UNIVERSITY
OF SUSSEX

Doctoral Thesis

Search for third generation scalar quarks in events with b-tagged jets with the ATLAS detector

A thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy

in the

Department of Physics and Astronomy
School of Mathematical and Physical Sciences

Author:
Giuseppe Lerner

Supervisor:
Dr. Iacopo Vivarelli

September 5, 2018

A Vittorio Cristoforetti

University of Sussex
School of Mathematical and Physical Sciences
Department of Physics and Astronomy

Doctoral Thesis

Search for third generation scalar quarks in events with *b*-tagged jets with the ATLAS detector

by Giuseppe Lerner

Abstract

The thesis presents the results of two searches for the direct pair-production of third generation scalar quarks, the stop and the sbottom, in proton-proton collisions at $\sqrt{s} = 13$ TeV delivered by the Large Hadron Collider (LHC) and recorded by the ATLAS detector. Third generation squarks are studied in the context of natural supersymmetric extensions of the Standard Model, highlighting their role in the solution of the Higgs hierarchy problem and considering both R-parity conserving and violating decay scenarios. The signal models of interest produce final states characterised by the presence of two bottom quarks, and the identification of the hadronic jets generated by their fragmentation plays a crucial role in the analyses. The performance of b-jet identification algorithms is studied in detail, and a novel approach for the estimate of the associated systematic uncertainties is presented. The first analysis in the thesis is a search for a pair-produced sbottom with two-body decays into Standard Model third generation quarks and quasi-degenerate electroweakinos, while the second targets the pair-production of the stop followed by R-parity violating decays into a bottom quark and a lepton. No evidence of SUSY particles is found, and exclusion limits are set on the relevant signal models using dedicated statistical tools.

I hereby declare that this thesis has not been and will not be, submitted in whole or in part to another university for the award of any other degree.

Brighton,
September 5, 2018

Giuseppe Lerner

Acknowledgements

The journey through my PhD has been long and full of challenges, and I feel thankful for having shared it with many people who contributed to make it an extraordinary experience.

I had the privilege to be part of the ATLAS group of the University of Sussex, in a warm and welcoming environment where I received constant support and inspiration since the very first days of work. Special thanks go to my supervisor, Iacopo Vivarelli, who provided me with exceptional guidance in my research and has always been available to help in any situation, from work-related problems to any kind of practical issue that I had. I don't think that I could have found a better mentor, and I cannot overestimate the impact that he has made on the quality of my PhD. I received great help and support also from Alex Cerri, who I should thank in particular for his deep and careful review of this thesis, but also for his unforgettable dinners at CERN and in Brighton. The Sussex group would simply not exist without the invaluable presence of Antonella De Santo and Fabrizio Salvatore, and the enthusiasm and creativity of Lily Asquith are also essential to make it an amazing place to be. I also wish to thank my Master degree supervisor from Milan, Tommaso Lari, who had a crucial role in developing my interest towards particle physics with his passionate attitude towards science and research.

A decisive figure that has helped me all along my PhD is Kerim Suruliz, with his spectacular ability to carry out a half dozen analyses simultaneously thanks to his legendary scripts and his inborn talent for physics. He is, without doubt, one of the brightest minds that I had the opportunity to meet in my PhD career, and at least a generation of Sussex students has benefited incredibly from his presence in the group. I am also deeply grateful to Carlos Chavez, who gave me decisive support in my operational work on the e-gamma trigger, and was always available to give his advice or simply to chat and share our time. Together with them, the other Sussex post-docs Nicky, Umberto, Mark, Benedict and Batool have also been a pleasure to be with during these years.

The list of current and former Sussex EPP students is long, and I genuinely don't know who to thank first. I will start from Fabio, who was brave enough to follow me at Sussex from Milan, and was as always a great company even if we almost scientifically managed to be permanently based in different places. The office would not have been the same

without *mbare* Fab and the countless laughs and interesting discussions that we shared, and Iannis and Mario S have been a great support during my thesis writing process in the last few months. A big thank you also to Zara, Suf, Nicola, Olly, Emma, Sam, Fab T, Dan and Mario G for always keeping a nice and enjoyable atmosphere in the group.

During my PhD I spent a huge amount of time working within the sbottom analysis team, and the results that we obtained would not have been possible without everyone's contribution. I am grateful to Monica D'Onofrio, Davide Costanzo and Martin Tripiana for their leading role in the analyses and in the paper editing, and to Silvia, John and Calum for sharing the heavy workloads and for participating to the efforts in a friendly and constructive way. Many thanks also to Leigh Schaefer, Jeff Dandoy and Evelyn Thomson for their outstanding work on the RPV stop analysis.

There are several people both in Brighton and at CERN who I wish to thank for the time that we spent together outside work, that was essential to distract me from the sometimes stressful PhD life. It would be impossible to mention everyone, but I am particularly grateful to Luigi and Manfredi in Brighton and Andrea and Carlo at CERN, without whom the PhD journey would have been way harder.

My life as a PhD student has often been very intense, with endless travels and commuting between Brighton and Geneva, and my main regret is that inevitably I didn't manage to spend as much time as I would have wanted with my friends and my family. Despite the distance, the love and support that I received from them have been constant also during these years, and it is needless to say that I am thankful for everything that they have done for me.

Lastly, there are no words to describe how lucky I feel for having Chiara by my side. She has never made me feel alone even when the circumstances were keeping us apart, and it is hard to express how much her presence has helped me throughout these years. Part of this thesis is also hers.

Contents

Introduction

When the ancient Greek philosopher and scientist Aristotle wrote in his "Metaphysics" book collection that

> *"The mathematical sciences particularly exhibit order symmetry and limitations; and these are the greatest forms of the beautiful"*

he may not have imagined that, 23 centuries later, an international collaboration of scientists would build a 27 km long circular accelerator, the Large Hadron Collider (LHC) [1], to investigate symmetry as a founding principle of nature at the scale of its elementary constituents.

In fact, while symmetric arguments have always been employed implicitly as a mean to obtain coherent descriptions of natural phenomena, it is only in modern times that the notion of symmetry was applied directly to the formulation of physical laws [2], using its geometrical definition of *invariance under a group of transformations*. The revolutionary theory of relativity published by Einstein in 1905 [3] was built on the assumption that the physical description of any phenomena must be invariant under spacetime (Lorentz) transformations, promoting spacetime symmetries to the role of primary principles of nature. Throughout the 20^{th} century, the Standard Model (SM) of the elementary particles [4] was developed by relating the dynamical properties of the interactions to specific types of local (gauge) symmetries, obtaining an accurate and predictive description of the electromagnetic, weak, and strong forces. The SM was further completed by the concept of spontaneous symmetry breaking, introduced to explain the finite mass values of the known particles without compromising the symmetric structure of the theory [5–7], leading to the recent discovery of the Higgs boson [8,9] by the ATLAS [10] and CMS [11] Collaborations at the LHC.

This remarkable sequence of successes led the scientific community to the idea that the existing open issues in high energy physics may be solved by a further extension of the symmetries of the SM. In particular, the core topic of the present thesis is the ex-

perimental search for Supersymmetry (SUSY) [12, 13], an extension of the spacetime symmetry group of the SM that introduces a transformation between different classes of elementary particles, bosons and fermions, such that each SM particle receives a superpartner with spin difference of $1/2$. The phenomenology of SUSY models is described in Chapter 1 under the assumption that the new symmetry, if it exists, must necessarily be broken, such that the SM particles and the corresponding superpartners can have different masses. More specifically, SUSY is presented as a possible solution to the Higgs hierarchy problem [14–16], that emerges in the SM due to the unnatural discrepancy between the electroweak scale (10^2 GeV) and the characteristic energy scale of gravity, the Planck scale (10^{19} GeV). In order to solve the hierarchy problem it is necessary to place stringent constraints on the masses of a subset of SUSY particles [17], including the superpartners of third generation squarks, the stop ($\tilde{t}$) and the sbottom ($\tilde{b}$), whose searches at the LHC are the main subject of the thesis.

After the theoretical introduction described above, Chapter 2 is dedicated to an overview of the experimental facilities of the CERN Organisation, which provide a unique opportunity to search for the production of SUSY particles in proton-proton collisions with multi-TeV centre of mass energy. The chapter also includes a technical description of the ATLAS detector, which is used to record the data analysed in this thesis. A discussion of the simulation and reconstruction techniques of proton-proton collision events in ATLAS is then presented in Chapter 3, covering the most important tools that are needed to perform the physics analyses.

The searches in this thesis target signal models where third generation squarks are produced in pairs and decay promptly into final states with two b-quarks and a variable number of extra hadrons, leptons and SUSY particles. An essential tool for the analyses is b-tagging, which is used to identify hadronic jets from b-quark fragmentation (b-jets), allowing to select events with signal-like properties whilst reducing the impact of SM backgrounds. A detailed study of the performance of the baseline b-tagging algorithm of ATLAS is presented in Chapter 4, where the impact of experimental sources of systematic uncertainty on the tagging efficiency is studied with Monte Carlo (MC) simulations as a function of the transverse momentum of the input jets, obtaining an overall estimate of the uncertainty which is then employed in the physics analyses. The study is also useful to measure the variations of the performance of the tagger when applied to different physics processes.

A general introduction to SUSY searches in ATLAS is presented in Chapter 5, describing the effort of the Collaboration to cover a variety of experimental signatures from

different SUSY scenarios, with focus on the case of third generation squarks. In addition, the chapter contains a discussion of the main aspects of the strategy followed for the two searches in the thesis, including the definition of signal-enriched regions, the background estimation techniques and the statistical tools for the interpretation of the results. The first major analysis, presented in Chapter 6 and included in a recent publication [18], is a search for a pair-produced $\tilde{b}$ that decays into a b-quark and a neutralino ($\tilde{\chi}_1^0$) or a t-quark and a chargino ($\tilde{\chi}_1^+$), with small mass splitting between $\tilde{\chi}_1^0$ and $\tilde{\chi}_1^+$ as predicted by naturalness arguments. This is followed in Chapter 7 by a second analysis, also published in a recent paper [19], that targets a pair-produced $\tilde{t}$ with direct decay into a b-quark and a lepton, with variable Branching Ratio (BR) into electrons, muons and τ leptons.

Finally, it is important to emphasise that during my PhD I carried out my research as a member of the ATLAS Collaboration, which is formed by thousands of physicists from (at the time of writing) over 180 institutes. The intense and constructive cooperation with other scientists, both at the scale of the Collaboration as a whole and within the smaller analysis teams, implied that often my individual work was strictly bound to collective efforts. To help identifying my contribution towards each of the results presented in the thesis, I collected the detailed list of my activities in Appendix A.

1 | The Standard Model and Supersymmetry

This chapter provides a theoretical introduction to the physics topics discussed in the thesis. Section 1.1 presents the main features of the Standard Model of elementary particles, while Section 1.2 is dedicated to a discussion of the open problems in particle physics that are calling for an extension of the theory. Supersymmetry is introduced in Section 1.3, which includes a discussion of the main motivations for light third generation squarks at the TeV scale.

1.1 The Standard Model

The Standard Model (SM) of elementary particles is the theory that describes the phenomenology of fundamental fields and interactions. It includes all the particles that have been directly observed in high energy physics experiments, classifying them based on their mass, spin and interaction properties.

1.1.1 Quantum Field Theories

The SM is a paradigm of a Quantum Field Theory (QFT), a major theoretical construction of the 20^{th} century that unifies the principles of special relativity and quantum mechanics. Elementary particles are described as excitations of quantum fields embedded in a four-dimensional Minkowski spacetime, so they can be classified based on their behaviour under Poincaré transformations: fermions are fields with half-integer spin in units of $\hbar$, while bosons have integer spin, 0 for scalars and 1 for vectors. Another relevant property of the fields is their mass, that determines the propagation of the particles in the spacetime in absence of interactions.
The theory makes use of the Lagrangian formalism, where all the necessary information

to derive the field dynamics is included in a Lagrangian density function:

$$\mathcal{L}_{tot} = \mathcal{L}_{free} + \mathcal{L}_{int}. \tag{1.1}$$

In the above equation the full expression of $\mathcal{L}_{tot}$ is split into a kinetic term $\mathcal{L}_{free}$ and an interaction term $\mathcal{L}_{int}$, which describe respectively the spacetime evolution of the free fields and the way in which they are coupled.

1.1.2 Fundamental interactions

Once the spacetime properties of the fields are defined, the following step is to examine their interactions. At the present state of knowledge, any interaction observed in nature can be reduced to four fundamental forces: weak and strong interactions, electromagnetism and gravity. Among these only gravity is not included in the SM, but its measured strength is extremely low compared to the others, so its impact is negligible in any realistic particle physics experiment.
As discussed in detail in Ref. [20], all the interaction terms of the SM can be derived by imposing that the Lagrangian is invariant under special local transformations of the fields known as gauge transformations. When the field content of the theory is established, the choice of a particular gauge symmetry fixes the structure of the couplings between the fermions and new spin-1 vector fields, commonly referred to as gauge fields or mediators of the force. The new fields are introduced in order to preserve the invariance of the Lagrangian under the gauge transformation, so there is a direct correspondence between the imposed symmetry and the properties of the interaction. The full gauge symmetry group of the SM is

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \tag{1.2}$$

where $SU(3)_C$ is the color symmetry group of the strong interaction and $SU(2)_L \otimes U(1)_Y$ is the electroweak symmetry group.
Having identified four fundamental forces allows to add an extra classification of the elementary fields. Fermions are referred to as leptons if they only interact via electromagnetic and weak (electroweak) force, while they are called quarks if they also experience the strong interaction. A crucial difference between the two types of particles, that emerges as a consequence of the different properties of electroweak and strong force, is that only elementary leptons can be observed as free particles, while quarks are always observed in bound states called hadrons. Hadrons are known as mesons if they are

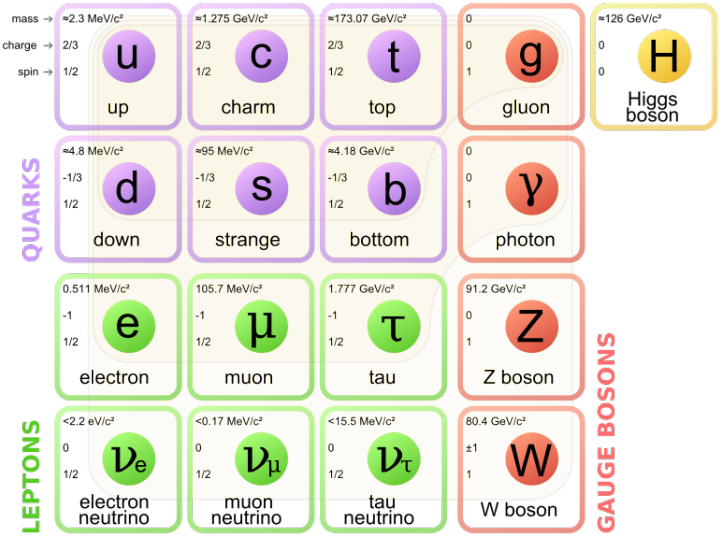


Figure 1.1: Elementary particles of the SM.

bosons, such as pions or kaons, while they are called baryons if they are fermions, with the most common examples being protons or neutrons. Figure 1.1 shows the presently known elementary particles of the SM, classifying them in quarks, leptons, gauge bosons and an additional scalar boson, the Higgs.

Electroweak force

The first fundamental interaction described by the SM is the electroweak force, which appears in the theory as a unified description of electromagnetic and weak forces. The electroweak force is based on the $SU(2)_L \otimes U(1)_Y$ symmetry group [4,21,22] introduced by Glashow, Weinberg and Salam to combine the features of Quantum Electrodynamics (QED) and of a formal description of the Fermi theory of weak interactions where the $SU(2)_L$ structure had first emerged. The L of the $SU(2)_L$ group stands for *left*: only the left chiral components of the fermion fields (spinors) carry a weak isospin charge and therefore experience the weak interaction.
When the $SU(2)_L \otimes U(1)_Y$ symmetry is imposed in the Lagrangian four vector bosons are obtained: three W_i^μ bosons originating from the $SU(2)_L$ part and one B^μ boson originating from $U(1)_Y$. Mass eigenstates are then derived by mixing the four gauge

bosons as follows:

$$\begin{aligned} W^{\mu}_{\pm} &= \frac{W^{\mu}_{1} \mp i W^{\mu}_{2}}{\sqrt{2}} \\ A^{\mu} &= B^{\mu} \cos\theta_W + W^{\mu}_{3} \sin\theta_W \\ Z^{\mu} &= W^{\mu}_{3} \cos\theta_W - B^{\mu} \sin\theta_W \end{aligned} \tag{1.3}$$

In the above equations the mediators of electromagnetic and weak interactions, the photon A, the Z and the two $W^{\pm}$ bosons, appear as linear combinations of the eigenstates of $SU(2)_L$ and $U(1)_Y$. The last two rows contain a fundamental parameter of the SM, the Weinberg angle θ_W, whose value has been accurately measured by various high energy physics experiments at different energy scales. In addition, to obtain the correct quantum numbers of each SM particle, the hypercharge Y introduced in the $SU(2)_L \otimes U(1)_Y$ group needs to be related to the electric charge Q and the weak isospin T_3 through the Gell Mann-Nishijima equation:

$$Q = \frac{Y}{2} + T_3 \tag{1.4}$$

As extensively tested in the experiments, the $W^{\pm}$ bosons carry an electric charge of $Q = \pm 1$ and are only coupled to left handed fermions through their weak isospin charge T_3. The Z boson is instead electrically neutral and couples differently to left and right handed fermions: for left handed fermions the coupling depends on both T_3 and Q, while for right handed fermions it is only proportional to Q.

Strong force

Together with the electroweak force, the SM provides a full description of the strong interaction. The corresponding field theory, Quantum Chromo Dynamics (QCD), is based on the $SU(3)_C$ symmetry group, which introduces eight vector boson mediators known as gluons [23]. The charge associated with the $SU(3)$ symmetry is the colour, C, and the elementary fermions that carry it are the quarks in Figure 1.1. Unlike photons in QED, gluons carry colour charge, reflecting the non-abelian character of $SU(3)_C$. This implies that they are allowed to couple with each other, with significant impact on the phenomenological properties of the interaction.

The QCD Lagrangian and the masses of the quarks can be used to compute the energy scale evolution of the coupling constant α_s, that can be expressed as a function of the transferred energy μ:

$$\alpha_s(\mu^2) = \frac{12\pi}{(33 - 2n_f) \log\left(\frac{\mu^2}{\Lambda^2_{QCD}}\right)} \tag{1.5}$$

where n_f is the number of quarks with mass below μ (see Ref. [24]) and $\Lambda_{QCD} \sim 200$ MeV is a characteristic scale of the theory. When the quark content of the SM is considered, Equation 1.5 shows that the coupling constant α_s decreases as a function of the characteristic scale of the interactions. On the other hand, the coupling diverges rapidly when μ decreases and approaches Λ. This implies that the theory is only perturbative at high energy scales, because the necessary condition $\alpha_s \ll 1$ is only true for $\mu \gg \Lambda_{QCD}$. In addition, three more essential features of QCD emerge:

- *Asymptotic freedom*: $\alpha_s \to 0$ as $\mu \to \infty$, so in the high energy limit the strength of the interaction becomes negligible and quarks behave as free particles [25].
- *Confinement*: at low energies (or large distances) the coupling strength increases rapidly, keeping the quarks strongly bound to each other and implying that only colour singlet states can be observed as free particles.
- *Hadronisation*: when two coloured objects are split, the extraction of a new pair of hadrons from the vacuum soon becomes energetically preferable with respect to a further increase in distance.

Hadronisation is responsible for the production of jets at high energy physics experiments, where quarks or gluons with high momentum form cascades of hadronic particles that terminate only when no more energy is left to extract more objects from the vacuum.

1.1.3 The Higgs mechanism

The theory of electroweak and strong interactions describes the experimental data with remarkable precision. However, a major limitation comes from the fact that its fundamental premises, gauge invariance and chiral symmetry, are not compatible with the presence of mass terms in the Lagrangian. If no extra feature is added, a Lagrangian with the interactions described in Section 1.1.2 predicts all the SM particles to be massless, in clear contrast with the experimental evidence.
The problem of introducing the masses of the particles without spoiling gauge and chiral symmetries can be solved by a spontaneous symmetry breaking mechanism [5–7]. The idea is to postulate the existence of an electrically neutral scalar complex field, the Higgs field, whose interaction with the other particles is responsible for their masses. The Lagrangian is expanded with an additional term

$$\mathcal{L}_\phi = (D^\mu \phi)^\dagger (D_\mu \phi) - V(\phi^\dagger \phi) \tag{1.6}$$

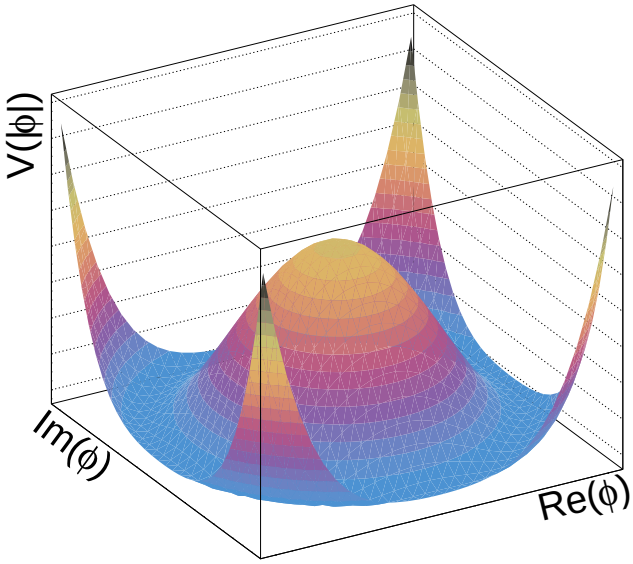


Figure 1.2: Shape of the Higgs potential for a one-dimensional complex field ϕ with $\lambda > 0$ and $\mu^2 < 0$.

where ϕ is the Higgs field, a weak isospin doublet with hypercharge $Y = 1$, and D_μ is a covariant derivative that contains the electroweak gauge fields W_i^μ and B^μ. The first component of ϕ has electric charge $+1$ while the second is neutral, and the full expression is given by

$$\phi = \begin{pmatrix} \phi^\dagger \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}. \tag{1.7}$$

The term V in the Lagrangian in Equation 1.6 is the Higgs potential:

$$V\left(\phi^\dagger\phi\right) = \mu^2\phi^\dagger\phi + \lambda\left(\phi^\dagger\phi\right)^2 \tag{1.8}$$

where μ and λ are new parameters that identify a mass term and a self-interaction term of the field ϕ. The potential is symmetric under rotations in the ϕ space and its shape depends on the sign of the two parameters. When considering $\lambda > 0$ and $\mu^2 < 0$, a characteristic shape with a continuum of minima is obtained, as shown in Figure 1.2. Due to the form of the potential, the field is forced to select an arbitrary minimum that breaks the rotational (gauge) invariance. The conventional choice is $\langle\phi_3\rangle = v$ and $\langle\phi_1\rangle = \langle\phi_2\rangle = \langle\phi_4\rangle = 0$, where v is a parameter with the dimension of an energy known as the Vacuum Expectation Value (VEV) of the Higgs field:

$$v \equiv \sqrt{\frac{-\mu^2}{\lambda}}. \tag{1.9}$$

The above choice allows to write a perturbation of the field ϕ around the minimum as

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \tag{1.10}$$

where three degrees of freedom are removed and the remaining one describes a scalar field with excitations around the VEV v. If the expression 1.10 is used in the Lagrangian 1.6, the mass terms of the bosons appear from the expansion of the covariant derivative D_μ coupled to the Higgs field, yielding:

$$m_W = \frac{gv}{2} \qquad m_Z = \frac{v}{2}\sqrt{g^2 + g'^2} \qquad m_A = 0 \tag{1.11}$$

where g and g' are the couplings strengths of the $SU(2)_L$ and $U(1)_Y$ groups, related to the Weinberg angle by

$$\sin(\theta_W) = \frac{g'}{\sqrt{g^2 + g'^2}}. \tag{1.12}$$

The spontaneous symmetry breaking mechanism predicts a relation between the masses of W and Z bosons, $m_W = m_Z \cos(\theta_W)$, that is consistent with the experimental measurements.

Once the masses of the gauge bosons are established, the following step is to consider the case of the fermions. A mass term in its simplest form, $m\overline{\psi}\psi = m(\overline{\psi}_L\psi_R + \overline{\psi}_R\psi_L)$, is not gauge invariant because the left handed components of the fermions are $SU(2)_L$ doublets while the right handed components are singlets, so they transform in a different way. The solution is to introduce the Higgs field in the term as

$$\mathcal{L}_f = -\lambda_f(\overline{\psi}_L\phi\psi_R + \overline{\psi}_R\phi^\dagger\psi_L) \tag{1.13}$$

where the gauge invariance is preserved thanks to the presence of two $SU(2)$ doublets. In the case of leptons doublets, (ν_ℓ, ℓ), the term can be rewritten using Higgs field expansion from Equation 1.10:

$$\begin{aligned} \mathcal{L}_\ell &= -\lambda_\ell \left[(\overline{\nu}_\ell\ \overline{\ell})_L \begin{pmatrix} 0 \\ v + H \end{pmatrix} \ell_R + \overline{\ell}_R\ (0, v + H) \begin{pmatrix} \nu_\ell \\ \ell \end{pmatrix}_L \right] \\ &= -\frac{\lambda_\ell v}{\sqrt{2}}\overline{\ell}\ell - \frac{\lambda_\ell H}{\sqrt{2}}\overline{\ell}\ell = -m_\ell\overline{\ell}\ell - \frac{m_\ell}{v}H\overline{\ell}\ell \end{aligned} \tag{1.14}$$

where the two terms that appear describe the mass of the leptons and their interaction with the Higgs, both related to the VEV v and the Yukawa couplings λ_ℓ. Equation 1.13 can also be used for the masses of down-type quarks, but additional terms are needed for

up-type quarks and possibly for neutrinos. The new terms can have the same structure of Equation 1.13, but they require the introduction of the complex conjugate of the Higgs field defined as:

$$\phi_c \equiv -i\tau_2\phi^\star = -\frac{1}{\sqrt{2}}\begin{pmatrix} v+H \\ 0 \end{pmatrix}. \tag{1.15}$$

In summary, the spontaneous symmetry breaking mechanism is based on the introduction of a new Higgs field with four degrees of freedom, whose definition is given in Equation 1.7. Three degrees of freedom are then used to generate the mass terms of the $W^\pm$ and Z bosons, as a result of the choice of vacuum (Equation 1.10). The remaining degree of freedom instead appears as a new scalar particle, the Higgs boson, with Yukawa couplings to the other SM particles and with mass given by the quadratic term of the potential in Equation 1.8:

$$m_H^2 = 2\mu^2 = 2\lambda v^2. \tag{1.16}$$

The observation of the Higgs boson

The discovery of a new scalar particle with mass around 125 GeV, compatible with the properties of the Higgs boson, was announced by the ATLAS [8] and CMS [9] Collaborations on July 4^{th}, 2012. Since then the Higgs has been measured in multiple decay channels, yielding no significant deviation with respect to the SM predictions. The mass is now known with a precision of about 0.2% [26, 27] and the signal strength in the different channels is consistent with the predictions for Yukawa couplings, as shown by the combined ATLAS and CMS summary plot in Figure 1.3. The couplings are compatible with a linear fit over several orders of magnitudes, providing strong evidence of the validity of the electroweak symmetry breaking mechanism. The precision of the measurements is limited by the capabilities of the LHC experiments, where some decay channels are particularly challenging due to the combination of low signal cross section and large backgrounds. Improving their accuracy is a major goal of the experimental community, because any deviation with respect to the predictions would indicate the presence of new physics Beyond the Standard Model (BSM).

1.2 The limitations of the Standard Model

The SM provides an accurate and self-consistent description of fundamental interactions, extensively tested in high energy physics experiments. Despite its major successes, how-

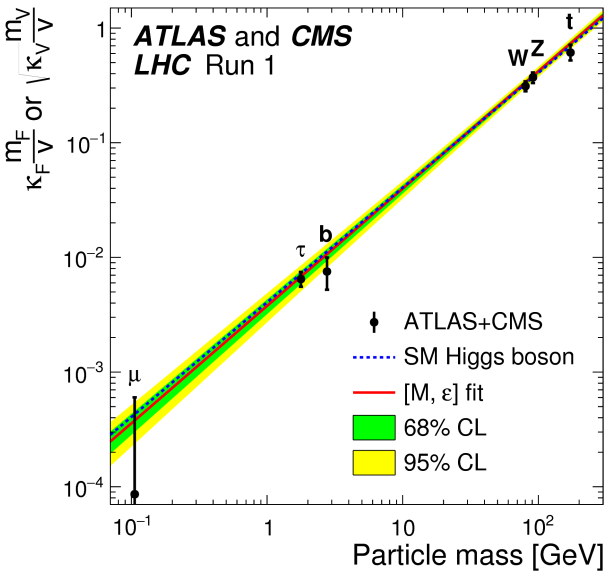


Figure 1.3: Strength of the Higgs boson couplings to the SM particles as a function of their masses [28]. The linear fit shows that the measured values are consistent with the expectations within uncertainties.

ever, there are several reasons to believe that the present theory is not complete. This section describes the most compelling issues that call for an extension of the SM, highlighting the problems related to the Higgs mass scale and the absence of a dark matter candidate.

1.2.1 The Higgs hierarchy problem

A major problem of the SM arises from the large difference between the scale of the Higgs boson mass, 10^2 GeV, and the Planck scale, 10^{19} GeV, at which quantum gravity effects are expected to dominate with respect to the other interactions [14–16]. In the theory, the value of the Higgs mass can be computed as the sum of two components:

$$m_H^2 = m_{H_0}^2 + \delta m_H^2 \tag{1.17}$$

where m_{H_0} is the bare mass, a free parameter of the Lagrangian, and δm_H is a radiative correction that depends on the Higgs couplings to other particles. Due to the structure of the Yukawa couplings, it is easy to show that a major correction is given by the diagram in Figure 1.4, where a fermion loop is added to the Higgs propagator. For a fermion with

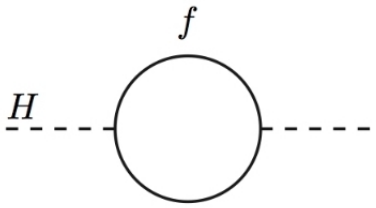


Figure 1.4: Fermion loop correction to the Higgs boson mass.

coupling λ_f the size of the correction is given by

$$\delta m_H^2|_f = -\frac{|\lambda_f|}{8\pi^2}\Lambda_{\rm NP}^2 + ... \tag{1.18}$$

where $\Lambda_{\rm NP}$ is the highest mass scale in the theory. Due to the quadratic dependence on $\Lambda_{\rm NP}$, the correction δm_H is strongly sensitive to any particle with mass beyond the electroweak scale. In particular, the reasonable assumption that new physics exists at the Planck scale implies that δm_H is of the order of 10^{19} GeV, pushing m_H far from the measured value of ~ 125 GeV. The only way to bring the value of m_H to the electroweak scale is to fine tune the bare parameter m_{H_0} in Equation 1.17 such that it almost perfectly cancels the large radiative correction. The Higgs mass is thus obtained as the difference between two terms of 10^{19} GeV each.

The large discrepancy between the measured scale of the Higgs boson mass and the Planck scale is known as the Higgs hierarchy problem. The SM solution based on the tuning of m_{H_0} is in contrast with the argument of naturalness, according to which these fine cancellations should not occur unless they arise from a specific feature of the theory. For example, the light masses of fermions and vector bosons are not concerning because their value is protected by chiral and gauge symmetries, which imply that any radiative correction can only be proportional the mass itself ($\delta m_\psi \propto m_\psi$ and $\delta A^2 \propto A^2$). The case of the scalar Higgs boson is special because no symmetry is protecting the mass, which can then receive larger corrections as the one in Equation 1.18. A possible solution to the problem is to postulate the existence of an extra symmetry that provides a natural explanation to the discrepancy between the electroweak and Planck scales. This is what happens in supersymmetric extensions of the SM, as described in Section 1.3.3.

1.2.2 Dark matter

Astrophysical observations have identified a variety of systems where the amount of ordinary matter, estimated through the light that it emits, is incompatible with the gravitational patterns. These results indicate the presence of an extra type of matter, referred to as dark matter, that must contribute to the gravitational effects without yielding any extra luminosity [29,30]. The existence of dark matter is further supported by additional observations, in particular by the measurements of the cosmic microwave background carried out by WMAP [31] and Planck [32]. A favoured hypothesis for the nature of dark matter is that it is composed of Weakly Interacting Massive Particles (WIMPs) [33], and the fact that none of the SM particles is able to provide a viable WIMP candidate is an indication that the present theory is incomplete. Searches for dark matter can be conducted with different approaches, including collider experiments where the new particles may be produced in exotic physics processes. In particular, Supersymmetric theories can embed dark matter candidates in several ways, making them extremely interesting to search for at the LHC.

1.2.3 More open issues

Along with the hierarchy problem and the absence of a dark matter candidate, there is more evidence that motivates the need for BSM physics. In the list below I briefly discuss some open issues that are being addressed by the scientific community:

- *Grand Unification:* Following the success of the electroweak theory, where weak force and electromagnetism are successfully described as the low energy limit of a single interaction, theorists have been exploring the possibility of a further unification that incorporates also the strong force. For this purpose, it is necessary to examine the evolution (*running*) of the coupling constants of each SM interaction as a function of the energy scale [34]. As seen in Section 1.1.2 in the case of α_s (Equation 1.5), the value of the constants has a non-trivial dependence on the energy scale determined by the number of degrees of freedom of the theory. In the SM, the evolution of the coupling constants of electromagnetic, weak and strong interactions is not converging to a common value, as shown by Figure 1.5a. This is in contrast with the idea of a unification of the interactions, that requires a unique coupling strength at a Grand Unification scale μ_{GUT}. The necessary behaviour can be recovered by introducing SUSY, which varies the particle content of the theory

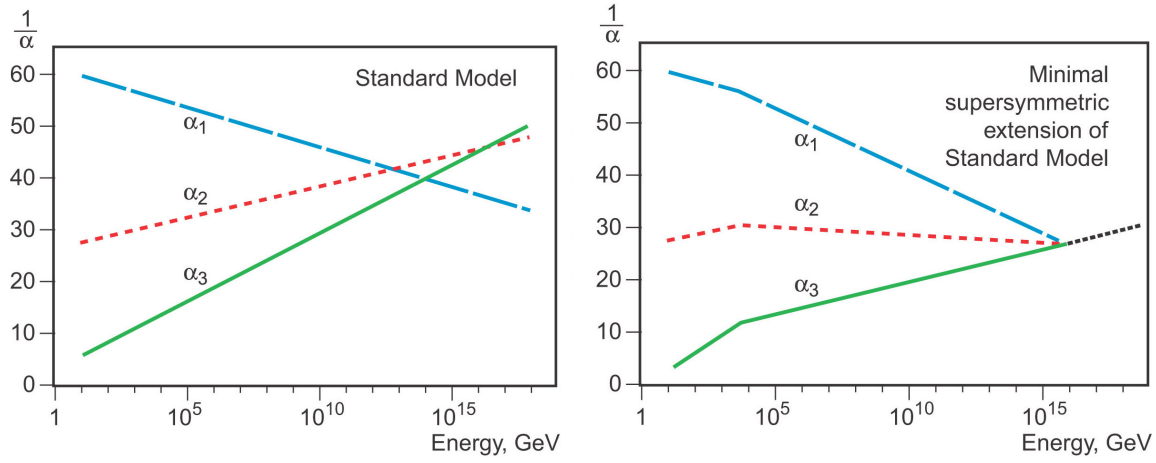


Figure 1.5: Evolution of the coupling constants of electromagnetic, weak and strong interactions as a function of energy in the SM (a) and in a supersymmetric extension (b) [35].

and modifies the scale dependence of the constants such that the three curves can intersect (Figure 1.5b).

- *Neutrino sector:* The SM predicts the existence of three types of massless neutrinos that can only convert into their corresponding charged lepton via electroweak interactions. However, the recent discovery of neutrino oscillations violates lepton flavour conservation and indicates that at least two of the three species must have non-zero mass [36]. Both these properties imply that the neutrino sector of the SM is not complete, calling for a deeper understanding of the underlying physics. Assuming a total of 3 neutrino flavours, the relation between mass and interaction eigenstates in the neutrino sector is described by the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) 3×3 matrix [37], which can be parametrised in terms of mixing angles between flavours (θ_{12}, θ_{13}, θ_{23}) and a variable number of extra phases depending on whether neutrinos are considered as Dirac or Majorana fermions [38]. Several experiments are targeting a full measurement of the neutrino parameters, and the most recent results are summarised by NuFIT-v3.2 [39, 40]. The ordering of the masses of the three neutrinos is still not known, and two possibilities referred to as normal hierarchy and inverted hierarchy remain viable based on the experimental data. Similarly, the parameters of the PMNS matrix are still not fully determined, and their values depend on the assumption on the mass hierarchy.

- *CP violation:* In the SM the only measured source of CP violation is a complex phase in the Cabibbo-Kobayashi-Maskawa (CKM) matrix [41], which is not sufficient to explain the relative abundance of matter and antimatter in the universe [42]. In order to solve this puzzle, additional sources of CP violation are expected to appear in new physics scenarios.

- *Gravity:* As previously discussed in this chapter, no description of gravity is included in the SM.

1.3 Supersymmetry

This section provides an introduction to supersymmetric extensions of the SM [12, 13], showing that they can solve many of the outstanding issues in high energy physics. The section begins with a theoretical introduction to supersymmetric Lagrangians in Section 1.3.1, while Section 1.3.2 introduces the Minimal Supersymmetric Standard Model (MSSM) and its particle content. Finally, Sections 1.3.3 and 1.3.4 describe in detail the classes of SUSY models that are relevant for the analyses presented in this thesis.

1.3.1 Supersymmetric theories

Supersymmetry (SUSY) is a type of spacetime symmetry that extends the Poincaré group of the SM with a new transformation between bosonic and fermionic states. In the SM two distinct types of symmetry are implemented:

- *Spacetime symmetries*, that form the Poincaré group and cause the conservation of four-momentum and angular momentum.

- *Gauge symmetries* or *local symmetries*, that result in the conservation of the associated charges.

A fundamental theorem by Coleman and Mandula [43] states that local and spacetime symmetries can only be combined as direct products, as it happens in the SM, otherwise the theory would be unable to predict scattering amplitudes with non-zero probability. This seems to imply that there is no way to introduce new hybrid types of symmetries, because the interactions between the fields would no longer be described. However, it can be shown that the Coleman-Mandula theorem does not apply to symmetries generated by fermionic transformations, where an operator Q acts on a state and modifies its

spin by $\frac{1}{2}$. These special types of symmetries are known as supersymmetries, and the corresponding transformations are given by:

$$\begin{aligned} Q\,|\text{boson}\rangle &= |\text{fermion}\rangle \\ Q\,|\text{fermion}\rangle &= |\text{boson}\rangle \end{aligned} \tag{1.19}$$

The spin difference between the initial and final state of the transformation implies that the operator Q must be a spin-$\frac{1}{2}$ object. If we let Q_α ($\alpha = 1, 2$) and $Q^\dagger_\beta$ be the components of the generator Q and of its hermitian conjugate $Q^\dagger$, their anticommutator is given by

$$\{Q_\alpha, Q^\dagger_\beta\} = 2\gamma^\mu_{\alpha\beta} p_\mu \tag{1.20}$$

where p_μ is the Lorentz four-momentum and γ^μ are the Dirac matrices. Other important commutators are:

$$\begin{aligned} [M^{\rho\sigma}, Q_\alpha] &= -i\,(\sigma^{\rho\sigma})^\beta_\alpha\, Q_\beta \\ \{Q_\alpha, Q_\beta\} &= \{Q^\dagger_\alpha, Q^\dagger_\beta\} = 0 \\ [p^\mu, Q_\alpha] &= \left[p^\mu, Q^\dagger_\alpha\right] = 0 \end{aligned} \tag{1.21}$$

where $M^{\rho\sigma}$ is the generator of the Lorentz transformations. From the last line it can be seen that the generator Q commutes with $p^2 = m^2$, which implies that fermion-boson pairs related by the transformation have the same mass. In supersymmetric extensions of the SM the commutators of Q with the generators of the gauge transformations must also vanish, in order to preserve the existing structure of the theory. Fermions and bosons related by Q must then have the same $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ quantum numbers, which is not the case for any pair of particles in the SM, indicating that none of the known particles can be regarded as the superpartner of any other. Since SUSY needs to apply to all fields in the theory, it follows that minimal extensions of the SM require the introduction of a new superpartner for each particle, incrementing the number of elementary states by at least a factor 2.

Superfields

The SM particles and their superpartners are organised in supermultiplets that contain two bosonic and two fermionic degrees of freedom. The fermions are described by Weyl spinors [44] with two polarisations, and for each of them there is a corresponding bosonic state. When the operators Q and $Q^\dagger$ act on the supermultiplets, they transform the bosonic components into the fermionic ones and vice versa. Depending on the spin of the particles involved, the supermultiplets can be classified as follows:

Chiral supermultiplets contain spin-$\frac{1}{2}$ fermions with left and right handed components, f_L and f_R, and two spin-0 partners known as scalar fermions or sfermions. The two scalars are conventionally labelled $\tilde{f}_L$ and $\tilde{f}_R$, where the tilde is used to identify SUSY counterparts of SM particles and the L and R labels are referred to the chirality of the partners.

Gauge supermultiplets contain vector bosons, assumed to be massless, and two spin-$\frac{1}{2}$ Weyl fermions known as gauginos.

Gravitational supermultiplets include the spin-2 graviton and a spin-$\frac{3}{2}$ partner, the gravitino.

1.3.2 The Minimal Supersymmetric Standard Model

This paragraph describes the simplest supersymmetric extension of the SM, known as Minimal Supersymmetric Standard Model (MSSM), where a new superpartner is added for each SM particle. The MSSM preserves the $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ gauge symmetry of the Lagrangian, so the quantum numbers of particles and superpartners are the same. The field content of the theory is the following:

- A chiral supermultiplet $q_{L,R}$-$\tilde{q}_{L,R}$ for each quark in the SM.
- A chiral supermultiplet $\ell_{L,R}$-$\tilde{\ell}_{L,R}$ for each lepton family.
- A gauge supermultiplet for each vector boson in the SM and its spin-$\frac{1}{2}$ partner: B_0-$\tilde{B}_0$, W_i-$\tilde{W}_i$, g-$\tilde{g}$.
- Two chiral supermultiplets, H_u-$\tilde{H}_u$ and H_d-$\tilde{H}_d$, that embed the SM Higgs and give mass to bosons and fermions through the usual spontaneous symmetry breaking mechanism. Among the total of eight degrees of freedom in H_u and H_d, only three are involved in the symmetry breaking, so the remaining five can appear as observable states:
  - $H^{\pm}$: two charged Higgs states.
  - A^0: a CP-odd neutral Higgs.
  - h^0 and H^0: two neutral CP-even Higgs fields, one of which must be the SM Higgs boson.

SUSY Lagrangian and R-parity

Once the field content is defined it is possible to examine the most general expression of the SUSY Lagrangian, that contains a larger number of terms with respect to the SM. Some of these terms violate the conservation of lepton and baryon number, which is instead prohibited in the SM, so they enable processes like proton decay that are heavily constrained by experimental data. In order to remove them from the SUSY Lagrangian, a solution is to introduce a new quantum number known as R-parity [45] that can be expressed as a function of the baryon and lepton numbers, B and L, and of the spin S:

$$P_R \equiv (-1)^{3(B-L)+2S}. \tag{1.22}$$

With the above definition, R-parity provides a simple classification of the particles:

- SM particles $\rightarrow P_R = 1$
- SUSY particles $\rightarrow P_R = -1$

In SUSY models where R-parity is conserved, referred to as R-Parity Conserving (RPC), the Lagrangian terms that do not conserve lepton and baryon number are automatically forbidden. The structure of the remaining terms implies that any Feynman diagram must only include vertices with an even number of SUSY particles, with relevant consequences on the phenomenology:

- SUSY particles must be produced in pairs.
- The decay of a SUSY particle must contain an odd number of SUSY particles.
- The Lightest Supersymmetric Particle (LSP) must be stable.

Finally, by analogy with the RPC models described above, SUSY models where R-parity is not conserved are referred to as R-Parity Violating (RPV).

SUSY breaking

As seen in Section 1.3.1, supersymmetric extensions of the SM require that particles and superpartners are mass-degenerate. This is clearly incompatible with the experimental data, because so far no evidence of SUSY particles has been found. The lack of observations indicates that SUSY, if it exists, must be broken in a way that makes the superpartners significantly heavier than the SM particles.

In the absence of experimental hints, the mechanism responsible for SUSY breaking can only be guessed based on generic arguments. As discussed in Ref. [46], a spontaneous breaking is not possible within the framework of the MSSM without the introduction of at least one extra field, with little or no indication about its properties. For this reason, a standard approach is to keep the field content of the theory to the minimum and introduce extra terms that break the symmetry explicitly, obtaining a Lagrangian of the form:

$$\mathcal{L} = \mathcal{L}_{\text{SUSY}} + \mathcal{L}_{\text{soft}} \tag{1.23}$$

where the *soft* label indicates that the new terms should only be regarded as a perturbation of the supersymmetric part of the Lagrangian. In practice, the soft term is parametrising our ignorance about the exact mechanism responsible for the SUSY breaking, describing its consequences without specifying its origin. The inclusion of the soft term increases the number of free parameters of the theory to 105 in addition to the ones of the SM, and the phenomenology of the models can vary dramatically depending on their values. A common approach is to identify classes of SUSY models based on assumptions that constrain specific sets of parameters, as it is done in the two cases described in Sections 1.3.3 and 1.3.4.

Mass eigenstates

As a result of SUSY breaking, the masses of the superpartners become free parameters and can be different from those of the SM particles. SUSY particles with the same quantum numbers can also mix, so the mass eigenstates are not necessarily the same as the interaction eigenstates introduced in the previous paragraph. The list below provides an overview of the mass eigenstates in the MSSM, that are also summarised in Table 1.1:

Sleptons and squarks Contrarily to what happens for standard particles, the left and right handed superpartners of quarks and leptons are not constrained to have equal mass. It is hence possible to define a mixing matrix that rotates the interactions eigenstates $\tilde{q}_L$-$\tilde{q}_R$ into mass eigenstates $\tilde{q}_1$-$\tilde{q}_2$, where by convention $\tilde{q}_1$ is the lightest. The structure of the mixing matrices is similar for squark and sleptons, and its expression in the case of third generation squarks [47] is given by:

$$\mathcal{M}^2_{\tilde{q}} = \begin{pmatrix} m^2_{\tilde{q}_L} & a_q m_q \\ a_q m_q & m^2_{\tilde{q}_R} \end{pmatrix} \tag{1.24}$$

with

$$m^2_{\tilde{q}_L} = M^2_{Q_3} + m^2_Z \cos 2\beta \left(I^{q_L}_3 - e_q \sin^2 \theta_W\right) + m^2_q, \quad (1.25)$$

$$m^2_{\tilde{q}_R} = M^2_{\{U,D\}_3} + m^2_Z \cos 2\beta \, e_q \sin^2 \theta_W + m^2_q, \quad (1.26)$$

$$a_q m_q = \begin{cases} (A_t - \mu \cot \beta)\, m_t \;\; (\tilde{q} = \tilde{t}) \\ (A_b - \mu \tan \beta)\, m_b \;\; (\tilde{q} = \tilde{b})\,. \end{cases} \quad (1.27)$$

In the above equations I^q_3 is the third component of the weak isospin, e_q the fractional electric charge of the quark q and m_Z is the mass of the SM Z boson. In addition, μ is the Higgsino mass parameter in the MSSM and $\tan\beta$ is the ratio between the VEVs of H_u and H_d. The remaining parameters are introduced in $\mathcal{L}_{\text{soft}}$: M_{Q_3,U_3,D_3} are the third components of the mass matrices of left and right handed squarks, corresponding to the third generation, while $A_{t,b}$ are the coefficients of their trilinear coupling terms (see Ref. [48]). The impact of the mixing is particularly relevant for the partners of the heavy flavour families, stop ($\tilde{t}_{1,2}$) and sbottom ($\tilde{b}_{1,2}$), due to the presence of the mass of the SM fermions in the off-diagonal terms of the matrix.

Gluinos Since gluons and gluinos carry colour charge, no mixing can occur and the mass eigenstates correspond to the interaction eigenstates.

Neutralinos and charginos The charged Higgs states mix with the superpartners of the gauge bosons, the gauginos, yielding two pairs of positively or negatively charged spin-$\frac{1}{2}$ mass eigenstates known as charginos ($\tilde{\chi}^{\pm}_i$, with $i = 1, 2$). The mixing matrix of charginos is given by:

$$\begin{pmatrix} M_2 & \sqrt{2} m_W \sin\beta \\ \sqrt{2} m_W \cos\beta & \mu \end{pmatrix} \quad (1.28)$$

in the $(\tilde{W}^{\pm}, \tilde{H}^{\pm})$ basis. In Equation 1.28 M_1 and M_2 are gaugino mass parameters from $\mathcal{L}_{\text{soft}}$, m_W is the mass of the W boson, and β and μ are defined as in Equations 1.24-1.27.

In a similar way, four neutral spin-$\frac{1}{2}$ particles called neutralinos ($\tilde{\chi}^0_i$, with $i = 1, 2, 3, 4$) result from the mixing of the neutral interaction eigenstates from the Higgs and gauge sectors. The 4×4 mixing matrix in the $(\tilde{B}, \tilde{W}^0, \tilde{H}_d, \tilde{H}_u)$ basis is

Names	Spin	P_R	Gauge Eigenstates	Mass Eigenstates
Higgs bosons	0	$+1$	$H_u^0\ H_d^0\ H_u^+\ H_d^-$	$h^0\ H^0\ A^0\ H^\pm$
squarks	0	-1	$\tilde{u}_L\ \tilde{u}_R\ \tilde{d}_L\ \tilde{d}_R$ $\tilde{s}_L\ \tilde{s}_R\ \tilde{c}_L\ \tilde{c}_R$ $\tilde{t}_L\ \tilde{t}_R\ \tilde{b}_L\ \tilde{b}_R$	(same) (same) $\tilde{t}_1\ \tilde{t}_2\ \tilde{b}_1\ \tilde{b}_2$
sleptons	0	-1	$\tilde{e}_L\ \tilde{e}_R\ \tilde{\nu}_e$ $\tilde{\mu}_L\ \tilde{\mu}_R\ \tilde{\nu}_\mu$ $\tilde{\tau}_L\ \tilde{\tau}_R\ \tilde{\nu}_\tau$	(same) (same) $\tilde{\tau}_1\ \tilde{\tau}_2\ \tilde{\nu}_\tau$
neutralinos	$1/2$	-1	$\tilde{B}^0\ \tilde{W}^0\ \tilde{H}_u^0\ \tilde{H}_d^0$	$\tilde{\chi}_1\ \tilde{\chi}_2\ \tilde{\chi}_3\ \tilde{\chi}_4$
charginos	$1/2$	-1	$\tilde{W}^\pm\ \tilde{H}_u^+\ \tilde{H}_d^-$	$\tilde{\chi}_1^\pm\ \tilde{\chi}_2^\pm$
gluino	$1/2$	-1	$\tilde{g}$	(same)

Table 1.1: SUSY particles and Higgs sector in the MSSM [49].

given by:

$$\begin{pmatrix} M_1 & 0 & -m_Z\cos\beta\sin\theta_W & m_Z\sin\beta\sin\theta_W \\ 0 & M_2 & m_Z\cos\beta\cos\theta_W & m_Z\sin\beta\cos\theta_W \\ -m_Z\cos\beta\sin\theta_W & m_Z\cos\beta\cos\theta_W & 0 & \mu \\ m_Z\sin\beta\sin\theta_W & -m_Z\sin\beta\cos\theta_W & -\mu & 0 \end{pmatrix} \tag{1.29}$$

where the same parameters of the previously defined matrices are appearing.

Phenomenological MSSM

The number of free parameters in the MSSM after SUSY breaking can be reduced under the following assumptions:

- The only source of CP violation in the theory is the CKM matrix.
- No flavour changing neutral currents exist at tree level, implying that the sfermion mass matrices and trilinear coupling matrices must be diagonal.
- Universality applies to the first and second generation of the sfermions, whose masses are assumed to be degenerate.

As a result, the free parameters of the theory can be reduced to 19: the ratio $\tan\beta$ between the VEVs of the two Higgs doublets, the mass of the pseudoscalar Higgs M_A,

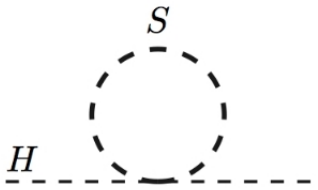


Figure 1.6: Scalar loop correction to the Higgs mass that cancels the fermion loop in Figure 1.4.

the Higgs-Higgsino mass parameter μ, five mass parameters for 1^{st} and 2^{nd} generation sfermions and five more for 3^{rd} generation sfermions, three mass parameters for gluino, Bino and Wino and three more parameters for third generation trilinear couplings. This simplified version of the theory is known as phenomenological MSSM (pMSSM).

1.3.3 SUSY as a solution to the hierarchy problem

The MSSM can address many of the outstanding issues of the SM, including the Higgs hierarchy problem [50]. As discussed in Section 1.2.1, the problem originates from the divergent radiative corrections to the Higgs mass shown in Figure 1.4, that make the value of m_H extremely sensitive to any new physics at high energy scales. SUSY introduces additional loops [49] involving the scalar partners of the fermions, as the one in Figure 1.6, so an extra radiative correction appears for both right and left handed superpartners:

$$\delta m_H^2|_s = +\frac{|\lambda_s|}{16\pi^2}\Lambda_{\rm NP}^2 + ... \tag{1.30}$$

where λ_s is the Yukawa coupling of the sfermions to the Higgs, and SUSY implies that $\lambda_s \equiv \lambda_f$ for each fermion-sfermion pair. As a result, the two scalar contributions cancel exactly the divergent correction caused by the fermion loop, whose expression is defined in Equation 1.18. This implies that even hypothetical fermions with mass of the order of the Planck scale would not have an impact on the Higgs mass, so the measured value $m_H \sim 125$ GeV can be obtained without any unnatural tuning of the parameters. In other words, SUSY is forcing λ_f and λ_s to be the same and, by doing so, it is removing the quadratic divergence in $\Lambda_{\rm NP}$.

At higher order in perturbation theory, it can be shown that both fermions and scalars coupled to the Higgs yield corrections of the form:

$$\delta m_H^2|_X = \pm\frac{\lambda_X}{16\pi^2} m_X^2 \log\left(\frac{\Lambda_{\rm NP}}{m_X^2}\right) + ..., \quad X = f, s \tag{1.31}$$

where there $\Lambda_{\rm NP}$ appears in a logarithm but a quadratic dependence on the mass is introduced, and the sign is again opposite for scalars and fermions. In realistic models where SUSY is broken the above terms do not cancel, yielding a non-zero correction to the Higgs mass due to the mass difference between the standard particles and their partners:

$$\delta m_H^2|_{s+f} = \frac{\lambda}{16\pi^2}\left[m_f^2 \log\left(\frac{\Lambda_{\rm NP}}{m_f^2}\right) - m_s^2 \log\left(\frac{\Lambda_{\rm NP}}{m_s^2}\right)\right] \tag{1.32}$$

In order to keep the size of the correction under control, the difference between the masses of standard and SUSY particles with large couplings λ to the Higgs must not be too large. The exact threshold on the mass difference is arbitrary, and is determined by the amount of fine tuning that one is willing to accept without regarding the model as unnatural.

In the case of the top quark and its partner, the stop ($\tilde{t}$), the correction in Equation 1.32 can be expressed as a function of the parameters of $\mathcal{L}_{\rm soft}$ [17]:

$$\delta m_H^2|_{stop} \simeq -\frac{3y_t^2}{8\pi^2}\left(m_{Q_3}^2 + m_{U_3}^2 + |A_t|^2\right)\log\left(\frac{\Lambda_{\rm NP}}{\rm TeV}\right) \tag{1.33}$$

where y_t is the Yukawa coupling of the top quark and the dependence on the mass in the logarithm is neglected. In addition, as introduced in Equations 1.24-1.27, the m_{Q_3} and m_{U_3} parameters are the third components of the mass matrices of left and right squarks, while A_t is a trilinear coupling coefficient.

In the MSSM, an even stronger constraint is placed on the Higgsino mass parameter μ, which is related to the Higgs mass at tree level and therefore should be close to the measured value of m_H:

$$\frac{m_H^2}{2} = -|\mu|^2 + ... + \delta m_H^2. \tag{1.34}$$

Finally, the gluino mass parameter M_3 is also constrained because it enters a 2-loop correction to the Higgs mass:

$$\delta m_H^2|_{gluino} \simeq -\frac{2y_t^2}{\pi^2}\left(\frac{\alpha_s}{\pi}\right)|M_3|^2 \log^2\left(\frac{\Lambda_{\rm NP}}{\rm TeV}\right). \tag{1.35}$$

The above arguments can be used to classify the mass eigenstates of the MSSM. Higgsinos, third generation squarks and gluinos form the natural SUSY spectrum, while the rest of the sparticles can in principle be heavier without spoiling the naturalness of the theory, so they are referred to as decoupled SUSY.

The full spectrum of a natural MSSM model is shown in Figure 1.7. If the parameter

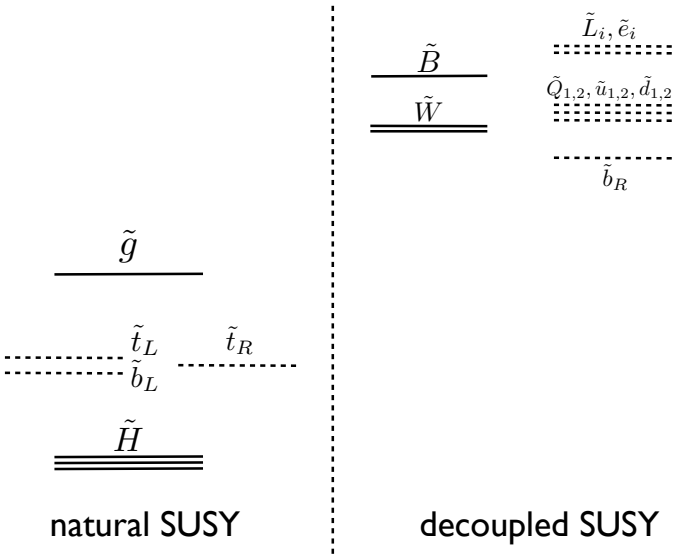


Figure 1.7: Natural MSSM mass spectrum [17]. The particles on the left side are constrained by naturalness arguments, while the remaining ones are decoupled.

μ in the mixing matrices of charginos and neutralinos (Equations 1.28-1.29) is assumed to be significantly smaller than M_1 and M_2, three light and semi-degenerate Higgsino-like mass eigenstates are predicted. Due to the structure of the squark mixing matrix (Equation 1.24) the constraints on m_{Q_3} m_{U_3} imply that the left handed third generation squarks and the right handed $\tilde{t}$ must also be light, with favoured decays into SM top or bottom quarks plus one of the three Higgsino-like states. Gluinos are light because of the constraint on M_3, while the lightest Higgsino-like state is assumed to be the LSP and provides a good candidate for dark matter.
Natural SUSY models are targeted by several SUSY searches in ATLAS, including the analysis presented in Chapter 6 of this thesis.

1.3.4 SUSY models with spontaneous R-parity breaking

The imposition of R-parity, as described in Section 1.3.2, is an *ad-hoc* adjustment to the theory which is needed to remove unwanted terms from the Lagrangian. A formal way to embed it in the MSSM is to enlarge the gauge symmetry group to $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \otimes U(1)_{B-L}$, where the additional $U(1)_{B-L}$ group naturally includes R-parity conservation. The caveat is that the $U(1)_{B-L}$ symmetry is not observed at the electroweak

scale, so a breaking mechanism must be assumed and RPV terms can appear as a consequence.
A minimal way to break $U(1)_{B-L}$ spontaneously, without spoiling experimental constraints such as the limits on the proton decay, is to assume that right-handed scalar neutrinos (sneutrinos) develop a non-zero VEV [51]. This gives rise to a class of models that can be referred to as minimal $B-L$ MSSM, with at least three major phenomenological features:

- The introduction of RPV terms implies that the LSP can carry colour or electric charge without causing a conflict with astrophysical data, because a prompt decay into SM particles is possible.
- Since the spontaneous breaking of $U(1)_{B-L}$ happens through the VEV of right handed sneutrinos, the branching ratio (BR) of the LSP decays can be directly related to the neutrino mass hierarchy and to the value of the mixing angle θ_{23}, as further described below.
- Once the LSP is chosen, its decays are fixed by the theory as a result of the coupling structure of the minimal MSSM.

The nature of the LSP is not determined by the properties of the minimal $B-L$ MSSM, but the distinctive trait of the theory is that it allows such particle to be colour charged. In this context, as described in detail in Ref. [51], third generation squarks are particularly interesting LSP candidates to be searched for at the LHC experiments. Considering the case of the stop ($\tilde{t}$), the decay mode determined by the structure of the MSSM couplings is the following:

$$\tilde{t}_1 \to b\ell, \quad \ell = e, \mu, \tau \tag{1.36}$$

where the BR into electrons, muons and τ leptons depends on the many free parameters of the theory.
As anticipated, it is possible to correlate the BR of the $\tilde{t}$ decay with the neutrino mass hierarchy and the mixing parameter $\sin^2(\theta_{23})$. Figure 1.8 (from Ref. [51]) shows a full scan of the $\tilde{t}$ BRs in the BR(e)-BR(τ) plane as a function of the free parameters of the theory, considering normal and inverted neutrino mass hierarchies and using two different values of $\sin^2(\theta_{23})$ derived from NuFIT-v1.2 [39, 52]. Interestingly, the results of the scan imply that the experimental measurement of the BR of a $\tilde{t}$ LSP within the minimal $B-L$ MSSM would provide useful indications to discriminate between the neutrino mass hierarchies and the different possible values of $\sin^2(\theta_{23})$.

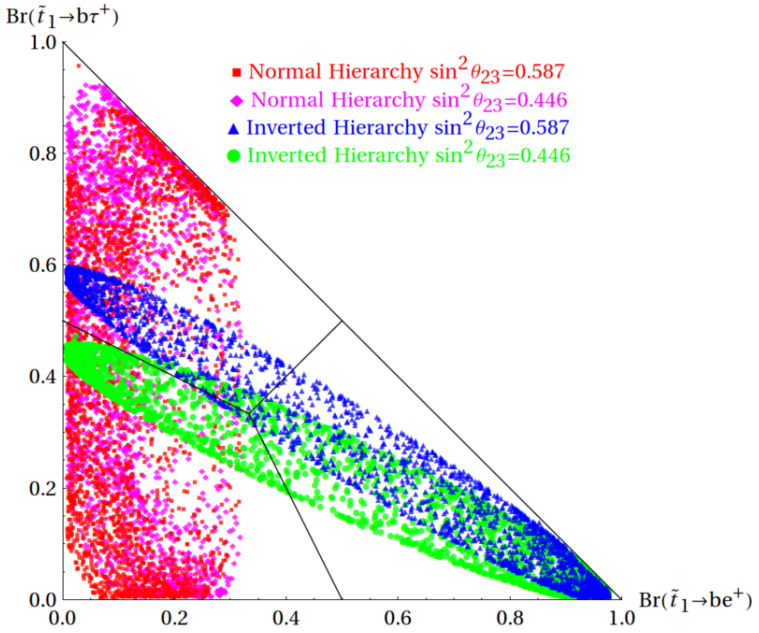


Figure 1.8: Scan of the BR of the $\tilde{t}$ decay into b-e, b-μ and b-τ as a function of free parameters of the minimal $B-L$ MSSM, for a fixed mass hierarchy and for a fixed value of $\sin^2(\theta_{23})$ [51]. The scan parameters in the plot include the mass of the $\tilde{t}$, ranging from 400 GeV to 1 TeV. The values of $\sin^2(\theta_{23})$ are taken from NuFIT-v1.2 [39, 52].

The signal model described in this paragraph is targeted by a dedicated search presented in Chapter 7.

2 | LHC and ATLAS

The present thesis focuses on the analysis of proton-proton collisions delivered by the LHC accelerator [1] at CERN during the 2015 and 2016 operation at $\sqrt{s} = 13$ TeV, and recorded by the ATLAS detector [10]. This chapter provides an overview of the experimental setup, starting from a discussion of the CERN accelerator complex (Section 2.1) and the experiments served by the LHC (Section 2.2), and continuing with a description of ATLAS and its components (Section 2.3).

2.1 The Large Hadron Collider

The European Organisation for Nuclear Research (CERN) hosts the largest particle accelerator in the world, the Large Hadron Collider (LHC), in a 27 km long underground tunnel situated at the border between Switzerland and France. The LHC is able to accelerate two beams of protons in opposite direction, delivering collisions at a maximum centre of mass energy $\sqrt{s} = 14$ TeV[1]. The collisions take place in four interaction points, where particle physics detectors are employed to reveal and measure their products.

2.1.1 Particle accelerators

In a circular collider, the trajectory of the beams is maintained primarily by an array of dipole magnets that produce a bending field orthogonal to their direction of motion. The relativistic relation

$$p = 0.3BR \tag{2.1}$$

expresses the momentum p of the particles (in GeV) as a function of the magnetic field B (in Tesla) and the radius R of the accelerator (in meters). The rate of physics collisions

[1]The design energy $\sqrt{s} = 14$ TeV has not been reached yet due to technical limitations. In the data taking period relevant for this thesis, the LHC operated at $\sqrt{s} = 13$ TeV.

delivered by the accelerator is measured by the instantaneous luminosity:

$$\mathcal{L}_{\text{inst}} \equiv \frac{1}{\sigma}\frac{dN}{dt} \tag{2.2}$$

where dN/dt is the rate of occurrence of a benchmark physics process and σ is its total cross section. The istantaneous luminosity is measured in inverse barn per second ($b^{-1}s^{-1}$ where $1\ b \equiv 10^{-24}$ cm^2) and its value is independent of the particular process considered, so it can be expressed as a function of beam parameters:

$$\mathcal{L}_{\text{inst}} = f\frac{Nn^2}{4\pi\sigma_x\sigma_y}. \tag{2.3}$$

In the above expression f is the revolution frequency, N is the number of particle bunches in which the two beams are divided (assumed to be equal for both), n is the number of particles in each bunch and $4\pi\sigma_x\sigma_y$ is the transverse area of the bunches at the interaction point, described by the gaussian widths σ_x and σ_y. The total amount of data delivered by the accelerator is given by the integrated luminosity:

$$\mathcal{L}_{\text{tot}} = \int_{t_1}^{t_2} \mathcal{L}_{\text{inst}}\, dt \tag{2.4}$$

with integral taken over the relevant period of operation. One of the advantages of circular colliders is that the beams are kept in their trajectory for long periods of time, so the individual proton bunches can be repeatedly used for collisions instead of decelerating them after a single crossing.

The main limitation of circular colliders is the synchrotron radiation that charged particles emit when they experience transverse acceleration. The associated rate of energy loss can be expressed as

$$\frac{dE}{dt} = k\frac{E^4}{m^4R^2} \tag{2.5}$$

where k is a dimesional constant, m is the mass of the accelerated particles and R is the radius of curvature of their trajectory. Since it is inversely proportional to m^4, the energy loss is extremely significant for light particles such as electrons or positrons, while for protons it is reduced by a factor $(m_e/m_p)^4 \sim 10^{-12}$. At electron-positron circular colliders, synchrotron radiation is the main effect that determines the maximum energy reach. Hadron colliders are less affected by this phenomenon, and the limiting factor for their energy is the maximum magnetic field that can be produced by the bending magnets.

2.1.2 Physics at hadron colliders

Circular accelerators are designed to provide head-on collisions of two beams, that are generally organised in particle bunches as already introduced in Section 2.1.1. Each intersection between two bunches, or *bunch crossing*, generates a variable number of interactions resulting in the production of an ensemble of collision remnants known as *final state* objects, that can be detected by the high energy physics experiments. The collection of final state objects corresponding to a given bunch crossing is commonly referred to as a physics *event*.

At highly energetic hadron colliders such as the LHC, the physics processes of interest are produced by the hard scattering of partons (quarks or gluons) that carry an unknown fraction x_1 and x_2 of the hadron momentum [53], as described by the scheme in Figure 2.1. If the two beams have the same energy E and the colliding particles are considered as a whole, the centre of mass frame of the interactions corresponds to the laboratory frame and the total energy of the collisions is simply given by $\sqrt{s} = 2E$. However, since the hard scattering interaction involves individual partons, the *partonic* centre of mass energy $\sqrt{\hat{s}}$ of each specific process is smaller than the energy of the full hadron-hadron system. The quantities s and $\hat{s}$ are related by:

$$\hat{s} = x_1 x_2 s \tag{2.6}$$

where x_1 and x_2 vary on a collision-by-collision basis following probability density functions $f_i(x_1, Q^2)$ and $f_j(x_2, Q^2)$. These quantities, known as Parton Distribution Functions (PDFs), are measured by fitting dedicated experimental observables to collision data from various physics processes at different energy scales [54]. The total cross section of a generic process with final state X can be expressed as:

$$\sigma_{pp\to X} = \sum_{ij} \int dx_1 dx_2 \; f_i(x_1, Q^2) f_j(x_2, Q^2) \; \hat{\sigma}_{ij\to X} \tag{2.7}$$

where the contribution of the PDFs and the partonic cross section $\hat{\sigma}$ are factorised, and the sum is taken over all initial state partons that produce the final state of interest.

It should be noted that, since x_1 and x_2 can have different values, the physics objects X produced by the hard interaction carry an unknown boost in the direction of the beam axis. In addition, the final state contains extra objects produced by the partons that are not involved in the hard scattering, characterised by low transverse momentum and small production angles with respect to the beam. These underlying objects are a major source of background at hadron colliders.

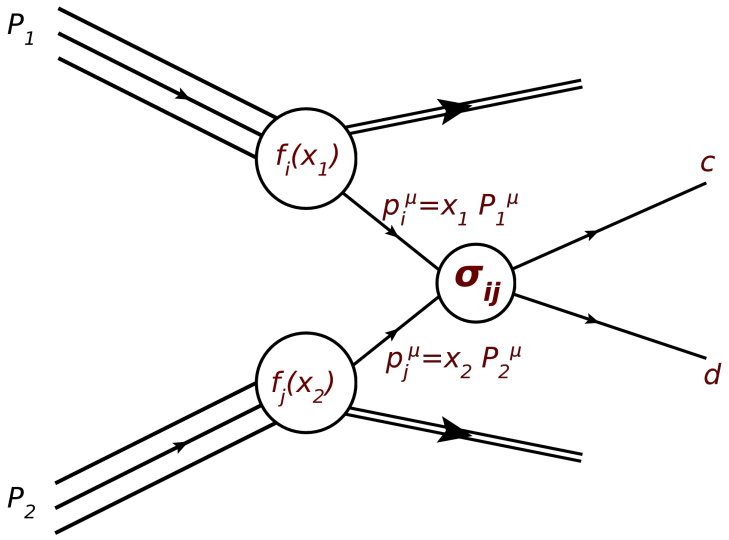


Figure 2.1: Hard scattering process involving two partons i and j whose momenta p_i and p_j are a fraction x_1 and x_2 of the total momenta of the protons. [55]

Underlying hadronic activity in the events can also be generated by proton-proton collisions where no hard scattering takes place. From a comparison between the total cross sections of inelastic collisions and of any other SM process (see Figure 5.2) it follows that the majority of the interactions yield only soft hadronic objects in the final state, while the hard scattering happens at lower rates. For every bunch crossing it is possible to compute the average number of interactions $\langle\mu\rangle$, commonly referred to as pileup, as a function of the instantaneous luminosity, the number of circulating bunches N, their frequency f and the total cross section of inelastic scattering of the protons σ_{inel}:

$$\langle\mu\rangle = \frac{\sigma_{\text{inel}}\mathcal{L}_{\text{inst}}}{Nf}. \tag{2.8}$$

As described in Section 2.1.4, the amount of pileup at the LHC is significant and its effects need to be kept under control when performing physics analyses.

2.1.3 The purpose of the LHC

The construction of the LHC at CERN was approved in December 1994 to replace the existing electron-positron accelerator, the Large Electron-Positron Collider (LEP) [56], after the completion of its physics programme. At the time, the largest machines in operation were the Tevatron [57], a proton-antiproton collider located at the Fermilab National Laboratory in the US, and the LEP itself, whose 27 km underground tunnel

is now used to host the LHC. The LEP experiments were designed to provide precise measurements of electroweak processes [58], thanks to the large number of Z and W bosons produced by the electron-positron collisions with low background rates. In parallel, the Tevatron was exploring the energy frontier, leading to the discovery of the top quark [59, 60] in proton-antiproton collisions at an energy of 1.8 TeV (subsequently increased to 1.96 TeV). Despite the reach of the Tevatron, however, the general consensus of the scientific community was that a more energetic machine would be needed to search for new physics up to the TeV scale [61], including the SM Higgs boson. Indeed, the Higgs was eventually not observed neither at the LEP [62] nor at the Tevatron [63,64], so its discovery became a primary goal of the LHC experiments.
In terms of operation, the new accelerator was expected to deliver collisions at an unprecedented centre of mass energy, larger by one order of magnitude with respect to the Tevatron and by two orders of magnitude with respect to the LEP. To achieve this goal, the only possible option was to build a hadron collider, because electron-positron machines are limited by the emission of synchrotron radiation (Equation 2.5). A relevant choice to make was whether to design a proton-antiproton collider, following the strategy of the Tevatron, or a proton-proton machine with two identical beams. The advantage of using antiproton beams is that in proton-antiproton collisions the cross section of processes initiated by quark-antiquark annihilation increases, due to the higher antiquark content of antiprotons. However, at the characteristic centre of mass energy of the LHC the dominant production modes of the most relevant physics processes are initiated by gluons, which are equally present in protons and antiprotons. Since antiproton beams are difficult to produce at high intensities, it was chosen to run the LHC with two proton beams.
Finally, the LHC was also designed to deliver collisions of lead ions, with a dedicated physics programme aimed at studying the thermodynamical properties of QCD and the quark-gluon plasma. This programme is beyond the scope of this thesis, so the operation with lead ions is not further described.

2.1.4 The CERN accelerator complex and the LHC

Figure 2.2 shows the full infrastructure of the CERN laboratory that prepares the proton beams for the injection in the LHC at an energy of 450 GeV [65]. Protons are produced through the ionisation of a hydrogen source, then they are accelerated to 50 MeV by a linear accelerator called LINAC 2. They are then injected in the first circular machine, the

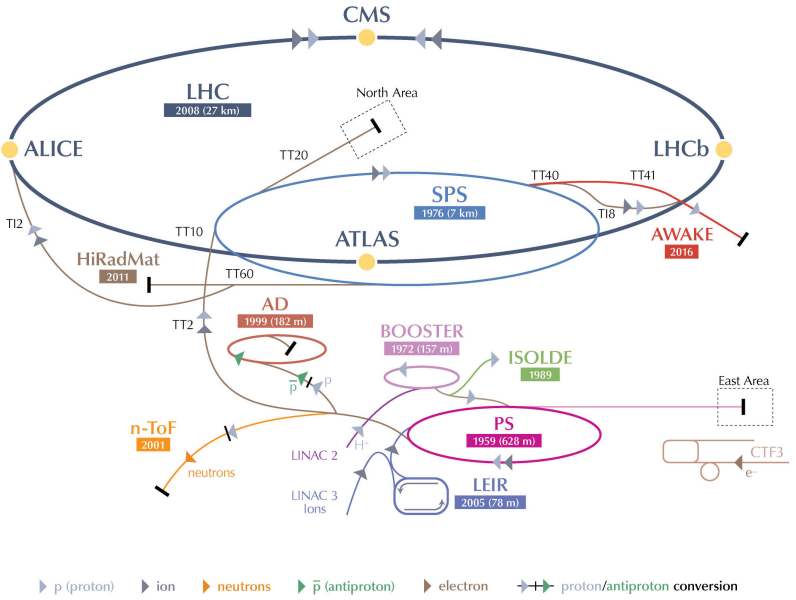


Figure 2.2: Scheme of the experimental facilities of CERN, showing the accelerator complex and the associated experiments.

Proton Synchrotron Booster (PSB), that increases their energy up to 1.4 GeV and prepares them for the following step, the Proton Synchrotron (PS). The PS is the oldest accelerator of CERN still in operation: it was built in 1959 and today it is used to accelerate the protons from 1.4 to 26 GeV. The next stage is the Super Proton Synchrotron (SPS), where the W bosons were discovered in 1983 by the UA1 and UA2 Collaborations [66, 67]. The SPS is now used to increase the energy of the protons up to 450 GeV, so that they can finally be injected in the LHC.

The LHC operation

The design report [68] published in 2004 describes the expected performance of the LHC before its construction, summarised in Table 2.1 in comparison with the 2015 and 2016 runs. The centre of mass energy is designed to be $\sqrt{s} = 14$ TeV with a maximum instantaneous luminosity $\mathcal{L}_{\text{inst}} = 10^{34}\text{cm}^{-2}\text{s}^{-1}$, to be obtained with a beam split into 2808 bunches of $\sim 10^{11}$ protons each. During the first years of operation some of the performance parameters have been reached or even exceeded, while others are still below the targets, including the centre of mass energy $\sqrt{s}$.
The first run of the LHC (Run 1) began when the first beams circulated in the machine on September 10^{th}, 2008. After 9 days of operation a serious incident was caused by

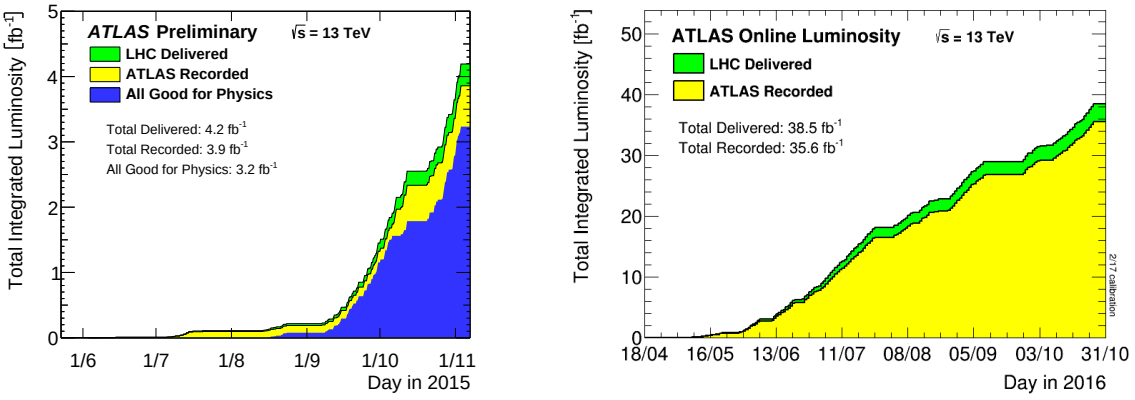
	Design	2015	2016
$\sqrt{s}$ [TeV]	14	13	13
Maximum n_b	2808	2244	2220
Average $\langle\mu\rangle$	-	13.7	24.9
Maximum $\langle\mu\rangle$	19	28.1	52.2
Peak $\mathcal{L}_{\text{inst}}$ [$10^{33}\text{cm}^{-2}\text{s}^{-1}$]	10	5.0	13.8
$\mathcal{L}_{\text{tot}}$ for physics [fb^{-1}]	-	3.2	32.8

Table 2.1: Performance parameters of the LHC during the 2015 and 2016 operation as measured by the ATLAS experiment [69], compared to the design values from Ref. [68].

the eccessive heating of a connection between two magnets, causing severe damage to a sector of the accelerator. After a necessary shutdown period of one year, the first proton collisions took place in fall 2009 at the minimum centre of mass energy $\sqrt{s} = 900$ GeV. In 2010 and 2011 the LHC delivered two extended periods of data taking at $\sqrt{s} = 7$ TeV reaching $\mathcal{L}_{\text{inst}} = 2 \times 10^{32}\text{cm}^{-2}\text{s}^{-1}$ and $\mathcal{L}_{\text{inst}} = 3.65 \times 10^{33}\text{cm}^{-2}\text{s}^{-1}$ respectively. In 2012 the energy of the collisions was raised to $\sqrt{s} = 8$ TeV, reaching a maximum $\mathcal{L}_{\text{inst}} = 8 \times 10^{33}\text{cm}^{-2}\text{s}^{-1}$ and accumulating a total integrated luminosity $\mathcal{L}_{\text{tot}} = 20\ \text{fb}^{-1}$. At the end of Run 1 the LHC had almost reached the design value of the instantaneous luminosity, while the centre of mass energy was still significantly below the target due to the limitations in the performance of the magnets that emerged after the incident of 2008.

After a long shutdown period of two years, the second run of the LHC (Run 2) began in spring 2015 at $\sqrt{s} = 13$ TeV, close to the design value of the centre of mass energy. Table 2.1 shows how the machine performed during the first two years of Run 2, during which the data analysed in this thesis were collected. The cumulative integrated luminosity versus time recorded by the ATLAS experiment is presented in Figures 2.3a and 2.3b for 2015 and 2016 respectively. Figure 2.4a shows the distribution of the average number of interactions per bunch crossing $\langle\mu\rangle$ (see Equation 2.8) during the 2015 and 2016 runs: the increase in instantaneous luminosity in 2016 caused a larger amount of pileup in the experiments. Finally, in Figure 2.4b the peak luminosity is plotted for every fill of the 2016 run, during which the performance of the LHC has significantly exceeded the design targets.

At the time of writing Run 2 is still ongoing. In 2017 the LHC delivered an additional



(a) (b)

Figure 2.3: Total integrated luminosity delivered by the LHC and recorded by the ATLAS detector in 2015 (a) and 2016 (b) [69]. For 2015 the figure includes the histogram obtained after further quality criteria are applied to the data to make them available for physics analyses.

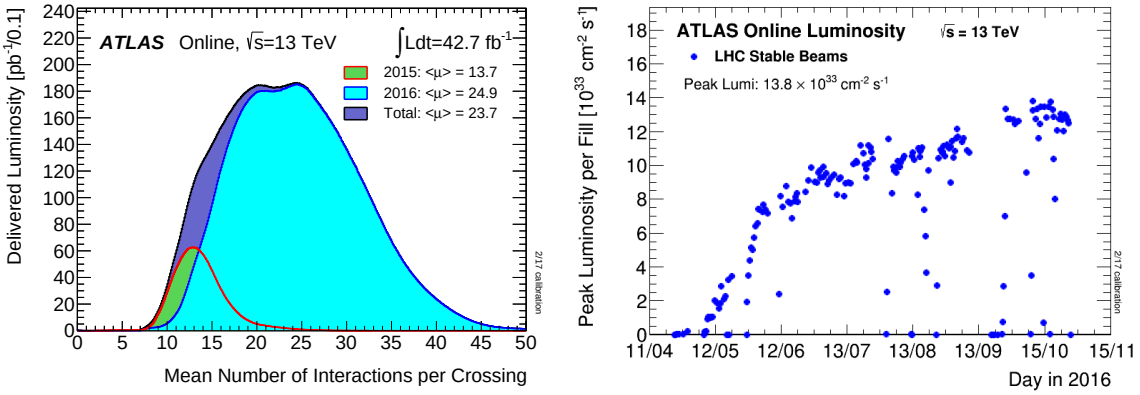
dataset comparable to the one of 2016, reaching new record values of instantaneous luminosity and $\langle\mu\rangle$ at $\sqrt{s} = 13$ TeV. Another year of data taking is scheduled for 2018, after which a second long shutdown will begin. Since the analyses presented in this thesis are only using the data from the 2015 and 2016 runs, the more recent datasets are not further discussed.

2.2 The LHC experiments

The LHC serves seven different experiments, all located along its ring. The main experiments, ATLAS [10], CMS [11], LHCb [70] and ALICE [71], are built in correspondence with the four interaction points where the collisions take place, as indicated by the yellow dots in Figure 2.2. In addition there are three smaller experiments, TOTEM [72], LHCf [73] and MoEDAL [74]. A short description of their main purpose is provided below:

ATLAS and CMS are the two largest detectors at the LHC. They are both general-purpose experiments, designed with different technologies to search for the Higgs boson and to explore the energy frontier at the TeV scale.

LHCb is a forward detector optimised to the study the decays of B and D mesons in



(a) (b)

Figure 2.4: Distribution of the average number of interactions per bunch crossing $\langle\mu\rangle$ (see Equation 2.8) during the 2015 and 2016 runs (a) and maximum instantaneous luminosity for every fill in the 2016 run (b) as measured by the ATLAS detector [69].

proton-proton collisions, providing precision tests of the SM parameters with special attention to CP violation.

ALICE is an asymmetric detector specifically designed to measure the products of lead ion collisions. Its purpose is to study the properties of QCD phase transitions and quark gluon plasma.

TOTEM is a smaller experiment mounted next to the CMS detector along the LHC tunnel. It is used to monitor the LHC luminosity by providing measurements of the total, elastic and diffractive cross-section of proton-proton collisions in the forward region, at small angles with respect to the beam.

LHCf is a second forward detector located in the LHC tunnel at both sides of the ATLAS cavern. Its purpose is to study cosmic ray shower processes by using particles scattered at small angles with respect to the beam axis.

MoEDAL is mounted next to the LHCb detector and is used to search for direct evidence of magnetic monopoles or stable and highly ionising massive particles.

2.2.1 General-purpose experiments

The purpose of the ATLAS and CMS detectors is to investigate the existence of new particles at heavy mass scales, with particular attention to the Higgs boson. Since the

experimental signature of BSM physics is unknown, both detectors are designed to identify as many different types of objects as possible, in order to maximise the chances of a discovery.
An essential feature of general-purpose detectors is hermeticity. Hermetic detectors are designed to observe all possible products of the collisions by covering the largest possible solid angle around the interaction region. They are split in multiple sub-systems that provide the necessary information to identify the different types of particles, with fine granularity in order to determine their position. If all particles with non-negligible interaction properties are measured with sufficient precision, hermetic detectors can reveal the presence of invisible objects such as neutrinos or dark matter candidates by reconstructing the resulting momentum imbalance in the transverse plane, known as missing transverse momentum (E_T^{miss}).
ATLAS and CMS must also satisfy general requirements common to all experiments at the LHC. The accelerator provides a maximum bunch crossing rate of about 40 MHz, setting the scale of the speed of response that the detector components must have to discriminate between subsequent events. A complex trigger and data acquisition system is needed to select the events where interesting physics objects are produced, saving the relevant information to permanent storage for offline analysis. Another important feature is resistance to radiation, especially for the detector components located at small distance from the interaction region.
The ATLAS and CMS detectors implement the above requirements using different technologies, in order to complement each other through independent measurements of the same physics phenomena. Their overall performance is similar and the agreement between their results is a fundamental test of their reliability.

2.3 The ATLAS detector

ATLAS is the biggest detector at the LHC. The full system weighs approximately 7.5 kTons and has a cylindrical symmetry around the beam axis, with a transverse diameter of 25 m and a longitudinal length of 44 m. The original design of the detector was proposed in 1994 [75], then its components were developed and assembled in parallel with the construction of the LHC.
A schematic representation of the detector is shown in Figure 2.5, including an overview of its main subsystems. The apparatus is designed to meet the requirements of a general-purpose experiment, with special attention to the products of hard scattering with large

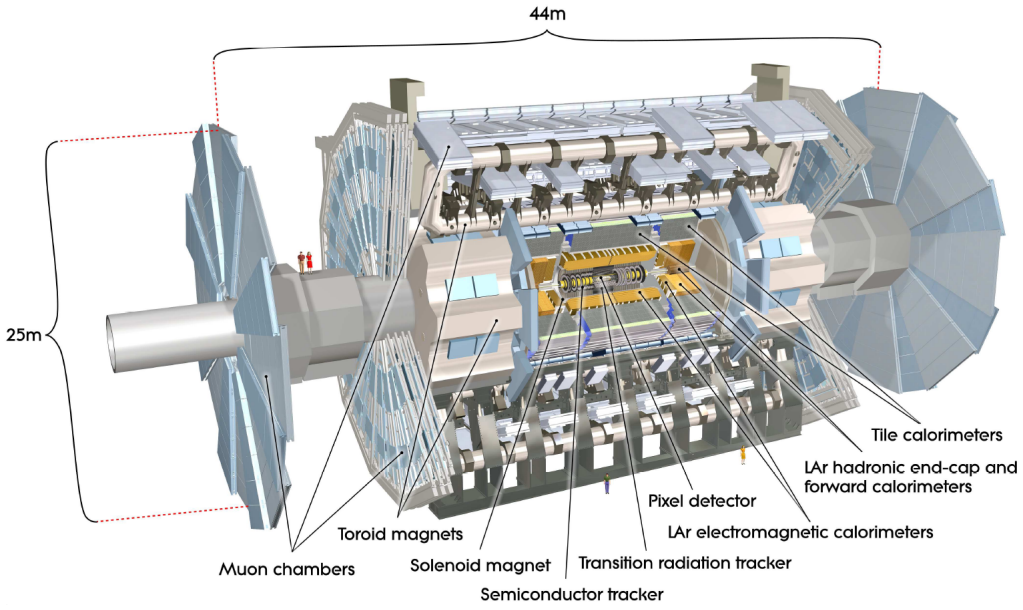


Figure 2.5: Overview of the ATLAS detector and its components [10].

transverse momentum (p_{T}) with respect to the beam axis. The detector consists of a set of concentric layers surrounding the interaction region, and can be divided into four major components. The Inner Detector (ID, Section 2.3.3) reconstructs the ionisation trajectories of charged particles and can be used to identify their production and decay vertices. Two calorimeters (Section 2.3.4) are then employed to measure the energy of electromagnetic and hadronic objects, while a Muon Spectrometer (MS, Section 2.3.5) is placed in the outermost layer to identify muons that travel through the detector. A magnet system (Section 2.3.2) made of two independent components is employed to bend the trajectories of charged particles in the ID and MS, allowing the measurement of their momenta. In order to maximise the hermeticity of the detector, each sub-system is made of components that are parallel to the beam axis (cylindrical barrels) and orthogonal to it (discoidal end-caps).
Figure 2.6 shows a scheme of the transverse section of the ATLAS detector, illustrating the characteristic signatures of the most relevant SM particles:

Photons are neutral objects, so they appear as showers in the electromagnetic calorimeter with no associated ionisation track in the ID.

Electrons are identical to photons in the calorimeter, but they also produce a track in the ID.

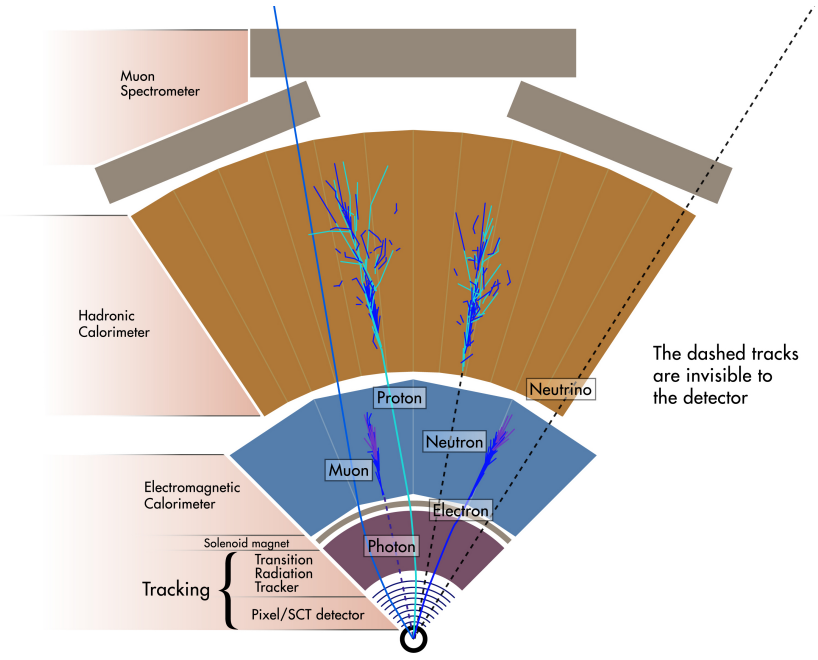


Figure 2.6: Simplified representation of the experimental signatures of the main types of particles in the ATLAS detector.

Neutrons leave no signature in the ID and generate hadronic showers that can start either in the hadronic calorimeter (as in Figure 2.6) or, with sizeable probability, also in the electromagnetic calorimeter (see Section 2.3.4).

Protons add an ID track to the characteristic signature of neutral hadrons.

Muons travel through each layer of the detector yielding minimum ionisation tracks in the ID and in the MS.

Neutrinos leave no signal in the detector, so their presence can only be inferred by reconstructing momentum imbalance in the transverse plane.

2.3.1 Coordinate system

In ATLAS the direction of the beam defines the z axis of a right-handed cartesian reference frame, with origin located at the centre of the interaction region. The transverse section of the detector is mapped by the x and y coordinates, with the x axis pointing towards the centre of the LHC ring and the y axis pointing in upward direction. Due to the symmetry of the system it is natural to introduce cylindrical coordinates, replacing x

and y with the radius r and the azimuthal angle ϕ. In addition, the polar angle θ defined with respect to the beam axis (with origin at the centre of the interaction region) can be used to describe the longitudinal position of objects in the detector.
An alternative to θ for particles with non-zero mass is the rapidity y:

$$y = \frac{1}{2} \log \frac{E + p_z}{E - p_z} \tag{2.9}$$

where E is the energy and p_z is the longitudinal projection of the momentum of the particle. The above expression is additive under Lorentz boosts in the z direction, so any difference Δy is a Lorentz invariant. This feature is particularly convenient at the LHC, where the particles are produced with an unknown longitudinal boost due to the variable fraction of momentum carried by the partons (see Section 2.1.2).
Finally, it is common to employ the pseudorapidity η, which is equivalent to the rapidity in the limit of massless objects and can be expressed simply as a function of θ:

$$\eta = -\log \left[\tan \left(\frac{\theta}{2} \right) \right] \tag{2.10}$$

with $|\eta| = 0$ corresponding to the centre of the detector ($\theta = \pi/2$) and $|\eta| = \pm\infty$ in the forward regions ($\theta = 0, \pi$).

2.3.2 Magnet system

The ATLAS magnet system [76], shown in Figure 2.7, provides the bending field that curves the trajectories of charged particles in the ID and in the MS. It is composed of two independent components: a solenoid, located around the ID, and an outer system consisting in a set of barrel and end-cap toroids, that provide the magnetic field to the MS. The layout of the magnet system distinguishes ATLAS from CMS, where a single solenoid magnet is employed [77], and has driven the design of the remaining parts of the detector. The solenoid and toroid magnets are briefly described below:

Solenoid The solenoid magnet surrounds the ID described in Section 2.3.3, and provides a magnetic field of 2 T using a nominal current of 7.73 kA that runs through 9 km of superconducting NbTi wires. The total length of the solenoid is 5.8 m, while the internal radius is 1.2 m and the thickness is only 4.5 cm, in order to minimise the amount of material placed in front of the calorimeters.

Toroids The barrel and end-cap toroids of ATLAS are both made of 8 identical magnets, that surround the calorimeter system with an azimuthal symmetry around the axis

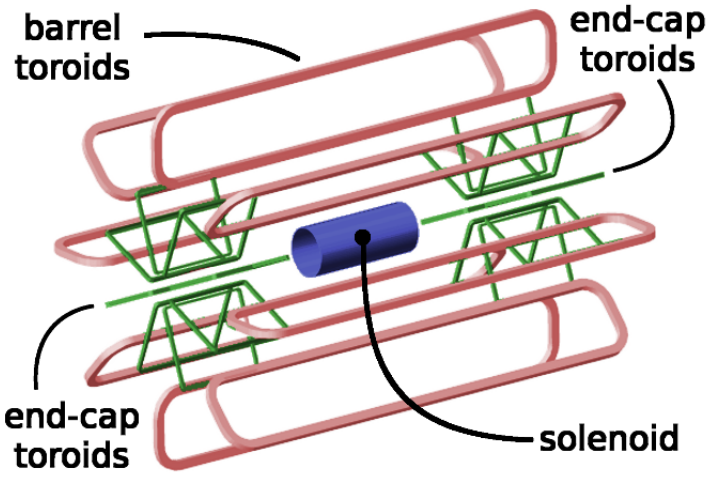


Figure 2.7: Schematic view of the ATLAS magnet system [76].

of the detector. The barrel system is 25.3 m long and has an outer diameter of 20.1 m, while the end-caps have a length of 5.3 m and 10.7 m of diameter. The peak field produced by both barrel and end-cap toroids is 4 T, obtained with a nominal current of 20.5 kA.

2.3.3 Inner Detector

The ID [78] is located at the core of the ATLAS experiment, immediately around the interaction region where the collisions take place. Its purpose is to reconstruct the trajectories of charged particles in the events, with a pseudorapidity coverage $|\eta| < 2.5$. As shown in Figure 2.8, the system is organised in cylindrical layers where the particles interact and release a localised signal. Due to the large multiplicity of particles produced by proton collisions, each layer needs to be finely segmented in order to provide an accurate measurements of the position of the hits, with enough precision to distinguish the individual tracks. Since the system is placed in proximity of the beam, it is also important to ensure that the hardware components are able to resist the high dose of radiation that they receive during the operation. Finally, the amount of material in the ID needs to be small in order to minimise the particle interactions before the calorimeter, that degrade the quality of the energy measurement.

The ID is split in three different subdetectors: a silicon Pixel detector that includes an Insertable B-Layer (IBL) installed during the shutdown period before the beginning of

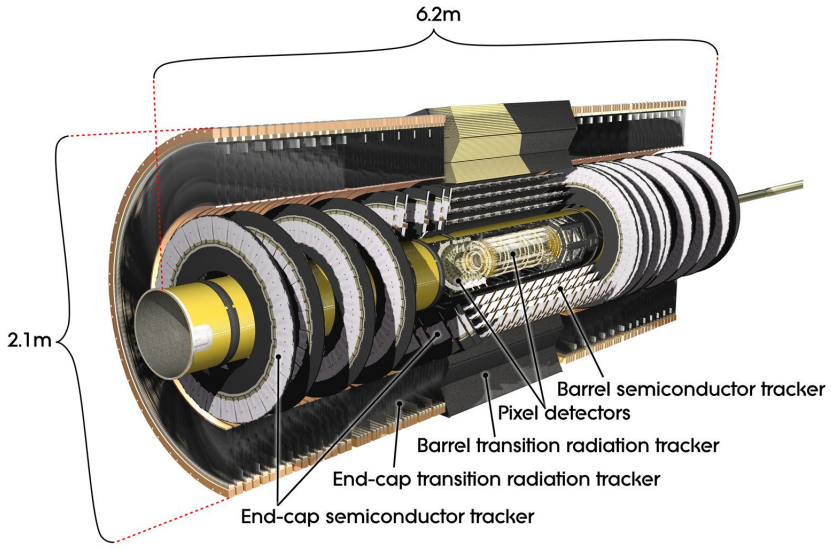


Figure 2.8: Transverse view of the barrel and end-cap components of the ATLAS ID system [10].

Run 2, a Semi-Conductor Tracker (SCT) and a Transition Radiation Tracker (TRT). Each of them has a cylindrical symmetry, with the usual barrel and end-cap structure. Overall, the ID measures 6.2 m in length and 2.1 m in diameter.

Pixel detector

The Pixel detector [79] is the innermost part of the ID, designed to provide tracking information with the finest level of granularity. As shown in Figure 2.9, the barrel part is made of four cylindrical layers of modules: the IBL [80, 81] (at 33.25 mm of distance from the beam axis), the b-layer or L0 (50.5 mm), the L1 (88.5 mm) and the L2 (122.5 mm). The end-caps consist of three circular disks, that contribute to achieve the desired longitudinal acceptance $|\eta| < 2.5$. The silicon pixels have a different segmentation in the R-ϕ and z directions: the smallest pixel size in the R-ϕ plane is 50 μm, while the minimum longitudinal pixel size is 250 μm in the IBL and 400 μm in the other layers. Due to the cost of the technology, it was chosen to use the pixels only in the innermost layers of the ID, where it is essential to maximise the spatial resolution.

Semi-Conductor Tracker

The Semi-Conductor Tracker SCT [82, 83] is the intermediate part of the ID, consisting of four barrel layers and two end-caps with nine disks each. The components of the SCT

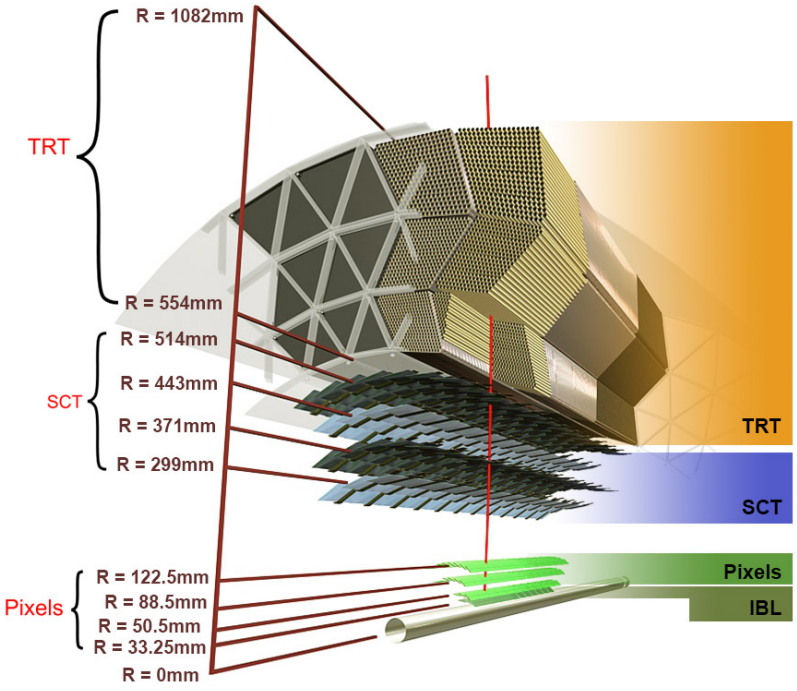


Figure 2.9: Transverse view of the barrels of the ATLAS Inner Detector, showing the radial distance of each layer from the beam axis [10].

are made of silicon, but the pixels are replaced by strips of 80 μm $\times 12$ cm to cover a larger area of space compared to the innermost layers. The strips are semi-parallel to the direction of the beam in the barrels, while they are disposed radially in the end-caps. The individual layers are made of a pair of strip sensors, arranged at a relative angle of 40 mrad. This geometrical solution allows to obtain a good measurement of the position of the hits in the longitudinal direction, despite the significant length of the strips. The SCT is able to provide a spatial resolution of 16 μm in the R-ϕ plane and of 580 μm in the z direction.

Transition Radiation Tracker

The TRT [84,85] occupies the external layer of the ID, with radius between 55.4 cm and 108.2 cm. Unlike the inner layers, this system is based on the use of straw detectors with 4 mm of diameter and maximum length of 144 cm, filled with a mixture of Xenon (70%), CO_2 (27%) and Oxygen (3%). Similarly to the SCT strips, the straw tubes are located in a barrel, where they are parallel to the beam direction, and in two end-caps, where they are arranged radially. Each straw is crossed by a thin tungsten wire that serves as anode, so

the device operates as a small drift chamber that measures the ionisation produced by the charged particles that cross it. A crucial feature of the TRT is that the region between the tubes is occupied by radiating material, where the particles emit x-rays proportionally to their relativistic γ factor. These photons are known as transition radiation, because they are generated when the particles cross the boundary between two mediums with different dielectric constants [86]. The transition radiation is absorbed by the Xenon atoms in the tubes and is significantly larger for electrons than for heavier particles, because the average γ factor of electrons is greater. As a result, the TRT is able to provide a useful measurement to discriminate between electrons and other types of particles. Thanks to its geometry, the TRT yields a large number of hits (~ 30) for each track in the R-ϕ direction, with a longitudinal acceptance $|\eta| < 2$ and a spatial resolution of about 130 μm.

2.3.4 Calorimeter system

The calorimeter system of ATLAS surrounds the ID and the solenoid magnet and provides a measurement of the energy of electrons, photons and hadrons, that generate showers of particles as a result of their interaction with the detector materials. The characteristic parameters describing the depth of electromagnetic and hadronic showers are the electromagnetic radiation length X_0 and the hadronic interaction length λ_I, with $\lambda_I \gg X_0$ for all common materials[2]. The calorimeter system is designed to provide a full containment of the showers, and each component is segmented in cells in order to obtain localised measurements of the energy deposits.

As shown in Figure 2.6, ATLAS employs two different types of calorimeters: the inner component is the Electromagnetic Calorimeter (ECAL) dedicated to electrons and photons, while the outer part is the Hadronic Calorimeter (HCAL) optimised for strongly interacting particles. The ECAL and HCAL are based on a sampling technology [87], in which the material that absorbs the incoming particles is distinct from the one that measures their energy. In addition, both calorimeters are non-compensating, which means that they have a different signal response to the electromagnetic and hadronic energy release, so a dedicated correction must be applied to calibrate the different types of showers at the same energy scale. Finally, in order to increase the longitudinal acceptance of

[2]The radiation length X_0 is defined as the mean distance over which an electron reduces its energy by a factor $1/e$, while the interaction length λ_I is the mean distance travelled by a hadron before being subjected to an inelastic nuclear interaction.

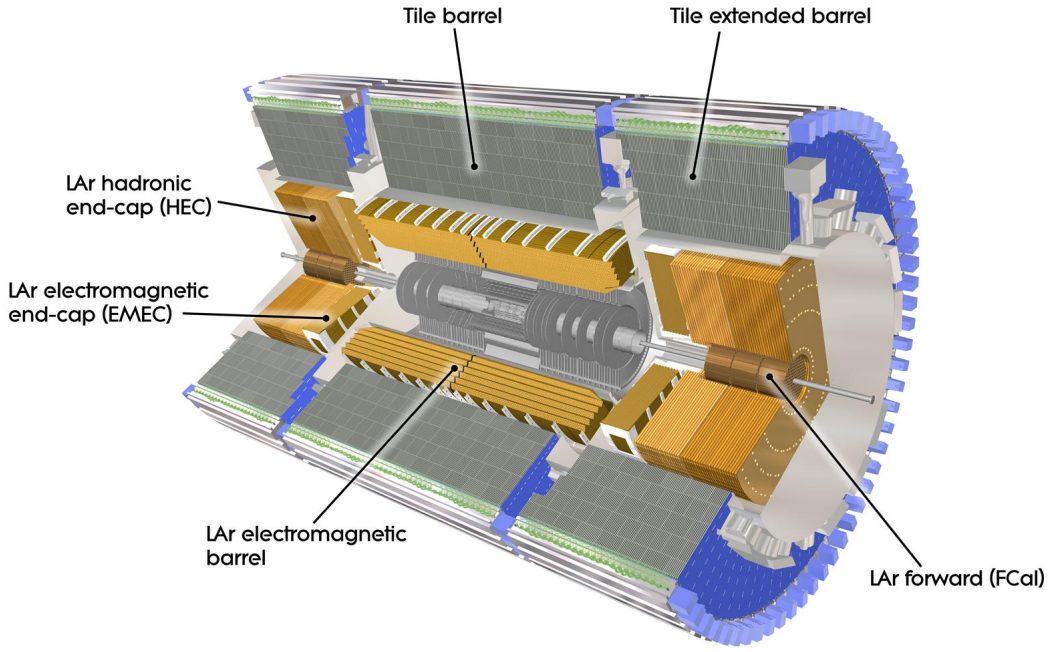


Figure 2.10: Section of the calorimeter system of ATLAS, highlighting the different components of the ECAL and the HCAL [10].

the detector, two Forward Calorimeters (FCal) are mounted on the opposite end-caps close to the beam pipe.

Electromagnetic calorimeter

The ECAL [88, 89] is a sampling calorimeter that employs lead (Pb) plates as absorbers and Liquid Argon (LAr) ionisation chambers with copper electrodes as active components. The system is divided into a central barrel, with acceptance $|\eta| < 1.475$, and two end-caps with range $1.375 < |\eta| < 3.2$. The lead plates are arranged radially with a characteristic accordion geometry, that provides a full azimuthal coverage and avoids the presence of radial cracks. The choice of LAr is motivated by its linear response combined with large signal yields and good resistance to radiation damage.
A limitation of the LAr chambers comes from the long duration (approximately 400 ns) of the analog signal pulse produced by the through-going particles, that has the characteristic triangular shape shown in Figure 2.11. A dedicated set of Front-End Boards (FEBs) is employed to process the triangular signals as described in detail in Ref. [10], obtaining a more complex signal shape also shown in Figure 2.11, optimised to suppress

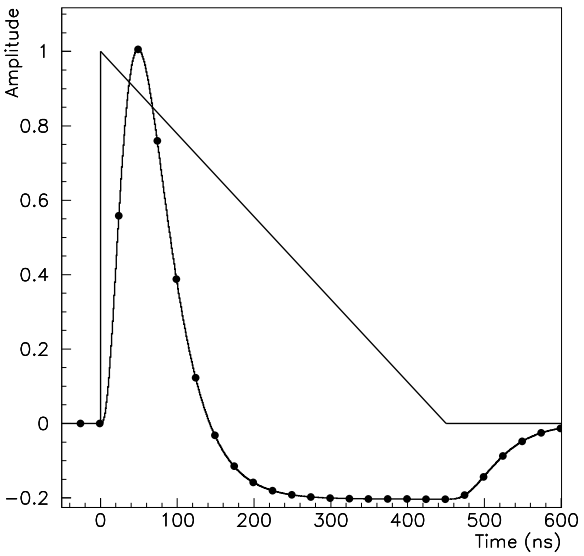


Figure 2.11: Signal amplitude as a function of time for a triangular ionisation pulse in a LAr cell of the ECAL barrel, and for a FEB output signal after the shaping [10]. The dots indicate the sampling points that are taken every 25 ns.

electronic and pileup noise. The shaped pulse is sampled every 25 ns in synchronisation with the LHC clock, obtaining a measurement of its amplitude and of the timing information (with resolution below 1 ns) that is necessary to assign the signal to the correct bunch crossing.

The ECAL is finely segmented in cells in the η-ϕ direction and is composed of three layers, each with specific properties and purposes. The first layer has a thickness of $4.3\ X_0$ and a fine granularity in the η-ϕ plane, which is needed to obtain a precise measurement of the initial development of the showers. Thanks to its segmentation, this layer plays a key role in the discrimination between showers produced by single photons and showers that originate from the decay of neutral pions (π_0) into two photons. The second layer measures $16\ X_0$ in thickness, and provides additional information which is used to determine the position and direction of the showers. The third layer has a depth of $2\ X_0$ and is used to measure the leakage of electromagnetic showers beyond the ECAL, contributing to the discrimination between electromagnetic and hadronic showers. The ECAL is completed by a presampler layer of LAr with no passive material, which is placed inside the solenoid magnet volume in order to measure the particle interactions in the inner parts of the detector.

In total, the ECAL has a thickness of $22\ X_0$ in the barrel for particles with $\eta = 0$, which grows as a function of pseudorapidity reaching $33\ X_0$ in the barrel and up to $38\ X_0$ in the end-caps. The longitudinal segmentation of the components of the ECAL produces a few crack regions where the nominal performance is deteriorated. This happens at $\eta = 0$ due to a 4 mm gap in the central barrel, at $1.37 < |\eta| < 1.52$ due to the transition between barrel and end-caps, and at $|\eta| = 2.5$ where there is a small gap between two end-cap discs.
As described in Ref. [89], the noise-subtracted energy resolution of the ECAL for electromagnetic showers can be expressed as:

$$\frac{\sigma_{\mathrm{E}}}{\mathrm{E}} = \frac{10\%}{\sqrt{\mathrm{E}\ [\mathrm{GeV}]}} \oplus 0.5\% \tag{2.11}$$

where the symbol $\oplus$ represents the sum in quadrature. The first term in Equation 2.11 is related to the stochastic fluctuations of the number of electromagnetic particles that contribute to the calorimeter signal, while the constant term (which increases up to 0.7% in the end-caps) originates from non-uniformities of the calorimeter response that are independent of the energy of the showers.

Hadronic calorimeter

The HCAL is dedicated to the measurement of hadronic showers, that are only partially contained by the ECAL. As shown in Figure 2.10, the full system consists of a barrel, two extended barrels and two Hadronic End-Caps (HEC).
The barrel and extended barrels form the Tile Calorimeter (TileCal) [90, 91], which uses steel tiles as absorber and plastic scintillators as active material. The barrel covers the central pseudorapidity region $|\eta| < 1.0$, while the extended barrels have a range $0.8 < |\eta| < 1.7$, with a partial overlap aimed at reducing the impact of the transition regions. Similarly to ECAL, both central and extended barrels of TileCal are divided into three layers, with an approximate thickness of 1.5, 4.1 and 1.8 interaction lengths in the central component and of 1.5, 2.6 and 3.3 interaction lengths in the extended ones. Finally, the HEC system is composed of two independent end-caps made of LAr and copper, that increase the pseudorapidity coverage of the detector to $1.5 < |\eta| < 3.2$.
Figure 2.12 shows a the cumulative amount of material as a function of the pseudorapidity η in the various layers of the detector, in units of interaction length λ_I. The total thickness is at least $10\ \lambda_I$ in all pseudorapidity regions, with peaks and dips caused by the transitions between different components.

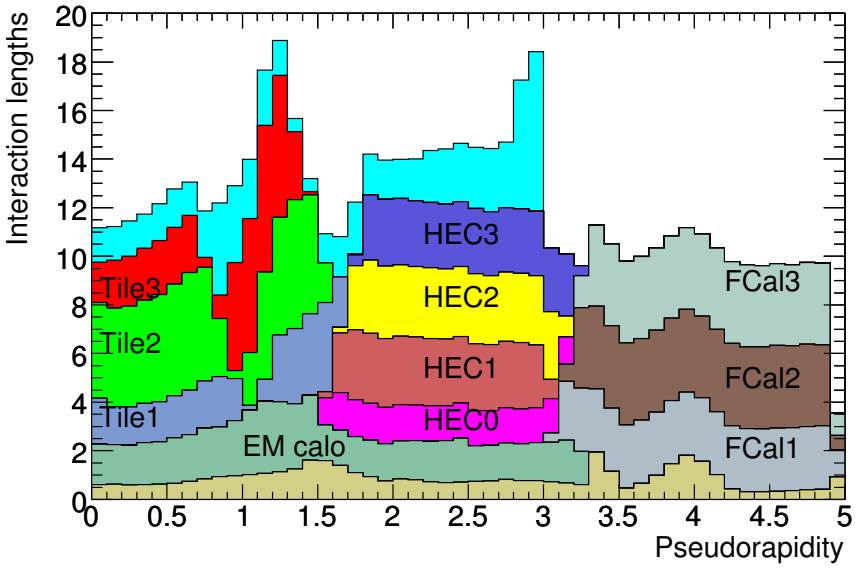


Figure 2.12: Depth of the major components of the detector in units of interaction length, as a function of the pseudorapidity $|\eta|$ [10]. The graph shows the material in front of the ECAL (in beige colour), the different layers of ECAL and HCAL including the forward components, and finally the total amount of material in front of the first active layer of the MS (in light blue) up to $|\eta| < 3.0$.

Finally, the jet energy resolution of the TileCal system in combination with the inner ECAL layers is measured to be:

$$\frac{\sigma_{\mathrm{E}}}{\mathrm{E}} = \frac{52\%}{\sqrt{\mathrm{E}\ [\mathrm{GeV}]}} \oplus 3\% \tag{2.12}$$

and a similar performance is obtained also by the HEC system, as further discussed in Ref. [10].

Forward calorimeters

The calorimeter system is completed by two forward detectors (FCal) [92] that provide electromagnetic and hadronic energy measurements in the pseudorapidity range $3.1 < |\eta| < 4.9$. The FCal uses LAr as active material and is divided in three layers of 45 cm each. The inner layer uses copper as absorber and is dedicated to electromagnetic showers, while the remaining two are made of tungsten and are optimised to measure the hadronic ones. The measurements performed by the FCal are not as accurate as those from the other components of the calorimeter system, due to a coarser segmentation and

to the large amount of background from underlying hadronic activity in the forward region. As discussed in Ref. [92], the electromagnetic and hadronic energy resolutions are measured to be approximately:

$$\frac{\sigma_{\mathrm{E}}}{\mathrm{E}}\,(\mathrm{EM}) = \frac{28.5\%}{\sqrt{\mathrm{E}\,[\mathrm{GeV}]}} \oplus 3.5\% \quad \text{and} \quad \frac{\sigma_{\mathrm{E}}}{\mathrm{E}}\,(\mathrm{had}) = \frac{94\%}{\sqrt{\mathrm{E}\,[\mathrm{GeV}]}} \oplus 7.5\%. \tag{2.13}$$

Despite the worse resolution, the FCal detectors give an crucial contribution to the hermeticity of ATLAS, which is important for the measurement of the missing transverse momentum (E_T^{miss}).

2.3.5 Muon Spectrometer

The outermost component of ATLAS is the MS [93] dedicated to the detection of muons, that travel through the inner layers of the detector as Minimum Ionising Particles (MIPs) and release little to no energy in the calorimeters. The muon trajectory in the MS is curved in the R-z plane by the bending field of the toroid magnets, so the reconstruction of the tracks can be used for the measurement of the momentum, similarly to what happens in the ID. The larger size of the MS allows accurate measurements of muons up to the TeV scale, which is essential for the physics programme of ATLAS.

The MS consists of a central barrel with three cylindrical layers with a radius of 5, 7.5 and 10 m, and four end-cap wheels at a longitudinal distance of 7.4, 10.8, 14 and 21.5 m from the centre of the detector. This structure hosts four different subsystems shown in Figure 2.13, each contributing to specific needs of the detector. Muon Drift Tube (MDT) chambers and Cathode Strip Chambers (CSCs) are dedicated to high-resolution measurements of the tracks in the pseudorapidity range $|\eta| < 2.7$, while Resistive Plate Chambers (RPCs) and Thin Gap Chambers (TGCs) are designed to provide trigger information with faster time of response.

A brief description of the four components of the MS is given below:

Monitored Drift Tubes The MDTs are chambers composed of drift tubes with 30 mm of diameter, filled with a mixture of Argon and CO_2. They cover the pseudorapidity range $|\eta| < 2.7$, except in the innermost end-cap layer where they only reach $|\eta| < 2.0$, providing a spatial resolution of approximately 80 μm per tube and 35 μm per chamber. They are disposed orthogonally with respect to the beam axis, and they only provide a measurement of the η coordinate of the hits. The major shortcoming of the MDTs is that their drift time can reach up to 700 ns, limiting the maximum rate of operation of the system.

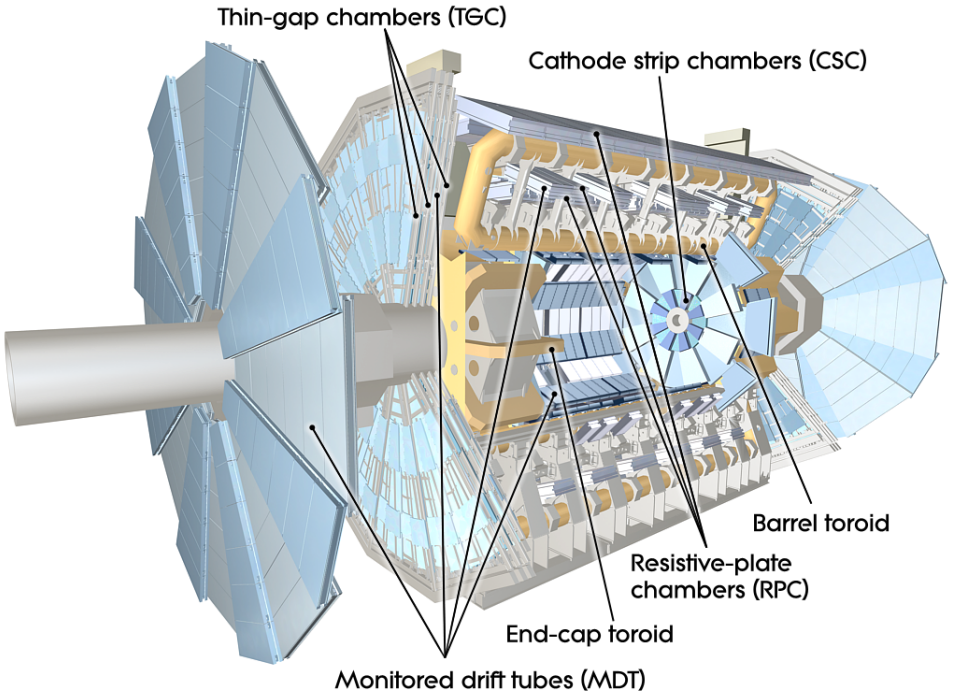


Figure 2.13: Scheme of the ATLAS outer layer showing the MS [10].

Cathode Strip Chambers The CSCs are multi-wire proportional chambers with a finer spatial resolution with respect to the MDTs. They are used in the first layer of the end-caps corresponding to $2.0 < |\eta| < 2.7$, where a higher rate of particles is expected.

Resistive Plate Chambers The RPCs are gas chambers mounted in the barrel system of the MS, in the pseudorapidity range $|\eta| < 1.05$. They are formed by two parallel plates with oppositely-charged anode and cathode, both made of a plastic material with high resistivity (Bakelite). The signal from ionising particles is collected by external metallic strips with a time resolution of the order of 1 ns, significantly shorter than the separation between two bunch crossings at the LHC (25 ns). RPC signals are used as inputs for the muon triggers thanks to the rapid speed of response, and they also provide a measurement of the ϕ coordinate of the muon hits, which is not measured by the MDTs.

Thin Gap Chambers The TGCs are multi-wire proportional chambers optimised to achieve a fast signal collection. They are mounted in the end-cap MS to improve the muon trigger capability in the $1.05 < |\eta| < 2.4$ region and to determine the ϕ

coordinate of muon trajectories in the forward direction up to $|\eta| = 2.7$.

2.3.6 Trigger and Data Acquisition

As discussed in Section 2.1, the LHC is designed to deliver bunch crossings up to a rate of 40 MHz, significantly higher than the readout capabilities of ATLAS. The physics processes of interest, however, are produced in a minor fraction of these events, while most bunch crossings yield only soft hadronic collisions that are not interesting for the physics analyses. The Trigger system of ATLAS [94] is designed to perform an online selection of events where interesting physics objects are identified, making a fast real-time decision based on the signals in the relevant sub-detectors. The accepted events are recorded by a Data Acquisition system, that channels the data from the sub-detectors to permanent storage.
A scheme of the Trigger and Data Acquisition (TDAQ) system is shown in Figure 2.14 where the interplay between the different sub-components is highlighted. A full trigger selection involves a chain of decisions taken first by a hardware-based Level 1 (L1) trigger and subsequently by a software-based High Level Trigger (HLT). If the output rate is too high, the chain of selections can be individually pre-scaled by a factor P, which means that only a fraction $1/P$ of the events that meet the selection criteria are accepted. Since this procedure causes a loss of interesting events, pre-scaled trigger chains are generally not used by the analyses in the present thesis.

Level-1 Trigger

The L1 trigger [96] performs a first selection of events that reduces the rate from the bunch crossing frequency (up to 40 MHz) to a maximum output of 100 kHz, using low granularity information from the calorimeters (L1Calo) and the dedicated sub-systems in the MS (L1Muon). The selections on the signals from the sub-detectors are optimised to identify events where interesting physics objects are produced, such as electrons, muons, τ leptons, jets or missing transverse momentum (E_T^{miss}). For Run 2, the system has been upgraded by adding a topological trigger processor (L1Topo) [97] that combines kinematical information from different L1Calo and L1Muon objects, allowing to define more complex selections based on the topology of the event.
The detector data is kept in memory buffers for a latency period of 2.5 μs, during which the Central Trigger Processor (CTP) has time to perform a decision about whether the event meets the selection criteria. If the event is accepted, the corresponding data is

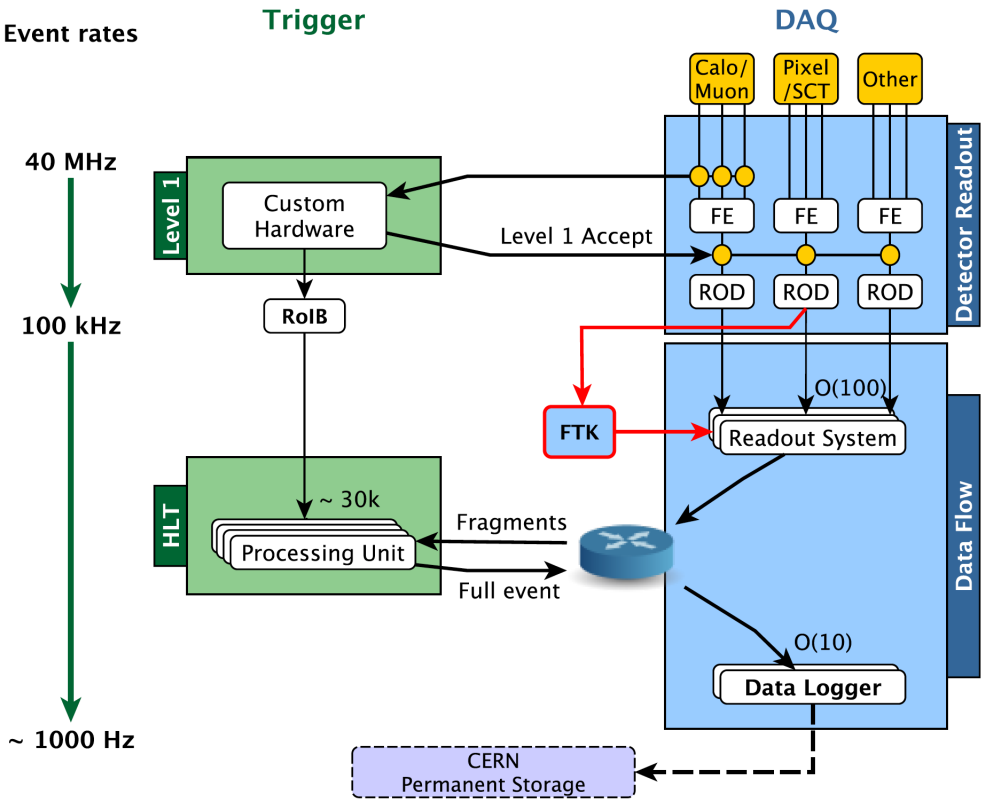


Figure 2.14: Scheme of the TDAQ system of the ATLAS detector [95].

sent to detector-specific Readout Drivers (RODs) and subsequently transferred to the Read-Out System (ROS). In addition, the Region of Interests (RoIs) with the η and ϕ coordinates of the interesting objects in the event are determined by a Region of Interest Builder (RoIB) and transferred to the HLT to act as seeds for further selections.

High Level Trigger

The HLT [98] performs additional selections that reduce the event rate from the output of L1 (100 kHz) to values of the order of 1 kHz. The system exploits software algorithms that use the full detector information, initially restricted to the RoIs defined by the L1 trigger in order to reduce the computational time. The events selected by the HLT algorithms are sent to mass storage by Sub-Farm Output (SFO) nodes. In Run 1 the HLT consisted of two separate levels of decisions, Level 2 (L2) and Event Filter (EF), running in separate server farms. Before the beginning of Run 2 the two levels were merged [99] into a single homogeneous farm, removing redundancies between them and simplifying the hardware and software structure. In addition, a new hardware-based Fast Tracker (FTK) [100] is

being incorporated, in order to provide tracking information with full detector coverage to the HLT.

Data Quality

The data collected by the ATLAS detector must satisfy standard quality criteria in order to be used for offline analysis. To provide an effective monitoring of the performance of the detector during the runs, the ATLAS data are recorded in luminosity blocks, short periods of time (with typical length of approximately one minute) during which the configuration of each sub-system is kept constant. A set of quality parameters can be monitored during the operation for each sub-detector, and a record of the performance is kept in order to compile a Good Run List (GRL) for each period of data taking. The GRL contains the list of luminosity blocks in each run where the detector has performed in line with the quality requirements.

3 | Simulation and reconstruction of ATLAS events

This chapter presents a comprehensive discussion of the experimental signature of the major physics processes in the ATLAS experiment. Section 3.1 is dedicated to a discussion of the event simulation techniques, from the proton-proton collisions to the interaction of the particles with the detector. The remaining part of the chapter (Section 3.2) describes the standard procedures for the reconstruction of different types of physics objects in ATLAS, based on the signal that they release in each detector sub-system.

3.1 Monte Carlo simulations

A crucial tool for the analysis of the ATLAS data is the simulation of physics processes, heavily used to predict their kinematical properties and their experimental signature in the detector. The standard approach is to employ Monte Carlo (MC) generators, where the events are simulated stochastically following probability distributions determined by the theory. The simulations include the particles interacting with the detector and the associated response of the detector sub-systems, so the results can be directly compared to ATLAS data.

3.1.1 Simulation of physics processes

The production of physics objects in proton-proton collisions can be viewed as a multi-stage process, as shown in Figure 3.1. The first step is the hard scattering of individual partons from the protons, already discussed in Section 2.1.2, where the PDFs play a key role. The hard scattering can be further split into a Matrix Element (ME) computation, which typically uses perturbative calculations truncated at a fixed order, and a Parton Shower (PS) stage which simulates the emission of extra soft objects in the event. Due to

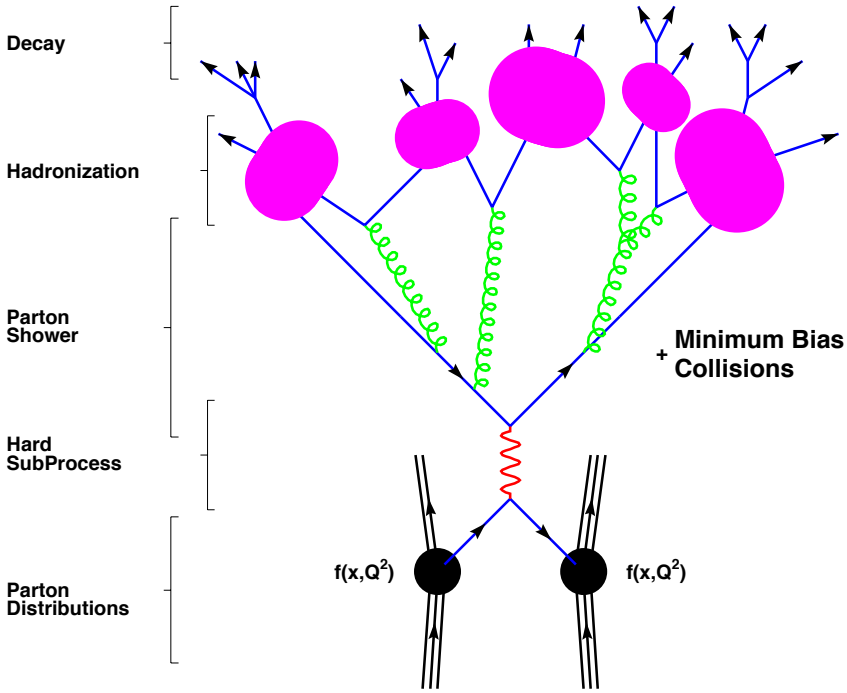


Figure 3.1: Scheme of a physics process at the LHC, showing the various stages from the two protons in the initial state to the decays of the produced particles [101].

the confinement properties of QCD, any quarks or gluons produced by the hard scattering are eventually converted into hadrons through a hadronisation process (see Section 1.1.2). If the particles are unstable, their decay can be simulated in different stages of the chain. A description of the individual steps of the MC simulation procedure is provided below:

Factorisation and PDFs As discussed in Section 2.1.2 the hard scattering in proton-proton collisions involves two partons (quarks or gluons) that carry an unknown fraction x_1 and x_2 of the the momentum of the protons. The cross section σ of the full physics process is given by Equation 2.7, where the probability of finding a parton with momentum fraction x is factorised with respect to the partonic cross section $\hat{\sigma}$. To take into account the effects of higher order corrections, it is necessary to introduce a factorisation scale μ_F and a renormalisation scale μ_R (see Ref. [53] for a full discussion) and Equation 2.7 becomes:

$$\sigma_{pp\to X} = \sum_{ij} \int dx_1 dx_2 \; f_i(x_1, \mu_F^2) f_j(x_2, \mu_F^2) \; \left[\hat{\sigma}_0 + \alpha_S(\mu_R^2)\hat{\sigma}_1 + ...\right]_{ij\to X} \quad (3.1)$$

where the perturbative corrections $\hat{\sigma}_n$ of order α_S^n are introduced. Formally the

total cross section is independent of μ_F and μ_R, because the scale dependance at infinite order is exactly compensated by the coefficients $\hat{\sigma}_n$. At fixed order, however, the numerical results are affected by the choice of the two scales, which is a source of theoretical uncertainty. The PDFs are determined by experimental measurements at fixed energy scales, and their evolution to different scales is determined through the DGLAP equations [102–104]. For the analyses discussed in this thesis, PDFs derived by the CTEQ [105, 106], NNPDF [107, 108] and MSTW [109] groups are used.

Matrix element (ME) An essential step of the MC simulations is the computation of the partonic cross section $\hat{\sigma}_n$. The processes of interest can be schematically represented as

$$ij \to X + \sum_i x_i$$

with an initial state ij and a final state that includes a main product X and a variable number of extra objects x_i. The associated scattering matrices, or S-matrices, can be computed at different orders in perturbation theory.

As an example, figure 3.2 shows three Feynman diagrams for the case of top pair production initiated by two quarks: the first (Figure 3.2a) is a Leading Order (LO) diagram, while the others (Figures 3.2b and 3.2c) are Next-to-Leading Order (NLO) diagrams that include a gluon emission and a loop correction respectively. While the diagrams in Figures 3.2b and 3.2c are both divergent, an important theorem by Kinoshita-Lee-Nauenberg [110, 111] states that at fixed orders in perturbation theory the infinities are exactly cancelled, yielding a finite result for the cross section. If a larger multiplicity of objects in the final state is included, however, the cancellations are spoiled and divergent contributions from soft collinear emissions can appear. The standard strategy to remove the divergencies is to introduce a cutoff in the computation of the ME, such that extra soft objects are excluded and a finite result can be obtained. These missing objects are subsequently added by dedicated PS generators.

Parton Shower (PS) Once the computation of the ME is completed, a PS generator is used to include the emission of extra soft objects. In QCD, the possible processes are the emission of a gluon by a quark ($q \to qg$) and the emission of a quark pair or a gluon pair by a gluon ($g \to qq$ and $g \to gg$). The MC generators make use of Sudakov form factors [113], that describe the probability for a virtual gluon at an

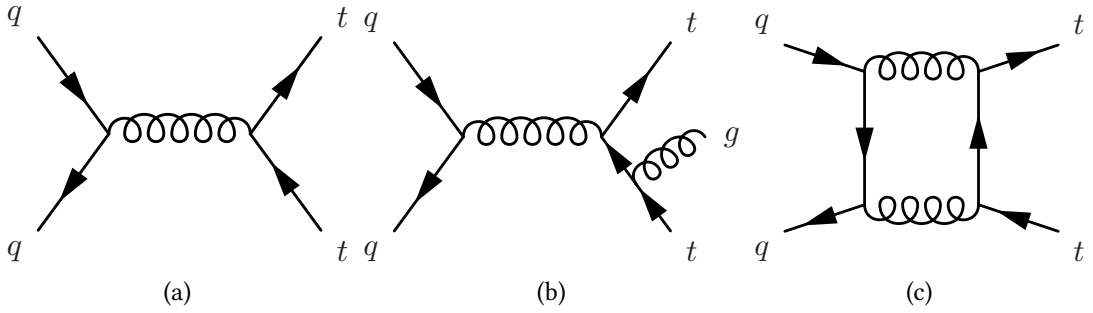


Figure 3.2: Feynman diagrams for top pair-production initiated by two quarks, where (a) is a LO diagram, (b) is a NLO diagram with a gluon emission and (c) is a NLO diagram with a loop correction [112].

energy scale Q_a to evolve to a scale Q_b without radiating.
The PS objects can appear as part of the Final State Radiation (FSR), as in Figure 3.2b, or of the Initial State Radiation (ISR). The latter case is implemented in the MC using a backward evolution technique, where the ME is computed using partons with fixed momentum and the initial state is determined by evolving them in reversed time order.

ME-PS matching In order to combine the ME and PS generators, it is necessary to define a procedure that determines the respective domains. This is done by dedicated matching algorithms that identify two orthogonal regions of phase space, one containing the products of the hard scattering and one with the additional soft objects in the event. The goal of the matching is to remove any potential overlap between the physics objects generated by the ME and by the PS, without causing significant discontinuities in observable spectra. The most common implementations are the Catani-Krauss-Kuhn-Webber (CKKW) [114] and the Michelangelo L. Mangano (MLM) [115] schemes, both widely used by the ATLAS Collaboration.

Hadronisation Quarks or gluons in the final state evolve until they reach a Q^2 value of the order of $\Lambda_{\mathrm{QCD}} \sim 200$ MeV, at which the confinement properties of QCD become relevant. As explained in Section 1.1.2, at this stage all the elementary hadrons in the event must recombine into colourless objects, mesons or baryons. The hadronisation process is regulated by non-perturbative QCD, so its simulation needs to rely on phenomenological models: the most common ones are the cluster model [116, 117] and the Lund string model [118, 119].

Underlying event and pileup While the hard scattering processes at the LHC involve only two quarks or gluons, the remaining partons from the initial state protons can yield an extra number of soft interactions at lower energy scale. As a result, the LHC processes can generally be regarded as Multiple Parton Interactions (MPIs) that produce a significant amount of underlying hadronic activity in the events (see Section 2.1.2). The underlying event is simulated using phenomenological models tuned on experimental data [120, 121].

Additional hadronic activity can originate from interactions between different proton pairs in the same bunch crossing (in-time pileup) or by remnants of collisions in different bunch crossings (out-of-time pileup). Both processes are simulated similarly to the underlying event, and are included in the MC samples used by the ATLAS Collaboration.

3.1.2 Monte Carlo generators

The scientific community has developed a variety of MC generators to implement the various steps of the simulation of physics processes at the LHC. A brief description of those that are used in this thesis is given below:

Pythia is a multi-purpose LO generator that can handle the full reaction chain for $2 \to n$ ($n \leq 3$) processes [122, 123], including hadronisation (with the Lund string model) and underlying event.

Herwig is a similar multi-purpose LO generator that simulates $2 \to 2$ processes [124] using the cluster model for hadronisation and underlying event.

MadGraph is a MC generator that handles the ME part of the simulation chain [125]. It can be used at LO or in its NLO version aMC@NLO, and it must be interfaced with another generator (typically Pythia or Herwig) for PS, hadronisation and underlying event.

Powheg is a NLO ME generator [126] that can be interfaced with Pythia or Herwig for PS , hadronisation and underlying event.

Sherpa is a multi-purpose generator that performs LO or NLO calculations for $2 \to n$ processes [127]. It includes ME, PS [128] and underlying event generator and it uses the CKKW method for the ME-PS matching.

3.1.3 Detector simulation

The output of the MC generators is a list of four-vectors of particles in the final state, that can be used to examine the kinematical features of the physics processes. In order to compare the MC samples with the ATLAS data, however, it is also necessary to simulate the interaction of the particles with the detector and the resulting signals from the sub-systems [129].
The detector simulation is implemented by the Geant 4 package [130], which is integrated in the ATLAS offline software. The full reconstruction chain includes a first stage dedicated to the interaction of the particles with the detector, and a second (digitisation) stage where the energy deposits are converted into voltages and currents in the sub-systems. The output of the simulation can be presented in a format identical to the output of the ATLAS TDAQ system, so the MC samples and the real ATLAS data can be processed by the same reconstruction software.
Finally, it is important to mention that the ATLAS Collaboration makes wide use of faster simulations such as ATLFAST-II (or AF2) [131], where a parametrised description of the detector response is implemented in order to reduce the CPU processing time.

3.2 Object reconstruction

An essential step of the analysis of ATLAS data is the reconstruction of physics objects in the events, based on their experimental signature in the detector. As discussed in Section 2.3.6, a first selection of events is performed by the TDAQ system, where the relevant features are identified by fast online algorithms. Once the events are recorded, the final observables to be used for the analyses are extracted by more refined reconstruction algorithms, less affected by timing limitations.
The searches presented in this thesis employ different types of trigger selections described in Section 3.2.1, requiring the presence of missing transverse momentum (E_T^{miss}) or single electrons, photons or muons in the events. At offline level, the relevant objects to be reconstructed include tracks and primary vertices (Section 3.2.2), electrons and photons (Section 3.2.3), muons (Section 3.2.4), jets from the hadronisation of quarks and gluons (Section 3.2.5) with highlight on the case of jets from b-quarks (Section 3.2.6) and, finally, missing transverse momentum (Section 3.2.7).

3.2.1 Trigger objects

The ATLAS TDAQ system performs the online selection of events based on the definition of *trigger chains*, sequences of algorithms optimised to identify a given type of physics object [94]. After the primary hardware-based selections applied at L1, the HLT chains are typically configured as a sequence of *Feature Extraction* algorithms, where the relevant physical quantities are reconstructed based on the signals in the subdetectors, and *Hypothesis* algorithms, where the trigger selections are implemented. The following naming convention is used for the chains:

$$[\text{LEVEL}]\ [\text{TYPE}]\ [\text{THRESHOLD}]\ [\text{IDENTIFICATION}]\ [\text{ISOLATION}] \tag{3.2}$$

where the trigger level (L1 or HLT), chain type (electron, muon, E_T^{miss} or others), energy threshold, identification and isolation criteria (if present) are listed.
The collection of chains used during the online operation of the detector is known as the ATLAS trigger menu, optimised to meet the requirements of a broad range of physics analyses and operational purposes. The list below provides a brief description of the triggers that are relevant for the searches presented in this thesis, namely electron, photon, muon and E_T^{miss} chains:

Electrons and photons The online trigger chains targeting electron and photon candidates are seeded by L1 algorithms that identify RoIs with significant energy deposits in the calorimeters (L1Calo). As shown in Figure 3.3, trigger towers measuring 0.1×0.1 in η and ϕ are formed in both ECAL and HCAL, with a pseudorapidity range $|\eta| < 2.5$. A sliding window procedure is used to select blocks of 4×4 towers where the total energy in at least one of the central pairs (horizontal or vertical) exceeds a fixed threshold. Isolation selections are applied by using the energy deposits in the external ring of the 4×4 blocks, and a veto on significant energy deposits in the HCAL is added to suppress hadronic showers.
The RoIs defined at L1 are passed to the HLT, that applies an additional sequence of selections using algorithms that closely resemble the offline ones (see Section 3.2.3). The requirement of a track matched to the calorimeter cluster is introduced for all electron chains, to distinguish them from photon candidates. During my PhD I have contributed to a variety of tasks related to the electron and photon triggers, including a study of the online rate of a single electron trigger chain as a function of the transverse energy threshold, which is discussed in Appendix B.

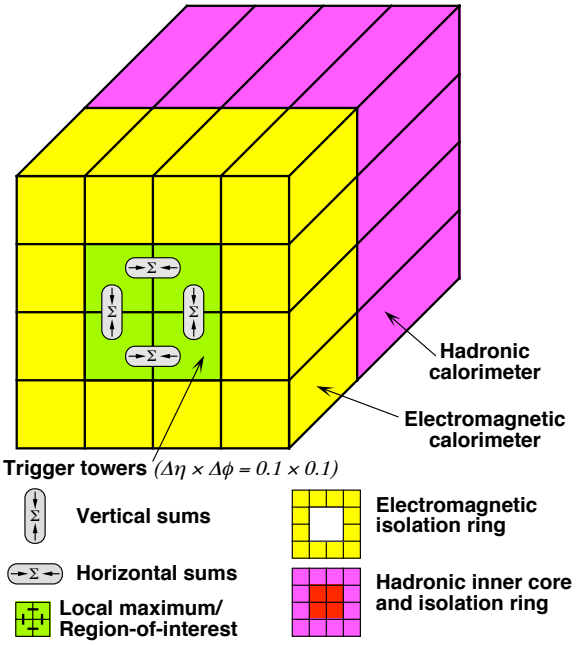


Figure 3.3: Scheme of a 4×4 block of trigger towers employed by the L1Calo trigger algorithms. [94].

Muons The L1 trigger identifies muon candidates by requiring the coincidence of hits on different layers of the RPC or TGC chambers, and an estimate of the transverse momentum of the particle is obtained by measuring the deviation of the trajectory with respect to a straight line. The HLT receives the RoI information and uses the precision chambers, MDTs and CSCs, to apply further selections on the candidate, with a two-stage approach similar to the one of the electron and photon chains. The tracks from the MS are matched and combined with those from the ID, and additional isolation variables can be used to reduce the number of fake candidates.

Missing transverse momentum The computation of the missing transverse momentum (E_T^{miss}) performed by the L1 trigger employs the same energy deposits in 0.1×0.1 towers used for the electron and photon chains. To maximise the coverage, all layers of the ECAL and HCAL for $|\eta| < 4.9$ are employed. The HLT algorithms use as input only the signals provided by the calorimeters, with no RoI information due to the global nature of the E_T^{miss} variable. Different methods are employed for the reconstruction of the E_T^{miss}, based on cells, jets or energy clusters, with dedicated techniques to correct for pileup effects.

Trigger efficiency and scale factors

The most relevant parameter to describe the performance of the trigger chains is the selection efficiency ϵ, defined as the fraction of events where the presence of a real object[1] is successfully identified by the online algorithms.
The efficiency of the chains is typically studied as a function of the transverse momentum of the associated offline objects in the event, as shown in Figure 3.4 for the cases of electrons, muons and E_T^{miss}. The efficiency of E_T^{miss} chains is measured in events satisfying lepton trigger requirements, while for electrons and muons the tag-and-probe method in $Z \to ee$ and $Z \to \mu\mu$ events is used [94]. Since the selections are applied based on online energy measurements, it is natural to obtain turn-on curves with a step-like behaviour close to the online threshold of the chain, and a smoothened profile due to finite resolution effects. Both electron (Figure 3.4a) and E_T^{miss} (Figure 3.4b) triggers have high efficiencies above the turn-on, although especially in the case of E_T^{miss} triggers a tight offline selection is necessary. On the contrary, the L1 muon trigger efficiency is limited in both barrel (Figure 3.4c) and end-cap (Figure 3.4d) regions, mainly due to geometrical acceptance effects.
Finally, by comparing the trigger efficiency in data and simulation it is possible to obtain calibration Scale Factors (SFs) for both electron and muon chains, essential to improve the data-MC agreement in the analysis selections.

3.2.2 Tracks and primary vertices

Electrically charged particles release localised ionisation signals (hits) in the ID, which is able to measure the positions of the hits with variable precision depending on the granularity of the corresponding layer (see Section 2.3.3). By combining the individual hits on different layers, it is hence possible to identify tracks that correspond to the trajectories of the particles [133]. The baseline track reconstruction algorithm in ATLAS follows an inside-out approach where the initial track candidates are formed by at least 3 hits in the silicon detectors, and are subsequently extrapolated to the outer silicon layers and the TRT. The track finding procedure is then completed by an outside-in algorithm, which uses TRT hits as seeds and extends them towards the inner layers.
Due to the magnetic field produced by the solenoid, the tracks follow a characteristic helicoidal trajectory whose curvature is inversely proportional to the momentum of the

[1] For the purpose of the definition of the trigger efficiency, the concept of physics object is extended to event-level quantities such as the E_T^{miss}.

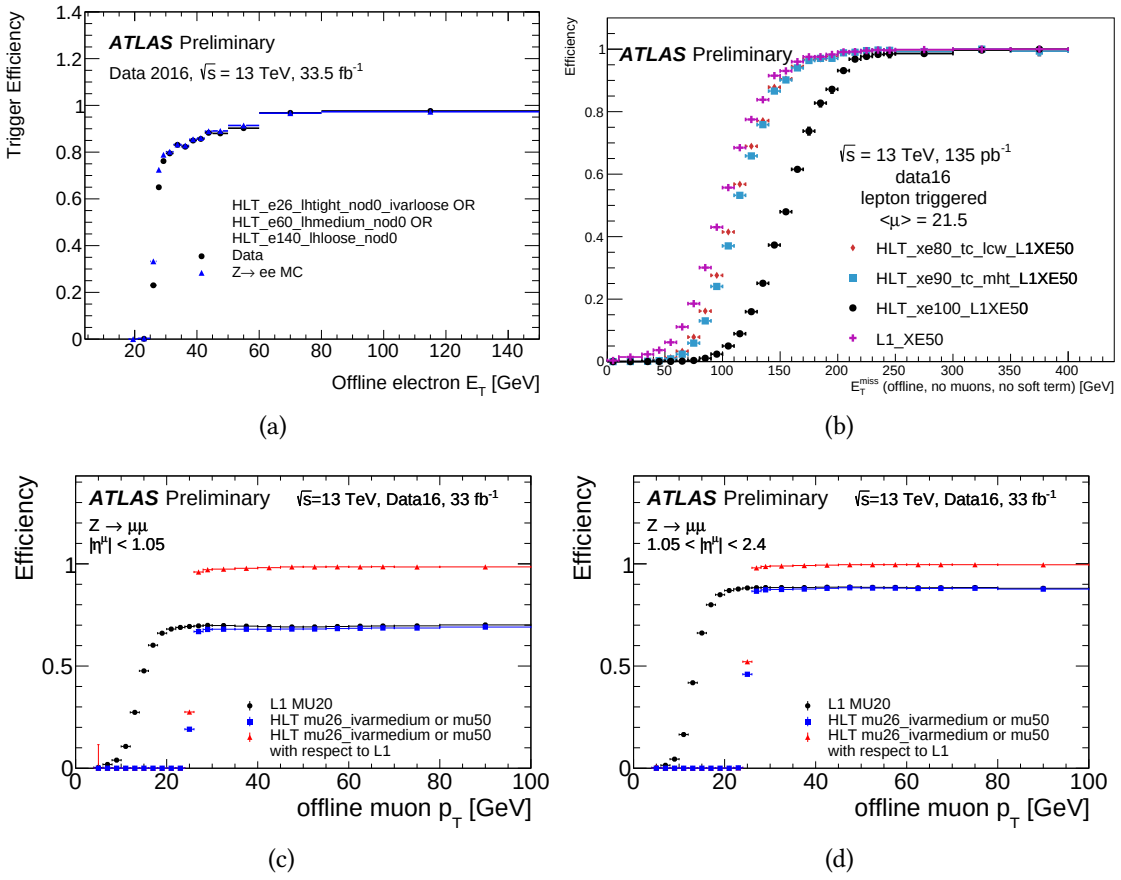


Figure 3.4: Trigger efficiency as a function of offline variables for electrons (a), E_T^{miss} (b), and muons in the barrel (c) and end-caps (d). Various types of online E_T^{miss} algorithms are compared in Figure (b), while for electrons and muons a logic OR of chains with different E_{T} thresholds and isolation requirements is used [132].

particle. A track is fully described by five perigee parameters measured at the point of closest approach to the z-axis:

$$(\theta, \phi, q/p, d_0, z_0)$$

where θ and ϕ are the longitudinal and azimuthal angles, q/p is the ratio between the charge and the momentum of the particle, and the quantities d_0 and z_0 are known as transverse and longitudinal impact parameters, representing the minimum distance between the track and the centre of the detector in the transverse plane and in the longitudinal direction respectively.

The tracks reconstructed by the above algorithms can be used to identify the presence of Primary Vertices (PVs) formed by collisions or decays of particles in the event. The

vertex reconstruction procedure [134] uses vertex finding algorithms, that associate the tracks to a set of vertex candidates, and fitting algorithms, that determine the position of the vertices and the associated uncertainties. The number of PVs in each event is correlated to the $\langle\mu\rangle$ parameter, which describes the average number of interactions per bunch crossing (see Section 2.1.2), and is used as a direct measurement of the amount of pileup for a variety of calibration purposes. The PV associated to the main hard scattering interaction in the event is selected as the one for which the sum of the squared momenta of the associated tracks ($\sum p_{\mathrm{T}}^2$) is the largest. After the main PV is identified, it is useful to express the impact parameters d_0 and z_0 of the tracks with respect to its position.

3.2.3 Electrons and photons

The characteristic experimental signature of electrons and photons in ATLAS consists of an electromagnetic shower in the ECAL and, for electrons, an associated track in the ID. In both cases the selection procedure [135] begins with a sliding window algorithm that searches for clusters of energy deposits in the calorimeter, scanning it in blocks of 3×5 cells with individual transverse size of 0.025×0.025 in η and ϕ.
The discrimination between electron and photon candidates is implemented by matching calorimeter clusters with ID tracks. As described in Refs. [136,137], the electron and photon identification is based on a dedicated tracking strategy where the standard ID tracks loosely matched to clusters are refitted with a Gaussian Sum Filter (GSF) algorithm [138] that takes into account the non-linear bremsstrahlung effects. This results in a new track collection which is used in the following stages of the identification procedure, yielding improvements up to 5% in the electron identification efficiency, particularly significant in the large pseudorapidity regions where the material thickness of the ID is larger.
The simplest photon candidates are identified as clusters with no corresponding track in the ID, but it is also possible to reconstruct converted photons that produce an electron-positron pair before reaching the ECAL. The latter case is characterised by the presence of two oppositely charged tracks matched to the cluster or, if the opening angle of the electron-positron pair is large, a single track with no hits in the innermost silicon layers [139]. Electron candidates are instead reconstructed as clusters matched to a track with a sufficient numbers of silicon hits. After the matching, the calorimeter cluster is rebuilt using blocks of 3×7 cells in the barrel and 5×5 cells in the end-caps.
The computation of the total energy of the clusters includes corrections from the pre-

shower sub-detector, to quantify the energy lost by the particles before the ECAL, and from the cells surrounding the bulk of the cluster, to estimate the lateral and longitudinal leakage of the shower. For electrons, the final energy is obtained by combining the cluster energy with the momentum of the track, while the η and ϕ coordinates are derived exclusively from the track[2]. The energy calibration is performed using a mixture of MC based algorithms and data driven corrections derived from $Z \to ee$ events [140, 141].
Once the electron and photon candidates are reconstructed, additional selections are applied in order to suppress the backgrounds from the misidentification of different types of objects (fakes). These include dedicated *identification* algorithms, that exploit variables related to the shape of the electromagnetic showers, track quality and matching requirements and, for electrons, information about the transition radiation released in the TRT. The electron identification is based on a *Likelihood* approach built from the probability distributions of discriminating variables in signal and background candidates. For both electrons and photons a dedicated set of identification Working Points (WPs) is defined, with different selection efficiency and purity[3]. Similarly, isolation variables and WPs are defined using information from both ID tracks and calorimeter clusters.
The offline selection efficiency of electrons and photons can be factorised as:

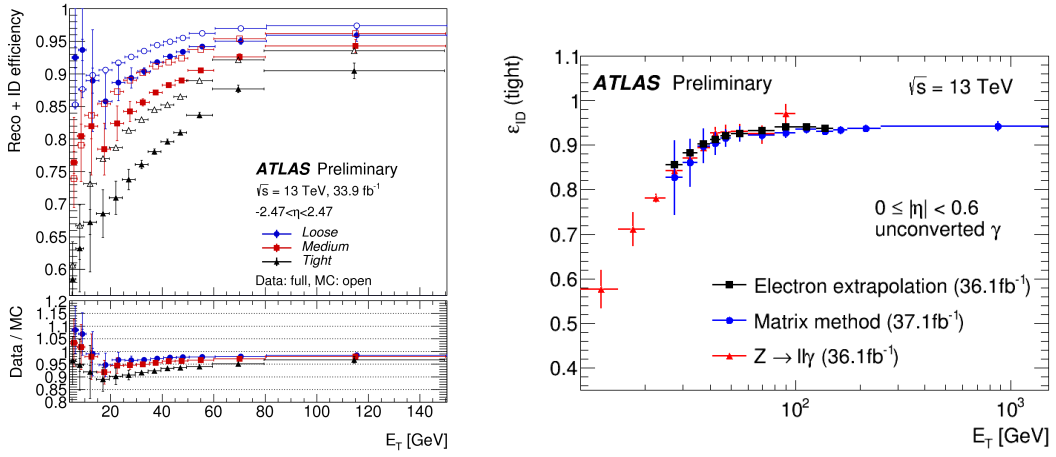
$$\epsilon_{\text{tot}} = \epsilon_{\text{reco}} \times \epsilon_{\text{ID}} \times \epsilon_{\text{iso}} \tag{3.3}$$

where each step of the selection procedure contributes separately to the total. The reconstruction and identification efficiency in 2016 data is shown in Figure 3.5 for electrons and unconverted photons, highlighting a data-MC discrepancy in the electron identification performance caused by a mis-modelling of the TRT conditions. Calibration corrections are derived in the form of efficiency SFs, that are applied to the MC to match the observed performance in data.

3.2.4 Muons

Muons appear as charged particle tracks in the ID and MS, with a finite curvature caused by the magnetic field of the solenoid and toroid systems. Their selection procedure [143] begins with two independent track fits in the ID and MS, that are then combined using different algorithms depending on the signature in each sub-system. The ATLAS software is able to identify the following types of muon candidates:

[2] If the track has no hits in the silicon detectors, the η coordinate is derived from the cluster.
[3] The purity is defined as the fraction of real electrons (photons) among the selected objects.



(a) (b)

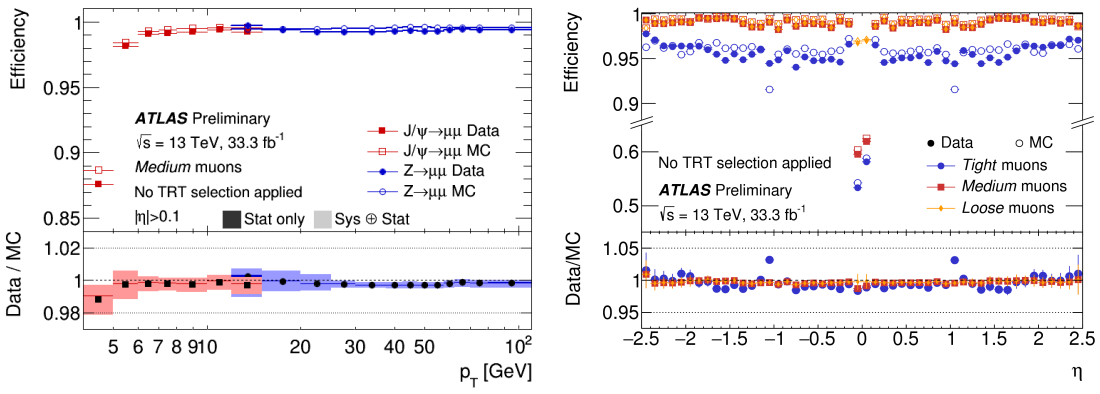
Figure 3.5: Reconstruction and identification efficiencies for electrons (a) and for unconverted photons (b) [142]. The electron efficiency measured in $Z \rightarrow ee$ events in both data and MC, while the photon efficiency is measured through different data-driven techniques [137] in the central pseudorapidity region ($|\eta < 0.6|$).

Combined When matching tracks are found in both ID and MS, a global refit is performed in order to obtain a combined muon track that describes the trajectory of the particle through the entire detector, taking into account the energy loss in the calorimeter. This signature provides a good momentum resolution in all p_{T} ranges, driven by the ID at low p_{T} and by the MS at high p_{T}. The longitudinal acceptance of combined muons is limited by the ID ($|\eta| < 2.5$).

Extrapolated Muons in the forward region ($2.5 < |\eta| < 2.7$) yield a track in the MS but lie outside the acceptance of the ID. Their trajectory is reconstructed from the MS track, with a loose requirement on its compatibility with the interaction point in the centre of the detector.

Segment tagged Tracks in the ID are classified as muons if a corresponding hit in the inner layer of the MS is found. In general, this type of signature occurs only for muons with low p_{T} or as a result of acceptance cracks in the MS.

Calorimeter tagged Finally, tracks in the ID are classified as muons if they can be matched with an energy deposit in the calorimeter compatible with the signature of a MIP, with no associated MS track. This happens primarily in the central pseudorapidity region ($|\eta| < 0.1$) due to the presence of a crack in the barrel of the



(a) (b)

Figure 3.6: Muon reconstruction and identification efficiency as a function of p_{T} (a) and η (b) in the ATLAS detector [144].

MS.

Additional identification selections are applied in order to separate prompt muons from various background sources, mostly pion and kaon decays. The selections employ track quality variables in both ID and MS, including the normalised χ^2 of the combined track fit, and other requirements on the compatibility between the charge and momentum measurements in the two sub-systems. In addition, track-based and calorimeter-based isolation variables are used to further select prompt muons, produced at significant angular distance from other objects in the event.
Similarly to the case of electrons and photons, dedicated sets of identification and isolation WPs are defined to meet the requirements of different ATLAS analyses. The performance is calibrated using the tag-and-probe method in $Z \to \mu\mu$ or $J/\psi \to \mu\mu$ events, and SFs are obtained for each step of the selection process. The muon reconstruction and identification efficiencies as a function of p_{T} and η are shown in Figure 3.6.

3.2.5 Jets

As a result of the confinement properties of QCD, elementary quarks and gluons from proton collisions generate collimated sprays of particles (jets) [145], which appear as showers in the calorimeter system associated to tracks in the ID. The analyses presented in this thesis employ Calorimeter jets, that reconstruct the four-vector of the original partons by measuring the energy of the showers in the ECAL and HCAL.

Jet reconstruction

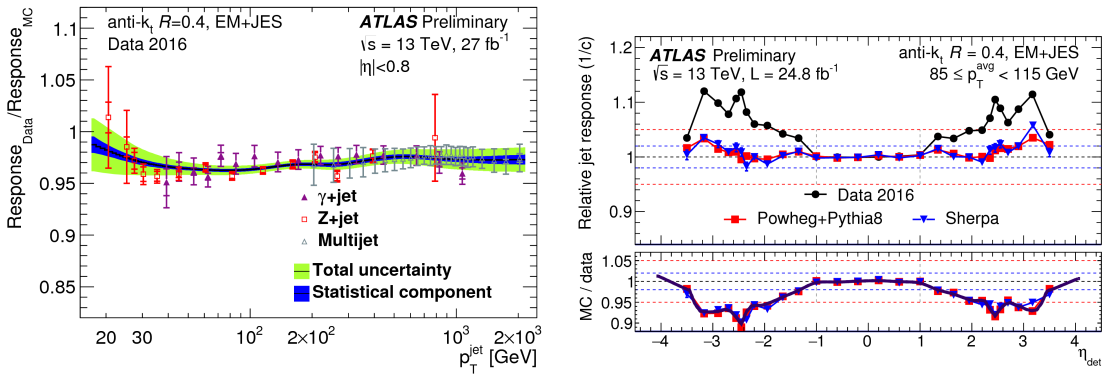
The energy of the showers is measured by forming clusters [146, 147] in the calorimeters, similarly to what is done for the triggers described in Section 3.2.1 but with a three-dimensional structure. The clusters are built through an iterative algorithm that begins with the identification of seed cells with energy deposits above a high threshold t_{seed} and continues by adding neighbouring cells with a lower threshold t_{cell}. Once the bulk of the cluster is formed, an external layer of cells is also added in order to include potential leaks of the shower.
The collection of clusters in each event is used as input for jet-finding algorithms, which treat them as massless objects with four-vectors of energy $E = \sum_i E^i_{\text{cell}}$ and angular coordinates determined by their position in the detector. The standard jet-finding algorithm of the ATLAS Collaboration is the anti-k_{T} algorithm [148], where a collection of jets is formed through an iterative recombination of the input clusters. The characteristic shape of the output jets is a cone with angular width driven by a free parameter R, which is set to 0.4 in all the analyses discussed in this thesis. As discussed in Ref. [145], the anti-k_{T} algorithm is infrared and collinear safe, meaning that the output jet collection remains unchanged if the event is modified by adding a collinear splitting of a parton or an extra soft emission in the final state.

Jet calibration

Jets obtained from calorimeter clusters are reconstructed at the characteristic energy scale of Electromagnetic (EM) objects (EM scale), which underestimates the hadronic energy deposits due to the non-compensation of the calorimeter and to other detector-related effects. The jet calibration procedure used for the analysis of 2015 and 2016 data follows the EM + Jet Energy Scale (JES) scheme [149], where a set of corrections parametrised by the p_{T} and η of the jets are applied to match the energy of the initial partons. The JES is defined by the following sequence of corrections:

Vertex correction The four-vectors of the jets reconstructed by the anti-k_{T} algorithm, which are initially pointing to the geometrical centre of the detector, are adjusted to point to the PV in the event.

Pileup correction The effect of pileup on the energy of the jets is mitigated through an area-based subtraction procedure [150], which removes the average energy deposits in the calorimeters produced by the underlying event.



(a) (b)

Figure 3.7: Momentum (a) and pseudorapidity (b) dependence of the jet response in the ATLAS detector [153]. The left plot shows the ratio of the jet response in data and MC obtained from the combination of three in situ calibration techniques, while in the right plot the relative jet response is shown as a function of η in di-jet events for data and two different MC samples. The discrepancy between data and MC for $|\eta| > 1.0$ is ascribed to the modelling of the material budget of the detector in the forward regions.

Jet energy and η correction Jets are further calibrated by applying energy and pseudorapidity corrections derived from MC simulation, where the reconstructed kinematical properties are compared to the true values in bins of reconstructed p_{T} and η.

Global sequential correction Additional corrections, known as Global Sequential Corrections [151], are applied based on individual jet properties, such as the fraction of energy deposited in different calorimeter layers or the number and type of associated tracks.

In-situ calibration A final set of in-situ corrections accounts for discrepancies between data and simulation in the measurement of the p_{T} of the jets [152]. The corrections are obtained from multijet and γ/Z+jet events, where the momentum of a probe jet is balanced against a well-measured reference object. The jet response function, defined as $p_{\mathrm{T}}^{\mathrm{probe}}/p_{\mathrm{T}}^{\mathrm{ref}}$, is shown in Figure 3.7 as a function of p_{T} and η.

Jet cleaning selections and Jet Vertex Tagger

Dedicated selections are applied to reduce the background of fake jets, that may arise from a variety of sources ranging from hardware effects, LHC beam-gas interactions or cosmic ray showers. Quality requirements are introduced for several parameters, including the fraction of energy deposited in the different layers of the calorimeter system and the fraction of jet p_{T} measured by the tracks in the ID. The above requirements define a "bad jet" condition, which is used in the analyses to suppress the fake contamination.
In addition, a Jet Vertex Tagger (JVT) [154] variable is developed in order to remove jets produced by pileup effects in the kinematical range $20 < p_{\mathrm{T}} < 60$ GeV and for $|\eta| < 2.5$. JVT is a multivariate discriminant built from variables related to the fraction of charged tracks in the jets that point to the PV, indicating the compatibility of the jet with originating from the hard scattering interaction. Particular efforts have been dedicated to ensure that JVT is stable as a function of the number of PVs in the event, to avoid the introduction of pileup-dependent biases. The performance of the JVT selections is calibrated using $Z \to \mu\mu$ + jets events, and efficiency SFs are derived in bins of jet p_{T} for a set of WPs.

3.2.6 *b*-tagging

The ability to identify the products of bottom quark hadronisation (b-hadrons) is crucial for the analyses presented in this thesis and for several more areas of research within the ATLAS Collaboration. Bottom quarks generate hadronic jets (b-jets) with the same mechanism of the other coloured particles, so they are reconstructed by the standard jet algorithms described in Section 3.2.5. Once the jets are formed, dedicated b-tagging algorithms [155] are employed to assess the presence of a b-hadron within their cone.

b-tagging algorithms

The characteristic properties of b-jets in ATLAS are shown in Figure 3.8, where a comparison with a standard light jet is presented. Due to their lifetime of the order of 10^{-12} s, b-hadrons can travel a measurable distance in the detector before decaying, yielding a secondary vertex inside the jet. The impact parameter of the tracks from the b-hadron decay is also expected to differ from zero, providing a useful tool for jet flavour discrimination. A similar experimental signature is obtained for jets from charm quark hadronisation (c-jets), with lower average displacement of the secondary vertex due to the

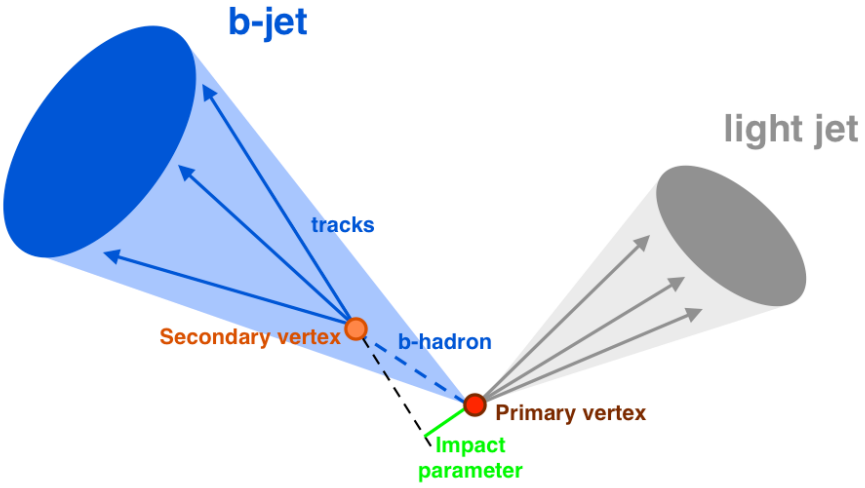


Figure 3.8: Comparison between the characteristic properties of light jets and b-jets in the ATLAS events. The trajectory of the b-hadron, the secondary vertex and the impact parameter of a track from its decay are highlighted.

shorter lifetime of c-hadrons. This implies that dedicated algorithms can be developed also to identify c-jets [156], but also that c-jets are a significant source of background for b-tagging.
The ATLAS Collaboration has developed three types of b-tagging algorithms that identify b-jets following alternative approaches [157]. The outputs of these algorithms are then combined into a single multivariate tagger, MV2, that provides the best discriminating power between the various jet flavour hypotheses. A brief overview of the major characteristics of each algorithm is given below:

IP2D and IP3D The first approach involves the use of the impact parameters of the tracks as main discriminating variables, since they are expected to be larger for tracks that originate from the decay of b-hadrons. The significance of the transverse impact parameter of the tracks, d_0/σ_{d_0}, is employed by both IP2D and IP3D taggers, and the latter uses also the longitudinal impact parameter significance, $z_0 \sin\theta/\sigma_{z_0 \sin\theta}$. The probability distributions of these quantities are employed to define ratios of b-jet and light-jet hypotheses, which are then combined into a single discriminant.

Secondary Vertex Finding Algorithm A second possibility is to perform an explicit reconstruction of secondary vertices within the cones of the jets. The Secondary

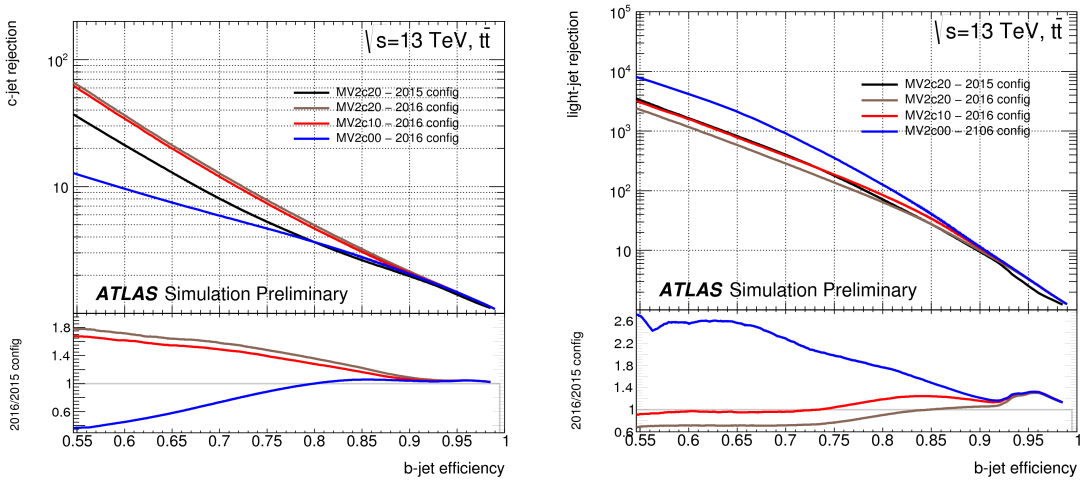
Vertex Finding algorithm uses as input all pairs of tracks associated to each jet, forming two-track vertices that are then discarded based on their compatibility with background sources such as hadronic interactions with the detector, photon conversions or decays of long-lived particle (kaons or Λ baryons). Additional quality requirements on the tracks are introduced to suppress the background from pileup vertices, and to reduce the probability of reconstructing fake vertices in dense track environments.

Decay Chain Multi-Vertex Algorithm (JetFitter) A third type of algorithm, JetFitter [158], performs a reconstruction of the full decay chain of b-hadrons inside the jets based on their expected topological structure. The trajectory of the b-hadron and of its decay products is extracted using a Kalman filter [159].

Mutivariate Algorithm (MV2) A Boosted Decision Tree algorithm [160], MV2, is implemented with the ROOT TMVA package [161] to combine the output of the three taggers described above. The algorithm is optimised in three different versions, MV2c00, MV2c10 and MV2c20, where the numbers are related to the fraction of c-jets (in $\%$) that are present in the background training sample in addition to a majority of light flavour jets. The analyses in this thesis employ the MV2c10 tagger in a configuration where the training sample contains 7% c-jets and 93% light jets[4] [157].

The performance of different configurations of the MV2 algorithm in $t\bar{t}$ events is shown in Figure 3.9, where the efficiency of correctly identifying b-jets is compared with the rejection probability of c-jets and light jets, defined as the inverse of the corresponding mis-tag rates. The versions of the algorithm with more c-jets in the background training sample yield a better c-jet rejection and a worse light jet rejection, while the opposite happens if no c-jets are included in the training. Benchmark cuts on the output of the MV2c10 algorithm are used to define a set of b-tagging WPs, optimised to provide specific efficiency values (60%, 70%, 77% and 85%) on a reference $t\bar{t}$ sample. In addition, to correct for a degradation of the b-tagging performance at high values of the jet p_{T}, a more sophisticated set of WPs is defined by modifying the cut threshold of MV2c10 as a function of the p_{T}, yielding a constant tagging efficiency in all kinematic regimes. More detail on the performance of b-tagging algorithms is provided in Chapter 4.

[4] In its first Run 2 configuration MV2c10 was trained using a background sample with 10% c-jets, as indicated by the name of the tagger. Its performance was then re-optimised in 2016 by changing the c-jet fraction to 7%, but the name was left unchanged.



(a) (b)

Figure 3.9: Performance of the MV2 algorithm in $t\bar{t}$ events, showing the rejection versus efficiency curves for the c-jet (a) and light jet (b) backgrounds [157]. The curves in each plot are obtained using alternative configurations of MV2 (2015 and 2016) and different fractions of c-jets in the background training sample. The results in this thesis are based on the MV2c10 tagger with the 2016 configuration.

b-tagging calibration

In order to apply b-tagging requirements in physics analyses, the performance of the algorithms must be accurately measured in data and MC simulation. The first parameter to be measured is the b-jet efficiency, but it is also important to determine the probability of mis-tagging c-jets (referred to as c-tagging efficiency) and light jets (mis-tag rate).
A variety of techniques have been developed to measure the above parameters in data, allowing to derive calibration corrections in terms of efficiency SFs for real b, c and light jets separately. The baseline b-jet calibration, performed by selecting $t\bar{t}$ events in di-leptonic final states [162], is able to achieve a precision of the order of a few $\%$ in the jet p_{T} range between 20 and 300 GeV. The reach can be extended to approximately 500 GeV by performing a similar analysis in semi-leptonic events, which benefits from more available statistics. The c-tagging efficiency is measured by selecting events with c-jets produced in association with a W boson decaying into a lepton and a neutrino [163]. The presence of a c-jet is determined by reconstructing a soft muon from the semi-leptonic decay of the c-hadron, and the purity of the sample is improved by correlating the charge of the two leptons in the event. Finally, a negative-tag method [164] is employed to

measure the mis-tag rate that originates mainly from tracking resolution effects. Several more calibration strategies have been studied within the ATLAS Collaboration, and a comprehensive overview of the most relevant ones is provided in Ref. [155].

3.2.7 Missing transverse momentum

Thanks to the hermetic coverage of the ATLAS detector, the presence of invisible particles in the events can be assessed by reconstructing a significant momentum imbalance in the transverse plane. Indeed, while the longitudinal boost of the hard scattering products varies on an event-by event basis, the net transverse momentum of the reconstructed objects in the final state is expected to be balanced within finite resolution effects, unless at least one particle has escaped the detection.
The momentum imbalance in the transverse plane is measured by the two-dimensional missing transverse momentum vector, $\mathbf{p}^T_{\text{miss}}$ [165], defined as the negative vectorial sum of the x-y components of the four-momenta of the visible objects in the event:

$$\boldsymbol{p}^T_{\text{miss}} = -\sum_{\text{all}} \mathbf{p}^T_{\text{obj}}. \tag{3.4}$$

The magnitude of $\mathbf{p}^T_{\text{miss}}$ is known as missing transverse momentum or missing transverse energy (E_T^{miss}), and the azimuthal coordinate of the vector is simply given by:

$$\phi_{E_T^{\text{miss}}} = \arctan(p^y_{\text{miss}}/p^x_{\text{miss}}) \tag{3.5}$$

where the x and y projections of $\mathbf{p}^T_{\text{miss}}$ are used.
The sum in Equation 3.4 can be split in a hard term that includes all the relevant physics objects reconstructed by the detector (electrons, muons, τ leptons, photons, jets) calibrated at their appropriate energy scale, and a soft term that contains the remaining detector signal objects in the event. The individual contributions to the hard term are added in sequence following a fixed order: electrons, photons, hadronically decaying τ leptons, jets and muons. During the reconstruction process, the algorithm solves the overlaps between different types of objects by keeping those that appear earlier in the chain. For the analyses presented in this thesis the hadronic τ term is not included in the computation, because explicit τ reconstruction is not employed, so any hadronic τ in the events is reconstructed as a jet and contributes to the jet term of the E_T^{miss}. The soft term can be reconstructed using ID tracks (track-based soft term) or calorimeter signals (calorimeter-based soft term): for the analysis of 2015 and 2016 data, the track soft term is used as a baseline choice due to its lower sensitivity to pileup effects [166].

The performance of the E_T^{miss} is studied in events where invisible particles are present, such as leptonically decaying W or $t\bar{t}$ samples, in order to validate the reconstruction algorithms and remove potential biases in the absolute scale. In parallel, a measurement of the resolution is obtained from the width of the E_T^{miss} distribution in $Z \to \ell\ell$ events where no real sources of momentum imbalance are expected.

4 | b-tagging performance studies with Monte Carlo simulations

This chapter presents a study of the performance of the baseline b-tagging algorithm in ATLAS, MV2c10, with focus on the boosted kinematic regime of the b-jets. Since the data-based calibration analyses have a limited reach at high p_{T} due to the small size of the available b-jet sample [155], the study is entirely performed using MC simulations, where b-jets with p_{T} up to a few TeV can be obtained. The workflow of the analysis is outlined in Section 4.1 together with a description of the technical aspects, while Section 4.2 is dedicated to a full discussion of the results.

4.1 Analysis strategy

As described in Section 3.2.6, the MV2c10 algorithm computes a multivariate b-tagging weight for each input jet, based on a set of variables related to the jet and to the associated tracks. Since these quantities are reconstructed by the detector with finite accuracy, their measured values carry an experimental uncertainty that affects the resulting MV2c10 weight, causing a systematic uncertainty on the performance of the tagger. The purpose of the study presented in this chapter is to evaluate the impact of various sources of systematic uncertainty on the performance of MV2c10, quantifying the uncertainty on the tagging efficiency as a function of the p_{T} of the jets. The study is particularly relevant for jets with p_{T} higher than a few hundred GeV, for which the data-based calibration analyses are unable to provide a direct measurement of the b-tagging performance[1].

[1]For reference, the baseline b-tagging efficiency calibration [155] has a b-jet p_{T} reach of 300 GeV.

4.1.1 Workflow and object definitions

The analysis is performed with a software that runs the b-tagging algorithm on a set of benchmark MC samples, computing the value of the output MV2c10 discriminant for all input jets in each event. The output discriminant is built first using the nominal track and jet collections, then using modified input collections according to the expected experimental uncertainties on tracks and jets (see Section 4.1.3).
The performance of the MV2c10 tagger is studied separately for real b, c and light jets (see Section 3.2.6) with flavour determined at truth level by performing an angular (ΔR) matching between jets and b or c-hadrons in the events. The b-tagging decision is taken by applying a cut on the MV2c10 outputs of each jet, and efficiency distributions as a function of the jet p_{T} are derived for each different source of systematic uncertainty. By comparing the efficiency values in each bin of p_{T}, it is hence possible to assess the impact of each variation on the performance of the MV2c10 tagger.
The studies are entirely based on Calorimeter jets reconstructed through an anti-k_{T} algorithm with radius $R = 0.4$, as introduced in Section 3.2.5. Tracks are required to satisfy the "Loose" quality criteria [167], which corresponds to the following selections:

- $p_{\mathrm{T}} > 400$ MeV and $|\eta| < 2.5$.
- At least 7 hits in Pixel and SCT layers.
- A maximum of 1 shared module (Pixel hit shared by two or more tracks, or double hit in a single SCT layer).
- A maximum of 2 silicon holes (Pixel or SCT layers with no hit in correspondence of the track trajectory) and 1 pixel hole.

4.1.2 Monte Carlo samples

The analysis of the performance of the MV2c10 algorithm requires large samples of b-jets covering a wide range of kinematic regimes, with p_{T} distributed from a minimum of 20 GeV up to a few TeV. In order to obtain them, the following MC samples are used:

Inclusive $\mathrm{t\bar{t}}$: The pair production of SM top quarks yields two bottom quarks in each event, originating from the $t \to Wb$ decay. The majority of the b-jets have p_{T} values below the top mass scale, but occasionally also b-jets with higher boosts are produced. The nominal $t\bar{t}$ sample in ATLAS is simulated using Powheg [126] as

event generator and Pythia 6 for the PS [122, 123], with the EvtGen [168] software regulating the decays of the b-hadrons. Two alternative samples that use Pythia 8 or Herwig [124] instead of Pythia 6 are also available for comparison.

$Z' \to bb$: The decay of a Z' resonance into a pair of bottom quarks provides a sample of b-jets with average p_{T} approximately given by $m_{Z'}/2$. The studies presented in this chapter employ a $Z' \to bb$ sample with $m_{Z'} = 5$ TeV, where a large amount of b-jets with high p_{T} is obtained. The sample is simulated with Pythia 8 interfaced with EvtGen.

$W' \to tb$: A third sample of b-jets is obtained by simulating the decay of a W' resonance into a top-bottom pair, with the top decaying into a W boson and a bottom quark. The peculiarity of this sample is that 50% of the b-jets originate directly from the decay of the W', while the remaining 50% is produced through the decay of the top quark. W' samples are generated with MadGraph [125] interfaced with Pythia 8 and EvtGen.

4.1.3 Systematic uncertainties

The performance of b-tagging algorithms is affected by a variety of systematic effects related to the input tracks [169] and to global jet properties. This paragraph presents the sources of systematic uncertainty that are considered for the present study, providing a description of their origin and their expected impact.

Impact Parameter Resolution

The transverse and longitudinal impact parameters of the tracks (d_0 and z_0) are essential for the computation of the multivariate MV2c10 weight, because the tracks from the decay of b-hadrons are distanced from the PV.
The intrinsic resolutions of d_0 and z_0 are measured in multi-jet and $Z \to \mu\mu$ events [170], and are found to be larger in data compared to MC simulations. A dedicated analysis is performed to extract d_0 and z_0 smearing corrections, to be applied to each track in order to reproduce the resolution observed in data. Alternatively, since the resolution in MC is influenced by the modelling of the inactive modules of the Pixel detector, a correction can also be derived by re-evaluating the track parameters after randomly disabling a 5% fraction of them. In both cases, an impact on the performance of the MV2c10 algorithm is expected.

Fake Tracks

Uncorrelated hits in the ID can cause the reconstruction of fake tracks as a result of combinatorial effects, particularly relevant in high pileup conditions. Under the assumption that the number of real (i.e. non-fake) tracks is proportional to the pileup parameter μ, the number of fakes can be estimated by measuring the total number of tracks as a function of μ, attributing any deviation from a linear behaviour to the fake component [169]. The difference between the number of fake tracks in data and MC is found to be around 30% independently of p_{T} and η, so a systematic variation is implemented by randomly removing 30% of them from the nominal input collection in the MC samples.

Track Reconstruction Efficiency

Another relevant parameter for the performance of b-tagging is the reconstruction efficiency of real tracks, which is affected by the knowledge of the amount of material in the ID [171]. The associated uncertainty is evaluated as a function of the track p_{T} and η by measuring the difference in tracking efficiency between the nominal sample and one where additional material is added. Four independent variations are implemented:

- Passive material in the ID scaled by 5%.
- Passive material in the IBL layer scaled by 40%.
- Passive material in the Patch Panel 0 (PP0) region of the ID scaled by 50%.
- Change in the Geant 4 physics model used for the event reconstruction.

Tracking in dense environments

Specific measurements of the track reconstruction efficiency are performed in the core of hadronic jets [172], where a deterioration of the tracking performance is expected due to the high density of the experimental environment. The fraction of lost tracks is measured in bins of jet p_{T} within $\Delta R(\mathrm{track}, \mathrm{jet}) < 0.1$ in data and MC, and the observed difference of approximately 0.4% is applied as a systematic uncertainty by randomly dropping 0.4% of the tracks inside the jet core.

Weak modes in the Inner Detector alignment

The alignment of the ID is performed with a track fit procedure [173] based on the minimisation of a χ^2 function, which depends on a set of alignment parameters. The de-

formations of the detector material that have little or no impact on the value of the χ^2 are known as weak modes, and represent a source of ambiguity in the alignment procedure. As a result, the following potential biases on the performance of the tracking algorithms [174] may occur:

- Charge-antisymmetric bias on the reconstructed q/p ratio of the tracks.
- Bias on the transverse and longitudinal impact parameters (d_0 and z_0).

The impact of the weak mode biases on the performance of MV2c10 is studied by shifting the tracking parameters within the expected systematic uncertainties, evaluating the corresponding efficiency variations.

Jet Energy Scale and Resolution

The calibration of the jets to the appropriate energy scale [175], described in Section 3.2.5, carries an uncertainty that influences their measured momentum, potentially causing a bias in the p_{T}-binned b-tagging efficiency. The JES uncertainty is implemented based on the results of the calibration studies, in the form of 3 variations that parametrise the effect of a larger number of corrections. Similarly, the resolution of the measured jet energy (Jet Energy Resolution, JER) also carries an experimental uncertainty [176], and the associated effect on the performance of the MV2c10 algorithm is studied.

Jet Angular Resolution

A final uncertainty is associated with the direction of the jet axis, which can be reconstructed either from the energy deposits in the calorimeter or from the tracks in the ID. The two measurements may differ due to experimental resolution effects, and a smearing factor can be applied to the MC simulation to match the resolution in data, similarly to what is done for the impact parameters of the tracks. The associated systematic uncertainty on the b-tagging is estimated by smearing the ϕ and η values of the jets by gaussian factors of 0.004 and 0.008 respectively.

4.2 Results on *b*-tagging performance

This section is dedicated to a full discussion of the performance of the MV2c10 b-tagging algorithm, including the analysis of the systematic effects described in the previous paragraphs. As anticipated in Section 3.2.6, the algorithm can be employed with two types

of WPs, implemented by defining benchmark selections on the value of the MV2c10 discriminant of the jets. In the fixed cut WPs the selection on the MV2c10 output is a constant value independent of the jet p_{T}, so the resulting performance may vary in different kinematic regimes. Alternatively, fixed efficiency WPs are defined by scaling the value of the MV2c10 cut as a function of the jet p_{T}, to ensure that the b-tagging efficiency remains constant over a wide range of momenta. The performance results in this section are presented for a fixed cut WP used for all the analyses in the next chapters, optimised to yield a b-tagging efficiency of 77% on a reference $t\bar{t}$ sample. The case of the fixed efficiency WP, for which similar conclusions can be drawn, is presented in Appendix C.

4.2.1 Nominal performance

Before studying the impact of systematic effects, it is useful to examine the nominal performance of the MV2c10 tagger for different physics processes and for different kinematic regimes of the jets.

The b-tagging efficiency as a function of the jet p_{T} is shown in Figures 4.1a and 4.1b for the 77% fixed cut and fixed efficiency WPs, comparing the performance in $t\bar{t}$ and $Z' \to bb$ samples. The general trend of the curves is similar for the two physics processes: the efficiency is relatively low for soft b-jets, it reaches a peak in the p_{T} range between 50 and 200 GeV and it decreases again for higher boosts of the jets. The performance in $t\bar{t}$ and Z' is comparable up to about 200 GeV, but a major deviation is observed in the high p_{T} region, where the efficiency in the Z' dataset is significantly lower. In particular, the fixed efficiency WP is able to yield a constant efficiency only when applied to the $t\bar{t}$ sample (which is used to optimise its p_{T}-dependent selections), while the Z' curve still shows a significant decrease at high p_{T}.

Similarly, Figures 4.1c and 4.1d show the nominal c-tagging and mis-tag efficiencies for the 77% fixed cut and fixed efficiency WPs, obtained using a $t\bar{t}$ sample. As expected, the comparison between the two plots indicates that the fixed efficiency WP yields a larger mis-identification efficiency at high p_{T} for both c-jets and light jets.

4.2.2 Relation between *b*-hadron and *b*-jet momentum

The discrepancy between the b-tagging performance in $t\bar{t}$ and Z' can be understood by analysing the different substructure of the b-jets in the two physics processes, focusing on the kinematical properties of the associated b-hadrons. In particular, the b-hadrons associated with the jets are extracted using the same truth-level ΔR matching procedure

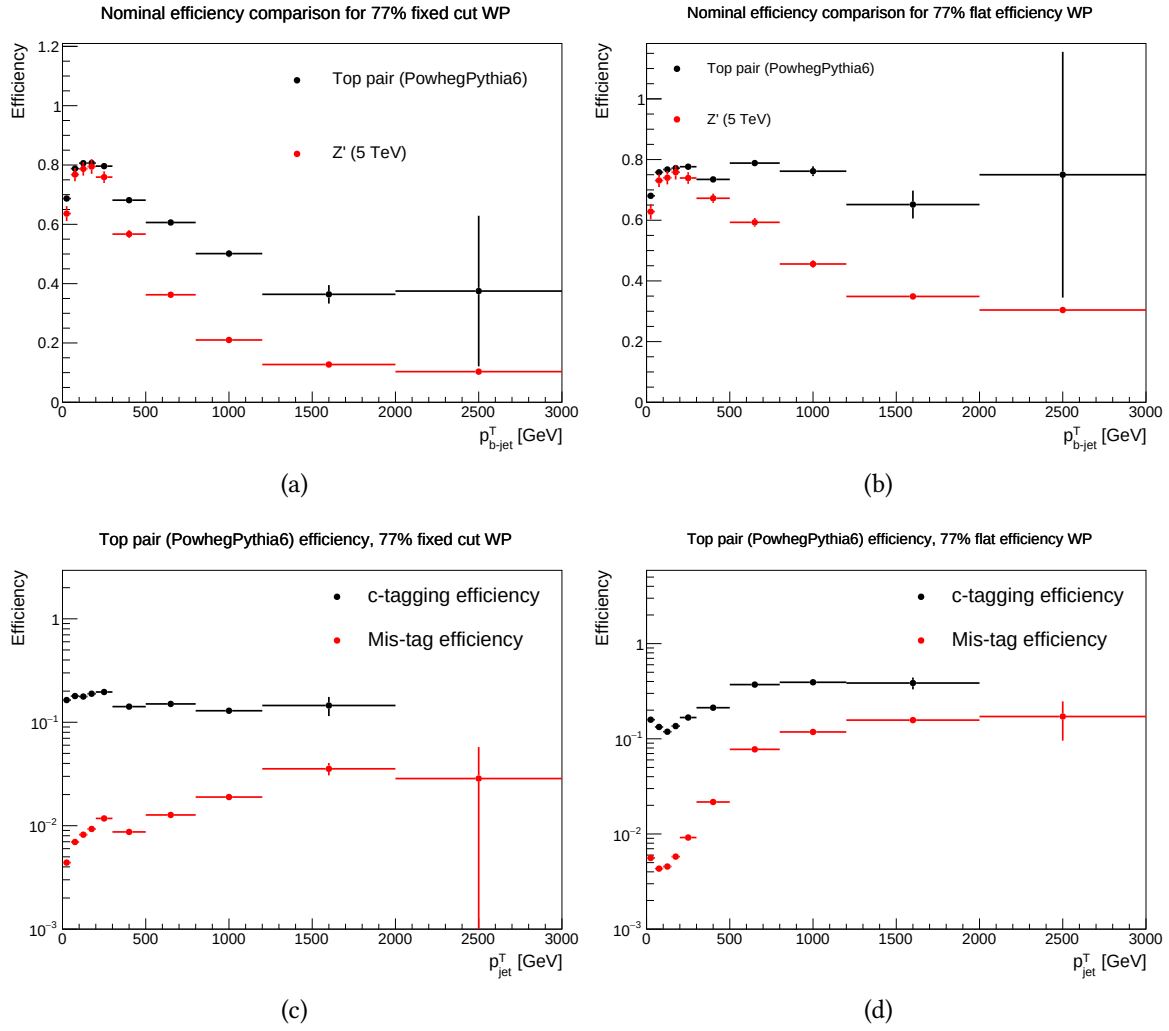


Figure 4.1: Nominal tagging efficiency for different jet types, physics processes and WPs. Figures (a) and (b) show the b-tagging efficiency as a function of the b-jet p_{T} in $t\bar{t}$ and Z' samples for the 77% fixed cut and fixed efficiency WPs, while the c-tagging and mis-tag efficiencies are compared in Figures (c) and (d) in a $t\bar{t}$ sample for same two WPs.

described in Section 4.1.1, which is already used as standard tool to determine the real flavour of the jets.

The average p_{T} of the b-hadrons as a function of the p_{T} of the associated b-jets is shown in Figure 4.2 for $t\bar{t}$ (with Pythia 6, Pythia 8 and Herwig as PS generators), Z' and W' samples. The relation is approximately linear for Z', while for $t\bar{t}$ the slope of the curve is lower and tends to decrease at high p_{T}. Lastly, the W' sample exhibits an intermediate behaviour between the the two. It is important to emphasise that the $t\bar{t}$ distributions

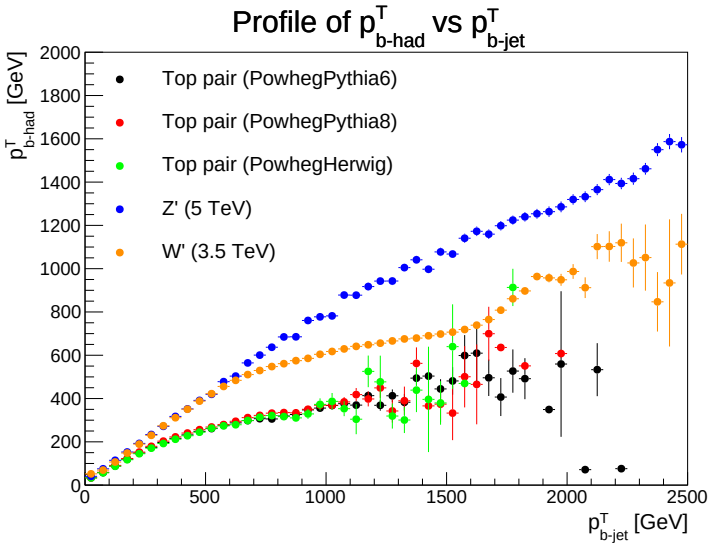


Figure 4.2: Average p_{T} of the b-hadrons as a function of the p_{T} of the associated b-jets in $t\bar{t}$ (with Pythia 6, Pythia 8 and Herwig as PS generators), Z' and W' samples. Both quantities in the plot are extracted from the truth information of the MC simulations. The observed discrepancies between the different physics processes are discussed in detail in Section 4.2.2.

are independent of the chosen PS generator, proving that the observed discrepancies are not related to different types of hadronisation or fragmentation models in the nominal $t\bar{t}$ and Z' samples, but rather to the intrinsic features of the physics processes.
Further details are provided in Figure 4.3, where the distribution of the ratio between b-hadron and b-jet p_{T} is shown separately in individual bins of b-jet p_{T} for the same four MC samples. When the 50-100 GeV bin is considered (Figure 4.3a) the distributions appear similar, while for higher p_{T} ranges the discrepancies between the different physics processes become progressively more relevant. In the 1000-1100 GeV bin (Figure 4.3d) the $t\bar{t}$ and Z' histograms have different peaks, while interestingly the W' histogram shows a hybrid structure, with one peak similar to the $t\bar{t}$ histogram and a second peak that follows the shape of the Z' distribution.
The observed features can be interpreted by analysing the different production mechanisms of b-hadrons in the physics processes under exam. In $t\bar{t}$ events the b-hadrons originate from top quark decays, so the transverse momenta of the associated b-jets in the rest frames of the top quarks are limited by $m_t \sim 170$ GeV. This is also the case for the majority of the b-jets in the laboratory frame, which populate the first bins in Figure 4.2 and yield the largest b-tagging efficiency in Figure 4.1. The b-jets with higher

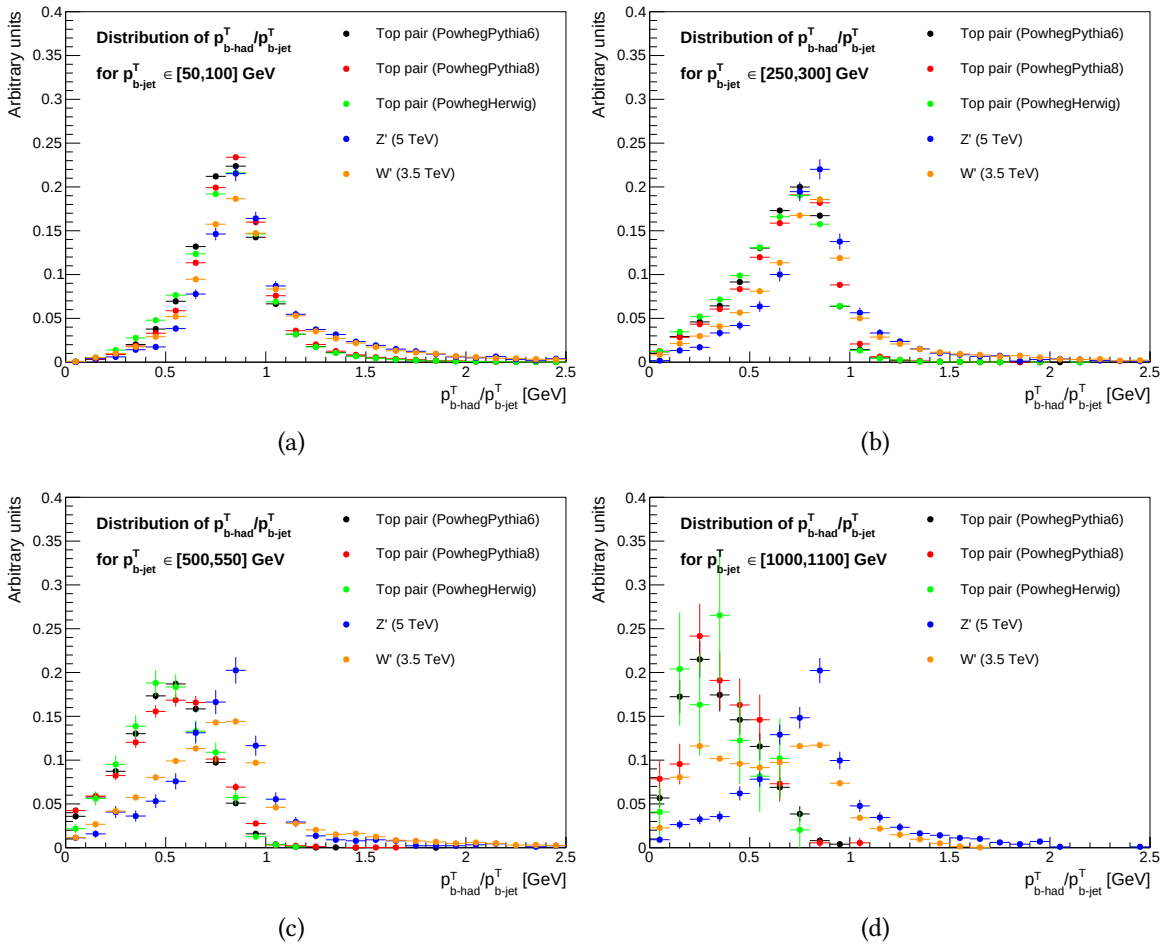


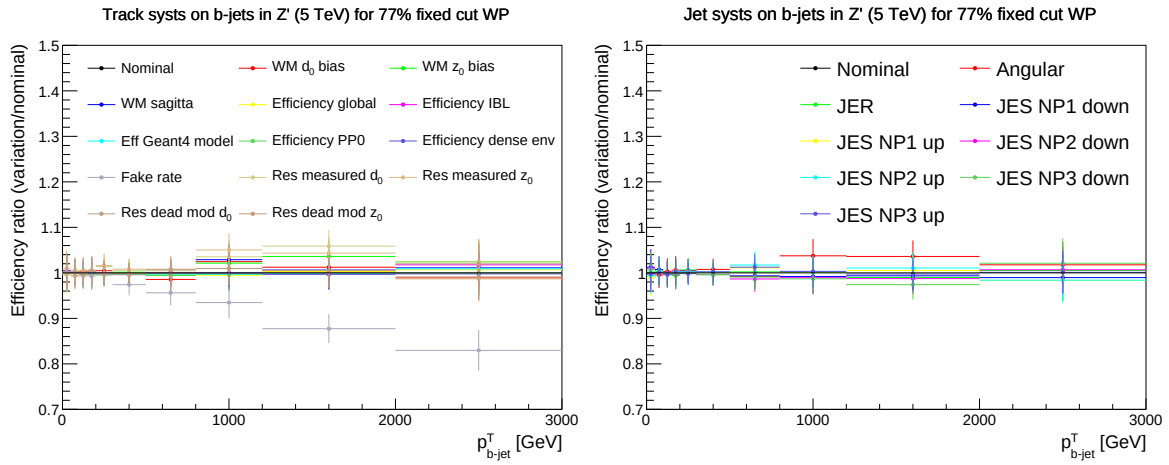
Figure 4.3: Distribution of the ratio of b-hadron and b-jet p_{T} in $t\bar{t}$ (with Pythia 6, Pythia 8 and Herwig as PS generators), Z' and W' samples in the b-jet p_{T} ranges 50-100 GeV (a), 250-300 GeV (b), 500-550 GeV (c) and 1000-1100 GeV (d). The histograms are obtained using the MC truth information and are normalised to unit area.

p_{T} must necessarily originate from top quarks with significant boost, where the decay products are collimated and can potentially overlap within the cone of the jets [177]. Indeed, Figures 4.2 and 4.3 show that the fraction of jet p_{T} carried by the b-hadrons tends to decrease in high jet p_{T} bins, implying that the reconstructed jets contain more objects from different sources. The situation is different in Z' events, where the b-hadrons originate directly from the two-body decay of the high-mass resonance, so the corresponding b-jets are well isolated and the relation between b-hadron and b-jet p_{T} is nearly independent of the p_{T} regime. Finally, in W' events a first b-hadron is produced directly in the decay of the resonance, while a second one originates from the decay of the top quark.

As a result, 50% of the jets show the same features of those from the Z' sample, while the others behave similarly to the $t\bar{t}$ ones, yielding the hybrid behaviour observed in the plots.
The larger b-tagging efficiency at high p_{T} in $t\bar{t}$ events compared to Z' can be explained by the fact that a given bin of b-jet p_{T} corresponds to different average boosts of the b-hadrons. The performance of the b-tagging algorithms is driven by the p_{T} of the b-hadron, which influences the average displacement of the secondary vertex (see Figure 3.8) and the associated quantities, such as the impact parameters d_0 and z_0. The MV2c10 algorithm is optimised for the bulk of $t\bar{t}$ events, where it achieves the highest efficiency, while the performance deteriorates when the kinematical properties of the b-hadrons are significantly altered.

4.2.3 Systematic uncertainties on *b*-tagging efficiency

After the discussion of the nominal performance of the MV2c10 tagger, it is now possible to analyse the impact of the uncertainties presented in Section 4.1.3. Since the number of systematic variations is large, it is convenient to consider separately the track-related and jet-related ones, so that the the individual contributions can be highlighted. In addition, due to the observed discrepancy in the nominal efficiency, it is useful to evaluate the size of the systematic variations separately for $t\bar{t}$ and Z' samples.
Figure 4.4 shows the uncertainty on the b-tagging efficiency as a function of the b-jet p_{T} caused by track and jet systematics in a 5 TeV Z' sample, for the 77% fixed cut WP. A remarkable feature of the Z' sample is the large amount of statistics in the boosted p_{T} regime of the b-jets, which allows to measure the uncertainty with high accuracy up to a 2-3 TeV p_{T} bin, while c-tagging efficiency and mistag rate cannot be studied due to the lack of c-jets and light jets in the final state. The size of the variations is generally low, with the largest contribution given by the track fake rate, which reduces the tagging efficiency in the highest p_{T} bins.
Analogous efficiency plots are shown in Figure 4.5 for $t\bar{t}$, where c-tagging efficiency and mis-tag rate can also be shown thanks to the sufficient amount of c-jets and light jets in the sample. Due to the lower available statistics at high p_{T}, the uncertainties are evaluated only up to a maximum p_{T} of 2 TeV. The observed uncertainties on the b-tagging efficiency (Figures 4.5a-4.5b) are generally compatible with the results obtained with the Z' sample: the dominant variations originate from track-related effects, while the jet systematics have a negligible impact. The uncertainties do not exceed $\sim 10\%$ even for



(a) (b)

Figure 4.4: Ratio of modified and nominal b-tagging efficiencies as a function of b-jet p_{T} in a 5 TeV Z' sample for the 77% fixed cut WP, showing the impact of the individual track (a) and jet (b) systematics defined in Section 4.1.3.

the highest p_{T} bin, and the size of the track fake rate uncertainty is slightly smaller than the same variation in the Z' sample.

The impact of the uncertainties on c-tagging efficiency and mistag rate from light jets in $t\bar{t}$ events is presented in Figures 4.5c to 4.5f. While the jet uncertainties are again causing minor effects, the track-related systematics produce significant variations of the efficiency in a broad range of jet p_{T}. In particular, the smearing of the resolution of both impact parameters (d_0 and z_0) causes a relevant increase of the mistag efficiency for both c-jets and light jets. This result is not surprising, since the misidentification of jets is mainly due to mis-measurements of tracking parameters [164].

The total impact of track-related systematic uncertainties on the efficiencies is shown in Figure 4.6 for b, c and light jets in the $t\bar{t}$ sample and for b-jets in the Z' sample. For simplicity, the histograms are obtained by summing in quadrature the contributions from the individual sources of uncertainty. This strategy neglects potential correlations between different systematics and may produce a double counting of similar effects, but it provides a good estimate of the overall size of the tracking uncertainties on the efficiency of the MV2c10 algorithm. The relative uncertainty on the b-tagging efficiency is below 10% for all b-jets up to 1.2 TeV, and remains below 20% even when b-jets with p_{T} higher than 2 TeV are considered. The size of the variations at high p_{T} is slightly more significant in Z' compared to $t\bar{t}$, due to the larger impact of the track fake rate uncer-

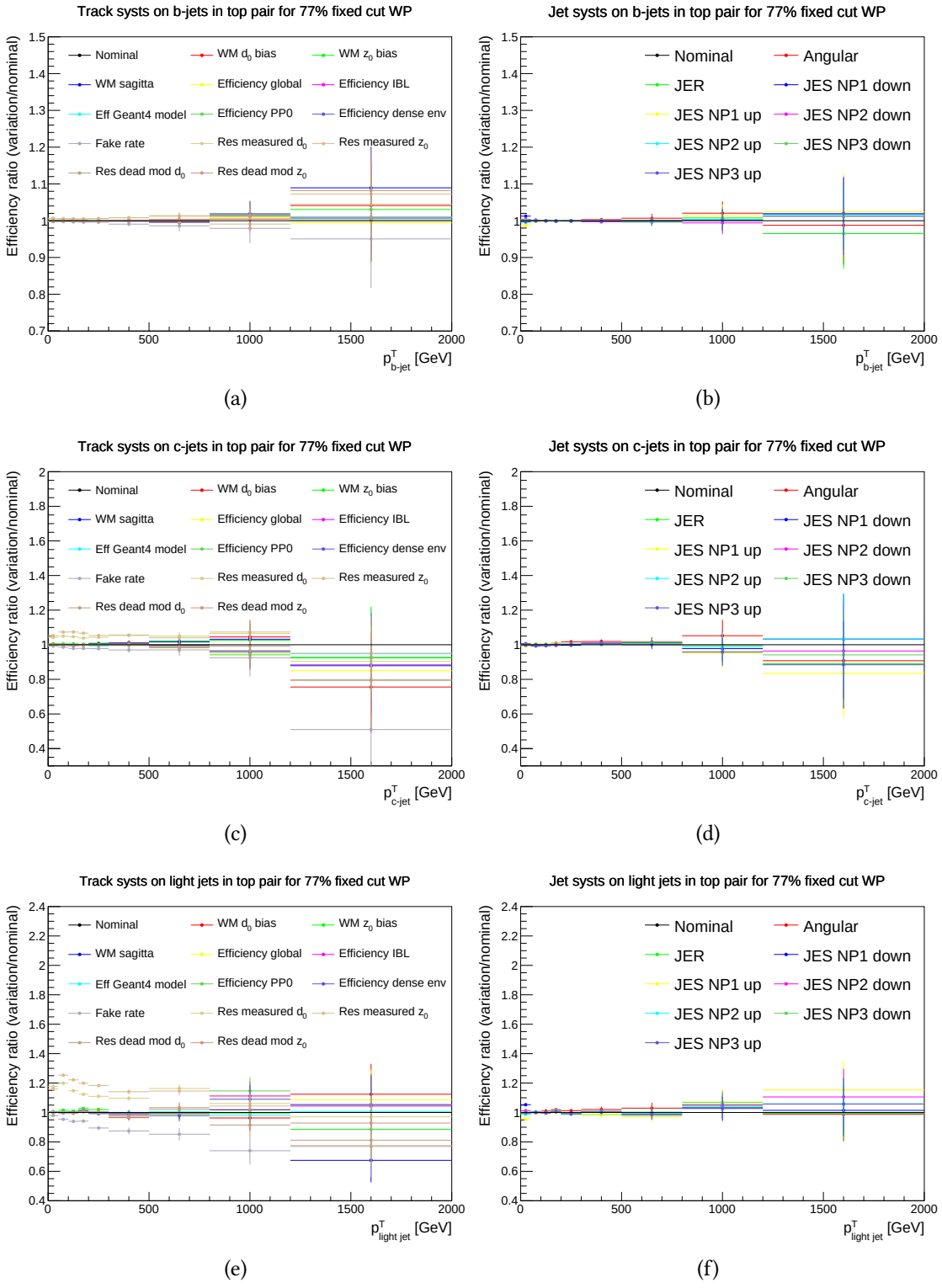


Figure 4.5: Ratio of modified and nominal b-tagging efficiencies vs jet p_{T} in a $t\bar{t}$ sample for the 77% fixed cut WP, showing the impact of the individual track (left) and jet (right) systematics defined in Section 4.1.3 on b (top), c (middle) and light (bottom) jets.

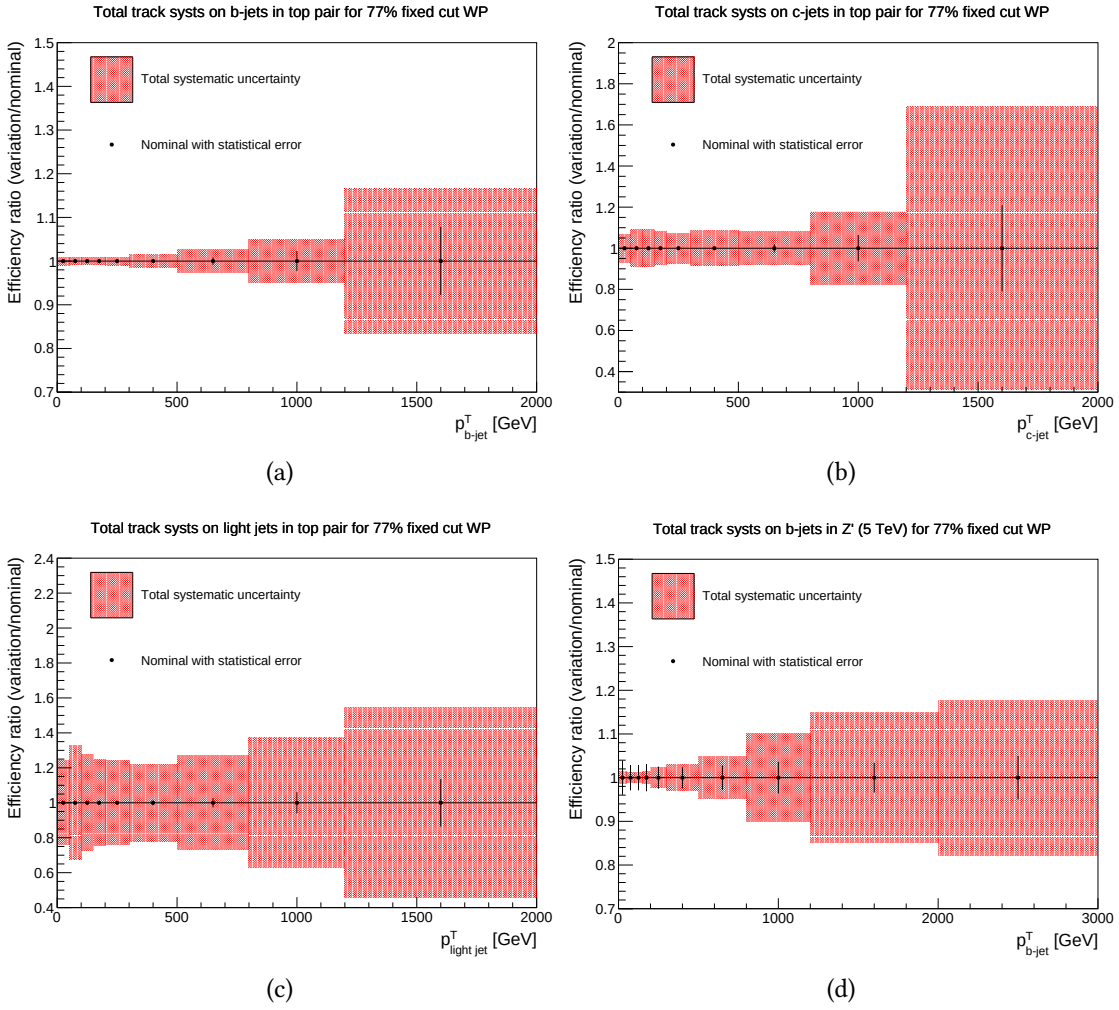


Figure 4.6: Relative uncertainty caused by tracking systematics defined in Section 4.1.3 on the tagging efficiency as a function of jet p_{T} for b (a), c (b) and light jets (c) in a $t\bar{t}$ sample, and for b-jets in a Z' sample (d), for the 77% fixed cut WP.

tainty. The uncertainty on the c-tagging efficiency is around 10% up to jet p_{T} values of 800 GeV, while the mis-tag uncertainty is stable around 30% and reaches 50% only in the highest-p_{T} bin. The corresponding plots for jet-related systematics are included in Appendix C.

4.2.4 Final remarks

The performance studies presented in this chapter are used to define a systematic uncertainty on the efficiency SFs of the MV2c10 algorithm, based on the total tracking

uncertainties shown in Figure 4.6. The uncertainty is particularly relevant in boosted regime of the jets, where the data-driven calibration techniques cannot be employed.
In addition, the comparison between the performance in $t\bar{t}$, Z' and W' samples is also showing a significant discrepancy between the b-tagging efficiencies in different physics processes, caused by the different sub-structure of the b-jets. This effect is not related to the modelling of hadronisation and fragmentation in the MC samples, and it must be taken into account when applying b-tagging selections on jets with p_{T} greater than a few hundred GeV.
Finally, the studies in this chapter are based on the 77% fixed cut WP, which is used for all analyses presented in this thesis, but similar results with the fixed efficiency WP are included in Appendix C. The analysis was also repeated for the 60%, 70% and 85% WPs, commonly used within the ATLAS Collaboration, reaching similar conclusions about both nominal performance and systematic uncertainties.

5 | Third generation SUSY searches in ATLAS

The core topic of this thesis is the search for third generation squarks in the framework of natural pMSSM models (introduced in Section 1.3.3) and models with spontaneous R-parity breaking (Section 1.3.4). While the two types of signals are targeted by independent analyses, the general approach presents several common aspects that are introduced in the present chapter before discussing the individual cases.
Section 5.1 describes the characteristic experimental signature of SUSY particles at the LHC, with emphasis on the case of third generation squarks. The dominant background processes are then introduced in Section 5.2 together with a description of the MC samples that are employed for the searches in Chapters 6 and 7. Section 5.3 presents the general workflow of the two analyses, which is common to several more SUSY searches in ATLAS, including a description of the 2015 and 2016 data samples, the online and offline event selection, the definition of signal-enriched regions and the baseline background estimation strategy. Finally, Section 5.4 is dedicated to the statistical tools that are employed for the evaluation and the interpretation of the results of the searches.

5.1 SUSY searches

The superpartners of SM particles are expected to be produced in the proton-proton collisions delivered by the LHC, yielding characteristic experimental signatures in the ATLAS detector. The production mechanism is generally assumed to be regulated by RPC couplings, that are less constrained compared to the RPV ones [178], so most ATLAS analyses are carried out by considering only pair-produced SUSY particles, with model-dependent exceptions far beyond the scope of this thesis. This section presents an overview of the typical approach to SUSY searches in ATLAS, highlighting the chal-

lenges that they pose and the strategies used to overcome them. The phenomenology of third generation squarks is discussed in detail, in order to provide the context of the analyses in Chapters 6 and 7.

5.1.1 Experimental challenges

The main difficulties faced by SUSY searches in ATLAS are related to the presence of a significant amount of background from SM processes, whose importance varies depending on the cross section and experimental signature of the targeted signals. The production cross section of the most relevant SUSY particles at the LHC [179] is shown in Figure 5.1 as a function of their mass, to be compared with the background cross sections in Figure 5.2: for SUSY particles with masses of the order of hundreds of GeV, the background cross sections are larger by several orders of magnitude. The signal cross sections are obtained by treating squarks and gluinos as decoupled particles, which means that squarks do not contribute to gluino production diagrams and vice versa. Under this assumption, the cross section of squark pair-production is found to be the same for all families, with a twofold degeneracy for each flavour due to the presence of left and right handed components. Gluinos have a relatively large cross-section, while neutralinos and charginos are disfavoured because their production is mediated by electroweak interaction.
The low cross section of SUSY particles implies that dedicated selections must be implemented to remove the majority of the events collected by the detector, identifying those with signal-like properties. A description of the baseline analysis strategy is presented in Section 5.3.

5.1.2 Benchmark models

The parametrisation of SUSY models can be challenging from a practical point of view, due to the large number of degrees of freedom in the full SUSY lagrangian. A common strategy [180] is to concentrate on specific signal processes extracted from the full models, ignoring any particle that does not contribute to them. For example, an analysis can be designed to target pair-produced squarks decaying into a SM quark and a neutralino, neglecting completely the rest of the SUSY mass spectrum. The signature-based benchmarks, referred to as simplified models, allow to minimise the number of free parameters that define the phenomenological properties of the signal: in the above example, the masses of squarks and neutralinos are sufficient. Their simplicity is also a benefit for the reinterpretation of the results of the analyses under different theoretical

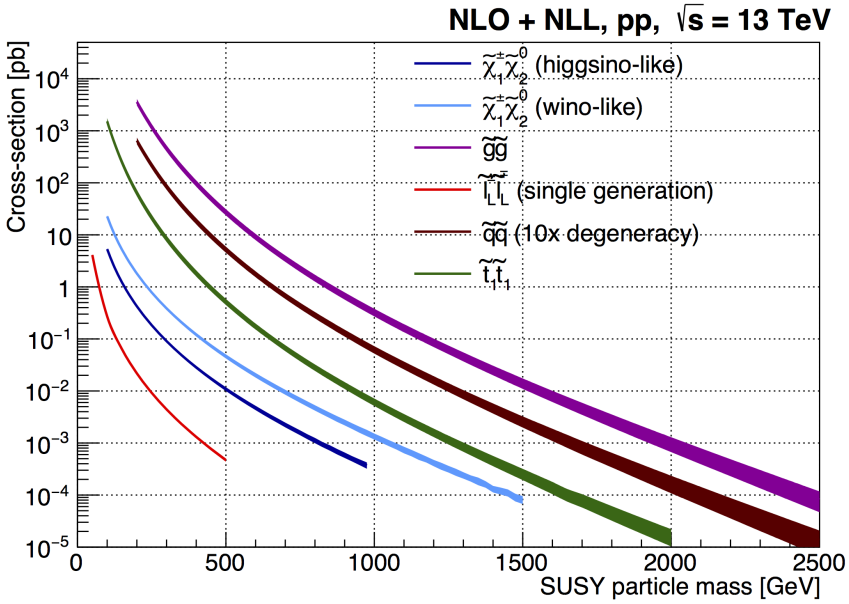


Figure 5.1: Cross sections for pair production of SUSY particles as a function of their mass in proton-proton collisions at $\sqrt{s} = 13$ TeV [179]. The cross sections are computed under the assumption that the different SUSY particles are decoupled. The brown line describes the production of $\tilde{q}\tilde{q}$ pairs for five mass-degenerate flavour families (excluding only the $\tilde{t}$) with two (left and right) degrees of freedom per family, yielding an overall ten-fold degeneracy.

assumptions, which is a powerful way to constrain new models.

When performing physics analyses with simplified models, it is important to keep track of the theoretical assumptions behind the specific features of the signals, to be able to interpret the results of the searches in broader contexts. For this purpose, after the completion of the analyses of the full Run 1 dataset at $\sqrt{s} = 8$ TeV, the ATLAS Collaboration has published a reinterpretation paper where the results of 22 searches are used to set limits on a 19-parameter phenomenological MSSM with a neutralino LSP, obtained from a multi-dimensional scan which takes into account previous electroweak and flavour measurements as well as dark matter constraints [182]. This result, which is presented in terms of limits on the masses of each SUSY particle, provides a useful comparison between the constraints on simplified models and those on realistic SUSY scenarios, identifying the signals that would have escaped the existing searches.

The analyses presented in this thesis are entirely based on simplified models of third generation squarks, with different decay modes inspired by the theoretical arguments

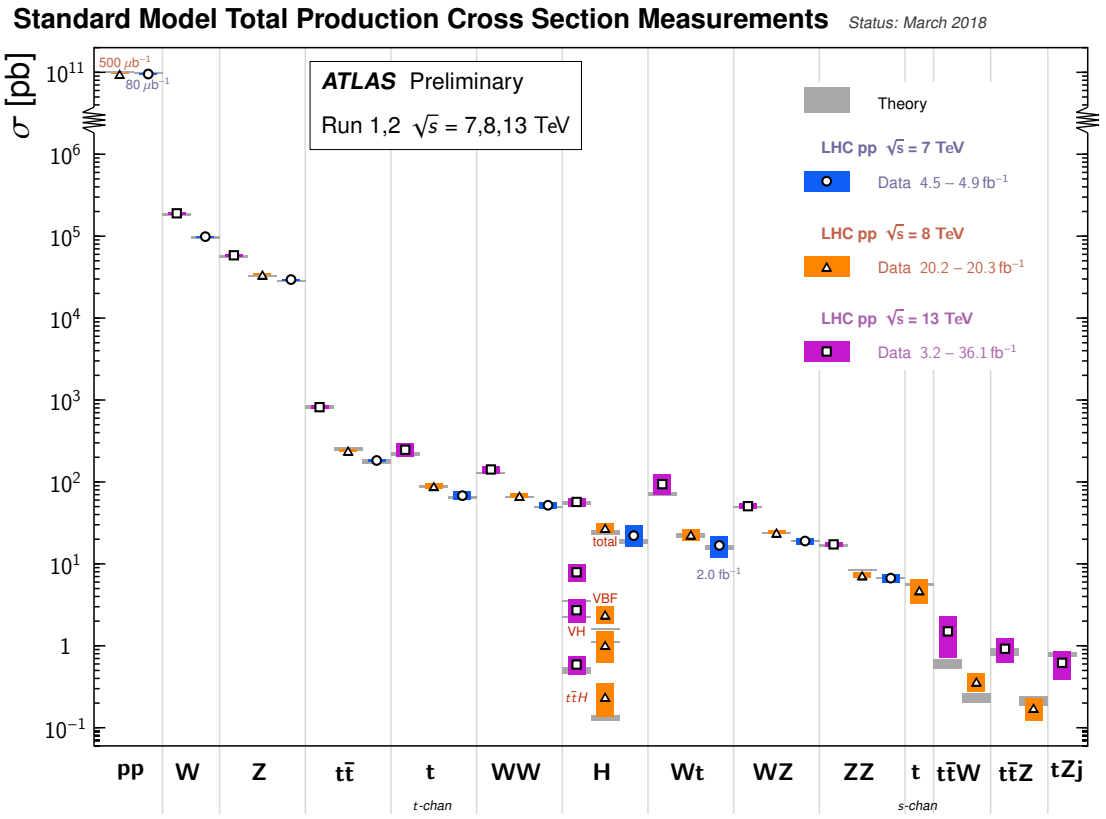


Figure 5.2: Summary of the cross sections of the most relevant SM processes measured by the ATLAS experiment at $\sqrt{s} = 7$, 8 and 13 TeV [181].

discussed in Chapter 1.

5.1.3 Phenomenology of third generation squarks

While the cross section of third generation squark pair-production with decoupled gluinos depends only on the mass of the sparticles, the experimental signature of the models varies significantly if different decay modes are considered. Lighter signals are generally easier to target thanks to the larger production rate, but in some situations the kinematical properties of the final states resemble those of the SM backgrounds, compromising the possibility to obtain an effective separation. For each SUSY model, it is hence important to optimise dedicated selections for different hypotheses on the masses and the decay modes.

In the case of third generation squarks, the experimental signature is generally similar for stop ($\tilde{t}$) and sbottom ($\tilde{b}$) production, so the analyses can often be interpreted in both

scenarios. The next pages provide an overview of the phenomenology studied within the ATLAS Collaboration, highlighting the differences between the two particles.

R-parity conserving decays

The ATLAS searches for $\tilde{t}$ or $\tilde{b}$ pair-production with RPC decays are usually done under the assumption of a neutralino ($\tilde{\chi}_1^0$) LSP [183]. In the simplest case the squarks decay directly into the LSP plus some extra SM particles, so the signal properties are basically defined by fixing the values of the masses of the two superpartners. More degrees of freedom appear if additional SUSY particles are included in the decay chain, such as heavier neutralinos or charginos.

The signal benchmark samples are typically generated as a grid of simplified models parametrised by the mass of the third generation squark and of one of the SUSY particles from the decay chain. If more than two SUSY particles are involved, dedicated assumptions are made on the masses of the extra ones in order to keep a total of two free parameters. For example, when working in the framework of the natural pMSSM spectrum shown in Figure 1.7, it is necessary to include an extra light chargino ($\tilde{\chi}_1^+$) and a next-to-lightest neutralino ($\tilde{\chi}_2^0$) in addition to the $\tilde{t}$ (or $\tilde{b}$) and the $\tilde{\chi}_1^0$. A two-dimensional signal grid can be obtained by making a specific assumption on the mass splitting between the extra particles and the LSP, leaving the values of $m_{\tilde{t}}$ and $m_{\tilde{\chi}_1^0}$ as the only free parameters.

The phenomenology of third generation SUSY signals is determined by the masses of the squarks and their potential decay products, as well as by other relevant parameters, such as the mixing between left and right components of the squarks (Equation 1.24) and the mixing between gaugino and Higgsino states (Equations 1.28 and 1.29) which result in different properties of the chargino-neutralino sector. Figure 5.3 illustrates four different classes of pMSSM models considered by a recent analysis targeting a pair-produced $\tilde{t}$ [184], that provide a good representation of the phenomenology studied within the ATLAS Collaboration:

(a) Pure bino ($\tilde{B}$) LSP model:
A first simple configuration is obtained by setting only M_1 to the electroweak scale in the neutralino mixing matrix in Equation 1.29, with the other parameters set to multi-TeV values. This yields a mass spectrum with a single light $\tilde{\chi}_1^0$, composed almost exclusively by the $\tilde{B}$ interaction eigenstate.

(b) Wino ($\tilde{W}$) Next to LSP (NLSP) model:

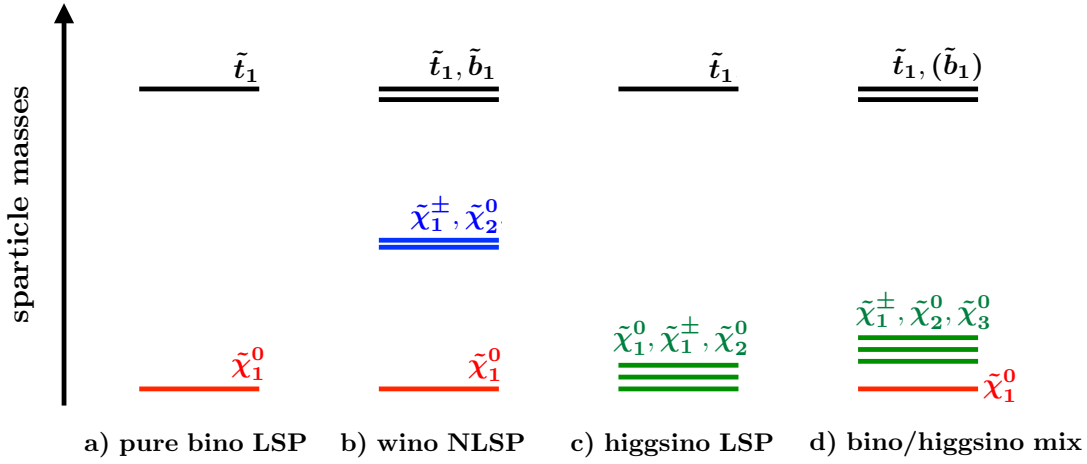


Figure 5.3: Schematic view of four possible mass spectra involving third generation squarks and light neutralinos and charginos [184]. The individual cases are discussed in the text.

A second model is designed to obtain two degenerate $\tilde{W}$-like mass eigenstates (a $\tilde{\chi}_2^0$ and a $\tilde{\chi}_1^+$) in addition to the $\tilde{B}$-like $\tilde{\chi}_1^0$. This is done in the pMSSM by setting $M_2 = 2M_1$ in Equation 1.29 while keeping the other parameters at higher scales, as suggested by models with gauge unification at the GUT scale [185–187].

(c) Higgsino ($\tilde{H}$) LSP model:
As discussed in Section 1.3.3, naturalness arguments predict a mass spectrum with three Higgsino-like light mass eigenstates at the electroweak scale (see Figure 1.7). This configuration is obtained by setting μ to a light scale, and a variable mass splitting between the three eigenstates is determined by the values of the M_1 and M_2 parameters, with complete degeneracy in the asymptotic limit ($M_1, M_2 \to \infty$). This is the reference model for the analysis in Chapter 6, where the pair production of the $\tilde{b}$ (not shown in Figure 5.3) is targeted.

(d) $\tilde{B}$-$\tilde{H}$ mix model:
Finally, a well-tempered $\tilde{\chi}_1^0$ model [188] is designed by imposing dark matter relic density constraints in the pMSSM together with the requirement of naturalness, resulting in a mass spectrum where the $\tilde{\chi}_1^0$ is a mixture of $\tilde{B}$ and $\tilde{H}$ and additional light states are present. In practice, this configuration can be obtained by setting $\mu \sim M_1$ in the pMSSM.

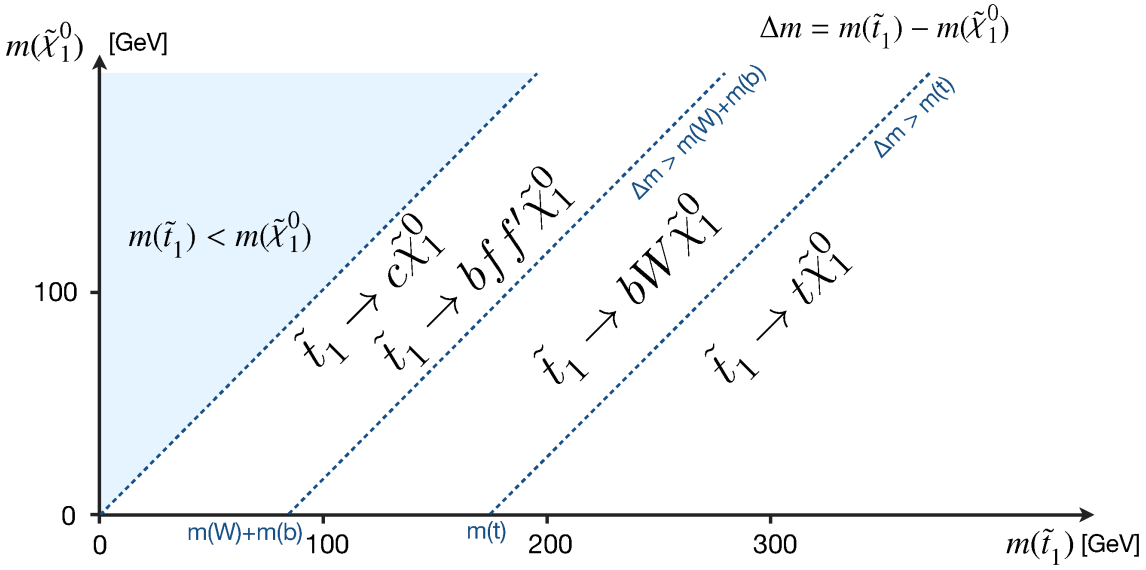


Figure 5.4: Decay modes of the $\tilde{t}$ in the $m_{\tilde{t}}$-$m_{\tilde{\chi}_1^0}$ plane [183].

The scheme in Figure 5.4 shows the possible decay modes of the $\tilde{t}$ in the $m_{\tilde{t}}$-$m_{\tilde{\chi}_1^0}$ plane when the phenomenology is limited to direct decays into a $\tilde{\chi}_1^0$, as in the pure $\tilde{B}$ LSP model described above. An important two-body decay of the $\tilde{t}$, whose diagram is shown in Figure 5.5a, is the following:

$$\tilde{t}_1 \to t\tilde{\chi}_1^0 \tag{5.1}$$

for which the kinematical constraint $m_{\tilde{t}_1} > m_t + m_{\tilde{\chi}_1^0}$ must be satisfied. Similar decays of the $\tilde{t}$ into heavier neutralinos can also occur in models where additional light states are present, yielding more objects in the events, but for simplicity only the $\tilde{\chi}_1^0$ is considered in the present discussion. When a two-body decay takes place, the boost of the quarks and LSPs in the final state depends on the mass splitting $\Delta m(\tilde{t}_1, \tilde{\chi}_1^0)$. If the splitting is large the objects are produced with high momentum, while compressed mass spectra yield softer decay products, that often require dedicated analysis techniques.
In signal models where an extra $\tilde{\chi}_1^+$ is included, the alternative two-body decay of the $\tilde{t}$ illustrated in Figure 5.5b is also possible:

$$\tilde{t}_1 \to b\tilde{\chi}_1^+ \tag{5.2}$$

with the kinematic requirement that $m_{\tilde{t}_1} > m_b + m_{\tilde{\chi}_1^+}$.
When designing the analysis, it is convenient to target the diagrams in Figures 5.5a and 5.5b under the assumption that the $\tilde{t}$ decays exclusively in the channels of interest, $t\tilde{\chi}_1^0$ and $b\tilde{\chi}_1^+$ with 100% BR respectively. While doing so, it is important to recall that both decays may occur with comparable BRs, yielding asymmetric final states as the one shown

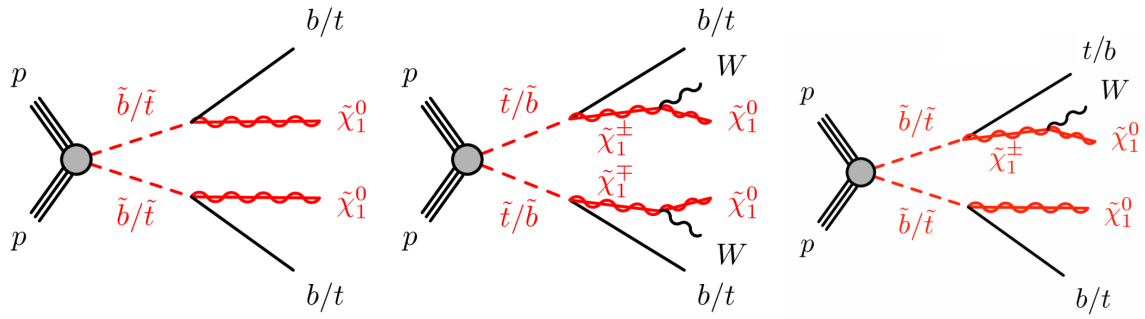


Figure 5.5: Pair-produced $\tilde{t}$ decaying directly into a $\tilde{\chi}_1^0$ pair (a), a $b\tilde{\chi}_1^+$ pair (b) or a mixture of the two final states (c). The grey circles in the production vertices indicate generic effective couplings, which represent un-specified production modes of the particles. The figures also show how each diagram changes when replacing the $\tilde{t}$ with the $\tilde{b}$.

in Figure 5.5c, that are interesting to examine in addition to the symmetric ones.
Moreover, as indicated in Figure 5.5, the diagrams associated with the above processes can be used to describe both $\tilde{t}$ or $\tilde{b}$ pair-production by exchanging accordingly top and bottom quarks in the final states. The replacement has a significant impact on the phenomenological properties of the signals, because top quarks have a more complex experimental signature, so the analyses need to be adapted to the chosen case. A comprehensive search where $\tilde{b}$ pair-production is assumed is presented in Chapter 6.
Returning to the discussion of Figure 5.4, another possible two-body final state [189] can be obtained through the following decay:

$$\tilde{t}_1 \rightarrow c\tilde{\chi}_1^0 \tag{5.3}$$

where no third generation quarks are produced. Since this process implies flavour violation, it is commonly expected to be relevant only when the competitive channels are forbidden by kinematic constraints, as in the compressed mass region with small $\Delta m(\tilde{t}, \tilde{\chi}_1^0)$.
When two-body decays are forbidden, the off-shell decays of the $\tilde{t}$ via virtual W bosons or top quarks may become relevant, yielding three-body or four-body final states. The decay via a virtual top quark (Figure 5.6a) is:

$$\tilde{t}_1 \rightarrow bW\tilde{\chi}_1^0 \tag{5.4}$$

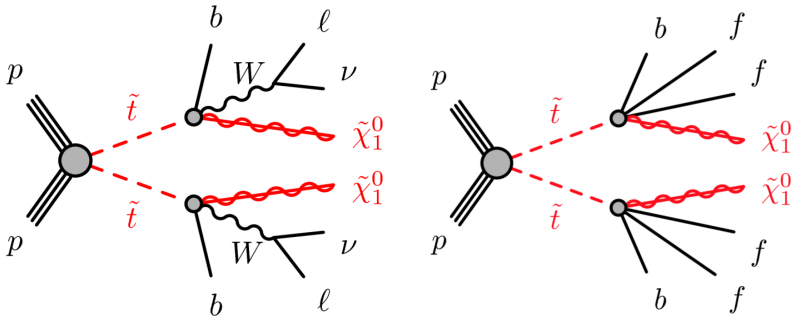


Figure 5.6: Three-body (a) and four-body (b) decays of the $\tilde{t}$. For the three-body case, the leptonic decay of W bosons is shown.

while the four-body decay where also the W is virtual gives:

$$\tilde{t}_1 \to bff'\tilde{\chi}_1^0 \tag{5.5}$$

As discussed in detail in Ref. [189], the four-body decay and the two-body decay into charm quarks (Equation 5.3) are predicted to be in competition in the same region of the $m_{\tilde{t}}$-$m_{\tilde{\chi}_1^0}$ plane, with a relative BR that depends on the values of the relevant parameters of the model.

R-parity violating decays

When the $\tilde{t}$ decays via RPV couplings the assumption of a $\tilde{\chi}_1^0$ LSP can be dropped, and final states containing only SM particles become possible. As discussed in Section 1.3.4, models with spontaneously broken $U(1)_{B-L}$ symmetry predict the two-body decay of the $\tilde{t}$ into a bottom quark and a lepton (see Equation 1.36) where the preferred lepton flavour is related to the neutrino mass hierarchy. This signal, shown in Figure 5.7a, yields two b-ℓ pairs with invariant mass peaking at $m_{\tilde{t}_1}$ and is targeted by the analysis presented in Chapter 7.
Finally, an alternative RPV coupling predicts the prompt decay of the $\tilde{t}$ into a pair of quarks:

$$\tilde{t}_1 \to qq' \tag{5.6}$$

The corresponding diagram is shown in Figure 5.7b, and the analysis is particularly challenging due to the overwhelming multi-jet background at the LHC.

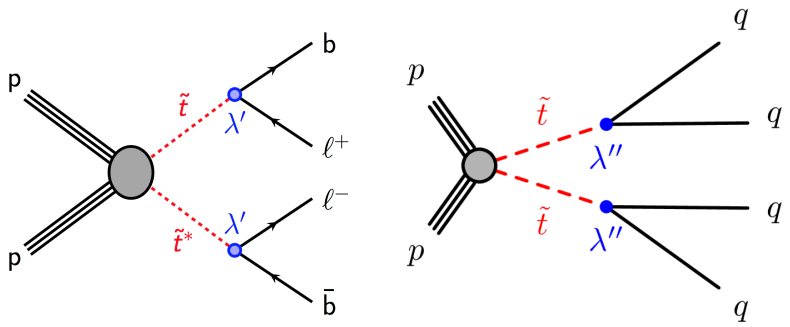


Figure 5.7: RPV decays of the $\tilde{t}$ into a b-ℓ pair (a) and two quarks (b).

Present status of ATLAS searches for third generation squarks

The signal models described in the last paragraphs have been targeted by several ATLAS analyses, which used data from all the past runs of the LHC. The most recent results, including those presented in this thesis, are based on the 2015+2016 dataset, which is discussed in detail in Section 5.3.1. In the absence of discoveries, strong exclusion limits have been set on a variety of SUSY models, including third generation squarks, using the statistical tools described in Section 5.4.

Figure 5.8 shows the present status of the exclusion limits on $\tilde{t}$ pair-production in the $m_{\tilde{t}}$-$m_{\tilde{\chi}_1^0}$ plane, where the $\tilde{\chi}_1^0$ is assumed to be the only SUSY particle produced by the decay of the squark, and the relevant final states are those described in Figure 5.4. The plot is available in the public webpage of the ATLAS Collaboration [190], and is progressively updated to include the results of the new searches. The limits are particularly stringent for large $\Delta m(\tilde{t}, \tilde{\chi}_1^0)$, where the $\tilde{t}$ is excluded up to approximately 1 TeV, but lighter squarks are still permitted in the compressed region of the mass plane. The fact that the limits are weaker in the compressed mass region is a recurrent feature in various SUSY searches, because signals with small mass splittings generate decay products with low transverse momenta, that are more difficult to separate from the SM backgrounds.

In addition to the limits in Figure 5.8, the public webpage [190] presents the results of the analyses that target different types of modes, including those with more neutralinos or charginos in the decay chains, or those where RPV decays of the squarks are considered. A detailed discussion of the constraints on the signal models studied in this thesis is provided in the individual analysis chapters (Chapter 6 and 7).

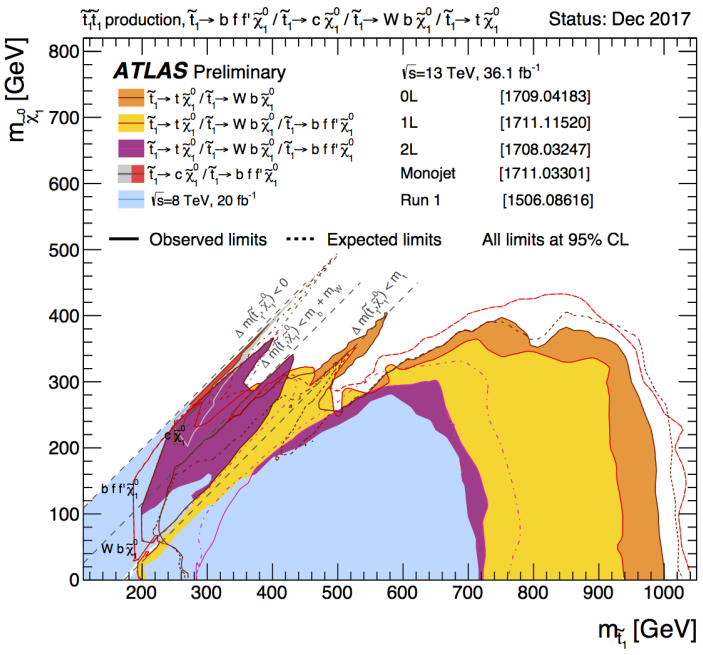


Figure 5.8: Summary of the exclusion limits of the ATLAS searches for $\tilde{t}$ pair production in the $m_{\tilde{t}}$-$m_{\tilde{\chi}_1^0}$ plane, based on the analysis of 36.1 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV [190]. The decay channels described in Figure 5.4 are considered.

5.1.4 Other SUSY searches

The physics programme of the ATLAS Collaboration is organised to cover a wide variety of SUSY signal models, in addition to the third generation signals discussed in the previous paragraphs. The most relevant classes of searches are briefly described below:

Gluino production The direct production of gluinos ($\tilde{g}$) benefits from relatively large cross sections compared to other SUSY processes, as already shown in Figure 5.1. The summary of the present constraints on $\tilde{g}$ pair production is shown in Figure 5.9 in the $m_{\tilde{g}}$-$m_{\tilde{\chi}_1^0}$ plane for different benchmark models, where the $\tilde{g}$ decays into a $\tilde{\chi}_1^0$ either directly or through intermediate states, including third generation squarks. The limits on $m_{\tilde{g}}$ are already extending beyond 2 TeV for specific signal models.

Electroweak production Dedicated searches are performed to target the direct production of neutralinos, charginos and sleptons, which is regulated by the electroweak interaction. The cross sections are generally smaller compared to processes mediated by the strong force, especially when the production of sleptons is considered. Significant differences are also present depending on the flavour com-

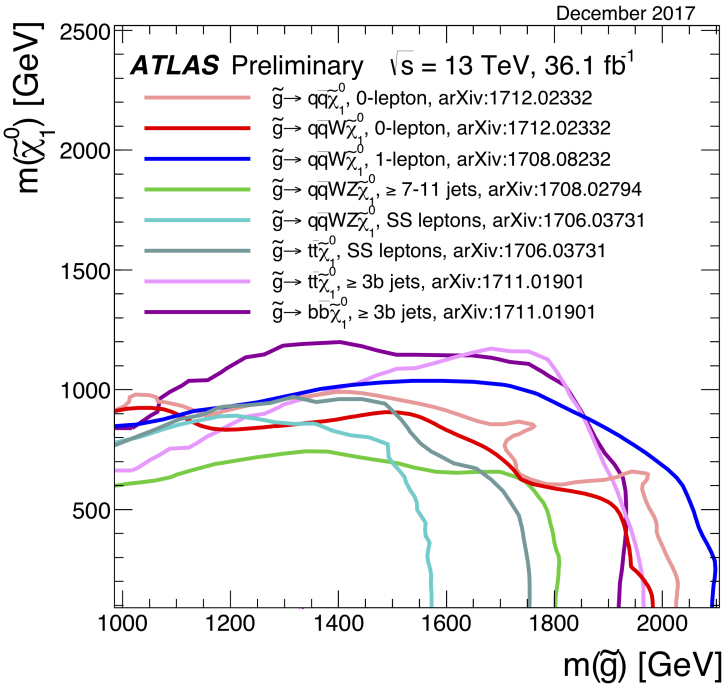


Figure 5.9: Summary of the exclusion limits on $\tilde{g}$ pair production in the $m_{\tilde{g}}$-$m_{\tilde{\chi}_1^0}$ plane for various simplified models where the $\tilde{g}$ decays into the $\tilde{\chi}_1^0$ either directly or through a cascade chain that includes other SUSY particles with intermediate mass [190].

position of neutralinos and charginos, with $\tilde{H}$-like states more suppressed compared to $\tilde{W}$ or $\tilde{B}$-like ones [179]. Figure 5.10 shows the present exclusion limits on $\tilde{\chi}_1^{\pm}$ pair production or associated production of $\tilde{\chi}_2^0$ and $\tilde{\chi}_1^{\pm}$ under the assumption of $\tilde{W}$-like cross sections, with decays mediated by SM vector bosons. The constraints are strongly model dependent, and a general overview of the results is available in the ATLAS public webpage [190].

RPV and long lived particles The characteristic signature of RPV SUSY models consists in the direct decay of gluinos, third generation squarks or electroweakly produced particles into SM objects, resulting in events with no real E_T^{miss} (unless neutrinos are produced in the final states). An additional class of models is obtained by dropping the assumption that SUSY particles (with the exception of the LSPs in RPC models) are promptly decaying, which is used in all the analyses presented in the previous paragraphs. The production of long lived states results in characteristic experimental signatures in the detector, that require dedicated analysis techniques. The present status of the ATLAS searches for RPV and long lived SUSY

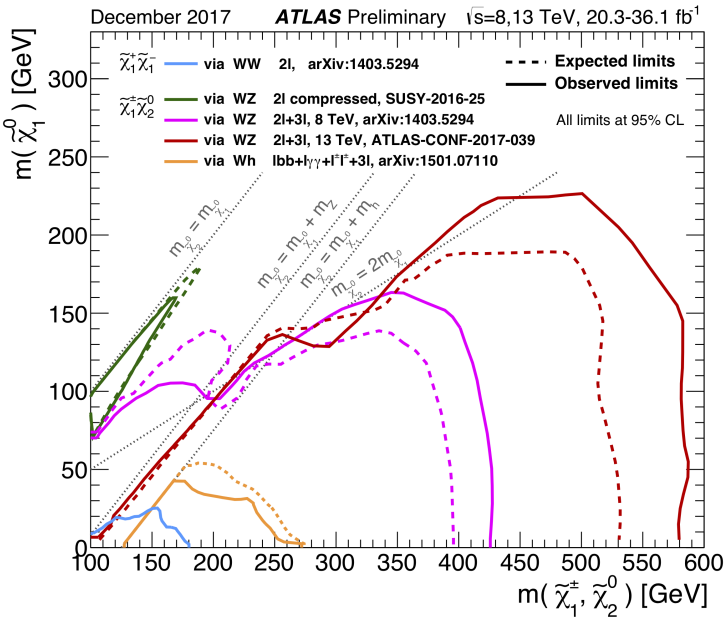


Figure 5.10: Summary of the limits on $\tilde{\chi}_1^+\tilde{\chi}_1^-$ and $\tilde{\chi}_1^\pm\tilde{\chi}_2^0$ production with decays mediated by SM bosons, in the $m_{(\tilde{\chi}_1^\pm,\tilde{\chi}_2^0)}$-$m_{\tilde{\chi}_1^0}$ plane [190].

particles is summarised in the public webpage [190], where a variety of relevant results are collected.

5.2 Backgrounds and Monte Carlo samples

As discussed in Section 5.1, the experimental signature of third generation squarks is similar to a variety of SM processes with significantly larger cross sections. The accurate modelling of these backgrounds is hence essential to implement robust physics analyses that are able to target the desired signals. In particular, the definition of Signal Regions (see Section 5.3.5) requires an accurate knowledge of the kinematical properties of both signals and backgrounds, and the final sensitivity of the searches depends on the ability to estimate the leftover backgrounds with high accuracy. This section is dedicated to a discussion of the properties of the most relevant SM backgrounds in ATLAS, together with a description of the MC samples that are used to model them in the analyses.

5.2.1 Background processes

The main sources of SM background, whose relative importance depends on the specific selections of each analysis, are described in detail the list below, which includes information about their properties obtained from the Particle Data Group [191]:

Z + jets The associated production of Z bosons and jets is a relevant source of background in 0-lepton plus E_T^{miss} or 2-lepton final states, due to the $Z \to \nu\nu$ decay (with BR $\sim 20\%$) and the $Z \to \ell\ell$ decay ($\ell = e, \mu, \tau$, with total BR $\sim 10\%$). On the contrary, the $Z \to qq$ decay, which occurs with the largest BR ($\sim 70\%$), generates a pair of quarks with relatively low boost, yielding a fully hadronic final state which is covered by the dominant multi-jet background and hence is not relevant for third generation SUSY searches. The simulated Z events are classified at truth level based on the flavour of the hadrons produced in association with the vector boson: the mutually exclusive b-filtered and c-filtered sub-samples are defined as containing at least one b-hadron or c-hadron (and no b-hadrons), while the remaining events with no heavy flavour hadrons are included in the light-filtered sub-sample. Since both analyses in this thesis select events with b-tagged jets, the Z+jets background is generally dominated by the b-filtered category, whose cross section predicted by the current MC simulations was found to be under-estimated in previous analyses [192]. A diagram corresponding to the production of a Z boson and a bottom quark pair is shown in Figure 5.11a.

W + jets Similarly, the production of W bosons and jets yields a relevant background in many ATLAS analyses, especially when final states with 1 lepton are considered, due to the $W \to \ell\nu$ decay ($\ell = e, \mu, \tau$) which happens with a total BR of $\sim 32\%$. The dominant $W \to qq'$ decay yields a multi-jet final state, which is again irrelevant for the analyses in this thesis. Similarly to Z+jets, W+jets events are also divided in the b, c and light-filtered categories, and the relative impact of the former is enhanced when b-tagged jets are required. Figure 5.11b shows a diagram corresponding to the associated production of a W boson and a bottom quark pair, closely resembling the Z boson one.

Top pairs A major background process at the LHC is the pair production of top quarks, already introduced for the b-tagging performance studies in Section 4.1.2. The dominant $t \to Wb$ decay, with $\sim 99.8\%$ BR, yields two b-jets and two oppositely charged W bosons that further decay as described above, yielding 0-lepton, 1-

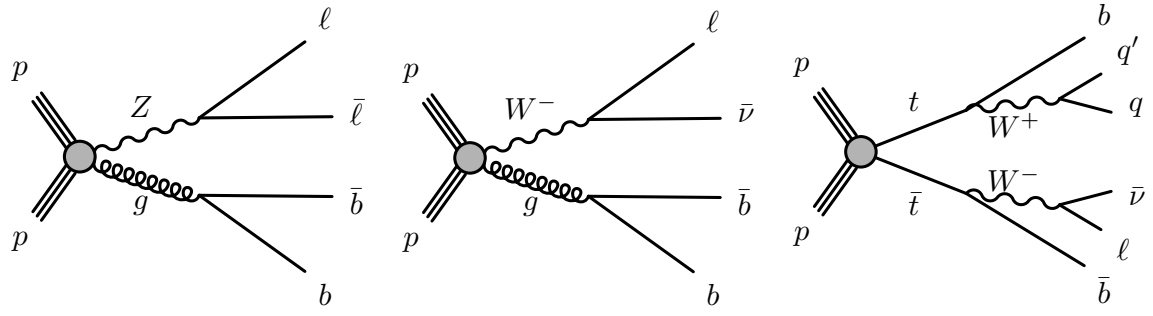


Figure 5.11: Diagrams for the associated production of a pair of bottom quarks and a Z (a) or W (b) boson, and for the pair production of top quarks (c) at the LHC.

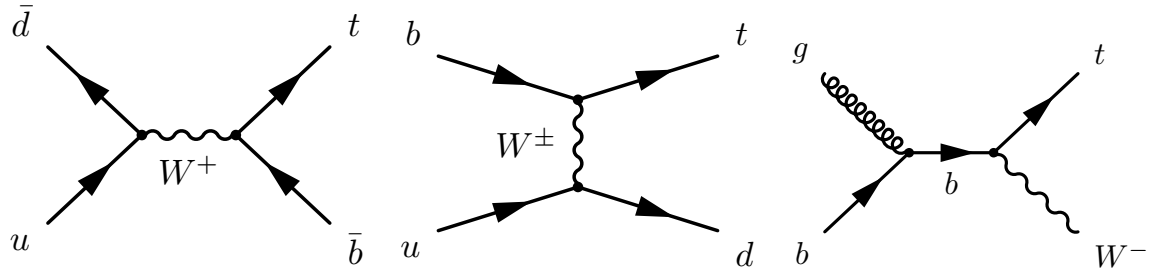


Figure 5.12: Diagrams for the production of a single top quark at the LHC, in the s-channel (a), t-channel (b) and Wt channel (c).

lepton and 2-lepton final states with 45.7%, 43.8% and 10.5% BRs respectively. The diagram corresponding to the semi-leptonic decay scenario is shown in Figure 5.11c.

Single top While top quarks are mainly produced in pairs, the single production is also possible at the LHC: the corresponding LO diagrams are shown in Figure 5.12 and are referred to as s-channel, t-channel and Wt channel. The latter case is generally the most relevant for third generation SUSY analyses, due to the presence of an extra W boson in the final state.

Top pair plus X The pair production of top quarks can also happen in association with an extra boson in the final state, that can be a vector boson (γ, W or Z) or a Higgs boson. Despite the lower cross section with respect to the processes considered so

far, $t\bar{t}$+X production can be a relevant background for third generation SUSY models, especially in the case of $t\bar{t}W$ and $t\bar{t}Z$ for which the cross section is relatively larger.

Di-bosons Another set of processes with low cross section but signal-like properties is the di-boson production, where pairs of vector bosons can appear in different combinations and the most relevant cases are WW, WZ or ZZ. Final states with variable lepton and jet multiplicities are obtained when different decays of the bosons are considered.

Multi-jet The majority of proton collisions at the LHC yield fully hadronic final states, which result in multi-jet events with no prompt leptons or invisible objects. These events may assume signal-like properties as a result of the mis-reconstruction of isolated leptons, or due to large measured values of the E_T^{miss} from detector resolution effects. While the probability of observing these circumstances in individual multi-jet events is low, the large overall rate of the process implies that a non-negligible net contribution may be found in the relevant channels of the analysis.

5.2.2 Monte Carlo samples

Signal and background processes are simulated through the procedures described in Section 3.1, and MC samples are employed for all of them with the exception of multi-jet production, which is estimated in the analyses using dedicated data-driven techniques. The details of the MC generators and the associated settings are included in Table 5.1, and a brief summary is provided below for the relevant physics processes:

Signal samples The benchmark SUSY models are generated using MadGraph [125] at LO and interfaced to Pythia [123] for the modelling of PS, hadronisation and underlying event.

***V*+jets** The associated production of W or Z bosons and jets is simulated with Sherpa [127], including jets from the hadronisation of b and c quarks. The MEs are calculated for up to two additional partons at NLO and four partons at LO, and the matching is performed using the Sherpa PS generator [128]. The simulations are split in b, c and light-filtered samples, and a larger amount of MC statistics is generated for the b-filtered ones, that are particularly relevant in the regions of phase space selected by the analyses in this thesis. A further splitting is implemented by

selecting sub-samples of events with different ranges of a kinematical variable defined as the maximum between H_{T}[1] and the p_{T} of the boson, and the sub-samples containing more boosted particles are simulated with higher precision.

Top production The $t\bar{t}$ and single top processes are simulated at NLO with the Powheg event generator [126], using Pythia [122] for the PS. The events are retained only if they contain at least one leptonically decaying W boson, while the fully hadronic decays are removed because they are a negligible source of background in the SUSY analyses. In addition, a dedicated slicing is implemented using E_T^{miss} and H_{T} as filtering variables, in order to populate the tails of the associated distributions with sufficient MC statistics.

Other backgrounds The production of top quark pairs plus electroweak vector bosons (W, Z) or Higgs bosons is simulated at NLO using MadGraph and Pythia. Diboson processes are instead simulated with Sherpa, similarly to the V+jets samples.

For all samples described above, with the exception of those generated with Sherpa, the EvtGen program [168] is employed to simulate the decays of b and c hadrons, while pileup is simulated with Pythia by overlaying additional pp collisions to the hard scattering events. The interaction of the particles in the final state with the detector is simulated with Geant4 as described in Section 3.1.3, using a full simulation for the background processes and an AF2 simulation for the signal samples.
Finally, additional samples are generated to estimate the theoretical uncertainties associated with the configuration of the MC simulations. These include variations of the renormalisation, factorisation or CKKW matching scales, different PDF sets or hadronisation models, as described in Section 5.3.4.

5.3 Analysis strategy

The analyses in this thesis are based on a cut and count strategy, where events are selected according to the experimental signature of the targeted signals and a dedicated set of discriminating variables is employed to define signal-enriched regions with reduced levels of SM background. The relevant observables of the analyses are the number of events passing each selection, and their values are predicted for both signal and background processes using MC samples, data-driven techniques or combinations of the two.

[1]The H_{T} variable is defined as the scalar sum of the p_{T} of all jets in the event.

Process	ME	PS	UE tune	PDF	Cross section
Signal	MadGraph 2.2.3	Pythia v8.186	A14	NNPDF2.3	NLO+NLL
$t\bar{t}$	Powheg v2	Pythia v6.428	P2012	CT10	NNLO+NNLL
Single top (s, Wt)	Powheg v2	Pythia v6.428	P2012	CT10	NNLO+NNLL
Single top (t)	Powheg v1	Pythia v6.428	P2012	CT10	NNLO+NNLL
W/Z+jets	Sherpa 2.2.1	Sherpa 2.2.1	Default	NNPDF3.0	NNLO
Diboson	Sherpa 2.2.1	Sherpa 2.2.1	Default	NNPDF3.0	NLO
$t\bar{t}+W/Z$	MG5_aMC@NLO 2.2.3	Pythia v8.212	A14	NNPDF2.3	NLO

Table 5.1: Details of the MC simulations used for each physics process.

This section introduces the general strategy of the analyses in Chapters 6 and 7, together with a description of the common technical aspects. These include the data sample, the trigger and event cleaning selections, the definition of physics objects and experimental quantities in the events with the related performance corrections in MC, and the dominant sources of systematic uncertainties. In addition, the standard strategies for the definition of Signal Regions (SRs) and the estimation of the leftover backgrounds are outlined.

5.3.1 Data sample and event selection

The analyses presented in this thesis are based on proton-proton collision data at $\sqrt{s} = 13$ TeV collected by the ATLAS detector during 2015 and 2016. The dataset corresponds to an integrated luminosity of 36.1 fb^{-1}, computed after requiring a standard set of data quality criteria as introduced in Section 2.3.6. The measured value of the integrated luminosity carries an uncertainty of 3.2%, derived through a procedure similar to the one described in Ref. [193] using x-y beam separation scans performed in August 2015 and May 2016. The average pileup parameter μ is 13.7 in 2015 and 24.9 in 2016, where a higher instantaneous luminosity was reached (see Section 2.1.4).
Due to the different operating conditions of the experiment in 2015 and 2016, a variety of selections are applied differently in the various periods of the data taking, including trigger requirements and calibration parameters. Since the same MC samples are used for all periods, the simulated events are assigned a random number that identifies a given ATLAS run, allowing to associate them with specific periods of operation so that their parameters can be modified consistently with what is done in data.

Triggers

As discussed in Sections 2.3.6 and 3.2.1, events are recorded by the ATLAS detector if they satisfy the requirements of pre-defined trigger selections based on the presence of specific objects or features in the events. The analyses in Chapters 6 and 7 employ E_T^{miss} triggers in 0-lepton events, single electron or muon triggers in channels with one or more leptons, and single photon triggers when the presence of a photon is required. The detail of the trigger chains that are used is given below:

E_T^{miss} triggers The online requirement on the E_T^{miss} is applied using different energy thresholds depending on the period of data taking. In 2015 data, events are selected by reconstructing the E_T^{miss} from an input jet collection and requiring that its measured value exceeds 70 GeV. During the 2016 operation, a similar algorithm is employed but the energy threshold is progressively increased to 90, 100 and 110 GeV due to the higher instantaneous luminosity, that has a significant impact on the trigger rate. As already seen in Figure 3.4b, these triggers reach their efficiency plateau above offline E_T^{miss} values of approximately $150 - 200$ GeV.

Single electron triggers In both 2015 and 2016 data a logic OR of three different chains is used to select electron events at trigger level. The first chain includes an isolation requirement, which allows to keep a low energy threshold of 24 GeV in 2015 and 26 GeV in 2016, where tighter identification and isolation selections are also applied. In the second chain the isolation requirement is dropped to reduce efficiency losses at high E_{T}, and the energy threshold is raised to 60 GeV (constant for the full dataset). Finally, the third chain has looser identification requirements and a higher energy threshold of 120 GeV in 2015 and 140 GeV in 2016. The efficiency plateau is reached by selecting offline electrons with $p_{\text{T}} > 27$ GeV.

Single muon triggers Similarly, events with muons are selected with an OR of two chains, where the first has a lower energy threshold and an isolation requirement. The first chain has an energy threshold of 20 GeV in 2015 and 26 GeV in 2016, where a tighter isolation selection is also included, while the second chain has a threshold of 50 GeV in the whole dataset. The trigger is fully efficient for offline muon selections with $p_{\text{T}} > 27$ GeV.

Single photon triggers Events with photons are selected using a dedicated chain with energy threshold of 120 GeV in 2015 and 140 GeV 2016. An offline selection of photons with $p_{\text{T}} > 145$ GeV is applied to ensure that the trigger is fully efficient.

Event cleaning

A set of event-level cleaning selections are applied in both data and MC samples to remove events where pathological behaviour is observed. A first essential requirement is the presence of at least one PV in the events, with a minimum of two associated tracks with $p_{\mathrm{T}} > 400$ MeV. In addition, data events in which the electromagnetic or hadronic calorimeters are flagged to be in error state are removed from the analysis, and a similar veto is applied in case of single event upsets in the SCT.
Further cleaning selections are applied if poorly reconstructed jets or muons are identified through standard ATLAS criteria. Events with bad jets (as defined in Section 3.2.5) passing the Overlap Removal (OR) selections (see Section 5.3.2) are removed from the analysis. In addition, a similar veto is applied if baseline muon candidates with relative uncertainty on q/p larger than 20% are found before the OR, or if a muon with cosmic-like properties survives the OR.

5.3.2 Object definition

In each event, the definition of physics objects for offline analysis is performed by employing the reconstruction algorithms described in Section 3.2, that are standard tools within the ATLAS Collaboration. While minor differences are present between different searches, including those in Chapters 6 and 7, it is convenient to present the object definition in the present paragraph, highlighting the specificities of each analysis when necessary.
A common approach followed by many SUSY searches is to define two categories of reconstructed objects in the events, referred to as *baseline* and *signal* objects. The former type is identified by a minimal set of conditions applied to the outputs of the reconstruction algorithms, to obtain a complete sample of candidates for each physics object, while the latter is selected by applying tighter requirements on the baseline object collections in order to reduce the background contamination.
As illustrated by the scheme in Figure 5.13, the baseline object collections are used as inputs of an OR procedure, which is introduced to solve potential ambiguities caused by the presence of close-by objects in the detector, or single physics objects reconstructed in different ways. This may occur, for example, if a lepton is produced within the cone of a b-jet, or if an isolated electron is identified also by the jet reconstruction algorithm. Signal objects are hence defined as the subset of baseline objects which pass the OR selection and satisfy an additional set of analysis-dependent criteria.

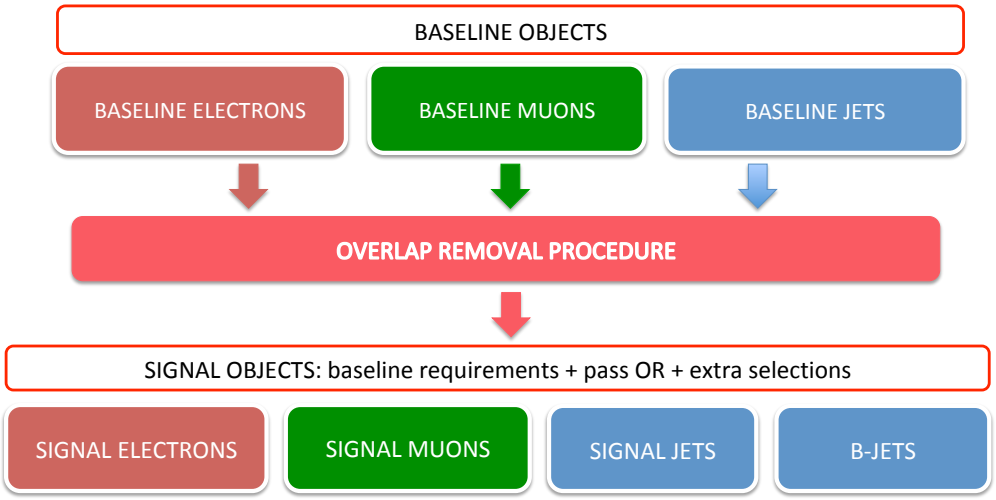


Figure 5.13: Scheme of the strategy used to define the relevant sets of physics objects in each event, in the simple case where only electrons, muons, jets and b-jets are reconstructed. The baseline collections are formed using the object reconstruction algorithms introduced in Section 3.2 and a loose set of selections, while the associated signal collections are obtained by applying tighter requirements on the baseline objects, including the OR selection described in Section 5.3.2. In the analysis channels requiring the reconstruction of photons, the corresponding baseline and signal collections are also defined, and baseline photons are included in the OR procedure.

The OR proceeds through the following sequence:

1. The overlap between electrons and jets within $\Delta R < 0.2$ is solved by removing the jets, unless they are b-tagged, in which case the electrons are removed because they are likely to originate from a b-hadron decay.

2. Leptons within $\Delta R < 0.4$ from a jet are removed, except if the jets have less than 3 tracks and are overlapping with muons, in which case they are removed instead.

3. If photons are included in the OR procedure, they are removed when they are within $\Delta R < 0.4$ from electrons or muons, while jets within $\Delta R < 0.4$ from a photon are removed.

In the analyses in Chapters 6 and 7 the OR is performed in the simple version with no photons, with the exception of the data driven estimate of Z+jets events discussed in Section 6.3.2, where photon events are used and the OR is modified accordingly.

The detailed definition of baseline and signal objects used for the two analyses is provided below:

Electrons The baseline electron collection is defined by requiring $p_{\mathrm{T}} > 10$ GeV, $|\eta| < 2.47$ and loose identification [135,140,141]. Signal electrons are further required to pass the OR and satisfy additional impact parameter cuts ($|d_0|/\sigma_{d_0} < 5$, $|z_0| \sin\theta < 0.5$), tight identification and loose isolation. The p_{T} threshold for signal electrons depends on the analysis: it is 20 GeV[2] for the search in Chapter 6 and 40 GeV for the one in Chapter 7.

Muons Baseline muons are reconstructed using the combined reconstruction chain (see Section 3.2.4), and are required to pass a medium track quality requirements [143] and to satisfy $p_{\mathrm{T}} >$10 GeV and $|\eta| < 2.5$. Signal muons are then selected after the OR with $p_{\mathrm{T}} > 20$ GeV[3], after applying impact parameter cuts ($|d_0|/\sigma_{d_0} < 3$, $|z_0| \sin\theta < 0.5$) and loose isolation selections. Finally, cosmic muons are defined through impact parameter selections ($|z_0| > 1$ mm or $|d_0| > 0.2$ mm) and are used for event cleaning purposes.

Photons Reconstructed photons are not used in the main channels of the two analyses in this thesis, but they are employed for an alternative estimate of the $Z \to \nu\nu$ background in the $\tilde{b}$ analysis (Chapter 6). Baseline photons are selected with $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.37$ (whilst being outside the $1.37 < |\eta| < 1.52$ range), and must satisfy the tight photon shower shape and electron rejection requirements [137]. Signal photons are further required to have $p_{\mathrm{T}} > 145$ GeV and to be isolated.

Jets Hadronic jets are reconstructed using the anti-k_{T} algorithm seeded by topological clusters, with cone size parameter $R = 0.4$. Baseline jets are selected with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 4.8$, while signal jets are required to pass the OR, have $|\eta| < 2.8$ and satisfy JVT requirements to suppress pileup (see Section 3.2.5). The signal jet p_{T} threshold is 20 or 35 GeV in the analysis in Chapter 6 (depending on the channel) and 60 GeV in the analysis in Chapter 7.

***b*-jets** The identification of b-jets is based on the MV2c10 algorithm and uses the 77% fixed cut WP as described in Chapter 4. The p_{T} threshold of signal jets is applied

[2] In events selected with a single electron trigger, the offline electron with highest p_{T} is required to have $p_{\mathrm{T}} > 27$ GeV to reach the efficiency plateau.

[3] Similarly to electrons, also signal muons are required to satisfy $p_{\mathrm{T}} > 27$ GeV in events triggered by single muon chains.

also to b-jets, while the pseudorapidity range is reduced to $|\eta| < 2.5$.

Missing transverse momentum The E_T^{miss} is reconstructed through the procedure described in Section 3.2.7, using as inputs the baseline objects described above, including photons. Hadronically decaying τ leptons are not explicitly reconstructed, so they enter the computation with the standard jet calibration.

5.3.3 Monte Carlo corrections

A dedicated set of MC weights are applied to the simulations in order to match the distributions of specific performance parameters in data. These include online and offline SFs, already introduced in Chapter 3, and ad-hoc reweighing procedures used to ensure that given quantities (such as the pileup parameter μ) are equally distributed in data and MC samples. The following list describes the MC weights used for the analyses in this thesis:

***b*-tagging SFs** In order to calibrate the performance of the MV2c10 algorithm, as discussed in Chapter 4, efficiency scale factors are defined separately for real b, c and light jets:

$$\text{SF}_i = \frac{\epsilon_i^{\text{data}}}{\epsilon_i^{\text{MC}}}, \quad i = b, c, l \tag{5.7}$$

and the corresponding inefficiency scale factors for non b-tagged jets are also defined:

$$\text{SF}_i = \frac{1 - \epsilon_i^{\text{data}}}{1 - \epsilon_i^{\text{MC}}}, \quad i = b, c, l. \tag{5.8}$$

The above SFs are jet-based quantities, and their values depend on the kinematics (p_{T} and η) of the jets. The total event-level weights are obtained as the product of the individual SFs for all jets to which the b-tagging decision is applied.

JVT SFs Additional SFs are used to calibrate the performance of JVT in signal jets, based on the same definition of the b-tagging SFs described above.

Lepton offline SFs Efficiency SFs are also obtained for offline electrons and muons in the events, as anticipated in Sections 3.2.3 and 3.2.4. Electron SFs are computed separately for the reconstruction, identification and isolation steps, while for muons the reconstruction and isolation SFs are employed.

Lepton trigger SFs The performance of the lepton triggers is calibrated using dedicated SFs, that are obtained for both single electron and single muon chains. In

events where at least one signal lepton is present, the total trigger efficiency is computed as the probability that at least one of them has fired the corresponding trigger:

$$\epsilon_{\text{trig}} = 1 - \prod_i (1 - \epsilon_i) \tag{5.9}$$

where ϵ_i is the trigger efficiency corresponding to the lepton i. The trigger efficiency SF for each event is hence obtained as:

$$\text{SF}_{\text{trig}} = \frac{\epsilon_{\text{trig}}^{\text{data}}}{\epsilon_{\text{trig}}^{\text{MC}}} = \frac{1 - \prod_i (1 - \epsilon_i^{\text{data}})}{1 - \prod_i (1 - \epsilon_i^{\text{MC}})} \tag{5.10}$$

where only trigger-matched leptons are considered.

Pileup weight Finally, an event-level weight is employed to correct the distribution of the pileup parameter μ in the MC samples, matching it to the one observed in the 2015+2016 dataset. The impact of the reweighing is expected to be negligible if the selections of the analysis are not sensitive to pileup.

5.3.4 Systematic uncertainties

In addition to the object definition described in the previous paragraph, it is useful to present an overview of the sources of systematic uncertainty that are relevant for both analyses in this thesis, as well as for several other searches in ATLAS. The list of experimental uncertainties is only partially overlapping with the one presented in Section 4.1.3 for the b-tagging performance studies, where tracking and jet related effects are included. In particular, the tracking uncertainties are not evaluated directly in the physics analyses, but their effects are implicitly taken into account by applying higher-level systematics associated with reconstructed objects or quantities in the events. In addition to the experimental effects, the modelling of signal and background processes introduces more sources of systematics, referred to as theory uncertainties, that are particularly relevant for the majority of SUSY searches.
Similarly to what was done in Chapter 4, the systematic uncertainties are implemented by modifying the MC samples according to the expected variations of the relevant quantities. The full list of systematics is presented in the next paragraph, describing separately the experimental and theoretical contributions.

Experimental systematics

The following experimental uncertainties are considered for the analyses in Chapters 6 and 7:

JES and JER The uncertainties on the Jet Energy Scale (JES) and Resolution (JER), already introduced in Section 4.1.3, are implemented with the same procedure in the physics analyses. The impact of the JES uncertainty is evaluated using a set of three uncorrelated variations, as detailed in Ref. [175], while for the JER uncertainty the variation of a single parameter is sufficient [176].

JVT The p_{T}-dependent JVT scale factors are modified within the uncertainties, obtained from dedicated measurements in $Z \to \mu\mu$ events [154].

***b*-tagging** The uncertainties on the performance of the MV2c10 algorithm are implemented by modifying the nominal values of the efficiency SFs, based on the result of data-driven calibration measurements and on the MC studies described in Chapter 4. The variations are applied separately to b-jets, c-jets and light jets, with flavour determined from the truth information in the MC samples. The following four independent contributions are evaluated:

- b-tagging efficiency uncertainty measured in $t\bar{t}$ events [162].
- c-tagging efficiency uncertainty from leptonic W events [163].
- Mis-tag rate uncertainty from a negative-tag method [164].
- Uncertainty on the tagging efficiency of high-p_{T} jets based on the studies in Chapter 4, referred to as extrapolation uncertainty.

Lepton and photon reconstruction The uncertainties on the calibration of the energy or momentum of leptons (electrons and muons) and photons, as well as on their energy or momentum resolutions, are implemented by varying the four-vectors within the measured variations [136, 143].

Lepton and photon efficiency The efficiency SFs associated with the reconstruction and calibration of leptons and photons are also modified within their expected uncertainties, based on the results of the corresponding performance studies [136, 139, 143].

Trigger efficiency Similarly, the trigger efficiency SFs are also carrying experimental uncertainties, that are implemented using the results of the dedicated measurements [94].

Missing transverse momentum The above systematic uncertainties on the physics objects are propagated to the E_T^{miss}, which is rebuilt using the modified input four-vectors. In addition, further variations are introduced to take into account the uncertainty on the track-based soft term described in Section 3.2.7.

Pileup The uncertainty on the total inelastic cross section of proton-proton collisions (of the order of 10%) affects the expected number of soft interactions in each event, and is hence propagated to the distribution of the pileup parameter μ in the MC samples. In the SUSY analyses, a systematic uncertainty on the pileup weights introduced in Section 5.3.3 is hence obtained by recomputing the weights using these modified μ profiles.

Luminosity The luminosity of the 2015+2016 dataset carries a 3.2% uncertainty [193], which is also taken into account in the analyses.

Theory systematics

The uncertainties on the modelling of signals and backgrounds are taken into account by adopting dedicated strategies for each process:

***V* + jets** The uncertainties on the production of vector bosons and jets are estimated by varying the Sherpa parameters related to the factorisation, renormalisation, resummation and CKKW matching scales, combining the individual variations into a total theoretical uncertainty. In addition, an extra 40% uncertainty is assigned to the heavy flavour jet content in the W+jets background by varying independently the normalisation of event samples with c-jets or b-jets, based on the results in Ref. [194].

Top production The theory systematics on the $t\bar{t}$ and single top (Wt) backgrounds are evaluated as the difference between the predictions of the nominal MC samples described in Section 5.2.2 and those of alternative samples with different generators or parameter settings. A first uncertainty concerns the amount of extra radiation emitted by the initial and final state of the scattering process, and is estimated using modified parameters in the nominal Powheg+Pythia generator. In parallel, the

hadronisation and PS uncertainties are obtained by comparing the nominal yields with those of alternative MC samples generated with Powheg and showered with Herwig, while the uncertainty on the event generator is estimated by comparing the above Powheg+Herwig sample with an alternative MadGraph+Herwig one. Finally, an extra source of uncertainty comes from the combined modelling of the $t\bar{t}$ and Wt processes, that are sharing the same final states and are hence subject to quantum mechanical interference effects. This is estimated using dedicated LO samples of $t\bar{t}$, Wt and inclusive $WWbb$ production generated with MadGraph, by comparing the sum of $t\bar{t}$ and Wt with the $WWbb$ predictions.

Dibosons and top pair plus *W/Z* The uncertainties on diboson and $t\bar{t} + W/Z$ backgrounds are estimated by modifying event generator parameters related to the factorization, renormalization, resummation, and CKKW matching scales, similarly to what done for the W/Z samples. Additional uncertainties on the cross sections are also considered, measuring around 6% for dibosons [195] and 13% for $t\bar{t} + W/Z$ [125].

Signals The theory systematics on the signal samples are calculated for each SUSY model. They are generally dominated by the uncertainties on the choice of the PDF set and on the renormalisation and factorisation scales [179].

Monte Carlo statistics Due to the finite size of the MC samples, the rates of the physics processes are predicted with finite precision. The impact of the uncertainty on MC statistics is particularly important when extreme regions of phase space are selected, as it often happens in SUSY analyses, and is relevant both for the nominal samples and for the samples that are used to estimate the theory systematics.

5.3.5 Signal Regions

A crucial step of the cut and count analyses is the optimisation of Signal Regions (SRs) which aims at enhancing the signal yield with respect to the dominant background sources. As a preliminary step, a basic selection of candidate events is performed by applying trigger and event cleaning cuts as described in Section 5.3.1, and by requiring the presence of a relevant set of physics objects based on the experimental signature of the targeted SUSY models. For example, in order to isolate the signal in Figure 5.7a (which is the target of the analysis in Chapter 7) it is reasonable to select events that pass a lepton trigger requirement and have a minimum of two jets and two isolated leptons.

Once the events with signal-like properties are identified, additional selections must be applied to suppress the leftover backgrounds whilst retaining the largest possible fraction of signal. This is done using dedicated sets of discriminating variables with different distributions in signal and background, that are strongly analysis-dependent and are hence not discussed in the present chapter.
The optimisation of SR selections is performed by maximising the value of a figure of merit that represents the discovery significance of the signal model of interest (see Section 5.4.2). A variety of definitions of the significance are possible [196], and the analyses in this thesis employ the Z_{N} formula [197], which is implemented in the RooStats package [198] of ROOT [199]. Alternatively, the following simplified expression can be used:

$$Z = \frac{N_{\mathrm{sig}}}{\sqrt{N_{\mathrm{bkg}} + (\sigma_{\mathrm{bkg}} N_{\mathrm{bkg}})^2}} \tag{5.11}$$

where N_{sig} and N_{bkg} are the signal and background yields, and σ_{bkg} is an assumption on the relative systematic uncertainty on the background. Equation 5.11 is useful to understand intuitively the essential properties of the significance, that can be expressed as the ratio of the signal yield in the SR and the total uncertainty on the background, defined as the sum in quadrature of the Poisson error $\sqrt{N_{\mathrm{bkg}}}$ and the expected systematic uncertainty $\sigma_{\mathrm{bkg}} N_{\mathrm{bkg}}$.
The above considerations serve as general guidelines for the definition of the SRs in the analyses in Chapters 6 and 7, which is done following the procedure described below:

1. A preliminary event selection is applied according to the basic experimental signature of the targeted signal models.

2. A set of signal models with fixed masses of the SUSY particles is chosen as benchmark for the SR definitions. Typically, each of these benchmark signals is representative of a subset of the targeted SUSY models which yields events with similar kinematical properties.

3. A set of discriminating variables providing a good separation between signal and background processes is identified. For each of them, a range of potential SR cut thresholds is chosen by comparing the kinematical distribution of the variable in signal and background events.

4. For every signal benchmark model, several possible SR definitions are tested by performing a multi-dimensional scan of selections on the discriminating variables,

varying the cut thresholds within the ranges determined in the previous stage. Each combination of cuts yields a different value of the discovery significance, which is computed using the signal and background yields predicted by the MC samples together with an estimate of the size of the systematic uncertainty on the background prediction.

5. The optimal SR definitions are chosen by selecting the combinations of cuts which maximise the discovery significance in each scan.

Finally, it is important to mention that the SUSY analyses are performed following a blinding procedure, which consists in hiding the number of observed events in the SRs until the selections are properly optimised and a robust background estimation strategy is in place. This is done to prevent the analysers from being biased during the definition of the search strategy, and is a widely adopted standard in the high energy physics experiments.

5.3.6 Background estimation

In order to determine the background yields in the SRs without relying exclusively on MC simulations, a semi data-driven method based on the definition of Control Regions (CRs) is implemented in several SUSY analyses, including those presented in this thesis. The CRs are defined as orthogonal selections to the SRs, such that their event yield is dominated by a specific background process and the signal is instead strongly suppressed. A normalisation scale factor for each relevant background is derived by rescaling the expected MC yield to the observed number of data events in the corresponding CR, and the rate prediction in the SR is rescaled accordingly. In a simplified case with one SR and one CR with 100% background purity, the expected background yield in the SR is given by:

$$N_{\mathrm{SR}}^{\mathrm{exp}} = \mu_{\mathrm{MC}} \cdot N_{\mathrm{SR}}^{\mathrm{MC}}, \qquad \mu \equiv \frac{N_{\mathrm{CR}}^{\mathrm{data}}}{N_{\mathrm{CR}}^{\mathrm{MC}}} \tag{5.12}$$

where the parameter μ is the normalisation scale factor described above. Alternatively, the SR yield can be expressed as a function of the number of observed data events in the CR and the MC Transfer Factor (TF):

$$N_{\mathrm{SR}}^{\mathrm{exp}} = TF \cdot N_{\mathrm{CR}}^{\mathrm{data}}, \qquad TF \equiv \frac{N_{\mathrm{SR}}^{\mathrm{MC}}}{N_{\mathrm{CR}}^{\mathrm{MC}}} \tag{5.13}$$

where it is shown explicitly that the result relies on the MC simulation only for the computation of the ratio between the SR and CR predictions. As a direct consequence, the

total uncertainty on $N_{\mathrm{SR}}^{\mathrm{exp}}$ is determined exclusively by the Poisson error on $N_{\mathrm{CR}}^{\mathrm{data}}$ and by the uncertainty on the MC extrapolation from the CR to the SR, while any variation in the absolute normalisation of the MC sample (which may be due, for example, to the uncertainty on the cross section of the physics process) has no impact on the final background estimate.

In realistic physics analyses, multiple background processes are normalised in an independent set of mutually exclusive CRs, and the background purity in each CR cannot reach 100%. The expected values of the normalisation factors can be computed by solving a system of equations with n free parameters and constraints:

$$\begin{cases} N_{\mathrm{CR,1}}^{\mathrm{data}} = \mu_1 N_{\mathrm{CR,1}}^{\mathrm{MC,1}} + \mu_2 N_{\mathrm{CR,1}}^{\mathrm{MC,2}} + ... + \mu_n N_{\mathrm{CR,1}}^{\mathrm{MC,n}} \\ N_{\mathrm{CR,2}}^{\mathrm{data}} = \mu_1 N_{\mathrm{CR,2}}^{\mathrm{MC,1}} + \mu_2 N_{\mathrm{CR,2}}^{\mathrm{MC,2}} + ... + \mu_n N_{\mathrm{CR,2}}^{\mathrm{MC,n}} \\ ... \\ N_{\mathrm{CR,n}}^{\mathrm{data}} = \mu_1 N_{\mathrm{CR,n}}^{\mathrm{MC,1}} + \mu_2 N_{\mathrm{CR,n}}^{\mathrm{MC,2}} + ... + \mu_n N_{\mathrm{CR,n}}^{\mathrm{MC,n}} \end{cases} \tag{5.14}$$

where μ_i is the normalisation factor of the process i, $N_{\mathrm{CR,j}}^{\mathrm{MC,i}}$ is the MC yield of the process i in the CR j and $N_{\mathrm{CR,k}}^{\mathrm{data}}$ is the number of data events in the CR k. In practice, the μ values are determined through a dedicated statistical procedure known as parameter estimation, described in detail in Section 5.4.

The accuracy of the background estimation strategy is verified in a dedicated set of Validation Regions (VRs), where the signal contamination is required to be low and the data are expected to match the SM predictions within uncertainties. As illustrated in Figure 5.14, the VRs are generally designed to lie within the CR-SR extrapolation region, in order to assess the impact of any source of bias that may affect the transfer factors defined in Equation 5.13.

The agreement between data and predictions in the VRs is a necessary condition to remove the blinding constraint introduced in Section 5.3.5 from the SRs of the analysis. Once this is done, the signal is expected to appear as an excess of events in the SRs with respect to the background only hypothesis, with no corresponding effect in the VRs.

5.4 Statistical analysis

The previous two sections have presented the general structure of the cut and count analyses of Chapters 6 and 7, that are based on the definition of different types of regions where the relevant signal and background parameters are measured. While a basic assessment of the value of the parameters can be obtained by solving systems as the one in

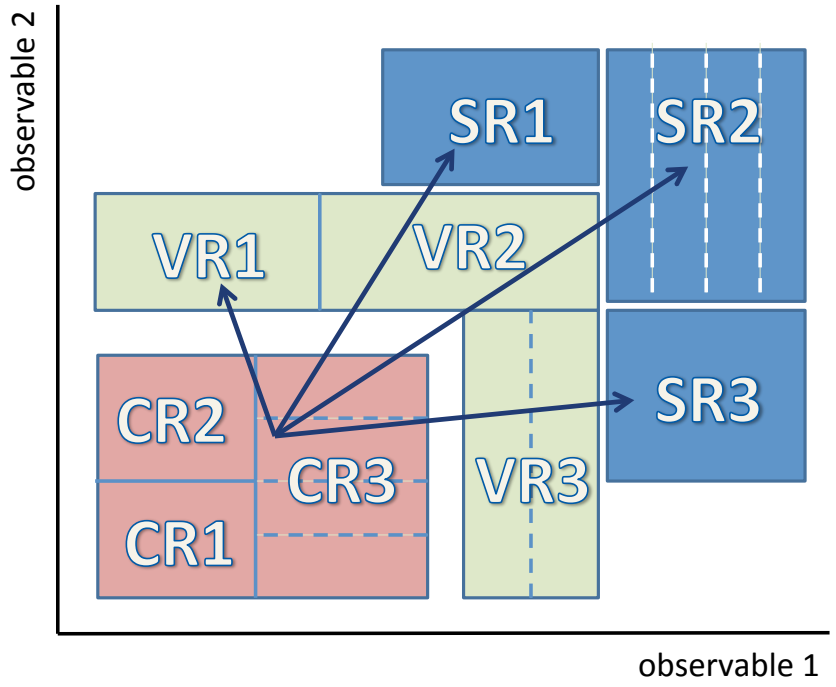


Figure 5.14: Scheme of the general structure of a cut and count analysis, where a set of CRs, VRs and SRs are defined using two observables [200]. The individual regions are required to be mutually orthogonal, and the VRs are located in an intermediate position to verify the quality of the CR-SR extrapolation.

Equation 5.14, the extraction of quantitative results requires appropriate statistical tools that take into account the inherent uncertainties. The analyses in this thesis employ two types of statistical procedures, known as parameter estimation and hypothesis testing: the former allows to measure a set of parameters from the input dataset, while the latter is used to verify the compatibility of specific models (hypotheses) with the observed data. Both procedures are implemented through the HistFitter framework [200] following a frequentist statistical approach, where the results of each analysis are interpreted as one of an infinite set of possible outcomes of the same experiment.

5.4.1 Parameter estimation

One of the primary goals of the physics analyses in this thesis is the measurement of the parameters through which the data in the CRs, VRs and SRs are interpreted. These include the normalisation factors μ_b of the SM backgrounds (see Equation 5.12) and an analogous quantity μ_s for the benchmark signal sample, commonly referred to as signal

strength. The expected rate in any region X of the analysis can be predicted as:

$$N_{\mathrm{X}} = \mu_s N_s + \sum_i \mu_b^i N_b^i \tag{5.15}$$

where N_b^i is the expected yield of the background process i and N_s is the expected signal yield, all determined from the corresponding MC simulations. In addition, a set of extra nuisance parameters θ can be introduced to describe the impact of the systematic uncertainties on both signal and background, modifying Equation 5.15 as:

$$N_{\mathrm{X}} = \mu_s N_s (1 + \sum_j \theta_s^j \sigma_s^j) + \sum_i \mu_b^i N_b^i (1 + \sum_j \theta_b^{ij} \sigma_b^{ij}) \tag{5.16}$$

where σ_b^{ij} and σ_s^j are the relative variations of background and signal yields respectively produced by the individual uncertainties. The above expression is configured such that the nominal yields are obtained for $\theta = 0$, while $\theta = \pm 1$ corresponds to the $\pm 1\sigma$ variations. The $\boldsymbol{\theta}$ parameters can be common to many physics processes, for example when representing an experimental uncertainty that applies simultaneously to all of them, but they can also be related to single processes if they describe specific uncertainties such as the theory systematics.

The values of the μ_s, $\boldsymbol{\mu}_b$ and $\boldsymbol{\theta}$ parameters are extracted from data by constructing a likelihood function L, which condenses the relevant information of the analysis into a unique expression. The likelihood is a function of the set of free parameters to be measured, and is constrained by the observed number of events in the various regions of the analysis. As discussed in Ref. [200], its general expression is the product of Poisson distributions of event counts in the relevant regions (typically CRs and/or SRs) and of additional distributions that implement the impact of the systematic uncertainties:

$$\begin{aligned} L(\boldsymbol{N}^{\mathrm{obs}}, \boldsymbol{\theta}^0 | \mu_s, \boldsymbol{\mu}_b, \boldsymbol{\theta}) &= P_{\mathrm{SR}} \times P_{\mathrm{CR}} \times C_{\mathrm{syst}} = \\ &= \prod_{i \in \mathrm{SR,CR}} P(N_i^{\mathrm{obs}} | N_i(\mu_s, \boldsymbol{\mu}_b, \boldsymbol{\theta})) \times C_{\mathrm{syst}}(\boldsymbol{\theta}^0, \boldsymbol{\theta}). \end{aligned} \tag{5.17}$$

The Poisson factors in Equation 5.17 reflect the stochastic nature of the event counts in the SRs and CRs, that are expressed as a function of the signal and background normalisations μ_s and $\boldsymbol{\mu}_b$ and the nuisance parameters $\boldsymbol{\theta}$, as described by Equation 5.16. The additional C_{syst} component is a function of auxiliary measurements $\boldsymbol{\theta}_0$ and nuisance parameters $\boldsymbol{\theta}$, which is introduced to constrain the systematic uncertainty in the fit, penalising values of $\boldsymbol{\theta}$ significantly shifted from $\boldsymbol{\theta}_0$. The constraint term C_{syst} can be implemented in different ways [200], and a common choice is a unit Gaussian where the

individual measurements θ_0^i are set to 0, such that the fitted values of θ^i are expected to be approximately 0 ± 1, reproducing the expected size of the systematic uncertainties through Equation 5.16.
Once the likelihood function in Equation 5.17 is built, the values of the parameters are obtained by maximising it following the Maximum Likelihood Estimation (MLE) procedure, illustrated in detail in Ref. [201].

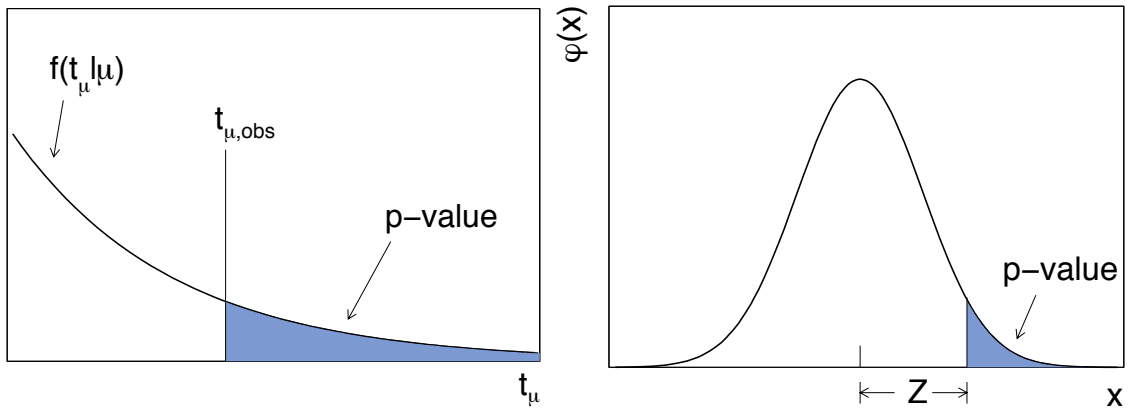
5.4.2 Hypothesis testing

The ultimate purpose of the physics analyses is to determine whether the data indicate or exclude the presence of a BSM signal. This is achieved through a statistical procedure known as hypothesis testing [202], and is implemented by defining a null hypothesis H0 to be tested against an alternative H1. When targeting the discovery of a signal of new physics, H0 and H1 are chosen as the background only and the signal plus background hypotheses respectively, while the roles are inverted when setting exclusion limits.
The hypothesis testing procedure is based on the definition of a test statistic t, a function of the observed data configured such that it assumes large values when the data are incompatible with H0, following a probability distribution $f(t)$. In the frequentist approach, a standard way to determine the shape of $f(t)$ is to generate a large number of pseudo-experiments (toys) in which the values of the observable quantities are randomly generated under the H0 hypothesis. Alternatively, for given types of test statistics and under given assumptions on the size of the statistical samples, the distribution can be obtained using asymptotic approximations [203].
Once the distribution $f(t)$ is known, the hypothesis H0 is tested through the computation of a p-value, which represents the probability of observing a larger incompatibility of the data with the predictions in an infinite number of repetitions of the experiment under the assumption that H0 is valid. The p-value is computed as:

$$p = \int_{t_{\text{obs}}}^{\infty} f(t)dt \tag{5.18}$$

where the integral is taken from the observed value of the test statistic to infinity, as shown in Figure 5.15a. For practical purposes, it is convenient to convert the p-value into an equivalent significance Z, defined as the number of standard deviations σ from the mean of a gaussian distribution for which the integral of the tail of the curve is equal to p:

$$Z = \Phi^{-1}(1 - p), \tag{5.19}$$



(a) (b)

Figure 5.15: Graphical representation of a p-value (a) obtained from a test statistic t for a given signal strength μ (see Equation 5.20) and of the corresponding significance Z (b) [203].

where Φ^{-1} is the inverse of the cumulative distribution of the gaussian (see Figure 5.15b). The threshold chosen by the particle physics community to declare a discovery is a significance $Z = 5$ against the background only hypothesis, which corresponds to $p = 2.87 \times 10^{-7}$, while $Z = 3$ ($p = 0.0013$) is regarded as evidence. Signal models are instead excluded using $p = 0.05$ ($Z = 1.64$).

The Profile Likelihood Ratio

From the above discussion it is clear that the choice of the appropriate test statistic plays a crucial role in the hypothesis testing procedure. The test statistics used by the LHC experiments are based on a Profile Likelihood Ratio (PLR) [202] obtained from the likelihood in Equation 5.17:

$$\lambda(\mu_s) = \frac{L(\mu_s, \hat{\hat{\boldsymbol{\theta}}})}{L(\hat{\mu}_s, \hat{\boldsymbol{\theta}})} \tag{5.20}$$

where the vector $\boldsymbol{\theta}$ includes the background normalisation factors and the nuisance parameters related to the systematic uncertainties. The denominator $L(\hat{\mu}_s, \hat{\boldsymbol{\theta}})$ corresponds to the absolute maximum of the likelihood function, while the numerator $L(\mu_s, \hat{\hat{\boldsymbol{\theta}}})$ is the maximum for a given value of μ_s.

Equation 5.20 implies that the range of λ extends from 0 to 1, with larger values corresponding to a better agreement of the data with the μ_s hypothesis under test. A test statistic with the range required by the definition of the p-value (Equation 5.18) can be

defined as:

$$t_\mu = -2 \ln \lambda(\mu) \tag{5.21}$$

where larger values of t correspond to a lower compatibility between the data and the hypothesis.
Finally, two versions of the test statistic are defined to perform different types of test:

Test statistic for discovery The discovery of a new signal is targeted by testing the background only hypothesis. This is done by using a PLR function with $\mu_s = 0$, with the following definition:

$$q_0 = \begin{cases} -2 \ln \lambda(0) & \hat{\mu}_s \geq 0 \ , \\ 0 & \hat{\mu}_s < 0 \ , \end{cases} \tag{5.22}$$

where q_0 is set to 0 for negative $\hat{\mu}_s$ to avoid the exclusion of the background hypothesis if a deficit of events is observed in the SRs. A set of possible analytical expressions of the discovery significance Z obtained from the above test statistic are discussed in Ref. [204], including the simplified expression in Equation 5.11.

Test statistic for exclusion When targeting the exclusion of a signal model, the test statistic is instead computed as follows:

$$q_\mu = \begin{cases} -2 \ln \lambda(\mu_s) & \hat{\mu}_s \leq \mu_s \ , \\ 0 & \hat{\mu}_s > \mu_s \ . \end{cases} \tag{5.23}$$

where a non-zero signal strength μ_s is assumed in H0.

Exclusion limits with the CL_s method

A problematic feature of the exclusion test statistic in Equation 5.23 is that it yields a non-negligible probability of excluding signal models to which the analysis is poorly sensitive. While qualitative arguments suggest that these models should not be constrained, low p-values for the signal plus background hypothesis may indeed be obtained in case of significant down-fluctuations of the observed events in the SRs. This pathological behaviour can be avoided by exploiting the fact that such signals yield similar distributions of discovery and exclusion test statistics, defining an alternative figure of merit for the exclusion [205] as:

$$\mathrm{CL}_s = \frac{p_{\mu_s}}{1 - p_b} \tag{5.24}$$

where p_b and p_{μ_s} are the p-values of the background only and signal plus background hypotheses, and the threshold for a $Z = 2$ (or 95% Confidence Level, CL) exclusion is $\mathrm{CL}_s < 0.05$. In situations where the exclusion and discovery test statistics have similar distributions, the numerator and denominator of Equation 5.24 are of the same order, so the signals are not excluded as expected intuitively. Thanks to this feature, the CL_s is adopted by the LHC experiments as the standard tool to set exclusion limits on signal models.

5.4.3 Statistical tools in SUSY searches

In summary, the following statistical tools are relevant for the SUSY analysis presented in this thesis:

Background only fit The background estimation strategy of the analysis, described in Section 5.3.6, is based on the evaluation of the normalisation factors of the backgrounds through the parameter estimation procedure described in Section 5.4.1. In particular, a likelihood of the form of Equation 5.17 is built including all the CRs of the analysis, and is then maximised to determine the values of the normalisation factors and of the nuisance parameters $\boldsymbol{\theta}$, taking into account the mutual correlations. The fitted parameters are then used to compute the yield predictions in each region using Equation 5.16.

Discovery test The results in each SR of the analysis are interpreted by computing a p-value for the background only hypothesis, following the hypothesis testing procedure presented in Section 5.4.2. A PLR function is built from a likelihood which includes all the CRs and the relevant SR under exam, where the yield predictions are determined only from the SM processes (or, equivalently, where the signal strength parameter μ_s is set to 0). The p-value and the associated significance are then computed as the integral of the distribution of the test statistic q_0 (Equation 5.22), using either toy experiments or the relevant asymptotic formula [203].

Limit setting If no excess is observed in the SRs, exclusion limits are set on a various signal hypotheses using the CLs method (see Equation 5.24). This requires the computation of the q_0 and q_μ test statistics (Equations 5.22 and 5.23) from the two corresponding PLR functions, constructed under the background only hypothesis and for a given signal strength μ_s. The evaluation of the PLRs requires the minimisation of several likelihood functions, performed as usual with the HistFitter

framework [200]: for this reason, in ATLAS SUSY jargon the limit setting procedure is often referred to as exclusion fit.

6 | Search for bottom squarks with two-body RPC decays

A major result of this thesis is the search for $\tilde{b}$ pair production in two-body RPC decay scenarios, inspired by the naturalness arguments outlined in Section 1.3.3. As discussed in detail in Appendix A, during my PhD I contributed as main analyser to a first publication with 3.2 fb^{-1} of data at $\sqrt{s} = 13$ TeV [206] and to a subsequent publication with 36.1 fb^{-1} [207], which expands the first result thanks to the larger available dataset and a more elaborated analysis strategy. The chapter focuses on the 36.1 fb^{-1} analysis, and begins by presenting the signal benchmark models of interest (Section 6.1) together with an overview of the results from previous searches. The event selection techniques are then discussed in Section 6.2, where the performance of the key discriminating variables on the signal and background processes is described in detail. Section 6.3 outlines the background estimation strategy employed in the different channels of the analysis, which includes a standard CR-based approach and additional data-driven techniques, while Section 6.4 present the full results of the search based on the statistical tools introduced in Section 5.4.

6.1 Signal model

The analysis in this chapter targets the pair production of the superpartner of the bottom quark, the $\tilde{b}$, in natural RPC scenarios, already introduced in Section 1.3.3. The next paragraph describes in detail the signals of interest (Section 6.1.1), the simplified benchmark models (Section 6.1.2) and the experimental constraints from previous searches (Section 6.1.3).

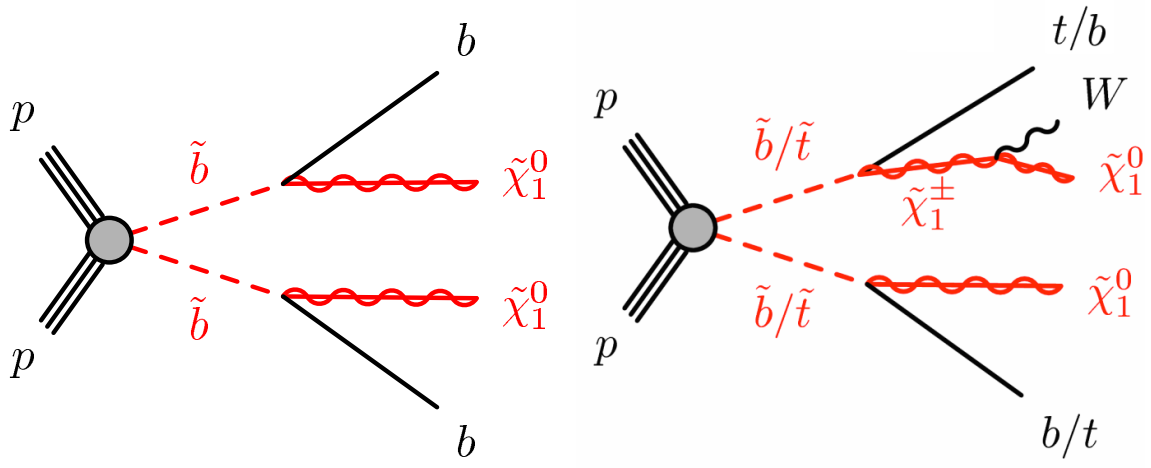
6.1.1 Bottom squarks in natural SUSY models

As introduced in Sections 1.3.3 and 5.1.3, naturalness arguments related to the Higgs hierarchy problem constrain the masses of a subset of SUSY particles, including Higgsino-like mass eigenstates, third generation squarks and gluinos, yielding the natural mass spectrum shown in Figure 1.7. The target of this analysis is a pair-produced $\tilde{b}$ extracted from the natural pMSSM spectrum, with relevant RPC decays into a stable $\tilde{\chi}_1^0$ LSP or a $\tilde{\chi}_1^+$ NLSP. The model is closely related to the Higgsino-like LSP scenario presented in Figure 5.3c for the case of the $\tilde{t}$ pair production, but in the present analysis the pair production of the $\tilde{b}$ is assumed.

Similarly to what is done in many other SUSY searches, the signal is generated as a simplified model which includes only the $\tilde{b}$, the $\tilde{\chi}_1^+$ and the $\tilde{\chi}_1^0$, while no additional SUSY particles are considered. An assumption of negligible mass splitting between the $\tilde{\chi}_1^0$ and the $\tilde{\chi}_1^+$, coherent with the natural SUSY spectrum, is implemented by setting $\Delta m(\tilde{\chi}_1^+, \tilde{\chi}_1^0) = 1$ GeV in the signal samples, such that the $\tilde{\chi}_1^+$ undergoes a prompt decay into $W^*\tilde{\chi}_1^0$ with a high degree of virtuality of the vector boson. The practical consequence is that the final state objects from the W^* emission are produced with low momentum, which prevents them from being reconstructed in the detector, so the experimental signature of any $\tilde{\chi}_1^0$ or $\tilde{\chi}_1^+$ from the $\tilde{b}$ decay chain is limited to the presence of E_T^{miss} in the events.

In full pMSSM models, the natural SUSY spectrum is obtained by setting the μ parameter in Equations 1.28 and 1.29 to the electroweak scale, while the M_1 and M_2 parameters are set to a few TeV or more. In addition to the $\tilde{\chi}_1^0$ and the $\tilde{\chi}_1^+$, this configuration predicts an extra light $\tilde{\chi}_2^0$, that can decay into a $Z^*\tilde{\chi}_1^0$ pair or possibly into a $W^*\tilde{\chi}_1^+$ pair if it is heavier than the $\tilde{\chi}_1^+$. The $\tilde{\chi}_2^0$ is not included in the benchmark model of the present analysis, under the assumption that it is almost mass degenerate with the other two states, such that its interaction properties and its experimental signature are both effectively equivalent to those of the $\tilde{\chi}_1^0$. This implies that the $\tilde{b}$ decays into a $\tilde{\chi}_1^0$ or a $\tilde{\chi}_2^0$ are experimentally indistinguishable, so the $\tilde{\chi}_1^0$ can be taken as representative of both scenarios.

From the above discussion it is clear that the assumption of negligible mass splitting between the three Higgsino-like states has important implications on the phenomenological properties of the signals. An alternative scenario with mass splittings of the order of a few GeV, still compatible with naturalness arguments, gives rise to additional soft objects in the final states, and is targeted by the analysis in Ref. [184] based on the model in Figure 5.3c.



(a) (b)

Figure 6.1: Benchmark signal models targeted by the analysis in this chapter. Diagram (a) shows a pair-produced $\tilde{b}$ with 100% BR decay into $b\tilde{\chi}_1^0$, which yields an experimental signature with two b-jets and E_T^{miss}. Diagram (b) shows an asymmetric decay of the $\tilde{b}$ (or $\tilde{t}$) into $b\tilde{\chi}_1^0$ and $t\tilde{\chi}_1^+$ (or $t\tilde{\chi}_1^0$ and $b\tilde{\chi}_1^+$) which yields a $tb + E_T^{\text{miss}}$ final state under the assumption of small $\Delta M(\tilde{\chi}_1^+, \tilde{\chi}_1^0)$. The latter signature represents 50% of the events in the mixed decay scenario where the $\tilde{b}$ decays into $b\tilde{\chi}^0$ and $t\tilde{\chi}_1^+$ with 50% BR each, which also includes 25% of $bb + E_T^{\text{miss}}$ and 25% of $tt + E_T^{\text{miss}}$ events.

6.1.2 Benchmark processes

The analysis in this chapter targets the two-body decays of the $\tilde{b}$ into $b\tilde{\chi}_1^0$ or $t\tilde{\chi}_1^+$, where $\tilde{\chi}_1^+$ and $\tilde{\chi}_1^0$ yield the same experimental signature due to the above assumptions on $\Delta m(\tilde{\chi}_1^+, \tilde{\chi}_1^0)$.

The first process of interest, illustrated by the diagram in Figure 6.1a, assumes the $\tilde{b} \to b\tilde{\chi}_1^0$ decay with 100% BR, which results in the $bb + E_T^{\text{miss}}$ final state with two bottom quarks and invisible particles. The analysis is also targeting a model with the mixed decay of the $\tilde{b}$ into $b\tilde{\chi}^0$ and $t\tilde{\chi}_1^+$ with 50% BR each, which gives rise to a more complex scenario where 25% of the events yield a $bb + E_T^{\text{miss}}$ final state, 50% of them yield the asymmetric $tb + E_T^{\text{miss}}$ final state in Figure 6.1b and the remaining 25% yields a pure $tt + E_T^{\text{miss}}$ final state with two top quarks and two invisibly decaying $\tilde{\chi}_1^+$. For simplicity, in the following paragraphs the mixed decay model is referred to as $tb + E_T^{\text{miss}}$ signal scenario, which includes both the asymmetric events with actual $tb + E_T^{\text{miss}}$ signature and the two symmetric components ($bb + E_T^{\text{miss}}$ and $tt + E_T^{\text{miss}}$) with 25% fraction each.

An additional consequence of the assumption on the small $\Delta m(\tilde{\chi}_1^+, \tilde{\chi}_1^0)$ is that the $bb + E_T^{\text{miss}}$, $tb + E_T^{\text{miss}}$ and $tt + E_T^{\text{miss}}$ final states can be obtained equivalently from $\tilde{b}$ or $\tilde{t}$ pair production. The $bb + E_T^{\text{miss}}$ scenario, which in the present analysis is interpreted as coming from the $\tilde{b} \to b\tilde{\chi}_1^0$ decay, can also be obtained from a pair-produced $\tilde{t}$ decaying into $b\tilde{\chi}_1^+$ with 100% BR, and similarly the $tt + E_T^{\text{miss}}$ final state can be obtained from both $\tilde{t} \to t\tilde{\chi}_1^0$ and $\tilde{b} \to t\tilde{\chi}_1^+$ decays. The same consideration is valid also for the $tb + E_T^{\text{miss}}$ scenario, which is produced by a combination of the two decays. The present analysis concentrates on the two diagrams in Figure 6.1, while the pure $tt + E_T^{\text{miss}}$ final state is targeted by a dedicated search for $\tilde{t}$ production [207].

6.1.3 Previous results

The analysis in this thesis is based on the full 2015+2016 dataset described in Section 5.3.1, but the same signal models were also targeted by previous ATLAS and CMS searches, with minor differences in the case of the asymmetric signal in Figure 6.1b. In particular, I was personally involved in an ATLAS search for $\tilde{t}$ pair production in the $tb + E_T^{\text{miss}}$ final state with $\sqrt{s} = 8$ TeV data, which is included in the third generation summary paper of Run 1 [183] and is discussed in my Master Degree thesis [208], and I also contributed to an early Run 2 result on $\tilde{b}$ pair production in the $bb + E_T^{\text{miss}}$ final state, which used 3.2 fb^{-1} of data at $\sqrt{s} = 13$ TeV collected during 2015 [209].
The first ATLAS searches for $\tilde{b}$ signals in the $bb + E_T^{\text{miss}}$ channel were performed during Run 1, first with 2.05 fb^{-1} of data at $\sqrt{s} = 7$ TeV [210] and then with 20.1 fb^{-1} of data at $\sqrt{s} = 8$ TeV [211]. These results are obtained with a baseline analysis strategy replicated also in the subsequent Run 2 searches to which I participated during my PhD, which include the first Run 2 publication with 3.2 fb^{-1} of data [209] and the 36.1 fb^{-1} analysis presented in this chapter [18]. Figure 6.2 shows the exclusion limits of the 3.2 fb^{-1} analysis [209] on the simplified signal model shown in Figure 6.1a: $\tilde{b}$ masses up to 800 (840) GeV are excluded for $\tilde{\chi}_1^0$ mass below 360 (100) GeV, while mass differences $\Delta m(\tilde{b}, \tilde{\chi}_1^0)$ below 100 GeV are excluded up to a $\tilde{b}$ mass of 500 GeV. The figure includes the reinterpretation of two results of mono-jet searches, based on 20.1 fb^{-1} of data at $\sqrt{s} = 8$ TeV [212] and on 3.2 fb^{-1} of data at $\sqrt{s} = 13$ TeV [206], that are sensitive to signals with small values of $\Delta m(\tilde{b}, \tilde{\chi}_1^0)$.
The $tb + E_T^{\text{miss}}$ final state was studied for the first time in ATLAS by the dedicated Run 1 analysis mentioned above [183,208], targeting a full natural pMSSM model parametrised by μ and m_{Q_3} (see Equation 1.24) and two simplified models with a pair-produced $\tilde{t}$ de-

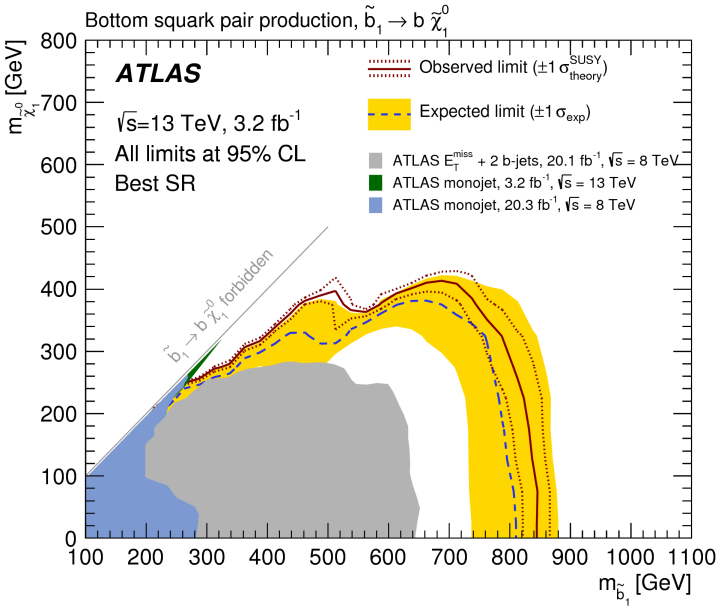


Figure 6.2: Exclusion limits on a pair-produced $\tilde{b}$ decaying into a $b\tilde{\chi}_1^0$ pair with 100% BR, based on the analysis of 3.2 fb^{-1} of data at $\sqrt{s} = 13$ TeV [209].

caying into $t\tilde{\chi}_1^0$ or $b\tilde{\chi}_1^+$ with $\Delta m(\tilde{\chi}_1^+, \tilde{\chi}_1^0)$ set to 5 and 20 GeV respectively, and with variable BR of the two decays. The two simplified models closely resemble the one considered by this analysis (Figure 6.1b) with the exception of the larger values of $\Delta m(\tilde{\chi}_1^+, \tilde{\chi}_1^0)$. The fact that the Run 1 analysis considered a pair-produced $\tilde{t}$ instead of a pair-produced $\tilde{b}$ is irrelevant for the comparison, because an equivalent experimental signature is obtained in both cases (as already explained in Section 6.1.2). In Ref. [183] the signals are targeted by four SRs in the one-lepton channel, and the exclusion limits are shown in Figure 6.3 for the case of $\Delta m(\tilde{\chi}_1^+, \tilde{\chi}_1^0) = 5$ GeV. The observed exclusion is limited by the presence of mild SR excesses, that are interesting to investigate in $\sqrt{s} = 13$ TeV data, as further discussed in Section 6.2.4.

Finally, the CMS Collaboration has also published multiple searches [213–215] targeting the pair production of the $\tilde{b}$ in the $bb + E_T^{\text{miss}}$ final state. The analyses are generally based on the definition of multiple SRs binned in key discriminating variables related to the kinematics of the jets and the E_T^{miss}, resulting in competitive exclusion limits on the benchmark signal model.

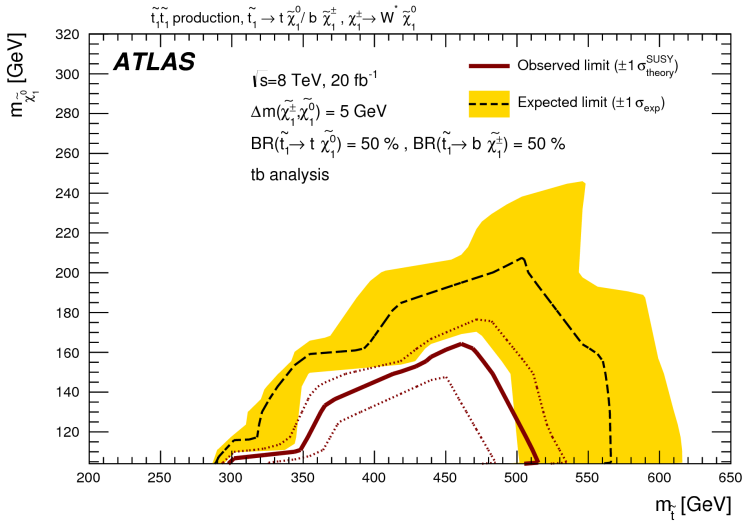


Figure 6.3: Exclusion limits on a pair-produced $\tilde{t}$ decaying into $t\tilde{\chi}_1^0$ or $b\tilde{\chi}_1^+$ with 50% BR each, based on the analysis of 20.1 fb^{-1} of data at $\sqrt{s} = 8$ TeV [183].

6.2 Event selection

The present section is dedicated to the description of the event selection techniques that are employed to isolate the targeted signals with respect to the dominant SM backgrounds. The search is based on the definition of two mutually exclusive sets of SRs with zero leptons and one lepton respectively, that are referred to the zero-lepton (b0L) and one-lepton (b1L) channels of the analysis. The former is optimised for the symmetric $\tilde{b}$ signal yielding the $bb + E_T^{\text{miss}}$ signature in Figure 6.1a, while the latter targets the mixed $\tilde{b}$ decays represented by the diagram in Figure 6.1b. The next paragraphs introduce the key discriminating variables that are used in both channels of the analysis (Section 6.2.1) showing the corresponding kinematical distributions after a preliminary set of event selections are applied (Section 6.2.2). The detailed definitions of the SRs in the zero-lepton and one-lepton channels is then presented in Sections 6.2.3 and 6.2.4.

6.2.1 Discriminating variables

The selection of events is performed based on multiple discriminating variables, built from the physics objects presented in Section 5.3.2. A summary of their definitions is provided in the present paragraph, where the signal jets are ordered in decreasing p_{T}:

H_{T} A first variable is defined as the scalar sum of the p_{T} of the signal jets in the event:

$$H_{\mathrm{T}} = \sum_i p_{\mathrm{T}}^{\mathrm{jet}_i}. \tag{6.1}$$

H_{T4} A subset of SRs also employ a modified version of H_{T}, referred to as H_{T4}, where the scalar sum of the p_{T} of the jet is computed starting only from the 4^{th} jet in each event (if present).

m_{eff} Similarly, the scalar sum of the H_{T} and the E_T^{miss} is used:

$$m_{\mathrm{eff}} = \sum_i p_{\mathrm{T}}^{\mathrm{jet}_i} + E_T^{\mathrm{miss}} = H_{\mathrm{T}} + E_T^{\mathrm{miss}}. \tag{6.2}$$

$\mathcal{A}$ Another useful quantity is the asymmetry of p_{T} of the of the leading two jets:

$$\mathcal{A} = \frac{p_{\mathrm{T}}(j_1) - p_{\mathrm{T}}(j_2)}{p_{\mathrm{T}}(j_1) + p_{\mathrm{T}}(j_2)}. \tag{6.3}$$

Events with large $\mathcal{A}$ are characterised by the presence of a single high-p_{T} jet plus extra soft activity.

E_T^{miss} and $E_T^{\mathrm{miss,corr}}$ The standard E_T^{miss} is used as a key discriminant between SUSY signals and SM backgrounds. A modified $E_T^{\mathrm{miss,corr}}$ version is defined in the di-lepton and single photon channels by removing the leptons and the photon from the baseline E_T^{miss} computation, which corresponds to treating them as invisible objects.

$\Delta\phi^{j}_{\mathrm{min}}$, $\Delta\phi^{j_{1\text{-}4}}_{\mathrm{min}}$, $\Delta\phi^{j_{1\text{-}2}}_{\mathrm{min}}$ These variables represent the minimum azimuthal distance between the jets and the E_T^{miss}:

$$\Delta\phi^{j}_{\mathrm{min}} = \min[\Delta\phi(\mathrm{jet}, E_T^{\mathrm{miss}})] \tag{6.4}$$

where $\Delta\phi^{j}_{\mathrm{min}}$ is computed with all signal jets in the event, while $\Delta\phi^{j_{1\text{-}4}}_{\mathrm{min}}$ and $\Delta\phi^{j_{1\text{-}2}}_{\mathrm{min}}$ employ only the leading 4 or 2 jets respectively. If the E_T^{miss} is arising from jet mis-measurements, the $\Delta\phi$ variables are expected to have a low average value.

$E_T^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}}$, $E_T^{\mathrm{miss}}/m_{\mathrm{eff}}$ These two ratios are also employed to reject events where the E_T^{miss} is likely to arise from jet mis-measurements, or from the presence of neutrinos produced in the decays of the hadrons within the jet cones. The $E_T^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}}$ variable is also known as E_T^{miss} significance and represents the ratio of the E_T^{miss} and the energy resolution of the calorimeter, while $E_T^{\mathrm{miss}}/m_{\mathrm{eff}}$ is a similar quantity that assumes low values if the E_T^{miss} is small compared to the total energy in the calorimeter.

m_{jj} and $m_{\ell\ell}$ The invariant masses of the leading two jets (m_{jj}) and leading two leptons ($m_{\ell\ell}$) in the events are used as key discriminants in the analysis. The m_{jj} variable has two possible variations: the first is m_{bb}, where the invariant mass is computed with the leading b-tagged jets in events where at least two of them are present, and the second is m_{bj}, which is computed in events with exactly one b-jet using the b-jet and the leading non b-tagged jet.

m_{T} The transverse mass m_{T} is computed in the single lepton channels as:

$$m_{\text{T}} = \sqrt{2p_{\text{T}}^{\text{lep}} E_T^{\text{miss}} - 2\boldsymbol{p}_{\text{T}}^{\text{lep}} \cdot \boldsymbol{p}_{\text{T}}^{\text{miss}}} \tag{6.5}$$

where $\boldsymbol{p}_{\text{T}}^{\text{miss}}$ is the bi-dimensional vector whose magnitude is the E_T^{miss} (see Equation 3.4), while $\boldsymbol{p}_{\text{T}}^{\text{lep}}$ and $p_{\text{T}}^{\text{lep}}$ are the two-vector with the transverse components of the lepton momentum and the associated magnitude. The m_{T} variable has a kinematical endpoint at a mass m_X in single-lepton events where the lepton and the entire $\boldsymbol{p}_{\text{T}}^{\text{miss}}$ are produced by the decay of the particle X. A particularly important case corresponds to events with a single leptonically-decaying W boson produced in association with visible objects (for example in W+jets or semi-leptonic $t\bar{t}$ events) resulting in an endpoint of m_{T} in correspondence with the W mass.

$m_{b\ell}^{\text{min}}$ In events with two b-jets and one lepton, the minimum invariant mass of one of the b-jets and the lepton is defined as:

$$m_{b\ell}^{\text{min}} = \text{min}_{i=1,2} \left(m_{\ell b_i} \right). \tag{6.6}$$

The remarkable feature of $m_{b\ell}^{\text{min}}$ is that it has a kinematical limit given by $\sqrt{m_t^2 - m_W^2}$ when computed in $t\bar{t}$ events.

$m_{\text{T}}^{\text{min}}(j_{\text{1-4}}, E_T^{\text{miss}})$ and $m_{\text{T}}^{\text{min}}(b_{\text{1-2}}, E_T^{\text{miss}})$ These two variables are defined as the minimum transverse masses of the E_T^{miss} and two types of jet collections in the event: $m_{\text{T}}^{\text{min}}(j_{\text{1-4}}, E_T^{\text{miss}})$ employs the four leading jets, while in $m_{\text{T}}^{\text{min}}(b_{\text{1-2}}, E_T^{\text{miss}})$ the two leading b-jets are used. These variables have a kinematical endpoint in events where one of the jets and the $\boldsymbol{p}_{\text{T}}^{\text{miss}}$ originate from the decay of an on-shell particle such as a top quark.

m_{CT} The contransverse mass (m_{CT}) [216] is a key discriminant in the zero-lepton channel of the analysis, which is optimised for the $bb + E_T^{\text{miss}}$ signal in Figure 6.1a (see Section 6.2.3). The contransverse mass targets processes where two pair-produced

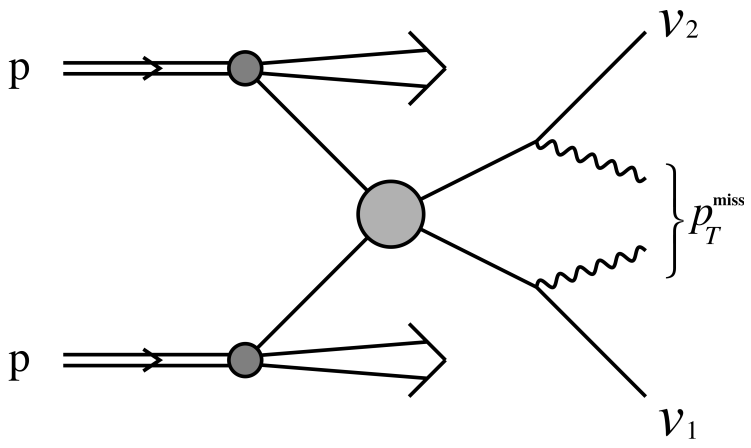


Figure 6.4: Diagram of a generic process targeted by the contransverse mass m_{CT} (Equation 6.7) and the stransverse mass m_{T2} (Equation 6.8), where pair-produced particles decay into visible legs (v_1 and v_2) and invisible objects [217].

particles decay into a semi-invisible final state, as described in Figure 6.4. It is computed using the four-momenta of the visible decay products v_1 and v_2 of the two particles, and its value is given by:

$$m_{\mathrm{CT}}^2(v_1, v_2) = [E_{\mathrm{T}}(v_1) + E_{\mathrm{T}}(v_2)]^2 - [\boldsymbol{p}_{\mathrm{T}}(v_1) - \boldsymbol{p}_{\mathrm{T}}(v_2)]^2, \tag{6.7}$$

with $E_{\mathrm{T}} = \sqrt{p_{\mathrm{T}}^2 + m^2}$. When applied to physics processes as the one in Figure 6.4, m_{CT} has a kinematical endpoint given by $(m_i^2 - m_X^2)/m_i$, where i and X are the pair-produced and invisible particles respectively.

am_{T2} An additional variable that targets semi-invisibly decaying pair-produced particles is am_{T2}, which is an asymmetric variation of the stransverse mass (m_{T2}) [218, 219] adapted to the one-lepton channel of the analysis (see Section 6.2.4). The baseline definition of m_{T2} [217] is:

$$m_{\mathrm{T2}}^2(\chi) = \min_{\boldsymbol{q}_{\mathrm{T}}^{(1)} + \boldsymbol{q}_{\mathrm{T}}^{(2)} = \boldsymbol{p}_{\mathrm{T}}^{miss}} \left[\max \left\{ m_{\mathrm{T}}^2(\boldsymbol{p}_{\mathrm{T}}(v_1), \boldsymbol{q}_{\mathrm{T}}^{(1)}; \chi), m_{\mathrm{T}}^2(\boldsymbol{p}_{\mathrm{T}}(v_2), \boldsymbol{q}_{\mathrm{T}}^{(2)}; \chi) \right\} \right] \tag{6.8}$$

where the minimum is taken over a range of possible decompositions of $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{miss}}$ in two transverse vectors, interpreted as the possible transverse momenta of the invisible particles produced by the two decay legs in Figure 6.4. In Equation 6.8 the vectors $\boldsymbol{p}_{\mathrm{T}}(v_i)$ are the transverse projections of the momentum three-vectors of the visible objects, $\boldsymbol{q}_{\mathrm{T}}^{(i)}$ are the two components of $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{miss}}$ and χ is a free parameter

that represents the mass of the invisible particles (set to 0 for this search). Similarly to m_{CT}, the m_{T2} variable also yields a kinematical endpoint when computed in events with semi-invisibly decaying pair-produced particles.
In the present analysis the am_{T2} variable is computed in the one-lepton channel using three visible objects in addition to the $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{miss}}$ vector, namely two b-jets and the lepton. The purpose is to obtain a kinematical endpoint driven by the top mass in semi-leptonic $t\bar{t}$ events, where the only source of E_T^{miss} is a neutrino and the visible particles are produced in the decay of the top quarks, while signal events are expected to yield longer am_{T2} tails, due to the extra E_T^{miss} from the $\tilde{\chi}_1^0$s and the different kinematics of the b-jets and the lepton. In practice, in this configuration an algorithm must be developed to combine the two b-jets and the lepton in order to form the two visible decay legs v_1 and v_2 of the targeted processes. While it is clear that in both signal and $t\bar{t}$ events the two b-jets are produced in different decay legs, the lepton may instead originate from any of the two. The ambiguity is solved through the following sequence of steps, based on the computation of the invariant mass $m_{b\ell}(n)$ of each of the two b-jet and lepton pairs:

- If $m_{b\ell}(1)$ and $m_{b\ell}(2)$ are both larger than the top mass (taken as 170 GeV) neither of the two pairings is compatible with a top decay, so am_{T2} is not computed and the event is rejected[1].
- If $m_{b\ell}(1) < 170$ GeV and $m_{b\ell}(2) > 170$ GeV, am_{T2} is computed by defining v_1 as the sum of b_1 and the lepton, because only this pairing yields an object compatible with the final state of a top quark decay.
- Reversely, if $m_{b\ell}(1) > 170$ GeV and $m_{b\ell}(2) < 170$ GeV, am_{T2} is computed by pairing b_2 and the lepton.
- If both $m_{b\ell}(1)$ and $m_{b\ell}(2)$ are below 170 GeV, am_{T2} is computed with both pairings and the minimum value among the two is chosen.

As further shown in Section 6.2.2, the am_{T2} variable computed with the above algorithm provides a powerful tool to suppress the dominant $t\bar{t}$ background in the one-lepton channel of the analysis.

[1] As further discussed in Section 6.2.4, a requirement of $m_{b\ell}^{\mathrm{min}} < 170$ GeV is applied in all regions where am_{T2} is used.

	$bb + E_T^{\text{miss}}$ channel	$tb + E_T^{\text{miss}}$ channel
Trigger	E_T^{miss}	Single lepton
Baseline leptons (e/μ)	$= 0$	$= 1$, $p_{\text{T}} > 10$ GeV
Signal leptons (e/μ)	$= 0$	$= 1$, $p_{\text{T}} > 27$ GeV
Offline E_T^{miss}	> 250 GeV	–
Signal jets	≥ 2, $p_{\text{T}} > 20(35)$ GeV	≥ 2, $p_{\text{T}} > 35$ GeV
b-jets (77% fixed cut WP)	$= 2$	$= 2$

Table 6.1: Preliminary event selections in the zero-lepton channel and one-lepton channel of the analysis, targeting the $bb + E_T^{\text{miss}}$ and $tb + E_T^{\text{miss}}$ final states.

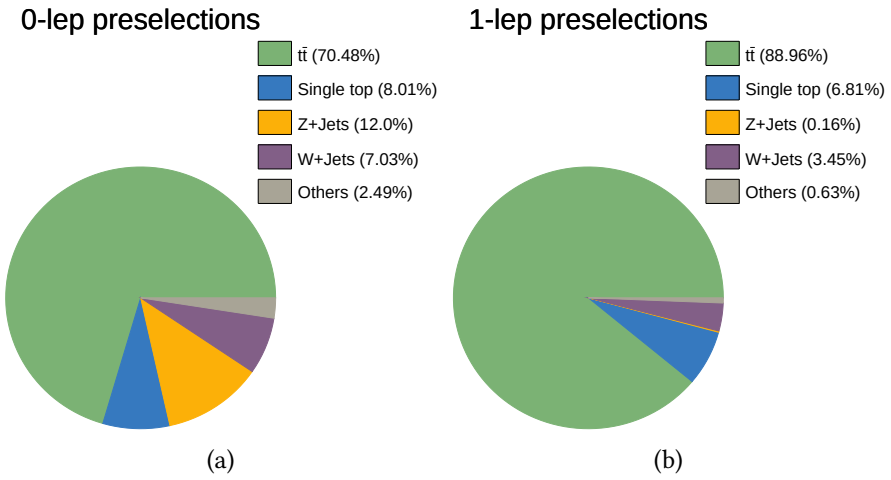
6.2.2 Preliminary selections and key distributions

As anticipated, the analysis in this chapter is divided in two main channels designed to target the $bb + E_T^{\text{miss}}$ final state in Figure 6.1a and the $tb + E_T^{\text{miss}}$ final state in Figure 6.1b. The preliminary event selections used in each channel are described in the list below, summarised in Table 6.1:

Zero-lepton channel The $bb + E_T^{\text{miss}}$ final state is selected with a E_T^{miss} trigger chain, as described in Section 5.3.1, and by vetoing the presence of baseline leptons (with $p_{\text{T}} > 10$ GeV) in the events. A least two signal jets are required, with $p_{\text{T}} > 35$ GeV or > 20 GeV for a specific SR (see Section 6.2.3), and exactly two of them are required to be b-tagged using the 77% fixed cut WP. An offline requirement of $E_T^{\text{miss}} > 250$ GeV is applied to select the efficiency plateau of the E_T^{miss} trigger.

One-lepton channel The $tb + E_T^{\text{miss}}$ final state may yield both zero-lepton or one-lepton events, but in the present analysis it was chosen to concentrate on the one-lepton channel, following the strategy of the Run 1 search [183]. The events are selected with single lepton (electron or muon) triggers, and by requiring the presence of exactly one baseline lepton that must also satisfy the signal requirements and have $p_{\text{T}} > 27$ GeV. The selections on jets and b-tagging are equivalent to the zero-lepton channel, and a signal jet p_{T} threshold of 35 GeV is used in all one-lepton regions.

Figure 6.5 shows the MC breakdown of the SM background yields in the zero-lepton and one-lepton channels, after the preliminary event selections described above. The dominant background in both cases is $t\bar{t}$ production, which is enhanced with respect



(a) (b)

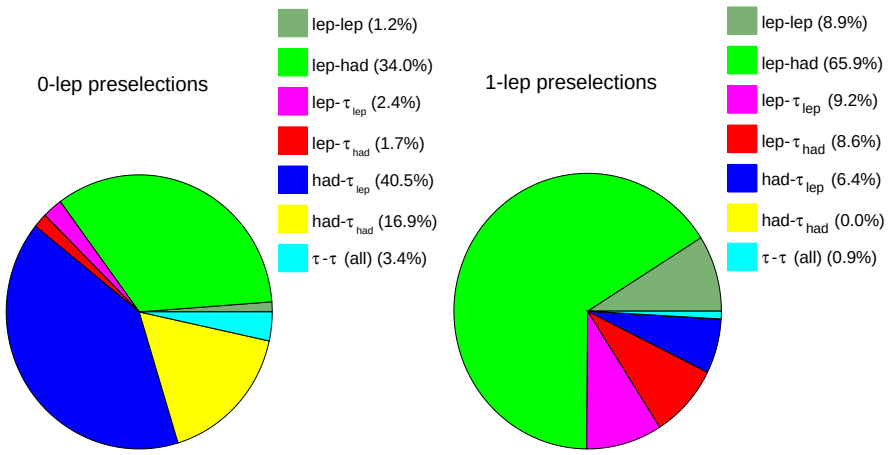
Figure 6.5: Background composition after the preliminary events selections described in Table 6.1 in the zero-lepton (a) and one-lepton (b) channels, determined using MC simulations.

to the other processes due to the requirement of two b-tagged jets. It is important to recall that the fully hadronic decay of $t\bar{t}$ is not included in the MC samples, because it is expected to be negligible in channels with leptons or large E_T^{miss} (see Section 5.2.1), so the fractions in Figure 6.5 represent only events with at least one lepton[2]. Remarkably, semi-leptonic $t\bar{t}$ accounts for 70% of the total rate even in the zero-lepton channel, despite the fact that a baseline lepton veto is applied.

Further information about the $t\bar{t}$ background is provided by Figure 6.6, which shows the breakdown of its rate into individual sub-processes based on the type of decay of the two top quarks: hadronic and leptonic decays of the quarks (see the $t\bar{t}$ diagram in Figure 5.11c) are classified separately, and the leptonic decays are split into light (e/μ) and τ leptons, which are further divided into hadronic or leptonic τ decays.

In the zero-lepton channel (Figure 6.6a) the dominant component of the $t\bar{t}$ rate consists of events with one τ (with either leptonic or hadronic decay), while semi-leptonic (e/μ) events represent a sub-dominant fraction of its yield. This is a result of a combined effect of the baseline lepton veto and the large E_T^{miss} requirement, that have a different impact on the various components of the original $t\bar{t}$ sample. The semi-leptonic (lep-had) and di-leptonic (lep-lep) events are strongly suppressed by the lepton veto, which is instead

[2]While in general only electron and muons are referred to as reconstructed leptons in ATLAS, lepton-filtered MC samples include also events where the top decays into τ leptons, which further decay either hadronically or leptonically.



(a) (b)

Figure 6.6: Breakdown of the $t\bar{t}$ rate into different types of decay of the two top quarks after the zero-lepton (a) and one-lepton (b) preselections, determined using the truth information from the MC samples. The leptonic decays include electrons or muons only, while τ decays are further split into hadronically-decaying (τ_{had}) and leptonically-decaying (τ_{lep}), except for the di-τ channel which includes both components.

less effective in had-τ_{lep} events where the average p_{T} of the lepton is lower and the E_T^{miss} is larger due to the presence of two extra neutrinos in the final state. In parallel, the had-τ_{had} events survive the lepton veto but are more suppressed than the had-τ_{lep} ones by the E_T^{miss} selection.
Finally, the one-lepton channel of the analysis (Figure 6.6b) is dominated as expected by semi-leptonic $t\bar{t}$, while a smaller fraction of events contain an additional electron, muon or τ lepton.

Key discriminating variables in zero-lepton and one-lepton channels

Having shown the composition of the background rates after the zero-lepton and one-lepton preselections, it is clear that the discriminating variables designed to suppress $t\bar{t}$ events play a crucial role in both channels of the analysis. Figure 6.7 shows the kinematical distributions of four major variables defined in Section 6.2.1, comparing the shape in signal and in the most relevant background processes.
The first variable under exam is m_{CT} (Figure 6.7a) which is shown in the zero-lepton channel for $t\bar{t}$, Z+jets and a $bb + E_T^{\mathrm{miss}}$ signal with $m_{\tilde{b}} = 600$ GeV and $m_{\tilde{\chi}_1^0} = 100$ GeV. As expected, the shape of m_{CT} is peaked at low values for $t\bar{t}$, which is a semi-invisibly

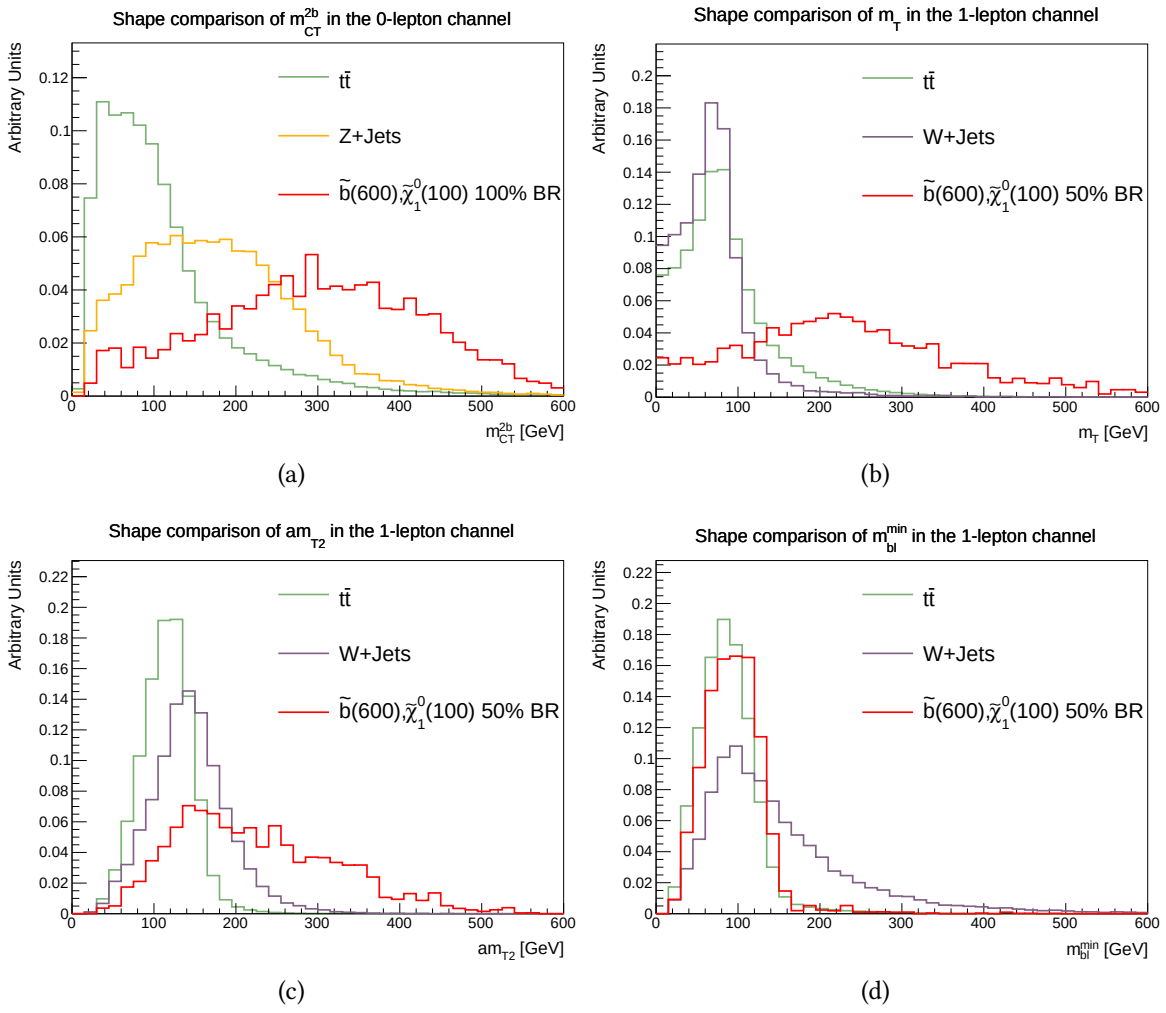


Figure 6.7: Distributions of key variables after the preliminary selections in Table 6.1: m_{CT} (a) is shown in the zero-lepton channel for $t\bar{t}$, Z+jets and a $bb + E_T^{\mathrm{miss}}$ signal with $(m_{\tilde{b}},m_{\tilde{\chi}_1^0}) = (600, 100)$ GeV, while m_{T} (b), am_{T2} (c) and $m_{b\ell}^{\mathrm{min}}$ (d) are shown in the one-lepton channel for $t\bar{t}$, W+jets and a $tb + E_T^{\mathrm{miss}}$ signal with the same masses.

decaying pair-produced background, while the distribution extends to larger values in signal and in the Z+jets background. This clearly suggests that a selection on m_{CT} can be used to reduce drastically the $t\bar{t}$ background, adding also a non-negligible suppression of Z+jets, provided that the targeted signal has a sufficiently large $\Delta m(\tilde{b}, \tilde{\chi}_1^0)$. Figures 6.7b and 6.7c show a similar situation in the one-lepton channel, where the distributions of m_{T} and am_{T2} are compared in an asymmetric $tb + E_T^{\mathrm{miss}}$ signal and in the $t\bar{t}$ and W+jets backgrounds, highlighting their strong discriminating power. In particular, am_{T2} is designed to have a kinematical endpoint for $t\bar{t}$, while m_{T} has it for both

background processes. Finally, Figure 6.7d shows the distribution of the invariant mass variable $m_{b\ell}^{\min}$, which is bounded below the mass of the top quark for both signal and $t\bar{t}$ background, while no such constraint is present in W events.

Classification of $t\bar{t}$ events beyond the kinematical endpoints

While a major fraction of $t\bar{t}$ events can be removed by applying appropriate selections on the discriminating variables in Figure 6.7, the distributions are also showing the presence of non-negligible tails which extend beyond the expected endpoints. Due to the large overall rate of $t\bar{t}$ after the preliminary selections of the analysis, these tails represent a significant source of background in the SRs (see Sections 6.2.3 and 6.2.4), so it is important to understand the mechanisms through which they are generated. A first answer comes from the finite resolution of the detector, that produces fluctuations in the measurement of jets and lepton momenta and E_T^{miss}, but there are also other effects that are particularly relevant for the analysis in this chapter.

An interesting insight can be obtained by studying the real flavour of the two b-tagged jets in $t\bar{t}$ events, that serve as input for the computation of the m_{CT} and am_{T2} variables, comparing the composition after the standard preselections and in the tails of the two distributions. The charts in Figure 6.8 show the breakdown of the $t\bar{t}$ rate into each possible combination of light, c, b or hadronic τ jet flavours, resulting in 10 possible configurations of the events. As expected, the flavour composition after the zero-lepton and one-lepton preselections (Figures 6.8a and 6.8b) is dominated by real b-jets, with a subdominant fraction of cases where one mis-tag is present. On the contrary, when events in the m_{CT} and am_{T2} tails are selected (Figures 6.8c and 6.8d), the fraction of cases with at least one mis-tag (mainly from c-jets) increases dramatically, indicating that at least one of the two jets used for the computation of m_{CT} or am_{T2} is not a b-jet from the decay of the top quark. This result is not surprising, because the kinematical endpoints of m_{CT} and am_{T2} should only be present if the visible decay legs of $t\bar{t}$ are correctly reconstructed, so the mis-tag of jets increases the likelihood of populating the tails of the distributions of both variables.

Additional information can be obtained by analysing the type of decay of the two top quarks in $t\bar{t}$ events, as already done in Figure 6.6. The breakdown of the $t\bar{t}$ rate is shown in Figures 6.9a and 6.9b in the tails of the m_{T} and am_{T2} distributions respectively, to be compared with the one-lepton preselection chart in Figure 6.6b. In case of m_{T}, the dominant component of the rate beyond the kinematical endpoint consists of events with

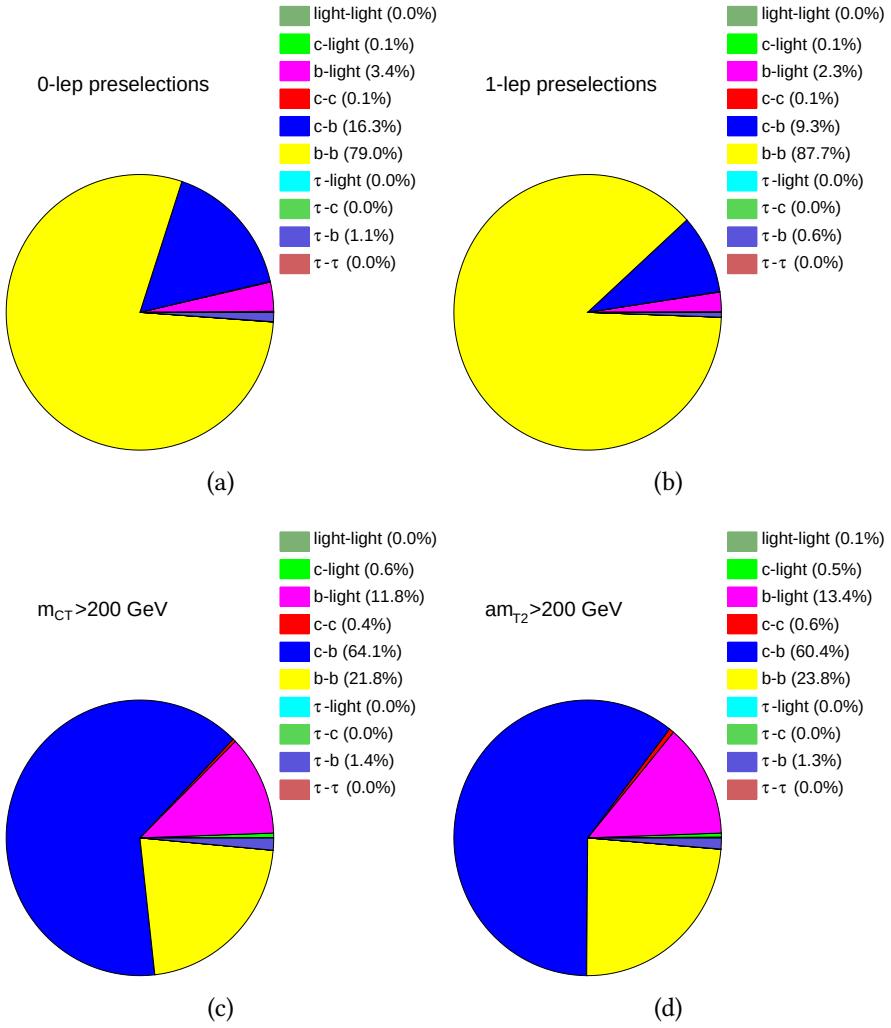


Figure 6.8: Breakdown of the $t\bar{t}$ rate into each possible combination of light, c, b or hadronic τ flavours of the two b-tagged jets, after the preliminary event selections in Table 6.1 in the zero-lepton (a) and one-lepton (b) channels, and with additional selections on $m_{\mathrm{CT}} > 200$ GeV (c) and $am_{\mathrm{T2}} > 200$ GeV (d) respectively.

one lepton (e or μ) plus a second lepton that can either be a τ (with hadronic or leptonic decay) or a second electron or muon which escapes the detection. This happens because the presence of an extra leptonic decay of the top is an additional source of E_T^{miss} in the event, implying that the transverse mass m_{T} is not necessarily bound below the W mass. On the other hand, the top decay composition beyond the am_{T2} endpoint remains similar to the inclusive preselection chart, indicating that the presence of additional leptons in the events is not significantly enhancing the am_{T2} tails.

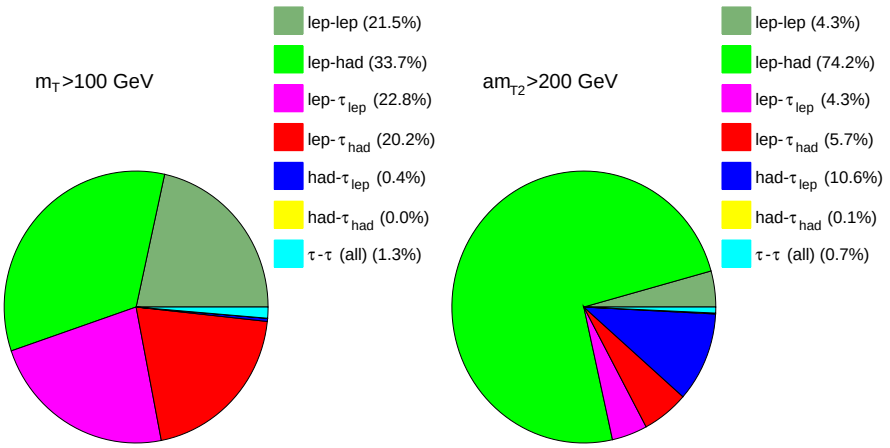


Figure 6.9: Breakdown of the $t\bar{t}$ rate into different types of decay of the two top quarks after the one-lepton preselections plus $m_{\mathrm{T}} > 100$ GeV (a) and $am_{\mathrm{T2}} > 200$ GeV (b), to be compared with the preselection composition in Figure 6.6b.

	channel	SR selections	Dominant source of tails in $t\bar{t}$
m_{CT}	zero-lepton	Lower bound	Mis-tag of c- or light jets
am_{T2}	one-lepton	Lower bound	Mis-tag of c- or light jets
m_{T}	one-lepton	Lower bound	Di-leptonic events
$m_{b\ell}^{\mathrm{min}}$	one-lepton	Upper bound	Not relevant

Table 6.2: Summary of the behaviour of the m_{CT}, am_{T2}, m_{T} and $m_{b\ell}^{\mathrm{min}}$ variables in $t\bar{t}$ events, highlighting the dominant source of kinematical tails in their distributions when relevant for the analysis in this chapter.

A summary of the analysis in this paragraph is presented in Table 6.2. The mis-tag of b-jets is identified as the dominant source of m_{CT} and am_{T2} tails in $t\bar{t}$ events in the zero-lepton and one-lepton channels, while the m_{T} tails are found to be generated mainly by di-leptonic $t\bar{t}$ events. These three variables play a crucial role in the definition of the SRs of the analysis, as further discussed in the next sections. Finally, the $m_{b\ell}^{\mathrm{min}}$ variable is peaked below the top quark mass in both signal and $t\bar{t}$ background, and the kinematical tails beyond the end-points are not a significant concern for the present analysis, because they are excluded from the SRs.

6.2.3 Zero-lepton channel SRs

The first channel of the analysis is designed to target the $bb + E_T^{\text{miss}}$ signal in Figure 6.1a in zero-lepton events, based on the preliminary selections in Table 6.1. In order to be sensitive to signals over a wide range of $\tilde{b}$ and $\tilde{\chi}_1^0$ masses, three mutually exclusive sets of SRs are defined, and the detail of the corresponding selections is presented in Table 6.3. Each SR is defined following the procedure described in Section 5.3.5, which includes a dedicated optimisation scan on the relevant set of discriminating variables.
A description of each type of SR is provided below:

b0L-SRAx Type-A SRs are designed to target signal models with large mass splittings between $\tilde{b}$ and $\tilde{\chi}_1^0$. The key discriminating variable is m_{CT}, whose properties have been discussed in detail in the previous paragraphs, and three overlapping SRs are defined by applying progressive cut thresholds ($m_{\text{CT}} > 350$, 450 and 550 GeV) indicated by the "x" label in the region name. The baseline E_T^{miss} threshold of 250 GeV is used to ensure full trigger efficiency, and the b-tagging requirement is applied to the two leading jets in p_{T}, because the b-jets are expected to be boosted in models with large $\Delta m(\tilde{b}, \tilde{\chi}_1^0)$. Signal jets are defined with p_{T} threshold of 35 GeV, and background processes with large jet multiplicities, such as $t\bar{t}$, are suppressed by selecting only events with total number of jets between 2 and 4, and with p_{T} of the fourth jet below 50 GeV. In order to reduce the impact of the multi-jet background, events where any signal jet is aligned with the E_T^{miss} in the transverse plane are removed by requiring $\Delta\phi_{\text{min}}^{j_{1\text{-}4}} > 0.4$, and an additional cleaning selection of $E_T^{\text{miss}}/m_{\text{eff}} > 0.25$ is included. Finally, a selection on $m_{\text{bb}} > 200$ GeV is applied in all SRs to reduce the background from Z+jets events with two b-jets produced by gluon splitting.

b0L-SRB A single SR of type B is designed for signals with intermediate mass splittings between $\tilde{b}$ and $\tilde{\chi}_1^0$, targeting the $\Delta m(\tilde{b}, \tilde{\chi}_1^0)$ range between 50 and 250 GeV. In these scenarios the discriminating power of m_{CT} and m_{bb} is limited, so these variables are replaced by $m_{\text{T}}^{\text{min}}(j_{1\text{-}4}, E_T^{\text{miss}})$ with a minimum threshold of 250 GeV, which provides a significant suppression of the $t\bar{t}$ background. The selection on the jet multiplicity is equivalent to the b0L-SRAs, with 2 to 4 jets with $p_{\text{T}} > 35$ GeV, but no upper bound is applied to the p_{T} of the fourth jet. Similarly, b0L-SRB events are selected with exactly two b-tags, but the two b-jets are not necessarily required to be leading in p_{T}. Finally, the baseline E_T^{miss} threshold of 250 GeV is used, and

dedicated selections on the azimuthal distance between the two b-jets and the E_T^{miss} are applied to reduce the background from Z+jets.

b0L-SRC Lastly, a SR of type C is optimised for compressed mass splittings between $\tilde{b}$ and $\tilde{\chi}_1^0$, characterised by $\Delta m(\tilde{b}, \tilde{\chi}_1^0) < 50$ GeV, where the b-quarks from the decay of the $\tilde{b}$ are produced with low p_{T}. In these conditions the variables computed from the four-momenta of the b-jets are less effective, so the signal is targeted by requiring a jet from ISR similarly to what is done in the mono-jet analyses [206, 212]. In particular, the selected events must contain a non b-tagged jet with $p_{\text{T}} > 500$ GeV, which recoils against a signal-like system yielding $E_T^{\text{miss}} > 500$ GeV (from the heavy neutralinos escaping the detection) and two soft b-jets, whose p_{T} threshold is lowered to 20 GeV to increase the signal acceptance. A large azimuthal distance between the ISR jet and the E_T^{miss} is selected by requiring $\Delta\phi(j_1, E_T^{\text{miss}}) > 2.5$, while a loose extra cleaning requirement of $\Delta\phi(j_2, E_T^{\text{miss}}) > 0.2$ is applied to minimise the impact on the acceptance. The number of signal jets with $p_{\text{T}} > 20$ GeV is required to be between 2 and 5, and an extra requirement of $H_{\text{T4}} < 70$ GeV is applied. Finally, a tight $m_{\text{eff}} > 1.3$ TeV cut is applied in combination with a large asymmetry $\mathcal{A} > 0.8$, to select signal-like events where any jet in addition to the ISR one is relatively soft.

The background composition in three of the zero-lepton SRs is shown in Figure 6.10 for the type A region with the lowest m_{CT} threshold (b0L-SRA350) and for the type B and C regions. While the event yield after the preliminary selections in Table 6.1 is dominated by $t\bar{t}$, the most relevant background in all SRs is Z+jets, whose experimental signature is less distinguishable from the signal processes. This is true in particular for the b0L-SRAx regions, thanks to the good performance of the m_{CT} variable in suppressing the $t\bar{t}$ background. The detail of the rate predictions in all SRs is shown in Section 6.4, where the full results of the analysis are presented.

The most effective SR for each signal benchmark in the $m_{\tilde{b}}$-$m_{\tilde{\chi}_1^0}$ plane is shown in Figure 6.11, superimposed to the expected exclusion limits at 95% CL obtained from the fit procedure described in Section 6.4. As expected, the b0L-SRAs have the best performance in the region with large mass splitting between the $\tilde{b}$ and the $\tilde{\chi}_1^0$, with larger m_{CT} thresholds for larger values of $\Delta m(\tilde{b}, \tilde{\chi}_1^0)$, while b0L-SRB dominates in the intermediate region and b0L-SRC is optimal in the compressed scenarios.

Finally, Table 6.4 shows the lists of the expected MC signal yields after each selection of b0L-SRA, b0L-SRB and b0L-SRC respectively, commonly referred to as cutflow tables,

	b0L-SRAx	b0L-SRB	b0L-SRC
Lepton veto	No e/μ with $p_{\mathrm{T}} > 10$ GeV after OR		
N_{jets} ($p_{\mathrm{T}} > 35$ GeV)	2–4	2–4	-
N_{jets} ($p_{\mathrm{T}} > 20$ GeV)	-	-	2–5
$p_{\mathrm{T}}(j_1)$ [GeV]	> 130	> 50	> 500
$p_{\mathrm{T}}(j_2)$ [GeV]	> 50	> 50	> 20
$p_{\mathrm{T}}(j_4)$ [GeV]	< 50	-	-
H_{T4} [GeV]	-	-	< 70
b-jets	j_1 and j_2	any 2	j_2 and (j_3 or j_4 or j_5)
E_T^{miss} [GeV]	> 250	> 250	> 500
$E_T^{\mathrm{miss}}/m_{\mathrm{eff}}$	> 0.25	-	-
$\Delta\phi_{\mathrm{min}}^{j_{1\text{-}4}}$	> 0.4	> 0.4	-
$\Delta\phi(b_1, E_T^{\mathrm{miss}})$	-	< 2.0	-
$\Delta\phi(b_2, E_T^{\mathrm{miss}})$	-	< 2.5	-
$\Delta\phi(j_1, E_T^{\mathrm{miss}})$	-	-	> 2.5
$\Delta\phi(j_2, E_T^{\mathrm{miss}})$	-	-	> 0.2
m_{jj} [GeV]	> 200	-	> 200
m_{CT} [GeV]	$> 350, 450, 550$	-	-
$m_{\mathrm{T}}^{\mathrm{min}}(j_{1\text{-}4}, E_T^{\mathrm{miss}})$ [GeV]	-	> 250	-
m_{eff} [GeV]	-	-	> 1300
$\mathcal{A}$	-	-	> 0.8

Table 6.3: Summary of the selections of each signal region in the zero-lepton channel of the analysis. The "x" in the b0L-SRA labels represents the associated m_{CT} threshold, and both light and b-tagged jets are labelled in decreasing order in p_{T}.

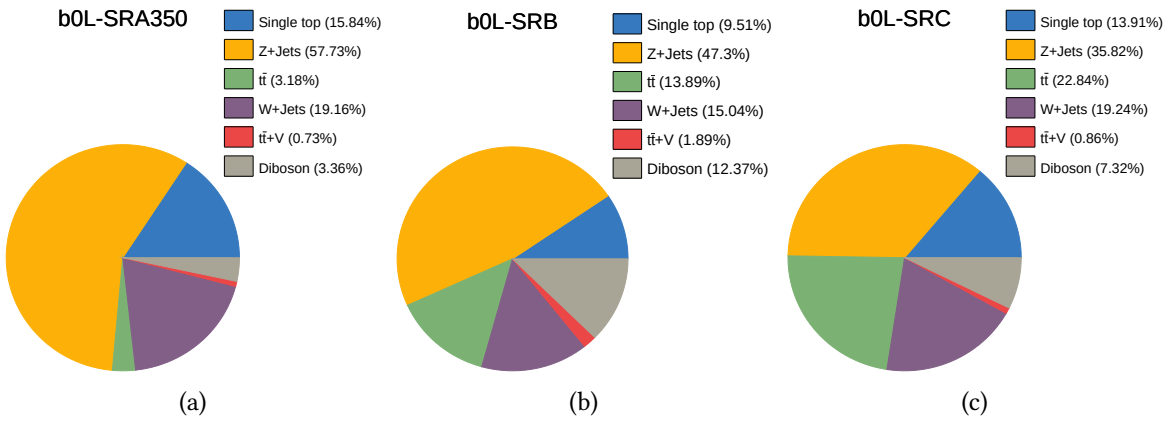


Figure 6.10: Background composition in three b0L SRs defined in Table 6.3: b0L-SRA350 (a), b0L-SRB (b) and b0L-SRC (c). The rates are predicted using MC simulations.

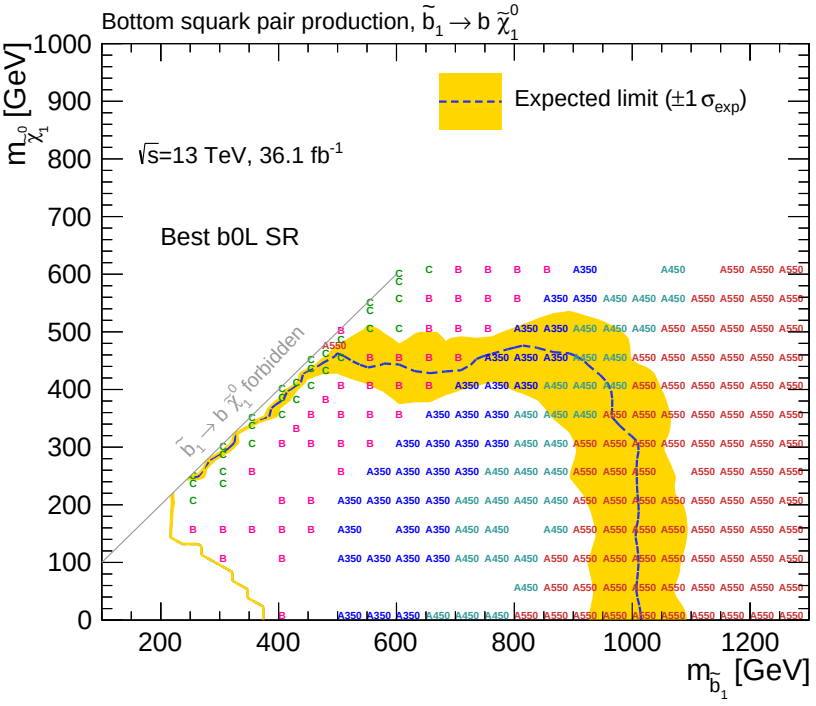


Figure 6.11: Expected exclusion limits of the zero-lepton SRs on a $bb + E_T^{\text{miss}}$ signal scenario in the $m_{\tilde{b}}$-$m_{\tilde{\chi}_1^0}$ plane, obtained from the fit procedure described in Section 6.4. The SR with the best expected CL$_s$ is indicated in the plot for every signal model considered.

for three reference $bb + E_T^{\text{miss}}$ signals with $(m_{\tilde{b}}, m_{\tilde{\chi}_1^0}) = (1000, 1)$ GeV, $(m_{\tilde{b}}, m_{\tilde{\chi}_1^0}) = (700, 450)$ GeV and $(m_{\tilde{b}}, m_{\tilde{\chi}_1^0}) = (450, 430)$ GeV respectively. The cutflows confirm that the overall selection efficiency (defined as the ratio of the signal yield in the SR and the signal yield in the whole dataset) of the b0L-SRAx regions for a signal with large $\Delta m(\tilde{b}, \tilde{\chi}_1^0)$ is significantly higher compared to the corresponding selection efficiencies of b0L-SRB and b0L-SRC for signals with intermediate and small $\Delta m(\tilde{b}, \tilde{\chi}_1^0)$. This is consistent with the pattern observed in Figure 6.11, where it is shown that the sensitivity of the b0L-SRAx regions reaches heavier $\tilde{b}$ signals with lower cross sections.

6.2.4 One-lepton channel SRs

The one-lepton channel of the analysis targets the $tb + E_T^{\text{miss}}$ final state obtained from the signal model in Figure 6.1b, based on the corresponding preliminary selections defined in Table 6.1. The selection criteria, summarised in Table 6.5, follow the same strategy of the Run 1 search described in Section 6.1.3, where minor excesses over the SM predictions were observed.

b0L-SRAx, $(m_{\tilde{b}}, m_{\tilde{\chi}_1^0}) = (1000, 1)$ GeV		b0L-SRB, $(m_{\tilde{b}}, m_{\tilde{\chi}_1^0}) = (700, 450)$ GeV		b0L-SRC, $(m_{\tilde{b}}, m_{\tilde{\chi}_1^0}) = (450, 430)$ GeV	
Total (36.1 fb^{-1})	222.1	Total (36.1 fb^{-1})	2420	Total (36.1 fb^{-1})	34235
Production filter	210.6 (94.9%)	Production filter	2216 (91.6%)	Gen. filt. ($E_T^{\text{miss}} > 200$ GeV)	5238 (15.3%)
E_T^{miss} trigger	204.1 (96.9%)	E_T^{miss} trigger	1958 (88.4%)	Production filter	3181 (60.7%)
Event cleaning	202.2 (99.1%)	Event cleaning	1934 (98.8%)	E_T^{miss} trigger	3180 (99.9%)
$E_T^{\text{miss}} > 250$ GeV	177.5 (87.8%)	$E_T^{\text{miss}} > 250$ GeV	1048 (54.2%)	Event cleaning	3143 (98.9%)
$\Delta\phi_{\text{min}}^{j_{1\text{-}4}} > 0.4$	146.6 (82.6%)	$\Delta\phi_{\text{min}}^{j_{1\text{-}4}} > 0.4$	919.6 (87.7%)	0 leptons	3105 (98.8%)
$E_T^{\text{miss}}/m_{\text{eff}} > 0.25$	136.1 (92.8%)	2-4 jets ($p_{\text{T}} > 35$ GeV)	771.6 (83.9%)	2-5 jets ($p_{\text{T}} > 20$ GeV)	2519 (81.1%)
2-4 jets ($p_{\text{T}} > 35$ GeV)	109.0 (80.1%)	$p_{\text{T}}(j_2) > 50$ GeV	750.7 (97.3%)	Leading light jet	2346 (93.1%)
$p_{\text{T}}(j_1) > 130$ GeV	108.9 (99.9%)	0 leptons	745.5 (99.3%)	$\Delta\phi(b_1, E_T^{\text{miss}}) > 2.5$	2247 (95.8%)
$p_{\text{T}}(j_2) > 50$ GeV	108.1 (99.2%)	2 b-jets	374.4 (50.2%)	$\Delta\phi(b_2, E_T^{\text{miss}}) > 0.2$	2101 (93.5%)
$p_{\text{T}}(j_4) < 50$ GeV	90.2 (83.5%)	$\Delta\phi(b_1, E_T^{\text{miss}}) < 2.0$	49.2 (13.1%)	Subleading jet b-tagged	696.9 (33.2%)
0 leptons	89.6 (99.3%)	$\Delta\phi(b_2, E_T^{\text{miss}}) < 2.5$	31.9 (64.8%)	2 b-jets	159.6 (22.9%)
2 b-jets	28.3 (31.6%)	$m_{\text{T}}^{\text{min}}(j_{1\text{-}4}, E_T^{\text{miss}}) > 250$ GeV	15.0 (47.1%)	$H_{\text{T4}} < 70$ GeV	154.6 (96.9%)
2 leading b-jets	23.6 (83.4%)			$E_T^{\text{miss}} > 500$ GeV	23.9 (15.5%)
$m_{\text{bb}} > 200$ GeV	23.1 (97.8%)			$p_{\text{T}}(j_1) > 500$ GeV	21.5 (90.1%)
$m_{\text{CT}} > 350$ GeV	16.2 (70.2%)			$m_{\text{eff}} > 1300$ GeV	14.5 (67.4%)
$m_{\text{CT}} > 450$ GeV	12.3 (75.9%)			$\mathcal{A} > 0.8$	11.5 (79.5%)
$m_{\text{CT}} > 550$ GeV	9.0 (73.4%)			$m_{\text{jj}} > 200$ GeV	11.2 (97.1%)

Table 6.4: b0L SR cutflows, showing the nominal signal yield after each step of the selection sequence and the corresponding relative efficiency [220].

	b1L-SRAx	b1L-SRA300-2j	b1L-SRB
Number of leptons (e, μ)	1	1	1
N_{jets} ($p_{\text{T}} > 35$ GeV)	≥ 2	= 2	≥ 2
b-jets	any 2	j_1 and j_2	any 2
E_T^{miss} [GeV]	> 200	> 200	> 200
$E_T^{\text{miss}}/\sqrt{H_{\text{T}}}$ [GeV$^{1/2}$]	> 8	> 8	> 8
$m_{b\ell}^{\text{min}}$ [GeV]	< 170	< 170	< 170
$\Delta\phi_{\text{min}}^{j}$	> 0.4	–	> 0.4
$\Delta\phi_{\text{min}}^{j_{1\text{-}2}}$	–	> 0.4	–
am_{T2} [GeV]	> 250	> 250	> 200
m_{T} [GeV]	> 140	> 140	> 120
m_{bb} [GeV]	> 200	> 200	< 200
m_{eff} [GeV]	$> 600, 750$	> 300	> 300
$m_{\text{T}}^{\text{min}}(b_{1\text{-}2}, E_T^{\text{miss}})$ [GeV]	–	–	> 200
$\Delta\phi(b_1, E_T^{\text{miss}})$	–	–	> 2.0

Table 6.5: Summary of the selections of each signal region in the one-lepton channel of the analysis. The "x" in the b1L-SRA labels represents the associated m_{eff} threshold, while the -2j suffix indicates the exclusive SR in number of jets. Both light and b-tagged jets are labelled in decreasing order in p_{T}.

A description of the different types of one-lepton SRs is provided below:

b1L-SRAx The one-lepton SRs of type A are designed for signals with large $\Delta m(\tilde{b}, \tilde{\chi}_1^0)$, similarly to the b0L-SRAs in the zero-lepton channel. The key discriminating variables are m_{T} and am_{T2}, with lower thresholds of 140 GeV and 250 GeV, while an upper bound on $m_{b\ell}^{\text{min}}$ is applied (< 170 GeV) based on the kinematical distributions presented in Section 6.2.2. These selections are particularly effective against the $t\bar{t}$ background, including di-leptonic events, but they also provide a strong suppression of W+jets and single top. Additional selections on E_T^{miss}, $E_T^{\text{miss}}/\sqrt{H_{\text{T}}}$ and $\Delta\phi_{\text{min}}^{j}$ are employed to reduce the background from multi-jet events to negligible levels. A requirement of large invariant mass of the two b-jets ($m_{\text{bb}} > 200$ GeV), also used in the b0L-SRAs, provides a further enhancement of the sensitivity to the targeted signals. Two overlapping SRs are defined using incremental thresholds on m_{eff} (600 and 750 GeV), as indicated by the "x" in the region labels, and with no restriction on the maximum number of jets in the events. Together with these, a third SR of type A is defined with the same selections except for a looser m_{eff} threshold (300 GeV) and an exclusive requirement on the number of signal jets,

that must be exactly equal to two. This SR, labelled as b1L-SRA-300-2j, targets models with small $\Delta m(\tilde{b}, \tilde{\chi}_1^0)$ and is kinematically similar to a Run 1 SR defined with a maximum of two jets with $p_{\mathrm{T}} > 50$ GeV [183].

b1L-SRB A SR of type B is designed for signals with small mass splittings between the $\tilde{b}$ and the $\tilde{\chi}_1^0$, following a similar strategy to b0L-SRB in the zero-lepton channel. The selection on m_{bb} is reverted to < 200 GeV with respect to the b1L-SRAx regions, while the lower thresholds of m_{T} and am_{T2} are relaxed to 120 GeV and 200 GeV respectively. The $m_{\mathrm{T}}^{\mathrm{min}}(b_{1\text{-}2}, E_T^{\mathrm{miss}})$ variable is used as key discriminant to suppress the $t\bar{t}$ background, using a threshold of 200 GeV, and the leading b-jet in p_{T} is required to have a large azimuthal distance from the E_T^{miss} ($\Delta\phi(b_1, E_T^{\mathrm{miss}}) > 2.0$). The multi-jet background is reduced using the same selections of the b1L-SRAx regions, and a minimal m_{eff} threshold of 300 GeV is used.

Figure 6.12 shows the background composition in three one-lepton SRs, namely b1L-SRA600, b1L-SRA300-2j and b1L-SRB. Similarly to what happens in the zero-lepton channel, the fraction of $t\bar{t}$ events is significantly reduced with respect to the preliminary selections in Table 6.1. The production of single top becomes a major source of background, particularly in the SRs of type A, and a significant contribution is also given by $t\bar{t}+V$. As for the zero-lepton regions, the detail of the rate predictions in each region is presented in Section 6.4.

The best b1L SR for each asymmetric $tb + E_T^{\mathrm{miss}}$ signal in the $m_{\tilde{b}}$-$m_{\tilde{\chi}_1^0}$ plane is shown in Figure 6.13, together with the expected limits at 95% CL from the associated exclusion fit (see Section 6.4). The general structure is similar to the $bb+E_T^{\mathrm{miss}}$ limits in Figure 6.11, with the SRs of type A yielding the best performance for large mass splittings between the $\tilde{b}$ and the $\tilde{\chi}_1^0$, and b1L-SRB dominating in compressed scenarios. The exclusive b1L-SRA300-2j has a lower sensitivity to the targeted signal grid, so it is not appearing in the figure, but it is included in the analysis to provide a cross-check of the corresponding region in the Run 1 search.

The cutflow tables of the b1L SRs are shown in Table 6.6 for a $tb + E_T^{\mathrm{miss}}$ signal model with $(m_{\tilde{b}}, m_{\tilde{\chi}_1^0}) = (700, 300)$ GeV, to which all b1L SRs have a non-negligible sensitivity. The overall selection efficiencies of these SRs are comparable and are penalised by the low efficiency of the lepton trigger selection, due to the presence of a significant fraction of events with no leptons in the signal samples.

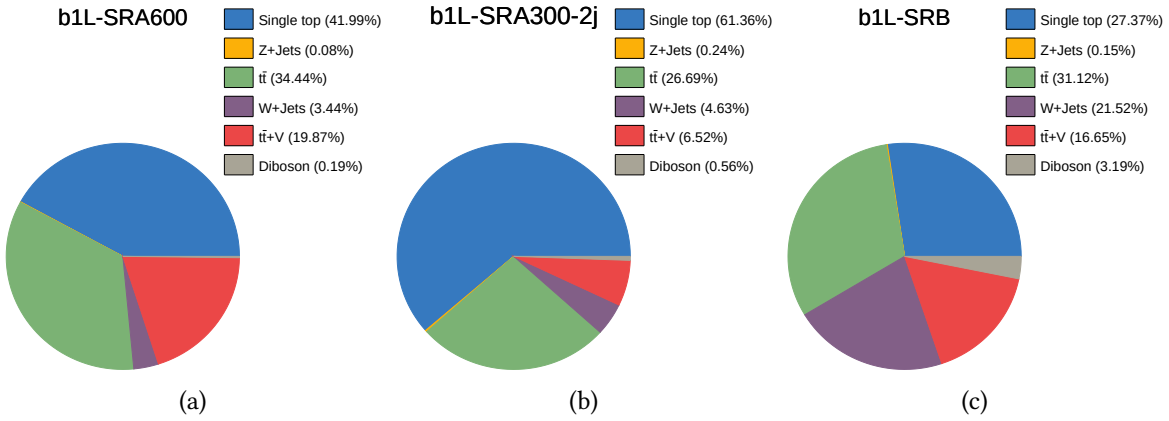


Figure 6.12: Background composition in three b1L SRs defined in Table 6.3: b1L-SRA600 (a), b1L-SRA300-2j (b) and b1L-SRB (c). The rates are predicted using MC simulations.

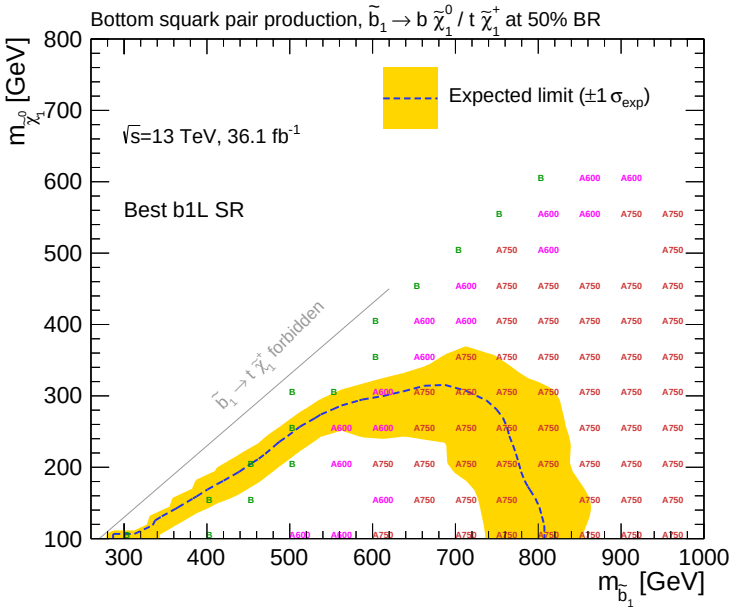


Figure 6.13: Expected exclusion limits of the one-lepton SRs on a $tb + E_T^{\text{miss}}$ signal scenario in the $m_{\tilde{b}}$-$m_{\tilde{\chi}_1^0}$ plane, obtained from the fit procedure described in Section 6.4. The SR with the best CL$_s$ is indicated in the plot for every signal model considered. The b1L-SRA300-2j region is not appearing in the figure because it is less effective than the others in the signal grid under exam.

b1L SRs cutflow, $tb + E_T^{\text{miss}}$ signal with $(m_{\tilde{b}}, m_{\tilde{\chi}_1^0}) = (700, 300)$ GeV						
Total (36.1 fb^{-1})	2420					
Production filter	2182 (90.1%)					
Lepton trigger	257.3 (11.8%)					
Event cleaning	254.6 (98.9%)					
1 lepton	235.1 (92.3%)					
$p_{\text{T}}(\ell_1) > 27$ GeV	235.1 (100.0%)					
≥ 2 jets ($p_{\text{T}} > 35$ GeV)	224.4 (95.4%)					
$\Delta\phi_{\text{min}}^{j1\text{-}4} > 0.4$	182.3 (81.2%)					
2 b-jets	73.4 (40.3%)					
$E_T^{\text{miss}} > 200$ GeV	54.1 (73.8%)					
$E_T^{\text{miss}}/\sqrt{H_{\text{T}}} > 8$ GeV$^{\frac{1}{2}}$	53.0 (97.9%)					
	b1L-SRAx (all)				b1L-SRB	
$m_{\text{T}} > 140$ GeV	44.1 (83.2%)				$m_{\text{T}} > 120$ GeV	45.5 (85.9%)
$m_{b\ell}^{\text{min}} < 170$ GeV	43.2 (97.9%)				$m_{b\ell}^{\text{min}} < 170$ GeV	44.6 (97.9%)
$am_{\text{T2}} > 250$ GeV	23.5 (54.4%)				$am_{\text{T2}} > 200$ GeV	33.1 (74.2%)
$m_{\text{bb}} > 200$ GeV	14.6 (62.3%)				$m_{\text{bb}} < 200$ GeV	13.1 (39.5%)
	b1L-SRAx (inclusive)		b1L-SRA300-2j		$\Delta\phi(b_1, E_T^{\text{miss}}) > 2.0$	12.0 (92.1%)
$m_{\text{eff}} > 450$ GeV	14.6 (100.0%)	$m_{\text{eff}} > 300$ GeV	14.6 (100.0%)		$m_{\text{T}}^{\text{min}}(b_{1\text{-}2}, E_T^{\text{miss}}) > 200$ GeV	11.4 (94.5%)
$m_{\text{eff}} > 600$ GeV	14.2 (96.9%)	< 3 jets ($p_{\text{T}} > 35$ GeV)	3.78 (25.8%)			
$m_{\text{eff}} > 750$ GeV	11.7 (82.5%)					

Table 6.6: b1L SR cutflows for a $tb + E_T^{\text{miss}}$ signal with $(m_{\tilde{b}}, m_{\tilde{\chi}_1^0}) = (700, 300)$ GeV, showing the nominal signal yield after each step of the selection sequence and the corresponding relative efficiency [220].

6.3 Background estimation

The estimate of the background yields in the SRs is based on the definition of CRs for the most relevant SM processes, and the accuracy of the predictions is tested in VRs following the strategy introduced in Section 5.3.6. A dedicated set of CRs and VRs is defined for each type of SR in each channel of the analysis, to ensure that the background estimation is carried out in regions that are kinematically close to the corresponding SR selections. The CRs associated with zero-lepton and one-lepton SRs are mutually orthogonal, with the only exception of the single top CR of type A, where the same selection is used for both channels: these choices simplify the statistical combination of the results of the individual channels of the analysis, as further discussed in Section 6.4.3. The detail of the CR and VR selections for the zero-lepton channel SRs of the analysis is described in Section 6.3.1, while the CRs and VRs for the one-lepton channel are presented in Section 6.3.4. In parallel to the CR-based approach, the Z+jets background in the zero-lepton channel is also estimated by two different data-driven methods, discussed in Section 6.3.2, while the impact of the multi-jet background is evaluated using a jet smearing technique presented in Section 6.3.3.

6.3.1 Control and Validation Regions in the zero-lepton channel

The background composition in the zero-lepton SRs from Figure 6.10 indicates that the dominant process is Z+jets, especially in the b0L-SRAx regions, but $t\bar{t}$, single top and W+jets are also significant. The background estimation strategy for the b0L-SRAx regions is based on four independent CRs for W, Z, $t\bar{t}$ and single top quark production, labelled as type-A CRs, while three CRs for W, Z and $t\bar{t}$ are employed for b0L-SRB and b0L-SRC. The detail of the CR definitions is provided in Tables 6.7, 6.8 and 6.9 for type A, B and C respectively, together with a schematic view of all the relevant regions of the channel, including VRs and SRs. The main features of the CRs for each background process are described below:

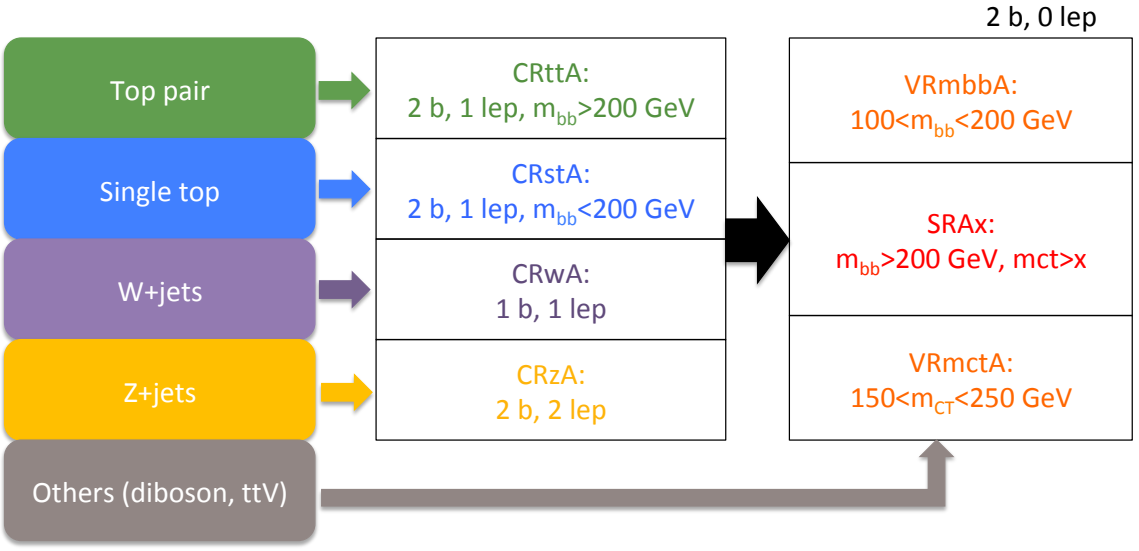
CRs for Z+jets Data samples dominated by Z+jets events are obtained by requiring two same-flavour opposite-sign (SFOS) signal leptons (electrons or muons) with invariant mass $m_{\ell\ell}$ close to the mass of the Z boson (76 GeV $< m_{\ell\ell} <$ 106 GeV). The events are selected with a single lepton trigger, and the p_{T} threshold of the reconstructed leptons is set to 27 and 20 GeV for the leading and sub-leading ones

in p_{T}. A dedicated b0L-CRz is defined for each zero-lepton SR, with specific selections on the relevant variables to reduce the extrapolation from CR to SR. A requirement of two b-tagged jets is applied in all Z CRs equivalently to the corresponding SRs, and the kinematics of the Z boson are reproduced by applying a selection on the $E_T^{\mathrm{miss,corr}}$ variable (see Section 6.2.1) which mimics the expected missing transverse momentum in $Z \to \nu\nu$ events.

CRs for $t\bar{t}$ and single top Events with one signal lepton and two b-tagged jets are selected to isolate samples of $t\bar{t}$ and single top, using the baseline single lepton trigger strategy described in Section 5.3.1. A CR for $t\bar{t}$ is implemented for all zero-lepton SRs, while a single top CR is defined only for the b0L-SRAx regions. In the latter case, the two CRs for $t\bar{t}$ and single top are separated by requiring $m_{\mathrm{bb}} > 200$ GeV and $m_{\mathrm{bb}} < 200$ GeV respectively, and the purity of the single top CR is increased by applying a selection on $m_{b\ell}^{\mathrm{min}} > 170$ GeV.

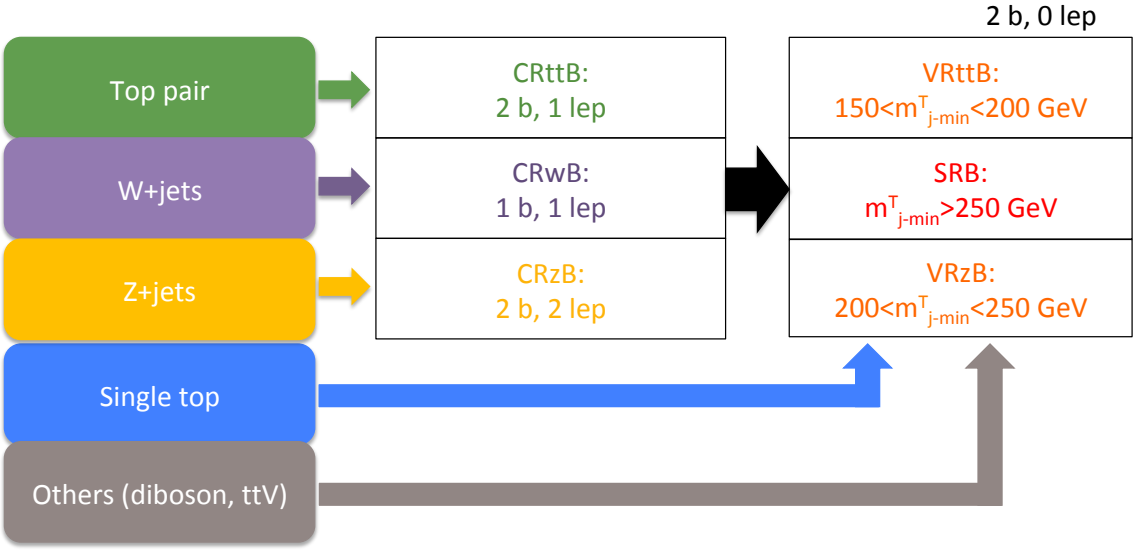
CRs for W+jets A dedicated CR for W production is defined for each zero-lepton SR by selecting events with one signal lepton, similarly to what is done for the $t\bar{t}$ and single top CRs. The peculiarity of the W CRs is that only one b-jet is required, contrarily to all the other regions of the analysis, where the b-jet multiplicity is always equal to 2. This choice allows to reduce the contamination from $t\bar{t}$ production, improving the purity of the CR. The disadvantage is that the different b-jet multiplicity increases the extrapolation uncertainty from CR to SR, introducing a significant source of systematic uncertainty related to the heavy flavour jet content in W+jets events.

The CRs presented in this paragraph are employed to constrain the normalisations of the associated backgrounds in the final fits of the analysis, described in Section 6.4. In particular, a normalisation factor is included in the fits for each background process for which a dedicated CR is defined, while the rates of the remaining backgrounds are estimated either directly from MC simulations, for dibosons or $t\bar{t}$ plus vector bosons, or from data driven techniques, for multi-jet production.
A first important test of the accuracy of the background predictions is obtained by examining the data-MC agreement of key kinematical distributions in the CRs, as shown in Figure 6.14 for a selection of variables in different b0L CRs. The normalisations of the MC samples are derived from the background only fits, and the data are following the predictions within the expected systematic uncertainties in all cases. More detail about



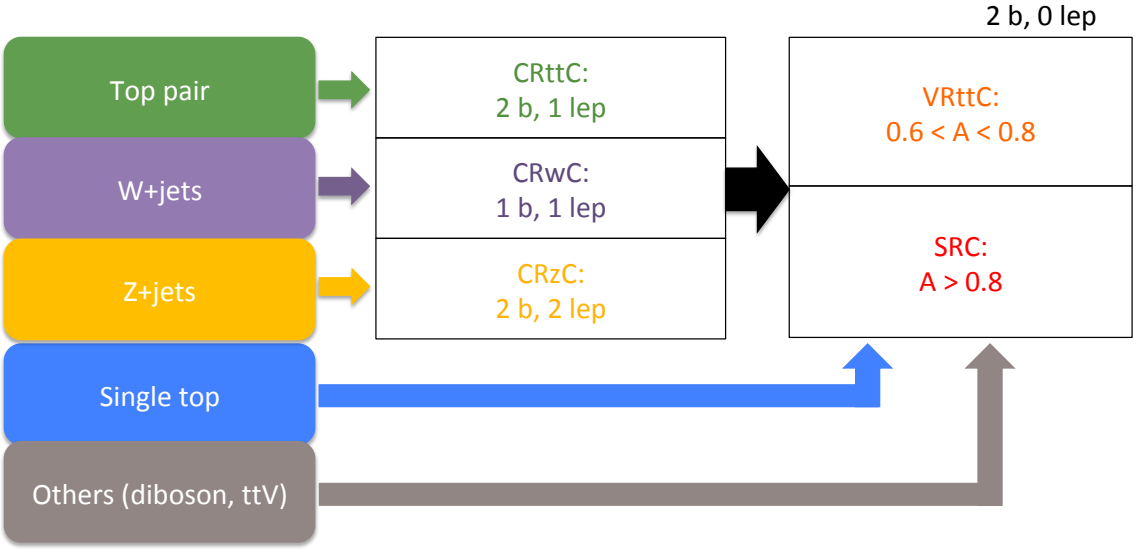
b0L-	CRzA	CRttA	CRstA	CRwA
Number of leptons ($\ell = e, \mu$)	2 SFOS	1	1	1
$p_{\mathrm{T}}(\ell_1)$ [GeV]	> 90	> 27	> 27	> 27
$p_{\mathrm{T}}(\ell_2)$ [GeV]	> 20	–	–	–
$m_{\ell\ell}$ [GeV]	[76 –106]	–	–	–
N_{jets} ($p_{\mathrm{T}} > 35$ GeV)	2–4	2–4	2–4	2–4
$p_{\mathrm{T}}(j_1)$ [GeV]	> 50	> 130	–	> 130
$p_{\mathrm{T}}(j_2)$ [GeV]	> 50	> 50	> 50	> 50
$p_{\mathrm{T}}(j_4)$ [GeV]	< 50	< 50	< 50	< 50
b-jets	j_1 and j_2	j_1 and j_2	j_1 and j_2	j_1
E_T^{miss} [GeV]	< 100	> 200	> 200	> 200
$E_T^{\mathrm{miss,corr}}$ [GeV]	> 100	–	–	–
$E_T^{\mathrm{miss}}/m_{\mathrm{eff}}$	> 0.25	> 0.25	> 0.25	> 0.25
$\Delta\phi_{\mathrm{min}}^{j_{1\text{-}4}}$	–	> 0.4	> 0.4	> 0.4
m_{T} [GeV]	–	–	–	> 30
m_{bb} [GeV]	> 200	< 200	> 200	$m_{\mathrm{bj}} > 200$
m_{CT} [GeV]	> 250	> 250	> 250	> 250
$m_{b\ell}^{\mathrm{min}}$ [GeV]	–	–	> 170	–

Table 6.7: Scheme of the background estimation strategy for the b0L-SRAx regions and summary of the CR definitions. Jets and leptons (electrons or muons) are labelled in decreasing order in p_{T}.



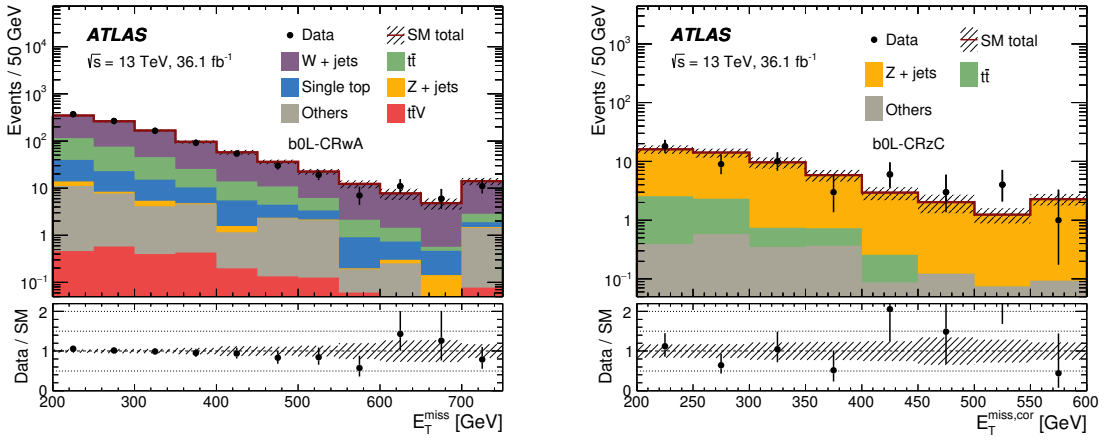
b0L-	CRzB	CRttB	CRwB
Number of leptons ($\ell = e, \mu$)	2 SFOS	1	1
$p_{\mathrm{T}}(\ell_1)$ [GeV]	> 27	> 27	> 27
$p_{\mathrm{T}}(\ell_2)$ [GeV]	> 20	–	–
$m_{\ell\ell}$ [GeV]	[76–106]	–	–
N_{jets} ($p_{\mathrm{T}} > 35$ GeV)	2–4	2–4	2–4
$p_{\mathrm{T}}(j_1)$ [GeV]	> 50	> 50	> 50
$p_{\mathrm{T}}(j_2)$ [GeV]	> 50	> 50	> 50
b-jets	any 2	any 2	any 1
E_T^{miss} [GeV]	< 100	> 100	> 100
$E_T^{\mathrm{miss,corr}}$ [GeV]	> 200	–	–
$\Delta\phi_{\mathrm{min}}^{j_{1\text{-}4}}$	> 0.4	> 0.4	> 0.4
m_{T} [GeV]	–	> 30	> 30
$m_{\mathrm{T}}^{\mathrm{min}}(j_{1\text{-}4}, E_T^{\mathrm{miss}})$ [GeV]	> 200	> 200	> 250
$\Delta\phi(b_1, E_T^{\mathrm{miss}})$	–	< 2.0	< 2.0
$\Delta\phi(b_2, E_T^{\mathrm{miss}})$	–	< 2.5	–

Table 6.8: Scheme of the background estimation strategy for b0L-SRB and summary of the CR definitions. Jets and leptons (electrons or muons) are labelled in decreasing order in p_{T}.



b0L-	CRzC	CRttC	CRwC
Number of leptons ($\ell = e, \mu$)	2 SFOS	1	1
$p_{\mathrm{T}}(\ell_1)$ [GeV]	> 27	> 27	> 27
$p_{\mathrm{T}}(\ell_2)$ [GeV]	> 20	–	–
$m_{\ell\ell}$ [GeV]	[76 –106]	–	–
N_{jets} ($p_{\mathrm{T}} > 20$ GeV)	2–5	2–5	2–5
Leading jet p_{T} [GeV]	> 250	> 500	> 500
b-jets	j_2 and (j_3 or j_4)	j_2 and (j_3 or j_4)	j_2
E_T^{miss} [GeV]	< 100	> 100	> 100
$E_T^{\mathrm{miss,corr}}$ [GeV]	> 200	–	–
m_{T} [GeV]	–	> 30	[30–120]
m_{eff} [GeV]	> 500	> 1300	> 500
m_{jj} [GeV]	> 200	> 200	> 200
H_{T4} [GeV]	< 70	< 70	< 70
$\mathcal{A}$	> 0.5	> 0.5	> 0.8
$\Delta\phi(j_1, E_T^{\mathrm{miss}})$	> 2.5	> 2.5	> 2.5

Table 6.9: Scheme of the background estimation strategy for b0L-SRC and summary of the CR definitions. Jets and leptons (electrons or muons) are labelled in decreasing order in p_{T}.



(a) (b)

Figure 6.14: Kinematical distributions of key variables in b0L-CRs: E_T^{miss} in b0L-CRwA (a) [18] and $E_T^{\text{miss,corr}}$ in b0L-CRzC (b) [18]. The MC samples are normalised using the results of the background fits, and good agreement is found between the data and the predictions. The uncertainty bands include the detector-related systematics described in Section 6.3.5, and the last bins of the distributions include overflow events.

the results of the background fits are provided in Section 6.4.1.
In addition, a set of VRs is defined to verify the quality of the extrapolation from the CRs to the SRs, as anticipated in Section 5.3.6. The selections of the VRs are closely resembling the corresponding SRs, with baseline lepton veto and E_T^{miss} trigger requirement, and are optimised to minimise the yields of the benchmark signal models, so that the data counts are expected to agree with the background only predictions. The detail of each VR definition is presented in Table 6.10, which provides a description of the modified selections with respect to the corresponding SRs in Table 6.3. The VR strategy for each type of b0L SR, also shown in the schemes in Tables 6.7-6.9, is briefly summarised below:

VRs for the b0L-SRAx regions Two VRs of type A are defined by inverting the selections on m_{bb} and m_{CT} with respect to the b0L-SRAx regions, using a low m_{CT} threshold (250 GeV) to reduce the signal contamination. In addition, an upper bound on $m_{\text{T}}^{\text{min}}(j_{\text{1-4}}, E_T^{\text{miss}})$ is included to ensure that the VRs are orthogonal to b0L-SRB.

VRs for b0L-SRB The extrapolation from the type B CRs to b0L-SRB is verified in

VR	Corresponding SR	Selection changes
b0L-VRmctA	b0L-SRA	$m_{\mathrm{T}}^{\mathrm{min}}(j_{1\text{-}4}, E_T^{\mathrm{miss}}) < 250$ GeV, $150 < m_{\mathrm{CT}} < 250$ GeV
b0L-VRmbbA	b0L-SRA	$m_{\mathrm{T}}^{\mathrm{min}}(j_{1\text{-}4}, E_T^{\mathrm{miss}}) < 250$ GeV, $100 < m_{\mathrm{bb}} < 200$ GeV
b0L-VRzB	b0L-SRB	$m_{\mathrm{CT}} < 250$ GeV, $200 < m_{\mathrm{T}}^{\mathrm{min}}(j_{1\text{-}4}, E_T^{\mathrm{miss}}) < 250$ GeV, $\mathcal{A} < 0.8$, no selection on $\Delta\phi(b_1, E_T^{\mathrm{miss}})$ and $\Delta\phi(b_2, E_T^{\mathrm{miss}})$
b0L-VRttB	b0L-SRB	$m_{\mathrm{CT}} < 250$ GeV, $150 < m_{\mathrm{T}}^{\mathrm{min}}(j_{1\text{-}4}, E_T^{\mathrm{miss}}) < 200$ GeV, $\mathcal{A} < 0.8$
b0L-VRttC	b0L-SRC	$m_{\mathrm{CT}} < 250$ GeV, $m_{\mathrm{T}}^{\mathrm{min}}(j_{1\text{-}4}, E_T^{\mathrm{miss}}) < 250$ GeV, $0.6 < \mathcal{A} < 0.8$

Table 6.10: Summary of the VRs used in the zero-lepton channel of the analysis. Each VR is defined with respect to a specific SR, indicated in the middle column, with modified selections to reduce the signal contamination, as described in the right column.

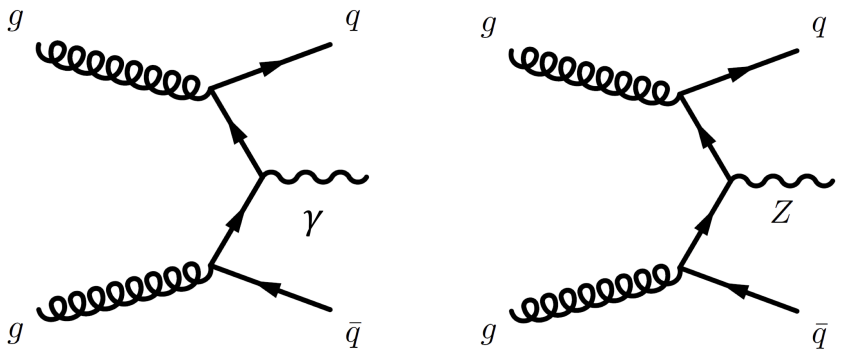
two VRs, that are dominated by $t\bar{t}$ and Z+jets events respectively. The $t\bar{t}$ VR is defined with $150 < m_{\mathrm{T}}^{\mathrm{min}}(j_{1\text{-}4}, E_T^{\mathrm{miss}}) < 200$ GeV, while the Z VR has $200 < m_{\mathrm{T}}^{\mathrm{min}}(j_{1\text{-}4}, E_T^{\mathrm{miss}}) < 250$ GeV and no selection on the azimuthal distance between the b-jets and the E_T^{miss}. Both VRs have $m_{\mathrm{CT}} < 250$ GeV and $\mathcal{A} < 0.8$ for orthogonality with the VRs and SRs of type A and C.

VR for b0L-SRC A single VR of type C is defined from b0L-SRC by selecting the asymmetry window $0.6 < \mathcal{A} < 0.8$. Upper bounds on m_{CT} and $m_{\mathrm{T}}^{\mathrm{min}}(j_{1\text{-}4}, E_T^{\mathrm{miss}})$ are also included to ensure that the VR is orthogonal to type A and B regions.

6.3.2 Data driven estimates of the Z background

The accuracy of the CR-based predictions of the Z background in the zero-lepton SRs is further verified by employing two independent data-driven estimates, that have been developed by C.Macdonald and J.Anders and are discussed in detail in their PhD theses [221, 222]. A brief description of the two techniques is provided below:

***Z* from γ** In the first method, the Z+jets yield is predicted by mimicking the SR selections in a channel where the E_T^{miss} is replaced with a photon, exploiting the similarity between γ+jets and Z+jets production illustrated by the diagrams in Figure 6.15 [223].



(a) (b)

Figure 6.15: Comparison between LO diagrams of γ+jets (a) and Z+jets (b) production, highlighting the similarity between the two processes.

***Z*+*bb* from *Z*+light** The second technique consists in the extrapolation of the shape of m_{CT} from Z+jets events with light jet flavour, where the variable is computed with the two leading jets in p_{T}, to the Z+bb events selected in the SRs. By construction, the measured shape can only be used in regions where the two b-jets are also leading in p_{T}, so the method is only employed for the SRs of type A.

6.3.3 Multi-jet background estimate from jet smearing

An additional data driven technique is employed to estimate the background from multi-jet production in the zero-lepton regions of the analysis, where large E_T^{miss} is required. The method, referred to as jet smearing, is widely used by the ATLAS Collaboration [224], and is described in detail in the PhD thesis of C.Macdonald [221] who implemented it in the present search. The underlying assumption is that large E_T^{miss} in multi-jet events is mainly caused by the mis-measurement of the p_{T} of the jets, which results in an overall momentum imbalance in the transverse plane. The procedure consists of the following steps:

1. A sample of seed events with well measured jets and low E_T^{miss} is selected in data.

2. The seed events are smeared n times using a jet response function determined in MC simulation and corrected with data, obtaining a sample of n smeared events with larger E_T^{miss} tails (pseudo-data) for each seed.

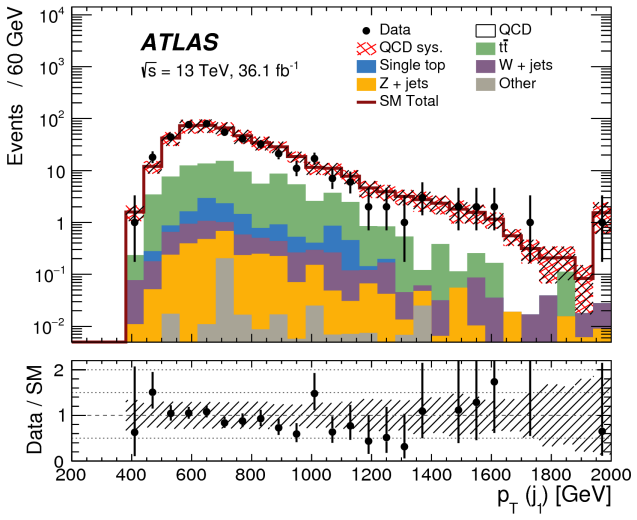


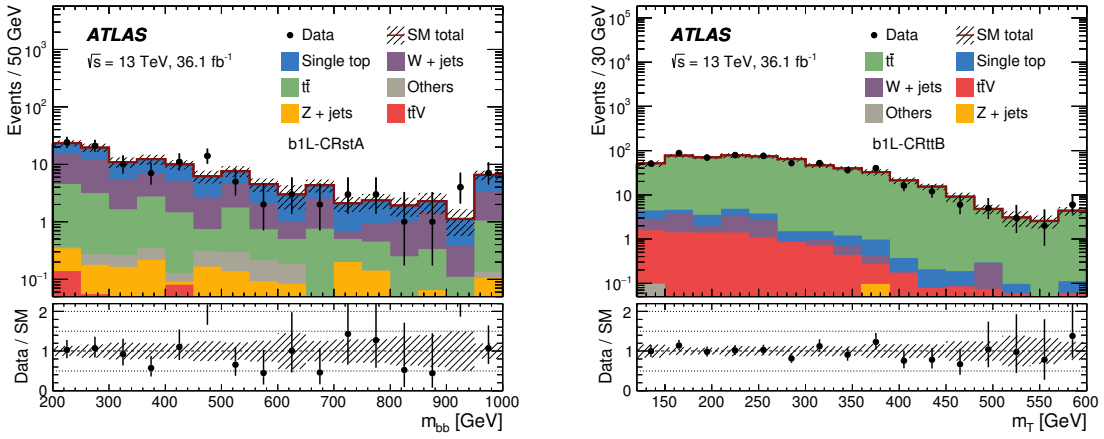
Figure 6.16: Distribution of the p_{T} of the leading jet in a multi-jet enriched CR, where the multi-jet yield is labelled as QCD and is estimated using the jet smearing technique described in Section 6.3.3 [18].

3. The pseudo-data are normalised in a multi-jet enriched CR, and are then used to predict the background yields in the relevant regions of the analysis.

Additional corrections are applied in the present search to take into account the E_T^{miss} from the decay of the b-hadrons, and systematic uncertainties on the pseudo-data are evaluated as described in Ref. [221]. The distribution of the p_{T} of the leading jet in a multi-jet enriched CR is shown in Figure 6.16, where good agreement is observed between data and predictions, while the expected multi-jet yields in the remaining regions of the analysis are found to be negligible [18].

6.3.4 Control and Validation Regions in the one-lepton channel

As anticipated in Figure 6.12, the dominant background processes in the one-lepton SRs are $t\bar{t}$ and single top production. Two CRs are defined to constrain $t\bar{t}$ in the b1L-SRAx and b1L-SRB regions by selecting events with am_{T2} below the top quark mass, and by inverting the m_{bb} and $m_{\mathrm{T}}^{\mathrm{min}}(b_{1\text{-}2}, E_T^{\mathrm{miss}})$ selections respectively. In order to allow the statistical combination of the results of the b0L-SRAx and b1L-SRAx regions, as well as of b0L-SRB and b1L-SRB (see Section 6.4.3), the associated CRs are made orthogonal by applying an upper bound on $m_{\mathrm{CT}} < 250$ GeV in b1L-CRttA. The single top normalisation in the b1L-SRAx regions is measured in the same CR used for the corresponding b0L



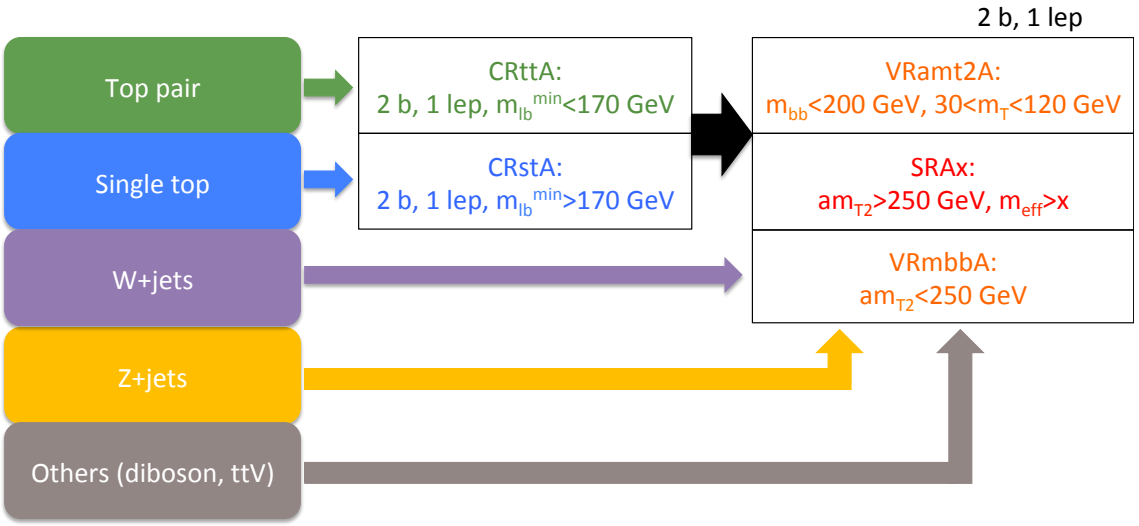
(a) (b)

Figure 6.17: Kinematical distributions of key variables in b1L-CRs: m_{bb} in b1L-CRstA (a) [18] and m_{T} in b1L-CRttB (b) [18]. The MC samples are normalised using the results of the background fits, and good agreement is found between the data and the predictions. The uncertainty bands include the detector-related systematics described in Section 6.3.5, and the last bins of the distributions include overflow events.

selections, while a dedicated single top CR is defined for b1L-SRB. A third CR is defined to estimate W+jets in b1L-SRB, using events with a single b-tagged jet similarly to what is done in the b0L regions.

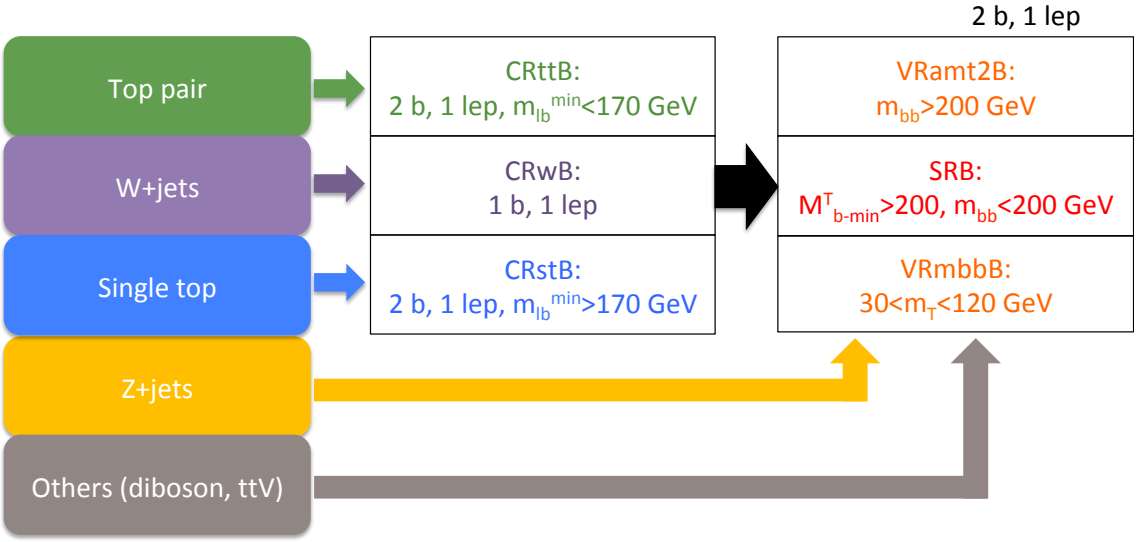
In summary, the backgrounds in the b1L-SRAx regions and in b1L-SRB are normalised in two and three CRs respectively, with selections detailed in Tables 6.11 and 6.12. The data-MC agreement in the b1L CRs is shown in Figure 6.17 after the normalisation factors from the background only fit are applied.

The VRs associated with the b1L SRs are presented in Table 6.13, and are also included in the schemes in Tables 6.11 and 6.12 together with the other regions of the corresponding channels. For the type A selections, two VRs with m_{eff} threshold of 300 GeV are implemented: the first is defined with inverted am_{T2} with respect to the b1L-SRAx regions, while in the second an upper bound is placed on m_{T} and m_{bb}. Finally, the accuracy of the background estimate in b1L-SRB is verified in two VRs with no selection on $\Delta\phi(b_1, E_T^{\mathrm{miss}})$, by reverting the m_{bb} and m_{T} selections.



b1L-	CRttA	CRstA
Number of leptons ($\ell = e, \mu$)	1	1
$p_{\mathrm{T}}(\ell)$ [GeV]	> 27	> 27
N_{jets} ($p_{\mathrm{T}} > 35$ GeV)	≥ 2	[2–4]
$p_{\mathrm{T}}(j_1)$ [GeV]	> 35	> 130
$p_{\mathrm{T}}(j_2)$ [GeV]	> 35	> 50
$p_{\mathrm{T}}(j_4)$ [GeV]	> 35	[35–50]
$\Delta\phi_{\mathrm{min}}^{j_{1\text{-}4}}$	> 0.4	> 0.4
b-jets	any 2	j_1 and (j_2 or j_3 or j_4)
m_{bb} [GeV]	< 200	> 200
$m_{b\ell}^{\mathrm{min}}$ [GeV]	< 170	> 170
E_T^{miss} [GeV]	> 200	> 200
m_{T} [GeV]	> 140	–
am_{T2} [GeV]	< 250	–
m_{eff} [GeV]	> 300	–
m_{CT} [GeV]	< 250	> 250
$E_T^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}}$ [GeV $^{1/2}$]	> 8	–
$E_T^{\mathrm{miss}}/m_{\mathrm{eff}}$	–	> 0.25

Table 6.11: Scheme of the background estimation strategy for the b1L-SRAx regions and summary of the CR definitions. Jets and leptons (electrons or muons) are labelled in decreasing order in p_{T}.



b1L-	CRttB	CRstB	CRwB
Number of leptons ($\ell = e, \mu$)	1	1	1
$p_{\mathrm{T}}(\ell)$ [GeV]	> 27	> 27	> 27
N_{jets} ($p_{\mathrm{T}} > 35$ GeV)	≥ 2	≥ 2	≥ 2
$p_{\mathrm{T}}(j_1)$ [GeV]	> 35	> 35	> 35
$p_{\mathrm{T}}(j_2)$ [GeV]	> 35	> 35	> 35
$\Delta\phi_{\mathrm{min}}^{j_{1\text{-}4}}$	> 0.4	> 0.4	> 0.4
b-jets	any 2	any 2	any 1
m_{bb} [GeV]	< 200	> 200	> 200
$m_{b\ell}^{\mathrm{min}}$ [GeV]	< 170	> 170	< 170
E_T^{miss} [GeV]	> 200	> 200	> 200
m_{T} [GeV] > 120	[30–120]	[30–120]	
am_{T2} [GeV]	< 200	–	> 200
$E_T^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}}$ [GeV $^{1/2}$]	> 8	> 8	> 8
$m_{\mathrm{T}}^{\mathrm{min}}(b_{1\text{-}2}, E_T^{\mathrm{miss}})$ [GeV]	< 200	> 200	> 200
$\Delta\phi(b_1, E_T^{\mathrm{miss}})$	> 2.0	> 2.0	> 2.0

Table 6.12: Scheme of the background estimation strategy for b1L-SRB and summary of the CR definitions. Jets and leptons (electrons or muons) are labelled in decreasing order in p_{T}.

VR	Corresponding SR	Selection changes
b1L-VRamt2A	b1L-SRA300	$30 < m_{\mathrm{T}} < 140$ GeV, $m_{\mathrm{bb}} < 200$ GeV
b1L-VRmbbA	b1L-SRA300	$am_{\mathrm{T2}} < 250$ GeV
b1L-VRamt2B	b1L-SRB	$\Delta\phi(b_1, E_T^{\mathrm{miss}}) > 2.0$, $m_{\mathrm{bb}} > 200$ GeV
b1L-VRmbbB	b1L-SRB	$\Delta\phi(b_1, E_T^{\mathrm{miss}}) > 2.0$, $30 < m_{\mathrm{T}} < 120$ GeV

Table 6.13: Summary of the VRs used in the one-lepton channel of the analysis. Each VR is defined with respect to a specific SR, indicated in the middle column, with modified selections to reduce the signal contamination, as described in the right column.

6.3.5 Systematic uncertainties

A variety of sources of systematic uncertainties are considered in the present analysis, as outlined in Section 5.3.4, and the final impact in the zero-lepton and one-lepton SRs is summarised in Table 6.14. The size of the uncertainties is strongly reduced by the presence of the CRs, which implies that the only relevant variations for the normalised background processes are those affecting the associated transfer factors (see Equation 5.13 in Section 5.3.6) rather than the SR yields.

The dominant sources of experimental uncertainty are the JES and JER, whose impact varies depending on the SRs reaching a maximum of 11% in b0L-SRC, and the b-tagging calibration SFs. In the latter case, due to the significant amount of mis-tags from charm in regions with large m_{CT} and am_{T2} (see Figure 6.8), the uncertainty on the c-tagging efficiency plays an important role. The high-p_{T} MC based uncertainty evaluated in Chapter 4 is also included, and contributes to the total flavour tagging uncertainty in Table 6.14. The impact of the uncertainties related to leptons, photons and E_T^{miss} on the final results is found to be negligible.

In addition, the modelling of the background processes in the extreme regions of phase space selected by the analysis is an important source of systematic uncertainty. The dominant theoretical uncertainty in the b0L regions is related to Z+jets production, while single top and $t\bar{t}$ (collectively referred to as top production in Table 6.14) have the largest impact in the one-lepton channel.

Finally, the experimental and theoretical uncertainties are also evaluated for the expected signal yields in the SRs. The experimental systematics on the $bb + E_T^{\mathrm{miss}}$ signals (Figure 6.1a) are between 15% and 30% depending on the masses, while they range between 10% and 25% for $tb + E_T^{\mathrm{miss}}$ models (Figure 6.1b) in the one-lepton channel, with dom-

Source \ SR	b0L-Ax	b0L-B	b0L-C	b1L-Ax	b1L-B	b1L-A300-2j
Experimental uncertainty						
JES	2.3-3.4%	5.7%	4.3%	1.2-1.5%	0.9%	6.9%
JER	0.9-3.3%	3.5%	11%	5.3-8.6%	0.9%	4%
Flavour tagging	3.3-4.3%	7.5%	4.7%	6.1-6.3%	2%	6.6%
Theoretical modelling uncertainty						
Z+jets	9.6-12.3%	13%	11%	-	-	-
W+jets	3.4-5.2%	4.7%	7.6%	1.3-1.6%	8.6%	7.9%
Top production	2.2-3.1%	6%	3.6%	19%	13%	22%

Table 6.14: Summary of the most relevant experimental and theoretical uncertainties on the SM background predictions in each SR of the analysis [18]. The uncertainties are relative to the total background yields, and a range is shown for the multiple SRs of type A. The individual variations can be correlated, so the total uncertainty on the background is not necessarily obtained by summing them in quadrature.

inant impact from flavour tagging systematics. The theory uncertainties are dominated by the variations of renormalisation and factorisation scales, and amount to 15%-25% for bottom squark masses between 400 GeV and 1100 GeV.

6.4 Results and interpretation

Having described all the relevant aspects of the analysis, it is now possible to present the final results obtained by performing the likelihood fits described in Section 5.4.3. Section 6.4.1 presents the outcome of the background only fits in the zero-lepton and one-lepton channels, including the data-MC comparisons in the VRs, while the SR results are discussed in Section 6.4.2. Since no significant excess beyond the SM predictions is observed, the results are interpreted by performing exclusion fits (Section 6.4.3) starting from individual fits in each channel to set limits on $bb + E_T^{\text{miss}}$ and $tb + E_T^{\text{miss}}$ models and continuing with a statistical combination of zero-lepton and one-lepton SRs that improves the limits on the $tb + E_T^{\text{miss}}$ signal.

6.4.1 Background only fits

A dedicated background only fit is performed for each type of selection in each channel of the analysis, resulting in three fits in the zero-lepton channel (for the b0L selections of type A, B and C) and two in the one-lepton channel (for the b1L regions of type A and B). Each fit provides an estimate of the normalisation factors of the relevant background processes using the CRs presented in Tables 6.7-6.9 and 6.11-6.12, yielding independent background predictions in the associated VRs and SRs.
The expected background yields from MC simulations are shown in Figure 6.18 for each CR in each channel, together with the observed events in data and the resulting normalisation factors obtained from the fits. The $\mu_{t\bar{t}}$ parameter is generally compatible with 1 within the associated uncertainties with the exception of b0L-CRttC, which requires the presence of an ISR jet with high-p_{T} and results in a normalisation factor of 0.66 ± 0.18. The μ_{st} parameter is significantly smaller than 1 in all cases, while the normalisation factors of W and Z+jets are always larger than 1, possibly due to a mis-modelling of the heavy flavour jet fractions in the Sherpa MC samples [225].
Figure 6.19 presents the full comparison between the expected and observed yields in the VRs of each channel of the analysis, using the normalisation factors of the backgrounds presented in Figure 6.18. The bottom panel of the figure shows the associated pulls, defined as the difference between the observed and predicted number of events in the VRs (n_{obs} and n_{pred}) divided by the total uncertainty on the background (σ_{tot}). The expected yields from the SM processes are compatible with the data within uncertainties: the largest discrepancies are observed in b0L-VRttC and b1L-VRamt2A, but the deviation from the background hypothesis is below 2σ in both cases, which corresponds to a sufficient level of agreement between data and predictions. In summary, the results in Figure 6.19 confirm the accuracy of the background estimation strategy in the targeted regions of phase space in all channels of the analysis.

6.4.2 Unblinded SRs

Having verified that the background modelling is accurate in the VRs, the observed yields in the SRs are compared to the corresponding background predictions in the zero-lepton and one-lepton channels of the analysis, as shown in Tables 6.15 and 6.16. The tables include the number of observed events in each SR, the total background prediction before and after the fit and the breakdown of the post-fit background yield in the individual SM processes, using the normalisation factors in Figure 6.18. The SR pulls, defined equiva-

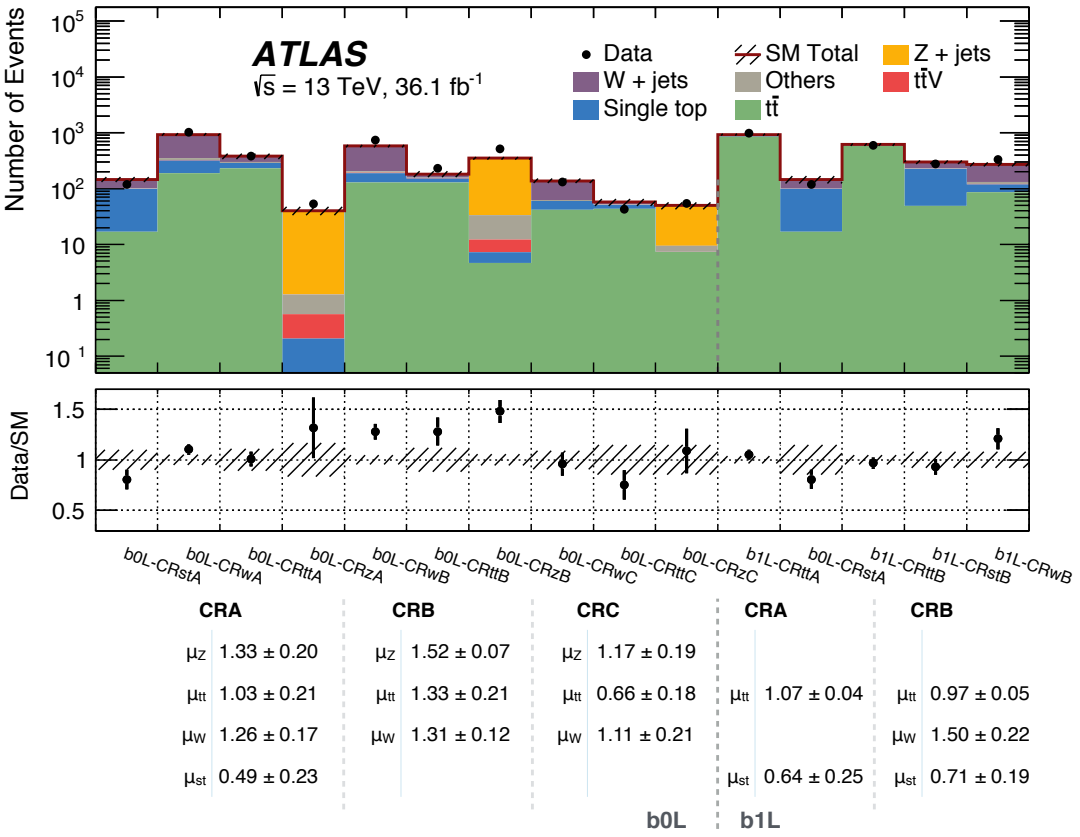


Figure 6.18: Results of the background only fits in the CRs associated with the zero-lepton and one-lepton channels of the analysis [18]. The panels in the figure present the comparison between the background yields predicted by the MC simulations and the observed counts in the CRs, highlighting the ratio of the observed and expected events and the post-fit values of the normalisation factors (μ) of each fitted background. The uncertainty on the MC predictions in the top panel includes only the MC statistics, while the normalisation factors are presented taking into account both statistical and experimental uncertainties.

lently to the VR ones and summarised in Figure 6.20, indicate that no significant excess beyond the SM predictions is present. The largest deviation is observed in b1L-SRA300-2j, which corresponds to a discrepancy of approximately 1.5σ between the background only hypothesis and the measured data, too low to suggest the presence of any signal. Together with the total SR yields presented in Tables 6.15 and 6.16, it is also useful to examine the distributions of the key discriminating variables in each region, as shown in Figure 6.21. The distributions are presented after applying all the SR selections except the

b0L- Signal Region	SRA350	SRA450	SRA550	SRB	SRC
Observed	81	24	10	45	7
Total background (fit)	70 ± 13	22 ± 5	7.2 ± 1.5	37 ± 7	5.5 ± 1.5
Z+jets	46 ± 12	13.6 ± 3.7	4.0 ± 1.2	20.0 ± 5.2	2.3 ± 0.8
$t\bar{t}$	2.0 ± 0.6	0.5 ± 0.2	0.16 ± 0.07	5.1 ± 2.7	0.8 ± 0.3
Single top	4.7 ± 3.4	1.2 ± 1.0	0.5 ± 0.3	2.6 ± 1.1	0.7 ± 0.3
W+jets	15 ± 5	5.0 ± 1.8	2.4 ± 1.0	5.5 ± 2.0	1.3 ± 0.8
Others	2.5 ± 1.7	1.4 ± 1.2	0.07 ± 0.03	4.0 ± 1.1	0.4 ± 0.1
Total background (MC exp.)	60.4	18.5	6.2	28	5.4

Table 6.15: Results of the background only fits in the b0L SRs [18], based on the corresponding normalisation parameters in Figure 6.18. The uncertainties on the background yield predictions include statistical, experimental and theoretical components.

b1L- Signal Region	SRA600	SRA750	SRB	SRA300-2j
Observed	21	13	69	12
Total background (fit)	24 ± 6	15 ± 4	53 ± 12	6.7 ± 2.3
$t\bar{t}$	10 ± 5	5.5 ± 2.7	16 ± 7	2.4 ± 1.3
Single top	7 ± 4	4.5 ± 2.8	10 ± 5	3.3 ± 2.0
W+jets	0.9 ± 0.5	0.6 ± 0.3	17 ± 8	0.4 ± 0.3
$t\bar{t}V$	5.4 ± 0.6	4.0 ± 0.5	9 ± 1	0.6 ± 0.1
Others	0.07 ± 0.02	0.07 ± 0.03	1.8 ± 0.3	0.07 ± 0.02
Total background (MC exp.)	27	17	52	8.4

Table 6.16: Results of the background only fits in the b1L SRs [18], based on the corresponding normalisation parameters in Figure 6.18. The uncertainties on the background yield predictions include statistical, experimental and theoretical components.

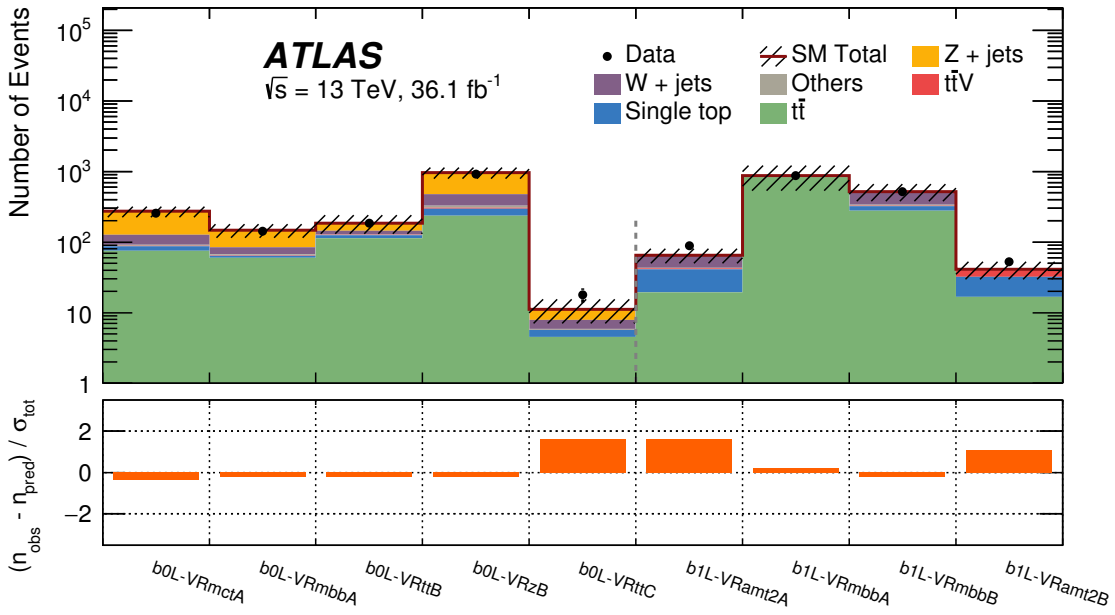


Figure 6.19: Results of the background only fits in the VRs of the analysis, using the background normalisation factors in Figure 6.18 [18]. The uncertainty band in the top panel includes all the variations presented in Section 5.3.4. The bottom panel shows the observed pulls in each VR, defined as the ratio of the difference between observed and predicted yields and the total uncertainty on the SM background, including the Poisson uncertainty.

ones on the displayed variable, and are commonly referred to as N-1 plots. Each figure includes a histogram corresponding to a relevant signal benchmark model in addition to the post-fit background distribution and the data, so that the discriminating power of each variable is highlighted: for this purpose, a $bb + E_T^{\text{miss}}$ signal is chosen for the zero-lepton SRs (Figures 6.21a-6.21c) while a $tb+E_T^{\text{miss}}$ signal is used in the one-lepton channel (Figures 6.21d-6.21f). An arrow in each plot indicates the cut threshold used to define the corresponding SR, whose position is optimised to maximise the signal to background ratio. A good level of agreement between data and SM predictions is observed over the entire kinematical ranges displayed in each plot, both where the background processes are expected to be dominant and where the targeted signals are expected to appear.
Finally, the baseline N-1 distribution of m_{CT} in the b0L-SRAx regions shown in Figure 6.21a is reproduced in Figure 6.22 using the Z+jets predictions from the two data-driven estimates described in Section 6.3.2. All three methods yield a good level of agreement between data and expected SM backgrounds, and the corresponding predictions of the

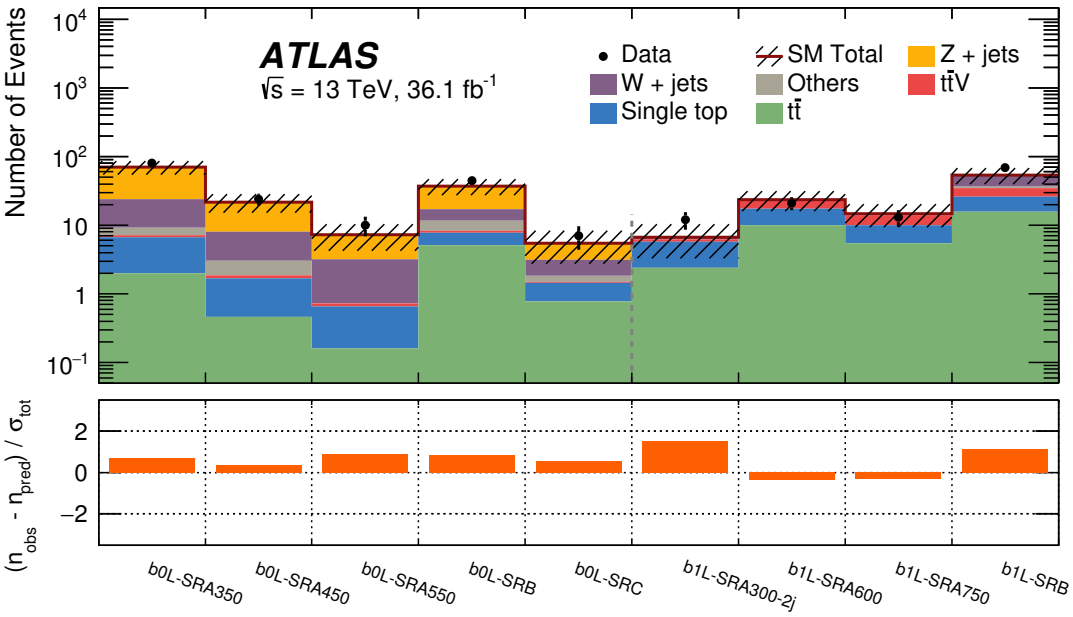


Figure 6.20: Results of the background only fits in the SRs of the analysis, using the background normalisation factors in Figure 6.18 [18]. The uncertainty band in the top panel includes all the variations presented in Section 5.3.4. The bottom panel shows the observed pulls in each SR, defined as the ratio of the difference between observed and predicted yields and the total uncertainty on the SM background, including the Poisson uncertainty.

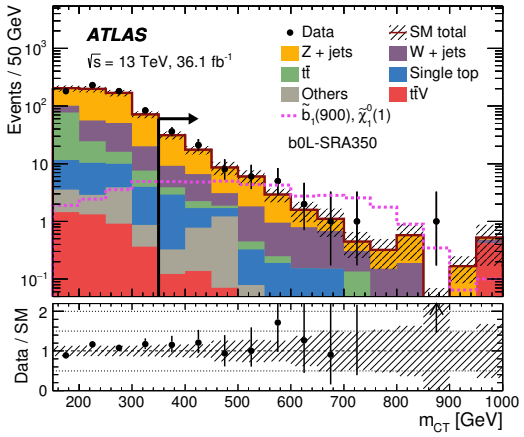
Z+jets yield in the relevant VRs and SRs are found to be consistent, confirming the robustness of the background estimation strategy of the analysis.

6.4.3 Exclusion limits

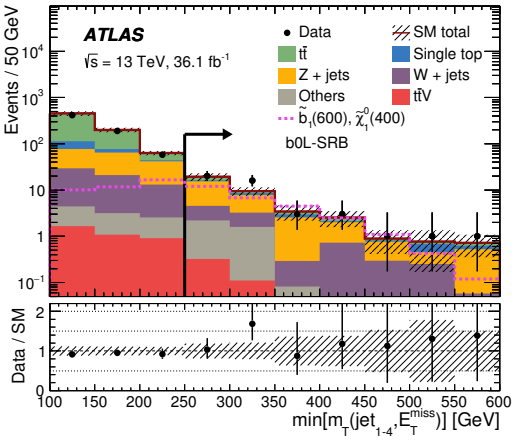
Since no significant excess is observed in any of the SRs, the results of the analysis are translated into exclusion limits, both on generic models of new physics and on the benchmark signal models shown in Figure 6.1. The limits are obtained by performing exclusion fits with the CL_s method, following the procedure outlined in Section 5.4.2.

Limits on generic signal models

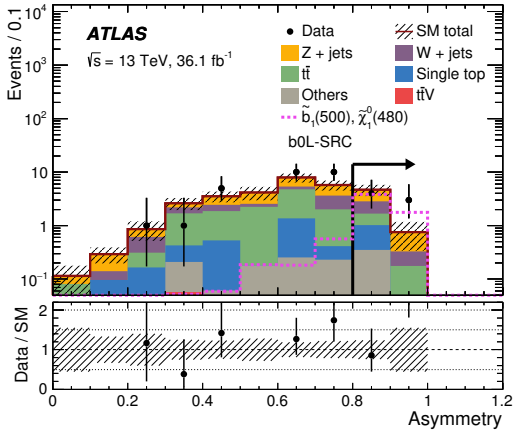
The model independent exclusion limits, which represent the constraints of the analysis on generic signal models that are able to enter the SRs, are presented in Table 6.17. In particular, the table presents the S^{95}_{obs} (S^{95}_{exp}) values, which indicate the observed (expected)



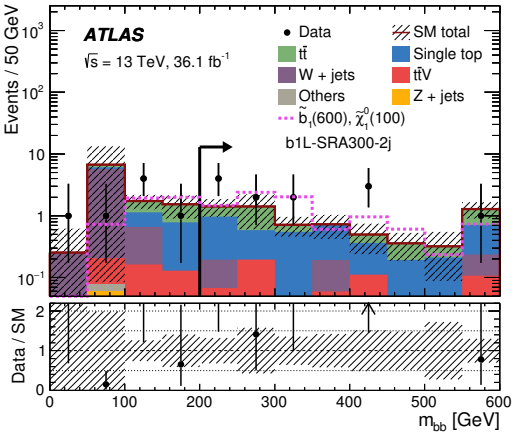
(a)



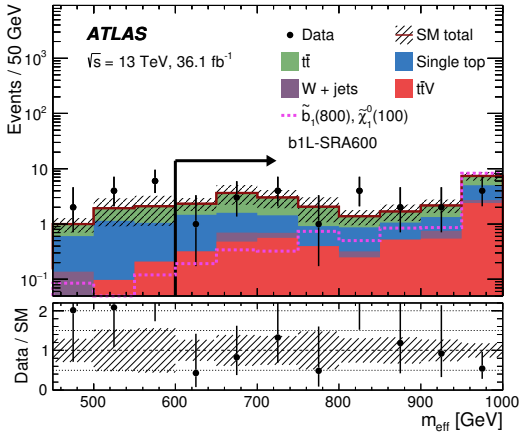
(b)



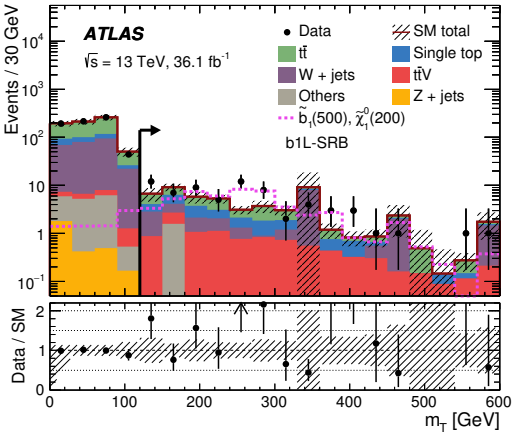
(c)



(d)

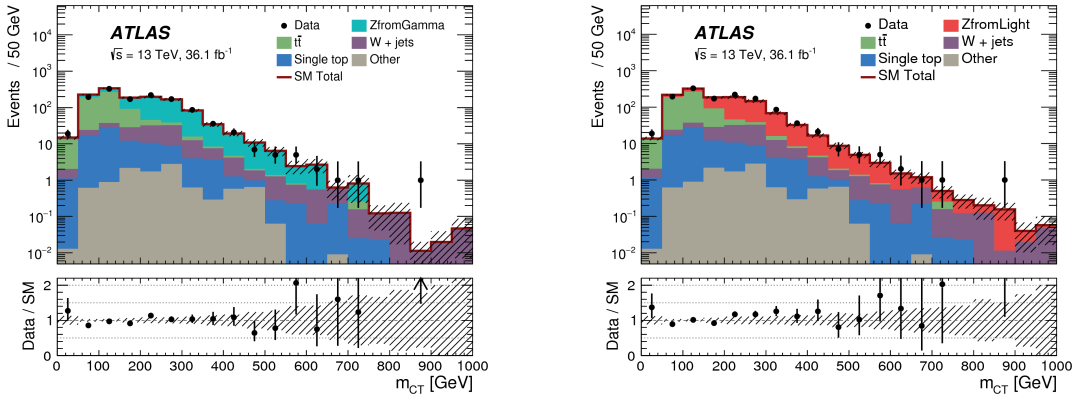


(e)



(f)

Figure 6.21: Post-fit distributions of m_{CT} in b0L-SRA (a), $m_{\mathrm{T}}^{\mathrm{min}}(j_{1\text{-}4}, E_T^{\mathrm{miss}})$ in b0L-SRB (b), $\mathcal{A}$ in b0L-SRC (c), m_{bb} in b1L-SRA300-2j (d), m_{eff} in b1L-SRA (e), m_{T} in b1L-SRB (f), where all selections are applied except those on the variables displayed in each plot [18]. The arrows indicate the thresholds of the SR selections.



(a) (b)

Figure 6.22: Post-fit N-1 distributions of m_{CT} in b0L-SRA where the prediction of the Z background is obtained from γ+jets (a) and from Z plus light jets events (b), as described in Section 6.3.2 [18].

limit at 95% CL on the number of events from BSM signals in each SR. In addition, by normalising S^{95}_{obs} to the total integrated luminosity of 36.1 fb^{-1}, the upper limit is also set on the visible cross-section of new physics models (σ_{vis}) defined as the product of the production cross-section, the SR acceptance and the detector efficiency[3]. Finally, the table presents the p_0 values of the background only hypothesis in each SR and the associated significance Z, as defined in Section 5.4.2.

Limits on the $\tilde{b}$ from the individual channels of the analysis

In addition, exclusion limits are specifically set on the two signal models targeted by the present analysis, which consist in a pair-produced sbottom that decays either exclusively to $b\tilde{\chi}^0_1$, yielding the $bb + E_T^{\mathrm{miss}}$ final state in Figure 6.1a, or to $b\tilde{\chi}^0_1$ and $t\tilde{\chi}^+_1$ with 50% BR each, as represented by the asymmetric $tb + E_T^{\mathrm{miss}}$ scenario in Figure 6.1b under the assumption of $\Delta m(\tilde{\chi}^+_1, \tilde{\chi}^0_1) = 1$ GeV. A variety of different mass hypotheses on the $\tilde{b}$ and the $\tilde{\chi}^0_1$ are considered for both types of models, covering a large area in the $m_{\tilde{b}}$-$m_{\tilde{\chi}^0_1}$ plane.

The $bb + E_T^{\mathrm{miss}}$ scenario is targeted exclusively by the SRs in the zero-lepton channel, because no lepton is produced in the decay chain. The exclusion limit in the $m_{\tilde{b}}$-$m_{\tilde{\chi}^0_1}$ plane is shown in Figure 6.23, where for each point in the parameter space the SR that

[3]The acceptance is defined as the fraction of the total signal events that satisfy the SR selections at truth level, while the detector efficiency is defined as the number of reconstructed signal events in the SR divided by the corresponding SR yield at truth level.

Signal channel	$\langle \epsilon A \sigma \rangle^{95}_{\text{obs}}$[fb]	S^{95}_{obs}	S^{95}_{exp}	p_0 (Z)
b0L-SRA350	1.06	38.2	$30.9^{+11.3}_{-8.4}$	0.28 (0.60)
b0L-SRA450	0.43	15.6	$13.9^{+5.6}_{-3.8}$	0.37 (0.34)
b0L-SRA550	0.30	10.7	$7.8^{+3.7}_{-1.6}$	0.20 (0.85)
b0L-SRB	0.72	26.1	$19.9^{+8.3}_{-5.4}$	0.23 (0.74)
b0L-SRC	0.24	8.7	$6.8^{+3.3}_{-1.3}$	0.30 (0.54)
b1L-SRA300-2j	0.39	14.1	$9.3^{+3.5}_{-3.1}$	0.08 (1.43)
b1L-SRA600	0.38	13.6	$14.8^{+5.4}_{-4.4}$	0.50 (0.00)
b1L-SRA750	0.27	9.9	$11.2^{+4.0}_{-2.3}$	0.50 (0.00)
b1L-SRB	1.12	40.3	$28.7^{+10.7}_{-8.2}$	0.21 (0.80)

Table 6.17: Model independent exclusion limits of the analysis [18]. The first two columns show the 95% CL observed upper limits on the visible cross-section ($\langle \epsilon A \sigma \rangle^{95}_{\text{obs}}$) and on the number of signal events (S^{95}_{obs}), while the third column shows the upper limit given the expected number of background events (S^{95}_{exp}) and its $\pm 1\sigma$ variations. Finally, the p_0 values of the background only hypothesis (with maximum value truncated at 0.5) and the associated significance Z are included.

yields the best expected CL_s is used. The exclusion contour of the present analysis is superimposed to the result of the previous searches described in Section 6.1.3, that are significantly extended thanks to the larger size of the dataset and the optimised analysis strategy. As previously shown in Figure 6.11, the best sensitivity to signals with large mass splitting between the $\tilde{b}$ and the $\tilde{\chi}^0_1$ is obtained with the b0L-SRA regions, while the intermediate and small mass splittings are targeted by b0L-SRB and b0L-SRC. The $\tilde{b}$ is excluded up to 950 GeV for $\tilde{\chi}^0_1$ masses below 420 GeV, while the constraints are less stringent in the compressed region of the mass plane.

Unlike the $bb+E_T^{\text{miss}}$, the $tb+E_T^{\text{miss}}$ scenario is significantly constrained by both channels of the analysis. The one-lepton SRs are sensitive to this model by construction, because they are specifically optimised for it, but the zero-lepton SRs are also able to provide a significant sensitivity by exploiting orthogonal subsets of events from the same signal samples. In particular, the sensitivity of the zero-lepton SRs comes primarily from the fraction of events with $bb + E_T^{\text{miss}}$ signature, that correspond to 25% of the total events in the signal with mixed decay, but also from the asymmetric $tb+E_T^{\text{miss}}$ events (and even from the $tt + E_T^{\text{miss}}$ ones) when the top quarks decay hadronically, despite the SR selec-

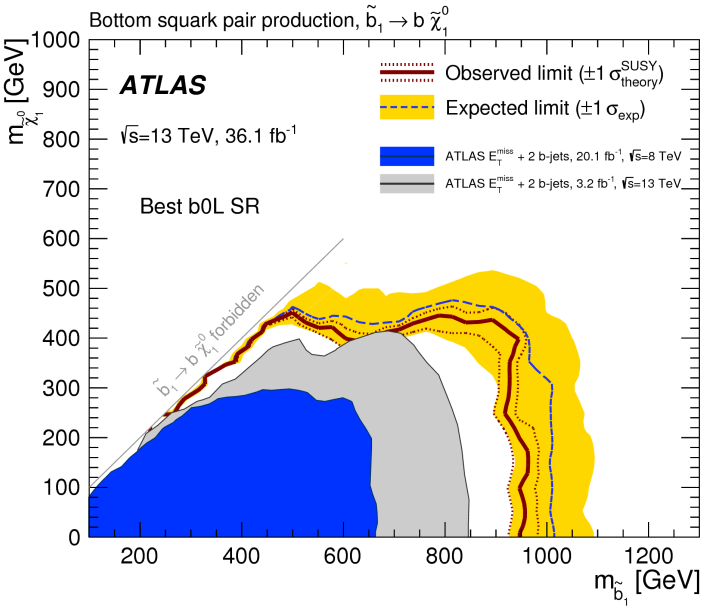
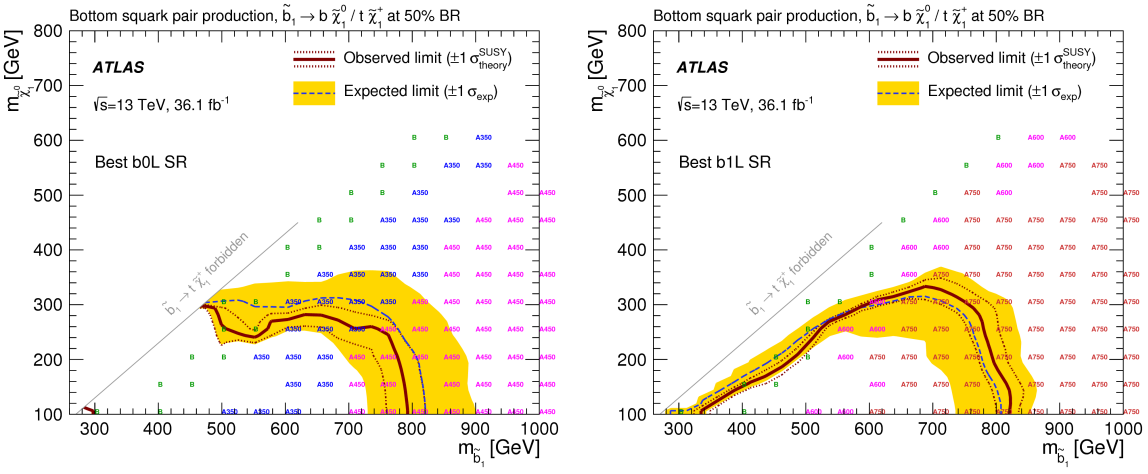


Figure 6.23: Exclusion limits on a pair-produced $\tilde{b}$ decaying into a $b\tilde{\chi}_1^0$ pair with 100% BR, based on the zero-lepton SR that yields the largest expected CL_s for each signal model [18].

tion on the maximum number of signal jets, which is not optimal for the experimental signature of this signal.

The exclusion limits of the zero-lepton and one-lepton SRs on the mixed signal scenario are shown in Figure 6.24, where the SR with the best expected CL_s is used for each benchmark model, as indicated in the plots. The zero-lepton SRs that yield significant constraints are the two b0L-SRA selections with low and intermediate m_{CT} threshold (b0L-SRA350 and b0L-SRA450) and b0L-SRB, while the acceptance of the mixed signal in b0L-SRA550 and b0L-SRC is too low to obtain a sufficient sensitivity. In the one-lepton channel, as already shown in Figure 6.13, each SR has an optimal performance in a specific region of parameter space, with the exception of b1L-SRA300-2j which is known to have a lower sensitivity to the model under exam. The exclusion contours obtained from the two channels are comparable for signals with large mass splitting between the $\tilde{b}$ and the $\tilde{\chi}_1^0$, while the zero-lepton selections set stronger limits in the compressed signal scenario, where the one-lepton SRs rapidly lose their sensitivity because the boost of the lepton from the signal decay chain is strongly reduced.



(a) (b)

Figure 6.24: Exclusion limits on a pair-produced $\tilde{b}$ decaying into $b\tilde{\chi}_1^0$ or $t\tilde{\chi}_1^+$ with 50% BR each, based on the zero-lepton (a) and one-lepton (b) SR that yields the largest expected CL_s for each signal model [18].

Statistical combination of zero-lepton and one-lepton SRs

Since the zero-lepton and one-lepton SRs have a similar sensitivity to the mixed decay scenario of the $\tilde{b}$, it is reasonable to perform a statistical combination of the two channels to further extend the exclusion limits of the analysis. In particular, the SRs to be combined are chosen based on the region of parameter space where they yield the best performance, as determined from the plots in Figure 6.24. The following three independent combinations are defined:

Comb-A-LowMass (A-LM) Statistical combination of b0L-SRA350 and b1L-SRA600, targeting intermediate mass splittings in the $m_{\tilde{b}}$-$m_{\tilde{\chi}_1^0}$ plane.

Comb-A-HighMass (A-HM) Statistical combination of b0L-SRA450 and b1L-SRA750, targeting large mass splittings in the $m_{\tilde{b}}$-$m_{\tilde{\chi}_1^0}$ plane.

Comb-B Statistical combination of b0L-SRB and b1L-SRB, targeting the compressed region in the $m_{\tilde{b}}$-$m_{\tilde{\chi}_1^0}$ plane.

Each of the above combinations is performed by running exclusion fits which include simultaneously the two input SRs and the associated sets of CRs. In practice, this is done by multiplying the likelihood functions of the exclusion fits of the individual channels,

maximising the resulting likelihood with respect to the full set of parameters.
The statistical combination can only be performed using orthogonal SR selections, while the CRs can be chosen to be either fully orthogonal or fully overlapping. In the former case two independent normalisation factors are used for each background process in the two channels, while in the latter case the same normalisation factors are employed in both SRs. The choice of common CRs would reduce the total number of parameters used for the background estimation, but it is not optimal if the SRs are defined with a different set of discriminating variables, because the resulting CR-SR extrapolation would increase. This is the reason why, in the present analysis, it was chosen to keep an orthogonal set of CRs for the b0L and b1L regions, with the only exception of b0L-CRstA, which is used in both channels of the analysis to constrain the normalisation of the single top background.
The A-LM and A-HM combinations are based on the regions defined in Tables 6.7 and 6.11 where, as anticipated, the CRs are all mutually orthogonal with the exception of b0L-CRstA. A minor ambiguity comes from the fact that a single μ_{st} must be used in both regions, while the corresponding values in the individual fits are different (0.49 ± 0.23 in the b0L fit and 0.64 ± 0.25 in the b1L fit, as shown in Figure 6.18) but the impact on the final results is found to be negligible. Finally, the implementation of the combined fit of b0L-SRB and b1L-SRB is straightforward, because the CRs and SRs (defined in Tables 6.8 and 6.12) are all mutually orthogonal.
The results of the statistical combination are shown in Figure 6.25, where the combined fit yielding the best expected CL$_s$ is used for each benchmark model in the $m_{\tilde{b}}$-$m_{\tilde{\chi}_1^0}$ plane. The combined limits extend the limits from the individual channels by approximately 50 GeV in $m_{\tilde{b}}$, and $\tilde{b}$ masses up to 860 GeV are excluded for $\tilde{\chi}_1^0$ masses below 250 GeV.

6.5 Summary

This chapter presented the results of a comprehensive search for $\tilde{b}$ pair production based on 36.1 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV. Events are selected with large E_T^{miss}, two b-tagged jets and either no lepton or one lepton, targeting the symmetric and mixed signals shown in Figure 6.1, and no significant excess is observed above the expected background from SM processes. The results are interpreted by setting stringent exclusion limits, that reach $\tilde{b}$ masses up to 950 (860) GeV for $\tilde{\chi}_1^0$ masses up to 420 (250) GeV for the symmetric and mixed signal models. As explained in Section 6.1.2, the

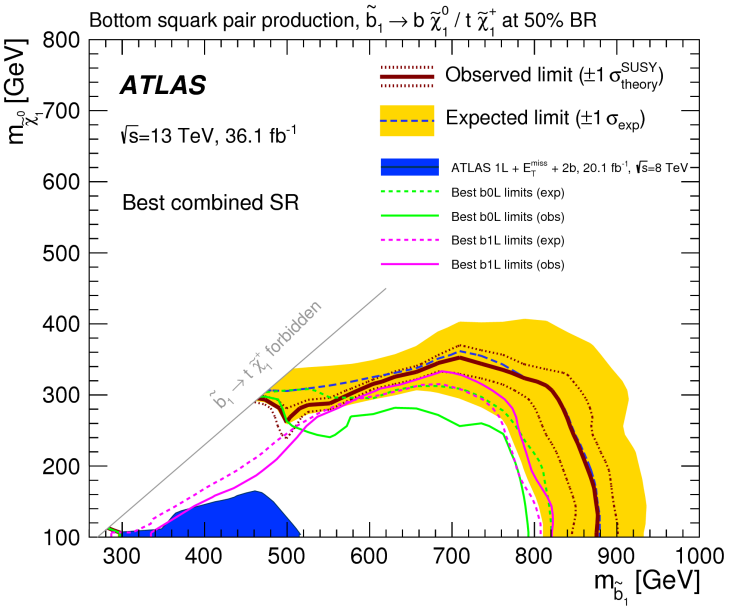


Figure 6.25: Exclusion limits on a pair-produced $\tilde{b}$ decaying into $b\tilde{\chi}_1^0$ or $t\tilde{\chi}_1^+$ with 50% BR each, using the statistical combination of zero-lepton and one-lepton SRs that yields the largest expected CL_s for each signal model [18].

constraints can be equivalently applied to $\tilde{t}$ pair production due to the assumption on $\Delta m(\tilde{\chi}_1^+, \tilde{\chi}_1^0) = 1$ GeV. The present analysis improves significantly the constraints on third generation squarks from previous Run 1 and Run 2 searches.

7 | Search for top squarks with RPV decays into b-ℓ pairs

The final topic covered in this thesis is a search for a pair produced $\tilde{t}$ in the context of the minimal $B - L$ MSSM models introduced in Section 1.3.4, where a leptoquark-like decay of the $\tilde{t}$ into a b-quark and a charged lepton (e, μ or τ) is assumed. As described in Appendix A, during my PhD I have been involved in the analysis with a leading role, participating to the development of all its major aspects and contributing to the publication of the results with 36.1 fb^{-1} of data at $\sqrt{s} = 13$ TeV [19]. This chapter begins by describing the signal process of interest in Section 7.1, and continues with a detailed discussion of the event selection techniques and the SR definitions in Section 7.2. The background estimation strategy, based on the definition of CRs for the dominant SM processes, is described in Section 7.3, while Section 7.4 presents the final results of the search.

7.1 Top squarks in the $B - L$ MSSM

As introduced in Section 1.3.4, the $\tilde{t}$ represents a viable LSP candidate in RPV MSSM models with broken $U(1)_{B-L}$ symmetry, with dominant decay into b-ℓ pairs. In particular, the BR of the decay into different lepton flavours (indicated as BR(e), BR(μ) and BR(τ) for the b-e, b-μ and b-τ decays) can be related to the neutrino mixing parameters, under the assumption that the symmetry breaking occurs through the VEV of right-handed sneutrinos [51]. The results of a scan of MSSM parameters, shown in Figure 1.8 in the BR(e)-BR(τ) plane, indicate that the BR of the $\tilde{t}$ decay into a b-e pair may reach 100% if an inverted mass hierarchy is assumed, while for a normal mass hierarchy the BR into b-τ may be as large as 90%.

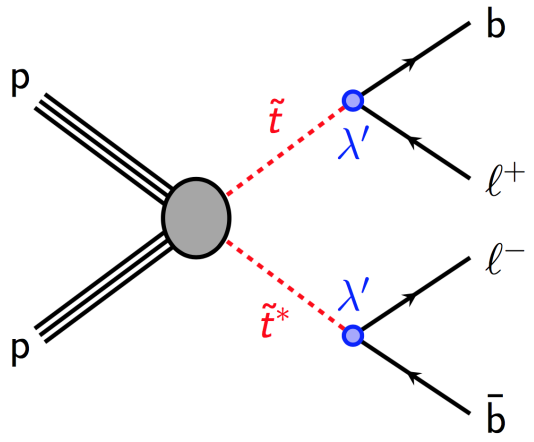


Figure 7.1: Benchmark signal model targeted by the analysis in Chapter 7, with a pair-produced stop decaying into b-ℓ pairs through an RPV coupling.

7.1.1 Signal benchmark

The experimental signature of the signal model is illustrated by the diagram in Figure 7.1, previously shown also in Figure 5.7a. The final state contains two oppositely charged leptons and two b-jets, and the b-ℓ pairs in each decay leg are expected to reconstruct the invariant mass of the $\tilde{t}$. Contrarily to the $\tilde{b}$ signals studied in Chapter 6, the present model yields no E_T^{miss} in the events, because no invisible particles are produced in the decay chain.

MC samples are generated for $\tilde{t}$ masses between 600 GeV and 1.6 TeV, based on a simplified model approach where the $\tilde{t}$ decays exclusively as in Figure 7.1 and no other SUSY particle is considered. The BR of the $\tilde{t}$ decays into b-e, b-μ and b-τ is set to $1/3$ each, so that a large sample of simulated events is available for all combinations of lepton flavours in the final state, with $\text{BR}(e) + \text{BR}(\mu) + \text{BR}(\tau) = 1$. Signal samples with any value of BR(e), BR(μ) and BR(τ) with unit sum are obtained through a dedicated reweighing procedure based on the truth information in the MC samples, where the relative normalisation of different sub-sets of events is rescaled according to the flavour of the leptons in the two decay legs. The selections of the present analysis are designed to target only final states with electrons and muons, split by flavour into the e-e, e-μ, and μ-μ channels, while signal events with τ leptons are only used for the interpretation of the results (see Section 7.4).

7.1.2 Previous results

The analysis in this chapter is the first ATLAS search targeting the signal in Figure 7.1 at $\sqrt{s} = 13$ TeV, while the only previous result on the same model is a Run 1 analysis performed with 20.1 fb^{-1} of data at $\sqrt{s} = 8$ TeV, made public only in the form of an ATLAS conference note [226]. Other searches by ATLAS [227] and CMS [228] have targeted the pair production of first, second and third generation leptoquarks (LQ) at $\sqrt{s} = 13$ TeV, but they have focused on the LQ decays into pairs of particles of the same generation (j-e, j-μ, b-τ, where j indicates a light jet) without covering the b-e and b-μ cases.
In summary, the only exclusion limits set by LHC experiments on the signal model targeted by the present analysis are those from the Run 1 ATLAS result in Ref. [226], and are shown in Figure 7.2 as a function of the $\tilde{t}$ mass in the BR(e)-BR(τ) plane. Similarly to what is done in the present search, the Run 1 result was optimised only for the b-e and b-μ decays of the $\tilde{t}$, while no SRs have been implemented for the b-τ decay. A mass limit of 500 GeV is obtained for BR(τ) values up to 80%, while the limit reaches 1 TeV for BR(e) $= 100\%$.
The Run 1 analysis is used as reference for the search described in the present chapter, which is designed to target $\tilde{t}$ masses between 600 GeV and 1.6 TeV. Mass values below 600 GeV are indeed strongly constrained by the results in Figure 7.2, while $\tilde{t}$ signals heavier than 1.6 TeV cannot be probed due to the low production cross section.

7.2 Event selection

This section presents a full discussion of the event selection strategy of the analysis, optimised for the signal benchmark process in Figure 7.1 with $\mathrm{BR}(e) = \mathrm{BR}(\mu) = 50\%$. A preliminary set of selections based on the expected experimental signature of the model is introduced in Section 7.2.1, while Sections 7.2.2 and 7.2.3 present the list of the relevant discriminating variables and the corresponding SR selections.

7.2.1 Preliminary selections and background composition

The decays of a pair-produced $\tilde{t}$ into b-e and b-μ pairs yield events with two b-jets and two leptons, that can be classified in the e-e, e-μ and μ-μ channels. A preliminary sample of candidate events is obtained using a logic OR of the single electron and muon triggers introduced in Section 5.3.1, together with the standard set of event cleaning selections

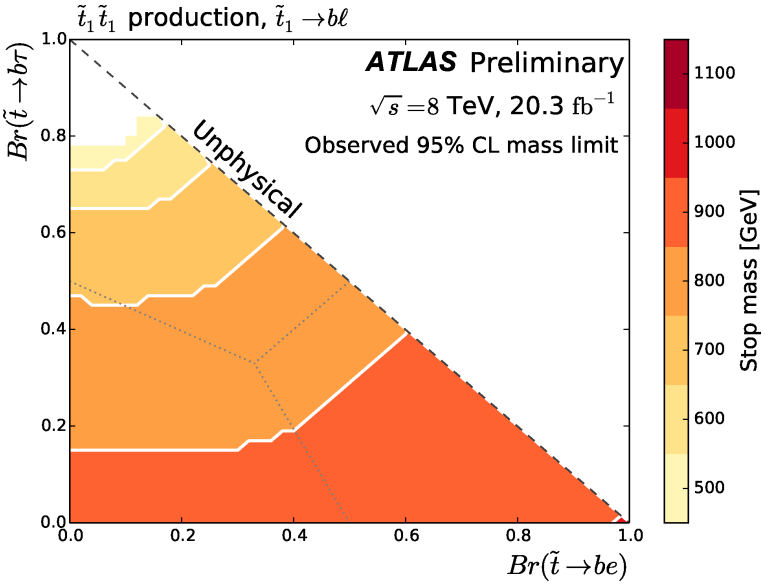


Figure 7.2: Observed limits on the mass of a pair produced $\tilde{t}$ decaying into b-ℓ pairs in the the BR(e)-BR(τ) plane with $\mathrm{BR}(e)+\mathrm{BR}(\mu)+\mathrm{BR}(\tau) = 1$, obtained from a dedicated Run 1 ATLAS search [226].

described in Section 5.3.1. In addition, the events are required to contain a minimum of 2 signal leptons (e or μ) with $p_{\mathrm{T}} > 40$ GeV, where the leading two must have opposite charge, and at least 2 signal jets with a p_{T} threshold of 60 GeV, that are needed to reproduce the experimental signature of the signal in Figure 7.1.
As further discussed in Section 7.2.3, events with more than two jets or leptons are not vetoed in the SRs to avoid an unnecessary reduction of the signal acceptance. In signal events it is reasonable to assume that the two leptons and the two jets from the decay of the pair-produced $\tilde{t}$ correspond to the leading two leptons and jets in p_{T}, because the decay products of the heavy $\tilde{t}$ signals under exam carry a significant momentum, whose average value is approximately given by $m_{\tilde{t}}/2$. Events with more than two signal leptons are rare and have a low impact on the yield predictions, while the jet multiplicity has a wider distribution due to extra hadronic activity produced in association with the objects in Figure 7.1. The assumption that the two leading jets in p_{T} represent the two b-jets produced in the $\tilde{t}$ decays is verified using the truth information from the MC samples, and is valid in 65-75% of signal events for $\tilde{t}$ masses between 600 GeV and 1.6 TeV, with larger percentages in heavier $\tilde{t}$ signals. Based on this, the relevant discriminating variables of the analysis are computed using the two leading leptons and jets in each event, as further described in Section 7.2.2.

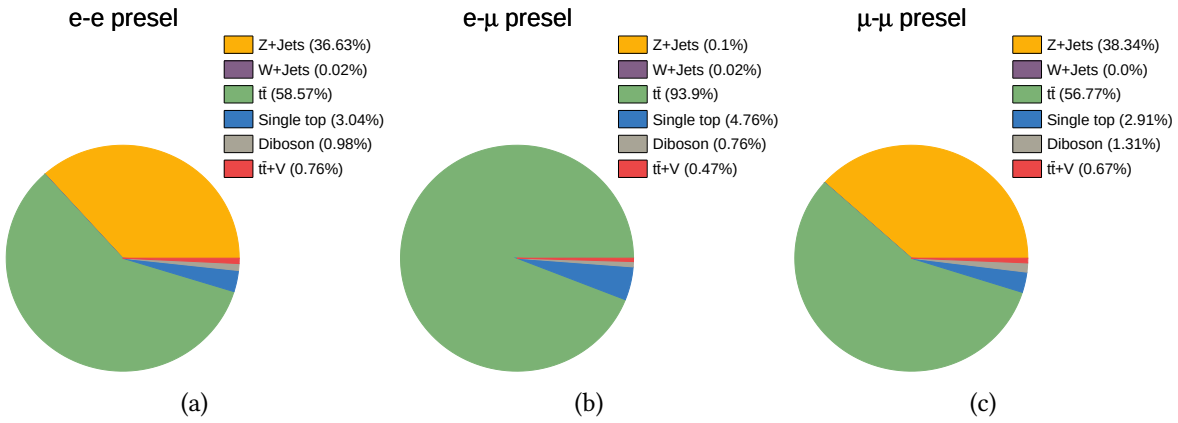
Analysis objects	Selections
Trigger	Single lepton (e or μ)
Signal leptons (e/μ)	≥ 2, $p_{\mathrm{T}} > 40$ GeV, leading 2 oppositely charged
Signal jets	≥ 2, $p_{\mathrm{T}} > 60$ GeV
b-jets (77% fixed cut WP)	≥ 1 among the two leading jets

Table 7.1: Preliminary event selections of the analysis, in addition to the event cleaning requirements described in Section 5.3.1.

Finally, despite the fact that the two signal jets from the $\tilde{t}$ decays originate from b-quark hadronisation, the b-tagging requirement at preselection level is limited to ≥ 1 b-tag among the two leading jets, because otherwise the low efficiency of the MV2c10 algorithm at high p_{T} (discussed in detail in Chapter 4) would cause a significant loss of signal acceptance.
The summary of the preliminary event selections of the analysis is presented in Table 7.1, while Figure 7.3 shows the breakdown of the expected background yield in the Same Flavour (SF) channels (e-e and μ-μ) and in the Different Flavour (DF) channel (e-μ). No significant discrepancy is observed between the background composition in the e-e and μ-μ channels, while in the DF channel the fraction of Z+jets events becomes negligible. The dominant SM process is $t\bar{t}$ production, plus a relevant fraction of Z+jets events in the SF channels, while the remaining background processes have a lower impact at preselection level.

7.2.2 Discriminating variables

The present section provides an overview of the key discriminating variables that are used to define the SRs of the analysis, based on the preliminary event selections in Table 7.1. All variables are implemented under the assumption that, in signal events, the two jets and the two leptons with highest p_{T} represent the four direct decay products of the $\tilde{t}$ pair, labelled as b_1, b_2, ℓ_1 and ℓ_2 in decreasing order in p_{T}. The list of discriminating variables includes the invariant masses of different b-ℓ pairs in the events, plus other event-level quantities that exploit the kinematical properties of the signal.



(a) (b) (c)

Figure 7.3: Background composition after the preliminary events selections described in Table 7.1 in the e-e (a), e-μ (b) and μ-μ (c) channels, determined using MC simulations.

Invariant masses of b-ℓ pairs

An effective separation between signal and background events is obtained using a set of variables related to the invariant masses of the b-ℓ pairs, formed by combining the two leading jets and leptons in the events. In order to employ these variables, it is necessary to implement a pairing algorithm that determines which of the two possible combinations of input objects ($b_1\ell_1$ and $b_2\ell_2$, or $b_1\ell_2$ and $b_2\ell_1$) corresponds to the two decay legs of the targeted signal. The algorithm exploits the fact that the invariant masses of the b-ℓ pairs from the two legs are expected to be similar, because in signal both pairs originate from the decay of a $\tilde{t}$, and the correct combination is selected by minimising the invariant mass asymmetry:

$$m_{b\ell}^{\text{asym}} = \frac{(m_{b\ell}^{0} - m_{b\ell}^{1})}{(m_{b\ell}^{0} + m_{b\ell}^{1})}. \tag{7.1}$$

where $m_{b\ell}^{0}$ and $m_{b\ell}^{1}$ represent the invariant masses of the two b-ℓ pairs for a given combination, labelled such that $m_{b\ell}^{0} > m_{b\ell}^{1}$.

The $m_{b\ell}^{\text{asym}}$ associated with the b-ℓ pairs selected by the algorithm is used as a key discriminating variable in the analysis, because signal events yield a smaller invariant mass asymmetry than the dominant SM backgrounds, as confirmed by the shape comparison in Figure 7.4a. In addition, two important variables are the leading invariant mass of the b-ℓ pair selected by the pairing algorithm, $m_{b\ell}^{0}$, and the sub-leading invariant mass of the rejected b-ℓ pair, $m_{b\ell}^{1}(\text{rej})$, whose distributions are shown in Figures 7.4b and 7.4c in the same signal and background processes. In particular, $m_{b\ell}^{0}$ is able to reconstruct the value of the $\tilde{t}$ mass in signal events with remarkable accuracy, while $m_{b\ell}^{1}(\text{rej})$ yields no

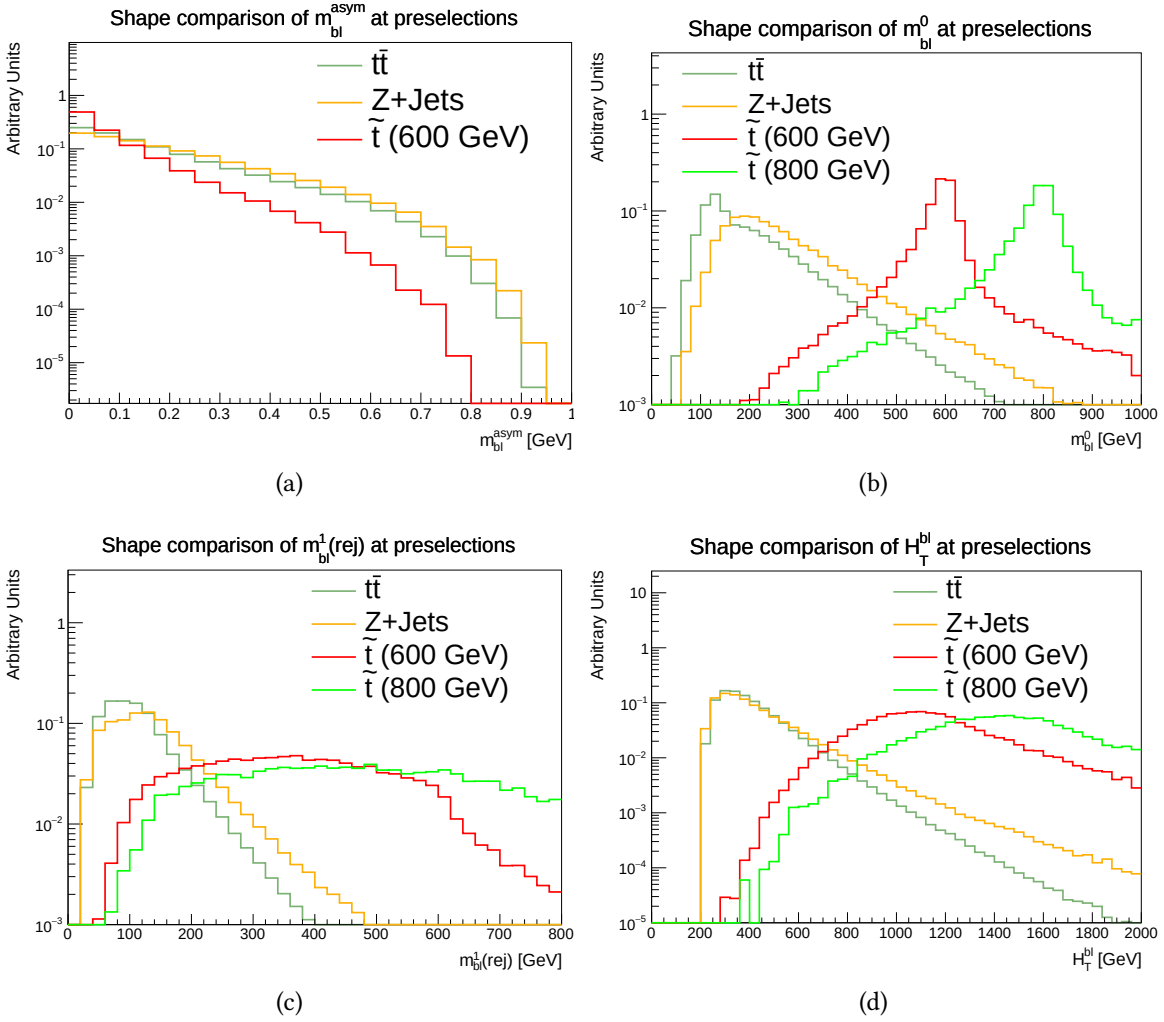


Figure 7.4: Distributions of $m_{b\ell}^{\text{asym}}$ (a), $m_{b\ell}^{0}$ (b), $m_{b\ell}^{1}(\text{rej})$ (c) and $H_{\text{T}}^{b\ell}$ after the preliminary event selections in Table 7.1 for $t\bar{t}$, Z+jets and two signals models with $\text{BR}(e) = \text{BR}(\mu) = 50\%$ and $\tilde{t}$ masses of 600 and 800 GeV.

peak in signal but can be used to obtain a significant suppression of the SM backgrounds, especially when applied to $t\bar{t}$.

Other variables

Together with the invariant mass variables described above, the following quantities are also employed:

$\boldsymbol{H_{\text{T}}^{b\ell}}$ A variation of the jet-based H_{T} introduced in Section 6.2.1 is defined as follows:

$$H_{\text{T}}^{b\ell} = p_{\text{T}}^{j1} + p_{\text{T}}^{j2} + p_{\text{T}}^{\ell 1} + p_{\text{T}}^{\ell 2} \tag{7.2}$$

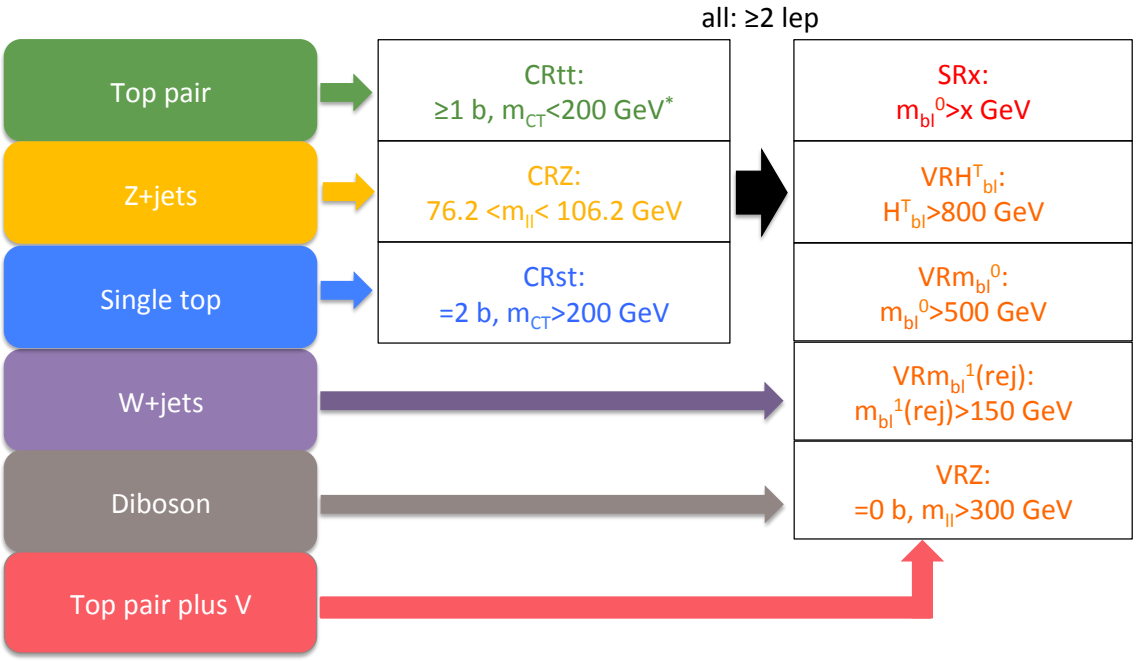
using the leading two jets and the leading two leptons in each event. As illustrated by the shape comparison between signal and background processes in Figure 7.4d, the values of $H_{\mathrm{T}}^{b\ell}$ are expected to be significantly larger in signal events, where the four objects included in the computation originate from the decay of heavy particles.

$\boldsymbol{m_{\ell\ell}}$ The invariant mass of the leading two leptons is employed in the e-e and μ-μ channels of the analysis to suppress Z+jets events, where it reconstructs the mass of the Z boson.

$\boldsymbol{m_{\mathrm{CT}}}$ The contransverse mass m_{CT}, previously defined in Equation 6.7, is computed in the present analysis using the two leading jets in the events, and is used for the definition of a CR for single top production as further described in Section 7.3.1.

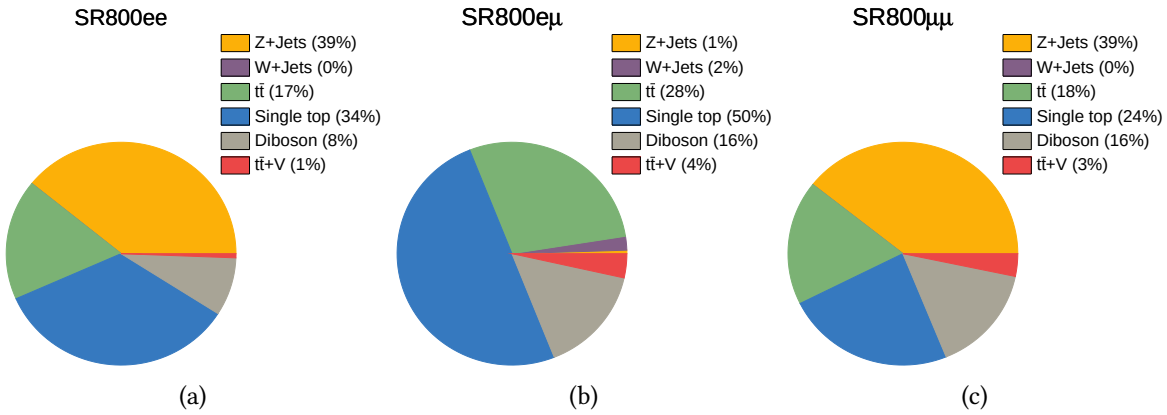
7.2.3 Signal Region definitions

The kinematical distributions in Figure 7.4 show that the set of discriminating variables introduced in Section 7.2.2 can be used to obtain a major suppression of the SM backgrounds whilst retaining a large fraction of signal events, allowing to define SRs sensitive to heavy $\tilde{t}$ models with low cross sections. In order to target these models, the SRs must necessarily be defined with large signal acceptance and detector efficiency (previously defined in Section 6.4.3), which means that any selection that would cause a non-negligible loss of signal events should be avoided even if it is coherent with the properties of the signal process of interest. For example, while it would seem natural to select events with two b-jets based on the experimental signature of the $\tilde{t}$ signal under exam, the b-tagging requirement in the SRs is relaxed to $\geq$1 b-tag (with 77% fixed cut WP) among the two leading jets to cope with the low efficiency of the MV2c10 algorithm at high jet p_{T}, particularly significant in processes where the b-jets are not initiated by the decay of a top quark (see Figure 4.1a).
The detail of the SR selections is summarised in Table 7.2, which includes also the CRs and VRs further defined in Sections 7.3.1 and 7.3.2. A common selection of $m_{b\ell}^{\mathrm{asym}} < 0.2$ is applied to all regions to select events with signal-like properties, while the dominant SM backgrounds are suppressed by requiring $H_{\mathrm{T}}^{b\ell} > 1$ TeV, $m_{b\ell}^{1}(\mathrm{rej}) > 150$ GeV and $m_{\ell\ell} > 300$ GeV, obtaining a sample of events where the leading two leptons and jets in p_{T} are strongly boosted. Finally, two overlapping SRs are defined with $m_{b\ell}^{0} > 800, 1100$ GeV, and are labelled as SR800 and SR1100.



Region	N_b	$m^0_{b\ell}$ [GeV]	$H_{\mathrm{T}}^{b\ell}$ [GeV]	$m^1_{b\ell}$(rej) [GeV]	$m_{\ell\ell}$ [GeV]	m_{CT} [GeV]
SR800	≥ 1	> 800	> 1000	> 150	> 300	–
SR1100	≥ 1	> 1100	> 1000	> 150	> 300	–
CRst	$= 2$	[200,500]	< 800	< 150	> 120	> 200
CRtt	≥ 1	[200,500]	[600,800]	< 150	> 300	$< 200^*$
CRZ	≥ 1	> 700	> 1000	–	[76.2,106.2]	–
VR$m^0_{b\ell}$	≥ 1	> 500	[600,800]	< 150	> 300	–
VR$m^1_{b\ell}$(rej)	≥ 1	[200,500]	[600,800]	> 150	> 300	–
VR$H_{\mathrm{T}}^{b\ell}$	≥ 1	[200,500]	> 800	< 150	> 300	–
VRZ	$= 0$	[500,800]	> 1000	> 150	> 300	–

Table 7.2: Summary of the definitions of the SRs, CRs and VRs of the analysis, after the preliminary event selections in Table 5.3.1. In addition to the above selections, a requirement of $m_{b\ell}^{\mathrm{asym}} < 0.2$ is applied to all regions. The $m_{\ell\ell}$ selection is used only in the e-e and μ-μ channels of the analysis, while the m_{CT} selection in CRtt is only applied to events with exactly two b-jets to ensure that the region is orthogonal to CRst, as indicated by the * in the tables.



(a) (b) (c)

Figure 7.5: Background composition in the e-e (a), e-μ (b) and μ-μ (c) channels of SR800, determined using MC simulations.

The breakdown of the background yields in the e-e, e-μ and μ-μ channels of SR800 is presented in Figure 7.5, to be compared with the corresponding preselection diagrams in Figure 7.3. The fraction of the single top and diboson events in SR800 is significantly enhanced with respect to the corresponding impact at preselection level, while the contribution from $t\bar{t}$ production is reduced. It is important to mention that the overall background prediction in SR800 amounts to ~ 5 events, and is further lowered to ~ 1 event in SR1100 (see Table 7.5), resulting in a relatively large uncertainty from MC statistics on the yield predictions of the individual background processes.
Table 7.3 presents the SR800 and SR1100 signal cutflows for three alternative values of the $\tilde{t}$ mass ($m_{\tilde{t}} = 800, 1200, 1500$ GeV) and for a mixed decay scenario with $\text{BR}(e) = \text{BR}(\mu) = \text{BR}(\tau) = 1/3$. The table confirms that the individual selections are highly efficient when applied to the signal models under exam, with the exception of the $m^0_{b\ell}$ selection, whose efficiency depends on the relative difference between the cut threshold and the $\tilde{t}$ mass. The large overall selection efficiency of the SRs has a positive impact on the sensitivity of the analysis to heavy $\tilde{t}$ signals with low production cross sections, as further discussed in Section 7.4.2.

7.3 Background estimation

Similarly to what is done in the $\tilde{b}$ search in Chapter 6, the normalisation of the dominant SM backgrounds is derived from a set of CRs, and the accuracy of the background predictions is tested in VRs. The summary of the CR and VR definitions is shown in

Selection	$m_{\tilde{t}} = 800$ GeV yield (eff)	$m_{\tilde{t}} = 1.2$ TeV yield (eff)	$m_{\tilde{t}} = 1.5$ TeV yield (eff)
Total (36.1 fb^{-1})	1022	57	9
Production filter	737 (72.0%)	42.6 (73.8%)	6.9 (74.7%)
Trigger	705 (95.7%)	41.0 (96.3%)	6.7(96.3%)
Jet cleaning	704 (99.9%)	40.9 (99.8%)	6.7 (99.8%)
Muon cleaning	704 (100.0%)	40.9 (100.0%)	6.6 (100.0%)
2 signal leptons	385 (54.7%)	21.7 (53.0%)	3.5 (52.2%)
Trigger matching	385 (99.8%)	21.6 (99.8%)	3.5 (99.8%)
Opposite-charge ℓ	375 (97.5%)	20.9 (96.8%)	3.3 (96.5%)
2 signal jets	363 (96.9%)	20.6 (98.2%)	3.3 (98.7%)
≥ 1 b-tagged jets	285 (78.5%)	14.3 (69.6%)	2.0 (62.3%)
$m_{b\ell}^{\text{asym}} < 0.2$	245 (85.8%)	12.2 (85.3%)	1.8 (86.3%)
$H_{\text{T}}^{b\ell} > 1000$ GeV	228 (92.9%)	12.1 (99.4%)	1.8 (99.7%)
$m_{\ell\ell} > 300$ GeV	199 (87.6%)	11.5 (94.6%)	1.7 (96.4%)
$m_{b\ell}^{1}(\text{rej}) > 150$ GeV	195 (97.8%)	11.4 (99.4%)	1.7 (99.9%)
$m_{b\ell}^{0} > 800$ GeV	81.2 (41.6%)	10.7 (93.4%)	1.6 (96.3%)
$m_{b\ell}^{0} > 1100$ GeV	4.4 (2.3%)	8.4 (73.6%)	1.5 (89.5%)

Table 7.3: Cutflow tables for the inclusive SR800 and SR1100 regions, showing the MC expected yields and the relative efficiencies at each selection stage for three signal samples with $m_{\tilde{t}} = 800, 1200, 1500$ GeV and for BR(e) = BR(μ) = BR$(\tau) = 1/3$ [229]. The production filter is a technical step that applies a basic trigger requirement and selects events with at least two electrons or muons with loose p_{T} and η thresholds (uncalibrated $p_{\text{T}} > 9$ GeV and $|\eta| < 2.6$). The relative efficiencies for the $m_{b\ell}^{0} > 1100$ GeV selection are shown with respect to the $m_{b\ell}^{1}(\text{rej}) > 150$ GeV requirement.

Table 7.2 together with the previously defined SRs, while a detailed description of their features is provided in Sections 7.3.1 and 7.3.2.

7.3.1 Control Regions

Based on the SR background composition shown in Figure 7.5, three CRs are employed to measure the normalisation factors of single top, $t\bar{t}$ and Z production, while the diboson, $t\bar{t} + V$, and W backgrounds are estimated directly from MC simulation. The CR definitions are described below:

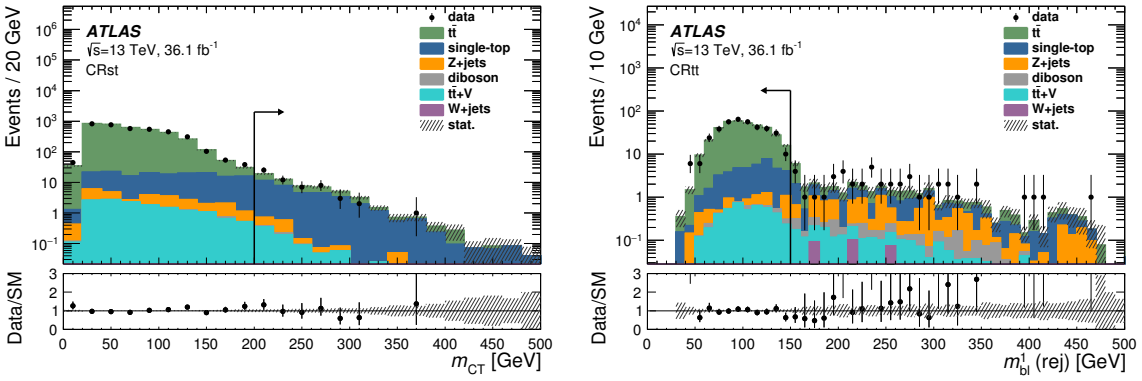
CRst A first CR is defined for the single top background, which is dominated by the Wt channel illustrated by the diagram in Figure 5.12c. The SR selections on $H_{\mathrm{T}}^{b\ell}$, $m_{b\ell}^{0}$ and $m_{b\ell}^{1}(\mathrm{rej})$ are reversed as detailed in Table 7.2, while the $m_{\ell\ell}$ threshold is lowered to 120 GeV, yielding a large data sample dominated by top production and with negligible contamination from signal events. In addition, due to the higher cross section of $t\bar{t}$ production compared to single top, an extra selection on the contransverse mass m_{CT} (introduced in Section 6.2.1) is necessary to improve the single top purity in CRst, by exploiting the known kinematical endpoint in $t\bar{t}$ events. Since the endpoint is only present if the variable is computed with the two b-jets from the top decay, the b-tagging selection in CRst is applied to both leading jets in the events, while the remaining jets are required to be light. The performance of m_{CT} is illustrated by its N-1 distribution in CRst, shown in Figure 7.6a: a fraction of single top events above 55% is observed for $m_{\mathrm{CT}} > 200$ GeV, which corresponds to the CR threshold.

CRtt A CR for $t\bar{t}$ production is defined by inverting the $H_{\mathrm{T}}^{b\ell}$, $m_{b\ell}^{0}$ and $m_{b\ell}^{1}(\mathrm{rej})$ requirements with respect to the SR, similarly to what is done in CRst. In addition, an upper bound on m_{CT} is applied specifically to the events where two b-tagged jets are present, to ensure that CRtt and CRst are mutually exclusive. Figure 7.6b shows the N-1 distribution of $m_{b\ell}^{1}(\mathrm{rej})$ in CRtt, which confirms that the majority of $t\bar{t}$ events lie below the CR threshold.

CRZ Finally, a CR for Z+jets production is defined by selecting events with invariant mass $m_{\ell\ell}$ close to the mass of the Z boson. The remaining SR selections are left unchanged, with the only exception of $m_{b\ell}^{0}$ which is relaxed to > 700 GeV to enhance the CRZ event yield.

7.3.2 Validation Regions

Four mutually exclusive VRs are implemented to verify the CR-SR extrapolation in different regions of phase space, as detailed in Table 7.2. Three of them are designed to test the extrapolation from CRst and CRtt to the SRs in the $m_{b\ell}^{0}$, $m_{b\ell}^{1}(\mathrm{rej})$, and $H_{\mathrm{T}}^{b\ell}$ variables, and are labelled accordingly as VR$m_{b\ell}^{0}$, VR$m_{b\ell}^{1}(\mathrm{rej})$ and VR$H_{\mathrm{T}}^{b\ell}$. The background yield in these VRs is generally dominated by $t\bar{t}$, with a non-negligible contribution of single top in VR$m_{b\ell}^{1}(\mathrm{rej})$, and the region definitions are optimised to minimise the contamination from signal events.



(a) (b)

Figure 7.6: Kinematical distribution of m_{CT}in CRst (a) and $m^1_{b\ell}(\mathrm{rej})$in CRtt (b) where all selections are applied except the one on the displayed variable [19]. The background normalisations are derived from the background only fit, and the arrows indicate the threshold of the CR selections.

In parallel, a fourth VR is defined to test the extrapolation on the di-leptonic invariant mass from CRZ, where the Z peak is selected, to the SRs, where a requirement of $m_{\ell\ell} > 300$ GeV is present. This is done by selecting events with no b-tagged jets, obtaining a simultaneous suppression of signal and top production processes, while keeping the remaining selections close to the corresponding ones in CRZ and in the SRs.
The kinematical distributions of $m^0_{b\ell}$, $m^1_{b\ell}(\mathrm{rej})$, $H_{\mathrm{T}}^{b\ell}$ and $m_{\ell\ell}$ are shown in Figure 7.7 in the four VRs described above, using the normalisation factors of $t\bar{t}$, single top and Z production derived from the corresponding CRs. Good agreement is observed between data and SM predictions in all VRs, confirming the accuracy of the background estimation in the regions of phase space targeted by the analysis.

7.3.3 Systematic uncertainties

The impact of the dominant experimental and theoretical systematics in the SRs of the analysis is presented in Table 7.4, based on the sources of uncertainty introduced in Section 5.3.4.
The size of the variations caused by experimental effects is found to be at most 10%, where the dominant sources include b-tagging, JES, JER and lepton reconstruction. The uncertainty on modelling of the backgrounds has a larger impact, up to 45% relative to the total background yield for $t\bar{t}$ in SR1100, due to the finite amount of MC statistics in

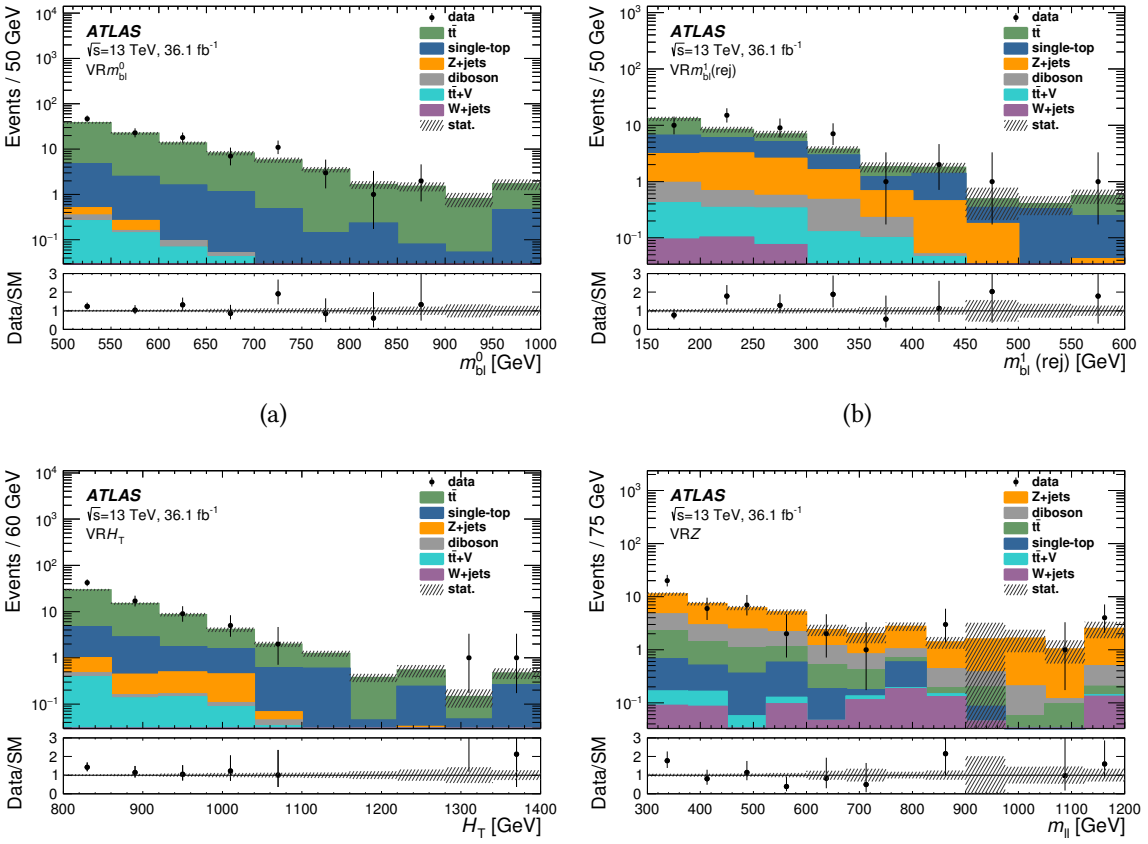


Figure 7.7: Kinematical distributions of $m^0_{b\ell}$ in VR$m^0_{b\ell}$ (a), $m^1_{b\ell}(\mathrm{rej})$ in VR$m^1_{b\ell}(\mathrm{rej})$ (b), $H^{b\ell}_{\mathrm{T}}$ in VR$H^{b\ell}_{\mathrm{T}}$ (c) and $m_{\ell\ell}$ in VRZ (d) after the background only fit [19].

the tight region of phase space selected by the SRs.

It is important to mention that, due to the low background predictions in both SRs of the analysis (see Table 7.5), the above variations do not exceed the size of the Poisson uncertainty on the expected SM yields, so the sensitivity of the search is not spoiled by their presence.

Finally, the dominant uncertainty on the expected signal yields is related to the cross section calculation, and ranges between 13% and 27% for $\tilde{t}$ masses between 600 GeV and 1.6 TeV [179].

Source \ Region	SR800	SR1100
Experimental uncertainty		
b-tagging	3%	5%
Jet energy resolution	2%	10%
Jet energy scale	1%	3%
Electrons	1%	4%
Muons	1%	3%
Theoretical modelling uncertainty		
MC statistical uncertainty	8%	17%
$t\bar{t}$	8%	45%
Single top	21%	22%

Table 7.4: Summary of the dominant experimental and theoretical systematics in SR800 and SR1100 before the background only fit, quoted relative to the total pre-fit MC background predictions [19]. The individual variations can be correlated, so the total uncertainty on the background is not necessarily obtained by summing them in quadrature.

7.4 Results and interpretation

Similarly to the $\tilde{b}$ search in Chapter 6, the results of the analysis are obtained by performing a background only fit (Section 7.4.1) and, in the absence of a significant excess beyond the SM predictions, by setting exclusion limits on generic BSM models and on the benchmark SUSY signal in Figure 7.1 (Section 7.4.2).

7.4.1 Background only fit

The SM predictions in the VRs and SRs of the analysis are derived by performing a background only fit in the three CRs defined in Section 7.3.1, which yields the following normalisation factors for the single top, $t\bar{t}$ and Z processes:

- $\mu_{st} = 1.10 \pm 0.27$.
- $\mu_{t\bar{t}} = 0.94 \pm 0.06$.
- $\mu_{Z} = 1.22 \pm 0.18$.

The final results are summarised in Table 7.5, where the observed counts and the post-fit SM predictions in SR800 and SR1100 are shown together with the breakdown of the yields in the different lepton flavour channels. No significant excess is observed in the SRs: only 2 events are found in SR800, both in the symmetric μ-μ channel, and one of them passes also the SR1100 selections.
The kinematical distributions of $m_{b\ell}^{0}$, $m_{b\ell}^{\mathrm{asym}}$, $H_{\mathrm{T}}^{b\ell}$, $m_{\ell\ell}$, and $m_{b\ell}^{1}(\mathrm{rej})$ in SR800 are shown in Figure 7.8 in the usual N-1 format, where the selection on the displayed variable is removed and is indicated by an arrow. The event observed in SR1100 is found to have a particularly large value of $H_{\mathrm{T}}^{b\ell}$, which is caused by the presence of a high-p_{T} muon whose measured momentum carries a large uncertainty.

7.4.2 Exclusion limits

Together with the expected and observed yields in each SR, Table 7.5 also shows the corresponding model independent limits on the number of events from generic BSM processes, as well as those on the visible cross section σ_{vis} previously defined in Section 6.4.3. In addition, exclusion limits are specifically set on the signal model in Figure 7.1 using the CL$_s$ method introduced in Section 5.4.2, by considering different $\tilde{t}$ masses and different values of BR(e), BR(μ) and BR(τ) with unit sum.
Figure 7.9 shows the exclusion contours in the BR(e)-BR(τ) plane for various $\tilde{t}$ mass hypotheses, obtained using the best expected CL$_s$ among the inclusive and individual e-e, e-μ and μ-μ channels of the SRs for each signal benchmark. Models with BR(τ) up to 80% or more are excluded for $\tilde{t}$ masses in the 600-1000 GeV range, while the limits reach $\tilde{t}$ masses up to 1.5 TeV for larger values of BR(e) and BR(μ). The weak constraints at large BR(τ) are due to the absence of SRs with τ leptons, while the limits are slightly stronger for large BR(e) compared to large BR(μ) due to the higher efficiency of the single electron triggers.
The remarkable performance of the analysis is confirmed by the ability to exclude signal models with $\tilde{t}$ masses up to 1.5 TeV, that have a total production cross section of 0.26 fb ($\pm 24\%$) [179] which translates into a nominal yield prediction of only 9 events in the full 36.1 fb^{-1} dataset. This means that the SR selections are able to remove almost completely the SM backgrounds with a small impact on the signal acceptance, as already indicated by the signal cutflows presented in Table 7.3. For reference, the analysis in Chapter 6 is sensitive to $\tilde{b}$ masses up to about 1 TeV (see Figure 6.23), corresponding to a cross section of 6.1 fb ($\pm 16\%$) and a total expected yield of around 220 events. The better sensitivity

SR800

	inclusive	ee	$e\mu$	$\mu\mu$
Observed yield	2	0	0	2
Total post-fit bkg yield	5.2 ± 1.4	1.8 ± 0.5	2.1 ± 0.8	1.35 ± 0.32
Post-fit single top yield	2.0 ± 1.3	0.6 ± 0.4	1.1 ± 0.7	0.32 ± 0.20
Post-fit Z yield	1.40 ± 0.33	0.80 ± 0.24	0.01 ± 0.01	0.59 ± 0.14
Post-fit $t\bar{t}$ yield	1.0 ± 0.5	0.27 ± 0.14	0.54 ± 0.25	0.21 ± 0.10
Post-fit diboson yield	0.64 ± 0.23	0.14 ± 0.05	0.31 ± 0.12	0.19 ± 0.08
Post-fit $t\bar{t}V$ yield	0.12 ± 0.03	0.01 ± 0.01	0.07 ± 0.02	0.04 ± 0.02
Post-fit W yield	0.03 ± 0.03	–	0.04 ± 0.04	–
Total MC bkg yield	4.9 ± 1.2	1.7 ± 0.4	2.0 ± 0.7	1.23 ± 0.28
S^{95}_{exp}	$6.4^{+3.0}_{-1.9}$	$4.1^{+1.8}_{-1.1}$	$4.0^{+2.2}_{-0.9}$	$3.9^{+1.6}_{-0.7}$
S^{95}_{obs}	4.0	3.0	3.0	4.8
σ_{vis}[fb]	0.11	0.08	0.08	0.13

SR1100

	inclusive	ee	$e\mu$	$\mu\mu$
Observed yield	1	0	0	1
Total post-fit bkg yield	$1.2^{+0.6}_{-0.5}$	$0.51^{+0.22}_{-0.20}$	$0.44^{+0.39}_{-0.33}$	0.22 ± 0.13
Post-fit single top yield	0.32 ± 0.29	0.11 ± 0.10	0.21 ± 0.19	–
Post-fit Z yield	0.47 ± 0.15	0.28 ± 0.10	–	0.19 ± 0.11
Post-fit $t\bar{t}$ yield	$0.21^{+0.55}_{-0.21}$	$0.06^{+0.16}_{-0.06}$	$0.13^{+0.34}_{-0.13}$	$0.01^{+0.03}_{-0.01}$
Post-fit diboson yield	0.13 ± 0.05	0.06 ± 0.03	0.07 ± 0.03	0.01 ± 0.01
Post-fit $t\bar{t}V$ yield	0.03 ± 0.01	–	0.01 ± 0.01	0.01 ± 0.01
Post-fit W yield	$0.01^{+0.02}_{-0.01}$	–	$0.01^{+0.02}_{-0.01}$	–
Total MC bkg yield	$1.1^{+0.6}_{-0.5}$	$0.46^{+0.21}_{-0.19}$	$0.43^{+0.40}_{-0.33}$	0.18 ± 0.10
S^{95}_{exp}	$3.9^{+2.4}_{-0.5}$	$3.0^{+1.3}_{-0.0}$	$3.0^{+1.3}_{-0.0}$	$3.1^{+0.6}_{-0.1}$
S^{95}_{obs}	3.9	3.0	3.1	4.1
σ_{vis}[fb]	0.11	0.08	0.08	0.11

Table 7.5: Summary of the results of the analysis in SR800 (top) and SR1100 (bottom) after the background only fit, for the inclusive and individual flavour channels [19]. The tables include the expected and observed limit on the number of signal events (S^{95}_{exp} and S^{95}_{obs}) and the observed limits on the visible cross sections (σ_{vis}) of generic BSM models.

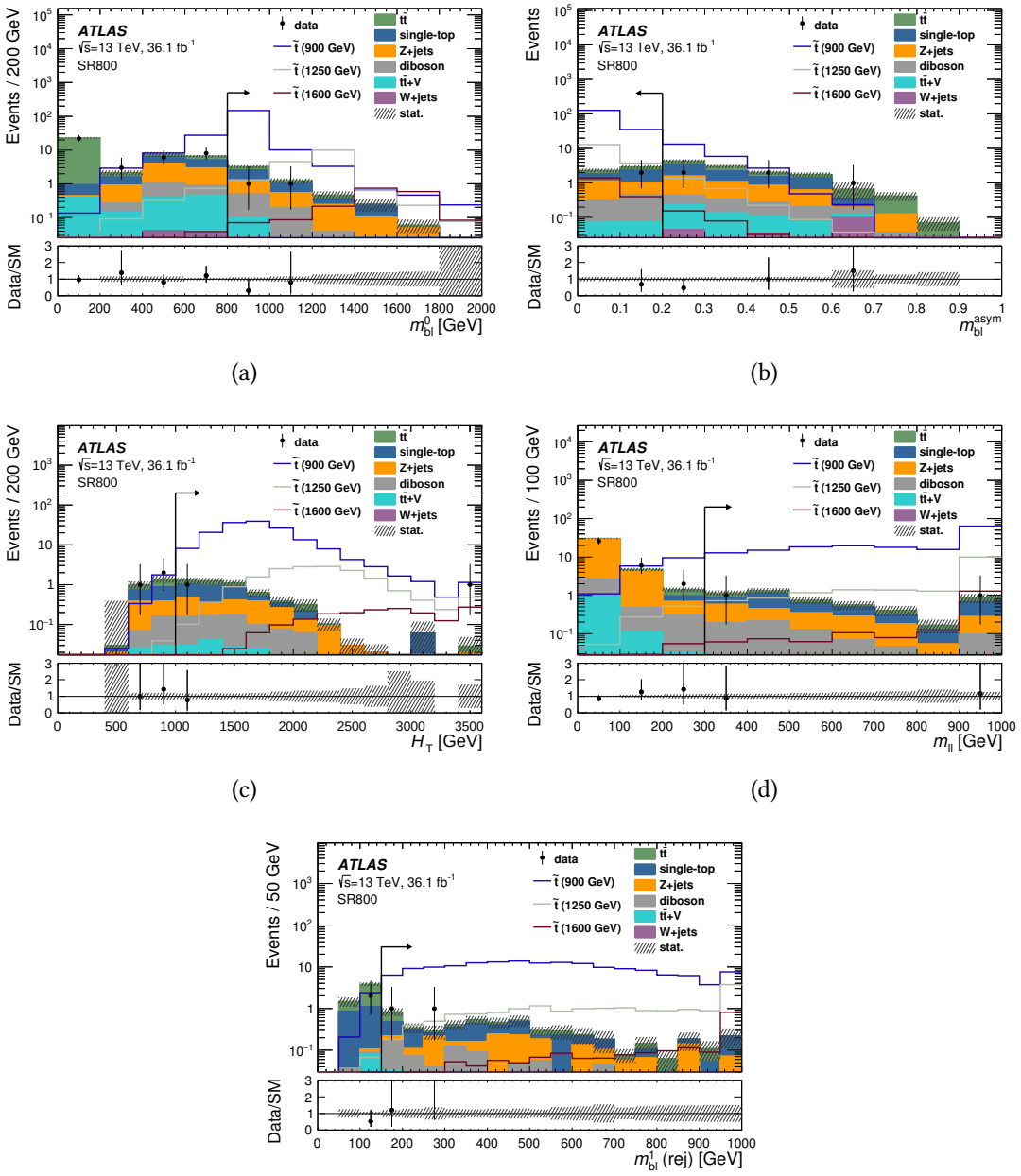


Figure 7.8: Kinematical distributions of $m^0_{b\ell}$ (a), $m^{\mathrm{asym}}_{b\ell}$ (b), $H^{b\ell}_{\mathrm{T}}$ (c), $m_{\ell\ell}$ (d), and $m^1_{b\ell}(\mathrm{rej})$ (e) in SR800 after the background only fit, where all selections are applied except the one on the displayed variable [19]. For comparison, the signal models generated with $\tilde{t}$ masses of 900, 1250, and 1600 GeV and with 50% BR into b-e and b-μ are included.

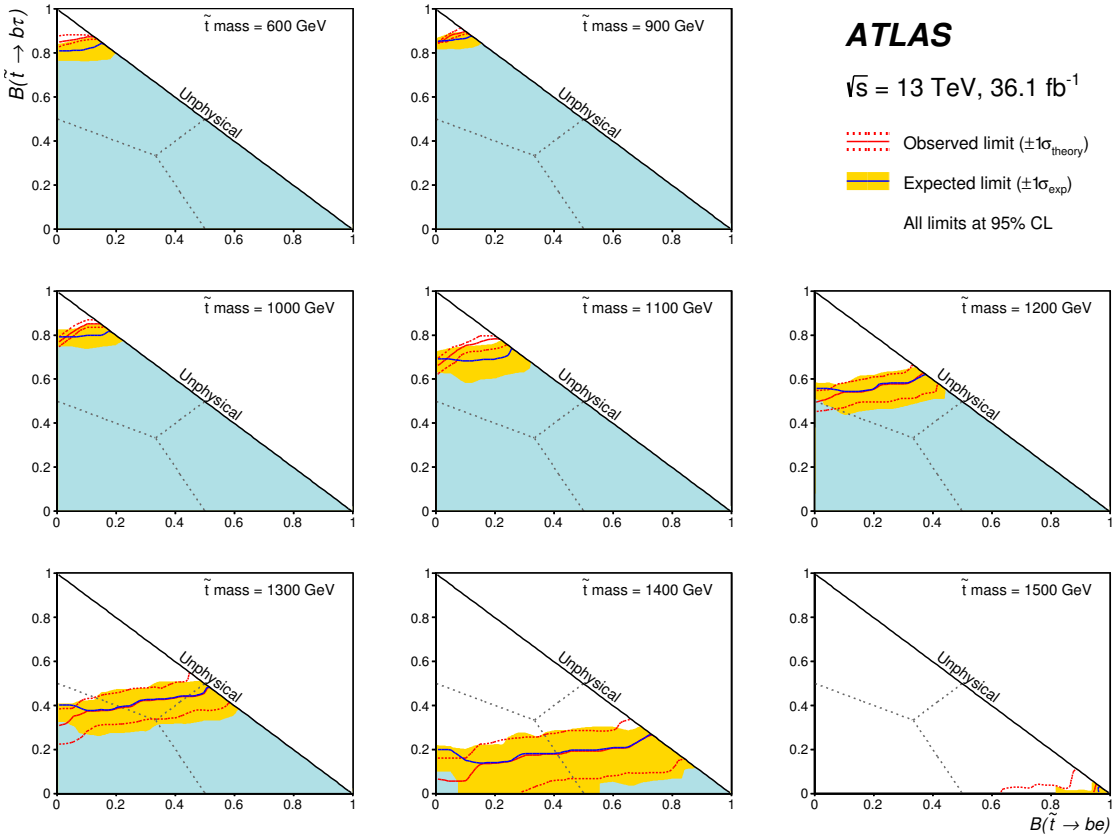


Figure 7.9: Expected and observed limits at 95% CL on the $\tilde{t}$ signal in Figure 7.1 in the BR(e)-BR(τ) plane with $\mathrm{BR}(e)+\mathrm{BR}(\mu)+\mathrm{BR}(\tau) = 1$, for $\tilde{t}$ masses between 600 GeV and 1.5 TeV [19]. For each combination of BR(e), BR(μ) and BR(τ) under exam, the limits are based on the best expected CL_s among the different flavour channels of the SRs.

of the present search is achieved thanks to the striking properties of the targeted signal, already highlighted by the kinematical distributions in Figures 7.4 and 7.8.
Finally, the information from the individual plots in Figure 7.9 is summarised in Figure 7.10, which shows the highest $\tilde{t}$ mass excluded for each reference point in the BR(e)-BR(τ) plane.

7.5 Summary

This chapter presented the results of the first ATLAS publication targeting a pair-produced $\tilde{t}$ that decays through an RPV coupling into b-ℓ pairs with different lepton flavours, based on 36.1 fb^{-1} of data at $\sqrt{s} = 13$ TeV. Two SRs are defined by selecting events with at

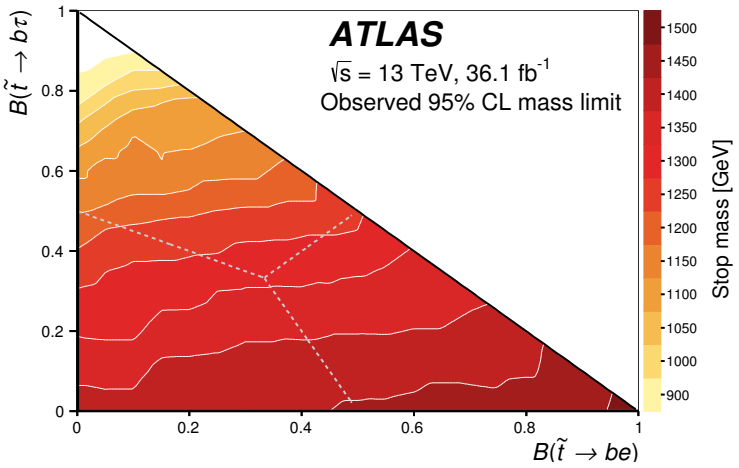


Figure 7.10: Observed lower limits [19] at 95% CL on the $\tilde{t}$ mass in the BR(e)-BR(τ) plane with $\mathrm{BR}(e) + \mathrm{BR}(\mu) + \mathrm{BR}(\tau) = 1$.

least two leptons (e or μ) and two jets, one of which is required to be b-tagged, and no significant excess is observed beyond the SM predictions. As a result, exclusion limits are placed on the signal model in Figure 7.1 for different $\tilde{t}$ masses and BR hypotheses, ranging between 600 GeV and 1 TeV up to large BRs into b-τ, and reaching up to 1.5 TeV for 100% BR into b-e. These results significantly extend the previous constraints from a preliminary ATLAS analysis at $\sqrt{s} = 8$ TeV [226] that targeted the same signal model.

8 | Conclusions and outlook

The core subject of the thesis is the search for the direct pair-production of third generation squarks, the $\tilde{t}$ and the $\tilde{b}$, in proton-proton collisions at $\sqrt{s} = 13$ TeV delivered by the LHC and recorded by the ATLAS detector. Third generation squarks are introduced in the context of natural SUSY models that are able to solve the Higgs hierarchy problem, and the corresponding phenomenology at the LHC is examined under different assumptions on their decay modes in both RPC and RPV scenarios.

The baseline b-tagging algorithm of the ATLAS experiment, MV2c10, plays an essential role in the searches, because the targeted signals are all characterised by the presence of two bottom quarks in the final states. The studies in Chapter 4 allow to measure the performance of MV2c10 in different physics processes for a wide range of kinematical regimes of the jets, and provide an estimate of the impact of several detector-related systematic uncertainties on the tagging efficiency as a function of the jet p_{T}.

The first major result of the thesis, presented in Chapter 6, is the search for a pair-produced $\tilde{b}$ decaying into a b-quark and a $\tilde{\chi}_1^0$ or a t-quark and a $\tilde{\chi}_1^+$, with negligible mass splitting between $\tilde{\chi}_1^0$ and $\tilde{\chi}_1^+$ as predicted by naturalness arguments. The analysis is divided into two mutually exclusive channels based on the signal lepton multiplicity: a set of zero-lepton SRs is designed for the $\tilde{b}$ decay into $b\tilde{\chi}_1^0$ with 100% BR, while additional one-lepton SRs are optimised for the mixed decay of the $\tilde{b}$ into $b\tilde{\chi}_1^0$ or $t\tilde{\chi}_1^+$ with 50% BR each. When targeting signals with large $\Delta m(\tilde{b}, \tilde{\chi}_1^0)$, the dominant background from $t\bar{t}$ production is suppressed using dedicated discriminating variables that yield a kinematical endpoint in $t\bar{t}$ events, namely m_{CT} [216] in the zero-lepton channel and am_{T2} [217] in the one-lepton channel. Variables related to the transverse mass of jets and E_T^{miss} are found to be optimal to target signals with intermediate $\Delta m(\tilde{b}, \tilde{\chi}_1^0)$, while the compressed region in the $\tilde{b}$-$\tilde{\chi}_1^0$ mass plane is probed by an additional zero-lepton SR that requires the presence of a jet from ISR recoiling against two b-jets with low p_{T} and the E_T^{miss}.

Subsequently, Chapter 7 presents a search for a pair-produced $\tilde{t}$ with RPV decay into a b-quark and a lepton, with variable BR into electrons, muons and τ leptons. The analysis

is optimised exclusively for final states containing electrons and muons, and a dedicated algorithm is developed to identify the b-ℓ pairs from the two decay legs in signal events, allowing to reconstruct the invariant mass of the $\tilde{t}$ with good accuracy. Additional selections on the kinematical properties of jets and leptons in the events are employed, obtaining a good sensitivity to heavy $\tilde{t}$ signals.
The results of the two searches described above are found to be consistent with the SM background predictions, and are used to set stringent exclusion limits on the relevant signal scenarios. In particular, the analysis in Chapter 6 excludes $\tilde{b}$ masses up to 950 GeV for $\tilde{\chi}_1^0$ masses below 420 GeV for a $\tilde{b}$ decay into b-$\tilde{\chi}_1^0$ at 100% BR, while the limit reaches 860 GeV in $\tilde{b}$ mass for $\tilde{\chi}_1^0$ masses below 250 GeV for the mixed decay of the $\tilde{b}$ into b-$\tilde{\chi}_1^0$ or t-$\tilde{\chi}_1^+$ with 50% BR each. Finally, the search in Chapter 7 is used to set limits on a pair-produced $\tilde{t}$ decaying into b-ℓ pairs: BRs into b-τ up to 80% or more are excluded for $\tilde{t}$ masses in the 600-1000 GeV range, while $\tilde{t}$ masses up to 1.5 TeV are excluded for 100% BR into b-e.

8.1 Future prospects of SUSY searches

Similarly to the analyses described in this thesis, none of the searches conducted by ATLAS and CMS so far were able to find evidence for signals compatible with the production of SUSY particles, despite an extensive effort by the two Collaborations aimed at covering a wide variety of signal hypotheses. Good agreement with the SM expectations was also observed in the measurements of the properties of the newly discovered Higgs boson [28], confirming the remarkable predictivity of the SM but yielding no clear indication about possible solutions to the shortcomings described in Chapter 1.
When interpreting the results of SUSY searches, it is important to remember that the limits on simplified benchmark processes are not necessarily representative of the constraints on full SUSY models with many free parameters, which may yield complex signal configurations that spoil the sensitivity of the existing searches. A reinterpretation of the limits of 22 SUSY analyses in the context of a 19-parameter pMSSM was published at the end of Run 1 [182], while no equivalent result is available yet for Run 2. For this reason, at the time of writing it is not trivial to determine to what extent the present constraints are stressing the idea of SUSY as a solution to the Higgs hierarchy problem or to other open issues in high energy physics. Moreover, it is clear that the exclusion limits on a finite number of signal models are not sufficient to rule out the existence of a broken SUSY in nature, because the theory is built with an arbitrary number of free

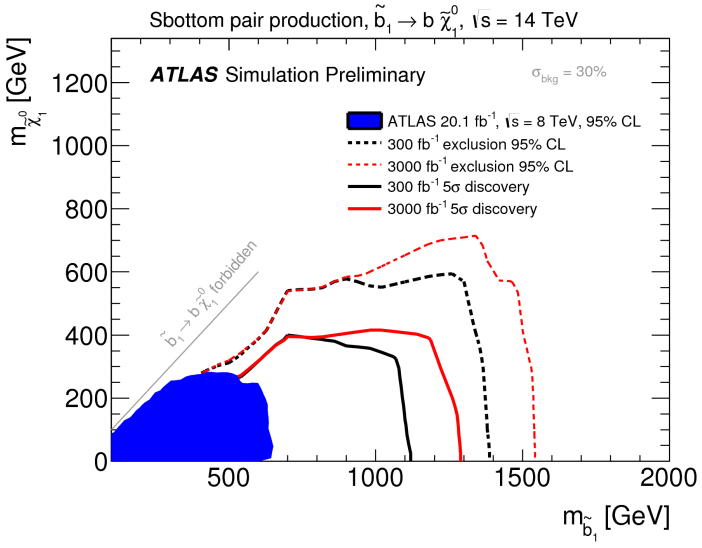


Figure 8.1: Expected exclusion limits and discovery reach for a pair-produced $\tilde{b}$ decaying into $b\tilde{\chi}_1^0$ with 100% BR, for an integrated luminosity of 300 fb^{-1} and 3000 fb^{-1} at $\sqrt{s} = 14$ TeV [230].

parameters that can be arranged to escape any level of constraints.

At present SUSY remains a beautiful theoretical framework that can embed elegant solutions to the shortcomings of the SM, and its broad phenomenology will continue to be tested at the LHC in the upcoming years. The data recorded by ATLAS during 2017 at $\sqrt{s} = 13$ TeV increased the size of the Run 2 dataset to a total of approximately 80 fb^{-1} and, together with the next periods of data taking, will allow to extend the sensitivity of the searches beyond the current exclusion limits. As an example, Figure 8.1 shows the exclusion and discovery prospects for the case of a pair-produced $\tilde{b}$ decaying into $b\tilde{\chi}_1^0$, based on early simulation studies carried out before the beginning of Run 2 [230]. The exclusion limits have already been significantly extended by the analysis in Chapter 6, but larger datasets will allow to probe additional regions of phase space, continuing the hunt for new physics at the energy frontier.

A | Summary of personal contributions

During my PhD I carried out my research activities as a member of the ATLAS Collaboration, working in close contact with inspiring teams of scientists from several countries and research institutes, and receiving constant support from the ATLAS group of the University of Sussex. My individual work has often been part of larger collective efforts which included contributions by many other researchers, and the results of the searches in Chapters 6 and 7 have been published on behalf of the full ATLAS Collaboration as envisaged by its policies. The purpose of this appendix is to provide the detailed list of my individual contributions to the results in the thesis:

Chapter 4 and Appendix C

b-tagging performance studies with Monte Carlo simulations

The analysis of the impact of experimental uncertainties on the efficiency of the MV2c10 b-tagging algorithm was performed in collaboration with researchers from the Tsinghua University, with guidance provided by the b-tagging calibration working group of the ATLAS Collaboration. I was involved in all phases of the research activity, developing essential parts of the software and producing the whole material presented in the thesis. The final estimate of the uncertainty was made available to the Collaboration as part of the official calibration file (CDI) for multiple WPs, and the studies are summarised in an ATLAS internal note.

Chapter 6

Search for bottom squarks with two-body RPC decays

The search for a pair-produced $\tilde{b}$ has been my core activity for all the duration of my PhD. I was one of the main analysers for two subsequent publications [18,209], working in close contact with researchers from the University of Liverpool, the

University of Sheffield and the IFAE research institute (Barcelona). I had a leading role in all major aspects of the work, including the simulation of signal benchmark samples, the SR optimisation, the definition of the background estimation strategy and the implementation of the software for the full analysis workflow. For both publications I followed the internal review of the analysis within the Collaboration through all its major steps, and for the 36.1 fb^{-1} analysis I personally produced the final results that are included in the paper together with their statistical interpretation, running all the background fits and setting all the exclusion limits on the signal models. I also prepared the auxiliary material of the paper, which is included in the HepData record [220].

Chapter 7

Search for top squarks with RPV decays into b-ℓ pairs

I also had a leading role in the search for a pair-produced $\tilde{t}$ decaying into b-ℓ pairs, in collaboration with researchers from the University of Pennsylvania. I developed the major aspects of the analysis strategy, working on the simulation of the signal benchmark samples and on the optimisation of SRs, CRs and VRs, and I implemented a software for the full analysis workflow, providing the necessary cross-checks to validate the final results that are included in the paper [19]. Finally, I implemented the truth-level version of the analysis for the HepData record associated with the publication [229].

Appendix B

Single electron trigger rate

In parallel with my core research activities described above, during my PhD I have been involved in various operational tasks within the electron-photon trigger working group of the ATLAS Collaboration. In particular, I have been the only author of the study of the single electron trigger rate as a function of the online E_{T} threshold presented in Appendix B, which resulted in a performance plot made public by the Collaboration [231].

B | Single electron trigger rate

This appendix presents a study of the rate of the isolated single electron trigger chain HLT_e26_lhtight_nod0_ivarloose as a function of its online E_{T} threshold, that I have carried out during my PhD as part of my involvement in the electron-photon trigger working group of the ATLAS Collaboration. The HLT_e26_lhtight_nod0_ivarloose chain is particularly relevant because, thanks to the background suppression provided by the HLT-level track-based isolation, it is the single electron chain with lowest E_{T} threshold that has been kept unprescaled during the entire 2016 run of the LHC, and is used by both searches in this thesis as well as by several more ATLAS analyses.

The rate of the chain is measured in a run during which the LHC has operated at a constant instantaneous luminosity value of $L_{\mathrm{inst}} = 8 \times 10^{33}$ cm^{-2}s^{-1} for a period of data taking of approximately 190 pb^{-1}, allowing precise measurements of the performance of the detector in stable conditions. The E_{T} threshold of the chain is varied from its minimum value of 26 GeV to a maximum of 72 GeV in steps of 2 GeV each, and the measured rate in data is compared to the predictions from the W+jets, Z+jets and multi-jet processes obtained from MC simulations. The multi-jet predictions are further tested in a fake-enriched region with inverted isolation requirement using the prescaled HLT_e26_lhvloose_nod0_L1EM20VH chain, and the measured bin-by-bin disagreement between data and MC is applied as a systematic uncertainty on the final prediction of the rate of multi-jet events passing the HLT_e26_lhtight_nod0_ivarloose selection.

The results of the analysis are summarised by the plot in Figure B.1, that has been made public by the ATLAS Collaboration [231]. The measured rate in data is in agreement with the MC predictions within uncertainties for all the E_{T} thresholds, and is dominated by the physics processes of interest (W+jets and Z+jets) with a significant but not major contribution from jets mis-identified as electrons.

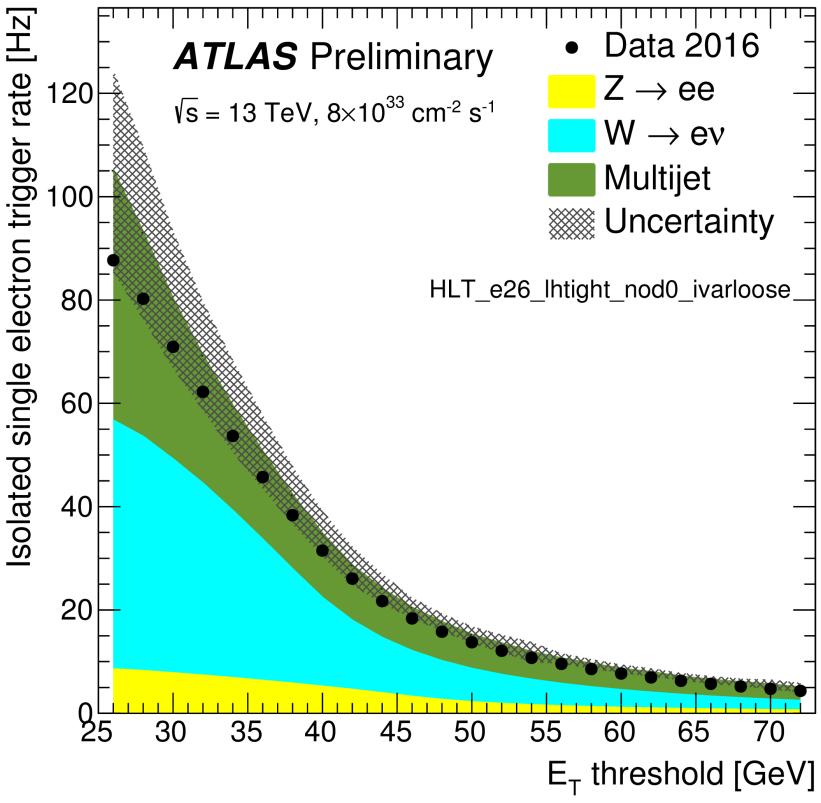
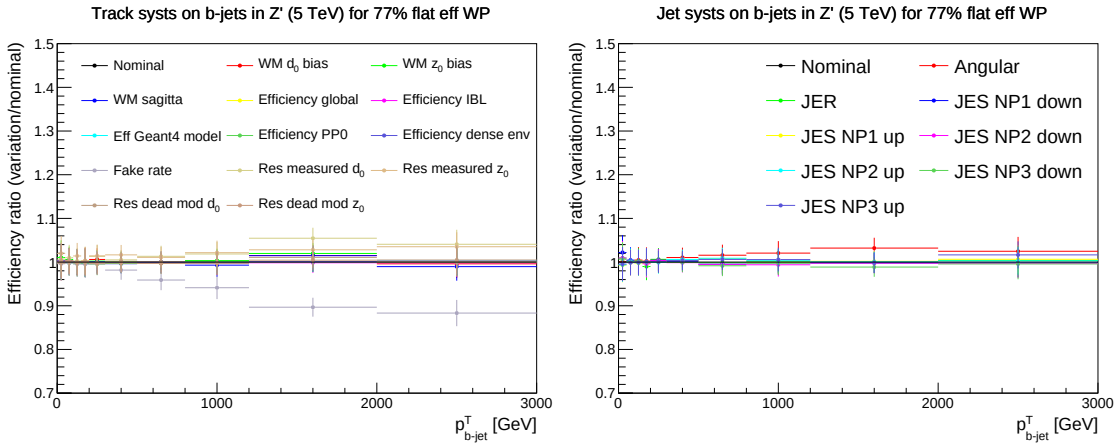


Figure B.1: Rate (in Hz) of the isolated single electron trigger as a function of the E_{T} threshold at the HLT in the $[26, 72]$ GeV range [231].

C | Additional plots of b-tagging efficiency

The present appendix includes a set of plots that complete the b-tagging performance studies described in detail in Chapter 4.
Figures C.1-C.2 show the uncertainty on b, c and light jets efficiency as a function of jet p_{T} in Z' and $t\bar{t}$ samples for the 77% fixed efficiency WP, to be compared with Figures 4.4-4.5 where the fixed cut WP is used. Figure C.3 is instead showing the total impact of jet uncertainties on the tagging efficiency of b, c and light jet efficiency in $t\bar{t}$, and of b-jets in Z', for the 77% fixed cut WP.



(a) (b)

Figure C.1: Ratio of modified and nominal b-tagging efficiencies vs b-jet p_{T} in a 5 TeV Z' sample for the 77% fixed efficiency WP, showing track (a) and jet (b) systematics defined in Section 4.1.3.

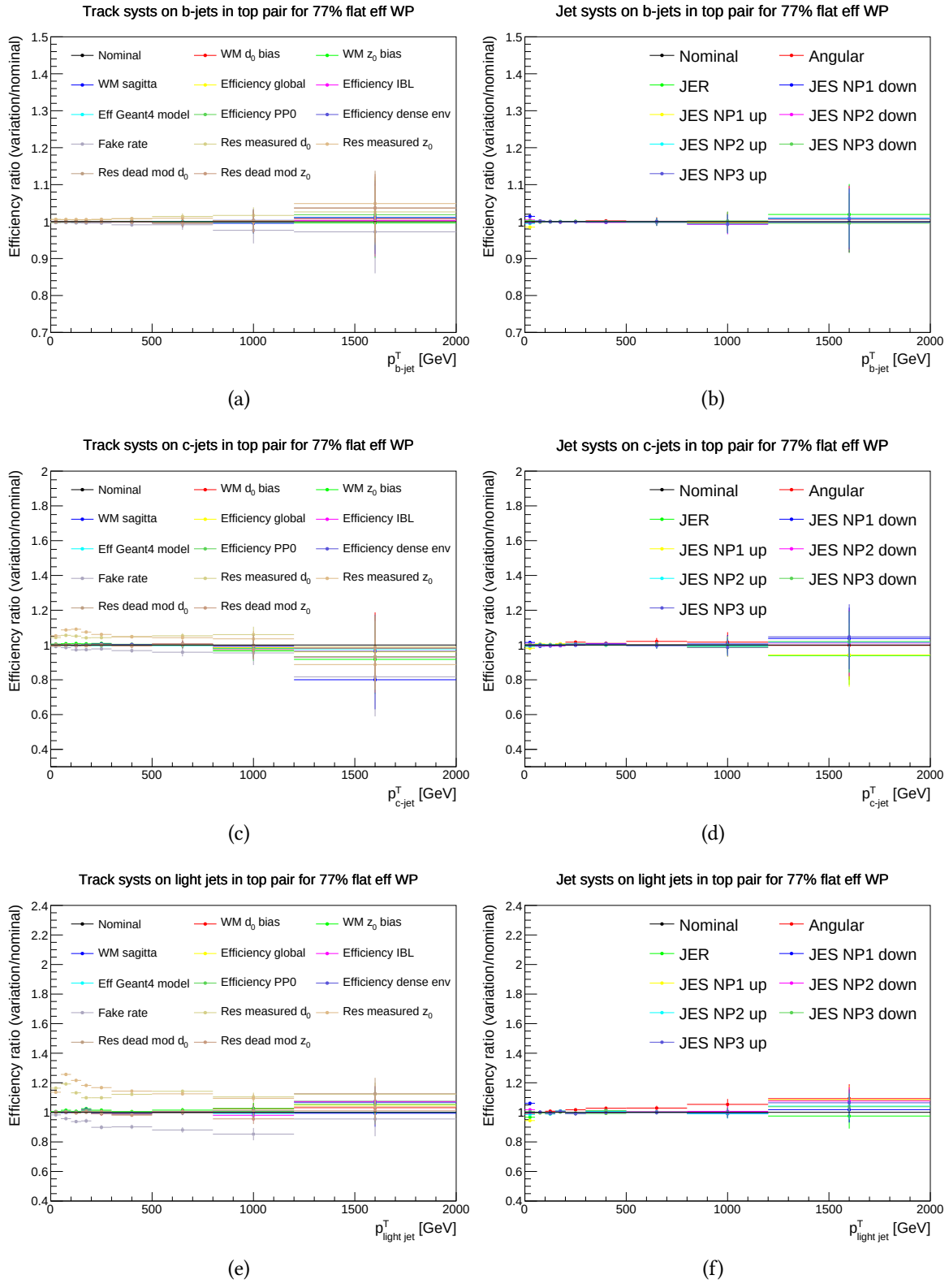
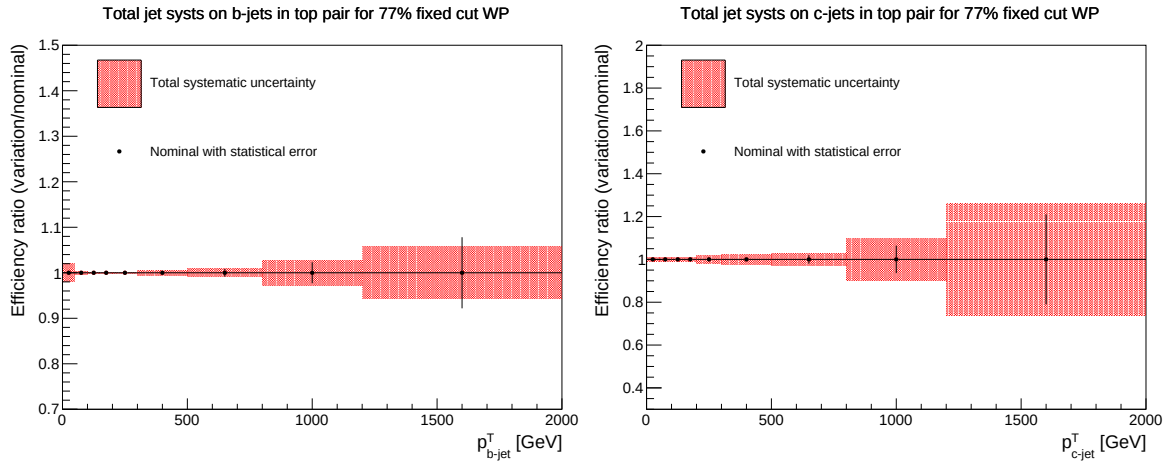
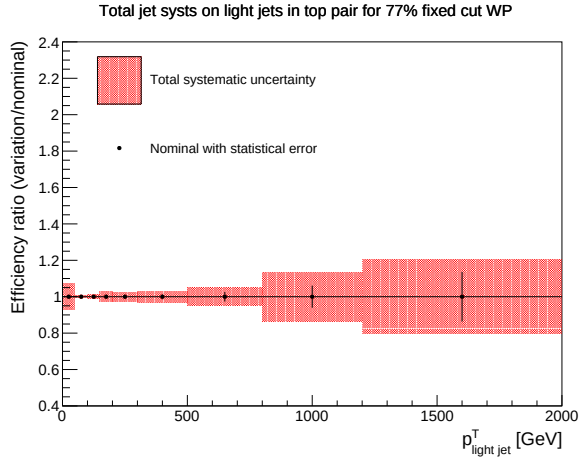


Figure C.2: Ratio of modified and nominal b-tagging efficiencies vs jet p_{T} in a $t\bar{t}$ sample for the 77% fixed efficiency WP, showing the impact of the individual track (left) and jet (right) systematics defined in Section 4.1.3 on b (top), c (middle) and light (bottom) jets.

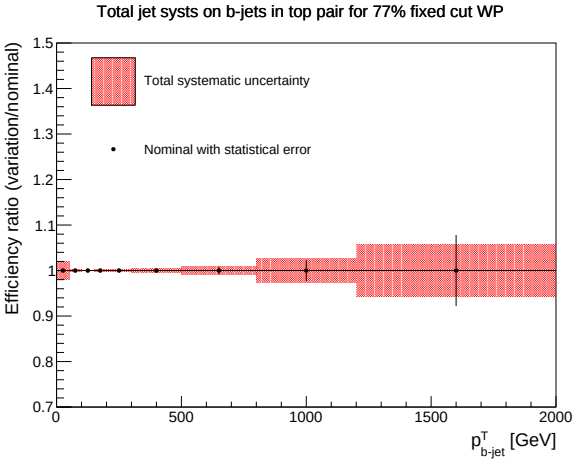


(a)

(b)



(c)



(d)

Figure C.3: Relative uncertainty caused by jet systematics defined in Section 4.1.3 on the tagging efficiency vs jet p_{T} for b (a), c (b) and light jets (c) in a $t\bar{t}$ sample, and for b-jets in a Z' sample (d), for the 77% fixed cut WP.

Acronyms

ALICE	A Large Ion Collider Experiment
ATLAS	A Toroidal LHC ApparatuS
BR	Branching Ratio
BSM	Beyond the Standard Model
CERN	European Organization for Nuclear Research
CKKW	Catani-Krauss-Kuhn-Webber
CKM	Cabibbo-Kobayashi-Maskawa
CL	Confidence Level
CMS	Compact Muon Solenoid
CR	Control Region
CSC	Cathode Strip Chamber
CTP	Central Trigger Processor
ECAL	Electromagnetic Calorimeter
EF	Event Filter
EM	Electromagnetic
FCal	Forward Calorimeter
FEB	Front-End Board
FSR	Final State Radiation
FTK	Fast Tracker

GRL Good Run List

GSF Gaussian Sum Filter

HCAL Hadronic Calorimeter

HEC Hadronic End-Caps

HLT High Level Trigger

IBL Insertable B-Layer

ID Inner Detector

ISR Initial State Radiation

JER Jet Energy Resolution

JES Jet Energy Scale

JVT Jet Vertex Tagger

L1 Level 1

L1Calo Level 1 Calorimeter

L1Muon Level 1 Muon

L1Topo Level 1 Topological

L2 Level 2

LAr Liquid Argon

LEP Large Electron-Positron Collider

LHC Large Hadron Collider

LHCb Large Hadron Collider beauty

LHCf Large Hadron Collider forward

LO Leading Order

LSP Lightest Supersymmetric Particle

MC Monte Carlo

MDT Muon Drift Tube

ME Matrix Element

MIP Minimum Ionising Particle

MLM Michelangelo L. Mangano

MoEDAL Monopole and Exotics Detector At the LHC

MPI Multiple Parton Interaction

MS Muon Spectrometer

MSSM Minimal Supersymmetric Standard Model

NLL Next-to-Leading Logarithm

NLO Next-to-Leading Order

NLSP Next to LSP

NNLL Next-to-Next-to-Leading Logarithm

NNLO Next-to-Next-to-Leading Order

OR Overlap Removal

PDF Parton Distribution Function

PLR Profile Likelihood Ratio

PMNS Pontecorvo-Maki-Nakagawa-Sakata

pMSSM phenomenological MSSM

PS Parton Shower

PS Proton Synchrotron

PSB Proton Synchrotron Booster

PV Primary Vertex

QCD Quantum Chromo Dynamics

QED Quantum Electrodynamics

QFT Quantum Field Theory

ROD Readout Driver

RoI Region of Interest

RoIB Region of Interest Builder

ROS Read-Out System

RPC Resistive Plate Chamber

RPC R-Parity Conserving

RPV R-Parity Violating

SCT Semi-Conductor Tracker

SF Scale Factor

SFO Sub-Farm Output

SM Standard Model

SPS Super Proton Synchrotron

SR Signal Region

SUSY Supersymmetry

TDAQ Trigger and Data Acquisition

TGC Thin Gap Chamber

TileCal Tile Calorimeter

TOTEM TOTal cross section, Elastic scattering and diffraction dissociation Measurement at the LHC

TRT Transition Radiation Tracker

VEV Vacuum Expectation Value

VR Validation Region

WIMP Weakly Interacting Massive Particle

WP Working Point

Bibliography

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

[2] K. Brading and E. Castellani, *Symmetries in physics: Philosophical reflections*, arXiv:quant-ph/0301097 (2003), arXiv:quant-ph/0301097 [quant-ph].

[3] A. Einstein, *Zur Elektrodynamik bewegter Körper. (German) [On the electrodynamics of moving bodies]*, Annalen der Physik **322** no. 10, (1905) 891–921.

[4] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264–1266.

[5] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321–323.

[6] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508–509.

[7] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12** (1964) 132–133.

[8] The 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. **B716** (2012) 1–29, arXiv:1207.7214 [hep-ex].

[9] The CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys.Lett. **B716** (2012) 30–61, arXiv:1207.7235 [hep-ex].

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

[11] The CMS Collaboration, *The CMS Experiment at the CERN LHC*, JINST **3** (2008) S08004.

[12] J. Wess and B. Zumino, *Supergauge Transformations in Four-Dimensions*, Nucl. Phys. **B70** (1974) 39–50.

[13] A. Salam and J. A. Strathdee, *Supersymmetry and Nonabelian Gauges*, Phys. Lett. **51B** (1974) 353–355.

[14] L. Susskind, *Dynamics of Spontaneous Symmetry Breaking in the Weinberg-Salam Theory*, Phys. Rev. **D20** (1979) 2619–2625.

[15] G. 't Hooft, *Naturalness, chiral symmetry, and spontaneous chiral symmetry breaking*, NATO Sci. Ser. B **59** (1980) 135–157.

[16] M. J. G. Veltman, *The Infrared - Ultraviolet Connection*, Acta Phys. Polon. **B12** (1981) 437.

[17] M. Papucci, J. T. Ruderman, and A. Weiler, *Natural SUSY Endures*, JHEP **09** (2012) 035, arXiv:1110.6926 [hep-ph].

[18] The ATLAS Collaboration, *Search for supersymmetry in events with b-tagged jets and missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **11** (2017) 195, arXiv:1708.09266 [hep-ex].

[19] The ATLAS Collaboration, *Search for $B - L$ R-parity-violating top squarks in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS experiment*, Phys. Rev. **D97** no. 3, (2018) 032003, arXiv:1710.05544 [hep-ex].

[20] M. E. Peskin and D. V. Schroeder, *An introduction to quantum field theory; 1995 ed.* Westview, Boulder, CO, 1995. https://cds.cern.ch/record/257493.

[21] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579–588.

[22] A. Salam and J. C. Ward, *Electromagnetic and weak interactions*, Phys. Lett. **13** (1964) 168–171.

[23] H. Fritzsch, M. Gell-Mann, and H. Leutwyler, *Advantages of the Color Octet Gluon Picture*, Phys. Lett. **47B** (1973) 365–368.

[24] A. Deur, S. J. Brodsky, and G. F. de Teramond, *The QCD Running Coupling*, Prog. Part. Nucl. Phys. **90** (2016) 1–74, arXiv:1604.08082 [hep-ph].

[25] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Nonabelian Gauge Theories*, Phys. Rev. Lett. **30** (1973) 1343–1346.

[26] The ATLAS Collaboration, *Measurement of the Higgs boson mass from the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels with the ATLAS detector using 25 fb^{-1} of pp collision data*, Phys. Rev. **D90** no. 5, (2014) 052004, arXiv:1406.3827 [hep-ex].

[27] The CMS Collaboration, *Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV*, Eur. Phys. J. **C75** no. 5, (2015) 212, arXiv:1412.8662 [hep-ex].

[28] The ATLAS and CMS Collaborations, *Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV*, JHEP **08** (2016) 045, arXiv:1606.02266 [hep-ex].

[29] V. C. Rubin and W. K. Ford, Jr., *Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions*, Astrophys. J. **159** (1970) 379–403.

[30] G. Bertone, D. Hooper, and J. Silk, *Particle dark matter: Evidence, candidates and constraints*, Phys. Rept. **405** (2005) 279–390, arXiv:hep-ph/0404175 [hep-ph].

[31] D. L. et al., *Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Power Spectra and WMAP-derived Parameters*, The Astrophysical Journal Supplement Series **192** no. 2, (2011) 16. http://stacks.iop.org/0067-0049/192/i=2/a=16.

[32] Planck Collaboration, P. A. R. e. a. Ade, *Planck 2013 results. XVI. Cosmological parameters*, Astron. Astrophys. **571** (2014) A16, arXiv:1303.5076 [astro-ph.CO].

[33] L. Roszkowski, E. M. Sessolo, and S. Trojanowski, *WIMP dark matter candidates and searches - current issues and future prospects*, arXiv:1707.06277 [hep-ph].

[34] C. Giunti, C. W. Kim, and U. W. Lee, *Running coupling constants and grand unification models*, Mod. Phys. Lett. **A6** (1991) 1745–1755.

[35] F. W. David J. Gross, H. David Politzer, *The 2004 Nobel Prize in Physics - Popular Information*, Nobel Media AB 2014 (2004) .

[36] W. M. Alberico and S. M. Bilenky, *Neutrino oscillations, masses and mixing*, Phys. Part. Nucl. **35** (2004) 297–323, arXiv:hep-ph/0306239 [hep-ph]. [Fiz. Elem. Chast. Atom. Yadra 35, 545 (2004)].

[37] Z. Maki, M. Nakagawa, and S. Sakata, *Remarks on the unified model of elementary particles*, Prog. Theor. Phys. **28** (1962) 870–880.

[38] S. M. Bilenky, J. Hosek, and S. T. Petcov, *On Oscillations of Neutrinos with Dirac and Majorana Masses*, Phys. Lett. **94B** (1980) 495–498.

[39] "Nu-FIT Website." http://www.nu-fit.org/. Retrieved 11$^{\text{th}}$ March, 2018.

[40] I. Esteban, M. C. Gonzalez-Garcia, M. Maltoni, I. Martinez-Soler, and T. Schwetz, *Updated fit to three neutrino mixing: exploring the accelerator-reactor complementarity*, JHEP **01** (2017) 087, arXiv:1611.01514 [hep-ph].

[41] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652–657.

[42] M. B. Gavela, P. Hernandez, J. Orloff, and O. Pene, *Standard model CP violation and baryon asymmetry*, Mod. Phys. Lett. **A9** (1994) 795–810, arXiv:hep-ph/9312215 [hep-ph].

[43] S. R. Coleman and J. Mandula, *All Possible Symmetries of the S Matrix*, Phys. Rev. **159** (1967) 1251–1256.

[44] H. Weyl, *Electron and Gravitation. 1. (In German)*, Z. Phys. **56** (1929) 330–352. [Surveys High Energ. Phys.5,261(1986)].

[45] G. R. Farrar and P. Fayet, *Phenomenology of the Production, Decay, and Detection of New Hadronic States Associated with Supersymmetry*, Phys. Lett. **76B** (1978) 575–579.

[46] M. Drees, R. Godbole, and P. Roy, *Theory and phenomenology of sparticles: An account of four-dimensional N=1 supersymmetry in high energy physics*. Hackensack, USA: World Scientific (2004) 555 p, 2004.

[47] K. Hidaka and A. Bartl, *Impact of bosonic decays on the search for the lighter stop and sbottom squarks*, Phys. Lett. **B501** (2001) 78–85, arXiv:hep-ph/0012021 [hep-ph].

[48] A. Schuessler and D. Zeppenfeld, *Unitarity constraints on MSSM trilinear couplings*, ArXiv e-prints (2007), arXiv:0710.5175 [hep-ph].

[49] S. P. Martin, *A Supersymmetry primer*, Adv.Ser.Direct.High Energy Phys. **21** (2010) 1–153, arXiv:hep-ph/9709356 [hep-ph].

[50] R. Barbieri and G. F. Giudice, *Upper Bounds on Supersymmetric Particle Masses*, Nucl. Phys. **B306** (1988) 63–76.

[51] Z. Marshall, B. A. Ovrut, A. Purves, and S. Spinner, *Spontaneous R-Parity Breaking, Stop LSP Decays and the Neutrino Mass Hierarchy*, Phys. Lett. **B732** (2014) 325–329, arXiv:1401.7989 [hep-ph].

[52] M. C. Gonzalez-Garcia, M. Maltoni, J. Salvado, and T. Schwetz, *Global fit to three neutrino mixing: critical look at present precision*, JHEP **12** (2012) 123, arXiv:1209.3023 [hep-ph].

[53] J. M. Campbell, J. W. Huston, and W. J. Stirling, *Hard Interactions of Quarks and Gluons: A Primer for LHC Physics*, Rept. Prog. Phys. **70** (2007) 89, arXiv:hep-ph/0611148 [hep-ph].

[54] R. Placakyte, *Parton Distribution Functions*, in *Proceedings, 31st International Conference on Physics in collisions (PIC 2011): Vancouver, Canada, August 28-September 1, 2011.* 2011. arXiv:1111.5452 [hep-ph].

[55] B. Isildak, *Measurement of the differential dijet production cross section in proton-proton collisions at $\sqrt{s} = 7$ TeV*. PhD thesis, Bogazici U., 2011. arXiv:1308.6064 [hep-ex].

[56] S. Myers and E. Picasso , *The design, construction and commissioning of the CERN large Electron-Positron collider*, Contemporary Physics **31** no. 6, (1990) 387–403.

[57] *Design Report Tevatron 1 project*, Tech. Rep. FERMILAB-DESIGN-1984-01, 1984. https://cds.cern.ch/record/1478620.

[58] The ALEPH, DELPHI, L3, OPAL, SLD Collaborations, the LEP Electroweak Working Group, the SLD Electroweak Group, the SLD Heavy Flavour Group, *Precision electroweak measurements on the Z resonance*, Phys. Rept. **427** (2006) 257–454, arXiv:hep-ex/0509008 [hep-ex].

[59] The CDF Collaboration, *Observation of top quark production in $\bar{p}p$ collisions*, Phys. Rev. Lett. **74** (1995) 2626–2631, arXiv:hep-ex/9503002 [hep-ex].

[60] The D0 Collaboration, *Observation of the top quark*, Phys. Rev. Lett. **74** (1995) 2632–2637, arXiv:hep-ex/9503003 [hep-ex].

[61] C. Llewellyn Smith, *Genesis of the Large Hadron Collider. Genesis of the Large Hadron Collider*, Philos. Trans. R. Soc. Lond. A **373** no. 2032, (2014) 20140037. http://cds.cern.ch/record/2288313.

[62] The ALEPH, DELPHI, L3, OPAL Collaborations, the LEP Working Group for Higgs Boson Searches, *Search for the standard model Higgs boson at LEP*, Phys. Lett. **B565** (2003) 61–75, arXiv:hep-ex/0306033 [hep-ex].

[63] The CDF Collaboration, *Combined search for the standard model Higgs boson decaying to a bb pair using the full CDF data set*, Phys. Rev. Lett. **109** (2012) 111802, arXiv:1207.1707 [hep-ex].

[64] The D0 Collaboration, *Combined Search for the Standard Model Higgs Boson Decaying to $b\bar{b}$ Using the D0 Run II Data Set*, Phys. Rev. Lett. **109** (2012) 121802, arXiv:1207.6631 [hep-ex].

[65] M. Benedikt, P. Collier, V. Mertens, J. Poole, and K. Schindl, *LHC Design Report v3.* CERN Yellow Reports: Monographs. CERN, Geneva, 2004. https://cds.cern.ch/record/823808.

[66] The UA1 Collaboration, *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540$ GeV*, Phys. Lett. **122B** (1983) 103–116. [,611(1983)].

[67] The UA2 Collaboration, *Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider*, Phys. Lett. **122B** (1983) 476–485.

[68] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, *LHC Design Report v1.* CERN Yellow Reports: Monographs. CERN, Geneva, 2004. https://cds.cern.ch/record/782076.

[69] The ATLAS Collaboration, "ATLAS Luminosity Public Results Run 2." https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2. Retrieved 5th March, 2018.

[70] The LHCb Collaboration, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005.

[71] The ALICE Collaboration, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002.

[72] The TOTEM Collaboration, *The TOTEM experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08007.

[73] The LHCf Collaboration, *The LHCf detector at the CERN Large Hadron Collider*, JINST **3** (2008) S08006.

[74] The MoEDAL Collaboration, *Technical Design Report of the MoEDAL Experiment*, Tech. Rep. MoEDAL-TDR-001, 2009. https://cds.cern.ch/record/1181486.

[75] The ATLAS Collaboration, *ATLAS: Technical proposal for a general-purpose p-p experiment at the Large Hadron Collider at CERN*, CERN-LHCC-94-43 (1994). https://cds.cern.ch/record/290968.

[76] The ATLAS Collaboration, *ATLAS magnet system: Technical design report*, CERN-LHCC-97-18 (1997). http://cds.cern.ch/record/338080.

[77] The CMS Collaboration, *CMS, the magnet project: Technical design report*, CERN-LHCC-97-10 (1997). http://cds.cern.ch/record/331056.

[78] The ATLAS Collaboration, *The ATLAS Inner Detector commissioning and calibration*, Eur. Phys. J. **C70** (2010) 787–821, arXiv:1004.5293 [physics.ins-det].

[79] The ATLAS Collaboration, N. Wermes and G. Hallewel, *ATLAS pixel detector: Technical Design Report.* Technical Design Report ATLAS. CERN, Geneva, 1998. https://cds.cern.ch/record/381263.

[80] The ATLAS Collaboration, M. Capeans, G. Darbo, K. Einsweiller, M. Elsing, T. Flick, M. Garcia-Sciveres, C. Gemme, H. Pernegger, O. Rohne, and R. Vuillermet, *ATLAS Insertable B-Layer Technical Design Report*, Tech. Rep. CERN-LHCC-2010-013. ATLAS-TDR-19, Sep, 2010. https://cds.cern.ch/record/1291633.

[81] The ATLAS Collaboration, M. Backhaus, *The upgraded Pixel Detector of the ATLAS Experiment for Run2 at the Large Hadron Collider*, Tech. Rep. ATL-INDET-PROC-2015-015, CERN, Geneva, Dec, 2015. http://cds.cern.ch/record/2110260.

[82] A. Abdesselam et al., *The barrel modules of the ATLAS semiconductor tracker*, Nucl. Instrum. Meth. **A568** (2006) 642–671.

[83] The ATLAS Collaboration, *The ATLAS semiconductor tracker end-cap module*, Nucl. Instrum. Meth. **A575** (2007) 353–389.

[84] The ATLAS TRT Collaboration, E. Abat et al., *The ATLAS TRT barrel detector*, JINST **3** (2008) P02014.

[85] E. Abat et al., *The ATLAS TRT end-cap detectors*, JINST **3** (2008) P10003.

[86] B. Dolgoshein, *Transition radiation detectors*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **326** no. 3, (1993) 434 – 469. http://www.sciencedirect.com/science/article/pii/016890029390846A.

[87] C. W. Fabjan and F. Gianotti, *Calorimetry for Particle Physics*, Rev. Mod. Phys. **75** no. CERN-EP-2003-075, (2003) 1243–1286. 96 p. http://cds.cern.ch/record/692252.

[88] The ATLAS Collaboration, *ATLAS liquid argon calorimeter: Technical design report*, CERN-LHCC-96-41 (1996). http://cds.cern.ch/record/331061.

[89] The ATLAS Collaboration, *Readiness of the ATLAS Liquid Argon Calorimeter for LHC Collisions*, Eur. Phys. J. **C70** (2010) 723–753, arXiv:0912.2642 [physics.ins-det].

[90] The ATLAS Collaboration, *ATLAS tile calorimeter: Technical Design Report.* Technical Design Report ATLAS. CERN, Geneva, 1996. https://cds.cern.ch/record/331062.

[91] The ATLAS Collaboration, *Readiness of the ATLAS Tile Calorimeter for LHC collisions*, Eur. Phys. J. **C70** (2010) 1193–1236, arXiv:1007.5423 [physics.ins-det].

[92] A. Artamonov et al., *The ATLAS forward calorimeters*, JINST **3** (2008) P02010.

[93] The ATLAS Collaboration, *ATLAS muon spectrometer: Technical Design Report.* Technical Design Report ATLAS. CERN, Geneva, 1997. https://cds.cern.ch/record/331068.

[94] The ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2015*, Eur. Phys. J. **C77** no. 5, (2017) 317, arXiv:1611.09661 [hep-ex].

[95] B. Abbott, R. Blair, G. Crone, B. Green, J. Love, J. Proudfoot, O. Rifki, W. P. Vazquez, W. Vandelli, and J. Zhang, *The evolution of the region of interest builder for the ATLAS experiment at CERN*, JINST **11** no. 02, (2016) C02080.

[96] The ATLAS Collaboration, *ATLAS level-1 trigger: Technical Design Report.* Technical Design Report ATLAS. CERN, Geneva, 1998. https://cds.cern.ch/record/381429.

[97] The ATLAS Collaboration, *Technical Design Report for the Phase-I Upgrade of the ATLAS TDAQ System*, CERN-LHCC-2013-018 (2013). https://cds.cern.ch/record/1602235.

[98] The ATLAS Collaboration, P. Jenni, M. Nessi, M. Nordberg, and K. Smith, *ATLAS high-level trigger, data-acquisition and controls: Technical Design Report.* Technical Design Report ATLAS. CERN, Geneva, 2003. https://cds.cern.ch/record/616089.

[99] M. zur Nedden, *The LHC Run 2 ATLAS trigger system: design, performance and plans*, Journal of Instrumentation **12** no. 03, (2017) C03024. http://stacks.iop.org/1748-0221/12/i=03/a=C03024.

[100] M. Shochet, L. Tompkins, V. Cavaliere, P. Giannetti, A. Annovi, and G. Volpi, *ATLAS Fast TracKer (FTK) Technical Design Report*, Tech. Rep. CERN-LHCC-2013-007. ATLAS-TDR-021, Jun, 2013. https://cds.cern.ch/record/1552953.

[101] M. Dobbs and J. B. Hansen, *The HepMC C++ Monte Carlo event record for High Energy Physics*, Comput. Phys. Commun. **134** (2001) 41–46.

[102] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, Nucl. Phys. **B126** (1977) 298–318.

[103] V. N. Gribov and L. N. Lipatov, *Deep inelastic e p scattering in perturbation theory*, Sov. J. Nucl. Phys. **15** (1972) 438–450. [Yad. Fiz.15,781(1972)].

[104] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.*, Sov. Phys. JETP **46** (1977) 641–653. [Zh. Eksp. Teor. Fiz.73,1216(1977)].

[105] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin, and C. P. Yuan, *New parton distributions for collider physics*, Phys. Rev. **D82** (2010) 074024, arXiv:1007.2241 [hep-ph].

[106] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. M. Nadolsky, and W. K. Tung, *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **07** (2002) 012, arXiv:hep-ph/0201195 [hep-ph].

[107] The NNPDF Collaboration, R. D. Ball et al., *Parton distributions for the LHC Run II*, JHEP **04** (2015) 040, arXiv:1410.8849 [hep-ph].

[108] R. D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. **B867** (2013) 244–289, arXiv:1207.1303 [hep-ph].

[109] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. **C63** (2009) 189–285, arXiv:0901.0002 [hep-ph].

[110] T. Kinoshita, *Mass singularities of Feynman amplitudes*, J. Math. Phys. **3** (1962) 650–677.

[111] Lee, T. D. and Nauenberg, M., *Degenerate Systems and Mass Singularities*, Phys. Rev. **133** (1964) B1549–B1562.

[112] J. A. Aguilar-Saavedra, D. Amidei, A. Juste, and M. Perez-Victoria, *Asymmetries in top quark pair production at hadron colliders*, Rev. Mod. Phys. **87** (2015) 421–455, arXiv:1406.1798 [hep-ph].

[113] V. V. Sudakov, *Vertex parts at very high-energies in quantum electrodynamics*, Sov. Phys. JETP **3** (1956) 65–71. [Zh. Eksp. Teor. Fiz.30,87(1956)].

[114] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber, *QCD matrix elements + parton showers*, JHEP **11** (2001) 063, arXiv:hep-ph/0109231 [hep-ph].

[115] M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani, *Matching matrix elements and shower evolution for top-quark production in hadronic collisions*, JHEP **01** (2007) 013, arXiv:hep-ph/0611129 [hep-ph].

[116] B. R. Webber, *A QCD Model for Jet Fragmentation Including Soft Gluon Interference*, Nucl. Phys. **B238** (1984) 492–528.

[117] G. Marchesini and B. R. Webber, *Monte Carlo Simulation of General Hard Processes with Coherent QCD Radiation*, Nucl. Phys. **B310** (1988) 461–526.

[118] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, *Parton Fragmentation and String Dynamics*, Phys. Rept. **97** (1983) 31–145.

[119] T. Sjostrand, *Jet Fragmentation of Nearby Partons*, Nucl. Phys. **B248** (1984) 469–502.

[120] P. Z. Skands, *Tuning Monte Carlo Generators: The Perugia Tunes*, Phys. Rev. **D82** (2010) 074018, arXiv:1005.3457 [hep-ph].

[121] The ATLAS Collaboration, *ATLAS Run 1 Pythia8 tunes*, Tech. Rep. ATL-PHYS-PUB-2014-021, CERN, Geneva, Nov, 2014. https://cds.cern.ch/record/1966419.

[122] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **05** (2006) 026, arXiv:hep-ph/0603175 [hep-ph].

[123] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852–867, arXiv:0710.3820 [hep-ph].

[124] G. Marchesini, B. R. Webber, G. Abbiendi, I. G. Knowles, M. H. Seymour, and L. Stanco, *HERWIG: A Monte Carlo event generator for simulating hadron emission reactions with interfering gluons. Version 5.1 - April 1991*, Comput. Phys. Commun. **67** (1992) 465–508.

[125] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, *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].

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

[127] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter, *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv:0811.4622 [hep-ph].

[128] S. Schumann and F. Krauss, *A Parton shower algorithm based on Catani-Seymour dipole factorisation*, JHEP **03** (2008) 038, arXiv:0709.1027 [hep-ph].

[129] The ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. **C70** (2010) 823–874, arXiv:1005.4568 [physics.ins-det].

[130] The GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A Simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250–303.

[131] The ATLAS Collaboration, *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim*, Tech. Rep. ATL-PHYS-PUB-2010-013, CERN, Geneva, Oct, 2010. https://cds.cern.ch/record/1300517.

[132] The ATLAS Collaboration, "ATLAS Trigger Public Results." https://twiki.cern.ch/twiki/bin/view/AtlasPublic/TriggerPublicResults. Retrieved 5th March, 2018.

[133] T. Cornelissen, M. Elsing, S. Fleischmann, W. Liebig, and E. Moyse, *Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)*, ATL-SOFT-PUB-2007-007, ATL-COM-SOFT-2007-002 (2007).

[134] The ATLAS Collaboration, *Vertex Reconstruction Performance of the ATLAS Detector at $\sqrt{s} = 13$ TeV*, Tech. Rep. ATL-PHYS-PUB-2015-026, CERN, Geneva, Jul, 2015. http://cds.cern.ch/record/2037717.

[135] The ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, Eur. Phys. J. **C74** no. 10, (2014) 3071, arXiv:1407.5063 [hep-ex].

[136] The ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*, Tech. Rep. ATLAS-CONF-2016-024, CERN, Geneva, Jun, 2016. https://cds.cern.ch/record/2157687.

[137] The ATLAS Collaboration, *Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run-1 data*, Eur. Phys. J. **C76** no. 12, (2016) 666, arXiv:1606.01813 [hep-ex].

[138] The ATLAS Collaboration, *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*, Tech. Rep. ATLAS-CONF-2012-047, CERN, Geneva, May, 2012. https://cds.cern.ch/record/1449796.

[139] The ATLAS Collaboration, *Photon identification in 2015 ATLAS data*, Tech. Rep. ATL-PHYS-PUB-2016-014, CERN, Geneva, Aug, 2016. https://cds.cern.ch/record/2203125.

[140] The ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using 2012 LHC proton-proton collision data*, Eur. Phys. J. **C77** no. 3, (2017) 195, arXiv:1612.01456 [hep-ex].

[141] The ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using data collected in 2015 at $\sqrt{s} = 13$ TeV*, Tech. Rep. ATL-PHYS-PUB-2016-015, CERN, Geneva, Aug, 2016. https://cds.cern.ch/record/2203514.

[142] The ATLAS Collaboration, "ATLAS Egamma Public Results." https://twiki.cern.ch/twiki/bin/view/AtlasPublic/ElectronGammaPublicCollisionResults. Retrieved 5^{th} March, 2018.

[143] The ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s}$ =13 TeV*, Eur. Phys. J. **C76** no. 5, (2016) 292, arXiv:1603.05598 [hep-ex].

[144] The ATLAS Collaboration, "ATLAS Muon Public Results." https://twiki.cern.ch/twiki/bin/view/AtlasPublic/MuonPerformancePublicPlots. Retrieved 5^{th} March, 2018.

[145] G. P. Salam, *Towards Jetography*, Eur. Phys. J. **C67** (2010) 637–686, arXiv:0906.1833 [hep-ph].

[146] W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, S. Menke, S. Rajagopalan, D. Rousseau, S. Snyder, and G. Unal, *Calorimeter Clustering Algorithms: Description and Performance*, Tech. Rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, CERN, Geneva, Apr, 2008. https://cds.cern.ch/record/1099735.

[147] The ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. **C77** (2017) 490, arXiv:1603.02934 [hep-ex].

[148] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti-k_{T} jet clustering algorithm*, JHEP **04** (2008) 063, arXiv:0802.1189 [hep-ph].

[149] The ATLAS Collaboration, *Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} = 13$ TeV*, Tech. Rep. ATL-PHYS-PUB-2015-015, CERN, Geneva, Jul, 2015. http://cds.cern.ch/record/2037613.

[150] The ATLAS Collaboration, *Pile-up subtraction and suppression for jets in ATLAS*, Tech. Rep. ATLAS-CONF-2013-083, CERN, Geneva, Aug, 2013. https://cds.cern.ch/record/1570994.

[151] The ATLAS Collaboration, *Jet global sequential corrections with the ATLAS detector in proton-proton collisions at sqrt(s) = 8 TeV*, Tech. Rep. ATLAS-CONF-2015-002, CERN, Geneva, Mar, 2015. https://cds.cern.ch/record/2001682.

[152] The ATLAS Collaboration, *Data-driven determination of the energy scale and resolution of jets reconstructed in the ATLAS calorimeters using dijet and multijet events at $\sqrt{s} = 8$ TeV*, Tech. Rep. ATLAS-CONF-2015-017, CERN, Geneva, Apr, 2015. https://cds.cern.ch/record/2008678.

[153] The ATLAS Collaboration, "ATLAS JetEtMiss Public Results." https://twiki.cern.ch/twiki/bin/view/AtlasPublic/JetEtmissPublicResults. Retrieved 5^{th} March, 2018.

[154] The ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2014-018, CERN, Geneva, May, 2014. https://cds.cern.ch/record/1700870.

[155] The ATLAS Collaboration, *Performance of b-Jet Identification in the ATLAS Experiment*, JINST **11** no. 04, (2016) P04008, arXiv:1512.01094 [hep-ex].

[156] The ATLAS Collaboration, *Performance and Calibration of the JetFitterCharm Algorithm for c-Jet Identification*, Tech. Rep. ATL-PHYS-PUB-2015-001, CERN, Geneva, Jan, 2015. https://cds.cern.ch/record/1980463.

[157] The ATLAS Collaboration, *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*, Tech. Rep. ATL-PHYS-PUB-2016-012, CERN, Geneva, Jun, 2016. http://cds.cern.ch/record/2160731.

[158] G. Piacquadio and C. Weiser, *A new inclusive secondary vertex algorithm for b-jet tagging in ATLAS*, J. Phys. Conf. Ser. **119** (2008) 032032.

[159] R. Fruhwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl. Instrum. Meth. **A262** (1987) 444–450.

[160] L. Breiman, J. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*. Chapman and Hall/CRC, 1984.

[161] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, arXiv:physics/0703039.

[162] The ATLAS Collaboration, *Calibration of b-tagging using dileptonic top pair events in a combinatorial likelihood approach with the ATLAS experiment*, Tech. Rep. ATLAS-CONF-2014-004, CERN, Geneva, Feb, 2014. https://cds.cern.ch/record/1664335.

[163] The ATLAS Collaboration, *Calibration of the b-tagging efficiency for c jets with the ATLAS detector using events with a W boson produced in association with a single c quark*, Tech. Rep. ATLAS-CONF-2013-109, CERN, Geneva, Dec, 2013. https://cds.cern.ch/record/1640162.

[164] The ATLAS Collaboration, *Calibration of the performance of b-tagging for c and light-flavour jets in the 2012 ATLAS data*, Tech. Rep. ATLAS-CONF-2014-046, CERN, Geneva, Jul, 2014. https://cds.cern.ch/record/1741020.

[165] The ATLAS Collaboration, *Performance of missing transverse momentum reconstruction for the ATLAS detector in the first proton-proton collisions at at $\sqrt{s}$= 13 TeV*, Tech. Rep. ATL-PHYS-PUB-2015-027, CERN, Geneva, Jul, 2015. http://cds.cern.ch/record/2037904.

[166] The ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s}$ = 13 TeV*, arXiv:1802.08168 [hep-ex].

[167] The ATLAS Collaboration, *Track Reconstruction Performance of the ATLAS Inner Detector at $\sqrt{s} = 13$ TeV*, Tech. Rep. ATL-PHYS-PUB-2015-018, CERN, Geneva, Jul, 2015. https://cds.cern.ch/record/2037683.

[168] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152–155.

[169] The ATLAS Collaboration, *Early Inner Detector Tracking Performance in the 2015 data at $\sqrt{s}$ = 13 TeV*, Tech. Rep. ATL-PHYS-PUB-2015-051, CERN, Geneva, Dec, 2015. https://cds.cern.ch/record/2110140.

[170] The ATLAS Collaboration, *Tracking Studies for b-tagging with 7 TeV Collision Data with the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2010-070, CERN, Geneva, Jul, 2010. https://cds.cern.ch/record/1281352.

[171] The ATLAS Collaboration, *Studies of the ATLAS Inner Detector material using $\sqrt{s}$ =13 TeV pp collision data*, Tech. Rep. ATL-PHYS-PUB-2015-050, CERN, Geneva, Nov, 2015. https://cds.cern.ch/record/2109010.

[172] The ATLAS Collaboration, *Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC Run 2*, Eur. Phys. J. **C77** no. 10, (2017) 673, arXiv:1704.07983 [hep-ex].

[173] The ATLAS Collaboration, *Alignment of the ATLAS Inner Detector with the initial LHC data at $\sqrt{s}$ = 13 TeV*, Tech. Rep. ATL-PHYS-PUB-2015-031, CERN, Geneva, Jul, 2015. https://cds.cern.ch/record/2038139.

[174] The ATLAS Collaboration, *Study of alignment-related systematic effects on the ATLAS Inner Detector tracking*, Tech. Rep. ATLAS-CONF-2012-141, CERN, Geneva, Oct, 2012. https://cds.cern.ch/record/1483518.

[175] The ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. **D96** no. 7, (2017) 072002, arXiv:1703.09665 [hep-ex].

[176] The ATLAS Collaboration, *Jet energy resolution in proton-proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector*, Eur. Phys. J. **C73** no. 3, (2013) 2306, arXiv:1210.6210 [hep-ex].

[177] The ATLAS Collaboration, *Identification of high transverse momentum top quarks in pp collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector*, JHEP **06** (2016) 093, arXiv:1603.03127 [hep-ex].

[178] A. Monteux, *New signatures and limits on R-parity violation from resonant squark production*, JHEP **03** (2016) 216, arXiv:1601.03737 [hep-ph].

[179] C. Borschensky, M. Krämer, A. Kulesza, M. Mangano, S. Padhi, T. Plehn, and X. Portell, *Squark and gluino production cross sections in pp collisions at $\sqrt{s}$ = 13, 14, 33 and 100 TeV*, Eur. Phys. J. **C74** no. 12, (2014) 3174, arXiv:1407.5066 [hep-ph].

[180] The LHC New Physics Working Group Collaboration, D. Alves et al., *Simplified Models for LHC New Physics Searches*, J.Phys. **G39** (2012) 105005, arXiv:1105.2838 [hep-ph].

[181] The ATLAS Collaboration, "ATLAS Standard Model Public Results Run 2." https://twiki.cern.ch/twiki/bin/view/AtlasPublic/StandardModelPublicResults. Retrieved 15th March, 2018.

[182] The ATLAS Collaboration, *Summary of the ATLAS experiment's sensitivity to supersymmetry after LHC Run 1 interpreted in the phenomenological MSSM*, JHEP **10** (2015) 134, arXiv:1508.06608 [hep-ex].

[183] The ATLAS Collaboration, *ATLAS Run 1 searches for direct pair production of third-generation squarks at the Large Hadron Collider*, Eur. Phys. J. **C75** no. 10, (2015) 510, arXiv:1506.08616 [hep-ex]. [Erratum: Eur. Phys. J.C76,no.3,153(2016)].

[184] The ATLAS Collaboration, *Search for top-squark pair production in final states with one lepton, jets, and missing transverse momentum using 36 fb^{-1} of $\sqrt{s} = 13$ TeV pp collision data with the ATLAS detector*, arXiv:1711.11520 [hep-ex].

[185] A. H. Chamseddine, R. Arnowitt, and P. Nath, *Locally Supersymmetric Grand Unification*, Phys. Rev. Lett. **49** (1982) 970–974.

[186] R. Barbieri, S. Ferrara, and C. Savoy, *Gauge models with spontaneously broken local supersymmetry*, Physics Letters B **119** no. 4, (1982) 343 – 347.

[187] G. L. Kane, C. F. Kolda, L. Roszkowski, and J. D. Wells, *Study of constrained minimal supersymmetry*, Phys. Rev. **D49** (1994) 6173–6210, arXiv:hep-ph/9312272 [hep-ph].

[188] N. Arkani-Hamed, A. Delgado, and G. F. Giudice, *The Well-tempered neutralino*, Nucl. Phys. **B741** (2006) 108–130, arXiv:hep-ph/0601041 [hep-ph].

[189] R. Gröber, M. M. Mühlleitner, E. Popenda, and A. Wlotzka, *Light Stop Decays: Implications for LHC Searches*, Eur. Phys. J. **C75** (2015) 420, arXiv:1408.4662 [hep-ph].

[190] The ATLAS Collaboration, "ATLAS Supersymmetry searches." https://twiki.cern.ch/twiki/bin/view/AtlasPublic/SupersymmetryPublicResults. Retrieved 5^{th} March, 2018.

[191] The Particle Data Group Collaboration, C. Patrignani et al., *Review of Particle Physics*, Chin. Phys. **C40** no. 10, (2016) 100001.

[192] The ATLAS Collaboration, *Evidence for the $H \to b\bar{b}$ decay with the ATLAS detector*, JHEP **12** (2017) 024, arXiv:1708.03299 [hep-ex].

[193] The ATLAS Collaboration, *Improved luminosity determination in pp collisions at sqrt(s) = 7 TeV using the ATLAS detector at the LHC*, Eur. Phys. J. **C73** no. 8, (2013) 2518, arXiv:1302.4393 [hep-ex].

[194] The ATLAS Collaboration, *Measurement of the cross section for the production of a W boson in association with b^- jets in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys. Lett. **B707** (2012) 418–437, arXiv:1109.1470 [hep-ex].

[195] The ATLAS Collaboration, *Multi-Boson Simulation for 13 TeV ATLAS Analyses*, Tech. Rep. ATL-PHYS-PUB-2016-002, CERN, Geneva, Jan, 2016. https://cds.cern.ch/record/2119986.

[196] J. T. Linnemann, *Measures of Significance in HEP and Astrophysics*, Talk from PhyStat2003, Stanford, Ca, USA (2003) 5 p. https://cds.cern.ch/record/691153.

[197] R. D Cousins and J. Tucker, *Evaluation of two methods for incorporating a systematic uncertainty into a test of the background-only hypothesis for a Poisson process*, arXiv:physics/0702156v1 (2007) .

[198] L. Moneta, K. Belasco, K. S. Cranmer, S. Kreiss, A. Lazzaro, D. Piparo, G. Schott, W. Verkerke, and M. Wolf, *The RooStats Project*, PoS **ACAT2010** (2010) 057, arXiv:1009.1003 [physics.data-an].

[199] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, Nucl. Instrum. Meth. **A389** (1997) 81–86.

[200] M. Baak, G. J. Besjes, D. Côte, A. Koutsman, J. Lorenz, and D. Short, *HistFitter software framework for statistical data analysis*, Eur. Phys. J. **C75** (2015) 153, arXiv:1410.1280 [hep-ex].

[201] G. Cowan, *Statistical Data Analysis*. Oxford science publications. Clarendon Press, 1998. https://books.google.co.uk/books?id=ff8ZyW0nlJAC.

[202] Cowan, Glen, *Statistics for Searches at the LHC*, Proceedings, 69th Scottish Universities Summer School in Physics, LHC Phenomenology (SUSSP69), St. Andrews, Scotland, August 19-September 1, 2012 (2013) 321–355, arXiv:1307.2487 [hep-ex].

[203] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. **C71** (2011) 1554, arXiv:1007.1727 [physics.data-an]. [Erratum: Eur. Phys. J.C73,2501(2013)].

[204] R. D. Cousins, J. T. Linnemann, and J. Tucker, *Evaluation of three methods for calculating statistical significance when incorporating a systematic uncertainty into a test of the background-only hypothesis for a Poisson process*, Nucl. Inst. and Methods in Physics Research Section A **595** no. 2, (2008) 480 – 501.

[205] A. L. Read, *Presentation of search results: The CL(s) technique*, J. Phys. **G28** (2002) 2693–2704.

[206] The ATLAS Collaboration, *Search for new phenomena in final states with an energetic jet and large missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector*, Phys. Rev. **D94** no. 3, (2016) 032005, arXiv:1604.07773 [hep-ex].

[207] The ATLAS Collaboration, *Search for a scalar partner of the top quark in the jets plus missing transverse momentum final state at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **12** (2017) 085, arXiv:1709.04183 [hep-ex].

[208] G. Lerner, *Search for scalar top and bottom in final states with one lepton and missing transverse momentum, using 8 TeV pp collisions recorded by the ATLAS detector*, Master's thesis, Università degli Studi di Milano, Italy, 2014.

[209] The ATLAS Collaboration, *Search for bottom squark pair production in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Eur. Phys. J. **C76** no. 10, (2016) 547, arXiv:1606.08772 [hep-ex].

[210] The ATLAS Collaboration, *Search for scalar bottom quark pair production with the ATLAS detector in pp Collisions at $\sqrt{s} = 7$ TeV*, Phys. Rev. Lett. **108** (2012) 181802, arXiv:1112.3832 [hep-ex].

[211] The ATLAS Collaboration, *Search for direct third-generation squark pair production in final states with missing transverse momentum and two b-jets in $\sqrt{s} = 8$ TeV pp collisions with the ATLAS detector*, JHEP **10** (2013) 189, arXiv:1308.2631 [hep-ex].

[212] The ATLAS Collaboration, *Search for new phenomena in final states with an energetic jet and large missing transverse momentum in pp collisions at $\sqrt{s}$ =8 TeV with the ATLAS detector*, Eur. Phys. J. **C75** no. 7, (2015) 299, arXiv:1502.01518 [hep-ex]. [Erratum: Eur. Phys. J.C75,no.9,408(2015)].

[213] The CMS Collaboration, *Searches for third-generation squark production in fully hadronic final states in proton-proton collisions at $\sqrt{s} = 8$ TeV*, JHEP **06** (2015) 116, arXiv:1503.08037 [hep-ex].

[214] The CMS Collaboration, *Search for new physics with the M_{T2} variable in all-jets final states produced in pp collisions at $\sqrt{s} = 13$ TeV*, JHEP **10** (2016) 006, arXiv:1603.04053 [hep-ex].

[215] The CMS Collaboration, *Search for the pair production of third-generation squarks with two-body decays to a bottom or charm quark and a neutralino in proton-proton collisions at $\sqrt{s} = 13$ TeV*, arXiv:1707.07274 [hep-ex].

[216] D. R. Tovey, *On measuring the masses of pair-produced semi-invisibly decaying particles at hadron colliders*, JHEP **04** (2008) 034, arXiv:0802.2879 [hep-ph].

[217] C. G. Lester and D. J. Summers, *Measuring masses of semiinvisibly decaying particles pair produced at hadron colliders*, Phys. Lett. **B463** (1999) 99–103, arXiv:hep-ph/9906349 [hep-ph].

[218] A. J. Barr, B. Gripaios, and C. G. Lester, *Transverse masses and kinematic constraints: from the boundary to the crease*, JHEP **11** (2009) 096, arXiv:0908.3779 [hep-ph].

[219] P. Konar, K. Kong, K. T. Matchev, and M. Park, *Dark Matter Particle Spectroscopy at the LHC: Generalizing M(T2) to Asymmetric Event Topologies*, JHEP **04** (2010) 086, arXiv:0911.4126 [hep-ph].

[220] The ATLAS Collaboration, *HepData Record: Search for supersymmetry in events with b-tagged jets and missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, 2017. http://dx.doi.org/10.17182/hepdata.79165.

[221] C. M. Macdonald, *Searches for supersymmetry in final states containing b-tagged jets with the ATLAS detector.* PhD thesis, University of Sheffield, September, 2017. http://etheses.whiterose.ac.uk/17990.

[222] J. K. Anders, *Searches for direct pair production of third generation squarks, and dark matter, in final states containing b-jets and $E_{\mathrm{T}}^{\mathrm{miss}}$ using the ATLAS detector at the LHC.* PhD thesis, University of Liverpool, May, 2017. https://cds.cern.ch/record/2291836. Presented 07 Jun 2017.

[223] S. Ask, M. A. Parker, T. Sandoval, M. E. Shea, and W. J. Stirling, *Using gamma+jets Production to Calibrate the Standard Model Z(nunu)+jets Background to New Physics Processes at the LHC*, JHEP **10** (2011) 058, arXiv:1107.2803 [hep-ph].

[224] The ATLAS Collaboration, *Search for squarks and gluinos with the ATLAS detector in final states with jets and missing transverse momentum using 4.7 fb^{-1} of $\sqrt{s} = 7$ TeV proton-proton collision data*, Phys. Rev. **D87** no. 1, (2013) 012008, arXiv:1208.0949 [hep-ex].

[225] The ATLAS Collaboration, *Studies on top-quark Monte Carlo modelling with Sherpa and MG5_aMC@NLO*, Tech. Rep. ATL-PHYS-PUB-2017-007, CERN, Geneva, May, 2017. http://cds.cern.ch/record/2261938.

[226] The ATLAS Collaboration, *A search for $B - L$ R-Parity violating scalar top decays in $\sqrt{s} = 8$ TeV pp collisions with the ATLAS experiment*, ATLAS-CONF-2015-015 (2015).

[227] The ATLAS Collaboration, *Search for scalar leptoquarks in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS experiment*, New J. Phys. **18** no. 9, (2016) 093016, arXiv:1605.06035 [hep-ex].

[228] The CMS Collaboration, *Search for third-generation scalar leptoquarks and heavy right-handed neutrinos in final states with two tau leptons and two jets in proton-proton collisions at* $\sqrt{s} = 13$ *TeV*, JHEP **07** (2017) 121, arXiv:1703.03995 [hep-ex].

[229] The ATLAS Collaboration, *HepData record: A search for* $B - L$ R*-parity-violating top squarks in* $\sqrt{s} = 13$ *TeV* pp *collisions with the ATLAS experiment*, 2017. http://dx.doi.org/10.17182/hepdata.78376.v2.

[230] The ATLAS Collaboration, *Search for Supersymmetry at the high luminosity LHC with the ATLAS experiment*, Tech. Rep. ATL-PHYS-PUB-2014-010, CERN, Geneva, Jul, 2014. https://cds.cern.ch/record/1735031.

[231] The ATLAS Collaboration, "ATLAS Egamma Trigger Plots." https://twiki.cern.ch/twiki/bin/view/AtlasPublic/EgammaTriggerPublicResults. Retrieved 5^{th} March, 2018.