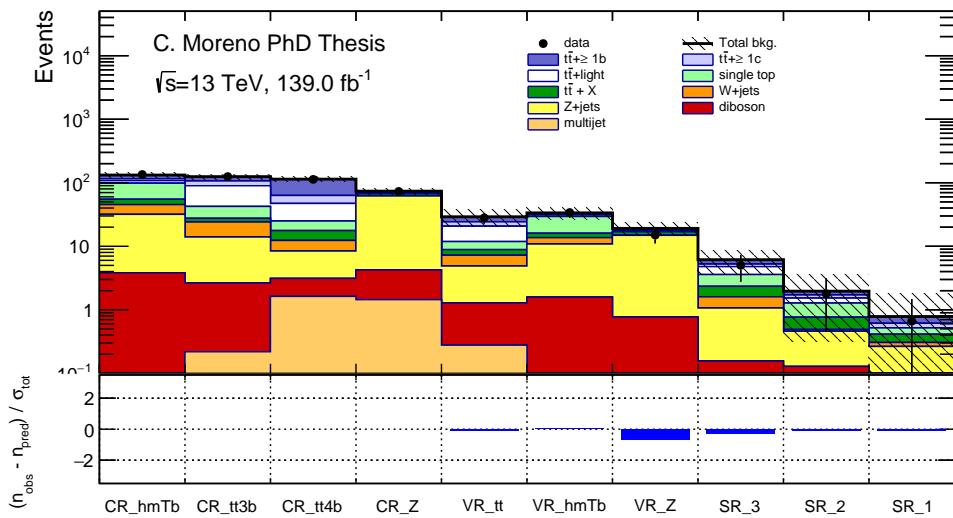


Figure B.13: Comparison between data and post-fit background prediction in all regions dedicated to the higgsino mass hypothesis  $m(\tilde{H}) = 700 \text{ GeV}$ , with the significance in each bin shown in the bottom panel.

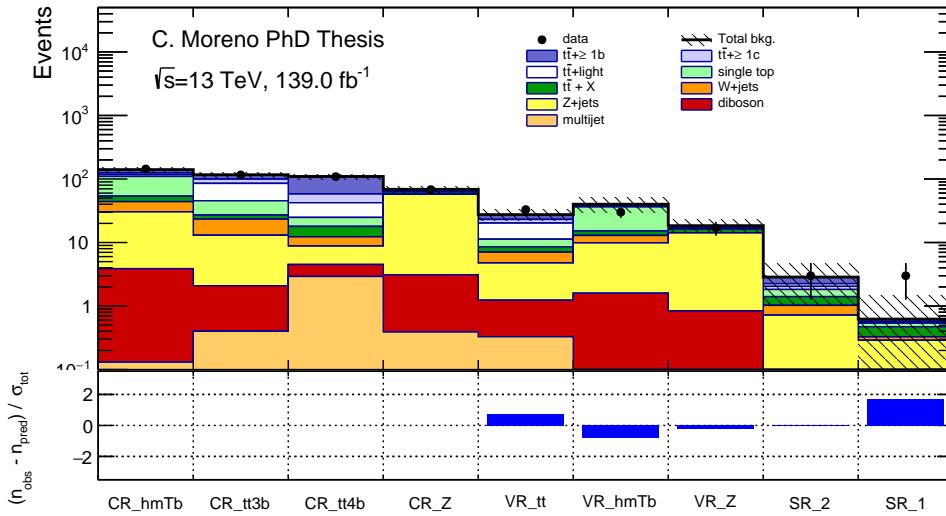


Figure B.14: Comparison between data and post-fit background prediction in all regions dedicated to the higgsino mass hypothesis  $m(\tilde{H}) = 800$  GeV, with the significance in each bin shown in the bottom panel.

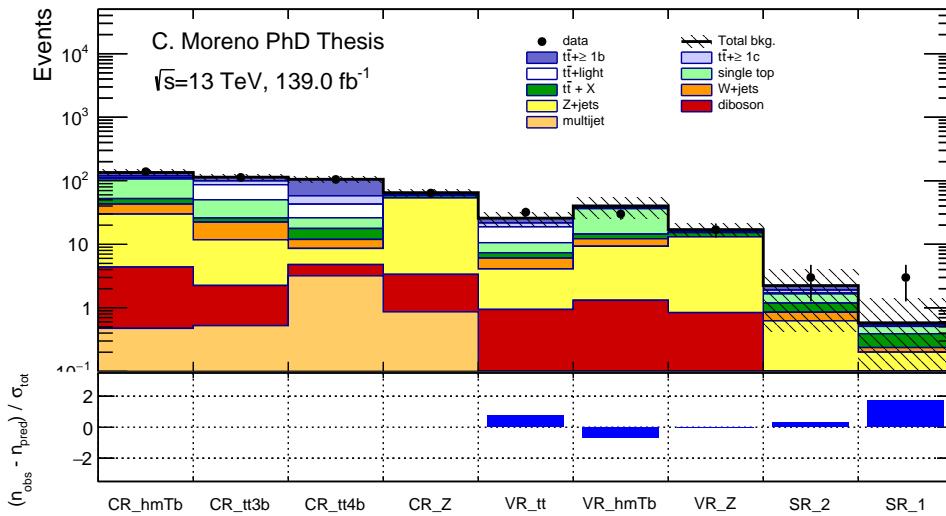


Figure B.15: Comparison between data and post-fit background prediction in all regions dedicated to the higgsino mass hypothesis  $m(\tilde{H}) = 900$  GeV, with the significance in each bin shown in the bottom panel.

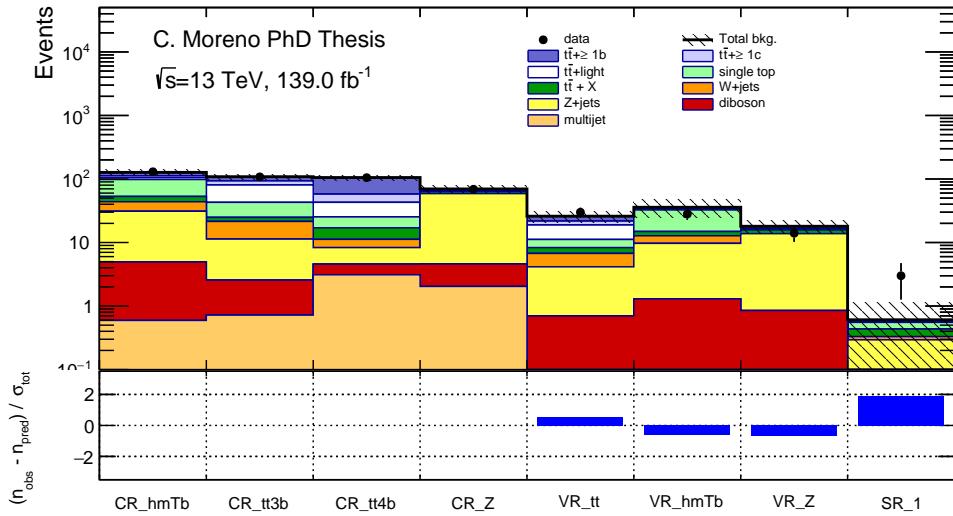


Figure B.16: Comparison between data and post-fit background prediction in all regions dedicated to the higgsino mass hypothesis  $m(\tilde{H}) = 1000 \text{ GeV}$ , with the significance in each bin shown in the bottom panel.

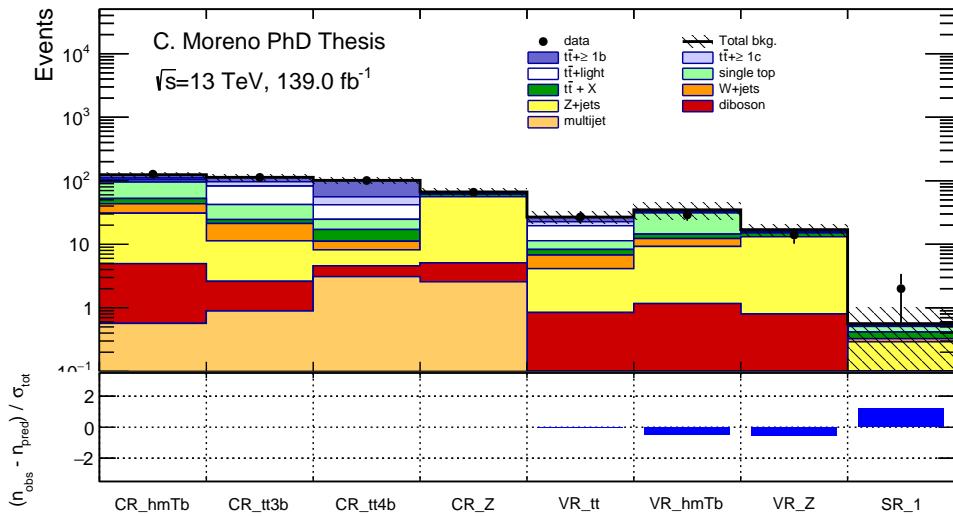


Figure B.17: Comparison between data and post-fit background prediction in all regions dedicated to the higgsino mass hypothesis  $m(\tilde{H}) = 1100 \text{ GeV}$ , with the significance in each bin shown in the bottom panel.

# Bibliography

- [1] M. Gell-Mann, *A schematic model of baryons and mesons*, Phys. Lett. **8** (1964) 214–215.
- [2] S. L. Glashow, *Partial-symmetries of weak interactions*, Nucl. Phys. **22** (1961) 579–588.
- [3] A. Pais and S. B. Treiman, *How Many Charm Quantum Numbers are There*, Phys. Rev. Lett. **35** (1975) 1556–1559.
- [4] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264–1266.
- [5] A. Purcell, *Go on a particle quest at the first CERN webfest. Le premier webfest du CERN se lance à la conquête des particules*, Tech. Rep. BUL-NA-2012-269 **35/2012** (2012), <https://cds.cern.ch/record/1473657>.
- [6] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1–29.
- [7] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012) 30–61.
- [8] Particle Data Group, *Review of Particle Physics*, Progress of Theoretical and Experimental Physics **2020** (2020), 083C01.
- [9] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508–509.
- [10] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321–323.
- [11] G. 't Hooft and M. Veltman, *Regularization and renormalization of gauge fields*, Nucl. Phys. B **44** (1972) 189–213.
- [12] K. G. Begeman et al., *Extended rotation curves of spiral galaxies: dark haloes and modified dynamics*, Monthly Notices of the Royal Astronomical Society **249** (1991) 523–537.
- [13] D. Clowe et al., *Weak-Lensing Mass Reconstruction of the Interacting Cluster 1E 0657-558: Direct Evidence for the Existence of Dark Matter*, ApJ **604** (2004) 596–603.
- [14] Planck Collaboration, *V. CMB power spectra and likelihoods*, A&A **641** (2020).
- [15] C. A. Baker et al., *Improved Experimental Limit on the Electric Dipole Moment of the Neutron*, Phys. Rev. Lett. **97** (2006) 131801.

- [16] ATLAS Collaboration, *ATLAS Standard Model Summary Plots June 2021*, <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2021-032>, Accessed: 14-03-2022.
- [17] ATLAS Collaboration and CMS Collaboration, *Combined Measurement of the Higgs Boson Mass in pp Collisions at  $\sqrt{s} = 7$  and 8 TeV with the ATLAS and CMS Experiments*, Phys. Rev. Lett. **114** (2015) 191803.
- [18] ATLAS Collaboration, *Combined measurements of Higgs boson production and decay using up to 80  $\text{fb}^{-1}$  of proton-proton collision data at  $\sqrt{s} = 13$  TeV collected with the ATLAS experiment*, Phys. Rev. D **101** (2020) 012002.
- [19] LHCb Collaboration, *Test of lepton universality in beauty-quark decays*, arXiv:2103.11769 [hep-ex].
- [20] Muon  $g - 2$  Collaboration, *Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm*, Phys. Rev. Lett. **126** (2021) 141801.
- [21] J.-L. Gervais and B. Sakita, *Field theory interpretation of supergauges in dual models*, Nucl. Phys. B **34** (1971) 632–639.
- [22] J. Wess and B. Zumino, *Supergauge transformations in four dimensions*, Nucl. Phys. B **70** (1974) 39–50.
- [23] H. Georgi and S. L. Glashow, *Unity of All Elementary-Particle Forces*, Phys. Rev. Lett. **32** (1974) 438–441.
- [24] A. Salam and J. Strathdee, *Super-symmetry and non-Abelian gauges*, Phys. Lett. B **51** (1974) 353–355.
- [25] S. Coleman and J. Mandula, *All Possible Symmetries of the S Matrix*, Phys. Rev. **159** (1967) 1251–1256.
- [26] S. P. Martin, *A Supersymmetry Primer*, Advanced Series on Directions in High Energy Physics (1998) 1–98.
- [27] S. Ferrara, B. Zumino, and J. Wess, *Supergauge multiplets and superfields*, Phys. Lett. B **51** (1974) 239–241.
- [28] J. Wess and B. Zumino, *Supergauge invariant extension of quantum electrodynamics*, Nucl. Phys. B **78** (1974) 1–13.
- [29] S. Deser and B. Zumino, *Broken Supersymmetry and Supergravity*, Phys. Rev. Lett. **38** (1977) 1433–1436.
- [30] R. Casalbuoni and others., *High-energy equivalence theorem in spontaneously broken supergravity*, Phys. Rev. D **39** (1989) 2281–2288.
- [31] J. Ellis, C. Smith, and G. Ross, *Will the universe become supersymmetric?*, Phys. Lett. B **114** (1982) 227–229.
- [32] S. A. Abel et al., *SUSY breaking by a metastable ground state: Why the early universe preferred the non-supersymmetric vacuum*, JHEP **01** (2007) 089, arXiv:hep-th/0610334.

- [33] W. Fischler et al., *Meta-stable supersymmetry breaking in a cooling universe*, JHEP **03** (2007) 107, [arXiv:hep-th/0611018](https://arxiv.org/abs/hep-th/0611018).
- [34] L. O’Raifeartaigh, *Spontaneous symmetry breaking for chirals scalar superfields*, Nucl. Phys. B **96** (1975) 331–352.
- [35] P. Fayet and J. Iliopoulos, *Spontaneously broken supergauge symmetries and goldstone spinors*, Phys. Lett. B **51** (1974) 461–464.
- [36] A. E. Nelson and N. Seiberg, *R-symmetry breaking versus supersymmetry breaking*, Nucl. Phys. B **416** (1994) 46–62.
- [37] G. Giudice and R. Rattazzi, *Theories with gauge-mediated supersymmetry breaking*, Phys. Rep. **322** (1999) 419–499.
- [38] R. Barbieri, S. Ferrara, and D. Nanopoulos, *From high energy supersymmetry breaking to low energy physics through decoupling*, Phys. Lett. B **116** (1982) 16–20.
- [39] C. R. Nappi and B. A. Ovrut, *Supersymmetric extension of the  $SU(3) \times SU(2) \times U(1)$  model*, Phys. Lett. B **113** (1982) 175–179.
- [40] H. Haber and G. Kane, *The search for supersymmetry: Probing physics beyond the standard model*, Phys. Rep. **117** (1985) 75–263.
- [41] S. Dimopoulos and D. Sutter, *The supersymmetric flavor problem*, Nucl. Phys. B **452** (1995) 496–512.
- [42] G. R. Farrar and P. Fayet, *Phenomenology of the production, decay, and detection of new hadronic states associated with supersymmetry*, Phys. Lett. B **76** (1978) 575–579.
- [43] M. Chemtob, *Phenomenological constraints on broken R parity symmetry in supersymmetry models*, Progress in Particle and Nuclear Physics **54** (2005) 71–191.
- [44] P. Draper and H. Rzehak, *A review of Higgs mass calculations in supersymmetric models*, Phys. Rep. **619** (2016) 1–24.
- [45] J. Alwall, P. C. Schuster, and N. Toro, *Simplified models for a first characterization of new physics at the LHC*, Phys. Rev. D **79** (2009) 075020.
- [46] D. Alves et al., *Simplified models for LHC new physics searches*, Journal of Physics G: Nuclear and Particle Physics **39** (2012) 105005.
- [47] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001–S08001.
- [48] *Proceedings, ECFA-CERN Workshop on large hadron collider in the LEP tunnel: Lausanne and Geneva, Switzerland, March 21-27*, CERN-YELLOW-84-10-V-1.
- [49] ATLAS Collaboration, *ATLAS: letter of intent for a general-purpose pp experiment at the large hadron collider at CERN*, <https://cds.cern.ch/record/291061>.
- [50] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002–S08002.
- [51] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003–S08003.

- [52] CMS Collaboration, *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08004–S08004.
- [53] LHCb Collaboration, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005–S08005.
- [54] P. Cruikshank et al., *Mechanical design aspects of the LHC beam screen*, Conf. Proc. C **970512** (1997) 3586.
- [55] J. Caron, *AC Collection. Legacy of AC. Pictures from 1992 to 2002*, May 1998.
- [56] E. Mobs, *The CERN accelerator complex - 2019. Complexe des accélérateurs du CERN - 2019*, <https://cds.cern.ch/record/2684277>, General Photo.
- [57] R. Cappi et al., *The CERN PS complex as part of the LHC injector chain*, Conference Record of the 1991 IEEE Particle Accelerator Conference (1991) 171–173 vol.1.
- [58] L. Di Lella and C. Rubbia, *The Discovery of the W and Z Particles*, Adv. Ser. Direct. High Energy Phys. **23** (2015) 137–163.
- [59] ATLAS Public Luminosity plots, <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2>, Accessed:14-03-2022.
- [60] O. Aberle et al., *High-Luminosity Large Hadron Collider (HL-LHC): Technical design report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2020.
- [61] J. Goodson <http://www.jetgoodson.com/images/thesisImages/magnetSystems.png>, Accessed: 14-03-2022.
- [62] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report*, tech. rep., Sep, 2010. <https://cds.cern.ch/record/1291633>.
- [63] ATLAS Collaboration, *New Small Wheel Technical Design Report*, tech. rep., Jun, 2013. <https://cds.cern.ch/record/1552862>. ATLAS-TDR-020.
- [64] ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2015*, Eur. Phys. J. C **77** (2017) 317, [arXiv:1611.09661 \[hep-ex\]](https://arxiv.org/abs/1611.09661).
- [65] ATLAS Collaboration, *Operation of the ATLAS trigger system in Run 2*, JINST **15** (2020) P10004–P10004.
- [66] ATLAS Collaboration, *ATLAS level-1 trigger: Technical Design Report*. Technical design report. ATLAS. CERN, Geneva, 1998. <https://cds.cern.ch/record/381429>.
- [67] ATLAS Collaboration, *ATLAS high-level trigger, data-acquisition and controls: Technical Design Report*. Technical design report. ATLAS. CERN, Geneva, 2003. <https://cds.cern.ch/record/616089>.
- [68] ATLAS Collaboration, *Trigger Menu in 2016*, ATL-DAQ-PUB-2017-001 (2017), <https://cds.cern.ch/record/2242069>.
- [69] ATLAS Collaboration, *Trigger Menu in 2017*, ATL-DAQ-PUB-2018-002 (2018), <https://cds.cern.ch/record/2625986>.
- [70] ATLAS Collaboration, *Trigger menu in 2018*, ATL-DAQ-PUB-2019-001 (2019), <https://cds.cern.ch/record/2693402>.

- [71] ATLAS Public DAQ plots,  
<https://twiki.cern.ch/twiki/bin/view/AtlasPublic/ApprovedPlotsDAQ>, Accessed: 14-03-2022.
- [72] ATLAS Collaboration, *Fast TrackEr (FTK) Technical Design Report*,  
<https://cds.cern.ch/record/1552953>, ATLAS-TDR-021.
- [73] ATLAS Collaboration, *The ATLAS Level-1 Calorimeter Trigger*, JINST **3** (2008) P03001–P03001.
- [74] ATLAS Collaboration, *Performance of the ATLAS muon triggers in Run 2*, JINST **15** (2020) P09015–P09015.
- [75] R. Simoniello, *The ATLAS Level-1 Topological Processor: from design to routine usage in Run-2*, 2018 IEEE Nuclear Science Symposium and Medical Imaging Conference Proceedings (NSS/MIC) (2018) 1–4.
- [76] ATLAS Collaboration, *The ATLAS central level-1 trigger logic and TTC system*, JINST **3** (2008) P08002–P08002.
- [77] ATLAS Collaboration, *Racks in USA15 (100m underground) hosting trigger electronics*, General Photo, Jun, 2015.
- [78] W. Lampl et al., *Calorimeter clustering algorithms: Description and performance*, ATL-LARG-PUB-2008-002 (2008), <https://cds.cern.ch/record/1099735>.
- [79] ATLAS Collaboration, *The Level-1 Trigger Muon Barrel System of the ATLAS experiment at CERN*, JINST **4** (2009) P04010–P04010.
- [80] R. Ichimiya et al., *Sector logic implementation for the ATLAS endcap level-1 muon trigger*, 8th Workshop on Electronics for LHC Experiments (2002).
- [81] ATLAS Collaboration, *The ATLAS Level-1 Muon Topological Trigger Information for Run 2 of the LHC*, JINST **10** (2015) C02027–C02027.
- [82] ATLAS Collaboration, *Operation of the upgraded ATLAS Central Trigger Processor during the LHC Run 2*, JINST **11** (2016) C02020–C02020.
- [83] ATLAS Collaboration, *Performance of the ATLAS Level-1 topological trigger in Run 2*, Eur. Phys. J. C **82** (2022) 7, [arXiv:2105.01416 \[hep-ex\]](https://arxiv.org/abs/2105.01416).
- [84] ATLAS Public L1Calo plots, <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/L1CaloTriggerPublicResults>, Accessed: 14-03-2022.
- [85] ATLAS Collaboration, *Measurement of the CP-violating phase  $\phi_s$  in  $B_s^0 \rightarrow J/\psi\phi$  decays in ATLAS at 13 TeV*, Eur. Phys. J. C **81** (2021) 342, [arXiv:2001.07115 \[hep-ex\]](https://arxiv.org/abs/2001.07115).
- [86] ATLAS Collaboration, *Technical Design Report for the Phase-I Upgrade of the ATLAS TDAQ System*, <https://cds.cern.ch/record/1602235>, Final version presented to December 2013 LHCC.
- [87] P. Skands, *QCD for Collider Physics*, 4, 2011. CERN-TH-2011-080, MCNET-11-13.
- [88] R. D. Ball et al., *Parton distributions from high-precision collider data*, Eur. Phys. J. C **77** (2017).

- [89] *HERA - A Proposal for a Large Electron Proton Colliding Beam Facility at DESY*. 1981. DESY-HERA-81-10.
- [90] X. Artru and G. Mennessier, *String model and multiproduction*, Nucl. Phys. B **70** (1974) 93–115.
- [91] A. Buckley et al., *General-purpose event generators for LHC physics*, Phys. Rep. **504** (2011) 145–233.
- [92] B. Webber, *A QCD model for jet fragmentation including soft gluon interference*, Nucl. Phys. B **238** (1984) 492–528.
- [93] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159–177.
- [94] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, Eur. Phys. J. C **76** (2016).
- [95] S. Alioli et al., *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010) 043, [arXiv:1002.2581 \[hep-ph\]](https://arxiv.org/abs/1002.2581).
- [96] S. Frixione et al., *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP **11** (2007) 070, [arXiv:0709.2092 \[hep-ph\]](https://arxiv.org/abs/0709.2092).
- [97] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) 079, [arXiv:1405.0301 \[hep-ph\]](https://arxiv.org/abs/1405.0301).
- [98] P. Freeman et al., *Sherpa: a mission-independent data analysis application*, Astronomical Data Analysis **4477** (2001) 76 – 87.
- [99] S. Agostinelli et al., *Geant4—a simulation toolkit*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **506** (2003) 250–303.
- [100] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823–874.
- [101] ATLAS Collaboration, *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim*, tech. rep., CERN, Geneva, Oct, 2010. <https://cds.cern.ch/record/1300517>.
- [102] K. Edmonds et al., *The Fast ATLAS Track Simulation (FATRAS)*, tech. rep., CERN, Geneva, Mar, 2008. <https://cds.cern.ch/record/1091969>.
- [103] T. Cornelissen et al., *The new ATLAS track reconstruction (NEWT)*, Journal of Physics: Conference Series **119** (2008) 032014.
- [104] ATLAS Collaboration, *Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC Run 2*, Eur. Phys. J. C **77** (2017) 673, [arXiv:1704.07983 \[hep-ex\]](https://arxiv.org/abs/1704.07983).
- [105] ATLAS Collaboration, *Number of tracks vs. mu with full 2016 data*, <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLLOTS/IDTR-2016-015/>, Accessed: 14-03-2022.

- [106] ATLAS Collaboration, *Number of reconstructed vertices versus number of interactions in 2016 data and simulation*, <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/IDTR-2016-013/>, Accessed: 14-03-2022.
- [107] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data*, JINST **14** (2019) P12006, arXiv:1908.00005 [hep-ex].
- [108] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at  $\sqrt{s} = 13$  TeV*, Eur. Phys. J. C **76** (2016) 292, arXiv:1603.05598 [hep-ex].
- [109] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. C **77** (2017) 490, arXiv:1603.02934 [hep-ex].
- [110] M. Cacciari, G. P. Salam, and G. Soyez, *The anti- $k_t$  jet clustering algorithm*, JHEP **04** (2008) 063, arXiv:0802.1189 [hep-ph].
- [111] ATLAS Collaboration, *Jet energy scale and resolution measured in proton–proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, Eur. Phys. J. C **81** (2021) 689, arXiv:2007.02645 [hep-ex].
- [112] ATLAS Collaboration, *Jet global sequential corrections with the ATLAS detector in proton-proton collisions at  $\sqrt{s} = 8$  TeV*, ATLAS-CONF-2015-002 (2015), <https://cds.cern.ch/record/2001682>.
- [113] ATLAS Collaboration, *Tagging and suppression of pileup jets*, ATLAS-CONF-2014-018 (2014), <https://cds.cern.ch/record/1700870>.
- [114] ATLAS Collaboration, *Identification and rejection of pile-up jets at high pseudorapidity with the ATLAS detector*, Eur. Phys. J. C **77** (2017) 580, arXiv:1705.02211 [hep-ex], [Erratum: Eur.Phys.J.C 77, 712 (2017)].
- [115] ATLAS Collaboration,, *Selection of jets produced in 13TeV proton-proton collisions with the ATLAS detector*, ATLAS-CONF-2015-029 (2015), <https://cds.cern.ch/record/2037702>.
- [116] ATLAS Collaboration, *ATLAS b-jet identification performance and efficiency measurement with  $t\bar{t}$  events in  $pp$  collisions at  $\sqrt{s} = 13$  TeV*, Eur. Phys. J. C **79** (2019) 970, arXiv:1907.05120 [hep-ex].
- [117] ATLAS Collaboration, *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*, ATL-PHYS-PUB-2017-013 (2017), <https://cds.cern.ch/record/2273281>.
- [118] ATLAS Collaboration, *Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data*, ATLAS-CONF-2011-102 (2011), <https://cds.cern.ch/record/1369219>.
- [119] G. Piacquadio and C. Weiser, *A new inclusive secondary vertex algorithm for b-jet tagging in ATLAS*, Journal of Physics: Conference Series **119** (2008) 032032.

- [120] ATLAS Collaboration, *Expected performance of the ATLAS b-tagging algorithms in Run-2*, ATL-PHYS-PUB-2015-022 (2015), <http://cdsweb.cern.ch/record/2037697>.
- [121] ATLAS Collaboration, *Identification of Jets Containing b-Hadrons with Recurrent Neural Networks at the ATLAS Experiment*, ATL-PHYS-PUB-2017-003 (2017), <http://cdsweb.cern.ch/record/2255226>.
- [122] ATLAS Collaboration, *Performance of 2019 recommendations of ATLAS Flavor Tagging algorithms with Variable Radius track jets.*, <http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLLOTS/FTAG-2019-006/>, Accessed: 14-03-2022.
- [123] ATLAS Collaboration, *ATLAS b-jet identification performance and efficiency measurement with  $t\bar{t}$  events in pp collisions at  $\sqrt{s} = 13$  TeV*, Eur. Phys. J. C **79** (2019) 970, arXiv:1907.05120 [hep-ex].
- [124] ATLAS Collaboration, *Measurement of b-tagging Efficiency of c-jets in  $t\bar{t}$  Events Using a Likelihood Approach with the ATLAS Detector*, ATLAS-CONF-2018-001 (2018), <https://cds.cern.ch/record/2306649>.
- [125] ATLAS Collaboration, *Measurement of the c-jet mistagging efficiency in  $t\bar{t}$  events using pp collision data at  $\sqrt{s} = 13$  TeV collected with the ATLAS detector*, Eur. Phys. J. C **82** (2022) 95, arXiv:2109.10627 [hep-ex].
- [126] ATLAS Collaboration, *Calibration of light-flavour jet b-tagging rates on ATLAS proton-proton collision data at  $\sqrt{s} = 13\text{-}14$  TeV*, ATLAS-CONF-2018-006 (2018), <https://cds.cern.ch/record/2314418>.
- [127] ATLAS Collaboration, *Flavour-tagging efficiency corrections for the 2019 ATLAS PFlow jet b-taggers with the full LHC Run II dataset*, FTAG-2020-001, <http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLLOTS/FTAG-2020-001/>.
- [128] ATLAS Collaboration, *Performance of algorithms that reconstruct missing transverse momentum in  $\sqrt{s} = 8$  TeV proton-proton collisions in the ATLAS detector*, Eur. Phys. J. C **77** (2017) 241, arXiv:1609.09324 [hep-ex].
- [129] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at  $\sqrt{s} = 13$  TeV*, Eur. Phys. J. C **78** (2018) 903, arXiv:1802.08168 [hep-ex].
- [130] ATLAS Collaboration, *Missing transverse momentum performance using the full run 2 pp data at 13 TeV.*, <http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLLOTS/JETM-2019-03/>, Accessed: 14-03-2022.
- [131] ATLAS Collaboration, *Object-based missing transverse momentum significance in the ATLAS detector*, ATLAS-CONF-2018-038 (2018), <https://cds.cern.ch/record/2630948>.
- [132] ATLAS Collaboration, *Search for electroweak production of supersymmetric particles in final states with two or three leptons at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, Eur. Phys. J. C **78** (2018).

- [133] ATLAS Collaboration, *Search for the direct production of charginos and neutralinos in final states with tau leptons in  $\sqrt{s} = 13$  TeV pp collisions with the ATLAS detector*, Eur. Phys. J. C **78** (2018).
- [134] ATLAS Collaboration, *Search for chargino and neutralino production in final states with a Higgs boson and missing transverse momentum at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, Phys. Rev. D **100** (2019) 012006, [arXiv:1812.09432 \[hep-ex\]](#).
- [135] ATLAS Collaboration, *Search for pair production of higgsinos in final states with at least three b-tagged jets in  $\sqrt{s} = 13$  TeV pp collisions using the ATLAS detector*, Phys. Rev. D **98** (2018) 092002, [arXiv:1806.04030 \[hep-ex\]](#).
- [136] A. Hoecker et al., *TMVA - Toolkit for Multivariate Data Analysis*, 2009. CERN-OPEN-2007-007.
- [137] L. Garrido and A. Juste, *On the determination of probability density functions by using Neural Networks*, Comput. Phys. Commun. **115** (1998) 25–31.
- [138] ATLAS Collaboration, *Performance of the missing transverse momentum triggers for the ATLAS detector during Run-2 data taking*, JHEP **08** (2020) 080, [arXiv:2005.09554 \[hep-ex\]](#).
- [139] M. Czakon and A. Mitov, *Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders*, Comput. Phys. Commun. **185** (2014) 2930, [arXiv:1112.5675 \[hep-ph\]](#).
- [140] P. Kant et al., *HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions*, Comput. Phys. Commun. **191** (2015) 74–89, [arXiv:1406.4403 \[hep-ph\]](#).
- [141] N. Kidonakis, *NNLL resummation for s-channel single top quark production*, Phys. Rev. D **81** (2010) 054028, [arXiv:1001.5034 \[hep-ph\]](#).
- [142] LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, [arXiv:1307.1347 \[hep-ph\]](#).
- [143] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, [arXiv:0811.4622 \[hep-ph\]](#).
- [144] S. Hoeche et al., *QCD matrix elements + parton showers: The NLO case*, JHEP **04** (2013) 027, [arXiv:1207.5030 \[hep-ph\]](#).
- [145] S. Catani et al., *Vector boson production at hadron colliders: a fully exclusive QCD calculation at NNLO*, Phys. Rev. Lett. **103** (2009) 082001, [arXiv:0903.2120 \[hep-ph\]](#).
- [146] ATLAS Collaboration, *Search for the Standard Model Higgs boson produced in association with top quarks and decaying into  $b\bar{b}$  in pp collisions at  $\sqrt{s} = 8$  TeV with the ATLAS detector*, Eur. Phys. J. C **75** (2015) 349, [arXiv:1503.05066 \[hep-ex\]](#).
- [147] ATLAS Collaboration, *Measurements of inclusive and differential fiducial cross-sections of  $t\bar{t}$  production with additional heavy-flavour jets in proton-proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, JHEP **04** (2019) 046, [arXiv:1811.12113 \[hep-ex\]](#).
- [148] T. Gleisberg and S. Hoeche, *Comix, a new matrix element generator*, JHEP **12** (2008) 039, [arXiv:0808.3674 \[hep-ph\]](#).

- [149] F. Cascioli, P. Maierhöfer, and S. Pozzorini, *Scattering Amplitudes with Open Loops*, Phys. Rev. Lett. **108** (2012).
- [150] *QCD matrix elements + parton showers*, JHEP **11** (2001) 063, [arXiv:hep-ph/0109231](https://arxiv.org/abs/hep-ph/0109231).
- [151] P. Baldi et al., *Parameterized neural networks for high-energy physics*, Eur. Phys. J. C **76** (2016) 235, [arXiv:1601.07913 \[hep-ex\]](https://arxiv.org/abs/1601.07913).
- [152] T. Chen and C. Guestrin, *XGBoost*, Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (2016).
- [153] L. Moneta et al., *The RooStats Project*, PoS **ACAT2010** (2010) 057, [arXiv:1009.1003 \[physics.data-an\]](https://arxiv.org/abs/1009.1003).
- [154] H. Cramér, *Mathematical Methods of Statistics*, Princeton University Press (March 23, 1999).
- [155] C. R. Rao, *Information and the Accuracy Attainable in the Estimation of Statistical Parameters*. Springer New York, New York, NY, 1992.
- [156] J. Neyman and E. S. Pearson, *On the Problem of the Most Efficient Tests of Statistical Hypotheses*, Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character **231** (1933) 289–337.
- [157] L. Lista, *Statistical Methods for Data Analysis in Particle Physics*. Springer, Cham, 2017.
- [158] T. Junk, *Confidence level computation for combining searches with small statistics*, Nucl. Instrum. Meth. A **434** (1999) 435–443.
- [159] A. Wald, *On Cumulative Sums of Random Variables*, The Annals of Mathematical Statistics **15** (1944) 283 – 296.
- [160] ATLAS Collaboration, *Luminosity determination in pp collisions at  $\sqrt{s} = 13$  TeV using the ATLAS detector at the LHC*, ATLAS-CONF-2019-021 (2019), <https://cds.cern.ch/record/2677054>.
- [161] M. Baak et al., *HistFitter software framework for statistical data analysis*, Eur. Phys. J. C **75** (2015) 153.
- [162] CMS Collaboration,, *Search for higgsinos decaying to two Higgs bosons and missing transverse momentum in proton-proton collisions at  $\sqrt{s} = 13$  TeV*, [arXiv:2201.04206 \[hep-ex\]](https://arxiv.org/abs/2201.04206).
- [163] ATLAS Collaboration, *Search for supersymmetry in events with four or more charged leptons in  $139\text{ fb}^{-1}$  of  $\sqrt{s} = 13$  TeV pp collisions with the ATLAS detector*, JHEP **07** (2021) 167, [arXiv:2103.11684 \[hep-ex\]](https://arxiv.org/abs/2103.11684).
- [164] ATLAS Collaboration, *Search for charginos and neutralinos in final states with two boosted hadronically decaying bosons and missing transverse momentum in pp collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, Phys. Rev. D **104** (2021) 112010, [arXiv:2108.07586 \[hep-ex\]](https://arxiv.org/abs/2108.07586).
- [165] R. D. Peccei and H. R. Quinn, *CP Conservation in the Presence of Pseudoparticles*, Phys. Rev. Lett. **38** (1977) 1440–1443.
- [166] S. Weinberg, *A New Light Boson?*, Phys. Rev. Lett. **40** (1978) 223–226.

- [167] F. Wilczek, *Problem of Strong P and T Invariance in the Presence of Instantons*, Phys. Rev. Lett. **40** (1978) 279–282.
- [168] G. Barenboim et al., *Implications of an axino LSP for naturalness*, Phys. Rev. D **90** (2014) 035020, arXiv:1407.1218 [hep-ph].
- [169] L. Covi et al., *Axinos as dark matter*, JHEP **05** (2001) 033, arXiv:hep-ph/0101009.
- [170] K. J. Bae et al., *Cosmology of the DFSZ axino*, JCAP **03** (2012) 013, arXiv:1111.5962 [hep-ph].
- [171] G. F. Giudice and R. Rattazzi, *Theories with gauge mediated supersymmetry breaking*, Phys. Rept. **322** (1999) 419–499, arXiv:hep-ph/9801271.
- [172] ATLAS Collaboration, *Reinterpretation of searches for supersymmetry in models with variable R-parity-violating coupling strength and long-lived R-hadrons*, ATLAS-CONF-2018-003 (2018), <https://cds.cern.ch/record/2308391>.
- [173] ATLAS Collaboration, *Performance of vertex reconstruction algorithms for detection of new long-lived particle decays within the ATLAS inner detector*, ATL-PHYS-PUB-2019-013 (2019), <https://cds.cern.ch/record/2669425>.
- [174] M. Chala et al., *Constraining Dark Sectors with Monojets and Dijets*, JHEP **07** (2015) 089, arXiv:1503.05916 [hep-ph].
- [175] T. Jacques et al., *Complementarity of DM searches in a consistent simplified model: the case of Z*, JHEP **10** (2016) 071, arXiv:1605.06513 [hep-ph], [Erratum: JHEP 01, 127 (2019)].
- [176] G. Busoni et al., *Making the Most of the Relic Density for Dark Matter Searches at the LHC 14 TeV Run*, JCAP **03** (2015) 022, arXiv:1410.7409 [hep-ph].
- [177] M. Duerr et al., *How to save the WIMP: global analysis of a dark matter model with two s-channel mediators*, JHEP **09** (2016) 042, arXiv:1606.07609 [hep-ph].
- [178] N. F. Bell et al., *Dark Forces in the Sky: Signals from Z' and the Dark Higgs*, JCAP **08** (2016) 001, arXiv:1605.09382 [hep-ph].
- [179] F. Kahlhoefer et al., *Implications of unitarity and gauge invariance for simplified dark matter models*, JHEP **02** (2016) 016, arXiv:1510.02110 [hep-ph].
- [180] M. Duerr et al., *Hunting the dark Higgs*, JHEP **04** (2017) 143, arXiv:1701.08780 [hep-ph].
- [181] ATLAS Collaboration, *Search for Dark Matter Produced in Association with a Dark Higgs Boson Decaying into  $W^\pm W^\mp$  or ZZ in Fully Hadronic Final States from  $\sqrt{s} = 13$  TeV pp Collisions Recorded with the ATLAS Detector*, Phys. Rev. Lett. **126** (2021) 121802, arXiv:2010.06548 [hep-ex].