Search for new physics in events with same sign leptons and missing energy with ATLAS at LHC

O.-A. Ducu

CPPM-T-2015-06
CERN-THESIS-2015-163

University Aix-Marseille
Faculty of Science of Luminy

University of Bucharest
Faculty of Physics

PhD THESIS

Domain : *Physics and Exact Sciences*

Search for new physics in events with same sign leptons and missing energy with ATLAS at LHC

presented by

Otilia-Anamaria DUCU

to obtain the degree of Doctor in Physics
of University of Aix-Marseille and University of Bucharest

sustained on 15 October 2015, with the jury formed by :

Dr Calin Alexa	Thesis supervisor	IFIN-HH (Romania)
Dr Philippe Bloch	Examiner	CERN (Switzerland)
Dr Irinel Caprini	Examiner	IFIN-HH (Romania)
Dr Andreas Hoecker	Examiner	CERN (Switzerland)
Dr Eric Kajfasz	Examiner	CPPM (France)
Dr Pascal Pralavorio	Thesis supervisor	CPPM (France)

Abstract

English version

The Standard Model of particle physics describes the phenomena in the infinitesimal with great success. However, it suffers from several inadequacies : it can not explain the neutrino masses, it has no candidate for the dark matter and it has no solution for the gauge hierarchy problem, which all require new physics and beyond the Standard Model theories. Such a theory is Supersymmetry (SUSY) and occupies a primer place in the LHC physics program. Here two proton beams are colliding at energies up to 14 TeV and gigantic detectors were built to reconstruct the collision events. For the presented studies only the data recorded with the ATLAS detector is analyzed. More details on the theoretical framework, the LHC collider and the ATLAS experiment are given in the first part of this thesis.

At hadron colliders, the lepton identification and reconstruction are crucial for precise SM cross sections and coupling measurements or for new physics searches. Therefore, the second part of my thesis is dedicated to performance studies : in-situ measurements of electron identification and reconstruction efficiencies. Both the methodology employed to obtain the results and the methods applied to estimate the background are detailed. The search for new physics is presented in the third part of the thesis. For these studies the channel with two same - sign leptons and missing transverse energy is considered. The results are obtained with $L = 20$ fb^{-1} of data recorded at $\sqrt{s} = 8$ TeV and with the very first data at $\sqrt{s} = 13$ TeV. At the end of Run-1 no significant excess in data over the Standard Model prediction is observed and stringent limits on the supersymmetric particle masses are set. With this analysis, the gluino mass is constrained to be above 1 TeV and the sbottom mass should be above 600 GeV when using simplified assumptions. These results provide new and very strong constraints on natural SUSY models.

With the two same - sign leptons and jets final state it is also possible to measure directly the top quark Yukawa coupling using the $t\bar{t} + H$ channel. This coupling is the largest in the Standard Model and it connects two of the heaviest particles predicted by the Standard Model, the top quark and the Higgs boson. At $\sqrt{s} = 8$ TeV the results are dominated by the systematic uncertainties and only upper limits can be obtained. In the fourth part of this thesis, a complete optimization is performed to reach the highest sensitivity for this signal in the two same - sign leptons channel at $\sqrt{s} = 13$ TeV with 100 fb^{-1} of data, corresponding to the end of Run-2. This prospect study is based on a simplified but realistic framework for systematics and the 8 TeV Monte Carlo samples are PDF re-weighted. Two signal regions are proposed and a precision of 60% on the signal strength is reached. When combined with other measurements performed in $t\bar{t} + H$ ($\rightarrow \gamma\gamma$) and $t\bar{t} + H$ ($\rightarrow b\bar{b}$) channels the top Yukawa coupling can be measured with a precision below 30%.

French version

Le Modèle Standard de la physique des particules décrit les phénomènes observés à l'échelle subatomique avec une grande précision. Cependant, certaines questions demeurent ouvertes : la masse non nulle des neutrinos est inexpliquée, de même que l'existence de la matière noire; de plus, hypothétiser la validité du modèle jusqu'à l'échelle de Planck entraîne un important problème de hiérarchie de jauge. Un nouveau modèle étendant le domaine de validité du Modèle Standard est ainsi nécessaire. Parmi différentes théories proposées à ces fins, la Supersymétrie (SUSY) occupe une place importante dans le programme de recherche de nouvelle physique du LHC. Cet accélérateur génère des collisions de faisceaux de protons à des énergies atteignant 14 TeV, qui sont analysées par de volumineux détecteurs. Dans les études présentées ici, seules les données collectées par le détecteur ATLAS sont analysées. La première partie du document décrit en détails le contexte théorique sous-jacent à ces études, ainsi que lecolllisioneur LHC et l'expérience ATLAS.

Le LHC étant un collisionneur hadronique, l'identification et la reconstruction des leptons est cruciale pour les mesures précises de section efficaces des processus du Modèle Standard, ou la recherche de nouvelle physique. C'est pourquoi la deuxième partie du document est dédiée aux mesures de performances : mesures in situ des efficacités de reconstruction et d'identification des électrons. La méthodologie employée pour obtenir ces résultats, et les méthodes d'estimation du bruit de fond associé sont détaillées. La recherche de nouvelle physique est présentée dans la troisième partie du document. Cette étude a été réalisée dans le canal incluant deux leptons de même charge et de l'énergie transverse manquante. Les résultats inclus ont été obtenus avec un ensemble de données collectées à $\sqrt{s} = 8$ TeV correspondant à une luminosité intégrée de $L = 20$ fb^{-1}, ainsi que les premières données collectées avec $\sqrt{s} = 13$ TeV. Après la première phase d'opération du LHC, aucun excès significant n'a été observé dans les données par rapport au préditions du Modèle Standard, et de fortes contraintes ont été placées sur les masses des partenaires supersymétriques. Cette étude en particulier contraint les masses du gluino et du squark bottom à des valeurs supérieures à respectivement 1 TeV et 600 GeV, sous certaines hypothèses. Ces résultats représentent des contraintes nouvelles et importantes sur les modèles SUSY naturels.

L'état final constitué d'une paire de leptons de même charge et de jets permet également la mesure directe du couplage de Yukawa pour le quark top, à l'aide du processus $pp \to t\bar{t}H$. Ce couplage est le plus élevé dans le Modèle Standard et associe les deux particules élémentaires les plus lourdes prédites par la théorie, le quark top et le boson de Higgs. A l'énergie de $\sqrt{s} = 8$ TeV les mesures sont dominées par les incertitudes systématiques et seules une limite supérieurs peut être placée sur le couplage. Dans la quatrième partie de ce document, une étude complète de la sensibilité pouvant être atteinte avec une énergie de $\sqrt{s} = 13$ TeV et une luminosité intégrée de 100 fb^{-1} of data, correspondant à la deuxième phase d'exploitation du LHC, est présentée pour le canal incluant deux leptons de même charge. Cette étude prospective est basée sur un modèle simple mais approché des incertitudes systématiques affectant la mesure. Deux régions de signal sont proposées et une précision de 60% sur la magnitude du signal mesuré est atteinte. En combinaison avec d'autres mesures réalisées dans les canaux $t\bar{t} + H$ ($\to \gamma\gamma$) et $t\bar{t} + H$ ($\to b\bar{b}$), le couplage de Yukawa du quark top pourrait être mesuré avec une précision de l'ordre de 30%.

Acknowledgments

First of all I want to thank Calin Alexa, Pascal Pralavorio and Emmanuel Monnier for giving me the great opportunity to follow a joint PhD affiliated both to the national Romanian Institute for R&D in Physics and Nuclear Engineering (IFIN-HH, Bucharest) and the Centre de Physique des Particules de Marseille (CPPM), and for the successful collaboration during the past years. I am grateful for all the support and encouragement, for the long discussions and good advice. Their dedication, enthusiasm and motivation inspired me during the entire PhD. In particular, I want to thank Calin for accepting me in the Romanian ATLAS group five years ago, introducing me the field of Particle Physics and learning me the mysteries of Pythia. To Emmanuel for *twisting hands* and giving me the chance to spend such a wonderful and fruitful period in the ATLAS group from CPPM, as a PhD student. Thank you Pascal for believing in me, for discussing with me and giving me all the invaluable advice! I am very grateful for all your help and support during the past three years. Thank you also for proposing me to work on the prospect studies and learning me how important is to change the research topic. Working closely with you was a great experience! I am looking forward a new productive collaboration with IFIN-HH and CPPM ATLAS physics groups.

I want to thank Julien Maurer (CPPM, IFIN-HH) for guiding me, and for all his invaluable advice and help. You learned me to code, to perform the data analysis and understand the mysteries of fake leptons. You saved me so many nights of sleep by helping me to debug and understand the puzzles from the 1-lepton and SS-leptons SUSY analyses. Collaborating with you (and even being supervised by you) since my second year of Master studies helped me to evolve as a scientist, to acquire a lot of knowledge and I am very grateful for this. Your ethic, professionalism and perseverance influenced me during the past years.

David Côté (CERN), Jean-Francois Arguin (U. Montreal) and Ximo Poveda (CERN) closely followed my activity in the SS-leptons SUSY group and helped me to improve my results, understand the statistical interpretation tool and gain more and more experience with the detector backgrounds. Your super fast feedback and the interesting discussions carried at CERN in B40 helped me to progress and succeed in my work. I will never forget the nights spent with you and Julien Maurer on skype until late in the morning to understand the problems with the fake lepton rates or the disagreement on the validation plots. This just because we had to send the note in time to be reviewed by the ATLAS collaboration... Your enthusiasm, dedication and motivation inspired and encouraged me. Thank you for everything!

Working in such a large experiment like ATLAS gave me the chance to collaborate with scientist around the world. In particular I want to thank Alex Koutsman (CERN) and Jeanette Lorenz (Freiburg) from 1-lepton SUSY group for the very useful discussions on HistFitter related issues. I worked in the SS-leptons SUSY group for few years and I enjoyed being part of the team! I want to thank all my colleagues from Cambridge, U. Montreal, CERN, CPPM, Freiburg and Chicago. In particular I want to express my gratitude to Tom Gillam (Cambridge) : working together on different topics of the analysis helped me to evolve, and your great activity on the dynamic matrix method just made our life easier. Good luck in all what you are doing! From U. Montreal group I want to thank especially Léa Gauthier. It was very nice to work close to you and fight against all the problems brought by the fake leptons! Hopefully we will see each other in Canada soon.

My colleagues from CPPM helped me to start the electron performance studies. As the period was quite busy, I appreciate the time they spent to explain me the concepts, the methodology, the framework and, most important, the motivation to perform such measurements. Thank you Thomas Serre and Elodie Tiouchichine!

I consider myself very lucky for having the opportunity to closely collaborate with Fabrice Hubaut for the prospect studies. I am grateful for all fruitful and interesting discussions, and for your help. I really hope I will have another occasion to learn from your experience.

During my qualification task I had the opportunity to work with Louise Heelan and Marie Helene Genest. I thank you for your encouragement and invaluable support.

I want to thank Lucien Heurtier (CPHT, Polytechnique-Paris) for the great collaboration on the Z' paper, and also for the nice period spent in Manchester. Your enthusiasm is contagious! A big thank

you also to all my colleagues and friends from IFIN-HH and CPPM, especially to Marina Rotaru, Gabriel Stoicea, Monica Dobre, Albert Olariu, Adam Jinaru, Alexandra and Valentina Tudorache, Walaa Kanso, Asma Hadef and Venugopal Ellajosyula.

I thank to my jury for accepting to be part of the evaluation, for reading my thesis and attending my defense!

Finally I would like to thank to my parents, to my brother and friends for being so patient with me, for encouraging me and offering me their unconditional support.

Contents

Introduction

The Standard Model of particle physics was developed in the second half of the XX^{th} century as a common effort of scientists from around the world. It is based on relativistic quantum mechanics and gauge theories and classifies all the elementary particles and their interactions. Its great success in describing the phenomena in the infinitesimal is shown by its high predictive power which culminated in July 2012 when the last predicted SM elementary particle was discovered by the ATLAS and CMS experiments, at the LHC. Beside all its accomplishments, several discrepancies with the experimental observations or problems in the mathematical description of the fundamental interactions exist. Therefore, a new theory built on top of the Standard Model should exist and associated new physics should be discovered. If new physics is at TeV scale (or below) it can be discovered at the LHC. As the new physics can *hide* in any region of the phase space, the LHC physics program includes searches for a wide spectrum of Beyond the Standard Model scenarios. Driven by the exciting motivation to look for new physics, this thesis is dedicated to search for signals predicted by Beyond the Standard Model theories, in particular by Supersymmetry.

In the first part overviews of the theoretical framework and the ATLAS experiment are presented, while in the following parts the studies performed with the LHC data are detailed. Chapter 1 is dedicated to the Standard Model. The Supersymmetry is introduced in Chapter 2 and includes the main motivation to consider it as an extension of the SM. At the end of the chapter the status of SUSY searches after LHC Run-1 is presented. The ATLAS experiment is discussed in Chapter 3 and the object reconstruction in ATLAS is detailed in Chapter 4.

The second part starts with Chapter 5, where the electron reconstruction and identification efficiencies measurements with $L = 20$ fb^{-1} of $\sqrt{s} = 8$ TeV data are presented. The precise measurement of these efficiencies plays a crucial role in the SM measurements and are used by all ATLAS analyses. In the third part of this thesis, the search for strongly produced SUSY particles in events with two same - sign leptons is described. Chapter 6 starts with a description of the considered SUSY models, which are either gluino mediated or squark pair production. Further the signal region definitions and their optimization are presented. The background categories and the methods used to estimate them are discussed in detail. Finally, in the absence of any significant excess, limits on the SUSY particles mass are placed in model independent and model dependent way. Chapter 7 describes a combination of this search with a three lepton final state. This effort maximizes the sensitivity of the previous mentioned SUSY search in several regions of the phase space. Therefore new models are included and a new signal region re-optimization is performed. Several improvements are performed and new methods are developed to estimate the background in the signal regions. As no significant excess is observed the limits on the sparticles masses are tightened. Generally, the gluino mass is excluded up to 1 TeV and the stop and sbottom masses up to 600 GeV. Chapter 8 presents the very first results from the analysis searching for a potential SUSY signal at $\sqrt{s} = 13$ TeV.

The fourth part is dedicated to prospect studies for the top quark Yukawa coupling measurement with the Run-2 LHC data in the channel where a pair of top quarks is produced in association with a Higgs boson (Chapter 9). From the possible Higgs boson decays the $t\bar{t} + H$ ($\rightarrow WW$, $\tau\tau$ and ZZ) mode is considered with a final state containing jets and two same - sign leptons, one of the most sensitive channels. A complete signal region optimization is performed to reach the highest sensitivity for this signal at $\sqrt{s} = 13$ TeV with 100 fb^{-1} of data. In addition, a dedicated study is carried to search and define control and validation regions for the main sources of background. The impact on the results in the signal regions and implicit consequences on the reached significance is presented. The thesis ends with general conclusions.

Chapter 1

The Standard Model

In particle physics the Standard Model (SM) provides an excellent description of the phenomena at the electroweak scale. As shown in Section 1.1, it has a total of seventeen fundamental particles : six quarks, six leptons, four gauge bosons and one Higgs boson. If the photons and gluons are massless, the electroweak bosons are massive and they are the propagators of the electroweak and strong interactions discussed in Section 1.2. The SM is a non-abelian gauge theory with the $SU(3)_C \times SU(2)_L \times U(1)_Y$ symmetry group as presented in Section 1.3. The mathematical tool used to break the electroweak symmetry and give masses to the involved particles is known as the Brout-Englert-Higgs mechanism and detailed in Section 1.4. Finally, Sections 1.5 and 1.6 are dedicated to the SM model limitations.

1.1 Standard Model constituents

In the Standard Model, the elementary particles are classified as fermions (half integer spin) and bosons (integer spin). The formers are quarks or leptons and form the fundamental structure of matter, while the vector bosons (spin 1) are the propagators of the fundamental interactions. Their classification is proposed in Table 1.1.

	Particle name and Particle mass		(Q, T_3, Y_W)	S	Color
Leptons	(e, μ, τ)	(0.5, 105.7, 1776.8) MeV	$(-1, -\frac{1}{2}, -1)$	$\frac{1}{2}$	No
	(ν_e, ν_μ, ν_τ)	$0 < m_\nu < 2$ eV ($<$ 2.2 eV, $<$ 0.2 MeV, $<$ 15.5 MeV) *	$(0, +\frac{1}{2}, -1)$	$\frac{1}{2}$	No
Quarks	U : (u, c, t)	(2.3 MeV, 1.3 GeV, 173.2 GeV)	$(+\frac{2}{3}, +\frac{1}{2}, +\frac{1}{3})$	$\frac{1}{2}$	(R, G, B)
	D : (d, s, b)	(4.8 MeV, 95 MeV, 4.2 GeV)	$(-\frac{1}{3}, -\frac{1}{2}, +\frac{1}{3})$	$\frac{1}{2}$	(R, G, B)
Gauge bosons	(g, γ, Z)	(0, 0, 91.2) GeV	(0, 0, 0)	1	(RGB, No, No)
	$W\pm$	80.4 GeV	$(\pm 1, \pm 1, 0)$	1	No
	graviton, G	unknown	(0,0,0)	2	No
Higgs boson	H	125.7 GeV	(0, -$\frac{1}{2}$, 1)**	0	No

(*) results obtained when considering the neutrino oscillations; (**) the complex scalar field Ψ in the SU(2) representation forms an isospin doublet (H^+, H^0), where H^+ has $T_3 = +\frac{1}{2}$ and H^0 has $T_3 = -\frac{1}{2}$. Its anti-doublet is (H^{0*}, H^-), where H^{0*} has $T_3 = +\frac{1}{2}$ and H^- has $T_3 = -\frac{1}{2}$.

Table 1.1: A presentation of the Standard Model elementary particles and some of their important properties [1, 2] : their mass, the electric charge (Q), the spin (S), the third component of the weak isospin (T_3) and the weak hypercharge (Y_W). These properties are mentioned only for the left-handed particles. For the right-handed fermions $T_3 = 0$ and $Y_W = 2(Q - T_3) = 2 \cdot Q$. Even if not included in the SM, the graviton properties are shown for completeness.

The **charged leptons** are elementary half-spin particles with the electric charge Q = -1, which are not affected by the strong interaction, only by the electromagnetic and weak forces. The right-handed leptons are not involved in the weak interactions ($T_3 = 0$).

The **neutral leptons** or neutrinos are weakly interacting particles with a mass equal to zero within

the Standard Model. However, the experimental results are showing that their masses are different from zero [3].

All the SM **quarks** are massive, carry fractional electric charge, weak isospin and color charge. They are the only elementary particles affected by all fundamental interactions. Due to the color confinement phenomenon [4] they are never directly observed in isolated states, only within hadrons. One exception is the top quark which decays before hadronizing.

The **gauge bosons** are vector bosons (spin 1) mediating the fundamental interactions. The photon is associated to the electromagnetic interaction and it is the $U(1)_Y$ gauge group generator. The Z and $W^{\pm}$ bosons are related to the weak interaction and are the $SU(2)_L$ generators. The 8 types of gluons carry the strong force and are the $SU(3)_C$ group generators.

The **Higgs boson** is a scalar boson (spin 0) associated to the Higgs field and observed at the ATLAS and CMS experiments for the first time in July 2012. Its mass is measured to be around 125 GeV [5–13]. After its discovery, a central role in the LHC physics program is occupied by the measurement of the Higgs boson properties, like the Yukawa couplings. Prospect studies performed to measure the top quark Yukawa coupling using $L = 100$ fb^{-1} of Run-2 data ($\sqrt{s} = 13$ TeV) are presented in Chapter 9.

1.2 Electroweak and strong interactions

The Standard Model describes the electromagnetic, strong and weak interactions. A short historical introduction is presented in this section and the mathematical description is shown in Section 1.3 [2, 14–19]. Note that the gravitational force is not included in the SM.

Electroweak interaction

The weak interaction was proposed for the first time by Fermi in 1933 to explain the β decay as a four fermions interaction [20]. It was therefore called *Fermi's interaction* and considered to be a direct contact between the involved fermions – where the spin of the electron and the spin of its associated neutrino are considered to be parallel. This form of interaction was also explaining the muon decay into an electron and the associated neutrinos. Later, Gamow and Teller proposed an extension of this interaction to include the case where the spins are anti-parallel [21].

If the parity conservation was known for the strong and electromagnetic interactions, several (ignored) papers were theoretically proving a non conservation of the weak interaction. This was studied more deeply by Lee and Yang [22] when they also proposed several experiments to show the parity violation. In 1957, Wu found a clear evidence of this hypothesis after monitoring the β ray emissions from a Co^{60} source of atoms cooled near zero absolute and placed in a uniform magnetic field; it was further confirmed by Garwin, Lederman and Weinrich [23]. Given the experimental results, in the same year Marshak and Sudarshan proposed a new theoretical approach of the weak theory describing it with a $V - A$ Lagrangian (vectorial - axial symmetries) [24]. They considered the observation to be a result of the interaction of the weak force mediators only with left-handed particles. This also led to the conclusion that the neutrinos are left-handed, while the anti-neutrinos are right-handed particles. A more robust theory was written in 1968 by Glashow, Salam and Weinberg, when they proposed a complete description of an unified electromagnetic-weak interaction using the Yang-Mills framework (which is described in Section 1.3) [25, 26]. It was known as the electroweak interaction (EW) and experimentally recognized in 1973 at the Gargamelle experiment when the neutral currents were discovered and in 1983 when the W and Z bosons were discovered.

Strong interaction

A first hint of strong interaction arrived with the discovery of neutron in 1932, as a force able to hold together the neutrons and protons forming the nucleus. In 1934 the nature of this interaction was described by Yukawa [27], where the carrier was a spin zero particle, a meson – the pion – discovered in 1947 by Powell, Lattes and Occhialini [28]. In the following period, several theories formulating the strong interaction were proposed. The field propagators were considered to be the **elementary** baryons and mesons. However, the experimental results obtained in the period after ∼1952, when the bubble and spark chambers were invented, raised doubts about the elementarity of the hadrons. A forward step in understanding the nature of the strong interaction was made in 1961, when Gell-Mann proposed the *Eightfold Way* or the SU(3) flavor theory [29]. It classifies the hadrons in families of particles with similar masses but with the constituents organized depending on their hypercharge and isospin. This organization was performed in multiplets, octets and decouples using the SU(3) group representation. The theory was experimentally confirmed in 1964 when the predicted Ω^- baryon was discovered. Given the obtained results, a first interpretation of the hadron classification was proposed by Gell-Mann and Zweig within the Quark Model [30, 31]. They

suggested that the baryons constituents are three elementary particles (quarks), while the mesons contain only two of them. At that time, only thee quark flavors were considered (u, d, s) to which several quantum numbers like isospin, strangeness, etc, were attributed.

The first experimental quark observation (strange quark) dates back to 1968 at the SLAC experiment when measuring the inelastic electron-proton scattering cross sections [32,33]. The observed cross section-transfered momentum dependency confirmed the Bjorken scaling hypothesis [34] and implicitly the existence of the proton structure. In that period the point-like constituents were called partons, given the Parton Model proposed by Feynman. Later they are identified as the up and down quarks. In 1970 the charm quark was predicted [35] and in 1973 the top and bottom quarks were mentioned for the first time in [36] increasing the number of quarks to six.

Another important contribution toward the hard interaction understanding was the introduction of a new quantum number associated to quarks : the color. It was mentioned for the first time in 1965 and linked with the Ω^- baryon formed by three strange quarks and thus violating the Pauli principle [37]. The Δ^{++} baryon has a similar composition, three up quarks. The theoretical solution was brought by Han, Nambu and Greenberg [38,39] after adding a new SU(3) degree of freedom. This solution implies that the quarks occur in different (color) states. They were also suggesting that the quarks are interacting via gauge bosons (the gluons). In 1972 the number of colors is set to three according to the experimental results obtained for the $\pi^0 \to \gamma\gamma$ decay amplitude and to the ratio of hadrons and muons production cross sections in e^+e^- collisions [40].

After the discovery of the asymptotic freedom in 1973 by Politzer, Wilczek and Gross, the quantum chromodynamics (QCD) field theory was considered to correctly describe the strong interaction [4,41–43]. They showed that in QCD the quarks are bound together when they are at distances comparable with the proton size, while at smaller distances and high energies the bounds are asymptotically weaker. The force carriers are the gluons (massless) and the first experimental evidence was obtained in 1979 at PETRA experiments in a three jet event [44,45]. An interesting property of the strong interaction is the gluon-gluon interactions, which can lead to composite particles called glueballs. They are predicted by QCD lattice calculations with a mass around 1 - 1.7 TeV, but not yet observed [46].

1.3 Gauge symmetry

The fundamental role of symmetries in particle physics has become visible in 1917 when the Noether theorem was published [47]. This theorem connects symmetries and conservation laws : each symmetry corresponds to a conservation law and each conservation law reveals a symmetry. In field theories the existence of a symmetry translates to the invariance of the Lagrangian under a continuous group of local field transformations. These transformations are called gauge transformations and form a Lie group. In such a group, each symmetry can be written as a multiplication of infinitesimal symmetries (it presents a continuous symmetry). They appeared for the first time in the Maxwell formulation of classical electrodynamics (CED), where it was shown that the magnetic and electric fields remain invariant when the potentials $\mathbf{A}$ and $\mathbf{\Psi}$ are changed by adding a gradient or a partial derivative of a scalar function, $\mathbf{A}' = \mathbf{A} + \nabla\ \Lambda$ and $\Psi' = \Psi - \frac{\partial \Lambda}{\partial t}$. The gauge properties are really explored only after the development of quantum mechanics and implicit quantum electrodynamics. The latter is the first quantified theory [48–53].

Electromagnetic interaction (QED)

The Lagrangian describing a free Dirac complex field Ψ[1]

$$\mathcal{L}_0 = \bar{\Psi}(i\gamma^\mu \partial_\mu - m)\Psi \tag{1.1}$$

is invariant under a U(1) global gauge transformation $\Psi \to \Psi' = e^{i\alpha}\Psi$, where $\alpha \in \mathbb{R}$. If a U(1) local transformation $\alpha(x)$ is considered, the Lagrangian is no longer invariant. However, the gauge principle implies a theory described by a locally invariant Lagrangian. In this context, a new massless spin 1 gauge field A_μ needs to be added (no terms of the form $\bar{\Psi}A_\mu\Psi$) and transforms as :

$$A_\mu(x) \to A'_\mu(x) = A_\mu(x) + \frac{1}{e}\partial_\mu \alpha(x) \tag{1.2}$$

To ensure the local invariance, the partial derivative ∂_μ is replaced by a covariant derivative $\mathcal{D}_\mu = \partial_\mu - ieA_\mu$. These changes are only mathematical and the physics is not altered. After implementing these substitutions

[1] There is an implied assumption of the twice repeated index $\mu = 0, 1, 2, 3$ and the ∂^μ are 4×4 matrices

in Equation 1.1 and adding the free term for the gauge field, the complete local invariant Lagrangian is :

$$\bar{\mathcal{L}}_{\text{QED}} = \underbrace{\bar{\Psi}(i\gamma^\mu\partial_\mu - m)\Psi}_{E_{kin} \text{ and mass of } \Psi} + \underbrace{e(\bar{\Psi}\gamma^\mu\Psi)A_\mu}_{\Psi \text{ - } A_\mu \text{ interaction}} \underbrace{-\frac{1}{4}F^{\mu\nu}F_{\mu\nu}}_{E_{kin} \text{ of } A_\mu} \tag{1.3}$$

where A^μ is the electromagnetic potential field, e is the electric charge, $e\bar{\Psi}\gamma^\mu\Psi A_\mu$ is the current produced by the Dirac particles and $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$ is the electromagnetic field tensor. Shortly, the Lagrangian expresses the interaction between the Dirac fields (Ψ) and the electromagnetic massless field (A_μ). Note that QED is an abelian group theory.

Yang - Mills theory

Using a field with two $\frac{1}{2}$ spin components, $\Psi = \begin{pmatrix} \Psi_1 \\ \Psi_2 \end{pmatrix}$, and inspired by the hypothesis of a nucleon with two states of the same mass (the proton and the neutron), Yang and Mills generalized in 1954 the QED formalism to non-abelian groups [54], i.e this consider multiple massless spin-one particles which can interact among themselves. Since then, this framework has proven to be extremely powerful to describe other gauge theories. Few examples are the quantum chromodynamics (QCD) based on SU(3) symmetry group and the grand unification theories. In the Yang-Mills theory, a unitary transformation $\Psi(x) \to \Psi'(x) = U(x)\Psi(x)$ under a non-abelian Lie-Group G of dimension N is considered. Note that compared to QED, the U(x) transformations do not commute and the group generators, T^a, are matrices of dimension $N \times N$. To ensure the local gauge invariance, a covariant derivative is used and n extra gauge fields $A_\mu = \sum_a^{N^2-1} A_\mu^a T^a$ are added. Generally, the global and local invariant Lagrangian can be written as :

$$\bar{\mathcal{L}}_{\text{YM}} = \bar{\Psi}(i\gamma^\mu\mathcal{D}_\mu - m)\Psi - \frac{1}{4}\,\text{Tr}\{F_{\mu\nu}F^{\mu\nu}\} \quad \text{where} \tag{1.4}$$

$$\mathcal{D}_\mu = \partial_\mu - ig\sum_{a=1}^{N^2-1} nA_\mu^a, \quad g \text{ is the coupling constant} \tag{1.5}$$

$$A_\mu = UA_\mu U^{-1} - \frac{i}{g}U\partial_\mu U^{-1} \tag{1.6}$$

The interaction between the extra gauge fields (boson) and the Dirac field Ψ is given by the $\bar{\Psi}\mathcal{D}_\mu\Psi$ term ($\bar{\Psi}A_\mu\Psi$). It can be observed that the n extra massless gauge fields introduce n kinetic terms. To fulfill the local invariance requirement, the $F^{\mu\nu}$ term should be also modified :

$$F^{\mu\nu} = [\mathcal{D}^\mu, \mathcal{D}^\nu] = \partial^\mu A^\nu - \partial^\nu A^\mu + ig[A^\mu, A^\nu] \tag{1.7}$$

The $[A^\mu, A^\nu]$ commutator is zero for abelian gauge theories because the photons are not self-interacting, while the $W^\pm$ or Z bosons are.

Weak and electroweak interaction

In the Yang-Mills theory the weak interaction can be described in terms of a SU(2) symmetry. In this case, the $U(x)$ matrices are written as $e^{i\alpha^a\tau^a}$. Here $a = \overline{1,3}$, α is a real constant and $\tau^a = \frac{1}{2}\sigma^a$. σ^a denotes the Pauli matrices. The Dirac field transformation is $\Psi(x) \to \Psi'(x) = e^{i\alpha^a\tau^a}\Psi(x)$. For such a symmetry the global and local invariant Lagrangian is simplified to :

$$\bar{\mathcal{L}}_{SU(2)} = \bar{\Psi}(i\gamma^\mu\partial_\mu - m)\Psi - g_w(\bar{\Psi}\gamma^\mu\boldsymbol{\tau}\Psi)\mathbf{A}_\mu - \frac{1}{4}\,\text{Tr}\{F_{\mu\nu}F^{\mu\nu}\} \tag{1.8}$$

where the first term can be identified with the Lagrangian of a free Dirac field. When it is compared with the QED Lagrangian from Equation (1.3) three new extra gauge field are observed : the $W^\pm$ and W^0 massless gauge bosons. Note that, the term *massless boson* is used as these bosons are massive only after the spontaneous symmetry breaking (Section 1.4). The second term can be interpreted as three currents generated by the Dirac field (here g_w is the coupling constant). The last term is the gauge invariant kinetic term corresponding to the A_μ field for QED. This Lagrangian is also the baseline of the electromagnetic and weak forces unification based on a U(1) $\times$ SU(2) group symmetry.

Strong interaction (QCD)

According to the colored quark model, each quark carry one of the three color states : red, green and blue. The color states are connected by a SU(3) symmetry group where the allowed gluon states are described

by the Gell-Mann matrices. These matrices are the eight generators of the group (λ^a, $a = \overline{1,8}$) and the gluon states are the propagators of the strong interaction. The free Lagrangian of such a system is :

$$\mathcal{L} = \sum_f \bar{q}_f (i\gamma^\mu \partial_\mu - m_f) q_f \tag{1.9}$$

where the quark field of color a and flavor f is denoted with q_f^a. It is invariant under a global transformation $q_f^a \to (q_f^a)' = U_\beta^\alpha q_f^\beta$, where $U = e^{-ig_s \frac{\lambda^a}{2} \alpha_a}$ is a 3×3 unitary matrix describing a rotation in the color space, with g_s the strong coupling constant and α_a some arbitrary group parameters. If an infinitesimal local transformation is considered, the quark field transforms as :

$$q_f^a(x) \to [1 - ig_s \frac{\lambda^a(x)}{2} \partial\alpha_a(x)] q_f^a(x) \tag{1.10}$$

To achieve the local invariance of the Lagrangian, a covariant derivative is used and eight external massless gauge (gluon) fields are added similarly to QED :

$$\mathcal{D}_\mu = \partial_\mu - ig_s \frac{\lambda^a}{2} G_\mu^a(x) = \partial_\mu - ig_s G_\mu(x) \tag{1.11}$$

$$G_\mu \to G_\mu^a - \frac{1}{g_s} \partial_\mu \alpha_a - g_s \alpha_b G_\mu^c \tag{1.12}$$

Equation (1.12) presents the the gauge transformation of the gluon flied. The final global and local Lagrangian describing the interaction between gluons and quarks is simply :

$$\bar{\mathcal{L}}_{QCD} = \sum_f \bar{q}_f (i\gamma^\mu \mathcal{D}_\mu - m_f) q_f - \frac{1}{4} G_a^{\mu\nu} G_{\mu\nu}^a \tag{1.13}$$

and symbolically it can be written as :

$$\bar{\mathcal{L}}_{QCD} = \text{"}\bar{q}q\text{"} + \text{“}g_s\text{"}\bar{q}qG\text{"} + g_s\text{"}G^3\text{"} + g_s^2\text{"}G^4\text{"} \tag{1.14}$$

where

- "$\bar{q}q$" is the kinematic terms for the quarks;
- g_s"$\bar{q}qG$" stands for the color interaction between quarks and gluons and involves the SU(3) λ^a matrices representing the gluons;
- in the g_s"G^3" + g_s^2"G^4" the term $G_a^{\mu\nu} G_{\mu\nu}^a$ generates the cubic and quadratic gluon self-interactions, with their strength given in terms of the universal coupling constant, g_s.

1.4 Brout-Englert-Higgs mechanism

To maintain the local invariance of the Lagrangian describing an interaction, the external gauge bosons must have a mass equal to zero. This requirement is fulfilled for QCD and QED but not for the weak interaction where massive ($W^\pm$ and Z) gauge bosons are needed. To generate the mass, a mathematical tool called the Brout-Englert-Higgs mechanism was proposed in 1962 - 1964 [55–57]. It is based on spontaneous symmetry breaking of the electroweak gauge sector and introduces a new scalar with a specific form of the potential able to maintain the entire Lagrangian invariant under the SU(2)×U(1) symmetry group. Another problem in the electroweak model is the requirement of massless fermions. If they are considered to be massive (as the experiments are showing), the local invariance is lost. Both *problems* can be explained with the Brout-Englert-Higg mechanism after adding the lepton and quark interactions with the Higgs boson [14–19,58,59].

Spontaneous discrete symmetry breaking

To illustrate the spontaneous discrete symmetry breaking, a Lagrangian describing a scalar particle Ψ with a potential V is considered :

$$\mathcal{L} = E_{kin} - V = \frac{1}{2}(\partial_\mu \Phi)^2 - (\frac{1}{2}\mu^2 \Phi^2 + \frac{1}{4}\lambda \Phi^4) \tag{1.15}$$

where μ^2 and λ are real constants and $\mu^2 < 0$[2]. The second term is similar to a mass term (only the sign is different), $\mathcal{L}$ is invariant under $\Phi \to -\Phi$ and the minimum potential energy is reached for $\Phi = \pm\sqrt{-\frac{\mu^2}{\lambda}} =$

[2]If $\mu^2 > 0$ the Lagrangian describes just a scalar field of mass μ. The $\mu^2 < 0$ choice is for the reflection symmetry of the Lagrangian.

$\pm v$. If a small perturbation around the minimum is performed, the scalar field Φ can be rewritten in terms of a new field variable η :

$$\Phi = v + \eta \quad \text{or} \quad \Phi = -v + \eta \tag{1.16}$$

Considering the first case[3], the Lagrangian from Equation (1.15) becomes :

$$\mathcal{L}' = \frac{1}{2}(\partial_\mu \Phi)^2 - \boldsymbol{\lambda v^2 \eta^2} - \lambda v \eta^3 - \frac{1}{4}\lambda \eta^4 + \text{const} \tag{1.17}$$

Note that $\mathcal{L}'$ is not invariant under a $\eta \to -\eta$ transformation because the symmetry was spontaneously broken by the choice of the ground state. If this Lagrangian is compared to Equation (1.1) the mass term can be identified with :

$$\frac{1}{2} m_\eta^2 \eta^2 = \lambda v^2 \eta^2 \Rightarrow m_\eta = \sqrt{2\lambda v^2} = \sqrt{-2\mu^2} \tag{1.18}$$

Conclusion : considering a case with a discrete symmetry with only two ground states, the Lagrangian symmetry is spontaneously broken (no external factor is responsible) and a mass term (with a correct sign) was generated.

Spontaneous continuous symmetry breaking

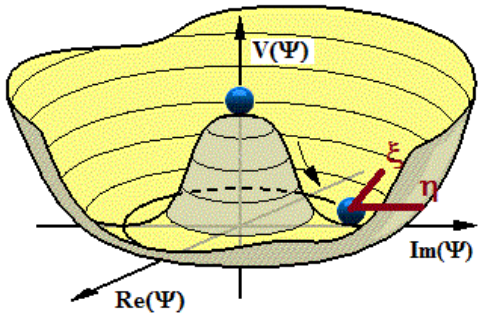
More interesting results are obtained when a continuous symmetry is considered. In this case, two real scalar fields combined into a complex scalar field are used, $\Phi = \frac{\Phi_1 + i\Phi_2}{\sqrt{2}}$, and the Lagrangian is written as :

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \Phi_1)^2 + \frac{1}{2}(\partial_\mu \Phi_2)^2 - \frac{1}{2}\mu^2(\Phi_1^2 + \Phi_2^2) - \frac{1}{4}\lambda(\Phi_1^2 + \Phi_2^2)^2 = (\partial_\mu \Phi)^*(\partial_\mu \Phi) - \mu^2 \Phi^* \Phi + \lambda(\Phi^* \Phi)^2 \tag{1.19}$$

$\mathcal{L}$ is invariant under a U(1) phase transformation, $\Phi \to e^{i\alpha}\Phi$: is having a U(1) global gauge symmetry.

The minimum potential energies are forming a circle of radius $v^2 = -\frac{\mu^2}{\lambda}$ in the (Φ_1, Φ_2) plane[4]. If a small perturbation around the ground state is performed, the complex scalar field Φ is written in terms of two real scalar fields η and ξ :

$$\Phi(x) = \sqrt{\frac{1}{2}}(v + \eta(x) + i\xi(x)) \tag{1.20}$$



The Lagrangian of the system will be written as :

$$\mathcal{L}' = \frac{1}{2}(\partial_\mu \eta)^2 + \frac{1}{2}(\partial_\mu \xi)^2 + \underbrace{\boldsymbol{\mu^2 \eta^2}}_{\text{mass term}} - \underbrace{\lambda v(\eta^3 + \eta\xi^2)}_{\text{field interactions}} - \underbrace{[\frac{\lambda}{4} \cdot (\eta^4 + \xi^4) - \frac{\lambda}{2}\eta^2\xi^2]}_{\text{Kinetic energy}} - \underbrace{\frac{1}{2}\mu^2 v^2}_{\text{const}} \tag{1.21}$$

and it is not invariant under the $\eta \to -\eta$ and $\xi \to -\xi$ transformations.

Conclusion : when the spontaneous breaking of a continuous symmetry (SSB) is considered for a system described by a complex field, a massive (η) and a massless (ξ) scalar fields are obtained. This is a consequence of the Goldstone's theorem and ξ is called the Goldstone boson.

The Brout-Englert-Higgs mechanism for a local invariant U(1) symmetry

The Brout-Englert-Higgs mechanism is needed to obtain a gauge theory for the electroweak interactions with massive gauge bosons. The previous results are not satisfactory because a massless boson appears after the SSB. This issue is solved if a Lagrangian invariant under a local gauge transformation is used (Section 1.3). In the case of a U(1) symmetry the SSB is obtained after considering a small perturbation around the ground state :

$$\Phi(x) = \sqrt{\frac{1}{2}}(v + h(x))e^{i\frac{i\alpha(x)}{v}} \tag{1.22}$$

In this case, the Lagrangian from Equation (1.15) takes the final form :

$$\bar{\mathcal{L}}' = [\frac{1}{2}(\partial_\mu h)^2 - \boldsymbol{\lambda v^2 h^2}] + [\frac{1}{2}\boldsymbol{e^2 v^2 A_\mu A^\mu} + \frac{1}{2}e^2 A_\mu^2 h^2 + v e^2 A_\mu^2 h] - [\lambda v h^3 - \frac{1}{4}\lambda h^4] - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{1.23}$$

and includes massive particles and massive gauge fields :

[3] For simplicity only $\Phi = v + \eta(x)$ is considered without loosing the generality.
[4] For simplicity only the case with $\Phi_1 = v$ and $\Phi_2 = 0$ is considered.

- The first bracket describes a scalar **massive particle** which can be identified with the Higgs boson ($M_H = v\sqrt{2\lambda}$). The λ parameter is the strength of the Higgs boson self-interaction in the Higgs potential and, unlike in Equation (1.21), the Goldstone boson is not present anymore.
- The second bracket describes the **massive** external gauge field A_μ ($e\mu/\sqrt{\lambda}$ is the mass term).
- The third bracket stands for the Higgs boson - external gauge field interaction terms.

The Higgs mechanism for a local invariant SU(2)×U(1) symmetry

Extending the symmetry group to SU(2)×U(1) and applying the Brout-Englert-Higgs mechanism, the masses for the $W^\pm$ and Z bosons are obtained. The Higgs field is considered to be a SU(2) doublet of complex scalar fields Φ (the Higgs doublet) with the weak isospin equal to $-\frac{1}{2}$. Given the gauge invariance it can be expanded around the vacuum state as:

$$\Phi = \begin{pmatrix} \Phi^+ \\ \Phi^0 \end{pmatrix} \rightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}$$

The electric charges of the two fields are chosen such that the hypercharge of the system is +1. The choice of the (neutral) vacuum state is assuring a symmetry breaking for $SU(2)_L \times U(1)_Y$ and keeps the $U(1)_{QED}$ transformation invariant. This is mandatory because the photon is massless. The covariant derivative is written as :

$$\mathcal{D} = \partial_\mu + ig\frac{1}{2}\boldsymbol{\tau}\boldsymbol{W}_\mu + ig'\frac{1}{2}YB_\mu \tag{1.24}$$

where $\boldsymbol{\tau}$ and $\boldsymbol{W}$ are vectors with three components. $\boldsymbol{\tau}$ is the SU(2) group generator (Pauli matrices), Y is the U(1) group generator and B_μ is the U(1) gauge boson. Rewriting only the first part of the Lagrangian (Equation (1.19)) and applying the Higgs mechanism, one can obtain the relation :

$$(\mathcal{D}_\mu\Phi)^\dagger(\mathcal{D}_\mu\Phi) = \frac{1}{8}v^2[\underbrace{g^2(W_1^2 + W_2^2)}_{g^2(W^+)^2 + g^2(W^-)^2} + \underbrace{(-gW_3 + g'YB_\mu)^2}_{(g^2+g'^2)Z_\mu^2 + 0\cdot\mu^2}] \tag{1.25}$$

Note that the photon and the Z boson are a mixing between the B and $\mathbf{W}$ fields. The obtained mass terms are identified with :

- $M_{W^+} = M_{W^-} = \frac{1}{2}vg$, where g is the gauge coupling constant.
- $M_Z = \frac{1}{2}v\sqrt{(g^2 + g'^2)} = \frac{m_{W\pm}}{\cos(\theta_W)}$, where $\frac{g'}{g} = \tan(\theta_W)$ or $\frac{M_W}{m_Z} = \cos(\theta_W)$.
- Computing the new potential term in the Lagrangian, the Higgs boson mass is $\sqrt{2\lambda v^2}$, similarly to Equation (1.18).

Fermion masses and the CKM matrix

The Higgs mechanism and the SSB are used to generate masses also for fermions, not only for the gauge bosons. To achieve this, a $SU(2)_L \times U(1)_Y$ gauge invariant term should be constructed and added in the Lagrangian[5]. For the SM fermions this can be easily proved : the left-handed fermions form an isospin doublet ($\chi_L \rightarrow \chi'_L e^{i\boldsymbol{W\tau} + i\boldsymbol{\alpha Y}}$) while the right-handed fermions form singlets ($\Psi_R \rightarrow \Psi'_R = \Psi_R e^{i\alpha Y}$). Hence the new term should be a singlet under all $SU(2)_L \times U(1)_Y$ transformations. Such a term is $-\lambda\bar{\Psi}\Phi\Psi$. Here λ stands for the Yukawa coupling and Φ is the previously defined Higgs doublet. Therefore the term used to give masses to all fermions (and implicit add the fermion - Higgs boson interactions) is gauge invariant and it can be written as :

$$\mathcal{L}_F = -\lambda(\bar{\Psi}_L\Phi\Psi_R + \bar{\Psi}_R\Phi\Psi_L) \tag{1.26}$$

To obtain the **lepton masses** the above relation is rewritten as :

$$\bar{\mathcal{L}}_l = -\frac{m_l\sqrt{2}}{v} \cdot \left[\begin{pmatrix} \bar{\nu}_l & \bar{l} \end{pmatrix}_L \begin{pmatrix} \Phi^+ \\ \Phi^0 \end{pmatrix} l_R + \bar{l}_R \begin{pmatrix} \Phi^- & \bar{\Phi}^0 \end{pmatrix} \begin{pmatrix} \nu_l \\ l \end{pmatrix}_L \right] \tag{1.27}$$

[5] Terms like $-m\bar{\Psi}\Psi = -m[\bar{\Psi}_L\Psi_R + \bar{\Psi}_R\Psi_L]$ cannot be used because they are not gauge invariant.

Note that the constant coupling parameter is chosen to obtain the proper mass term after SSB. To retrieve the **quark masses** the new term needed in Lagrangian is :

$$\bar{\mathcal{L}}_q = -G_D^{ij} \begin{pmatrix} \bar{U}_i & \bar{D'}_i \end{pmatrix}_L \begin{pmatrix} \Phi^+ \\ \Phi^0 \end{pmatrix} D_{jR} - G_U^{ij} \begin{pmatrix} \bar{U}_i & \bar{D'}_i \end{pmatrix}_L \begin{pmatrix} -\bar{\Phi}^0 \\ \Phi^- \end{pmatrix} U_{jR} + \text{ h.c} \tag{1.28}$$

where the $i, j = \overline{1, N}$ with N number of quark doublets. Note that the arbitrary coupling constant from Equation (1.27) is replaced by complex matrices that connect the flavor eigenstates. They cannot be predicted and the U (D) symbol stands for the up (down) quark generation.

After the SSB, the physical mass terms for the leptons and quarks are generated. For the latter, the G matrices should be diagonalized and only the $i = j$ terms must be retained. This is performed by multiplying each of them with a phase-factor unitary matrix : $\Phi^0 G_D^{ij} = \frac{v}{\sqrt{2}} U_{DL} G_D^{ij} U_{DR}^\dagger = \text{diag}(m_d, m_s, m_b)$ and $\Phi^0 G_U^{ij} = \frac{v}{\sqrt{2}} U_{UL} G_U^{ij} U_{UR}^\dagger = \text{diag}(m_u, m_c, m_t)$. Accordingly, the true quark fields (the mass matrices are diagonal) are $D'_{Li} = U_{DLij} D_{Lj}$, $D'_{Ri} = U_{DRij} D_{Rj}$, etc (the unitary matrices U_U and U_D are different). Therefore the couplings between the electroweak charge current interaction associated to the U and D-type quarks are changed leading to a mixing between the different quark flavors, i.e. $(U_D U_U^\dagger)_{ij} = V_{ij}$ where $i \neq j$. The involved term in the Lagrangian can be written as $i\bar{\Psi}(\mathcal{D}^\mu \gamma_\mu)\Psi$, and the mixing appears when the D and U weak eigenstates are replaced by the quark mass eigenstates (change of basis). This mixing is given by the CKM matrix which provides the transition from a D-type quark into an U-type quark. The matrix elements are constructed after applying the weak interaction universality constraint and are expressed as a function of four parameters : three rotation angles and the phase responsible for the CP- violation in flavor-changing SM process. According to present measurements, the magnitudes of the elements are [1] :

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} 0.97427 \pm 0.00015 & 0.22534 \pm 0.00065 & 0.00351^{+0.00015}_{-0.00014} \\ 0.22520 \pm 0.00065 & 0.97344 \pm 0.00016 & 0.0412^{+0.0011}_{-0.0005} \\ 0.00867^{+0.00029}_{-0.00031} & 0.0404^{+0.0011}_{-0.0005} & 0.999146^{+0.000021}_{-0.000046} \end{pmatrix} \tag{1.29}$$

The diagonal terms are bringing the highest contribution, showing that the favorite mixing is obtained with quarks from the same family. The mixing between generation is not applied for leptons, as the neutrinos are considered to be massless.

1.5 Running coupling constants and anomalies

The contribution of loop diagrams with divergent diagrams can be minimized if the energy scale is chosen to be close to the actually exchanged energy in the interaction. With this approach, the physics is not changed, only the contribution from the finite loop diagrams considered for a given scale. A consequence is that the coupling constants (α) are now dependent on the cutoff energy (i.e. not constant anymore). This phenomenon is known as the running of the coupling constants and it is described within the renormalization group theory [15]. The function encoding the running of the different couplings, at a given energy scale μ, is a beta function. If $\beta(\alpha(\mu)) = \frac{b_1}{2\pi} \cdot \alpha^2(\mu)$, i.e. the higher order corrections are ignored, the running coupling constant is simply :

$$\alpha(\mu_2) = \alpha(\mu_1) \frac{1}{1 - \frac{b_1}{2\pi} \alpha(\mu_1) \ln \frac{\mu_2}{\mu_1}} \tag{1.30}$$

where μ_1 is the transfered momentum $\sim 2m_e$ and μ_2 the cutoff scale. In the case of QED, the β function is positive, $\alpha^2(0) = \frac{e^2}{4\pi}$ from Coulomb law in natural units and b_1 is set to $\frac{2}{3}$. The coupling constant dependency on the squared cutoff scale is written as :

$$\alpha(\mu_2^2) = \frac{e^2(\mu_1^2)}{4\pi} = \frac{e^2}{4\pi} \frac{1}{1 - \frac{e^2}{12\pi^2} \ln \frac{\mu_2^2}{4m_e^2}} \tag{1.31}$$

where e is the elementary charge. It can be interpreted as the electron charge screening by virtual e^+e^- pairs created from the vacuum. Thus, at short distances the screening is reduced and the effective charge is greater. If an energy reference around the MeV scale is considered, $\alpha(\mu_2 = 0)$ is simply equal to $\frac{1}{137}$. At the scale of the Z mass it measures $\frac{e^2}{4\pi} = \frac{1}{127}$.

Screening effects are present also in QCD : the quark - antiquark virtual pairs created in the vacuum tend to screen the color. Compared to QED, the gluons are charged and the vacuum is not screening the color field but it is antiscreening it. This can be understood as an increase in the field due to the virtual gluon and quark pairs or by possible changes of its color. Therefore, when going closer and closer to the quark, the antiscreening effect is reduced : the asymptotic freedom property of QCD. This translates into a smaller effective constant when decreasing the distances (Figure 1.1).

Generally, the strong coupling constant can be written as :

$$\alpha_s(\mu_2^2) = \frac{g^2(\mu_2^2)}{4\pi} = \frac{g^2}{4\pi}\frac{1}{1-\frac{g^2}{8\pi^2}[-\frac{11n_c}{3}+\frac{2n_f}{3}]\ln\frac{\mu_2^2}{\lambda^2}}$$

where λ is a parameter replacing the electron mass from QED and the b_1 term is :

$$b_1 = (-\frac{11n_c}{3}+\frac{2n_f}{3}) \tag{1.32}$$

The n_c (n_f) variable stands for the number of colors (flavors) considered for the SU(3) group symmetry. In contrast with QED the β function is negative. The $\alpha_s(\mu_2^2)$ relation can be written only as a function of μ_2 (further denoted also with Q) and one Λ parameter which includes the dependency on λ and g : $\approx ln\frac{Q^2}{\Lambda^2}$.

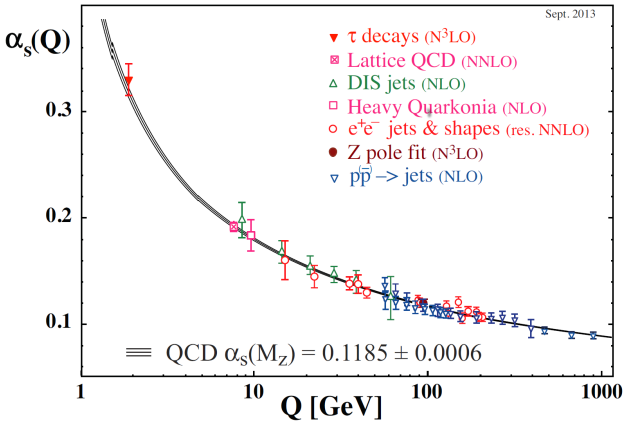


Figure 1.1: Running of the strong coupling constant in QCD [1].

Anomalies

Anomalies appear when the classical field theories are quantified. They are represented by a non conservation or a failure in obtaining a symmetry as in the classical theory. A first evidence is found by Bell, Jackiw and Adler [60,61] when analyzing the axial vector current associated to chiral symmetries : its conservation is violated when requiring the electric charge conservation and maintaining the QED as a local invariant gauge symmetry. However, it can be exactly calculated when considering the one loop diagrams contributing to the pion decays as $\pi^0 \to \gamma\gamma$ or $\pi^0 \to ee\gamma$. Note that these processes are forbidden in the classical chiral theory. In the Standard Model, the anomalies in the electroweak currents are similar to the axial anomaly : the gauge bosons couple with currents that have axial and vectorial components leading to anomalies, i.e $j_e^\mu = e_L^\dagger \bar{\sigma}^\mu \nu_L = \frac{1}{2}\bar{e}\gamma^\mu(1-\gamma^5)\nu_e$. It was shown that in a theory with both lepton and quark families the anomalies cancel exactly, given the same number of generations. This is not the case for a theory considering to not have either leptons or quarks.

1.6 Problems of the Standard Model

The impressive successes in explaining the laws of nature and the high predictive power of the Standard Model at the electroweak scale is shown by the experimental results. All the nineteen parameters of the SM are known and the impressive consistency between the different observables is detailed in [62]. For illustration, Figure 1.3 presents the agreement between the direct and indirect mass measurements of the top quark, W and H bosons. Beside all the accomplishments, clear discrepancies with the experimental observations or problems in the mathematical description of the fundamental interactions exist in the Standard Model. The most obvious is the absence of the **gravity** and implicitly of its mediator, the graviton. They could not be considered in the SM due to the difficulty of translating the general relativity (GR) into a quantum field theory (eg. presence of infinities in the computation).

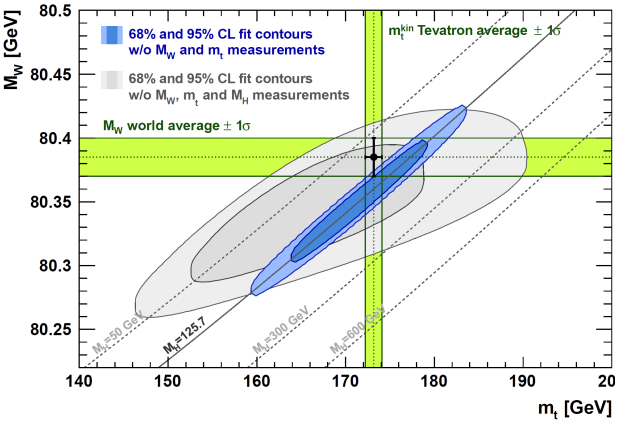


Figure 1.2: Contour plots showing the consistency between several direct and indirect Standard Model physical observables measurements [62].

Another problem is related to the amount of **dark energy and dark matter** in the universe which is not predicted by the Standard Model. Studies of the Cosmic Microwave Background are showing that the universe contains ∼27% dark matter and ∼68% dark energy, while the rest is ordinary matter formed by the elementary particles as we know them [63]. The dark matter was postulated by Oort and Zwicky [64,65] to explain the *fast* orbital velocities of the stars and galaxies. At that time, the result of the whirling of galaxies computation was found to be much higher than expected in order to prevent the starts from

flying away from the galactic center as a result of the centrifugal force. The equilibrium state of the galaxies could be explained by assuming that they are formed by more matter than the visible stars. The first indirect evidence of dark matter was observed by Rubin in 1970 when analyzing the spiral galaxy rotation curves [66]. The next observations were also confirming its existence. Currently several direct and indirect searches for dark matter candidates are performed using the modern technology. A direct detection experiment is usually carried in underground laboratories to reduce the background from cosmic rays. Here, the signal recorded with cryogenic or noble liquid detectors is coming from the dark matter particles which scatter off the nuclei – the XENON and EDELWEISS experiments [67, 68]. The indirect detection targets the signal coming from the annihilation or the decay of dark matter particles. Such experiments are PAMELA, FermiLAT, IceCube or AMS [69–72]. Other techniques to look for dark matter are based on cosmology instruments such as Planck [73]. Finally, the dark matter may also be produced in the laboratory and studied at experiments like the Large Hadron Collider. Until now, no dark matter candidate have been observed experimentally and only indirect evidences exist (i.e. galaxies rotations, bullet clusters). Some dark matter candidates are considered to be either weakly interactive massive particles (WIMPs, mass around 100 GeV), or stable light particles (axions, mass around eV scale), or more exotic candidates (gravitinos, neutralinos, etc) [74]. However, none have been observed so far.

The second missing piece in the SM is the **dark energy** and the first evidence came from the supernovae studies based on the relation between the distance to an object and its redshift [75]. They were suggesting that the expansion of the universe is an indirect source of dark energy. It can be explained by a cosmological constant in the GR equation (Lambda - CDM model) [76–78] and interpreted as a measure of the energy density of the vacuum. However, it does not work properly with known results from particle physics : this is the so-called "cosmological constant problem" [79].

The **matter - antimatter asymmetry** is another unexplained feature in the Standard Model which predicts almost equal amounts of particles and antiparticles. If this would be true, no matter would exist as its constituents would annihilate in pairs. A small asymmetry is present in the SM in the quark sector, but it cannot explain the observed asymmetry [80].

In the SM, the **neutrinos masses** are considered to be zero. However in the lepton sector the neutrino oscillations imply the existence of a mass different of zero for all neutrinos [81, 82] (Table 1.1). Mathematically the mass can be explained by the seesaw mechanism [58], where a right-handed neutrino partner is added and considered to be sterile (i.e. not participating at any of the SM interactions). Its mass is not given by the Higgs mechanism and thought to be related to processes involving high energy scales which are expected to be suppressed at low (i.e. electroweak) energies. The corrections given to this suppression leads to the left-handed neutrino masses which are inverse proportional to the right-handed partner masses. Hence this theory, not confirmed yet, can explain the small associated masses of the left-handed neutrinos, given the existence of some very massive right-handed neutrinos. As the neutrinos are mixing, the mass terms depend on the neutrino flavors and they can be written as a PMNS[6] matrix, the equivalent of the CKM matrix. The matrix can be parametrized by three angles and one complex phase, which are not present in the SM [83, 84].

Beside the incoherence with the experimental results, several theoretical problems are present in the Standard Model. One of the most important is the **hierarchy problem** : since the Higgs boson couples to all massive particles, the quantum correction loops increase dramatically when going to the Planck scale. These corrections are related to virtual particles carrying a large transverse momentum (**k**, Figure 1.3). Hence, an infinite number of diagrams should be added in the Higgs mass term :

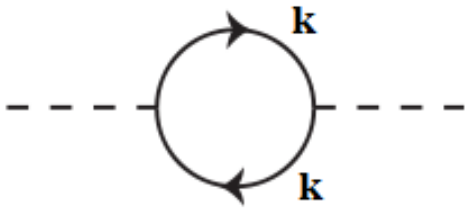


Figure 1.3: One loop fermionic diagram of the first-order quantum correction to the Higgs boson mass.

$$m^2(\Lambda) = m^2(\mu) + \delta m^2 = (125)^2 + \delta m^2$$
$$\delta m^2 \sim \sum (-1)^{2s} m_{f,b}^2 \Lambda^2 \tag{1.33}$$

where s is the spin index, $m_{f,b}$ the boson or fermion mass and Λ is the UV cutoff energy. Therefore a high fine-tunning ($\sim 10^{17}$) between $m^2\Lambda^2$ and δm^2 is required to accommodate the 125 GeV Higgs boson mass, making the SM un-natural at the Planck scale. From Equation (1.33) one can see that the highest contribution is brought by the heavy fermions or bosons (t, W, H and Z). This problem applies only to spin-0 bosons because the fermion masses are protected by the chiral symmetry and the spin-1 bosons are protected by the gauge invariance.

[6] Pontecorvo–Maki–Nakagawa–Sakata

Another important theoretical problem is the non-violation of the **CP symmetry** in the strong interactions. The QCD Lagrangian is globally gauge invariant under a $\mathrm{U(N)}_L \times \mathrm{U(N)}_R$ chiral transformation with N number of fermions and it can be decomposed in the corresponding vectorial (V) and axial (A) symmetries, $\mathrm{U(1)}_V \times \mathrm{SU(N)}_V \times \mathrm{U(1)}_A \times \mathrm{SU(N)}_A$. The conservation of the baryon number is an exact symmetry under the $\mathrm{U(1)}_V$ subgroup, regardless of the quark masses, and an exact isospin symmetry can be obtained under the $\mathrm{SU(2)}_V$ subgroup (hadron organization in multiplets). The axial components have no such corresponding symmetries. The $\mathrm{SU(2)}_A$ global symmetry is spontaneously broken by a $\langle \bar{q}_R q_L \rangle$ quark condensate and the associated Goldstone bosons are the $\pi^{0,\pm}$ mesons, with a vanishing mass in the limit of massless up and down quarks. If a $\mathrm{U(1)}_A$ symmetry is considered, the $q \to q' = e^{i\theta} q$ transformation is not conserved because of the presence of an explicit quantum anomaly. If this anomaly is combined with the corresponding CP violation in the weak interactions, the QCD Lagrangian contains a term which naturally violates the CP symmetry :

$$\mathcal{L} \sim \theta \cdot G_a^{\mu\nu} G_a^{\nu\mu} \tag{1.34}$$

It implies a very high neutron electric dipole moment, if the θ parameter is not set to be very small ($\theta < 10^{-9}$). The reason for this value being so small is an open question in the Standard Model. A possible solution is brought by Peccei and Quinn [85,86] : after considering the entire QCD Lagrangian invariant under an additional global chiral symmetry, the θ parameter can be set to 0 through a given chiral transformation. Same results could be obtained also under SSB. The resulted Goldstone boson is the axion.

In the Standard Model, the **gauge coupling constants** (electromagnetic, weak and strong) do not unify at a given energy when running the coupling constants to high energies (Figure 1.4, dotted black lines).

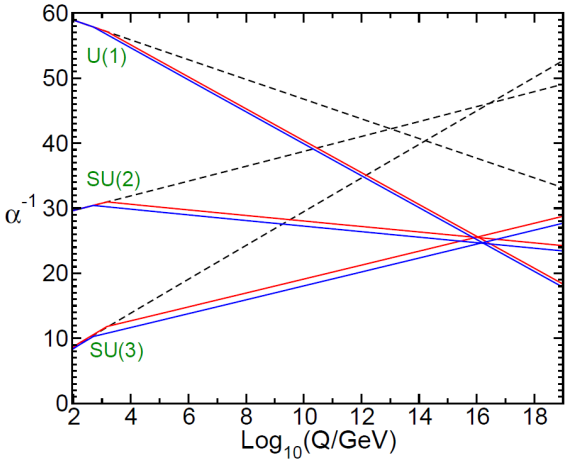


Figure 1.4: Running of the inverse gauge coupling constants. The dotted lines reflects the distribution in the Standard Model and the colored continuous lines in one SUSY model.

The list with Standard Model problems can continue with the required number of generations or vector bosons, etc. Several BSM theories bring robust solutions to these issues. One of them is Supersymmetry, a theory which groups new scalars with the SM fermions and new fermions with the SM bosons (Chapter 2).

Chapter 2

Supersymmetry

From all the existing theories able to solve the Standard Model problems, the Supersymmetry, or simply SUSY, occupies a central place even if it is not (yet) discovered after analyzing the entire $\sqrt{s} = 8$ TeV data. An important reason for its important place is that the hierarchy problem is solved by adding superpartners to the SM particles. Moreover, at energies close to the Planck scale, the weak, strong and electromagnetic interactions precisely unify and a dark matter candidate is present in the list of SUSY particles. This chapter is dedicated to a general presentation of the SUSY theory and its phenomenological constraints. It starts with a short history in Section 2.1 and further presents the mathematical description of the theory in Section 2.3. SUSY phenomenology is shortly described in Section 2.4 and the status of SUSY after the LHC Run-1 is mentioned in Section 2.5. Section 2.6 closes this chapter.

2.1 History

Supersymmetry was first mentioned in 1966 by H. Miyazawa as a symmetry connecting the mesons and baryons after extending the SU(6) symmetry group to SU(6, 21) supergroup[1] [87,88]. It was also the first time when a classification of the superalgebras is given. The results were largely ignored and Supersymmetry reappeared in '70 in two dimensional models, in the context of the string theory [89–91]. SUSY, as a graded Lie algebra, is proposed by Goland, Likthman and others in 1971 after extending the Poincare group to a super-Poincare group [92,93]. This was performed with the introduction of four anticommuting generators which are later identified with the supercharges. In 1974, SUSY was written as a four dimensional quantum field theory (QFT) by Wess and Zumino [94,95], leading to its application in the particle physics field.

The major difference between SUSY and the Standard Model as QFTs is the absence of any discrimination between fermions and bosons in the former : they are grouped in the same supermultiplet, combining their spin and internal degrees of freedom. This is possible because the SUSY theory presents a superalgebra and is defined in terms of superfields and supermultiplets. Note that the Coleman-Mandula theorem [96] – which states that the space-time (spin, mass, momentum) and internal symmetries (quantum numbers associated to the gauge symmetries like electric charge, color, weak isospin) cannot be combined – does not apply here. This was shown by Haag, Lopuszanski and Sohnius in 1975 [97]. Generally, by allowing both commuting (bosons) and anticommuting (fermions) symmetry generators the space-time - internal symmetries are mixed and Supersymmetry becomes the unique possible extension of the known space-time symmetries. The combination is performed using a Lie superalgebra and it requires the introduction of the $\mathbb{Z}_2$ grading and it includes transformations that change one type of particle into another one by incrementing its spin with half-unit. Now, every fermion is associated to a superpartner boson and each boson to a superpartner fermion.

As a consequence, in the minimal extension of the Standard Model (MSSM) the number of particles present in the theory doubled [98], as illustrated in Table 2.1. The simplest example of a supermultiplet is formed by a Standard Model fermion and its scalar superpartner, $\{e_L, \tilde{e}_L\}$, while the gauge supermultiplets are built from SM gauge (vector) bosons and their gauginos partners. One important remark is the existence of two Higgs boson supermultiplets, which is mandatory to give mass to all quarks : after the symmetry breaking, only a Higgs boson with $Y = -\frac{1}{2}$ (H_d) can give mass to the U-type quarks and one with $Y = +\frac{1}{2}$ (H_u) to the D-type quarks. The charginos $(\tilde{\chi}^{\pm})$ and neutralinos $(\tilde{\chi}^0)$ are ordered mass eigenstates formed by superposition of the SUSY partners of the SM Higgs and EW gauge bosons.

Given the experimental results, no SUSY particles were discovered yet. This strongly points to a theory with a spontaneously broken supersymmetry and much heavier sparticles than their SM partners. The

[1] Six is the number of mesons and 21 the number of baryons.

<table>
<tr><td></td><td></td><td colspan="2">Proper states in</td><td rowspan="2">Spartners</td><td colspan="2">Proper states in</td></tr>
<tr><td></td><td></td><td>interaction term</td><td>mass term</td><td></td><td>interaction term</td><td>mass term</td></tr>
<tr><td rowspan="2">$\begin{pmatrix}\nu\\L\end{pmatrix}$, l</td><td rowspan="2">Leptons
$S=1/2$</td><td colspan="2" rowspan="2">$\begin{pmatrix}\nu_e\\e_L\end{pmatrix}, e_R$
$\begin{pmatrix}\nu_\mu\\\mu_L\end{pmatrix}, \mu_R$
$\begin{pmatrix}\nu_\tau\\\tau_L\end{pmatrix}, \tau_R$</td><td rowspan="2">Sleptons
$S=0$</td><td colspan="2">$\begin{pmatrix}\widetilde{\nu}_e\\\widetilde{e}_L\end{pmatrix}, \widetilde{e}_R$
$\begin{pmatrix}\widetilde{\nu}_\mu\\\widetilde{\mu}_L\end{pmatrix}, \widetilde{\mu}_R$</td></tr>
<tr><td>$\begin{pmatrix}\widetilde{\nu}_\tau\\\widetilde{\tau}_L\end{pmatrix}, \widetilde{\tau}_R$</td><td>$\widetilde{\tau}_1, \widetilde{\tau}_2, \widetilde{\nu}_\tau$</td></tr>
<tr><td rowspan="2">$\begin{pmatrix}U\\D\end{pmatrix}$, u, d</td><td rowspan="2">Quarks
$S=1/2$</td><td colspan="2" rowspan="2">$\begin{pmatrix}u_L\\d_L\end{pmatrix}, u_R, d_R$
$\begin{pmatrix}c_L\\s_L\end{pmatrix}, c_R, s_R$
$\begin{pmatrix}t_L\\b_L\end{pmatrix}, t_R, b_R$</td><td rowspan="2">Squarks
$S=0$</td><td colspan="2">$\begin{pmatrix}\widetilde{u}_L\\\widetilde{d}_L\end{pmatrix}, \widetilde{u}_R, \widetilde{d}_R$
$\begin{pmatrix}\widetilde{c}_L\\\widetilde{s}_L\end{pmatrix}, \widetilde{c}_R, \widetilde{s}_R$</td></tr>
<tr><td>$\begin{pmatrix}\widetilde{t}_L\\\widetilde{b}_L\end{pmatrix}, \widetilde{t}_R, \widetilde{b}_R$</td><td>$\widetilde{t}_1, \widetilde{t}_2, \widetilde{b}_1, \widetilde{b}_2$</td></tr>
<tr><td rowspan="2">H_u, H_d</td><td>Gauge
Bosons
$S=1$</td><td>$W^\pm, W^0, B, g$</td><td>$W^\pm, Z^0, \gamma, g$</td><td>Gauginos
$S=1/2$</td><td>$\widetilde{W}^\pm, \widetilde{W}^0, \widetilde{B}, \widetilde{g}$</td><td rowspan="2">Gluino $\widetilde{g}$
Neutralinos
$\widetilde{\chi}_1^0, \widetilde{\chi}_2^0, \widetilde{\chi}_3^0, \widetilde{\chi}_4^0$
Charginos
$\widetilde{\chi}_1^\pm, \widetilde{\chi}_2^\pm$</td></tr>
<tr><td>Higgs
Boson
$S=0$</td><td>$\begin{pmatrix}H_u^+\\H_u^0\end{pmatrix}, \begin{pmatrix}H_d^0\\H_d^-\end{pmatrix}$</td><td>$h^0, H^0, A^0, H^\pm$</td><td>Higgsinos
$S=1/2$</td><td>$\begin{pmatrix}\widetilde{H}_u^+\\\widetilde{H}_u^0\end{pmatrix}, \begin{pmatrix}\widetilde{H}_d^0\\\widetilde{H}_d^-\end{pmatrix}$</td></tr>
<tr><td>G</td><td>Graviton
$S=2$</td><td colspan="2">G</td><td>Gravitino
$S=\frac{3}{2}$</td><td colspan="2">$\tilde{G}$</td></tr>
</table>

Table 2.1: Superpartners of the Standard Model particles, organized as $\mathrm{SU}(2)_L$ multiplets. The spin-0 fields are complex scalars and the spin 1/2 fields are left-handed two-component Weyl spinor.

mechanism breaking the symmetry is still unknown and it cannot be performed with any of the particles presented in the Table 2.1. Therefore, a new sector is responsible for the breaking. To avoid a high fine-tunning (as in the Standard Model) the SUSY particle masses should not exceed the TeV scale.

2.2 Motivation for SUSY

It was already mentioned in Chapter 1 that the hierarchy problem can be solved by Supersymmetry. Indeed, **both** the quadratic corrections (due to the boson self-interactions) and the logarithmic corrections (due to fermion loop interaction) can be canceled [94]. The Higgs mass, after considering the loop corrections for a Higgs-SM fermion interaction, can be written as :

$$\text{SM} : m_H^2 = (m_H^2)_0 \qquad -n_c\frac{\lambda_f^2}{8\pi^2}\Lambda^2 + n_c\frac{\lambda_f^2}{8\pi^2}m_f^2\ln\frac{\Lambda^2}{m_f^2} + ... \tag{2.1}$$

$$\text{adding the SUSY corrections :} \qquad +n_c\frac{2\lambda_s}{16\pi^2}\Lambda^2 - n_c\frac{\lambda_s}{16\pi^2}\left[m_{sL}^2\ln\frac{\Lambda^2}{m_{sL}^2} + m_{sR}^2\ln\frac{\Lambda^2}{m_{sR}^2}\right]... \tag{2.2}$$

where m represents the particle mass, n_c accounts for the number of colors (3 for quarks, 1 for scalars), λ for the couplings and Λ for the cut-off scale. Here $\lambda_s = \lambda_f^2$. An important observation is that the opposite sign between the SM and SUSY corrections. Note that the cancellation of the quadratic divergence occurs independently on the fermion-scalar mass term, as these masses are not present in the Higgs mass expression. This is also the reason why the (soft) SUSY breaking terms can be introduced in the Lagrangian *by hand* without destroying the exact cancellation. The logarithmic corrections are sensitive to the masses of the heaviest fermion particles (i.e. the top quark) coupling with the Higgs boson. Generally, it will cancel only if the new SUSY partners have masses close to the SM particles. Therefore, the SUSY theory can solve also the SM fine-tunning problem **only** if the SUSY partners appear below or close to the TeV scale. Considering this hypothesis, SUSY can be discovered at LHC and the ATLAS BSM physics program is well motivated.

The other reasons to favor SUSY are :

- Predicts a light Higgs boson ($m_H < 130$ GeV), in accordance with the ATLAS and CMS observation.

- Its algebra contains a space translation generators as needed in the quantum gravity.
- It can provide a dark matter candidate, a weakly interacting particle not present in the Standard Model : the lightest neutralino ($\tilde{\chi}_1^0$) or the gravitino.

2.3 Theoretical framework

A short introduction of the SUSY theoretical framework [99–105] is presented in this section. Firstly, the SUSY algebra is presented and the superspace and superfield notions are introduced in Section 2.3.1. The section continues with the superfield strength and the gauge transformations relations. Finally, the MSSM Lagrangian and the superpotentials are presented and the section closes with the R parity law.

2.3.1 The SUSY algebra and gauge transformations

Generally, a SUSY transformation [100, 101] can be regarded as $\phi \to \phi + \xi\psi$, where ϕ stands for a bosonic field and ψ for a fermionic field[2]. It changes a boson into a fermion and vice versa and it is performed with a combination of two fermionic (spin 1/2) operators, the spinorial charge (supercharge) and the anti-charge (Q_α and $\bar{Q}_\alpha = Q_\alpha^\dagger \gamma^0$). Generally :

$$Q_\alpha |j> = |j \pm \frac{1}{2}> \tag{2.3}$$

where $\alpha = \overline{\{1,2\}}$ denotes the spinor component. Note that these generators, which transform as spinors under the Lorentz transformations, are needed to combine the space-time and internal symmetries (to *extend* the Coleman-Mandula theorem). Beside the usual conservation relations involving the 4-momentum (P_μ) and the angular momentum, the SUSY superalgebra includes new commuting and anticommuting relations :

$$\begin{aligned} \{Q_\alpha, Q_{\dot{\alpha}}^\dagger\} = 2\sigma^\mu_{\alpha\dot{\alpha}} P_\mu \quad ; \quad [Q_\alpha, P_\mu] = [Q_\alpha^\dagger, P_\mu] = 0 \\ \{Q_\alpha, Q_{\dot{\alpha}}\} = 0 \quad ; \quad \{Q_\alpha^\dagger, Q_{\dot{\alpha}}^\dagger\} = 0 \end{aligned} \tag{2.4}$$

where $(Q_\alpha)^* = \bar{Q}_{\dot{\alpha}}$. The first relation is fundamental in SUSY and points to a space-time translation generated by two SUSY consecutive transformations. The second relation is obvious because SUSY is an *internal* symmetry[3], therefore independent of the space-time position.

Superspaces and superfields

In SUSY, the 4D Minkowski space is extended to include also the spin directions – definition of a superspace. More precisely, beside the time-space coordinates x^μ the anti-commuting Grassmann two-component Weyl spinors θ_α and $\bar{\theta}_{\dot{\alpha}} = \theta_\alpha^\dagger$ are also included. Knowing that :

$$\{\theta_\alpha, \theta_{\dot{\alpha}}\} = 0 \quad ; \quad \theta\theta = \theta\theta\theta = ... = 0 \quad ; \quad \{\theta Q, \bar{\theta}\bar{Q}\} = 2\theta\bar{\theta}\sigma^\mu P_\mu \tag{2.5}$$

a general $N = 1$ SUSY superfield[4] can be constructed as a Taylor expansion :

$$\begin{aligned} S(x^\mu, \theta_\alpha, \bar{\theta}_{\dot{\alpha}}) \quad &= f(x) + \theta^\alpha \phi_\alpha(x) + \bar{\theta}^{\dot{\alpha}} \bar{\chi}_{\dot{\alpha}}(x) + \theta^\alpha \theta_\alpha m(x) + \theta^{\dot{\alpha}} \theta_{\dot{\alpha}} n(x) + \\ &\bar{\theta}^{\dot{\alpha}} (\sigma^\mu \theta)_{\dot{\alpha}} A_\mu(x) + \theta^\alpha \theta_\alpha \bar{\theta}^{\dot{\alpha}} \bar{\lambda}_{\dot{\alpha}}(x) + \bar{\theta}^{\dot{\alpha}} \bar{\theta}_{\dot{\alpha}} \theta^\alpha \psi_\alpha(x) + \theta^\alpha \theta_\alpha \bar{\theta}^{\dot{\beta}} \bar{\theta}_{\dot{\beta}} d(x) \end{aligned} \tag{2.6}$$

where the components f, m, n and d are complex scalar fields, A_μ is a complex vector field and ϕ, $\bar{\chi}$, $\bar{\lambda}$ and ψ are fermionic fields (Weyl spinors). The latter are anticommuting with each other and with the superspace spinorial coordinates. To construct the SUSY irreducible representations, a certain set of supercovariant constraints are imposed. Note that in MSSM, the SUSY transformations do not change the $SU(3)_L$, $SU(2)_L$ or $U(1)$ quantum numbers. Therefore, each SM field and its SUSY partner in a supermultiplet must preserve them.

The superspace (SUSY infinitesimal) global transformations are performed using the supertranslation and translation generators acting on the superspace :

$$Q_\alpha = -i\frac{\partial}{\partial\theta^\alpha} - (\sigma^\mu_\alpha \bar{\theta})^{\dot{\alpha}} \partial_\mu \quad ; \quad \bar{Q}_{\dot{\alpha}} = i\frac{\partial}{\partial\bar{\theta}^{\dot{\alpha}}} + (\theta\sigma^\mu_{\dot{\alpha}})_\alpha \partial_\mu \quad ; \quad Q_\alpha^\dagger = \bar{Q}_{\dot{\alpha}} \quad ; \quad P_\mu = i\partial_\mu \tag{2.7}$$

[2] Here, the ξ variable is just an infinitesimal parameter and it has to be a spinor.

[3] SUSY is not really an internal symmetry as the generators are spinors and not scalars. It can be rather regarded as an extension of the Poincare space-time symmetries.

[4] N=1 : only one spinor generator, Q_α and its hermitic conjugate are considered in the theory.

where the differential generators satisfy the SUSY algebra. To construct the SUSY Lagrangian and impose constraints on the general superfield, the superspace covariant derivatives must be also included. This is mainly because the simple derivative is not supersymmetric covariant. The relations are :

$$D_\alpha = \frac{\partial}{\partial\theta^\alpha} - i\sigma^\mu_{\alpha\dot\alpha}\bar\theta^{\dot\alpha}\partial_\mu \quad ; \quad D^\dagger_{\dot\alpha} = -\frac{\partial}{\partial\bar\theta^{\dot\alpha}} + i\theta^\alpha\sigma^\mu_{\alpha\dot\alpha}\partial_\mu \quad ; \quad D^\dagger_\alpha = \bar D_{\dot\alpha} \quad ; \quad P_\mu = i\partial_\mu \tag{2.8}$$

The covariant derivatives anticommute with the supercharges :

$$\{Q_\alpha, D_\beta\} = 0 \quad ; \quad \{Q^\dagger_\alpha, D_\beta\} = 0 \quad ; \quad \{Q_\alpha, D^\dagger_\beta\} = 0 \quad ; \quad \{Q^\dagger_\alpha, D^\dagger_\beta\} = 0 \tag{2.9}$$

Therefore, the subspace of superfields satisfying the $\boldsymbol{\bar D_{\dot\alpha} S = 0}$ relation is stable under supersymmetric transformations (in fact it is an irreducible representation). Such a superfield is called **chiral superfield**, and depends on a complex scalar field ϕ, a complex Weyl spinor ψ^α and an auxiliary (non-propagating) scalar complex field F :

$$\Phi(x^\mu, \theta_\alpha, \bar\theta_{\dot\alpha} = 0) = \phi(y) + \sqrt{2}\theta^\alpha\psi_\alpha(y^\mu) + \theta^\alpha\theta_\alpha F(y^\mu) \tag{2.10}$$

$$\Phi(y^\mu, \theta_\alpha, \bar\theta_{\dot\alpha}) = e^{-i\partial_\mu\theta\sigma^\mu\bar\theta}\Phi(y^\mu, \theta_\alpha, \bar\theta_{\dot\alpha} = 0) \tag{2.11}$$

where $y^\mu = x^\mu + i\bar\theta\bar\sigma^\mu\theta$. The auxiliary scalar field F has a mass dimension of two as needed to supplement the two degrees of freedom in Φ – note that Ψ has a mass dimension of four – and ensures the invariance of the SUSY transformations for each field in a multiplet. The left- and right-handed SM fermions are described by two chiral superfields and are associated to two spin-0 bosons. Under an infinitesimal $N = 1$ SUSY transformation $\varepsilon Q + \bar\varepsilon\bar Q$ the components of the superfield transform as :

$$\partial_\varepsilon\phi = \varepsilon^\alpha\psi_\alpha \quad ; \quad \partial_\varepsilon\psi_\alpha = i(\sigma^\mu\bar\varepsilon)_\alpha\partial_\mu\phi - \varepsilon_\alpha F \quad ; \quad \partial_\varepsilon F = i(\sigma^\mu\partial_\mu\psi)_{\dot\alpha}\bar\varepsilon^{\dot\alpha} \tag{2.12}$$

where ε is a constant Grassmann complex Weyl spinor. This relates naturally the spinorial and scalar fields. Note that any chiral superfield can be written as a double covariant derivative of a general superfield, $\Phi = \bar D_{\dot\alpha}\bar D^{\dot\alpha} S$, and the products of two chiral superfields is another chiral superfield.

Together with the chiral superfields, additional **vector (or real) superfields** (V) are present in the theory. Each vector field must satisfy the $\boldsymbol{V = V^*}$ relation and under the Wess-Zumino (WZ) gauge (see next section) its expression can be written as :

$$V_{WZ}(x^\mu, \theta_\alpha, \bar\theta_{\dot\alpha}) = -\theta^\alpha(\sigma^\nu_\alpha\bar\theta)_{\dot\alpha}A_\nu(x^\mu) + i\theta^\alpha\theta_\alpha\bar\theta^{\dot\alpha}\bar\lambda_{\dot\alpha} - i\bar\theta^{\dot\alpha}\bar\theta_{\dot\alpha}\theta^\alpha\lambda_\alpha + \theta^\alpha\theta_\alpha\bar\theta^{\dot\alpha}\bar\theta_{\dot\alpha}D \tag{2.13}$$

It contains one massless real vector field A_μ (gauge boson), its fermionic partner λ_α (gauginos) and one gauge auxiliary (non-propagating) real scalar field D. The latter has one degree of freedom, needed to supplement the three degrees of freedom associated to A_μ, and a mass dimension of two. As the F term in the chiral superfield, the D term is needed to ensure the invariance of the SUSY transformations. The Standard Model gauge bosons are introduced by vector superfields, which are associated to gauginos (Majorana fermions). Generally, under an infinitesimal SUSY transformation, the superfield components transform as :

$$\partial_\varepsilon A^\alpha_\mu = -\bar\varepsilon^{\dot\alpha}\bar\sigma_{\mu\dot\alpha}\lambda^{\dot\alpha} - \bar\lambda^\alpha\bar\sigma_{\mu\alpha}\varepsilon^\alpha \quad ; \quad \partial_\varepsilon\lambda_\alpha = \frac{i}{2}\sigma^\mu_\alpha\bar\sigma^\nu_\alpha(\partial_\mu A^\alpha_\nu - \partial_\nu A^\alpha_\mu) + \varepsilon_\alpha D \quad ; \quad \partial_\varepsilon D = i\varepsilon^\alpha\sigma^\mu_\alpha(\partial_\mu\bar\lambda)_\alpha - i\bar\varepsilon^\alpha\bar\sigma^\mu_\alpha(\partial_\mu\lambda)_{\dot\alpha} \tag{2.14}$$

However, in the WZ gauge, A_μ transforms as in Equation (1.2) and λ_α and D are invariant. Note that any of the relations $\Phi\Phi^*$, $\Phi + \Phi^*$ and $i(\Phi - \Phi^*)$, involving chiral superfields, is a vectorial superfield.

Superfield strength

The superfield strength – similarly to the SM $F_{\mu\nu}$ – is a chiral superfield derived from the vectorial superfield. It is noted with $\mathcal{W}$ and is gauge-invariant by construction :

$$\mathcal{W}_\alpha = -\frac{1}{4}\bar D_{\dot\alpha}\bar D^{\dot\alpha}D_\alpha V \quad ; \quad \bar{\mathcal{W}}_\alpha = -\frac{1}{4}D_\alpha D^\alpha\bar D_{\dot\alpha}V \tag{2.15}$$

When expanded in terms of component superfields it can be written as :

$$\mathcal{W}^\alpha(x, \theta, \bar\theta = 0) = -i\lambda^\alpha + (\theta\sigma_{\mu\nu})^\alpha(\partial_\mu A^\alpha_\nu - \partial_\nu A^\alpha_\mu) + \theta^\alpha D - \theta^\alpha\theta_\alpha(\bar\sigma^\mu D_\mu\bar\lambda)^\alpha \tag{2.16}$$

Gauge transformations

The SM gauge transformations cannot be applied to superfields because of the dependency on the fermionic directions. Therefore, a gauge transformation is performed by adding the gauge parameter to a chiral superfield Λ. In this way the transformations are generalized and include both *pure* gauge and SUSY transformations. For the abelian case (invariance for the U(1) case), the transformations are :

$$\partial\Phi = \mathrm{e}^{2igq\Lambda}\,\Phi \quad ; \quad \partial\Phi^* = \mathrm{e}^{-2igq\Lambda}\,\Phi^* \quad ; \quad \partial V = V + i(\Lambda^* - \Lambda) \tag{2.17}$$

where $\Lambda = \alpha(x) + \theta\xi + \theta^2 A$ (α is a complex parameter), g is the coupling constant and q_i are the U(1) charges. They are also called "supergauge transformations" or "gauge transformations" with a supersymmetric form. Note that such a transformation is applied to obtain the V_{WZ} superfield in Equation 2.13.

For non-abelian Lie superalgebras, the gauge transformations are performed using the T_a generators and g_a coupling constants. Compared to the previous case, it is necessary to introduce a vectorial superfield V_a and a gauge chiral superfield Λ_a. The general gauge transformations are :

$$\begin{gathered}\partial\Phi_i = (e^{i\Lambda})_i^j\Phi_j \quad ; \quad \partial\Phi^{*i} = \Phi^{*j}(e^{-i\Lambda^\dagger})_j^i \quad ; \quad \partial e^V = e^{i\Lambda^\dagger}e^V e^{-i\Lambda} \\ \text{given : } V = \sum_a 2g_a T^a V^a \quad ; \quad \Lambda = \sum_a 2g_a T^a \Lambda^a\end{gathered} \tag{2.18}$$

Given the gauge transformations defined in Equation (2.19), the field strength of the vector superfield from Equation (2.15) is rewritten as :

$$\mathcal{W} = -\frac{1}{4}D^\dagger D^\dagger(e^{-V}De^V) \quad \text{and transforms as : } \partial\mathcal{W} = e^{i\Lambda}\mathcal{W}_a e^{-i\Lambda} \tag{2.19}$$

It implies that the trace over $(\mathcal{W}^a\mathcal{W}_a)$ is a chiral superfield invariant under the supergauge transformations - similarly to the last term from Equation (1.4) standing from non-supersymmetric Yang-Mills theory.

2.3.2 The MSSM Lagrangian

The MSSM Lagrangian is formed by chiral superfields describing the fermions (Equation (2.11)) and vector superfields describing the bosons (Equation (2.13)). The gauge group is the same as in the Standard Model. It was already pointed out that each component of the superfield depends on the fermionic degrees of freedom. This is not the case for the Lagrangian and the action $\mathcal{S}$ (the integral of the Lagrangian over the superspace). Therefore, in an intermediate step the dependency on the fermionic coordinates *is removed*. This is straightforward due to the properties of the Grassmann integration, which allow the selection of the desired components only :

$$\Phi|_{F+F*} = \int d^2\theta d^2\bar\theta[\delta^{(2)}(\bar\theta)\Phi + \delta^{(2)}(\theta)\Phi] = F + F^* \tag{2.20}$$

$$V|_D = \int d^2\theta d^2\bar\theta V = \frac{1}{2}D + \frac{1}{4}\partial^\mu\partial_\mu f \tag{2.21}$$

which are invariant under the SUSY transformations. Also, the action written as $\mathcal{S} = \int d^2x \int d^2\theta \int d^2\bar\theta$ is invariant under any SUSY transformation. This is mainly due to the supercharge definition, which leads to $(\varepsilon Q + \bar\varepsilon\bar Q)\mathrm{S} = 0$ before integration. The massless free Lagrangian density describing a chiral supermultiplet, under the WZ gauge, can be written as :

$$\mathcal{L}' = \Phi^*\Phi|_D = \int d2\theta d^2\theta^\dagger\Phi^*\Phi = \partial_\mu\phi_i^\dagger\partial^\mu\phi_i + i\psi_i^\dagger\bar\sigma^\mu\partial_\mu\psi_i + F_i^\dagger F_i \tag{2.22}$$

where the summing index i accounts for all internal degrees of freedom and the spinorial field index i for the spinor component respectively. The interaction and the mass terms in the Lagrangian can be easily introduced, given the fact that any multiplication of some superfields is also a superfield :

$$\mathcal{L} = \Phi^{*i}\Phi_i|_D + [W(\Phi_i)|_F + cc] = \Phi^{*i}\Phi_i|_D + \Big[\underbrace{\frac{1}{2}M^{ij}\Phi_i\Phi_j}_{\text{mass term}} + \underbrace{\frac{1}{6}y^{ijk}\Phi_i\Phi_j\Phi_k}_{\text{interaction}} + c.c\Big] \tag{2.23}$$

The $W(\Phi_i)$ variable denotes the superpotential, a (any) homomorphic function[5] depending only on chiral superfields. It contains the mass and all the allowed interactions. M^{ij} has the significance of mass and is symmetric in i and j. Here y^{ijk} are the Yukawa couplings, as well symmetric in i, j and k.

[5]Note that the homomorphic functions preserve the operations between two groups (or two algebraic structure).

Given the component-field evolution under a SUSY transformation, the $\mathcal{L}$ Lagrangian is invariant. However, it is not invariant under gauge transformations (Equation (2.19)). To ensure this invariance, the Lagrangian should be rewritten as :

$$\mathcal{L} = [\Phi^{*i}(e^V)_i^j \Phi_j]|_D + \frac{1}{4}[\mathrm{Tr}(\mathcal{W}^a \mathcal{W}_a) + hc]_F + \frac{1}{4}[W(\Phi_i) + hc]_F \tag{2.24}$$

where

$$[\mathcal{W}^\alpha \mathcal{W}_\alpha]_F \sim D^2 + 2i\lambda\sigma^\mu \partial_\mu \lambda^\dagger - \frac{1}{2}F^{\mu\nu}F_{\mu\nu} \tag{2.25}$$

This term is similar to $F^{\mu\nu}F_{\mu\nu}$ kinetic term from the Standard Model and is present in any Yang-Mills gauge theory.

Equation (2.24) stands for general renormalizable Lagrangian for a supersymmetric gauge theory. This Lagrangian and, implicitly, the action are invariant under both SUSY and gauge transformations. Moreover, due to the renormalizability condition the dimension of all terms in the Lagrangian is at most 4. The auxiliary fields D and F appearing in the gauge and chiral superfields are eliminated after requiring $\partial\mathcal{L}/\partial F_i = \partial\mathcal{L}/\partial D_i = 0$. The gauge bosons are included in the covariant derivatives and in the $F_{\mu\nu}$ term, while their gaugino partners are described by the λ_a fermionic fields. Note that each massless gauge bosons is accompanied by a Weyl spinor in the supermultiplet. Finally, the Higgs bosons and their superpartners are present as chiral fields and the only difference between them is that the scalar components acquire a vacuum expectation value (VEV).

The MSSM superpotential

Given the general form presented in Equation (2.23), the MSSM superpotential [100] can be written as :

$$W_{MSSM} = uY_u(UH_u^0 - DH_u^+) - dY_d(UH_d^- - DH_d^0) - lY_l(\nu H_d^- - LH_d^0) + \mu H_u H_d \tag{2.26}$$

where Y_u, Y_d and Y_l terms are the Yukawa couplings (3×3 matrices) which gives the quark and lepton masses and the CKM mixing angles. The μ term is the Higgsino mass, similarly to the Higgs mass in the SM. Note that both H_u and H_d (with opposite hypercharges) are needed to give mass to all up and down quarks and leptons (no $dY_d DH_u^*$, ... terms are allowed [100]).

The couplings in SUSY implies the squark-Higgsino-quark and slepton-Higgsino-lepton interactions, beside the SM Higgs-quark-quark and Higgs-lepton-lepton ones. Some diagrams involving the top Yukawa coupling are illustrated in Figure 2.1. Note that they should have the same coupling strength y_t, while the scalar quadratic interactions are characterized by y_t^2 strength.

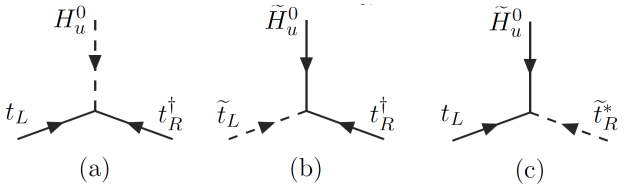


Figure 2.1: Top Yukawa coupling in the SM (a) and in SUSY (b,c) [100].

The scalar potential associated to the Higgs boson squared-mass results from the superpotential and has the expression :

$$V = |\mu|^2(|H_u^0|^2 + |H_u^+|^2 + |H_d^0|^2 + |H_d^-|^2) \tag{2.27}$$

Similarly, for the Higgsino mass term it is :

$$|\mu|(\widetilde{H}_u^+ \widetilde{H}_d^- - \widetilde{H}_u^0 \widetilde{H}_d^0) \tag{2.28}$$

Equation (2.27) has a minimum at $H_u^0 = H_d^0 = 0$. Therefore, the EW symmetry breaking cannot appear and some negative soft symmetry breaking terms for the Higgs scalar should be introduced in MSSM. The mass of this soft term must be around μ^2 and much smaller than the Planck scale to correctly cancel the divergences. This introduces another fine-tunning, known as *the μ problem.*

The R parity

The superpotential presented in Equation (2.26) can be generalized by adding new couplings :

$$W_{\Delta\mathrm{L},\Delta\mathrm{B}} = \underbrace{\frac{1}{2}\lambda_{ijk}(\nu^i L^j - L^i \nu^j)l^k + \lambda'_{ijk}(\nu^i D^j - L^i U^j)d^k + \mu'(\nu H_u^0 - LH_u^+)}_{\Delta L=1} + \underbrace{\frac{1}{2}\lambda''_{ijk}u^i d^j d^k}_{\Delta B=1} \tag{2.29}$$

where i = 1,2,3 is the generation index, B is the baryon number and L the lepton number[6]

[6] B = $+(-)\frac{1}{3}$ for the quark$_{L(R)}$ supermultiplets and zero otherwise. L = +(-)1 for the lepton$_{L(R)}$ supermultiplets and zero otherwise.

It can be seen that Equation (2.29) is a (gauge invariant) term violating the total lepton number (λ_{ijk}, λ'_{ijk}, μ_i) and the total baryon number (λ''_{ijk}). Such processes are not observed (yet) experimentally but several constraints exist. The most important arise from the proton decay non-conservation : if the λ' and λ'' couplings are not suppressed, the proton lifetime is very short (Figure 2.2). However its lifetime is estimated to be greater than 10^{34} years. This can be avoided if the couplings are very small, or if the sfermions are very heavy. This can be achieved by introducing a new symmetry, the matter parity :

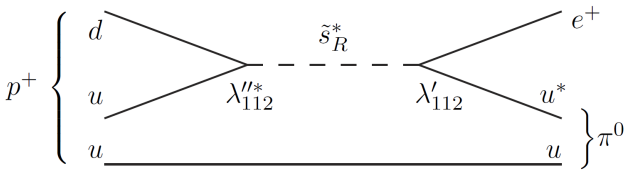


Figure 2.2: Illustrative Feynman diagram for the proton decay.

$$P_M = (-1)^{3(B-L)} \tag{2.30}$$

It is defined with respect to the baryon and total leptonic numbers and it is introduced for each particle present in the theory. Note that the quark / lepton supermultiplets have P_M = -1, while the Higgs supermultiplets, gauge bosons and gauginos are characterized by P_M = +1. More generally, only the terms satisfying the P_M = +1 condition are allowed in the Lagrangian, hence only the ones present in Equation (2.26). This symmetry is favored in the grand unification theories or in the cases when the $U(1)_{B-L}$ symmetry is spontaneously broken at a very high energy scale.

If the matter parity is combined with the angular momentum conservation :

$$P_R = (-1)^{3(B-L)+2s} \quad \text{where } s \text{ stands for the spin particle.} \tag{2.31}$$

This discrete symmetry is characterizing the particle interactions and it considered to be conserved in MSSM. All the SM particles have $P_R = 1$ and all the SUSY partners P_R = -1. Some important consequences are the existence of an even number of SUSY particles in the cascade decays, all λ, λ' and λ'' Yukawa couplings introduced in Equation (2.29) are set to zero, the lightest SUSY particle (LSP) must be stable and therefore present in all final states involving a sparticle decay.

2.3.3 SUSY breaking in the MSSM

In MSSM, the Standard Model hierarchy problem is solved before any symmetry breaking. Therefore the SUSY breaking can appear at energies around 1 TeV (*low energy* scenario), and can be performed by a term with a positive mass dimension (M_{SUSY}). It can be classified in :

- **Spontaneous SUSY breaking** : the Langrangian of the physical system is invariant under the SUSY transformations but the ground state is not like a field acquiring a VEV close to M_{SUSY} breaks the symmetry. To preserve the symmetry, the vacuum energy should be zero. This approach cannot be generalized for SU(N) group psymmetry.
- **Soft (explicit) SUSY breaking** : in the SUSY Lagrangian some (non-symmetric) breaking terms are introduced *by hand*.

Only the soft SUSY breaking mechanism is described and later used in the interpretation sections of the SUSY analysis chapters. In order to not spoil the UV properties of the SUSY theory, the breaking terms should have positive mass dimension (to not be important in the far UV region). The most general such terms are :

$$\mathcal{L}_{softB} = -\Big(\frac{1}{2}M_a\lambda^a\lambda^a + \frac{1}{6}a^{ijk}\phi_i\phi_j\phi_k + \frac{1}{2}b^{ij}\phi_i\phi_j\Big) + c.c. - (m^2)^i_j\phi^{*i}\phi_j \tag{2.32}$$

where M_a is the gaugino (a=1 gluino, a=2 wino and a=3 bino) masses, m^i_j the non-holomorphic scalar (sfermions) mass, a^{ijk} the trilinear scalar couplings – of form $(a_u)_{ij}\tilde{u}^j_R(\tilde{u}^i_L H^0_u - \tilde{d}^i_L H^+_u)$ – and b^{ij} the holomorphic biliniar scalar mass terms – of form $b(H^+_u H^-_d - H^0_d H^0_d)$. Some Feynman diagrams illustrating these terms are shown in Figure 2.3. The (soft) mass terms are around the TeV scale and the $\mathcal{L}_{soft}$ term is compatible with the general Lagrangian presented in Equation (2.24). It also preserve the gauge invariance and conserve the R parity. Moreover, the soft SUSY Lagrangian will not introduce any new divergences, in all the orders of the perturbation theory. The MSSM can be considered to be a supersymmetric Standard Model extended with soft SUSY breaking terms and its phenomenology is driven by the value of the μ term (Equation (2.26)).

With the soft symmetry breaking terms mentioned in Equation (2.32), 105 new parameters (masses, phases, mixing angles, etc) are introduced in the theory. Some of them present tight constraints, as they

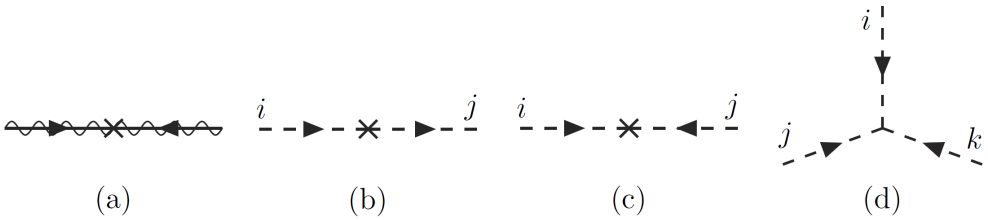


Figure 2.3: Soft SUSY breaking terms : M_a (a), $(m^2)^i_j$ (b), b^{ij} (c) and a^{ijk} (d).

contain problems like flavor changing neutral currents and CP violating processes well studied experimentally. For the non-diagonal squared slepton masses, where the sleptons mixing is present and the internal lepton numbers are not conserved anymore, the highest limits arise from the $\mu \to e\gamma$ BR measurements. Concerning the squark (squared) mass matrices, the tightest constraints are imposed by the neutral kaon and $D^0\bar{D}^0$ systems. Potential CP violation effects are bypassed if the soft parameters are not including new complex phases.

To reduce the number of parameters and bring solutions to the just mentioned problems, several hypotheses can be made - this approach is usually named the soft supersymmetry-breaking universality. One assumption is that the scalar and the corresponding Yukawa couplings are proportional ($a_{u,d,l} = A_{mu0}Y_{u,d,l}$), ensuring that only the third generation squarks and sleptons have large couplings. Moreover, the $(m^2)^i_j$ mass matrices are all flavor diagonal (flavor blind as a solution to the flavor problem[7]), and the third generation sfermions are mass degenerate. The last constraints are on the gaugino masses, and the A_{mu0} term is real. In this framework, the running coupling constants in MSSM do unify [106] and the MSSM is described only by fifteen parameters : four mass terms for the gauginos, four mass terms for sfermions, the third generation trilinear couplings and four[8] mass terms in the Higgs sector. This number increases to 19 for the phenomenological MSSM (PMSSM), as the constrain on the third generation sfermion mass degeneracy is relaxed (and hence 5 more parameters).

Going even further, the number of free parameters is drastically reduced in the gravity mediated SUSY breaking scenarios, where the supersymmetry breaking is performed in a hidden sector and transmitted using the gravitational interaction. In the minimal extension, mSUGRA/CMSSM [106–109], only five parameters appear : the common mass of the scalars at the grand unification (GUT) scale ($m_0^2 = m_{\tilde{Q}}^2 = m_{\tilde{U}}^2 = m_{\tilde{D}}^2 = m_{\tilde{L}}^2$), the common mass of the fermions at the GUT scale ($m_{1/2} = M_3 = M_2 = M_1$), the common trilinear coupling ($A_0 = A_u = A_d = A_l$), the ratio of the VEV of the two Higgs doublets ($\tan\beta$) and the sign of the higgsino mass parameter ($\mathrm{sign}(\mu)$). The GUT scale is considered to be 2×10^{16} GeV, where the mentioned universality conditions are defined. The soft parameters values are run down to the EW scale by the renormalization equations.

Origin of the supersymmetry breaking

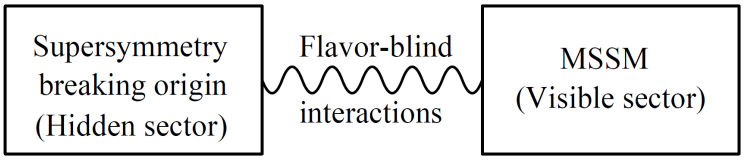


Figure 2.4: An illustration of the SUSY breaking in a Hidden sector (Q scale) and its transmission to the MSSM sector (TeV scale) through certain interactions (like gravity).

The most favorable hypothesis is a spontaneous SUSY breaking in a hidden sector of particles (Q scale) which are not coupling directly with the MSSM particles when some of the hidden sector presents $\langle F \rangle \neq 0$. The effects of the SUSY breaking are brought to MSSM (visible sector) by some messengers mediating given interactions (Figure 2.4). If the mediators are flavor blind, the soft SUSY breaking terms of the MSSM will also be - as favored by experimental results. Such mediators arise from gravity and the usual gauge interactions. The most favorable are the *gravity-mediated* or *Planck-scale-mediated* scenarios. Here the soft terms in the visible sector are inversely proportional to the Planck scale, $M_{soft} \sim \frac{\langle F \rangle}{M_P}$, because it should vanish in the limit $\langle F \rangle \to 0$ (no SB) and $M_P \to \infty$ (negligible gravity). For $M_{soft} \sim$ TeV the SB in the hidden sector appears at a scale of $\sqrt{\langle F \rangle} \sim 10^{10-11}$ GeV. Within such scenarios, the

[7] An alternative solution is to push the scalar masses (in particular of the first and second squak generation) at values higher than few TeV. Though, this is not really wanted as the hierarchy problem will not be entirely solved.

[8] Note that only three parameters are usually used.

gravity is explicitly included. The associated gauge fields are the gravitons and the gravitino (spin $2-\frac{1}{2}$). The goldstino is absorbed by the gravitino which acquires mass and longitudinal polarization though a super-Higgs mechanism.

Another scenario is the *gauge-mediated SUSY breaking* or GMSB [110], where the breaking is performed by the usual EW and QCD gauge interactions and the renormalizability of the theory is completely preserved. The soft terms arise from loop diagrams and the messengers are new chiral supemultiplets coupling to $\langle F \rangle$ (where F is a component of the new chiral superfield S having a non-zero vacuum expectation value). Unlike the gravity-mediated scenarios, the typical soft term mass is $M_{soft} \sim \frac{\alpha_a}{3\pi}\frac{<F>}{M_{mess}}$ [9] and $\sqrt{<F>} \sim 10^4$ GeV for a comparable M_{mess} mass. The messengers mass is acquired by the coupling with the S multiplet and the connection between the hidden sector and the MSSM is performed through the charges.

These one-loop couplings (Figure 2.5) lead to massive gauginos, $M_a = \frac{\alpha_a}{3\pi}\Lambda$ (where $\Lambda = \frac{<F>}{<S>}$). All the gauge bosons are protected by the gauge invariance and they don't get any contributions. The SUSY scalars acquire (positive) mass at two-loops level and are comparable with the gaugino ones :

$$m_i^2 = 2\Lambda\left[\left(\frac{\alpha_3}{4\pi}\right)^2 C_3^i + \left(\frac{\alpha_2}{4\pi}\right)^2 C_2^i + \left(\frac{\alpha_1}{4\pi}\right)^2 C_1^i\right] \tag{2.33}$$

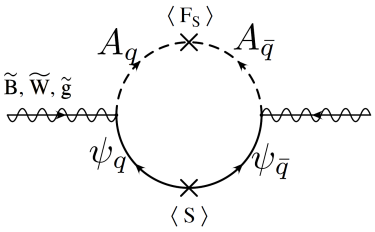


Figure 2.5: Illustration of the messenger - gauginos coupling.

where C_j^i are the quadratic Casimir invariants. This relation is obtained from the running Renormalization Group General Equations (RGE). The mass hierarchy of the sparticles is related to the strength of the gauge interactions and the lightest sparticle tends to be bino- or higgsino-like (unless not $N = 1$ SUSY). The goldstino is the gravitino and is the LSP. Its mass is very small and varies between few eV and few keV :

$$m_{3/2} \sim \frac{\langle F \rangle}{M_P} \ll M_{soft} \tag{2.34}$$

This has an important impact on the models conserving the R-parity, as the interactions between the gravitino and other particles cannot be neglected all the time. This is due to the two degrees of freedom of the gravitino participating in the interactions. Note that this simple GMSB model can be generalized by considering N copies of the messenger sector.

Finally, there are a lot of other symmetry breaking mechanisms. Some of them are assuming the existence of extra dimensions, which can be up to six as the string theory is proposing. In this case, the MSSM chiral supermultiplets are confined into a (4D) brane and the breaking appears into a parallel (hidden) brane. If between the two membranes only the gravity is propagated, the model is called *anomaly-mediated SUSY breaking* (AMSB). The soft term masses are given as a function of renormalization group parameters and the sleptons are predicted to have negative squared masses. In the case when the gauge fields are propagating between the branes, they are the messengers, and then the model is called *gauge-mediated SUSY breaking*.

2.3.4 The Brout-Englert-Higgs mechanism and the mass spectrum

As in the Standard Model, the sfermion and gauge bosons mass occur when the $< H > \neq 0$, which corresponds to the minimum of the potential. Moreover, the potential should break the $SU(2)_L \times U(1)_Y$ EW symmetry down to the electromagnetism scale (EM). In MSSM, this scalar potential is obtained from Equation (2.32) and it can be written as :

$$V = (|\mu|^2 + m_{H_u}^2)|H_u^0|^2 + (|\mu|^2 + m_{H_d}^2)|H_d^0|^2 - (bH_u^0 H_d^0 + c.c.) + \frac{1}{8}(g^2 + g'^2)(|H_u^0|^2 - |H_d^0|^2)^2 \tag{2.35}$$

where the $|\mu|^2$ terms arise from the F auxiliary fields and the g^2 and g'^2 terms from the D fields. Here the terms involving squarks and sleptons fields are not considered, as they are assumed to have vanishing VEV. Also H_u^+ and H_d^+ are set to zero, given the properties of the gauge transformations. This last choice also ensures that the potential is not broken, because the two components cannot get a VEV. The b, $< H_u^0 >$ and $< H_u^0 >$ terms can be chosen real and positive. This requirement is assured by two bounding conditions :

$$2b < 2|\mu|^2 + m_{H_u}^2 + m_{H_d}^2 \quad ; \quad b^2 > (|\mu|^2 + m_{H_u}^2)(|\mu|^2 + m_{H_d}^2) \tag{2.36}$$

[9] Here the first term is a one loop factor for diagrams with gauge interactions and M_{mess} is the messenger mass

If the second inequality is not satisfied the potential has a stable minimum, hence the symmetry is not broken. If $H_u^{0\,2} = H_d^{0\,2}$, as in mSUGRA/CMSSM or GMSB, Equations (2.36) are not satisfied. Therefore, the EWSB appears at tree level (when the conditions are satisfied) and it is driven by the quantum corrections which introduce a mass difference between the two Higgs terms. It is named the *radiative electroweak symmetry breaking*.

Once the non-zero VEV of the H_u^0 and H_d^0 terms is ensured, their mean values satisfy the relations :

$$< H_u^0 >^2 + < H_d^0 >^2 = v^2 = \frac{2m_Z^2}{g^2 + g'^2} \sim (174 \text{ GeV})^2 \tag{2.37}$$

$$\tan\beta = \frac{< H_u^0 >}{< H_d^0 >} \quad \text{or} \quad = \frac{v_u}{v_d} \sim \frac{v_t}{v_b} \tag{2.38}$$

where the β value is bonded only by the experimental results : $0 < \beta < \pi/2$. Further :

$$\sin 2\beta = \frac{2b}{m_{H_u}^2 + m_{H_d}^2 + 2|\mu|^2} > 0 \quad \text{and} \quad m_Z^2 = \frac{|m_{H_u}^2 + m_{H_d}^2|}{\sqrt{1 - \sin^2(2\beta)}} - m_{H_u}^2 - m_{H_d}^2 - 2|\mu|^2 \tag{2.39}$$

The last equation ensures that all the input parameters have a mass of the order of the Z boson squared mass. These equations also highlight the μ *problem*, as there is no reason to have the same masses for the μ term (appearing in the potential) and for the symmetry breaking terms (b, $m_{H_u}^2$ and $m_{H_d}^2$).

The Higgs boson

It was already mentioned that in SUSY the Higgs scalar fields are defined by two complex $\mathrm{SU}(2)_L$ doublets or by 8 real scalar degrees of freedom. After the SB, three degrees of freedom (the Goldstone bosons) are observed in the longitudinal components of the Z^0 and $W^\pm$ SM bosons, while the other degrees of freedom correspond to five real scalar fields : two CP-even h^0 and H^0, one CP-odd A^0 and two charged H^+ and H^- states. The masses of these different bosons are :

$$\begin{aligned}
m_{A_0}^2 &= 2|\mu|^2 + m_{H_u}^2 + m_{H_d}^2 \\
m_{h_0}^2 &= \frac{1}{2}\left(m_{A_0}^2 + m_Z^2 - \sqrt{(m_{A_0}^2 - m_{Z_0}^2)^2 + 4m_Z^2 m_{A^0}^2 \sin^2(2\beta)}\right) \\
m_{H_0}^2 &= \frac{1}{2}\left(m_{A_0}^2 + m_Z^2 + \sqrt{(m_{A_0}^2 - m_{Z_0}^2)^2 + 4m_Z^2 m_{A^0}^2 \sin^2(2\beta)}\right) \\
m_{H^\pm}^2 &= m_{A^0}^2 + m_W^2
\end{aligned} \tag{2.40}$$

(2.41)

Note that the Higgs sector of the MSSM involves only two parameters, m_{A_0} and $\tan\beta$. Whereas m_{A_0} is not constrained in MSSM, the second parameter must satisfy $1 \leq \tan\beta \leq m_t/m_b$. The masses of H^0 and $H^\pm$ bosons are not constrained either and are arbitrary large. By convention the h^0 mass is lighter than H^0 and is constrained at three-level by $m_{h_0} < m_Z |\cos(2\beta)|$. This relation implies large radiative quantum corrections, with the largest contribution coming from the top quark and top squark loops. At one loop order, beside the two parameters, the Higgs mass depends on additional parameters like top quark and top squark masses or top squark mixing angle (θ_t), entering through loop effects. The correction applied to the h_0 bosons mass is :

$$\begin{aligned}
\Delta m_h^2 &= m_Z^2 \cos^2(2\beta) + \frac{3}{4\pi^2} \sin^2\beta y_t^2 \Big\{ m_t^2 \ln \frac{m_{\tilde{t}_1} m_{\tilde{t}_2}}{m_t^2} + \cos^2\theta_t \sin^2\theta_t (m_{\tilde{t}_2}^2 - m_{\tilde{t}_1}^2) \ln \frac{m_{\tilde{t}_2}^2}{m_{\tilde{t}_1}^2} + \\
&\quad \frac{\cos^4\theta_t \sin^4\theta_t}{m_t^2} \Big[(m_{\tilde{t}_2}^2 - m_{\tilde{t}_1}^2)^2 - \frac{1}{2}(m_{\tilde{t}_2}^4 - m_{\tilde{t}_2}^4) \ln \frac{m_{\tilde{t}_2}^2}{m_{\tilde{t}_1}^2} \Big] \Big\}
\end{aligned} \tag{2.42}$$

With these corrections included (and other minor ones) an upper bound on the Higgs boson mass of 135 GeV (compatible with the observation) is obtained, assuming that the sparticles contributing are not heavier than 1 TeV. Note that this equation has some interesting features : (i) the m_{h_0} mass can be increased by increasing the $\tan\beta$ parameter, if its value is smaller than 13 [111] (when saturates) and (ii) the Higgs boson mass can be highly increased by large top squark mixing or by heavy top quarks ($\ll m_t$). This leads to the well known *Higgs mass tension problem* : as to accommodate a light Higgs boson mass (126 GeV) a large stop mass is needed, while to solve the hierarchy problem a light stop mass (close to the SM top quark) is necessary.

Neutralinos and charginos

After the SB of the gauge symmetry, the gauginos mix with the Higgsinos (terms like $\sqrt{2}gA_i^* T_a \psi_i \lambda^a$ in the Lagrangian, with $A_i = H_i$, $\lambda^a = \tilde{W}^a, \tilde{B}$ and $\psi_i = \tilde{H}_i$). Therefore, the chargino states arise from wino - charged Higgsinos mixing and the neutralinos from bino - neural Higgsinos mixing respectively. The neutralino (N) mass terms entering in the Lagrangian is :

$$\mathcal{L}_N = -\frac{1}{2}\begin{pmatrix}\tilde{B} & \tilde{W}^0 & \tilde{H}_d^0 & \tilde{H}_d^0\end{pmatrix}\mathbf{M_N}\begin{pmatrix}\tilde{B}\\ \tilde{W}^0\\ \tilde{H}_d^0\\ \tilde{H}_d^0\end{pmatrix}$$

$$\text{where}\quad \mathbf{M_N} = \begin{pmatrix} M_1 & 0 & -\cos\beta\sin\theta_W m_Z & \sin\beta\sin\theta_W m_Z \\ 0 & M_2 & \cos\beta\cos\theta_W m_Z & -\sin\beta\cos\theta_W m_Z \\ -\cos\beta\sin\theta_W m_Z & \cos\beta\cos\theta_W m_Z & 0 & -\mu \\ \sin\beta\sin\theta_W m_Z & -\sin\beta\cos\theta_W m_Z & -\mu & 0 \end{pmatrix} \tag{2.43}$$

M_1 and M_2 are defined in the MSSM soft Lagrangian, $-\mu$ is the higgsino (h_0) mass terms and the terms proportional with g and g' result from the Higgs-higgsino-gaugino couplings. A similar system can be obtained for the charginos (C) :

$$\mathcal{L}_C = -\frac{1}{2}\begin{pmatrix}\tilde{W}^+ & \tilde{H}_u^+ & \tilde{W}^- & \tilde{H}_d^-\end{pmatrix}\underbrace{\begin{pmatrix}0 & \mathbf{X}^T\\ \mathbf{X} & 0\end{pmatrix}}_{M_C}\begin{pmatrix}\tilde{W}^+\\ \tilde{H}_u^+\\ \tilde{W}^-\\ \tilde{H}_d^-\end{pmatrix} \quad \text{where} \quad \mathbf{X} = \begin{pmatrix} M_2 & \sqrt{2}\sin\beta m_W \\ \sqrt{2}\cos\beta m_W & \mu \end{pmatrix} \tag{2.44}$$

The mass eigenstates are obtained after diagonalizing the mass matrices defined previously by an unitary matrix : $M_N^{Diag} = \mathbf{Z^* M_N Z^{-1}}$ and $M_C^{Diag} = \mathbf{V^* M_C U}$. Therefore, the neutralino mass eigenstates are $N_i = Z_{ij}\psi_j^0$ and for chargino $C_i^+ = U_{ij}\psi_j^+$ and $C_i^- = V_{ij}\psi_j^-$.

Gluinos

Because the gluino is a $SU(3)_C$ color octet fermion, it cannot mix with other particles as is not having the appropriate quantum numbers. This is true also in models with R-parity violation. Therefore, at tree level, its mass is only given by the M_3 parameter in the soft SB Lagrangian. The quantum corrections to the mass can be as large as 25% (depends on the squark masses) and tend to increase the tree level prediction mass.

Sfermions

As in the Standard Model, all the fermion partners can mix if they satisfy certain conditions : same electric charge, R-parity and color quantum numbers. Hence, the mass eigenstates of squarks and sfermions are obtained by diagonalizing the 6 squared mass matrices for up type squarks, down-type squarks and charged leptons. For sneutrinos, the 3×3 mass matrix is diagonalized. Under the assumption that the soft breaking terms are flavor-blind, the first two squark generations and the sleptons have negligible Yukawa couplings. Hence, seven almost degenerate, unmixed pairs are formed as shown in Table 2.1. This leads to a small virtual sparticles contribution to the flavor-changing processes, avoiding the flavor problem mentioned in Section 2.3.3. Concerning the third generation, their masses are very different compared to the other two, because of the large Yukawa couplings and soft coupling in the renormalization group equations. This translates in large mixing between the left- and right-handed particles.

Discussion on sparticles mass spectrum

As the SB mechanism is not known yet, the SUSY particles masses are a priori not known. However, the experimental observations and the Standard Model problems are setting some useful constraints and guidelines. To solve the hierarchy problem, the spartner masses should not be heavier than few TeV (the so caller *natural SUSY* discussed in the next section). If this is considered to not be an issue, the masses could be very high as defined in the Split-SUSY scenarios [112, 113]. Here only the gauginos and Higgsinos have masses at the TeV scale and therefore discoverable at LHC, whereas the other particles are very heavy are not in the LHC range. In Figure 2.6 the running of the SUSY particles in a certain mSUGRA framework is given for illustration. The effects of the large Yukawa couplings are visible, as the $\mu^2 + m_{H_u}^2$ runs negative, and therefore it is responsible for the EWSB. Several SUSY mass spectra can be obtained

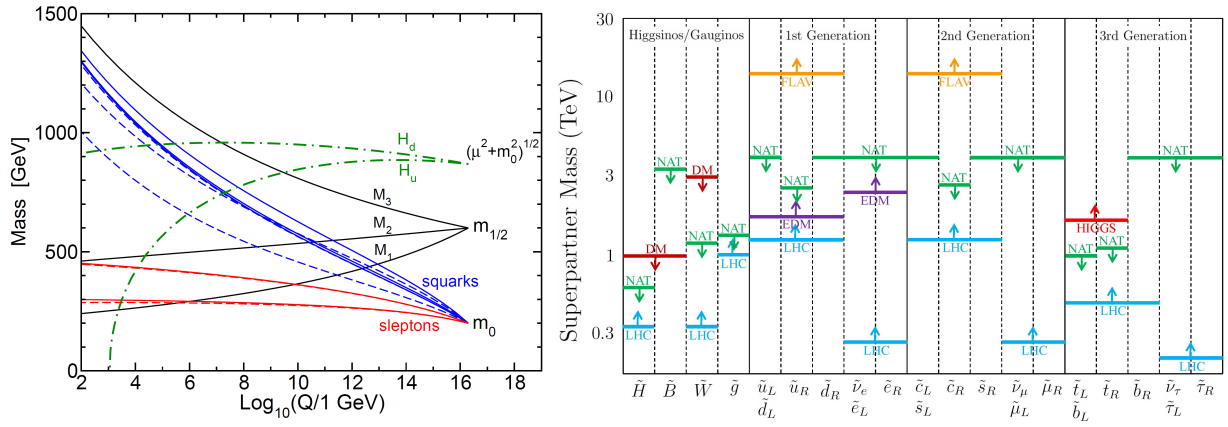


Figure 2.6: Left : Evolution of the scalar and gaugino mass parameters in MSSM, when the boundary conditions are imposed at the GUT scale [100]. Plot made in the mSUGRA model framework with $\tan\beta = 10$, m_0 = 200 GeV, $m_{1/2}$ = -A_0 = 600 GeV and $\sin(\mu) = 1$; Right : A natural SUSY spectrum given the constraints mentioned in the text [114].

under a considered SB mechanism (mSUGRA, GMSB, etc), and implicit with a certain set of constraints on the parameters of the considered model.

2.4 SUSY phenomenology

In this section a short description of the phenomenology in the MSSM framework is presented. It starts with an introduction of the *natural* SUSY spectrum in Section 2.4.1, detailing the expected range of sparticle masses and also the various constraints applied on these masses. In Section 2.4.2 the most interesting and favorite decay modes of the Standard Model superpartners are illustrated.

2.4.1 A natural SUSY spectrum

The scenarios where SUSY provides a solution to the hierarchy problem are known as *natural SUSY*. It was already mentioned that, for this to happen some stringent constraints on the sparticle masses must be imposed, as presented in Figure 2.6 right (NAT, green). For generality, the constraints obtained from the Dark Matter searches (DM, dark red), LHC[10] (light blue), Higgs boson observation (light red), flavor-violation (FLAV, orange) and Electric dipole moments (EDM, dark magenta) are also shown.

When considering scenarios with R-parity conservation, the LSP is considered to be the lighest neutralino ($\tilde{\chi}_1^0$). In the case when $M_1 < M_2$, μ the $\tilde{\chi}_1^0$ is likely to be mainly formed by the $\tilde{B}$ state, with a mass half of $\tilde{\chi}_1^\pm$. Also, the $\tilde{\chi}_1^\pm$ and $\tilde{\chi}_2^0$ are quasi-degenerate in mass. In the opposite case, $\mu < M_1$, M_2, the LSP is more Higgsino-like and the $\tilde{\chi}_2^0$ and $\tilde{\chi}_1^\pm$ masses are close to $\tilde{\chi}_1^0$. In both situations the interaction is governed by the gauge couplings. If $M_2 \ll M_1$, μ the LSP is wino-like and $\tilde{\chi}_1^\pm$ is slightly heavier [115]. All three cases are illustrated in Figure 2.7.

As showed in Figure 2.6 right, the charginos and neutralinos (M_i, i=1,2,3) are much lighter than the gluinos. In the mSUGRA case, this can be simply obtained from the one-loop RG equation for the running of strong coupling constant g_i :

$$\frac{\mathrm{d}\, g_i}{\mathrm{d}\, t} = -\frac{b_i}{2\pi} g_i^2 \tag{2.45}$$

and for the gaugino mass parameters :

$$\frac{\mathrm{d}\, M_i}{\mathrm{d}\, t} = -\frac{b_i}{2\pi} g_i M_i \tag{2.46}$$

where the b_i coefficients for a SU(N) gauge theory are generally written as $b_N = \frac{11}{3}N - \frac{1}{3}n_f - \frac{1}{6}n_s$ (see also Section 1.5 for the running of the coupling constants in the SM). Here n_f (n_s) is the number of fermions

[10] Note that these results are well improved after the entire 7 and 8 TeV data are analyzed as shown in Section 2.5

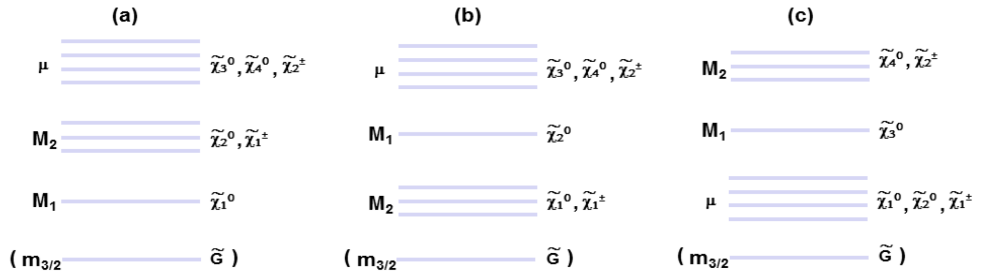


Figure 2.7: EW mass spectrum for the R-parity conservation case : (a) bino-wino case, i.e. mSUGRA, open spectra, (b) Higgsino case, i.e. natural, compressed spectra and (c) Wino case, i.e. AMSB, degenerate spectra [104].

(scalars) multiplets which couples to the gauge bosons. From these two equations, one can obtain :

$$\frac{\mathrm{d}}{\mathrm{d}\,t}\left(\frac{M_i}{g_i}\right) = 0 \quad \rightarrow \quad \frac{M_i(Q)}{g_i(Q)} = \frac{m_{1/2}}{g_i(m_P)} \quad \text{, P = Planck scale} \tag{2.47}$$

as all the gaugino masses are unified at GUT energies. If Q is considered to be the electroweak scale (∼91 GeV), where all the couplings are known, Equation 2.47 can be rewritten as :

$$\underbrace{M_3}_{bino} : \underbrace{M_2}_{wino} : \underbrace{M_1}_{gluino} \simeq 7:2:1 \tag{2.48}$$

Generally, the increase of the running gluino mass is explained by a larger QCD coupling with respect to the EW gauge ones. Given the $-\frac{8\alpha_s}{3\pi} M_3^2 \log\frac{\Lambda}{m_3}$ term, a gluino not heavier than few TeV is needed to protect the top squark mass (Figure 2.8).

$$\frac{M_1}{g_1^2} = \frac{M_2}{g_2^2} = \frac{M_3}{g_3^2} = \frac{m_{1/2}}{g_U^2} \tag{2.49}$$

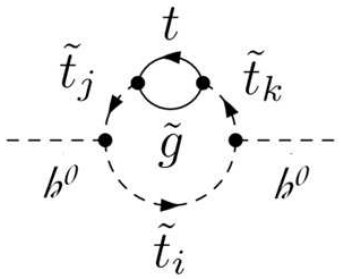


Figure 2.8: Constraints on $\tilde{g}$ mass.

Concerning the squarks, the mass difference between the first two generations and the third generation is significant. The formers are much larger than the masses of third generation constituents due to the absence of strong constraints on their masses. A light constraint requires that their mass should be smaller than the gluino - i.e. $m_{\tilde{q}} > 0.8 m_{\tilde{g}}$ in the mSUGRA framework. Moreover, the masses are quasi degenerate, to induce virtual contributions and obtain meson flavor changing during the decays via neutral currents.

The third generation squark are probably the lightest squarks, the stop in particular. This is due to the high Yukawa coupling and mixing effects between the left- and righ-handed constituents. The other constraints on the top squark mass - which should be lighter than few TeV - comes from the Higgs boson mass (Section 2.2). Beside them, notable are also the constraints coming from several pMSSM studies where the allowed spectrum of stop masses is given after considering the restrictions coming from experimental results like dark matter, flavor physics, search for SUSY particles, etc [116]. From the indirect constraints on the gluino mass, the stop should be lighter than 2 TeV. The sbottom mass is also pulled down by the stop mass, given the common mass term. Though, the constraints are not so tight as for the stop mass and a scenario with an sbottom mass below few TeV is still considered to be natural.

Finally, the slepton masses are lighter that the first two generation squarks as the radiative corrections are less important as for the strong couplings. The case when the stau is lighter than the other sleptons is very probable at large values of tanβ, due to the radiative effects of the large tau Yukawa coupling.

2.4.2 Decay channels

Only the most probable sparticles decay channels are presented in this section :

- The **gluino** decay can be either via an on-shell or off-shell squark and when the $\tilde{g} \rightarrow \bar{q}\tilde{q}$ two body decays are allowed they will dominate. If the third generation squarks are the lightest, the 2 body decays involving its constituents ($\tilde{g} \rightarrow \bar{t}\tilde{t}_1$ or $\tilde{g} \rightarrow \bar{b}\tilde{b}_1$) can be the only possible decay modes. In

the opposite situation, when the squarks are heavier than the gluino, its decay will be via off-shell squarks presenting a neutralino or a chargino in the cascade decays ($\tilde{g} \to \bar{q}q\tilde{\chi}^0$ or $\tilde{g} \to \bar{q}q\tilde{\chi}^{\pm}$). One important remark, is that if leptons are present in the decay, the probability of having a positive or negative charge is the same, because of the gluino Majorana nature.

- The dominant **squark** decay mode is $\tilde{q} \to q\tilde{q}$. If this channel is not allowed the squarks can decay into a quark and one chargino or neutralino ($\tilde{q} \to q'\tilde{\chi}^{\pm}$ or $\tilde{q} \to q\tilde{\chi}^0$). The decay channel involving $\tilde{\chi}_1^0$ is kinematically preferred and even dominant for the right-handed squarks (because LSP is mostly bino-like). The left-handed squarks are decaying mainly into charginos or heavier neutralinos as the squark-quark-wino is higher than the squark-quark-bino couplings. The decay into higgsino-like charginos for the first two generations is negligible, while for the third generation is not due to the high Yukawa couplings. If the decays $\tilde{t}_1 \to t\tilde{g}$ and $\tilde{t}_1 \to t\tilde{\chi}_1^0$ are kinematically forbiden, the lighter stop can decay as $\tilde{t}_1 \to t\tilde{\chi}_1^{\pm}$ or $\tilde{t}_1 \to bW\tilde{\chi}_1^0$. In the most extreme case, when all the mentioned decay modes are closed, the stop can decay as $\tilde{t}_1 \to c\tilde{\chi}_1^0$ and $\tilde{t}_1 \to bff'\tilde{\chi}_1^0$. The lifetimes of such decays are high enough, ensuring a quasi-stable stop hadronizing into bound states.

- The two-body **slepton** decays are $\tilde{\ell} \to l\tilde{\chi}^0$, $\tilde{\ell} \to \nu\tilde{\chi}^{\pm}$, $\tilde{\nu} \to \nu\tilde{\chi}^0$ and $\tilde{\nu} \to l\tilde{\chi}^{\pm}$. The case with LSP $= \tilde{\chi}_1^0$ in the cascade decay is always kinematically allowed. The other cases, with $\tilde{\chi}_2^0$ or $\tilde{\chi}^{\pm}$ in the decay are becoming important if the sleptons are heavy enough. Because the righ-handed sleptons are not coupling with SU(2)$_L$ gauginos, they are preferentially decaying into a bino-like $\tilde{\chi}_1^0$. If $\tilde{\chi}^{\pm}$ and $\tilde{\chi}_2^0$ are wino-like the favorite left-handed sleptons decay mode is having these gauginos in the final state, rather than $\tilde{\chi}_1^0$. This is because the slepton-lepton-wino coupling, proportional to the SU(2)$_L$ gauge coupling g, is higher than the slepton-lepton-bino coupling, which is proportional to the (smaller) U(1)$_Y$ coupling g'.

- Depending on the mass spectrum, the **gauginos** can decay into the lighter gauginos. Some examples are ${\tilde{\chi}^0}_i \to W^{\mp}{\tilde{\chi}^{\pm}}_j$, or ${\tilde{\chi}^0}_i \to Z^0{\tilde{\chi}^0}_j$ or ${\tilde{\chi}^0}_i \to h^0{\tilde{\chi}^0}_j$ for neutralino, and ${\tilde{\chi}^{\pm}}_i \to W^{\pm}{\tilde{\chi}^0}_j$ or ${\tilde{\chi}^{\pm}}_i \to Z^0{\tilde{\chi}^{\pm}}_j$ or ${\tilde{\chi}^{\pm}}_i \to h^0{\tilde{\chi}^{\pm}}_j$ for chargino. If the gauginos are too heavy the SM bosons are produced off-shell. In the case of light sleptons, there are other allowed channels, like ${\tilde{\chi}^0}_i \to \tilde{\ell}^{\pm}\tilde{\ell}^{\mp}$, ${\tilde{\chi}^0}_i \to \tilde{\nu}\nu$, ${\tilde{\chi}^{\pm}}_i \to \tilde{\ell}^{\pm}\nu$ and ${\tilde{\chi}^{\pm}}_i \to \tilde{\nu}l^{\pm}$. The (s)leptons can be replaced by quarks : $\tilde{\chi}_1^{\pm} \to qq\tilde{\chi}_1^0$ or $\tilde{\chi}_2^0 \to qq\tilde{\chi}_1^0$.

- The models with a gravitino LSP are well motivated, as its longitudinal component, the goldstino, couples non-gravitationally to all sparticles - particles pairs (S, P). Even if all the MSSM constituents S can appear as NLSP, in almost all gauge-mediated models this role is occupied by a neutralino or a charged slepton.

2.5 Experimental results

As the LHC is a hadron collider, the highest production cross section is reached for the gluino pair production, squark pair production or squark-gluon pair production as it can be seen in Figure 2.9. To search inclusively for such signals, given the decay channels presented in Section 2.4.2 and the R parity conservation, the target signature is characterized by several energetic (b) jets and high missing transverse energy. Note that requesting leptons (arising from gauginos or W/Z bosons leptonic decays) in the cascade helps to improve the sensitivity when the gluino and LSP mass are close (compressed spectra). The main sources of SM background are the production of a pair of top quarks and W or Z ($\to \nu\nu$) bosons associated to jets, processes which are accompanied by large $\not{E}_T$ (i.e. with neutrinos in the final state). Beside them, several sources of detector background are making the measurements more challenging, like the *QCD* (formed by fake (b) jets) or the fake $\not{E}_T$ and leptons backgrounds. The ATLAS results obtained for the inclusive searches are shown in the upper box of Figure 2.10 and point to strong limits on gluino (1.3 - 1.7 TeV) and weaker limits on 1^{st} and 2^{nd} squark generations (250 - 850 GeV). It is very important to emphasize that these results were obtained using also simplified models, with several stop decay channels, different sparticle masses hierarchies.

The searches for the third quark generation can be considered as one of the most motivated by the natural spectrum and *discoverable* at the LHC due to the light expected mass. When compared to other searches, the biggest experimental challenge is to remove the very high $t\bar{t}$ background, characterized by a cross section of ~240 pb at $\sqrt{s} = 8$ TeV. From the several scenarios considered, the highest sensitivity is reached with the stop analysis having one lepton, four jets and at least one b jet signature [119]. All the stop ATLAS results are presented in Figure 2.11 (left) and are pointing to an almost full coverage of the allowed (by the naturalness conditions) phase space. At the end of LHC Run-1, stop masses are excluded up to 530 - 640 GeV, constraining even more the existence of R-parity conserved SUSY. Note that before the LHC results no constraints were imposed on the stop by the natural spectrum. Generally, the same

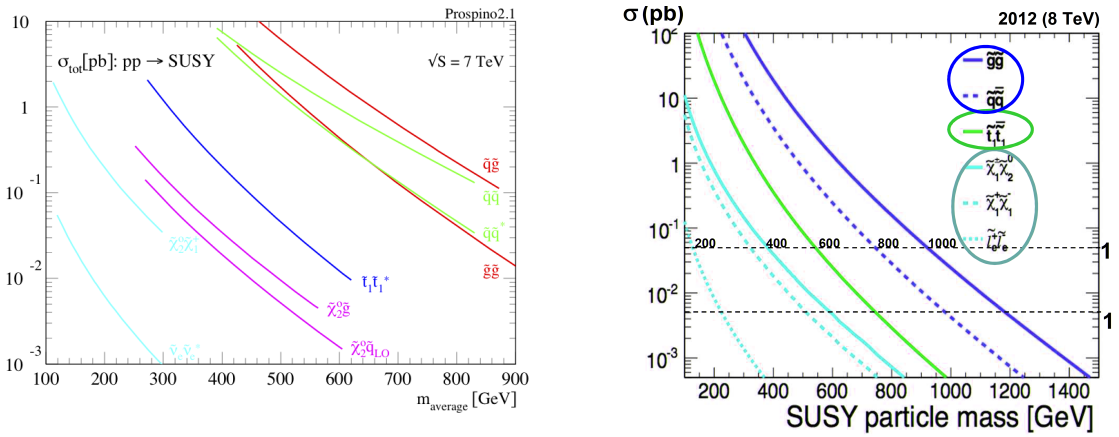


Figure 2.9: Production cross section for SUSY particles in proton-proton collisions at an energy of 7 TeV (left) and 8 TeV (right) in the center of mass [117].

sensitivity is reached also by the CMS collaboration. Regarding the sbottom searches, the dominant SM background comes from the top pair production in association with H or Z bosons decaying into b jets or in association with one or two b jets. From the detector background, an essential contribution arises from $t\bar{t}$ in association with tau leptons or c jets misidentified as b jets. The limits presented in Figure 2.10 are showing tight limits on the allowed masses, up to 630 GeV. Given these results, natural stop and sbottom scenarios are almost excluded but more complex scenarios can still hide.

Another important searches performed to ensure a full scan and coverage of the SUSY phase space are the direct production of gauginos and sleptons. Even if these signatures are characterized by small cross sections - like 0.1 pb for the $\tilde{\chi}_1^{\pm}\tilde{\chi}_1^{\pm}$ ($\tilde{\chi}_1^{\pm}\tilde{\chi}_2^0$) pair production, when an equal mass of 300 GeV (350 GeV) is considered - several signatures can be studied at LHC. The studied final state consists of multi-leptons (from the W and Z bosons) and missing transverse energy. For the charginos and neutralinos pair production, four main signal categories can be distinguished : (i) $\tilde{\chi}_1^{\pm}\tilde{\chi}_1^{\pm}$ with the main electroweak background formed by WW ($\rightarrow l\nu$) SM processes, (ii) $\tilde{\chi}_1^{\pm}\tilde{\chi}_1^0$ dominated by the leptonic W SM background, (iii) $\tilde{\chi}_1^{\pm}\tilde{\chi}_2^0$ with leptonic WZ SM background and (iv) $\tilde{\chi}_2^0\tilde{\chi}_1^0$ signatures dominated by leptonic ZZ SM background. However, only (i) and (iv) can be probed at LHC with the available luminosity of 20 fb^{-1}. The ATLAS results are shown in Figure 2.10, placing mass limits up to 700 GeV for both $\tilde{\chi}_2^0$ and $\tilde{\chi}_1^{\pm}$ in the most probable scenarios. For the direct slepton production the obtained mass limits are up to 350 GeV.

The SUSY Higgs sector [122] requests 5 Higgs bosons, h, H, A^0 and $H^{\pm}$, and plays an important role in all SUSY searches. Note that h is the most probable candidate for the just discovered SM Higgs, but the hypothesis of H being the SM Higgs is not formally excluded yet. Generally, two different approaches are considered : study the h boson couplings given the SM Higgs boson production and decay modes [10] and perform a direct search for H, A^0 and $H^{\pm}$. The former will be discussed in this thesis in Chapter 9 for the $t\bar{t} + H$ production. With the second approach, the strongest limits on the neutral bosons are coming from the decay modes including a pair of τ leptons [123]. The results points to heavy boson and the remaining allowed phase space is compatible with the SM Higgs. For the charged Higgs, the main production mode is via $t \rightarrow bH^{\pm}$, if $H^{\pm}$ is lighter than the top quark, otherwise is $t(b)H^{\pm}$. The main decay channel for $H^{\pm}$ is one tau lepton and the associated neutrino [124–126]. Another exotic search is considering the charge Higgs decay channel to be $H^{\pm} \rightarrow \bar{c}s$ [127]. In all scenarios no sign of new Higgs bosons is observed and the mass limits are pushed to values typically larger than 130 GeV.

The search for long lived particles (LLP) - other than LSP - are characterized by very specific signatures. If the LLP are very weakly couplings with the gravitino (GMSB models) they lead to a signal with non-pointing photons or Z bosons. In R-parity violating models, the lifetime is proportional to the λ^{-2}, λ'^{-2} or λ''^{-2} couplings and if at least one of them is of order 10^{-5}, the signature is a displaced vertex. Another case, when considering an AMSB model, is a small mass difference (100 MeV) between the $\tilde{\chi}_1^{\pm}$ and $\tilde{\chi}_1^0$ with a signature of twisted track due to the low momentum emitted. The LLP can be also stable and massive, leading to stable R hadrons - composed by squarks or sgluons - or sleptons. In ATLAS all these cases are considered (Figure 2.10) and they can explain the missing excess in the just mentioned SUSY searches. However, these are not the only possible scenarios and the topic cannot be considered to be entirely covered by ATLAS and CMS.

Finally, the models in which the R-parity is violated and implicit by the leptonic and baryonic numbers are also considered. The $\tilde{\chi}_1^0$ can decay into a pair of leptons and the associated neutrino, leading to a

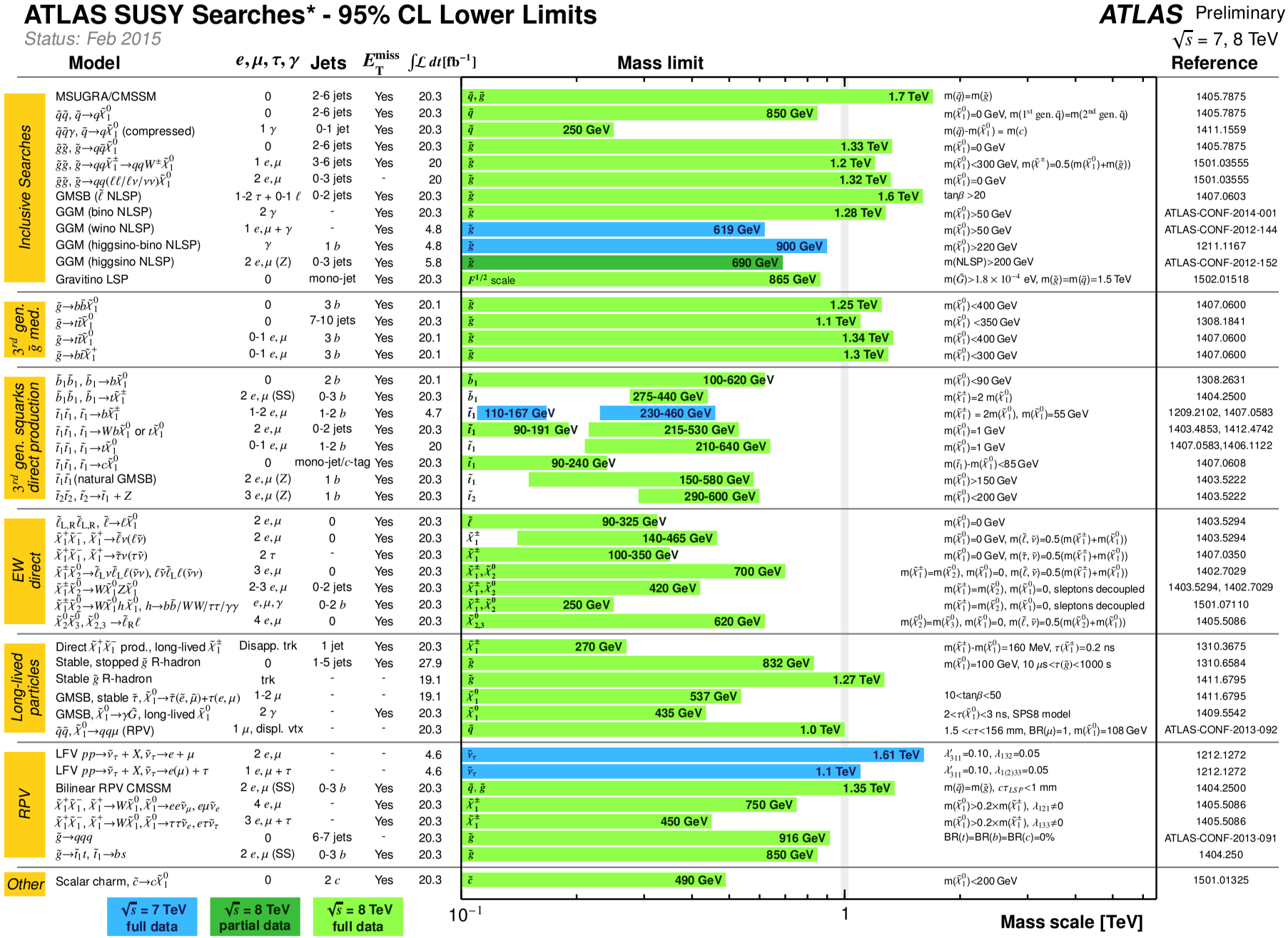


Figure 2.10: The ATLAS SUSY search physics program results [118].

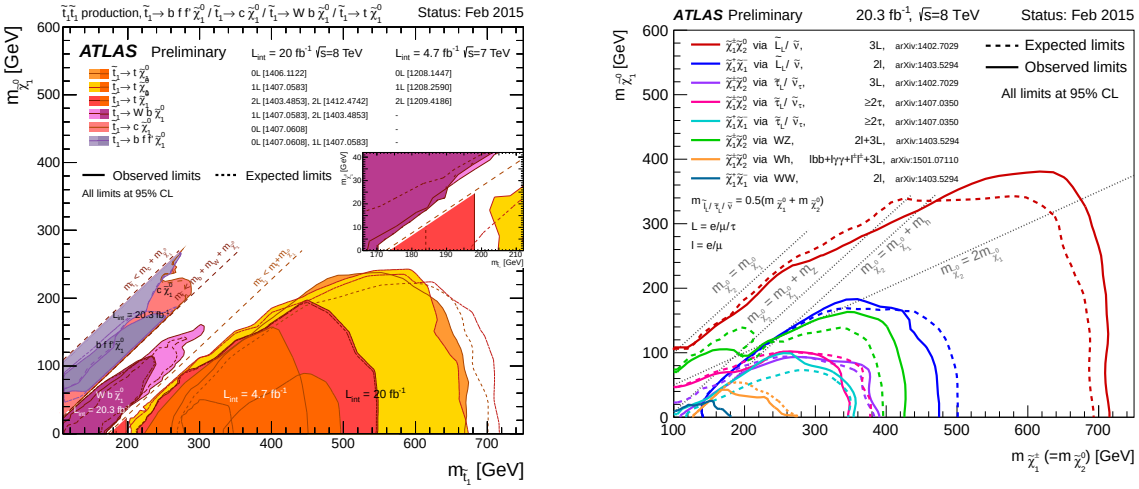


Figure 2.11: Left : ATLAS limits on top squark using $\sqrt{s} = 7$ and 8 TeV of data [120]. Right : Summary of ATLAS results in the searches for electroweak production of charginos and neutralinos using $\sqrt{s} = 8$ TeV of data [121].

signature with up to four leptons if the leptonic number is violated. Other signatures with up to six jets, dilepton or dijet resonances are all characterized by a lower background and $\not{E}_T$ when compared to R-parity conservation scenarios. The obtained mass limits in ATLAS are presented in Figure 2.10, and again are tighter than in the standard SUSY searches. As in the LLP search case a large spectrum is still uncovered.

Constraints from other indirect searches

Other SUSY constraints can be set with the Standard Model precision measurements, especially for the rare decays (small Branching ratio – BR) and even with SM forbidden decays. These signatures can be enhanced with SUSY particles in the loops and therefore contribute through loop corrections. From the first category, one can mention the W mass measurement with a precision of 2×10^{-4} and the anomalous muon magnetic moment measurement ($a_\mu = (g\text{ -}2)/2$) has a precision of 10^{-10} [128]. The latter disagrees at 3.6σ with the SM prediction and it can be explained by the presence of uncolored SUSY particles with masses of 100 - 400 GeV. In the second category an important role is played by B physics, in particular the studies of $B \to \tau\nu$, $B \to X_s\gamma$ or $B_S \to \mu^+\mu^-$ processes [129, 130]. The first process is sensitive at tree level to a messenger like the charged Higgs boson and the experimental results point to small difference between the expected BR in the SM $(0.97\pm0.22)\times10^{-4}$ and the observed value of $(1.16\pm0.22)\times10^{-4}$. The second decay mode is also interesting as the loops can present SUSY particles like light Higgsinos, stops or gauginos. The BR in the SM is $(3.15\pm0.23)\times10^{-4}$ and it is close to the experimental value of $(3.34\pm0.22)\times10^{-4}$. Finally the last decay, $B_S \to \mu^+\mu^-$, is a flavor changing neutral current process and highly suppressed in the SM as it violates the flavor conservation. The SM prediction of the BR is $(3.32\pm0.17)\times10^{-9}$ and the most recent observation is a BR of $(3.2^{+1.4+0.5}_{-1.2-0.3})\times10^{-9}$, excluding the very large variations in the loops. The main contribution is expected to be through the exchange of heavy neutral scalars - the so called *penguin* diagrams with a h Higgs boson and a loop of squarks and charginos. The h mediated processes can lead to potential large enhancements at large $\tan\beta$, the contribution being proportional with $\tan^3\beta$. Using these LHC results [131, 132] two intervals at 95% CL are obtained [130]:

$$1.1 \times 10^{-9} < BR(B_S \to \mu^+\mu^-) < 6.4 \times 10^{-9} \quad ; \quad 3.1 \times 10^{-9} < BR(B_S \to \mu^+\mu^-) < 4.6 \times 10^{-9} \qquad (2.50)$$

These relations are used to put limits in several MSSM models.

Figure 2.12 illustrates the effect of imposing the constraints presented in Equation (2.50) on β parameter phase space in two models, mSUGRA/CMMSM (left plot) and pMSSM (right plot). In mSUGRA/CMSSM, when the Higgs mass limit constraints are not considered, the obtained BRs can be 1.5×10^{-9} lower than the SM expectation (but still higher than the observed value). However, when the mass limits are imposed, the number of allowed points is drastically decreasing. This suggests the impossibility of having a smaller value for the SM expectation **and** be in agreement with the SUSY and Higgs boson results. When the results on the pMSSM $\tan\beta$ parameter are considered (Figure 2.12, right plot) it is visible that they do agree with the experimental observation. It can be concluded that more relevant parameters are used in pMSSM, compared to mSUGRA/CMSSM where some unrealistic *universality constraints* are used.

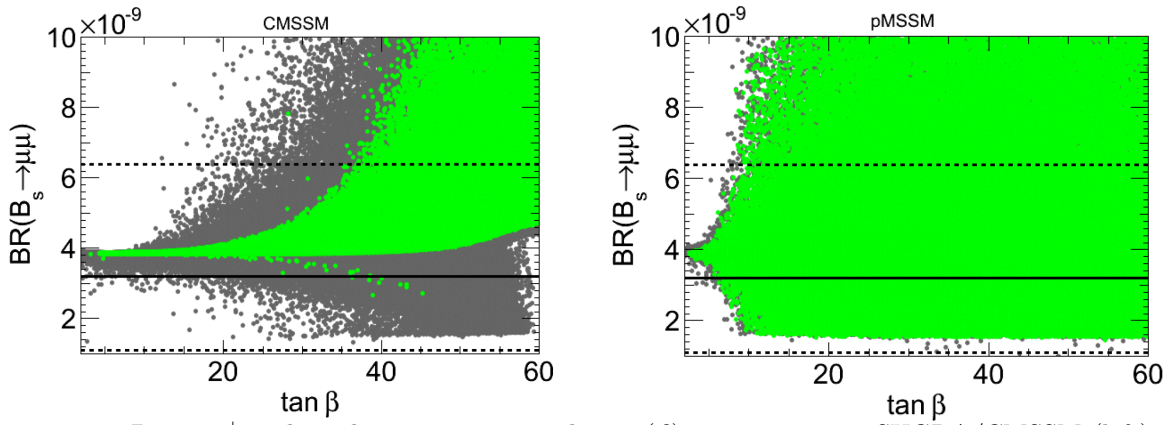


Figure 2.12: $B_S \to \mu^+\mu^-$ branching ratio versus the tan(β) parameter in mSUGRA/CMSSM (left) and pMSSM (right). The measured value is showed with a black line, while the $\pm 2\sigma$ variation is represented with the dashed lines. The points in agreement with the SM Higgs boson constraints are shown with green [130].

Constraints from dark matter searches

Beside the mentioned constraints on the SUSY MSSM parameters coming from indirect searches, additional constraints come from the dark matter searches. Moreover, the Supersymmetry theory can provide some dark matter candidates if the R-parity is conserved, as the LSP is stable and weakly interacting. Depending on the considered model, the most probable SUSY candidates are either the lightest neutralino, or the gravitino or the lightest neutrino. Hence, the observation of such a particle – besides neutrino – can be a sign of SUSY.

Several constraints on the SUSY particles in MSSM model are coming from different cosmological parameters. One of them is the density of the cold dark matter, which is obtained from astrophysical and cosmological measurements [133] and has the value of 0.1099±0.0062 [134]. It is considered to give the most precise constraint from all the indirect and dark matter searches. Another measurements, like relic density (Planck [73, 135]) are also considered.

In the regular dark matter searches, the WIMP signal detection is performed either directly or indirectly. The first category includes the experiments where the atoms resulting from WIMP-nucleus collisions are studied. It is using techniques either based on cryogenic germanium or silicon crystals (like CDMS [136] or EDELWEISS [68]), or based on noble gas scintillators (like Xenon100 [67]). As none of the existing experiments claim a significant WIMP signal, only limits on its nuclear scattering cross section are placed. Regarding the second category, the indirect searches are considering the detection of the decay products of the potential WIMP candidates. These signals appear either as an excess of positrons in the cosmic rays (PAMELA [69], AMS-02 [72]) or as an excess in the energy spectrum of gamma rays (FERMI-LAT [70]). Other experiments are looking for a WIMP signal coming from the Sun, which is characterized by highly energetic neutrinos [71]. When analyzing the data a small disagreement between the registered data and the background expectation is observed [72]. It can be explained in some dark matter annihilation scenarios, or if the quasar contribution is considered.

2.6 Search for SUSY in events with same - sign leptons

Even if the signatures including leptons are characterized by small branching ratios, they are well motivated as the level of SM and detector backgrounds is low. In this thesis I will consider mainly signatures with two leptons with the same electric charge or three leptons, (b) jets and missing transverse energy. The low SM background is mainly formed by diboson and $t\bar{t}$ processes in association with a W/Z boson or top + Z. This signature allows reaching some uncovered regions of the phase space or an increase in sensitivity in the compressed regions where the regular *no-lepton* signatures are not performing very well. The latter is due to the looser cuts imposed on several kinematic variables like jet transverse momentum, transverse mass or $\not{E}_T$. Moreover, in some scenarios involving leptons from gluino cascade decays, the probability to have a same - sign (SS) leptons final state is the same as for the two opposite (OS) leptons due to the Majorana nature of gluinos. The multi-lepton final states are ensured by signals including several Z and/or W bosons (arising from H bosons, $\tilde{\chi}_1^{\pm}$ or $\tilde{\chi}_2^0$ gauginos, etc). Several SUSY scenarios are predicting such a signature. The third generation searches are the main target and the considered signature is ensured by the presence of the top quarks (and further W bosons) in the intermediate states. Beside the SUSY signatures, such an

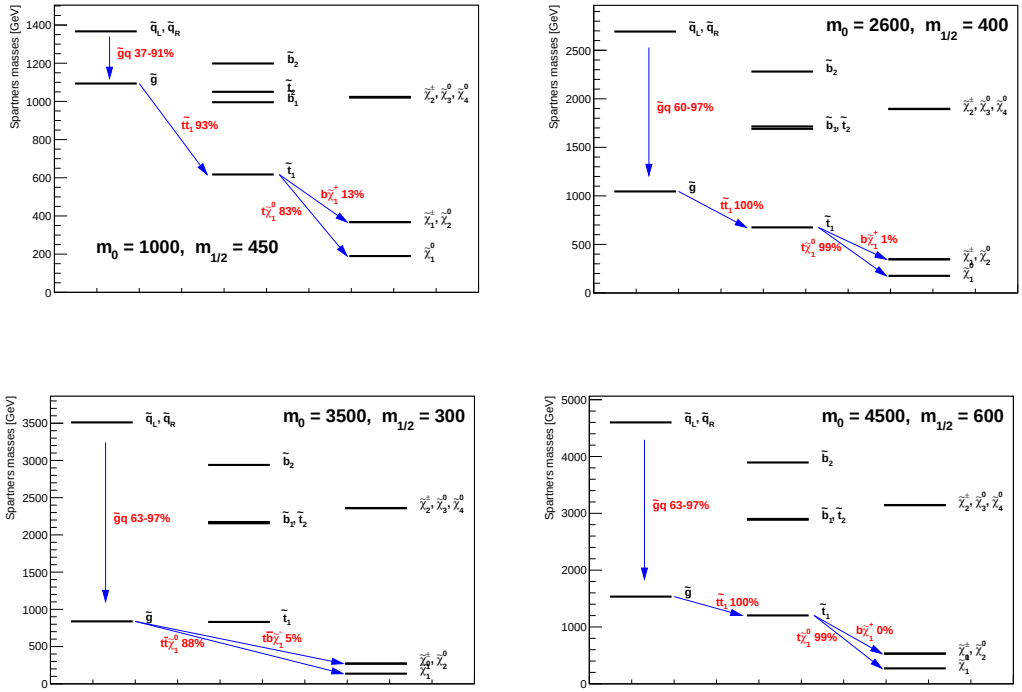


Figure 2.13: Some typical mSUGRA/CMSSM mass spetrum, for four points : top left $m_0 = 1$ TeV and $m_{1/2} = 450$ GeV, top right $m_0 = 2.6$ TeV and $m_{1/2} = 400$ GeV, bottom left $m_0 = 3.5$ TeV and $m_{1/2} = 300$ GeV and bottom right $m_0 = 4.5$ TeV and $m_{1/2} = 600$ GeV [2].

analysis is sensitive also to other beyond the Standard Model physics (not SUSY) like black holes, double charged Higgs, fourth quark generation, same sign top quarks, etc.

SUSY models

Because the SUSY breaking mechanism is not known, in MSSM a number of 105 new parameters are included. However, to perform an interpretation of the results with the SS leptons channel (or perform any other phenomenological interpretation) on a model with so many parameters is not realistic. Therefore, several approaches are considered to reduce the number of free parameters. Under some *reasonable assumptions* two main model categories were built : complex [137–141] and simplified [142–144] . The breaking mechanism is either gravity mediated, or gauge mediated or anomaly mediated. Among the complex SUSY models, the most popular are the mSUGRA/CMSSM (5 parameters) and pMSSM (19 parameters), which are described by the free parameters presented in Section 2.3.3. Some typical mass spectrum for mSUGRA/CMSSM and decay modes are shown in Figure 2.13 for four representative points.

Even fewer parameters are used in the simplified models. Here only a limited set of particle production and decay modes is allowed and the particle masses (and other) parameters are free. Therefore, certain signal topologies can be more easily studied and the optimization is performed in the given regions of the phase space to increase the sensitivity to any potential signal. When considering the results interpretation one should take them only as approximations. Both categories were considered in this thesis and Sections 6.1 and 7.1 are dedicated to their introduction.

2.7 Conclusions

In this chapter several theoretical motivations were presented in order to consider the Supersymmetry as an alternative theory to the Standard Model extension. In the context of *natural SUSY*, where the stop quark and gauginos have masses close to the EW scale, these SUSY particles can be observed at LHC. Several channels are possible to discover these new particles. In this thesis I am focusing my study on same - sign leptons, several (b) jets and missing transverse momentum final state. This analysis is promising because the SM and the detector backgrounds are low when compared to other channels with no leptons (for the same set of cuts). Moreover, beside the SUSY scenarios several non-SUSY topologies can be also studied and the considered parameters constrained. The entire $\sqrt{s} = 8$ TeV data is considered for this analysis.

Chapter 3

LHC and ATLAS detector

3.1 The Large Hadron Collider

The largest and most powerful hadron collider in the world is LHC, a research program approved in December 1994 [145, 146]. It is built and operated by the European Organization for Nuclear Research (CERN) and placed into a tunnel of 3.8 m in diameter, at a depth of 50 - 175 m, near Geneva. It consists of two rings, the superconducting accelerator and the collider. It was designed to collide proton-proton beams and reach an energy of 14 TeV in the center of mass. The main research program goals are the SM problems and the new physics beyond the Standard Model. Prioritized are the discovery of the Higgs boson and the searches for Supersymmetry. The LHC project has three big experiments : CMS and ATLAS aiming at a instantaneous luminosity of 10^{34} cm^{-2}s^{-1} and LHCb aiming at $L_{inst} = 10^{32}$ cm^{-2}s^{-1}. Beside the proton beams, LHC accelerates lead nuclei, targeting an energy of 5.5 TeV in the center of mass and a luminosity L_{inst} of 10^{27} cm^{-2}s^{-1}; the resulted physics is analyzed at the ALICE experiment.

The proton beams

Prior to the collision, the two proton beams are accelerated up to the nominal energy, in opposite directions, in two different LHC beam pipes. Before injecting the proton beams in the LHC ring, its constituents are accelerated up to 450 GeV by a series of consecutive accelerators (Figure 3.1). First, the linear accelerator LINAC 2 generates protons with an energy of 50 MeV, which are injected in the PSB boosted synchrotron and accelerated up to 1.4 GeV. Further, they are accelerated in the PS synchrotron up to 26 GeV and finally in the SPS, the super proton synchrotron, where they reach the energy of 450 GeV. Once the proton beam is injected in the LHC machine, it is captured and accelerated at a frequency of 400 MHz. For each $\sqrt{s} = 7$ TeV beam are expected a total of 2808 proton bunches. Each bunch is composed by $\sim 10^{11}$ protons. This implies a required maximum electric field of 16 MV/beam (it was 8 MV at the 450 GeV injection). The nominal bunch spacing is 25 ns.

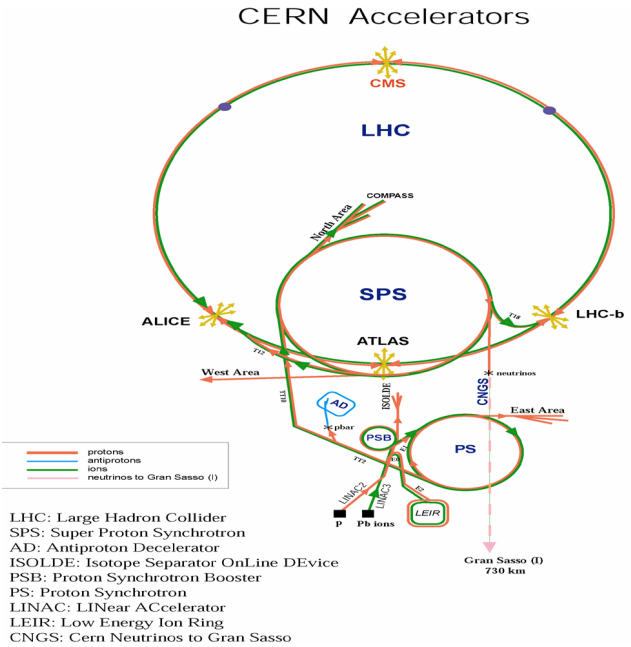


Figure 3.1: Proton beam accelerating system [147].

The magnetic and cryogenic systems

A very high magnetic field is required to direct and hold together the protons in the circular beam pipe. To reach such high intensities, superconducting magnets are used. The magnet coils are based on niobium-titan superconductor cables, which are able to conduct the electricity without resistance or energy loss. With the superconductivity facilities, the LHC magnetic system is able to operate at required intensities of up to 8.4 T, given an electric current of 11850 A. Of course this implies also an adequate cryogenic system, able to provide a temperature below 1.9 K. Highest constraints on the cryostat system design are coming from the reuse of the LEP tunnel and its facilities to construct the LHC collider. Hence, only five cryogenic islands - grouped in one cryogenic distribution line - are used to distribute over distances of up to 3.3 km

the cooling superfluids. A total of eight refrigeration plants, filled with liquid Helium, are supplying the required cooling superfluid.

The proton beam bending is ensured by 1232 dipole magnets, each piece weighting 27.5 tonnes, a length of 16.5 m and a diameter of 570 mm. Each magnet has two apertures, for the two proton beams tubes (two-in-one magnet design, Figure 3.2) and provides a magnetic field of 8.33 T. The required temperature is most of the time 1.9 K, but it can also go up to 4.5 K. As the protons are charged particles, it is important to focus the beam and retain it in the ultrahigh vacuum chamber. This is performed using a total of 858 quadrupole magnets arranged in the FODO common lattice. Each is 5 - 7 m long and designed to generate a field gradient of 220 Tm^{-1} at 1.9 K. Other thousands of multipoles magnets are used to improve the beam focusing and also to reduce the unwanted interactions between the bunch constituents. Finally, inner triplet magnets are used to squeeze the protons in the collision points. The bunch size varies between 0.2 mm and 16 μm.

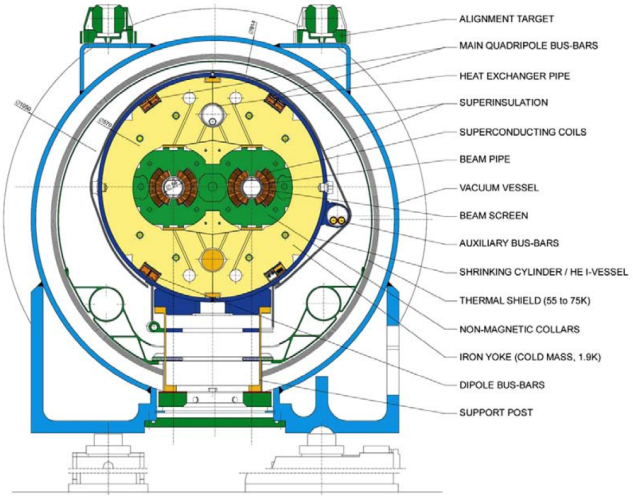


Figure 3.2: Cross section of the LHC dipole [145].

Beside the mentioned magnets, many more normal magnets are used in order to reach the target beam parameters, leading to more than 9000 magnets surrounding the LHC ring. They represent around 90% (40000 tonnes) of the cold mass.

Magnet powering and protection

The magnet powering is performed in eight independent and symmetric sectors, to ensure a limitation of the stored energy in a magnet. Each sector is having only one electrical circuit connecting in series the main beam bending dipole magnets. The quadrupole magnets are supplied by two separate circuits, to focus and defocus the beam. The list continues with the circuits needed to connect the other magnets needed at LHC. It leads to a total of 1612 electrical circuits, with a stored energy up to 1.22 GJ for the main dipole and 20 MJ for each main quadrupole circuit, in each sector. This high energy can damage the superconducting magnets and the bus-bars. If such a magnet is superheated, it can loose the superconductivity properties and return to the normal conducting state (quench). In the worse case, the magnet can be destroyed. To avoid this, a sophisticated energy subtraction system is implemented. Hence, if a resistive voltage appears, it is rapidly detected and the energy is dissipated via a dummy resistor (cold diodes) connected in series with the magnet. The superconducting bus-bars are as well protected against overheating.

The importance of the magnets and bus-bars protection can be highlighted by the September 2008 accident. During the powering tests of the main dipole magnet circuit, an electric arc appeared due to a wrong electrical connection between two magnets. This happened after the electrical current was increased from 7 kA to 9.3 kA. When the intensity reached 8.7 kA, a resistance of the bus-bar appeared and activated (correctly) the dump resistors. Even so, an electrical arc developed and perforated the cryostat. It lead to an explosion which destroyed around thirty superconducting magnets, several connections and the vacuum. This accident delayed the LHC operations for at least one year [148].

Vacuum system

Beside the powerful magnetic system, the LHC machine is characterized by the biggest vacuum system in the world. It is divided in three subsystems. First are the insulation vacuum for cryogenically cooled magnets and the insulation vacuum for the helium distribution (50 km of pipes). They reach a vacuum pressure of 10^{-6} mbar at cryogenic temperatures (5 - 20 K). Its role is to maintain a constant temperature around the cryogenic parts. More stringent requirements are imposed on the third subsystem, the beam vacuum. This is mandatory, as the interactions between the protons and the gas molecules should remain negligible. Therefore, an ultra-high vacuum is reached using 48 (6) km of arc (straight) sections. In the arc sections the vacuum pressure is lower than 10^{-9} mbar and the temperature is 1.9 K. The straight sections are held at the room temperature, with a vacuum pressure of 10^{-10} - 10^{-11} mbar. Here a lower pressure is needed because these sections are located near the interaction point, where the collisions are produced.

LHC performance during Run-1

The first proton-proton collisions at LHC occurred in November 2009, at an energy of 900 GeV in the

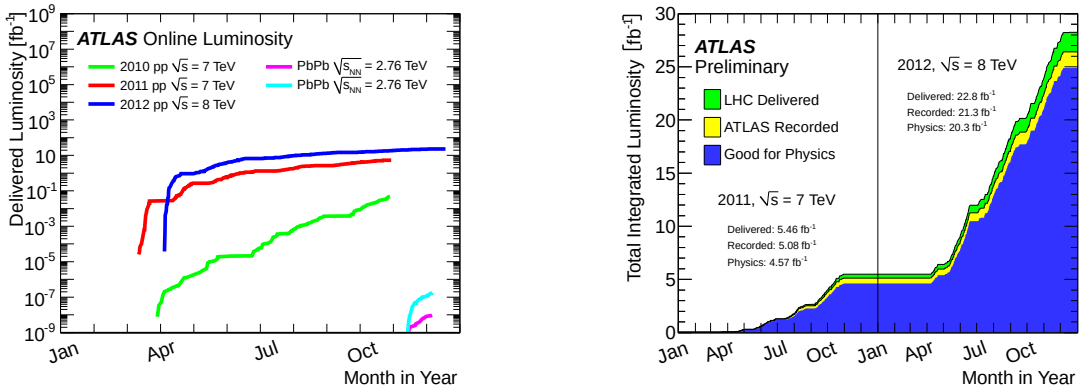


Figure 3.3: Cumulative distribution showing the delivered luminosity during 2010, 2011 and 2012 data taking (left) and the total integrated luminosity and data quality in 2011 and 2012 [149].

center of mass system and increased to 2.37 TeV within few days. After the test period was closed, in April 2010 the $\sqrt{s} = 7$ TeV collisions started. In a first step, only an integrated luminosity of 47 pb^{-1} was collected (the so called 2010 data) as the machine was still under validation. It increased rapidly during the eight months of data taking as shown in Figure 3.3 left. The 2010 run finished with 368 bunches, having around $2{\cdot}10^{11}$ protons per bunch. After analysing the data, several Standard Model particles were re-observed (i.e. the top quark, W and Z bosons), cross sections were measured (i.e. $t\bar{t}$, W^+W^-) and the mass limits on the beyond the Standard Model particles were tightened (i.e. on several SUSY particles). In March 2011 the second $\sqrt{s} = 7$ TeV data taking began with an increased number of bunches in the beam. The *tests* were performed using 50 ns bunch spacing. At the end of June 2011 the number of bunches increased to 1380. By the end of October - the end of the 2011 run - the integrated luminosity was around 4.7 fb^{-1}. Many publications improving the SM measurements and tightening the mass limits on BSM particles are released after analyzing this data. The last Run-1 data taking started in May 2012, with an increase in energy in the center of mass to 8 TeV. The number of bunches remained at 1380 and the bunch spacing at 50 ns. For this 2012 run, a tighter β^* was used and a final integrated luminosity of 23 fb^{-1} was recorded. The instantaneous luminosity reached $7.7{\cdot}10^{33}$ cm^{-2}s^{-1}, a value close to the target [150]. The greatest achievement of this run was the Higgs boson discovery and the high increase in sensitivity of physics BSM which implies the production of heavy particles. One example can be considered to be the production of SUSY particles (Chapter 2). After analysing the 2012 data a lot of results are publicly released in articles and CONF notes. Notable is the precision reached for the SM measurements and the very tight mass limits on the particles considered in the *natural SUSY spectrum*.

In Figure 3.3 right a comparison between the delivered LHC luminosity and the registered luminosity by the ATLAS detector during the $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV data taking is performed. The presented luminosities for the three runs are delivered to both ATLAS and CMS experiments. Compared to 2011, during the 2012 run the number of simultaneous interactions increased considerably (Figure 3.4). As already mentioned in the introduction, beside the proton collisions, the LHC program includes also lead-lead runs. The energy in the center of mass is set to 2.76 TeV. In 2010 the delivered luminosity was 9.7 μb^{-1}, while in 2011 166 μb^{-1}. In this thesis, none of the heavy nuclei runs were considered, only the proton-proton collisions data collected at $\sqrt{s} = 8$ TeV.

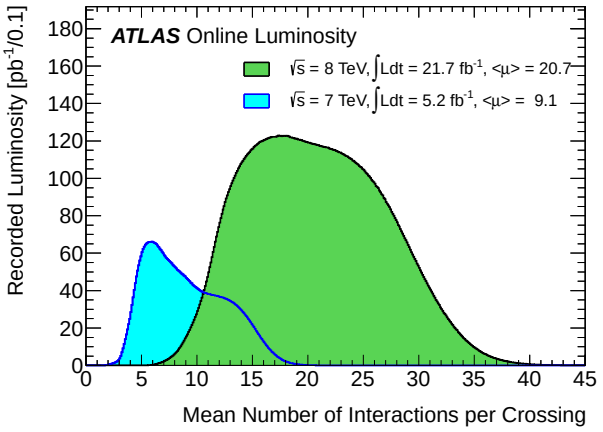


Figure 3.4: Distribution of mean number of interactions per bunch crossing.

The four main experiments at LHC

Given the large physics topics areas targeted by the LHC program, four main experiments were installed and developed. The ATLAS and CMS are the biggest experiments and are designed to understand and discover the physics implying heavy (unknown) particles around the TeV scale. The b physics is primarily studied at the LHCb experiment, which covers the main subjects of heavy flavor and electroweak physics. The recorded luminosity was 1 fb^{-1} in 2011 and 2 fb^{-1} in 2012. It allowed the study of $B_s \to \mu^+\mu-$ rare SM processes (i.e. flavor changing neutral current) and the CP violation in different processes involving B_s, D_0 and Kaon meson states. It also lead to the discovery of two new baryon resonances, Ξ_b^- [151]. A large

ion collider experiment (ALICE) is studying the quark confining mechanism, by creating a quark-gluon plasma. The considered collisions are lead-lead or lead-proton.

3.2 The ATLAS experiment

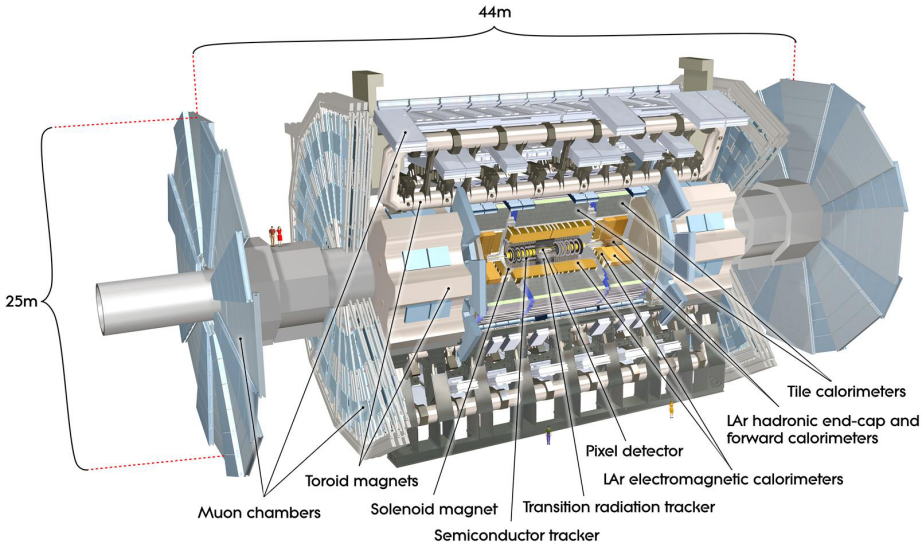


Figure 3.5: The ATLAS detector and its sub-detectors [152]. Is having 7000 tons, with an overall length of 42 m and a diameter of 22 m.

The ATLAS detector [153,154] is designed to cover a large range of physics involving new heavy particles which could not be discovered in the previous experiments. An important target is the Higgs boson, which was used as a benchmark to establish the performances of the ATLAS sub-detectors. A good resolution is needed in particular for the $H \to \gamma\gamma$ channel. To study the $H \to b\bar{b}$ decay channels a high b-tagging performance is mandatory. For any precise measurement in the channels including leptonic final state, like $H \to ZZ^* \to llll$, the lepton reconstruction, identification and coverage are critical. These performances are needed also for any BSM search involving extra Higgs bosons, like H^{++}, A or H^-. Potential heavy bosons, like W' or Z', can also be discovered in ATLAS. Their search can be performed in the leptonic channels, which requires high resolution momentum measurements at the TeV scale. These signals can present very energetic jets, for which a high resolution in the TeV range is needed. Supersymmetric or extra-dimensional models, are dominated by signatures which can lead to high missing momentum, which should be precisely determined. This is one of the reasons why the detector has a coverage close to 4π solid angle (Figure 3.5). To achieve all the requirements, the ATLAS detector is composed of four main sub-systems. The tracking system, measuring the charged particles trajectory, is the closest to the interacting point. An electromagnetic calorimeter stops and determines the energy deposited by electrons and photons. It also provides all the information necessary for the electron and photon reconstruction and identification. The hadronic calorimeter has the same purpose for the jets. The two calorimeters cover the η range up to 4.9. As the tracking system is functioning in the $|\eta| < 2.5$ region, the fine granularity of the electromagnetic calorimeter provides all the information needed for the precision measurements of electron and photons. The hadronic calorimeter has a coarser granularity, suited for the jet reconstruction and transverse momentum computation. Finally, the muon spectrometers are forming the *covering* of the detector. A short description of all sub-detector systems is given in this section.

The ATLAS coordinate system is considered to be right-handed, with the z axis along the beam pipe, the x axis pointing to the center of LHC (interaction point), and the y axis upward. The azimuthal angle is $\tan\varphi = \frac{y}{x}$, and the polar angle is $\tan\theta = \frac{\sqrt{x^2+y^2}}{z}$. The pseudorapidity is calculated as η = -ln($\frac{\tan\theta}{2}$). The angular separation between two objects is $\Delta R = \sqrt{\Delta\eta^2 + \Delta\varphi^2}$.

3.2.1 The tracking system

The tracking system, or the inner detector (ID), is composed of three sub-detectors : the pixel detector, the silicon microstrip tracker (SCT) and the transition radiation tracker (Figure 3.6). It is surrounded by a magnetic field with an intensity of 2 T and the first two sub-detectors cover the region with $|\eta| < 2.5$. The

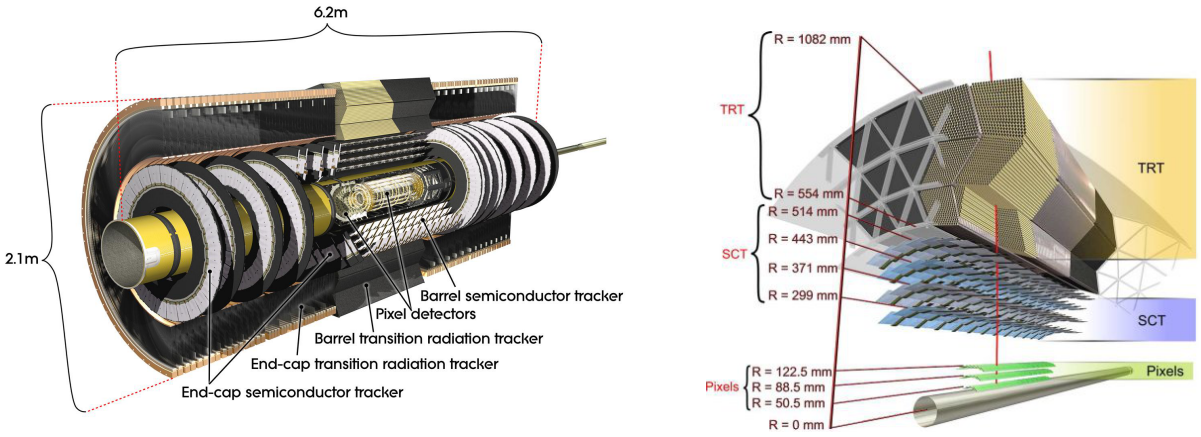


Figure 3.6: Cross-section of the Inner Detector [152].

TRT capabilities are up to $|\eta| = 2$. All tracks with $p_T > 0.5$ GeV are considered. Its performance consists in a robust pattern recognition, a very good momentum resolution and both primary and secondary vertices measurements of the charged particles tracks.

The pixel detector

It is situated at 5 - 10 cm distance from the interaction point, has the highest granularity and provides an excellent spatial resolution as needed for the momentum and impact parameter measurements and vertices identification. Note that the tracks density present in the detector after the collisions is huge - around 1000 particles are expected to emerge for every bunch crossing (every 25 ns). The pixel detector is formed by three concentric cylindrical layers around the beam axis in the barrel region of the detector and three perpendicular disks on each side in the end-cap region. Same arrangement is valid also for the other sub-detectors. The first layer - the vertexing layer - is at a radius of 51 mm, while the second and the third at a radius of 89 mm and 123 mm respectively. They are segmented in η - φ and z, providing a high resolution of 3D space point measurements. A total of 1744 silicon pixel sensors - diodes - were installed, each measuring 50×400 μm^2, with a thickness of 250 μm. They will start to operate at 150 V and after few years of operation the voltage will be increase to 600 V. The operation temperature is -5 / -10°C. Each sensor is formed by 47232 pixels, leading to 46080 readout channels. The total number of pixels is 80.4 millions. Only 5% were damaged at the end of 2012 run [155], mainly due to failures of the optical transmitters or breaks of the high voltage supplies. During the 2012 run, the pixel detector efficiency during the data taking was 99.9% [156].

The SCT detector

Together with the Pixel detector, it contributes to the momentum, impact parameter and vertex position determination. It has a good resolution in the bending plane and provides 3D space point measurements. In the barrel region it consists of four cylinders (eight layers), while in the end-caps of 18 discs (9×2). On them are attached 4088 (2112 + 1976) silicon sensors, initially operating at a voltage of 150 V, which can increase to 250 - 350 V after few year of functioning. Note that the pixels used for the Pixel detector were replaced by strips, given the larger covered area of 63 m^2. Same operating temperature is needed. The sensor has a thickness of 285 μm and is formed by 780 readout strips. Each strip step is 80 μm large and up to 6 cm long. Given the arrangement in strips, the total number of channels is 6.3 million. For a better identification of the position of the hit, the four sides of the cylinder layers are rotated and oriented at an angle of $\pm$20 mrad. Same rotation angle is retained also for the end - cap discs. The approximate operational fraction of the SCT detector was 99.3% during the data taking [155]. In 2012, the data acquisition efficacy was 99.1% [156]. The main reasons for the observed inefficiency are the read-out errors and the busy states [157].

The TRT detector

With respect to the interaction point, it is situated at radii of 55 cm and 1.1 m, at the edge of the ID. Hence, it assures a continuous tracking at large radii for all particles increasing the momentum resolution. In particular, it provides around 36 measurements (hits), leading to a precise particle identification and track reconstruction in the $\eta - \varphi$ plane. Its design is a combination of straw trackers and transition radiation detectors and operates at the room temperature. The straws are polyimide drift tubes with a diameter of 4 mm and a length of 144 cm for the barrel and 37 cm for the end-cap. The wall of the straw is a 25 μm

thick polyimide film. Each straw anode is a gold plated tungsten wire, with a diameter of 30 μm. On the cathode, a voltage of 1530 V is applied, and the straw is filled with a mixture of Xe (70%), CO_2 (27%) and O_2 (3%) gas. This combination assures an increase of ionized charges by a factor of 25000. In the barrel, the TRT is formed by 73 layers of straws surrounded with propylene fibers serving as transition radiation material and in the end-cap by 160 straw planes surrounded by propylene foils.

The particle identification capability of the TRT is driven by the performance of detecting the transition radiation (TR). The TR photons are emitted when the relativistic charged particles traverse the boundary between the two different dielectric mediums (the gas and the propylene). They are absorbed in the gas, producing larger electric signals than the minimum ionizing particles. Given the performance of the TRT detector, light particles as the electrons can be discriminated from charged pions, ensuring the high electron identification efficiency. If the Xe-based mixture is replaced by an Ar-based mixture, the tracking capacities would remain unaffected, but the cross section for absorbing the TR would decrease by a factor 10. This reduces dramatically the particle identification capabilities. This hypothesis was deeply studied at the end of 2012 data taking, as during this run several leaks (200 liters per day at the end of the run) were present in the output pipes, at the gas exit from the TRT modules. Because the Xe mixture is very expensive, for the Run-2 data taking, several regions of the TRT detector will be filled with the Ar-based mixture. Studies showed that if the most pessimistic scenario is considered, namely all the TRT detector regions are filled with the Ar-based mixture, a high increase of fake electrons in the low p_T and high η regions will appear. Beside the electron reconstruction and identification, also the photon reconstruction is highly affected [1]. In particular, I was involved in studying the impact on the analyses presented in Chapters 6 and 7. These results were documented in two internal notes [158, 159].

At the end of Run-1 (2010 - 2012 period) the number of non-operational TRT channels was 2.5% [155]. Around 1% was due to mechanical problems, while the rest of 1.5% due to electrical problems. Overall, the TRT efficiency during the 2012 run data taking was 99.8% [156].

ID track reconstruction and its performance

The tracks associated to the charged particles are reconstructed in the ID if $p_T > 0.5$ GeV and $|\eta| < 2.5$. For the soft particles, the efficiency is limited by the large amount of material (around 4.5 tons, Figure 3.7). In a first instance, the raw data from the pixel and SCT detectors is clustered and the TRT timing information is giving the particle *position*. This information is used as input to different track fitter algorithms. The track reconstruction procedure includes global χ^2, Kalman-filter and Gaussian filter techniques in order to discriminate against the noise, to minimize the fake tracks and to remove the cluster-track ambiguities. In the last step, the primary vertices are reconstructed.

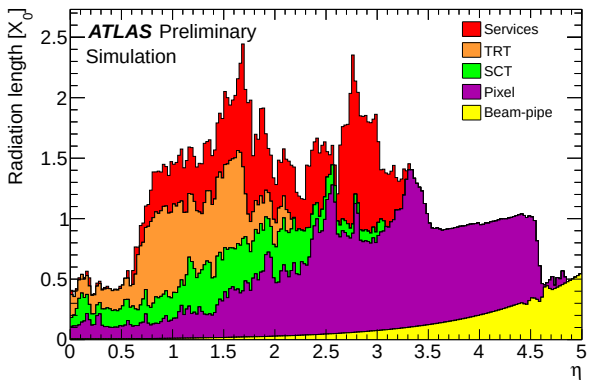


Figure 3.7: Amount of material up to the ID boundaries, in units of radiation length X0, traversed by a particle as a function of η [160].

A reconstructed track in the central region of the detector is considered to be of good quality if presents at least 3 hits in the pixel detector, 4 hits in the SCT and around 30 in the TRT. The precision of one hit, in the azimuthal and longitudinal plane is 10×115 μm in the pixel detector, 17×580 μm in the SCT and 130 μm in the TRT. The transverse momentum relative resolution for muons with $p_T = 100$ GeV is around 4% in $|\eta| < 1.5$ region, increasing up to 16% in for $|\eta| = 2.5$. For pions with p_T of 100 GeV, the resolution on the d_0 impact parameter is around 1%, for the $z_0 \times \sin\theta$ it varies between 5% (in the barrel) and 10% in the end-cap. The associated uncertainty is 10%. The electron charge determination is very important for new physics searches performed in the same - sign channel, as as well as for the charge asymmetry measurements. For muons this is precisely determined using the muon spectrometer, while for electrons only the information from the ID is used. This measurement is challenging given the conversions of bremsstrahlung photons, which leads to pattern recognition problems. More details on the wrong charge measurements are given in the analysis chapters of this thesis.

For each event a reconstructed primary vertex (PV) with at least three tracks is assigned. In events with multiple vertex candidates the PV is considered to be the vertex with the highest $\Sigma p_{T\,track}^2$ value, while in events with multiple simultaneous interactions this is not necessarily true. Therefore, the PV candidate is identified among the previously reconstructed PV candidates in the event. The primary vertex reconstruction efficiency is 96% for $H \to \gamma\gamma$, and 99% for $t\bar{t}$ events.

[1] No order of magnitude is given as no results were made public.

3.2.2 The ATLAS calorimeter systems

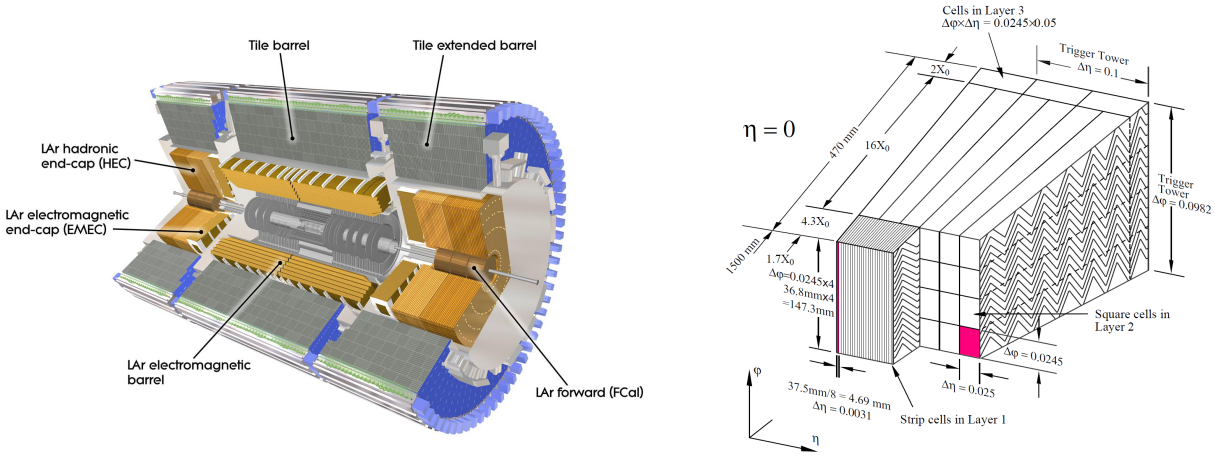


Figure 3.8: Illustration of the ATLAS sub-systems (left). The calorimeter segmentation in the central region is presented in the right figure [152].

The ATLAS calorimetry is separated in one barrel ($|\eta| < 1.475$), and two end-cap regions (Figure 3.8 left). The former is divided in the LAr electromagnetic barrel and the Tile barrel sub-systems. The two are separated by a small gap of 4 mm at $z = 0$. Each end-cap calorimeter includes the LAr electromagnetic end-cap (EMEC), the LAr hadronic end-cap (HEC) and the LAr forward (FCal) subsystems. The HEC calorimeter is situated behind the EMEC, while the FCal is placed near the inner detector.

The LAr electromagnetic calorimeter

The LAr electromagnetic calorimeter is split in two *half-barrels* and two end-cap coaxial wheels covering the $|\eta| < 3.2$ range. As the name suggests, it is a calorimeter with liquid argon as active medium. The absorbers and electrodes have an accordion shape allowing a full coverage in φ. The absorbers are 1.1 - 2.2 mm thick and are made of lead and stainless steel. This combination ensures large electromagnetic (EM) shower shapes for the charge particles at the passage through the calorimeter. The electrode material is cooper and their isolation is assured by polyimide sheets. To collect the signal, a high potential difference is applied between the electrodes and absorbers. This geometry and material composition allows a high granularity and a high $\eta - \varphi$ resolution as needed to identify as precisely as possible the charged particles (and implicitly the photons) crossing and stopping in this calorimeter.

Each half-barrel calorimeter presents a **pre-sampler** situated near the inner detector and covers the $|\eta| < 1.52$ range. It has only one layer of 11 mm depth (and no absorber). It is constructed from 32 longitudinal sections ($\Delta\eta \times \Delta\varphi = 1.52{\times}0.2$), where one sector has a length of 3.2 m and a thickness of 0.28 m. To assure a high and constant η granularity of $\Delta\eta = 0.2$ - unless at the end of the barrel, where it is around 0.12 - it is divided in eight modules. The granularity of each module is $\Delta\eta = 0.025$ and $\Delta\varphi = 0.1$. The role of the pre-sampler is to detect the shower produced by the charged particle interactions with the material in front of the LAr calorimeters (Figure 3.7), in order to correct for the lost electron and photon energies. Moreover, this information is also used to improve the energy measurement in the $1.37 < |\eta| < 1.52$ region, where the amount of material is very high. Note that this region - denoted also by *crack* or the *transition region* - is situated between the barrel and end-cap regions of the detector.

The **barrel region** presents a very fine granularity in φ, as needed for the electron and photon reconstruction and identification. It collects almost all the incident charged particle energy. Each half-barrel has a weight of 57 tonnes, a length of 3.2 m and the inner and outer radii are of 1.4 m and 2 m respectively. It consists of 1024 absorbers and is divided in 16 modules. The thicknesses of one module is at least 22 radiation lengths (X_0), increasing up to 33 X_0 in $0.8 < |\eta| < 1.3$ interval. One module is divided in three layers as illustrated in Figure 3.8 right. The first layer is having the highest granularity in η, providing a precise pseudorapidity measurement and the shower shape variables needed to reconstruct and identify electrons and photons. The latter are key discriminants against the fake electrons, mandatory for the accurate electron and photon reconstruction and identification (Chapters 4 and 5). The second layer accumulates most of the EM shower energy, while the third layer, having a coarser granularity, collects only the tail shower energy. Figure 3.9 (middle column) presents the granularity of each layer, given the η interval.

Compared to the LAr barrel, the **end-cap regions** have a smaller granularity and are positioned at a distance of 3.6 m from the interaction point. They are extending the measurement up to $|\eta| < 3.2$, starting

	Barrel		End-cap	
EM calorimeter				
Number of layers and $\|\eta\|$ coverage				
Presampler	1	$\|\eta\| < 1.52$	1	$1.5 < \|\eta\| < 1.8$
Calorimeter	3	$\|\eta\| < 1.35$	2	$1.375 < \|\eta\| < 1.5$
	2	$1.35 < \|\eta\| < 1.475$	3	$1.5 < \|\eta\| < 2.5$
			2	$2.5 < \|\eta\| < 3.2$
Granularity $\Delta\eta \times \Delta\phi$ versus $\|\eta\|$				
Presampler	0.025×0.1	$\|\eta\| < 1.52$	0.025×0.1	$1.5 < \|\eta\| < 1.8$
Calorimeter 1st layer	$0.025/8 \times 0.1$	$\|\eta\| < 1.40$	0.050×0.1	$1.375 < \|\eta\| < 1.425$
	0.025×0.025	$1.40 < \|\eta\| < 1.475$	0.025×0.1	$1.425 < \|\eta\| < 1.5$
			$0.025/8 \times 0.1$	$1.5 < \|\eta\| < 1.8$
			$0.025/6 \times 0.1$	$1.8 < \|\eta\| < 2.0$
			$0.025/4 \times 0.1$	$2.0 < \|\eta\| < 2.4$
			0.025×0.1	$2.4 < \|\eta\| < 2.5$
			0.1×0.1	$2.5 < \|\eta\| < 3.2$
Calorimeter 2nd layer	0.025×0.025	$\|\eta\| < 1.40$	0.050×0.025	$1.375 < \|\eta\| < 1.425$
	0.075×0.025	$1.40 < \|\eta\| < 1.475$	0.025×0.025	$1.425 < \|\eta\| < 2.5$
			0.1×0.1	$2.5 < \|\eta\| < 3.2$
Calorimeter 3rd layer	0.050×0.025	$\|\eta\| < 1.35$	0.050×0.025	$1.5 < \|\eta\| < 2.5$
Number of readout channels				
Presampler	7808		1536 (both sides)	
Calorimeter	101760		62208 (both sides)	

Figure 3.9: Granularity in $\Delta\eta \times \Delta\varphi$ of the Lar calorimeter [153].

with $|\eta| > 1.375$. In the region filled with dense material, the energy measurement is improved by a LAr pre-sampler positioned between the end-cap and barrel sub-systems. It covers the region of $1.5 < |\eta| < 1.8$. Given the geometry of the end-cap calorimeter, the absorbers are arranged radially with an accordion shape. To ensure the folding angle of 60° - 120°, each end-cap region is divided in two coaxial wheels. The outer wheel covers the $1.475 < |\eta| < 2.5$ interval, while the inner wheel covers the $2.5 < |\eta| < 3.2$ region. The gap region between the wheels (3 mm) corresponds to $|\eta| = 2.5$ and is filled with a non-dense material. One wheel weighs 27 tonnes, has a thicknesses of 63 cm and a radius varying from 33 cm to 2 m. It is divided in eight modules and the number of absorbers and implicit electrodes is 768 (256) in the outer (inner) wheel. The thickness of one module is greater than 24 X_0, apart for $|\eta| < 1.475$ region. It increases up to 38 X_0 in the outer wheel and up to 36 X_0 in the inner wheel. As in the barrel, the precision in the $1.5\ |\eta| < 2.5$ region is ensured by three layers (Figure 3.9, last column). In the $|\eta| < 1.5$ and $2.5 < |\eta| < 3.2$ intervals, only two layers are considered.

The charged particles produced in the proton-proton interactions produce secondary particles when passing through the liquid argon medium. To collect the signal (ionized electrons) in the barrel region and pre-sampler, a high voltage (HV) of 2000 V is applied. The drift gap - the distance between the absorber and the electrode - is 2.1 mm. The granularity in the $\eta - \varphi$ plane is 0.2×0.2 and corresponds to 32 electrodes powered simultaneously (constant HV). In the end-cap region, to assure an uniform calorimeter response, the HV varies between 1000 V and 2300 V, while in the pre-sampler is -2000 V. This variation is due to the alignment of the absorbers and accounts for the dependency of the end-cap wheels radius. The drift gap changes between 2.8 and 0.9 mm in the outer wheel and between 3.1 - 1.8 mm in the inner wheel. The granularity is the same as in the barrel region.

Energy reconstruction in the LAr electromagnetic calorimeter

The ionization signal, when drifting toward the electrodes, produce electric currents proportional to the deposited energy. The signal produced by each electrode is summed and characterized by a triangular shape. It has around 1 ns rise time and around 460 ns decay (drift) time in the barrel, while in the end-cap varies with the cell gap size. The drift time is much longer than the time between two crossing. This signal is amplified and shaped, sampled every 25 ns in 32 points and digitized. Given the large amount of data obtained during the p-p collisions, only the first five points defining the shape signal are retained (and sent via the optical fibers to the read out driver boards). The difference in the predicted signal shape and the measured signal shape is found to be less than 4%. The cell energy is computed from the digitized signal and accounts for the noise in the cell :

$$E_{cell} = F_{\mu A \to \mathrm{MeV}} \times F_{DAC \to \mu A} \times \left(\frac{M_{phys}}{M_{cell}}\right)^{-1} \times G \times \sum_{j=1}^{5} a_j (s_j - p) \tag{3.1}$$

The first term, which is computed from test-beam studies, is converting the ionized signal into a total energy deposited in the EM calorimeter, at the EM scale, while the second term converts the DAC (digital

to analog converter) signal into an electrical current (μA) respectively. The $(M_{phys}/M_{cell})^{-1}$ fraction of the maximum of the physical and calibration pulses is a correction to the electronic gain factor (G). In the sum over the number of points the following variables are used : the samples of the shaped signal (digitized in the selected electronic gain s_j), the read-out electronic pedestal (specific for the considered electronic gain and computed in specific calibration runs, p) and the a_j weights. The latter are the optimal filtering coefficients and are derived from the predicted ionization signal and accounts for the noise present in the cell.

The timing stability is essential for the energy measurement : the read-out clock of the LAr cell must correspond to the LHC bunch crossing (difference smaller than 1 ns). This ensures a correct energy measurement for each event. Compared to 2010 and 2011 data taking, the 2012 data quality was improved. It is due to the reduction of HV trips and noise bursts. The former appears when the power supply cannot maintain a constant voltage, while the latter are due to electronic noise in the calorimeter affecting several channels. The overall fraction of operational channels during 2012 run is 99.9% [155].

LAr electromagnetic calorimeter performance

The fractional electromagnetic calorimeter energy resolution is parametrized as :

$$\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \tag{3.2}$$

where a and b are the sampling (stochastic) and noise parameters and c is a constant. The a term is around 10% and reflects the sampling frequency. The electronic noise term b is around 200 - 400 MeV and affects the calorimeter performance at energies below 20 GeV. Compared to the previous parameters, the constant term c is more sensitive to the detector construction and operation. The main contributions are coming from a local term (associated to the cluster size, accordion shape, etc), module mechanics (material, distance between absorbers, etc), temperature inhomogeneities... It is retained below 0.7%.

The computation of the energy deposited in the calorimeter cells must account for several effects. Given the accordion dependency, the amount of material crossed by the particle is φ dependent. This induces a φ modulation of the energy. Moreover, the calorimeter granularity is finite and the shower is spread in many cells, leading to a bias in the particle position. Also the z and η position dependency should be considered. After taking into account these effects, the constant term for one cell is reduced to 0.43% [161]. Studying an energy range of $10 < E < 180$ GeV, the reconstruction energy response is found to be linear (within $\pm 1\%$). The overall constant term is measured to be 0.43% in the barrel and around 0.62% in the end-cap [161]. Using 2010 data to measure the electron energy resolution, the overall constant term is 1.2% $\pm$ 0.6 in the barrel region and 1.8% $\pm$ 0.6% in the end-cap [162]. The energy response stability with time and with the background given by the overlap of multiple interactions is linear, within $\pm 0.1\%$ [163, 164]. The resolution of the φ polar angle (or the axis direction of the EM shower development) can be precisely measured given the high granularity of the LAr electromagnetic calorimeter. The resolution varies between 50 - 75 mrad/$\sqrt{E}$ [153].

Hadronic calorimeters

This calorimeter complements the measurement of the energy deposited by the jets and localize the particles. Compared with the electromagnetic calorimeter the resolution is poorer. It is composed by the tile, LAr hadronic end-caps and LAr forward calorimeters. As illustrated in Figure 3.8, the **tile calorimeter** is situated behind the LAr electromagnetic barrel calorimeter and is formed by a central (barrel) region and two extended barrel regions. The barrel is having a length of 5.8 m ($|\eta| < 1.0$), while each extended region is 2.6 m ($1.0 < |\eta| < 1.7$). The tile calorimeter is situated at a distance of 2.28 m from the inner detector and has a radius of 4.25 m. The absorbers are made of steel and the active medium are scintillators made of polystyrene doped with fluor. The tile calorimeter is formed by 64 modules (the module structure is shown in Figure 3.10) and divided in three layers. The absorbers are 5 mm thick and separated by a 4 mm space plates hosting the scintillator tiles. The tiles

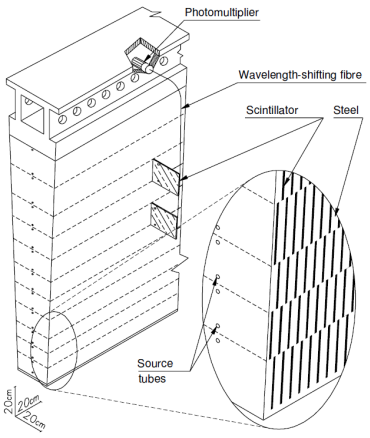


Figure 3.10: Illustration of the tile calorimeter structure [153].

are 3 mm thick and 200 - 400 mm long, with a radius varying between 97 and 187 mm. When a charged particle passes through the calorimeter, the light produced in the scintillator is propagated via optical fibers to the photo-multiplier. Two optical fibers are connected at the edge of a tile and several of them at one photo-multiplier. This assures a 3D dimensional cell structures of dimension $\Delta\eta \times \Delta\varphi = 0.1{\times}0.1$ in the first layer or $0.2{\times}0.1$ in the last layer. A total of 9852 channels (5760 in the central region and 4092 in the extended region) are formed, given the 460000 scintillating tiles. During 2012 data taking, around 98.3% channels were operational [155]. The efficiency of the data acquisition is estimated to be around 99.6% [156].

In end-cap regions the **LAr hadronic end-caps** or HEC sub-systems are used. They are covering the $1.5 < |\eta| < 3.2$ region. Similarly to the LAr electromagnetic end-cap, it is formed by two coaxial cylindrical wheels with an inner radius of 20.3 cm. Each wheel is divided in two longitudinal sections and is constructed of 32 identical modules. For the front wheel the modules are made of 24 copper plates (22 mm thick), while for the the other wheels only 16 plates (50 mm thick) were used. The distance between the plates is around 8.5 mm, contains three electrodes and is filled with liquid argon. When a minimum signal of 1800 V is applied the electron drift time is around 430 ns. This configuration leads to a number of 5632 channels and cells of dimension $\Delta\eta \times \Delta\varphi = 0.1{\times}0.1$ ($0.2{\times}0.2$) in the $|\eta| < 2.5$ (otherwise) region. During 2012 data taking, the fraction of operational channels was 99.6% [155].

The **LAr forward** calorimeter is extending the coverage in the $3.2 < |\eta| < 4.9$ region and is divided in 3 layers. It is located in the same cryostat with the end-cap calorimeters, just near the inner detector at a distance of 4.7 m from the interaction point. This arrangement minimizes the energy loss in the crack regions and the level of background in the muon spectrometers. Each Fcal sub-system is divided in one electromagnetic (FCal1) and two hadronic (FCal2 and FCal3) modules, with a depth of 45 cm. The active medium is liquid argon and the absorbers are either copper in FCal1, or tungsten otherwise. The electrodes are made of copper and grouped in sets of four, six or nine placed near (FCal1) or away of the interaction point. A total of 1762 channels are obtained in the FCal calorimeter. Around 99.8% channels were functional during 2012 data taking [155].

Noise in the calorimeter cells and topological clusters

As described in Section 4.5, the calorimeter jets are formed by 3D topological clusters built from the calorimeter cells, with positive energy. The clusters follow the cell signal significance patterns in the ATLAS calorimeters, measured as the absolute ratio between the cell signal and the energy equivalent noise in the cell. The ratio is $|E_{cell}|/\sigma_{noise} > 4$ for the seed, > 2 for the neighbors and > 0 for the bound. The total cell noise is the sum in quadrature of the cell noise due to readout electronics and the pile-up. It is dominated by the pile-up noise, represented by the additional low transverse momentum pp collisions accompanying the signal (hard-scatter pp interactions). It is divided in out-of-time and in-time pile-up. The former arise from additional pp interactions in the current bunch-crossing. The latter refers to the energy deposited in the calorimeter from the previous and following bunch-crossing, relative to the triggered event. The pile-up noise is measured as the total width of energy distribution per cell for a given bunch spacing, Δt, and for a given number of expected pp collisions per bunch, $< \mu >$:

$$< \mu >= \frac{L \times \sigma_{inel}^{pp}}{N_{bunch} \times f_{LHC}} \tag{3.3}$$

where L is the luminosity in the considered time period, σ_{inel}^{pp} is the cross section of inelastic pp collisions (71.5 mb), N_{bunch} is the number of colliding bunches and f_{LHC} is the LHC revolution frequency.

Each LAr calorimeter cell is sensitive to pile-up and the use of a fast signal shaping leads to a cancellation of in-time and out-time pile-up in a given cell. Given the run conditions and the used methodology, a full pile-up suppression is not possible, especially in the central region of the calorimeter. To reduce the remaining pileup, a pile-up correction method is applied to reconstructed jets (Section 4.5). The Tile calorimeter shows very little sensitivity to pile-up, as the LAr calorimeter absorbs most of the energy associated to the soft produced particles.

As the pileup contribution could not be entirely suppressed, the value of σ_{noise} is adjusted to prevent an increase in the number of clusters due to too large pile-up while retaining as much signal as possible.

Performance of the hadronic calorimeter

To study the performance [153] of the hadronic calorimeter, some of its sub-systems were exposed to muons, electrons and pion beams with energies up to 200 GeV. The energy resolution is quantified using Equation (3.2). The results obtained with the pion beams yield to a stochastic term of 56.4% $\sqrt{\text{GeV}}$ and a noise term of 5.5%. When using the electron beam and the end-cap calorimeter, the a term is found to be 21.4% $\sqrt{\text{GeV}}$ and b around zero. The results obtained after analyzing the pion beams lead to values of a of 70.6% $\sqrt{\text{GeV}}$ and for the constant term of 5.8%. The performance of the FCal calorimeter is

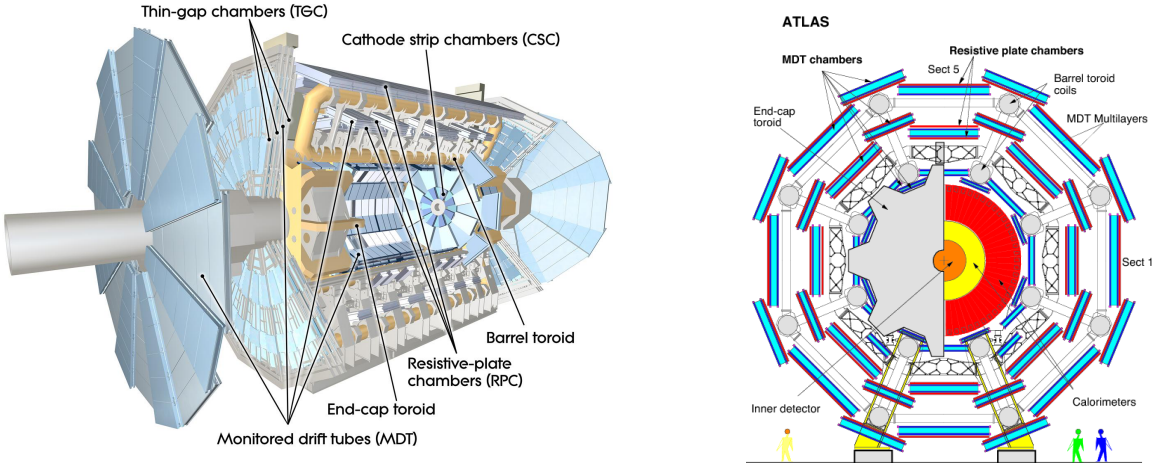


Figure 3.11: Schematic view of the muon spectrometer [152, 153].

studied with electron and pion beams at energies varying from 10 to 200 GeV. The obtained values for the parameters after analyzing the electron beams are : 28.5% for a and 3.5% for c. When muon data is analyzed, with improved methods, the resolution is better (from 10% to 5.8%) : the stochastic term is 70% and the constant term is 3%.

The expected resolution of the hadronic calorimeter is 50% for the a term and 3% for the constant term. To improve the results obtained only with the information contained in the hadronic calorimeter, the electromagnetic calorimeter was also considered. This is well motivated as the particles stopped by the hadronic calorimeter are depositing a fraction of their energy also in the electromagnetic sub-systems. The combined information leads to a stochastic term of 52%, to a noise term of 1.6 GeV and to a constant term of 3%.

3.2.3 Muon spectrometer

The muon spectrometer is designed to measure the momentum and identify muons tracks with a high precision - the design resolution is around 10% for 1 TeV tracks. It covers the $|\eta| < 2.7$ range, while for triggering purposes only the $|\eta| < 2.4$ interval reached. The measurement can be performed for muons with a momentum varying between 3 GeV and 3 TeV. The spectrometer is formed by a magnetic system and several tracking chambers.

The **magnetic system** (Figure 3.12) is composed of three toroids : one barrel ($|\eta| < 1.1$), and two end-caps ($1.1 < |\eta| < 2.7$). The barrel toroid is 25 m long with an inner (outer) radius of 9.4 m (20.1 m), and is formed by eight superconducting coils grouped in a torus shape. Each coil has an area of 5×26 m^2. Each end-cap toroid is placed inside the barrel toroid, at the end of the central solenoid. It is 5 m long, the inner (outer) radius is 1.65 m (10.7 m) and is made of 8 coils. Like the central solenoid, the barrel and end-cap supermagnets are cooled using helium (4.8 K). They are operated at a current of 20.5 kA and the generated magnetic field can reach 4 T.

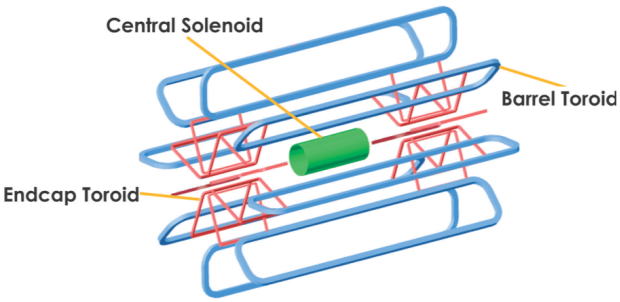


Figure 3.12: Magnet system [152].

The **precision tracking chambers** are organized in three stations, one in the barrel region situated between and on the magnet toroids and two on the end-caps. The latter are situated in front and behind the magnet system. The chambers in the barrel are arranged in cylindrical layers around the beam axis at radii of 5, 7.5 and 10 m. In the end-cap regions, the tracking chambers form large wheels perpendicular on the beam axis, at radii of 7.4, 10.8 and 14 m, and situated at 21.5 m from the interaction point (Figure 3.11). The **monitored drift tube chambers** (MDT) are used for the momentum measurement in $|\eta| < 2.7$ range, unless in the deepest end-cap layer where the range is restricted to $|\eta| < 2.0$. The structure of the chambers is rectangular in the barrel and trapezoidal in the end-cap. They are formed by 30 - 70 aluminum drift tubes with 3 to 8 layers (diameter of 3 cm and length varying from 1 to 6 m). A total of 1088 chambers (350000 tubs) are defining the spectrometer. They operate with $Ar + CO_2$ gas and are retained at a pressure of 3 bars. The applied potential is 3080 V and the maximum drift time is 700 ns.

Such long times arise from the used gas mixture. The achieved resolution is of ∼80 μm per tube (∼35 μm per chamber), and is strongly dependent on the drift distance. From the total number of channels, a fraction of 99.7% were operational during the 2012 data taking, while efficiency of the data acquisition was 99.6% [155, 156].

The charged particles momentum in the deepest layer of the end-cap spectrometer ($2.0 < |\eta| < 2.7$) is measured using the **cathode-strip chambers** (CSC). Each CSC is segmented, in the φ plane, in 16 large and small chambers with strips with width of 1.5 mm and 1.6 mm respectively. It is located at a distance of 7 m from the interaction point and the drift time is only 3.6 ns. Compared to the MDT chambers they have a better resolution, around 40 μm and 5 mm in the bending and transverse plane respectively. This systems is built by two disks, formed by eight chambers with four planes per chamber. This translated in four independent measurements in $\eta - \varphi$ for one track. Form a total of 31000 channels, only 96% were operational during the 2012 data taking, while the efficiency of data acquisition was 100% [155, 156].

The **alignment of the MDT and CSC precision tracking chambers** is playing an important role in particle position identification and momentum resolution. In a first instance, the chambers are installed with a precision of 5 mm - 2 mrad. However, the precision on the chamber location reconstruction must be smaller than 30 μm. To reach this, an alignment system based on optical elements is designed.

The **trigger chambers** ensures the muons detection with $p_T > 6$ GeV in the $|\eta| < 2.4$ and full φ range, by reconstructing the tacks (approximately) pointing to the interaction point. This is performed using two different technologies, because the radiation level is 10 times higher in the end-cap than in the barrel. Therefore, in the central region ($|\eta| < 1.05$) resistive plate chambers (RPC) are used, while in the end-caps are thin gap chambers (TGC). The RPC is divided in three concentric layers (trigger stations) around the beam axis, at radii of 7.8, 8.4 and 10.2 m. Each layer is a made of sets of two parallel plastic resistive plates covered with copper based bands (the electrode). The distance between the 2 plates is of 2 mm (filled with gas) and the electric field is around 49000 V/mm. When the charged particles interacts with the RPC *detector* they produce a signal of 5 ns. This arrangement allows the selection of muons with p_T varying from 6 to 35 GeV. If a particle is passing through all the three stations, its track will have six associated hits in the $\eta - \varphi$ plane. A total of 370000 channels are characterizing the RPC *detector*. During the 2012 data taking, a fraction of 97.1% channels were operational (the 2012 data acquisition efficiency was 99.8%) [155, 156]. The TGC are multi-wire chambers, with a wire - cathode distance (1.4 mm) smaller than the wire - wire distance (1.8 mm). This leads to a very high resolution for most of the tracks, as required by the end-cap geometry (trigger layers outside of the magnetic field, 10 times larger radiation levels than in the central region, etc). From a total of 320000 channels, 98.2% were efficient during the 2012 data taking [155]. The efficiency to acquire the 2012 data was 99.6% [156].

Performance of the muon spectrometer

After analyzing the 2011 and 2012 data, using $Z \to \mu\mu$ events, the muon momentum resolution varies between 1.7% in the central region and $p_T < 10$ GeV. In the large η region and $p_T > 100$ GeV it is around 4% [165].

3.2.4 Forward detectors

Beside the mentioned detector sub-systems, the ATLAS detector has three additional detectors aiming at a good coverage of the very forward rapidity region : LUCID, ZDC and ALFA. The **LUCID** detector – LUminosity measurement using Cerenkov Integrating Detector – is dedicated to monitoring online instantaneous luminosity and to measuring the integrated luminosity with a precision better than 5%. It is located at a distance of ±17 m from the interaction point and covers $5.6 < |\eta| < 5.9$ range. The detector must record all the particles resulted from the p-p inelastic interactions, as this is a measure of the total number of interactions needed for the luminosity computation. To achieve this, the detector is formed by an array of 20 Cerenkov aluminum tubes (cones) of 1.5 m length, surrounding the beam pipe and pointing to the interaction point. Each tube has a diameter of 15 mm and is filled with C_4F_{10} gas maintained at a pressure of 1.2 - 1.4 bar. This assures a Cerenkov threshold of 10 MeV for electrons and 2.8 GeV for pions. When the charged particles cross the tube, the emitted Cerenkov cone light at 3° is measured by photomultiplier tubes (15 mm diameter). A single bunch crossing detection is possible giving the fast timing response (few ns) of the LUCID detector.

The **ALFA** detector - Absolute luminosity for ATLAS - aims to measure the absolute luminosity via elastic scattering at small angles (3 μrad). These angles can be reached using a high-beta (β^*) optics and using special beam conditions as the scattering angles are larger than the nominal beam divergence. The measurement is based on the Roman Pot technique and implies its emplacement very close to the interaction point. Therefore, four Roman Pot spectrometers[2] (two in both sides of the interaction point)

[2] A Roman Pot is a detector volume, separated from the vacuum by a thin window, which can be moved very close to the

are placed at a distance of ±240 mm from the interaction point, at a distance of 1 mm from the beam. The 1500 scintillating fiber trackers included in the 10 double-sided modules are assuring a spatial resolution of 30 μm. Another essential requirements are the minimal sensitivity to the radio frequency noise and the ability to work at the vacuum present in the Pot. The LUCID detector is calibrated with the ALFA detector, allowing a precision on the luminosity measurement better than 5%.

Another forward detector is **ZDC** - Zero Degree Calorimeters - providing a coverage of the $|\eta| > 8.3$ region, for the neutral particles (in particular neutrons and photons). It is located at ±140 m from the interaction point and formed by 2×4 modules : one electromagnetic (EM) and three hadronic. The EM module (29 X_0 deep) is made of 11 tungsten plates, parallel to the beam pipe, with their height extended by steel plates. The tungsten plates presents 96 quartz rods, each connected to one pixel of the multi-anode phototube, capturing the Cerenkov light produced by the incident particles. The rod(s) containing the signal gives the position of the incident particle, while the intensity of the light reflects its energy. To obtain another improved measurement, between the tungstens and steel plats some quartz strips are positioned. They are connected to the photomultiplier tubes by optical fiber. The three hadronic modules are very similar to the EM ones. The main difference is the methodology used to connect the rods with the photomultiplier. In the latter case, the clusters of 4 rods are mapped into individual phototubes and they have only one optical fiber. During the proton-proton collisions, the ZDC detector will enhance the acceptance for the diffractive processes (needed for the MC tunning) and also will provide additional mini-bias trigger for ATLAS. When the LHC will reach luminosities above 10^{33} cm^{-2}s^{-1}, the the detector will be removed for proton-proton collisions and installed only for the heavy-ion data taking.

3.2.5 Trigger system and data acquisition

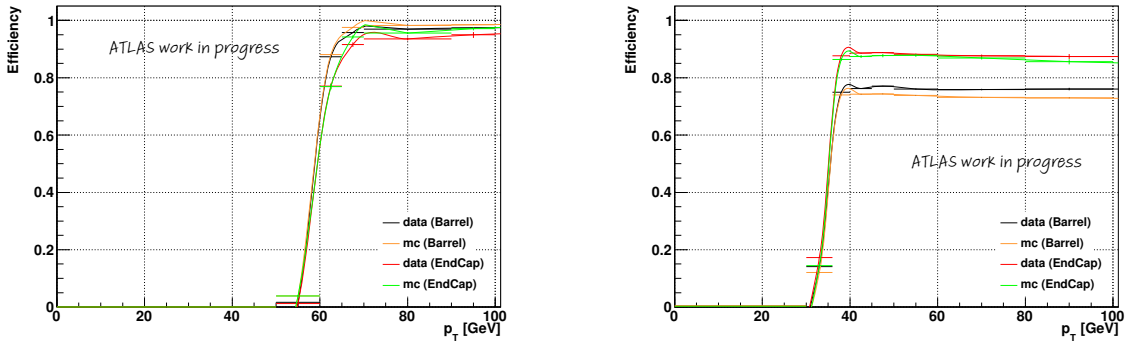


Figure 3.13: Efficiency of `Eff_EF_e60_medium1` (left) and `Eff_EF_mu36_tight` (right) triggers in data and MC. The results are shown separately for the barrel and end-cap regions of the ATLAS detector [166].

Given the very high bunch crossing rate, of 20 MHz during the 2010 - 2012 data taking (but it is designed to reach 40 MHz), it is not possible to record all the produced events. Therefore, a dedicated trigger system is designed to record events at a rate of 400 Hz and is based on the physics objects identification. It has three levels : L1, L2 and event filter (EV) - the last two are forming the High Level Trigger (HLT).

The Level1 trigger is selecting events at a rate of 75 kHz using the information from the calorimeters (L1Calo) and muon spectrometers. The L1Calo is formed by 7000 analogue trigger towers with a granularity of $\Delta\eta \times \Delta\varphi = 0.1 \times 0.1$ and identifies electrons, photons, jets or hadronic taus with a transverse momentum higher than a given value. It has a compact architecture, ensuring a very fast response (latency of 2.1 μs). The L1 muon trigger is relaying on the RPC and TGC detectors, characterized by a fine granularity. A total of 256 distinct L1 trigger items, with several selection criteria on the p_T, lepton and jet flavor or isolation, are defining the L1 trigger system. All the selected events are transmitted to the central trigger processor in less than 2.5 μs. Beside the first event selection, the L1 triggers identifies regions of interest (RoI), used by the HLT system. These zones, defined in the $\eta - \varphi$ plane of the ATLAS calorimeters and spectrometers, can be regarded as regions with interesting features from where the event was selected.

If L1 is a hardware implemented in some given requested electronics, the HLT is a software based trigger system using a lot of computer clusters. The L2 trigger selection is performed using all (and only) the data from the RoI and all the detector sub-systems information (i.e. the inner detector is included). This reduces the time needed to perform the interesting event selection. Considering the entire detector granularity and precision, the L2 events are sorted until the trigger rate is 35 kHz with a processing time of 40 ms. This reduction is possible due to a set of complex reconstruction algorithms. If the events is

beam pipe.

kept, the data information is sent to the event filter. Here the event is built using a the standard (on-line Athena) ATLAS reconstruction algorithm. It has the same selection principle as L2, at a filtering rate of 400 Hz and an event processing time around 4 seconds.

Trigger performance

To reflect the very good ATLAS trigger performance, in Figure 3.13 the efficiency obtained for two main triggers used in the search for new physics is shown. More details are given in Chapter 6.

3.3 Conclusions

The great success of LHC experiment was reflected in the discovery of the Higgs boson and other baryonic states, as well as the multitude of the Standard Models and Beyond the Standard Model papers. Concerning the ATLAS detector, its performance is given by the multitude of systems designed to achieve the goals of the experiment under unprecedented amount of radiation. One important characteristic is the high precision reached for the particle energy, momentum, position, etc. The methodology to reconstruct these parameters is presented in Chapter 4.

Chapter 4

Object reconstruction in ATLAS

In this chapter the object reconstruction and identification in ATLAS are presented. More information is provided for the objects used in this thesis to measure the electron reconstruction and identification efficiencies and to search for new physics in events with two same - sign leptons, several jets and missing transverse energy. The chapter is organized as followings. The electron object reconstruction and identification are presented in Section 4.1. For the electron identification both cut-based and multivariate analysis techniques are briefly described and the proposed identification menus are mentioned. The techniques used to reconstruct and identify the photon objects are detailed in Section 4.2 and for muons in Section 4.3. The tau leptons are briefly mentioned in Section 4.4. The jet reconstruction and calibration, together with the associated systematic uncertainties, is presented in Section 4.5. In the same section, the algorithms used to tag the b jets are described, as well as their efficiencies and the probabilities to mistag a c or a light jet as a b jet. The chapter closes with the missing transverse momentum reconstruction (Section 4.7) and some conclusions.

4.1 Electrons

Electrons are reconstructed and identified for $p_T > 5$ GeV in $|\eta| < 4.9$ range of the detector [154,167,168]. These measurements are challenging at LHC since the pp collisions are dominated by jets which could fake an electron. More precisely, the background electrons arise from heavy flavor or Dalitz decays, misidentified hadrons, photon conversions from pion decays, etc. Events of this type are highly reduced with the available electron reconstruction algorithm or by imposing tight cuts during the electron identification, as shown later in this section. Compared to $\sqrt{s} = 7$ TeV, the pileup dependency is well diminished by the improvements brought to the tracking algorithm [160] and electron identification menu (Section 5). Only the information regarding the electrons in the central region of the detector ($|\eta| < 1.5$) is presented, as the forward electrons are not used in this thesis.

4.1.1 Reconstruction

In ATLAS, in order to reconstruct the electrons in the $|\eta| < 2.5$ region seed clusters are built in the electromagnetic calorimeter and matched to the reconstructed tracks in the inner detector, using the `AuthorElectron` algorithm. A seed cluster is defined as a 2D window of 3×5 cells ($\Delta\eta\times\Delta\varphi = 0.025\times0.025$) in the η and φ space in the EM layers. Only clusters with energies above 2.5 GeV are considered. To remove the (duplicate) clusters which are close to each other, the energy in the 2D windows is compared and only the cluster with the highest energy is retained. The clusterization methodology is found to be very efficient and the duplicate cluster rates are at sub-percent level [167]. The cluster reconstruction efficiency is computed using simulated samples, after applying a cut on the angular separation between the generated and reconstructed electrons. For real electrons[1] it is found to be around 95% for electrons with $p_T \sim 7$ GeV, 99% $p_T \sim 15$ GeV and up to 99.9% for $p_T > 45$ GeV. If a match between the generated and reconstructed energy is added, the efficiency drops in $1.5 < |\eta| < 1.65$ due to the large amount of material in front of EM calorimeter [169].

The electron trajectory is determined using information from the inner detector, like for the trajectory of any other charged particle. This involves the measurement of the track associated parameters : the position in the transverse (d_0) and longitudinal (z_0) planes of the perigee, the particle direction (φ, θ) and the parameter which provides the inverse track momentum multiplied by the charge (q/p). The track parameters and the associated uncertainties are obtained from the track fitting procedure performed with

[1] Electrons arising from Z, W, etc.

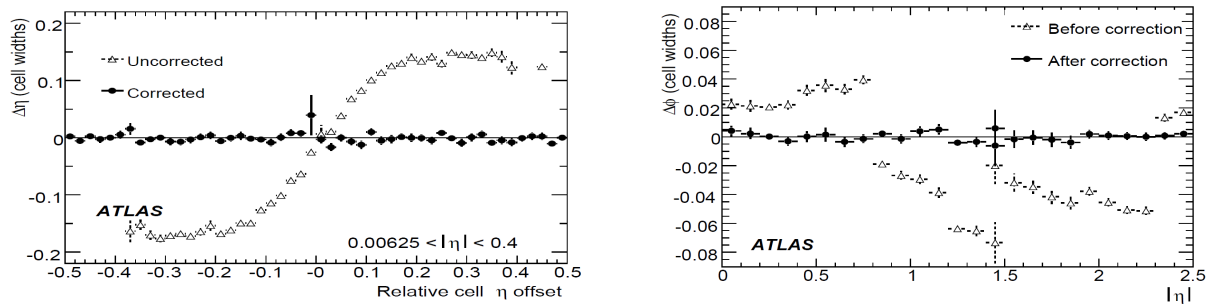


Figure 4.1: The bias on the η position reconstruction as a function of the impact point of the incident particle in the layer one barrel cells (left). In the right plot, the bias on the φ position measurement as a function of $|\eta|$ is presented. Distribution are shown before and after applying the corrections [154].

the ATLAS Global χ^2 Track Fitter [170]. Within this process all the tracks with $p_T > 400$ MeV and $|\eta| < 2$ are considered. An electron is reconstructed only if at least one reconstructed track in the ID is associated to the seed cluster. For the matching, the tracks are extrapolated from the last measurement point to the second EM layer. If the difference in η and φ positions between the impact point and the seed cluster position is lower than a typical value of 0.01, the matching is considered to be successful.[2]. Once it is performed, the electron cluster is rebuilt using towers of cells of dimension 3×7 and 5×5 second sampling cells (middle layer of the EM calorimeter) in the barrel and end-cap regions. The dimensions of the cell are optimized to account for the different geometry and granularity in the two regions of the EM calorimeter.

The tracks attached to the EM clusters are classified with respect to their quality (i.e. at least four hits in the SCT) and distance to the EM cluster in the η - φ plane (ΔR). High quality tracks are chosen to reject electrons coming from photon conversions. Electrons from photon conversion are found to be below the percent level as measured for $Z \to ee$ and $J/\varphi \to ee$ events and further reduced by applying the electron-photon overlap removal (Section 5.1.2). During the electron identification the two *collections* are well separated.

To improve the electron reconstruction efficiencies, a new algorithm Gaussian Sum Filter (GSF) [160] was implemented starting in 2012. It aims to recover the energy loss due to bremsstrahlung processes in the detector regions with large amount of material. Using all tracks already associated to the cluster electrons, around each seed electron cluster a region of interest (RoI) with a cone $\Delta R = 0.3$ is built. If no track is found, within this RoI the GSF is applied to recover tracks with a large momentum loss. The resulting tracks are corrected for the bremsstrahlung and together with the corresponding electromagnetic clusters are used as inputs to the `AuthorElectron` algorithm. A higher and a more uniform electron reconstruction efficiency over η is obtained, which varies between 95% and 99% depending on η and p_T range. More details are given in Chapter 5 where the electron reconstruction efficiency measurement is described and the obtained results are presented.

4.1.2 Calibration

To study and compute the corrections needed to account for the detector geometry and material distribution two sets of MC samples are used. For the first set an ideal geometry (no misalignments) with the best knowledge of the dead material is implemented. For the second scenario, the dead material between the tracker and calorimeters is increased and the misalignments are included. The latter is used to assign the systematic uncertainties.

Position measurement

The electron cluster position (η, φ) is calculated as the energy-barycenter of all cells in the cluster, using the information from the first and second layers of the EM calorimeter. Given the LAr calorimeter accordion geometry and the finite granularity, several correction are considered. These are applied separately for each layer and for the barrel and end-cap regions of the detector. The η position is estimated in each cell of the cluster as a function of the impact point of the incident particle and is biased toward the center of the cell as illustrated in Figure 4.1 left (open marker). This is simply because the cells have a finite granularity and the shower energy, which is not fully inside the cell, is computed with respect to the cell center. To account for this bias, an η correction is applied. Due to the detector geometry, the correction is computed

[2]If the tracks are not presenting silicon hits, the matching is performed only as a function of the φ coordinate (because the η coordinate measured in the TRT is not precisely known) and it requires only a rough match, i.e. the track in the barrel should match the cluster in the barrel.

for several η regions and depends on the distance from the interaction point of the incident particle to the center of the cell. It is described by the following function :

$$\Delta\eta(\nu) = A(\eta)\tan^{-1} B(\eta)\nu + C(\eta)\nu + D(\eta)|\nu| + E(\eta) \tag{4.1}$$

The distance ν can vary from $-\frac{1}{2}$ to $+\frac{1}{2}$ and the A, B, ... parameters are polynomials of the form :

$$A = \sum_i a_i|\eta|^i \text{ and similarly for B, C, ..} \tag{4.2}$$

The correction parameters, a_i, b_i, etc, are determined for several cluster size and energies. Once the corrections are considered, the bias is completely removed (Figure 4.1 left, filled marker). For the φ position this effect is lower, as accordion geometry leads to more energy sharing between the cells. Therefore, the bias is small as shown in Figure 4.1 right (open marker). It can be seen that, after applying the corrections, the bias is removed. Once the corrections are applied, the *per layer* measurements are combined to obtain the final (η, φ) position of the electron cluster.

Energy measurement

Most of the incident electron energy is deposited in the clusters built in the EM calorimeters and calculated as shown in Section 3.2.2. However, a small energy is deposited in the non-instrumented section, as the inner detector, cryostat, solenoid, cables, etc. To account for these effects, the energy cell is calibrated [171]. In Figure 4.2 the steps followed by the calibration procedure are presented. Firstly, the EM cluster energy

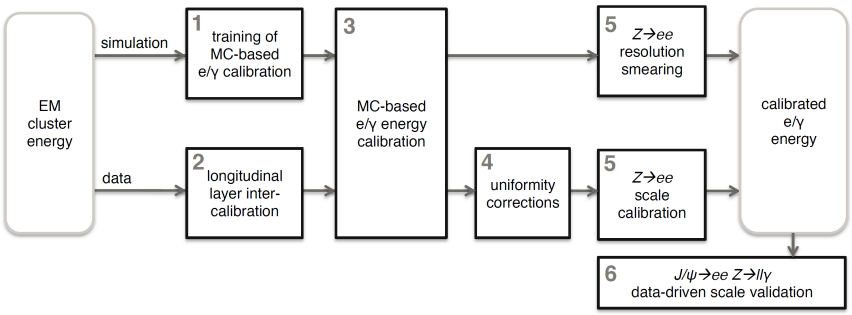


Figure 4.2: The steps used in the electron energy calibration [171]. This procedure is valid also for photons.

is calibrated to the initial electron energy using the special MC samples described above. This implies the usage of a multivariate analysis trained on MC as well as the inter-calibration of the scales of the first and second longitudinal EM calorimeter layers. The former ensures an accurate description of the detector geometry and electron interactions, using as input parameters the total energy measured in the calorimeter, the ratio between the energy in the presampler and calorimeter, the shower depth ($X = \sum_i X_i E_i / \sum_i E_i$, where X_i is the calorimeter thickness and E_i the cluster energy in the layer i) and the cluster barycenter in η and φ. The inter-calibration of the different calorimeter layers - including the presampler - is mandatory as the same scale in data and MC should be ensured prior to the overall energy calibration. An additional set of corrections are imposed to account for the different detector response non-uniformity, not included in the calibration. Once the cluster energy is calibrated, the overall electron energy resolution measured in data is adjusted as a function of η, using $Z \to ee$ events (resolution smearing). The correction is obtained using the assumption that the resolution curve is well modeled in MC up to a constant term :

$$\left(\frac{\sigma(E)}{E}\right)^{data} = \left(\frac{\sigma(E)}{E}\right)^{MC} \oplus c \tag{4.3}$$

where c is the scale correction (as a function of η). As $Z \to ee$ events are considered, the E term is replaced by the invariant mass of the 2 electrons (m_{ee}), which depends on the η value of each electron. In the last step, to account for the smaller resolution in data, additional scale factors are applied to MC electron energies to match the data (scale calibration). They are obtained using the relation $m_{ee}^{data} = m_{ee}^{MC}(1+\alpha)$, where α is the deviation from the optimal calibration. Note that, similarly to the resolution smearing, the invariant mass distribution is used to compute the correction and the η dependency is considered. The goodness of the electron calibration is shown in Figure 4.3.

The central values of these two parameters are α = -0.0002 $\pm$ 0.007 and c = 0.0197 $\pm$ 0.0023. Two sets of uncertainties are assigned and later used in the two analyses searching for SUSY signatures in events with two same sign leptons ($\sqrt{s}$ = 8 TeV) :

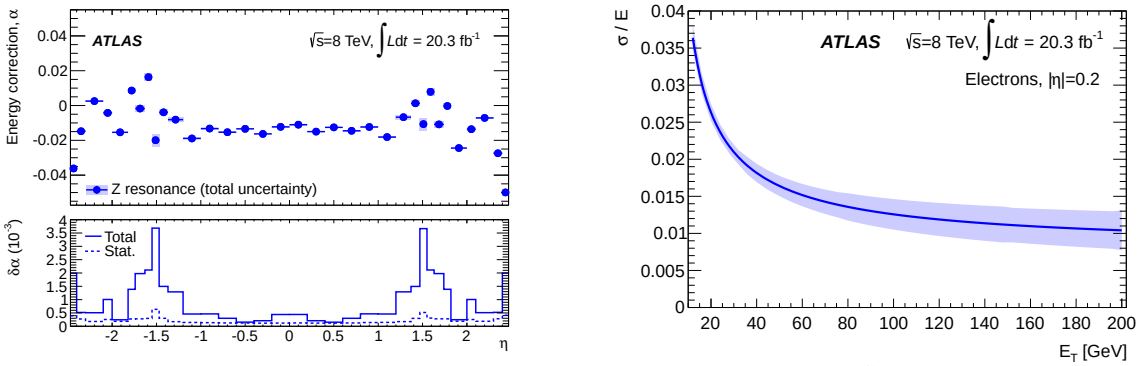


Figure 4.4: Top : energy scale corrections, α defined as $E^{data} = (1+\alpha)E^{MC}$ and derived from $Z \to ee$ events are shown as a function of η. The uncertainty bands includes statistical and systematic uncertainties. In the bottom panel the statistical and total energy scale uncertainty as a function of η are shown. Right : reconstructed electron resolution curve and its uncertainty as a function of p_T [171].

- The uncertainty on the electron energy scale is on average 0.04% in the $|\eta| < 1.37$ and $|\eta| > 1.82$ regions and around 0.2% in the $1.37 < |\eta| < 1.82$ region when considering electrons with p_T ∼40 GeV. For less energetic electrons ($p_T \sim 11$ GeV) the uncertainty increases to 1.1% in the $|\eta| < 1.82$ interval and is around 0.4% for $|\eta| > 1.82$ (Figure 4.4, left).
- The uncertainty on the electron energy scale resolution is better than 10% for electrons with $p_T < 50$ GeV and asymptotically increases up to 40% at higher electron energies (Figure 4.4, right).

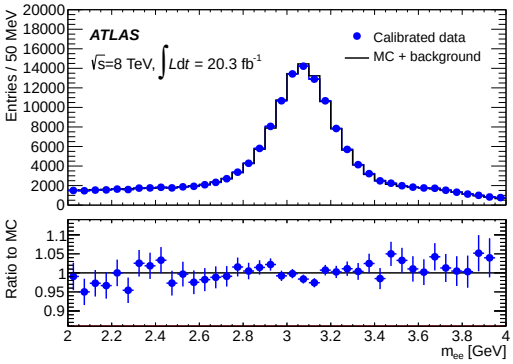


Figure 4.3: Performance of the electron calibration using $J/\Psi \to ee$ events [171].

4.1.3 Identification

The reconstructed electrons can be isolated[3] or fakes from hadronic jets, heavy flavor decays, photon conversions, Dalitz decays etc. To reject as many as possible of these fake electrons, while keeping the signal electron efficiency as high as possible, a series of discriminant variables is used [168, 172] :

- Hadronic calorimeter :
  - $R_{had} = \frac{Ehad}{p_T}$ and $R_{had1} = \frac{Ehad1}{p_T}$ (Figure 4.5 left), where $Ehad$ is the transverse energy in the hadronic calorimeters behind the cluster, $Ehad1$ is the transverse energy in the first layer of the hadronic calorimeter behind the cluster and p_T is the ratio between the EM cluster energy and $cosh(\eta_2)$, with η_2 defined as the pseudorapidity of the cluster in the second EM layer.
- First layer of the electromagnetic calorimeter :
  - $f1 = \frac{E_1}{E}$: fraction of energy reconstructed in the first EM layer, where E_1 is the energy summed over all strips of the cluster and E is the total reconstructed energy in the EM calorimeter cluster.
  - E_{ratio} : ratio of the energy difference between the largest and second largest energy deposits in the cluster and the sum of these energies (Figure 4.5 middle).
  - $W_{stot} = \sqrt{\sum_i E_i \cdot (i - i_{max})^2 / \sum_i E_i}$: the shower width (Figure 4.5 right) calculated within a window of $\Delta\eta \times \Delta\phi \simeq 0.06 \times 2$ (∼40 strips in η), where i is the index of the strip, i_{max} is the first local maximum and E_i is the deposited energy in the strip.
- Second layer of the electromagnetic calorimeter :
  - $R_\varphi = \frac{e233}{e237}$ and $R_\eta = \frac{e237}{e277}$, where $e233$ ($e237$, $e277$) is the uncalibrated energy summed over all cells in the middle layer, in a rectangle of size 3×3 (3×7, 7×7).

[3] In the following, the truth electron classification is : *isolated* if it matches a true electron coming from a Z or W boson, *hadron fake* if it does not match a true electron or tau or muon, *non-isolated* if it matches a true electron originating from b or c mesons and background electron if it matches a true electron coming from Dalitz decays or conversions.

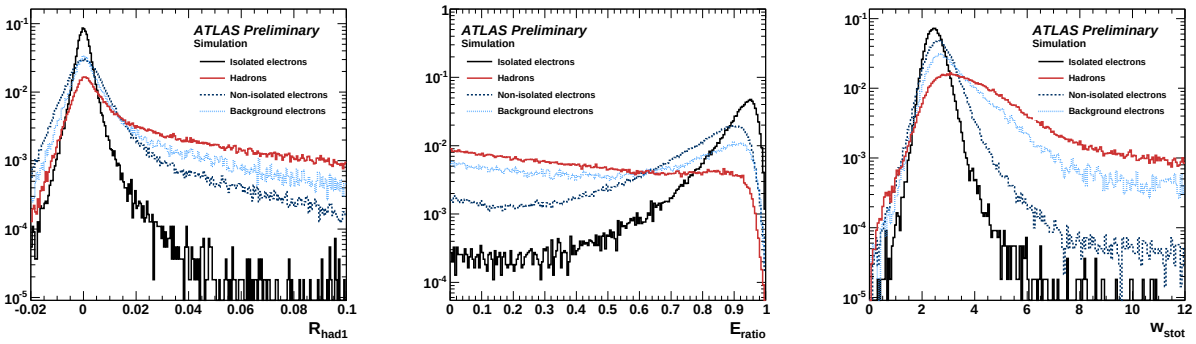


Figure 4.5: Distributions of R_{had} (left), E_{ratio} (middle) and W_{stot} (right) discriminant variables for isolated and fake electrons [167].

  - $W_{\eta 2} = \sqrt{(\sum_i (E_i \cdot \eta_i^2)/(\sum_i E_i) - ((\sum_i E_i \cdot \eta_i)/(\sum_i E_i))^2}$: lateral width, calculated within a 3×5 window, where i is the index of the cell, E_i and η_i are the associated energy and pseudorapidity.

- Third layer of the electromagnetic calorimeter :
  - $f_3 = E_3/E$: fraction of energy in the third layer of the EM calorimeter, where E_3 is the sum of energy contained in a 3×3 window around the most energetic cell.

- Inner detector, track quality cuts :
  - n_{Pixel} : number of hits in the pixel detector.
  - n_{Si} : number of hits in the pixel and SCT detectors.
  - n_{Blayer} : number of hits in the B layer.
  - d_0 and σ_{d_0} : transverse impact parameter and its significance.
  - $\Delta q/p$: momentum lost by the track between the perigee and the last measured point divided by the original momentum.
  - n_{TRT} : total number of hits in the TRT detector.
  - F_{HT} : ratio of the number of high-threshold hits and the total number of hits in the TRT.

- Cluster-track match variables :

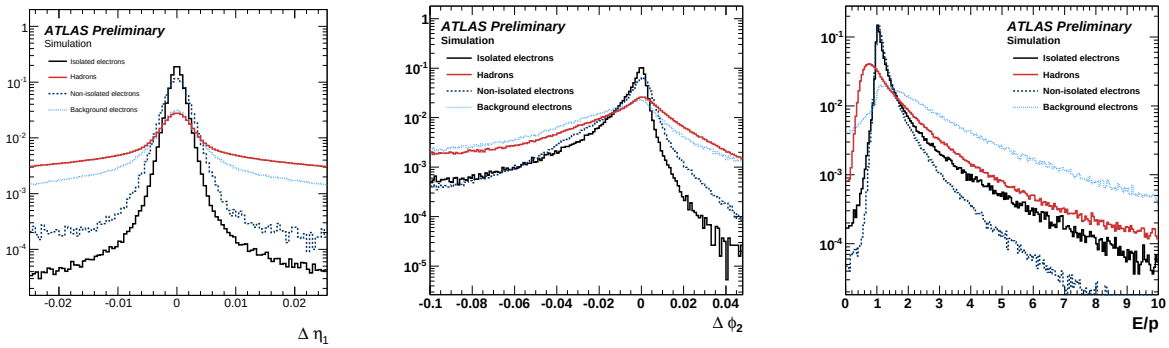


Figure 4.6: Distributions of $\Delta\eta_1$ (left), $\Delta\varphi_2$ (middle) and E/p (right) discriminant variables for isolated and fake electrons [167].

  - $\Delta\eta_1 = |\eta_{cluster}^{layer1} - \eta_{ID}|$, the difference between the cluster η in the first layer and the track η extrapolated to the calorimeter (Figure 4.6 left).
  - $\Delta\varphi_2 = |\varphi_{cluster}^{layer2} - \varphi_{ID}|$, the difference between the cluster φ in the second layer and the azimuth of the track extrapolated to the calorimeter (Figure 4.6 middle).
  - $\Delta\varphi_{res}$: same as Δ_{φ_2}, but the track momentum is rescaled to the cluster energy before extrapolating the track to the second layer.
  - E/p : the ratio between the cluster energy and track momentum (Figure 4.6 right).

- Additional cuts :
  - $!isConv$: no candidates matched to reconstructed photon conversion.

	Cut-based				MVA - likelihood function (LLH)		
Name	*loose++*	*medium++*	*tight++*	*multilepton*	*loose* LLH	*medium* LLH	*very tight* LLH
$R_{\text{had}(1)}$	✓	✓	✓	✓	✓	✓	✓
f_3		✓	✓	✓	✓	✓	✓
$W_{\eta 2}$	✓	✓	✓	✓	✓	✓	✓
R_η	✓	✓	✓	✓	✓	✓	✓
R_φ					✓	✓	✓
W_{stot}	✓	✓	✓	✓	✓	✓	✓
E_{ratio}	✓	✓	✓	✓	✓	✓	✓
f_1					✓	✓	✓
n_{Blayer}		✓	✓	✓	✓	✓	✓
n_{Pixel}	✓	✓	✓	✓	✓	✓	✓
n_{Si}	✓	✓	✓	✓	✓	✓	✓
d_0		✓	✓			✓	✓
σd_0						✓	✓
$\Delta q/p$				✓	✓	✓	✓
n_{TRT}		✓	✓	✓	✓	✓	✓
F_{HT}		✓	✓	✓	✓	✓	✓
Δ_{η_1}	✓	✓	✓	✓	✓	✓	✓
Δ_{φ_2}			✓				
$\Delta\varphi_{res}$				✓	✓	✓	✓
E/p			✓				
!isConv			✓				✓

Table 4.1: The electron identification menus [168].

Cuts applied in the hadronic calorimeter and second layer of the electromagnetic calorimeter reject the jets from high energetic pions with wide showers. Given the fine granularity in the $|\eta| < 2.35$ region, first layer variables are used to detect the shower substructure and separate π^0 from γ. The cut on number of hits in the B layer is a photon background discriminant. The cuts on the track cluster matching are ensuring a good track quality. Finally, requirements on the presented variables are combined in different menus using both cut based and multivariate analysis (MVA) techniques (Table 4.1).

Cut based electron identification

Using the cut based methodology four identification menus are defined with increasing background rejection power : *loose++*[4], *multilepton*, *medium++* and *tight++*. The cuts optimization for each menu is performed in 10 η bins and 11 p_T bins (first p_T bin is starting at 5 GeV and the last bin is > 80 GeV) to treat correctly the correlations, account for the different material distribution in the detector and the different distributions of the discriminant variables. This is essential as the shower shape variables are changing with the crossed amount of material, the track variables, the electron energy, etc. To improve the electron identification with 2012 data [167], several changes with respect to previous menus [154] are performed given the performances of the new reconstruction and the $\sqrt{s} = 8$ TeV run conditions (higher pileup, higher luminosity, etc.). The requirements on the track-cluster matching and the impact parameter variables are loosened. Tighter cuts are imposed on all variables from the first EM layer to account for the uncertainties associated to the material in front of the calorimeter and for the crossing dependency. To recover the loss in efficiency, the requirements on the lateral shower shape variables in the second EM layer and hadronic leakage are relaxed. Moreover, a new operating point (*multilepton*) is added and optimized for soft electrons in the $H \to ZZ \to llll$ analysis. When comparing to *loose++* criterion, the background rejection is increased for these electrons and the signal efficiency is unchanged. This performance is achieved after the shower shape variables are loosened and new variables (i.e. measuring the bremsstrahlung effects) are added.

[4] *Plus Plus* name is given for historical reasons.

Likelihood base identification

Using the multivariate analysis techniques a likelihood based function (LLH) is used to perform an improved electron identification optimization [173]. The signal and the background probability distribution functions of the discriminant variables[5] (Table 4.1) are obtained from data. The new LLH menus are optimized to have a better rejection, with respect to the cut-based menu, of the light flavor jets and conversions at a given electron efficiency. Hence the proposed *loose* LLH criterion has the signal electron efficiency of the *multilepton* criterion, *medium* LLH of the *medium++* and *very tight* LLH of the *tight++*. Studies based on MC simulations show that hadrons and conversions are reduced by a factor two when comparing with the cut-based identification. The identification efficiencies for the mentioned menus obtained with $Z \to ee$ data and background MC events are presented in Table 4.2 (for more details see Section 5.3).

criteria	$20 < p_T < 50$ GeV	
	ID efficiency Data [%] $Z \to ee$	ID efficiency MC [%] background
loose++	95.7 ± 0.2	4.76 ± 0.04
multilepton	92.9 ± 0.2	1.64 ± 0.02
medium++	88.1 ± 0.2	1.11 ± 0.02
tight++	77.5 ± 0.2	0.46 ± 0.01
loose LLH	92.8 ± 0.2	0.94 ± 0.02
medium LLH	87.8 ± 0.3	0.51 ± 0.01
very tight LLH	77.0 ± 0.3	0.29 ± 0.01

Table 4.2: Electron identification efficiencies for different menus computed with $Z \to ee$ data and background MC events [168].

4.2 Photons

In this section, the photon reconstruction and identification in the $|\eta| < 2.37$ and $p_T > 15$ GeV range are presented. They play an important role at LHC as the photon final states are used as a probe of perturbative QCD. Moreover, these states can be resonant in the SM physics, i.e. $H \to \gamma\gamma$ [7], and BSM, i.e. $G^* \to \gamma\gamma$ [174], or non-resonant, i.e. $\tilde{\chi}_1^0 \to \gamma\tilde{G}$ [175]. For my thesis, the photons are crucial for the electron reconstruction efficiency measurement.

Photon reconstruction and calibration

In a first instance, both electrons and photons are treated similarly by the `AuthorElectron` algorithm. As detailed in Section 4.1, an electron is reconstructed if at least one track is matched to the electromagnetic cluster. For photons, the reconstruction is performed separately for converted and unconverted photons [154, 176]. The former is characterized by the existence of at least one track originating from a vertex inside the tracker matched to the electromagnetic cluster, while the later is not having such a match.

The conversion vertices reconstructed in the inner detector are classified according to the associated number of tracks (single- or double-track). If a track is matched in η and φ to a reconstructed EM cluster, the object is added to the converted photon collection. The matching is considered to be successful if the impact parameter associated to the track, after extrapolation from its last measurement point to the second EM layer, is inside an (η, φ) window of radius 0.05 from the cluster center. It is extended to 0.1 in φ, in regions where the bremsstrahlung losses are expected during the track extrapolation. As for electrons, sometimes more than one conversion vertex candidate is matched to an electromagnetic cluster. This is possible when the EM showers have more than one vertex inside a small (η, φ) interval. Therefore, during the matching, the double-track candidates have priority over single-track ones, and, among all the double (single) - track candidates, the one with the smallest radial position of the corresponding vertex are prioritized. Finally, if a reconstructed electromagnetic cluster cannot be associated to a track, it is considered to be an unconverted photon candidate.

Note that almost all converted photons are also regarded as electrons. The existence of a track associated to a conversion vertex is used to separate the two categories. If the track associated to the EM cluster coincides with a track originating from a conversion vertex the electron is treated as a converted photon

[5] The variables counting the hits on the track are used as simple cuts and no PDFs are built.

object[6]. If the track cannot be associated to a conversion vertex, the object is classified as an electron. In the case of a initially reconstructed electron with a matched track having only the TRT information (and the usual requirements on energy and track p_T are fulfilled), the candidate is considered to be a converted photon even if no conversion vertex is associated to the EM cluster. During this procedure, several unconverted photons are also recovered : if the reconstructed electron has the best track candidate with only the TRT information or the converted photon candidate condition is not passed and the track p_T is smaller than a typical value of 2 GeV, the candidate is considered to be an unconverted photon.

The gain of recovery procedure on photon reconstruction efficiency is studied using $H \to \gamma\gamma$ processes ($p_{T\gamma} > 20$ GeV, $\eta < 1.37$ without the crack region) [176]. Before the recovery it is 56.61% ± 0.12% (5.71% ± 0.09% for converted photons and 85.89% ± 0.10% for unconverted photons), while after the recovery it increases to 97.82% ± 0.03% (94.33% ± 0.09% for converted photons and 99.33% ± 0.01% for uncovered photons). Only a fraction of 0.06% ± 0.01% of photons is not reconstructed.

The photon calibration is similar to electrons, as presented in section 4.1.2, and the most important sources of experimental systematic uncertainties are :

- Photon energy scale : the associated uncertainty is around 2 - 3% for $|\eta| < 1.37$ and $|\eta| > 1.82$ for both photon types. In $1.52 < |\eta| < 1.82$ region it is ∼0.9% for unconverted photons and ∼0.2% for converted photons [171].
- Photon energy scale resolution : similar to electrons [171].

Photon identification

The main photon discriminant variables are related to the associated electromagnetic shower shapes and isolation. Signal photons can arise from neutral hadron decays or from radiative decays of other particles, while the fake photons arise from QCD jets.

Name	Description	Loose ID	Tight ID
-	$\|\eta\| < 2.37$, $1.37 < \|\eta\| < 1.52$ excluded	✓	✓
R_{had}	Used over the range $\|\eta\| < 0.8$ and $\|\eta\| > 1.37$	✓	✓
R_{had1}	Used over the range $0.8 < \|\eta\| > 1.37$	✓	✓
R_{η}	-	✓	✓
$W_{\eta 2}$	-	✓	✓
R_{φ}	-		✓
W_{s3}	Shower width for three strips around strip with maximum energy deposit		✓
W_{stot}	-		✓
F_{side}	Energy outside core of three central strips but within seven strips divided by energy within the three central strips		✓
ΔE	Difference between the energy associated with the second maximum in the strip layer and the energy reconstructed in the strip with the minimal value found between the first and second maxima		✓
E_{ratio}	-		✓

Table 4.3: Variables used for the *loose* and *tight* photon identification cuts [177]. See Section 4.1.3 for variable definition.

As shown in Table 4.3 two photon identification criterias, *loose* and *tight*, are defined. The *loose* photon selection is built for trigger purposes, while the tight selection is optimized for physics analyses and provides a good background rejection. The photon identification menu optimization is performed separately for converted and unconverted photons in order to account for the differences in the shower variables, i.e. a wider shower for a pair of electrons arising from a photon conversion. The *tight* photon identification efficiency is around 85% for photon candidates with $p_T > 40$ GeV, while the corresponding background

[6]There is one exception, if there is a double-track conversion candidate with the coinciding track characterized by a B layer hit and the second track lacks one.

rejection is around 5000; for photons with p_T inside the [15, 40] GeV interval, the identification efficiency varies in the (65, 85)% range [177].

To further reject the fake photons, on top of the identification criteria additional cuts on isolation variables are considered. These variables are based on the transverse energy deposited in the electromagnetic calorimeter, in a cone of a given radius around the photon candidate. Some converted and unconverted photon identification efficiencies are presented in Figure 4.7 for several isolation requirements [177].

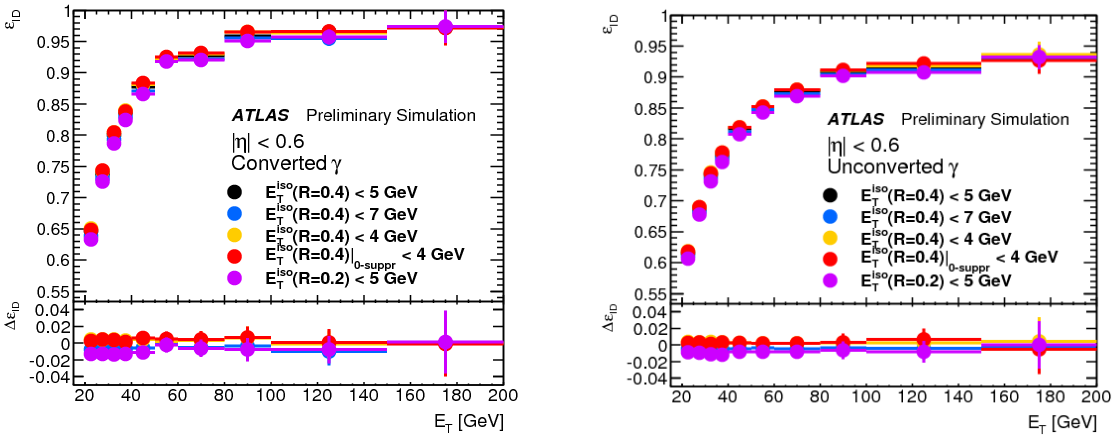


Figure 4.7: Identification efficiency for converted (left) and unconverted (right) photons in MC [177], for alternative isolation requirements, as a function of photon p_T in $|\eta| < 0.6$ region ($\sqrt{s} = 7$ TeV, $L = 4.9$ fb^{-1}).

4.3 Muons

In ATLAS the muon reconstruction and identification is performed in the $|\eta| < 2.7$ and $p_T > 4$ GeV interval, using the information from the Inner Detector (ID) and Muon Spectrometers (MS) [154,165,178–181]. The muon identification is carried out using the information provided by three reconstruction chains, `STACO`, `MUID` and `MUONS`. Note that the importance of these measurements is reflected in the SM results like $H \to llll$ [7], and in all the searches for BSM searches with leptonic channels.

Muon reconstruction

During the $\sqrt{s} = 8$ TeV data taking the chains used to reconstruct muons are `STACO` and `MUID`. For a further improvement the two chains are combined into only one, the so called `MUONS` or third muon chain. The four different stages considered during the muon reconstruction with the third chain are shortly described.

The segment finding in the Muon Spectrometer (MS) is based on the `Muonboy` algorithm and uses regions of activity (ROA) in the (η, φ) space which are identified with the information present in the trigger chambers. These regions are of size 0.04×0.04 and give at least one hit. All muons chambers intersecting with the ROAs are considered for the muon track reconstruction.

Straight segment candidates are built using Monitored Drift Tubes (MDT) hits. It is a segment track at station level. To obtain one, two outer MDT hits from the same or adjacent station are considered and stringent requirements on the number of hits around the segment must be fulfilled. After the segment is found the trigger hit(s) from the ROA, corresponding to the pattern, are associated to it. Another set of segments are reconstructed in the inner chamber of the MS, namely the Cathode Strip Chambers (CSC). For this step, 3D CSC clusters are used.

The tracks are then built in the Muon Spectrometer by assembling the final muon trajectory from the MDT and CSC segments, using the `More` algorithm. First, seed segments formed by all segments in a road, satisfying some minimum quality criteria, are selected and grouped per station layer. The seed segments are combined with segments from another station. If such a pair passes, the comparison of position and angles, it is fitted using a fast track fitter. If the fit is successful, a track candidate is considered. These candidates are extended to other stations to built the full track candidates and fitted with the ATLAS global χ^2 fitter [170] using the measurement constraints from the individual hits. The fit accounts for the material effects and track bending in the magnetic field. After cleaning, the hit recovery adds the hits in the layers crossed by the track that was not yet associated and a final fit is performed. The final step is the removal of duplicate and overlapping tracks. Note that, compared to CMS, ATLAS has a standalone MS track reconstruction.

The last stage is the muon track reconstruction in the Inner Detector with the *inside-outside* approach [182]. The final muon objects are built after merging the inner tracker and the muon spectrometer candidates, as presented the muon identification section. Note that in Sections 6 and 7, where the Run-1 analyses in which I am involved are presented, the `STACO` chain is used, while for Run-2 preparation studies and Run-2 data analysis the `MUONS` chain is considered. The `STACO` chain contains the muons found by the three different algorithms :

- `Muonboy` : builds standalone segments and tracks with the hit information found in the MS.
- `STACO` : produce combined muons using the information from the ID and MS.
- `Mutag` : tracks from ID are associated to segments found by `Muonboy` algorithm. To avoid any overlap, only ID tracks and segments not used by `STACO` are used.

Muon momentum scale and resolution

To accounts for the muon energy loss in the calorimeters and in the detector material before reaching the MS several corrections should be applied to MC events. They are measured using a sample enriched in $Z \to \mu\mu$ events and are at most 0.1% for ID and MS tracks reconstruction. Exceptions are for some specific regions in the MS where the corrections are higher (0.3% or 0.4%). Depending on the p_T region, total resolution smearing corrections of $\geq$10% - 15% are needed for the simulated ID and MS tracks. The uncertainty associated to the momentum scale for combined muons is $\pm$0.05% for $|\eta| < 1$ and $\pm$0.2 for $|\eta| > 2.3$ for $Z \to \mu\mu$ events. [165].

Muon identification

A combined muon identification is preferred to optimally use the information retrieved from the ATLAS sub-detector systems. Hence, the final muon object is a collection of different identified muon types used depending of the muon position (see later) :

- Stand-alone muons (SA) : the muon tracks are reconstructed only in the MS using the `Muonboy` algorithm.
- Combined muons (CB) : the muon track is reconstructed separately in the MS and ID.
- Segment-tagged muons (ST) : using the `Mutag` algorithm, the track reconstructed in the ID is identified as a muon only if that track extrapolated to the MS presents nearby hits.
- Calorimeter-tagged muons (CaloTag) : a track reconstructed in the ID is considered to be a muon only if the track can be associated to the energy deposited in the calorimeter by a minimum ionizing particle.

When comparing the different types, the combined muons have the highest muon purity, while calorimeter-tagged muons have the lowest purity. The latter recovers acceptance in the instrumented regions of the MS.

Muon reconstruction efficiency

As the three chains have similar performance, only the `STACO` results are mentioned. In Figure 4.8 the reconstruction efficiencies as a function of η, for CB, ST and CaloTag muon types are presented. Generally, the efficiency is around 99% with a precision at the 1 per-mille level and the agreement between data and MC results is well within 1%. The drop in $1.1 < |\eta| < 1.3$ regions is due to MS chambers which were not installed for Run-1, but they will be installed during the 2013-2014 LHC shutdown. The efficiency in this η interval is recovered by the use of ST muons. In the region $\eta \approx 0$, where MS is partially equipped with muon chambers, the efficiency is fully recovered by the use of CaloTag muons. In the so called foward region, above $|\eta| > 2.5$, SA muons is the only type able to provide information. In this particular region, the reconstruction efficiency is around 92% for muons with p_T inside [10, 20] GeV and around 96 - 98% otherwise.

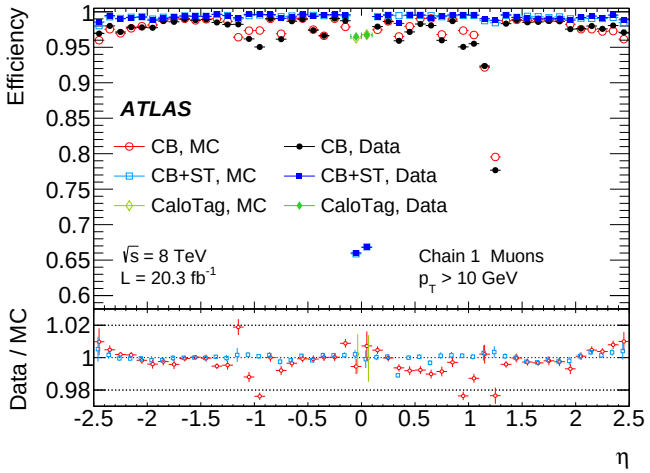


Figure 4.8: Muon reconstruction efficiency measured for different muon types in $Z \to \mu\mu$ events, as a function of η [165]. CaloTag muons are shown only in the region $|\eta| < 0.1$ as needed in the physics analyses. Only the statistical uncertainty is shown. The bottom panel shows the ratio between data and MC results and the associated total uncertainty.

4.4 Taus

With its high mass (1.8 GeV) and short decay length ($87\mu m$), the tau lepton decays before reaching the detector. Hence, its reconstruction and identification can be performed only with the decay products, which can be either leptons or hadrons [183–185]. The leptonic decay occur via an electron or a muon and two neutrinos, $\tau \to e\bar{\nu}_e\nu_\tau$ or $\tau \to \mu\bar{\nu}_\mu\nu_\tau$, in 35% of the cases. The leptonic tau decay cannot be used for the identification, as the decay products cannot be distinguished from prompt electrons and muons. The hadronic tau decay represents the remaining 65% and the intermediate state is characterized by the associated neutrino, charged and neutral pions and rarely by kaons. The hadrons are reconstructed as jets in the detector (denoted by τ_{vis}). To distinguished them from gluon- or quark-initiated jets, several discriminant variables as the (narrow) shower shape and number of tracks (1 or 3) are used.

4.5 Jets

Jets are playing a fundamental role in many physics measurements and in the searches for new physics at LHC. Their reconstruction starts with 3D topological clusters built in the ATLAS calorimeters, which are matched to tracks of charged particles measured in the inner detector. The algorithm used is anti-k_t and the energy calibration is performed using MC simulations [186–201].

Jet reconstruction

From all available algorithms, the anti-k_t [198] algorithm - with a distance parameter R set to 0.6 or 0.4 - is selected to reconstruct jets in ATLAS using the `FastJet` software package [202]. Inputs are stable simulated particles (= truth jets), reconstructed tracks in the inner detector (= track jets) or calorimeter clusters (= calorimeter jets). In this thesis only calorimeter jets with $R = 0.4$ are used. After the jet is reconstructed (Figure 4.9), its four momentum is defined as the four momentum sum of its constituents.

The anti-k_t algorithm is a simple, infrared and collinear (IRC) safe and soft- resilient jet algorithm. The latter particularity implies a certain rigidity ability in the algorithm to adapt a jet to a successive branching nature of QCD radiation. The two distances used to construct the jets are defined as :

- $d_{ij} = min(k_{ti}^{2p}, k_{kj}^{2p})\frac{\Delta_{ij}^2}{R^2}$: the distance between entities (particles or pseudo-jets) i and j;
- $d_{iB} = k_{ti}^{2p}$: the distance between entity i and beam, B.

Here $\Delta_{ij}^2 = (y_i - y_j)^2 + (\varphi_i - \varphi_j)^2$ and k_{ti}, y_i and φ are the transverse momentum, rapidity and azimuth of the particle i. R is the radius parameter and p is an index used to master the power of the energy versus geometrical (Δ_{ij}) scales. For this jet clustering algorithm, p is set to -1.

To built a jet, the smallest of the distances is identified : if it corresponds to d_{ij} the i and j entities are grouped,and if it corresponds to d_{iB} the i entity is called a jet. This procedure continues until no entities are left. In the case of an event with few well separated hard particles of transverse momenta k_{t1}, k_{t2}, ...

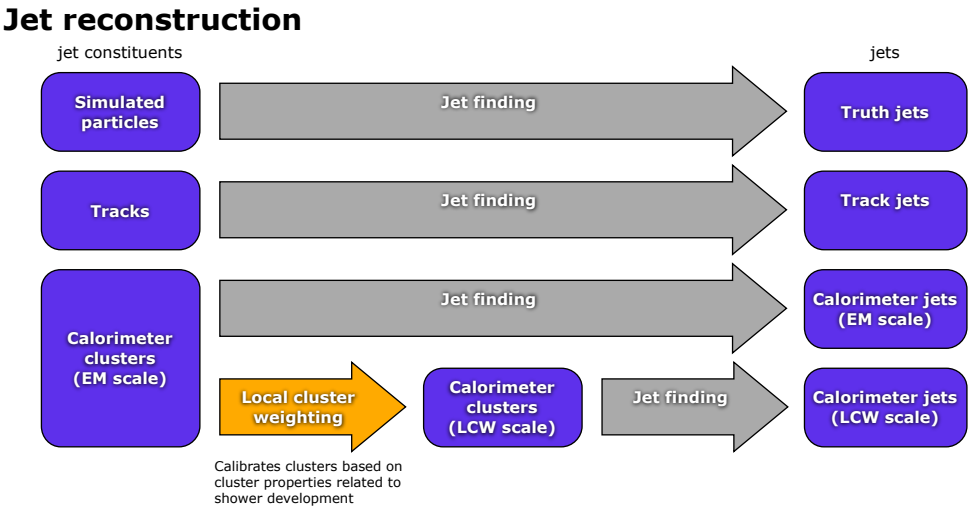


Figure 4.9: Illustration of jet reconstruction in ATLAS [187].

and many soft particles, the algorithm proceeds as following. The soft particles tend to group with the hard ones much before they group with other softer particles. If no other hard particle is within a distance of $2R$, the procedure will group with all soft particles withing a circle of radius R. The result is a perfectly conical jet. In the opposite case, if $R < \Delta_{12} < 2R$ the output will consist of two jets. If $k_{t1} \gg k_{t2}$, the first jet is conical while the second jet is partially conical because it is missing the part overlapping with the first hard jet. If $k_{t1} = k_{t2}$ both jets are partially conical and the overlapping part is divided by a straight line. Similarly, if $\Delta_{12} < R$ the two hard particles will form a jet. Note that the soft particles are not influencing the shape of the jet, while the hard particles do. Therefore the jet boundary is resilient with respect to soft radiation and flexible for the hard radiation.

Jet cleaning

Main sources of background jets are coming from proton beam collisions with the residual gas in the beam pipe, interactions in the tertiary collimators, cosmic muons overlapping in-time with collision events, calorimeter noise, etc. To separate the real jets from these backgrounds, four selections are defined with different levels of fake-jets rejection : *looser*, *loose*, *medium* and *tight*. The *looser* criteria has the highest jet efficiency, while the *tight* one has the highest background rejection. Their definitions are based on the reconstructed energy at the cell level, on the jet energy deposited in the direction of the shower development and on the number of reconstructed tracks matched to the jets.

The *looser* jet selection efficiency is greater than 99.8% in all considered p_T and η ranges. The *loose* selection efficiency is 1 - 2% lower, especially at low p_T and for $2.5 < |\eta| < 3.6$. The *medium* and *tight* selections efficiency are reduced due to cuts on the jet vertex fraction (JVF[7]). Typically, for *medium* and *tight* jets with p_T around 25 GeV the selection efficiencies are 96 and 85%, while for jets with $p_T > 50$ GeV the selection efficiency is 99% and 98% respectively[8].

Jet calibration

Initially the clusters are calibrated at the EM scale, which do not reflect the real composition of the hadronic shower. Then a local calibration scheme (LCW scale [203]) is applied on topological clusters. This is performed after calibrating the calorimeter cells such that the response of the calorimeter to hadrons is correctly reconstructed. This calibration uses the local cell signal weighting method which separates the hadronic signal from energy deposits and has a better resolution than the EM calibration. Jets are classified as electromagnetic or hadronic, using the measured energy density and the longitudinal shower depth. The energy corrections are obtained from single charged and neutral pion MC samples, according to the two jet classes.

Several corrections are considered for the reconstructed calorimeter jets, to account for noise threshold effects, energy loss in non-instrumented regions, etc [187]. Once applied, the jet energy scale (JES) is restored to that of jets reconstructed from stable simulated particles. These corrections are :

[7]This variable measures the ratio between the total momentum of tracks in the jet associated to the primary vertex and the total momentum of tracks associated to the jet, $\sum_{tracks_{jets},PV} p_T / \sum_{tracks_{jet}} p_T$.

[8]These results are obtained with $\sqrt{s} = 7$ TeV and $L = 4.7$ fb^{-1} of data.

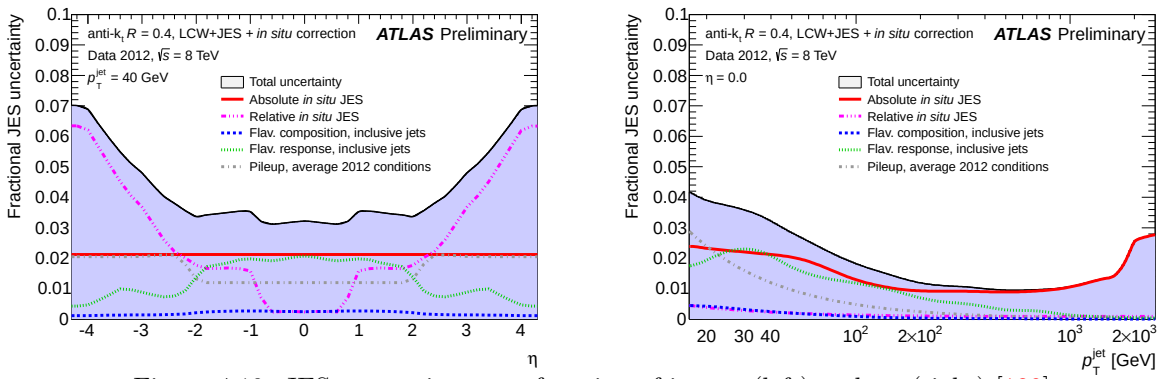


Figure 4.10: JES uncertainty as a function of jet eta (left) and p_T (right) [189].

- Pile-up corrections : all calorimeter jets formed from topological clusters are calibrated to account for energy offset due to in-time and out-of-time pile-up. For $\sqrt{s} = 7$ TeV data, the jet p_T calibration is based on an average correction, $< O >$, which measures the amount of p_T generated by pile-up in a jet. It is subtracted from the reconstructed jet p_T at any given scale.

$$p_T^{corr} = p_T^{jet} - < O > \tag{4.4}$$

  The measurement of $< O >$ parameter was performed in MC and parametrized as a function of number of reconstructed primary vertices N_{PV}, expected average number of interactions $< \mu >$ and η. This approach does not account for the pile-up energy fluctuations added to the calorimeter on an event-by-event basis and no jet-by-jet fluctuations or jet shape are considered. Therefore, for 8 TeV data an improved methodology is used. The p_T correction is estimated on an event-by-event basis, using a jet area (A^{jet}) in η - φ space and the pile-up energy density (ρ) :

$$O^{jet} = \rho \times A^{jet} \tag{4.5}$$

  The parameter ρ of an event is estimated as the median of the distribution of the density of many k_t jets [204] in the event. Jets obtained with the k_t algorithm are used instead of anti-k_t jets, as the former algorithm is more sensitive to soft jets. The measurement of the second parameter, A^{jet} is based on ghost association. To a real event, "ghost entities" of infinitesimally small p_T are thrown in the η - φ space to be clustered with the real jet signal during jet reconstruction. The associated number of ghosts to one jet will be a measure of its area. It is obtained after summing over all ghosts four-momentum belonging to a jet.

- Origin corrections are applied to the jet direction, to point back to the PV in the event and not to the nominal center of ATLAS.

- Jet calibration based on MC simulations : the corrections applied to reconstructed jet energy and η are ratios with respect to truth jets. The obtained collection is called calibrated jets with the EM + JES or LCW + JES scheme. Only the latter is used for the analyses presented in this thesis.

- Residual in situ corrections are applied to all jets to remove the remaining data-to-MC differences.

To measure these corrections several *in-situ* calibration methodologies involving different types of events like dijets, Z + jet or γ + jet, are used. These corrections are applied on jets with $p_T > 20$ GeV and $|\eta| < 4.5$ passing the considered quality selection. The final JES calibration and the associated uncertainty is a combination of the mentioned measurements. The results obtained for JES systematic uncertainty at the LCW scale are presented in Figure 4.10. These values are used to obtain the results presented in Chapters 6, 7 and 9. An uncertainty of 3 - 4% is found for jets in the $20 \leq p_T <$50 GeV interval, while it decreases up to 1% for more energetic jets (0.2 - 1 TeV) . For the central jets it is around 3.5% and it increases for forward jets up to 7% at $|\eta| = 4.5$.

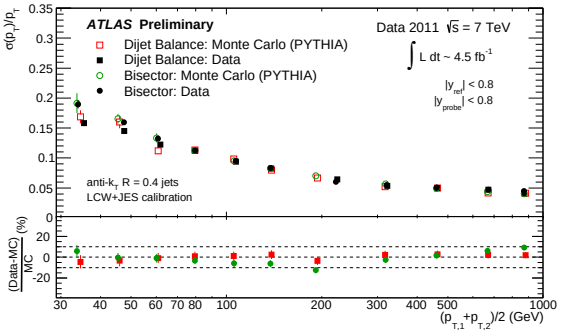


Figure 4.11: JER measured with the di-jet (squares) *in-situ* calibration at the LCW scale [189].

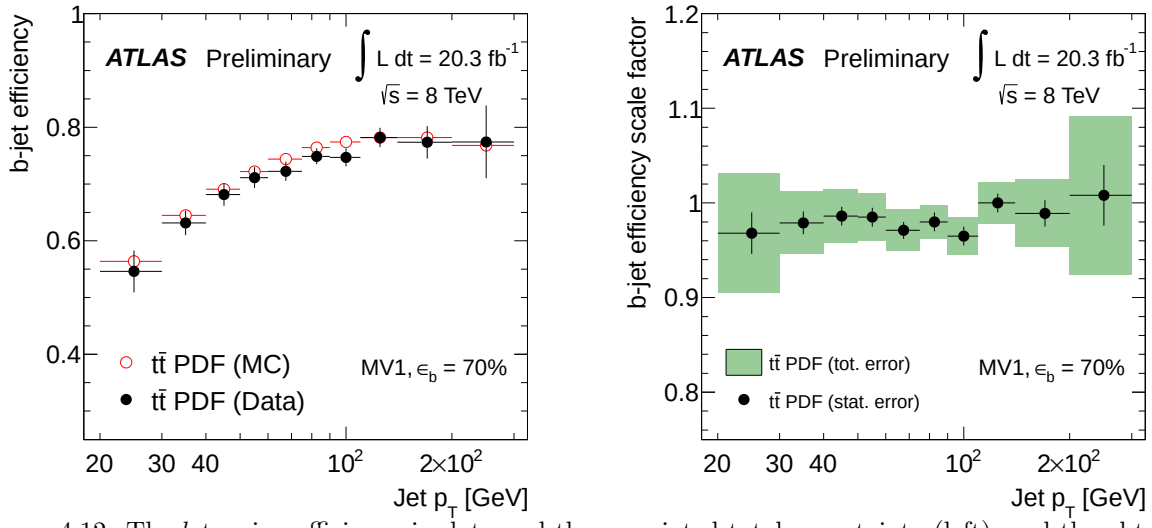


Figure 4.12: The b tagging efficiency in data and the associated total uncertainty (left), and the obtained scale factors together with the statistical and total uncertainty(right) for the MV1 b-tagging tool at 70% b jet efficiency working point [207].

Jet energy resolution

The precise measurement of the jet energy resolution is crucial for di-jets cross section measurements, inclusive jets or vector boson production in association with jets and for top quark cross section and mass measurements. It plays a fundamental role also in searches for resonances with jet final states. To measure it, similar methods as for JES measurement are used. As an example, the results obtained with $\sqrt{s} =$ 7 TeV 4.5 pb^{-1} of data are presented in Figure 4.11 for $\eta < 0.8$. The fractional jet energy resolution is around 17% for jets with p_T of 30 GeV decreasing down to 5% for more energetic jets ($p_T > 500$ GeV). The relative difference between the data and MC results is less than 10% (Figure 4.11, lower panel).

4.6 b-jet tagging

The ability of identifying b jets from c and u, d, s or gluons (called *light jets* from here on) plays a crucial role in the ATLAS physics program. Different b-tagging algorithms were developed to achieve a high b-tagging efficiency (ε_b) for real b jets while retaining the c and light flavor jets misidentification efficiency at very low levels. They are either relatively simple and based on impact parameters (IP3D), or secondary vertices (SV1), or more complicated,they are exploiting the topology of b and c hadron decays (JetFitter). The most discriminant variables associated to these algorithms are grouped in artificial neutral networks. The output weight probability densities are evaluated separately for b, c and light jets. They are used as input for the multivariate tagging algorithms to further increase the tagging performance. The algorithm used in ATLAS and in this thesis is MV1 [205–208].

The MV1 algorithm efficiencies are measured in data and the calibration is provided as data-to-MC efficiency ratios (scale factors). Several working points (WP = 60%, 70% and 80%) are defined with tight or loose b-tagging rates. The main objects used for the b-tagging are the calorimeter jets, the tracks reconstructed in the ID and the PV. Only jets with $p_T > 20$ GeV and $|\eta| < 2.5$ are used and a JVF cut is applied on jets with $p_T > 50$ GeV and $\eta < 2.4$ to reduce the pile-up contribution. The rates are obtained in $t\bar{t}$ events, where the W boson is required to decay leptonically.

The b-tagging efficiency versus p_T of the jet (for an average b-tagging efficiency of 70%), the associated uncertainties and the scale factors, obtained with $\sqrt{s} = 8$ TeV, $L = 20.3$ fb^{-1}, are shown in Figure 4.12. It increases from 55% for soft jets to ∼80% for jets with $p_T > 200$ GeV. The light flavor jet rejection rate (defined as the inverse of the mistag rate) is shown in Figure 4.13, as a function of the b-tagging efficiency [207]. Table 4.4 shows the benchmarks values obtained with MV1 algorithm. Several working points (WP) are shown together with the b-tagging efficiency, purity, c and light flavor quark rejection efficiency (RF) together with the tau lepton rejection efficiency. These values are

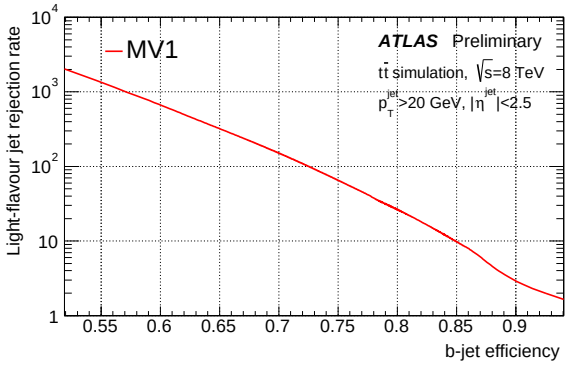


Figure 4.13: The light flavor rejection as a function of the b-tagging efficiency for WP = 70% [208].

obtained using $t\bar{t}$ samples.

WP [%]	ε_b [%]	purity [%]	c RF	tau RF	light jets RF
80	80	85.41	3.08	5.49	25.18
70	70	92.28	4.97	13.24	136.66
60	55.99	94.87	7.93	23.69	584.94

Table 4.4: b-tagging benchmarks obtained in $t\bar{t}$ for MV1 algorithm [209]. The considered jet collection is calorimeter reconstructed at LCW scale.

4.7 Missing transverse energy

The missing transverse momentum ($\vec{\not{E}}_T$) represents the energy arising from weakly or non interacting particles, defined as the momentum imbalance in the plane transverse to the beam axis. The transverse momentum imbalance is the negative vector sum of all particles p_T detected in the detector. Its magnitude is denoted with $\not{E}_T$. The $\not{E}_T$ reconstruction and calibration is optimized to suppress as much as possible the pile-up. The achieved performance is presented in terms of resolution, response and tails of the reconstructed $\not{E}_T$ [210–213].

$\not{E}_T$ reconstruction and calibration

For the $\not{E}_T$ reconstruction, the energy deposited in the calorimeters and the reconstructed muons in the MS are used. To recover the soft muons, the reconstructed muons using segments matched to ID tracks and extrapolated to MS are also considered. Contributions from soft particles are recovered by adding tracks. The $\vec{\not{E}}_T$ variable is defined as :

$$\not{E}_{x(y)} = \not{E}^{e}_{x(y)} + \not{E}^{\gamma}_{x(y)} + \not{E}^{\tau}_{x(y)} + \not{E}^{jets}_{x(y)} + \not{E}^{SoftTerm}_{x(y)} + \not{E}^{\mu}_{x(y)} \tag{4.6}$$

with each term computed as the negative sum of the reconstructed objects projected on the corresponding direction, x or y. Note that only reconstructed and calibrated objects are used. The energy deposited in the calorimeter is associated to reconstructed and identified parent in the following order : electrons, photons, hadronic tau, jets and muons. The energy not associated to such an object (like low p_T tracks) is considered in the $\not{E}^{SoftTerm}_{x(y)}$ term.

To parametrize the $\not{E}_T$ performance, the total transverse energy in the calorimeters, $\sum E_T^{Event}$ is computed as :

$$\sum E_T^{Event} = H_T + \sum p_T^{\mu} = \sum E_T^{e} + \sum E_T^{\gamma} + \sum E_T^{\tau} + \sum E_T^{jets} + \sum E_T^{SoftTerm} + \sum p_T^{\mu} \tag{4.7}$$

Pile-up suppression

The terms in the $\not{E}_T$ reconstruction equation, which are most affected by pile-up are $\not{E}^{jets}_{x(y)}$ and $\not{E}^{SoftTerm}_{x(y)}$. To suppress the pile-up in the jet term, beside the pile-up corrections already applied during the jet reconstruction (Section 4.5), a cut on JVF is added. Therefore, any jet with $p_T < 50$ GeV and $|\eta| < 2.4$ which does not fulfill the $|\text{JVF}| > 0$ requirement is discarded. Jets outside the inner detector with no associated tracks are retained ($\text{JVF} = 1$). To suppress the pile-up in the soft term, several methods are developed. Only the one based on the use of tracks is applied in the SUSY analysis presented in Chapter 6. It relies on the track information, as tracks can be associated to the primary vertex from the hard-scattering collision. The pile-up suppression is performed by scaling $\not{E}^{SoftTerm}_{x(y)}$ with the soft term vertex fraction, STVF. This variable is similar with JVF and represents the fraction of momenta of tracks matched to the soft term which are associated with the hard scattering vertex. The $\not{E}_T$ calculated with this correction is called STVF.

$\not{E}_T$ performance

The $\not{E}_T$ performance depends on the event topology, i.e. the presence of true $\not{E}_T$, leptons, jet activity, etc. For illustration, results obtained with $Z \to \mu\mu$ events before and after pile-up suppression with STVF are shown in Figure 4.14, left. This channel is considered as the high $\not{E}_T$ reconstruction is a direct result of imperfections in the reconstruction process or in the detector response. A good agreement between data and MC expectation is observed and the tails in $\not{E}_T$ are compatible either with signal candidates or with backgrounds like $t\bar{t}$ or diboson. This demonstrates that instrumental effects are well described.

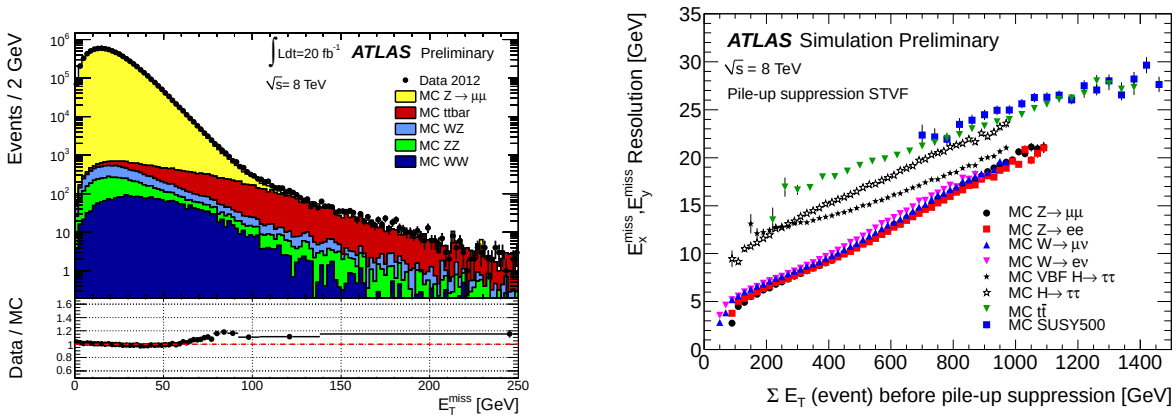


Figure 4.14: Left : $\not{E}_T$ distribution as measured in data with $Z \to \mu\mu$ events without pile-up suppression [211]. Right : $\not{E}_T$ resolution summary plot [211].

$\not{E}_T$ resolution

The $\not{E}_T$ resolution performance is shown in Figure 4.14 right, in several MC samples after STVF pile-up suppression. If before the pile-up suppression the resolution was similar for all the samples, after the suppression the resolution is very dependent on the contribution brought by the soft term.

Sources of systematic uncertainties

The systematic uncertainties associated to each individual object entering in the $\not{E}_T$ definition is evaluated and combined to obtain the overall systematic uncertainty. The uncertainties assigned to electrons, photons, taus, jets and muons were already discussed in previous sections. The systematic uncertainty on the soft term arises from MC modeling and effects of pile-up. It is computed after comparing the data and $Z \to \mu\mu$ MC results obtained with two methods. In the first method, the $\not{E}_T$ projection onto the Z boson transverse direction is compared in data and MC to obtain the scale uncertainty. Similar comparisons are performed using the $\not{E}_{x(y)}$ terms to obtain the resolution uncertainty. The second method exploits the balance between the soft term and the total transverse momentum of the hard objects (μ, e, γ, neutrinos and jets) in the event. The results are showing a systematic uncertainty of the order of a few percents and its impact is visible only before pile-up suppression. After the pile-up suppression the soft term is strongly reduced and the associated uncertainty becomes negligible.

4.8 Conclusions

In this section the electron, photon, muon and tau objects reconstruction and identification in ATLAS were presented. The jet reconstruction and b-tagging procedure were also considered and described. The benchmarks obtained in $t\bar{t}$ MC for MV1 b-tagging algorithm are shown for reference, as the b-jets signatures are of a great interest in this thesis. The reconstruction and calibration of missing transverse energy, a key variable in many BSM searches, is detailed together with the achieved performance and response. The main systematic uncertainties associated to these measurements were also presented. All these sources of uncertainties are considered to obtain the results in the physics analysis chapters.

Chapter 5

Electron reconstruction and identification efficiencies measurements

The precise knowledge of the electron reconstruction and identification efficiencies plays a crucial role in the Standard Model measurements, as a significant difference between the theoretical and measured cross section can be a sign of new physics. As an example, Figure 5.1 shows the ATLAS combined Standard Model cross sections summary plot, where for most of the processes the systematic uncertainty dominates. This clearly demonstrates the crucial role of object performance measurements.

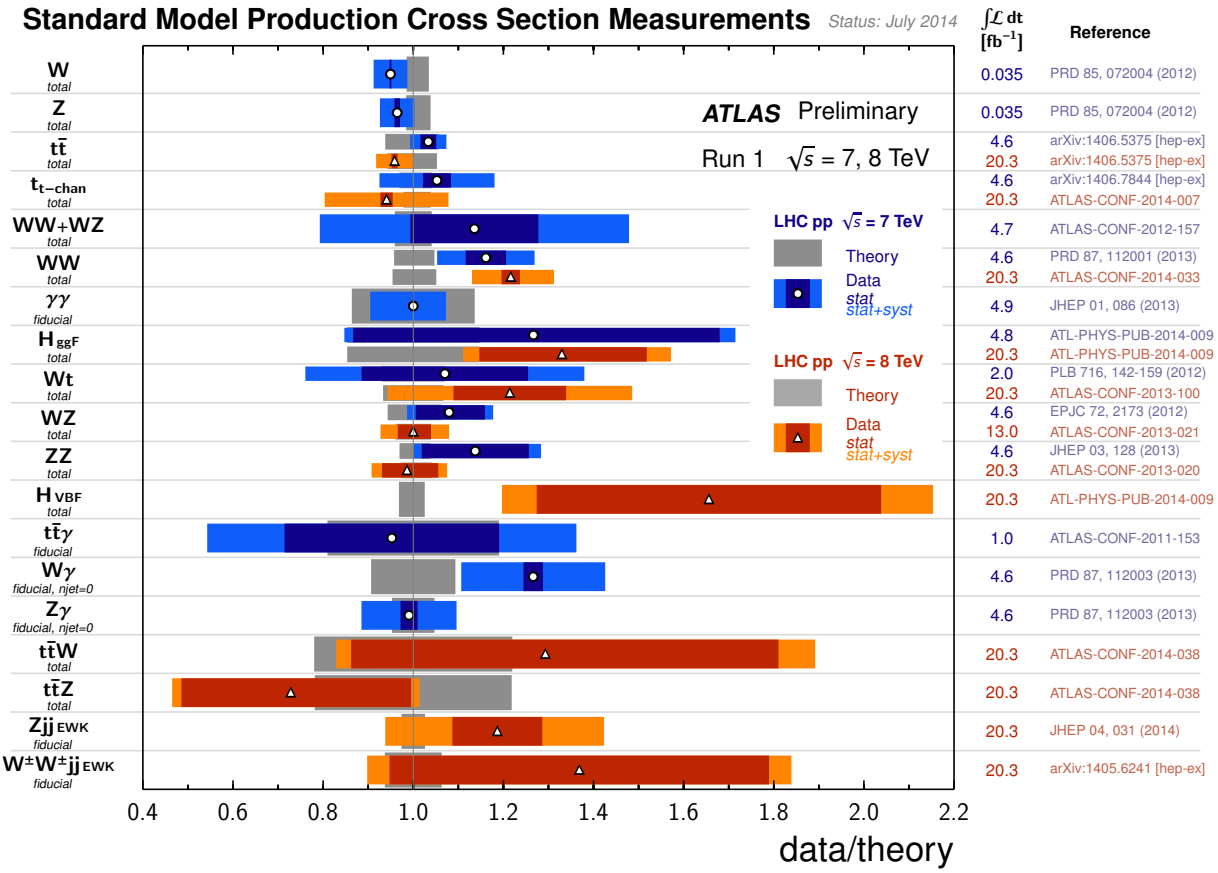


Figure 5.1: The data/theory ratio for several Standard Model total and fiducial production cross section measurements [214]. The dark color bar represents the statistical uncertainly, while the lighter color bar stands for the total uncertainty.

To illustrate the role of the electron performance, Table 5.1 details the systematic uncertainty sources in the electron channels for (di)boson Standard Model production cross section measurements. For the ZZ

and inclusive W/Z cross sections measurement the dominant sources of experimental systematic uncertainty are related to the electron reconstruction and identification efficiencies, eg. in $ZZ \to eeee$ channel it is 6.2%, given a total of 6.7%, while for the $W^{\pm}W^{\pm}$ and WZ measurements it is replaced by other sources, eg. JES or JER.

W^+W^- cross section measurement at $\sqrt{s} = 8$ TeV [215]		
Uncertainty source	$e\mu$	ee
Electron efficiency [%]	0.9	2.0
Jet E scale [%]	2.6	2.6
Jet E resolution [%]	2.3	2.2
$\not{E}_T$ soft term scale [%]	2.3	4.2
Total experimental unc. [%]	3.7	6.3
Total theoretical unc. [%]	3.3	3.7
$\sigma^{fiducial}$ [pb]	$377.8^{+6.9/25.1}_{-6.8/22.2}$	$68.5^{+4.2/7.7}_{-4.1/6.6}$
$\sigma^{measured}$ [pb]	$71.4^{+1.3/5.0}_{-1.3/4.4}$	$68.6^{+4.2/7.8}_{-4.1/6.7}$

WZ cross section measurement at $\sqrt{s} = 8$ TeV [216]			
Uncertainty source	$e\mu\mu$	$ee\mu$	eee
Ele reco. efficiency [%]	$^{+0.76}_{-0.81}$	$^{+1.52}_{-1.48}$	$^{+2.39}_{-2.29}$
Ele ID efficiency [%]	$^{+1.23}_{-1.26}$	$^{+2.18}_{-2.10}$	$^{+3.95}_{-3.62}$
Jet E scale [%]	$^{+0.76}_{-0.81}$	$^{+0.21}_{-0.21}$	$^{+0.47}_{-0.46}$
Jet E resolution [%]	$^{+1.00}_{-1.04}$	$^{+0.11}_{-0.11}$	$^{+0.12}_{-0.11}$
Total unc. (no lumi) [%]	$^{+6.80}_{-6.97}$	$^{+6.40}_{-6.34}$	$^{+10.38}_{-10.20}$
$\sigma^{fiducial}$ [pb]	$26.2^{+2.2/1.7}_{-2.1/1.7}$	$26.8^{+2.1/1.6}_{-2.0/1.6}$	$22.7^{+2.5/2.3}_{-2.3/2.3}$
$\sigma^{measured}$ [pb]	$21.4^{+1.9/1.5}_{-1.7/1.5}$	$21.9^{+1.8/1.4}_{-1.6/1.4}$	$18.6^{+2.1/1.9}_{-1.9/1.9}$)

ZZ cross section measurement at $\sqrt{s} = 8$ TeV [217]		
Uncertainty source	$eeee$	$ee\mu\mu$
Lepton reco. and ID eff. [%]	6.2	3.1
Total reconstruction unc. [%]	6.4	3.4
Total unc. (no lumi) [%]	6.7	3.3
$\sigma^{fiducial}$ [fb]	$4.6^{+0.8/0.4}_{-0.7/0.4}$	$11.1^{+1.0/0.5}_{-0.9/0.5}$

Z/W cross section measurement at $\sqrt{s} = 7$ TeV [218]		
Unc. source	$Z \to ee$	$W^{\pm} \to e\nu$
Ele reco. efficiency [%]	1.6	0.8
Ele ID efficiency [%]	1.8	0.9
Total experimental unc. [%]	2.7	1.8
Total unc. (no lumi) [%]	3.3	2.3
$\sigma^{fiducial} \times BR$ (no lumi) [nb]	0.426±0.004±0.012	4.79±0.014±0.089

Table 5.1: Fiducial and/or measured SM cross sections together with the total uncertainty. If available, the experimental and/or theoretical sources are also presented. For illustration, the sources of systematic uncertainties associated to electron and the dominant sources are shown as well. Top left : sources of relative systematic uncertainties [%] on the W^+W^- signal fiducial acceptance estimations. Top right : sources of relative systematic uncertainty [%] on the total WZ cross section; here the total uncertainty includes sources like signal and MC statistic and systematic on the data driven bkg estimation. Bottom left : relative systematic uncertainties associated to lepton reconstruction (reco.) and identification (ID) efficiencies measurement [%] on the ZZ signal fiducial acceptance estimations; to compute the total reconstruction uncertainty, besides the lepton reconstruction uncertainty, the lepton energy and momentum, lepton isolation and impact parameter and trigger efficiency are considered as well. Bottom right : relative systematic uncertainties on the inclusive W/Z measured integrated cross section in the electron channels.

Another crucial role of the electron performance is played in the measurement of the Higgs coupling to bosons [5, 7, 10]. For $H \to ZZ \to llll$ channel, results using $\sqrt{s} = 8$ TeV data improve over the $\sqrt{s} = 7$ TeV measurement [10]. Apart from the increase in luminosity, the improvements are coming from the new electron reconstruction algorithm, the new likelihood electron identification and the sub-percent systematic uncertainties assigned to electron efficiencies.

This chapter is dedicated to the electron reconstruction and identification efficiency measurements using $\sqrt{s} = 8$ TeV data with $L = 20.3$ fb^{-1}. All the Standard Model and Beyond Standard Model analyses involving electrons are based on these results. Generalities about the efficiency measurement and the methodology are described in Section 5.1. In Section 5.2, the electron reconstruction measurement with $Z \to ee$ events are detailed. Section 5.3 is dedicated to electron identification measurements with $Z \to ee$ events. All these results represent a part of my contribution for the ATLAS collaboration and they are documented in an ATLAS internal document [169] and a public conference note [168]. Complementary electron identification measurements relaying on J/Ψ or $Z \to ee\gamma$ events are presented together with the overall combined electron identification measurements. Section 5.4 presents the overall electron efficiencies measurement. Finally, Section 5.5 summarizes the main conclusions.

5.1 Generalities on electron efficiency measurement

In ATLAS, the efficiency to find an electron is measured at different stages :

- Electron reconstruction : the reconstruction efficiency, ε_{reco}, reflects the capability of the track reconstruction and its association with the electromagnetic cluster (Section 4.1.1).

- Electron identification : the identification efficiency, ε_{id}, reflects the probability of each reconstructed electron to pass a given identification requirement (Section 4.1.3).

- Trigger : $\varepsilon_{trigger}$ efficiencies are computed for each electron identification definition. It is the efficiency of one identified electron to satisfy the trigger requirements
- Other : ε_{other} is the efficiency of one identified electron to pass some additional requirements such as the track or calorimeter isolation. It is computed in each physics analysis as it is highly dependent on the imposed electron selection.

The final efficiency is just the product of the terms above :

$$\varepsilon_{total} = \varepsilon_{reco} \cdot \varepsilon_{id} \cdot \varepsilon_{trigger} \cdot \varepsilon_{other} \tag{5.1}$$

As mentioned in the introduction, only the electron reconstruction (reco.) and the identification (ID) efficiency measurements are presented in this thesis (Figure 5.2). To account for the (strong) dependency on electron η and p_T, the measurement is performed in fine (η, p_T) bins as presented in Table 5.2. This also allows a correct treatment of the background which is having a different composition and distribution in each bin. For example, the background fraction is ~80% in the $p_T < 15$ GeV and end-cap η regions, and smaller than 1% for electrons with $p_T > 45$ GeV.

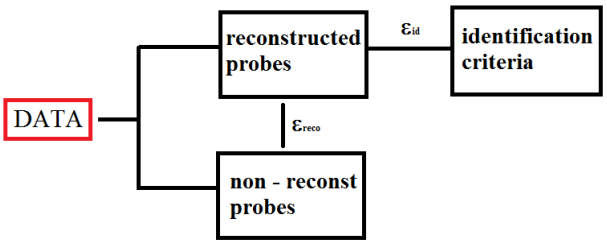


Figure 5.2: Illustration of different stages used for electron efficiency measurement.

p_T boundaries [GeV]													
Electron reconstruction			15	20	25	30	35	40	45	50	60	80	>80
Electron identification	7	10	15	20	25	30	35	40	45	50	60	80	>80

η boundaries											
Using $J/\Psi \to ee$ events	0			0.8		1.37	1.52		2.01		2.47
Using $Z \to ee(\gamma)$ events	0	0.1	0.6	0.8	1.15	1.37	1.52	1.81	2.01	2.37	2.47

Table 5.2: Boundaries of the (η, p_T) bins used to measure the electron reconstruction and identification efficiencies. Measurements with $J/\Psi \to ee$ events are performed as a function of $|\eta|$ while measurements with $Z \to ee$ events are performed in independent bins for $\eta < 0$ or $\eta > 0$, defined in a symmetric way.

The electron reconstruction efficiencies are measured at reconstruction plus track quality stage, where the track quality requirement is added to remove the beam background. For the electron identification, to cover the needs for all the ATLAS Standard Model and Beyond Standard Model analyses, a total of 7 electron working points are supported : *loose++*, *medium++* and *tight++* cut based and *loose*, *medium*, *tight* and *very tight* using a likelihood function (Section 4.1.3).

The method used to perform the electron reconstruction and identification measurements is presented in Section 5.1.1. The event selection is detailed in Section 5.1.2. The background estimation procedure is described for electrons with an associated track in Section 5.1.3 and for electrons with no associated track in Section 5.1.4.

5.1.1 The *Tag&Probe* method with $Z \to ee$ events

To obtain the reconstruction and identification efficiencies, real electron[1] enriched data samples must be defined. A method is developed to take full advantage of the precise known Z boson mass (91.1876 ± 0.0021 GeV). It is called *Tag&Probe* and it requires at least two electrons in the event on which different identification selections are imposed :

- Tag electron : very strict selection criteria to assure that it is a real electron from the Z boson.
- Probe electron : loose selection to not bias the probe (it can be a real or a background electron).

[1] Here electrons from W, Z or H bosons decays.

Figure 5.3 shows the typical *Tag&Probe* selection for a $Z \to ee$ event. The invariant mass of the probe and tag electrons should be inside [80, 100] GeV interval (peak region). After defining the tag electron and subtracting the background, the identification efficiency is simply the ratio between the number of events for which the probe is passing the imposed selection (numerator) and the total number of events in the selected region (denominator). For the reconstruction efficiency the numerator is formed by reconstructed electrons, namely clusters with associated tracks, while the denominator is populated by all clusters. In the latter, the cluster can be associated to a track or not, in which case the reconstructed algorithm classifies it respectively as an electron or a photon. In the case of identification efficiency, the denominator is the numerator of the reconstruction efficiency (no photons).

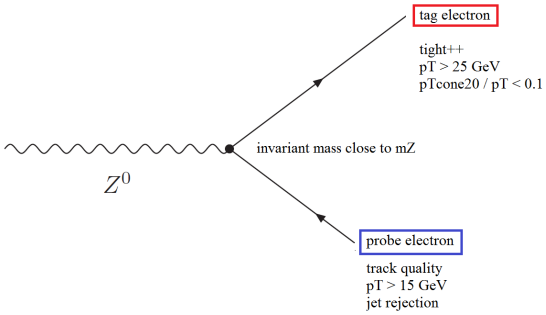


Figure 5.3: Illustration of a *tag* and *probe* electron selection in a $Z \to ee$ event.

When using this method to measure the electron efficiency in data, the main challenge is the background subtraction under the Z peak which could be very high even after the probe sample selection. The main sources of fake electrons are heavy flavor decays, photon conversions, jets accompanying the W processes faking electrons. To estimate them, dedicated methods based on combined fits or background templates were developed and validated with $\sqrt{s} = 7$ TeV data [2, 162, 219–221]. As the background electrons are not perfectly simulated the techniques rely only on data. The background subtraction is performed independently for the baseline selection enriched in probe electrons and for the final selection for which the efficiency needs to be measured.

For illustration, Figure 5.4 shows the invariant mass of the two leading reconstructed (probe) electrons in data and simulation in several p_T ranges. Here the component with two light jets is missing, giving the white area between the data and prediction. In the sector with soft electrons, $p_T < 25$ GeV, the background level is very high (more than ∼20,000 events in the [85, 95] GeV m_{ee} interval) making any measurement more difficult in this range.

5.1.2 Event selection for the *Tag&Probe* method

Events are selected with two lepton triggers, `e24vhi_medium1` or `e60_medium1`, and required to have at least two electrons using the standard algorithm, `AuthorElectron`, in the central region of the detector, $|\eta| < 2.47$. In addition, each event should have one reconstructed primary vertex with at least three tracks.

Electron p_T threshold is 15 GeV and the associated track is having at least 1 pixel hit and at least 7 silicon hits in the SCT and pixel detectors (track quality cuts) [2]. All electrons should be separated from any jet[3] with $p_T > 20$ GeV present in a cone with an opening angle $\Delta R_{jet,e} = 0.4$ around the electron. This cut helps in rejecting the jets faking electrons type of background.

Besides the mentioned requirements additional cuts are applied on the tag electron. The p_T threshold is increased to 25 GeV and for the identification efficiency measurements it should have an opposite electric charge with respect to the probe (no cut on the charge is applied for reconstruction efficiency measurements). The electron should be outside the crack region and pass the *tight++* identification criteria. It should also satisfy cuts on the track isolation, $p_T^{\text{cone30}}/p_T < 0.1$, and calorimeter isolation, $E_T^{\text{cone40}}/p_T > 0.05$.

Photons are required to be reconstructed by the standard *egamma* algorithm and be well separated from any electron within $\Delta R_{\gamma,e}$[4] > 0.4. The latter cut is rejecting photons reconstructed from a conversion vertex. No cut on the electric charge is applied, hence the background level is two times higher making the reconstruction efficiency measurement very challenging, especially at low p_T.

The invariant mass build with the tag and probe electrons satisfying the mentioned selection will define the so called *baseline distribution*.

[2] This requirement is removed when measuring the efficiency at reconstruction stage, Appendix A.2.
[3] The jet reconstructed from the electron information in the calorimeter is removed from counting.
[4] Here ΔR is the distance in η and Φ between the electron and photon clusters in the calorimeter.

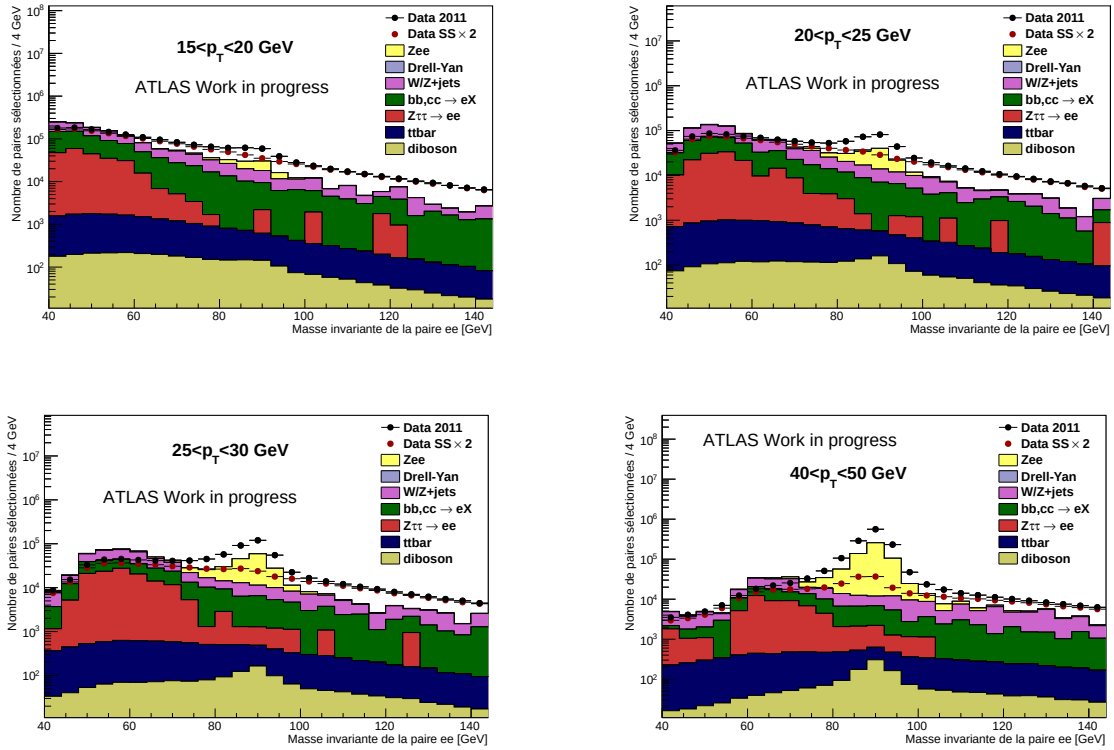
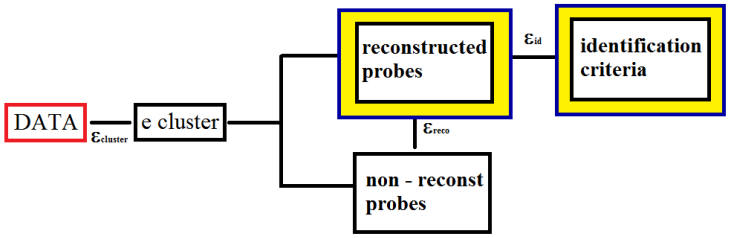


Figure 5.4: Illustrative distributions to show the level of background and its compositions after reconstructed electrons selection in several p_T ranges at $\sqrt{s}$ =7 TeV with $L = 4.7$ fb^{-1} of data [2]. The light jets sample is missing.

5.1.3 Background estimation for electrons with an associated track



The electron reconstruction and identification efficiencies are measured in the Z invariant mass peak region. To estimate the background in this region, in data, a *template* method is used since it has been proved to be more stable in the low p_T regime compared to fits of the invariant mass with an analytic description of the background (i.e. exponential, Gaussian, Breit-Wigner + Crystal Ball functions) [162,219]. The principle of this method is to isolate a sample enriched in fake electrons with the same composition as in the measurement region, normalize it and subtract it from the baseline distribution under the Z peak, assuming the shape of the sample is unchanged.

To build this background template, tag-probe pairs are selected with the probe failing a particular identification criterion. Thanks to the high efficiency and good background rejection of the existing electron identification menus, reverting some cuts in the electron definition allows obtaining a sample where most of the probe candidates are fake electrons. For example, using the *!loose++* criteria, only 5% of the signal electrons are selected, while 95% of the background electrons are retained. As it can be seen in Table 4.2 when failing tighter electron identification more background events can be selected with increased signal contamination. Therefore, the optimal criterion is the best compromise between high background selection and high signal rejection.

Several background templates are used to obtain the electron reconstruction and identification efficiencies :

- *Early Reco* : used to estimate the background for the early reconstruction efficiencies measurement. The probe electron defined in Section 5.1.2 should fail at least two cuts considered in the *loose++* criterion and satisfy the reverse calorimeter isolation requirement, $E_T^{\text{cone40}}/p_T > 0.05$ or 0.20.

- *Early ID* : same as *Early Reco* template. It is used to estimate the background for the early electron identification efficiencies measurement.
- *Preliminary Reco* : if the probe electron is in the $15 < p_T < 30$ GeV region it should fail at least two cuts defining the *multilepton* criteria, otherwise at least two cuts defining the *loose++* criteria. Additional requirements on the calorimeter isolation are considered as well (Table 5.3). This template is used for the preliminary reconstruction efficiency measurement.
- *Preliminary ID* : same definition as *Preliminary Reco* template. It is used to estimate the background for the preliminary electron identification efficiency measurement.

	$p_T < 30$ GeV	$15 \text{ GeV} > p_T > 30$ GeV
1^{st} variation	!2*multileptons* $E_T^{\text{cone30}}/p_T > 0.02$ high edge normalization, 120 GeV $< m_{ee} <$ 250 GeV	!2*loose* $E_T^{\text{cone40}}/p_T > 0.05$ high edge normalization, 120 GeV $< m_{ee} <$ 250 GeV
2^{nd} variation	!2*multileptons* $E_T^{\text{cone30}}/p_T > 0.02$ low edge normalization, 60 GeV $< m_{ee} <$ 70 GeV	!2*loose* $E_T^{\text{cone40}}/p_T > 0.20$ high edge normalization, 120 GeV $< m_{ee} <$ 250 GeV

Table 5.3: Preliminary electron background template definitions and the used normalization.

For the early results same template definitions as in [219] are used. For the preliminary results, more time allows the optimization of the background template to improve the precision of the electron efficiency measurements. As a result, a new template definition requiring the probe electron to fail two cuts defining the *multilepton* working point instead of *loose++* is favored for electrons with p_T inside the [15, 30] GeV interval. This new definition is found to retain more background events and have a smaller bias.

Requirements on the calorimeter isolation are added to minimize the signal contamination (estimated using the *tight++* $Z \to ee$ MC events present in the template[5]). This minimization helps to reduce the systematic uncertainty. Using $E_T^{\text{cone30}}/p_T > 0.02$, the signal contamination is at most 3% in the crack region and smaller than 1% otherwise for electrons with $p_T < 30$ GeV, while for the region with $p_T > 30$ GeV it is smaller than 0.5%.

After the signal subtraction the template is normalized in the side-band (tail) region of the tag and probe electron pair invariant mass which is enriched in fake electrons. As shown in Table 5.3 for the final results in the [15, 30] GeV p_T range the normalization is performed using the low edge of the distribution, 60 GeV $< m_{ee} <$ 70 GeV, or the upper edge, 120 GeV $< m_{ee} <$ 250 GeV. Both intervals are considered to assign the systematic uncertainty. For electrons with $p_T > 30$ GeV and for the early results only the upper tail is used as the signal contamination in the low tail is much larger (and the statistics is much lower as shown in the next section on the m_{ee} distributions).

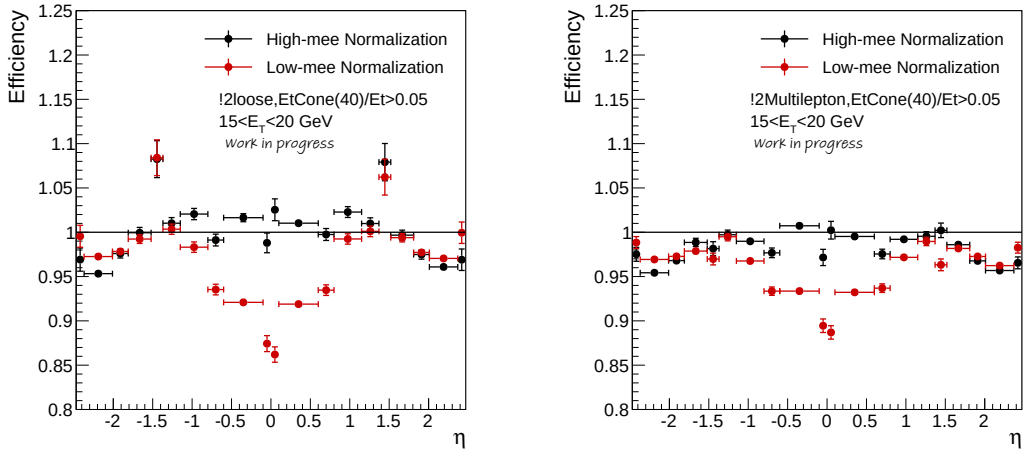


Figure 5.5: Impact of the background template definition on the electron identification efficiency in the $15 < p_T < 20$ GeV interval using $\sqrt{s} = 8$ TeV data. The measured efficiency corresponds to *very loose* LLH criterion [168].

The method described above is used for the electron reconstruction and identification measurements. However, for the identification efficiency measurement, a cut on the electric charge of the electron pair is

[5]The *tight++* MC distribution must be normalized to data luminosity and cross section. The normalization factor is the ratio between the number of *tight++* events in data and in MC, $\frac{N_{Data}^{tight++}}{N_{MC}^{tight++}}$.

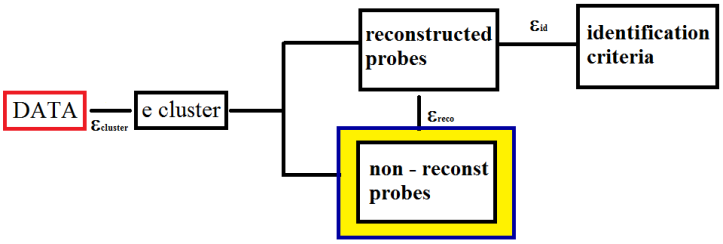
added to further reduce the signal contamination. At the denominator of the efficiency formula the probe pair in the template is opposite - sign, while at the numerator the probe pair is same - sign as the signal contamination is much higher[6].

To estimate the signal in the tail region the number of *tight++* data events, after the background subtraction, is multiplied by the inverse of *tight++* identification efficiency. Figure 5.5 shows the difference in *very loose* likelihood identification efficiency obtained with the preliminary and final template definitions, not passing at least two cuts defining the *loose++* or *multilepton* IDs. Also a comparison between the results obtained with the normalization done using the upper edge and the lower edge of the probe invariant mass is shown. The improvement brought by the changes done in the low p_T range is visible (i.e. no efficiency above 100%). Also the bias in the background template is reduced.

Shared templates When measuring the electron identification efficiencies as a function of the electron η and p_T, a dedicated background template is used in each (η, p_T) interval. As the binning is quite small, sometimes it could lead to a very low statistic in the template. It appears more often in the tail region, where the normalization is performed.

In order to solve this issue, shared background templates are used. These templates use a more coarser binning than the one presented in Table 5.2 (11 bins instead of 20). The results obtained with the coarser binning are used in each (η, p_T) bin within the finer binning. This can be safely applied as the studies using 7 TeV data showed a same background shape (distribution) in close η bins, hence the results will not be changed. This is valid and applied only for probe electrons with $p_T > 30$ GeV.

5.1.4 Background estimation for electrons with no associated track



For the reconstruction efficiency measurement, the background estimation is performed separately for electrons with an associated track and for electrons with no associated track (reconstructed as photons), as the latter component included in the efficiency denominator is not negligible at low p_T. Table 5.4 shows the fraction of photon probes in the baseline distribution before the background subtraction. In the [15, 20] GeV p_T interval it varies between 5% and 7.5%, while at higher intervals it decreases to 0.6 - 1%. Given the high level of background at low p_T, it is clear that a precise estimation is mandatory for accurate measurements.

η / p_T [GeV]	[15,20]	[20,25]	[25,30]	[30,35]	[35,40]	[40,45]	[45,50]	[50,60]	[60,80]	[80,100]
[-0.10,0.10]	0.073	0.048	0.028	0.017	0.011	0.008	0.007	0.009	0.010	0.007
[0.10,0.80]	0.075	0.046	0.028	0.017	0.011	0.008	0.007	0.009	0.009	0.008
[0.80,1.37]	0.053	0.032	0.022	0.015	0.011	0.008	0.007	0.008	0.009	0.008
[1.37,1.52]	0.069	0.040	0.024	0.018	0.013	0.011	0.011	0.011	0.010	0.007
[1.52,2.01]	0.050	0.029	0.021	0.018	0.014	0.013	0.011	0.012	0.011	0.008
[2.01,2.47]	0.063	0.039	0.028	0.023	0.019	0.016	0.015	0.016	0.016	0.010

Table 5.4: Fraction of photons probes in data at denominator before background subtraction, in different bins in η (rows) and p_T (columns) of the probe cluster.

To obtain the number of photons in the baseline distribution, the invariant mass of the tag electron-probe photon pair is fitted with a 3^{rd} degree polynomial[7] :

$$f(x) = A_1x^3 + A_2x^2 + A_3x + A_4 \tag{5.2}$$

[6]Studies based on 7 TeV data showed a smaller contamination in signal electrons, up to 0.1% in the low p_T region and much smaller otherwise, if the opposite electric charge sign (OS) pair is replaced by same electric charge pairs (SS), retaining the other requirements unchanged with respect to the baseline selection.

[7]No component for the signal is included as the function aims at photon background modeling.

The fit is performed in the side band region of the invariant mass distribution enriched in background events. To probe the stability of the fit, four different fit intervals are defined (as it will be discussed in Section 5.2). The fit range is divided in intervals of 5 GeV.

The corresponding χ^2 used to fit the data[8] is written as :

$$\chi^2 = \Sigma(N_i - f(x_i))^2 \tag{5.3}$$

where x_i is the value of the invariant mass at the center of (η, p_T) bin with index i ; N_i is the average number of events per GeV in this bin and corresponds to the difference between the observed data and estimated number of signal events in MC[9] :

$$N_i = \frac{N_i^{Data} - N_i^{MC}}{5 \text{ GeV}} \tag{5.4}$$

The four coefficients used in Equation (5.2) are computed when the χ^2 is minimized. Since χ^2 depends quadratically on the mentioned parameters, the system can be written in term of linear equation. Using the following notation :

$$M := \begin{pmatrix} S_{60} & S_{50} & S_{40} & S_{30} \\ S_{50} & S_{40} & S_{30} & S_{20} \\ S_{40} & S_{30} & S_{20} & S_{10} \\ S_{30} & S_{20} & S_{10} & S_{00} \end{pmatrix} \quad \text{with} \quad S_{jk} := \sum_{j,k} x_i^j N_i^k, \quad j = \overline{0,6}, \quad k = 0,1. \tag{5.5}$$

the system becomes :

$$\begin{pmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \end{pmatrix} = M^{-1} \cdot \begin{pmatrix} S_{31} \\ S_{21} \\ S_{11} \\ S_{01} \end{pmatrix} \tag{5.6}$$

The number of background events in a given interval $[x^{\min}, x^{\max}]$ is obtained by integrating the polynomial function, f, in that range :

$$N_{bkg} = \sum_{k=1}^{4} \frac{1}{k} A_k (x^{\max} - x^{\min}) \tag{5.7}$$

Given the choice of χ^2, the matrix M depends only on the interval in which the fit is performed. Hence, only 4 (exactly the number of fit intervals) instances of the matrix are defined and shared in all bins and for all systematic variations. This translates into a fast computation, in terms of running time. Also, the uncertainties can be written according to the uncertainty propagation law (Section 5.2.1).

Figure 5.6 presents the fit performance for the photon background estimation. Here the invariant mass of the tag electron-probe photon pair is shown for data and data minus signal MC events. The fit range is displayed with blue filled area. The obtained results present a stable fit in each $m_{e\gamma}$ fit interval.

The estimated photon background fraction is shown for three selections on the tag electron and invariant mass of the tag electron-probe photon pair:

- Nominal : *tight++*, 80 GeV $< m_{e\gamma} <$ 100 GeV, Table 5.5.
- Loose : *medium++* + $E_T^{\text{cone40}} < 5$ GeV, 70 GeV $< m_{e\gamma} <$ 105 GeV, Table 5.6.
- Tight : *tight++* + $E_T^{\text{cone40}} < 5$ GeV, 85 GeV $< m_{e\gamma} <$ 95 GeV, Table 5.7.

Results for the last two selections are shown for comparison and used to assign the systematic uncertainty. The background fraction varies between 85% and 100% for the nominal and loose selection in the region with soft probes, p_T <20 GeV, decreasing to 12% in the [40,45] GeV p_T range. For the tightest selection, it varies between 70% and 100% in the [15,20] GeV p_T interval and decreases down to 6% in the high p_T ranges.

5.2 Electron reconstruction efficiency

The electron reconstruction efficiency is measured in [15, 150] GeV p_T range using the $Tag\&Probe$ method with $Z \to ee$ events. The methodology followed is similar to the one used for the final $\sqrt{s} = 7$ TeV measurements [219].

[8] Name given by analogy, hence the variable is not following the χ^2 law.

[9] The number of events in MC is obtained as in data, i.e. not using the truth information, and are normalized to data luminosity and cross section.

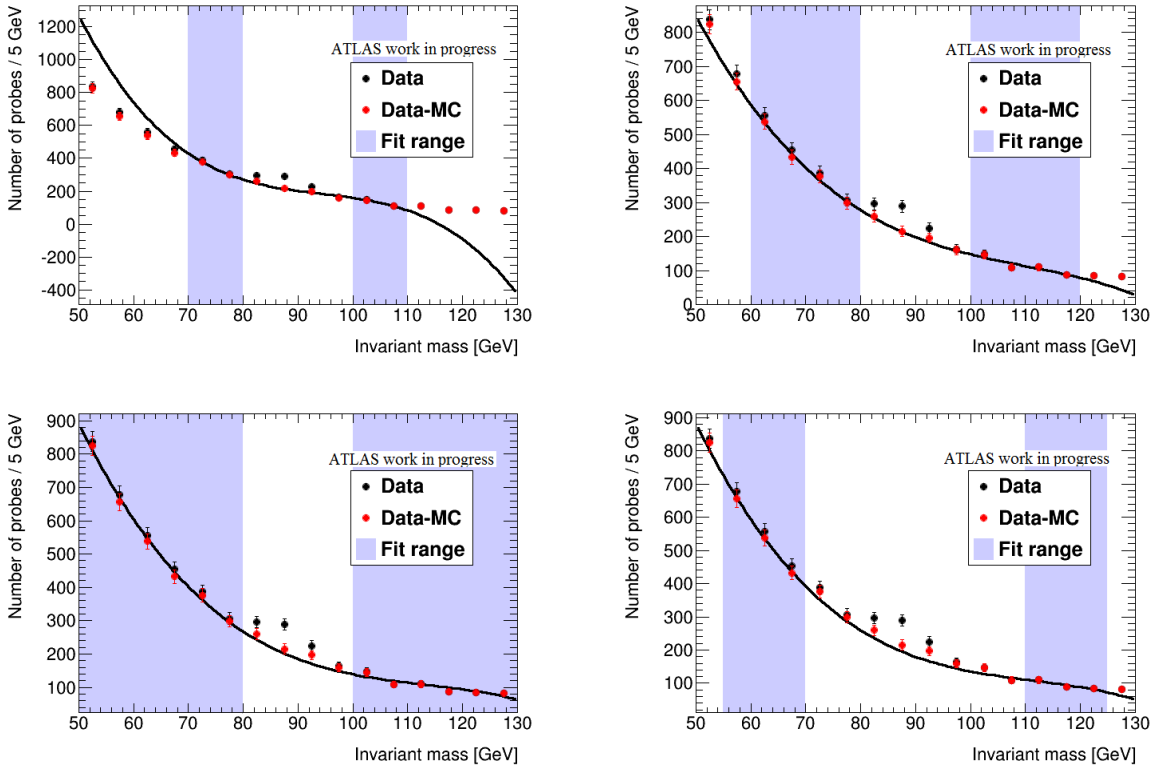
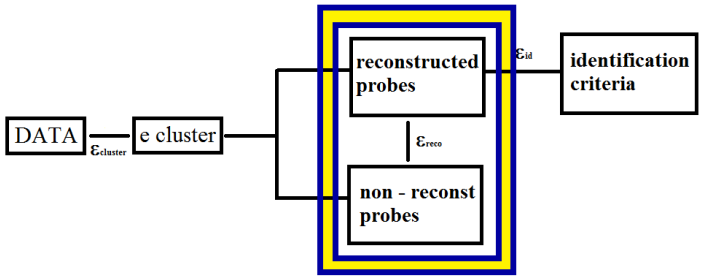


Figure 5.6: Performance for the photon background estimation: fits with 4 configurations in the [15,20] GeV E_T and [1.52,2.01] η bin. $\sqrt{s} = 8$ TeV and $L = 20.3$ fb^{-1}. The fit is only carried out and constrained in the blue shaded areas. This explains the deviations of the polynomial fit from the smoothly falling invariant mass shape for the data above and below the fit ranges. Within the signal region of an invariant mass of 80 to 100 GeV the fit is however well behaved due to the constraints on both sides of the signal region. Only the statistical uncertainty is shown.



The efficiencies are computed using already reconstructed electron and photon clusters (denominator), and measure the electron track reconstruction performance and its association with the cluster. This assumption is validated by the fact that, in MC, the cluster reconstruction is ∼100% for electrons with p_T >15 GeV and any η. The formula used for the measurement is :

$$\varepsilon_{reco} = \frac{N^{Pass} - (B^{Pass})_{nume}}{(N^{Pass} + N^{Fail}) - [(B^{Pass})_{deno} + (B^{Fail})_{deno}] + \Gamma} \tag{5.8}$$

where N^{Pass} and N^{Fail} stand for the number of reconstructed probe electrons in the peak region that pass or fail the track quality cut and the B^{Pass} and B^{Fail} terms correspond to the number of background electrons. The tag electron fulfills the selection defined in Section 5.1.2. The number of signal photons, Γ term, is the difference between the number of events in the peak region of the electron background template and the number of background photons estimated using the polynomial fit (Section 5.1.4).

It should be noted that the electron background is estimated separately in the sample with the probe electrons passing the required level cut and in the sample with probe electrons failing the cut.The photon background subtraction is performed independently on the fake electron subtraction. This approach aims at handling correctly the correlations between the numerator and the denominator.

In Section 5.2.1 the measurement is presented, while the results obtained with *preliminary Reco* background template are shown in Section 5.2.2.

η / p_T [GeV]	[15,20]	[20,25]	[25,30]	[30,35]	[35,40]	[40,45]
[-0.10,0.10]	100.00	96.17	77.72	67.22	42.71	35.71
[0.10,0.80]	92.59	89.42	79.46	63.89	42.64	31.69
[0.80,1.37]	94.65	76.73	62.95	45.66	38.85	25.15
[1.37,1.52]	94.16	77.53	52.88	28.17	31.16	18.60
[1.52,2.01]	84.29	74.43	48.98	29.39	21.77	12.79
[2.01,2.47]	84.87	61.41	48.29	35.90	18.44	19.06

Table 5.5: *tight++* tag, 80 GeV $< m_{e\gamma} <$ 100 GeV: estimated photon background fraction (in %) in data in different bins in η (rows) and p_T (columns) of the probe cluster.

η / p_T [GeV]	[15,20]	[20,25]	[25,30]	[30,35]	[35,40]	[40,45]
[-0.10,0.10]	100.00	92.69	83.94	71.62	52.61	43.03
[0.10,0.80]	96.39	97.97	83.68	74.37	50.08	39.58
[0.80,1.37]	95.87	83.28	74.82	55.46	47.12	30.39
[1.37,1.52]	94.45	89.99	63.88	37.70	36.92	26.49
[1.52,2.01]	86.45	80.42	56.99	41.76	32.05	20.14
[2.01,2.47]	88.09	63.55	55.63	41.46	26.39	23.85

Table 5.6: *medium++* + $E_T^{\text{cone40}} < 5$ GeV on the tag, 70 GeV $< m_{e\gamma} <$ 105 GeV: estimated photon background fraction (in %) in data in different bins in η (rows) and p_T (columns) of the probe cluster.

η / p_T [GeV]	[15,20]	[20,25]	[25,30]	[30,35]	[35,40]	[40,45]
[-0.10,0.10]	100.00	98.26	68.62	46.81	30.37	19.58
[0.10,0.80]	90.33	86.22	68.54	50.75	27.14	17.88
[0.80,1.37]	89.74	69.69	52.17	32.59	26.84	13.01
[1.37,1.52]	87.79	81.91	42.70	13.17	23.49	9.62
[1.52,2.01]	69.94	56.89	32.94	19.82	14.36	6.13
[2.01,2.47]	71.54	42.26	30.57	20.13	10.66	8.91

Table 5.7: *tight++* + $E_T^{\text{cone40}} < 5$ GeV on the tag, 85 GeV $< m_{e\gamma} <$ 95 GeV: estimated photon background fraction (in %) in data in different bins in η (rows) and p_T (columns) of the probe cluster.

5.2.1 Measurement

To measure the electron efficiency, Equation (5.8) is used and the number of background electrons in data is estimated using a slightly modified *preliminary Reco* electron background template :

- Probe electron failing the track quality cuts : the probe electron should fail any two cuts defining the *multilepton* or *loose++* working point and also the track quality cuts ($TemplF$).
- Probe electron passing the track quality cuts : the probe electron should fail any two cuts defining the *multilepton* or *loose++* working point but it should pass the track quality cuts ($TemplP$).

Changing the initial definitions helps in creating more stable measurements when changing the level of background, which translates into smaller systematic uncertainties.

The number of background electrons defined by B^{Pass} terms at the numerator and denominator of reconstruction efficiency formula is obtained using the relations :

$$(B^{Pass})_{nume} = B^{TemplP} \cdot \frac{B_{tail}^{BaseP}}{B_{tail}^{TemplP}} \tag{5.9}$$

$$(B^{Pass})_{deno} = B^{TemplP} \cdot \frac{B_{tail}^{BaseP} - N_{tail}^{tight++} \cdot 1/\varepsilon_{tight++}}{B_{tail}^{TemplP}} \tag{5.10}$$

and similar for B^{Fail} where B_{tail}^{BaseP} is replaced by B_{tail}^{BaseF} and B^{TemplP} by B^{TemplF}.

- The B^{TemplX} and B_{tail}^{TemplX} ($X = P$ of F) terms represent the number of background electrons passing or failing the track quality cut in the corresponding template under the peak or in the tail region.

- B_{tail}^{BaseX} ($X = P$ of F) term stands for the number of events in the baseline distribution (Section 5.1.2) for which the probe is passing or failing the track quality cut in the tail region.

- $N_{tail}^{tight++}$ is the number of *tight++* data events in the tail region after the background subtraction; $\varepsilon_{tight++}$ stands for the *tight++* identification efficiency, measured as explained in the next section.

Further, only results using the *preliminary Reco* electron background template are presented, as they are well improved compared to the *early Reco* results (Section 5.1.3). Using the selections presented in Section 5.1.4 the fractions of electron background events at the numerator and electron + photons background events at the denominator are shown in Appendix A, Section A.1. Generally, depending on the (η, p_T) interval and on the selection type, the background fraction at the numerator or denominator varies between 84% and 1.5%. A high level of background (between 50 and 80%) is obtained especially at low p_T (< 25 GeV). If the track quality cut is removed a similar level of background is obtained (Appendix A.2).

The good performance of the electron and photon background estimation is shown in Figure 5.7 for EM cluster probes passing the track quality cut[10]. Distributions are shown for several (η, p_T) intervals. The invariant mass computed with the tag electron and EM cluster probe is presented in data with a yellow area and in MC with a red area. The invariant mass computed with the tag electron and EM cluster probe satisfying the *tight++* ID criteria (cross markers) is also shown for comparison. Finally the electron background template is shown with a blue line and the distribution of the background template + signal $Z \to ee$ MC with a red line.

Figure 5.8 shows the data distribution of the invariant mass computed with the tag electron and a EM cluster probe (yellow area) or with a photon probe (green area), in several η and p_T bins at the denominator of the reconstruction efficiency. The photon background fit in the side-band band regions is displayed with blue lines. The peak above the blue lines is formed by electrons which were reconstructed as photons. They are considered as electrons in the measurement. This set of plots illustrates the high level of background at denominator and underlines the importance of a good background estimation for the efficiency measurements.

Statistical uncertainty

The statistical uncertainty is computed using the error propagation formula. For this, Equation (5.8) is rewritten in terms of independent (**uncorrelated**) variables :

$$\begin{aligned}
\varepsilon_{reco} &= \frac{N^{Pass} - B^{TemplP} \cdot B_{tail}^{BaseP}/B_{tail}^{TemplP}}{N^{Pass} - B^{TemplP} \cdot B_{tail}^{BaseP}/B_{tail}^{TemplP} + N^{Fail} - B^{TemplF} \cdot B_{tail}^{BaseF}/B_{tail}^{TemplF} + \Gamma} \\
&= \frac{(N^P B_{tail}^{TemplP} - B^P B^{TemplP}) \cdot B_{tail}^{TemplF}}{(N^P B_{tail}^{TemplP} - B^P B^{TemplP}) \cdot B_{tail}^{TemplF} + (N^F B_{tail}^{TemplF} - B^F B^{TemplF}) \cdot B_{tail}^{TemplP} + \Gamma B_{tail}^{TemplP} B_{tail}^{TemplF}} \qquad (5.11) \\
&= (N^P B_{tail}^{TemplP} - B^P B^{TemplP}) \cdot B_{tail}^{TemplF}/\text{deno} \qquad (5.12)
\end{aligned}$$

where $N^P = N^{Pass} - B^{TemplP}$, $B^P = B_{tail}^{BaseP} - B_{tail}^{TemplP}$, and the corresponding *fail* terms $N^F = N^{Fail} - B^{TemplF}$, $B^F = B_{tail}^{BaseF} - B_{tail}^{TemplF}$.

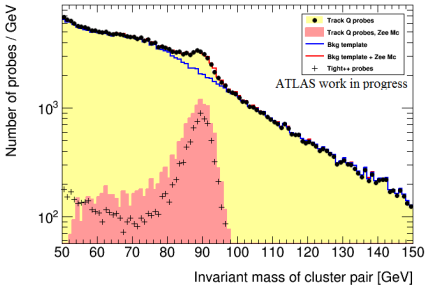
The statistical uncertainty associated to ε_{reco} is :

$$\begin{aligned}
\Delta\varepsilon_{reco}^2 = &\frac{[B_{tail}^{TemplP} \cdot (N^P B_{tail}^{TemplP} - B^P B^{TemplP})]^2}{deno^4} \times \{(B^F B^{TemplF} \Delta B_{tail}^{TemplF})^2 + \\
&(B_{tail}^{TemplF})^2 \cdot [(B_{tail}^{TemplF} \Delta B^F)^2 + (B_{tail}^{TemplF} \Delta\Gamma)^2] + (B^{TemplF} \Delta B^F)^2 + (B^F \Delta B^{TemplF})^2\} + \\
&\frac{[B_{tail}^{TemplF} \cdot (N^F B_{tail}^{TemplF} - B^F B^{TemplF} + \Gamma B_{tail}^{TemplF})]^2}{deno^4} \times \{(B^P B^{TemplP} \Delta B_{tail}^{TemplP})^2 + \\
&(B_{tail}^{TemplP})^2 \cdot [(B_{tail}^{TemplP} \Delta N^P)^2 + (B^{TemplP} \Delta B^P)^2 + (B^P \Delta B^{TemplP})]^2\} \qquad (5.13)
\end{aligned}$$

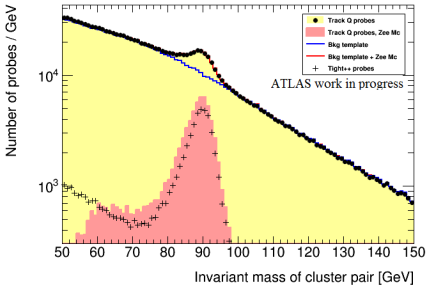
where the corresponding ΔY ($Y = B_{tail}^{TemplF}$, ...) is $\sqrt{\sum_i w_i^2}$, where i is the event index and w_i the corresponding weight. For the photon term, w_i is the number of observed events in each invariant mass bin (N_i in Section 5.1.4).

[10] See Figure A.2 in Appendix A.2 for the EM cluster probes failing the cut.

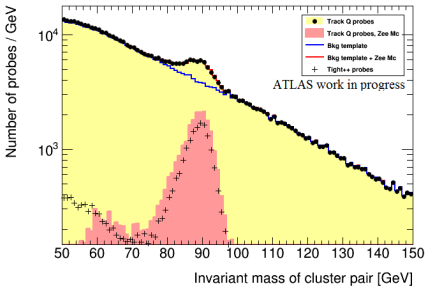
(a) $15 < p_T < 20$ GeV, η bin [0.0,0.1]



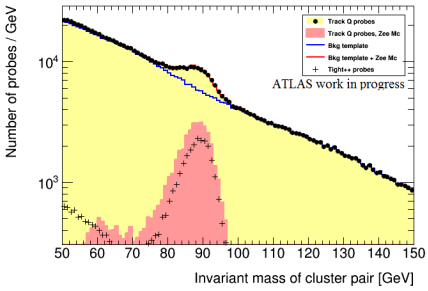
(b) $15 < p_T < 20$ GeV, η bin [0.1,0.6]



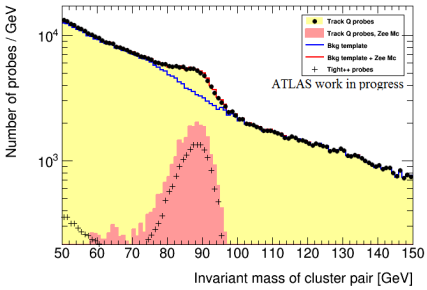
(c) $15 < p_T < 20$ GeV, η bin [0.6,0.8]



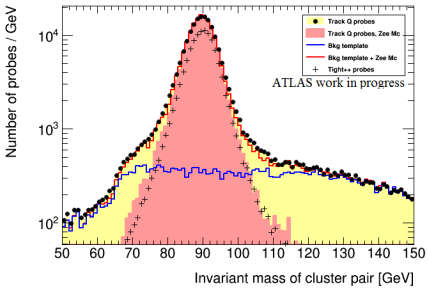
(d) $15 < p_T < 20$ GeV, η bin [0.8,1.15]



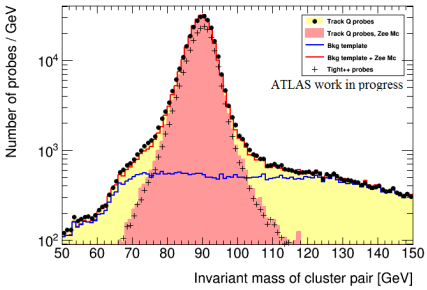
(e) $15 < p_T < 20$ GeV, η bin [1.15,1.37]



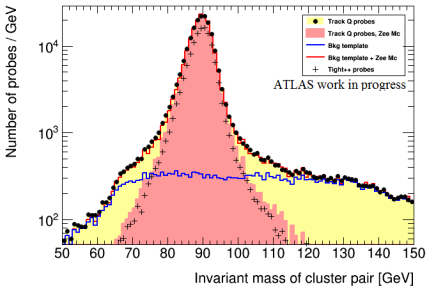
(f) $40 < p_T < 45$ GeV, η bin [1.37,1.52]



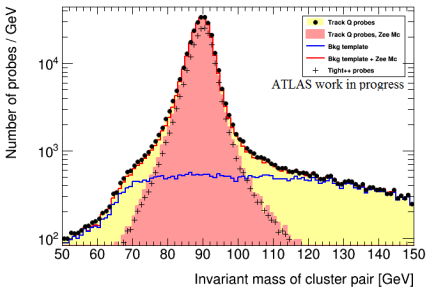
(g) $40 < p_T < 45$ GeV, η bin [1.52,1.81]



(h) $40 < p_T < 45$ GeV, η bin [1.81,2.01]



(i) $40 < p_T < 45$ GeV, η bin [2.01,2.37]



(j) $40 < p_T < 45$ GeV, η bin [2.37,2.47]

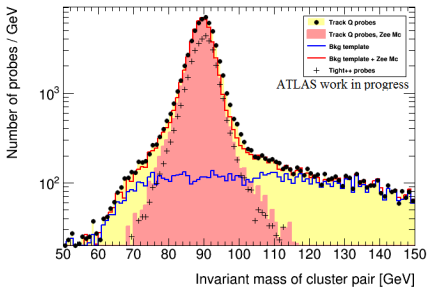
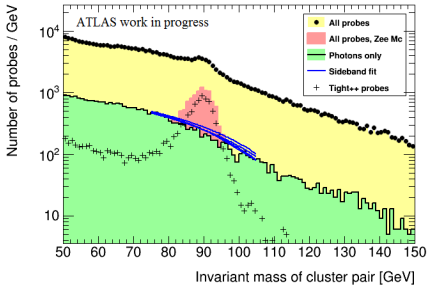
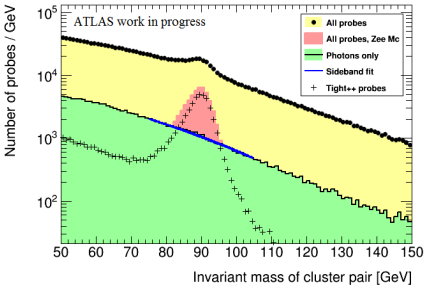


Figure 5.7: Invariant mass distribution computed with the tag electrons and EM cluster probe passing the track quality cuts (numerator level) in several (η, p_T) bins. $\sqrt{s} = 8$ TeV and $L = 20.3$ fb^{-1}.

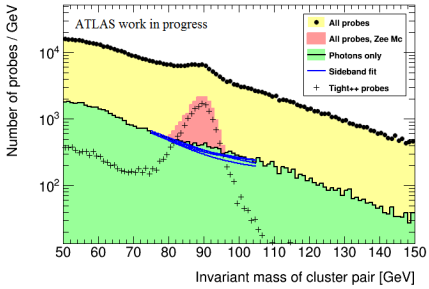
(a) $15 < p_T < 20$ GeV, η bin [0.0,0.1]



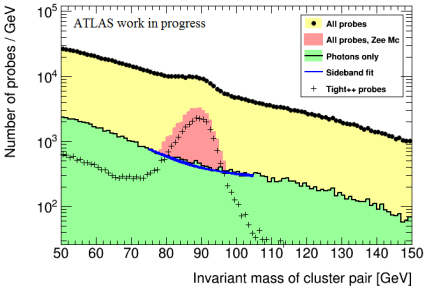
(b) $15 < p_T < 20$ GeV, η bin [0.1,0.6]



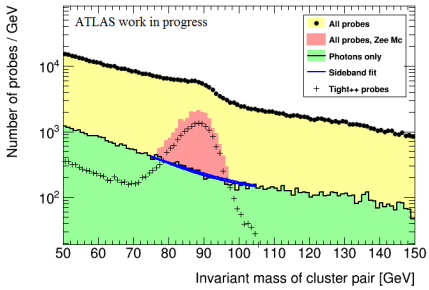
(c) $15 < p_T < 20$ GeV, η bin [0.6,0.8]



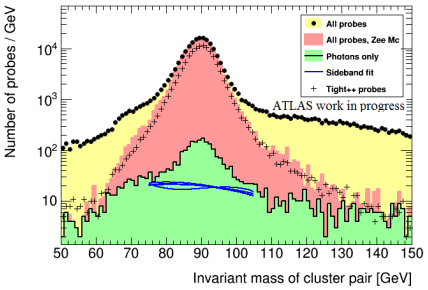
(d) $15 < p_T < 20$ GeV, η bin [0.8,1.15]



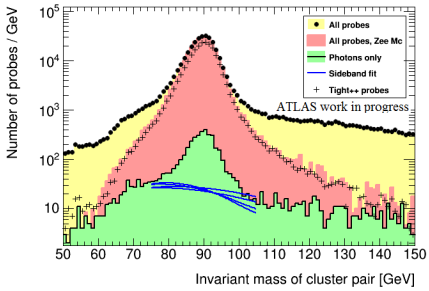
(e) $15 < p_T < 20$ GeV, η bin [1.15,1.37]



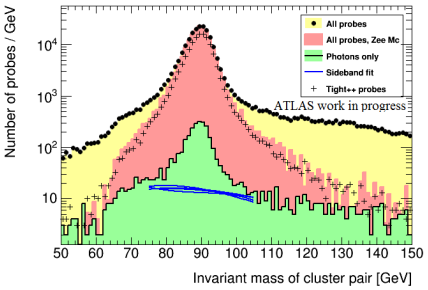
(f) $40 < p_T < 45$ GeV, η bin [1.37,1.52]



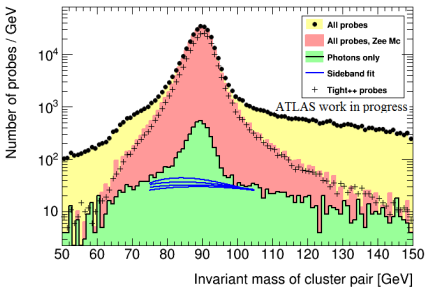
(g) $40 < p_T < 45$ GeV, η bin [1.52,1.81]



(h) $40 < p_T < 45$ GeV, η bin [1.81,2.01]



(i) $40 < p_T < 45$ GeV, η bin [2.01,2.37]



(j) $40 < p_T < 45$ GeV, η bin [2.37,2.47]

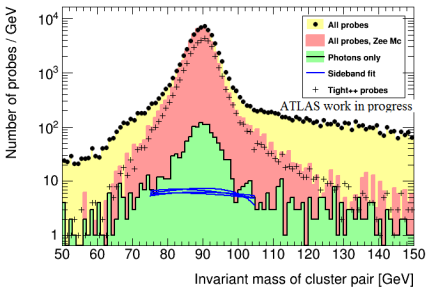


Figure 5.8: Invariant mass distribution computed with the tag electron and EM cluster probe (denominator level) in several (η, p_T) bins. $\sqrt{s} = 8$ TeV and $L = 20.3$ fb^{-1}.

Systematic uncertainties

The sources of systematic uncertainties are assessed by varying the relevant parameters of the measurement. The global systematic uncertainty is the RMS of the distribution of variations obtained with each configuration. The central value of the reconstruction efficiency is the average of this distribution. The considered variations for the electron terms are :

- Tag identification : *tight++* changed to *medium++* + $E_T^{\text{cone40}} < 5$ GeV or *tight++* + $E_T^{\text{cone40}} < 5$ GeV.
- The invariant mass, peak region : [80,100] GeV changed to [75,105] GeV or [70,110] GeV.
- The electron background template definitions (Table 5.3).

As already explained, the pairs in the tail region of the invariant mass at the denominator of the electron identification efficiency formula should have the same electric charge. This might yield a different composition with respect to the baseline distribution, which is formed by pairs with opposite electric charge. To account for this difference, an additional source of systematic uncertainty (10%) is added. This value is the difference between the number of OS and SS pairs of probe electrons without any requirement on the identification criteria. Studies performed with 7 TeV data [2, 219] showed a negligible impact on the total systematic uncertainty. At low p_T, the side-band region used to normalize the background template is changed to account for the variation in background composition and shape. At high p_T, the low edge could not be used anymore as the level of background is small and the signal contamination is highly increased, leading to a higher bias in the template.

The considered variations for the photon term are :

- The fit range : [70,80] ∪ [100,110] GeV, [60.80] ∪ [100,120] GeV, [50,80] ∪ [100,130] GeV and [55,70] ∪ [110,125] GeV.

A total of 72 ($3 \times 3 \times 2 \times 4$) variations are obtained for data and 9 (3×3) for MC. For the latter the variations on the template are not considered, as no background subtraction is done. In MC the particle truth information is used to discriminate the background from the signal electrons.

5.2.2 Results using $Z \to ee$ events

Figure 5.9 shows the efficiency to reconstruct an electron with a good track quality in data and MC. For illustration, results are shown in all (η, p_T) bins so that the evolution of the total uncertainties and central values is more visible.

The efficiency to reconstruct an electron associated to a track of good quality varies from 95% to 99% between end-cap and barrel regions for electrons with $p_T < 20$ GeV. Lower efficiencies are obtained in the crack region, where they decrease to 94%. For very energetic electrons ($p_T > 80$ GeV) the efficiency is around 99% over the whole η range. Same conclusions are valid also for the results at the reconstruction stage with no track quality cut, with one difference : higher efficiencies are obtained in the $1.37 < |\eta| < 2$ region where the amount of material is very large (Figure 3.7). These results are presented in Appendix A and are obtained to see the impact of adding the requirement of a good track quality on the electron.

The good agreement between the results at reconstruction and reconstruction + track quality obtained in data and in MC gives confidence in the MC description of the detector response for electrons with $p_T < 15$ GeV. The total associated uncertainty is smaller than 5% for electrons with $p_T < 20$ GeV. For more energetic electrons it varies between 0.5% and 1.5% except in the transition region where it could reach up to 3.3%. The statistical and systematic uncertainties are of the same order.

For 2012 data taking, a new tracking fitting algorithm was developed in order to improve the reconstruction of electrons subject to hard bremsstrahlung [160]. Figure 5.10 left shows the integrated over p_T results obtained with the old (2011, $\sqrt{s} = 7$ TeV) and new algorithms (2012, $\sqrt{s} = 8$ TeV). An overall improvement of 5% is observed when using $Z \to ee$ events, especially visible in the low p_T range (7%). This leads to an important gain in the SM and BSM measurements, especially for soft electrons or very high η ranges, see eg. the results of $H \to ZZ \to llll$ channel [10]. Figure 5.10 right shows the reconstruction efficiency integrated over the full pseudo-rapidity for both algorithms. The drop in the $1.37 < |\eta| < 2$ region, where more bremsstrahlung occurs due to a higher amount of material, is well reduced compared to 7 TeV results.

Other quantities needed by the physics groups are the corrections which should be applied to the simulation to reproduce the data. These, often called scale factors, are the ratio between the electron efficiency in data and the efficiency in simulations. Results are shown in Table A.10 (Appendix A) using 50 bins in η for a detailed view on the scale factors dependency[11]. Generally, the scale factors are close to unity (Figure 5.9 lower panel and Figure A.4 in Appendix A) with a total uncertainty smaller than 0.3% for electrons with $p_T > 30$ GeV and varying between 0.03 and 2% in the low p_T intervals.

[11] For the final recommendations released to physics groups the binning presented in Table 5.2 is used.

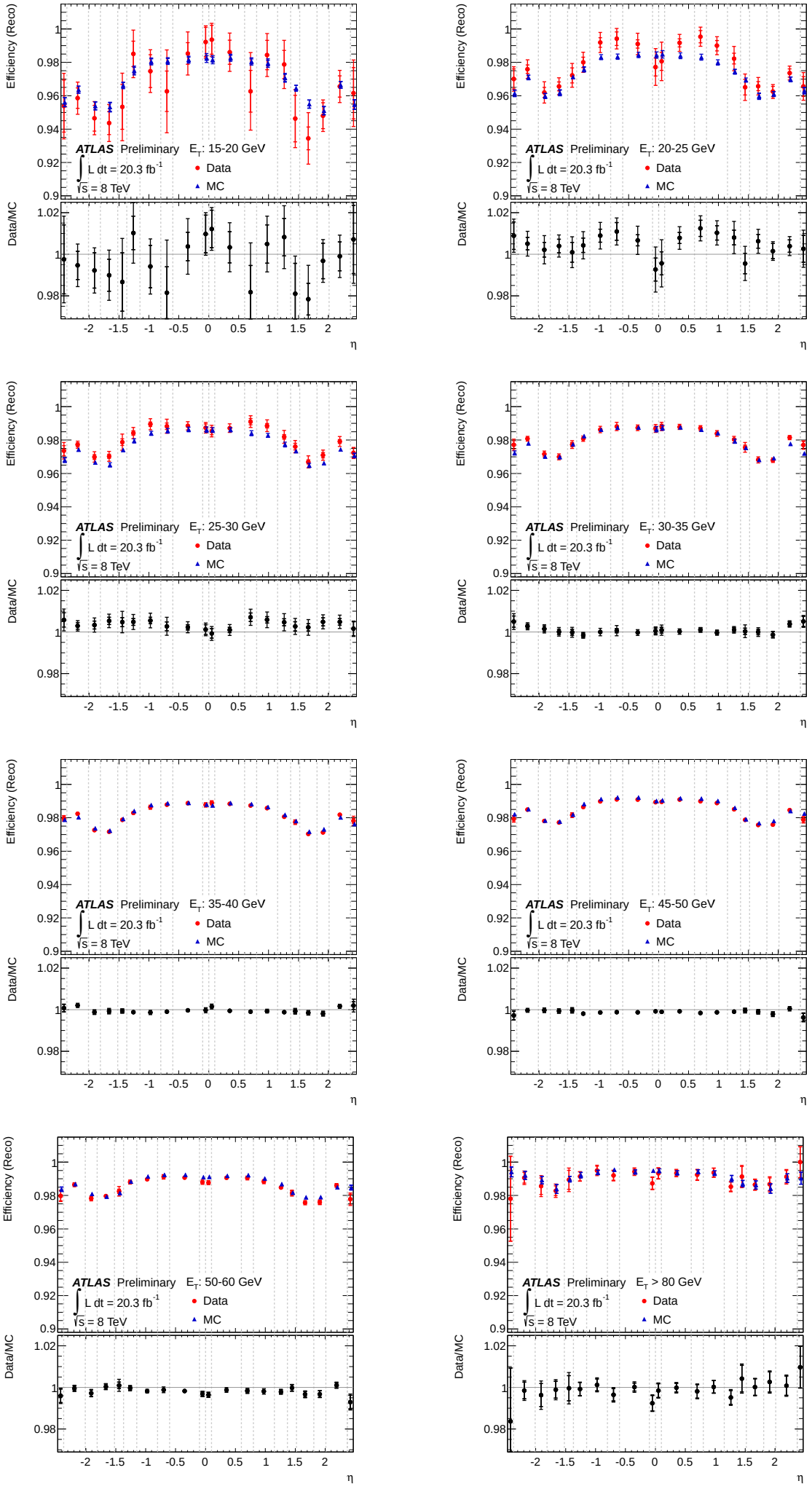


Figure 5.9: Electron reconstruction + track quality efficiencies in Data (red markers) and MC (blue markers) in all considered (η, p_T) bins. Only combined statistical and systematic uncertainties are shown. In the lower panel the data to MC ratio is shown.

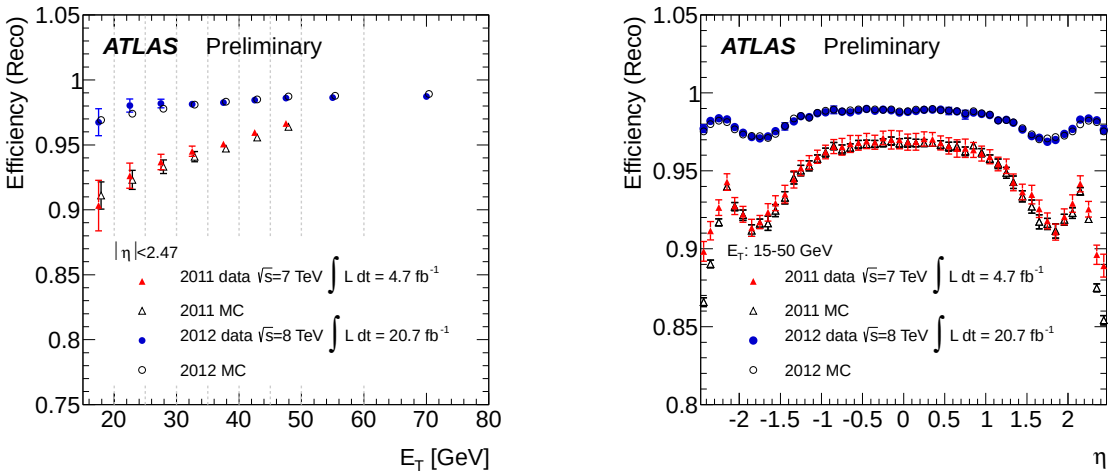


Figure 5.10: Electron reconstruction + track quality efficiencies in Data (filled markers) and MC (empty markers) as function of electron η in [15,50] GeV p_T range (left) and integrated over the full pseudorapidity (right). Results obtained with 2011 data ($\sqrt{s}$ = 7 TeV, red markers) and with 2012 ($\sqrt{s}$ = 8 TeV, blue markers) are shown. Total uncertainty is considered.

Pileup dependency

Figure 5.11 shows the reconstruction efficiency, with and without a track quality cut, as a function of the number of reconstructed primary vertices for $15 < p_T < 50$ GeV and integrated over η. Data results are compared to MC prediction and a flat distribution can be observed in both cases (no dependency of the electron performance to pileup).

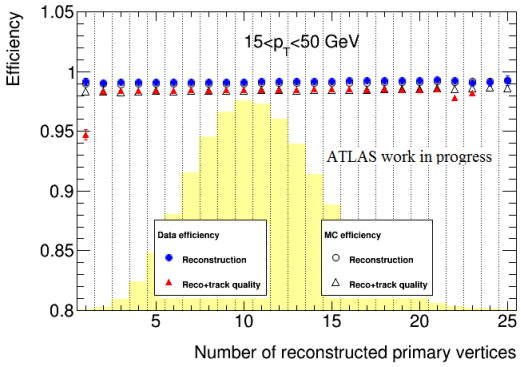
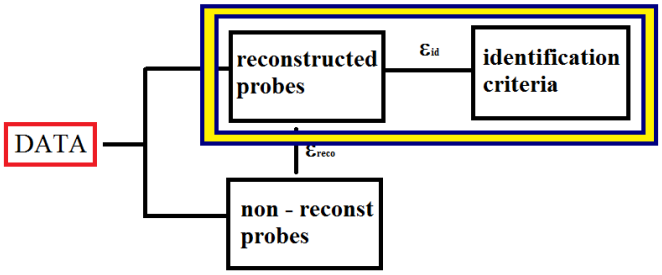


Figure 5.11: Reconstruction efficiencies as a function of number of reconstructed primary vertices with their total uncertainties measured in data and MC for $15 < p_T < 50$ GeV and integrated over η. The yellow shaded area shows the distribution of primary vertices for the full 2012 data set.

5.3 Electron identification efficiency



Electron identification efficiencies are measured in different sample sources depending on the p_T interval : the $J/\Psi \to ee$ processes are used for $7 < p_T < 20$ GeV range and the $Z \to ee$ events for $p_T > 10$ GeV. This division is mandatory because in the low p_T interval the jets faking electrons background is very high and a clean soft electron Z sample cannot be designed. Section 5.3.1 presents the measurement performed

with $Z \to ee$ events and in Section 5.3.2 the preliminary results are shown. Complementary measurements used to cover the low p_T range or used to assign the systematic uncertainties are detailed in Section 5.3.3. The combined results obtained with the different methods are shown in Section 5.3.4.

5.3.1 Identification efficiency measurement using $Z \to ee$ events

In this section only the measurements performed with $Z \to ee$ events are presented (also called the *Z mass method*). The key variable is the invariant mass of the 2 selected electrons and the relation used to compute the electron identification efficiency, ε_{id}, is :

$$\varepsilon_{id} = \frac{N^{Pass} - (B^{Base})_{nume}}{N^{Base} - (B^{Base})_{deno}} \tag{5.14}$$

where N^{Pass} is the number of events in the peak region of the invariant mass for which the probe is passing the required definition, N^{Base} is the number of events in the peak region, in the baseline distribution (the numerator of reconstruction + track quality measurement), and B^{Base} is the number of background events at the numerator (*nume*) and denominator (*deno*) in the peak region, in the baseline distribution. The electron identification efficiency is measured only with respect to reconstructed electrons with an associated track, hence no photon estimation is necessary.

Compared with the reconstruction efficiency measurements, the background estimation is performed independently at the numerator and denominator. This strategy holds because the background level at numerator is not significant, hence the correlations between the estimation at numerator and denominator are negligible.

Electron background estimation

Using the same approach as for the reconstruction efficiencies measurement, the background terms are written as :

$$(B^{Base})_{nume} = B^{Templ} \cdot \frac{B^{Base,SS}_{tail}}{B^{Templ}_{tail}} \tag{5.15}$$

$$(B^{Base})_{deno} = B^{Templ} \cdot \frac{B^{Base,OS}_{tail} - N^{tight++}_{tail} \cdot 1/\varepsilon_{tight++}}{B^{Templ}_{tail}} \tag{5.16}$$

To obtain the B^{Templ} term, which is the number of background events in the template, the *Early ID* or *Preliminary ID* (Section 5.1.3) template definition is considered. For the identification measurements, both early and preliminary results are presented. The difference at low p_T (< 30 GeV) between the two sets is within the uncertainties, while at high p_T no improvement is reached when the same template definition is used.

The good background template performance is illustrated in Figure 5.12 for one particular η - p_T bin, 20 GeV $< p_T <$ 25 GeV and $0.1 < \eta < 0.6$. Here only the final template definition is considered and the normalization is performed using the upper edge of the invariant mass. In the left plot, the invariant mass distribution computed with the electron probe pair at the denominator is shown in data (black points) and MC (blue line). The electron background template distribution is presented with magenta markers. The expected distribution (template + MC) shown with a red line match very well the data points. In the right plot, the distribution at the numerator is shown for all *tight++* probes in data and MC. The expected distribution (template + MC) is also in agreement with the data.

Statistical uncertainty

Using the error propagation formula, the statistical uncertainty is written as :

$$\Delta\varepsilon_{id} = \frac{1}{\Delta N_{deno}} \sqrt{(1 - 2\varepsilon_{id})\Delta N^2_{nume} + \varepsilon^2_{id}\Delta N^2_{deno}} \tag{5.17}$$

where N_{nume} and N_{deno} stands for the numerator and denominator terms of Equation (5.14) respectively. They can be written as a function of uncorrelated variables :

$$N_{nume} = N^{Pass} - B^{Templ} \cdot \frac{B^{Base}_{tail}}{B^{Templ}_{tail}} \tag{5.18}$$

$$N_{deno} = N^{Base} - B^{Templ} \cdot \frac{B^{Base}_{tail}}{B^{Templ}_{tail}} = N - B^{Templ} \cdot \frac{B}{B^{Templ}_{tail}} \tag{5.19}$$

where

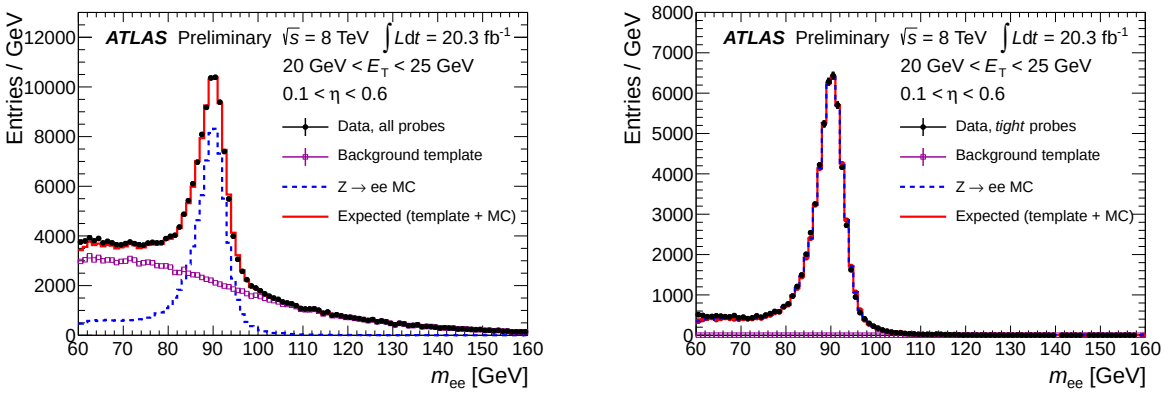


Figure 5.12: Performance of the final electron background template used for the electron identification measurements in one particular bin, 20 GeV $< p_T <$ 25 GeV and $0.1 < \eta < 0.6$. The invariant mass distribution at the denominator is shown on the left, while in the right side is shown at the numerator [168].

- $N = N^{Base} - B^{Templ}$ is the difference between the number of events in the baseline distribution and in the background template under the Z mass peak.
- $B = B^{Base} - B^{Templ}_{tail}$ is the difference between the number of events in the baseline distribution and in the template in the side band region of the invariant mass distribution.

Note that the terms are rewritten only at the denominator, as no correlation is present at the numerator. Indeed, the overlap between the baseline distribution and the template is zero because the probe electron in the template are failing the *loose++* or *multilepton* identification criterion, whereas in the baseline distribution they should pass at least the loose selection used to define it. Furthermore, the template is normalized in the tail region with same sign lepton pairs, while the baseline distribution is populated with opposite sign lepton pairs.

Finally, the uncertainty on the input parameters is :

$$\Delta N_{deno} = \sqrt{N + (\frac{B^{Templ} B}{B^{Templ}_{tail}})^2 \cdot (\frac{1}{B^{Templ}} + \frac{1}{B^{Templ}_{tail}} + \frac{1}{B})} \tag{5.20}$$

And similarly :

$$\Delta N_{nume} = \sqrt{N^{Pass} + (N_{bkg})^2 \cdot (\frac{1}{B^{Templ}} + \frac{1}{B^{Templ}_{tail}} + \frac{1}{B^{baseline}_{tail}})} \tag{5.21}$$

where N_{bkg} is the number of estimated background events using the template with SS leptons.

Systematic variations

Same variations as presented in Section 5.2.1 for the reconstruction efficiency are retained. A total of 18 data efficiencies and scale factors are computed in each (η, p_T) bin. For MC a total of 9 sets of results are obtained, as no background template variation is performed.

5.3.2 Early efficiencies at the identification stage

The results obtained with the *Early ID* background template definition are now discussed. In Figure 5.13, the electron identification efficiencies are shown in data and MC for *loose++*, *medium++* and *tight++* working points as a function of electron η and p_T. The data efficiency central values and the associated statistical and systematic uncertainties are presented in Appendix A.4 for *tight++*, *loose++* and *medium++* (Table A.11, Table A.13 and A.12) :

- Efficiencies for the *tight++* identification is at minimum 60% in the crack region for electrons with $p_T < 25$ GeV, otherwise it varies between 65% and 90%. In almost all (η, p_T) intervals the systematic uncertainties are dominating. For electrons with $p_T < 30$ GeV and $p_T > 80$ GeV the total uncertainty is above 1%, reaching 5% in few η bins.
- The *loose++* efficiencies are above 95% in all (η, p_T) range and the systematic uncertainty dominates over the statistical component. The total uncertainties reached for electrons with $p_T < 30$ GeV vary

between 1 - 8% and for the electrons with $p_T > 80$ GeV between 1 - 5%. In the remaining range, they vary between 0.2 and 1%. In few bins the efficiencies are higher than 100%, due to the background subtraction.

- For the *medium++* identification, the efficiencies are reduced to 75%, mainly in $2.37 < |\eta| < 2.47$ interval for electrons with $p_T > 50$ GeV, otherwise they vary between 77% and 97%. The systematic uncertainties are above the statistical component in almost all [η, p_T] intervals. For the electrons with $p_T < 35$ GeV and $p_T > 80$ GeV, the total uncertainties are between 1% and 7% and otherwise lower, down to 0.2%.

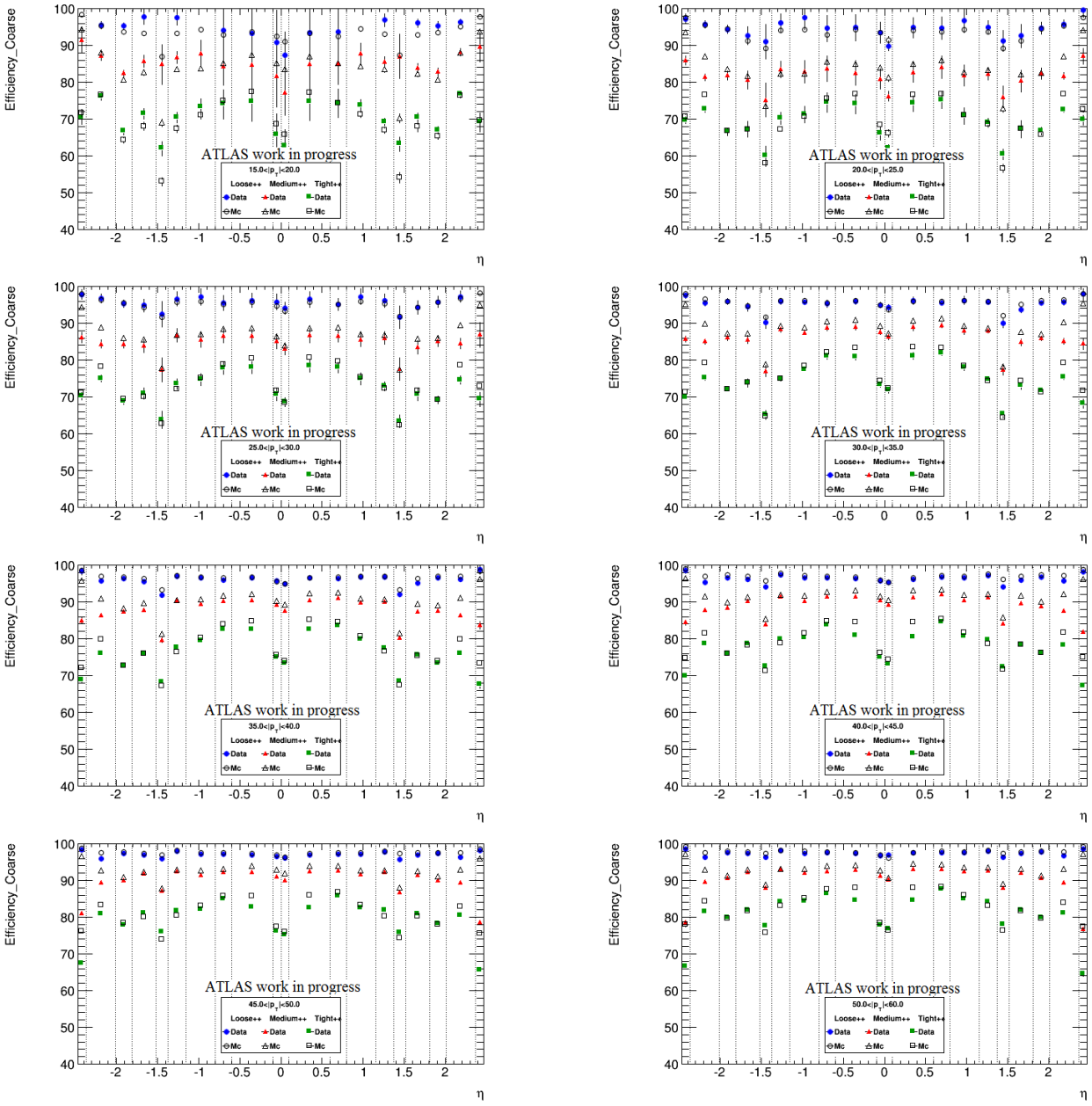


Figure 5.13: Early results : electron identification efficiency as a function of electron η and p_T, in data and MC for *loose++*, *medium++* and *tight++* IDs ($L = 20.3$ fb^{-1} and $\sqrt{s} = 8$ TeV). Only the total uncertainty is shown.

5.3.3 Complementary methods for electron identification efficiency

First, the electron identification efficiency measurements using alternative Z methods are presented. As the main systematic uncertainty is coming from the background subtraction, two methods were developed :

- Measurements with $Z \to ee$ events using the isolation distribution.

- *Tag&Probe* with $Z \to ee\gamma$ events in 10 GeV $< p_T <$ 20 GeV.

Measurements with $Z \to ee$ events using the isolation distribution

There are many similarities with the $Z \to ee$ invariant mass method described in Section 5.3, the most important changes concerns the background subtraction. The same event selection is used and the invariant mass of the tag and probe electrons is broadened to [70, 110] GeV. The background templates, subsets of the baseline distribution, are defined with reverse identification cuts (fail E_{ratio} or $\Delta\eta_1$ or Ws_3). The template definition is changed depending on the η and p_T of the probe electron minimizing the bias[12] in the background shape.

The calorimeter isolation is the main discriminant variable, $E_T^{\text{cone0.3}}$ and $E_T^{\text{cone0.4}}$. Figure 5.14 (left) shows the isolation distribution in data for all probes and for the background template (magenta points) which is normalized in the upper edge of the distribution. The scaled template is subtracted from the baseline distribution in the side band region in each (η, p_T) bin, leading to a distribution with only signal events. The threshold for the side band subtraction is varied from E_T^{cone} / 25 GeV = 0.4 to 0.6.

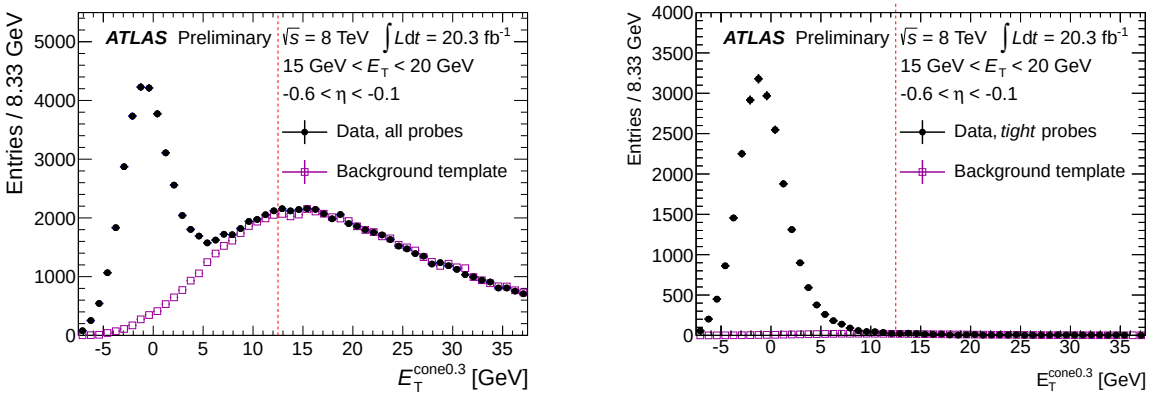


Figure 5.14: Background estimation using Z_{iso} method in the 20 $< p_T <$ 30 GeV and -0.6 $< \eta <$ -0.1 interval. The left figure shows the calorimeter isolation distribution for all selected probes, while in the right figure only for the probes passing the *tight++* identification are shown [168].

The identification efficiency is the ratio between the number of events with probes passing a certain cut (i.e. *tight++* probes in Figure 5.14, right) and the total number of events after the background subtraction. Results are obtained for electrons with 15 $< p_T <$ 80 GeV and are dominated by the systematic uncertainty. The total uncertainty is of the order of few percent in the region with electrons with 15 $< p_T <$ 30 GeV otherwise a per mille level precision is reached in the rest of the domain.

Tag&Probe **with $Z \to ee\gamma$ events in 10** GeV $< p_T <$ **20** GeV

Electrons which loose energy due to photon radiation through Final State Radiation (FSR) are targeted, as their identification efficiency is different from the one computed with the $Z \to ee$ events. The electron p_T is lowered due to the bremsstrahlung radiation. This method can be used only in the 10 $< p_T <$ 15 GeV interval (the statistic is lower at higher p_T due to the *tight* photons requirement).

Compared with $Z \to ee$ measurements, few changes are done in the event selection. One of them is the cut on the invariant mass of the probe pairs : $m_{e_T e_P} < 90$ GeV selecting FSR events and $m_{e_T\gamma} < 80$ GeV. Here e_T is the tag and e_P the probe electron. The latter cut rejects events with the probe electron wrongly reconstructed as a photon. To ensure that they come from a Z boson, the p_T of the probe and of the photon should be smaller than 30 GeV. This make sense as, in the interval used for the measurement, the Z boson is at rest.

The background subtraction is done using the same templates as described in Section 5.1.3. To increase the statistics, the upper edge of the invariant mass is decreased to 100 GeV instead of 120 GeV.

Identification efficiency using the J/Ψ method

$J/\Psi \to ee$ events are used to measure the electron identification efficiency in the [5, 25] GeV p_T range. The *Tag&Probe* method is applied and it relies on the J/Ψ meson mass (3.1 GeV). Only events produced from direct proton-proton collisions are targeted (prompt production). To discriminate these events from non-prompt sources which are biasing the measurement, ie. B hadrons decay, the pseudo-proper time[13]

[12] The bias can be due to the reverted cuts applied for the probe electrons, or due to signal contamination.
[13] The life time cannot be measured as the mother B hadron is not fully reconstructed.

based on the distance between the primary vertex and the displaced J/Ψ decay vertex (L_{xy}) is used :

$$\tau = \frac{L_{xy} \cdot m^{J/\Psi}}{p_T^{J/\Psi}} \tag{5.22}$$

where $m^{J/\Psi}$ and $p_T^{J/\Psi}$ stand for the mass and reconstructed p_T of the J/Ψ particle respectively.

Two methods are used to obtain the identification efficiencies (see Appendix A.4 for more details) as a function of the electron η and p_T. For the primer (short-τ method) the -1 $< \tau <$ 0.2 ps interval is retained while for the second (τ-fit method) the interval is enlarged to -1 $< \tau <$ 3 ps. Figure 5.15 shows the invariant mass distribution in data (black points) and the fit results (red line) for the two different measurements. A good agreement is observed, demonstrating the good understanding of the background subtraction.

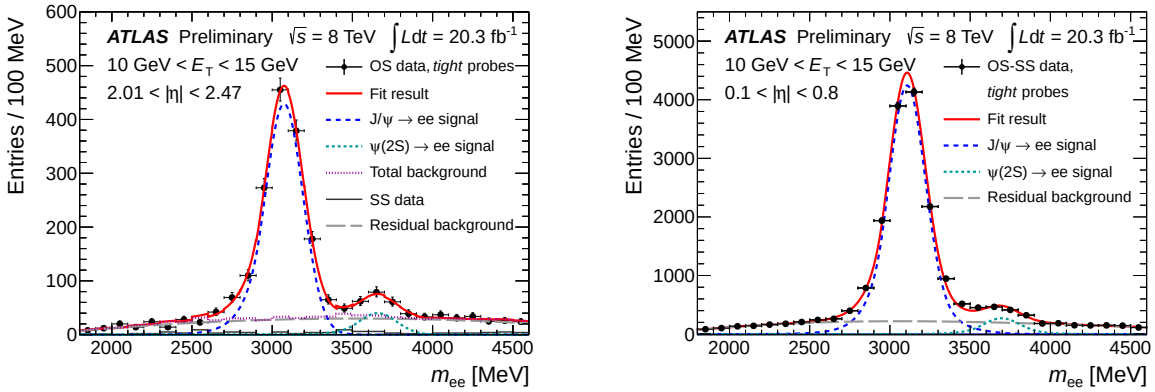


Figure 5.15: Background subtraction using the short τ (left) and τ fit (right) methods. Here the fitted invariant mass distribution for all probe electrons passing the *tight++* ID in 10 $< p_T <$ 15 GeV interval [168].

5.3.4 Combined results

Electron identification efficiencies are measured in different channels using J/Ψ or $Z \to ee$ events. Several methods are developed, mainly differing in the background subtraction. The signal over background discriminant variable is the J/Ψ lifetime, the invariant mass of the electron pair, or the electron calorimeter isolation.

J/Ψ and Z methods are statistically independent. This is not true for the J/Ψ short-τ and τ-fit, or Z mass and isolation methods because the same data is analyzed. Hence, for the final scale factors calculations they are considered to be variations of the same measurement. This ensures a good coverage of possible sources of systematic uncertainties. The combination is possible because the results obtained with the different methods agree.

Using the results of all systematic variations, the central value for each measurement, J/Ψ or Z, is obtained as :

$$\bar{x}_i = \sum_{v=0}^{V} \frac{x_{vi}}{N_v} \tag{5.23}$$

where x is the electron efficiency or the scale factor, i is the (η,p_T) bin and N_v is the number of variations.

The statistical uncertainty is just the average of the statistical uncertainties associated to each variation, while the systematic source is the RMS of the the variations :

$$RMS_{x_i} = \sqrt{\sum_{v=0}^{N_v} \frac{(x_i - \bar{x}_i)}{N_v}} \tag{5.24}$$

For the different p_T bins the methods are combined to obtain the central value of the efficiency or the scale factor :

- 7 - 10 GeV: $J/\Psi \to ee$ short-τ and τ-fit methods;
- 10 - 15 GeV: $J/\Psi \to ee$ together with $Z \to ee\gamma$ methods;
- 15 - 20 GeV: $J/\Psi \to ee$ together with $Z \to ee$ mass and isolation methods;

- 20 - 150 GeV: $Z \to ee$ methods.

The combination of the independent variations is done accounting for any existing correlation between the systematic uncertainties [222]. The combination is divided in two regions : $5 < p_T < 20$ GeV and $20 < p_T < 150$ GeV, to separate the J/Ψ from the Z measurements, as the latter will dominate due to the small uncertainties in the high p_T range.

The combination procedure [169] can be shortly detailed as :

- Calculate the covariance and correlation matrices for a measurement channel;
- Diagonalize the covariance matrix;
- Combine the statistical independent measurement channels;
- Diagonalize the covariance matrix for the combined measurement to extract the uncertainties of the combination.

For step 3 a χ^2 of the form :

$$\chi^2 = \sum_{i,c} \frac{[m_i - \sum_\alpha \gamma^c_{\alpha i} m_i b_\alpha - \bar{x}^c_i]^2}{(\delta^c_{i,stat})^2 \bar{x}^c_i (m_i - \sum_\alpha \gamma^c_{\alpha i} m_i b_\alpha) + (\delta^c_{i,uncor} m_i)^2} + \sum_\alpha b^2_\alpha, \tag{5.25}$$

is used to perform a global fit (over all bins) and obtain the combined values. Here i and c are the bin and channels indices respectively, m_i the combined value, b_α nuisance parameters, x^c_i the average results of the systematic variation and $\delta^c_{i,stat}$, $\delta^c_{i,uncorr}$, $\gamma^c_{\alpha i}$ are relative statistical, uncorrelated and correlated systematic uncertainties respectively. The m_i and b_α parameters are found after minimizing the χ^2 fit.

Results

Efficiency results using the final background template definitions and normalization methods are shown in Figures 5.16 and 5.17. Several cut based criterion are illustrated [14]. This will translate in smaller data to simulation corrections.

Preliminary data electron identification efficiencies integrated over the full η (right) and p_T (left) are shown in Figure 5.18 for *loose++*, *medium++*, *tight++*, *multilepton* and *loose*, *medium*, *tight* and *very tight* likelihood based identification criteria. When tightening the electron identification and, implicitly, the selection requirements, the overall efficiency is decreasing while the dependency on the electron η and p_T is increasing.

As expected, an efficiency increase in the high p_T regime is present as the separation between background and signal electrons is higher and the identification cuts can be relaxed. In [7, 10] GeV and [80, 150] GeV p_T intervals the efficiency is even higher and not smooth as the imposed selection is relaxed (i.e. probe p_T in the low p_T interval, isolation, no E/p cut in the high p_T range). The dependency on η is due to the detector design and the required selection which is based on the signal to background electron ratio. The drop at $\sim$0 is explained by the gap between the electromagnetic calorimeter wheels and TRT. The lower efficiency in $1.37 < |\eta| < 1.52$ interval is due to the existence of the transition area between the barrel and end-cap regions of the calorimeter. At high η the decrease in efficiency is explained by the large amount of material in front of the end-cap calorimeters (see also Figure 3.7).

The mentioned dependencies are well modeled also in simulation which shows only small differences when compared to data. This translates in smaller corrections which should be applied to MC during the physics analysis. Some of the scale factors are presented in Appendix A (Section A.4, Tables A.14 and A.15) for the *tight++* electron identification criterion. More generally, the overall scale factors for all 7 electron identification criteria are close to one. For low and very high p_T regions the corrections reach at maximum 10% for the most stringent selection, ie. *very tight* likelihood. The uncertainties vary between 5% and 10%, the highest being in the crack region.

Studies showed that the remaining discrepancies are due to differences in the variables distribution used for electron identification [162, 219]. Here enters the shower shape and the TRT high threshold hit ratio variables which are incorrectly modeled by the simulation. When comparing data to MC, the shower variables are wider and shifted to background distributions which translates into higher MC efficiencies. Contrary, the TRT high threshold hit ratio is underestimated in simulations for $|\eta| > 1$ leading to higher efficiencies in data.

[14] Comparing with the early results, lower efficiencies are obtained in $p_T < 30$ GeV range as the bias in the template is lowered (see also Figure 5.5). However, this difference is within the systematic uncertainty.

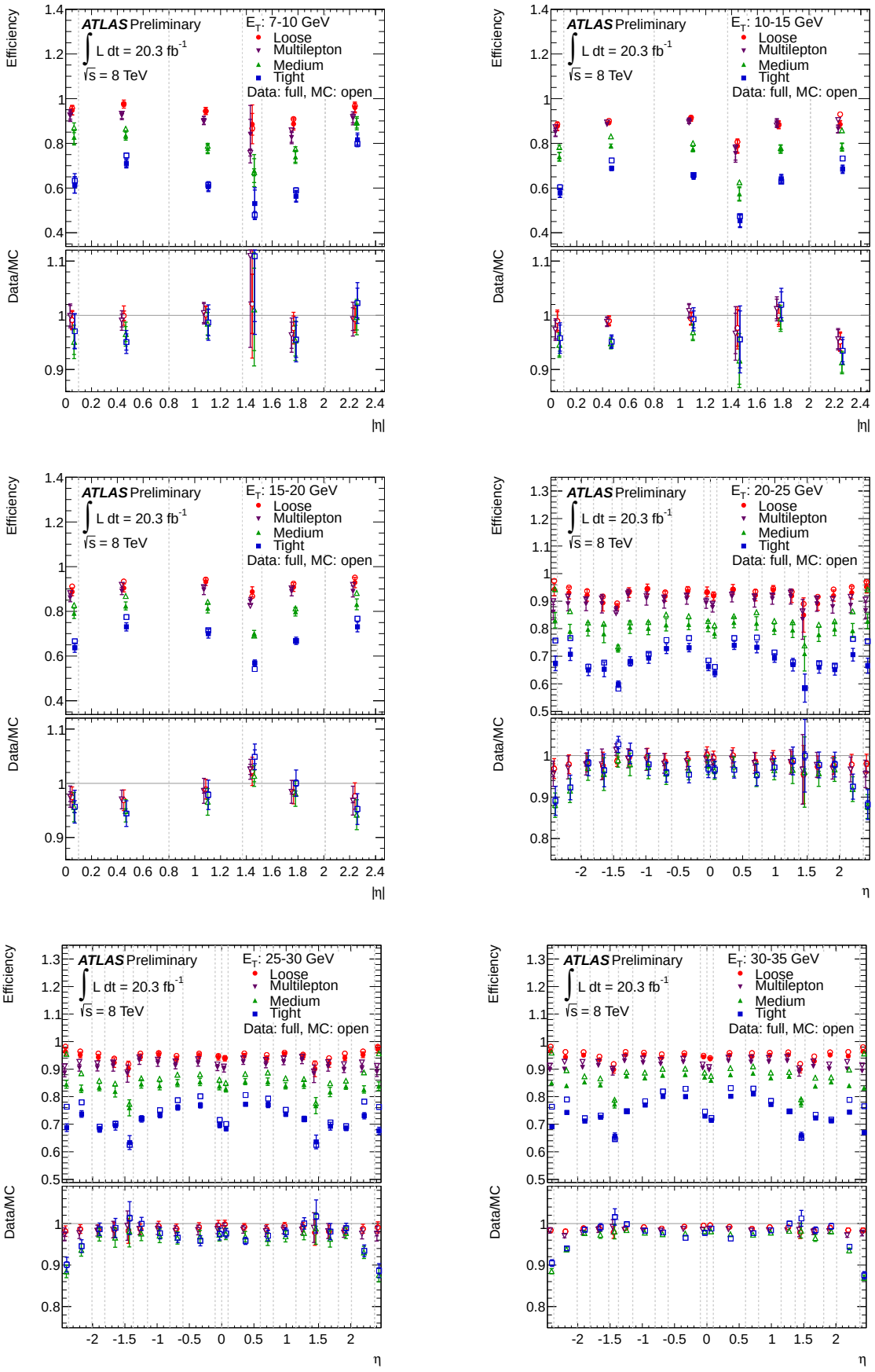


Figure 5.16: Preliminary data electron identification efficiencies using the combination of all presented methods. Results are presented for [7, 35] GeV p_T range for several working points. The uncertainties shown are statistical (inner bars) and statistical + systematic (outer bars).

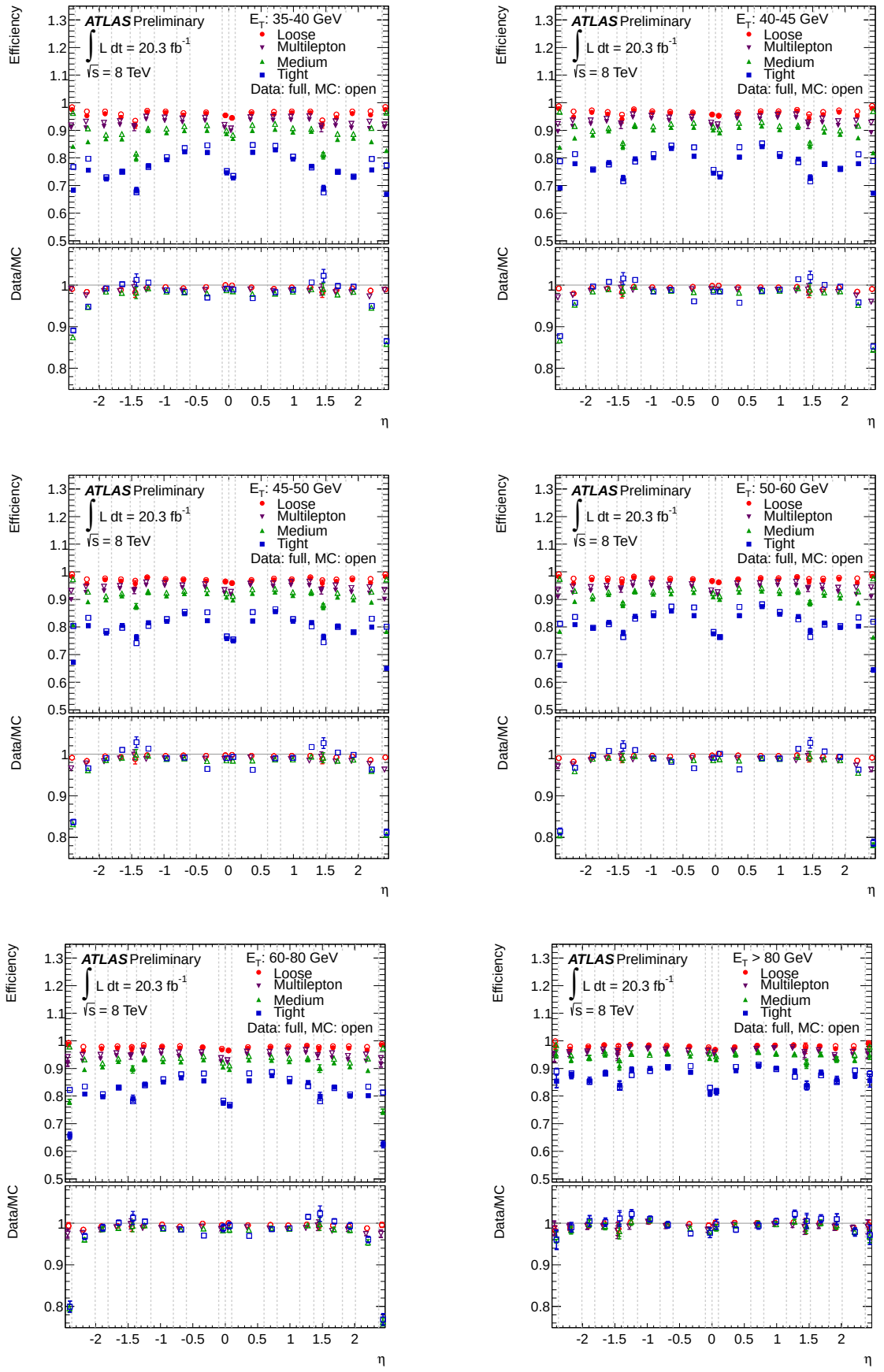


Figure 5.17: Preliminary data electron identification efficiencies using the combination of all presented methods. Results are presented for [35, >80] GeV p_T range for several working points. The uncertainties shown are statistical (inner bars) and statistical + systematic (outer bars).

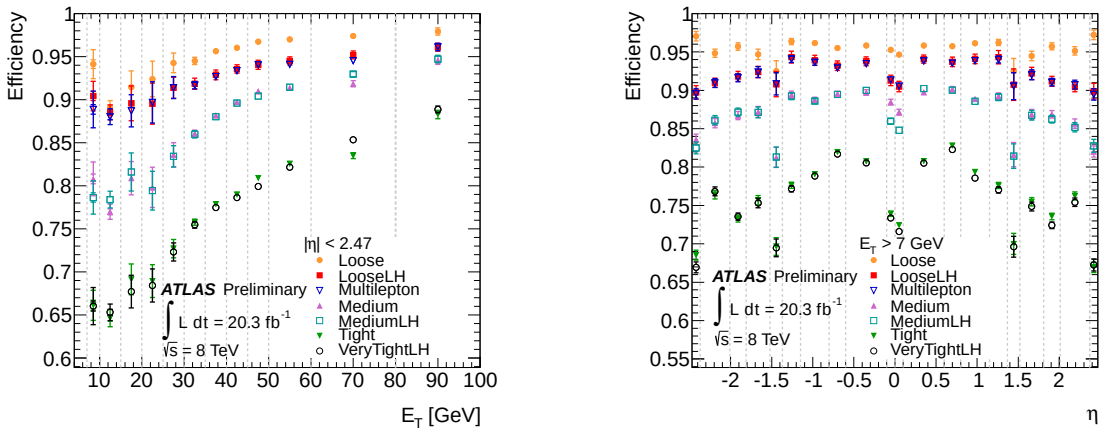


Figure 5.18: Electron identification efficiencies as function of electron p_T (left) and η (right). The uncertainties shown are statistical (inner bars) and statistical + systematic (outer bars).

Pileup dependency

In Figure 5.19 the electron identification efficiency as a function of number of primary vertices is shown. Compared to $\sqrt{s} = 7$ TeV (open markers) data results, the pileup dependency is well reduced given the improved `AutorElectron` algorithm. When comparing to MC results, the data to simulation ratio is either flat (i.e *loose++*) or decreases by 1.5 - 2% (i.e *medium++*, *tight++*). This difference is due to the incorrect modeling of R_{had}, R_{had1} and w_{stot} variables.

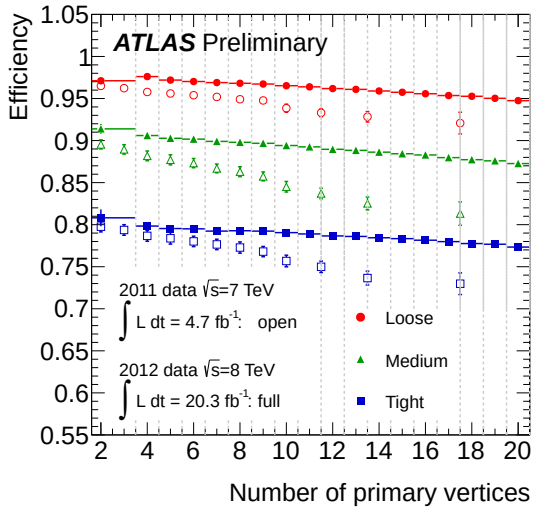


Figure 5.19: Electron identification efficiencies as a function of the number of primary vertices for several electron identification criteria ($p_T > 15$ GeV). Both $\sqrt{s} = 7$ TeV (open markers) and $\sqrt{s} = 8$ TeV (filled markers) data results are considered. The uncertainties shown are statistical (inner bars) and statistical + systematic (outer bars).

5.4 Overall electron efficiencies

As shown in Equation (5.1) the final electron efficiency used in physics analyses is the product of individual electron efficiencies measured at different steps. The efficiency to reconstruct and identify an electron is shown as a function of electron p_T (left) and η (right) in Figure 5.20. For loose working points the average efficiency is around 91% for electrons with p_T ~25 GeV, while for tight working points the value decreases to 71%. As the reconstruction efficiency is not strongly dependent on η or p_T, the trend of the combined distribution is driven by the identification efficiencies : it increases with electron p_T and presents a stronger η dependency for tighter identification criterion.

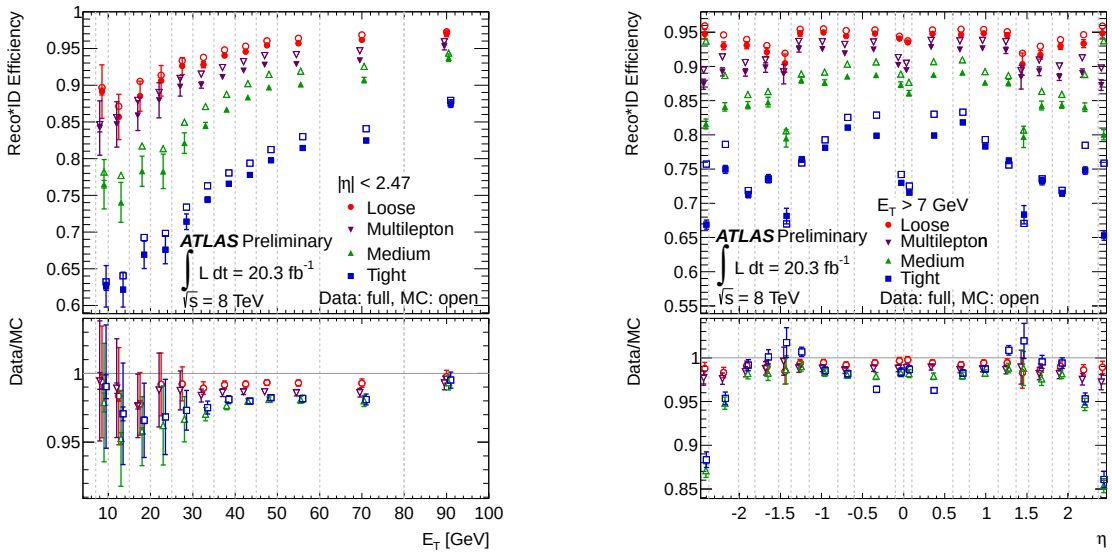


Figure 5.20: Combined electron and identification efficiencies as a function of electron p_T (left) and $|\eta|$ (right) in data (filled markers) and MC (open markers). The uncertainties shown are statistic (inner bars) and statistic + systematic (outer bars).

5.5 Conclusions

In this chapter the electron reconstruction and identification efficiency measurement techniques were presented together with the results obtained for the electron correction factors using 2012 $\sqrt{s} = 8$ TeV, $L = 20.3$ fb^{-1} of proton-proton collisions data. These corrections are released by the Electron/Photon performance group, as recommendations for the physics groups. They should be applied to MC electrons in order to account for the differences between simulation and data.

In 2012 a new tracking algorithm was developed to recover the loss efficiency due to hard bremsstrahlung (it improves the track-cluster matching). Integrated over η, the electron reconstruction efficiency varies between 97% at low p_T and 99% at high p_T. The improvement compared with $\sqrt{s} = 7$ TeV measurements is ∼6% in [15, 20] GeV p_T range and ∼3% in [30, 50] GeV p_T range for $Z \to ee$ events. At higher p_T, the results are similar. When the results are integrated over p_T, the electron reconstruction efficiency varies between 99% at low η and 95% at high η. When comparing to $\sqrt{s} = 7$ TeV results the gain is around 4 - 8% in $1.37 < \eta < 2.47$ range and around 3% otherwise. The global uncertainty associated to reconstruction efficiency measurement is below 0.5% for electrons with $p_T > 25$ GeV and in the [0.5, 1.5]% interval otherwise.

For the identification efficiencies, an improvement is reached after loosening the requirements on the electron shower shapes variables from the electromagnetic calorimeter which are more affected by the increase in luminosity (between 2011 and 2012). To recover the initial rejection of light jets, other cuts are added and the pileup dependent variables are tightened. The gain in statistics, the new electron *likelihood* working points and the improvement done on the background template definition is reflected in smaller uncertainties and closer to unity scale factors with respect to 7 TeV measurements. The identification efficiency strongly depends on the electron p_T and at tight identification criterion on the electron η. When calculated for *loose++* criterion it averages to 96%, while for *very tight* likelihood based function it is in average 78% for electrons with $p_T > 15$ GeV. As it will be shown in the next chapters these efficiencies decrease for processes involving many jets or boosted tops. The associated uncertainty is ∼5 - 6% for $p_T < 25$ GeV and $\sim 1 - 2\%$ otherwise.

The global improvement on the electron identification and reconstruction efficiencies increases the precision for SM cross section measurements. It is particularly noticeable for the systematic uncertainty on the electron reconstruction and identification efficiencies on $H \to ZZ \to eeee$ [10] where the signal strength measurement decreases from 9% to 4%. Non-negligible improvement is observed also in the BSM measurements involving electron channels.

Two other crucial efficiency measurements involving reconstructed or identified electrons are needed to measure the SM cross sections or search for BSM physics : the fake electron efficiency and the probability to reconstruct an electron with a wrong charge. The last measurement is important only in the channels involving two electrons with same electric charge. Given the strong dependency on the electron definition, these measurements are presented in Chapters 6 and 7 for the search for strongly produced SUSY particles in events with two same electric charge leptons and jets.

Chapter 6

Search for new physics in events with two same - sign leptons

New physics searches are of great interest and occupy a prime place in the ATLAS and CMS physics program. These searches look for extra-dimensions, supersymmetry, new quark and fermion families, new gauge bosons... Given the different final topologies there are a lot of specific analyses, but in essence with the same purposes : define dedicated regions enriched in signal, measure the background as precisely as possible and quantify any excess above the SM expectation.

Among the various signatures studied in ATLAS we highlight here the interesting channel associated to events with 2 leptons of same electric charge (same - sign leptons), jets and missing transverse energy. As already mentioned in Chapter 2, this topology is promising as the Standard Model processes are rare, and the Majorana nature of gluinos and the presence of top quarks in the cascade decays is assuring a high rate. The main targets driving the study are the strong production processes involving large mass gluinos and third generation squarks. Assumptions are the conservation of R-parity and $\tilde{\chi}_1^0$ is the LSP, as this approach is well motivated by the *natural* SUSY spectrum (Chapter 2). To show the versatility of this analysis, more models with violated R-parity or gravitino LSP are included as well.

In the following two chapters the results are obtained with $\sqrt{s} = 8$ TeV and $L = 20.3$ fb^{-1}. They were made public in one CONF note [223] and one paper [224]. There are more preliminary public versions : with 2 fb^{-1} of 7 TeV data [225], and with 5.8 fb^{-1} of 8 TeV data [226]. Even if I contributed significantly to [226], notably for the background estimates and the interpretation of the observations, this work is not presented here as it has been largely superseded and improved by [223] and [224]. In this chapter, preliminary results using 21 fb^{-1} of 8 TeV data are discussed and an inclusive selection with a pair of same - sign leptons and no special requirement on the third lepton is considered. In the next chapter the final results of the 8 TeV analysis are presented. To increase the sensitivity and the coverage of this signature, the leptons are further split in exclusive same - sign and three leptons channels. No τ leptons are considered in any version, i.e. leptons are only electrons and muons.

This chapter starts with a short presentation of the considered SUSY models (Section 6.1). The generators and filters used to simulate the Standard Model processes and the SUSY signal are described in Section 6.2. The event selection and object definitions are detailed in Section 6.3. In the same section, the signal region optimization strategy and the final signal region definitions are also presented. In Sections 6.4 - 6.6 the reducible and irreducible background categories are discussed, together with the methods used for their estimation. The validation is shown in Section 6.7. Section 6.8 presents the preliminary results in the signal regions and their interpretation in terms of exclusion limits in the absence of an excess. Final conclusions are presented in Section 6.9.

6.1 Description of SUSY models

To ensure a good coverage of the SUSY phase space, a very large number of SUSY simplified and phenomenological models are used to interpret the results, in the absence of signal. Figure 6.1 presents a summary of these models. Compared with phenomenological or complex SUSY models, the simplified models [142–144] allow to optimize the coverage over all SUSY phase space in an affordable way for experimentalists as the production mode, the branching ratios and kinematics are well known and controlled. Another reason is that the theorist favorite model can be constrained provided a signature with the same topology.

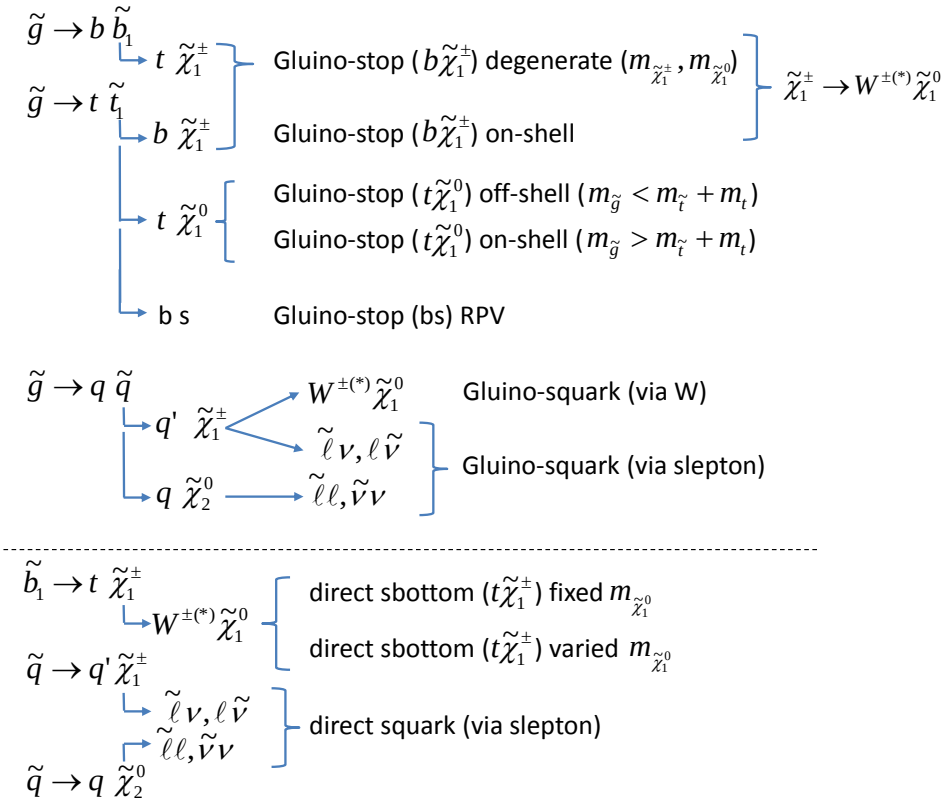


Figure 6.1: Overview of the simplified models considered in the analysis.

6.1.1 Gluino pair production models

Several gluino mediated simplified models are used to optimize the dedicated signal regions and interpret the results. A high sensitivity can be reached with optimized signal regions involving 0, 2 or more b jets and many light jets, depending on whether top quarks are present in the intermediate state. If not otherwise specified the LSP is considered to be $\tilde{\chi}^0$and the R parity is conserved.

In the **gluino stop off-shell model (Gtt),** $\tilde{g} \to tt\bar{\tilde{\chi}}_1^0$**,** the gluinos are produced in pairs and are assumed to be lighter than all squarks. The top squark is the lighest squark and its mass is set to 2.5 TeV. The first 2 squark generations masses are assumed to be 30 TeV (decoupled). Hence the contribution to gluino production through t-channel squark exchange is suppressed and gluinos are decaying with 100% branching ratio (BR) in a top quark and LSP via a virtual (off-shell) top squark. The kinematic bound is $m_{\tilde{g}} - m_{\tilde{\chi}_1^0} \geq 2m_t$[1]. The final state is characterized by 4 bottom quarks and 4 W bosons.

Even if the assumption on the top squark mass is not *natural* the model is still retained in both ATLAS and CMS experiments as it is a benchmark to compare all the analyses which are sensitive to a final state dominated by 4 top quarks and $\not{E}_T$.

The **gluino stop on-shell model, BR**$(\tilde{t}_1 \to t\tilde{\chi}_1^0) = 100\%$**,** is similar to the previous one, but the top squark is lighter than $m_{\tilde{g}}$ - m_t. The gluinos decays via an on-shell top squark, i.e $\tilde{g} \to t\tilde{t}_1 \to tt\tilde{\chi}_1^0$, and the LSP mass is set to 60 GeV. The final state contains 4 real W bosons and 4 b jets from the top quark and 2 $\tilde{\chi}_1^0$. Results are presented in the gluino-top squark mass plane.

In the **gluino stop on-shell model, BR**$(\tilde{t}_1 \to b\tilde{\chi}_1^{\pm}) = 100\%$**,** the $\tilde{t}_1$ is the lightest quark and the other quarks are lighter than gluino. The gluino are pair produced and decays into a top quark and top squark with 100% BR. The top squarks are decaying like $\tilde{t}_1 \to b\tilde{\chi}_1^{\pm}$. The kinematic bound is $m_{\tilde{g}} > m_t m_b$. The $\tilde{\chi}_1^0$ mass is set to 60 GeV and the $\tilde{\chi}_1^{\pm}$ mass to 120 GeV. The later is assumed to decay through a virtual W boson and LSP, $\tilde{\chi}_1^{\pm} \to W^{*\pm}\tilde{\chi}_1^0$. The final state is characterized by 4 $W^{(*)}$ bosons and 4 b quarks from the top quark decay, 2 $\tilde{\chi}_1^{\pm}$ and 2 $\tilde{\chi}_1^0$. Limits are expresses in the gluino-top squark mass plane.

In the **gluino stop and sbottom off-shell,** $\tilde{g} \to tb\tilde{\chi}_1^0$**, degenerate** $(\mathbf{m}_{\tilde{\chi}_1^{\pm}}, \mathbf{m}_{\tilde{\chi}_1^0})$ model gluinos are assumed to be lighter than all squarks. The lightest squarks are the $\tilde{t}_1$ and $\tilde{b}$. Gluinos pair production is the

[1] In [227] the three body decay is extended to a 7 body decay, $\tilde{g} \to f\bar{f}'bf''\bar{f}'''\bar{b}\tilde{\chi}_1^0$, with no constrain on the top quark or W boson virtuality. This was investigated to check if the present ATLAS analyses sensitivity can cover a more compressed spectra. This also helps in increasing the coverage of the analysis with two same - sign leptons for Run-2, by optimizing new signal regions targeting a very compressed SUSY phase space.

only considered process and decays through virtual top and bottom squarks with 100% BR for $\tilde{t}_1^* \to b\tilde{\chi}_1^\pm$ and $\tilde{b}^* \to t\tilde{\chi}_1^\pm$. A small mass difference of 2 GeV is assumed between the $\tilde{\chi}_1^\pm$ and $\tilde{\chi}_1^0$. This ensures that $\tilde{\chi}_1^\pm$ decays into $\tilde{\chi}_1^0$ and very soft (undetectable) fermions. The final state is characterized by 2 W bosons, 4 b jets and 2 LSPs. Results are shown as a function of the gluino and LSP masses.

The **gluino stop on-shell,** $\tilde{g} \to t\tilde{t}_1$ model with tops squarks decaying via $\tilde{t} \to c\tilde{\chi}_1^0$ is also considered. The neutralino mass is fixed to $m_{\tilde{\chi}_1^0} = (m_{\tilde{t}_1}$ - 20) GeV. The final state contains 2 real W bosons and 2 b jets from the top quark decay, two light jets and two LSPs. Limits are shown in the gluino-top squark mass plane.

The **gluino stop RPV,** $\tilde{g} \to bs$**,** model is considered to show the versatility of this search. Gluinos are produced in pairs and decay as $\tilde{g} \to t\tilde{t}_1$. The $\tilde{t}$ decays not conserving the R-parity and the baryon number ($\lambda''_{323} = 1$) in $\tilde{t} \to bs$. No other squarks or processes are considered. The final state consists in two top quarks, two b jets and two light jets. Limits are set in the gluino-LSP mass plane.

In the **gluino squark via** W **boson,** $\tilde{g} \to q\tilde{q}$**,** gluinos are decaying in a quark and a squark which goes as $\tilde{q} \to q\tilde{\chi}_1^\pm$ with $\tilde{\chi}_1^\pm \to W^\pm\tilde{\chi}_1^0$. The chargino mass is set twice the LSP mass. The final state is formed by 4 light jets, 2 W bosons and 2 LSPs. No b jets are expected. The limits are set in the gluino-LSP mass plane.

The **gluino squark via sleptons,** $\tilde{g} \to q\tilde{q}$**,** 2 step decay model. The produced squark decays either via charginos, either via neutralinos with equal probability. The gauginos are decaying with same probability into a slepton or sneutrino and the associated SM particle. In the cascade decay $\tilde{\ell} \to l\tilde{\chi}_1^0$ or $\tilde{\nu} \to \nu\tilde{\chi}_1^0$ processes are present. Squark masses are assumed to be much lighter than the gluino masses. The intermediate sparticle masses are set such that $\tilde{\chi}_1^\pm$ and $\tilde{\chi}_2^0$ masses are the same and sits exactly in the middle of gluino (squark) and LSP masses. All sleptons and sneutrinos are considered in the simulation and are degenerate in mass, which should be exactly in the middle between $\tilde{\chi}_1^\pm$ or $\tilde{\chi}_2^0$ and LSP masses. The decay chain can lead up to 4 light jets, four charged leptons and $\not{E}_T$. Limits are set in the gluino-LSP mass plane.

6.1.2 Squark pair production models

In the **direct bottom squark model** the considered decay mode is $\tilde{b}_1 \to t\tilde{\chi}_1^\pm$, with $\tilde{\chi}_1^\pm$ decaying as $\tilde{\chi}_1^\pm \to W^\pm\tilde{\chi}_1^0$. If $m_{W^\pm} < m_{\tilde{\chi}_1^\pm} - m_{\tilde{\chi}_1^0}$ the $\tilde{\chi}_1^\pm$ decay is a 2 body decay, while if $m_{W^\pm} > m_{\tilde{\chi}_1^\pm} - m_{\tilde{\chi}_1^0}$ it is a 3 body decay. Two grids are considered : one with $\tilde{\chi}_1^0$ mass set to 60 GeV and another one with $\tilde{\chi}_1^\pm$ twice the $\tilde{\chi}_1^0$ mass. The rest of SUSY mass spectrum is set to 4.5 TeV. The final state is characterized by 4 W bosons, 2 b jets and 2 LSPs. The results are interpreted as a function of $\tilde{\chi}_1^\pm$ and $\tilde{b}_1$ masses.

The **direct squark via sleptons** model is very similar to gluino squark via sleptons model and same mass constraints are applied. Here the considered production is a pair of first and second light squark generation, followed by $\tilde{q} \to q\tilde{\chi}_1^\pm$ or $\tilde{q} \to q\tilde{\chi}_2^0$ decays. Sleptons appear in the cascade decays of gauginos. Final states are characterized by 2 light jets (first and second generation), up to 4 charged leptons and $\not{E}_T$. Results are presented in the parameter space of squark and LSP masses.

6.1.3 Complete SUSY models

In addition to the simplified models a complete SUSY model is also considered. It is the **mSUGRA/CMSSM** model [106–109] with the parameters set as $\tan\beta = 30$, A_0 = -2m_0 and $\mu > 0$, and constraints compatible with the Higgs mass of 125 GeV[2]. Only strong production ($\tilde{q}\tilde{q}$, $\tilde{q}q^*$, $\tilde{q}\tilde{g}$, $\tilde{g}\tilde{g}$) and associated electroweak ($\tilde{g}\chi$, $\tilde{q}\chi$) processes are considered, and no sleptons or gaugions pair production. It is a typical mSUGRA scenario used to compare results with CMS and Tevatron experiments. Exclusion limits are presented as a function of m_0 and $m_{1/2}$ parameters.

6.2 Background and signal simulation

The search for SUSY in the same - sign lepton channel is characterized by two types of backgrounds. First, the Standard Model processes which leads to 2 same - sign leptons and several jets in the final state, ie. $t\bar{t}$ in association with vector bosons. In the second category enter the detector backgrounds which can be studied by looking at $t\bar{t}$ processes. To estimate some of these backgrounds and compute the systematic uncertainties Monte Carlo samples are used. The ATLAS detector simulation is performed with GEANT 4 or ATLFAST[3] [228–230]. To model the SUSY signal mentioned in the previous section dedicated MC

[2] Assume universal masses at GUT scale : m_0 = common of scalars (squarks, sleptons, Higgs boson); $m_{1/2}$ = common mass of gauginos and higgsinos; A_0 = common trilinear coupling; $\tan\beta$ = ratio of Higgs vacuum expectation values; sign(μ) = $\pm$ 1 = sign of μ, SUSY conserving Higgsino mass parameter

[3] ATLFAST is used only for few $t\bar{t}$ samples.

samples are generated. For these samples the ATLAS detector was simulated with ATLFAST. More details on the used generator, showering and hadronization, filters, etc, are given in this section.

Standard Model processes

Table 6.1 summaries the main sources of background in the SRs. For comparison a signal point from gluino mediated top squark off-shell model with $m_{\tilde{g}} = 1.$ TeV, $m_{\tilde{\chi}_1^0} = 200$ GeV is also given.

Process	Generator	Hadronization	$\sigma \times BR$ [pb], $\sqrt{s} = 8$ TeV	L [fb^{-1}]
$t\bar{t} + W$ $(W \to l\nu)$	MADGRAPH	PYTHIA	0.104	3300
$t\bar{t} + WW$ $(W \to l\nu)$	MADGRAPH	PYTHIA	0.00092	11000
$t\bar{t} + Z$ $(Z \to ll)$	MADGRAPH	PYTHIA	0.07	4400
WZ $(W/Z \to l\nu/ll)$	SHERPA	SHERPA	9.75	260
WW $(W \to l\nu)$	SHERPA	SHERPA	5.49	460
$ZZ \to llll$	POWHEG	PYTHIA	8.74	75400
Gtt ($m_{\tilde{g}} = 1$ TeV)	HERWIG++	PYTHIA	0.023	4347

Table 6.1: MC samples information for the main sources of background in the signal regions. Here l stands for electron or muon. For comparison, a signal point from Gtt model is shown.

Processes like $t\bar{t} + W$ and Z are generated using MADGRAPH [231] generator with up to 2 extra partons in the matrix elements. The particle hadronization and showering are performed using PYTHIA 6 generator [232]. The used parton distribution function (PDF) is CTEQ6L1 [233]. Additional samples $t\bar{t} + WW$ are produced using same generators, but with 0 extra partons and MSTW2008LO [234] PDF set. Alternative $t\bar{t} + V$ samples used for the systematic uncertainties evaluation are generated with ALPGEN [235] interfaced with JYMMY [236], adopting same PDF set.

Diboson processes, WZ and WW are modeled using SHERPA [237] with the PDF set CT10 [238]. Up to 3 jets in the matrix elements and parton showering are considered. Lepton p_T is above 5 GeV and a cut on the invariant mass of the 2 leptons from the $Z^{(*)}$ boson (γ^*) is added. For ZZ, the generation is performed with POWHEG [239] interfaced with PYTHIA 8 [240] using CT10 PDF set. The same lepton p_T cut is retained and the invariant mass cut threshold is increased to 4 GeV. To obtain the associated modeling systematic uncertainties, as explained in one of the next sections, two set of MC@NLO [241] samples with CT10 PDF set are used. For these samples, next to leading order (NLO) matrix elements are matched to PYTHIA 6 or HERWIG++ [242] with JYMMY parton showers. For VVV ($V = Z$ or W), MADGRAPH interfaced with PYTHIA 6 is used and the PDF set is CTEQ6L1.

Samples with $t\bar{t}$ processes used for data driven estimation cross checks are produced using 4 different generators. POWHEG interfaced with PYTHIA 6, POWHEG interfaced with HERWIG++ and JYMMY, MC@NLO interfaced with HERWIG++ and ALPGEN interfaced with HERWIG++ and JYMMY. For the first three samples CT10 PDF set is used, while for the latter CTEQ6L1.

For single top + Z and $t\bar{t}t\bar{t}$ processes, MADGRAPH interfaced with PYTHIA 6 generators are used. The PDF sets are MSTW2008LO and CTEQ6L1. Other exotic processes like $t\bar{t} + H$, or $H + Z, W$, for which the H boson mass is set at 125 GeV, are simulated with PYTHIA 8 assuring CTEQ6L1 PDF set.

SUSY signal

The SUSY signal is generated using HERWIG++ or MADGRAPH interfaced with PYTHIA 6, fixing the PDF set to CTEQ6L1. The cross sections includes next to leading order QCD corrections and the resummation of soft gluon emission at the next to leading logarithmic accuracy (NLL). The NLL results allow a significant reduction of the theoretical uncertainties from unknown higher perturbative orders [243]. It should be noted that only NLO predictions are used for models involving sleptons and gauginos, ie. electroweak production[4] [244, 245].

The cross sections central values for pair production processes following the strong production interactions (ie. $pp \to \tilde{q}\tilde{q}, \tilde{q}\tilde{q}^*, \tilde{q}\tilde{g}, \tilde{g}\tilde{g} + X$) are calculated using NLLFAST computer program [246]. For models including squarks, as input only the average mass of the first and second generations is given, because the masses are considered to be degenerate in almost all SUSY models. Same convention is adopted when using PROSPINO [247].

To obtain the systematic uncertainties, several sources are considered : the theoretical uncertainty and the parton distribution functions for NLL corrections. NLLFAST or PROSPINO is used to account for the

[4] As the NLL corrections for these processes are very small.

choice of the renormalization and factorization scales. To combine them, an overall uncertainty is computed taking $\pm$ 1 σ variation of CTEQ6L1 and MSTW PDFs, and α_s coupling constant. The later one is found to be negligible.

6.3 Signal regions

The trigger strategy, object definitions and event selection (Section 6.3.1) presentation are opening the description of the same sign leptons analysis. This part is in general common to all strong SUSY searches with leptons, jets and $\not{E}_T$ [248–251]. In Section 6.3.2 a short presentation of the optimization and the key variables in the background subtraction are presented. The section closes with the definition of the signal regions and their motivation. Finally, in Section 6.3.3 the importance of defining control and validation regions is presented.

6.3.1 Trigger, object definition and event selection

Two different trigger strategies are considered. In the region with $\not{E}_T > 150$ GeV only the $\not{E}_T$ triggers are used, otherwise single and dilepton triggers. The main advantage is the gain in event selection efficiency and the avoidance of isolated triggers. The isolation at trigger stage would introduce a bias in the background electron estimation. The efficiency of the $\not{E}_T$ triggers (`EF_xe80T_tclw_loose or EF_xe80_tclw_loose` depending on the data period) is increasing from 92%, if $\not{E}_T > 80$ GeV, to $\sim$ 100% when is reaching the plateau at 140 GeV.

The lepton triggers are applied depending on the leptons p_T and flavor for regions with low $\not{E}_T$. No cut on the electric charge is considered at this step.

- Events with two leading pre-selected electrons : if the leading electron has $p_T < 70$ GeV then a di-electron trigger, `EF_2e12Tvh_loose1`, is considered, otherwise a single lepton trigger, `EF_e60_medium1` is chosen[5]. Both triggers reach the efficiency plateau at 95% for an offline p_T cut of 20 and 60 GeV respectively.

- Events where the first two leading pre-selected leptons are an electron and a muon : if the leading lepton is an electron with $p_T > 70$ GeV the `EF_60_medium1` trigger must be fired. If it is a muon with $p_T > 40$ GeV `EF_mu36_tight` is chosen, otherwise `EF_e12Tvh_medium1_mu8`. For the electron trigger the efficiency plateau is at 95% starting with $p_T \sim 20$ GeV, while for the muon trigger the efficiency plateau is only 75% (90%) for the barrel (end-cap) at $p_T \sim 40$ GeV. This decrease is due to a drop of the efficiency in the barrel. The dilepton trigger reaches the plateau at 95% for leptons with $p_T > 20$ GeV.

- Events with two leading pre-selected muons : if the leading lepton p_T is higher than 40 GeV, `EF_mu18_tight_mu8_EFFS` is selected. The single muon trigger, `EF_mu36_tight` is considered otherwise. This dimuon trigger was chosen to avoid the drop in efficiency. This is possible as the `EF_mu8_EFFS` is not based on *Level1* trigger decision (were the drop in the efficiency appears) but only on the *EventFilter* level and hence fully efficient.

Given this selection the 100% efficient $\not{E}_T$ triggers ensures a very high event selection rate in the signal region.

Scale factors are calculated and applied to all lepton triggers. This accounts for any data - MC discrepancy in the trigger efficiency. For this study, $t\bar{t}$ MC samples are used since the event topology is similar to the SUSY signal. The estimation is performed as a function of lepton η (considering 2 regions : barrel, $|\eta| < 1.5$, and end-cap, with $|\eta| > 1.5$) and p_T using the *Tag and Probe* method. These corrections are $\pm$2 - 4% for the $\not{E}_T$ triggers and leptons with $p_T > 20$ GeV. A 1.09 scale factor is obtained if the electron p_T is inside the [15, 20] GeV interval and is firing the dielectron trigger.

Object definition

The ATLAS object definitions presented in Chapter 4 are optimized for a higher background reduction in the SUSY searches.

[5] As the name suggests the electrons p_T must be greater that 12 GeV for the dielectron trigger and the leading electron $p_T > 60$ GeV for the single lepton trigger

- The pre-selected (baseline) electrons were reconstructed using the standard algorithm, `AutorElectron`, and should have $p_T > 10$ GeV, $|\eta| < 2.47$ and pass the *medium++* identification criteria (see Section 4.3). To reject the electrons from hadron decays, the electron candidates must be separated from any pre-selected jet by $\Delta R_{e,jet} > 0.4$. For the signal electrons, a tighter identification criteria is imposed, *tight++*, and the the p_T is increased to 20 GeV. They should also pass the impact parameter and the isolation cuts to ensure that the electrons are promptly produced. To all *tight++* electrons the scale factors measured in Section 5.3.4 are applied.

  The transverse impact parameter of the track, d_0, its significance (uncertainty) and the longitudinal impact parameter are calculated with respect to the primary vertex [6]. Compared with the previous version of the analysis [226] the $d_0/\sigma(d_0)$ cut was tighten from 0.5 to 0.3 to further reduce the background of same - sign pairs for which the charge of one electron is wrong. For the same reasons, a $d_0/\sigma(d_0) < 5$ cut was also added for the baseline electrons. A tight cut on z_0 parameter, $z_0 sin\theta < 0.4$ mm, is applied to reduce the electrons from heavy flavor decays.

  The track isolation variable (p_T^{cone30}) is defined as the scalar sum of the transverse momenta of the tracks inside a cone with a radius $\Delta R < 0.3$ around the lepton track. The considered tracks in the sum must have $p_T > 0.4$ GeV and be compatible with the electron vertex. The calorimeter isolation variable (E_T^{cone30}) is computed from topological clusters which use noise suppression and gives more robustness against the pileup. Electron energy leakage in the cone and pileup corrections are applied. The calorimeter isolation uses the transverse energy deposited in the calorimeter, in a cone of radius $\Delta R = 0.3$ around the electron.

  The cuts on relative isolations are used to reduce the fake electron background. During the optimization the favored values are found to be $p_T^{\text{cone30}}/p_T < 0.16$ and $E_T^{\text{cone30}}/p_T < 0.18$. It is generally applied in SUSY analyses with energetic leptons and several energetic jets in the final state.

- All the pre-selected (baseline) muons are reconstructed using the `STACO` algorithm which uses both combined and segment tagged muons. They should have $p_T > 10$ GeV and $|\eta| < 2.4$. The *loose* identification defined in Section 4.3 is considered with few additional quality cuts on the inner detector tracks quality (ie. number of pixel, SCT, TRT, b layer hits). This working point allows to keep a high signal muon efficiency and a good background muon rejection. For the same reasons mentioned for electrons, all the muons should be well separated from any jet within a cone with $\Delta R_{\mu,jet} > 0.4$. To reject cases where a fake electron is very close to a real muon (eg. $Z \to \mu\mu$ when one muon radiates a photon later reconstructed as an electron), the event is rejected if $\Delta R_{e,\mu} < 0.2$. For the signal muons the p_T threshold is increased to 20 GeV and cuts on relative isolation ($p_T^{\text{cone30}}/p_T < 0.12$ and $E_T^{\text{cone30}}/p_T < 0.12$) and impact parameter variables ($d_0/\sigma(d_0) < 3$) are added as for the electrons.

- Jets are reconstructed using the *FastJet* implementation of the $anti-k_T$ algorithm with a distance parameter $R = 0.4$ (see Section 4.5, calorimeter jets with LCW scale). They should satisfy $p_T > 20$ GeV and $|\eta| < 2.8$ cuts. An overlap removal between all the pre-selected electrons with $p_T > 10$ GeV (before applying the mentioned $\Delta R_{e,jet}$ cut) and jets is applied since fake jets are reconstructed from the energy deposited by electrons in the calorimeter. If a jet is within $\Delta R_{jet,e} < 0.2$ the event is retained, but the jet is discarded. For the signal jets the p_T threshold is increased to 40 GeV.

- b jets are tagged using the `MV1` algorithm, at 70% efficiency operating point. The jet p_T is decreased to 20 GeV and the η region is restrained to < 2.5.

  No cut on the jet vertex fraction (JVF) is applied since the number of (b) jets with $p_T >$ (20 GeV) 40 GeV is not sensitive to pileup when a pair of leptons and several jets are required in the event, as shown in Figure 6.2.

- The $\not{E}_T$ definition described in Section 4.7 without the tau leptons and with STVF additional corrections to suppress pileup is used in this analysis. It should be mentioned that the objects entering in the computation must pass all the pre-selections described previously with one exception, the overlap removal between muons and jets which is not applied.

Event pre-selection

Beside the mentioned cleaning requirements, some additional ones are applied to remove some *unwanted* events. The impact from the noise bursts and data corruption in the LAr and Tile calorimeters is reduced

[6] d_0 ($z_0 sin\theta$) is the closest distance from the track to the primary vertex in the transverse (longitudinal) plane.

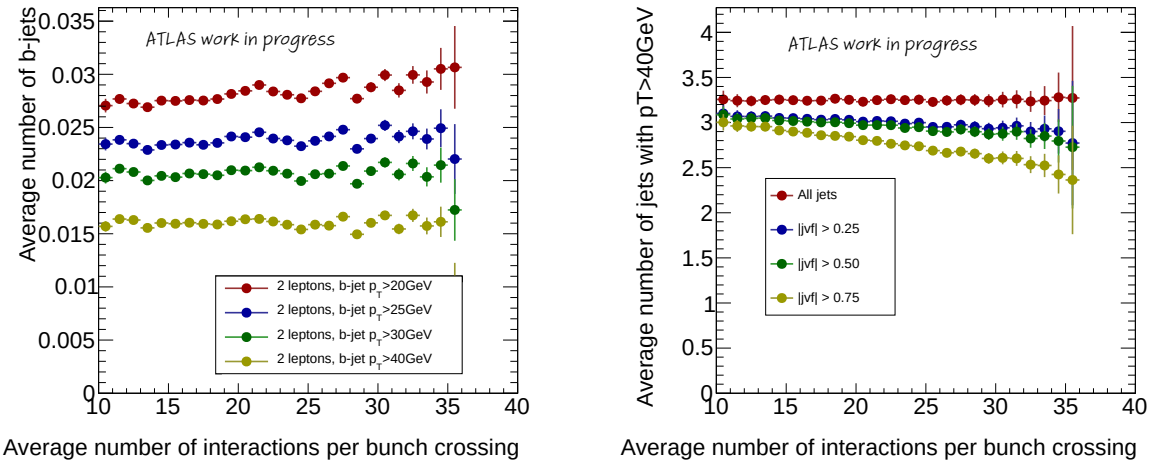


Figure 6.2: Average number of signal b jets (left) and signal jets (right) with respect to average number of interactions per bunch crossing with and without a cut on the jet vertex fraction (JVF). Event selection requires a pair of leptons and at least three signal jets.

after employing quality checks (*larError* $\neq$ *2, tileError* $\neq$ *2*). Cosmic muons and events with fake $\not{E}_T$ are also discarded.

Events should have at least two same-sign leptons and are divided in three channels (ee, $e\mu$, $\mu\mu$) depending on the flavor of the first two leading pre-selected leptons. The same - sign pair is defined by the electric charge of the first two leading pre-selected leptons. Finally the invariant mass of the lepton pair (M_{ll}) must be above 12 GeV, to remove heavy flavor decay chains giving same - sign leptons, trident events for which the original electron charge is wrongly measured, and resonances (ie. upsilon).

6.3.2 Signal regions definitions

The signal regions defined for the 7 TeV analysis with 2 fb^{-1} [225] were re-used for the first CONF note at $\sqrt{s}$ = 8 TeV and L = 5.8 fb^{-1} [226]. For the entire $\sqrt{s}$ = 8 TeV luminosity, more time allows a new optimization. The two main motivations are :

- The increase of center of mass energy, from 7 to 8 TeV. This leads to an increase in the sparticles production cross section as shown in Figure 2.9. As an example from 13 fb to 24 fb for bottom squark pair production ($m_{\tilde{b}}$ = 600 GeV), from 10 fb to 24 fb for gluino pairs ($m_{\tilde{g}}$ = 1 TeV).

- More complete exploration of the SUSY phase space. Gluino mediated top squark or direct first two generation squarks models are added. To increase the sensitivity to potential SUSY signatures, a reasonable and motivated approach is to define the SRs depending on the b jet multiplicity and not keeping only inclusive SRs. Namely, a signal region with 0 b jets targeting first and second squarks generations and a signal regions with more than 1 b jet for models including third squarks generation.

The signal regions optimization is done using a wide variety of SUSY models with R-parity conserved and $\tilde{\chi}_1^0$ considered to be the LSP. The main targets are models with direct gluino or direct sbottom production. This is well motivated by the *natural* spectrum showed in Figure 2.6 and the cross section production at 8 TeV. For a larger coverage of the phase space, models with direct first two generation squarks production are included as well.

A complete scan of different sets of cuts on the kinematic variables representative for the final state is considered. In total around 58000 signal region configurations are probed and the entire procedure is completely reported in [2]. Here only the main features are discussed, even if I brought an important contribution to the signal region optimization.

Several kinematic variables are used to discriminate the signal over the background :

- $\not{E}_T$: it is a key variable as the SUSY signatures are characterized by a large $\not{E}_T$ coming from the energy of the un-detected LSP candidates.

- Inclusive **effective mass**, m_{eff}: it is defined by summing the two leading leptons p_T, all signal jets and $\not{E}_T$. It represents the total energy in the transverse plane and is highly dependent on the initial particle mass.

- **Transverse mass**, m_T: it is computed with the leading lepton information and $\not{E}_T$, $m_T = \sqrt{2 \cdot p_T^l \cdot \not{E}_T \cdot (1 - cos(\Delta\Phi(l, \not{E}_T)))}$.
  It is a characteristic variable for single heavy particles (M_{hp}) decaying in two particles, one of which is invisible. One example can be the W boson decay, where the m_T is computed with the lepton produced in the decay. The end-point is the value of the heavy particle mass. Hence asking $m_T > M_{hp}$ generally does not affect to much the signal, while it removes most of the W + jets background where 1 jet is faking the lepton.

- Number of signal jets : several jet p_T thresholds are considered.

- Number of signal b jets : the p_T threshold is varied. An implicit requirement on the number of b jets is added, eg. $N_b = 0$, $N_b > 0$, $N_b > 2$, etc.

In Figures 6.3 and 6.4 some of the mentioned kinematic variables and the number of b jets distributions are illustrated. For reference, a SUSY signal from the gluino pair production via a virtual top squark, $\tilde{g} \to tt\tilde{\chi}_1^0$, model is also added ($m_{\tilde{g}}$ = 1 TeV, $m_{\tilde{t}}$ = 2.5 TeV and m_{LSP} = [300, 350,..., 550] GeV).

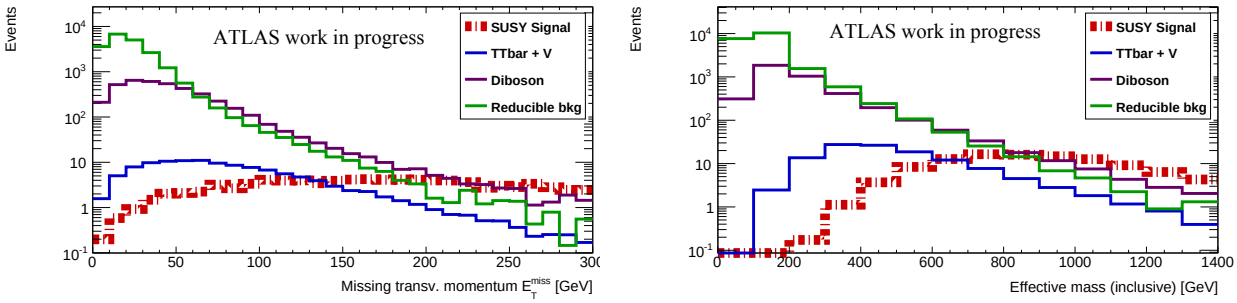


Figure 6.3: $\not{E}_T$ (left) and meff (right) distributions after requiring at least two same - sign signal leptons in the event.

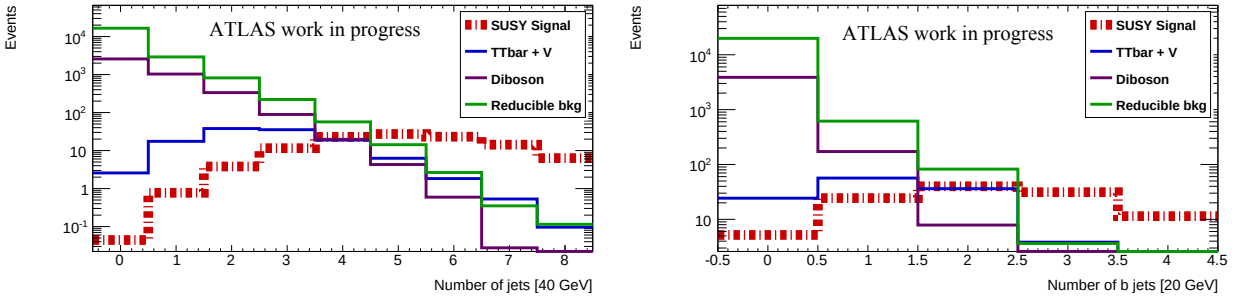


Figure 6.4: Number of signal jets (left) and b jets (right) distributions after requiring at least two same - sign signal leptons in the event.

As a result of the optimization three signal regions are defined as shown in Table 6.2. Two different sets of cuts will be considered for the discovery and exclusion cases. In the former we look for any excess over the SM in a model-independent way, whereas we use the latter one to exclude specific models, relying in particular on the shape of the m_{eff} distribution and statistically combining the orthogonal SRs. We will come back at this in Section 6.8.1.

Signal region	$\mathbf{N_{b-jets}}$	**Signal cuts (discovery)**	**Additional cuts (exclusion)**
SR0b	0	$N_{\text{jets}} \geq 3$, $\not{E}_T > 150$ GeV, $m_T > 100$ GeV, $m_{\text{eff}} > 400$ GeV	$m_{\text{eff}} > 300$ GeV
SR1b	≥ 1	$N_{\text{jets}} \geq 3$, $\not{E}_T > 150$ GeV, $m_T > 100$ GeV, $m_{\text{eff}} > 700$ GeV	$m_{\text{eff}} > 300$ GeV
SR3b	≥ 3	$N_{\text{jets}} \geq 4$	$N_{\text{jets}} \geq 5$, !SR1b

Table 6.2: Definition of the signal regions. The cuts for the discovery and exclusion cases are shown separately. For all signal regions, two isolated leading same - sign leptons with $p_T > 20$ GeV are required; (b) jets are selected with $p_T > 40$ (20) GeV.

In these SRs, the background formed by a pair of bosons is rejected by the high jet multiplicity requirement, while the p_T threshold of 40 GeV reduces the fake lepton background. Since the SUSY signal

is characterized by large $\not{E}_T$, m_T and m_{eff}, cutting on these variables suppress the real and reducible background. A b jet veto is added in one of the SRs to highly suppress the $t\bar{t}$ like background which dominates in regions with many jets.

The SR0b signal region is defined to gain sensitivity in models with gluino mediated squarks and direct squarks. The SR1b signal region is orthogonal to SR0b and targets the signatures with 1 or 2 b jets in the cascade decays. It should be underlined that for the signal region SR3b, beside the requirement on number of leptons and b jets only a cut on the number of jets is added, as this signature is very rare in the SM and the expected background is low. It is designed to cover the compressed spectra (regions with small mass difference) in models involving the third squark generation. It is also used for RPV SUSY searches characterized by small $\not{E}_T$.

To have some illustrative numbers of the performance of these signal regions, the number of signal events passing the combined SR0b, SR1b and SR3b region cuts are shown in Figure 6.5 (left) together with the acceptance times the efficiency ($A \times \epsilon$, right) for gluino pair production via virtual top squark model (in the $\tilde{g}$ - LSP mass plane) and direct bottom squarks model (in the $\tilde{b}$ - $\tilde{\chi}_1^{\pm}$ mass plane). Here the acceptance gives the number of events passing the selection cuts at generator level and the efficiency accounts for reconstruction losses like lepton identification, jet energy resolution, jet tagging efficiency, $\not{E}_T$ resolution, etc.

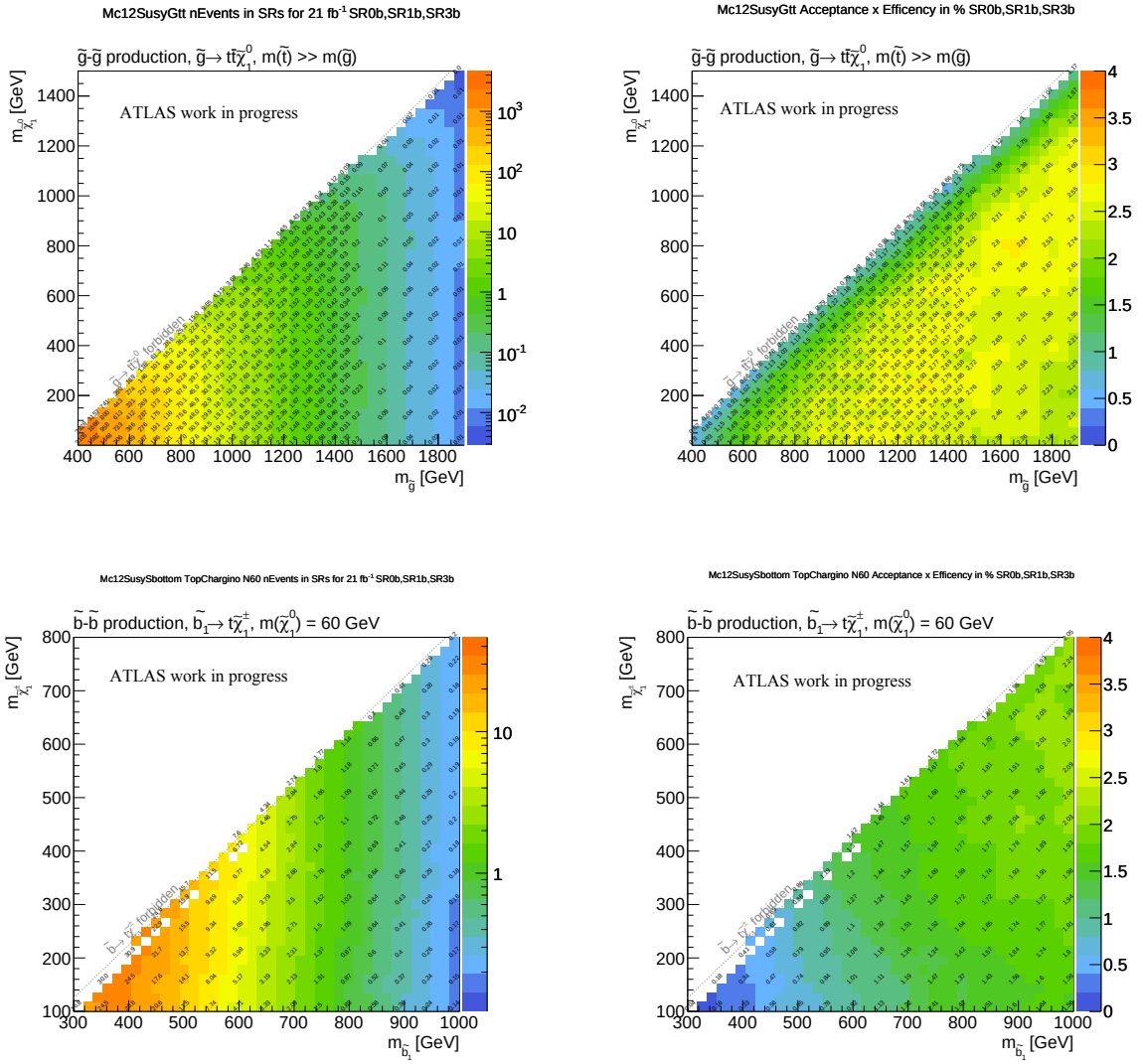


Figure 6.5: Number of signal events passing the combined SRs cuts (left) together with the $A \times \epsilon$ (right) in two models: gluino pair production via virtual top squark (top) and direct bottom squarks production (bottom).

For the gluino pair production via virtual top squark model (top), the number of events with massless LSP decreases from $\sim$ 150 at $m_{\tilde{g}}$ = 700 GeV to 10 at $m_{\tilde{g}}$ = 1 TeV and even < 1 at masses higher than 1.3 TeV. Values of $A \times \epsilon$ are around 1 - 2% near the diagonal and around 2.5% otherwise. Lower values obtained near the diagonal are due to the signal acceptance (the signal reconstruction efficiency is rather

constant in all the phase space). In this region the spectra is compressed and less energetic jets are more likely to fail the high cut from the SRs. The lower $\not{E}_T$ might bring some contribution as well. The dominant signal region is SR3b. When considering the direct bottom squark pair production model (bottom), after applying the combined signal region cuts, the number of events for massless $\tilde{\chi}_1^{\pm}$ decreases from 27 events at $m_{\tilde{b}}$ masses of 350 - 400 GeV to 4.5 events at masses of 600 - 650 GeV. The acceptance times efficiency is increasing from 0.2 to 1 for $\tilde{b}$ masses varying between 300 and 600 GeV. The low values in the compressed spectra are explained by softer jets and smaller $\not{E}_T$. From these results, one can expect this analysis to be sensitive up to ∼1 TeV in gluino mass and ∼550 GeV in bottom squark mass and not also at higher masses due to the small expected number of signal events.

6.3.3 Control and validation regions

The defined same - sign leptons signal regions are characterized by low Standard Model background and the detector background is at max half of the total estimated background. The methods used to estimate the background are validated by looking at the agreement between the data observation and SM expectation in regions dominated by one type of background. These regions are called validation regions (VR). Their definition is a balance between a high purity, a large statistics and a low signal contamination. The kinematics in these regions should be similar to the signal regions, i.e. in order to be as close as possible to the signal region. If no validation region can be designed, the background is checked by looking at different kinematic variables distributions. For a complete validation several selections are considered :

- Very loose : requiring at least two signal leptons in the event;
- Intermediate : adding soft cuts i.e at least one or two signal jets or b jets;
- Hard : increase the number of jets in the event.

To decrease the associated systematic uncertainty on the SM background in the signal region, control regions (CR) enriched in a particular type of background and orthogonal to the signal regions are preferred. However, given the rare SM processes which dominate in the SRs, the statistics obtained with $\sqrt{s} = 8$ TeV and $L = 20.3$ fb^{-1} of data in potential control regions is too small to be considered. Hence, these CR are used only as VR (Section 6.4).

6.4 Standard Model background

As already mentioned, the SM background with same - sign leptons in the final state is very low compared to other searches. Therefore, less strict requirements on the various kinematic variables can be imposed. The background composition, estimated with MC, is given below.

W or Z bosons production in association with a pair of top quarks ($t\bar{t}$ + V), where the W or Z boson and at least one W from the top quarks are decaying leptonically, providing the same signature in the final state as the SUSY signal. For illustration, few Feynman diagrams are shown in Figure 6.6 at leading and next to leading orders. The production cross sections are given in Table 6.3, together with the best experimental measurements [252–255]. The 30% systematic uncertainty on the theoretical cross section accounts for the scale (22%) and PDF (8%) variations. In the signal regions with b jet requirements, the $t\bar{t}$ + W or Z processes dominates. Due to the very low statistics (only few expected events at 8 TeV) no control regions could be defined, only validation regions.

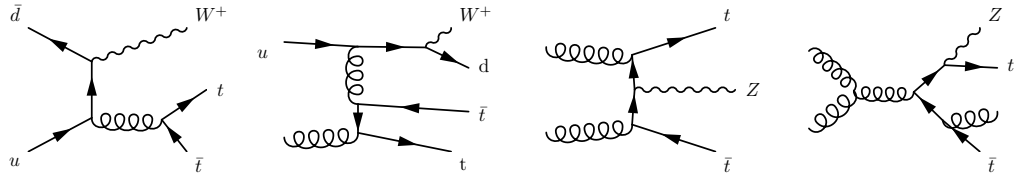


Figure 6.6: Examples of Feynman diagrams at leading and next to leading orders (LO and NLO) for $t\bar{t}$ + W and $t\bar{t}$ + Z backgrounds. Same final state as for SUSY signal with b jets is obtained, as the top quark is decaying (most of the time) into a b quark and a W boson.

The diboson background consists of **ZZ, WZ and same - sign WW processes (VV)**. It is dominant in the signal regions without b jets. Triboson WWW^*, WZW^* and ZZZ^* production is also considered, but their contribution is very small. Few Feynman diagrams are shown in Figure 6.7. The production cross

Process	Theoretical cross - section [fb]	Measured cross - section [fb]	Observed σ	Expected σ
$t\bar{t} + W$	232 ± 67	300^{+140}_{-110}	3.1	3.7
$t\bar{t} + WW$	1 ± 0.3	-	-	-
$t\bar{t} + Z$	206 ± 62	150^{+58}_{-54}	3.1	2.3

Table 6.3: Theoretical and measured cross sections [fb] at $\sqrt{s} = 8$ TeV [252,253,255] together with its total uncertainty of the individual $t\bar{t} + W$ and $t\bar{t} + Z$ signals. For the experimental results also the observed and expected significances are shown.

sections [256–259] are presented in Table 6.5. An overall systematic uncertainty on the theoretical cross section of 7% is used. These backgrounds are reduced by the high jet multiplicity required in the SRs and remain important only in SR0b.

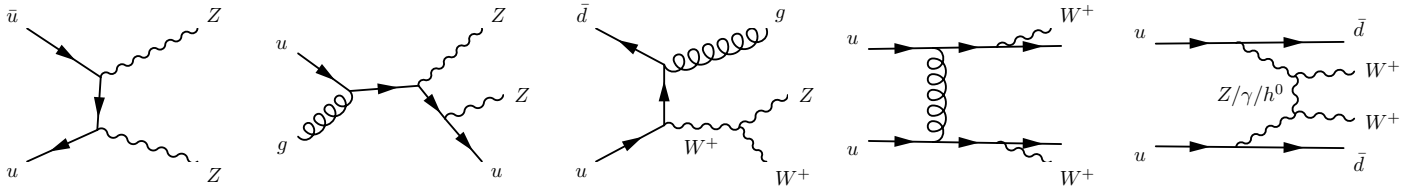


Figure 6.7: Examples of Feynman diagrams at LO and NLO for ZZ, WZ and $W \pm W\pm$ backgrounds. Same final state as for SUSY signal without b jets is obtained, as is characterized by leptons, jets and $\not{E}_T$ (when W bosons are present).

Process	$ZZ \to llll$	$WZ \to lll\nu$	$W^{\pm}W^{\pm} \to ll\nu\nu$	WWW^*, WZW^*, ZZZ^*
σ [pb]	8.74	9.75	5.49	0.4

Table 6.4: Production cross sections[pb] at $\sqrt{s} = 8$ TeV [256–259] of the individual $ZZ \to llll$, $WZ \to lll\nu$, $W^{\pm}W^{\pm} \to ll\nu\nu$, WWW^*, WZW^* and ZZZ^* signals.

Exotic processes. In this category enter sources as $t\bar{t} + H$, WH, ZH, 4 top quarks and tZ. They are all expected to be negligible close-by and in the signal regions. Table 6.5 shows the cross sections of these processes. An inclusive systematic uncertainty on the cross sections of 100% is used. For the preliminary results presented in this chapter none of the backgrounds are considered, as the samples where not yet available.

Process	ttbar + H	WH	ZH	4 t	tZ
σ [pb]	0.03	0.28	0.28	0.0006	0.035

Table 6.5: Production cross sections [pb] at $\sqrt{s} = 8$ TeV [260–262] of the individual $t\bar{t} + H$, WH, ZH, 4 top quarks and tZ signals. The Higgs boson is assumed to decay in a pair of W or Z bosons and the final cross section is accounting for the corresponding branching ratios.

Validation regions

Several validation regions are defined for the main Standard Model background processes :

- A $t\bar{t} + W$ validation region is defined with exactly two same - sign muons, at least one jet with $p_T > 30$ GeV and exactly 2 b jets with $p_T > 20$ GeV, $m_T > 80$ GeV and $\not{E}_T$ within (20, 150) GeV. To suppress the fake background contribution no $e\mu$ or ee pairs are considered and tight cuts on m_T and $\not{E}_T$ variables are applied. The cut on number of jets reduces the diboson contamination.
- For the $t\bar{t} + Z$ validation region, the number of leptons is increased to three (two expected leptons from the Z boson and another lepton from the W boson) and events with electrons are allowed, as the reducible background is lower in this case. Third signal lepton p_T is > 10 GeV and additional cuts on number of jets (≥ 2, $p_T > 40$ GeV), b jets (≥ 1 and < 3, $p_T > 20$ GeV), $\not{E}_T$ ($20 \leq \not{E}_T \leq 120$ GeV) and m_T (>100 GeV) are added. The increase in number of leptons is reducing almost all contamination from $t\bar{t} + W$ and same - sign WW, and the number jet and b jet cuts discarded the remaining diboson processes. The fake backgrounds are reduced by the m_T, $\not{E}_T$ and three lepton cuts.

Two validation regions are designed for the diboson background, with and without jets in the event :

- A first diboson validation region is defined using only same sign muons, at least 2 jets with $p_T > 20$ GeV, no b jets, $m_T > 100$ GeV and $\not{E}_T$ within (20,120) GeV range. Events with leading ee or $e\mu$ pairs are not considered due to the high fake background contamination. The tight cuts on $\not{E}_T$ and m_T is rejecting further this component, and the jet and b jet cuts are removing mainly the $t\bar{t} + V$ processes which otherwise dominates over the other SM backgrounds.

- Requiring no jets and b jets in the event, a second diboson validation region is defined with same - sign muons and same $\not{E}_T$ interval. The m_T cut is replaced by one on the m_{eff} variable ($150 \leq m_{\text{eff}} \leq 400$ GeV) for a better fake background rejection.

Figure 6.8 shows the results in the diboson and combined $t\bar{t} + Z$ and $t\bar{t} + W$ validations regions. The methodology used to estimate the detector backgrounds, quite small in this regions, is described later in this chapter. The agreement between the observed events in data and the estimated background is within 1.5σ and limited by the data statistics. Monte Carlo simulations show than in general the expected contamination from SUSY signal processes is found < 10 GeV, but for some specific models like direct bottom squark it can increase up to 25%.

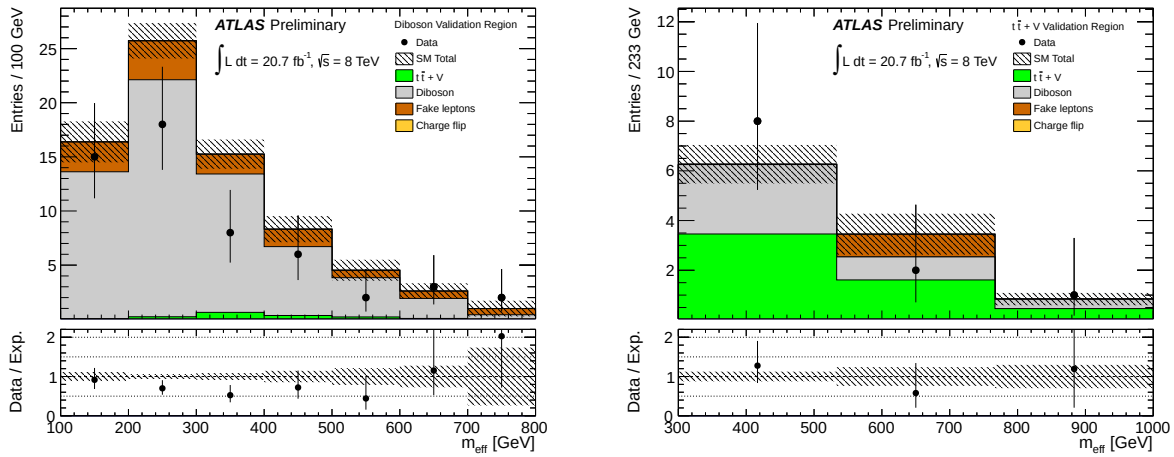


Figure 6.8: Results in diboson (left) and $t\bar{t} + V$ (right) validation regions.

As already mentioned, the small statistics and/or the very different kinematics prevent from using these validation regions as control regions. However, the background checks in these VR provide confidence that the MC simulations predict the SM background with same - sign leptons in events with several jets and b jets within a factor $\sim$1.5 given the available data statistic.

6.5 Charge flip background

Electron charge flip, or wrong charge measurement, can be due to a hard bremsstrahlung or a low curvature of the track :

- Hard bremsstrahlung. Due to the interaction with the material an electron radiates a hard photon that converts into an electron-positron pair (trident event, Figure 6.9). During the reconstruction, the final energy deposited in the calorimeter can be matched with the radiated positron track if the energy of the radiated electron is too small. If this happens, the charge of the initial electron is wrongly measured. Of course, same process can appear also for positron candidates. The generic term *electron* will be used both for electrons and positrons.

- The electron charge is measured using the curvature of the track : if the curvature is small, due the high number of reconstructed tracks and the limited resolution, the high momentum electrons can have a wrong charge measurement. Due to the improved performance of the electron identification algorithm this source is minimized for 8 TeV data.

Studies based on MC simulations using the truth particles information showed that this background occurs mainly in $t\bar{t}$ events in which both W bosons from the top quarks decay leptonically. Figure 6.10 left shows the distribution of the opposite - sign and same - sign invariant mass of the leading signal electrons in data in a region with $|\eta| > 2.1$. A significant peak is observed for the SS leptons distribution, corresponding to 1% of the total events.

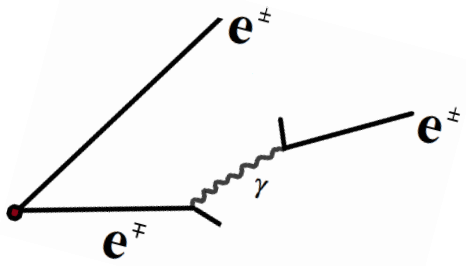


Figure 6.9: Example of process leading to electron wrong charge measurement.

For muons this background is very small due to the very rare muon bremsstrahlung and the usage of inner detector and muon spectrometer for the charge measurement. This is illustrated in Figure 6.10 right, where the opposite - sign and same - sign invariant mass distributions are shown for signal muons in a region with $|\eta| > 2.1$. No same - sign peak is visible.

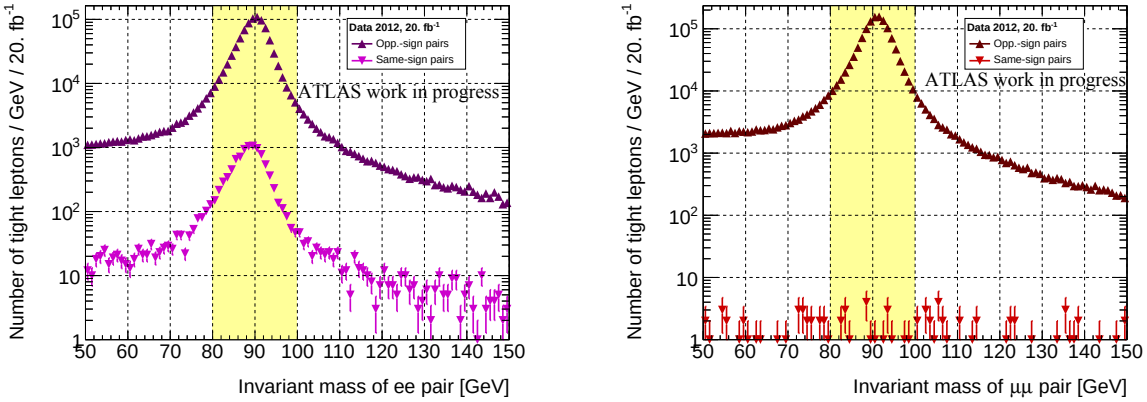


Figure 6.10: Opposite - sign and same - sign invariant mass distribution for signal electrons (left) and muons (right) in the a region with $|\eta| > 2.1$ ($\sqrt{s} = 8$ TeV, $L = 21$ fb^{-1}). For muons, no Z boson mass peak is seen when using SS events while for electrons a clear peak is visible.

6.5.1 Electron charge flip background estimation

To reduce the charge flip rate for the baseline electrons, beside the *medium++* identification an additional d_0/σ_{d_0} cut (< 5) is applied as mentioned in Section 6.3.1. For the signal *tight++* electrons the cut is tighten, from 5 to 3. A further reduction is thus obtained, as shown in Figure 6.11. At large η, where the number of trident events is higher, the reduction is sizable for both baseline and signal electrons. The impact of this new d_0/σ_{d_0} cut on SUSY signal is small (loss < 5 % in acceptance).

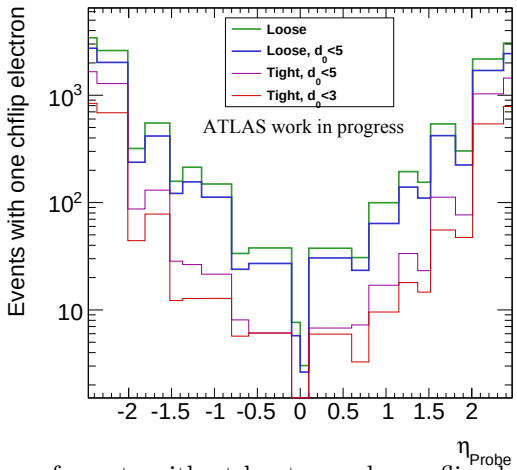


Figure 6.11: Impact on number of events with at least one charge flip electron when changing the d_0/σ_{d_0} cut threshold, after same - sign lepton selection at baseline level (green and blue lines) and tight level (red and magenta). Distributions are shown in data, where $\sqrt{s} = 8$ TeV and $L = 21$ fb^{-1}.

The remaining charge flip background is evaluated using a fully data driven approach. In the region of interest, opposite - sign pairs are weighted with the charge flip probability (rate). The probability is considered to be per selected electron. The event weight (w_{os}) is obtained with Equation (6.1) :

$$w_{\text{os}\to\text{ss}} = \zeta_1 + \zeta_2 - 2\zeta_1\zeta_2 \tag{6.1}$$

where the ζ_1 and ζ_2 are the probabilities for the leading and/or subleading electrons to have a wrong charge; for the electron-muons pairs the ζ_μ is zero. The event weight is obtained in the region of interest after applying same cuts as for the SS events estimation (the only difference being the cuts on the electric charge, which should be OS instead of SS).

Electron charge flip rate measurement

The electron charge flip rate (ζ) measurement is performed using a kind of $Tag\&Probe$ method (close to the one presented in Section 5.1.1) in a data sample enriched in $Z \to ee$ events, with no cut on the electric charge of the tag and probe pair to increase the statistic and avoid any bias. The dependency on probe electron η and p_T is considered, as the bremsstrahlung and conversions dominates in the high η region for high energetic electrons. To minimize the contamination of light jets faking electrons, the invariant mass of the electron pair is restricted to [75, 105] GeV interval.

The charge flip background estimation must be performed for signal electrons and also for electrons failing the *tight++* and isolation cuts. For this, two sets of charge flip rate measurements are implemented, with two signal electrons in the events and with one signal and one baseline electron. If the probability of having a charge flipped electron is the same for the two electrons then the measurement will be simply the ratio between number of SS events and (SS + OS) events in the (i, k) electron η and p_T bin :

$$\zeta_i = N^{ik}_{SS}/(N^{ik}_{SS} + 2 \cdot N^{ik}_{OS}) \tag{6.2}$$

This ideal case is not applicable as the electrons do not necessary have the same kinematics. Three methods are considered and compared to obtain the final numbers in a given (η,p_T) interval. Other methods can be found in [166].

In the first method ($M0$) both electrons are required to come from the same p_T and η bin. Equation (6.2) is used, as the rate is considered to be uncorrelated between the two electrons. The major inconvenience is the lack of statistics due to the strong inefficiency of this requirement. To avoid this, an alternative method, $M1$, considers only electrons coming from the same η regions (since the rate dependency is much stronger for η than for p_T).

The last method is a $Tag\&Probe$ method, $M2$. The tag is always the electron with the correct charge. To fulfill this requirement in data the tag should be in the barrel ($|\eta| < 1.37$) since the charge flip is significantly smaller in this region. The probe should be in the end-cap region ($|\eta| > 1.37$) to increase the probability of being the electron with a wrong charge. If the probe electron is coming from the barrel region then the method is expected to fail. With this method the charge flip rate is measured as :

$$\zeta_i = N^{ik}_{SS}/(N^{ik}_{SS} + N^{ik}_{OS}) \tag{6.3}$$

To chose the correct method among the three, a comparison between the truth rate and the results obtained after the $M1$ and $M2$ methods are applied on $Z \to ee$ MC sample is performed. The one leading to the smallest bias with respect to the truth, gives the final results. To measure the truth rate, the information at the truth level is used to define the SS and OS pairs. In true SS pairs, one electron is matched to a conversion or to a primary electron for which the reconstructed charge differs from the generated charge. In true OS pairs reconstructed electrons are matched to primary electrons for which the reconstructed charge is the same as the generated charge.

Given the output of this comparison, $M1$ is chosen in the barrel region (<1.37) due to the lack of statistics after requiring both electrons to be in the same η bin, while $M2$ is used in the high eta ($1.52 < |\eta| < 2.47$) and crack regions ($1.37 < |\eta| < 1.52$).

Results using $Z \to ee$ data events.

The measured charge flip probability varies between 10^{-4} and 0.02 in the range $0 \leq |\eta| < 2.47$ and $20 < p_T < 200$ GeV. For electrons with $p_T > 200$ GeV the same probability as in the last p_T bin is assumed. Results in different (p_T, η) intervals are shown in Figure 6.12 for electrons passing the signal definition (left) and failing the *tight++* and isolation requirements (right).

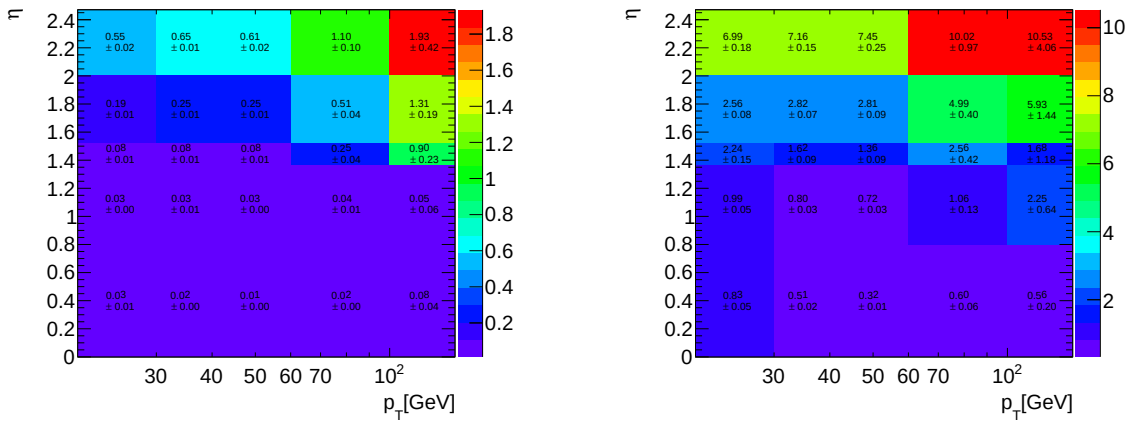


Figure 6.12: Charge flip rate [%] for electrons passing the signal definition (left) and failing the *tight++* and isolation requirements (right). Only the statistical uncertainties are shown.

Systematic uncertainty

The systematic uncertainty associated to the measurement is the bias between the rate obtained with the chosen method and the truth rate. For the signal electrons it is 40% for $|\eta| < 1.4$ and 10% for the other region, while for the baseline electrons the systematic uncertainty in the barrel is decreasing to 10%. Even if the uncertainty is rather high in the barrel the impact on the background estimates is small since the charge flip rate dominates in the high η region.

No background subtraction was performed for any of the methods, but the region used to extract the rate is expected to be background free in the high p_T region. For region with low p_T electrons the background can be higher, due to the light jet contamination. The final systematics are considered to cover the difference in the rates with and without the background subtraction.

6.5.2 Muon charge flip background estimation

No muon charge flip rate estimation is done for this analysis. Studies from exotics and SUSY searches [263], shows that the upper limit on the expected yield of events with one charge flipped muon is negligible compared to the electron case. To validate this assumption, applying the present object and event selection, the truth information from $t\bar{t}$ MC simulations is used. The number of same - sign muons with a flipped charge is found to be very small (< 0.1 %) near by and in the signal regions.

6.6 Fake lepton background

The fake lepton background is characterized by non isolated leptons arising from semileptonic heavy quarks decay, photon conversion into a e^-e^+ pair, trident electrons or hadron decays. For the later, muons comes from pion or kaon decays in the muon spectrometer.

Using $t\bar{t}$, W + jets MC samples and baseline leptons (instead of tight to increase the statistics) the truth composition is displayed in Figure 6.13 : prompt with a wrong reconstructed charge, conversion (here the converted electron is not coming necessarily from a boson), b and c quarks decays and other processes. The last category includes decays from hadrons, light mesons, etc. Two cases are considered : the leading lepton is prompt and the subleading is a fake (dark blue), and the leading is a fake and the subleading is the prompt lepton (light blue). Processes in which both leptons are fakes are not showed because they are negligible. Note that no separation between fake leptons and charge flip events is considered in this set of figures.

As expected, the case with a prompt leading and a fake subleading lepton is dominating. Note that, in the conversion category the leading lepton has a higher probability to be of wrong charge, so this is not true. The principal source of fake leptons comes from the decay of a b quark. For regions with a b jet veto, dominant contributions are also coming from photon conversions and hadron decays.

The fake leptons are highly decreased after requiring the *tight++* identification for electrons and adding the isolation cuts. In the signal region, the high cuts on $\not{E}_T$ and number of jets are further reducing it. However, the dominant source remains the $t\bar{t}$ processes with one real lepton from the W boson and one fake lepton from the b quark.

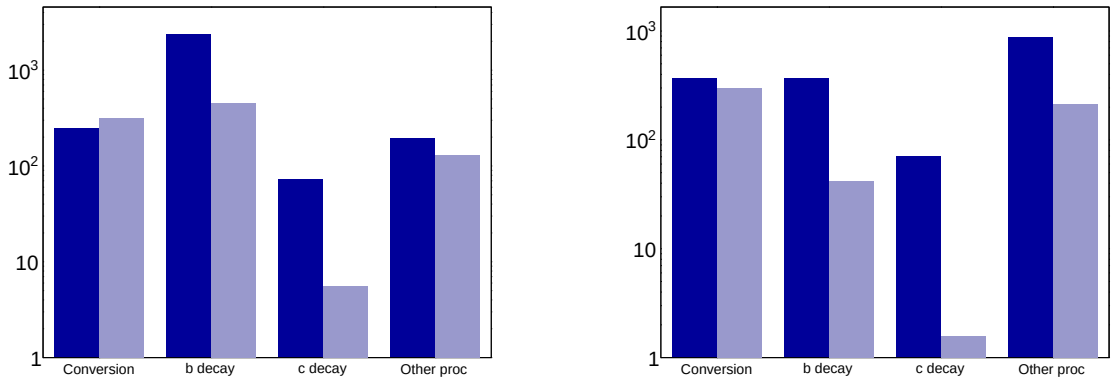


Figure 6.13: Truth composition after two same - sign baseline leptons with $p_T > 20$ GeV selection in $t\bar{t}$ (left) and W + jets MC (right). On the x axis the origin of the fake lepton is shown. Two cases were considered: prompt leading and fake subleading lepton (dark blue), and fake leading and prompt subleading lepton (light blue).

6.6.1 The matrix method

To estimate this type of background a fully data driven method, so called *Matrix Method* is used since the MC is not reliable. This method relies on the classification of the same - sign leptons candidates according to their isolation and identification criteria : real leptons are expected to be well isolated compared to jets faking leptons. Therefore a sample with more relaxed identification cuts is used and divided in four events categories :

- *Tight - Tight, TT* : both leptons are satisfying the signal definition (Section 6.3.1).
- *Tight - Loose, TL* : the leading lepton is tight and the subleading lepton is failing any of the requirements on the tracking isolation, calorimeter isolation, $z_0 sin\theta$ and/or *tight++*[7].
- *Loose - Tight, LT* : the leading lepton is loose and the subleading lepton is tight.
- *Loose - Loose, LL* : both leptons are loose in the event.

In the following the real efficiency, ε, is the probability of one real lepton to pass the signal lepton cuts, while the fake efficiency, ξ, is the probability of a fake lepton to pass the signal lepton cuts. Using these two parameters, the number of observed events with tight / loose leptons (N_{TT}, N_{TL}, ...) can be expressed as a function of the number of event with real / fake baseline leptons (N_{RR}, N_{RF}, ...) :

$$\begin{pmatrix} N_{TT} \\ N_{TL} \\ N_{LT} \\ N_{LL} \end{pmatrix} = \underbrace{\begin{pmatrix} \varepsilon_1\varepsilon_2 & \varepsilon_1\xi_2 & \xi_1\varepsilon_2 & \xi_1\xi_2 \\ \varepsilon_1(1-\varepsilon_2) & \varepsilon_1(1-\xi_2) & \xi_1(1-\varepsilon_2) & \xi_1(1-\xi_2) \\ (1-\varepsilon_1)\varepsilon_2 & (1-\varepsilon_1)\xi_2 & (1-\xi_1)\varepsilon_2 & (1-\xi_1)\xi_2 \\ (1-\varepsilon_1)(1-\varepsilon_2) & (1-\varepsilon_1)(1-\xi_2) & (1-\xi_1)(1-\varepsilon_2) & (1-\xi_1)(1-\xi_2) \end{pmatrix}}_{\Sigma} \begin{pmatrix} N_{RR} \\ N_{RF} \\ N_{FR} \\ N_{FF} \end{pmatrix}_{\text{baseline}} \tag{6.4}$$

where ε_i is the probability of a tight lepton i to be real, $(1-\varepsilon_i)$ is the probability of a loose lepton i to be real, ξ_i is the probability of a tight lepton i to be fake and $(1-\xi_i)$ is the probability of a loose lepton to be fake.

To obtain the number of events with real / fake signal leptons the system is reverted and the number of events with tight / loose leptons is weighted with the probability of being signal leptons, ie. $(N_{TT})_{signal} = \varepsilon_1\varepsilon_2 N_{RR}$:

$$\begin{pmatrix} N_{RR} \\ N_{RF} \\ N_{FR} \\ N_{FF} \end{pmatrix}_{\text{signal}} = \underbrace{\begin{pmatrix} \varepsilon_1\varepsilon_2 & 0 & 0 & 0 \\ 0 & \varepsilon_1\xi_2 & 0 & 0 \\ 0 & 0 & \xi_1\varepsilon_2 & 0 \\ 0 & 0 & 0 & \xi_1\xi_2 \end{pmatrix}}_{\Sigma'} \times \Sigma^{-1} \times \begin{pmatrix} N_{TT} \\ N_{TL} \\ N_{LT} \\ N_{LL} \end{pmatrix} \tag{6.5}$$

[7] In the muon case the last two cuts are not considered.

and the final number of fake signal leptons the region of interest is simply:

$$N_{fakes} = (N_{RF} + N_{FR} + N_{FF})_{\text{signal}} = \Sigma' \times \Sigma^{-1} \times \begin{pmatrix} N_{TT} \\ N_{TL} \\ N_{LT} \\ N_{LL} \end{pmatrix} \tag{6.6}$$

The matrix method is applied to data ($MxM.D$) which is formed by prompt SS leptons (PP), fake leptons (FL) and charge flip electrons (CF) :

$$MxM.D = MxM.(PP + FL + CF) = MxM.FL + MxM.CF \tag{6.7}$$

By construction the number of prompt SS leptons is equal to zero.

It should be mentioned that the charge flip electrons are less likely to pass the signal cuts than the electrons with a correct charge. The number of charge flip events must be removed at the pre-selection level before applying the matrix method, as we are interested in estimating only the fake lepton component. In reality this cannot be done as the matrix method is applied event by event. As a results, the matrix method is applied to the data including the charge flipped electrons ($MxM.D_{OS+SS}$) and to the data used for the charge flip background estimation ($MxM.D_{OS}$) [8]. The final fake lepton background yield is simply :

$$N_{fakes} = MxM.D - w_{OS \to SS} \times MxM.DOS \tag{6.8}$$

Discussion.

In Equation (6.8), N_{fakes} is an event weight and the final estimation is the sum over all the event weights passing the region selection requirements. In this approach, the charge flip background must be evaluated for all tight - loose leptons combinations. This is the reason of extracting charge flip electron rates also for the electrons failing the signal leptons definitions.

To have a viable fake lepton background estimation, the real and fake efficiencies, ε_i and ξ_i, must be measured independently, in regions having the kinematics of the signal regions. This is quite difficult in the search with same - sign leptons as the statistic is very low. Asking for a high $\not{E}_T$ cut or high jet multiplicities reduces severely the statistics. To avoid this, a sample with loose or no cuts on number of (b) jets is therefore considered to derive the lepton efficiencies. The final systematic uncertainties are computed to cover this difference in leptons kinematics.

6.6.2 Real lepton efficiency, ε

It is measured in a high purity data sample with the *Tag&Probe* method applied to leptons in $Z \to e^{\pm}e^{\mp}/\mu^{\pm}\mu^{\mp}$ events. To assure the tag is a real lepton it must fulfill the signal leptons requirements and any bias in the choice is removed by considering alternatively both leptons as tags. The measurement is performed on the probe inside an invariant mass window of $80 < M_{ll} < 100$ GeV, as function of p_T and η. Note that the results obtained for the real electron identification efficiency are different compared to Section 5.3.4, because now the it is measured with *medium++* electron probes and *tight++* plus tight isolation electron tags.

For $p_T < 40$ GeV, the real electron efficiency is found to vary between 86% and 91% increasing up to ∼95% for more energetic electrons. As expected, the muon efficiency is higher for the entire p_T range, from ∼95% at low p_T to 99% at high p_T. No background subtraction is performed given the high rejection available for the baseline leptons. Instead the difference between real efficiency in data and MC is taken as a systematic uncertainty, which is conservative as it includes also the difference between the real and simulated efficiencies.

The real lepton efficiency, extracted in the region dominated by well isolated leptons from the Z bosons, is higher than in the signal regions. This degradation is explained by the high hadronic activity present in the SRs, a characteristic of all strong production SUSY searches. A final overall systematic uncertainty of 3% for both electron and muon real efficiencies is considered by studying the dependency on the jet multiplicity and m_{eff} in $t\bar{t}$ MC.

6.6.3 Fake lepton efficiency, ξ

The fake lepton efficiency is measured in samples dominated by two same - sign leptons with $p_T > 20$ GeV for which one of the lepton is a fake. To remove a good fraction of diboson events, a veto on third lepton with

[8] The event weight obtained after applying the matrix method on OS data must be multiplied with the charge flip weight. This gives the estimation for the charge flip component.

$p_T > 10$ GeV is considered. The remaining contamination with real leptons is removed using MC samples; in the electron case also a non-negligible charge flip contamination is present and corrected using data estimation (as explained in Section 6.5). The SUSY signal contamination is found to be negligible, maximum 9% (0.3%) in the region used for the muon (electron) fake rate extraction. The different approaches used to obtain the fake lepton efficiencies are described and only the overall results are given since several improvements are presented in the next chapter.

Fake electron efficiency measurement

As the ee channel is dominated by charge flip electrons, the measurement is performed using $e\mu$ pairs for which the muon is assured to be tight. To fulfill this condition the signal muon must have $p_T > 40$ GeV. This assumption is reasonable as the fake signal muons are negligible above this p_T threshold, Figure 6.14. A clear gap between the data observation and prompt lepton estimation is present for $p_T < 40$ GeV.

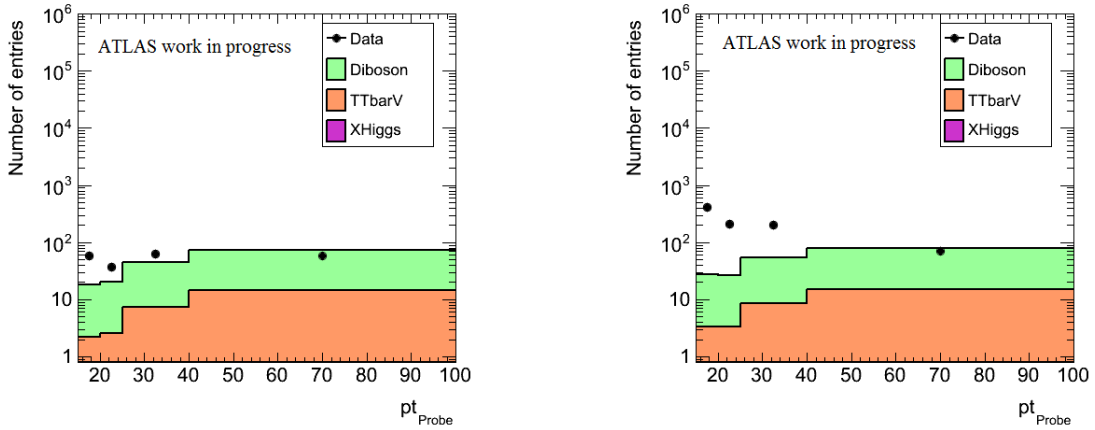


Figure 6.14: Transverse momentum distribution of the probe tight muon (left) and of the probe tight or loose muon (right) after requiring at least two muons and at least two jets with $p_T > 20$ GeV in the event ($L = 21$ fb^{-1} and $\sqrt{s} = 8$ TeV). Below 40 GeV events with fake muons dominates, whereas they are negligible at higher p_T.

The fake electron efficiency is highly sensitive to the heavy flavor component in the event : it is small if the source of the fake electrons comes from the b quark decay and high when electron conversion or light jet sources dominates. As the signal regions are defined depending on the b-jet multiplicity, the fake rate must be measured in regions with and without b jets.

After the prompt and charge flip background subtraction, the rate is simply measured using a *Tag&Probe* method, as the ratio between the number of tight (N_T) and tight + loose ($N_T + N_L$) electrons :

$$\xi = \frac{N_T}{N_T + N_L} \tag{6.9}$$

The available statistic with electrons with $p_T < 35$ GeV allow to perform the measurement in 2 η bins (0 - 1.5 and 1.5 - 2.47). As the fake lepton composition varies with the p_T range, this dependency is considered as well in the measurement.

Fake muon efficiency measurement

Di-muon same - sign pairs and at least 2 jets with $p_T > 25$ GeV are used for the measurement. Due to the poor statistic only the p_T dependency is considered. No extraction using data for muons with $p_T > 40$ GeV could be performed as the diboson contamination is very high. Therefore the same fake rate as in the last bin, $p_T < 40$ GeV, is used[9]. A similar *Tag&Probe* method as for electrons is employed in the $p_T < 40$ GeV interval. Both muons are considered alternatively as tag to avoid any bias in the measurement. Note that, as fake muons mainly comes from b jet decays, the baseline values are obtained with no cut on the number of b jets.

Corrections for regions with three b jets

In regions with ≥ 3 b jets the statistic is too poor in data and MC to perform a reliable fake lepton efficiency measurement. The small statistics available in MC allows only for the measurement of an extrapolation factor from regions with 2 b jets to regions with 3 b jets. This factor, equal to 1.5, is applied to

[9]This rate is corrected with a factor obtained using the truth information from $t\bar{t}$ MC. This factor is found to be 1.16 ± 0.14. In $t\bar{t}$ MC less than 4% muons are found to be fakes, hence this correction has a very small effect on the total fake muons estimation in any region.

the fake leptons background yield in the regions with at least 3 b jets. As the number of events with three b jets and one fake lepton is very small, this correction has a small impact on the results in SR3b.

Sources of systematic uncertainties on the ξ parameter measurement

For the fake lepton efficiency measurements, several sources of systematics are considered :

- Statistical uncertainty due to real-real lepton subtraction : it is varied by 30%, to cover the uncertainty on the production cross section, MC statistics, etc.
- Assumptions on the fake rate measurements : the method used to derive the efficiencies neglects the cases where the tag muon is fake and the baseline lepton, on which the measurement is performed, is real. The associated systematic uncertainty is computed as the ratio between the number of events with loose - tight leptons over tight - loose leptons ordered by p_T. It is found to be negligible with respect to other sources (2 - 3%).
- Muon fake rate measurement for $p_T > 40$ GeV: 66% systematic uncertainties resulted from the MC based correction. It is found to not affect too much the results in the SRs as the dominant source is coming from the low statistic in the signal regions (Section 6.8).
- Electron fake efficiency extrapolation for the b jet veto case : 50% for 40 $<p_T <$ 60 GeV bin and 90% $p_T \geq 40$ GeV.
- Lepton kinematics : to assure the fake rate extrapolation from regions with low jet multiplicity and $\not{E}_T$ to the SRs, several measurements are performed with several jet multiplicities and p_T thresholds, or different cuts on m_T, m_{eff} and $\not{E}_T$. The full symmetrized maximum deviation is taken as a systematic uncertainty. For electrons, it varies in the range 15% - 90% for the rates extracted with a b jet veto requirement (the high systematic is mainly for electrons with $p_T > 40$ GeV) and in the range 20% - 70% when at least one b jet is considered. For the later case the systematics are higher because of the low data available statistics. For muons, it varies between 27% and 50% in $p_T < 40$ GeV range. As a cross check, the fake efficiencies are computed with the truth information from MC for high jet multiplicities and m_{eff}. When comparing to data results, the difference is within the uncertainties.
- Corrections for regions with 3 b jets : as not enough statistic is available to extrapolate to this region a conservative uncertainty of 100% is applied. This value covers the possibility of a further increase with the number of b jets.

Results

Tables 6.6 - 6.8 summarizes the electron and muon fake rates obtained, together with their statistical and systematical uncertainties. In all cases the systematic component dominates over the total statistical uncertainty. Electron fake rate in a region with at least one b jet in the event varies between 5% and 46% depending on the electron p_T and η range (Table 6.6). The data statistical uncertainty is between 11% and 40% because of the tight same - sign lepton and b jet requirement.

Region	Fake rate [%]	Total uncert. [%]	Stat. (data) [%]	Stat. (MC subt.) [%]	Syst. [%]
p_T =20-30 GeV, $\|\eta\| < 1.5$	5	72	40	0	60
p_T =20-30 GeV, $1.5 < \|\eta\| < 2.5$	14	77	29	0	71
p_T =30-40 GeV	17	39	18	0	35
p_T =40-60 GeV	27	30	15	0	26
$p_T \geq 60$ GeV	46	20	11	0	17

Table 6.6: Measured electron fake rate including relative statistical and systematic uncertainties [%], in the presence of at least one b-jet.

Electron fake rate in a region without b jets in the event varies between 4% and 20% (Table 6.7). The same efficiency as in the range 30 $<p_T <$ 40 GeV is used for electrons with $p_T > 40$ GeV due to inherent difficulty to have a reliable extraction. The data statistical uncertainty is ∼5% - 25% and dominates over the MC statistical uncertainty.

Finally, the muon fake rate varies in 11% - 32% range and the data statistic uncertainty is around 25% for $p_T <$ 40GeV, as presented in Table 6.8. This source dominates over the statistical uncertainty associated to the MC subtraction.

Region	Fake rate [%]	Total uncert. [%]	Stat. (data) [%]	Stat. (MC subt.) [%]	Syst. [%]
p_T =20-30 GeV, $\|\eta\| < 1.5$	4	56	25	0	50
p_T =20-30 GeV, $1.5 < \|\eta\| < 2.5$	15	28	7	0	27
p_T =30-40 GeV	20	16	5	0	15
p_T =40-60 GeV	20	50	5	0	50
$p_T \geq 60$ GeV	20	90	5	0	90

Table 6.7: Measured electron fake rate including relative statistical and systematic uncertainties [%], for a veto on b-jets.

Region	Fake rate [%]	Total uncert. [%]	Stat. (data) [%]	Stat. (MC subt.) [%]	Syst. [%]
p_T =20-25 GeV	11	38	27	0	27
p_T =25-40 GeV	20	56	25	0	50
$p_T > 40$ GeV	32	66	–	–	66

Table 6.8: Measured muon fake rate including relative statistical and systematic uncertainties [%]. For $p_T > 40$ GeV, the statistical uncertainty from the fake rate for p_T =25-40 GeV is propagated and the systematic uncertainty on the MC-based correction is included.

6.7 Validation of background estimation

The methods used to estimate the SM and detector backgrounds are validated in regions with more busy and energetic events compared to the measurement regions. For completeness, several kinematic variables are probed individually. To avoid any bias, the overlap with the signal regions is minimized by rejecting events with $\geq$ three signal jets and $\not{E}_T > 150$ GeV.

Therefore, the data and background estimations are compared in ee, $e\mu$ and $\mu\mu$ channels for regions with and without a b jet in the event. The split by channel is considered as the background composition varies with respect to the dominant lepton flavor in the event. In the ee channel the charge flip background is validated in regions with low $\not{E}_T$ and m_T. The fake electrons are verified in the $e\mu$ channel, while the fake muons are checked in the $\mu\mu$ channel at low and high jet multiplicity respectively. The real diboson processes are checked in the $\mu\mu$ channel in the regions with a b jet.

Some key variables distributions, the number of jets and b jets, are shown in Figure 6.15, in all three channels after the same - sign lepton selection. The magenta shadowed region represents the total uncertainty on the background estimates, while the black lines represents the data associated statistical uncertainty. A good agreement is observed between data and SM background. Note that in the $e\mu$ channel the discrepancy is not covered by the total uncertainty in the bin with 7 jets with $p_T > 40$ GeV, but it is considered to be only a data fluctuation in a region with very limited statistics (3 observed events). Another disagreement in this channel, around 2σ, is present in the number of b jets distribution. Here the statistic in data is higher (10 events). As the charge flip and the diboson contributions are expected to be very low, the source of this discrepancy can be an underestimation of the fake lepton background. On the other distributions, $\not{E}_T$, m_T, m_{eff}, no significant discrepancy is observed.

To demonstrate that the b tagging efficiencies in data are well modeled by MC, a region with at least 3 b jets and opposite - sign leptons are considered to increase the statistics. The backgrounds sources, different than in SS region, are the diboson production, $t\bar{t}$, $t\bar{t} + V$, single top, Z plus jets and fake leptons. When comparing the agreement with observed data, a good compatibility is found for a large set of kinematic distributions. For illustration the leading and subleading b jet p_T distribution is shown in Figure 6.16. The agreement is very important since the 3^{rd} tagged b jet is generally a fake. Having the confirmation of a good simulation of the tagging efficiency in MC, gives confidence on the results obtained in SR3b.

6.8 Preliminary results

Given the good understanding and validation of the detector and SM backgrounds, the $\sqrt{s} = 8$ TeV data corresponding to 21 fb^{-1} integrated luminosity is used to obtain preliminary results in the search for supersymmetric particles with the two same - sign leptons channel. The simultaneous fit method is described in Section 6.8.1, while in Section 6.8.2 the considered systematic uncertainties are detailed. In Section 6.8.3 the results in the signal regions are presented. General limits on BSM physics are mentioned in Section 6.8.4, while in Section 6.8.5 the interpretation in several simplified SUSY models is shown.

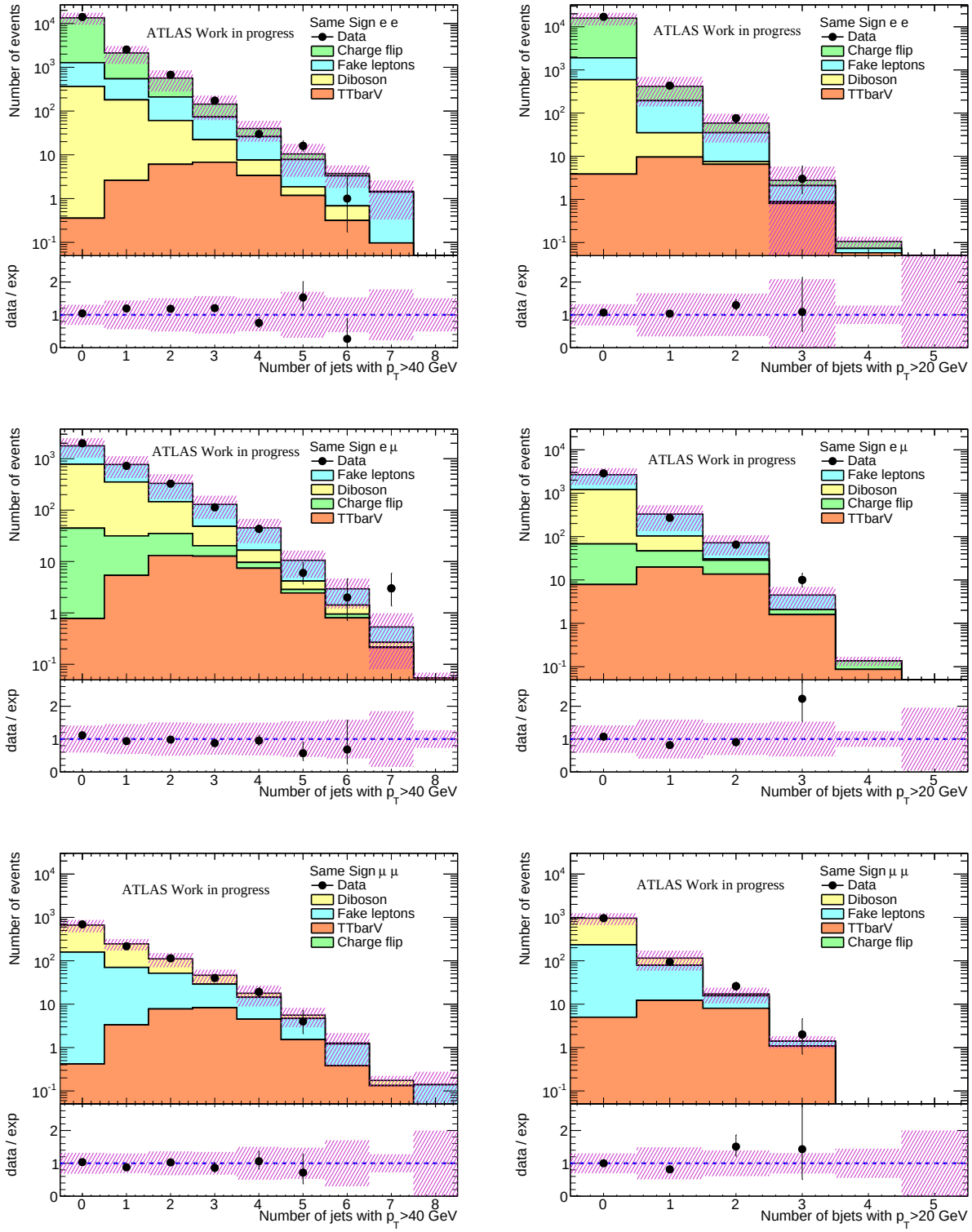


Figure 6.15: Distribution of number of jets with $p_T > 40$ GeV (left) and number of b jets (right) after two same - sign leptons selection in ee (top), $e\mu$ (middle) and $\mu\mu$ (bottom) channels. Hashed area shows the total uncertainty on the background estimation. $L = 21$ fb^{-1} and $\sqrt{s} = 8$ TeV.

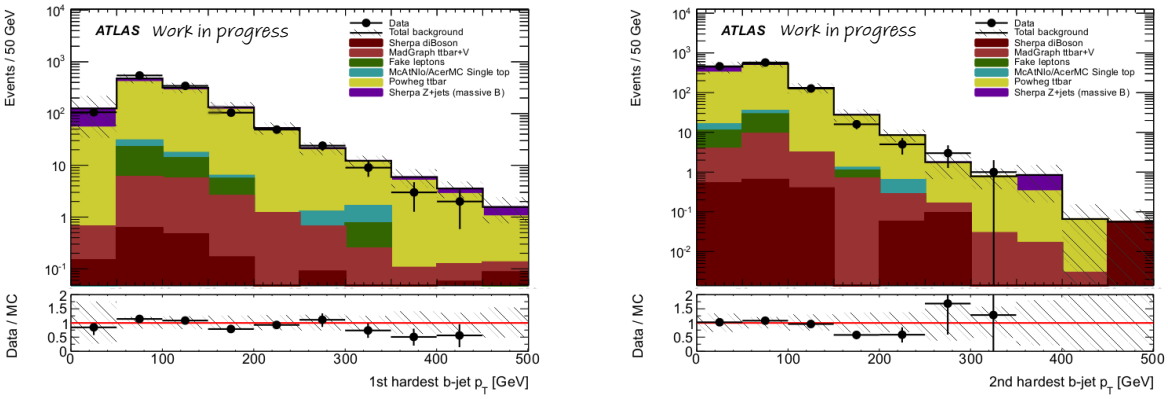


Figure 6.16: Leading (left) and subleading (right) b jet p_T distribution after requiring two opposite - sign signal leptons and three b jets. $L = 21$ fb^{-1} and $\sqrt{s} = 8$ TeV.

6.8.1 Simultaneous fit

In this section the simultaneous fit used to compute the final signal and background uncertainty estimate and to quantify the agreement between the observed data and background estimation in the signal regions is explained. The method is based on a profile-likelihood ratio test and is implemented in the `Histfitter` software tool [264, 265]. The latter uses the `RooStat` [266] and `HistFactory` [267] software packages.

The signal and background components (or samples) are described by binned Probability Density Functions (PDF) defined in statistical independent regions (= signal regions, control regions, validation regions). They are implemented as histograms. The PDF describes the parameter(s) of interest like the rate of a signal process, the normalization factors for the background processes if CRs are used, etc, and the free and nuisance parameters :

- The free parameters : scaling factors which are not constrained, used to adjust the relative contribution of the main background and signal components (μ_{XX}, where XX stands for signal or background).
- The nuisance parameters : scaling factors with external constraints (typically Gaussian) used to model all the statistical and systematical uncertainties. Their correlations are considered. The systematic uncertainty on each PDF can be treated in several ways, as defined in Appendix B.1. Shortly, each systematic uncertainty i is described by a nuisance parameter θ_i, obtained from the up/down variations of one systematic and the nominal histogram.

A discovery and an exclusion fit are used in the analysis to statistically interpret the results in the three signal regions.

- For the discovery fit only one signal region and only one bin are considered. The signal strength μ is changed model-independently in the SRs to obtain the general limits on BSM physics (Section 6.8.4).
- For the exclusion fit a given signal model is included in the fit and the statistical test decide if it is included or not. It is performed using all signal regions and considering several bins. It is used to obtain the exclusion limits (Section 6.8.5).

The likelihood function and its components

The likelihood function describing the analysis is a product of Poisson distributions for the signal region. Optional distributions for CRs and distributions applying constraints on the systematic uncertainties can be also added. Considering the general case when control regions are considered in the analysis, it can be written as :

$$\begin{aligned} L(\mu_{sig}, \vec{b}, \vec{\theta}|\vec{n}) &= P_{SR} \times P_{CR} \times C_{Syst} \\ &= \prod_{i=1}^{N_{SR}} \prod_{j=1}^{N_{Bins}} P(n_{SR}^{ij}|\lambda_{SR}(\mu_{sig}, \vec{b}, \vec{\theta})) \times \prod_{i=1}^{N_{CR}} \prod_{j=1}^{N_{Bins}} P(n_{CR}^{ij}|\lambda_{CR}(\mu_{sig}, \vec{b}, \vec{\theta})) \times C_{Syst}(\vec{\theta}) \end{aligned} \quad (6.10)$$

The first two terms in the product are the Poisson measurements of the number of observed events in the signal region (n_{SR}^{ij}) and the number of events in each control region (n_{CR}^{ij}). The i and j indexes are

accounting for the number of signal (control) regions and the number of considered bins of the discriminant variable. The Poisson expectations λ_{SR} and λ_{CR} can be written as :

$$\lambda_x(\mu_{sig}, \vec{b}, \vec{\theta}) = \mu_{sig} \cdot n^{ij}_{sig}(\vec{\theta}) + n^{ij}_{bkg}(\vec{\theta}, \vec{b}), \quad x \text{ stands for SR or CR} \tag{6.11}$$

where the λ function parameters are :

- μ_{sig} : it is a free parameter used to scale the signal component during the fit. It is also called the signal strength. The condition $\mu_{sig} = 0$ stands for no signal (SM analyses). If $\mu_{sig} = 1$ the signal is considered (BSM analyses) and its expectation is the nominal value of the model under consideration. Only $\mu_{sig} = 1$ is considered.
- $\vec{b} = \sum_i b_i$ are the background normalization factors for the background sources i. They are calculated using dedicated control regions.
- $\vec{\theta} = \{\theta^{(1)}, ..., \theta^{(N_\theta)}\}$: stands for the nuisance parameters used to parametrize the systematic uncertainties like luminosity, JES, JES, etc.

The n^{ij}_{sig} and n^{ij}_{bkg} terms are the numbers of signal events and the number of background events in the i signal (control) region and bin index j. The prediction of signal and background sources are forced to be positive for any values of the nuisance parameters and in all histograms bins. If control regions are not included in the fit the second term is dropped and the $\vec{b}$ parameters are not considered anymore. For this analysis the number of expected background events in the signal region is simply :

$$n^{ij}_{bkg}(\vec{\theta}) = n^{ij}_{bkg} + \theta^{(1)} + \theta^{(2)} + \theta^{(3)} + ... \tag{6.12}$$

The different sources of systematic uncertainties are included using the PDF $C_{Syst}(\vec{\theta})$. It is a product of the probability distributions describing each source of systematic uncertainty. Most of the time the PDF is a Gaussian with a width equal to 1 :

$$C_{Syst}(\vec{\theta}) = \prod_{k=1}^{N_\theta} G(\theta_j) \tag{6.13}$$

where index k runs over the number of the nuisance parameters. The impact of any change in nuisance parameters on the expectation values are described completely by the the λ_x functions.

The profile likelihood ratio test

A profile likelihood ratio (PLR) is used to perform the statistical test for exclusion of signal hypothesis. For one hypothesized signal rate μ_{sig} the PLR is given by the statistical test defined as [264, 268] :

$$q_{\mu_{sig}} = \begin{cases} -2 \ln \dfrac{L(\mu_{sig}, \hat{\hat{\vec{\theta}}}(\mu_{sig})|\vec{n})}{L(\hat{\mu}_{sig}, \hat{\vec{\theta}}|\vec{n})} & \text{if } \hat{\mu}_{sig} < \mu_{sig} \text{ with } \begin{cases} \hat{\hat{\vec{\theta}}}(\mu_{sig}) := \text{argmax}\, L(\mu_{sig}, \vec{\theta}|\vec{n}) \\ \hat{\mu}_{sig}, \hat{\vec{\theta}} := \text{argmax}\, L(\tilde{\mu}_{sig}, \vec{\theta}|\vec{n}) \end{cases} \\ 0 \text{ otherwise} \end{cases} \tag{6.14}$$

- if $\hat{\mu}_{sig} < 0$, $L(\hat{\mu}_{sig}, \vec{\theta}|\vec{n})$ is replaced by $L(0, \hat{\hat{\vec{\theta}}}(0)|\vec{n})$.
- PLR Numerator (tested μ_{sig}) : the likelihood function is maximized with $\hat{\hat{\vec{\theta}}}$ parameter, for a specific fixed value of μ_{sig}.
- PLR Denominator(best fitting μ_{sig}) : the likelihood function is maximized with all parameters free ($\hat{\mu}_{sig}$ and $\hat{\vec{\theta}}$), and basically gives an estimate of the best signal that fits the observed data.

In the PLR, the likelihood function at the denominator is always greater than at the numerator. This is true because at the denominator all the parameters are adjusted to fit at best the data, while at the numerator the signal is fixed. To better understand this, two extreme cases are given : *i*) the data observation is very compatible with the existence of a signal, hence the two likelihood functions are similar and their ratio close to 1; *ii*) the data observation is not compatible with the existence of a signal and the likelihood functions ratio is very small. Therefore, the obtained values for the PLR test are inside the $0 < q_{\mu_{sig}} < \infty$ interval. A very high value is interpreted as a **very improbable signal hypothesis**, while close to 0 as a **signal hypothesis compatible with the data observation**.

The p value

To quantify the level of disagreement between data and a tested hypothesis ($\mu_{sig} = 1$), the p-value (Figure 6.17) is used :

$$p_{\mu_{sig}} = \int_{q^{obs}_{\mu_{sig}}}^{\infty} f(q_{\mu_{sig}}|\mu_{real} = \mu_{sig}) dq_{\mu_{sig}} \quad (6.15)$$

where $q^{obs}_{\mu_{sig}}$ is the value of the statistic test observed in data. The function f denotes the PDF of $q_{\mu_{sig}}$ under the signal hypothesis i.e. the real signal strength is $\mu_{real} = \mu_{sig}$. It can be obtained with pseudo-experiments according to the likelihood function defined for the analysis (the $\vec{n}$ and $\vec{\theta}$ parameters are changed to obtain the dummy data corresponding to the background + signal prediction).

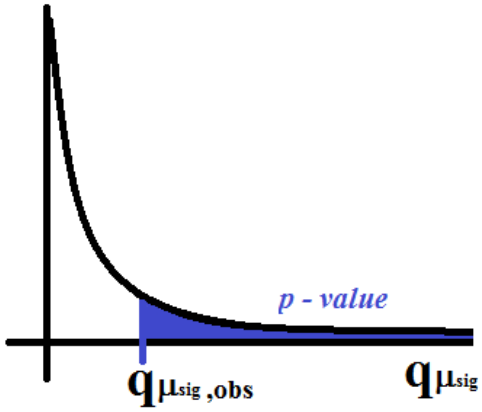


Figure 6.17: Illustration of the p-value computation under the signal hypothesis $\mu_{real} = \mu_{sig}$.

Excluding the signal hypothesis

To exclude a signal hypothesis the probability $P(q_{\mu_{sig}} \geq q^{obs}_{\mu_{sig}}|\mu_{real} = \mu_{sig})$ is computed[10]. If it is small enough, the signal is excluded. Generally, using the CL_{s+b} formalism the signal hypothesis is excluded if the p-value is below a chosen threshold which is set to 5% (95% CL).

This method cannot be used in the case of a small sensitivity to a hypothetical signal, as the signal hypothesis is rejected when it should not be. This happens because with the CL_{s+b} method, the statistical test is performed using a hypothetical signal. Therefore the CL_{s+b} method is replaced by the CL_s method [269]. It is the ratio between CL_{s+b} and CL_b :

$$CL_s = \frac{CL_{s+b}}{CL_b} = \frac{P(q_{\mu_{sig}} \geq q^{obs}_{\mu_{sig}}|\mu_{real} = \mu_{sig})}{P(q_{\mu_{sig}} \geq q^{obs}_{\mu_{sig}}|\mu_{real} = 0)} < 0.05 \quad (6.16)$$

If the data observation is close to the background expectation the CL_s and CL_{s+b} methods are giving approximately the same results. In the extreme case with big fluctuations (i.e 80 observed events for 130 expected events) the CL_{s+b} method results are corrected because the observed data is not compatible with the background only hypothesis either. Therefore, the signal hypothesis is excluded only if $CL_s < 5\%$.

6.8.2 Systematic uncertainties

The systematic uncertainties considered in the fit include all the sources associated to the reconstructed objects and to the SM and detector backgrounds estimation :

- Luminosity : the normalization of simulated background samples is scaled by 2.8% [270].
- Trigger efficiency : it is estimated by varying the correction factors, associated to data - MC differences, randomly within their uncertainties event by event and applying the results as a weight to the simulation. Additional 2% are added to account for the small inefficiencies in the plateau of $\not{E}_T$ triggers.
- Simulation of pile-up : MC events are re-weighted according to the ratio between MC and data distribution [271]. The factor was found to be 1 / (1.09$< \mu >$), where $< \mu >$ is the average number of interactions per crossing. The associated uncertainty is estimated by varying the $< \mu >$ variable by 10%.
- Object reconstruction : all the systematic uncertainties associated to jet energy scale, jet energy resolution, b tagging, missing energy, electron and muon energy scale, resolution and selection efficiencies are described in Chapter 4.
- Standard Model background : the systematic uncertainties described in Section 6.4.
- Data driven background : all the systematic uncertainties associated to data driven charge flip and fake leptons backgrounds are described in Section 6.5 - 6.6.

[10] The p-value is obtained after integrating over all probabilities in ($q^{obs}_{\mu_{sig}}$,∞) interval.

The break down of the systematic uncertainties is presented in Table 6.9 for all three signal regions. The systematic source on the fake lepton background is bringing the highest contribution in all SRs, followed by JES, JER and the theory uncertainty associated to $t\bar{t}$ + V and diboson. In the signal regions with b jets, an important contribution comes from the uncertainty associated to $b$jet tagging procedure.

Systematic uncertainties	SR3b	SR0b	SR1b
Fake-lepton background	$^{\pm 1.36}_{-0.93}$	± 3.08	$^{\pm 1.05}_{-0.28}$
Jet and $\not{E}_T$ scale and resolution	± 0.32	± 0.94	± 0.92
b-jet tagging	± 0.44	± 0.61	± 0.05
Total syst on top + V	± 0.40	± 0.18	± 0.76
Total syst on diboson	< 0.1	± 0.24	± 0.24
luminosity and pile-up	± 0.06	± 0.15	± 0.12
Charge-flip background	± 0.10	± 0.05	± 0.15
Lepton identification	< 0.10	< 0.10	< 0.10
Total uncertainty	± 1.61	± 3.34	± 1.64

Table 6.9: The breakdown of the systematic uncertainties on the expected backgrounds, expressed in units of events. The uncertainty on the fake lepton estimation is including the statistical source. The individual uncertainties are correlated and therefore do not necessarily add up in quadrature to the total systematic uncertainty.

6.8.3 Results in the signal regions

Using the discovery set of cuts mentioned in Table 6.2, the results in all three signal regions are obtained and presented in Table 6.10. The main backgrounds in SR0b are the diboson and the fake leptons and in SR1b and SR3b the top + V processes.

# events	SR3b	SR0b	SR1b
observed	4	5	8
expected background	3.06 ± 1.61	7.46 ± 3.34	3.67 ± 1.64
top+V	1.70 ± 0.83	0.54 ± 0.37	2.23± 0.98
diboson	0.07 ± 0.05	3.37 ± 1.02	0.70 ± 0.44
fake leptons	0.93 ± 0.92 ± 1.00	3.41 ±1.02 ± 2.9	0.28 ± 0.70 ± 0.78
charge flip	0.37 ± 0.05 ± 0.08	0.14 ±0.03 ± 0.04	0.45 ± 0.06 ± 0.14

Table 6.10: Number of events together with the total or statistical and systematic uncertainties in the three signal regions using the *discovery* set of cuts. Top + V category included processed like top + Z, $t\bar{t}$ + V, etc. Numbers correspond to $L = 20.1$ fb^{-1} and $\sqrt{s} = 8$ TeV.

No significant excess is observed for SR0b and SR3b while in SR1b a small excess below 2σ is present. This excess is present only in the channels with electrons : 4 vs. 0.95±0.32 in ee channel and 4 vs. 2.64 ± 1.47 in $e\mu$ channel, and it is considered to be only a fluctuation since detailed investigations exhibit no particular problems. As an illustration, few details about the investigation in ee channel are given. The numbers of tight - tight leptons is 4, loose - loose 1 and events tight - loose 0[11]. As the associated weight to these events is negative, the total fake background estimation is also negative, - 0.81 ± 0.52 (well known behavior of the matrix method in regions with low statistic). Applying a very tight cut on m_{eff} (700 GeV) only very energetic and isolated electrons are passing the selection. After removing some of the baseline electron requirements ($d_0/\sigma(d_0)$) 4 additional tight - loose are passing the SR1b selection and the background is increased to 1 - 2 events instead of the negative estimation. These checks are pointing to inappropriate electron definition, allowing a contamination with fake electrons in the SRs. The origin of the fake electrons can be either high energetic hadrons or a photon conversions. Discussions on this possibility is envisaged in details in the next chapter.

[11] This later category being populated by fakes

6.8.4 General limits on BSM physics

As no significant excess above the Standard Model prediction is observed, upper limits at 95% CL on the number of Beyond Standard Model (BSM) events for each signal region are derived using the CL_s formalism. An upper limit on the visible BSM cross section can be obtained after normalizing the number of observed BSM events to the data sample luminosity. This cross section is defined as the product of acceptance, reconstruction efficiency and production cross section.

	SR3b	SR0b	SR1b
$< \sigma_{vis} >^{95}_{obs}$	0.34	0.33	0.53
S^{95}_{obs}	7.0	6.7	11.0
S^{95}_{exp}	$5.9^{+2.4}_{-1.3}$	$7.9^{+2.6}_{-2.0}$	$6.8^{+2.6}_{-1.5}$

Table 6.11: 95% CL upper limits on visible BSM cross section, on observed and expected number of BSM signal events.

Using pseudo-experiments, the 95% CL upper limits on the observed (S^{95}_{obs}) and expected (S^{95}_{exp}) number of BSM events are set and presented in Table 6.11. The 95% CL upper limit on visible cross section is also shown. To obtain these results, the discovery set of cuts with minimum requirements on m_{eff} (Table 6.2), to keep a high acceptance, are used. Only one m_{eff} bin is retained and no fit is performed as its distribution is unknown for any BSM process. This information is important especially for theorist, as it can be used to probe any BSM model (for few examples see [272]).

6.8.5 Interpretation in SUSY models

The obtained results are interpreted as exclusion limits at 95% confidence level in the context of the phenomenological or simplified models described in Section 6.1. Both expected and observed limits are presented. The former is showing what would be excluded if the observed data correspond exactly to the background prediction, while the latter what can be excluded given the data observation. Note that, a combined simultaneous fit to all signal regions, using a the binned effective mass distribution, is performed. This m_{eff} distribution is shown in Figure 6.18, and the bin intervals are presented in Table 6.12.

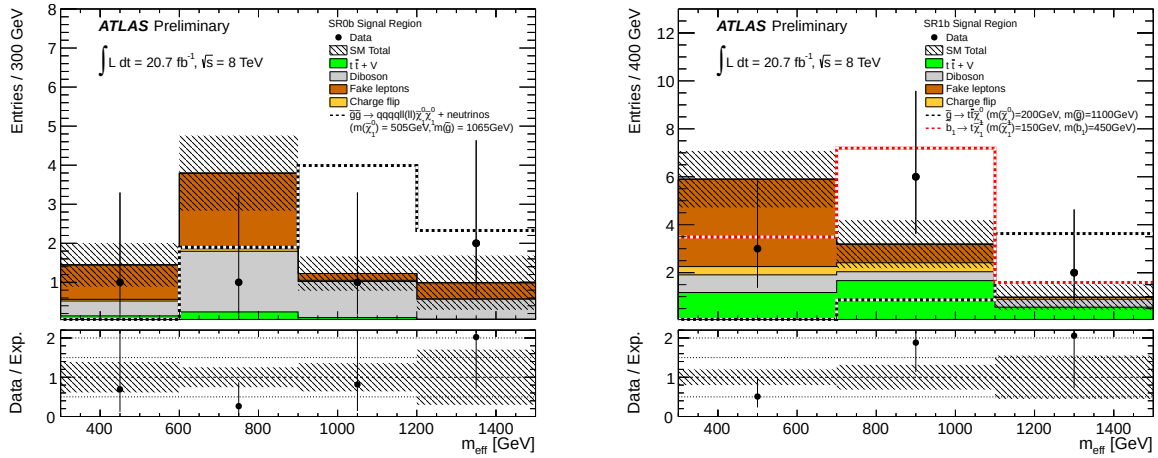


Figure 6.18: Effective mass distribution in SR0b (left) and SR1b (right) signal regions using the exclusion set of cuts. The last bin includes overflow.

SR	bin 1	bin 2	bin 3	bin 4
SR0b	300 - 600	600 - 900	900 - 1200	> 1200
SR1b	300 - 700	700 - 1100	> 1100	

Table 6.12: Intervals in m_{eff} [GeV] used for the exclusion fit.

The first model considered is the search for direct gluino production with the gluino decays as $\tilde{g} \to t\bar{t}\tilde{\chi}^0_1$ via a virtual top squark. The results are shown in the $m_{\tilde{g}}$ - $m_{\tilde{\chi}^0_1}$ mass plane, in Figure 6.19. The observed and exclusion limits are shown with bold red and dashed gray lines. The $\pm 1\sigma^{SUSY}_{theory}$ band around the observed

lines are obtained by varying the SUSY cross section by one standard deviation, $\pm 1\sigma$. All the mass limits on SUSY particles quoted are derived from the $-1\sigma^{SUSY}_{theory}$ line. The yellow band around the expected limit shows the $\pm 1\sigma$ uncertainty and accounts for all sources of statistic and systematic uncertainties, except the theoretical uncertainty on the SUSY cross section. 95% CL upper limits on production cross section are shown in each point of the SUSY phase space and they can serve as limits on cross section on theoretical models having same topology in the final state.

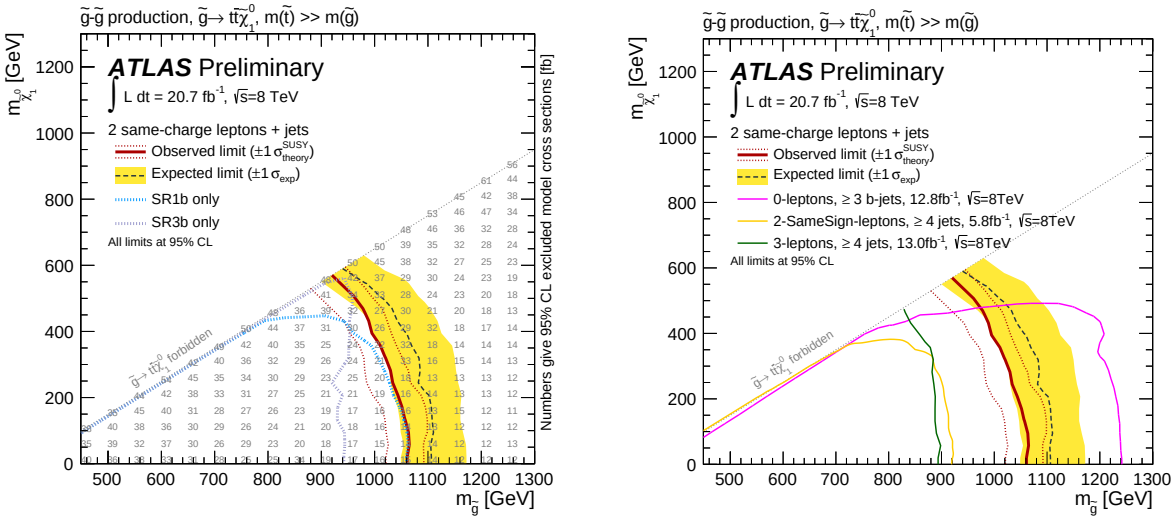


Figure 6.19: Expected and observed limits for gluino pair production via a virtual top squark, $\tilde{g} \to t\bar{t}\tilde{\chi}^0_1$ model. The importance of each signal region is shown in the left plot. A comparison with all ATLAS analyses available at the time of publication is shown on the right side plot. With dark orange are shown the results of the of this search using 5.8 fb^{-1} of 8 TeV data.

In Figure 6.19 left, the importance of each signal region can be appreciated. As expected, SR3b is performing very well in the compressed region with small mass difference between $\tilde{g}$ and $\tilde{\chi}^0_1$, while SR1b is sensitive to regions with large gluino masses and small LSP masses. With the current configuration the analysis excludes gluino masses bellow 900 GeV for LSP masses below 550 GeV. The limit on gluino mass increases up to 1 TeV for massless LSP. To evaluate the improvement brought by this analysis, Figure 6.19 right shows the exclusion limit obtained with the previous version of the analysis using $\sqrt{s} = 8$ TeV and $L = 5.8$ fb^{-1} [226] data is shown (orange line). For masses LSP the improvement is around 100 GeV. When comparing with $\sqrt{s} = 7$ TeV and $L = 2$ fb^{-1} [225] results the improvement is up to 300 GeV for massless LSP. The comparison with other ATLAS analyses targeting this scenario is also done in the same Figure 6.19 right. The highest sensitivity at large gluino masses is obtained by the channel with 0 leptons and 3 b jets [248] which improves the excluded region at high gluino masses, >1 TeV. It can be seen that the channel with two same - sign leptons is more sensitive to compressed spectra.

The second model considered is the direct bottom squarks pairs production, where the bottom squark decays as $\tilde{b} \to t\tilde{\chi}^{\pm}_1$. Only the case with a fixed LSP mass to 60 GeV is shown in Figure 6.20, in the $m_{\tilde{b}}$ - $m_{\tilde{\chi}^{\pm}_1}$ mass plane. Figure 6.20 left shows the importance of SR1b. Bottom squark masses below 470 - 480 GeV are excluded for chargino masses up to 250 GeV. In the right plot a comparison with a three lepton ATLAS analysis [273] using 13 fb^{-1} of 8 TeV data is performed. Hence, the same - sign search is improving the existing ATLAS limits for this model by up to 30 GeV.

The last model considered is the gluino direct production which decays via sleptons. The results are shown in the parameter space of $m_{\tilde{g}}$ - $m_{\tilde{\chi}^0_1}$ masses in Figure 6.21. The left plots shows the sensitivity reached with each signal region. The limits are driven by SR0b. Gluino masses below 1 - 1.1 TeV are excluded for LSP masses up to 650 GeV. Comparing with other ATLAS analyses [251], a significant improvement in all the phase space is visible.

6.9 Conclusions

The analysis searching for new physics in events with a pair of same - sign leptons, (b) jets and $\not{E}_T$ using 20.3 fb^{-1} of 8 TeV data is a very powerful analysis which covers a broad spectrum of SUSY models. In the two benchmark models including third generation squarks up to 40 GeV improvements are observed in the mass limit of gluinos and sbottoms, tightening the existing ATLAS limits. For the third model with gluinos decays via sleptons the obtained limits are well above the ones obtained with other channels.

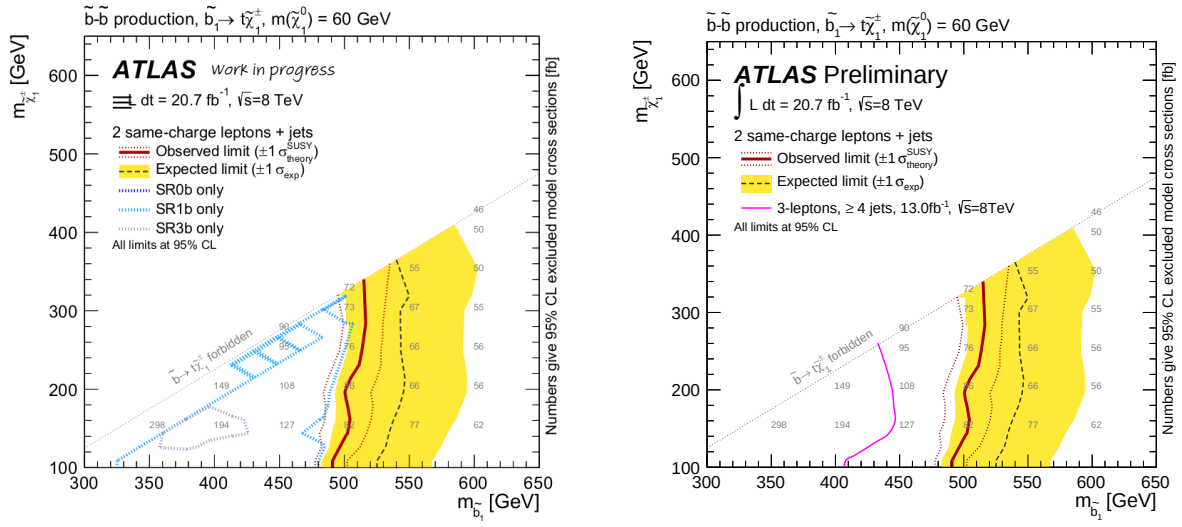


Figure 6.20: Expected and observed limits for direct bottom squarks pairs production, $\tilde{b} \rightarrow t\tilde{\chi}_1^{\pm}$ model. The importance of each signal region is shown in the left plot. A comparison with all ATLAS analyses available at the time of publication is shown on the right side plot.

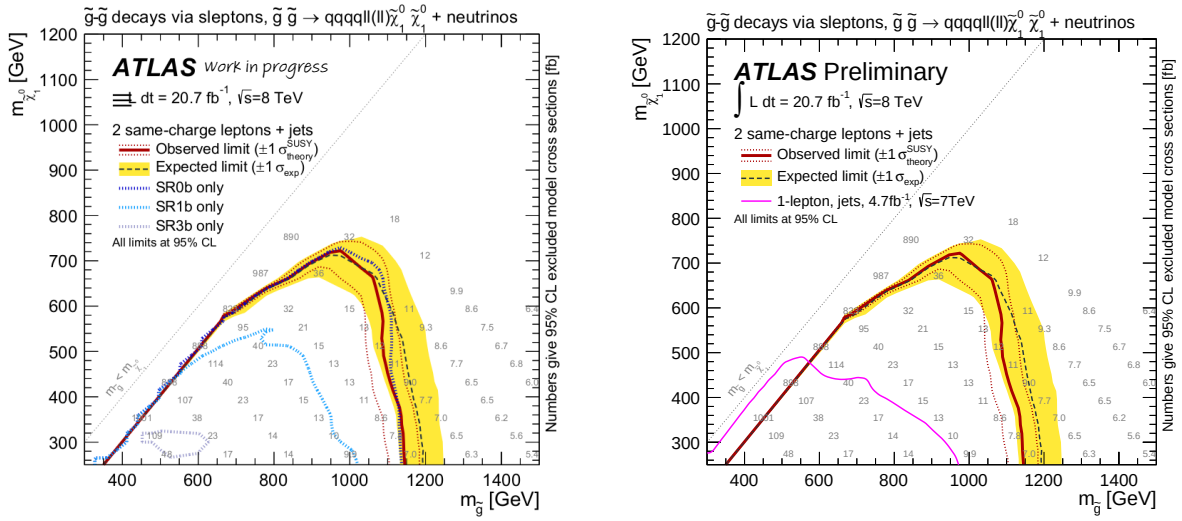


Figure 6.21: Expected and observed limits for a SUSY model where gluinos decay via sleptons, $\tilde{g} \rightarrow qql(ll)\tilde{\chi}_1^0$. The importance of each signal region is shown in the left plot. A comparison with all ATLAS analyses available at the time of publication is shown on the right side plot.

Similar conclusions are valid as well for the other models presented in Section 6.1, setting the most stringent constraints to date from ATLAS at the moment of release of the results.

These improvements are the results of the signal regions optimization using a machinery able to scan around 58000 configurations, assuring a well coverage of the SUSY phase space. Adopting SR defined as function of b jet multiplicity allowed an impressive extension of the targeted SUSY scenarios and therefore a better usage of the same - sign signature. At the background level, important enhancements are done for a better estimation of the fake lepton background : the fake lepton efficiencies measurements takes into account the p_T, η and b jet dependency, compared to only one overall factor with no dependency considered in the previous analysis. Similarly the charge flip estimation is fully data driven, while before OS MC events were used for the event re-weighting. To assure a better validation of SM background several VRs are added giving more confidence in the preliminary results. These changes lead to promising results, even if no significant excess over the SM predictions was observed, and also to a better understanding of the SM and detector background.

For the final results, to gain in precision and sensitivity several improvements are considered. One of the most important being the combination of SS and 3 leptons searches and the signal lepton isolation optimization. At the background level, several methods are added or improved, as detailed in the next chapter.

Chapter 7

Combination and improvement of searches with two same - sign or three leptons

To obtain the preliminary results with 8 TeV data, presented in Chapter 6, all signal regions [225, 226] were re-optimized. This effort maximizes the sensitivity of two same - sign lepton search in several regions of the phase space where SUSY or other new physics could hide. To improve even more, an extension of the same - sign inclusive channel is performed by including explicitly the third lepton in the event selection and re-optimizing the signal regions definitions. Other important changes or improvements are the usage of softer leptons, optimization of signal lepton isolation variables, new methods to cross - check the background estimates and signal models to interpret the results. All these changes with respect to Chapter 6 are implemented in the final publication of 8 TeV results [224] and are detailed in this Chapter.

This chapter is organized as follows. The new considered signal models are presented in Section 7.1. In Section 7.2 the signal region re-optimization is detailed and the final definitions are discussed. The changes in object definition and event selection are enumerated in Section 7.3. The methods used to estimate and cross check the SM and detector background estimation and validation are presented in Section 7.4. The results in the signal regions, general limits on BSM physics and the model dependent exclusion limits are shown in Section 7.5. In Section 7.6 a short comparison with the results obtained in the two leptons same - signal channel are presented and finally the chapter is closed with Section 7.7 where some general conclusions are enunciated.

7.1 Additional SUSY models

Besides the models presented in Section 6.1, a two step decay simplified model, **gluino mediated and direct production of 1^{st} and 2^{nd} generation squarks decaying to** $WZWZ$, is considered to optimize the signal regions definitions. The allowed decays are $\tilde{g} \to q\bar{q}\tilde{\chi}_1^{\pm}$ (or $\tilde{q} \to q\tilde{\chi}_1^{\pm}$) with $\tilde{\chi}_1^{\pm} \to W\tilde{\chi}_2^0$ and $\tilde{\chi}_2^0 \to Z\tilde{\chi}_1^0$ (Figure 7.1). In the final state are from two to four light jets from $\tilde{q}$ and $\tilde{g}$, four W and Z bosons and no b jets. The sparticles masses are set such that $\tilde{\chi}_1^{\pm}$ sits exactly in the middle between $\tilde{g}$ ($\tilde{q}$) and LSP, and such that $\tilde{\chi}_1^{\pm}$ sits exactly in the middle between the $\tilde{\chi}_2^0$ and LSP[1].

For illustration purposes, three other *real* SUSY models are added to interpret the results in the signal regions, as the same - sign plus three lepton channel is found to be competitive with other ATLAS searches. A first model is **gauge mediated SUSY breaking, GMSB**, where SUSY is broken through the SM gauge sector. This model is characterized by six parameters : the SUSY breaking mass scale (Λ), the messenger mass (M_{mess}), the number of SU(5) messenger fields (N_5), the ratio of the vacuum expectation values of the two Higgs doublets ($\tan\beta$), the sign of Higgs sector mixing parameter μ and the scale factor for the gravitino mass (C_{grav}). The free parameters are Λ and $\tan\beta$, and the other parameters are set such that the main production processes involves only gluinos and squarks ($M_{mess} = 250$ TeV, $N_5 = 3$, $\mu > 0$ and $C_{grav} = 1$). In the cascade decays, the next-to-lightest-SUSY particle (NLSP) is either of $\tilde{\chi}_1^0$, $\tilde{\tau}$, $\tilde{e}_{\mathrm{R}}$ or $\tilde{\mu}_{\mathrm{R}}$, depending on the mass hierarchy. The LSP is the gravitino with a mass of order keV.

A more exotic model with R parity violated, **bRPV**, is considered to show the generality of this study. The R parity violating couplings are embedded in the general mSUGRA/CMSSM model (Section 6.1).

[1] If the chargino masses are reverted, the final state would be characterized by four W bosons. This case is not considered yet.

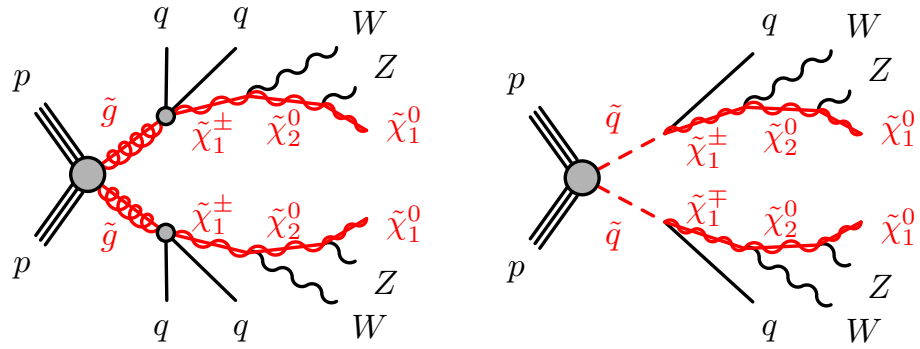


Figure 7.1: Feynman diagrams representative for the gluino mediated and direct production of 1^{st} and 2^{nd} generation squarks decaying to $WZWZ$ model.

The LSP is the lightest neutralino and its decay includes neutrinos. Results are expressed as function of m_0 and $m_{1/2}$.

A non-SUSY exotic model is **mUED**. It is the minimal extension of the SM with one additional universal spatial dimension. The KK quark decays to the lightest KK particle (KK photon) giving a signature very similar to SUSY particles cascade decay of a squark to LSP. The properties of the model depend on two parameters, the compactification radius R and the cut-off scale Λ.

7.2 Signal regions re-optimization

For the re-optimization the starting points are the signal regions definitions presented in Chapter 6 and the three leptons plus multijets analysis [273]. For a complete scan, three different approaches are considered and compared :

- SS leptons selection with no cut on third lepton, as in previous chapter : already defined SR0b, SR1b and SR3b are used as baseline to quantify the improvement of including a third lepton in the selection.
- Inclusive SS leptons selection : the same - sign leptons selection presented in Chapter 6 remains unchanged and requirements on the third lepton are included. The event is retained if the SS leptons selection is failed, but there is a third lepton passing the signal region cut[2].
- Combined SS leptons exclusive and three leptons selection : two exclusive signal regions, using only two same - sign leptons and only three leptons, are combined as the background contamination can be significantly reduced.

Background estimation : when only events with same - sign leptons are retained, same techniques as in Sections 6.4, 6.5 and 6.6 are employed. When the third lepton is included, no charge flip evaluation is needed and the fake lepton background is estimated using a matrix method applied on second and third lepton[3]. The SM background is estimated using MC samples.

Signal significance : during the optimization, the signal sensitivity is evaluated through its potential discovery significance :

$$Z_0 = \sqrt{2((s+b)\ln(1+s/b) - s)} \tag{7.1}$$

where s and b represent the expected number of signal and background events. This formula [274] derives from the logarithm of the likelihood ratio testing a 0 signal hypothesis using Poisson distributions. If s << b the significance is simply $\frac{s}{\sqrt{b}}$.

This ideal significance must be corrected to account for uncertainties on the background prediction. To do so, the following substitution was applied $b \to b + \Delta b$, where Δb represents the total uncertainty on the background. The later gives a more realistic description and denotes the real significance or Z_r. During the optimization, Δb is fixed to 40% as this optimistic approach is considered to be feasible. To combine

[2]No cut on the charge is considered for the three leptons selection, as there are at least two leptons with same electric charge within the three leptons retained in the event.

[3]This approach assumes that the leading lepton is always real and the fake lepton is either the subleading or thirdleading. Cases with two fake leptons or leading fake and subleading or thirdleading real are very rare, as shown in studies based on truth MC.

different signal regions or different channels, the combined significance Z_c is calculated as $Z_r^{SR_i} \oplus Z_r^{SR_j}$, where $SR_{i,j}$ is the index of the region (ie. inclusive SR1b or SR3b, exclusive three leptons signal region).

The optimization procedure is presented in Section 7.2.1, while the final signal regions definitions are defined in Section 7.2.2. This section is closed with an illustration of the improvements reached by the combination of the same - sign leptons and three leptons analyses.

7.2.1 Optimization of kinematic variable cuts

Before any optimization, the gain of combining the two same - sign and three leptons signatures is studied for a large number of simplified SUSY models with gluinos decays to 3^{rd} and $1^{st}/2^{nd}$ generation squarks, direct sbottom and $1^{st}/2^{nd}$ generation squars production (Section 6.1). For this study, already defined signal regions SR0b, SR1b, SR3b (Section 6.3.2) and SR3Lep (Tables 7.1) are used.

Signal region	Signal cuts
SR3Lep	$N_{jets} \geq 4$, $!(81 < M_{ll} < 101)$ GeV, $\not{E}_T > 50$ GeV, lepton $p_T > [20,15,15]$ GeV

Table 7.1: Three leptons signal region used as starting point and defined in Ref. [273]

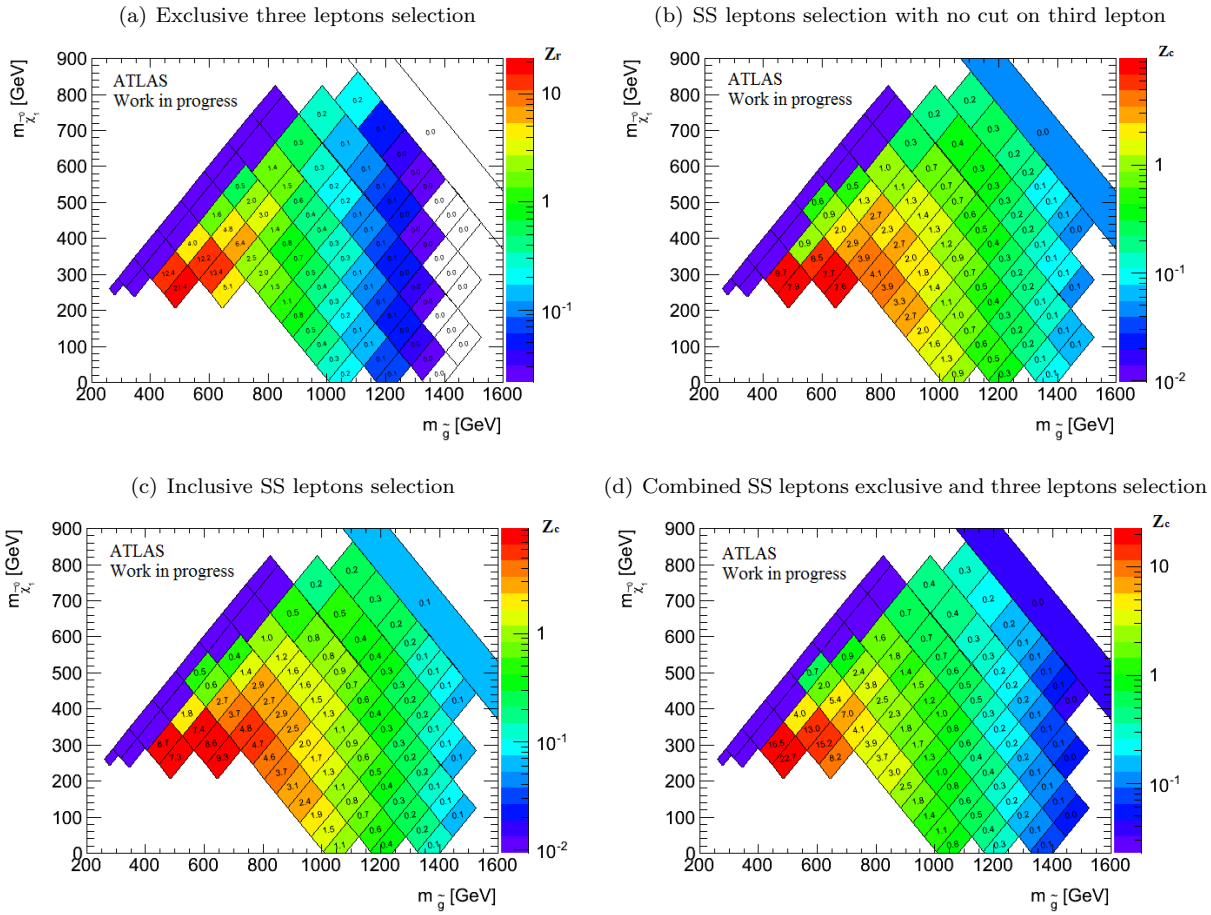


Figure 7.2: Example of sensitivity gain in $\tilde{g}\tilde{g}$ production in 2-step decay simplified model after combining the SS and 3 leptons analyses ($\sqrt{s} = 8$ TeV, $L = 20.3$ fb^{-1}). On the z axis the real or combined significance is shown.

An example of the procedure is shown in Figure 7.2 for the gluino mediated 1^{st} and 2^{nd} generation squarks decaying to $WZWZ$ model. It shows the combined significance for the exclusive three leptons selection (SR3Lep, Figure 7.2 a), for the SS leptons selection with no cut on third lepton (SR0b, SR1b and SR3b, Figure 7.2 b), for the inclusive SS leptons selection (SR0b, SR1b and SR3b, Figure 7.2 c) and the combined SS leptons exclusive and three leptons selection (SR0b, SR1b, SR3b and SR3Lep, Figure 7.2 d). Generally, these studies show the importance of a combined search and the motivation of the re-optimization of the existing signal regions to assure an entire coverage of the SS and 3 leptons phase space.

During the optimization, several kinematic variables motivated by the SUSY signature are used to discriminate the background over the signal. In Figures 7.3 - 7.5 cumulative distributions of these variables

are shown for few points extracted from three SUSY scenarios together with all SM and detector background sources, after two same - sign or three signal leptons selection. From the gluino mediated virtual stop Gtt model several points are considered : $m_{\tilde{g}}$ = 1 TeV, $m_{\tilde{t}}$ = 2.5 TeV and m_{LSP} = [300, 350,..., 550] GeV. The second model is gluino two step decay via chargino. The considered points are $m_{\tilde{g}}$ = [865, 905,..., 985] GeV, m_{LSP} = [545, 505,..., 425] GeV and $m_{\tilde{\chi}_1^\pm}$ = 705 GeV. The last model is direct sbottom pair production. The considered points are $m_{\tilde{b}}$ = 450 GeV, m_{LSP} = 60 GeV and $m_{\tilde{\chi}_1^\pm}$ = [100, 150,.., 250] GeV. The following variables are considered for the optimization :

- Missing transverse energy. Requiring a $\not{E}_T > 150$ GeV cut rejects most of charge flip electrons, fake leptons and diboson backgrounds and retains ∼15% of $t\bar{t} + V$ and 70% - 35% of the considered SUSY signal (Figure 7.3). The $\not{E}_T$ cut is varied with steps of 50 GeV. The optimization is performed separately in ee, $e\mu$ and $\mu\mu$ channels, since the muon channels are less affected by the detector backgrounds.

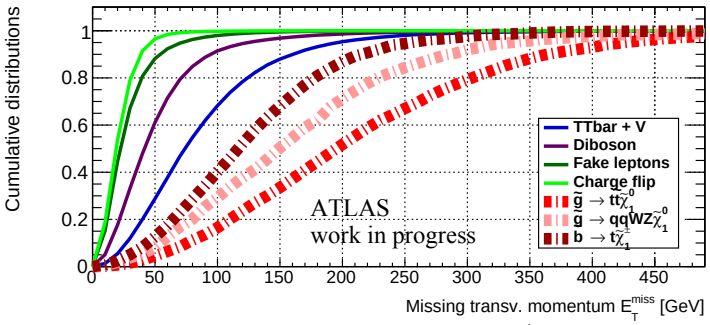


Figure 7.3: Cumulative distribution showing the importance of $\not{E}_T$ discriminant variable. Distributions shown after requiring two same - sign or three signal leptons in the event ($\sqrt{s}$ = 8 TeV, L = 20.3 fb^{-1}).

- Transversal mass. Figure 7.4 left shows that when adding $m_T > 100$ GeV, around 5 - 10% detector background is retained together with 40 - 50% SM background and 60 - 80% signal. For the optimization, the m_T cut is varied in steps of 25 GeV (even removed in some channels).

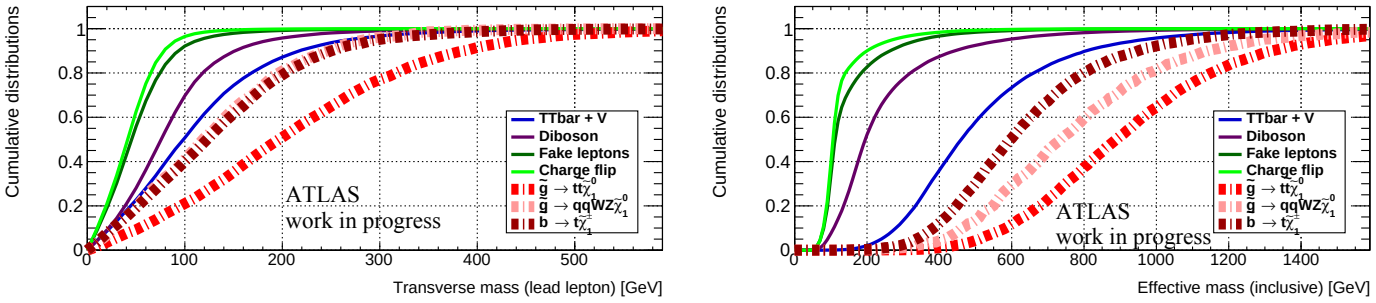


Figure 7.4: Cumulative distribution showing the importance of m_T (left) and m_{eff} (right) discriminant variable. Distributions shown after requiring two same - sign or three signal leptons in the event ($\sqrt{s}$ = 8 TeV, L = 20.3 fb^{-1}).

- Effective mass. Two different directions are followed : for the discovery case a minimum cut is applied to reject (mainly) backgrounds from fake leptons, while for exclusion a shape fit in several m_{eff} intervals is performed. The approach used for the exclusion case is having the same discriminant capability as considering a different cut for each sparticle mass combination. In Figure 7.4 right cumulative distributions of m_{eff} variable are shown for the background and considered SUSY signal. A cut at 400 GeV is highly reducing the detector and diboson backgrounds, and retains 60% of $t\bar{t} + V$ background and almost all signal. To further reduce the $t\bar{t} + V$ component tighter thresholds are well motivated. For the m_{eff} cut threshold optimization, several trials steps of 50 GeV are considered.

- Number of jets. As the signal is characterized by high jet multiplicity from production of squarks ($\tilde{q} \to q'\tilde{\chi}$), gluinos ($\tilde{g} \to q\bar{q}'$) and electroweak or top quark decays produced in cascade, this variable is reducing mainly the detector background and diboson processes, as shown in Figure 7.5 left. A requirement of $N_{jets} \geq 3$ removes almost all detector and diboson backgrounds and will retain ∼60% of $t\bar{t} + V$ processes and 90 - 80% of SUSY signal. For the optimization the jet p_T threshold and

number of jets multiplicity are relaxed or tighten to further reduce the $t\bar{t} + V$ background or to recover some of loss signal.

- Number of b jets. The importance of defining signal regions depending on b jet multiplicity was already discussed in Chapter 6. Figure 7.5 right shows the importance of an inclusive signal region with respect to number of b jets in the event for models with multilepton signature. As many other models are characterized by b jets in the event it make sense to consider during the optimization also signal regions with 3 leptons and b jets to gain sensitivity. Because some scenarios are characterized by two b jets and > 2 leptons, several combinations are considered for a signal region with exactly two b jets. Due to the high contamination in $t\bar{t} + V$ processes the gain in significance is not major and no SR2b signal region is defined.

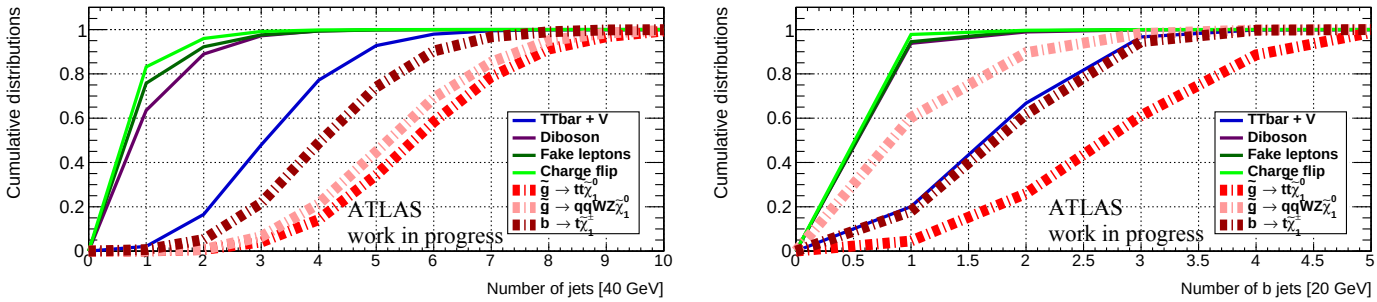


Figure 7.5: Cumulative distribution showing the importance of a cut on number of jets (left) and b jets (right). Distributions shown after requiring two same - sign or three signal leptons in the event ($\sqrt{s} =$ 8 TeV, $L = 20.3$ fb^{-1}).

- Lepton and jet p_T threshold : to increase the sensitivity in regions with small mass difference between sparticles, the lepton p_T thresholds are re-optimized. Only leptons with $p_T > 12$ GeV are considered, limited by the trigger threshold. An optimization depending on the leptons flavour is considered as well.

- Jet p_T threshold optimization is important for the signal regions including a third lepton. For the same - sign lepton selection requiring few energetic jets the sensitivity is not significantly increased and it is not considered during the re-optimization.

- Other discriminant variables are considered for the regions with 3 leptons, ie. a veto on the Z boson invariant mass. Their importance and description is mentioned in the following.

For the re-optimization, several simplified and physics models are used and the signal regions[4] are combined differently to reach the maximum sensitivity :

- An inclusive versus exclusive SR1b and SR3b are studied for gluino mediated top squark or direct bottom squark pair production, as it leads to many b jets and W bosons in the cascade decays.

- An inclusive versus exclusive SR0b is refined for gluino mediated or direct squark production via gauginos or sleptons, as no b jets are expected but many multileptons from sleptons or W or Z bosons.

- Exclusive signal regions are combined with SR3Lep where no explicit cut on fourth lepton is added.

Using the same approach as for the previous optimization (Section 6.3.2), but with a more basic setup, an important number of combinations are tested. The choice of the trial signal region is performed by looking at the extension of the excluded region. In Appendix B.2, Tables B.1-B.4 give the best or most motivated SRs trials considered for the final choice of the SS and 3 leptons combination, as well as their index numbers. As an example, Figure 7.6 shows for two simplified models, gluino two steps decay via sleptons $\tilde{g} \to qq\tilde{\chi}_1^{\pm} \to qq + \tilde{l}\nu/\tilde{\nu}l$ and gluino mediated virtual top squark $\tilde{g} \to t\bar{t}\tilde{\chi}_1^0$, the signal region index number which gives the best signal significance.

For gluino mediated virtual top squark Gtt model, a large number of signal regions are selected. To retain a reasonable number of regions, the significance obtained for the best combination in each point is compared to the most favored signal region to assure a minimal loss in sensitivity. A comparison example

[4]Note that, if not specified otherwise SR0b, SR1b, SR3b and SR3Lep are the signal regions considered during the optimization.

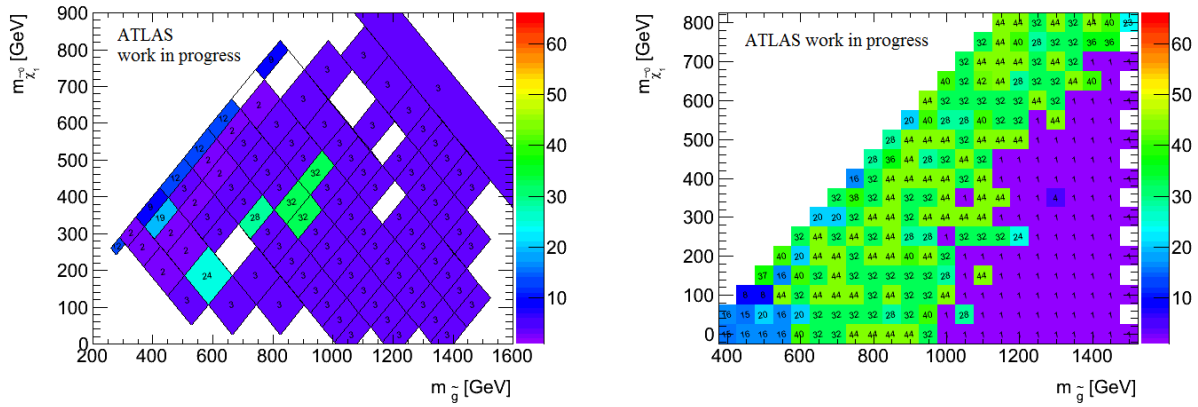


Figure 7.6: Best signal region index number defined in Appendix B.2 (z axis) giving the best significance for 2 typical SUSY models, gluino 2 steps decay via sleptons $\tilde{g} \to qq\tilde{\chi}_1^{\pm} \to qq + \tilde{l}\nu/\tilde{\nu}l$ (left) and gluino mediated virtual top squark $\tilde{g} \to t\bar{t}\tilde{\chi}_1^0$ (right). The empty boxes are just missing points due to technical problems.

is illustrated in Figure 7.7. The chosen combination, here index number 3 (see Table B.1 of Appendix B.2), is only mildly reducing the significance associated to the best SR combination. Similarly for other models, the final signal region is a compromise between an optimal coverage of the SUSY phase space and a reasonable number of signal regions. The later is well motivated as this search is aiming for a minimum model dependency.

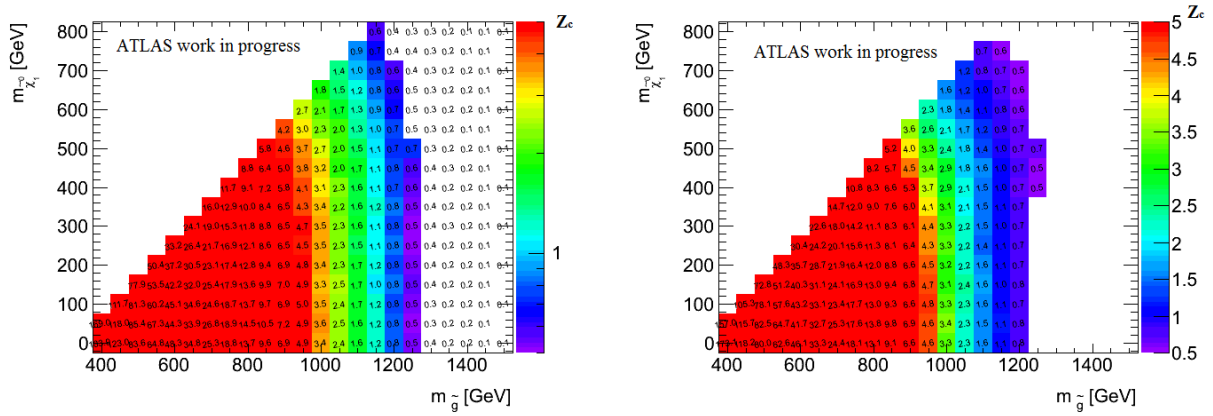


Figure 7.7: Considering the gluino mediated virtual top squark $\tilde{g} \to t\bar{t}\tilde{\chi}_1^0$ model, the signal regions combination giving the best significance in each point (left) and the significance obtained with the signal regions combination corresponding to index 3 (right) are shown. On the z axis the combined significance, Z_c is showed ($\sqrt{s} = 8$ TeV, $L = 20.3$ fb^{-1}).

7.2.2 Signal regions definition

As a result of the optimization five signal regions are chosen and presented in Table 7.2 :

- SR3b : it is an inclusive same - sign leptons selection with at least 5 signal jets with $p_T > 40$ GeV. The three b jets must pass the 70% b tagging efficiency and have $p_T > 20$ GeV.
- SR0b : events are required to have at least three signal jets with $p_T > 40$ GeV, no b jet with $p_T > 20$ GeV and a veto on third lepton with $p_T > 15$ GeV. High cuts on $\not{E}_T$ and m_T reduce the total background. The veto on b jets mainly reduce the $t\bar{t} + V$ background.
- SR1b : events should have at least three signal jets with $p_T > 40$ GeV, at least 1 signal b jet and exactly two same - sign leptons. The same cuts on $\not{E}_T$ and m_T as SR0b are applied. To make it orthogonal to SR3b, the events selected in this region are discarded.
- SR3Lep_high$\not{E}_T$: events must have three signal leptons, at least four signal jets and large $\not{E}_T$. The later cut rejects the Z + jets background. No cut on the number of b jets is applied. The main advantage over the SS lepton signal regions is the reduced detector background.

- SR3Lep_low$\not{E}_T$: this signal region has a similar motivation as the previous one and aims particularly at the compressed spectra. Events are required to have at least three signal leptons, at least four signal jets with $p_T > 40$ GeV and low $\not{E}_T$. A veto on the Z boson invariant mass is added to avoid the overlap with an existing ATLAS analysis searching for direct stop pair production in events with a Z boson, b-jets and missing transverse energy [250].

To maximize the sensitivity of SR3b, this signal region is prioritized over SR1b and SR3Lep_high$\not{E}_T$: any event with two same - sign or three leptons and at least three b-jets is counted in SR3b and vetoed in SR1b and SR3Lep_high$\not{E}_T$ (!SR3b in Table 7.2). After the signal regions are defined, a new optimization of the m_{eff} cut is performed for the discovery fit. The final value is chosen after the evaluation of trial signal regions sensitivity on several signal points, using similar techniques as for the other kinematic variables optimization.

Signal region	$N_{leptons}$	N_{b-jets}	Signal cuts (exclusion)	Additional cuts (discovery)
SR3b	2 SS or ≥ 3	$\geq$3	$N_{jets} \geq 5$	m_{eff} >350 GeV
SR1b	exactly 2 SS	$\geq$1	$N_{jets} \geq 3$, $\not{E}_T > 150$ GeV, m_T >100 GeV, !SR3b	m_{eff} >700 GeV
SR0b	exactly 2 SS	exactly 0	$N_{jets} \geq 3$, $\not{E}_T > 150$ GeV, $m_T > 100$ GeV, !SR3b	m_{eff} >400 GeV
SR3Lep low$\not{E}_T$	≥ 3	-	$N_{jets} \geq 4$, $50 < \not{E}_T < 150$ GeV, Z m_{invm} veto, !SR3b	m_{eff} >400 GeV
SR3Lep high$\not{E}_T$	≥ 3	-	$N_{jets} \geq 4$, $\not{E}_T > 150$ GeV, !SR3b	m_{eff} >400 GeV

Table 7.2: Definition of the signal regions for the two same - sign and three leptons analysis. The cuts for the discovery and exclusion cases are shown separately. For all signal regions, two same - sign or three leptons with $p_T >$ [20,15,15] GeV are required. (b) Jets are selected with $p_T > 40$ GeV (20 GeV).

To quantify the improvement brought by two same - sign and three leptons combination, Figure 7.8 shows an example of the significance ratio between the signal regions defined in Table 7.2 and the ones used in Chapter 6. Two models yielding no b jets and more than 3 leptons in the cascade decay are considered. In the left plot, results are shown for gluino mediated squark production via charginos, $\tilde{g} \to qq\tilde{\chi}_1^{\pm} \to qqW\tilde{\chi}_2^0 \to qqWZ\tilde{\chi}_1^0$, and on the right one, direct squark production via charginos, $\tilde{q} \to q\tilde{\chi}_1^{\pm} \to qW\tilde{\chi}_2^0 \to qWZ\tilde{\chi}_1^0$. These plots show an increase in significance from $\sim$2 to $\sim$4 depending on the phase space region. More generally this is true for all the models with no b jets. Note that for gluino mediated models via top squarks the impact of adding three leptons signal region is not very significant, as the limit is driven by SR3b. However for the models with one or 2 b jets the improvement is significant as illustrated in Figure 7.9 for direct bottom squark production model.

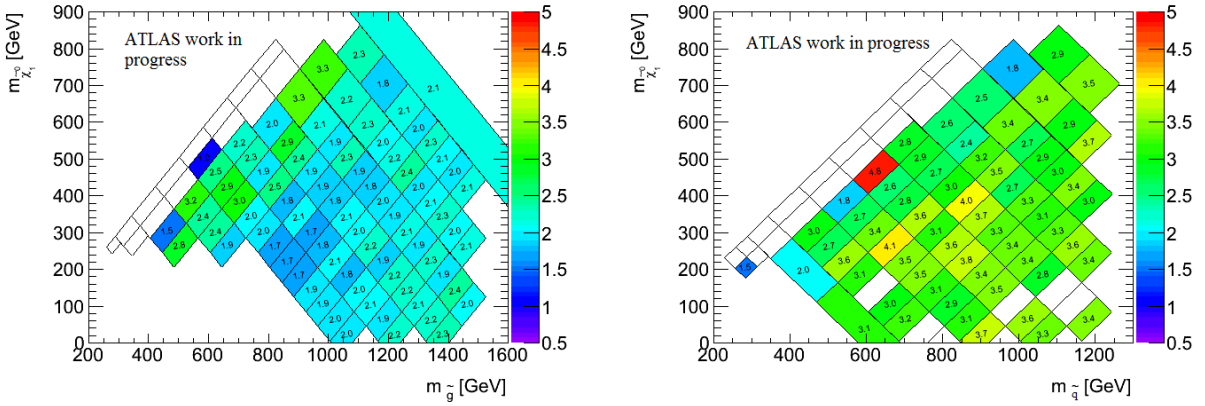


Figure 7.8: Example of the significance ratio between the final signal regions and the ones defined in Chapter 6 for the two models explained in the text ($\sqrt{s} = 8$ TeV, $L = 20.3$ fb^{-1}). In the compressed regions, this search has no sensitivity. The empty boxes at low LSP masses (< 200 GeV) are missing points due to technical problems.

7.3 Changes in object definition and event selection

Together with the signal regions re-optimization, few changes are done in the object definitions and event selection compared to the ones already described in Chapter 6. First, a logical OR between lepton and $\not{E}_T$ triggers is used to increase the lepton efficiency. Also, the $\not{E}_T$ definition without the additional corrections to suppress pileup is now used to harmonize with the other SUSY analyses. The overlap removal between pre-selected electrons and muons is also improved. If $\Delta R_{e\mu} < 0.1$, the electron is removed from the event

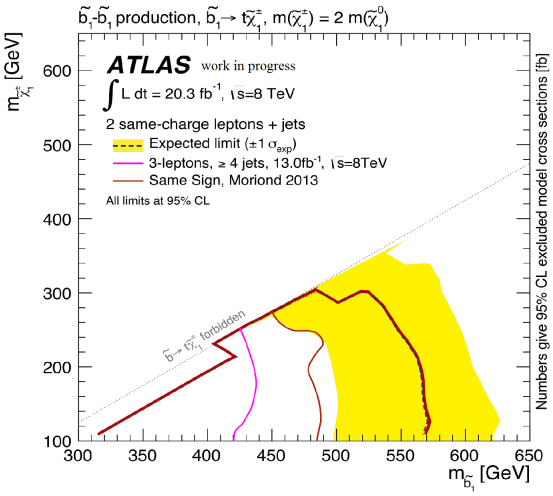


Figure 7.9: Expected exclusion limit for direct bottom squark production model after the signal regions re-optimization ($\sqrt{s} = 8$ TeV, $L = 20.3$ fb^{-1}). Previous limits (Moriond 2013) are shown with dark orange line.

instead of discarding the event. Finally, softer leptons are used to increase the sensitivity in the compressed regions and the subleading and subsubleading leptons p_T thresholds are decreased to 15 GeV. Note also that, if the event is a three-lepton event all possible opposite-charge same-flavor lepton combinations are used to calculate the M$_{ll}$ mass. In SR3Lep_low$\not{E}_T$ the event is rejected if one of the calculated mass is close to the Z-mass, i.e $84 < M_{ll} < 98$ GeV.

Finally, to further reduce the fake lepton background, new lepton isolations are also reconsidered. Lepton calorimeter and track isolations (*cone*20 and *cone*30) are scanned separately and combined using [0.04, 0.06, 0.08, ..., 0.20] cut values interval as detailed in Appendix B.2. The conclusions of this study are that tighter relative isolation definitions for electrons are needed to improve the signal significance ($p_T^{\text{cone20}}/p_T < 0.06$ and $topoE_T^{\text{cone20}}/p_T < 0.06$). For muons, same definitions as in Chapter 6 should be retained because no significant gain can be reached with a different isolation. These definitions are valid only for leptons with $p_T < 60$ GeV. Above this threshold, to further reduce the increase of the fake lepton efficiency, the relative isolation is replaced by a fixed cut on the absolute value of the $p_T^{\text{cone20,30}}$ and $topoE_T^{\text{cone20,30}}$ isolation variables : 3.6 GeV (0.6×60 GeV) for electrons and 7.2 GeV (0.12×60 GeV) for muons. The impact of this requirement is found to be negligible on the SUSY signal, while the fake rate is found to decrease by ∼10% in this range.

7.4 Background estimation and validation

Because of the addition of three leptons signal regions, several developments are needed to improve the background estimation techniques. To increase the precision on the estimation, the charge flip background method is changed (Section 7.4.1). For the fake lepton background, the method described in Section 6.6.1 is expanded to the number of loose leptons in the event (Section 7.4.2). Several methods are developed to verify the data driven methods and the detector background in SR3b, a particular critical item for Chapter 6 (Section 7.4.3).

7.4.1 Electron charge flip background

To improve the electron charge flip rate measurement a new method is developed, making use of all available statistics. Similarly to Section 6.5, this rate is obtained in several η and p_T bins using both same - sign and opposite - sign electron pairs without making any assumption on the two leptons. Because the Z boson invariant mass distribution is shifted and smeared when using same - sign electrons, an asymmetric interval is now used ($75 < M_{ee} <$100 GeV).

Making the assumption that no correlation exists between the different η and p_T bins, the number of measured same - sign events where one electron is in the (i, k) η and p_T bin and the other electron in the (j, l) bin can be expressed as :

$$N_{SS}^{ik,jl} = N_{OS}^{ik,jl}(\zeta_{ik} + \zeta_{jl}), \tag{7.2}$$

where $N_{OS}^{ik,jl}$ is the total number of events in the (i,k) and (j,l) bin, ζ is the electron charge flip rate. The number of SS events in a given bin is described by a Poisson distribution:

$$f(N_{SS}^{ik,jl}; N_{OS}^{ik,jl}(\zeta_{ik}+\zeta_{jl})) = \frac{[N_{OS}^{ik,jl}(\zeta_{ik}+\zeta_{jl})]^{N_{SS}^{ik,jl}} e^{-N_{OS}^{ik,jl}(\zeta_{ik}+\zeta_{jl})}}{N_{SS}^{ik,jl}!} \tag{7.3}$$

Taking the negative log likelihood ($-ln(L)$) for the probability of one of the two electrons to have a wrong charge and simplifying, one can obtain :

$$-ln[L(\zeta|N_{SS},N_{OS})] \approx -\sum_{ik,jl} ln[(N_{OS}^{ik,jl})(\zeta_{ik}+\zeta_{jl})]N_{SS}^{ik,jl} - N_{OS}^{ik,jl}(\zeta_{ik}+\zeta_{jl}) \tag{7.4}$$

This function is minimized to find the charge flip rate and the corresponding uncertainty in each (η,p_T) bin.

When applying the method, due to limited statistics in $Z \to ee$ events, an upper limit in p_T is considered (120 GeV). For electrons above this threshold same value and uncertainty as in the last p_T bin is used. The background subtraction (i.e light jets faking electrons) is done using the mass sidebands (25 GeV windows above and below the Z boson invariant mass). The total expected background contamination is the average of the observed events in the two intervals.

Sources of systematic uncertainties

- The difference in the rates with and without the background subtraction is taken as a source of systematic uncertainty. This uncertainty is calculated as a function of electron η and p_T and it is found to be around 90% in the low p_T and η range, decreasing to few % for high η and p_T.
- A closure test using MC is performed to check the reliability of data measurements. For this, the invariant mass of the same - sign and weighted opposite - sign electron pairs are compared and a good agreement is found. The difference of 1.4% for tight electrons (5% for loose electrons) is considered as a source of systematic uncertainty.
- One last source of systematic uncertainty was cumulated to the mentioned ones, the statistical uncertainty in the derivation sample. This occurs due to limited statistic in the $Z \to ee$ selected data sample (Figure 7.10).

Results

Compared to previous results, a slight increase in the electron charge flip rate is observed due to the new signal electrons isolation. Figure 7.10 shows the obtained rates, which varies between 10^{-5} and 0.05 for $|\eta| < 2.47$. The total relative systematic uncertainty is around 70% for electrons with low p_T and decreases to 5% for higher intervals. These results are in good agreement with the charge flip rate extracted using the previous *Tag&Probe* method (Section 6.5).

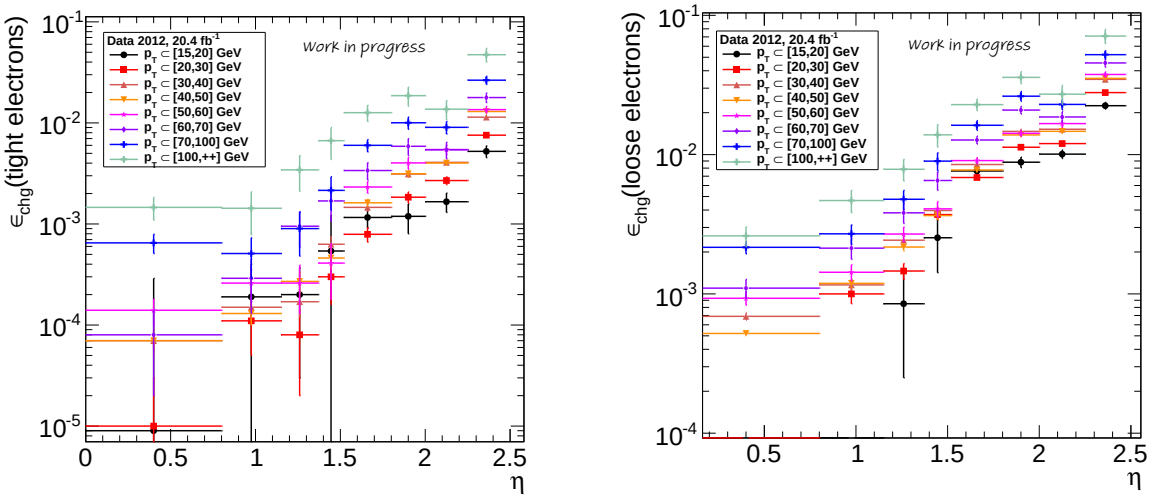


Figure 7.10: Charge flip rate [%] for electrons passing the signal definition (left) and failing the *tight++* and isolation requirements. Only the statistical uncertainties are shown.

The charge flip background estimations in all three SS leptons signal regions are shown in Table 7.3, when using the rates computed with the new and previous methods. However, it can be see that the likelihood method is giving slightly smaller uncertainties compared to the *Tag&Probe* method. For the signal regions with at least three leptons no charge flip estimation is done as it is negligible.

Method	SR0b	SR1b	SR3b
Likelihood	0.2 ± 0.1	0.5 ± 0.1	0.2 ± 0.1
Tag & Probe	0.2 ± 0.1	0.5 ± 0.2	0.2 ± 0.1

Table 7.3: Comparison of estimated charge flip background in the signal regions using the two described methods. Both statistical and systematic uncertainties are included.

7.4.2 Fake lepton background

To estimate this background in regions with three leptons in the event, the previous 4x4 matrix method has to be expanded to consider at most 2 fake leptons. For this, a dynamic matrix method is developed as it ensures a more precise estimation and discards any assumptions on the choice of the real lepton in the event. This method takes as input all the preselected leptons in the event and assigns a weight for each tight / loose combination. The event is processed separately for each weight allowing arbitrary channel selection based on the leading tight leptons. Therefore the bias of the event level variables like m_{eff} is avoided.

For example, for an event with three loose and tight leptons, $e^+e^-\mu^+$, with configuration TLL, then the matrix method will produce the following weight :

$$
\begin{array}{cc}
\textbf{Input} & \textbf{Output} \\
e^+e^-\mu^+, TLL \longrightarrow & \left\{ \begin{array}{llll}
LLL & w_{LLL} & e_L^+e_L^-\mu_L^+ & \text{Fails cuts} \\
\cdots & \cdots & & \\
TTL & w_{TTL} & e_T^+e_T^-\mu_L^+ & \text{Fails cuts} \\
TLT & w_{TLT} & e_T^+e_L^-\mu_T^+ & \text{2 leptons SS} \\
LTT & w_{LTT} & e_L^+e_T^-\mu_T^+ & \text{Fails cuts} \\
TTT & w_{TTT} & e_T^+e_T^-\mu_T^+ & > \text{2 leptons}
\end{array} \right.
\end{array}
$$

Of the possible combinations, only two pass the selection cuts (presuming trigger matching and other requirements are also satisfied). The other event combinations which are not passing the tight requirement are retained as loose by the dynamic matrix method.

For the method to work same input parameters, real ε and fake ξ efficiencies are needed as detailed in Section 6.6. The relation between the number of real N_R and fake N_F leptons in the event and the number of reconstructed tight N_T and loose N_L in a given (p_T, η) j bin can be written as :

$$
t = \begin{pmatrix} N_T \\ N_L \end{pmatrix} \quad \phi_j = \begin{pmatrix} \varepsilon_j & \xi_j \\ 1-\varepsilon_j & 1-\xi_j \end{pmatrix} \quad r = \begin{pmatrix} N_R \\ N_F \end{pmatrix} \quad \Rightarrow \quad t_\beta = \phi_{j\,\beta}{}^{\alpha} r_\alpha, \tag{7.5}
$$

where α takes two values corresponding to R or F and β for T or L.

From Equation (6.5), the number of events with one fake lepton is :

$$
\begin{pmatrix} 0 \\ N_F \end{pmatrix} = \frac{1}{\varepsilon - \xi} \begin{pmatrix} 1-\xi & -\xi \\ \varepsilon - 1 & \varepsilon \end{pmatrix} \begin{pmatrix} N_T \\ N_L \end{pmatrix}_{signal} \tag{7.6}
$$

which is used as a weight assigned to each event. Using the previous notation it can be rewritten as :

$$
t_\nu = \phi_\nu{}^{\mu} \omega_\mu{}^{\beta} \phi^{-1}{}_{\beta}{}^{\alpha} t_\alpha. \tag{7.7}
$$

where ω is selecting only the expected fake component and in this case it is :

$$
\omega = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.
$$

Generalizing to N leptons the above relation becomes :

$$
t_{\nu_1 \cdots \nu_N} = \phi_{\nu_1}^{\mu_1} \cdots \phi_{\nu_N}^{\mu_N} \; \omega_{\mu_1 \cdots \mu_N}^{\beta_1 \cdots \beta_N} \; \phi^{-1}{}_{\beta_1}^{\alpha_1} \cdots \phi^{-1}{}_{\beta_N}^{\alpha_N} t_{\alpha_1 \cdots \alpha_N}.
$$

with each ϕ computed with the appropriate fake and real efficiencies in the (p_T, η) bin for the lepton index contracted over. The generalized ω parameter is :

$$\omega_{\mu_1 \cdots \mu_N}^{\beta_1 \cdots \beta_N} = \delta_{\mu_1}^{\beta_1} \cdots \delta_{\mu_N}^{\beta_N} \; f(\beta_1, \ldots, \beta_N, \nu_1, \ldots, \nu_N),$$

and selects the expected fake or real component. The f function takes 0 or 1 value and selects the set of indices β_i that corresponds to components counted as fake background.

To conclude, for each N leptons event and a measured tight / loose combination, α, the method gives a weight for each tight / loose output combination, ν. For each combination, different leptons are defining the event, as explained in the above example. To account for this, each combination is therefore propagated separately through the final channel selection and trigger matching procedure, and the leptons are treated as tight or loose according to the output of the method. It should be underlined that one event cannot be in more than one signal region if all the selection criteria are passed.

Real and fake leptons efficiencies

To measure the input parameters, ε and ξ, same methods are employed as in Section 6.6. Due to tighter isolation cuts applied on signal electrons the real efficiency is smaller at low p_T, around 75% for $p_T > 20$ GeV and increasing up to 88% in the $p_T > 50$ GeV range. For electrons in [15, 20] GeV p_T interval the real efficiency is around 69%. For muons, similar rates as in Section 6.6.2 are obtained, from ∼95% for muons with $p_T > 20$ GeV to 99% at high p_T. For soft muons, in the [15, 20] GeV p_T interval the real efficiency is around 93%. A good agreement is found between the results obtained in data and MC. Same global uncertainty of 3% is considered to account for real efficiency extrapolation to regions with a more busy environment.

For the fake lepton efficiencies measurements, a finner binning in p_T is considered to better account for the p_T strong dependency. Electron fake efficiency is up to 80% smaller than in the previous results due to the signal electron definition. For regions with at least one b jet in the event, the rate varies between 5% and 10% depending on the p_T and η region. When a veto is imposed on the number of b jets, the rate is in the (3, 13)% range. Muon fake efficiency is found be in the (8, 13)% range. Compared to previous results, it decreases due to $e\mu$ overlap removal procedure and the new improved even selection[5]

Table 7.4 shows a comparison of the 4x4 and dynamic matrix method in SR0b and SR1b. A good agreement is observed, validating the use of dynamical matrix method.

Method	SR0b	SR1b
4x4 MxM	1.5 ± 0.3 (stat) ± 1.6 (syst)	0.8 ± 0.1 (stat) ± 1.1(syst)
dynamic MxM	$1.2^{+1.5}_{-1.2}$ (total)	$0.8^{+1.2}_{-0.8}$ (total)

Table 7.4: Comparison of estimated fake lepton background in two SS leptons signal regions using the two described methods.

7.4.3 Background method validation

Given the hypothesis of a problematic background estimation raised in the Chapter 6, different methods are implemented to validate the matrix method results for the fake lepton background and the correct simulation of fake b jets. These methods are shortly presented in this section.

Comparisons of data driven methods with MC

To validate the data driven estimates a MC based method is also developed. This method uses the MC samples to extrapolate the background prediction from regions with low $\not{E}_T$, m_{eff} and jet multiplicity to signal regions. For this study, the event selection requires at least two same - sign signal leptons and $\not{E}_T > 25$ GeV. The later cut rejects events with two or more fake leptons which are characterized by small $\not{E}_T$. To derive data to MC scale factors for the fake lepton background, six non-overlapping control regions[6] are defined depending on the b jet multiplicity (= 0 or ≥ 1), lepton flavor (ee, $e\mu$ and $\mu\mu$) and jet flavor (light or heavy). The required factors are :

- The correction factor for light flavor (LH) jet reconstructed as an electron (muon);

[5] In the previous results a bug was found in the overlap removal procedure, leading to unreliable results in the electron fake rate measurement without a b jet veto at high p_T. It was also slightly impacting on the muon fake rate.

[6] Orthogonal to the signal regions; normalized to luminosity and cross section.

- The correction factor for heavy flavor (HF) hadron decaying into an electron (muon);
- The charge flip rate.

The charge flip rates are the ones obtained in Section 7.4.1, and are applied to all electrons independent on the MC sample. To obtain the remaining scale factors, the $\not{E}_T$ and number of jets distributions in the control regions are fitted using a likelihood function. This function is computed as a negative sum of logarithms of Poisson probabilities. Once obtained, they are applied to all events depending on the lepton truth information : real, fake or charge flipped leptons.

The results in the signal regions are given in Table 7.5. A good agreement is obtained between the nominal data driven dynamic matrix method and the MC based one. It should be observed that higher uncertainties are obtained for the later one, apart for SR3Lep low$\not{E}_T$.

Method	SR0b	SR1b	SR3b	SR3L low$\not{E}_T$	SR3L high$\not{E}_T$
Monte Carlo-based	5±5	$0.6^{+1.4}_{-0.6}$	$2.0^{+1.4}_{-1.3}$	$1.0^{+0.8}_{-0.7}$	<0.1
dynamic MxM	$1.2^{+1.5}_{-1.2}$	$0.8^{+1.2}_{-0.8}$	0.7 ± 0.6	1.6 ±1.6	< 0.1

Table 7.5: Comparison of estimated fake lepton background in all signal regions using the nominal dynamic matrix method and the MC based one. Total uncertainty is shown.

Alternative method for SR3b detector background

For SR3b a further check is performed given the importance of this SR. The signal region is dominated by background including $t\bar{t}$ events with 2 real and at least one fake b jets, and one fake lepton. The fake b jet is coming either from a c quark either from the associated light jet.

Therefore, the generalized b jet matrix method developed by the group searching for SUSY in events with at least 3 b jets [248] is adapted to estimate the number of fake b jets. It is very similar with the method used for the fake lepton background described in Section 7.4.2. The differences between the two methods are : the loose leptons category is replaced by no signal b jets and the tight leptons by signal b tagged jets. The input parameters are the real and fake b jet efficiencies, and are estimated in data and $t\bar{t}$ MC. For the former the rates provided by Flavor Tagging performance group are used, while for the latter proper estimations are performed.

Two regions with a single lepton or two opposite - sign leptons in the event are defined for a complete measurement of the fake b jets efficiencies. In addition, events are requested to have at least two b jets, where two of the tagged b jets are most likely to be real and any additional jet, fake. The contamination with real three b jets, ie. $t\bar{t}b\bar{b}$, is subtracted using MC predictions. Events with a fake lepton reconstructed as a fake b jet is found to be negligible. The *Tag&Probe* method is applied, where the tag is the two real b jets pair and the probe is the additional b jet present in the event. In MC the fake b jet is identified using the truth information, while in data it is randomly selected and is generally taken to be the lowest p_T b-tagged jet. Both η and p_T dependencies are considered.

Results show a good agreement between data and MC measurements. The rates measured in the OS leptons region, which suffered from low statistic, vary between 1.5% at p_T ∼50 GeV and 3% at higher p_T. The one-lepton sample gives higher efficiencies, from ∼2% at low p_T to 6% in the [60, 200] GeV p_T range and 4% at higher p_T.

Expected background in SR3b, after applying the nominal and b jet matrix methods, is presented in Table 7.6. Here *nominal* stands for the estimation done using MC and *b jet MxM* for the fake lepton dynamic matrix method respectively. The b jet fake efficiencies are the ones measured in data in the region with at least one lepton. A good agreement is found. It can be observed that the uncertainty associated to the b jet matrix method is slightly higher and are dominated by the statistical component. Hence the b jet matrix method is used only as a cross check.

Method	SR3b
b jet MxM	2.1 ± 0.7
nominal	2.9 ± 0.9

Table 7.6: Comparison of estimated background in SR3b using the nominal (MC estimation + dynamic matrix method) and the b jet matrix method. Only the total uncertainty is shown.

Conclusions

Generally a very good compatibility between the results obtained with all the mentioned methods used to cross-check the background estimation is achieved. This affirmation is valid for the signal regions and for regions very close to them. Given this outcome the understanding of the detector background is considered to be satisfactory.

7.4.4 Validation regions

Several regions kinematically close to the signal regions are defined to validate the background. Their definition depends on lepton and jet flavor (ee, $e\mu$ and $\mu\mu$; with/without b jets). To avoid any bias, the overlap with the signal regions is minimized by asking an upper cut on $\not{E}_T$ of 200 GeV.

In Figure 7.11 the number of signal b jets and jets distributions are shown. A very good agreement is obtained between data and expectation. This affirmation is valid also at high jet multiplicities as required in the signal regions, with at least three jets (SR1b and SR0b), four (SR3Lep) and five (SR3b). Compared to Figure 6.15, the agreement in regions with 2 or 3 b jets is much better in $e\mu$ and $\mu\mu$ channels. This is due to new fake factors used for the extrapolation from regions with at least 1 b jet to regions with higher jet multiplicities.

Figure 7.12 shows the $\not{E}_T$ distribution after the lepton selection and at least three jets requirement. A good agreement between the data and SM expectation is observed in the regions with $\not{E}_T < 80$ GeV. At higher values the noticed disagreement is considered to be a fluctuation due to the limited statistic in data. An underestimation of the fake lepton background is not a priori excluded as a possible explanation.

Similarly to Section 6.7, the validation of miss-tagged jets modeling in MC is performed in regions with opposite - sign leptons and the agreement between the observed events and expected background is found to be satisfactory. Same conclusions are valid for all the studied kinematic variables distributions in the various regions kinematically close to the signal regions, proving that all methods used for the background estimation are robust. More distributions are shown in Appendix B.2.

Finally the diboson and $t\bar{t} + V$ sources of SM backgrounds are checked in appropriate validation regions, as defined in Section 6.4. This remains a very important verification as the dominant SM background is found to be the diboson production, especially the WZ processes in the regions with a b jet veto, and the $t\bar{t} + V$ processes in the regions with b jets and same - sign or three leptons. The results are presented in Figure 7.13, and the agreement between data expectation is much better than in Chapter 6. This is explained by the changes in event selection, like lepton p_T threshold of 15 GeV and the inclusion of events with three leptons.

7.5 Final results using 20.3 fb^{-1} of 8 TeV data

With the complete understanding of the SM and detector backgrounds, results in all five signals regions are presented in Section 7.5.1. A short discussion on the level of systematic uncertainties is presented in the same section. General limits on BSM physics are shown in Section 7.5.2, while model dependent limits are presented in Section 7.5.3. This sections closes with some general conclusions on the mass limits reached at the moment of publication.

7.5.1 Results in the signal regions

Using the discovery set of cuts detailed in Table 7.2 the obtained results are shown in Table 7.7 for all five signal regions. For the SM background only the total uncertainty is given, while for the detector backgrounds both statistic and systematic sources are presented.

No excess above the Standard Model expectation is present in SR3b, SR3lep_low$\not{E}_T$ and SR3lep_high$\not{E}_T$, while in the other signal regions small excesses are observed : 1.8σ in SR0b and 1.5σ in SR1b. If the two signal regions are combined the significance of the excess becomes 2.1σ. Differences with the previous results in SR3b (Table 6.10) comes from the cut on the number of light jets which was tightened from 4 to 5. The results in SR0b and SR1b are significantly different from the preliminary results. Detailed investigations show that the difference comes from new object definitions (i.e. $\not{E}_T$, electron isolation) and the inclusion of a third lepton veto.

Other investigations are carried on the data excess as it is found to be only in the channels including electrons (Table B.6) : seven events are observed in the ee channel, five in the $e\mu$ channel and only two in the $\mu\mu$ channel, while the corresponding standard model expectations are 1.48 ± 0.92, 3.12 ± 1.17 and 1.87 ± 0.68 events respectively. Almost all electrons are well isolated and very energetic, their p_T increasing up to 330 GeV where the statistic is too limited to perform any investigations with data. These events are

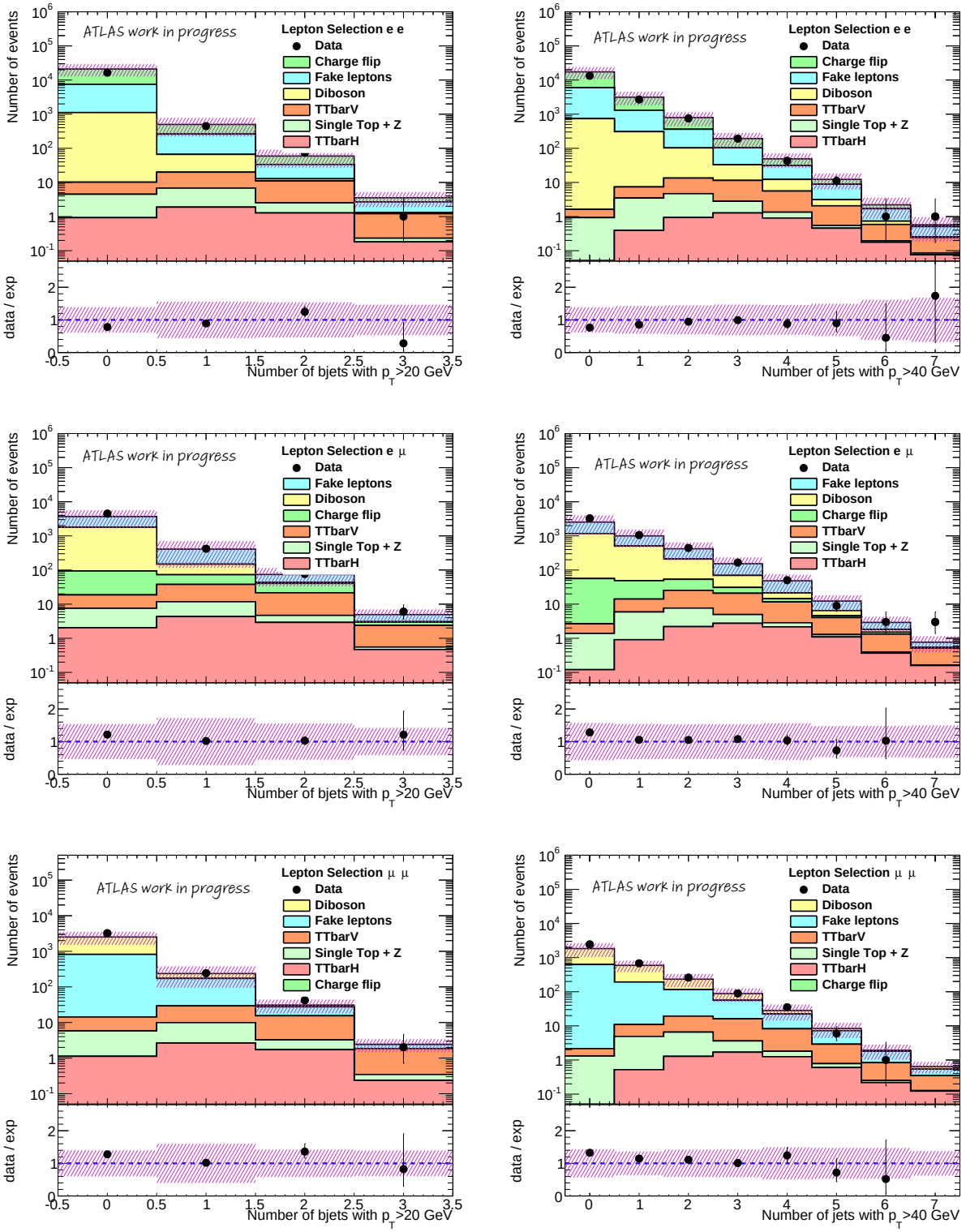


Figure 7.11: Background validation : number of signal b jets with $p_T > 20$ GeV (left) and jets with $p_T > 40$ GeV (right) distributions in ee (top), $e\mu$ (middle) and $\mu\mu$ channels. The hashed magenta area represents all the associated uncertainties on the background estimates. The black lines on the data points represent the associated statistical uncertainty. Data corresponding to Ł = 20.3 fb^{-1} and $\sqrt{s} = 8$ TeV.

also characterized by very energetic jets, hence by high m_{eff} as illustrated in Figure 7.14 for one particular event : $m_{\text{eff}} = 1.7$ TeV, the highest of all signal regions, three very energetic b jets with $p_T = 222$ GeV, 349 GeV and 711 GeV, and $\not{E}_T = 233$ GeV. More details are given in Appendix B.2.

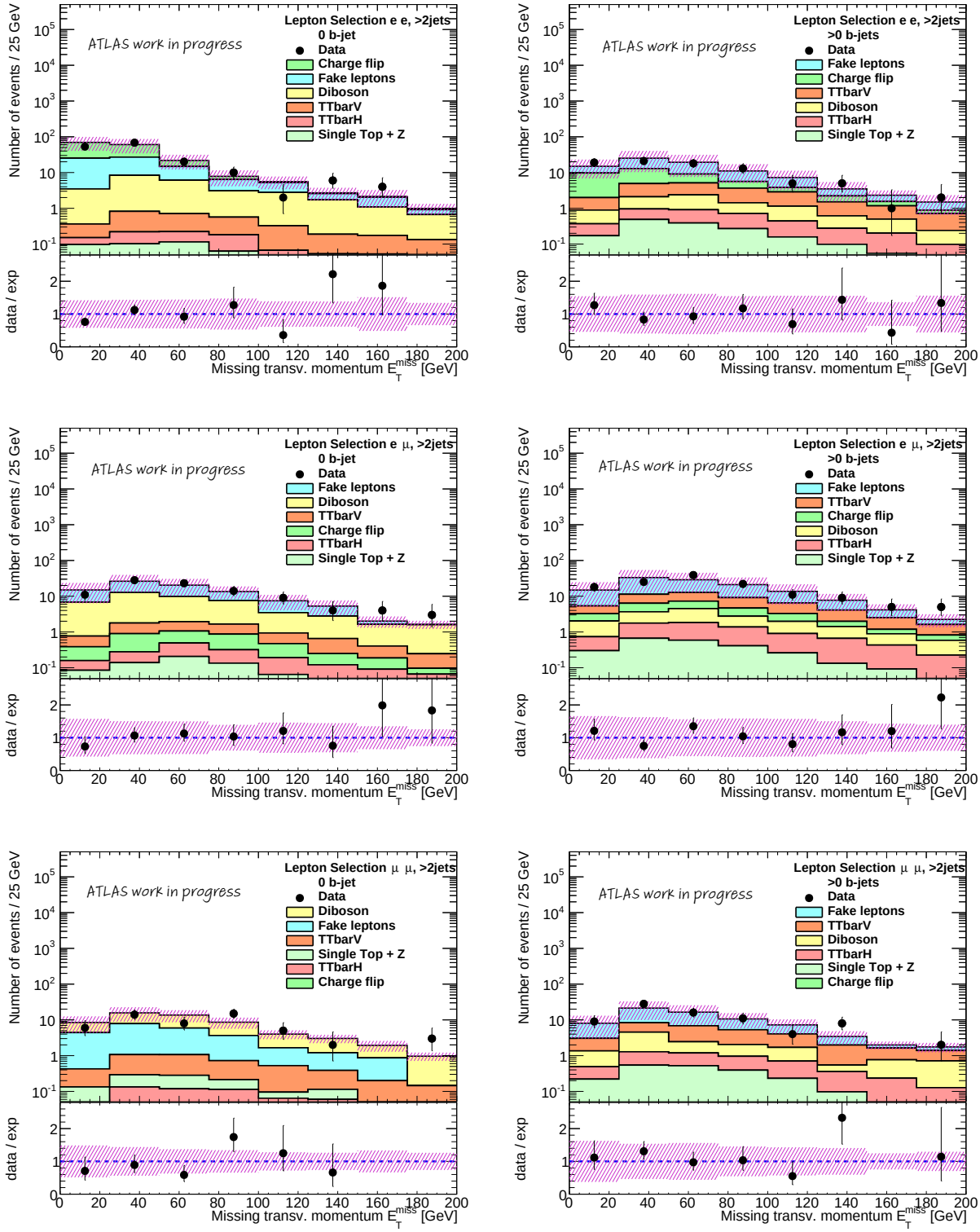


Figure 7.12: Background validation : missing transverse momentum distribution in regions with a veto on number of b jets (left) and in regions with at least 1 b (right) in ee (top), $e\mu$ (middle) and $\mu\mu$ channels. The hashed magenta area represents all the associated uncertainties on the background estimates. The black lines on the data points represent the associated statistical uncertainty. Data corresponding to $\mathrm{L} = 20.3\ \mathrm{fb}^{-1}$ and $\sqrt{s} = 8$ TeV.

Systematic uncertainties

The various sources of systematic uncertainties were already presented in Section 6.8.2. To improve the systematic uncertainty assigned to $t\bar{t} + V$ and diboson processes, several studies are carried on as detailed in Appendix B.2. The breakdown of the systematic uncertainties on the expected background estimates, expressed in units of events, is presented in Table 7.8. Generally, similar performance as for the preliminary

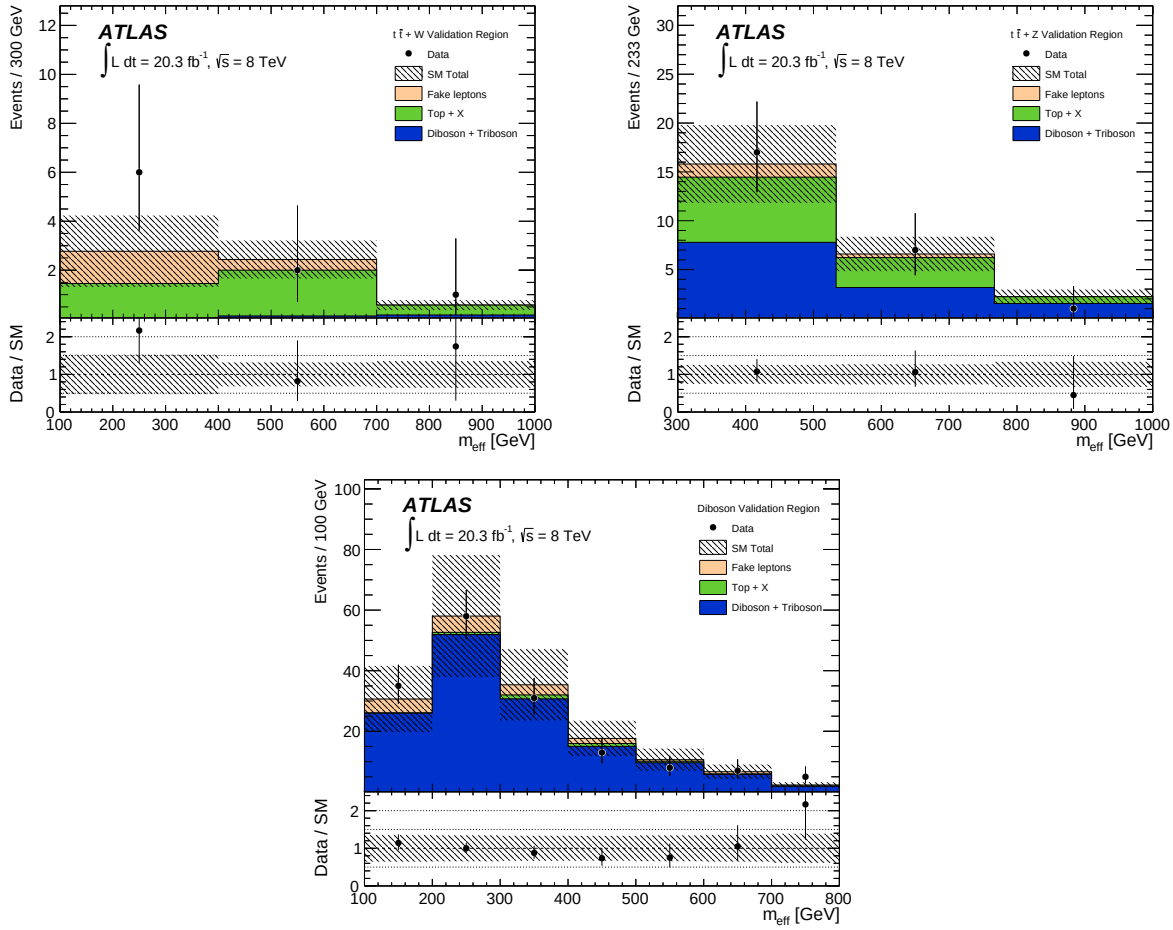


Figure 7.13: Results in $t\bar{t} + W$ (top left), $t\bar{t} + Z$ (top right) and diboson (bottom) validation regions. Data corresponding to Ł = 20.3 fb^{-1} and $\sqrt{s} = 8$ TeV.

# events	SR3b	SR0b	SR1b	SR3lep_low$\not{E}_T$	SR3lep_high$\not{E}_T$
observed	1	14	10	6	2
expected background	2.19 ± 0.80	6.46 ± 2.34	4.68 ± 2.14	4.33 ± 2.06	2.53 ± 0.92
top+V	1.28 ± 0.51	0.85 ± 0.37	2.45 ± 1.71	1.59 ± 1.04	1.26 ± 0.66
diboson	0.03 ± 0.02	4.21 ± 1.70	0.91 ± 0.42	1.19 ± 0.62	1.24 ± 0.59
fake leptons	$0.70 \pm 0.26 \pm 0.58$	$1.18 \pm 0.53 \pm 1.48$	$0.77 \pm 0.34 \pm 1.66$	$1.55 \pm 0.49 \pm 0.38$	$0.03 \pm 0.08 \pm 0.10$
charge flip	$0.18 \pm 0.05 \pm 0.03$	$0.22 \pm 0.05 \pm 0.04$	$0.54 \pm 0.07 \pm 0.12$	<0.01	<0.01

Table 7.7: Number of events together with the total or statistical and systematic uncertainty in the five signal regions using the discovery set of cuts. Top + V category includes four tops, top + Z, $t\bar{t} + V$ processes.

results is achieved. In SR0b, SR1b and SR3Lep_high$\not{E}_T$ the uncertainty on the Standard Model background dominates, while in SR3b and SR3Lep_low$\not{E}_T$ the uncertainty on the fake lepton background respectively. The largest systematic uncertainties are from the estimation of the fake lepton rate and from the theoretical predictions for diboson and/or $t\bar{t}$ + V processes. Uncertainties are quoted as symmetric, except where the negative prediction is truncated to zero.

7.5.2 General limits on BSM physics

As no significant excess above the Standard Model prediction is observed, upper limits at 95% CL on the number of Beyond Standard Model events for each signal region are derived using the CL$_s$ formalism. The same procedure as in Section 6.8.4 is followed. The number of observed, S^{95}_{obs}, and expected, S^{95}_{exp}, 95% CL upper limits on the number of BSM events are presented in Table 7.9. The 95% CL upper limits on visible cross section is also shown. To obtain these results, the discovery set of cuts with minimum requirements on m_{eff} (Table 7.2) are used.

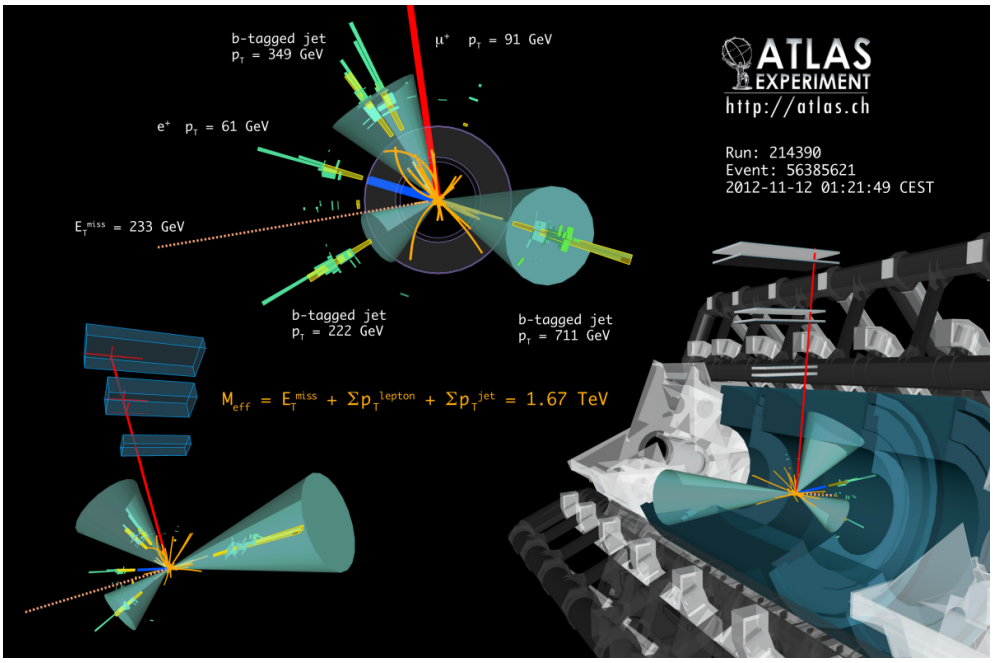


Figure 7.14: Observed event in SR1b. It is characterized by the highest m_{eff} of all signal regions, 1.7 TeV. The green cones shown the three leading jets with are tagged as b jets. The same events is shown with three different views.

Systematic uncertainties	SR3b	SR0b	SR1b	SR3lep_low$\not{E}_T$	SR3lep_high$\not{E}_T$
Fake-lepton background	± 0.6	$^{+1.5}_{-1.2}$	$^{+1.2}_{-0.8}$	$\mathbf{\pm 1.6}$	< 0.1
Theory unc. on dibosons	< 0.1	$\mathbf{\pm 1.5}$	± 0.3	± 0.4	± 0.4
Jet and $\not{E}_T$ scale and resolution	± 0.1	$\mathbf{\pm 0.7}$	± 0.4	± 0.4	± 0.3
Monte Carlo statistics	± 0.1	± 0.5	± 0.2	± 0.4	± 0.4
b-jet tagging	± 0.2	± 0.5	± 0.1	< 0.1	± 0.1
Theory unc. on $t\bar{t}V$, $t\bar{t}H$, tZ and $t\bar{t}t\bar{t}$	± 0.4	± 0.3	$\mathbf{\pm 1.7}$	$\mathbf{\pm 1.0}$	± 0.6
Trigger, luminosity and pile-up	< 0.1	± 0.1	± 0.1	± 0.1	± 0.1
Charge-flip background	± 0.1	± 0.1	± 0.1	–	–
Lepton identification	< 0.1	± 0.1	< 0.1	± 0.1	± 0.1
Total uncertainty	± 0.80	± 2.34	± 2.14	± 2.06	± 0.92

Table 7.8: The breakdown of the systematic uncertainties on the expected backgrounds, expressed in units of events. The uncertainty on the fake lepton estimation is including the statistical source. The individual uncertainties are correlated and therefore do not necessarily add up in quadrature to the total systematic uncertainty.

	SR3b	SR0b	SR1b	SR3lep_lowMET	SR3lep_highMET
$< \sigma_{vis} >^{95}_{obs}$	0.19	0.80	0.65	0.42	0.23
S^{95}_{obs}	3.9	16.3	13.3	8.6	4.6
S^{95}_{exp}	$4.4^{+1.7}_{-0.6}$	$8.9^{+3.6}_{-2.0}$	$8.0^{+3.3}_{-2.0}$	$7.2^{+2.9}_{-1.3}$	$5.0^{+1.6}_{-1.1}$

Table 7.9: 95% CL upper limits on visible BSM cross section, on observed and expected number of BSM signal events.

7.5.3 Interpretation, model dependent limits

Using the phenomenological and simplified models presented in Sections 6.1 and 7.1 exclusion limits are set to further constrain the SUSY phase space. For a better coverage all the five signal regions are combined during the fit and the m_{eff} distribution is divided in several bins (Table 7.10). Figure 7.15 shows the binned m_{eff} distribution with few signal points of particular sensitivity to each signal region : gluino-

mediated top squark $\rightarrow bs$ (RPV) for SR3b, gluino-mediated squark $\rightarrow q'W\tilde{\chi}_1^0$ for SR0b, gluino-mediated top squark $\rightarrow c\tilde{\chi}_1^0$ for SR1b and gluino-mediated squark $\rightarrow$ sleptons for SR3lep_low$\not{E}_T$.

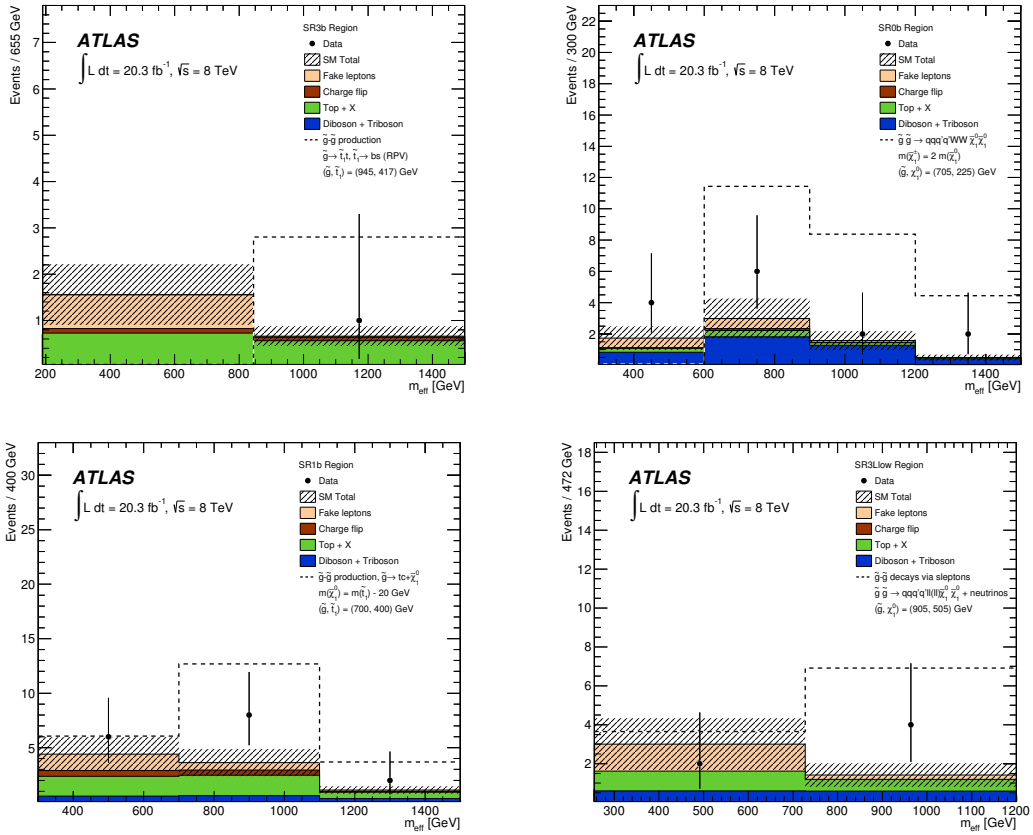


Figure 7.15: Binned m_{eff} distribution in four signal regions, SR3b (top - left), SR0b (top - right), SR1b (bottom - left) and SR3lep_low$\not{E}_T$ (bottom - right). The uncertainty band includes both statistical and systematic sources on the background estimation.

SR	bin 1	bin 2	bin 3	bin 4
SR0b	300 - 600	600 - 900	900 - 1200	>1200
SR1b	300 - 700	700 - 1100	>1100	
SR3b	190 - 890	>890		
SR3Lep_low$\not{E}_T$	255 - 728	>728		
SR3Lep_high$\not{E}_T$	355 - 1078	>1078		

Table 7.10: Intervals in m_{eff} [GeV] used for the exclusion fit.

Gluino pair production

Limits on gluino pair production via a virtual top squark Gtt model are shown in Figure 7.16 left. Compared with the preliminary results, the analysis excludes SUSY signal for gluino masses lighter than ~1.1 TeV (mainly with SR3b). Update results from the analysis looking at 0 leptons and 3 b jets channel are available and shown in Figure 7.16 right. It can be seen that its sensitivity is well above the one reached with same - sign leptons. Though, the limits in the compressed region did not really improved.

Results for other two gluino mediated top squark models are shown in Figure 7.17. In the left plot the model with $\tilde{t} \rightarrow c\tilde{\chi}_1^0$ is shown. The highest sensitivity is reached with SR1b and the limits are presented in $m_{\tilde{g}}$ - $m_{\tilde{t}}$ mass plane. Gluinos are excluded for masses between 660 and 840 GeV for $\tilde{t}$ masses between 500 and 100 GeV. In the right plot a RPV model with $\tilde{t} \rightarrow bs$ is considered to set the exclusion limits in $m_{\tilde{g}}$ - $m_{\tilde{t}}$ mass plane with the same - sign leptons channel. Gluinos are excluded with masses up to 890 GeV for $\tilde{t}$ masses of 300 - 1000 GeV. SR3b is driving exclusion limits as the final state is characterized by many b jets and very low $\not{E}_T$ from SM neutrinos coming from the W boson. These limits are competitive with the ones obtained by the analysis using 0 lepton and 7 - 10 jets [249] shown in the right plot with a blue line.

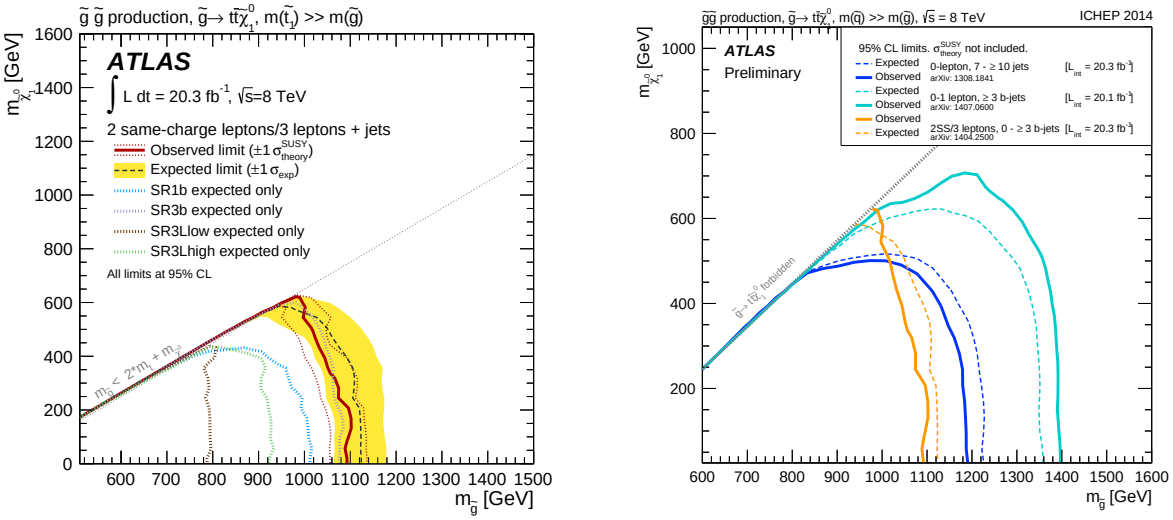


Figure 7.16: Expected and observed limits for gluino pair production via a virtual top squark, $\tilde{g} \rightarrow t\bar{t}$ model. The importance of each signal region is shown in the left plot. A comparison with all ATLAS analyses available at the time of publication is shown on the right side plot [275].

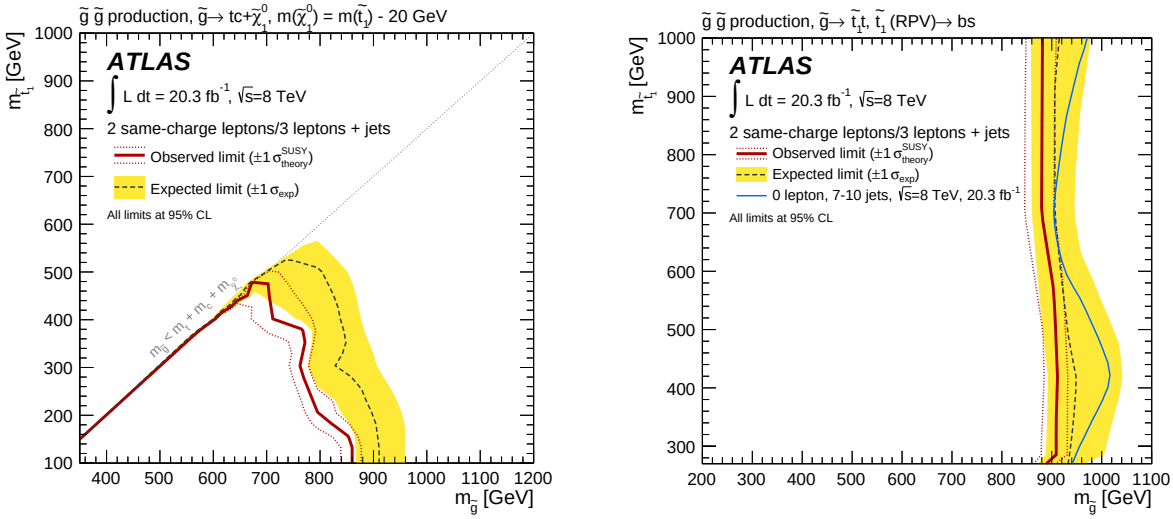


Figure 7.17: Expected and observed limits for two gluino mediated top squark models, $\tilde{t} \rightarrow c\tilde{\chi}_1^0$ (left) and $\tilde{t} \rightarrow bs$ (right).

Several gluino mediated or direct first and second generation squarks models with 1-step or 2-step decays are probed as well and results are shown in Figure 7.18. As explained in Section 6.1 the squarks are assumed to be left handed and degenerate in mass. A first presented model (top plot) is $\tilde{g} \rightarrow qq'W\tilde{\chi}_1^0$ via $\tilde{\chi}_1^\pm$ with $m_{\tilde{\chi}_1^\pm} = 2\ m_{\tilde{\chi}_1^0}$. This scenario excludes gluinos up to 860 GeV for LSP masses up to 400 GeV. SR0b is driving the sensitivity as no b jets and only 2 same - sign leptons are present in the cascade decays. In the second model (middle plots) the squarks are assumed to decay in $q'WZ\tilde{\chi}_1^0$, via $\tilde{\chi}_1^\pm$. In the direct squarks production model the final state is characterized by 2 light quarks and several leptons from the W and Z bosons. When the gluino mediated production model is considered the final state is formed by 4 light jets and several leptons from the W and Z bosons decays. Results are interpreted in the parameter space of the gluino (squark) and $\tilde{\chi}_1^0$ masses, and the sensitivity is dominated by SR3lep_low$\not{E}_T$ in the compressed region, otherwise by SR3lep_high$\not{E}_T$. With these scenarios gluino masses are excluded up to 1040 GeV for $\tilde{\chi}_1^0$ masses up to 520 GeV and squarks masses up to 670 GeV for $\tilde{\chi}_1^0$ masses up to 300 GeV. The last considered topologies (bottom plots) are the gluino mediated or direct squark via sleptons model, with $\tilde{l} \rightarrow l\tilde{\chi}_1^0$, that lead to final states with missing transverse momentum, two or four jets and up to four charged leptons. Results are interpreted in the gluino (squark) and $\tilde{\chi}_1^0$ mass plane, are the scenarios are excluded for gluino (squark) masses up to 1200 (780) GeV and $\tilde{\chi}_1^0$ masses up to 660 (460) GeV. The sensitivity is driven by SR3lep_high$\not{E}_T$.

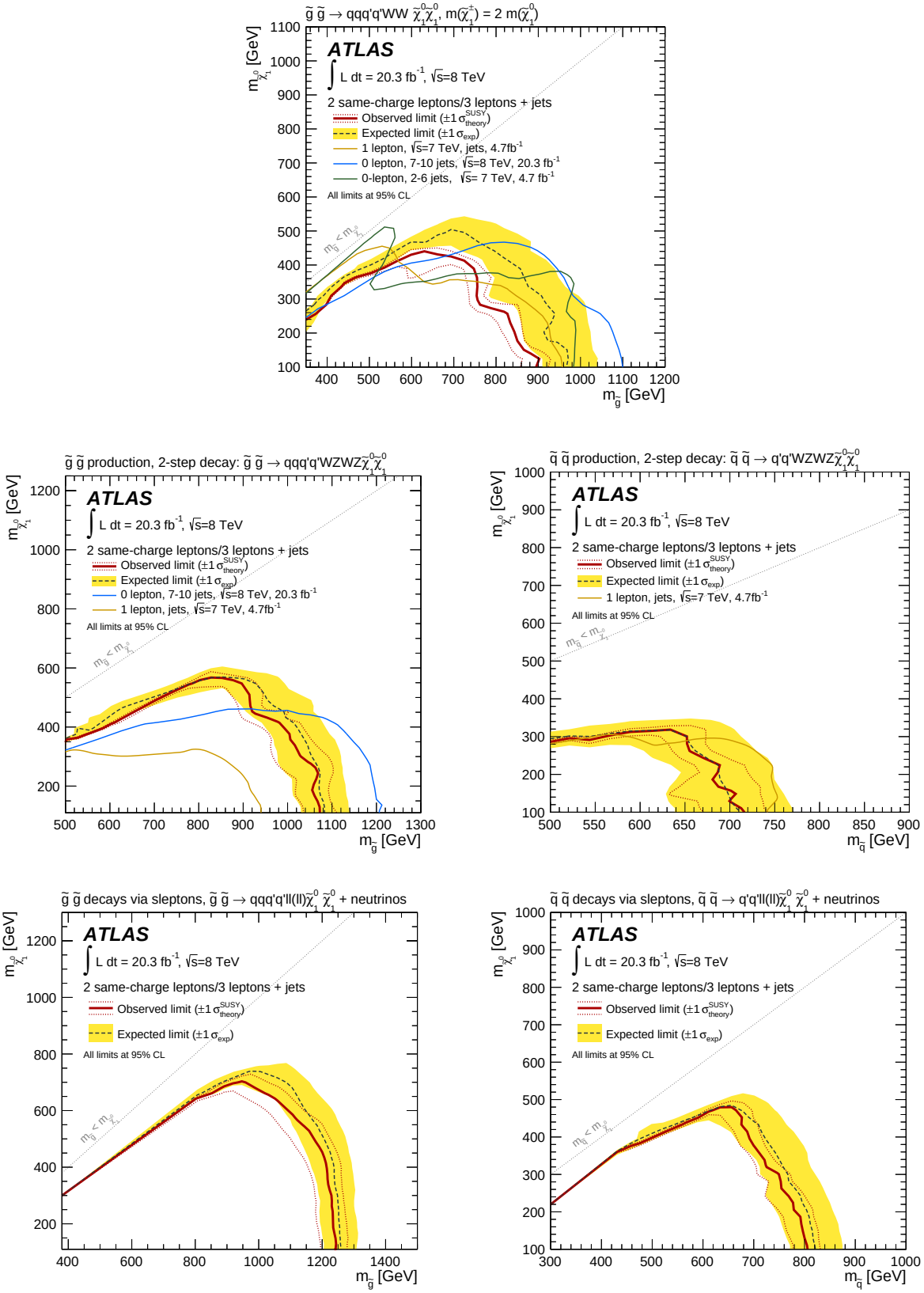


Figure 7.18: Expected and observed limits for two gluino mediated or direct first and second generation squarks.

Bottom squark pair production

In Figure 7.19 exclusion limits for direct bottom squark production simplified models with $\tilde{b} \rightarrow t\tilde{\chi}_1^0$ are shown. The final state is formed by 2 b jets, 4 W bosons and 2 $\tilde{\chi}_1^0$. SR1b and SR3lep are driving the sensitivity. When the LSP is set to 60 GeV, results are interpreted in $m_{\tilde{b}}$ - $m_{\tilde{\chi}_1^{\pm}}$ mass plane as shown in Figure 7.19 left. When the $\tilde{\chi}_1^{\pm}$ and $\tilde{\chi}_1^0$ masses related by $m_{\tilde{\chi}_1^{\pm}} = 2m_{\tilde{\chi}_1^0}$ limits are set in the $m_{\tilde{b}}$ - $m_{\tilde{\chi}_1^0}$ mass

plane, Figure 7.19 right. In this framework bottom squarks are excluded up to 440 GeV for any chargino or neutralino masses. Compared to preliminary limits showed in Figure 6.20, the exclusion is weaker due to the small excess observed in SR1b signal region.

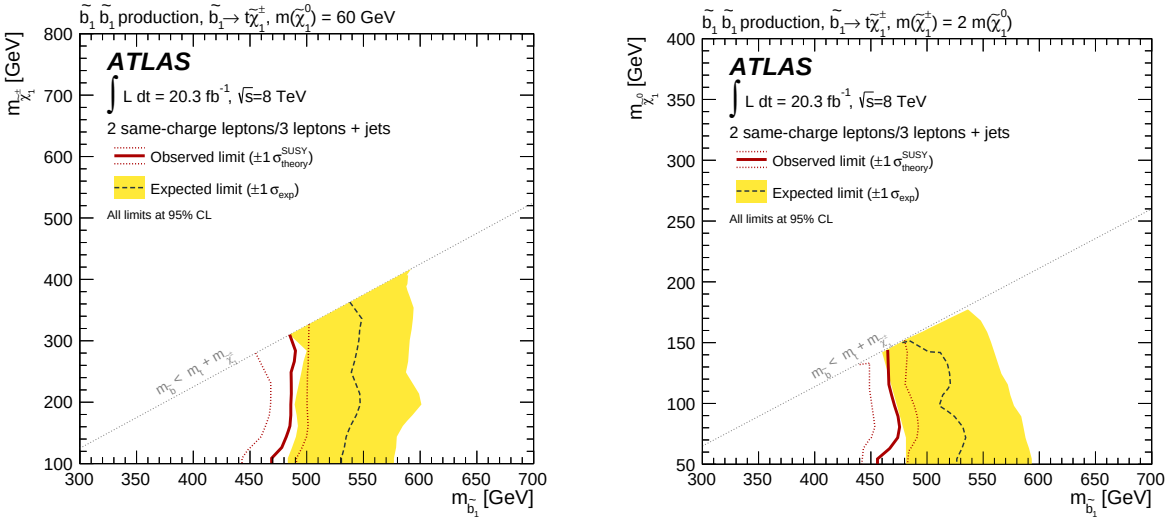


Figure 7.19: Expected and observed limits for two direct bottom squark models, $\tilde{b} \to t\tilde{\chi}_1^\pm$, with $m_{\tilde{\chi}_1^0}$ = 60 GeV (left) and $m_{\tilde{\chi}_1^\pm} = 2m_{\tilde{\chi}_1^0}$ (right).

SUSY benchmark models

In Figure 7.20 exclusion limits for the four standard benchmark models are shown. Results on mSUGRA/CMSSM model are set as function of m_0 and $m_{1/2}$ parameters (top left) and the sensitivity is found to be lead by SR3b, SR3lep_highMET and SR1b signal regions. Values of $m_{1/2}$ are excluded below 360 GeV for m_0 values up to 6 TeV. In the top right plot results on bRPV mSUGRA/CMSSM model with R-parity violating terms and unstable $\tilde{\chi}_1^0$ are obtained in m_0 and $m_{1/2}$ parameters plane. Within this model values of $m_{1/2}$ are excluded between 200 and 490 GeV for m_0 values up to 2.1 TeV, and the sensitivity is driven by SR3b. Models with $m_{1/2} < 200$ GeV are not considered as the lepton acceptance is highly reduced due to LSP lifetime increase in this region. GMSB model results are presented in the left bottom plot in the SUSY breaking mass scale, Λ, and $\tan(\beta)$ parameters space. With this analysis tighter limits are set, mainly driven by SR3lep_high$\not{E}_T$, and Λ values below 75 TeV are excluded for $\tan(\beta)$ below 60 corresponding to a lower limit on the gluino mass of 1.7 TeV. A last model, mUED is used for interpretation (bottom right), and limits are expressed as a function of the inverse compactification radius, $1/R$, and cut-off scale, Λ, multiplied with R. Values of $1/R$ below 900 GeV are excluded for $\Lambda R > 19$. The drop in sensitivity with the decrease of ΛR parameters is explained by the reduced mass splitting between the Kaluza-Klein states. SR3lep_high$\not{E}_T$ and SR0b are dominating the coverage.

Conclusions

The gluino pair production top squark mediated models exclude gluino masses below ∼1 TeV independently of the top squark mass. This is possible through a combination of SR3b, SR1b and SR3Lep. The three leptons signal regions and SR0b are highly sensitive to two step decay of first/second generation squarks and gluino pair production with R-parity conserved. The reach in gluino and $\tilde{\chi}_1^0$ masses varies by more than 300 GeV between the most favorable case, production via sleptons with the highest branching ratio, and the most difficult case, $q'W\tilde{\chi}_1^0$ in the cascade decays with the lowest branching ratio. In the very compressed spectra, with mass difference smaller than 100 GeV, none of the signal regions are sensitive. At high gluino masses the limits obtained with 0 lepton and 7 - 10 jets channel [249] are much tighter, this analysis benefiting from higher branching ratios. Compared to preliminary results, the observed limits are worse due to the small excesses observed in several signal regions.

7.6 Other searches in the two same - sign leptons channel

In this section a comparison with the results obtained by the CMS collaboration with $\sqrt{s} = 8$ TeV, $L = 20.3$ fb^{-1} of data in the same - sign lepton channel is performed. To shows the versatility of this signature and its coverage at the LHC experiment, all the ATLAS published analyses targeting this particular final state are shortly described.

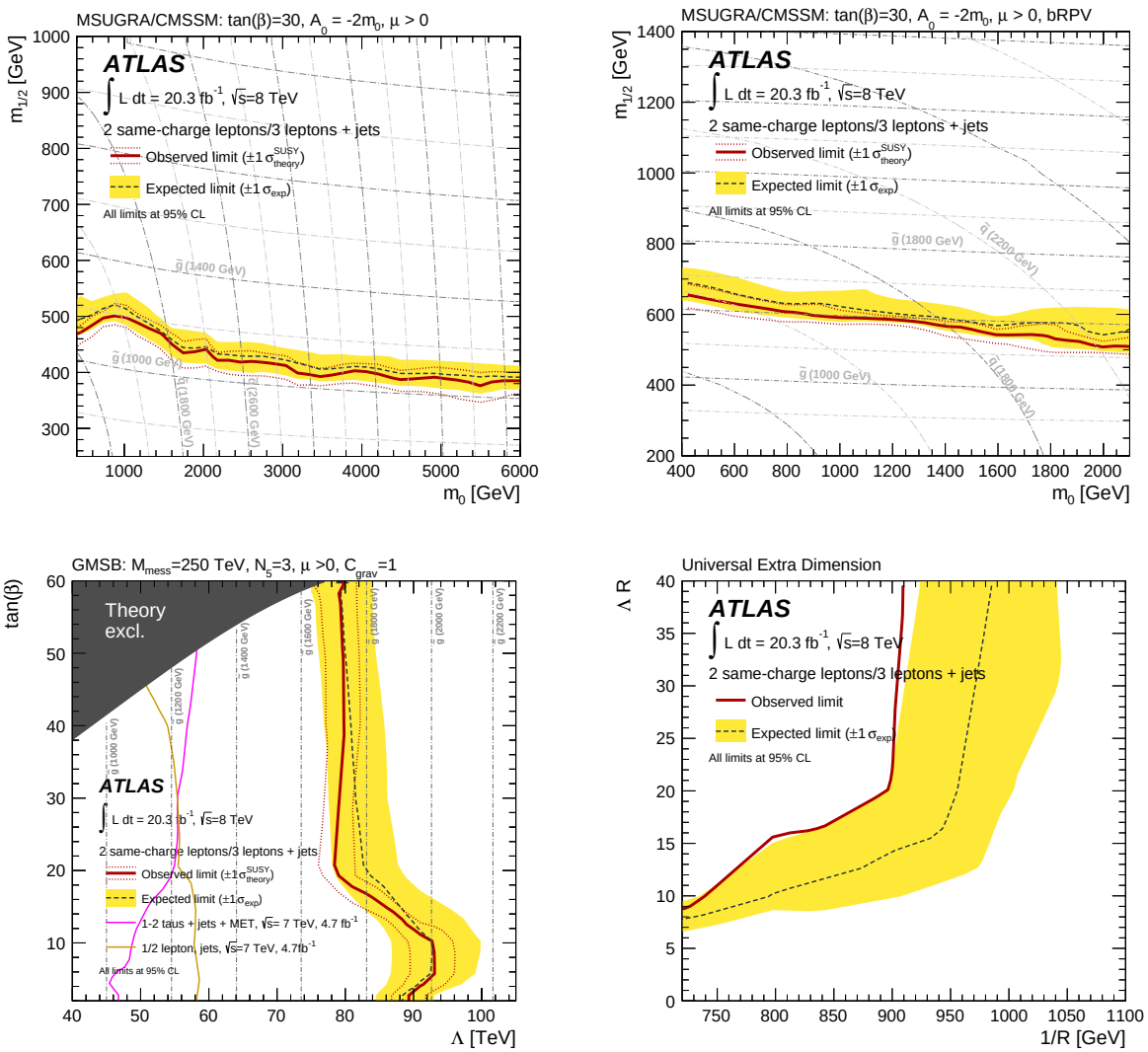


Figure 7.20: Expected and observed limits for the mSUGRA/CMSSM, bRPV, GMSB and mUED models.

Comparison with CMS

First, a short comparison with the results obtained by the CMS collaboration with a similar SUSY analysis is presented [276]. Compared to ATLAS, the CMS choses to perform a full scan of the kinematic phase space, by defining a total number of 50 signal regions with soft ($p_T > 10$ GeV) and energetic ($p_T > 20$ GeV) leptons. An overlap between these regions exist. The exclusion limits are obtained considering the results from the best signal region in each point of the SUSY phase space in the considered model. As in ATLAS, the discriminant variables are $\not{E}_T$, H_T, number of jets and b jets.

The SM background estimation is performed in a similar way as in ATLAS (MC simulations and no CRs due to the low available statistics). For the detector background different methods are employed. The fake leptons are estimated from events with one or two loose (but not tight) leptons which are weighed by the tight-to-loose ratio. The measurement is performed as a function of lepton type, p_T and η in a data sample of events with a single lepton candidate and a well separated jet. This region is defined with a veto on Z bosons candidates and small $\not{E}_T$ to suppress the leptons from W decays.

The systematic uncertainty on the fake lepton background is obtained in two steps :

- After testing the ability of this method to predict same - sign leptons background in simulated $t\bar{t}$ events.
- After quantifying the variation of the tight-to-loose ratio as function of the away jet[7] p_T threshold and the addition of the b tag requirement on this jet.

An overall systematic uncertainty of 50% is obtained.

[7] Away jet = a well separated jet of the lepton candidate.

For the charge flip background estimation both MC and data samples are used. The charge flip rate is measured in $Z \to ee$ MC and tested in a data sample using ee events with invariant mass consistent with the Z boson mass. An overall systematic uncertainty of 20% is taken after comparing the rates in $Z \to ee$ and $t\bar{t}$ MC. The estimation in the region of interest is done in data, after weighting OS events as in our analysis. The disadvantage of this approach, compared with ATLAS, is the reliability on MC simulation, which are not perfectly describing the data and suffers from low statistics.

As no excess was found several models were used to interpret the results in the signal regions. A detailed comparison between ATLAS and CMS common exclusion limits is done in Ref. [277] and shows that similar sensitivity to SUSY models is reached.

Other ATLAS results using $L = 20.3$ fb^{-1} of $\sqrt{s} = 8$ TeV data

Several other searches with two same - sign lepton signature are performed in ATLAS. First the **study of same - sign W boson production in association with 2 jets**, $W^{\pm}W^{\pm}jj$ [278] is discussed. This search is presenting the first evidence of massive vector boson scattering (VBS) in the pure W leptonically decay channel. It plays a key role in probing the BSM physics, as many scenarios predict an intensification in VBS due to i.e. the presence of additional resonances. The considered W boson decay products are either electrons or muons (ee, $e\mu$ and $\mu\mu$). The production cross section is measured in two fiducial regions. The former, called also the inclusive region, includes exactly two same - sign leptons with $p_T > 25$ GeV and $|\eta| < 2.5$. Beside the leptons, at least two jets with $p_T > 30$ GeV and $|\eta| < 4.5$ are required. As these jets arise from the VBS processes, they are expected to be quite forward ($m_{jj} > 500$ GeV). To remove the $t\bar{t}$ background, no b jets are required in the event. The second measurement region is included in the previous region. To increase the purity of electroweak $W^{\pm}W^{\pm}jj$ production, by minimizing the strong $W^{\pm}W^{\pm}jj$ processes, the two leading jets are required to be separated in rapidity ($|\Delta y_{jj}| > 2.4$). The prompt SS background (mainly WZ/γ^*) is estimated using MC after applying correcting factors to account for differences with the data. The charge flip background estimation (mainly from $t\bar{t} \to l\nu l\nu b\bar{b}$) is purely data driven and the uncertainty on the rate varies between 15% and 32% depending on the region. Note that in the analyses presented in this chapter it varies between 13 - 28%. The fake lepton background (arising from processes like W+jets, $t\bar{t}$, single top or multijets) is as well data driven, using TL events re-weighted with the fake rate. The latter is measured in a dijet sample and the achieved uncertainties are between 39% and 52%. They are dominated by the prompt lepton contamination and fake rate extrapolation to the SRs. The measured cross sections in the two fiducial regions (σ_{fid}) are in good agreement with the SM prediction (σ_{SM}). In the inclusive region σ_{fid} for the strong and electroweak production is 2.1 ± 0.5 (stat) ± 0.3 (syst) fb, compared to σ_{SM} of 1.3 ± 0.11 fb. In the second region, σ_{fid} computed for electroweak $W^{\pm}W^{\pm}jj$ production and including strong production interferences is found to be 1.3 ± 0.4 (stat) ± 0.2 (syst) fb, while σ_{SM} is 0.95 ± 0.06 fb.

Other well motivated models leading to same - sign leptons signature are studied in the exotic analyses. The first discussed is the **search for anomalous production of prompt same - sign lepton pairs and pair-produced doubly charged Higgs boson** [279]. The considered channels are ee, $e\mu$ and $\mu\mu$, and any combination of two *tight* [219] and isolated same - sign leptons is included. The event selection requires a minimum p_T threshold of 25 GeV (20 GeV) for the leading (subleading) lepton. The invariant mass of the SS leptons (m_{ll}) is required to be above 15 GeV and to reduce the background from the Z boson decay the mass difference between m_{ll} and m_Z should be below 10 GeV. The charge flip component is data-driven estimated, with the rate measured using a Likelihood method as presented in this chapter. The assigned systematic uncertainty on the rate varies between 6% and 20% as a function of electron η. For the fake lepton background, the estimation is as well purely data driven. It is based on a factor measured as the ratio between the number of events passing and failing a given selection, using a same - sign di-muon or dijet (electron factor) data sample. The uncertainty on muon factor is 17% for p_T ∼20 GeV and increases to 23% for p_T ∼60 GeV. For electrons, the uncertainty varies between 40% at p_T ∼20 GeV and 13% for p_T ∼100 GeV. For leptons with $p_T > 100$ GeV, no reliable measurement can be performed and the same factor as in the last measurement interval is used, with an uncertainty of 100%. Results are obtained after looking at the data-background estimation agreement on the m_{ll} distributions. No excess above SM expectation is observed and the results are used to set upper limits on BSM processes.

A second exotic analysis is the **search for microscopic black holes in a like-sign dimuon final state using large track multiplicity** [280]. It is known that the black holes evaporate by Hawking radiation which determines the energy and multiplicity of the emitted particles. In this search only a pair of SS isolated muons is considered. No veto on the third muon is imposed and the leading (subleading) lepton must have $p_T > 40$ GeV (15 GeV). They arise either directly from the back hole, or form the produced SM processes. Each event is required to have a primary vertex with at least five tracks with $p_T > 400$ GeV. The track multiplicity (N_{trk}) is the key variable used to discriminate the signal over

background and it is computed using the ID tracks with $p_T > 10$ GeV and $|\eta| < 2.5$. One signal region is defined after increasing the leading muon p_T to 100 GeV and requiring $N_{trk} > 30$. The background is classified as correlated and uncorrelated. In the first category enter processes like diboson or $t\bar{t}$ and are estimated using MC. The uncorrelated sources are the fake muons (W/Z +jets), and estimated from data. The muon fake rate (the probability to reconstruct a track as a muon) is measured in a photon + jet sample and found to be around 1%. The number of observed events in the SR is 0, compatible with the SM prediction (0.6 ± 0.2). These results are used to place upper limits on BSM sources.

The last presented analysis [281] is targeting events with b jets and a pair of same - sign leptons in the final state and is very similar with the SUSY analysis presented in this chapter. The main target signals are the production of vector-like quarks or pairs of chiral b' quarks, enhanced four-top-quarks production or pairs of positive charged top quarks. A total of eleven signal regions are designed. The event selection requires two leptons with same electric charge, large $\not{E}_T$ (between 40 and 100 GeV or >40 GeV or >100 GeV) and from two to four jets. As in the analysis I am involved, the SRs are defined depending on b jet multiplicity : inclusive and exclusive $1b$, exclusive $2b$ and inclusive $3b$. Both SM background and detector backgrounds are estimated from MC. The charge flip rate is measured using the same Likelihood method applied on a $Z \to ee$ events sample. In the high p_T regime, a sample of $t\bar{t}$ events is used to reduce the statistical uncertainty. For the fake lepton background, the matrix method is used, and the fake and real rates are measured in dedicated *single lepton* regions. The real lepton identification is measured using a region with $\not{E}_T > 150$ GeV targeting $W \to l\nu$ processes. The electron fake rate is measured in a region enriched in multijet production (m_T <20 GeV, $\not{E}_T + m_T < 60$ GeV), and the muon fake rate in a region enriched in heavy-flavor decays (cuts on the muon impact parameter). These variables are parameterized with respect to lepton η and p_T, and number of b jets. The uncertainty on the fake lepton estimation in the signal regions varies between 1.5% and 33%, while on the charge flip it is in the 5 - 11% range. In few of the defined signal regions, a 2.5σ excess is observed. This excess was deeply studied, within both Exotic and SUSY ATLAS groups. Overall, an excess is obtained also in the analysis I am involved in when implementing these few signal regions. Note that given the tighter isolation used in the SS SUSY analysis, this excess is reduced when using the object definitions presented in Section 7.3.

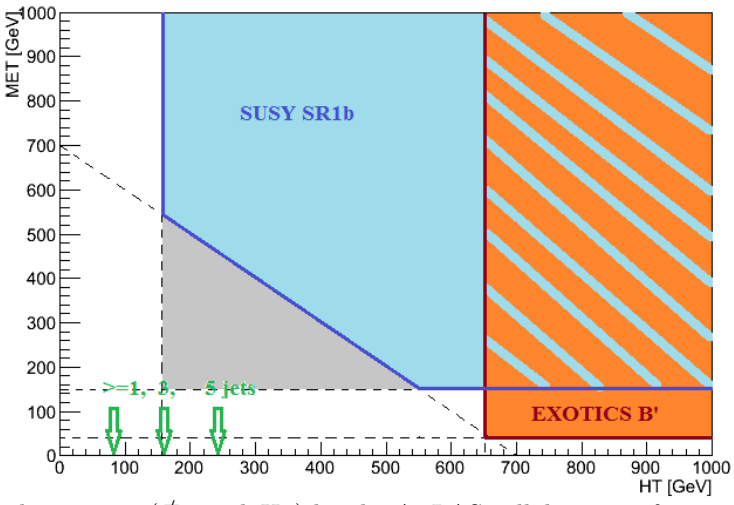


Figure 7.21: Covered phase space ($\not{E}_T$ and H_T) by the ATLAS collaboration for generic searches [2]. Note that the anti-diagonal can be read as a cut on m_{eff}.

Discussion

Given all the ATLAS public results, the same - sign lepton signature is intensively studied and covers a wide rage of the BSM physics phase space, as roughly illustrated in Figure 7.21. Generally same methods are used to estimate the detector backgrounds. The samples used to measure the needed input parameters are defined such that the obtained statistics is prioritized over the closeness to the signal regions. The main gain is the smaller uncertainties when compared to the approached used in the SS SUSY analysis presented in the thesis.

7.7 Conclusions

To improve the sensitivity to SUSY phase space, with respect to the preliminary results discussed in Chapter 6, the signal regions combining two same - sign leptons and three leptons final states have been

re-optimized to obtain the final results with $\sqrt{s} = 8$ TeV and $L = 20.3$ fb^{-1} of data. As a result, a total of five signal regions depending on lepton and b jet multiplicities are defined. The methods used to estimate the background have been improved and the confidence on the background estimation increased by the addition of new methods. For the fake lepton background, the 2×2 matrix method was generalized at the N leptons present in the event, avoiding any assumption on the leading lepton. For the charge flip rate, a likelihood based method using all the available statistics and no assumption on the two opposite - sign leptons is developed. Compared to the preliminary analysis, a b jet matrix method is used to validate the fake b simulation. This brings a high confidence on the background estimation in SR1b and SR3b, where events with more than 2 b jets are present. In the absence of any significant excess above the Standard Model expectation, the results in the SRs are used to set upper limits on BSM physics. To interpret them a total of 14 SUSY models and one mUED model are used. The exclusion limits are placed using a binned fit performed simultaneously in the five signal regions. At the time of publication, these results put new or significantly improved limits on the SUSY particles masses and considered parameters. This is valid in regions of the SUSY phase space where the R-parity can be conserved or not, where the lightest particle can be an squark or gluino, and where the mass difference between the SUSY particles can be large or compressed. When comparing to the analysis performed by the CMS collaboration in the channel with two same - sign leptons, a similar sensitivity is observed. In ATLAS this signature is intensively studied and at the end of $\sqrt{s} = 8$ TeV data analysis no significant excesses (above 3σ) was observed.

Generally, in ATLAS all the Run-1 data was deeply analyzed looking for any potential BSM signal. The achieved results are pointing to an unprecedented coverage of a very well motivated region of the SUSY (but not only) phase space. In particular, the studied channel plays a key role in the search for gluino or third generation quarks. This is mainly because a high sensitivity can be achieved in the compressed spectra, as softer cuts on kinematic variables can be imposed due to low SM background expectation.

Chapter 8

First results with 13 TeV data

The LHC Run-2 collisions started in June 2015 at the unprecedented energy in the center of mass of 13 TeV. This high energy leads to a large increase in the production cross sections, improving the discovery potential for beyond the Standard Model physics like Supersymmetry, Dark Matter, black holes, etc. It ensures also the observation and the study of rare processes like $t\bar{t}$+H or $t\bar{t}$+V, and higher precision Standard Model measurements. Beside the changes in the collision energy, several upgrades were brought to the LHC and detector machines, and improvements to the online and offline reconstruction and analysis software. Some of these improvements are shortly mentioned in this chapter.

The time is great to search for strong production of gluinos and squarks with same - sign leptons (and not only) at the weak-scale. Thanks to the new beam energy, the expected gluino production cross section increases by a factor 30 to 130 in the mass range of 1.5 - 2 TeV and the sensitivity of the analysis presented in Chapter 7 is well above the Run-1 results when performing projections for a luminosity L of 3-4 fb^{-1} [282], as expected by the end of 2015. Given the tight limits set on the sparticles masses at the end of Run-1, Supersymmetry is expected to be just at the corner and performing a first analysis with the first 13 TeV data is very exciting. Thus I choose to close the series of SUSY chapters of my thesis with the results obtained with the very first 407 pb^{-1} of Run-2 data, available in due-time before finishing the manuscript.

In the Section 8.1 of this chapter the improved LHC and ATLAS detector are shortly presented. Section 8.2 is dedicated to the physics gain when searching for Supersymmetry, with the new beam energy. The search for supersymmetry in events with two same - sign leptons is presented in Section 8.3. Here the inputs of the analysis, like the used data and MC samples or the object definition, are presented. The (still preliminary) optimized signal regions used at beginning of Run-2 are defined in Section 8.3.2, and the sensitivity reached with 3 fb^{-1} is also shown. In Section 8.3.3 the main background estimation and validation in Run-2 are shortly described. Finally the chapter ends with some proposals for future improvements of this analysis (Section 8.4) and with few conclusions (Section 8.5).

8.1 LHC and ATLAS in Run-2

During the two-years Long-Shut-Down-1 (LS1) the LHC machine was highly improved to target an energy of 6.5 TeV per proton beam and a bunch spacing of 25 ns. The main consolidations between 2013 and 2014 are illustrated in Figure 8.1 [283]. The 2015 Run-2 strategy started with the initial beam commissioning in April, for a period of 2 months, with an excellent and improved system performance. It refers to the beam instrumentation and collimation, injection and beam dump systems, vacuum and magnetic systems and the machine protection. The software and the analysis tools were also highly improved. Starting with April 5^{th} the first stable beams at 6.5 TeV were circulated in the LHC ring. Just after, a period of scrubbing, or of electron bombardment of the beam surface, was mandatory to highly reduce the electron cloud around the beams. Finally, on June 3^{rd} the LHC started the first collisions at $\sqrt{s} = 13$ TeV. The LHC Run-2 physics proton-proton period started with around 30 days of collisions with 50 ns bunch spacing, and ATLAS collected an integrated luminosity ~0.1 fb^{-1}. Other 70 days with 25 ns bunch spacing will follow and the expected integrated luminosity to be collected is estimated at ~3-4 fb^{-1} [284]. The transition between 25 ns to 50 ns happened at the end of August.

The ATLAS detector was also consolidated and several upgrades were performed during the LS1. It has a new beam-pipe and improved magnet and cryogenic systems. Compared to Run-1, the muon chambers were completed in the [1.1, 1.3] η region and, when needed, repaired. The dead pixel modules and the calorimeter electronics were also reconditioned. To improve the tracking and the vertex reconstruction performance at high luminosities, a fourth pixel layer or the "insertable B-layer" (IBL) was added. It is located just near the beam pipe, at a distance $R = 3.3$ cm, and presents a 3D silicon pixel technology. The

Figure 8.1: The main 2013-2014 LHC consolidations toward Run-2 data taking [283].

purpose of this layer is to ensure a high quality tracking performance and an improved b-tagging at high luminosities, as expected for the LHC Run-2 [285]. It is fully operational, and the significant gain with respect to Run-1 in the impact parameters resolution is illustrated in Figure 8.2. For the b-tagging, it is expected to have at least 4 times better rejection of the light flavor jets, highly increasing the sensitivity to BSM physics for the SUSY search with two same - sign leptons.

Overall, the ATLAS detector status shows more than 97% of channels working in each system, with significant improvements with respect to the LHC Run-2 in the Pixel sub-detector, Tile calorimeter, etc [286]. The total integrated luminosity recorded until 1^{st} of October during the stable beams of pp collisions at 13 TeV is around 1.31 fb^{-1}, as illustrated in Figure 8.3 (left) [287]. The integrated luminosity used for the physics studies presented in this chapter is 400 pb^{-1}. The mean number of interactions per crossing in the 25 ns and 50 ns bunch spacing data is also shown for reference in Figure 8.3, right. Almost the same number of simultaneous interactions as in the $\sqrt{s} = 8$ TeV data is observed (Figure 3.4, Chapter 3).

8.1.1 Performance of the ATLAS detector

The obtained performance of the object reconstruction and identification in ATLAS are shortly summarized in the following. For the anti-k$_T$ jets (R = 0.4), the uncertainties on the energy scale and calibration are estimated to vary from 6% when $p_T = 20$ GeV, decreasing to 1% for jets with p_T in the [200, 1800] GeV interval and increasing to 3% for more energetic jets [289]. The performance and the reconstruction of the $\not{E}_T$ with the first 13 TeV was measured in minimum bias and $Z \to \mu\mu$ events and show a reasonable agreement when compared to simulations [290].

To evaluate the electron energy scale with the first 13 TeV data, J/Ψ and Z events were selected and overall the obtained resolution was well understood (no public documentation yet). The electron identification efficiency measurements in the central region of the detector are obtained for the *loose*, *medium* and *tight* likelihood based criteria and are around 95%, 90% and 85% when integrating over η and p_T. The obtained scale factors are close to unity and the associated total uncertainties vary between 3% and 5%. No strong pile-up dependency is observed [291].

The medium muons reconstruction efficiency vary between 75% and 99%, and it is well modeled by simulations. The associated uncertainties are below 1% for muons with $p_T > 20$ GeV. The muon momentum

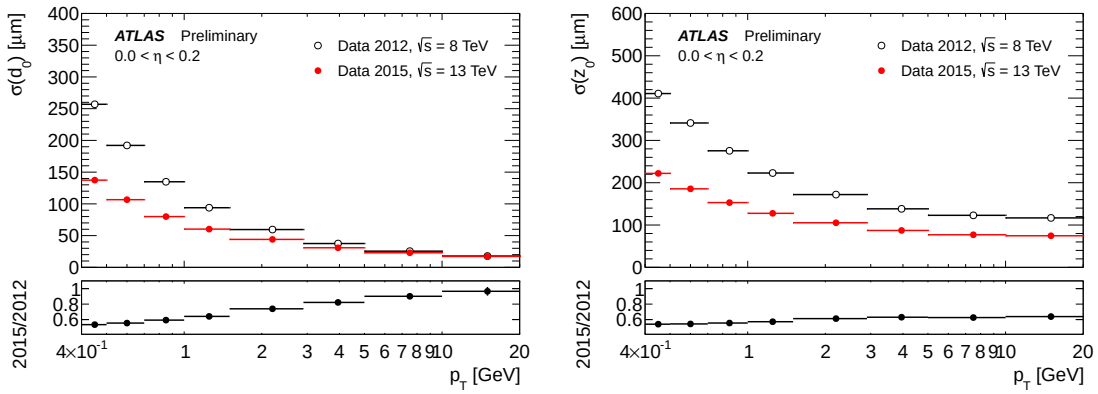


Figure 8.2: The $\sigma(d_0)$ (left) and $\sigma(z_0)$ (right) impact parameters resolution in $\sqrt{s} = 13$ TeV, with the IBL included as a function of p_T [288]. The resolution obtained in 2012 data is also shown.

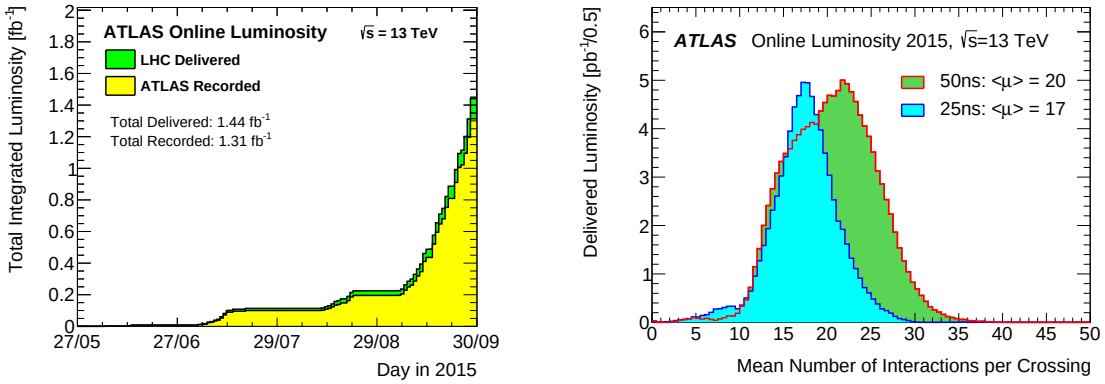


Figure 8.3: Right plot : cumulative (green) and recorded (yellow) ATLAS luminosity during Run-2 LHC, in 2015. Left plot : mean number of interactions per crossing in the 25 ns (cyan) and 50 ns (green) bunch spacing data [287].

is measured with a combined track information from the inner detector and muon spectrometers, the first dominating in the low p_T region. The associated resolution is similar to the Run-1 performance, in the limit of the associated uncertainties. The precision reached for the momentum scale is already very good, around 0.2%, as well as the precision on the muon resolution, which is within 5% in the $p_T < 100$ GeV range [292].

8.2 Gain in Supersymmetry at $\sqrt{s} = 13$ TeV

With the LHC Run-2 energy the production cross sections highly increased for all supersymmetric particles as illustrated in Figure 8.4 (right), greatly increasing the chances to find SUSY. The prime candidates remain the gluinos and the squarks. For completeness, the ratios of the LHC parton luminosities at 13 TeV and 8 TeV as a function of the SUSY particle mass are shown in the left plot : an increase up to a factor 2 is observed in the low mass region and up to a factor 8 in the 1 TeV range, for gluino pair production processes.

The prospect studies performed for the strong production SUSY search in the same - sign channel are showing that the analysis has an improved discovery potential even with $\sim$3 fb^{-1} [282]. Some of the most recent results are summarized in this chapter and documented in [293].

8.3 Search for Supersymmetry in the same - sign leptons channel

To search for Supersymmetry in events with same - sign leptons and jets with the 13 TeV data, the analysis presented in Chapter 7 is fully optimized in the view of a fast publication to account for the high increase in the collision energy, and further in the signal and background production cross-sections. For the signal regions optimization same models as in Run-1 are used, targeting mainly scenarios predicting gluinos and third squark generation, as highly motivated by Natural SUSY theoretical framework.

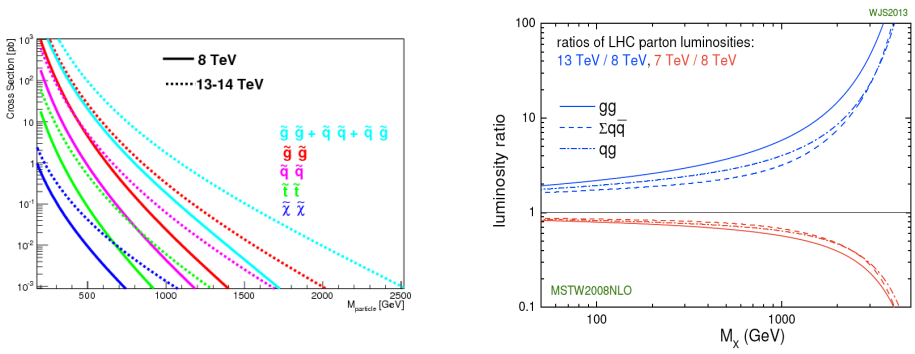


Figure 8.4: Left plot : SUSY particles production cross section at 8 TeV (bold line) and 13 TeV [294]. Right plot : Ratios of LHC parton luminosities [295].

8.3.1 Inputs for the analysis

In this section the inputs needed for this version of the analysis are presented. It starts with brief introduction of the amount of collected Data and of the signal and background Monte Carlo samples. The differences in the object definitions and the reached improvements are also summarized. The section ends with a short description of the methodology used to perform the lepton truth matching and with the event selection.

Data and Monte Carlo samples

The Data used in this analysis is collected in pp collisions at an energy of 13 TeV in the center of mass during the beginning of 2015 data-taking. It corresponds to period C ($L = 78.2$ pb^{-1}, 50 ns bunch spacing), D and E ($L = 322.4$ pb^{-1}, 25 ns bunch spacing). Period A is not included. The uncertainty on the integrated luminosity is ± 9 fb^{-1}.

The signal samples are generated using MADGRAPH interfaced with Pythia, with LO matrix elements with up to two extra partons. All the signal MC samples were generated using a 25 ns bunch spacing configuration. The background samples were obtained with the same generators as in LHC Run-1. They all have 50 ns bunch spacing configuration, although the differences between 25 ns and 50 ns bunch spacing were found to be very small for the relevant distribution in the analysis. For some prospect studies presented here the MC was normalized to a luminosity of 3 fb^{-1}, otherwise to the luminosity corresponding to the accumulated data.

Object definition

All jets are required to have $p_T > 20$ GeV and the pseudorapidity is restrained to $|\eta| < 2.8$. They are required to be well separated by electron and muon objects. For the b-jets, the p_T threshold remains 20 GeV but the acceptance is in the range $|\eta| < 2.5$. For the b-tagging an improved algorithm is used, the so called `MV2c20`. It is based on a neural network using the output weights of the `JetFitter`+`IP3D`, `IP3D` and `SV1` algorithms as input. The working point is set to 70%, as favored during the optimization studies. Unlike in Run-1, a cut on the jet vertex tagger (JVT[1]) is applied, rejecting pile-up jets with $p_T < 50$ GeV, $|\eta| < 2.4$ and JVF < 0.64, after the overlap removal. The gain in stability with respect to pile-up after applying this cut is illustrated in Figure 8.5.

Generally, when compared to Run-1, the pile-up rejection is much better in Run-2. This ensures less pile-up jets in the signal regions (smaller tails), and a higher sensitivity to find BSM signal. This is illustrated in Figure 8.6 right where the number of jets with $p_T > 25$ GeV distributions are shown – to be compared with Figure 7.11 (Chapter 7) for Run-1. As it was mentioned in the introduction, another great improvement in Run-2 is the b-tagging methodology which ensures a much higher light jets rejection. For this analysis, it translates again into a higher signal sensitivity and a lower b-jet fake rate when compared to Run-1 results. Several distributions of number of b-jets are shown in Figure 8.6 left.

All baseline electrons should satisfy *loose* LLH identification criteria and have $p_T > 10$ GeV and $|\eta| < 2$; the crack region ($1.37 < |\eta| < 1.52$) is excluded. The reduced electron acceptance is motivated by the fact that the largest contributions from charge-flip and fake lepton background are in these regions, due to poorer detector resolution or increased amount of material. Dedicated studies are showing that reducing the acceptance to $|\eta| < 2.0$ would reduce the contribution from processes with prompt SS leptons by 6.5% or less, but the reduction in non-prompt (11-21%) and charge flip electrons (30- 37%) is quite larger. Reducing

[1] A similar variable to the jet vertex fraction, having a much higher performance as requested for the Run-2 conditions.

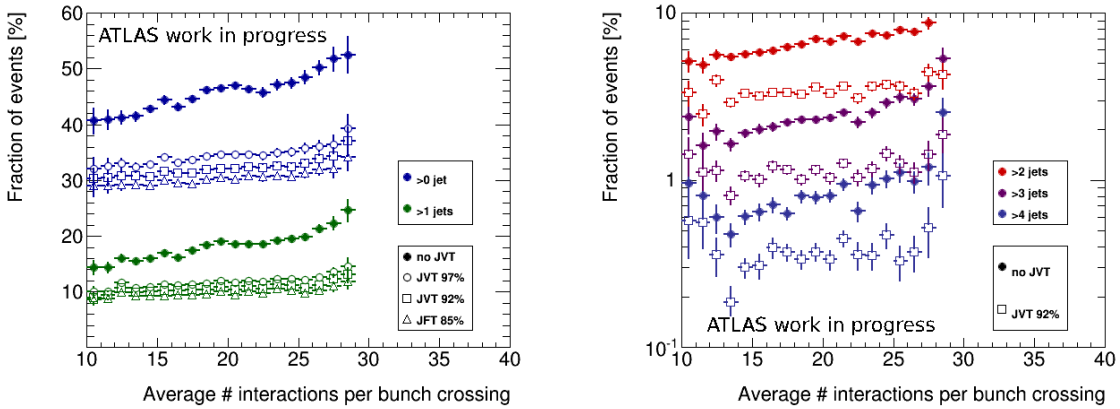


Figure 8.5: Fraction of events [%] in data with at least 1 or 2 jets (left) and with at least 3, 4 or 5 jets (right) with respect to average number of interactions per bunch crossing with and without a cut on the JVT. Event selection requires a pair of leptons, and all jets have $p_T > 25$ GeV. $L = 159$ pb^{-1} and $\sqrt{s} = 13$ TeV.

the acceptance more severely for the two leading electrons to $|\eta| < 1.37$ would reduce the non-prompt and charge flip components by 28-44% and ∼69%, respectively, with a reduction of only 13-18% for prompt SS processes. The latter cut is applied only in some of the defined validation regions. A requirement on the transverse impact parameter of $|d_0/\sigma(d_0)| < 5$ is also applied to baseline electrons to reduce the contribution from charge mis-identification. The signal leptons are additionally required to pass the *tight* LLH identification, a minimum cut on the longitudinal impact parameter ($|z_0 \times \sin(\theta)| < 0.4$ mm) and relative cuts on the track and calorimeter isolations ($p_T^{\text{cone20}}/p_T < 0.06$ and $E_T^{\text{cone20}}/p_T < 0.06$). The latter cut ensures a high reduction of the fake electron background.

In Run-2 the muons were reconstructed with the MUON chain (or the so called *third chain*). All baseline muons must pass *medium* selection working point, they must have $p_T > 10$ GeV and $|\eta| < 2.5$. Any event containing a cosmic or a *bad* muon is vetoed. The signal leptons must also satisfy the standard impact parameter cuts ($|z_0 \times \sin(\theta)| < 0.4$ mm and $|d_0/\sigma(d_0)| < 3$) and relative track and calorimeter isolation cuts ($p_T^{\text{cone30}}/p_T < 0.06$), ensuring a high fake muon background rejection.

Lepton truth matching

In several studies presented in this chapter, a lepton truth matching strategy is applied to distinguish the lepton originating from prompt decays of gauge boson or SUSY particles from non-prompt leptons originating from semileptonic b-decays, photon conversions or fakes. This strategy largely relies on the information from the `MCTruthClassifier` tool, although it may not be sufficient to fully classify electrons originating from photon conversions, in which case a matching with truth electrons within a cone of $\Delta R < 0.1$ is also used. This truth matching information is also necessary to filter out events with fake leptons or charge flip electrons from the MC samples used either to estimate the non-reducible background or to predict SUSY signal yields, to avoid double-counting with the data-driven background estimates.

Event selection

The even selection is inclusive in terms of leptons ("at least two same-sign leptons", see Chapter 6). It is found that for the early results no substantial gain would be achieved by considering trilepton final states separately as was done in the final Run-1 analysis.

8.3.2 Signal region definition

The definitions of the signal regions have been studied to provide an optimal performance for $\sqrt{s} = 13$ TeV collisions and a low integrated luminosity (2-4 fb^{-1}). We chose to categorize the signal regions based on their b-jet multiplicity, in continuation of the approach sustained in the Run-1 analysis. Their definition is shown in Table 8.1. The SR0b selections were subdivided into two overlapping selections ($\geq$3 or $\geq$5 jets, also denoted as SR0b5j and SR0b3j) to cover various signal scenarios that lead to differently jet-enriched final states.

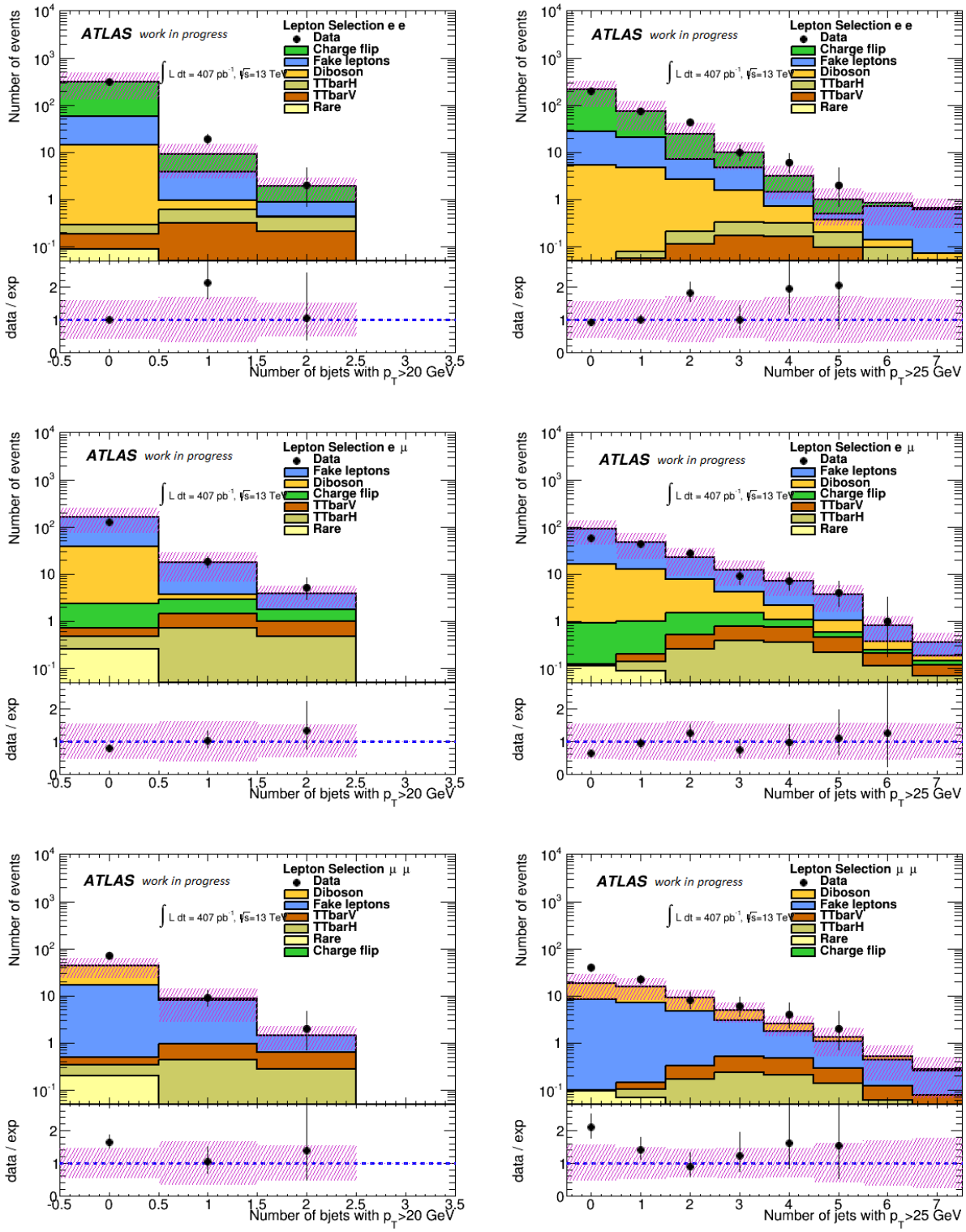


Figure 8.6: Background validation : number of signal b jets with $p_T > 20$ GeV (left) and jets with $p_T > 25$ GeV (right) distributions in ee (top), $e\mu$ (middle) and $\mu\mu$ channels. Only the statistical uncertainties on the background estimation are considered (hashed magenta area). The black lines on the data points represent the associated statistical uncertainty. Data corresponds to Ł = 407 pb^{-1} and $\sqrt{s} = 13$ TeV, and only a small overlap with the signal regions is present.

Signal region	N_{lept}	N^{20}_{b-jets}	Other variables
SR3b	$\geq$2	$\geq$3	-
SR1b	$\geq$2	$\geq$1	$N^{50}_{jets} \geq 4$, $\not{E}_T > 150$ GeV, $m_{\text{eff}} > 500$ GeV
SR0b5j	$\geq$2	==0	$N^{50}_{jets} \geq 5$, $\not{E}_T > 100$ GeV, $m_{\text{eff}} > 400$ GeV
SR0b3j	$\geq$2	==0	$N^{40}_{jets} \geq 3$, $\not{E}_T > 200$ GeV, $m_{\text{eff}} > 400$ GeV

Table 8.1: Definition of the signal regions for the early two same - sign leptons analysis. For all signal regions, two same - sign or three leptons with $p_T >$ [20,20] GeV are required.

Sensitivity of the signal regions

The sensitivity of the proposed signal regions is evaluated using the `HistFitter` framework to perform discovery hypothesis tests for the different signal scenarios of interest. The event yields assumed in these tests are the ones predicted by the Monte-Carlo. All the results in this section are presented for an integrated luminosity of 3 fb^{-1}, as expected for the first publication.

The fit configuration is setup in the exclusion mode, i.e. to perform hypothesis testing with a known BSM signal. Of course here the rejected hypothesis is the background-only hypothesis. Only one signal region is fitted at the time; furthermore, there is no control region defined in the analysis. Distinct global systematic uncertainties are assigned to each category of background, amounting to 30% for $t\bar{t}$+ V processes, 35% for diboson, 100% for minor processes such as $t\bar{t}$+H and 50% for the reducible background ($t\bar{t}$, V + jets, $W^{\pm}W^{\pm}$).

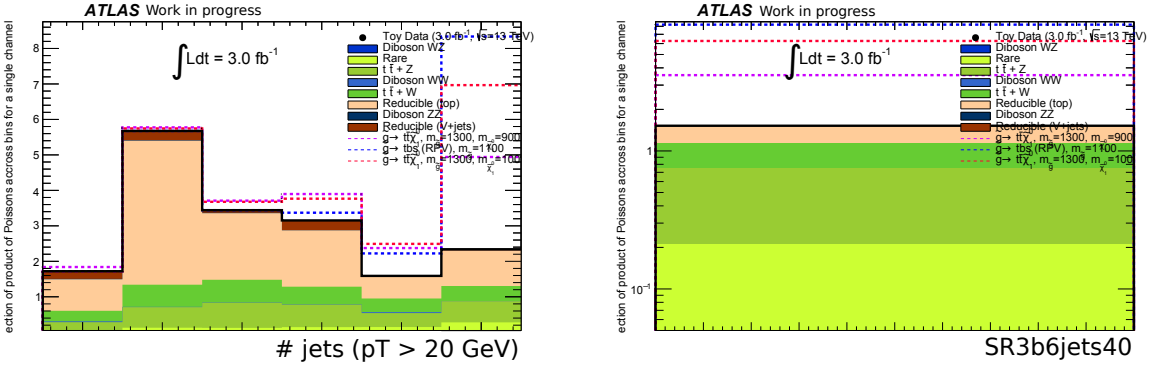


Figure 8.7: Expected yields for Standard Model processes and a few characteristic signal scenarios, in selections with at least two same-sign leptons and three b-jets. The selection on the right side further requires at least six jets with $p_T > 40$ GeV. MC, Ł = 3 fb^{-1}.

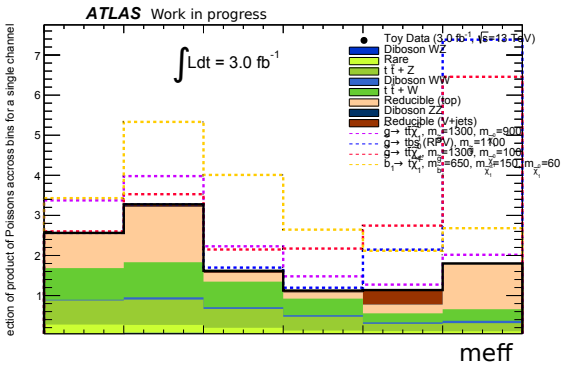
Background and signal yields

Figures 8.7 and 8.8 present the event yields of the background and a few characteristic signal scenarios, in various signal region candidates. The regions that have enough statistics to allow it are binned in the most relevant variable to discriminate between signal and background, either the effective mass or the number of jets (the latter for the 3 b-jets selection). Binnings are chosen such that the background yield in one bin remains at the level of around one event. One should note that they correspond to the binning used when performing shape fits as described in the next paragraph.

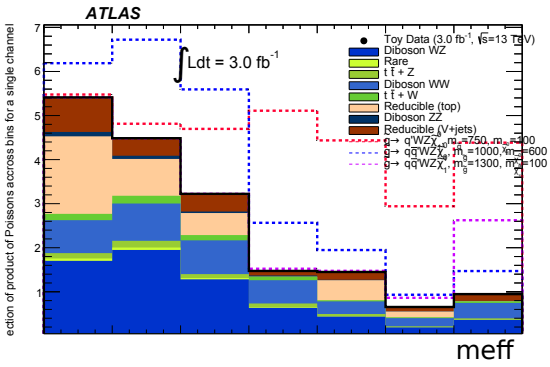
Sensitivity for the gluino stop offshell model

Figure 8.9 presents the signal significance obtained with several signal region candidates, for the gluino offshell stop model. For some of the regions, a shape fit of the effective mass or jet multiplicity distributions is performed. While it can provide additional benefit over simple cuts, the main reason for its use here was rather simply to avoid tuning cuts for each signal point. One can see indeed on the figure that for each mass, the significance obtained with an effective mass shape fit is rather close to the best of the significances computed with different choices of fixed cuts. As expected, at large neutralino mass signal regions based on ≥ 3 b-jets perform better than the ≥ 1 b-jet signal regions. There is not much difference in significance between the $\geq 3b$ selection requiring 6 hard jets and the one based on softer jets, possibly due to the fact that the mass spectra of the available grid points are not very compressed. One would probably see a larger difference for points lying on the diagonal $m_{\tilde{g}} = 2m_t + m_{\tilde{\chi}^0_1}$.

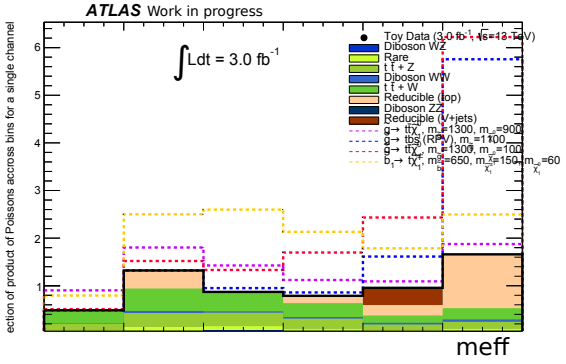
(a) $\not{E}_T > 100$, ≥ 4 jets ($p_T > 50$)



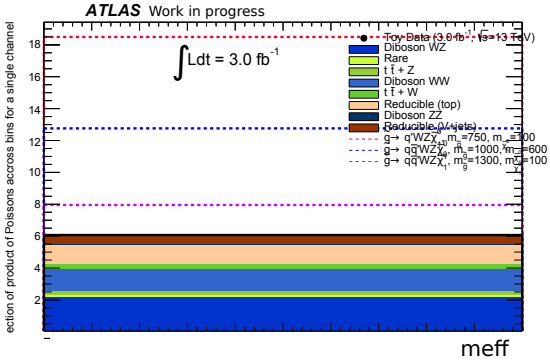
(b) $\not{E}_T > 100$, ≥ 3 jets ($p_T > 40$)



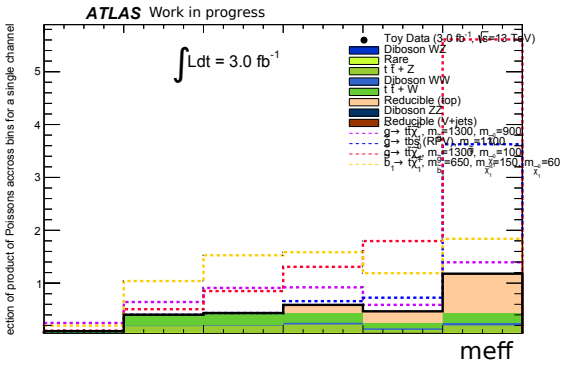
(c) $\not{E}_T > 150$, ≥ 4 jets ($p_T > 50$



(d) $\not{E}_T > 100$, ≥ 4 jets ($p_T > 40$)



(e) $\not{E}_T > 200$, ≥ 4 jets ($p_T > 50$



(f) $\not{E}_T > 100$, ≥ 5 jets ($p_T > 40$)

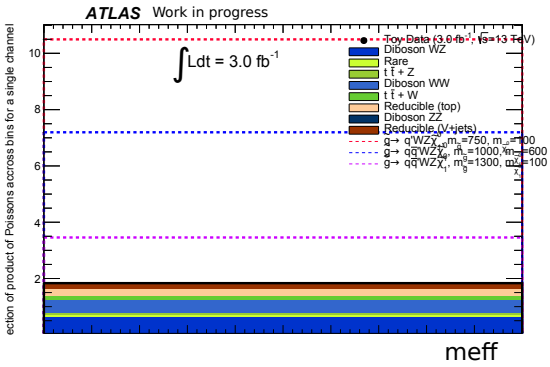


Figure 8.8: Expected yields for Standard Model processes and a few characteristic signal scenarios, in selections with at least two same-sign leptons, at least one (left) or no b-jet (right), and the other cuts that define each of these signal region candidates. MC, $\mathrm{L} = 3$ fb^{-1}.

As only few signal samples have been generated, the projections Figure 8.9 (right) provides an estimate of the evolution of the sensitivity in a finer way, by simply using the kinematic distributions and yields obtained in a reference sample (gluino mass 1300 GeV, neutralino mass 100 GeV) weighted by the ratio of inclusive cross-sections between the mass to be extrapolated to, and the reference. It has been checked that the significance obtained from the next available sample (at a gluino mass of 1500 GeV) matches rather closely the extrapolated value.

One can conclude from these plots that at low neutralino mass and with such as low integrated luminosity, there is no sensitivity beyond the Run-1 exclusion. At higher neutralino mass however, the analysis would be sensitive to an area not yet excluded.

Sensitivity for the direct sbottom model

Figure 8.10 (left) shows the signal significance for several sbottom masses and a light neutralino. Here, the advantage of keeping a moderate $\not{E}_T$ cut in the definition of the $\geq 1b$ signal region is clear. One can see that 2σ sensitivity up to masses of 650 GeV might be expected, which is quite a nice improvement over the Run-1 reach (450 GeV, although the expected limit was 50 GeV higher).

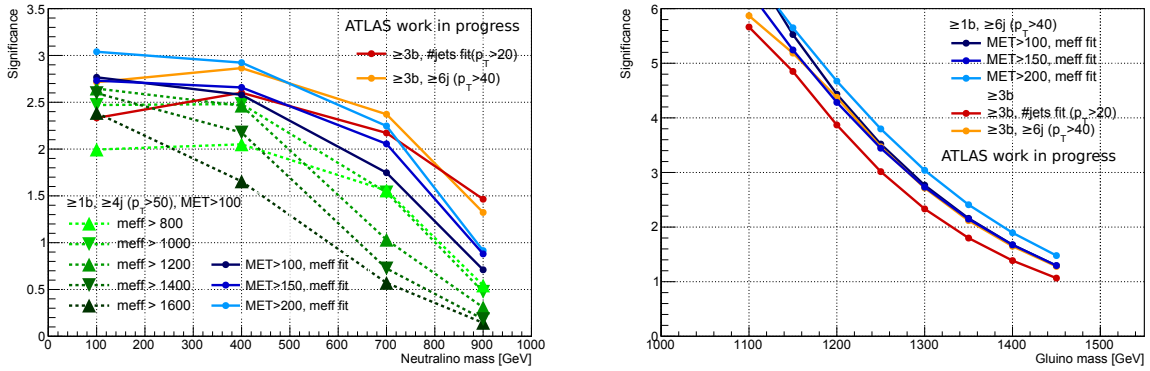


Figure 8.9: Expected signal significance for 3 fb^{-1} for the gluino offshell stop model $\tilde{g} \to t\bar{t}\tilde{\chi}_1^0$, varying the neutralino mass with a fixed gluino mass of 1300 GeV (left) or extrapolating to different gluino masses from a reference point with gluino and neutralino masses of 1300 and 100 GeV respectively (right). MC, Ł = 3 fb^{-1}.

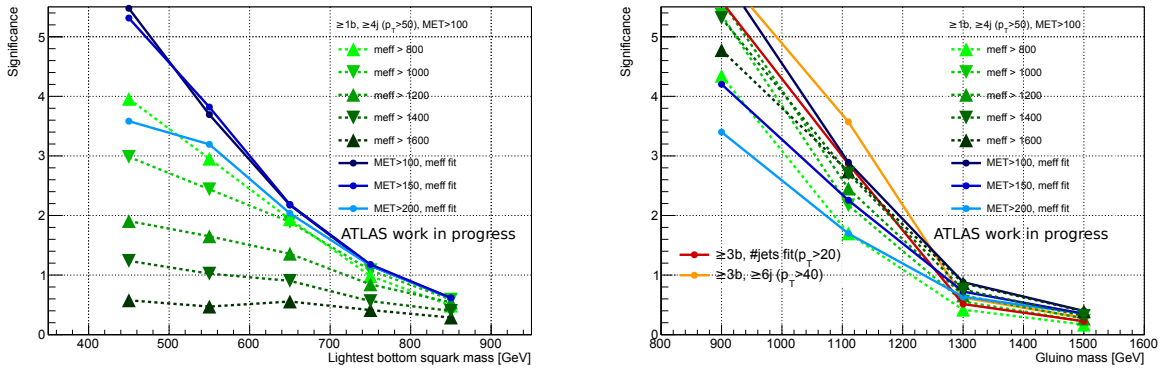


Figure 8.10: Expected signal significance for 3 fb^{-1} for the direct sbottom $\tilde{b} \to t\chi^{\pm}$ (left) or the gluino RPV $\tilde{g} \to tbs$ (right) models. The neutralino mass is fixed to 100 GeV in the sbottom scenario. MC, Ł = 3 fb^{-1}.

Sensitivity for the gluino model with RPV stop decay

The expected sensitivity for this model can be found on Fig. 8.10 (right). The most powerful signal region for this model is as expected the selection with at least 3 b-jets and 6 hard jets. One can notice however that the signal regions with moderate $\not{E}_T > 100$ GeV cut are quite sensitive as well : the absence of neutralinos is compensated by the presence of boosted tops decaying leptonically.

Sensitivity for the gluino and squark models with 2-step decays

Finally, the performances of signal region definitions vetoing b-jets are probed with the gluino and squarks models with 2-step decays in gauge bosons. The signal significances for varied masses of gluino/squark or neutralino are provided on Fig 8.11. The signal region requiring 5 hard jets appears as the best compromise for both models, and in different regions of the phase space. For the gluino model, not much sensitivity beyong the Run-1 exclusion is gained at high gluino mass; at high neutralino mass, however, a significant non-excluded region up to neutralino masses of 650 GeV seems to be at reach. For the squark model, on the other hand, one will be able to probe a new region of the phase space at high squark mass, up to 900 GeV, which was not excluded in Run-1.

8.3.3 Background estimation and validation

Generally the same background estimation strategy as in Run-1, and detailed in Chapter 7, is employed and in this section only the main results are shortly presented. The background validation in several regions near the defined signal regions is shown at the end of the section.

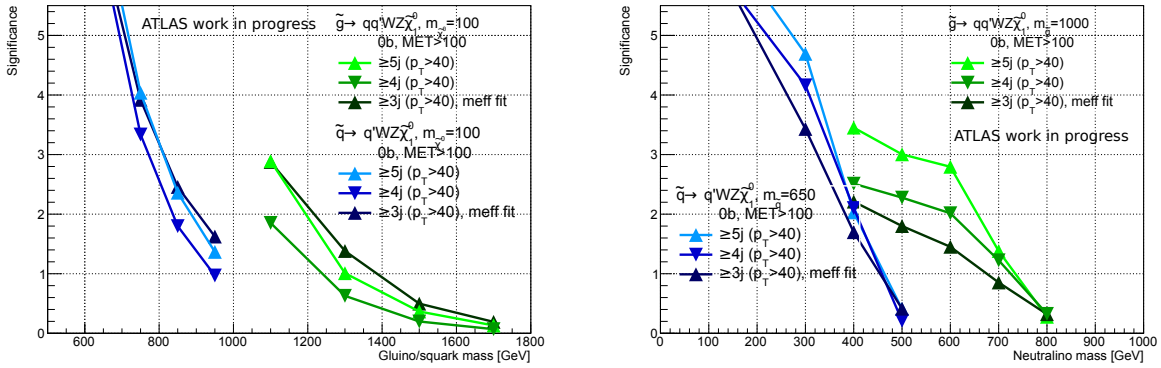


Figure 8.11: Expected signal significance for 3 fb^{-1} for the gluino and squark 2-step decay models $\tilde{g}/\tilde{q} \to q(q')WZ\tilde{\chi}_1^0$, varying the gluino or squark masses for a fixed neutralino mass of 100 GeV (left), or varying the neutralino mass for fixed gluino and squark masses of respectively 1000 and 650 GeV. MC, Ł = 3 fb^{-1}.

Prompt SS processes

The various prompt SS Standard Model processes to the signal regions are estimated by relying on the Monte-Carlo predictions, normalized with the best known theoretical cross-sections: these processes are too rare to allow use of control regions until a significant integrated luminosity will be collected. For example, in [296] (and in Chapter 9) which establishes projections of the sensitivity to the ttH signal strength in a final state very similar to this analysis, the impact of adding control regions for $t\bar{t}$+V processes was found to be favorable only after collecting >100 fb^{-1} of data.

Electron charge flip

The electron charge flip rates needed to re-weight the OS data events are measured with the *Tag&Probe* presented in Chapter 6. The results obtained in data are show in Table 8.2. Compared to Run-1, the rates are higher in the end-cap region of the calorimeter, at low p_T (<30 GeV). Dedicated studies are pointing that the main difference is coming from the $|d_0/\sigma(d_0)|$ cut, which is computed differently than in Run-1 (with respect to the beam line, and not with the primary vertex position as in Run-1).

η / p_T	10 GeV < p_T < 15 GeV	15 GeV < p_T < 30 GeV	p_T > 30 GeV
$0 < \eta < 1.37$	0.0035 ± 0.0043	0.0007 ± 0.0002	0.0004±0.0001
$1.37 < \eta < 2.$	0.0044 ± 0.0009	0.0044 ± 0.0009	0.0063±0.0006

Table 8.2: Absolute electron charge flip rate measured in data. Only the absolute statistical uncertainty is shown.

Electron fake rate

As in Run-1 the electron fake rate is measured with the *Tag&Probe* method using $e\mu$ pairs where the muon is required to be the tag. To ensure this, it must have $p_T > 40$ GeV and satisfy the signal definition. A veto on the third lepton is applied to remove the contamination from diboson or $t\bar{t}$+V Standard Model processes.

Truth composition studies showed that the defined SR0b and SR1b signal regions are mainly dominated by non-prompt electrons from heavy flavor decays (i.e. bottom and charmed mesons). A control region with a similar composition can be defined by asking at least one b-jet in the event. Unlike in Run-1, only this region is used to measure the electron fake rate which will be equally applied in regions with and without b-jets. The results in data are showing that the fake rate is 14% ± 6 (statistical uncertainty), generally similar to Run-1 values. The statistics is not high, and therefore only an inclusive measurement is performed. The p_T, η, number of jets, etc. dependencies will be considered when more data will be available.

Muon fake rate

Same-sign dimuon pairs are used for the measurement in data, with no further requirements given the low available statistics. To avoid any bias in the measurement both muons are considered alternatively as

tag. The tag muon definition is the same as for the electron fake efficiency measurement. The measured muon fake rate is 15% ± 8 (statistical uncertainty), as in Run-1.

Background validation

As one of the most crucial tasks at beginning of Run-2 in this analysis is the correct fake lepton background estimation, its validation is essential. Thus, as in Run-1, the agreement between the data observation and background estimation is checked in several validation regions with similar kinematics as in the SRs. For illustration, the effective mass and missing momentum distributions are shown after the lepton selection and at least 0, 1, 2 or 3 jets in the event in Figure 8.12 - 8.13. If no jet is required in the event, the distribution is dominated by charge flip electrons, otherwise by fake leptons. Generally a good agreement is observed, even if several rates needed for the background estimation are measured in an inclusive way.

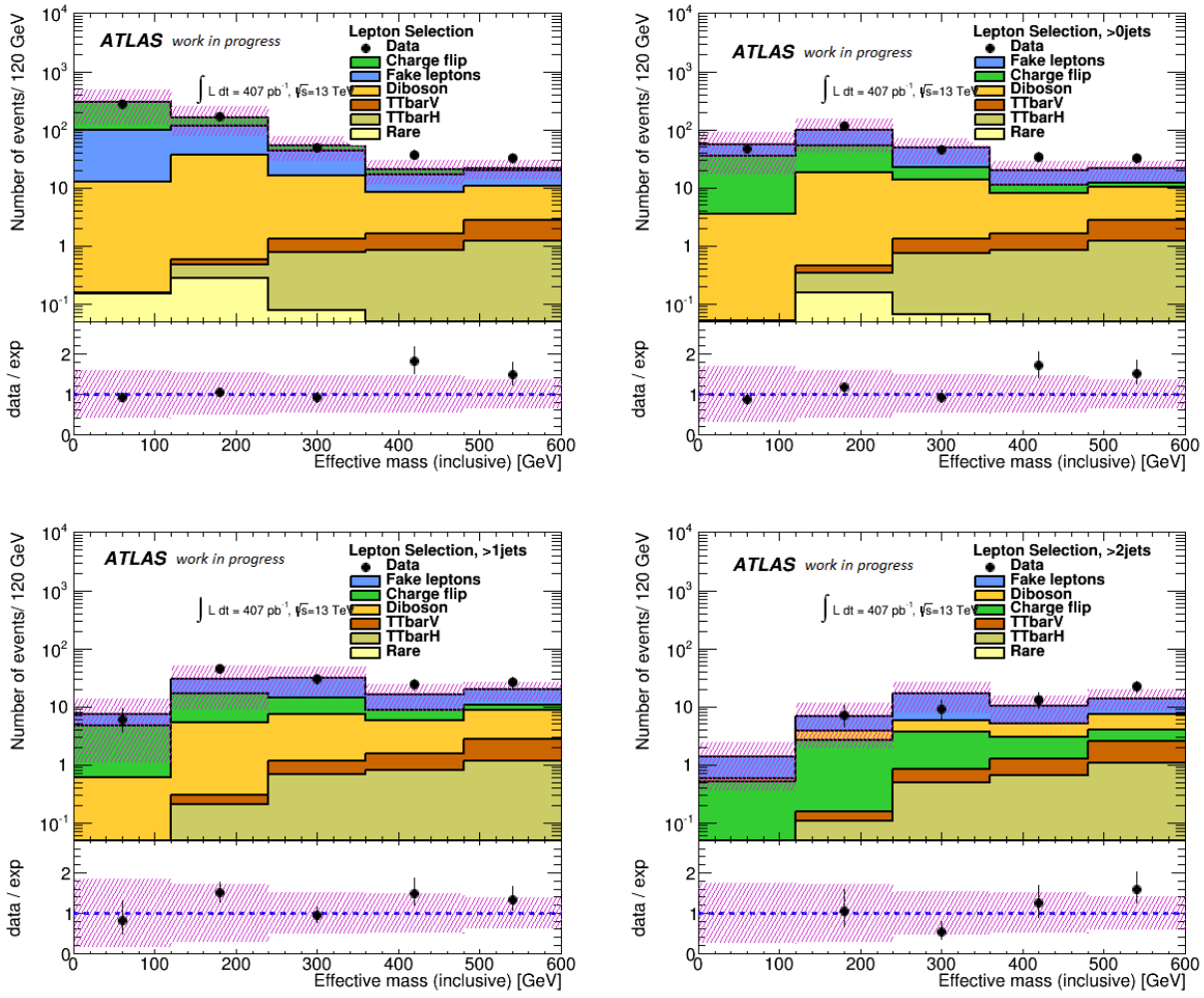


Figure 8.12: $ee + e\mu + \mu\mu$ channel: Effective mass distribution after lepton selections with at least 0, 1, 2 and 3 jets with $p_T > 40$ GeV (from top-left to bottom-right). Only the statistical uncertainties on the background estimates are shown (magenta hashed area). The black lines on the data points represent the associated statistical uncertainty.

Background estimation in the signal regions

Currently the methodology used to assign all the sources of systematic uncertainties on the background estimates is work in progress and not yet finalized. Moreover, until the methods used to estimate the background are not completely validated with more data – to ensure a higher statistics to perform the validation per leptonic channel and with at least 1 b jet or exactly 0 b jets in the event – we cannot look at the data observation in the signal regions due to ATLAS blinding policy. Also, with only 400 pb^{-1} of data, we expect very few background events in the signal regions – this is suggested also by the validation plots shown in Figures 8.12 - 8.13. Given all these reasons, no results in the signal regions are presented.

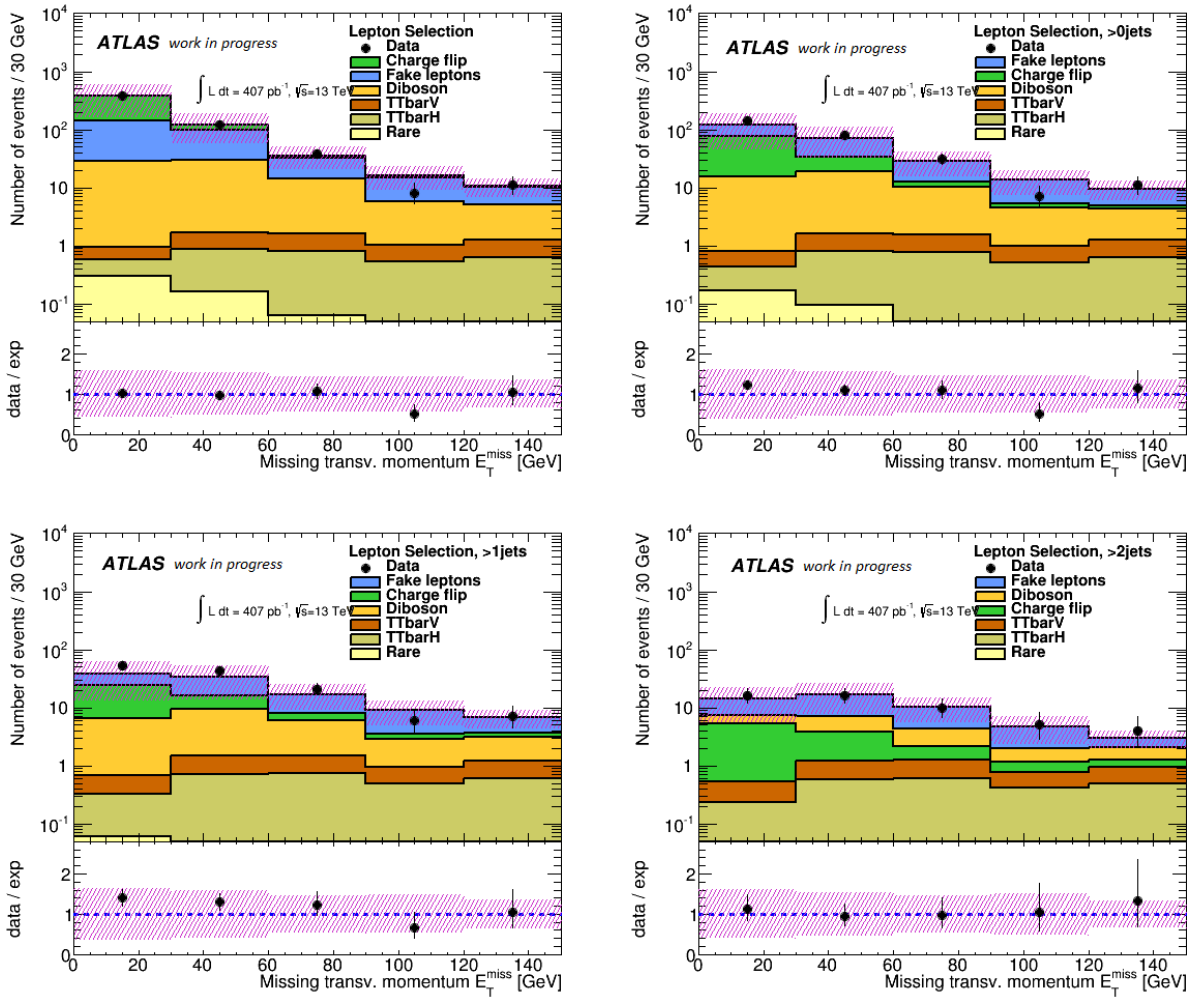


Figure 8.13: $ee + e\mu + \mu\mu$ channel: Missing transverse energy distribution after lepton selections with at least 0, 1, 2 and 3 jets with $p_T > 40$ GeV (from top-left to bottom-right). Only the statistical uncertainties on the background estimates are shown (magenta hashed area). The black lines on the data points represent the associated statistical uncertainty.

8.4 Future improvements for the analysis

The large increase in reachable gluino mass (up to 2-3 TeV) in Run-2 brings new experimental challenges. In particular, by energy conservation, the large mass difference between the initial particles (gluinos) and the final state particles (top quarks and DM candidates) implies that the rest of the available energy should be used to provide a high momentum to the particles present in the final state. Therefore, the top quarks are more and more energetic (boosted) and their products (the W bosons and the b quarks) will become increasingly collimated (Figure 8.14). In this energy regime, the usual techniques used to identify the top quarks become inefficient. As an example, if the W boson decays into a pair of jets, the final state will be formed by three superimposed jets that cannot be reconstructed separately. In the second case, if the W boson decays into a pair of leptons, the charged lepton will be very close to the jet originating from the b quark (non-isolated) and very hard to reconstruct.

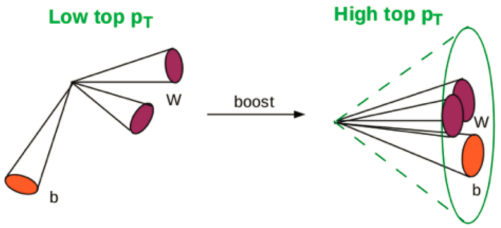


Figure 8.14: Illustration of boosted top quark [297].

This can be improved by using boosted top quark identification (or top-tagging) techniques, a new development in high energy physics experiments [298, 299]. Until now, the top-tagging was never used in any search with gluino-induced decay chains. It has been pioneered in the search for new massive resonances decaying into a pair of top quarks [299]. Compared to traditional techniques, this improved method reconstructs the hadronically decaying top quarks as a single jet, instead of three separate objects

and it leads to a high increase in the top quark reconstruction efficiency [299]. Generally the highest improvement in the discovery potential is observed for resonance masses above 1 TeV. The gluino-mediated stop production mode has a similar topology to the aforementioned analysis : the gluino mass is expected to be above 1. TeV and the decay mode is a pair of top quarks (and one neutralino). We thus expect that the application of the top tagging technique to the SUSY search presented in this thesis can lead to an increase in discovery potential.

Top-tagging techniques applied when the top quark decays leptonically has been little studied so far since the field of boosted top quark reconstruction has largely focused on hadronically decaying top quarks. A very recent preliminary study targeting gluino-stop-$\tilde{\chi}_1^0$ scenarios [300] suggests that the sensitivity can be increased up to 50% when using the top-tagging technique applied to top quarks decaying to muons. There are strong experimental reasons to believe that the discovery potential can be further increased by applying such techniques also to top quarks decaying to electrons. One of the main challenges in leptonic top tagging will be to improve the electron selection efficiency, as standard discriminants are not well suited to such non-isolated environments. Simply loosening the identification criteria would not suffice, as one wishes to keep a high rejection of other background sources. Therefore, it is strongly believed that all SUSY analyses searching for boosted particles, and in particular the SUSY analysis with two same - sign leptons and jets final state will highly benefit from using the bosted-top quark techniques.

8.5 Conclusions

With the restart of LHC operation in 2015 at the unprecedented energy of $\sqrt{s} = 13$ TeV in the center of mass, the chances to find SUSY, or any sign of Beyond the Standard Model physics, are greatly enlarged. The high increase in the LHC proton beam energy (from 8 TeV to 13 TeV) represents a very important jump in sensitivity to the production of heavy particles : for example, the gluino production cross-section will raise by a factor 30 to 130 in the mass range of 1.5 - 2 TeV. Thus, the coming years will therefore constitute a milestone in the direct searches for new physics and I am heavily involved in the preparation of the SUSY analysis in the same-sign leptons channel, in particular in the signal regions optimization studies and fake lepton background estimation. These results are expected to be published by the end of the year, and they are expected to correspond to a luminosity of 3-4 fb^{-1}.

Chapter 9

Top quark Yukawa coupling in the $t\bar{t}H$ channel with same - sign leptons

After the recent discovery of the Standard Model-like Higgs boson, precision Higgs boson studies started to be widely performed at LHC [9,10,301–303]. A special place is occupied by the top quark Yukawa coupling measurement which connects two of the heaviest Standard Model particles, the top quark (173 GeV) and the Higgs boson (125 GeV) : $\lambda_t = \frac{\sqrt{2}m_t}{v}$. Its value is predicted to be equal to unity and any deviation might give an insight on the possible existence and the energy scale of the new physics. In the absence of a deviation, one can constrain and even exclude BSM theories. A precise measurement can also lead to a better understanding of the origin of the Electroweak Symmetry Breaking in the Standard Model.

Given the heavy particles involved, it is not possible to constrain the top quark Yukawa coupling before the LHC Run-2, as shown in this chapter. Therefore, I perform prospects for measuring the top Yukawa coupling in Run-2 ($\sqrt{s}$ = 13 TeV) with an integrated luminosity of 100 fb^{-1}. The considered channel is the production of a Higgs boson in association with a pair of top quarks and the studied final state is formed by two same - sign leptons (2lSS) and several (b) jets. These studies are documented in [296]. The chapter starts with a brief introduction of the Standard Model-like Higgs boson properties (Section 9.1), emphasizing the latest mass, spin and parity measurements by the ATLAS collaboration. It also includes the status of the Higgs boson-SM vectors and fermions coupling strengths measurement after LHC Run-1. Section 9.2 is dedicated to the top quark Yukawa coupling results obtained by the ATLAS and CMS collaborations. Further, the prospects studies in the mentioned channel are discussed. Initially they are based on developments performed for the SUSY searches discussed in Chapters 6 and 7. However, an effort has been made to reduce as much as possible the difference with the published $t\bar{t}+H\rightarrow$ same - sign leptons search [304]. The considered data is formed by 8 TeV Monte Carlo samples re-weighted at $\sqrt{s}$ = 13 TeV (Section 9.3). The statistical interpretation is built on a simplified but realistic framework for systematics as detailed in Section 9.4. Before starting any optimization toward LHC Run-2, the analysis set-up is validated by comparing the signal sensitivity obtained with the 8 TeV signal regions defined in [304], with the public results. The output is presented in the same section. Section 9.5 is dedicated to prospect results. First, optimization is performed to define the signal, control and validation regions and the obtained definitions are summarized. The section continues with prospects as a function of luminosity, with and without considering control regions to constrain the SM background. Finally, possible improvements on the signal sensitivity are also presented. The chapter ends with Section 9.6 where some general conclusions are presented.

9.1 Higgs boson in the Standard Model

It was discussed in Section 1.4 that the mathematical tool used to generate mass to all elementary particles in the Standard Model predicts a massive Higgs boson. The understanding of this mechanism and the study of the Higgs boson properties are some of the prime objectives of the LHC physics program. Before its recent discovery, several limits were set on the Higgs bosson mass. Indirect constraints are obtained from the precision measurements in the electroweak sector of the Standard Model, requiring a mass lower than 158 GeV [306]. Direct mass constraints are set at 95% CL by LEP [307], Tevatron and LHC [308,309] collaborations, excluding a Higgs boson in the mass range of 114.4 GeV $< m_H <$ 560 GeV, apart from some regions in the 116.6 GeV $< m_H <$ 129.2 GeV interval. After the discovery of the Standard Model-like Higgs boson in July 2012 by the ATLAS and CMS collaborations [5], its mass is experimentally measured to be 125.09$\pm$0.21(stat)$\pm$0.11(syst) [305] when combining the $H\rightarrow\gamma\gamma$ and $H\rightarrow ZZ\rightarrow llll$ channels (Figure 9.1).

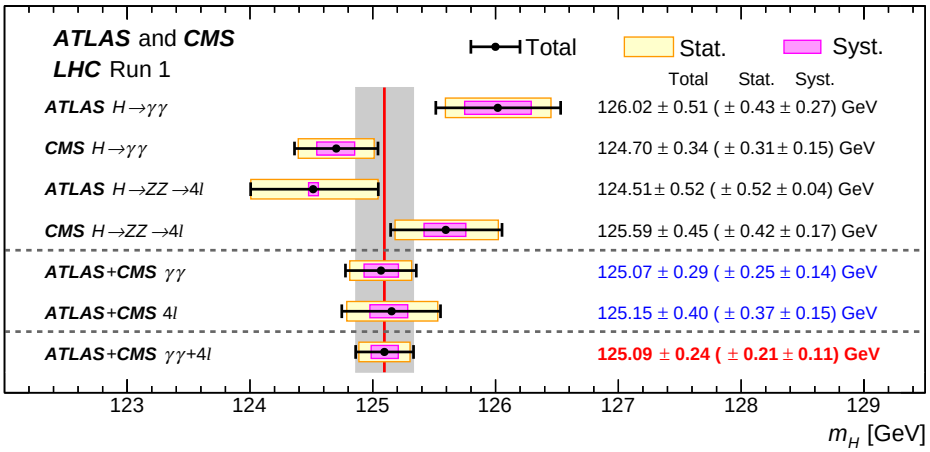


Figure 9.1: Results for the Higgs boson mass measurement obtained by ATLAS and CMS. The considered uncertainties are systematic (magenta band), statistical (yellow band) and total (black lines). The central value of the combined measurements is shown with a red line, while the associated uncertainty with a hashed gray area [305].

Beside the mass measurement, other experimental studies of the properties of the new particle are performed in ATLAS, as shortly described in the following. In the Standard Model, the Higgs boson is a neutral scalar particle with CP-even $J^{CP} = 0^{++}$ state arising from a $SU(2)_L$ doublet with $Y = +1$. The experimental spin and parity results published in [11, 303, 310] are showing that several non-SM $J^P = 0^-$ and BSM $J^P = 0_h^+$ models are excluded at more than 99% CL in the favor of a SM spin-0 and positive parity hypothesis. Given the known SM-like Higgs boson mass, the production cross sections and the branching ratios can be computed as well. The dominant production mode at LHC is the gluon fusion $gg \to H$ (ggF) followed by the vector boson fusion processes $gg/q\bar{q} \to qq'H$ (VBF). Other sizable contributions are brought by processes like Higgs boson production in association with a vector boson $q\bar{q}/gg \to Z/W + H$, a pair of top quarks $q\bar{q}/gg \to t\bar{t}H$, a pair of bottom quarks $b\bar{b} \to H$ or a single top. The latter production mode is $qb \to tHq'$ and $gb \to WtH$. Figure 9.2 illustrates Feynman Diagrams for few of the mentioned processes and Table 9.1 summarizes the theoretical and measured cross section values for each mentioned process [311]. The dominant Higgs boson decay mode is $H \to b\bar{b}$ with BR = 57.1%, followed by $H \to WW^*$, Hgg, etc as illustrated in Table 9.2.

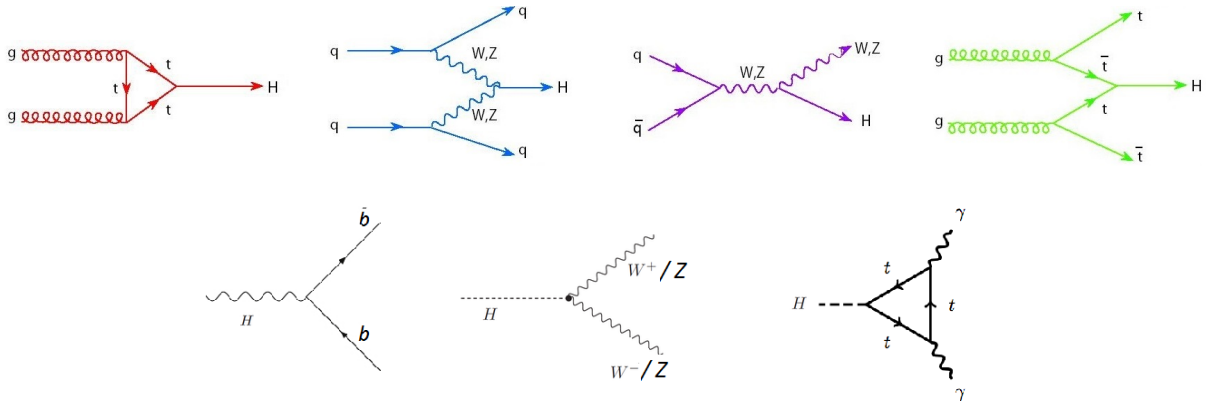


Figure 9.2: Feynman diagrams illustrating four Higgs boson production modes (top), and several decay modes (bottom).

The latest ATLAS Run-1 results of the Higgs boson signal strengths (μ) and the reached precision are summarized in Figure 9.3 [311]. The μ parameter is a measure of potential deviations from the Standard Model prediction, defined as the ratio between the measured and expected signal yield :

$$\mu = \frac{\sigma \times BR}{(\sigma \times BR)_{\mathrm{SM}}} \tag{9.1}$$

In Figure 9.3, the μ parameter is quoted for each Higgs boson production and decay mode (blue color). For completeness, the overall signal strength per decay mode and its associated uncertainty, are also presented

Process	Theoretical cross section [pb]	Measured cross section [pb]
$gg \to H$	19.2 ± 2.0	23.9 ± 3.6
$gg/qq' \to qq'H$	1.57 ± 0.04	2.43 ± 0.58
$q\bar{q}/gg \to WH$	0.698 ± 0.018	1.03 ± 0.53
$q\bar{q}/gg \to ZH$	0.412 ± 0.013	
$b\bar{b} \to H$	0.202 ± 0.028	-
$q\bar{q}/gg \to t\bar{t}H$	0.128 ± 0.014	0.24 ± 0.11
$qb \to tHq'$ $gb \to WtH$	0.018 ± 0.001	-
Total	22.3 ± 2.0	-

Table 9.1: Dominant Higgs boson production mode at LHC and the corresponding theoretical and measured cross sections [pb] at $\sqrt{s} = 8$ TeV. The quoted uncertainties on the theoretical cross section includes sources as QCD scales, parton distribution functions and α_s, while the measured cross section includes the statistical, experimental and theoretical components. Higgs boson mass is considered to be 125.36 GeV [311].

Process	BR (%)
$H \to bb$	57.1 ± 1.9
$H \to WW^*$	22.0 ± 0.9
$H \to gg$	8.53 ± 0.85
$H \to \tau\tau$	6.26 ± 0.35
$H \to c\bar{c}$	2.88 ± 0.35
$H \to ZZ*$	2.73 ± 0.11
$H \to \gamma\gamma$	0.228 ± 0.011
$H \to Z\gamma$	0.157 ± 0.014
$H \to \mu\mu$	0.022 ± 0.001

Table 9.2: Dominant Higgs boson decay modes at LHC and the branching ratio (%) [311].

in each channel (black color). The overall coupling can be constrained especially in the $H \to \gamma\gamma$ and $H \to WW^*$ channels, with a precision around 30%. However, also the $H \to ZZ^*$ and $H \to \tau\tau$ channels are sensitive a precision around 40% on μ can be obtained. Compared to $VH \to Vb\bar{b}$ (precision around 70%) in the $t\bar{t} + H$ channel the precision is even higher (though no combination exists yet), as presented in Section 9.2. Looking at the results in the $H \to \mu\mu$ and $H \to Z\gamma$ the reached precision is very poor – the uncertainty on μ is higher than 160%. Note that the precision reached with individual measurements, focusing on a certain Higgs boson production or decay mode, is less good when compared to the overall measurement results. When combining all the Higgs boson production and decay modes the global value of the signal strength is compatible with the Standard Model prediction, and the reached precision is around 15% – much better than the overall precision obtained per channel, emphasizing once more the importance of the combination :

$$\mu = 1.18^{+0.15}_{-0.14} = 1.18 \pm 0.10(\text{stat}) \pm 0.07(\text{exp})^{+0.08}_{-0.07}(\text{theory}) \tag{9.2}$$

To obtain these results, it is assumed that the relative contributions from the different Higgs boson production and decay modes are as predicted in the Standard Model (= 1). The theoretical uncertainty includes sources as the uncertainty on the cross sections, branching ratios and signal modeling. It is smaller than the uncertainty on the SM Higgs boson production due to the combination : contributions from VBF or $H + W/Z$ processes with small theoretical uncertainties have a higher weight in the combination than in the total cross section computation. The uncertainty on the global (combined) μ parameter is smaller than in the stand-alone measurements and the statistical and systematical components have the same order of magnitude.

Generally, at the end of Run-1, the precision reached for the Higgs boson-SM vector bosons and Higgs boson-SM fermions coupling strengths are ±7% and ±16% respectively and their value is compatible with the SM. These results corresponds to the case where no BSM contributions are included and only two global coupling strengths for all fermions and bosons are considered [311]. To measure these coupling strengths the production and the decay modes should be combined given the existence of one or more couplings in the production-to-decay chain. This measurement can be either indirect or direct. For example, an indirect measurement of the top Yukawa coupling can be performed when targeting the $gg \to H$ production mode or the $H \to \gamma\gamma$ decays. Here contributions from the top and bottom quarks appears in the loop processes, leading to an indirect coupling with the Higgs boson. As beyond the Standard Model physics can also bring contributions to these loops, the direct measurement of the top quark Yukawa coupling is mandatory. With the LHC Run-1 luminosity it can be directly measured using the $t\bar{t} + H$ processes as shown in Section 9.2.

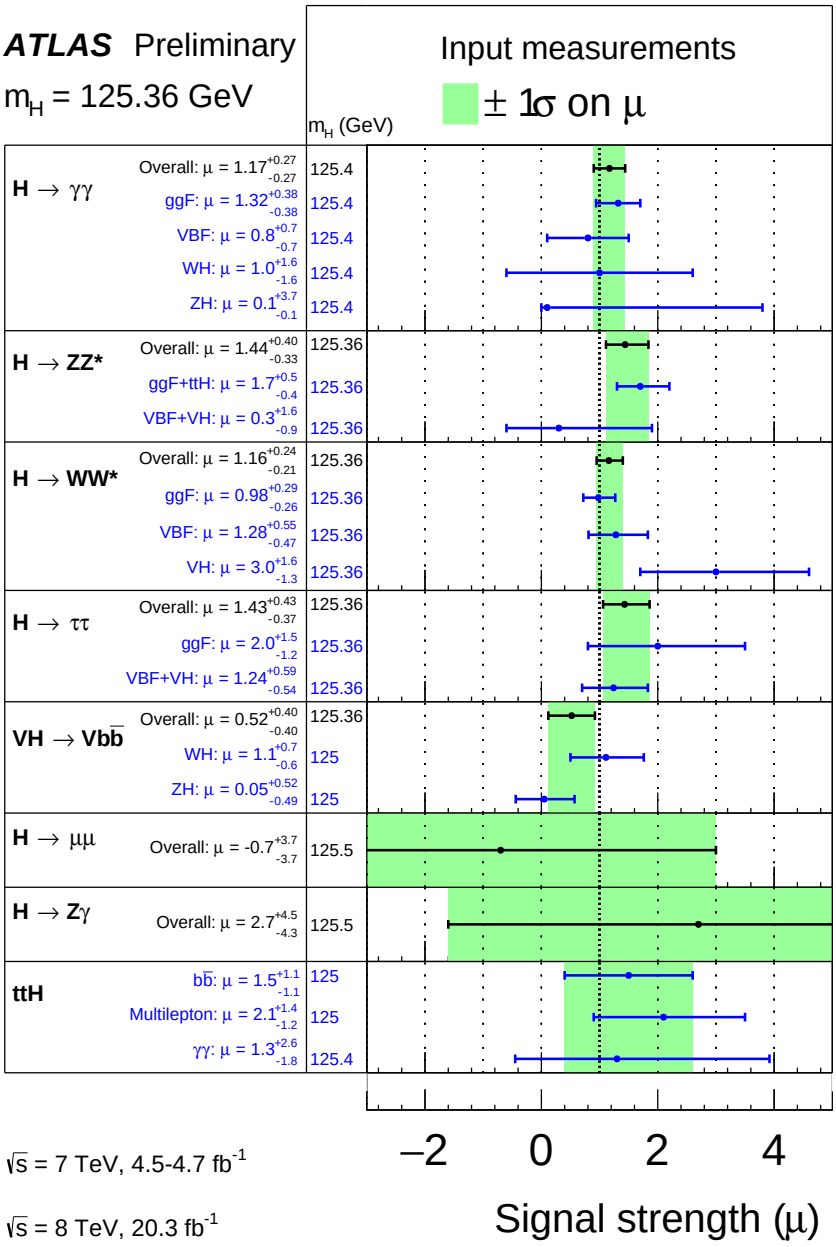


Figure 9.3: Results of the observed signal strength in different channels [311]. The uncertainty bar is the $\pm 1\sigma$ total uncertainty, while the green hashed area is the uncertainty of the overall signal strength in a given channel.

9.2 Top quark Yukawa coupling in Run-1

In ATLAS and CMS several Higgs boson decay channels were considered in Run-1 to measure the top quark Yukawa coupling using $t\bar{t}+H$ production processes [304, 312–318]. The obtained results in the $H \to \gamma\gamma$, $H \to b\bar{b}$ and $H \to$ multileptons final states are summarized in Table 9.3. Combining all the channels, the expected limit is 1.7 and the reached expected uncertainty on the Higgs signal strength is of the order of 100%. However, the $t\bar{t}+H \to$ two same - sign leptons standalone measurement leads to an uncertainty of 180%, showing the importance of the combination. It is interesting to note that all these channels are complementary given the exclusive number of leptons, b-jets and photons. Therefore it makes sense to separately optimize all of them, as performed in this thesis for the two same - sign leptons channel targeting $t\bar{t}+H(\to WW)$ decays (Figure 9.4). The extraction of the top quark Yukawa coupling from this channel is independent of the poorly constrained b-quark Yukawa coupling and of new BSM particles in the Higgs vertices, contrarily to $t\bar{t}+H(\to b\bar{b})$ and $t\bar{t}+H(\to \gamma\gamma)$, respectively.

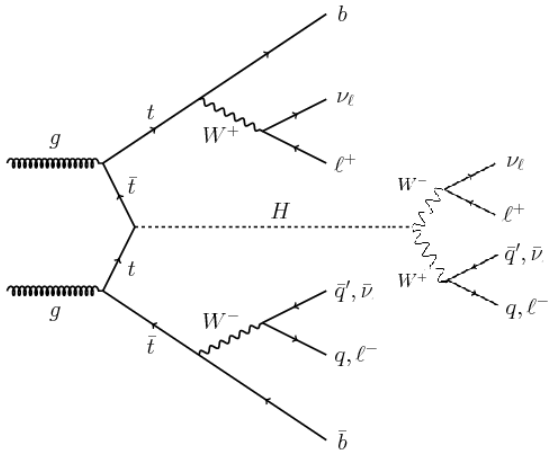


Figure 9.4: Feynman diagram illustrating the $t\bar{t} + H\ (\rightarrow WW) \rightarrow$ two same - sign leptons process.

Final state	Expected 95% CL limit on $\mu_{t\bar{t}H}$		Expected precision on $\mu_{t\bar{t}H}$	
	CMS	ATLAS	CMS	ATLAS
$H \rightarrow \gamma\gamma$ and di-leptonic $t\bar{t}$		6.6		
$H \rightarrow \gamma\gamma$ and hadronic $t\bar{t}$		10.1		
Combination of the above two	4.7	4.9		
$H \rightarrow b\bar{b}$ and semi-leptonic $t\bar{t}$	(4.2)	2.6		$\pm$1.2
$H \rightarrow b\bar{b}$ and di-leptonic $t\bar{t}$	(6.9)	4.1		$\pm$1.8
Combination of the above two	3.5 (3.3)	2.2		$\pm$1.1
=2 same - sign e/μ+0τ_{had} (2lSS)	3.4	3.9		$\pm$1.8
=3 e/μ+0τ_{had}	4.1	3.8		$\pm$1.8
=2 e/μ+1τ_{had}	NA	8.4		$\pm$3.3
=1 e/μ+2τ_{had}	14.2	17.6		$\pm$10.4
=4 leptons no had. tau	8.8	14.9		$\pm$3.6*
Combination of the above five	2.4	2.4		$\pm$1.2
All channels combined	1.7	1.4		

Table 9.3: Expected 95% CL limit on $t\bar{t} + H$ signal strength measurement $\mu_{t\bar{t}H}$ (first two columns where $\mu_{t\bar{t}H} = 0$ is assumed) and the precision on $\mu_{t\bar{t}H}$ (last two columns where $\mu_{t\bar{t}H}$=1 is assumed) published by ATLAS [13,304,312,313,315–319] and CMS [320,321] with Run-1 data. For CMS, the numbers in brackets correspond to Ref. [321]. * For the 4l Z-depleted category, $\mu_{t\bar{t}H} < -0.17$ results in a negative expected total yield so the lower uncertainty is truncated at this point.

9.3 Inputs for prospect studies

In this section the inputs for the prospect studies are discussed. It starts with a description of the used signal and background Monte Carlo samples, and the methodology employed to re-weight them at the energy $\sqrt{s}$ of 13 TeV (Section 9.3.1). It continues with a short presentation of the trigger strategy, the objects used in the analysis and the event selection. The section ends with the differences present in the object definitions with respect to the published $\sqrt{s} = 8$ TeV 2lSS analysis [304] (Section 9.3.2).

9.3.1 Signal and background samples

Final states with same - sign leptons in $t\bar{t} + H$ events are obtained when one lepton originates from one of the top and the other lepton from $H \rightarrow W^+W^-, ZZ$ or $\tau\tau$ decays. When considering leptons with p_T higher than 20 GeV, the events come at 80% from $t\bar{t} + H(\rightarrow W^+W^-)$ events. Exclusive same - sign leptons are present in 4.1% of $t\bar{t} + H(\rightarrow W^+W^-)$ events. The $t\bar{t} + H$ signal samples used in this study are described in Table 9.4. The low cross section Standard Model processes leading to a final sate formed by

two same - sign leptons were presented in Section 6.4.

Process	Generator	Hadronization	$\sigma\times$ BR [fb]	eff.
$t\bar{t}+H$ $(H\to\tau\tau\to ll)$	Powheg	Pythia8	130×0.063	0.123
$t\bar{t}+H$ $(H\to\tau\tau\to lh)$	Powheg	Pythia8	130×0.063	0.458
$t\bar{t}+H$ $(H\to ZZ$ inclusive)	Powheg	Pythia8	130×0.026	1.00
$t\bar{t}+H$ $(H\to WW$ inclusive)	Powheg	Pythia8	130×0.215	1.00

Table 9.4: Overview of $t\bar{t}+H$ MC samples. The used cross section for $t\bar{t}+H$ production at 8 TeV ($m_H=125$ GeV) is 130 fb at NLO [322]. The last column shows the filter efficiency.

Figure 9.5 illustrates the evolution of the cross section as a function of the center of mass energy $\sqrt{s}$ for the main SM processes. At $\sqrt{s}=13$ TeV, the NLO cross sections are 509 fb for $t\bar{t}+H$ [323], 419 fb for $t\bar{t}+W^+$ and 227 fb for $t\bar{t}+W^-$ [324] leading to a value of 646 fb for $t\bar{t}+W$ and 757 fb for $t\bar{t}+Z$ [1]. The NNLO+NNLL cross section for $t\bar{t}$ is 832 pb [325] [2]. This gives the following ratios with respect to 8 TeV cross sections : 3.9 for $t\bar{t}+H$, 2.6 for $t\bar{t}+W^+$, 3.2 for $t\bar{t}+W^-$, 2.8 for $t\bar{t}+W$, 3.7 for $t\bar{t}+Z$ and 3.3 for $t\bar{t}$. These numbers show that the signal sensitivity will increase at a higher center-of-mass energy for a given integrated luminosity. For the Run-2 prospect studies, the `PDF re-weighting tool` [326] is applied to the considered $\sqrt{s}=8$ TeV Monte Carlo samples. Cross sections at 13 TeV, obtained from re-weighted the 8 TeV samples, are similar below 10% level to the theoretical expectations just mentioned. Note that because of the sample re-weighting, the Run-1 pile-up conditions are assumed by default.

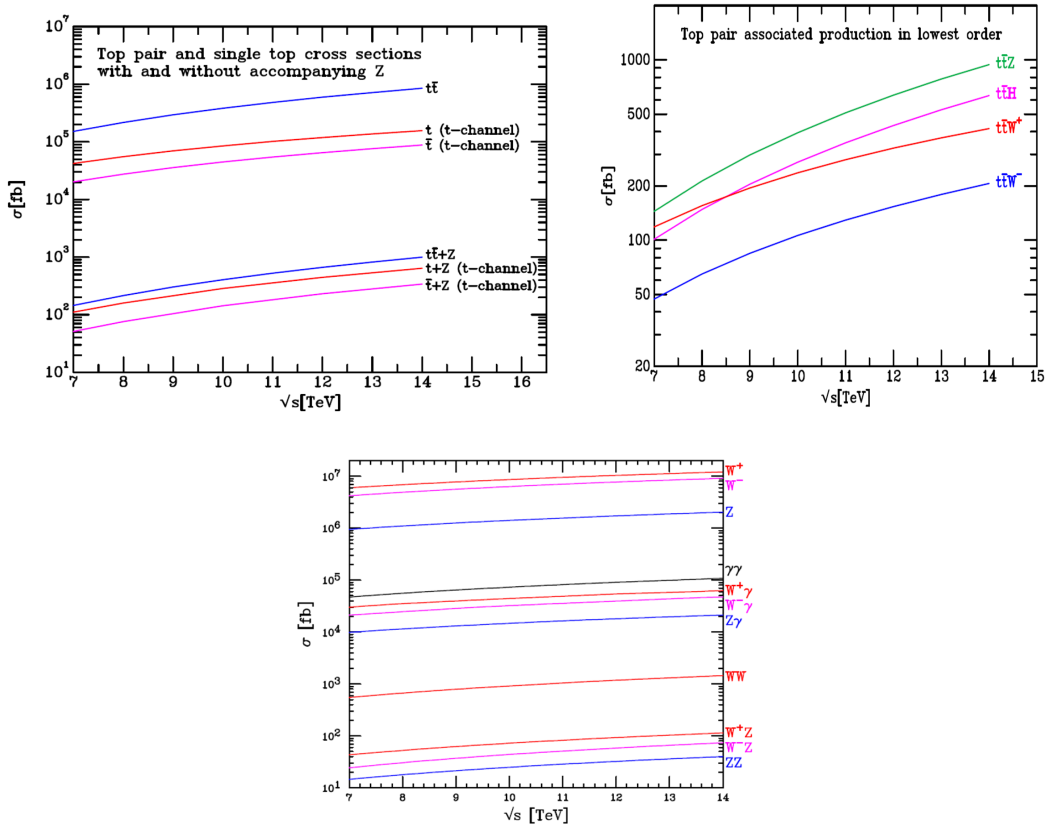


Figure 9.5: Evolution of SM cross section as a function of the center-of-mass energy for top [327] (top left), $t\bar{t}+V$ [324] (top right) and (di-)bosons [328] (bottom). A top mass of 172.5 GeV is assumed except for the top right plot where 173.2 GeV is assumed.

9.3.2 Trigger strategy and object definitions

All events should fire one of the non-isolated dilepton (`EF_2e12Tvh_loose1`, `EF_e12Tvh_medium1_mu8`, `EF_mu18_tight_mu8_EFFS`) or high p_T electron or muon (`EF_e60_medium1`, `EF_mu36_tight`) trigger items.

[1] The 13 TeV numbers for $t\bar{t}+W$ and 13 TeV number for $t\bar{t}+Z$ have been derived privately.
[2] A top mass of 172.5 GeV is assumed for $t\bar{t}$.

Contrarily to the ATLAS published analysis, the main unprescaled single lepton trigger items (`EF_e24_medium1`, `EF_mu24_tight`) are not considered, as they contain a track isolation requirement at HLT, which could bias the fake estimates with the matrix method. The impact of this trigger choice compared to the published analysis has been checked to be below 10% on signal and $t\bar{t}+V$ processes.

The object definitions used for Run-2 prospect studies follow the SUSY analysis presented in Chapter 7 if not otherwise specified. The **pre-selected leptons** [3] must satisfy the $p_T >$ 10 GeV requirement and η should be within the detector acceptance ($|\eta_{\text{cluster}}| < 2.47$ for electrons and $|\eta| < 2.5$ for muons). Leptons should be of *medium* (*medium* LLH for electrons) quality, separated from jets and, in case of one electron and muon in the event, satisfy the ΔR(e, μ) > 0.1 condition. The **signal leptons** must satisfy further requirements to decrease the contamination from charge flip and fake leptons. In particular, only the barrel electrons ($|\eta| < 1.37$) satisfying the *very tight* LLH criteria are considered at 8 TeV, like in the published $t\bar{t}+H$ 2lSS analysis. For Run-2 prospect studies, electrons falling in the crack region ($1.37 < |\eta| < 1.52$) or at large η region ($|\eta| > 2.3$) are removed. Further cuts involve the track and calorimeter isolations computed in a cone of $\Delta R < 0.2$ (0.3) radius around the electron (muon) and the impact parameter significance in the transversal and longitudinal planes :

- Calorimeter isolation : $E_T^{\text{cone2}}0 < 0.05 \times E_T$.
- Track isolation : $p_T^{\text{cone2}}0 < 0.05 \times p_T$.
- Impact parameters : $|z_0 \cdot sin(\theta)| < 0.4$ mm and $|d_0/\sigma(d_0)| < 3.0$.

The used **jets** are reconstructed using the `FastJet` implementation of the anti-k$_t$ jet algorithm with four-momentum recombination and a distance parameter $R = 0.4$. They are required to have $p_T >$20 GeV and $|\eta| < 2.5$. In case of b-tagging, the threshold is $p_T > 20$ GeV and the operating point of the MV1 tagging algorithm gives a 70% inclusive efficiency for jets identified as originating from a b quark. To remove pile-up jets, the Jet Vertex Fraction should be greater than 0.5 for jets with $p_T < 50$ GeV and $|\eta| = < 2.4$. Same definition for the missing transverse energy as defined in Section 7.3 is employed.

The few differences in the object definition with respect to the published analysis are :

- Muon definition: the published analysis uses Muid algorithm with tight identification instead of MUONS and medium identification as in this analysis.
- Electron / Muon $|z_0 sin\theta|$: the published analysis considers 1 mm upper cut while it is tighter (0.4 mm) in this study.
- Overlap removal between objects (baseline pre-selected objects here, signal objects in the published analysis) : in [304] the overlap removal between an electron and a jet (muons and jet) occurs if $\Delta R < 0.3$ $(0.04 + 10$ GeV$/p_T)$ instead of 0.2 (0.4) as imposed in this study. Also any two signal electrons should be separated by $\Delta R > 0.1$ in the published analysis while nothing is required in this analysis setup.

Generally, all events should contain exactly two signal leptons with $p_T > 10$ GeV and same electric charge. One or two of these leptons are matched ($\Delta R < 0.15$) to the trigger object depending if the single or the dilepton trigger is fired, respectively. The event is then filtered into one of the three channels depending on the flavor of the two signal leptons which form the same - sign pair (ee, $\mu\mu$, $e\mu$). Events with an invariant mass of the two leptons $M_{ll} < 12$ GeV are rejected to avoid contamination from heavy flavor resonances.

9.4 Analysis set-up

The methods applied to estimate the SM and detector backgrounds and their validation are summarized in Section 9.4.1. The statistical tool and the fit set-up are presented in details in Section 9.4.2. To give confidence to our results, the set-up is validated in Section 9.4.3 by comparing with the published Run-1 results.

9.4.1 Background estimation and validation

The $t\bar{t}+V$, dibosons+b and other SM background processes are estimated using the MC samples discussed in section 9.3.1. The detector backgrounds are determined using the methods applied on 8 TeV data, as described in Chapter 6. For the Run-2 prospects studies, the 8 TeV detector background estimation

[3] Taus are not considered in these prospect studies, i.e. no tau veto is applied.

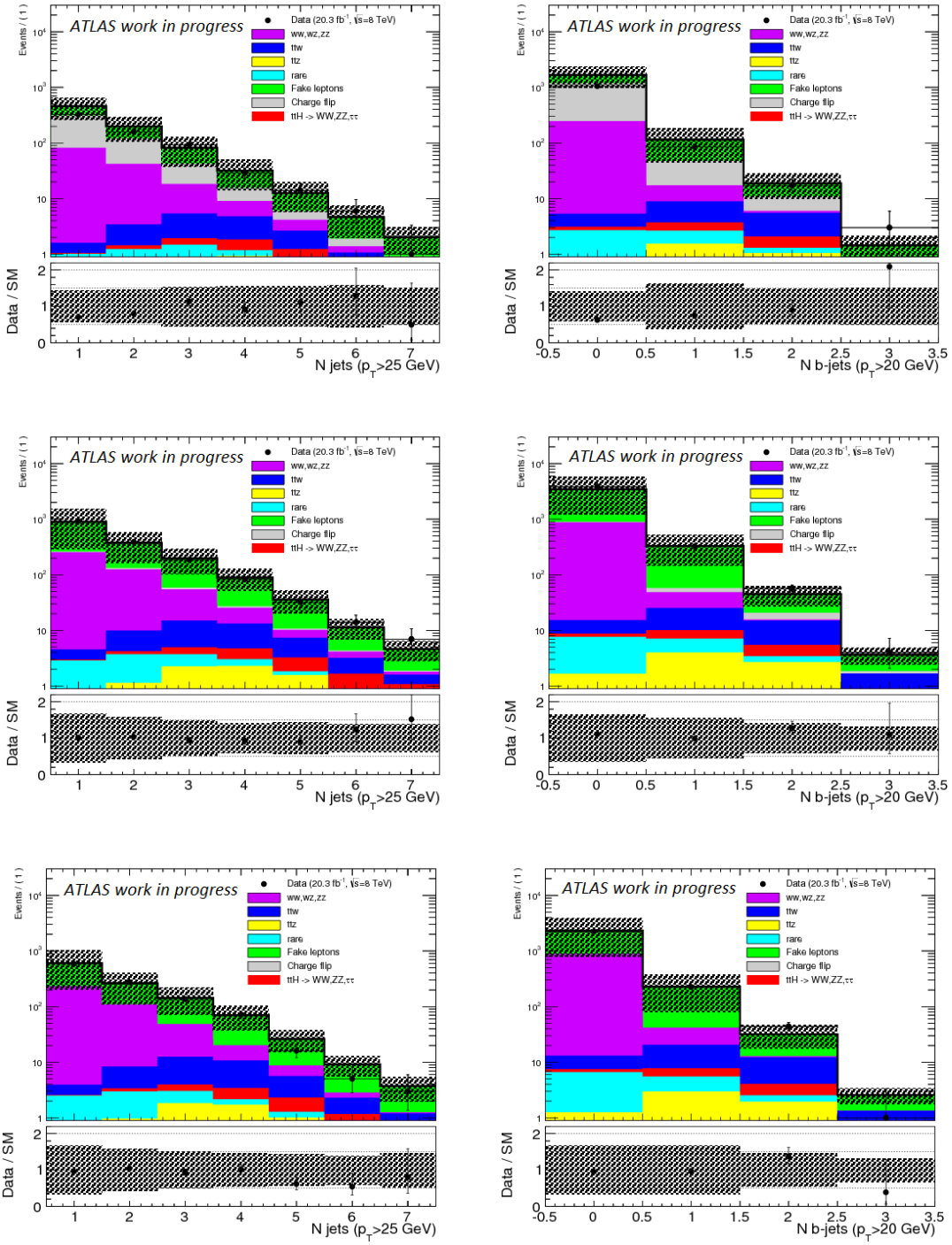


Figure 9.6: ee (top), $e\mu$ (middle) and $\mu\mu$ channels: distributions of the signal jet multiplicity ($p_T \geq 25$ GeV) (left) and b-jet multiplicity ($p_T \geq 20$ GeV) (right) after two same - sign signal lepton selection. Statistical and all systematic errors (section 9.4.2) are included.

is re-weighted to 13 TeV using a factor of 3.3 (Section 9.3.1, $t\bar{t}$ processes). First, the charge flip rate is estimated in a $Z \rightarrow ee$ data sample with the Tag&Probe method (Section 6.5). The rates are measured in 40 $|\eta|$-p_T bins and are typically in the range 10^{-5} - 10^{-2}. No charge flip estimation is computed for muons since it is known to be negligible. Second, the fakes leptons originating from jets, photon conversions or genuine leptons from heavy-flavor meson decays [4] are estimated with the 2×2 matrix method presented in Section 6.6. The two input parameters: real and fake loose-to-tight lepton efficiencies are measured in dedicated regions built with 8 TeV data. Figure 9.6 shows the number of jets and b-jets distributions in ee, $e\mu$ and $\mu\mu$ channels, after the pre-selection. At this stage, the fake leptons and charge flip represent a sizeable amount of the background. A good agreement is observed between 8 TeV data and the sum of SM

[4] This category also includes $t\bar{t} + \gamma$ and $W + \gamma$ backgrounds.

and detector backgrounds, validating the data-driven estimates.

9.4.2 Fit set-up

To perform the statistical interpretation of the analysis and to assess the final sensitivity to $t\bar{t}+H$ signal, the `Histfitter` tool is used. As detailed in Section 6.8.1, it is based on a likelihood function built with the number of background and signal events in the signal regions. This function is constructed as the product of Poisson probability distributions for event counts in the dedicated signal regions and potentially in each of the control regions (CRs), and of constraints for systematic uncertainties on the expected yields, the so called nuisance parameters (Equation (6.10), Section 6.8.1). The nuisance parameters are modeled by a Gaussian probability density function with a width given by the size of the uncertainty. The Poisson probability density functions also include free (μ) parameters, to scale the expected contributions from the major backgrounds and the signal to data. In this study, the parameter labeled with $\mu_{t\bar{t}H}$ corresponds to the $t\bar{t}+H$ signal strength in the signal region (=1 in the Standard Model). If no control region is used for a given background, no scaling exists for this background and the nuisance parameters are associated to the theoretical uncertainties on the cross section and acceptance as detailed for the SUSY searches. If a control region is included in the fit, the background scaling (μ_{bkg}) is defined as:

$$N_{SR}^{Bkg} = \frac{(N_{CR}^{data} - N_{CR}^{MC,others})}{N_{CR}^{MC}} \times N_{SR}^{MC} = \mu_{bkg} \times N_{SR}^{MC} \tag{9.3}$$

and one nuisance parameter is associated to the transfer factor (TF) from CR to SR defined as :

$$N_{SR}^{Bkg} = \frac{N_{SR}^{MC}}{N_{CR}^{MC}} \times (N_{CR}^{data} - N_{CR}^{MC,others}) = TF_{CR\to SR} \times (N_{CR}^{data} - N_{CR}^{MC,others}) \tag{9.4}$$

During the statistical test, the free parameters and the nuisance parameters are adjusted to maximize the likelihood function. The nuisance parameters for the statistical uncertainties are attached to each background, per signal region. In this study a total of six signal regions are considered, in which six SM background sources ($t\bar{t}+W$, $t\bar{t}+Z$, WZ+b, WW+b, ZZ+b and others) and two detector backgrounds (fake lepton and charge flip) are estimated. The error from MC statistics is typically below 10%. However the statistical error coming from the fake lepton or charge flip measurement can reach up to 40% at 8 TeV with 20 fb^{-1} of data. This gives a total of $6 \times (1+6+2) = 54$ nuisance parameters [5] when considering the six signal regions.

Several sources are considered for the systematic uncertainties to built a simplified set-up. The theory uncertainties considered in this study are shown in Table 9.5, and the choices made derive from the published $t\bar{t}$+H 2lSS analysis ones, which are recalled for comparison. Similarly, the detector background and detector systematics are given in Tables 9.6 and 9.7. This gives a total of 18 nuisance parameters, to be compared with 89 in the published $t\bar{t}+H$ 2lSS analysis. The main difference comes from the treatment of jet energy scale (JES) and b-tagging, over-simplified for the prospects studies.

Set-up	$t\bar{t}+H$	$t\bar{t}+W$	$t\bar{t}+Z$	WZ+b	WW+b	ZZ+b	others
Publication [304]*							
N(Nuisance Parameters)	4	4	4	4	4	4	0
Input error in the fit	+9/-12%	15-20%	22%	50%	25%	25%	–
This study							
N(Nuisance Parameters)	1	1	1	1	1	1	1
Input error in the fit	10%	20%**	20%**	35%	35%	35%	20%

Table 9.5: Number of nuisance parameters and size of the relative uncertainty as input for the fit in % for the theory uncertainties associated to cross-section and acceptance for standard model processes. * In [304], four nuisance parameters are included (scale/pdf on cross-section/acceptance) and the quadratic sum is given there. ** When $t\bar{t}+V$ control regions are included in the fit, these systematics are removed and a 5% systematics on the transfer factor from control to signal region is added (section 9.5.4).

The fit results are computed following the CLs prescription [269]. They are based on Asimov datasets with background only hypothesis ($\mu_{t\bar{t}H}=0$) for 8 TeV limit settings and signal plus background hypothesis ($\mu_{t\bar{t}H}=1$) for signal strength measurements at 8 TeV and 13 TeV.

[5] For detector backgrounds, these parameters are directly related to the statistics of the control regions used by the data-driven method, and are accounted for as systematic uncertainties.

Set-up	Fake Leptons	Charge Flip
Publication [304]		
N(Nuisance Parameters)	2 (e and μ transfer factor)	1
Input error in the fit	30-50% (Transfer factor)	40%
This study		
N(Nuisance Parameters)	2 (e and μ fake efficiency)	1
Input error in the fit	30-50%	40%

Table 9.6: Number of nuisance parameters and size of the relative uncertainties as input for the fit in % for the detector backgrounds systematics. The uncertainty on fake leptons includes both statistical and systematic errors.

Syst	Lumi	JES	JER	b-tag	Electron	Muon	$\not{E}_T$	JVF	Trigger
Publication [304]									
N(Nuisance Parameters)	1	23	1	26	4*	5*	0	1	1
This study									
N(Nuisance Parameters)	1	1	1	1	1	1	2	0	0

Table 9.7: Number of nuisance parameters considered for each detector systematics. The luminosity uncertainty is set to 2.8%. * Identification, isolation, energy scale and resolution.

9.4.3 Validation of the analysis set-up at 8 TeV

In order to validate the analysis set-up, the 8 TeV signal regions defined in [304] are implemented to compare the signal sensitivity obtained with the public results. These regions requires that $\sum Q_{lep} = \pm 2$, $p_T(\text{jet}) > 25$ GeV, $p_T(l_1, l_2) > (25, 20)$ GeV, $|\eta_{electron}| < 1.37$ and $N_{b-jets} \geq 1$. Further they are split in six separate regions with N_{jets}=4 (SR4j-ee, SR4j-$e\mu$, SR4j-$\mu\mu$) and $N_{jets} \geq$5 (SR5j-ee, SR5j-$e\mu$, SR5j-$\mu\mu$). The expected event yields of the published analysis, before the fit, are recalled in Table 9.8. Only the total uncertainties are shown. It can be seen that ee and $e\mu$ channels are dominated by detector backgrounds, while $\mu\mu$ is dominated by the $t\bar{t}+V$ background. Summing ee, $e\mu$ and $\mu\mu$ together, SR4j and SR5j have very similar background composition.

Category	q Flip	Fake Lept.	$t\bar{t}+W$	$t\bar{t}+Z$	Diboson+b	Exp. Bkg	$t\bar{t}+H$ ($\mu_{t\bar{t}H}$=1)	($t\bar{t}$ MC)
ee+≥5j	1.1±0.5	2.3±1.2	1.4±0.4	0.98±0.26	0.47±0.29	6.5±1.8	0.73±0.14	(1.3 ± 0.7)
$e\mu$+≥5j	0.85±0.35	6.7±2.4	4.8±1.2	2.1±0.5	0.38±0.3	15±3	2.1±0.41	(2.6 ± 0.8)
$\mu\mu$+≥5j	–	2.9±1.4	3.8±0.9	0.95±0.25	0.69±0.39	8.6±2.2	1.4±0.28	(1.8 ± 0.7)
ee + 4j	1.8±0.7	3.4±1.7	2.0±0.4	0.75±0.2	0.74±0.42	9.1±2.1	0.44±0.06	(5.0 ± 1.2)
$e\mu$ + 4j	1.4±0.6	12±4	6.2±1	1.5±0.3	1.9±1	24±5	1.16±0.14	(8.2 ± 1.4)
$\mu\mu$ + 4j	–	6.3±2.6	4.7±0.9	0.8±0.22	0.53±0.3	12.7±2.9	0.74±0.1	(3.7 ± 1.0)

Table 9.8: Expected event yields, before the fit, with 20 fb^{-1} of data at $\sqrt{s}$ = 8 TeV for the published analysis [304]. Other backgrounds (tZ, $t\bar{t}$+WW, triboson, $t\bar{t}t\bar{t}$, tH) are not shown as a separate column but they are included in the total background estimate. Uncertainties on $t\bar{t}$ MC are statistical only [316].

Category	q Flip	Fake Lept.	$t\bar{t}+W$	$t\bar{t}+Z$	Diboson+b	Exp. Bkg	$t\bar{t}+H$ ($\mu_{t\bar{t}H}$=1)	($t\bar{t}$ MC)
ee+≥5j	0.8±0.3	2.7±1.2	1.6±0.4	0.7±0.2	0.3±0.2	6.4±1.3	0.7±0.1	(1.5 ± 0.3)
$e\mu$+≥5j	0.6±0.2	4.9±1.8	4.5±1.0	1.7±0.4	0.4±0.2	12.5±2.2	1.8±0.2	(3.3 ± 0.5)
$\mu\mu$+≥5j	–	2.2±1.1	3.5±0.8	1.1±0.3	0.7±0.4	7.7±1.5	1.3±0.2	(1.6 ± 0.4)
ee + 4j	1.5±0.6	4.4±1.9	2.2±0.5	0.7±0.2	0.6±0.3	9.6±2.1	0.4±0.1	(2.9 ± 0.5)
$e\mu$ + 4j	1.2±0.5	7.9±2.4	5.8±1.2	1.5±0.3	1.1±0.5	17.8±2.8	1.1±0.1	(6.3 ± 0.7)
$\mu\mu$ + 4j	–	2.9±1.5	4.9±1.0	1.0±0.2	0.7±0.3	9.9±1.9	0.8±0.1	(3.1 ± 0.5)

Table 9.9: Expected event yields, before the fit, with 20 fb^{-1} of data at $\sqrt{s}$ = 8 TeV for the published analysis signal regions with the present analysis set-up. Other backgrounds (tZ, $t\bar{t}$+WW, triboson, $t\bar{t}t\bar{t}$, $H + W/Z$) are not shown as a separate column but they are included in the total background estimate. Uncertainties on $t\bar{t}$ MC are statistical only.

Table 9.9 shows the same event yield when implementing the published analysis signal regions in the framework described in Section 9.4.2. A detailed comparison with Table 9.8 is now proposed, from left to right. First, the event yields for detector backgrounds (charge flip + fake leptons) have both relative uncertainties in the 30 - 50% range. The different event yield values are explained by the different data-driven methods used for their estimation. It is interesting to note that in both cases, the $t\bar{t}$ MC samples are underestimating the detector background in average by a factor around two. Second, the SM process yields ($t\bar{t}+V$, dibosons+b and signal) are in good agreement. More precisely, the $t\bar{t}+H$ and $t\bar{t}+V$ event yields and their relative uncertainties are within 15% in almost all signal regions. The remaining differences come from the different samples used for $t\bar{t}+V$ background and the differences in the object definition. Overall, the uncertainties on the total background yields in our set-up are largely dominated by systematics errors : the dominant contribution comes from the fake lepton systematics (muon and then electron) and to a lesser extent from $t\bar{t}+W$ theory uncertainty, jet energy scale and b-tagging. This is similar to what is found in the published analysis.

Table 9.10 shows the uncertainties on $\mu_{t\bar{t}H}$ for different fit configurations. Using the present set-up a very good agreement is obtained with the published results for the statistical component of the 95% CL expected limit (consistently in ee, $e\mu$ and $\mu\mu$ channels). Similarly, when including the systematic uncertainties, both the expected limit and the precision on $\mu_{t\bar{t}H}$ are in excellent agreement and gives in both cases :

$$\mu_{t\bar{t}H} = 1.0^{+1.9}_{-1.7} = 1.0^{+1.3}_{-1.2}(\text{stat})^{+1.4}_{-1.2}(\text{syst}) \tag{9.5}$$

It is interesting to note that the statistical and systematic components contribute equally to the signal sensitivity. This comparison validates the present analysis set-up and allows to confidently perform Run-2 prospect studies.

	Expected Limit at 95% CL ($\mu_{t\bar{t}H}=0$)					SM value ($\mu_{t\bar{t}H}=1$)
Errors	Stat only				Stat⊕Syst	Stat+Syst=Tot
Channel	ee4-5j	$e\mu$4-5j	$\mu\mu$4-5j	SR4-5j	SR4-5j	SR4-5j
Publication [304]	7.4	3.8	4.5	2.6	3.9	$^{+1.3}_{-1.2}$ $^{+1.4}_{-1.2}$ $=^{+1.9}_{-1.7}$
This note	7.6	4.0	4.4	2.6	3.8	$^{+1.3}_{-1.2}$ $^{+1.4}_{-1.2}$ $=^{+1.9}_{-1.7}$

Table 9.10: Uncertainties on $\mu_{t\bar{t}H}$, after the fit, for different configurations with 20 fb^{-1} of data at $\sqrt{s}=8$ TeV for the published analysis signal regions SR4j-5j [304] and with the present fit set-up.

9.5 Prospects for full Run-2 statistics

This section is dedicated to prospects for Run-2, given a total luminosity of 100 fb^{-1}. The optimized signal region definitions for the analysis are defined in Section 9.5.1. Given the high luminosity available, a first attempt to build control and validation regions is discussed in section 9.5.2. Finally, fit results without any control region are presented in Section 9.5.3 and possible improvements are mentioned in Section 9.5.4. For all these studies, the signal strength is assumed to be $\mu_{t\bar{t}H}=1$.

9.5.1 Signal regions proposal for Run-2 LHC data

As discussed in section 9.4.3, the signal sensitivity of the published results is mainly driven by the high level of the fake lepton background and its associated uncertainty. For the prospect study, we assume, conservatively, that the error on this background will be the same, as quoted in Table 9.6. From section 9.4.3, it is also clear that SR5j carries most of the signal sensitivity. With the increase of the center-of-mass energy and luminosity in Run-2, a signal region optimization based on SR5j and further split in two regions, one with exactly one b-jet (SR1b) and another one with at least 2 b jets (SR2b) is considered in this section. To illustrate the approach, Figure 9.7 (left) shows a sketch of the phase space covered by SR4j, SR5j, SR1b and SR2b.

Figure 9.7 also shows the distribution of the number of jets with $p_T > 20$ GeV after event pre-selection for $= 1$ b-jet later called $1b$ (center) and > 1 b-jet later called $2b$ (right) regions. For $1b$, the background event yield is dominated by the detector one – fake lepton (76.8%), SM background (18.6%) and charge flip (4.6%). The signal over background ratio (S/B) is 3.9%. In the $2b$ region, S/B increases to 11.3% and the fake lepton background drops to 52.1%, charge flip increases to 10.3% while SM background raises to 37.6%. The decrease in the fake lepton background, by a factor four, is explained by the (tight) requirement of two b jets in the event : fewer b jets which could fake leptons are remaining in the event. For all the

other samples (apart from dibosons+b, negligible in 2b), similar event yields are observed in 1b and 2b since they all have two real b-jets in their final states and the b-tagging working point efficiency is high (70%). To reduce the dominant detector backgrounds and especially the fake lepton component, below the level of the irreducible SM background, it is therefore interesting to split the signal region in SR1b and SR2b and apply harder cuts in SR1b than in SR2b.

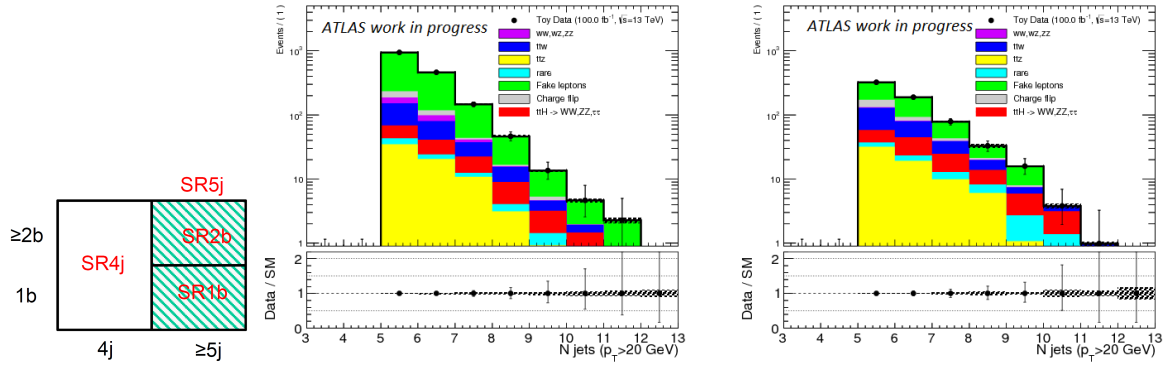


Figure 9.7: Left: Phase space occupied by the published ATLAS signal regions (SR4j, SR5j) and the new proposed signal regions (SR1b, SR2b). Right: Number of jets with $p_T > 20$ GeV for $= 1$ b-jet (center) and > 1 b-jet (right) at 13 TeV with 100 fb^{-1} of data. Distributions are obtained after event pre-selection, applying a Z-veto in ee channel, $\not{E}_T > 20$ GeV (see section 9.5.1) and at least five jets with $p_T > 20$ GeV. Only statistical uncertainties are shown.

Optimization strategy

For the signal region optimization, the SM background estimates are based on MC and the detector background estimates on the data-driven methods discussed in section 9.4.1. The variables considered to discriminate the signal and the backgrounds are :

- Jet and lepton p_T;
- Reconstructed masses: m_T, the transverse mass built with the hardest lepton and $\not{E}_T$ and M_{ll}, the invariant mass of the two same - sign leptons. Both can help to discriminate against $t\bar{t}+H$ and $t\bar{t}+V$ processes because of the differences in production and decay;
- Event activity and missing energy: as no mass peak can be reconstructed because of the presence of two neutrinos and high combinatorics between jets, only event-based variables are considered. These variables are the scalar transverse sum H_T of leptons and jets [6], the vectorial sum $\not{E}_T$ and m_{eff} $=H_T+\not{E}_T$.

The signal sensitivity is evaluated through its discovery significance, Z_r obtained from Equation 7.1. The considered relative background uncertainty is shown in Table 9.11 and represents a realistic situation where $\Delta B/B$ takes the expected values for the different background (see section 9.4.2). It is used by default in the following. To evaluate the sensitivity reached when combining several regions, a combined significance is computed as $Z^{ee}_{\text{SRx}} \oplus Z^{e\mu}_{\text{SRx}} \oplus Z^{\mu\mu}_{\text{SRx}}$, where x is the index of the signal region.

Background type	Z_i	Z_r
$t\bar{t}+W$, $t\bar{t}+Z$, Dibosons+b	0	20%
fakes	0	40%
charge flip	0	40%

Table 9.11: Background relative uncertainty (Δb) for two definitions of the signal significance: ideal (i) and real (r).

[6] Jet and lepton should satisfy p_T (jet) > 40 GeV and p_T (lepton) > 15 GeV.

Signal region definition for Run-2

Before optimizing any signal region, two general cleaning cuts are applied, a Z-veto in ee channel and $\not{E}_T > 20$ GeV to remove some of the detector backgrounds. The cuts for SR1b and SR2b are chosen such as the signal sensitivity Z_r is maximal with a sizable statistics for the signal. In SR1b, the optimization of the number of jets and jet p_T cuts points to at least five jets with $p_T > 30$ GeV. However, given the kinematics of the selected events, the two leading jets have almost always $p_T > 40$ GeV. Further, the p_T cut of the leading (Figure 9.8 left) and sub-leading leptons are optimized. The optimal p_T thresholds rises from 10 GeV to 30 GeV, as illustrated in Figure 9.9 (left). This allows a substantial reduction of detector background which drops below 40% of the total event yield. Finally, to remove the $t\bar{t}+V$ background, an upper cut of 150 GeV is applied on M_{ll} (Figure 9.8, right). Generally, the reached signal efficiency is 20%, a factor of 4.4 is gained on S/B and the detector background drops to 37% of the background event yield.

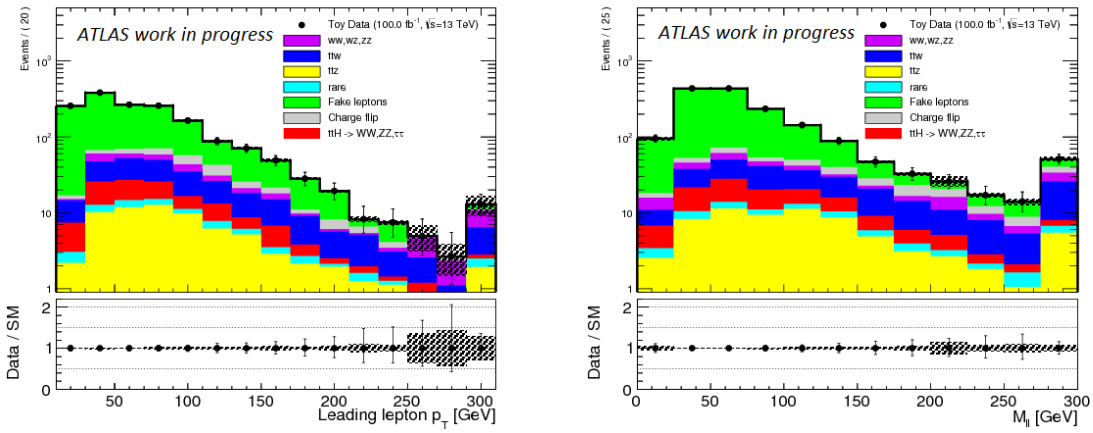


Figure 9.8: Distributions of discriminant variables for SR1b optimization combining ee, $e\mu$ and $\mu\mu$ channels: leading lepton p_T (left) and M_{ll} (right). Distributions are obtained at 13 TeV with 100 fb^{-1} of data after event pre-selection, a Z-veto in the ee channel, $\not{E}_T > 20$ GeV and one b-jet requirements. Only the statistical uncertainties are shown.

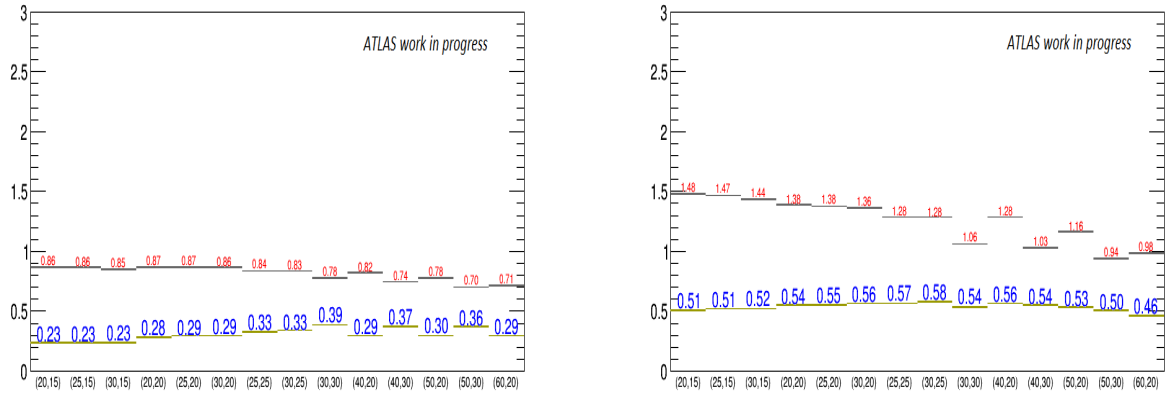


Figure 9.9: Optimization of the lepton p_T threshold cuts in SR1b (left) and SR2b (right). Numbers for the Z_r (Z_0) significance shown in bold blue (red) are obtained at 14 TeV for 20 fb^{-1} of data.

Figure 9.10 shows the distributions of similar discriminating variables used for SR2b optimization. The situation is much more favorable than for SR1b (less background) and therefore the cuts can be relaxed to increase the signal yield. Figure 9.9 (right) shows the optimization of the lepton p_T cut. In this case, the optimization points to at least five jets with $p_T > 25$ GeV, $M_{ll} < 150$ GeV, and $p_T(l_1, l_2) > (25, 25)$ GeV. Generally, the reached signal efficiency is 42%, a factor of 2.4 is gained on S/B and the detector background drops to 35% of the background event yield.

A summary of the signal regions definitions is proposed in Table 9.12, and the corresponding cut-flow is shown in Table 9.13.

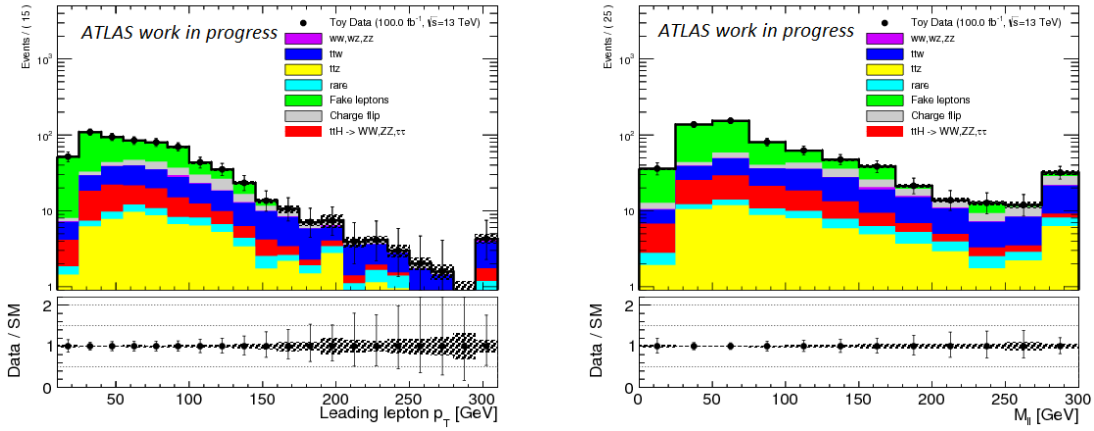


Figure 9.10: Distributions of discriminant variables for SR2b optimization combining ee, $e\mu$ and $\mu\mu$ channels: leading lepton p_T (left) and M_{ll} (right). Distributions are obtained at 13 TeV with 100 fb^{-1} of data after event preselection, a Z-veto in the ee channel, $\not{E}_T > 20$ GeV and at least 2 b-jet requirements. Only the statistical uncertainties are shown.

Region	$N_{b-\text{jets}}$	N_{jets}	Other variables
SR2b	≥ 2	≥ 5, $p_T > 25$ GeV	$\not{E}_T > 20$ GeV, $M_{ll} < 150$ GeV , $p_T(\text{lepton}) > 25$ GeV, Z-veto-ee
SR1b	$= 1$	≥ 5, $p_T > 30$ GeV	$\not{E}_T > 20$ GeV, $M_{ll} < 150$ GeV, $p_T(\text{lepton}) > 30$ GeV, Z-veto-ee

Table 9.12: Signal regions definition proposal for Run-2 prospects (13 TeV, 100 fb^{-1} of data).

9.5.2 Control and validation regions

Implementing control regions (CR) in the fit is a way to reduce the systematics coming from the SM background cross section uncertainties. Definitions are proposed in Section 9.5.2 for $t\bar{t}+W$ and Section 9.5.2 for $t\bar{t}+Z$SM components. For the detector backgrounds several validation regions are defined in Section 9.5.2.

$t\bar{t}+W$ control region

To better constrain $t\bar{t}+W$, two tentative control regions are defined depending on the b-jet multiplicity and shown in Table 9.14. The validation region used by the published 8 TeV analysis, where the purity is 30%, is also shown for comparison. It is known to be very hard to obtain a pure control region for $t\bar{t}+W$ with high statistics since $t\bar{t}+W$ cross section is very small. Moreover, the signature of $t\bar{t}+W$ processes is very close to $t\bar{t}+H$. To ensure the orthogonality with the signal region, less than five jets are required. Therefore the obvious choice is to ask for exactly 4 jets – for $\mu\mu$ channel the 3-jet bin is also considered since the purity is higher. The $t\bar{t}+W$ purity is quite poor, 52% for CR2b_ttW and 34% for CR1b_ttW. For this reason, only CR2b_ttW will be considered in the fit. Note also that the $t\bar{t}+H$ signal contamination is quite low, 6.7% and 6.5% in CR2b_ttW and CR1b_ttW respectively.

$t\bar{t}+Z$ control region

Control regions for $t\bar{t}+Z$ are defined in Table 9.15. The $t\bar{t}+Z$ process only contributes if Z decays leptonically. Therefore a Z mass cut can be applied to enhance $t\bar{t}+Z$ contribution. As for $t\bar{t}+W$ CR, less than five jets are required to ensure the orthogonality with the signal region. The control region chosen by the published 8 TeV analysis, where the purity is 62%, is also given for comparison. A higher purity can be obtained than in $t\bar{t}+W$ CR, 81% and 62% for CR2b_ttZ and CR1b_ttZ respectively. For the fit set-up only CR2b_ttZ is considered since it has the highest purity. Overall, the signal contamination is quite low, 2.9% in CR2b_ttZ and 1.7% in CR1b_ttZ.

Fake lepton validation regions

For fake leptons, three validation regions orthogonal to the signal regions, are designed by reverting some cuts compared to the signal and/or control regions, depending on b jet multiplicity. Table 9.16 details the

Region	Cuts	$t\bar{t}+H$	$t\bar{t}+W$	$t\bar{t}+Z$	Det. background ($t\bar{t}$ MC)	All background
SR2b	$\sigma \times BR \times L$	1860	6890	7420	– (2.310^8)	–
	$\oplus$ "signal" objects, $\not{E}_T > 20$ GeV, $N_{b-\text{jets}} \geq 2$, $N_{\text{jets}} \geq 5$, Zvetoee	65	127	68	360 (335)	577
	$\oplus$ p_T (jets) ≥ 25 GeV	54.1	92.7	52.0	241.1 (229.5)	404.0
	$\oplus$ p_T (leptons)> 25 GeV	33.8	69.0	38.7	57.5 (67.2)	178.9
	$\oplus$ $M_{ll} < 150$ GeV	27.3	38.7	22.3	35.9 (30.3)	102.8
SR1b	$\sigma \times BR \times L$	1860	6890	7420	– (2.310^8)	–
	$\oplus$ "signal" objects, $\not{E}_T > 20$ GeV, $N_{b-\text{jets}} = 1$, $N_{\text{jets}} \geq 5$, Zvetoee	60	146	70	1263 (924)	1552
	$\oplus$ p_T (jets) ≥ 30 GeV	35.0	72.1	34.0	463.3 (348.1)	597.4
	$\oplus$ p_T (leptons)> 30 GeV	16.7	46.4	19.5	40.4 (38.1)	126.5
	$\oplus$ $M_{ll} < 150$ GeV	12.0	20.5	10.4	26.4 (17.5)	70.6

Table 9.13: SR2b and SR1b cut flow for Run-2 prospects (13 TeV, 100 fb^{-1} of data). Channels ee, $e\mu$ and $\mu\mu$ are merged.

Region	$N_{b-\text{jets}}$	N_{jets}	Other variables
CR2b_ttW	≥ 2	≥ 3, $p_T > 25$ GeV ($\mu\mu$) ≥ 4, $p_T > 30$ GeV ($e\mu$, ee) < 5 , $p_T > 25$ GeV (all)	$\not{E}_T > 35$ GeV, $m_{\text{eff}} > 350$ GeV, $p_T(l_1, l_2) > (30, 30)$ GeV, Z-veto-ee
CR1b_ttW	$= 1$	≥ 4, $p_T > 30$ GeV < 5, $p_T < 25$ GeV	$\not{E}_T > 25$ GeV, $m_{\text{eff}} > 450$ GeV, $p_T(l_1, l_2) > (40, 30)$ GeV, Z-veto-ee
VR_ttW (8 TeV)	≥ 2 [$p_T > 25$ GeV]	2 or 3, $p_T > 25$ GeV	$p_T(l_1, l_2) > (20, 20)$ GeV

Table 9.14: $t\bar{t}+W$ control region definition for Run-2 prospects. Exactly two same - sign lepton are asked and leptons are ranked as leading (l_1) and sub-leading (l_2). For information, the definition of the validation region used by the published 8 TeV analysis is also given.

Region	$N_{b-\text{jets}}$	N_{jets}	Other variables
CR2b_ttZ	≥ 2	≥ 4, $p_T > 25$ GeV < 5, $p_T > 25$ GeV	$\not{E}_T > 20$ GeV, $m_{\text{eff}} > 300$ GeV, $p_T(l_1, l_2, l_3) > (30, 20, 10)$ GeV, Z cut ($80 < M_{ll} < 100$ GeV)
CR1b_ttZ	=1	≥ 4, $p_T > 30$ GeV with ≥ 2, $p_T > 50$ GeV < 5 , $p_T > 25$ GeV	$\not{E}_T > 20$ GeV, $m_{\text{eff}} > 450$ GeV, $p_T(l_1, l_2, l_3) > (40, 30, 10)$ GeV, Z cut ($80 < M_{ll} < 100$ GeV)
VR_ttZ (8 TeV)	≥ 1 [$p_T > 25$ GeV]	≥ 3, $p_T > 25$ GeV	$p_T(l_1, l_2, l_3) > (20, 20, 20)$ GeV, Z cut ($86 < M_{ll} < 96$ GeV)

Table 9.15: $t\bar{t}+Z$ control regions definition for Run-2 prospects. Leptons are ranked as leading (l_1), sub-leading (l_2), third-leading (l_3). For information, the definition of the validation region used by the published 8 TeV analysis is also given.

definition of the regions. A good purity is obtained in all cases ($> 80\%$), apart from VR_fakes_0b for muons (52%), with a minimal signal contamination (below 1%).

Overview of validation and control regions

Table 9.17 shows the purity of all pre-defined regions together with the signal contamination. The transfer factor $TF_{CR \to SR}$ (Section 9.4.2) is also given for the CRs. A visual summary is proposed in Figure 9.11.

Region	$N_{b-\text{jets}}$	N_{jets}	Other variables
VR_fakes_2b	≥ 2	1-2, $p_T > 30$ GeV	$20 < \not{E}_T < 100$ GeV, $m_{\text{eff}} < 300$ GeV, $m_T < 100$ GeV, Z-veto-ee, $= 2$ leptonsSS, p_T(leptons) > 15 GeV
VR_fakes_1b	$= 1$	1-2, $p_T > 25$ GeV	$20 < \not{E}_T < 100$ GeV, $m_{\text{eff}} < 300$ GeV, $m_T < 100$ GeV, Z-veto-ee, $= 2$ leptonsSS, p_T(leptons) > 15 GeV
VR_fakes_0b	$= 0$	1-2, $p_T > 25$ GeV	$20 < \not{E}_T < 100$ GeV, $m_{\text{eff}} < 300$ GeV, $m_T < 100$ GeV Z-veto-ee, $= 2$ leptonsSS, p_T(leptons) > 15 GeV

Table 9.16: Definition of fake validation regions for Run-2 prospects.

Region	N(Total)	Purity P_{CR}[%]	Signal cont. [%]	Transfer factor (TF)
CR2_ttW	114	52	7	0.66
CR1_ttW	107	34	7	0.56
CR2_ttZ	65	81	3	0.43
CR1_ttZ	55	62	2	0.30
VR_fakes_2b_ee	74	96*	0.	-
VR_fakes_2b_em	107	92*	0.	-
VR_fakes_2b_mm	38	89	0.	-
VR_fakes_1b_ee	660	95*	0.	-
VR_fakes_1b_em	1120	91*	0.	-
VR_fakes_1b_mm	435	84	0.	-
VR_fakes_0b_ee	7200	89*	0.	-
VR_fakes_0b_em	6500	74*	0.	-
VR_fakes_0b_mm	2300	52	0.	-

Table 9.17: Total number of events, purity, signal contamination and transfer factors for all control and validation regions for Run-2 prospect studies (13 TeV and 100 fb^{-1} of data). * In this case, fake and charge flip leptons are added.

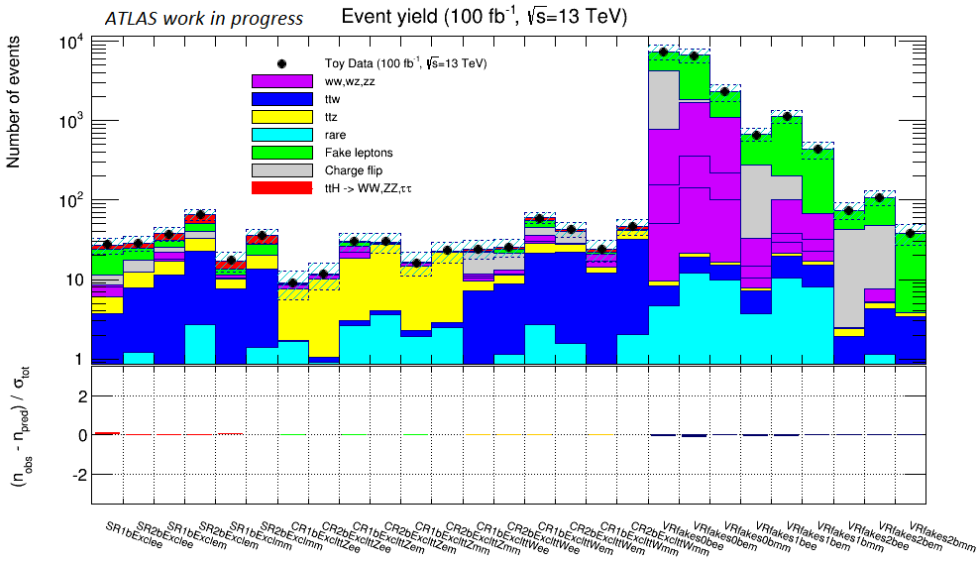


Figure 9.11: Event yields in signal, control and validation regions split in ee, $e\mu$ and $\mu\mu$ for Run-2 prospects (13 TeV and 100 fb^{-1} of data). Statistical and all systematic errors are included.

9.5.3 Signal sensitivity with no control regions

The event yields before fit for full Run-2 statistics using the signal regions optimised in section 9.5.1 are shown in Table 9.18. The level of the detector background is well below the irreducible background ($t\bar{t}+V$ and dibosons+b) in all $e\mu$ and $\mu\mu$ channels. These channels have the best signal sensitivity and among them SR2b is the best.

Category	q Flip	Fake Lept.	$t\bar{t}+W$	$t\bar{t}+Z$	Diboson+b	Exp. Bkg	$t\bar{t}+H(\mu$=1)	($t\bar{t}$ MC)
ee+>= 2b	5.0±2.0	6.9±3.5	6.7±1.8	4.6±1.3	0.1±0.1	24.5±5.0	4.2±0.9	(6.9 ± 2.7)
$e\mu$+>=2b	7.3±2.9	9.9±4.5	19.8±4.9	10.8±2.9	0.3±0.3	50.8±9.4	14.9±2.7	(22.5 ± 6.1)
$\mu\mu$+>=2b	–	6.9±2.9	12.2±3.3	6.9±1.9	0.1±0.1	27.5±5.8	8.2±1.6	(1.0 ± 1.0)
ee+=1b	3.0±1.2	12.2±5.0	3.2±0.9	2.3±0.7	3.9±3.1	25.0±6.2	2.54±0.6	(7.7 ± 4.2)
$e\mu$+=1b	3.6±1.4	5.5±3.5	10.3±2.6	5.5±1.5	5.6±3.3	31.2±6.1	6.2±1.1	(8.7 ± 4.3)
$\mu\mu$+=1b	–	2.2±1.2	7.0±1.9	2.6±1.0	2.0±1.8	14.2±3.4	3.3±0.7	(1.1 ± 1.1)

Table 9.18: Expected event yields with 100 fb^{-1} of data at $\sqrt{s}$ = 13 TeV after all selection in SR1b and SR2b. Other backgrounds (tZ, $t\bar{t}$+WW, triboson, $t\bar{t}t\bar{t}$, tH) are not shown as a separate column but they are included in the total background estimate. Uncertainties on $t\bar{t}$ MC are statistical only.

The precision on the signal strength after the fit (without control region) is :

$$\mu_{t\bar{t}H} = 1.0^{+0.81}_{-0.68} = 1.0^{+0.39}_{-0.37}(\text{stat})^{+0.71}_{-0.57}(\text{syst}) \tag{9.6}$$

Contrarily to the 8 TeV results of Eq.(9.5), the precision is now completely dominated by the systematic uncertainties. As an illustration, Table 9.19 shows the evolution of the signal sensitivity after the fit as a function of the luminosity when combining the six signal regions. While the statistical error is scaling as inverse of the square root of luminosity ratio, the systematics is almost constant around 0.65. Therefore the sensitivity starts to be dominated by systematics above 25 fb^{-1} – which roughly corresponds to the lower estimation for the second year of Run-2. It is thus clear that a dedicated optimisation is needed for 2015 and probably 2016 and that the signal regions of section 9.5.1 are only suited for full statistics of Run-2.

Year	2015		2016		2017
Luminosity (fb^{-1})	5	10	25	50	100
Stat Only	±1.62	±1.15	±0.73	±0.52	±0.38
Syst Only	±0.71	±0.69	±0.67	±0.66	±0.64
Stat⊕Syst	±1.77	±1.34	±0.99	±0.84	±0.75

Table 9.19: Symmetrized uncertainties on $\mu_{t\bar{t}H}$ assuming the SM ($\mu_{t\bar{t}H}$ = 1) for SR1b and SR2b signal regions at $\sqrt{s}$ = 13 TeV for different luminosities. The fit set-up described in section 9.4.2 is used, with no control region.

9.5.4 Possible improvements on the signal sensitivity

Figure 9.12 (left) shows the ranking of nuisance parameters after the fit. All pulls are close to one, showing the good behavior of the fit. The main systematics come from $t\bar{t}+W$ theory, b-tagging, jet energy scale and fake lepton uncertainties. The lower contribution from fake systematics compared to 8 TeV is due to the harder kinematics cuts used in the signal region. As the only significant correlation (20%) is observed between electron fakes and charge flip, as expected, it is possible to split the systematics contributions from theory, fakes/q-flip, experimental sources (fully dominated by b-tagging and JES) and luminosity:

$$\mu_{t\bar{t}H} = 1.0 \pm 0.38(\text{stat}) \pm 0.38(\text{theory}) \pm 0.36(\text{fakes, q}-\text{flip}) \pm 0.38(\text{b}-\text{tag, JES}) \pm 0.11(\text{lumi}) \tag{9.7}$$

In this section several possible improvements which can further lead to a reduction of the three main systematics are discussed : the $t\bar{t}+V$ theory uncertainties – either by using dedicated control regions, or by further splitting the signal regions in the lepton charge sum – and for the fakes and experimental uncertainties.

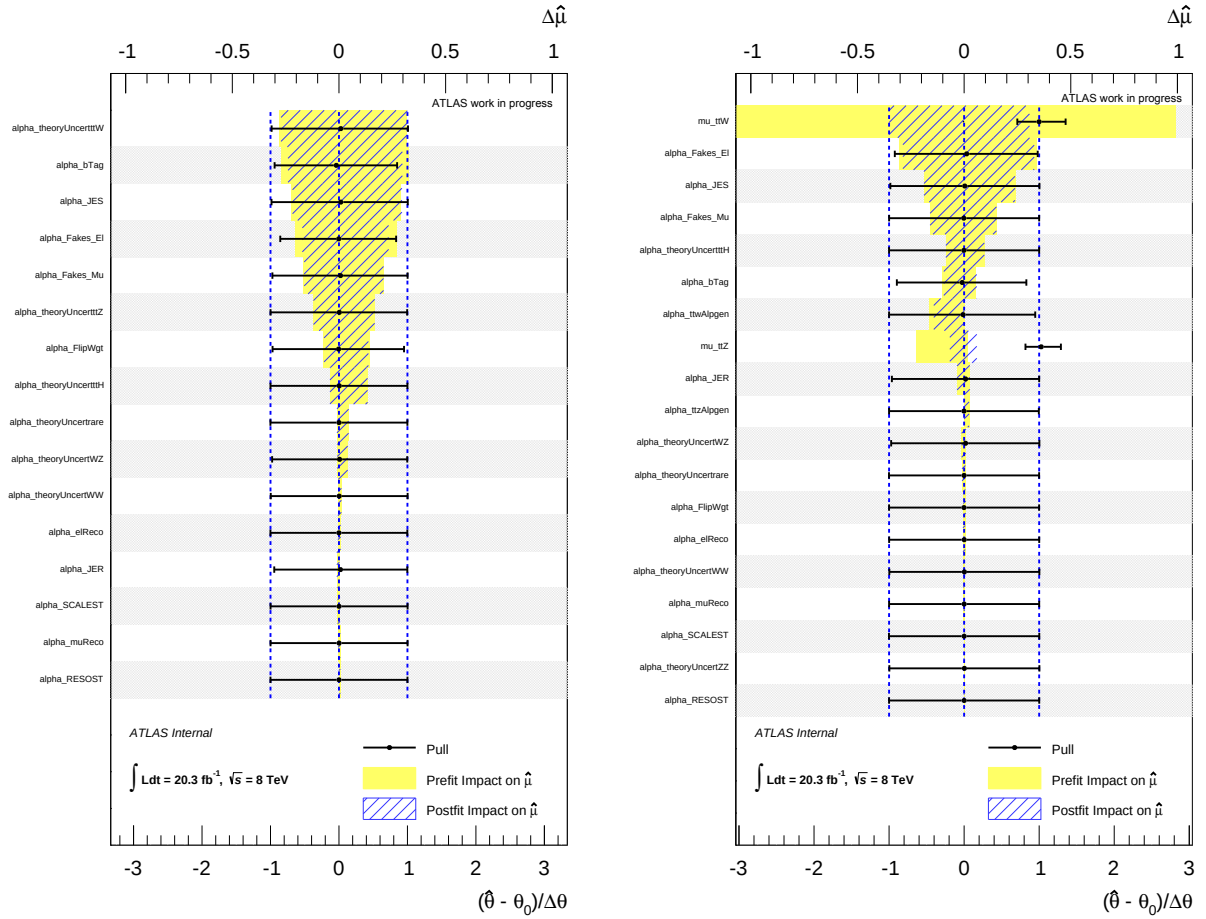


Figure 9.12: Ranking of all considered Nuisance Parameters (NP) with 100 fb^{-1} of data at $\sqrt{s} = 13$ TeV after the fit to an Asimov data set with $\mu_{t\bar{t}H} = 1$ without control region (left) and with $t\bar{t} + V$ control regions (right). The pulling of NP are measured in number of standard deviations with respect to their uncertainties. Yellow (resp. blue hatched) bands give the impact of the pre-fit (resp. post-fit) on the $\mu_{t\bar{t}H}$ error.

Reducing the theoretical uncertainty: adding a $t\bar{t} + V$ control region

When the $t\bar{t} + V$ control regions (CR2b_ttW and CR2b_ttZ) are added in the fit, two additional nuisance parameters are introduced to model the systematics associated to the transfer factor from CR to SR. Following a comparison between two different generators for the $t\bar{t} + V$ MC samples, the uncertainties on these NP are fixed to $\Delta TF = 5\%$. Furthermore, since the $t\bar{t}+V$ cross sections are normalized in the control regions ($\mu_{t\bar{t}+W}$ and $\mu_{t\bar{t}+Z}$ scale factors), the $t\bar{t} + V$ theory uncertainties are removed. Finally, to simplify the comparison with the fit without control regions, ee, $e\mu$ and $\mu\mu$ are merged for SR1b and SR2b.

Figure 9.12 (right) shows the impact of nuisance parameter after the fit with two CRs – in this case only the CR2b_ttW and CR2b_ttZ are considered, as they are the best control regions for the $t\bar{t} + W$ and $t\bar{t} + Z$ backgrounds. Apart from b-tagging, no pulls are observed. The precision on the signal strength after the fit is $\mu_{t\bar{t}H} = 1.00^{+0.80}_{-0.75}$ instead of $\mu_{t\bar{t}H} = 1.00^{+0.88}_{-0.74}$ with no control region [7]. The uncertainty on the $t\bar{t} + W$ scale factor and b-tagging dominates now the total uncertainty. The scale factors for $t\bar{t} + W$ and $t\bar{t} + Z$ are $\mu_{t\bar{t}W} = 1.0 \pm 0.32$ and $\mu_{t\bar{t}Z} = 1.0 \pm 0.24$ respectively. The relative small size of the $t\bar{t} + V$ theory (20%) uncertainty and the poor statistics/purity of the $t\bar{t} + W$ CR explain that no real gain is observed when adding the $t\bar{t} + V$ control regions.

Splitting signal region in Q(lepton)=+2 and -2

The main background in Table 9.18 comes from $t\bar{t} + W$. A further reduction of this background is possible by splitting the two signal regions in 2 same - sign leptons with a total charge Q of +2 and -2. Indeed

[7] This is slightly worse compared to $\mu_{t\bar{t}H} = 1.0^{+0.81}_{-0.68}$ given in section 9.5.3 because ee, $e\mu$ and $\mu\mu$ channels are merged.

the tree-level $t\bar{t}+W$ production process is the t-channel $u\bar{d} \to W^+g(\to t\bar{t})$, which therefore gives 1/3 of $Q=-2$ and 2/3 of $Q=+2$ because of the proton PDF. As a consequence, in the signal regions with $Q=-2$, the $t\bar{t}+H$, $t\bar{t}+W$, $t\bar{t}+Z$ and fake lepton backgrounds are at the same level. This increases the signal significance, even at the price of a lower statistics and gives $\mu_{t\bar{t}H} = 1.0^{+0.81}_{-0.72}$. In the complementary signal regions with $Q=+2$, the precision is lower, as expected : $\mu_{t\bar{t}H} = 1.0^{+0.99}_{-0.83}$. Combining all signal regions gives a signal sensitivity of :

$$\mu_{t\bar{t}H} = 1.0^{+0.74}_{-0.66} \tag{9.8}$$

which represents a 10% improvement compared to Eq.(9.6), due to the splitting of the signal regions in the lepton charge sum.

Reducing the fake uncertainty and experimental uncertainty (JES, b-tagging)

For Run-2 prospect, the 40% error on the fake and charge-flip estimations, obtained with 8 TeV analysis, is assumed. This could probably be reduced with the full data set available and should give a sub-dominant contribution.

The main experimental uncertainties come from the knowledge of b-tagging and JES. In this study these have been over-simplified, as presented in Table 9.7. A more correct treatment of this systematics will reduce their overall contribution to the signal sensitivity.

Overall possible improvement

Considering all the presented improvements, the total uncertainty on $\mu_{t\bar{t}H}$ with full Run-2 statistics could probably be reduced from ± 0.75 down to $\sim \pm 0.6$.

9.6 Conclusions

In this chapter, prospects studies for measuring the top quark Yukawa coupling with Run-2 full statistics ($\sqrt{s}$ = 13 TeV, 100 fb^{-1}) are discussed. As a results of the optimization, it was found that the highest sensitivity in the $t\bar{t}$+H channel with two same - sign leptons can be achieved by splitting a 5-jets signal region in =1b-jet and $\geq$2b-jets signal regions. This prospect study is based on a simplified but realistic framework for systematics, validated by a detailed comparison with the published ATLAS results at 8 TeV. Monte-Carlo samples are PDF re-weighted and an overall uncertainty of 40% is assigned to fake lepton estimate (as in 8 TeV published results). A precision on $\mu_{t\bar{t}H}$ of $1.0 \pm 0.75 = 1.0 \pm 0.38(\text{stat}) \pm 0.64(\text{syst})$ is obtained, completely dominated by systematics. The dominant uncertainties come from theory, experimental sources (b-tagging and jet energy scale) and fake lepton estimation. They all can be reduced with less conservative hypotheses which could lead to a final sensitivity on $\mu_{t\bar{t}H}$ around 60%. When compared to Run-1 results in this channel, an overall improvement by a factor at least three is obtained. Therefore, a first real test of this coupling can be performed only at the end of Run-2.

Conclusions

The LHC is the most powerful hadron collider in the world designed to collide two proton beams at energies up to 14 TeV in the center of mass. The main goals of the LHC research program are the measurement of the Higgs boson properties and the search for new physics. Both can be performed with the ATLAS experiment. The results presented in this thesis are based on the LHC run-1 data at $\sqrt{s} = 8$ TeV. Few results with the very first (available) runs at 13 TeV are also included for completeness.

In the first part of this thesis the theoretical framework and the ATLAS experiment are introduced, while the following parts are dedicated to the results obtained with the ATLAS data.

The second part starts with in-situ measurements of the electron identification and reconstruction efficiencies which are used to compute electron correction factors propagated to all ATLAS analyses.

The third part of this thesis is dedicated to the search for new physics in the channel with two same - sign or three leptons, jets and missing transverse energy. Dedicated studies are performed to increase the sensitivity of the search by optimizing the signal region definitions and the event selection. On the background side, the methods implemented to estimate the detector component are well improved compared to the first version of the analysis performed with the very first $\sqrt{s} = 7$ TeV data. At the end of Run-1, with this analysis a very wide SUSY space phase is scanned to search for potential *hidden* BSM signals. In the absence of a significant excess, tight mass limits are set on BSM particles. Generally the gluino masses are excluded below $\sim$1 TeV independently on the top squark mass. The bottom squark mass is constrained to be above 600 GeV, with simplified assumptions. These results provide new or very tight constraints on the natural SUSY models.

The fourth part of this thesis is dedicated to prospect studies for the top quark Yukawa coupling measurement with $L = 100$ fb^{-1} of Run-2 data. This coupling is the largest in the Standard Model as it connects two of the heaviest SM particles, the Higgs boson and the top quark. It is very sensitive to beyond the Standard Model physics : any deviation might give an insight on the energy scale of BSM physics. A precise and direct measurement is therefore mandatory and it can be performed through the $t\bar{t} + H$ production mode. With the entire Run-1 data, this coupling could not be constrained given the low production cross section ($\sigma = 130$ fb). However, at $\sqrt{s} = 13$ TeV the cross section increases to 508 fb allowing a precise measurement for the first time. From the possible Higgs boson decays the $H \to WW$ mode is considered with a two same - sign leptons and jets final state as it is one of the most sensitive channels. The prospect studies are performed using MC samples PDF re-weighted and a simplified, but realistic, framework for systematics. For the detector backgrounds similar techniques as for the SUSY analysis are considered. The highest sensitivity is obtained by splitting a 5-jet signal region in ==1b-jet and $\geq$2 b-jets 5 jests signal regions. The final results are obtained with an uncertainty of 60% on μ_{ttH}, improving the Run-1 results by a factor at least three. This individual measurement is essential and after combining the results in all ten possible channels a final precision of 20% - 30% on μ_{ttH} can be reached.

Selected list of publications and conference talks

Journal publications

[1] ATLAS Collaboration (2887 authors), "Search for supersymmetry at $\sqrt{s}$=8 TeV in final states with jets and two same-sign leptons or three leptons with the ATLAS detector", JHEP **1406** (2014) 035, [hep-ex: 1404.2500],

[2] ATLAS Collaboration (2876 authors), "Search for squarks and gluinos in events with isolated leptons, jets and missing transverse momentum at$\sqrt{s} = 8$ TeV with the ATLAS detector", JHEP **1504** (2015) 116, [hep-ex: 1501.03555],

[3] ATLAS Collaboration (2898 authors), "Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton-proton collision data", Eur. Phys. J. **C74** (2014) 2941, [hep-ex: 1404.2240],

[4] ATLAS Collaboration (2855 authors), "Summary of the searches for squarks and gluinos using $\sqrt{s} = 8$ TeV pp collisions with the ATLAS experiment at the LHC", accepted by JHEP (JHEP_207P_0715), [hep-ex: 1507.05525],

[5] ATLAS Collaboration (2856 authors), "ATLAS Run 1 searches for direct pair production of third-generation squarks at the Large Hadron Collider", accepted by Eur. Phys. J., [hep-ex: 1506.08616],

[6] ATLAS Collaboration (2866 authors), "Summary of the ATLAS experiment's sensitivity to supersymmetry after LHC Run 1 - interpreted in the phenomenological MSSM", accepted by JHEP (JHEP_217P_0715), [hep-ex: 1508.06608],

[7] O. Ducu, L. Heurtier and J. Maurer, "LHC signatures of a Z' mediator between dark matter and the SU(3) sector", submitted to JHEP, [hep-ex: 1509.05615].

ATLAS conference notes

[8] ATLAS Collaboration, "Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data", ATLAS-CONF-2014-032, http://cds.cern.ch/record/1706245.

[9] ATLAS Collaboration, "Constraints on promptly decaying supersymmetric particles with lepton-number- and R-parity-violating interactions using Run-1 ATLAS data", ATLAS-CONF-2015-018, http://cds.cern.ch/record/2017303.

[10] ATLAS Collaboration, "Search for strongly produced superpartners in final states with two same sign leptons with the ATLAS detector using 21 fb^{-1} of proton-proton collisions at $\sqrt{s} = 8$ TeV", ATLAS-CONF-2013-007, http://cds.cern.ch/record/1522430.

[11] ATLAS Collaboration, "Search for Supersymmetry in final states with two same-sign leptons, jets and missing transverse momentum with the ATLAS detector in pp collisions at $\sqrt{s} = 8$ TeV", ATLAS-CONF-2012-105, http://cds.cern.ch/record/1472674.

[12] ATLAS Collaboration, "Search for supersymmetry at $\sqrt{s} = 8$ TeV in final states with jets, missing transverse momentum and one isolated lepton", ATLAS-CONF-2012-104, http://cds.cern.ch/record/1472673.

[13] ATLAS Collaboration, "Further search for supersymmetry at $\sqrt{s} = 7$ TeV in final states with jets, missing transverse momentum and one isolated lepton", ATLAS-CONF-2012-041, http://cds.cern.ch/record/1435195.

[14] ATLAS Collaboration, "Search for supersymmetry at $\sqrt{s} = 7$ TeV in final states with large jet multiplicity, missing transverse momentum and one isolated lepton with the ATLAS detector", ATLAS-CONF-2012-140, http://cds.cern.ch/record/1483511.

Presentations in international conferences

[15] O. Ducu [conference talk], "Inclusive searches for squarks and gluinos with the ATLAS detector", Phenomenology Symposium (2014), Pittsburgh (USA).

[16] O. Ducu [poster], "Search for supersymmetry with same-sign leptons and jets with the ATLAS detector", International Conference on new Frontiers in Physics (2013), Kolymbari (Greece).

[17] O. Ducu [poster], "Search for supersymmetry with same-sign leptons and jets with the ATLAS detector using 21 fb^{-1} of data", EPS Conference on High Energy Physics (2013), Stockholm (Sweden).

Conference proceedings

[18] O. Ducu, "SUSY search with two same-sign leptons and jets with the ATLAS detector using 21 fb^{-1} of pp collisions at 8 TeV", EPJ Web of Conferences **71** (2014) 00043, http://dx.doi.org/10.1051/epjconf/20147100043.

[19] O. Ducu, "Supersymmetry search with two same-sign leptons and jets with the ATLAS detector in $\sqrt{s} = 8$ TeV pp collision data", PoS (2013) 052, http://pos.sissa.it/archive/conferences/180/052/EPS-HEP%202013_052.pdf.

ATLAS internal documentation

[20] O. Ducu *et al* (26 authors), "Search for strongly produced superpartners in final states with same-sign leptons or three leptons and jets: preparing for 2015 analyses", ATL-COM-PHYS-2015-329, http://cds.cern.ch/record/2012029.

[21] O. Ducu *et al* (6 authors), "Run-2 prospects for top Yukawa coupling measurement in the ttH channel with two same-sign leptons", ATL-COM-PHYS-2015-479, http://cds.cern.ch/record/2021084.

[22] O. Ducu *et al* (84 authors), "Summary of the ATLAS Run-I searches for direct pair production of third generation squarks at the Large Hadron Collider", ATL-COM-PHYS-2015-102, http://cds.cern.ch/record/1992301.

[23] O. Ducu *et al* (43 authors), "Summary of the searches for squarks and gluinos using $\sqrt{s} = 8$ TeV pp collisions with the ATLAS experiment at the LHC", ATL-COM-PHYS-2014-929, http://cds.cern.ch/record/1746381.

[24] O. Ducu *et al* (8 authors), "Extension of the $\tilde{g} \to t\bar{t}\tilde{\chi}_0^1$ simplified model to include decays with off-shell tops, and exclusion limits from the SS/3 leptons + jets analysis", ATL-COM-PHYS-2014-1422, http://cds.cern.ch/record/1966089.

[25] O. Ducu *et al* (14 authors), "Studies of events with three leptons and three b-tagged jets observed in the full 2012 ATLAS dataset", internal ATLAS documentation ATL-PHYS-INT-2014-006, http://cds.cern.ch/record/1698061.

[26] O. Ducu *et al* (39 authors), "Supporting document on electron efficiency measurements using the 2012 LHC proton-proton collision data", ATL-PHYS-INT-2014-024, http://cds.cern.ch/record/1951872.

[27] O. Ducu *et al* (6 authors), "Impact of TRT transition radiation detection in the search for SUSY with two same-sign leptons or three leptons at $\sqrt{s} = 8$ TeV", ATL-PHYS-INT-2014-003, http://cds.cern.ch/record/1645924.

[28] O. Ducu *et al* (17 authors), "Search for strongly-produced superpartners in final states with two same-sign leptons or three leptons at $\sqrt{s} = 8$ TeV", ATL-COM-PHYS-2013-887, http://cds.cern.ch/record/1558979.

[29] O. Ducu *et al* (19 authors), "Impact of Operating the TRT with Argon-Based Gas Mixtures in ATLAS Performance and Physics Analyses", ATL-INDET-INT-2013-010, http://cds.cern.ch/record/1614495.

[30] O. Ducu *et al* (18 authors), "Search for strongly-produced Supersymmetry with two same-sign leptons at $\sqrt{s} = 8$ TTeV", ATL-PHYS-INT-2013-005, http://cds.cern.ch/record/1528585.

[31] O. Ducu *et al* (7 authors), "Search for Supersymmetry with jets and missing transverse momentum and two same-sign leptons at 8 TeV", ATL-PHYS-INT-2012-067, http://cds.cern.ch/record/1482152.

[32] O. Ducu *et al* (14 authors), "Search for Supersymmetry with jets and missing transverse momentum and one lepton at $\sqrt{s} = 8$ TeV", ATL-COM-PHYS-2012-1033, http://cds.cern.ch/record/1461392.

[33] O. Ducu *et al* (43 authors), "Search for supersymmetry at $\sqrt{s} = 7$ TeV in final states with large jet multiplicity, missing transverse momentum and one isolated lepton with the ATLAS detector", ATL-PHYS-INT-2012-080, http://cds.cern.ch/record/1483520.

[34] O. Ducu *et al* (38 authors), "Search for Supersymmetry with jets and missing transverse momentum and one or more leptons at $\sqrt{s} = 7$ TeV", ATL-PHYS-INT-2012-082, http://cds.cern.ch/record/1488897.

Appendix A

Electron reconstruction and identification efficiency in ATLAS

A.1 Background estimation for electrons with and without associated track

η / p_T [GeV]	[15,20]	[20,25]	[25,30]	[30,35]	[35,40]	[40,45]	[45,50]	[50,60]	[60,80]	[80,100]
[-0.10,0.10]	77.11	52.00	27.35	14.13	6.53	3.49	3.92	7.91	12.59	14.07
[0.10,0.80]	77.30	56.00	30.73	13.84	6.65	3.81	4.10	8.20	13.27	15.36
[0.80,1.37]	76.43	54.89	30.88	14.07	6.38	3.67	4.00	8.19	13.26	14.69
[1.37,1.52]	79.86	56.13	33.75	16.34	7.33	4.34	4.58	8.35	13.54	15.62
[1.52,2.01]	70.85	47.67	27.09	13.47	6.62	3.66	3.91	7.49	12.24	12.91
[2.01,2.47]	68.15	44.88	25.05	12.97	6.88	3.90	3.98	7.28	11.01	10.17

Table A.1: *tight++* tag, 80 GeV $< m_{e\gamma} <$ 100 GeV: estimated electron background fraction at reco + track quality numerator (in %) in data in different bins in η (rows) and p_T (columns) of the probe electron.

η / p_T [GeV]	[15,20]	[20,25]	[25,30]	[30,35]	[35,40]	[40,45]	[45,50]	[50,60]	[60,80]	[80,100]
[-0.10,0.10]	83.44	60.79	33.04	17.17	7.98	4.14	4.42	8.40	13.87	17.31
[0.10,0.80]	83.61	64.45	37.10	17.05	8.13	4.47	4.57	8.89	14.90	19.44
[0.80,1.37]	81.94	62.02	36.88	17.26	7.76	4.37	4.67	9.27	15.35	19.23
[1.37,1.52]	83.15	60.87	38.69	19.04	9.01	5.34	5.15	9.05	15.31	22.71
[1.52,2.01]	74.60	52.42	31.46	16.06	8.02	4.36	4.45	8.12	14.16	17.44
[2.01,2.47]	72.85	50.38	29.86	15.57	8.15	4.46	4.36	8.05	13.10	13.66

Table A.2: *medium++* + $E_T^{\text{cone40}} < 5$ GeV on the tag, 70 GeV $< m_{e\gamma} <$ 105 GeV: estimated electron background fraction at reco + track quality numerator (in %) in data in different bins in η (rows) and p_T (columns) of the probe electron.

η / p_T [GeV]	[15,20]	[20,25]	[25,30]	[30,35]	[35,40]	[40,45]	[45,50]	[50,60]	[60,80]	[80,100]
[-0.10,0.10]	61.84	33.60	14.42	6.88	2.97	1.40	1.59	3.46	5.76	6.72
[0.10,0.80]	61.20	36.88	16.76	6.48	2.92	1.56	1.72	3.61	6.39	8.01
[0.80,1.37]	61.48	37.08	17.18	6.63	2.75	1.47	1.60	3.65	6.43	7.74
[1.37,1.52]	65.64	41.49	20.57	9.08	3.38	2.02	2.17	4.41	7.38	9.41
[1.52,2.01]	54.93	30.27	15.02	6.51	2.92	1.43	1.55	3.19	5.95	7.06
[2.01,2.47]	46.61	24.99	12.00	5.41	2.64	1.40	1.43	2.88	4.63	4.77

Table A.3: *tight++* + $E_T^{\text{cone40}} < 5$ GeV on the tag, 85 GeV $< m_{e\gamma} <$ 95 GeV: estimated electron background fraction at reco + track quality numerator (in %) in data in different bins in η (rows) and p_T (columns) of the probe electron.

Using the selections presented in Section 5.1.4 the and *preliminary Reco* electron background template, the fractions of electron background events at the numerator and electron + photons background events at

η / p_T [GeV]	[15,20]	[20,25]	[25,30]	[30,35]	[35,40]	[40,45]	[45,50]	[50,60]	[60,80]	[80,100]
[-0.10,0.10]	79.29	54.83	29.34	15.33	7.05	3.80	4.27	8.25	13.09	14.59
[0.10,0.80]	79.15	58.40	32.87	15.03	7.23	4.12	4.42	8.76	13.89	16.04
[0.80,1.37]	78.29	56.67	32.57	15.03	6.97	3.96	4.20	8.78	13.76	15.18
[1.37,1.52]	80.95	57.99	35.28	17.04	7.90	4.64	4.79	8.95	14.10	15.22
[1.52,2.01]	72.60	49.52	28.50	14.09	7.01	3.87	4.17	7.80	12.55	13.31
[2.01,2.47]	69.61	45.91	26.04	13.66	7.17	4.18	4.11	7.77	11.25	10.98

Table A.4: *tight++* tag, 80 GeV $< m_{e\gamma} <$ 100 GeV: estimated electron + photon background fraction at denominator (in %) in data in different bins in η (rows) and p_T (columns) of the probe electron.

η / p_T [GeV]	[15,20]	[20,25]	[25,30]	[30,35]	[35,40]	[40,45]	[45,50]	[50,60]	[60,80]	[80,100]
[-0.10,0.10]	85.46	63.58	35.69	18.78	8.76	4.58	4.92	8.78	14.61	18.00
[0.10,0.80]	85.35	67.38	39.81	18.78	8.93	4.92	4.99	9.63	15.60	20.25
[0.80,1.37]	83.53	63.92	39.13	18.64	8.59	4.76	4.97	9.97	16.12	19.83
[1.37,1.52]	83.92	62.73	40.57	20.08	9.73	5.78	5.44	9.76	16.11	22.91
[1.52,2.01]	76.23	54.46	33.12	17.04	8.65	4.70	4.84	8.46	14.73	17.88
[2.01,2.47]	74.31	51.41	31.03	16.43	8.63	4.86	4.61	8.68	13.29	14.58

Table A.5: *medium++* + $E_T^{\text{cone40}} < 5$ GeV on the tag, 70 GeV $< m_{e\gamma} <$ 105 GeV: estimated electron + photon background fraction at denominator (in %) in data in different bins in η (rows) and p_T (columns) of the probe electron.

η / p_T [GeV]	[15,20]	[20,25]	[25,30]	[30,35]	[35,40]	[40,45]	[45,50]	[50,60]	[60,80]	[80,100]
[-0.10,0.10]	65.06	36.61	15.85	7.54	3.27	1.54	1.78	3.57	6.10	7.10
[0.10,0.80]	64.23	39.52	18.37	7.19	3.21	1.71	1.85	3.88	6.72	8.46
[0.80,1.37]	64.09	38.96	18.52	7.22	3.09	1.59	1.67	3.95	6.71	8.05
[1.37,1.52]	68.01	44.00	21.86	9.53	3.74	2.17	2.25	4.79	7.86	9.56
[1.52,2.01]	57.07	32.13	16.16	6.93	3.16	1.53	1.69	3.29	6.09	7.20
[2.01,2.47]	48.45	25.88	12.67	5.80	2.82	1.53	1.49	3.14	4.70	5.23

Table A.6: *tight++* + $E_T^{\text{cone40}} < 5$ GeV on the tag, 85 GeV $< m_{e\gamma} <$ 95 GeV: estimated electron + photon background fraction at denominator (in %) in data in different bins in η (rows) and p_T (columns) of the probe electron.

the denominator are :

- Nominal selection : Table A.1 and Table A.4.
- Loose selection : Table A.2 and Table A.5.
- Tight selection : Table A.3 and Table A.6.

A.2 Measurements at the reconstruction stage

To see the impact of adding the track quality requirement at reconstruction stage the efficiency is measured for reconstructed electrons only. For this measurement Equation (5.8) is simplified to :

$$\varepsilon_{reco} = \frac{N^{baseline} - (B^{baseline})_{nume}}{N^{baseline} - (B^{baseline})_{deno} + \Gamma} \tag{A.1}$$

where $N^{baseline}$ is the number of events under the peak for which the tag fulfills the imposed selection (Section 5.1.2), and the probe electrons are from the baseline selection, with no further cut applied. The terms N^{Fail} and B^{Fail} in Equation (5.8) are zero since the electromagnetic clusters which fails the track association are reconstructed as photons and treated in the Γ term.

The $B^{baseline}$ terms at the numerator (*nume*) and denominator (*deno*) of reconstruction efficiency formula are defined as :

$$(B^{baseline})_{nume} = B^{Templ} \cdot \frac{B^{baseline}_{tail}}{B^{Templ}_{tail}} \tag{A.2}$$

$$(B^{baseline})_{deno} = B^{Templ} \cdot \frac{B_{tail}^{baseline} - N_{tail}^{tight++} \cdot 1/\varepsilon_{tight++}}{B_{tail}^{Templ}} \quad \text{(A.3)}$$

where B^{Templ} and B_{tail}^{Templ} terms are the number of background electrons in the background template (Section 5.1.3), in the peak and tail regions. $B_{tail}^{baseline}$ term is the number of events in the baseline distribution (Section 5.1.2) in the tail region.

η / p_T [GeV]	[15,20]	[20,25]	[25,30]	[30,35]	[35,40]	[40,45]	[45,50]	[50,60]	[60,80]	[80,100]
[-0.10,0.10]	77.66	52.74	27.95	14.42	6.66	3.56	3.99	8.06	12.74	14.21
[0.10,0.80]	78.07	56.90	31.51	14.19	6.82	3.90	4.19	8.37	13.48	15.50
[0.80,1.37]	77.38	56.00	31.89	14.56	6.61	3.78	4.11	8.39	13.55	14.93
[1.37,1.52]	79.97	57.18	34.85	16.83	7.59	4.48	4.70	8.58	13.68	15.61
[1.52,2.01]	71.99	48.77	28.06	13.81	6.80	3.75	3.98	7.63	12.40	13.05
[2.01,2.47]	68.58	45.27	25.39	13.14	6.96	3.94	4.02	7.36	11.08	10.26

Table A.7: *tight++* tag, 80 GeV $< m_{e\gamma} <$ 100 GeV: estimated electron background fraction at reco numerator (in %) in data in different bins in η (rows) and p_T (columns) of the probe electron.

η / p_T [GeV]	[15,20]	[20,25]	[25,30]	[30,35]	[35,40]	[40,45]	[45,50]	[50,60]	[60,80]	[80,100]
[-0.10,0.10]	83.91	61.57	33.81	17.54	8.16	4.24	4.50	8.55	14.04	17.46
[0.10,0.80]	84.26	65.40	38.08	17.51	8.36	4.59	4.69	9.07	15.12	19.63
[0.80,1.37]	82.75	63.14	38.13	17.90	8.06	4.52	4.82	9.51	15.73	19.56
[1.37,1.52]	83.08	61.51	39.89	19.70	9.33	5.52	5.29	9.33	15.56	22.79
[1.52,2.01]	75.67	53.58	32.53	16.54	8.28	4.48	4.55	8.30	14.40	17.70
[2.01,2.47]	73.28	50.85	30.24	15.77	8.25	4.52	4.40	8.14	13.21	13.80

Table A.8: *medium++* + $E_T^{\text{cone40}} < 5$ GeV on the tag, 70 GeV $< m_{e\gamma} <$ 105 GeV: estimated electron background fraction at reco numerator (in %) in data in different bins in η (rows) and p_T (columns) of the probe electron.

η / p_T [GeV]	[15,20]	[20,25]	[25,30]	[30,35]	[35,40]	[40,45]	[45,50]	[50,60]	[60,80]	[80,100]
[-0.10,0.10]	62.62	34.41	14.82	7.04	3.05	1.43	1.62	3.52	5.86	6.85
[0.10,0.80]	62.34	37.80	17.33	6.67	3.00	1.60	1.76	3.67	6.51	8.09
[0.80,1.37]	62.81	38.13	17.94	6.92	2.86	1.51	1.64	3.74	6.58	7.88
[1.37,1.52]	66.57	42.67	21.44	9.47	3.52	2.09	2.21	4.57	7.51	9.76
[1.52,2.01]	56.41	31.41	15.81	6.70	3.01	1.47	1.58	3.26	6.04	7.16
[2.01,2.47]	47.10	25.30	12.22	5.50	2.68	1.42	1.45	2.92	4.67	4.83

Table A.9: *tight++* + $E_T^{\text{cone40}} < 5$ GeV on the tag, 85 GeV $< m_{e\gamma} <$ 95 GeV: estimated electron background fraction at reco numerator (in %) in data in different bins in η (rows) and p_T (columns) of the probe electron.

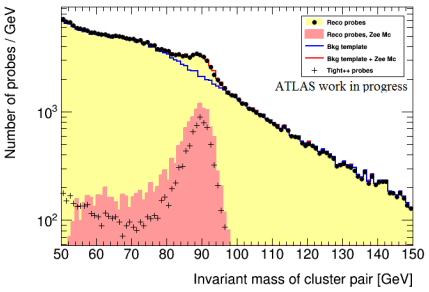
Further, only results using the *preliminary Reco* electron background template are presented, as they are well improved compared to the early results. Using the selections presented in Section 5.1.4 the fractions of electron background events at the numerator and electron + photons background events at the denominator are shown in :

- Nominal selection : Table A.7 and Table A.4.
- Loose selection : Table A.8 and Table A.5.
- Tight selection : Table A.9 and Table A.6.

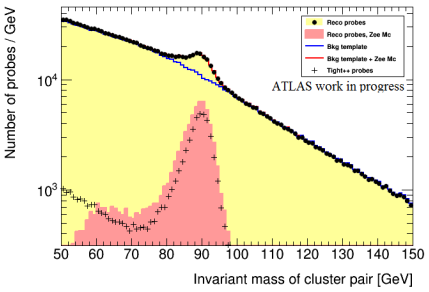
Depending on the (η, p_T) interval and on the selection type the background fraction varies between 85% and 1.5%. The high level of background is obtained especially at low p_T (< 25 GeV). The level of background is comparable with the one obtained at the reconstruction + track quality stage.

The good performance of the electron background templates is shown in Figures A.1 for several (η, p_T) intervals, at the numerator. Here the invariant mass distribution computed with the tag electron and the EM cluster probe is show in data (yellow area). To emphasize the high amount of background, the invariant mass distribution computed with the tag electron and the real probe is shown in MC (red area). The invariant mass computed with the tag and the *tight++* probe is also presented to illustrate the amount of signal and its distribution (cross markers). The electron background template is displayed with a blue line, and the background template + signal $Z \to ee$ MC distribution with a red line.

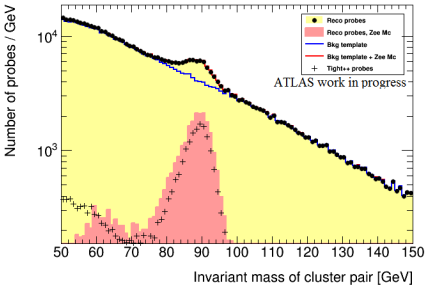
(a) $15 < p_T < 20$ GeV, η *bin [0.0,0.1]*



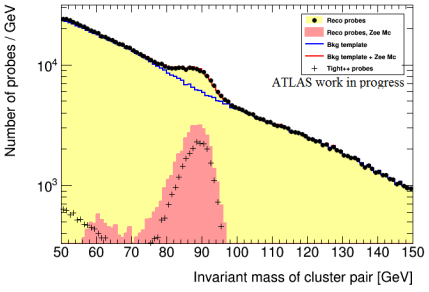
(b) $15 < p_T < 20$ GeV, η *bin [0.1,0.6]*



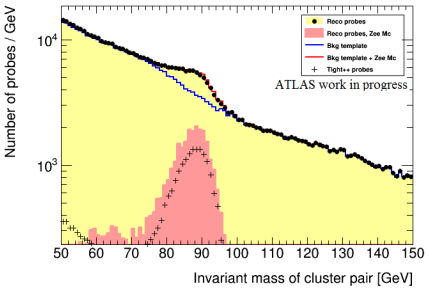
(c) $15 < p_T < 20$ GeV, η *bin [0.6,0.8]*



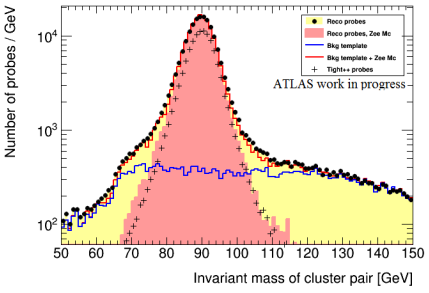
(d) $15 < p_T < 20$ GeV, η *bin [0.8,1.15]*



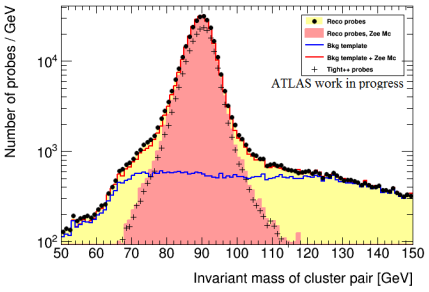
(e) $15 < p_T < 20$ GeV, η *bin [1.15,1.37]*



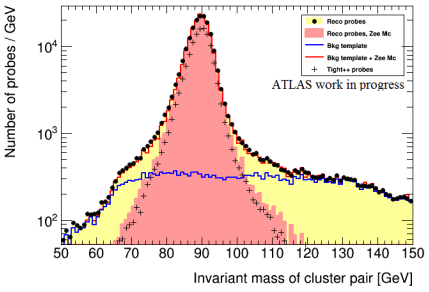
(f) $40 < p_T < 45$ GeV, η *bin [1.37,1.52]*



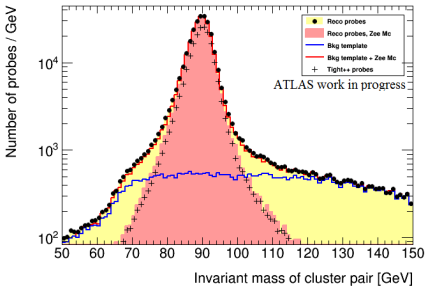
(g) $40 < p_T < 45$ GeV, η *bin [1.52,1.81]*



(h) $40 < p_T < 45$ GeV, η *bin [1.81,2.01]*



(i) $40 < p_T < 45$ GeV, η *bin [2.01,2.37]*



(j) $40 < p_T < 45$ GeV, η *bin [2.37,2.47]*

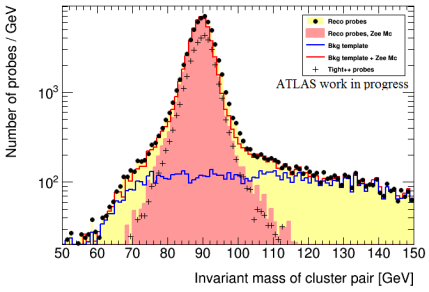


Figure A.1: Invariant mass distribution computed with the tag electron and the EM cluster probe (at the numerator level) in several (η, p_T) bins. $\sqrt{s} = 8$ TeV and $L = 20.3$ fb^{-1}.

Statistical uncertainty

The statistical uncertainty is computed using the error propagation formula. For this, the efficiency Equation (A.1) is rewritten in terms of independent (**uncorrelated**) variables :

$$\varepsilon_{reco} = \frac{N^{baseline} - B^{Templ} \cdot B_{tail}^{baseline}/B_{tail}^{Templ}}{N^{baseline} - B^{Templ} \cdot B_{tail}^{baseline}/B_{tail}^{Templ} + \Gamma} = \frac{NB_{tail}^{Templ} - B^{Templ}B}{NB_{tail}^{Templ} - B^{Templ}B + \Gamma B_{tail}^{Templ}} \tag{A.4}$$

where $N = N^{baseline} - B^{Templ}$ and $B = B_{tail}^{baseline} - B_{tail}^{Templ}$.

The associated statistical uncertainty is :

$$\begin{aligned}\Delta\varepsilon_{reco}^2 = & \frac{1}{(NB_{tail}^{Templ} - BB^{Templ} + \Gamma B_{tail}^{Templ})^4} \times \{[(B^{Templ}B - NB_{tail}^{Templ})B_{tail}^{Templ}\Delta\Gamma]^2 + \\ & (\Gamma B^{Templ}B\Delta B_{tail}^{Templ})^2 + (\Gamma B_{tail}^{Templ})^2 \cdot [(B_{tail}^{Templ}\Delta N)^2 + (B^{Templ}\Delta B)^2 + (B\Delta B^{Templ})^2]\}\end{aligned} \tag{A.5}$$

where the corresponding ΔY is $\sqrt{\sum_i w_i^2}$, where i is the event index and w_i the corresponding weight. For the photon term, w_i is the number of observed events in each invariant mass bin (N_i in Section 5.1.4).

Systematic uncertainty

The same approach and variations as for the reconstruction efficiency measurements are retained (Section 5.2.1).

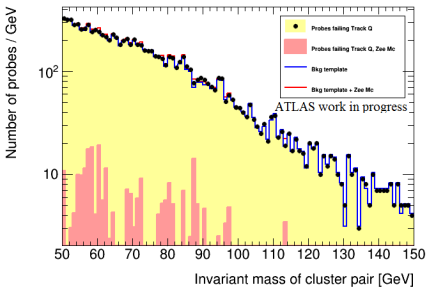
A.3 Results at the reco and reco + track quality stages

In this section the results obtained at reconstruction and reconstruction + track quality stages are presented. Efficiencies are showed in Figure A.3 as a function of η and p_T of the probe electron in data and $Z \to ee$ MC. The scale factors comparison is shown in Figure A.4.

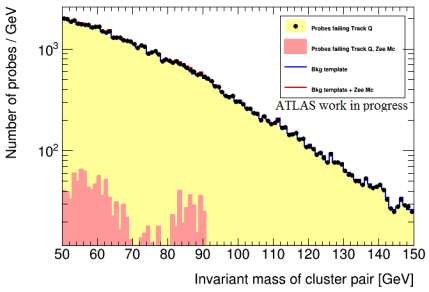
η / p_T [GeV]	[15,20]	[20,25]	[25,30]	[30,35]	[35,40]	[40,45]	[45,50]	[50,60]	[60,80]	[80,100]
[-2.47,-2.40]	$1.000 \pm 0.018 \pm 0.013$	$1.002 \pm 0.008 \pm 0.003$	$1.002 \pm 0.004 \pm 0.003$	$1.012 \pm 0.004 \pm 0.002$	$1.004 \pm 0.003 \pm 0.001$	$1.002 \pm 0.002 \pm 0.001$	$1.005 \pm 0.003 \pm 0.001$	$0.997 \pm 0.004 \pm 0.002$	$1.000 \pm 0.007 \pm 0.004$	$0.992 \pm 0.028 \pm 0.005$
[-2.40,-2.30]	$1.014 \pm 0.014 \pm 0.010$	$1.022 \pm 0.007 \pm 0.005$	$1.006 \pm 0.003 \pm 0.003$	$1.006 \pm 0.003 \pm 0.002$	$1.001 \pm 0.002 \pm 0.002$	$1.001 \pm 0.002 \pm 0.001$	$1.004 \pm 0.002 \pm 0.001$	$1.002 \pm 0.003 \pm 0.001$	$0.996 \pm 0.005 \pm 0.002$	$1.013 \pm 0.014 \pm 0.004$
[-2.30,-2.20]	$0.996 \pm 0.013 \pm 0.011$	$1.000 \pm 0.006 \pm 0.007$	$1.006 \pm 0.003 \pm 0.002$	$1.004 \pm 0.002 \pm 0.001$	$1.001 \pm 0.002 \pm 0.002$	$1.001 \pm 0.002 \pm 0.001$	$0.998 \pm 0.002 \pm 0.001$	$0.999 \pm 0.002 \pm 0.001$	$0.999 \pm 0.005 \pm 0.003$	$1.007 \pm 0.010 \pm 0.003$
[-2.20,-2.10]	$0.981 \pm 0.013 \pm 0.012$	$1.003 \pm 0.006 \pm 0.007$	$0.999 \pm 0.003 \pm 0.003$	$1.003 \pm 0.002 \pm 0.002$	$1.001 \pm 0.002 \pm 0.001$	$1.003 \pm 0.002 \pm 0.001$	$1.001 \pm 0.002 \pm 0.001$	$1.000 \pm 0.003 \pm 0.001$	$1.000 \pm 0.004 \pm 0.002$	$1.003 \pm 0.013 \pm 0.005$
[-2.10,-2.01]	$0.987 \pm 0.014 \pm 0.014$	$1.003 \pm 0.007 \pm 0.009$	$1.000 \pm 0.004 \pm 0.002$	$1.001 \pm 0.003 \pm 0.002$	$1.000 \pm 0.002 \pm 0.001$	$1.002 \pm 0.002 \pm 0.001$	$1.000 \pm 0.002 \pm 0.001$	$1.001 \pm 0.003 \pm 0.001$	$0.995 \pm 0.000 \pm 0.003$	$1.015 \pm 0.016 \pm 0.003$
[-2.01,-1.90]	$0.994 \pm 0.013 \pm 0.010$	$0.997 \pm 0.006 \pm 0.012$	$1.008 \pm 0.003 \pm 0.003$	$1.005 \pm 0.002 \pm 0.001$	$0.996 \pm 0.002 \pm 0.001$	$1.000 \pm 0.002 \pm 0.000$	$0.998 \pm 0.002 \pm 0.002$	$1.001 \pm 0.003 \pm 0.002$	$1.001 \pm 0.005 \pm 0.002$	$1.007 \pm 0.009 \pm 0.003$
[-1.90,-1.80]	$0.982 \pm 0.012 \pm 0.008$	$1.007 \pm 0.006 \pm 0.005$	$1.002 \pm 0.003 \pm 0.003$	$0.998 \pm 0.002 \pm 0.003$	$1.002 \pm 0.002 \pm 0.001$	$0.996 \pm 0.002 \pm 0.000$	$1.000 \pm 0.002 \pm 0.001$	$0.994 \pm 0.003 \pm 0.001$	$0.996 \pm 0.005 \pm 0.002$	$0.992 \pm 0.009 \pm 0.005$
[-1.80,-1.70]	$0.982 \pm 0.013 \pm 0.013$	$1.002 \pm 0.006 \pm 0.004$	$1.002 \pm 0.003 \pm 0.004$	$0.999 \pm 0.003 \pm 0.001$	$0.996 \pm 0.002 \pm 0.001$	$0.997 \pm 0.002 \pm 0.000$	$0.999 \pm 0.002 \pm 0.001$	$0.997 \pm 0.002 \pm 0.001$	$0.999 \pm 0.004 \pm 0.002$	$0.993 \pm 0.006 \pm 0.003$
[-1.70,-1.60]	$1.006 \pm 0.013 \pm 0.012$	$1.002 \pm 0.006 \pm 0.004$	$1.006 \pm 0.003 \pm 0.003$	$0.999 \pm 0.002 \pm 0.002$	$1.003 \pm 0.002 \pm 0.001$	$1.000 \pm 0.002 \pm 0.000$	$1.000 \pm 0.002 \pm 0.001$	$0.996 \pm 0.002 \pm 0.002$	$0.997 \pm 0.004 \pm 0.002$	$1.000 \pm 0.009 \pm 0.004$
[-1.60,-1.52]	$0.987 \pm 0.018 \pm 0.015$	$1.004 \pm 0.008 \pm 0.008$	$1.007 \pm 0.004 \pm 0.004$	$1.007 \pm 0.003 \pm 0.003$	$0.997 \pm 0.002 \pm 0.002$	$1.000 \pm 0.002 \pm 0.001$	$0.999 \pm 0.002 \pm 0.001$	$0.999 \pm 0.003 \pm 0.001$	$0.999 \pm 0.004 \pm 0.001$	$0.993 \pm 0.000 \pm 0.006$
[-1.52,-1.37]	$0.985 \pm 0.016 \pm 0.010$	$1.002 \pm 0.006 \pm 0.006$	$1.005 \pm 0.003 \pm 0.005$	$1.001 \pm 0.002 \pm 0.003$	$0.999 \pm 0.001 \pm 0.001$	$1.001 \pm 0.001 \pm 0.000$	$1.000 \pm 0.001 \pm 0.001$	$0.999 \pm 0.002 \pm 0.001$	$0.997 \pm 0.003 \pm 0.002$	$1.005 \pm 0.007 \pm 0.003$
[-1.37,-1.30]	$1.004 \pm 0.017 \pm 0.015$	$1.003 \pm 0.007 \pm 0.007$	$1.005 \pm 0.003 \pm 0.005$	$0.998 \pm 0.002 \pm 0.003$	$1.000 \pm 0.002 \pm 0.001$	$0.999 \pm 0.001 \pm 0.001$	$0.998 \pm 0.002 \pm 0.001$	$0.998 \pm 0.002 \pm 0.002$	$0.999 \pm 0.004 \pm 0.002$	$0.994 \pm 0.007 \pm 0.003$
[-1.30,-1.20]	$1.012 \pm 0.013 \pm 0.016$	$1.003 \pm 0.006 \pm 0.005$	$1.007 \pm 0.003 \pm 0.004$	$0.997 \pm 0.002 \pm 0.002$	$0.998 \pm 0.001 \pm 0.001$	$1.000 \pm 0.001 \pm 0.000$	$0.999 \pm 0.001 \pm 0.001$	$0.999 \pm 0.002 \pm 0.001$	$1.000 \pm 0.003 \pm 0.002$	$0.994 \pm 0.005 \pm 0.003$
[-1.20,-1.10]	$0.995 \pm 0.017 \pm 0.019$	$1.010 \pm 0.006 \pm 0.009$	$1.004 \pm 0.003 \pm 0.004$	$0.998 \pm 0.002 \pm 0.001$	$1.001 \pm 0.001 \pm 0.001$	$0.999 \pm 0.001 \pm 0.001$	$1.000 \pm 0.001 \pm 0.000$	$0.998 \pm 0.002 \pm 0.001$	$0.999 \pm 0.003 \pm 0.002$	$1.001 \pm 0.006 \pm 0.002$
[-1.10,-1.00]	$0.988 \pm 0.016 \pm 0.019$	$1.012 \pm 0.006 \pm 0.004$	$1.005 \pm 0.003 \pm 0.005$	$1.003 \pm 0.002 \pm 0.001$	$0.999 \pm 0.001 \pm 0.001$	$0.999 \pm 0.001 \pm 0.001$	$0.998 \pm 0.001 \pm 0.001$	$0.999 \pm 0.002 \pm 0.001$	$0.998 \pm 0.003 \pm 0.002$	$1.001 \pm 0.005 \pm 0.003$
[-1.00,-0.90]	$0.993 \pm 0.018 \pm 0.016$	$1.015 \pm 0.006 \pm 0.005$	$1.006 \pm 0.003 \pm 0.003$	$0.998 \pm 0.002 \pm 0.003$	$1.000 \pm 0.001 \pm 0.001$	$0.998 \pm 0.001 \pm 0.001$	$0.999 \pm 0.001 \pm 0.000$	$0.998 \pm 0.002 \pm 0.001$	$0.993 \pm 0.002 \pm 0.001$	$1.000 \pm 0.006 \pm 0.002$
[-0.90,-0.80]	$0.995 \pm 0.016 \pm 0.016$	$1.006 \pm 0.006 \pm 0.010$	$1.003 \pm 0.003 \pm 0.002$	$1.001 \pm 0.002 \pm 0.002$	$1.001 \pm 0.001 \pm 0.001$	$0.998 \pm 0.001 \pm 0.000$	$0.999 \pm 0.001 \pm 0.001$	$0.998 \pm 0.001 \pm 0.001$	$0.998 \pm 0.003 \pm 0.002$	$1.001 \pm 0.006 \pm 0.002$
[-0.80,-0.70]	$0.992 \pm 0.018 \pm 0.017$	$1.009 \pm 0.006 \pm 0.006$	$1.006 \pm 0.003 \pm 0.005$	$1.000 \pm 0.002 \pm 0.002$	$0.998 \pm 0.001 \pm 0.001$	$0.998 \pm 0.001 \pm 0.000$	$0.999 \pm 0.001 \pm 0.001$	$0.997 \pm 0.001 \pm 0.001$	$1.001 \pm 0.003 \pm 0.001$	$0.997 \pm 0.005 \pm 0.002$
[-0.70,-0.60]	$0.970 \pm 0.018 \pm 0.026$	$1.012 \pm 0.006 \pm 0.008$	$1.001 \pm 0.003 \pm 0.004$	$1.002 \pm 0.002 \pm 0.002$	$0.999 \pm 0.001 \pm 0.000$	$1.000 \pm 0.001 \pm 0.000$	$0.998 \pm 0.001 \pm 0.000$	$0.998 \pm 0.002 \pm 0.001$	$0.996 \pm 0.002 \pm 0.002$	$0.995 \pm 0.005 \pm 0.002$
[-0.60,-0.50]	$0.976 \pm 0.017 \pm 0.026$	$1.004 \pm 0.006 \pm 0.013$	$1.003 \pm 0.003 \pm 0.002$	$0.999 \pm 0.002 \pm 0.001$	$1.001 \pm 0.001 \pm 0.001$	$0.999 \pm 0.001 \pm 0.001$	$1.000 \pm 0.001 \pm 0.001$	$0.997 \pm 0.001 \pm 0.001$	$0.996 \pm 0.002 \pm 0.002$	$0.995 \pm 0.004 \pm 0.004$
[-0.50,-0.40]	$1.001 \pm 0.014 \pm 0.018$	$1.008 \pm 0.006 \pm 0.006$	$1.000 \pm 0.003 \pm 0.004$	$1.001 \pm 0.002 \pm 0.001$	$1.002 \pm 0.001 \pm 0.001$	$0.999 \pm 0.001 \pm 0.000$	$1.000 \pm 0.001 \pm 0.000$	$0.998 \pm 0.001 \pm 0.001$	$0.999 \pm 0.002 \pm 0.001$	$0.997 \pm 0.005 \pm 0.002$
[-0.40,-0.30]	$0.997 \pm 0.014 \pm 0.019$	$1.005 \pm 0.007 \pm 0.005$	$1.002 \pm 0.003 \pm 0.003$	$0.999 \pm 0.002 \pm 0.001$	$0.999 \pm 0.001 \pm 0.000$	$0.999 \pm 0.001 \pm 0.000$	$0.998 \pm 0.001 \pm 0.001$	$1.000 \pm 0.002 \pm 0.001$	$0.995 \pm 0.002 \pm 0.001$	$0.993 \pm 0.005 \pm 0.003$
[-0.30,-0.20]	$1.013 \pm 0.011 \pm 0.005$	$1.008 \pm 0.005 \pm 0.006$	$1.003 \pm 0.003 \pm 0.002$	$0.999 \pm 0.002 \pm 0.003$	$0.999 \pm 0.001 \pm 0.001$	$1.000 \pm 0.001 \pm 0.000$	$0.997 \pm 0.001 \pm 0.001$	$1.000 \pm 0.002 \pm 0.001$	$0.998 \pm 0.003 \pm 0.001$	$1.006 \pm 0.005 \pm 0.002$
[-0.20,-0.10]	$1.009 \pm 0.013 \pm 0.006$	$0.992 \pm 0.006 \pm 0.007$	$1.003 \pm 0.003 \pm 0.003$	$1.000 \pm 0.002 \pm 0.002$	$1.000 \pm 0.001 \pm 0.001$	$1.000 \pm 0.001 \pm 0.000$	$0.999 \pm 0.001 \pm 0.000$	$0.998 \pm 0.002 \pm 0.001$	$0.996 \pm 0.002 \pm 0.002$	$0.998 \pm 0.004 \pm 0.002$
[-0.10,0.00]	$1.011 \pm 0.011 \pm 0.005$	$0.991 \pm 0.006 \pm 0.009$	$1.006 \pm 0.003 \pm 0.002$	$1.000 \pm 0.002 \pm 0.002$	$1.000 \pm 0.001 \pm 0.001$	$1.000 \pm 0.001 \pm 0.000$	$0.998 \pm 0.001 \pm 0.000$	$0.996 \pm 0.001 \pm 0.001$	$0.998 \pm 0.003 \pm 0.002$	$0.997 \pm 0.005 \pm 0.003$
[0.00,0.10]	$1.013 \pm 0.010 \pm 0.005$	$0.999 \pm 0.005 \pm 0.009$	$0.999 \pm 0.003 \pm 0.002$	$1.001 \pm 0.002 \pm 0.001$	$1.001 \pm 0.001 \pm 0.001$	$1.000 \pm 0.001 \pm 0.001$	$0.999 \pm 0.001 \pm 0.001$	$0.996 \pm 0.002 \pm 0.001$	$0.998 \pm 0.003 \pm 0.001$	$1.000 \pm 0.004 \pm 0.001$
[0.10,0.20]	$1.010 \pm 0.010 \pm 0.009$	$0.996 \pm 0.006 \pm 0.005$	$0.996 \pm 0.003 \pm 0.002$	$1.001 \pm 0.002 \pm 0.002$	$1.000 \pm 0.001 \pm 0.001$	$1.000 \pm 0.001 \pm 0.000$	$1.000 \pm 0.001 \pm 0.001$	$0.998 \pm 0.002 \pm 0.001$	$0.998 \pm 0.003 \pm 0.001$	$1.001 \pm 0.005 \pm 0.002$
[0.20,0.30]	$1.007 \pm 0.015 \pm 0.011$	$1.000 \pm 0.006 \pm 0.005$	$1.005 \pm 0.003 \pm 0.003$	$1.000 \pm 0.002 \pm 0.002$	$0.999 \pm 0.001 \pm 0.001$	$0.998 \pm 0.001 \pm 0.000$	$1.000 \pm 0.001 \pm 0.000$	$0.997 \pm 0.001 \pm 0.001$	$1.000 \pm 0.003 \pm 0.001$	$1.005 \pm 0.005 \pm 0.003$
[0.30,0.40]	$0.986 \pm 0.015 \pm 0.029$	$1.014 \pm 0.005 \pm 0.003$	$1.003 \pm 0.003 \pm 0.002$	$1.000 \pm 0.002 \pm 0.001$	$1.000 \pm 0.001 \pm 0.001$	$1.000 \pm 0.001 \pm 0.000$	$0.999 \pm 0.001 \pm 0.001$	$0.998 \pm 0.001 \pm 0.001$	$0.999 \pm 0.003 \pm 0.001$	$0.997 \pm 0.005 \pm 0.002$
[0.40,0.50]	$1.004 \pm 0.016 \pm 0.009$	$1.010 \pm 0.005 \pm 0.005$	$1.000 \pm 0.003 \pm 0.002$	$1.002 \pm 0.002 \pm 0.001$	$1.001 \pm 0.001 \pm 0.002$	$0.999 \pm 0.001 \pm 0.000$	$1.000 \pm 0.001 \pm 0.001$	$1.000 \pm 0.002 \pm 0.002$	$0.999 \pm 0.003 \pm 0.002$	$0.998 \pm 0.005 \pm 0.002$
[0.50,0.60]	$0.963 \pm 0.016 \pm 0.027$	$1.008 \pm 0.006 \pm 0.007$	$1.003 \pm 0.003 \pm 0.005$	$1.001 \pm 0.002 \pm 0.002$	$1.001 \pm 0.001 \pm 0.001$	$1.001 \pm 0.001 \pm 0.000$	$1.000 \pm 0.001 \pm 0.001$	$0.998 \pm 0.002 \pm 0.001$	$1.003 \pm 0.003 \pm 0.001$	$0.996 \pm 0.004 \pm 0.002$
[0.60,0.70]	$0.989 \pm 0.018 \pm 0.019$	$1.009 \pm 0.006 \pm 0.008$	$1.006 \pm 0.003 \pm 0.003$	$1.000 \pm 0.002 \pm 0.001$	$1.000 \pm 0.001 \pm 0.001$	$0.999 \pm 0.001 \pm 0.000$	$0.998 \pm 0.001 \pm 0.001$	$0.999 \pm 0.002 \pm 0.001$	$1.002 \pm 0.003 \pm 0.002$	$0.992 \pm 0.005 \pm 0.002$
[0.70,0.80]	$0.978 \pm 0.019 \pm 0.018$	$1.013 \pm 0.005 \pm 0.003$	$1.003 \pm 0.003 \pm 0.004$	$1.002 \pm 0.002 \pm 0.001$	$1.000 \pm 0.001 \pm 0.000$	$1.000 \pm 0.001 \pm 0.000$	$0.999 \pm 0.001 \pm 0.001$	$0.998 \pm 0.002 \pm 0.001$	$1.001 \pm 0.003 \pm 0.001$	$0.998 \pm 0.005 \pm 0.002$
[0.80,0.90]	$1.004 \pm 0.013 \pm 0.011$	$1.010 \pm 0.006 \pm 0.010$	$1.009 \pm 0.003 \pm 0.004$	$1.001 \pm 0.002 \pm 0.001$	$0.999 \pm 0.001 \pm 0.001$	$0.999 \pm 0.001 \pm 0.000$	$0.999 \pm 0.001 \pm 0.000$	$1.000 \pm 0.002 \pm 0.001$	$0.996 \pm 0.003 \pm 0.002$	$0.999 \pm 0.006 \pm 0.003$
[0.90,1.00]	$0.987 \pm 0.017 \pm 0.014$	$1.007 \pm 0.006 \pm 0.009$	$1.003 \pm 0.003 \pm 0.003$	$1.001 \pm 0.002 \pm 0.001$	$1.000 \pm 0.001 \pm 0.001$	$1.000 \pm 0.001 \pm 0.001$	$0.998 \pm 0.001 \pm 0.000$	$0.996 \pm 0.002 \pm 0.001$	$0.997 \pm 0.003 \pm 0.002$	$1.003 \pm 0.006 \pm 0.004$
[1.00,1.10]	$1.001 \pm 0.017 \pm 0.014$	$1.007 \pm 0.006 \pm 0.006$	$1.002 \pm 0.003 \pm 0.004$	$0.998 \pm 0.002 \pm 0.001$	$1.001 \pm 0.001 \pm 0.001$	$0.998 \pm 0.001 \pm 0.001$	$0.998 \pm 0.001 \pm 0.000$	$0.999 \pm 0.002 \pm 0.001$	$0.997 \pm 0.003 \pm 0.002$	$1.003 \pm 0.005 \pm 0.002$
[1.10,1.20]	$1.030 \pm 0.013 \pm 0.008$	$1.006 \pm 0.006 \pm 0.009$	$1.006 \pm 0.003 \pm 0.003$	$0.997 \pm 0.002 \pm 0.002$	$1.001 \pm 0.001 \pm 0.001$	$0.999 \pm 0.001 \pm 0.001$	$0.999 \pm 0.001 \pm 0.001$	$0.999 \pm 0.002 \pm 0.002$	$0.998 \pm 0.003 \pm 0.002$	$0.993 \pm 0.007 \pm 0.003$
[1.20,1.30]	$0.994 \pm 0.016 \pm 0.016$	$1.012 \pm 0.006 \pm 0.004$	$1.006 \pm 0.003 \pm 0.005$	$1.003 \pm 0.002 \pm 0.001$	$1.000 \pm 0.001 \pm 0.001$	$0.998 \pm 0.001 \pm 0.001$	$0.999 \pm 0.001 \pm 0.001$	$0.998 \pm 0.002 \pm 0.001$	$1.006 \pm 0.004 \pm 0.001$	$0.997 \pm 0.006 \pm 0.001$
[1.30,1.37]	$1.004 \pm 0.017 \pm 0.013$	$1.002 \pm 0.007 \pm 0.009$	$1.006 \pm 0.004 \pm 0.003$	$0.997 \pm 0.003 \pm 0.002$	$1.000 \pm 0.002 \pm 0.001$	$1.000 \pm 0.001 \pm 0.001$	$1.001 \pm 0.002 \pm 0.001$	$0.994 \pm 0.002 \pm 0.001$	$0.998 \pm 0.004 \pm 0.001$	$0.994 \pm 0.007 \pm 0.003$
[1.37,1.52]	$0.988 \pm 0.017 \pm 0.015$	$0.992 \pm 0.006 \pm 0.008$	$1.003 \pm 0.003 \pm 0.004$	$0.999 \pm 0.002 \pm 0.004$	$1.001 \pm 0.001 \pm 0.001$	$1.001 \pm 0.001 \pm 0.001$	$0.999 \pm 0.001 \pm 0.001$	$1.004 \pm 0.002 \pm 0.001$	$0.998 \pm 0.003 \pm 0.002$	$1.008 \pm 0.009 \pm 0.006$
[1.52,1.60]	$0.964 \pm 0.018 \pm 0.025$	$1.001 \pm 0.007 \pm 0.006$	$0.997 \pm 0.004 \pm 0.005$	$1.003 \pm 0.003 \pm 0.001$	$1.000 \pm 0.002 \pm 0.001$	$1.001 \pm 0.002 \pm 0.000$	$1.003 \pm 0.002 \pm 0.002$	$0.996 \pm 0.003 \pm 0.001$	$0.993 \pm 0.005 \pm 0.003$	$0.993 \pm 0.008 \pm 0.004$
[1.60,1.70]	$1.001 \pm 0.014 \pm 0.014$	$1.002 \pm 0.006 \pm 0.003$	$1.008 \pm 0.004 \pm 0.003$	$0.996 \pm 0.002 \pm 0.003$	$0.999 \pm 0.002 \pm 0.001$	$0.995 \pm 0.002 \pm 0.001$	$0.998 \pm 0.002 \pm 0.001$	$1.000 \pm 0.003 \pm 0.001$	$0.991 \pm 0.004 \pm 0.003$	$1.000 \pm 0.011 \pm 0.003$
[1.70,1.80]	$0.970 \pm 0.012 \pm 0.007$	$1.005 \pm 0.006 \pm 0.007$	$1.002 \pm 0.003 \pm 0.002$	$0.996 \pm 0.003 \pm 0.002$	$0.997 \pm 0.002 \pm 0.001$	$0.998 \pm 0.002 \pm 0.001$	$0.996 \pm 0.002 \pm 0.001$	$0.995 \pm 0.003 \pm 0.001$	$1.001 \pm 0.004 \pm 0.003$	$0.992 \pm 0.006 \pm 0.004$
[1.80,1.90]	$0.994 \pm 0.012 \pm 0.013$	$1.001 \pm 0.006 \pm 0.005$	$1.006 \pm 0.003 \pm 0.003$	$0.998 \pm 0.003 \pm 0.001$	$0.997 \pm 0.002 \pm 0.000$	$0.998 \pm 0.002 \pm 0.001$	$0.995 \pm 0.002 \pm 0.001$	$0.994 \pm 0.003 \pm 0.001$	$0.990 \pm 0.004 \pm 0.002$	$1.004 \pm 0.008 \pm 0.002$
[1.90,2.01]	$1.007 \pm 0.014 \pm 0.007$	$1.007 \pm 0.006 \pm 0.004$	$1.007 \pm 0.004 \pm 0.004$	$1.007 \pm 0.003 \pm 0.002$	$1.002 \pm 0.002 \pm 0.001$	$0.999 \pm 0.002 \pm 0.000$	$0.998 \pm 0.002 \pm 0.001$	$0.997 \pm 0.003 \pm 0.002$	$1.005 \pm 0.005 \pm 0.002$	$0.987 \pm 0.008 \pm 0.007$
[2.01,2.10]	$0.981 \pm 0.014 \pm 0.011$	$1.000 \pm 0.007 \pm 0.006$	$1.006 \pm 0.004 \pm 0.002$	$1.005 \pm 0.003 \pm 0.003$	$1.003 \pm 0.002 \pm 0.001$	$1.002 \pm 0.002 \pm 0.001$	$1.003 \pm 0.002 \pm 0.001$	$1.001 \pm 0.003 \pm 0.002$	$0.988 \pm 0.004 \pm 0.002$	$1.004 \pm 0.009 \pm 0.004$
[2.10,2.20]	$1.013 \pm 0.013 \pm 0.008$	$1.003 \pm 0.006 \pm 0.003$	$1.007 \pm 0.003 \pm 0.004$	$1.004 \pm 0.002 \pm 0.001$	$1.001 \pm 0.002 \pm 0.001$	$1.001 \pm 0.002 \pm 0.001$	$1.000 \pm 0.002 \pm 0.001$	$1.000 \pm 0.003 \pm 0.001$	$0.992 \pm 0.005 \pm 0.003$	$1.005 \pm 0.012 \pm 0.005$
[2.20,2.30]	$0.997 \pm 0.014 \pm 0.013$	$1.004 \pm 0.006 \pm 0.007$	$1.001 \pm 0.003 \pm 0.001$	$1.003 \pm 0.002 \pm 0.002$	$1.002 \pm 0.002 \pm 0.001$	$1.002 \pm 0.002 \pm 0.001$	$1.001 \pm 0.002 \pm 0.001$	$0.999 \pm 0.003 \pm 0.001$	$0.993 \pm 0.004 \pm 0.002$	$0.988 \pm 0.007 \pm 0.005$
[2.30,2.40]	$1.021 \pm 0.014 \pm 0.006$	$1.004 \pm 0.006 \pm 0.005$	$1.003 \pm 0.003 \pm 0.002$	$1.004 \pm 0.002 \pm 0.001$	$1.004 \pm 0.002 \pm 0.001$	$1.003 \pm 0.002 \pm 0.001$	$1.000 \pm 0.002 \pm 0.001$	$1.002 \pm 0.003 \pm 0.002$	$0.997 \pm 0.005 \pm 0.002$	$1.000 \pm 0.017 \pm 0.005$
[2.40,2.47]	$1.013 \pm 0.017 \pm 0.023$	$1.003 \pm 0.008 \pm 0.009$	$1.015 \pm 0.005 \pm 0.002$	$1.005 \pm 0.003 \pm 0.001$	$1.005 \pm 0.003 \pm 0.002$	$1.004 \pm 0.002 \pm 0.001$	$0.995 \pm 0.003 \pm 0.001$	$1.000 \pm 0.005 \pm 0.003$	$1.002 \pm 0.009 \pm 0.004$	$1.008 \pm 0.013 \pm 0.004$

Table A.10: Scale factors at reconstruction + track quality level, in several (η, p_T) bins. Both statistical and systematic uncertainties are shown.

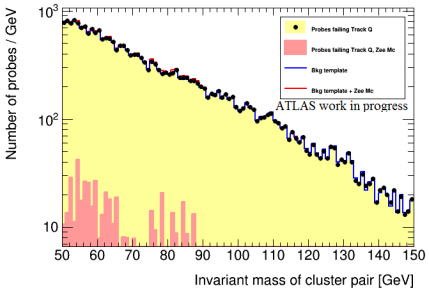
(a) $15 < p_T < 20$ GeV, η *bin [0.0,0.1]*



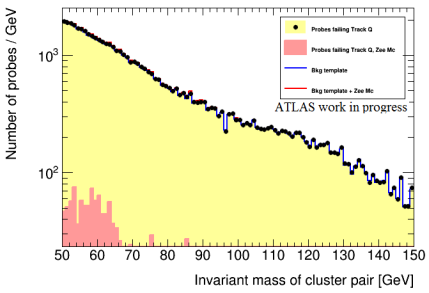
(b) $15 < p_T < 20$ GeV, η *bin [0.1,0.6]*



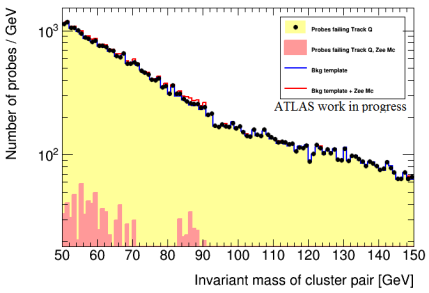
(c) $15 < p_T < 20$ GeV, η *bin [0.6,0.8]*



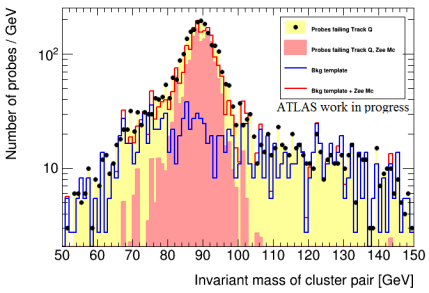
(d) $15 < p_T < 20$ GeV, η *bin [0.8,1.15]*



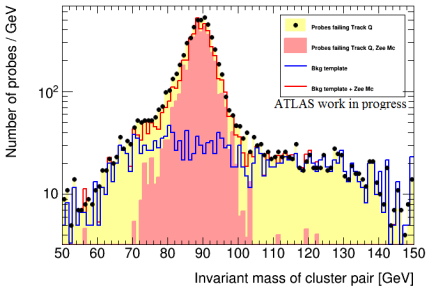
(e) $15 < p_T < 20$ GeV, η *bin [1.15,1.37]*



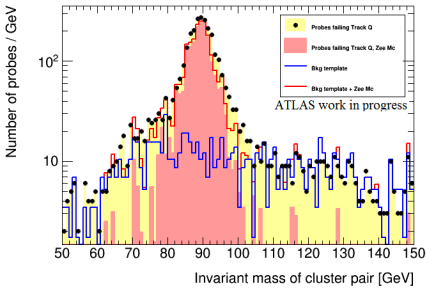
(f) $40 < p_T < 45$ GeV, η *bin [1.37,1.52]*



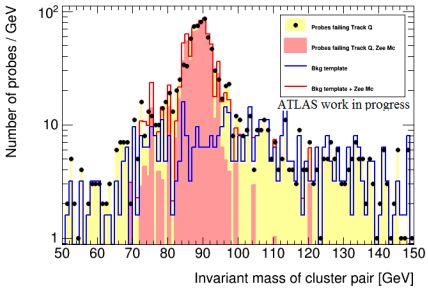
(g) $40 < p_T < 45$ GeV, η *bin [1.52,1.81]*



(h) $40 < p_T < 45$ GeV, η *bin [1.81,2.01]*



(i) $40 < p_T < 45$ GeV, η *bin [2.01,2.37]*



(j) $40 < p_T < 45$ GeV, η *bin [2.37,2.47]*

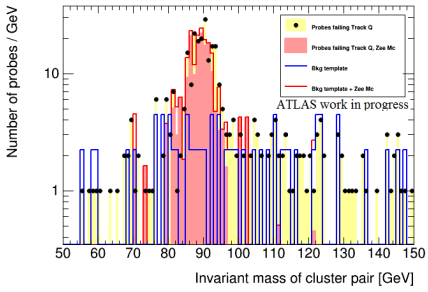


Figure A.2: Invariant mass distribution computed with the tag electron and EM cluster probe failing the track quality cut (at the numerator level) in several (η, p_T) bins. $\sqrt{s} = 8$ TeV and $L = 20.3$ fb^{-1}.

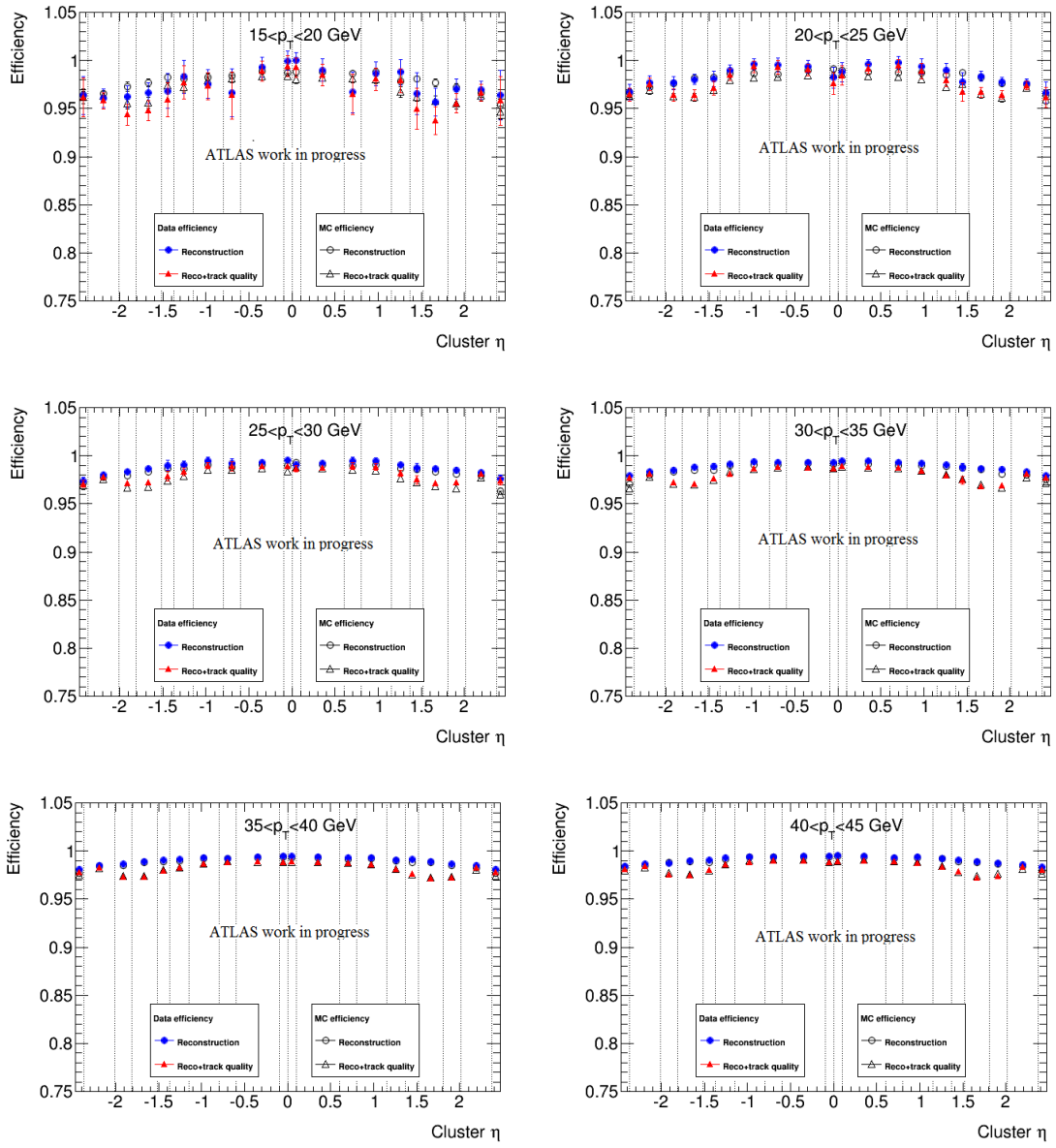


Figure A.3: Electron reconstruction and reconstruction + track quality efficiencies in data (filled markers) and MC (empty markers) in several (η, p_T) bins. $\sqrt{s} = 8$ TeV and $L = 20.3$ fb^{-1}. Only combined statistical and systematic uncertainties are shown.

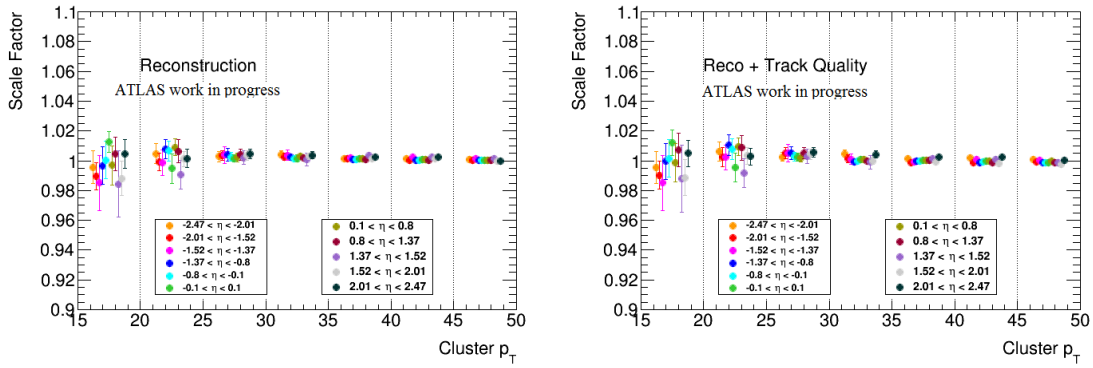


Figure A.4: Electron scale factors at reconstruction (left) and reconstruction + track quality (right) stages. $\sqrt{s} = 8$ TeV and $L = 20.3$ fb^{-1}. Only the total uncertainty is considered.

A.4 Electron identification measurements

Electron identification efficiencies using $Z \to ee$ events

In this appendix the electron identification efficiencies in data obtained with the Z mass method are presented in Tables A.13 - A.12 for *tight++*, *medium++* and *loose++* electron identification criteria.

Electron identification measurements using J/Ψ processes

Two methods are used to obtain the identification efficiencies [168]:

- For the first, only events with short pseudo-life time, -1 ps $< \tau <$ 0.2 ps, are considered. This sub-sample is dominated by prompt J/Ψ processes production. The amount of non-prompt events is smaller than 20% decreasing with the p_T of the probe electron. In data, a clear difference between the two processes cannot be achieved. For MC measurements the fraction of non-prompt J/Ψ processes, f_{NP}, is an important variable as it is used to combine the prompt and non-prompt efficiencies. This combination is essential as each efficiency is more or less dependent on the requirements imposed to define the tag or the probe electron, ie. isolation, p_T range. It is measured using the number of prompt and non-prompt simulated events, in different probe electron p_T intervals before applying any identification cuts.

  As for the $Z \to ee$ mass method, the shape of the background should be well described. Its composition is given by Ψ_{2s}, χ_c states or random combination of two leptons not coming from one decay. The estimation is done fitting the invariant mass of the electron pair, using a sum of analytical and Crystal-Ball functions[1].

  The final MC efficiency in each (η, p_T) bin is a combination between the prompt (P) and non-prompt (NP) efficiencies, written as $\varepsilon_P \cdot (1 - f_{NP}) + \varepsilon_{NP} \cdot f_{NP}$. The efficiencies are the ratio between the number of events for which the probe is passing the required selection (ie. electron identification, isolation) and total number of events in the selected sample. In the region with low p_T electrons the two ε_{NP} and ε_P are very close, leading to a final efficiency similar to ε_{NP}, as f_{NP} is ∼10 - 16%. In the opposite region they vary more and the amount of NP processes increase leading to differences between the total and prompt efficiencies.

- Lifetime-fit method implies longer pseudo-life times, -1 ps $< \tau <$ 3 ps. This looser interval is used to discriminate between the prompt and non-prompt J/Ψ processes after fitting the τ variable. Other tighter intervals are used to constrain the region used for the measurement and assign the systematic uncertainty.

  The total signal yield is obtain after the invariant mass of the probe electron pair is fitted using Crystal-Ball and Gaussian functions. To decrease the number of random combination of two electrons not coming from a decay to each event a weight is applied : +1 for all OS electrons combinations and -1 to all SS electrons pairs. The remained background is estimated with 1^{st} or 2^{nd} order Chebychev polynomial functions.

  Results are obtained for 12 different configurations with and without an isolation cut on the tag electron, and looser ranges for the τ variable.

[1] For OS leptons background 1^{st} or 2^{nd} order Chebychev or exponential functions are used. Crystal-Ball functions are implemented to model the signal and the Ψ_{2s} contribution.

p_T [GeV] / η	[-2.47,-2.37]	[-2.37,-2.01]	[-2.01,-1.81]	[-1.81,-1.52]	[-1.52,-1.37]	[-1.37,-1.15]	[-1.15,-0.8]	[-0.8,-0.6]	[-0.6,-0.1]	[-0.1,0]	[0,0.1]	[0.1,0.6]	[0.6,0.8]	[0.8,1.15]	[1.15,1.37]	[1.37,1.52]	[1.52,1.81]	[1.81,2.01]	[2.01,2.37]	[2.37,2.47]
tight++ data efficiency measurements																				
[15,20]	70.50	76.18	66.68	71.44	61.99	70.44	73.27	73.99	74.67	65.79	62.56	74.64	74.23	73.67	69.25	63.18	70.36	67.00	76.65	69.09
[20,25]	69.68	72.66	66.49	67.24	60.07	70.23	71.29	74.55	74.02	66.13	62.01	74.24	75.07	70.86	68.93	60.40	67.28	66.74	72.54	69.76
[25,30]	70.13	74.91	68.68	70.84	63.70	73.48	74.74	77.71	77.97	70.70	68.52	78.38	77.99	74.91	72.95	63.33	70.72	69.23	74.57	69.30
[30,35]	69.72	75.18	71.78	73.79	65.17	74.87	77.37	81.04	80.77	73.30	71.78	81.05	81.90	77.97	74.80	65.24	72.99	71.61	75.34	68.14
[35,40]	68.71	75.97	72.73	75.82	68.07	77.63	79.49	82.56	82.36	74.97	73.18	82.39	83.42	79.85	77.35	68.27	75.47	73.28	76.01	67.45
[40,45]	69.79	78.55	75.63	78.56	72.37	79.79	80.17	83.67	80.90	74.81	73.12	80.47	84.40	80.67	79.69	72.28	78.10	75.91	78.25	67.24
[45,50]	67.37	80.75	77.75	80.98	75.96	81.73	82.12	84.98	82.56	76.04	75.14	82.45	85.65	82.37	81.85	75.73	80.73	78.12	80.41	65.43
[50,60]	66.48	81.41	79.75	81.93	77.48	84.16	84.29	86.31	84.56	77.74	76.66	84.58	87.47	84.88	84.00	78.00	81.91	79.82	80.94	64.51
[60,80]	65.12	81.57	79.90	83.75	79.29	85.23	85.77	86.69	86.38	78.11	76.73	86.31	88.28	85.63	85.25	79.66	84.13	80.71	80.14	63.18
[80,8000]	85.15	88.37	87.74	90.22	83.73	87.04	89.68	90.37	89.20	78.79	80.77	89.23	90.69	89.10	89.63	84.83	89.25	86.32	90.81	88.69
tight++ statistical uncertainty																				
[15,20]	1.42	0.46	0.61	0.60	1.53	0.71	0.68	0.75	0.47	1.00	0.96	0.47	0.74	0.67	0.72	1.70	0.58	0.62	0.48	1.33
[20,25]	0.77	0.31	0.41	0.36	0.70	0.42	0.37	0.48	0.26	0.52	0.48	0.26	0.48	0.36	0.41	0.73	0.36	0.41	0.31	0.83
[25,30]	0.57	0.24	0.32	0.29	0.49	0.32	0.25	0.33	0.19	0.39	0.37	0.19	0.33	0.25	0.31	0.50	0.28	0.32	0.25	0.61
[30,35]	0.46	0.21	0.27	0.22	0.35	0.23	0.17	0.21	0.13	0.31	0.31	0.14	0.21	0.18	0.23	0.34	0.22	0.27	0.21	0.50
[35,40]	0.38	0.15	0.20	0.15	0.23	0.15	0.10	0.13	0.08	0.20	0.20	0.08	0.13	0.10	0.14	0.23	0.14	0.19	0.16	0.41
[40,45]	0.35	0.13	0.16	0.12	0.19	0.11	0.08	0.10	0.06	0.16	0.16	0.06	0.10	0.08	0.11	0.18	0.12	0.15	0.13	0.32
[45,50]	0.39	0.15	0.20	0.14	0.23	0.14	0.10	0.13	0.08	0.20	0.21	0.08	0.13	0.10	0.14	0.22	0.14	0.19	0.15	0.38
[50,60]	0.58	0.22	0.29	0.21	0.36	0.20	0.15	0.19	0.11	0.28	0.29	0.11	0.18	0.15	0.20	0.31	0.21	0.27	0.23	0.62
[60,80]	1.09	0.46	0.64	0.37	0.78	0.38	0.29	0.33	0.22	0.47	0.51	0.21	0.34	0.27	0.38	0.77	0.41	0.58	0.46	1.31
[80,8000]	2.55	1.15	1.65	1.27	2.51	0.89	0.73	0.85	0.54	1.22	1.25	0.58	0.81	0.66	1.03	2.70	1.17	1.51	1.90	5.51
tight++ systematic uncertainty																				
[15,20]	1.26	0.89	0.67	1.32	1.16	1.20	2.19	3.95	5.18	5.63	3.75	5.04	3.82	1.56	0.70	0.97	1.22	0.61	0.56	2.14
[20,25]	0.42	0.67	0.87	2.05	2.51	1.53	2.23	2.61	2.65	1.71	0.99	2.09	2.25	2.22	1.17	1.24	2.32	1.00	0.66	1.07
[25,30]	0.85	0.89	0.72	1.48	2.30	1.46	1.62	1.58	1.54	1.57	1.19	1.62	1.32	1.52	1.17	1.60	1.61	1.00	1.15	1.86
[30,35]	0.36	0.61	0.54	1.05	1.18	0.57	0.41	0.80	0.65	0.52	0.62	0.88	0.62	0.85	0.47	0.79	0.89	0.51	0.68	1.04
[35,40]	0.55	0.14	0.30	0.34	0.46	0.28	0.16	0.13	0.25	0.09	0.18	0.21	0.28	0.22	0.16	0.31	0.20	0.25	0.28	0.75
[40,45]	0.46	0.17	0.07	0.08	0.29	0.08	0.02	0.11	0.12	0.06	0.12	0.05	0.16	0.07	0.09	0.16	0.16	0.08	0.15	0.13
[45,50]	00.09	0.07	0.08	0.06	0.18	0.14	0.06	0.08	0.14	0.07	0.21	0.08	0.05	0.10	0.10	0.16	0.08	0.08	0.10	0.23
[50,60]	0.17	0.07	0.30	0.11	0.15	0.08	0.09	0.18	0.13	0.26	0.26	0.21	0.10	0.15	0.10	0.18	0.19	0.06	0.06	0.32
[60,80]	0.27	0.28	0.16	0.31	0.47	0.36	0.39	0.24	0.23	0.18	0.13	0.25	0.28	0.10	0.08	0.64	0.32	0.71	0.33	0.52
[80,8000]	1.39	2.07	1.71	1.95	0.79	1.59	0.39	0.36	0.41	1.17	0.23	0.19	0.47	0.60	0.80	1.49	0.70	0.56	3.80	5.03

Table A.11: *tight++*: Early electron identification efficiency together with statistical and systematic uncertainties. All values are given in %.

	[-2.47,-2.37]	[-2.37,-2.01]	[-2.01,-1.81]	[-1.81,-1.52]	[-1.52,-1.37]	[-1.37,-1.15]	[-1.15,-0.8]	[-0.8,-0.6]	[-0.6,-0.1]	[-0.1,0]	[0,0.1]	[0.1,0.6]	[0.6,0.8]	[0.8,1.15]	[1.15,1.37]	[1.37,1.52]	[1.52,1.81]	[1.81,2.01]	[2.01,2.37]	[2.37,2.47]
medium++ data efficiency measurements																				
[15,20]	91.42	87.20	82.40	85.63	84.88	86.69	87.81	84.20	84.74	81.55	77.23	84.84	84.81	87.67	85.61	87.00	83.90	82.81	88.14	89.60
[20,25]	85.97	81.37	81.83	80.63	75.19	83.38	82.88	83.76	82.52	80.86	76.19	82.62	84.13	81.83	82.35	75.97	80.37	82.26	81.67	87.09
[25,30]	86.20	84.37	84.29	83.80	77.39	86.73	85.59	86.56	86.51	85.01	83.05	86.71	86.50	85.53	86.02	77.47	83.55	85.09	84.43	86.90
[30,35]	85.69	85.08	86.18	85.62	77.03	88.29	87.34	88.85	88.92	87.66	86.30	89.05	89.49	87.90	87.93	77.29	84.82	85.88	85.14	84.41
[35,40]	84.95	86.28	87.40	87.73	79.69	90.37	89.42	90.19	90.46	89.17	87.63	90.39	90.97	89.71	89.95	80.11	87.35	87.65	86.39	83.71
[40,45]	84.41	87.74	88.41	90.16	83.92	91.51	90.19	91.39	91.43	90.43	89.17	91.30	91.98	90.47	91.19	84.00	89.67	88.75	87.64	81.75
[45,50]	80.98	89.44	89.96	91.80	87.23	92.55	91.47	92.29	92.33	91.12	90.08	92.46	92.69	91.65	92.29	86.75	91.43	90.01	89.49	78.57
[50,60]	78.63	89.69	90.60	92.18	87.87	92.91	92.09	92.50	92.92	91.24	90.31	93.05	93.08	92.42	92.61	88.04	91.96	90.55	89.43	76.70
[60,80]	77.90	90.40	90.74	93.22	90.05	94.10	93.29	92.48	94.12	91.08	89.97	93.85	93.93	92.95	93.44	89.46	93.15	91.30	89.28	75.91
[80,8000]	93.77	94.35	95.83	96.78	90.22	93.73	96.46	95.31	95.87	92.63	93.79	95.80	95.88	95.02	95.90	94.11	95.72	93.83	96.56	97.07
medium++ statistical uncertainty																				
[15,20]	2.21	0.50	0.69	0.74	3.30	0.91	0.90	0.89	0.56	1.28	1.18	0.56	0.88	0.87	0.94	3.73	0.71	0.71	0.53	2.05
[20,25]	0.99	0.32	0.43	0.40	1.18	0.48	0.43	0.53	0.28	0.59	0.53	0.28	0.54	0.42	0.48	1.24	0.40	0.43	0.32	1.11
[25,30]	0.68	0.25	0.33	0.31	0.75	0.35	0.28	0.34	0.20	0.42	0.40	0.20	0.34	0.28	0.34	0.77	0.30	0.34	0.25	0.81
[30,35]	0.52	0.21	0.28	0.22	0.46	0.24	0.17	0.21	0.13	0.32	0.32	0.14	0.21	0.18	0.24	0.46	0.23	0.27	0.21	0.56
[35,40]	0.40	0.14	0.18	0.14	0.27	0.13	0.10	0.12	0.08	0.18	0.19	0.07	0.12	0.09	0.13	0.27	0.13	0.17	0.15	0.43
[40,45]	0.34	0.11	0.13	0.10	0.20	0.09	0.07	0.09	0.05	0.12	0.13	0.05	0.09	0.07	0.09	0.19	0.10	0.13	0.12	0.30
[45,50]	0.35	0.13	0.16	0.11	0.25	0.12	0.09	0.11	0.06	0.14	0.16	0.06	0.11	0.08	0.11	0.23	0.11	0.16	0.13	0.35
[50,60]	0.56	0.20	0.26	0.18	0.43	0.17	0.13	0.17	0.10	0.23	0.25	0.10	0.16	0.14	0.18	0.35	0.18	0.23	0.21	0.64
[60,80]	1.36	0.43	0.62	0.34	1.00	0.36	0.29	0.31	0.20	0.42	0.48	0.20	0.32	0.25	0.36	0.90	0.41	0.56	0.45	1.97
[80,8000]	2.26	1.13	1.70	1.25	2.59	0.89	0.76	0.80	0.52	1.10	1.19	0.56	0.77	0.65	1.06	3.64	1.14	1.42	1.87	5.20
medium++ systematic uncertainty																				
[15,20]	2.69	1.09	0.63	1.44	4.27	1.38	3.43	5.00	6.88	8.14	6.02	7.09	5.28	2.73	0.92	4.69	0.89	0.84	0.79	3.48
[20,25]	0.67	0.86	1.07	2.42	4.41	2.19	3.02	3.42	3.36	2.68	1.24	2.64	3.01	2.90	1.65	2.70	2.91	1.47	0.84	1.77
[25,30]	1.29	1.14	0.89	1.74	3.20	1.86	2.09	1.89	1.85	1.95	1.51	1.95	1.58	1.90	1.68	2.96	1.83	1.39	1.40	3.31
[30,35]	0.62	0.70	0.71	1.16	1.52	0.72	0.47	0.92	0.78	0.65	0.76	1.01	0.74	1.02	0.60	1.22	0.97	0.57	0.84	1.33
[35,40]	0.71	0.17	0.39	0.41	0.57	0.36	0.19	0.13	0.30	0.10	0.22	0.25	0.35	0.26	0.20	0.53	0.19	0.34	0.34	1.00
[40,45]	0.61	0.18	0.06	0.10	0.31	0.10	0.03	0.13	0.15	0.07	0.14	0.06	0.17	0.08	0.12	0.24	0.17	0.12	0.17	0.15
[45,50]	0.07	0.06	0.10	0.07	0.22	0.16	0.06	0.09	0.16	0.06	0.25	0.09	0.05	0.12	0.09	0.20	0.07	0.09	0.12	0.24
[50,60]0.21	0.09	0.34	0.13	0.25	0.08	0.11	0.19	0.14	0.29	0.33	0.24	0.09	0.16	0.10	0.23	0.22	0.10	0.07	0.27	
[60,80]	0.30	0.22	0.22	0.20	0.42	0.44	0.46	0.28	0.26	0.20	0.21	0.24	0.26	0.09	0.11	0.43	0.42	0.78	0.43	1.25
[80,8000]	0.71	2.17	1.81	1.85	1.06	1.61	0.46	0.32	0.46	1.54	0.37	0.23	0.54	0.68	0.97	1.63	0.49	0.64	3.71	3.50

Table A.12: *medium++*: Electron identification efficiency together with statistical and systematic uncertainties. All values are given in %.

	[-2.47,-2.37]	[-2.37,-2.01]	[-2.01,-1.81]	[-1.81,-1.52]	[-1.52,-1.37]	[-1.37,-1.15]	[-1.15,-0.8]	[-0.8,-0.6]	[-0.6,-0.1]	[-0.1,0.0]	[0,0.1]	[0.1,0.6]	[0.6,0.8]	[0.8,1.15]	[1.15,1.37]	[1.37,1.52]	[1.52,1.81]	[1.81,2.01]	[2.01,2.37]	[2.37,2.47]
loose++ data efficiency measurements																				
[15,20]	102.39	95.59	95.25	97.77	103.45	97.65	100.06	94.02	93.32	90.85	87.37	93.42	93.78	100.14	96.88	107.08	96.10	95.34	96.41	100.88
[20,25]	97.24	95.66	94.40	92.67	91.01	96.12	97.48	94.63	94.82	93.41	89.90	94.85	94.74	96.69	94.99	91.22	92.62	94.64	95.81	99.69
[25,30]	97.84	96.72	95.36	94.84	92.48	96.58	97.11	95.54	96.21	95.80	94.06	96.46	95.11	97.24	96.18	91.57	94.30	95.80	97.21	100.46
[30,35]	97.46	95.52	95.92	94.46	90.23	95.91	95.70	95.36	95.95	94.98	94.22	96.16	95.63	96.17	95.69	90.02	93.66	95.58	95.79	97.91
[35,40]	98.29	95.66	96.41	95.44	91.90	97.02	96.64	95.84	96.57	95.70	94.95	96.49	96.35	96.67	96.65	91.95	95.18	96.59	96.15	98.76
[40,45]	98.62	95.28	96.56	96.06	94.10	97.30	96.48	96.45	96.35	95.77	95.24	96.14	96.75	96.49	97.14	94.11	95.89	96.73	95.63	98.08
[45,50]	98.33	95.90	97.27	97.02	95.89	97.99	97.18	97.22	97.00	96.52	96.15	96.98	97.23	97.19	97.76	95.74	96.95	97.35	96.39	98.08
[50,60]	98.57	96.30	97.65	97.40	96.25	98.15	97.46	97.59	97.45	96.83	96.97	97.61	97.50	97.64	98.03	96.25	97.46	97.67	96.65	98.64
[60,80]	99.27	97.19	97.44	98.40	97.86	98.97	98.67	97.72	98.70	97.39	97.31	98.28	98.57	98.06	98.44	96.97	98.48	98.11	96.74	100.28
[80,8000]	97.64	98.48	100.43	100.47	97.47	97.38	100.12	98.34	99.29	96.34	97.43	98.91	98.71	98.89	99.15	100.26	98.71	98.34	101.22	101.05
loose++ statistical uncertainty																				
[15,20]	2.33	0.48	0.68	0.77	3.85	0.94	0.94	0.91	0.57	1.33	1.22	0.57	0.89	0.92	0.96	4.42	0.73	0.70	0.51	2.17
[20,25]	1.02	0.29	0.41	0.40	1.36	0.48	0.44	0.53	0.29	0.59	0.54	0.29	0.54	0.43	0.48	1.42	0.40	0.41	0.30	1.16
[25,30]	0.69	0.23	0.31	0.31	0.84	0.35	0.27	0.34	0.19	0.41	0.39	0.19	0.34	0.28	0.34	0.86	0.29	0.32	0.24	0.85
[30,35]	0.52	0.19	0.26	0.22	0.50	0.23	0.16	0.20	0.13	0.31	0.31	0.13	0.20	0.17	0.23	0.49	0.22	0.26	0.19	0.56
[35,40]	0.36	0.12	0.16	0.12	0.27	0.11	0.08	0.10	0.07	0.16	0.17	0.06	0.11	0.08	0.11	0.27	0.11	0.15	0.13	0.41
[40,45]	0.29	0.09	0.10	0.08	0.18	0.07	0.05	0.07	0.04	0.10	0.10	0.04	0.07	0.05	0.08	0.17	0.08	0.09	0.09	0.22
[45,50]	0.23	0.10	0.12	0.09	0.22	0.09	0.06	0.08	0.05	0.11	0.13	0.05	0.09	0.06	0.08	0.20	0.09	0.12	0.10	0.21
[50,60]	0.43	0.17	0.22	0.16	0.42	0.14	0.11	0.14	0.08	0.19	0.21	0.08	0.14	0.12	0.15	0.34	0.16	0.18	0.17	0.55
[60,80]	1.37	0.39	0.59	0.31	1.03	0.33	0.27	0.28	0.19	0.37	0.44	0.18	0.29	0.23	0.33	0.92	0.38	0.52	0.41	2.19
[80,8000]	2.16	1.08	1.68	1.23	2.70	0.84	0.74	0.77	0.50	1.05	1.15	0.54	0.74	0.62	1.02	3.77	1.11	1.38	1.88	4.87
loose++ systematic uncertainty																				
[15,20]	2.43	0.90	0.56	1.96	5.29	1.93	3.77	5.36	6.83	8.61	6.22	7.03	5.29	2.86	1.42	6.43	0.86	1.02	0.57	3.09
[20,25]	0.71	1.01	1.12	2.32	4.46	2.49	3.21	3.59	3.57	2.77	1.26	2.75	3.05	3.24	1.69	2.42	3.18	1.57	1.03	1.76
[25,30]	1.11	1.12	0.93	1.66	3.28	1.90	2.25	2.01	1.97	2.08	1.58	2.10	1.61	2.04	1.74	2.93	1.79	1.46	1.52	3.53
[30,35]	0.57	0.75	0.70	1.07	1.46	0.69	0.44	0.96	0.81	0.65	0.78	1.05	0.73	1.07	0.56	1.20	0.83	0.47	0.83	1.40
[35,40]	0.74	0.14	0.40	0.40	0.58	0.39	0.18	0.11	0.31	0.11	0.20	0.25	0.36	0.26	0.20	0.49	0.13	0.34	0.31	1.19
[40,45]	0.72	0.18	0.04	0.08	0.26	0.08	0.03	0.14	0.15	0.08	0.14	0.06	0.19	0.08	0.13	0.21	0.17	0.13	0.17	0.16
[45,50]	0.06	0.03	0.07	0.06	0.17	0.16	0.05	0.08	0.16	0.07	0.27	0.10	0.06	0.12	0.09	0.15	0.07	0.09	0.10	0.26
[50,60]	0.24	0.14	0.41	0.11	0.21	0.08	0.10	0.16	0.15	0.27	0.34	0.25	0.06	0.14	0.08	0.20	0.23	0.08	0.07	0.25
[60,80]	0.26	0.30	0.30	0.27	0.44	0.46	0.44	0.27	0.26	0.17	0.26	0.27	0.29	0.09	0.13	0.27	0.41	0.82	0.44	1.32
[80,8000]	0.72	2.31	2.02	2.02	0.85	1.47	0.56	0.31	0.43	1.57	0.43	0.26	0.47	0.65	1.01	1.50	0.47	0.81	4.34	2.54

Table A.13: *loose++*: Electron identification efficiency together with statistical and systematic uncertainties. All values are given in %.

p_T [GeV] / η	[0.00,0.10]	[0.10,0.80]	[0.80,1.37]	[1.37,1.52]	[1.52,2.01]	[2.01,2.47]
[5,7]	$0.920 \pm 0.031 \pm 0.034$	$0.961 \pm 0.016 \pm 0.015$	$0.908 \pm 0.028 \pm 0.022$	$0.874 \pm 0.101 \pm 0.075$	$0.980 \pm 0.037 \pm 0.034$	$0.988 \pm 0.024 \pm 0.025$
[7,10]	$0.937 \pm 0.032 \pm 0.029$	$0.961 \pm 0.014 \pm 0.018$	$0.967 \pm 0.024 \pm 0.026$	$1.030 \pm 0.114 \pm 0.087$	$1.005 \pm 0.037 \pm 0.032$	$1.028 \pm 0.028 \pm 0.032$
[10,15]	$0.952 \pm 0.025 \pm 0.010$	$0.946 \pm 0.010 \pm 0.006$	$0.970 \pm 0.015 \pm 0.010$	$0.961 \pm 0.049 \pm 0.031$	$0.990 \pm 0.022 \pm 0.010$	$0.949 \pm 0.020 \pm 0.016$
[15,20]	$0.949 \pm 0.014 \pm 0.019$	$0.943 \pm 0.005 \pm 0.018$	$0.973 \pm 0.006 \pm 0.023$	$1.024 \pm 0.016 \pm 0.017$	$1.000 \pm 0.007 \pm 0.020$	$0.956 \pm 0.006 \pm 0.027$

(a)
h!

Table A.14: Combined scale factors for *tight++* in [5,20] GeV p_T interval. Both statistic and systematic uncertainties are shown.

p_T [GeV] / η	[-2.47, -2.37]	[-2.37, -2.01]	[-2.01, -1.81]	[-1.81, - 1.52]	[-1.52, -1.37]	[-1.37, - 1.15]	[-1.15, -0.80]	[-0.80, -0.60]	[-0.60, -0.10]	[-0.10, 0.00]
[20,25]	$0.961 \pm 0.015 \pm 0.039$	$0.920 \pm 0.006 \pm 0.026$	$0.968 \pm 0.010 \pm 0.030$	$0.963 \pm 0.008 \pm 0.047$	$1.004 \pm 0.014 \pm 0.019$	$1.015 \pm 0.010 \pm 0.027$	$0.974 \pm 0.007 \pm 0.027$	$0.957 \pm 0.009 \pm 0.023$	$0.950 \pm 0.005 \pm 0.020$	$0.972 \pm 0.013 \pm 0.019$
[25,30]	$0.964 \pm 0.011 \pm 0.013$	$0.942 \pm 0.004 \pm 0.014$	$0.980 \pm 0.007 \pm 0.014$	$0.985 \pm 0.006 \pm 0.023$	$0.987 \pm 0.009 \pm 0.039$	$0.995 \pm 0.006 \pm 0.013$	$0.975 \pm 0.005 \pm 0.013$	$0.966 \pm 0.006 \pm 0.011$	$0.953 \pm 0.003 \pm 0.012$	$0.976 \pm 0.008 \pm 0.011$
[30,35]	$0.966 \pm 0.009 \pm 0.011$	$0.940 \pm 0.004 \pm 0.005$	$0.983 \pm 0.006 \pm 0.004$	$0.984 \pm 0.004 \pm 0.009$	$0.996 \pm 0.007 \pm 0.019$	$0.991 \pm 0.005 \pm 0.006$	$0.979 \pm 0.003 \pm 0.003$	$0.976 \pm 0.004 \pm 0.003$	$0.961 \pm 0.002 \pm 0.003$	$0.979 \pm 0.006 \pm 0.004$
[35,40]	$0.945 \pm 0.008 \pm 0.006$	$0.950 \pm 0.003 \pm 0.003$	$0.994 \pm 0.005 \pm 0.003$	$0.996 \pm 0.004 \pm 0.004$	$1.010 \pm 0.006 \pm 0.014$	$1.010 \pm 0.003 \pm 0.003$	$0.987 \pm 0.002 \pm 0.002$	$0.987 \pm 0.003 \pm 0.002$	$0.969 \pm 0.002 \pm 0.001$	$0.992 \pm 0.005 \pm 0.004$
[40,45]	$0.929 \pm 0.007 \pm 0.002$	$0.961 \pm 0.003 \pm 0.001$	$0.996 \pm 0.004 \pm 0.002$	$1.003 \pm 0.003 \pm 0.003$	$1.011 \pm 0.005 \pm 0.015$	$1.010 \pm 0.003 \pm 0.002$	$0.982 \pm 0.002 \pm 0.002$	$0.984 \pm 0.002 \pm 0.001$	$0.957 \pm 0.001 \pm 0.001$	$0.981 \pm 0.004 \pm 0.002$
[45,50]	$0.874 \pm 0.009 \pm 0.003$	$0.969 \pm 0.003 \pm 0.003$	$0.999 \pm 0.005 \pm 0.003$	$1.011 \pm 0.004 \pm 0.004$	$1.029 \pm 0.006 \pm 0.013$	$1.013 \pm 0.004 \pm 0.002$	$0.985 \pm 0.002 \pm 0.001$	$0.987 \pm 0.003 \pm 0.001$	$0.961 \pm 0.002 \pm 0.002$	$0.986 \pm 0.005 \pm 0.003$
[50,60]	$0.851 \pm 0.012 \pm 0.004$	$0.961 \pm 0.005 \pm 0.003$	$0.996 \pm 0.007 \pm 0.004$	$1.004 \pm 0.005 \pm 0.005$	$1.020 \pm 0.007 \pm 0.011$	$1.011 \pm 0.004 \pm 0.003$	$0.990 \pm 0.003 \pm 0.001$	$0.980 \pm 0.004 \pm 0.002$	$0.961 \pm 0.002 \pm 0.003$	$0.989 \pm 0.007 \pm 0.004$
[60,70]	$0.848 \pm 0.021 \pm 0.004$	$0.974 \pm 0.009 \pm 0.006$	$0.981 \pm 0.010 \pm 0.003$	$1.003 \pm 0.007 \pm 0.004$	$1.019 \pm 0.012 \pm 0.013$	$1.003 \pm 0.007 \pm 0.003$	$0.986 \pm 0.005 \pm 0.003$	$0.985 \pm 0.006 \pm 0.006$	$0.970 \pm 0.004 \pm 0.004$	$0.997 \pm 0.011 \pm 0.009$
[70,80]	$0.848 \pm 0.021 \pm 0.004$	$0.974 \pm 0.009 \pm 0.006$	$0.981 \pm 0.010 \pm 0.003$	$1.003 \pm 0.007 \pm 0.004$	$1.019 \pm 0.012 \pm 0.013$	$1.003 \pm 0.007 \pm 0.003$	$0.986 \pm 0.005 \pm 0.003$	$0.985 \pm 0.006 \pm 0.006$	$0.970 \pm 0.004 \pm 0.004$	$0.997 \pm 0.011 \pm 0.009$
[80,100]	$1.053 \pm 0.047 \pm 0.042$	$0.986 \pm 0.015 \pm 0.003$	$0.994 \pm 0.017 \pm 0.017$	$1.001 \pm 0.012 \pm 0.009$	$1.000 \pm 0.021 \pm 0.014$	$1.027 \pm 0.013 \pm 0.005$	$0.987 \pm 0.008 \pm 0.010$	$0.992 \pm 0.010 \pm 0.005$	$0.971 \pm 0.006 \pm 0.006$	$1.003 \pm 0.021 \pm 0.013$

p_T [GeV] / η	[0.00,0.10]	[0.10,0.60]	[0.60,0.80]	[0.80,1.15]	[1.15,1.37]	[1.37,1.52]	[1.52,1.81]	[1.81,2.01]	[2.01,2.37]	[2.37,2.47]
[20,25]	$0.955 \pm 0.013 \pm 0.014$	$0.958 \pm 0.005 \pm 0.019$	$0.950 \pm 0.009 \pm 0.026$	$0.968 \pm 0.007 \pm 0.020$	$0.973 \pm 0.009 \pm 0.028$	$0.996 \pm 0.014 \pm 0.090$	$0.961 \pm 0.008 \pm 0.040$	$0.987 \pm 0.010 \pm 0.027$	$0.926 \pm 0.006 \pm 0.030$	$0.917 \pm 0.013 \pm 0.040$
[25,30]	$0.992 \pm 0.009 \pm 0.006$	$0.955 \pm 0.003 \pm 0.011$	$0.965 \pm 0.006 \pm 0.011$	$0.976 \pm 0.005 \pm 0.011$	$0.999 \pm 0.006 \pm 0.013$	$0.996 \pm 0.009 \pm 0.038$	$0.968 \pm 0.006 \pm 0.016$	$0.992 \pm 0.008 \pm 0.013$	$0.936 \pm 0.005 \pm 0.013$	$0.930 \pm 0.010 \pm 0.016$
[30,35]	$0.985 \pm 0.006 \pm 0.005$	$0.959 \pm 0.002 \pm 0.003$	$0.973 \pm 0.003 \pm 0.003$	$0.985 \pm 0.003 \pm 0.003$	$0.999 \pm 0.005 \pm 0.005$	$1.010 \pm 0.007 \pm 0.018$	$0.974 \pm 0.004 \pm 0.009$	$0.997 \pm 0.006 \pm 0.006$	$0.940 \pm 0.004 \pm 0.004$	$0.933 \pm 0.009 \pm 0.009$
[35,40]	$0.985 \pm 0.005 \pm 0.003$	$0.966 \pm 0.002 \pm 0.001$	$0.981 \pm 0.003 \pm 0.002$	$0.988 \pm 0.002 \pm 0.002$	$1.010 \pm 0.004 \pm 0.003$	$1.009 \pm 0.005 \pm 0.014$	$1.000 \pm 0.004 \pm 0.004$	$0.990 \pm 0.005 \pm 0.002$	$0.947 \pm 0.003 \pm 0.003$	$0.901 \pm 0.008 \pm 0.006$
[40,45]	$0.982 \pm 0.004 \pm 0.002$	$0.953 \pm 0.001 \pm 0.001$	$0.989 \pm 0.002 \pm 0.002$	$0.988 \pm 0.002 \pm 0.001$	$1.013 \pm 0.003 \pm 0.002$	$1.013 \pm 0.005 \pm 0.014$	$0.996 \pm 0.003 \pm 0.003$	$1.000 \pm 0.004 \pm 0.002$	$0.956 \pm 0.003 \pm 0.001$	$0.900 \pm 0.007 \pm 0.003$
[45,50]	$0.986 \pm 0.005 \pm 0.002$	$0.959 \pm 0.002 \pm 0.001$	$0.986 \pm 0.003 \pm 0.001$	$0.987 \pm 0.002 \pm 0.001$	$1.019 \pm 0.004 \pm 0.001$	$1.023 \pm 0.006 \pm 0.013$	$1.005 \pm 0.004 \pm 0.002$	$0.998 \pm 0.005 \pm 0.003$	$0.968 \pm 0.004 \pm 0.002$	$0.870 \pm 0.009 \pm 0.003$
[50,60]	$0.988 \pm 0.007 \pm 0.002$	$0.958 \pm 0.002 \pm 0.001$	$0.988 \pm 0.003 \pm 0.002$	$0.987 \pm 0.003 \pm 0.002$	$1.006 \pm 0.004 \pm 0.002$	$1.024 \pm 0.007 \pm 0.012$	$1.003 \pm 0.005 \pm 0.003$	$1.005 \pm 0.007 \pm 0.003$	$0.961 \pm 0.005 \pm 0.002$	$0.840 \pm 0.012 \pm 0.005$
[60,70]	$0.992 \pm 0.011 \pm 0.002$	$0.969 \pm 0.004 \pm 0.005$	$0.990 \pm 0.006 \pm 0.003$	$0.996 \pm 0.005 \pm 0.003$	$1.013 \pm 0.007 \pm 0.006$	$1.024 \pm 0.012 \pm 0.017$	$1.013 \pm 0.008 \pm 0.004$	$0.999 \pm 0.011 \pm 0.006$	$0.965 \pm 0.009 \pm 0.006$	$0.803 \pm 0.020 \pm 0.016$
[70,80]	$0.992 \pm 0.011 \pm 0.002$	$0.969 \pm 0.004 \pm 0.005$	$0.990 \pm 0.006 \pm 0.003$	$0.996 \pm 0.005 \pm 0.003$	$1.013 \pm 0.007 \pm 0.006$	$1.024 \pm 0.012 \pm 0.017$	$1.013 \pm 0.008 \pm 0.004$	$0.999 \pm 0.011 \pm 0.006$	$0.965 \pm 0.009 \pm 0.006$	$0.803 \pm 0.020 \pm 0.016$
[80,100]	$1.012 \pm 0.019 \pm 0.005$	$0.989 \pm 0.018 \pm 0.006$	$0.977 \pm 0.006 \pm 0.004$	$1.010 \pm 0.010 \pm 0.004$	$1.007 \pm 0.009 \pm 0.007$	$1.023 \pm 0.012 \pm 0.004$	$1.004 \pm 0.012 \pm 0.009$	$1.017 \pm 0.019 \pm 0.012$	$0.992 \pm 0.015 \pm 0.005$	$1.034 \pm 0.035 \pm 0.018$

Table A.15: Combined scale factors for *tight++* in [20, 150] GeV p_T interval. Both statistic and systematic uncertainties are shown.

Appendix B

Search for SUSY particles in events with two same sign leptons

B.1 Simultaneous fit method

As mentioned in Section 6.8.1 the simultaneous fit method is used to statistically test the compatibility between data and background estimation in the defined signal regions. Several input histograms (nominal) corresponding to different samples of the PDF in each region are used. To correctly account for any correlation, it is possible to treat the systematic uncertainties on each input histogram in different ways [265]. For some of these methods used to apply the systematics, two other histograms are needed, containing the distributions after applying the upward (Up) and downward (Down) variations of a systematic. Further the used methods are enunciated :

- *HistoSys* : it applies the systematic as a shape, correlated across the bins of the nominal histogram, retaining the original scaling histograms for the $+/-1$ sigma variation of the nominal (the Up and Down histograms). It can change both the normalization and the shape. In this analysis it is used for JES, JER, etc. systematics.
- *OverallSys* : it is treated with a Gaussian constrain, calculating the relative affect of the $+/-1$ sigma variations of the systematic source on the integral of the histogram. The systematic is applied as a global scaling factor, fully correlated across the histogram bins, but without using (changing) the shape. This method is used in the analysis for the systematics sources associated to lepton reconstruction, identification, pile-up, b-tagging, fake lepton background, etc.
- *UserOverallSys* : it is used with the up and down variations numerically defined by the user. It is further applied as an overall scaling factor, and acts as the OverallSys type. In this analysis it is used for the theory uncertainty sources associated to $t\bar{t}+ V$, VV, etc.
- *OverallHistoSys* : it combines the affects of *overallSys* and *histoSys* as one systematic. The histograms representing the $+/-1$ sigma variation of the systematic source are normalized to the integral of the nominal histogram, and the scaling factors required to perform the normalization are applied as an overallSys.
- *NormHistoSys*: similar treatment of systematics as for the *HistoSys*type, but the normalization is retained constant while the shape of the histogram can be varied.
- *ShapeSyst* : this option calculates the maximum absolute relative error from the $+/-1$ sigma histograms for each bin and then applies this as a pure shape systematic without correlations across the bins. This systematic can be also applied via merging different samples, i.e ALPGEN and MADGRAPH for $t\bar{t} + V$ background.
- *histoSysOneSideSym* : it is a Gaussian systematic error of mean γ_i and width σ_i. It is treated independently in each bin, similar to the *ShapeStat* method.
- *ShapeStat*: it is a Poissonian statistical error used to treat MC statistics for the sum of all components of mean γ_i and width σ_i. Since it is a statistical uncertainty it is applied independently in each bin. In this analysis it is used for all MC samples.

  The last two methods have an important property : the systematic is intrinsically un-constrained and cannot be profiled.

B.2 Combination and improvement of searches with two same - sign and three leptons

Signal regions re-optimization

In Tables B.1 - B.4 the best signal region trials together with some well motivated trials are shown for different SS / 3 leptons combinations.

Signal lepton optimization

To further reduce the fake lepton background, new lepton isolations are also reconsidered. Lepton calorimeter and track isolations (*cone*20 and *cone*30) are scanned separately and combined using [0.04, 0.06, 0.08, ..., 0.20] cut values interval. To quantify the improvement, the isolation efficiency obtained for each set of cuts is compared to the performance obtained with the previous signal lepton definition (Section 6.3.1). For this study the SM background and the signal response is based on MC. For the detector background a combination of data driven methods and MC is used due to the time consuming procedure of the data driven estimation. The impact of the isolation cuts is tested on several SUSY models applying similar procedure as for the signal regions optimization. The most favorable scenarios are selected (Table B.5). For illustration in Figure B.1 the results obtained for $\tilde{g} \rightarrow qq\tilde{\chi}_1^{\pm} \rightarrow qq\tilde{l}\nu, \tilde{\nu}l$ simplified model are shown. The significance using previous isolation cuts for electrons and muons is compared to the selected ones. A higher significance is found for 4, 5, 7 and 9 indexes, and the improvement is up to 10% in the phase space of interest. The conclusions of this study are shown in the main body of this thesis.

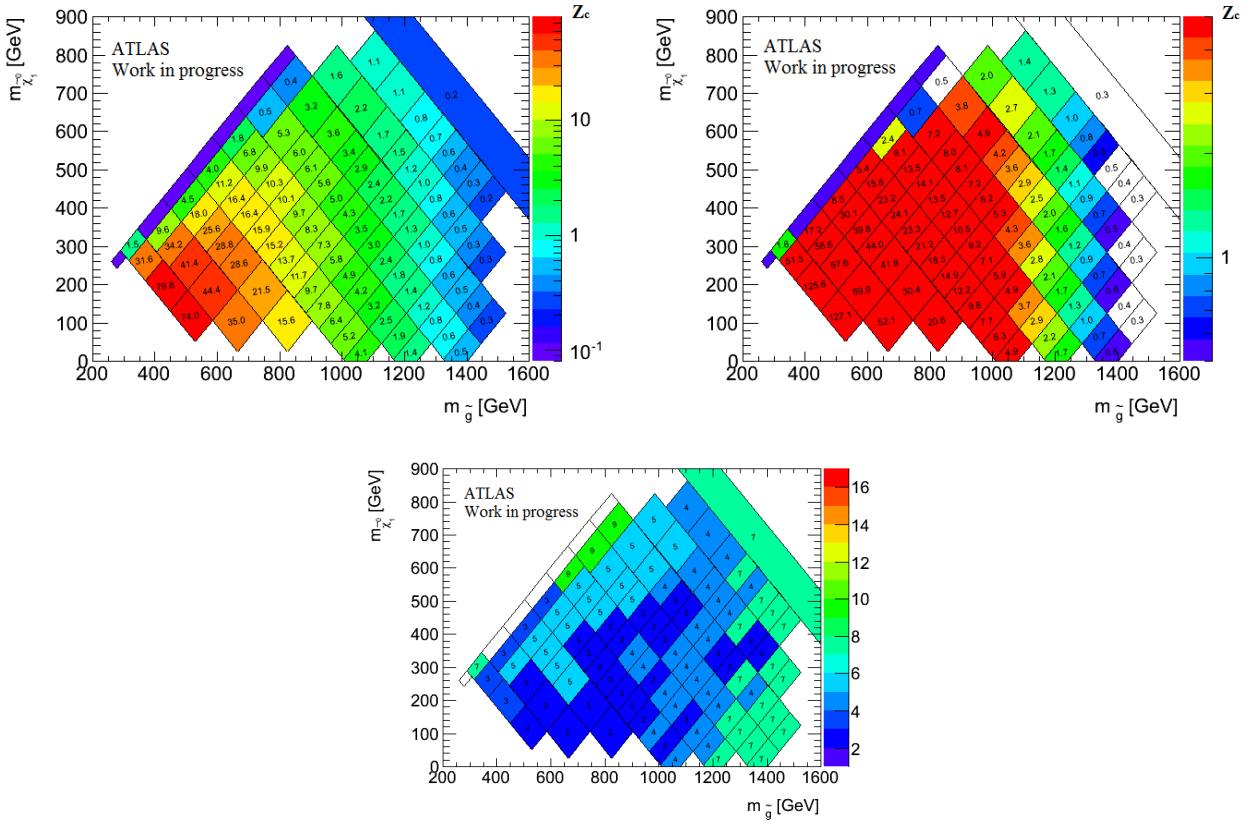


Figure B.1: Example of lepton isolation optimization using $\tilde{g} \rightarrow qq\tilde{\chi}_1^{\pm} \rightarrow qq\tilde{l}\nu, \tilde{\nu}l$ simplified model ($\sqrt{s}$ = 8 TeV, L = 20.3 fb^{-1}). Significance using Chapter 6 isolation (top - left) and best trials (top - right). The index of best trial is shown in the lower plot.

Background validation

In Figure B.2 the m_T distributions are shown in regions with and without b jets. A very good agreement can be observed for the regions with a b jet veto. The discrepancies at low m_T are due to the used ZZ MC samples (when the generator is changed, SHERPA instead of POWHEG, the discrepancy vanishes[1]). On

[1]The exiting SHERPA samples are not used due to the incorrect modeling of hard photon radiations

the distributions with at least one b jet requirement the agreement is not very good, especially in the $\mu\mu$ channel. This might be explained by an underestimation of the fake lepton background which dominated in this region of the phase space.

To finish, the last set of distributions is shown in Figure B.3 for m_{eff} variable. At least three jets are requested in the event. A fair agreement, sometimes limited by the data statistic is obtained.

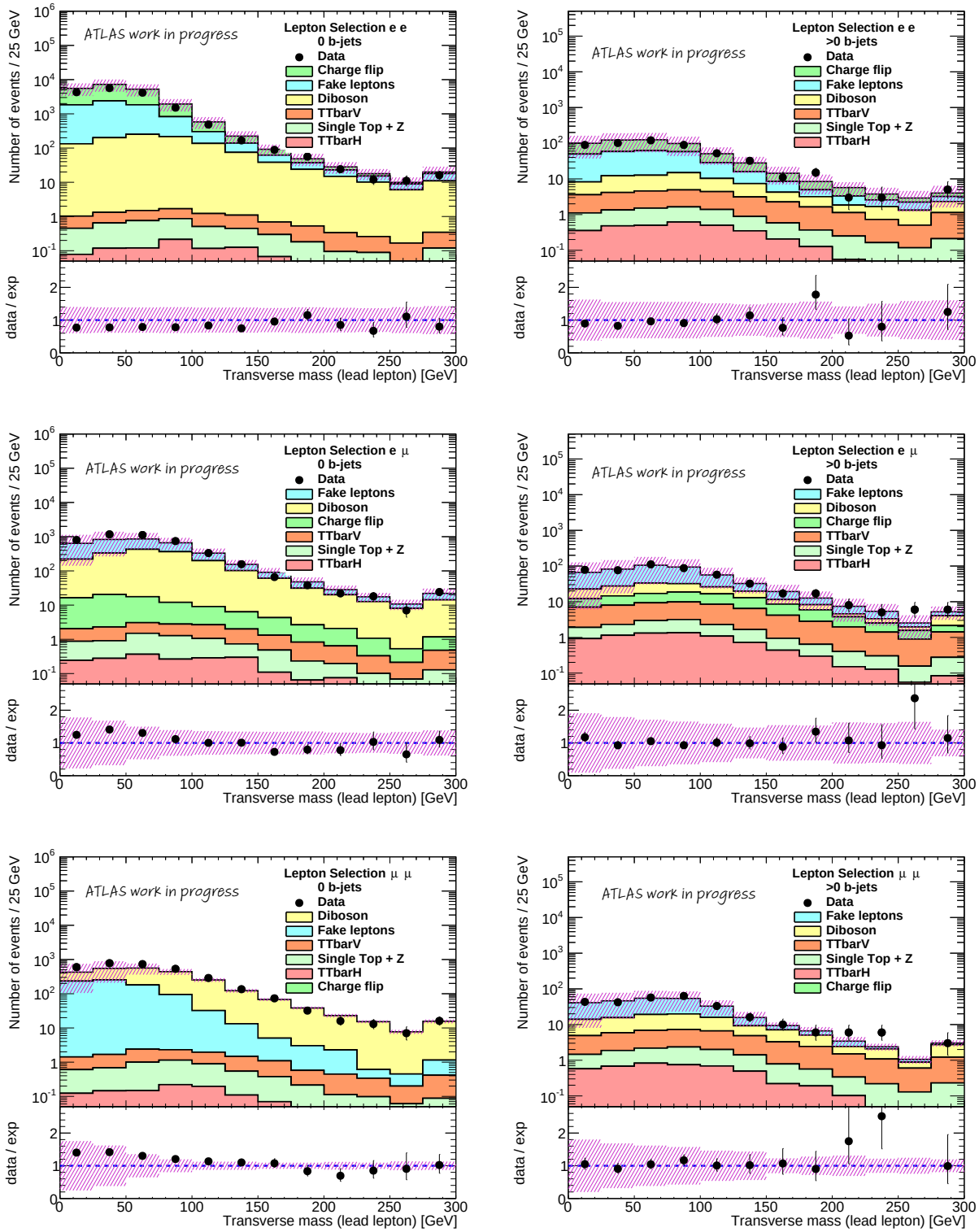


Figure B.2: Background validation : transverse mass distribution in regions with a veto on number of b jets (left) and in regions with at least 1 b (right) in ee (top), $e\mu$ (middle) and $\mu\mu$ channels. The hashed magenta area represents all the associated uncertainties on the background estimates. The black lines on the data points represent the associated statistical uncertainty. Data corresponding to $\text{L} = 20.3$ fb^{-1} and $\sqrt{s} = 8$ TeV.

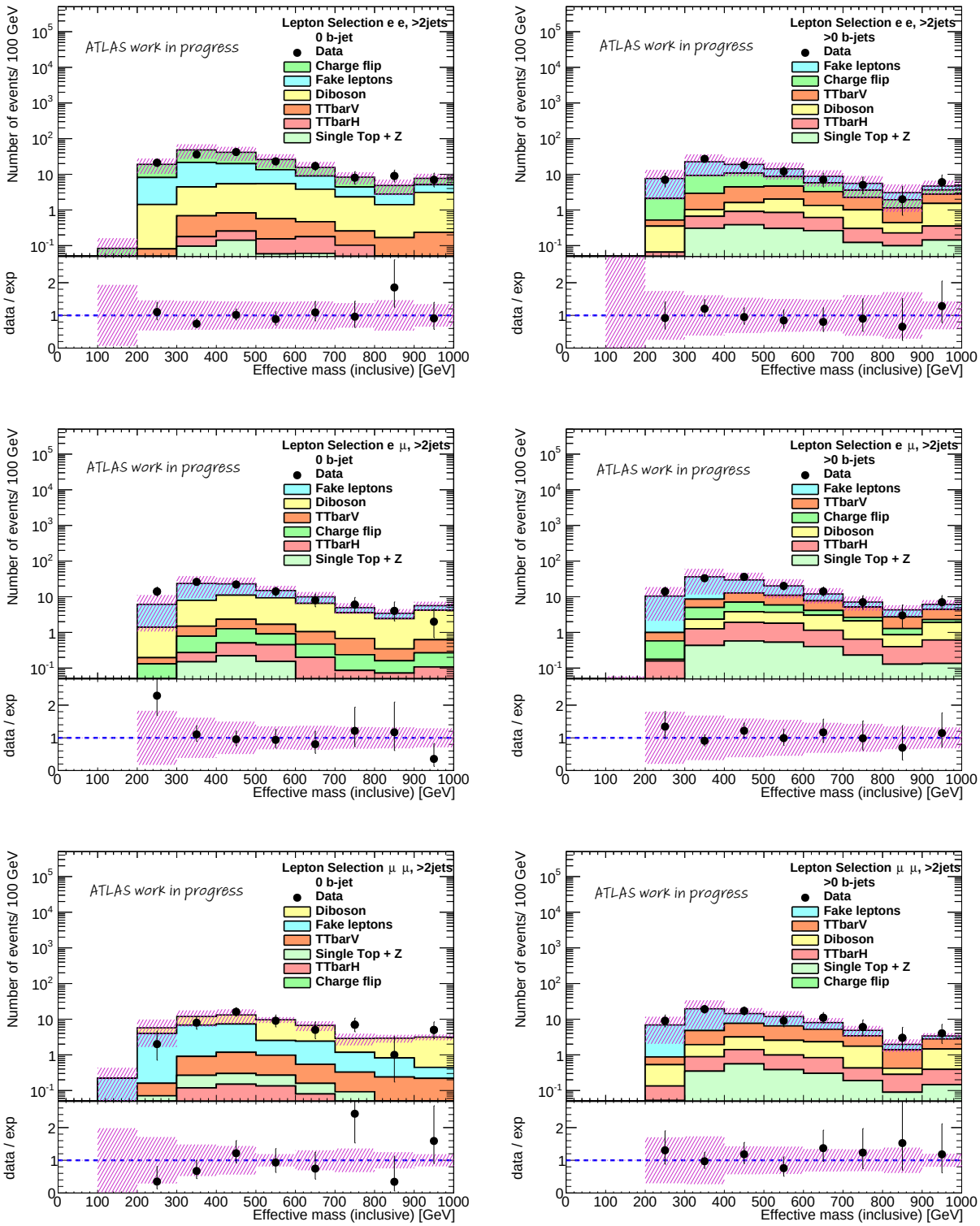


Figure B.3: Background validation : effective mass distribution in regions with a veto on number of b jets (left) and in regions with at least 1 b (right) in ee (top), $e\mu$ (middle) and $\mu\mu$ channels. The hashed magenta area represents all the associated uncertainties on the background estimates. The black lines on the data points represent the associated statistical uncertainty. Data corresponding to Ł = 20.3 fb^{-1} and $\sqrt{s} = 8$ TeV.

Index	SR	N_{b-jets}	N^{40}_{jets}	$\not{E}_T$	m_T	m_{eff}	Other
1, SS inclusive	SR0b	$== 0$	≥ 3	> 150 GeV	> 100 GeV	> 400 GeV	Zveto_3Lep, p_T Lep [20,15,15] GeV
	SR1b	≥ 1	≥ 3	> 150 GeV	> 100 GeV	> 700 GeV	$-\|\|-$
	SR3b	≥ 3	≥ 4	-	-	-	!SR1b, p_T Lep [20,15,15] GeV
2, SS exclusive	SR0b	$== 0$	≥ 3	> 150 GeV	> 100 GeV	> 400 GeV	!has3rdLep15, p_T Lep [20,15] GeV
	SR1b	≥ 1	≥ 3	> 150 GeV	> 100 GeV	> 700 GeV	$-\|\|-$
	SR3b	≥ 3	≥ 4	-	-	-	!has3rdLep15, !SR1b, p_T Lep [20,15] GeV
	SR3Lep_Low$\not{E}_T$	-	≥ 3	> 50 GeV, < 150 GeV	-	-	Zveto_3Lep, has3rdLep15, p_T Lep [20,15,15] GeV
	SR3Lep_High$\not{E}_T$	-	≥ 3	> 150 GeV	-	-	has3rdLep15, p_T Lep [20,15,15] GeV
3, SS inclusive	same as trial index 2, but $N^{40}_{jets} \geq 3$ becomes $N^{40}_{jets} \geq 4$ for SR3Lep						
4, Moriond	SR0b	$== 0$	≥ 3	> 150 GeV	> 100 GeV	> 400 GeV	p_T Lep [20,20] GeV
	SR1b	≥ 1	≥ 3	> 150 GeV	> 100 GeV	> 700 GeV	$-\|\|-$
	SR3b	≥ 3	≥ 4	-	-	-	!SR1b, p_T Lep [20,20] GeV

Table B.1: Definition of the best combined signal regions trails and their index numbers. All signal regions considered for one combination are orthogonal.

Index	SR	N_{b-jets}	N^{40}_{jets}	$\not{E}_T$	m_T	m_{eff}	Other
5, SS inclusive	SR3b	≥ 3	≥ 4	-	-	-	p_T Lep [20,15,15] GeV
	SR1b	≥ 1	≥ 3	> 150 GeV	> 100 GeV	> 700 GeV	!SR3b, p_T Lep [20,15,15] GeV
	SR0b	$== 0$	≥ 3	> 150 GeV	> 100 GeV	> 400 GeV	$-\|\|-$
	SR3Lep_Low$\not{E}_T$	-	≥ 3	> 50 GeV, < 150 GeV	-	-	Zveto_3Lep, has3rdLep15, !SR3b, p_T Lep [20,15,15] GeV
	SR2b	≥ 2	-	> 50 GeV, < 150 GeV	-	-	!has3rdLep15, !SR3b, p_T Lep [20,15] GeV
6, SS inclusive	same as trial index 5, but $N^{40}_{jets} \geq 3$ becomes $N^{40}_{jets} \geq 4$ for SR3Lep						
7, SS inclusive	SR3b	≥ 3	≥ 4	-	-	-	p_T Lep [20,15,15] GeV
	SR1b	≥ 1	≥ 3	> 150 GeV	> 100 GeV	> 700 GeV	!SR3b, Zveto_3Lep, p_T Lep [20,15,15] GeV
	SR0b	$== 0$	≥ 3	> 150 GeV	> 100 GeV	> 400 GeV	$-\|\|-$
	SR3Lep_Low$\not{E}_T$	-	≥ 3	> 50 GeV, < 150 GeV	-	-	Zveto_3Lep, has3rdLep15, !SR3b, p_T Lep [20,15,15] GeV
	SR2b	≥ 2	-	> 50 GeV, < 150 GeV	-	-	!has3rdLep15, !SR3b, p_T Lep [20,15] GeV
8, SS inclusive	same as trial index 7, but $N^{40}_{jets} \geq 3$ becomes $N^{40}_{jets} \geq 4$ for SR3Lep						
9, SS exclusive	SR3b	≥ 3	≥ 4	-	-	-	!has3rdLep15, p_T Lep [20,15] GeV
	SR1b	≥ 1	≥ 3	> 150 GeV	> 100 GeV	> 700 GeV	! has3rdLep15, !SR3b, p_T Lep [20,15] GeV
	SR1b	≥ 1	-	> 50 GeV, < 150 GeV	-	-	$-\|\|-$
	SR0b	$== 0$	≥ 3	> 150 GeV	> 100 GeV	> 400 GeV	$-\|\|-$
	SR3Lep_Low$\not{E}_T$	-	≥ 3	> 50 GeV, < 150 GeV	-	-	Zveto_3Lep, has3rdLep15, !SR3b, p_T Lep [20,15,15] GeV
	SR2b	≥ 2	-	> 50 GeV, < 150 GeV	-	-	!has3rdLep15, !SR3b, p_T Lep [20,15] GeV
10, SS exclusive	SR3b	≥ 3	≥ 4	-	-	-	!has3rdLep15, p_T Lep [20,15] GeV
	SR1b	≥ 1	≥ 3	> 150 GeV	> 100 GeV	> 700 GeV	!SR3b,!has3rdLep15, p_T Lep [20,15] GeV
	SR0b	$== 0$	≥ 3	> 150 GeV	> 100 GeV	> 400 GeV	$-\|\|-$
	SR0b	> 0	-	> 50 GeV, < 150 GeV	-	-	$-\|\|-$
	SR3Lep_Low$\not{E}_T$	-	≥ 3	> 50 GeV, < 150 GeV	-	-	Zveto_3Lep, has3rdLep15, !SR3b, p_T Lep [20,15,15] GeV
	SR2b	≥ 2	-	> 50 GeV, < 150 GeV	-	-	!has3rdLep15, !SR3b, p_T Lep [20,15] GeV
11, SS exclusive	SR3b	≥ 3	≥ 4	-	-	-	!has3rdLep15, p_T Lep [20,15] GeV
	SR1b	≥ 1	≥ 3	> 150 GeV	> 100 GeV	> 700 GeV	!SR3b, !has3rdLep15, p_T Lep [20,15] GeV
	SR1b	< 1	-	> 50 GeV, < 150 GeV	-	-	$-\|\|-$
	SR0b	$== 0$	≥ 3	> 150 GeV	> 100 GeV	> 400 GeV	$-\|\|-$
	SR3Lep_Low$\not{E}_T$	-	≥ 3	> 50 GeV, < 150 GeV	-	-	Zveto_3Lep, has3rdLep15, !SR3b, p_T Lep [20,15,15] GeV
	SR2b	≥ 2	-	> 50 GeV, < 150 GeV	-	-	!has3rdLep15, !SR3b, p_T Lep [20,15] GeV
12, SS exclusive	SR3b	≥ 3	≥ 4	-	-	-	p_T Lep [20,15] GeV
	SR1b	≥ 1	≥ 3	> 150 GeV	> 100 GeV	> 700 GeV	!SR3b, p_T Lep [20,15] GeV
	SR0b	$== 0$	≥ 3	> 150 GeV	> 100 GeV	> 400 GeV	$-\|\|-$
	SR0b	$== 0$	-	$> >50$ GeV, < 150 GeV	-	-	$-\|\|-$
	SR3Lep_Low$\not{E}_T$	-	≥ 3	> 50 GeV, < 150 GeV	-	-	Zveto_3Lep, has3rdLep15, !SR3b, p_T Lep [20,15,15] GeV
	SR2b	≥ 2	-	> 50 GeV, < 150 GeV	-	-	!has3rdLep15, !SR3b, p_T Lep [20,15] GeV
13, SS inclusive	SR3b	≥ 3	≥ 4	-	-	-	p_T Lep [20,15,15] GeV
	SR1b	≥ 1	≥ 3	> 150 GeV	> 100 GeV	> 700 GeV	!SR3b, p_T Lep [20,15,15] GeV
	SR0b	$== 0$	≥ 3	> 150 GeV	> 100 GeV	> 400 GeV	$-\|\|-$
	SR3Lep_Low$\not{E}_T$	-	≥ 3	> 75 GeV, < 150 GeV	-	-	Zveto_3Lep, has3rdLep15, !SR3b, p_T Lep [20,15,15] GeV
	SR2b	≥ 2	-	> 75 GeV, < 150 GeV	-	-	!has3rdLep15, !SR3b, p_T Lep [20,15] GeV
14, SS inclusive	same as trial index 13, but $N^{40}_{jets} \geq 3$ becomes $N^{40}_{jets} \geq 4$ for SR3Lep						
15, SS inclusive	SR3b	≥ 3	≥ 4	-	-	-	p_T Lep [20,15,15] GeV
	SR1b	≥ 1	≥ 3	> 150 GeV	> 100 GeV	> 700 GeV	!SR3b, Zveto_3Lep, p_T Lep [20,15,15] GeV
	SR0b	$== 0$	≥ 3	> 150 GeV	> 100 GeV	> 400 GeV	$-\|\|-$
	SR3Lep_Low$\not{E}_T$	-	≥ 3	> 75 GeV, < 150 GeV	-	-	Zveto_3Lep, has3rdLep15, !SR3b, p_T Lep [20,15,15] GeV
	SR2b	≥ 2	-	> 75 GeV, < 150 GeV	-	-	!has3rdLep15, !SR3b, p_T Lep [20,15] GeV
16, SS inclusive	same as trial index 15, but $N^{40}_{jets} \geq 3$ becomes $N^{40}_{jets} \geq 4$ for SR3Lep						

Table B.2: Definition of the best combined signal regions trails and their index numbers.

Index	SR	N_{b-jets}	N^{40}_{jets}	$\not{E}_T$	m_T	m_{eff}	Other
17, SS inclusive	SR3b	≥ 3	≥ 4	-	-	-	p_T Lep [20,15,15] GeV
	SR1b	≥ 1	≥ 3	>150 GeV	>100 GeV	>700 GeV	!SR3b, p_T Lep [20,15,15] GeV
	SR0b	$== 0$	≥ 3	>150 GeV	>100 GeV	>400 GeV	$-\|\|-$
	SR3Lep_Low$\not{E}_T$	-	≥ 3	>100 GeV, <150 GeV	-	-	Zveto_3Lep, has3rdLep15, !SR3b, p_T Lep [20,15,15] GeV
	SR2b	≥ 2	-	>100 GeV, <150 GeV	-	-	!has3rdLep15, !SR3b, p_T Lep [20,15] GeV
18, SS inclusive	same as trial index 17, but $N^{40}_{jets} \geq 3$ becomes $N^{40}_{jets} \geq 4$ for SR3Lep						
19, SS inclusive	SR3b	≥ 3	≥ 4	-	-	-	p_T Lep [20,15,15] GeV
	SR1b	≥ 1	≥ 3	>150 GeV	>100 GeV	>700 GeV	!SR3b, Zveto_3Lep, p_T Lep [20,15,15] GeV
	SR0b	$== 0$	≥ 3	>150 GeV	>100 GeV	>400 GeV	$-\|\|-$
	SR3Lep_Low$\not{E}_T$	-	≥ 3	>100 GeV, <150 GeV	-	-	Zveto_3Lep, has3rdLep15, !SR3b, p_T Lep [20,15,15] GeV
	SR2b	≥ 2	-	>100 GeV, <150 GeV	-	-	!has3rdLep15, !SR3b, p_T Lep [20,15] GeV
20, SS inclusive	same as trial index 19, but $N^{40}_{jets} \geq 3$ becomes $N^{40}_{jets} \geq 4$ for SR3Lep						
21, SS inclusive	SR3b	≥ 3	≥ 4	-	-	-	p_T Lep [20,20,15] GeV
	SR1b	≥ 1	≥ 3	>150 GeV	>100 GeV	>700 GeV	!SR3b,p_T Lep [20,20,15] GeV
	SR0b	$== 0$	≥ 3	>150 GeV	>100 GeV	>400 GeV	$-\|\|-$
	SR3Lep_Low$\not{E}_T$	-	≥ 3	>50 GeV, <150 GeV	-	-	Zveto_3Lep, has3rdLep15, !SR3b, p_T Lep [20,20,15] GeV
	SR2b	≥ 2	-	>50 GeV, <150 GeV	-	-	!has3rdLep15, !SR3b, p_T Lep [20,20] GeV
22, SS inclusive	same as trial index 21, but $N^{40}_{jets} \geq 3$ becomes $N^{40}_{jets} \geq 4$ for SR3Lep						
23, SS inclusive	SR3b	≥ 3	≥ 4	-	-	-	p_T Lep [20,20,15] GeV
	SR1b	≥ 1	≥ 3	>150 GeV	>100 GeV	>700 GeV	!SR3b, Zveto_3Lep, p_T Lep [20,20,15] GeV
	SR0b	$== 0$	≥ 3	>150 GeV	>100 GeV	>400 GeV	$-\|\|-$
	SR3Lep_Low$\not{E}_T$	-	≥ 3	>50 GeV, <150 GeV	-	-	Zveto_3Lep, has3rdLep15, !SR3b, p_T Lep [20,20,15] GeV
	SR2b	≥ 2	-	>50 GeV, <150 GeV	-	-	!has3rdLep15, !SR3b, p_T Lep [20,20] GeV
24, SS inclusive	same as trial index 23, but $N^{40}_{jets} \geq 3$ becomes $N^{40}_{jets} \geq 4$ for SR3Lep						
25, SS inclusive	SR3b	≥ 3	≥ 4	-	-	-	p_T Lep [20,20,15] GeV
	SR1b	≥ 1	≥ 3	>150 GeV	>100 GeV	>700 GeV	!SR3b,p_T Lep [20,20,15] GeV
	SR0b	$== 0$	≥ 3	>150 GeV	>100 GeV	>400 GeV	$-\|\|-$
	SR3Lep_Low$\not{E}_T$	-	≥ 3	>75 GeV, <150 GeV	-	-	Zveto_3Lep, has3rdLep15, !SR3b, p_T Lep [20,20,15] GeV
	SR2b	≥ 2	-	>75 GeV, <150 GeV	-	-	!has3rdLep15, !SR3b, p_T Lep [20,20] GeV
26, SS inclusive	same as trial index 25, but $N^{40}_{jets} \geq 3$ becomes $N^{40}_{jets} \geq 4$ for SR3Lep						
27, SS inclusive	SR3b	≥ 3	≥ 4	-	-	-	p_T Lep [20,20,15] GeV
	SR1b	≥ 1	≥ 3	>150 GeV	>100 GeV	>700 GeV	!SR3b, Zveto_3Lep, p_T Lep [20,20,15] GeV
	SR0b	$== 0$	≥ 3	>150 GeV	>100 GeV	>400 GeV	$-\|\|-$
	SR3Lep_Low$\not{E}_T$	-	≥ 3	>75 GeV, <150 GeV	-	-	Zveto_3Lep, has3rdLep15, !SR3b, p_T Lep [20,20,15] GeV
	SR2b	≥ 2	-	>75 GeV, <150 GeV	-	-	!has3rdLep15, !SR3b, p_T Lep [20,20] GeV
28, SS inclusive	same as trial index 27, but $N^{40}_{jets} \geq 3$ becomes $N^{40}_{jets} \geq 4$ for SR3Lep						
29, SS inclusive	SR3b	≥ 3	≥ 4	-	-	-	p_T Lep [20,20,15] GeV
	SR1b	≥ 1	≥ 3	>150 GeV	>100 GeV	>700 GeV	!SR3b,p_T Lep [20,20,15] GeV
	SR0b	$== 0$	≥ 3	>150 GeV	>100 GeV	>400 GeV	$-\|\|-$
	SR3Lep_Low$\not{E}_T$	-	≥ 3	>100 GeV, <150 GeV	-	-	Zveto_3Lep, has3rdLep15, !SR3b, p_T Lep [20,20,15] GeV
	SR2b	≥ 2	-	>100 GeV, <150 GeV	-	-	!has3rdLep15, !SR3b, p_T Lep [20,20] GeV
30, SS inclusive	same as trial index 29, but $N^{40}_{jets} \geq 3$ becomes $N^{40}_{jets} \geq 4$ for SR3Lep						
31, SS inclusive	SR3b	≥ 3	≥ 4	-	-	-	p_T Lep [20,20,15] GeV
	SR1b	≥ 1	≥ 3	>150 GeV	>100 GeV	>700 GeV	!SR3b, Zveto_3Lep, p_T Lep [20,20,15] GeV
	SR0b	$== 0$	≥ 3	>150 GeV	>100 GeV	>400 GeV	$-\|\|-$
	SR3Lep_Low$\not{E}_T$	-	≥ 3	>100 GeV, <150 GeV	-	-	Zveto_3Lep, has3rdLep15, !SR3b, p_T Lep [20,20,15] GeV
	SR2b	≥ 2	-	>100 GeV, <150 GeV	-	-	!has3rdLep15, !SR3b, p_T Lep [20,20] GeV
32, SS inclusive	same as trial index 31, but $N^{40}_{jets} \geq 3$ becomes $N^{40}_{jets} \geq 4$ for SR3Lep						

Table B.3: Definition of the best combined signal regions trails and their index numbers. All signal regions considered for one combination are orthogonal.

Index	SR	N_{b-jets}	N^{40}_{jets}	$\not{E}_T$	m_T	m_{eff}	Other
33, SS inclusive	SR3b	≥ 3	≥ 4	-	-	-	p_T Lep [20,20,20] GeV
	SR1b	≥ 1	≥ 3	$>$ 150 GeV	$>$ 100 GeV	$>$ 700 GeV	!SR3b,p_T Lep [20,20,20] GeV
	SR0b	$== 0$	≥ 3	$>$ 150 GeV	$>$ 100 GeV	$>$ 400 GeV	$-$||$-$
	SR3Lep_Low$\not{E}_T$	-	≥ 3	$>$ 50 GeV, $<$ 150 GeV	-	-	Zveto_3Lep, has3rdLep15, !SR3b, p_T Lep [20,20,20] GeV
	SR2b	≥ 2	-	$>$ 50 GeV, $<$ 150 GeV	-	-	!has3rdLep15, !SR3b, p_T Lep [20,20] GeV
34, SS inclusive	same as trial index 33, but $N^{40}_{jets} \geq 3$ becomes $N^{40}_{jets} \geq 4$ for SR3Lep						
35, SS inclusive	SR3b	≥ 3	≥ 4	-	-	-	p_T Lep [20,20,20] GeV
	SR1b	≥ 1	≥ 3	$>$ 150 GeV	$>$ 100 GeV	$>$ 700 GeV	!SR3b, Zveto_3Lep, p_T Lep [20,20,20] GeV
	SR0b	$== 0$	≥ 3	$>$ 150 GeV	$>$ 100 GeV	$>$ 400 GeV	$-$||$-$
	SR3Lep_Low$\not{E}_T$	-	≥ 3	$>$ 50 GeV, $<$ 150 GeV	-	-	Zveto_3Lep, has3rdLep15, !SR3b, p_T Lep [20,20,20] GeV
	SR2b	≥ 2	-	$>$ 50 GeV, $<$ 150 GeV	-	-	!has3rdLep15, !SR3b, p_T Lep [20,20] GeV
36, SS inclusive	same as trial index 35, but $N^{40}_{jets} \geq 3$ becomes $N^{40}_{jets} \geq 4$ for SR3Lep						
37, SS inclusive	SR3b	≥ 3	≥ 4	-	-	-	p_T Lep [20,20,20] GeV
	SR1b	≥ 1	≥ 3	$>$ 150 GeV	$>$ 100 GeV	$>$ 700 GeV	!SR3b,p_T Lep [20,20,20] GeV
	SR0b	$== 0$	≥ 3	$>$ 150 GeV	$>$ 100 GeV	$>$ 400 GeV	$-$||$-$
	SR3Lep_Low$\not{E}_T$	-	≥ 3	$>$ 75 GeV, $<$ 150 GeV	-	-	Zveto_3Lep, has3rdLep15, !SR3b, p_T Lep [20,20,20] GeV
	SR2b	≥ 2	-	$>$ 75 GeV, $<$ 150 GeV	-	-	!has3rdLep15, !SR3b, p_T Lep [20,20] GeV
38, SS inclusive	same as trial index 37, but $N^{40}_{jets} \geq 3$ becomes $N^{40}_{jets} \geq 4$ for SR3Lep						
39, SS inclusive	SR3b	≥ 3	≥ 4	-	-	-	p_T Lep [20,20,20] GeV
	SR1b	≥ 1	≥ 3	$>$ 150 GeV	$>$ 100 GeV	$>$ 700 GeV	!SR3b, Zveto_3Lep, p_T Lep [20,20,20] GeV
	SR0b	$== 0$	≥ 3	$>$ 150 GeV	$>$ 100 GeV	$>$ 400 GeV	$-$||$-$
	SR3Lep_Low$\not{E}_T$	-	≥ 3	$>$ 75 GeV, $<$ 150 GeV	-	-	Zveto_3Lep, has3rdLep15, !SR3b, p_T Lep [20,20,20] GeV
	SR2b	≥ 2	-	$>$ 75 GeV, $<$ 150 GeV	-	-	!has3rdLep15, !SR3b, p_T Lep [20,20] GeV
40, SS inclusive	same as trial index 39, but $N^{40}_{jets} \geq 3$ becomes $N^{40}_{jets} \geq 4$ for SR3Lep						
41, SS inclusive	SR3b	≥ 3	≥ 4	-	-	-	p_T Lep [20,20,20] GeV
	SR1b	≥ 1	≥ 3	$>$ 150 GeV	$>$ 100 GeV	$>$ 700 GeV	!SR3b,p_T Lep [20,20,20] GeV
	SR0b	$== 0$	≥ 3	$>$ 150 GeV	$>$ 100 GeV	$>$ 400 GeV	$-$||$-$
	SR3Lep_Low$\not{E}_T$	-	≥ 3	$>$ 100 GeV, $<$ 150 GeV	-	-	Zveto_3Lep, has3rdLep15, !SR3b, p_T Lep [20,20,20] GeV
	SR2b	≥ 2	-	$>$ 100 GeV, $<$ 150 GeV	-	-	!has3rdLep15, !SR3b, p_T Lep [20,20] GeV
42, SS inclusive	same as trial index 41, but $N^{40}_{jets} \geq 3$ becomes $N^{40}_{jets} \geq 4$ for SR3Lep						
43, SS inclusive	SR3b	≥ 3	≥ 4	-	-	-	p_T Lep [20,20,20] GeV
	SR1b	≥ 1	≥ 3	$>$ 150 GeV	$>$ 100 GeV	$>$ 700 GeV	!SR3b, Zveto_3Lep, p_T Lep [20,20,20] GeV
	SR0b	$== 0$	≥ 3	$>$ 150 GeV	$>$ 100 GeV	$>$ 400 GeV	$-$||$-$
	SR3Lep_Low$\not{E}_T$	-	≥ 3	$>$ 100 GeV, $<$ 150 GeV	-	-	Zveto_3Lep, has3rdLep15, !SR3b, p_T Lep [20,20,20] GeV
	SR2b	≥ 2	-	$>$ 100 GeV, $<$ 150 GeV	-	-	!has3rdLep15, !SR3b, p_T Lep [20,20] GeV
44, SS inclusive	same as trial index 43, but $N^{40}_{jets} \geq 3$ becomes $N^{40}_{jets} \geq 4$ for SR3Lep						

Table B.4: Definition of the best combined signal regions trails and their index numbers. All signal regions considered for one combination are orthogonal.

index	e, track iso (cone, p_T^{cone} / p_T)	e, calorimeter iso (cone, $topoE_T^{\text{cone}}$ / p_T)	μ, track iso (cone, p_T^{cone} / p_T)	μ, calorimeter iso (cone, $topoE_T^{\text{cone}}$ / p_T)
$\Delta R(\mu,\text{jet}) = 0.1$				
1	(30, 0.16)	(30, 0.18)	(30, 0.04)	(30, 0.06)
2	(30, 0.16)	(30, 0.18)	(20, 0.04)	(30, 0.04)
3	(30, 0.16)	(30, 0.18)	(20, 0.16)	(30, 0.16)
$\Delta R(\mu,\text{jet}) = 0.1$				
4	(20, 0.04)	(20, 0.04)	(30, 0.12)	(30, 0.12)
5	(20, 0.06)	(20, 0.06)	(30, 0.12)	(30, 0.12)
6	(20, 0.16)	(20, 0.16)	(30, 0.12)	(30, 0.12)
7	(30, 0.04)	(30, 0.04)	(30, 0.12)	(30, 0.12)
8	(30, 0.16)	(30, 0.18)	(20, 0.04)	(20, 0.04)
9	(30, 0.16)	(30, 0.18)	(20, 0.1)	(20, 0.16)
10	(30, 0.16)	(30, 0.18)	(30, 0.04)	(30, 0.04)
11	(30, 0.16)	(30, 0.18)	(30,0.04)	(30, 0.06)
Chapter 6 isolation, $\Delta R(\mu,\text{jet}) = 0.4$				
12	(30, 0.16)	(30, 0.18)	(30, 0.12)	(30, 0.12)

Table B.5: Preferred isolation cut combinations during the optimization. [1,..., 12] gives the index of the lepton isolation trials.

Results in the signal regions

Looking at Table B.6 it can be seen that a total of 5 electrons (marked with $*$) from 35 are reconstructed in the crack region, $1.37 < |\eta| < 1.52$. It translated in 2 events in SR0b (given 14 observed and 6.46±2.34 expected) and 3 events in SR1b (given 10 observed and 4.68±2.14 expected. Again in the absence of any study on the fake electron behavior and a measurement of fake electron rate in this region, the hypothesis of a contamination in fake electrons cannot be excluded. This will be pursued further for Run 2 preparation studies.

Run nr.	Event nr.	Lep1 flavor	Lep2 flavor	Lep1 p_T [GeV]	Lep2 p_T [GeV]	Lep1 η	Lep2 η	Lep1 φ	Lep2 φ	# b jets	m_{eff} [GeV]	$\not{E}_T$ [GeV]
* 211867	114475838	e	μ	217.06	22.04	1.49	-1.49	-0.82	-1.85	0	655.04	182.00
213695	137372266	e	e	183.71	15.81	0.50	0.19	0.13	-0.71	0	883.38	156.27
* 205112	28127952	e	e	251.73	157.38	1.45	1.83	1.02	-2.80	0	864.47	168.99
200967	65768254	μ	μ	117.87	19.92	-1.15	2.04	-0.33	2.55	0	570.02	201.94
204853	32974439	μ	μ	150.4	26.33	0.73	1.4	-0.99	1.16	0	610.48	190.58
212172	124625192	e	e	28.82	25.31	-1.55	-0.47	2.47	1.63	0	660.86	280.715
209736	101857477	e	e	41.45	32.61	-1.32	-0.18	2.75	-1.27	0	953.87	239.00
204769	84547980	e	μ	40.71	37.37	1.60	1.18	-0.30	-1.98	0	1219.21	156.45
212815	94433206	e	e	99.88	56.88	-0.70	0.83	1.57	-1.80	0	600.26	203.456
214758	180440317	e	e	78.45	19.75	0.15	-1.11	1.50	0.26	0	1026	156.18
205055	90779676	μ	e	93.6	21.89	-0.92	0.95	3.02	0.51	0	550.04	157.62
203335	10282394	e	μ	71.0	35.5	-1.11	0.75	-2.05	-3.03	0	593.0	176.29
204265	244454423	e	e	36.40	31.8	-1.60	-0.25	0.169	-2.38	0	493.59	158.112
204158	23402386	e	μ	160.85	66.4	0.63	-1.54	-0.77	2.61	0	1229.44	206.00
* 203027	13345240	e	e	210.00	50.45	1.49	1.70	0.38	-0.45	1	827.336	268.63
* 212034	89226677	e	e	330.81	21.84	-1.50	-1.22	-2.11	-0.72	1	889.75	161.22
214390	8419976	e	e	208.40	112.508	0.52	0.85	2.67	-0.97	2	1126.64	264.02
214390	56385621	μ	e	91.53	61.07	0.06	-0.56	1.72	2.87	2	1668.45	233.05
208982	1184747217	e	e	249.63	57.92	1.09	0.19	-1.40	2.79	1	1072	190.35
207772	3095631	e	e	159.64	131.07	-1.13	0.02	1.55	-1.30	2	879.50	194.56
*204910	96095322	e	μ	204.63	40.63	-1.45	-1.50	0.54	3.13	1	8120.85	325.40
210308	300382237	μ	e	94.42	40.90	0.53	-0.32	-1.43	-1.83	2	1000.65	153.29
207046	56765172	e	e	102.00	36.10	0.88	0.77	1.38	1.41	1	957.22	215.28
206409	69687314	e	μ	109.67	32.12	-0.33	-1.73	-0.83	-1.35	2	892.39	185.37

Table B.6: Event details in SR0b and SR1b, using discovery set of cuts. The (*) marker is explained in the text.

Systematic uncertainties

More details on the systematic uncertainties applied on MC background :

- $t\bar{t} + V$ **processes**. Theoretical and shape systematic uncertainties includes several sources and are taken from [252] for $t\bar{t} + Z$ and [253] for $t\bar{t} + W$. Their values are 21% for the renormalization scale, 8% for the PDF and 22% leading to an overall uncertainty of 22%. Generator shape uncertainty brings a very important contribution as the cross section is known at NLO (the MC generation was done with MADGRAPH at LO). To estimate the systematics, the renormalization and factorization scales, the initial and final state radiation and jet matching between MADGRAPH and PYTHIA generator parameters are varied. As these sources are uncorrelated, they are combined in quadrature. To asses the final uncertainties several selections are considered. One region is defined with looser cuts on lepton selection and at least 3 jets in the event to gain in statistic. A second region requires tighter cuts on jets multiplicity (≥5) and $\not{E}_T > 150$ GeV. It is designed to asses the uncertainties in the compressed region. For the third region, m_T is > 100 GeV, $\not{E}_T > 150$ GeV and $m_{\text{eff}} > 400$ or 700 GeV. It is considered for open spectra. An uncertainty of 12% is obtained, and represents the relative deviation from the nominal sample, for each variation, and added in quadrature. The overall total uncertainty of 25% accounts for all mentioned sources.

- **diboson processes**. The associated theoretical uncertainty on the cross section is given in the literature as already mentioned in Section 6.4. To account for the shape uncertainty, two separate procedures are performed as the MC generator is different for WZ processes, SHERPA, and ZZ

processes, POWHEG. For WZ processes, the following generator parameters are varied : the scale at which the parton shower and matrix element description are separated, and the factorization and normalization scales. The same regions as for $t\bar{t}$ + V were considered. But, as these samples are generated with two additional partons, the jet cut was relaxed by one. To account for a different number of additional jets requested in the signal regions, which is higher than 4, the impact of having 4 additional partons is studied. For this study, the number of jets requested in the mentioned regions is increased by one. For ZZ processes, the uncertainty on the shape is evaluated by changing the renormalization and factorization scales used for the generation. An additional systematic on the MC generator is obtain after comparing three MC samples generated in different conditions : MC@NLO + HERWIG++, MC@NLO + PYTHIA and POWHEG + PYTHIA. The obtained uncertainties, taken as the largest difference between any of the results

The final theoretical uncertainties are 35% for WW, WZ and VVV, 80% for ZZ, 30% for $t\bar{t}$ + V and 50% for top + Z on normalization. In addition, the symmetrized difference between MADGRAPH and ALPGEN samples is considered for $t\bar{t}$ + V to account for additional sources of shape uncertainties. The summary of the other sources of theoretical uncertainties is given in Table B.7.

Background	**Theoretical systematic uncertainty**	**Signal regions**		
		SR0b/SR1b	SR3b	SR3L low/high
$t\bar{t}+V$	Uncertainty on cross section	⟵ 22% ⟶		
	Shape uncertainty	⟵ 12% ⟶		
WZ	Uncertainty on cross section	⟵ 7% ⟶		
	Shape uncertainty	17%	29%	23%
	Finite no. of parton uncertainty	30%	163%	56%
ZZ	Uncertainty on cross section	⟵ 5% ⟶		
	Shape uncertainty	47%	23%	7%
	MC generator uncertainty	37%	78%	82%

Table B.7: $t\bar{t}$ + V and diboson : additional theoretical systematic uncertainties [224].

- **$t\bar{t}$ + H and $t\bar{t}t\bar{t}$ processes.** 100% theoretical systematic uncertainties are considered for these backgrounds. This conservative uncertainty is applied without other improvements or studies as this background is very small in this analysis.

- Uncertainties on signal MC : to assign the systematic uncertainties the CTEQ and MSTW PDF sets are varied, the renormalization and factorization scales with CTEQ PDF central value, the renormalization and factorization scales with MSTW PDF central value, vary strong couplings using CTEQ PDF. One asymmetric uncertainty is considered and vary between 0 and 50%.

Bibliography

[1] K. Olive *et al.*, "Review of Particle Physics," *Chinese Physics C*, vol. 38, p. 090001, 2014.

[2] J. Maurer, "Mesure des performances de reconstruction des électrons et recherche de Supersymétrie dans les canaux avec deux leptons de même charge dans les données du détecteur ATLAS." PhD, CERN-THESIS-2013-151 ; CPPM-T-2013-02, 2013.

[3] Y. Fukuda *et al.*, "Evidence for oscillation of atmospheric neutrinos," *Phys.Rev.Lett.*, vol. 81, p. 1562, 1998. arXiv:hep-ex/9807003.

[4] H. Politzer, "Reliable Perturbative Results for Strong Interactions?," *Phys.Rev.Lett.*, vol. 30, p. 1346, 1973.

[5] ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC," *Phys.Lett. B*, vol. 716, p. 1, 2012. arXiv:hep-ex/1207.7214.

[6] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," *Phys.Lett. B*, vol. 716, p. 30, 2012. arXiv:hep-ex/1207.7235.

[7] 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.*, vol. D90, p. 052004, 2014. arXiv:hep-ex/1406.3827.

[8] ATLAS Collaboration, "Measurement of Higgs boson production in the diphoton decay channel in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector," *Phys. Rev.*, vol. D90, p. 112015, 2014. arXiv:hep-ex/1408.7084.

[9] ATLAS Collaboration, "Measurements of the properties of the Higgs-like boson in the $WW^{(*)} \to \ell\nu\ell\nu$ decay channel with the ATLAS detector using 25 fb^{-1} of proton-proton collision data." ATLAS-CONF-2013-030, 2013.

[10] ATLAS Collaboration, "Measurements of Higgs boson production and couplings in the four-lepton channel in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector," *Phys.Rev.*, vol. D91, p. 012006, 2015. arXiv:hep-ex/1408.5191.

[11] ATLAS Collaboration, "Evidence for the spin-0 nature of the Higgs boson using ATLAS data," *Phys.Lett.*, vol. B726, p. 120, 2013. arXiv:hep-ex/1307.1432.

[12] ATLAS Collaboration, "Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC," *Phys.Lett.*, vol. B726, p. 88, 2013. arXiv:hep-ex/1307.1427.

[13] ATLAS Collaboration, "Search for $H \to \gamma\gamma$ produced in association with top quarks and constraints on the Yukawa coupling between the top quark and the Higgs boson using data taken at 7 TeV and 8 TeV with the ATLAS detector," *Phys.Lett. B*, vol. 740, p. 222, 2015. arXiv:hep-ex/1409.3122.

[14] H. Leutwyler, "On the history of strong interaction," 2012. Lectures given at the International School of SubNucl.Phys. at Eurice.

[15] W. N. Cottingham and D. A. Greenwood, *An Introduction to the Standard Model of Particle Physics*. Cambridge University Press, 2001.

[16] D. Griffiths, *Introduction to Elementary Particles*. Physics textbook, Wiley, 2008.

[17] F. Halzen and A. D. Martin, *Quarks and Leptons : An Introductory Course in Modern Particle Physics*. New York, USA: Wiley, 1984.

[18] V. Baran, "Quantum field theories." Lectures at the University of Bucharest, Oct 2010 - June 2011.

[19] C. Alexa, "Introduction to Standard Model of Elementary Particles." Lectures at the University of Bucharest, Oct 2010 - June 2011.

[20] E. Fermi, "Tentativo di una teoria dei raggi β," *Nuo. Cim.*, vol. 11, p. 1, 1934.

[21] G. Gamow and E. Teller, "Selection Rules for the β desintegration," *Phys.Rev.*, vol. 49, p. 895, 1936.

[22] T. D. Lee and C. N. Yang, "Question of Parity Conservation in Weak Interactions," *Phys.Rev.*, vol. 104, p. 254, 1956.

[23] C. S. Wu *et al.*, "Experimental Test of Parity Conservation in Beta Decay," *Phys.Rev.*, vol. 105, p. 1413, 1957.

[24] R. Marshak and E. Sudarshan, "Chirality Invariance and the Universal Fermi Interaction," *Phys.Rev.*, vol. 109, p. 1860, 1958.

[25] S. Glashow, "Partial symmetries of weak interactions," *Nucl.Phys.*, vol. 22, p. 579, 1960.

[26] A. Salam and S. Weinberg, "Broken Symmetries," *Phys.Rev.*, vol. 127, p. 965, 1962.

[27] H. Yukawa, "On the Interaction of Elementary Particles. I.," *Proceedings of the Physico-Mathematical Society of Japan*, vol. 17, p. 48, 1935.

[28] C. Lattes, H. Muirhead, G. Occhialini, and C. Powell, "Report of the discovery of the pion," *Nature*, vol. 159, p. 694, 1947.

[29] M. Gell-Mann, "Symmetries of Baryons and Mesons," *Phys.Rev.*, vol. 125, p. 1067, 1962.

[30] M. Gell-Mann, "A schematic model of baryons and mesons," *Phys.Lett.*, vol. 8, p. 214, 1964.

[31] G. Zweig, "An SU(3) model for strong interaction symmetry and its breaking II." CERN-TH-412, NP-14146, PRINT-64-170, 1964.

[32] E. Bloom *et al.*, "High-Energy Inelastic e–p Scattering at 6° and 10°," *Phys.Rev.Lett.*, vol. 23, p. 930, 1969.

[33] M. Breidenbach *et al.*, "Observed Behavior of Highly Inelastic Electron–Proton Scattering," *Phys.Rev.Lett.*, vol. 23, p. 935, 1969.

[34] J. D. Bjorken, "Asymptotic Sum Rules at Infinite Momentum," *Phys.Rev.*, vol. 179, p. 1547, 1969.

[35] S.L. Glashow and J. Iliopoulos and L. Maiani, "Weak Interactions with Lepton-Hadron Symmetry," *Phys.Rev.*, p. 1285, 1970.

[36] M. Kobayashi and T. Maskawa, "CP-Violation in the Renormalizable Theory of Weak Interaction," *Progress of Theoretical Physics*, vol. 49, p. 652, 1973.

[37] N. Bogoliubov, B. Struminsky, and A. Tavkhelidze, "On composite models in the theory of elementary particles," *JINR Preprint D*, 1968.

[38] Y. Nambu and M.-Y. Han, "Three-Triplet Model with Double SU(3) Symmetry," *Phys.Rev. B*, vol. 139, p. 1006, 1965.

[39] O. Greenberg, "Spin and Unitary-Spin Independence in a Paraquark Model of Baryons and Mesons," *Phys.Rev.Lett.*, vol. 13, p. 598, 1964.

[40] A. Bardeen, H. Fritzsch, and M. Gell-Mann, "Light cone current algebra, π^0 decay, and e^+ e^- annihilation," 1972. arXiv:hep-ph/0211388.

[41] D. Gross and F. Wilczek, "Asymptotically Free Gauge Theories. I," *Phys.Rev.*, vol. D8, p. 3633, 1973.

[42] D. Gross and F. Wilczek, "Ultraviolet Behavior of non-Abelian Gauge Theories," *Phys.Rev.Lett.*, vol. 30, p. 1343, 1973.

[43] H. Politzer, "Asymptotic Freedom: An Approach to Strong Interactions," *Phys. Reports*, vol. 14, p. 129, 1974.

[44] D. Barber *et al.*, "Discovery of Three-Jet Events and a Test of Quantum Chromodynamics at PETRA," *Phys.Rev.Lett.*, vol. 43, p. 830, 1979.

[45] B. Stella and H. Meyer, "Υ(9.46 GeV) and the gluon discovery (a critical recollection of PLUTO results)," *Eur.Phys.J.*, p. 203, 2011.

[46] W. Ochs, "The Status of Glueballs," *J. Phys. G*, vol. 40, p. 043001, 2013. arXiv:hep-ph/1301.5183.

[47] E. Noether, "Invariant Variation Problems," *Gott.Nachr.*, vol. 1918, p. 235, 1918. arXiv:physics/0503066.

[48] S. Tomonaga, "On a Relativistically Invariant Formulation of the Quantum Theory of Wave Fields," *Progress of Theoretical Physics*, vol. 1, p. 27, 1946.

[49] J. Schwinger, "On Quantum-Electrodynamics and the Magnetic Moment of the Electron," *Phys.Rev.*, vol. 73, p. 416, 1948.

[50] J. Schwinger, "Quantum Electrodynamics. I. A Covariant Formulation," *Phys.Rev.*, vol. 74, p. 1439, 1948.

[51] R. Feynman, "Space–Time Approach to Quantum Electrodynamics," *Phys.Rev.*, vol. 76, p. 769, 1949.

[52] R. Feynman, "The Theory of Positrons," *Phys.Rev.*, vol. 76, p. 749, 1949.

[53] R. Feynman, "Mathematical Formulation of the Quantum Theory of Electromagnetic Interaction," *Phys.Rev.*, vol. 80, p. 440, 1950.

[54] R. Mills and C. Yang, "Conservation of Isotopic Spin and Isotopic Gauge Invariance," *Phys.Rev.*, vol. 91, p. 191, 1954.

[55] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons," *Phys.Rev.Lett.*, vol. 13, p. 321, 1964.

[56] F. Englert and R. Brout, "Broken Symmetries and the Masses of Gauge Bosons," *Phys.Rev.Lett.*, vol. 13, p. 508, 1964.

[57] F. Englert and R. Brout, "Global Conservation Laws and Massless Particles," *Phys.Rev.Lett.*, vol. 13, p. 585, 1964.

[58] V. Martin, "Lectures on Particle Physics." http://www2.ph.ed.ac.uk/~vjm/Lectures/SHParticlePhysics2012.html, 2012.

[59] I. Vulpen and I. Angelozzi, "The Standard Model Higgs Boson." http://www.nikhef.nl/~ivov/HiggsLectureNote.pdf. Part of the Lecture Particle Physics II, 2013-2014.

[60] S. Adler, "A PCAC puzzle: $\pi^0 \to \gamma\gamma$ in the σ-model," *Il Nuovo Cimento A*, vol. 60, p. 47, 1969.

[61] S. Adler, "Axial-Vector Vertex in Spinor Electrodynamics," *Phys.Rev.*, vol. 177, p. 2426, 1969.

[62] M. Baak, M. Goebe, J. Haller, A. Hoecker, *et al.*, "The Electroweak Fit of the Standard Model after the Discovery of a New Boson at the LHC," *Eur.Phys.J.*, vol. C72, p. 2205, 2012. arXiv:hep-ph/1209.2716.

[63] P. Ade *et al.*, "Planck 2013 results. I. Overview of products and scientific results," *Astron.Astrophys.*, vol. 571, p. A1, 2014. arXiv:astro-ph/1303.5062.

[64] J. Oort, "The force exerted by the stellar system in the direction perpendicular to the galactic plane and some related problems," *Bulletin of the Astronomical Institutes of the Netherlands*, vol. 6, p. 249, 1932.

[65] F. Zwicky, "Die Rotverschiebung von extragalaktischen Nebeln," *Helvetica Physica Acta*, vol. 6, p. 110, 1933.

[66] V. C. Rubin and W. K. Ford, Jr., "Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions," *Astrophys.J*, vol. 159, p. 379, 1970.

[67] E. Aprile *et al.*, "Dark Matter Results from 225 Live Days of XENON100 Data," *Phys.Rev.Lett.*, vol. 109, p. 181301, 2012. arXiv:astro-ph.CO/1207.5988.

[68] E. Armengaud *et al.*, "Final results of the EDELWEISS-II WIMP search using a 4-kg array of cryogenic germanium detectors with interleaved electrodes," *Phys.Lett.*, vol. B702, p. 329, 2011. arXiv:astro-ph.CO/1103.4070.

[69] O. Adriani *et al.*, "Cosmic-Ray Positron Energy Spectrum Measured by PAMELA," *Phys.Rev.Lett.*, vol. 111, p. 081102, 2013. arXiv:astro-ph.HE/1308.0133.

[70] M. Ackermann *et al.*, "Fermi LAT search for dark matter in gamma-ray lines and the inclusive photon spectrum," *Phys. Rev.*, vol. D86, p. 022002, 2012. arXiv:astro-ph.HE/1205.2739.

[71] R. Abbasi *et al.*, "Limits on a muon flux from Kaluza-Klein dark matter annihilations in the Sun from the IceCube 22-string detector," *Phys.Rev.*, vol. D81, p. 057101, 2010. arXiv:astro-ph.CO/0910.4480.

[72] M. Aguilar *et al.*, "First Result from the Alpha Magnetic Spectrometer on the International Space Station: Precision Measurement of the Positron Fraction in Primary Cosmic Rays of 0.5-350 GeV," *Phys. Rev. Lett.*, vol. 110, p. 141102, 2013.

[73] P. Ade *et al.*, "Planck 2013 results. XVI. Cosmological parameters," *Astron.Astrophys.*, vol. 571, p. A16, 2014. arXiv:astro-ph/1303.5076.

[74] G. Bertone, D. Hooper, and J. Silk, "Particle dark matter: Evidence, candidates and constraints," *Physics Report*, vol. 405, p. 279, 2005. arXiv:hep-ph/0404175.

[75] E. Hubble, "A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae," *Proceedings of the National Academy of Science*, vol. 15, p. 168, 1929.

[76] S. M. Carroll, "The Cosmological constant," *Living Rev.Rel.*, vol. 4, p. 1, 2001. arXiv:astro-ph/0004075.

[77] S. Perlmutter *et al.*, "Measurements of Omega and Lambda from 42 high redshift supernovae," *Astrophys.J*, vol. 517, p. 565, 1999. arXiv:astro-ph/9812133.

[78] A. Riess *et al.*, "Observational evidence from supernovae for an accelerating universe and a cosmological constant," *Astronomical Journal*, vol. 116, p. 1009, 1998. arXiv:astro-ph/9805201.

[79] S. Weinberg, "The cosmological constant problem," *Rev. Mod. Phys.*, vol. 61, p. 1, 1989.

[80] S. M. Barr, G. Segre, and H. A. Weldon, "The Magnitude of the Cosmological Baryon Asymmetry," *Phys.Rev.*, vol. D20, p. 2494, 1979.

[81] B. Cleveland *et al.*, "Measurement of the solar electron neutrino flux with the Homestake chlorine detector," *Astrophys.J*, vol. 496, p. 505, 1998.

[82] Q. Ahmad *et al.*, "Measurement of the Rate of $\nu_e + d \rightarrow p + p + e^-$ Interactions Produced by ^8B Solar Neutrinos at the Sudbury Neutrino Observatory," *Phys.Rev.Lett.*, vol. 87, p. 071301, 2001. arXiv:nucl-ex/0106015.

[83] F. An *et al.*, "Improved Measurement of Electron Antineutrino Disappearance at Daya Bay," *Chin. Phys.*, vol. C37, p. 011001, 2013. arXiv:hep-ex/1210.6327.

[84] K. Nakamura *et al.*, "Review of Particle Physics," *J.Phys. G*, vol. 37, p. 1, 2010.

[85] R. Peccei and H. Quinn, "CP Conservation in the Presence of Pseudoparticles," *Phys.Rev.Lett.*, vol. 38, p. 1440, 1977.

[86] R. Peccei, "QCD, strong CP and axions," *J.Korean Phys.Soc.*, vol. 29, p. 199, 1996. arXiv:hep-ph/9606475.

[87] H. Miyazawa, "Baryon Number Changing Currents," *Progress of Theoretical Physics*, vol. 36, p. 1266, 1966.

[88] H. Miyazawa, "Spinor Currents and Symmetries of Baryons and Mesons," *Phys.Rev.*, vol. 170, p. 1586, 1968.

[89] P. Ramond, "Dual Theory for Free Fermions," *Phys. Rev.*, vol. 3, p. 2415, 1971.

[90] A. Neveu and J. Schwarz, "Factorizable dual model of pions," *Nucl.Phys. B*, vol. 31, p. 86, 1971.

[91] J.-L. Gervais and B. Sakita, "Field theory interpretation of supergauges in dual models," *Nucl.Phys. B*, vol. 34, p. 632, 1971.

[92] Y. Golfand and E. Likhtman, "Extension of the Algebra of Poincare Group Generators and Violation of p Invariance," *JETP Lett.*, vol. 13, p. 323, 1971.

[93] D. Volkov and V. Akulov, "Is the Neutrino a Goldstone Particle?," *Phys.Lett. B*, vol. 46, p. 109, 1973.

[94] J. Wess and B. Zumino, "A Lagrangian Model Invariant Under Supergauge Transformations," *Phys. Lett.*, vol. B49, p. 52, 1974.

[95] J. Wess and B. Zumino, "Supergauge Transformations in Four-Dimensions," *Nucl. Phys.*, vol. B70, p. 39, 1974.

[96] S. Coleman and J. Mandula, "All Possible Symmetries of the S Matrix," *Phys.Rev.*, vol. 159, p. 1251, 1967.

[97] R. Haag, M. Sohnius, and J. Lopuszanski, "All possible generators of supersymmetries of the S-matrix," *Nucl.Phys. B*, vol. 88, p. 257, 1975.

[98] P. Fayet, "Supersymmetry and Weak, Electromagnetic and Strong Interactions," *Phys.Lett. B*, vol. 64, p. 159, 1976.

[99] M. Sohnius, "Introducing Supersymmetry," *Physics Report*, vol. 128, p. 39, 1985.

[100] S. P. Martin, "A Supersymmetry primer," *Adv.Ser.Direct.High Energy Phys.*, vol. 21, pp. 1–153, 2010. arXiv:hep-ph/9709356.

[101] I. Aitchison, "Supersymmetry in particle physics," 2007. Cambridge Books Online.

[102] F. Larsan, "Introduction to supersymmetry," 2010. Lectures.

[103] M. Carena, "Lectures on Supersymmetry," 2007. ICTP Summer School on Particle Physics.

[104] I. Melzer-Pellmann and P. Pralavorio, "Lessons for SUSY from the LHC after the first run," *Eur.Phys.J.*, vol. C74, p. 2801, 2014. arXiv:hep-ex/1404.7191.

[105] C. Alexa, "Introduction to Supersymmetry," 2015. Lectures at the University of Bucharest.

[106] A. Chamseddine, R. Arnowitt, and N. Pran, "Locally Supersymmetric Grand Unification," *Phys.Rev.Lett.*, vol. 49, p. 970, 1982.

[107] R. Barbieri and G. Giudice, "Upper Bounds on Supersymmetric Particle Masses," *Nucl.Phys.*, vol. B306, p. 63, 1988.

[108] R. Arnowitt and P. Nath, "Supersymmetric mass spectrum in SU(5) supergravity grand unification," *Phys. Rev. Lett.*, vol. 69, p. 725, 1992.

[109] G. L. Kane, C. Kolda, L. Roszkowski, and J. D. Wells, "Study of constrained minimal supersymmetry," *Phys. Rev.*, vol. D49, p. 6173, 1994.

[110] M. Dine and W. Fischler, "A Phenomenological Model of Particle Physics Based on Supersymmetry," *Phys.Lett. B*, vol. 110, p. 227, 1982.

[111] J. L. Feng, "Naturalness and the Status of Supersymmetry," *Ann.Rev.Nucl.Part.Sci.*, vol. 63, p. 351, 2013. arXiv:hep-ph/1302.6587.

[112] N. Arkani-Hamed and S. Dimopoulos, "Supersymmetric unification without low energy supersymmetry and signatures for fine-tuning at the LHC," *JHEP*, vol. 0506, p. 073, 2005. arXiv:hep-th/0405159.

[113] G. Giudice and A. Romanino, "Split supersymmetry," *Nucl.Phys.*, vol. B699, p. 65, 2004. arXiv:hep-ph/0406088.

[114] J. Feng, "Constraints on Supersymmetry." http://hep.ps.uci.edu/~jlf/research/presentations/presentations.html, 2013.

[115] S. P. Martin and P. Ramond, "Sparticle Spectrum Constraints," *Phys.Rev.*, vol. D48, p. 5365, 1993. arXiv:hep-ph/9306314.

[116] C. Berger, J. Gainer, J. Hewett, and T. Rizzo, "Supersymmetry Without Prejudice," *JHEP*, vol. 0902, p. 023, 2009. arXiv:hep-ph/0812.0980.

[117] W. Beenakker, R. Hopker, and M. Spira, "PROSPINO: A Program for the production of supersymmetric particles in next-to-leading order QCD." arXiv:hep-ph/9611232, 1996.

[118] ATLAS Collaboration, "ATLAS Susy Searches." https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/SUSY/ATLAS_SUSY_Summary/ATLAS_SUSY_Summary.png, 2015.

[119] ATLAS Collaboration, "Search for top squark pair production in final states with one isolated lepton, jets, and missing transverse momentum in $\sqrt{s}$= 8 TeV pp collisions with the ATLAS detector," *JHEP*, vol. 1411, p. 118, 2014. arXiv:hep-ex/1407.0583.

[120] ATLAS Collaboration, "ATLAS Limits on stop mass." https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/SUSY/ATLAS_SUSY_Stop_tLSP/ATLAS_SUSY_Stop_tLSP.png, 2015.

[121] ATLAS Collaboration, "ATLAS results from the electroweak production of charginos and neutralinos searches." https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/SUSY/ATLAS_SUSY_EWSummary/ATLAS_SUSY_EWSummary.pdf, 2015.

[122] A. Djouadi, "The Anatomy of electro-weak symmetry breaking. II. The Higgs bosons in the minimal supersymmetric model," *Physics Report*, vol. 459, p. 1, 2008. arXiv:hep-ph/0503173.

[123] ATLAS Collaboration, "Search for the neutral Higgs bosons of the Minimal Supersymmetric Standard Model in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector," *JHEP*, vol. 1302, p. 095, 2013. arXiv:hep-ex/1211.6956.

[124] ATLAS Collaboration, "Search for charged Higgs bosons decaying via $H^+ \to \tau\nu$ in top quark pair events using pp collision data at $\sqrt{s} = 7$ TeV with the ATLAS detector," *JHEP*, vol. 1206, p. 039, 2012. arXiv:hep-ex/1204.2760.

[125] ATLAS Collaboration, "Search for charged Higgs bosons in the τ + jets final state with pp collision data recorded at $\sqrt{s} = 8$ TeV with the ATLAS experiment." ATLAS-CONF-2013-090, 2013.

[126] ATLAS Collaboration, "Search for charged Higgs bosons through the violation of lepton universality in $t\bar{t}$ events using pp collision data at $\sqrt{s} = 7$ TeV with the ATLAS experiment," *JHEP*, vol. 1303, p. 076, 2013. arXiv:hep-ex/1212.3572.

[127] ATLAS Collaboration, "Search for a light charged Higgs boson in the decay channel $H^+ \to c\bar{s}$ in $t\bar{t}$ events using pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector," *Eur.Phys.J.*, vol. C73, p. 2465, 2013. arXiv:hep-ex/1302.3694.

[128] G. Bennett *et al.*, "Final Report of the Muon E821 Anomalous Magnetic Moment Measurement at BNL," *Phys.Rev.*, vol. D73, p. 072003, 2006. arXiv:hep-ex/0602035.

[129] W. Altmannshofer, M. Carena, N. Shah, and F. Yu, "Indirect Probes of the MSSM after the Higgs Discovery," *JHEP*, vol. 1301, p. 160, 2013. arXiv:hep-ph/1211.1976.

[130] A. Arbey, M. Battaglia, F. Mahmoudi, and D. M. Santos, "Supersymmetry confronts $B_s \to \mu^+\mu^-$: Present and future status," *Phys.Rev.*, vol. D87, p. 035026, 2013. arXiv:hep-ph/1212.4887.

[131] LHCb Collaboration, "Measurement of the $B_s^0 \to \mu^+\mu^-$ branching fraction and search for $B^0 \to \mu^+\mu^-$ decays at the LHCb experiment," *Phys.Rev.Lett.*, vol. 111, p. 101805, 2013. arXiv:hep-ex/1307.5024.

[132] CMS Collaboration, "Measurement of the B(s) to mu+ mu- branching fraction and search for B0 to mu+ mu- with the CMS Experiment," *Phys.Rev.Lett.*, vol. 111, p. 101804, 2013. arXiv:hep-ex/1307.5025.

[133] G. Hinshaw *et al.*, "Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results," *Astro. J. Supp.*, vol. 208, p. 19, 2013. arXiv:astro-ph/1212.5226.

[134] J. Ellis and K. A. Olive, "Supersymmetric Dark Matter Candidates. Part of the book Particle Dark Matter: Observations, Models and Searches," 2010. arXiv:astro-ph.CO/1001.3651.

[135] P. Ade *et al.*, "Planck 2013 results. XXIII. Isotropy and statistics of the CMB," *Astron.Astrophys.*, vol. 571, p. A23, 2014. arXiv:astro-ph/1303.5083.

[136] Z. Ahmed *et al.*, "Results from a Low-Energy Analysis of the CDMS II Germanium Data," *Phys.Rev.Lett.*, vol. 106, p. 131302, 2011. arXiv:astro-ph.CO/1011.2482.

[137] J. Fan, M. Reece, and J. T. Ruderman, "A Stealth Supersymmetry Sampler," *JHEP*, vol. 1207, p. 196, 2012. arXiv:hep-ph/1201.4875.

[138] S. Chatrchyan *et al.*, "Search for supersymmetry in events with photons and low missing transverse energy in pp collisions at $\sqrt{s} = 7$ TeV," *Phys.Lett.*, vol. B719, p. 42, 2013. arXiv:hep-ex/1210.2052.

[139] R. Barbier, C. Berat, M. Besancon, M. Chemtob, A. Deandrea, *et al.*, "R-parity violating supersymmetry," *Phys.Rept.*, vol. 420, p. 1, 2005. arXiv:hep-ph/0406039.

[140] C. Csaki, Y. Grossman, and B. Heidenreich, "MFV SUSY: A Natural Theory for R-Parity Violation," *Phys.Rev.*, vol. D85, p. 095009, 2012. arXiv:hep-ph/1111.1239.

[141] G. Krnjaic and D. Stolarski, "Gauging the Way to MFV," *JHEP*, vol. 1304, p. 064, 2013. arXiv:hep-ph/1212.4860.

[142] N. Arkani-Hamed, P. Schuster, N. Toro, J. Thaler, L. T. Wang, *et al.*, "MARMOSET: The Path from LHC Data to the New Standard Model via On-Shell Effective Theories," 2007.

[143] J. Chakrabortty, P. Ghosh, and W. Rodejohann, "Lower limits on $\mu \to e\gamma$ from new measurements on U_{e3}," *Phys. Rev.*, vol. D86, p. 075020, 2012. arXiv:hep-ph/1204.1000.

[144] D. Alves *et al.*, "Simplified Models for LHC New Physics Searches," *J.Phys.*, vol. G39, p. 105005, 2012. arXiv:hep-ph/1105.2838.

[145] L. Evans and P. Bryant, "LHC Machine," *JINST*, vol. 3, p. S08001, 2008.

[146] O. Brüning, P. Collier, *et al.*, "LHC Design Report." CERN-2004-003-V-1, 2004.

[147] "CERN accelerators." http://project-beam-instr-sw.web.cern.ch.

[148] "19th september 2008 incident at the lhc." http://indico.cern.ch/event/44986/material/slides/1?contribId=12.

[149] "LHC Luminosity." https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults.

[150] "The LHC's first long run." http://cerncourier.com/cws/article/cern/54381.

[151] LHCB, "Observation of two new Ξ_b^- baryon resonances," *Phys.Rev.Lett.*, vol. 114, p. 062004, 2015. arXiv:hep-ex/1411.4849.

[152] "ATLAS." http://www.atlas.ch/.

[153] ATLAS Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider," *JINST*, vol. 3, p. S08003, 2008.

[154] ATLAS Collaboration, "Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics," 2009. arXiv:hep-ex/0901.0512.

[155] ATLAS Collaboration, "ATLAS Detector Status." http://twiki.cern.ch/twiki/bin/view/AtlasPublic/ApprovedPlotsATLASDetector, 2012.

[156] ATLAS Collaboration, "Fraction of good quality data in 2012." http://atlas.web.cern.ch/Atlas/GROUPS/DATAPREPARATION/PublicPlots/DQ/DQ-eff-table2012pp-AprilDecember2012.png, 2013.

[157] G. Barone, "ATLAS Silicon Microstrip Tracker Operation and Performance." ATL-INDET-PROC-2013-004, 2013.

[158] O. Ducu, J. Poveda, *et al.*, "Impact of Operating the TRT with Argon-Based Gas Mixtures in ATLAS Performance and Physics Analyses." ATL-INDET-INT-2013-010, 2013.

[159] O. Ducu, J. Poveda, *et al.*, "Impact of TRT transition radiation detection in the search for SUSY with two same-sign leptons or three leptons at $\sqrt{s} = 8$ TeV." ATL-COM-PHYS-2013-1550, 2014.

[160] ATLAS Collaboration, "Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung." ATLAS-CONF-2012-047, 2012.

[161] M. Aharrouche, J. Colas, L. di Ciaccio, *et al.*, "Response uniformity of the ATLAS liquid argon electromagnetic calorimeter," *Nucl.Instrum.Meth*, vol. 582, p. 429, 2007. arXiv:physics.ins-det/0709.1094.

[162] ATLAS Collaboration, "Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data," *Eur.Phys.J.*, vol. C72, p. 1, 2012. arXiv:hep-ex/1407.3935.

[163] ATLAS Collaboration, "Electron energy response stability with time in 2012 data." ATL-COM-PHYS-2012-1593, 2012.

[164] ATLAS Collaboration, "Electron energy response with respect to pileup in 2012 data." ATL-COM-PHYS-2012-1668, 2012.

[165] ATLAS Collaboration, "Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton–proton collision data," *Eur.Phys.J.*, vol. C74, 2014. arXiv:hep-ex/1407.3935.

[166] J F. Arguin and M. Fiascaris and J. Maurer and O. Ducu and others, "Search for strongly produced Supersymmetry with two same-sign leptons at $\sqrt{s} = 8$ TeV." ATL-COM-PHYS-2012-1824, 2012.

[167] ATLAS Collaboration, "Expected electron performance in the ATLAS experiment." ATL-PHYS-PUB-2011-006, 2011.

[168] ATLAS Collaboration, "Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data." ATLAS-CONF-2014-032, 2014.

[169] O. Ducu, J. Maurer, T. Serre, and all, "Supporting document on electron efficiency measurements using the 2012 LHC proton-proton collision data." ATL-COM-PHYS-2013-1295, 2014.

[170] T G Cornelissen and M Elsing and I Gavrilenko and J-F Laporte and W Liebig and M Limper and K Nikolopoulos and A Poppleton and A Salzburger, "The global χ^2 track fitter in ATLAS," *J.Phys.: Conference Series*, vol. 119, p. 032013, 2008.

[171] ATLAS Collaboration, "Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data," *Eur.Phys.J.*, vol. C74, 2014. arXiv:hep-ex/1407.5063.

[172] ATLAS Collaboration, "ATLAS Electron identification and reconstruction: discriminant variables." https://twiki.cern.ch/twiki/bin/view/AtlasProtected/EgammaAOD, 2014.

[173] ATLAS Collaboration, "Description and Performance of the Electron Likelihood Tool at ATLAS using 2012 LHC Data." ATL-COM-PHYS-2013-378, 2013.

[174] ATLAS Collaboration, "Search for Extra Dimensions in diphoton events using proton-proton collisions recorded at $\sqrt{s} = 7$ TeV with the ATLAS detector at the LHC," *New J.Phys.*, vol. 15, p. 043007, 2013. arXiv:hep-ex/1210.8389.

[175] ATLAS Collaboration, "Search for nonpointing and delayed photons in the diphoton and missing transverse momentum final state in 8 TeV pp collisions at the LHC using the ATLAS detector," *Phys.Rev.*, vol. D90, p. 112005, 2014. arXiv:hep-ex/1409.5542.

[176] ATLAS Collaboration, "Expected photon performance in the ATLAS experiment," 2011. ATL-PHYS-PUB-2011-007.

[177] ATLAS Collaboration, "Measurements of the photon identification efficiency with the ATLAS detector using 4.9 fb^{-1} of pp collision data collected in 2011." ATLAS-CONF-2012-123, 2012.

[178] R. Nicolaidou, L. Chevalier, S. Hassani, J. F. Laporte, E. L. Menedeu, and A. Ouraou, "Muon identification procedure for the ATLAS detector at the LHC using Muonboy reconstruction package and tests of its performance using cosmic rays and single beam data," *JP. Conference Series*, vol. 219, p. 032052, 2010.

[179] ATLAS Collaboration, "Muon Reconstruction Integration." ATL-COM-SOFT-2011-003., 2011.

[180] ATLAS Collaboration, "Staco muon reconstruction chain." https://twiki.cern.ch/twiki/bin/view/AtlasProtected/StacoMuonCollection, 2014.

[181] ATLAS Collaboration, "Third muon reconstruction chain." https://twiki.cern.ch/twiki/bin/view/AtlasProtected/MuonsCollection, 2014.

[182] ATLAS Collaboration, "Concepts, Design and Implementation of the ATLAS New Tracking." ATL-COM-SOFT-2007-002, 2007.

[183] ATLAS Collaboration, "Performance of the Reconstruction and Identification of Hadronic Tau Decays in ATLAS with 2011 Data." ATLAS-CONF-2012-142, 2012.

[184] ATLAS Collaboration, "Identification of Hadronic Decays of Tau Leptons in 2012 Data with the ATLAS Detector." ATLAS-CONF-2013-064, 2013.

[185] ATLAS Collaboration, "Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment in pp cllisions at $\sqrt{s}$=8 TeV," 2014. Submitted to Eur.Phys.J.C, arXiv:hep-ex/1412.7086.

[186] ATLAS Collaboration, "Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV," *Eur.Phys.J.*, vol. C73, p. 2304, 2013. arXiv:hep-ex/1112.6426.

[187] ATLAS Collaboration, "Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector," *Eur.Phys.J.*, vol. C75, p. 17, 2015. arXiv:hep-ex/1406.0076.

[188] ATLAS Collaboration, "Jet energy scale and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$TeV with ATLAS 2011 data." ATLAS-CONF-2013-004, 2013.

[189] ATLAS Collaboration, "Uncertainty on Jet Energy Scale." https://twiki.cern.ch/twiki/bin/view/AtlasPublic/JetEtmissApproved2013JESUncertainty, 2013.

[190] ATLAS Collaboration, "Jet energy resolution in proton-proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector," *Eur.Phys.J.*, vol. C73, p. 2306, 2013. arXiv:hep-ex/1210.6210.

[191] ATLAS Collaboration, "Jet Energy Resolution." https://twiki.cern.ch/twiki/bin/view/AtlasPublic/JetEtmissApproved2011JetResolution, 2011.

[192] ATLAS Collaboration, "Pile-up corrections for jets from proton-proton collisions at $\sqrt{s} = 7$ TeV in ATLAS in 2011." ATLAS-CONF-2012-064, 2012.

[193] ATLAS Collaboration, "Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector," 2015. Not public yet.

[194] ATLAS Collaboration, "Jet area based pile-up subtraction in ATLAS." https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JetAreas, 2014.

[195] ATLAS Collaboration, "Pile-up impact on Jet Energy Scale." https://twiki.cern.ch/twiki/bin/view/AtlasPublic/JetEtmissApproved2013Pileup1, 2013.

[196] ATLAS Collaboration, "Tagging and suppression of pileup jets with the ATLAS detector." ATLAS-CONF-2014-018, 2014.

[197] ATLAS Collaboration, "Selection of jets produced in proton-proton collisions with the ATLAS detector using 2011 data." ATLAS-CONF-2012-020, 2012.

[198] M. Cacciari, G. Salam, and G. Soyez, "The Anti-k(t) jet clustering algorithm," *JHEP*, vol. 0804, p. 063, 2008. arXiv:hep-ph/0802.1189.

[199] "Jetetmiss group, `jet uncertainties` package." https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/JetUncertainties.

[200] "Jetetmiss group, jet energy resolution provider, `jetresolution` package." https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/JetEnergyResolutionProvider.

[201] "Flavour tagging performance group, calibration data interface." https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/BTaggingCalibrationDataInterface.

[202] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet, a software package for jet finding in pp collisions." http://fastjet.fr/.

[203] Cojocaru *et al.*, "Hadronic calibration of the ATLAS liquid argon end-cap calorimeter in the pseudorapidity region $1.6 < |\eta| < 1.8$ in beam tests," *Nucl.Instrum.Meth*, vol. 531, p. 481, 2004. arXiv:physics.ins-det/0407009.

[204] S. D. Ellis and D. E. Soper, "Successive combination jet algorithm for hadron collisions," *Phys.Rev.*, vol. D48, p. 3160, 1993. arXiv:hep-ph/9305266.

[205] G. Piacquadio and C. Weiser, "A new inclusive secondary vertex algorithm for b-jet tagging in ATLAS," *J.Phys.: Conference Series*, vol. 119, p. 032032, 2008.

[206] ATLAS Collaboration, "Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data." ATLAS-CONF-2011-102, 2011.

[207] ATLAS Collaboration, "Calibration of b-tagging using dileptonic top pair events in a combinatorial likelihood approach with the ATLAS experiment." ATLAS-CONF-2014-004, 2014.

[208] ATLAS Collaboration, "Calibration of the performance of b-tagging for c and light-flavor jets in the 2012 ATLAS data." ATLAS-CONF-2014-046, 2014.

[209] ATLAS Collaboration, "b-tagging Benchmarks, $\sqrt{s} = 8$ TeV analyses." https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/BTaggingBenchmarks, 2014.

[210] ATLAS Collaboration, "Performance of Missing Transverse Momentum Reconstruction in ATLAS with 2011 Proton-Proton Collisions at $\sqrt{s} = 7$ TeV." ATLAS-CONF-2012-101, 2012.

[211] ATLAS Collaboration, "Performance of Missing Transverse Momentum Reconstruction in ATLAS studied in Proton-Proton Collisions recorded in 2012 at $\sqrt{s} = 8$ TeV." ATLAS-CONF-2013-082, 2013.

[212] ATLAS Collaboration, "ETmiss performance in 2012 data." https://twiki.cern.ch/twiki/bin/view/AtlasPublic/JetEtmissApproved2013EtMiss, 2013.

[213] "Jetetmiss group, multi-jet jes uncertainty provider package." https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/MissingETUtility.

[214] ATLAS Collaboration, "ATLAS combined Standard Model cross section measurements summary plots." https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/SM/index.html, 2014.

[215] ATLAS Collaboration, "Measurement of the W^+W^- production cross section in proton-proton collisions at $\sqrt{s}$=8 TeV with the ATLAS detector." ATLAS-CONF-2014-033, 2014.

[216] ATLAS Collaboration, "A Measurement of WZ Production in Proton-Proton Collisions at $\sqrt{s}$=8 TeV with the ATLAS detector." ATLAS-CONF-2013-021, 2013.

[217] ATLAS Collaboration, "Measurement of the total ZZ production cross section in proton-proton at $\sqrt{s}$=8 TeV with the ATLAS detector." ATLAS-CONF-2013-020, 2013.

[218] ATLAS Collaboration, "Measurement of the inclusive $W^{\pm}$ and Z/gamma cross sections in the electron and muon decay channels in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector," *Phys.Rev.*, vol. D85, p. 072004, 2012. arXiv:hep-ex/1109.5141.

[219] ATLAS Collaboration, "Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton-proton collision data," *Eur.Phys.J.*, vol. C74, p. 2941, 2014. arXiv:hep-ex/1404.2240.

[220] ATLAS Collaboration, "Supporting document on electron performance measurements using the 2011 LHC proton-proton collisions." ATL-COM-PHYS-2012-1023, 2012.

[221] T. Serre, "Electron identification efficiency measurement and search for SUSY in the 2 lepton channel with the ATLAS detector." PhD, CERN-THESIS-2014-207, 2014.

[222] "Measurement of the Inclusive ep Scattering Cross Section at Low Q^2 and x at HERA," *Eur.Phys.J.*, vol. C63, p. 625, 2009. arXiv:hep-ex/0904.0929.

[223] ATLAS Collaboration, "Search for strongly produced superpartners in final states with two same sign leptons with the ATLAS detector using 21 fb^{-1} of proton-proton collisions at $\sqrt{s} = 8$ TeV." ATLAS-CONF-2013-007, 2013.

[224] "Search for supersymmetry at $\sqrt{s}$=8 TeV in final states with jets and two same-sign leptons or three leptons with the ATLAS detector," *JHEP*, vol. 1406, p. 035, 2014. arXiv:hep-ex/1404.2500.

[225] ATLAS Collaboration, "Search for Gluinos in Events with Two Same-Sign Leptons, Jets, and Missing Transverse Momentum with the ATLAS Detector in pp Collisions at $\sqrt{s} = 7$ TeV," *Phys. Rev. Lett.*, vol. 108, p. 241802, 2012. arXiv:hep-ex/1203.5763.

[226] ATLAS Collaboration, "Search for Supersymmetry in final states with two same-sign leptons, jets and missing transverse momentum with the ATLAS detector in pp collisions at $\sqrt{s} = 8$ TeV." ATLAS-CONF-2012-105, 2012.

[227] J. Maurer, O. Ducu, *et al.*, "Extension of the $\tilde{g} \to t\bar{t}\tilde{\chi}_0^1$ simplified model to include decays with off-hell tops, and exclusion limits from the SS/3 leptons + jets analysis." ATL-COM-PHYS-2014-1422, 2014.

[228] ATLAS Collaboration, "The ATLAS Simulation Infrastructure," *Eur. Phys. J*, vol. C70, p. 823, 2010. arXiv:ph.ins-det/1005.4568.

[229] S. Agostinelli *et al.*, "GEANT4: A simulation toolkit," *Nucl. Instrum. Meth.*, vol. A506, p. 250, 2003.

[230] ATLAS Collaboration, "The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim." ATLAS-CONF-2013-004, 2010.

[231] J. Alwall *et al.*, "MadGraph/MadEvent v4: The New Web Generation," *JHEP*, vol. 09, p. 028, 2007. arXiv:hep-ph/0706.2334.

[232] T. Sjostrand, S. Mrenna, and P. Skands, "PYTHIA 6.4 physics and manual," *JHEP*, vol. 05, p. 026, 2006. arXiv:hep-ph/0603175.

[233] J. Pumplin *et al.*, "New generation of parton distributions with uncertainties from global QCD," *JHEP*, vol. 07, p. 012, 2002. arXiv:hep-ph/0201195.

[234] A. Martin, W. Stirling, R. Thorne, and G. Watt, "Parton distributions for the LHC," *Eur.Phys.J.*, vol. C63, p. 189, 2009. arXiv:hep-ph/0901.0002.

[235] M. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. Polosa, "ALPGEN, a generator for hard multiparton processes in hadronic collisions," *JHEP*, vol. 0307, p. 001, 2003. arXiv:hep-ph/0206293.

[236] J. Butterworth, J. R. Forshaw, and M. Seymour, "Multiparton interactions in photoproduction at HERA," *Z.Phys.*, vol. C72, p. 637, 1996. arXiv:hep-ph/9601371.

[237] T. Gleisberg *et al.*, "Event generation with SHERPA 1.1," *JHEP*, vol. 0902, p. 007, 2009. arXiv:hep-ph/0811.4622.

[238] H. La, M. Guzzi, J. Huston, Z. Li, P. Nadolsky, *et al.*, "New parton distributions for collider physics," *Phys.Rev.*, vol. D82, p. 074024, 2010. arXiv:hep-ph/1007.2241.

[239] 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*, vol. 1006, p. 043, 2010. arXiv:hep-ph/1002.2581.

[240] T. Sjostrand, S. Mrenna, and P. Z. Skands, "A Brief Introduction to PYTHIA 8.1," *Comput.Phys.Commun.*, vol. 178, p. 852, 2008. arXiv:hep-ph/0710.3820.

[241] S. Frixione, B. R. Webber, and P. Nason, "The MC and NLO 3.4 Event Generator," 2008. arXiv:hep-ph/0812.0770.

[242] M. Bahr *et al.*, "Herwig++ Physics and Manual," *Eur. Phys. J*, vol. C58, p. 639, 2008. arXiv:hep-ph/0803.0883.

[243] M. Kramer, A. Kulesza, R. V. der Leeuw, *et al.*, "Supersymmetry production cross sections in pp collisions at $\sqrt{s} = 7$ TeV," 2012. arXiv:hep-ph/1206.2892.

[244] W. Beenakker, S. Brensing, M. Kramer, A. Kulesza, E. Laenen, *et al.*, "Soft-gluon resummation for squark and gluino hadroproduction," *JHEP*, vol. 0912, p. 041, 2009. arXiv:hep-ph/0909.4418.

[245] W. Beenakker, S. Brensing, M. Kramer, A. Kulesza, E. Laenen, *et al.*, "Squark and Gluino Hadroproduction," *Int.J.Mod.Phys.*, vol. A26, p. 2637, 2011. arXiv:hep-ph/1105.1110.

[246] LHC SUSY Cross Section Working Group, "Susy cross sections." https://twiki.cern.ch/twiki/bin/view/LHCPhysics/SUSYCrossSections, 2013.

[247] T. Plehn, "Prospino." http://www.thphys.uni-heidelberg.de/~plehn/index.php?show=prospino&visible=tools, 2013.

[248] ATLAS Collaboration, "Search for gluino pair production in final states with missing transverse momentum and at least three b-jets using 12.8 fb^{-1} of pp at $\sqrt{s} = 8$ TeV with the ATLAS Detector." ATLAS-CONF-2012-145, 2012.

[249] ATLAS Collaboration, "Search for new phenomena in final states with large jet multiplicities and missing transverse momentum at $\sqrt{s}$=8 TeV proton-proton collisions using the ATLAS experiment," *JHEP*, vol. 1310, p. 130, 2013. arXiv:hep-ex/1308.1841.

[250] J. F. da Costa, J. Poveda, and T. Yamanaka, "Search for direct stop pair production in events with a Z boson, b-jets and missing transverse energy with the ATLAS detector using 20fb^{-1} of proton - proton collisions at $\sqrt{s} = 8$ TeV." ATL-PHYS-INT-2014-007, 2014.

[251] ATLAS Collaboration, "Further search for supersymmetry at $\sqrt{s}$=7 TeV in final states with jets, missing transverse momentum, and isolated leptons with the ATLAS detector," *Phys. Rev. D*, vol. 86, p. 092002, 2012. arXiv:hep-ex/1208.4688.

[252] M. Garzelli, A. Kardos, C. Papadopoulos, and Z. Trocsanyi, "t $\bar{t}$ W^{+-} and t $\bar{t}$ Z Hadroproduction at NLO accuracy in QCD with Parton Shower and Hadronization effects," *JHEP*, vol. 1211, p. 056, 2012. arXiv:hep-ph/1208.2665.

[253] J. M. Campbell and R. K. Ellis, "$t\bar{t}W^{+-}$ production and decay at NLO," *JHEP*, vol. 1207, p. 052, 2012. arXiv:hep-ph/1204.5678.

[254] CMS Collaboration, "Measurement of associated production of vector bosons and top quark-antiquark pairs at $\sqrt{(s)} = 7$ TeV," *Phys.Rev.Lett.*, vol. 110, p. 172002, 2013. arXiv:hep-ex/1303.3239.

[255] ATLAS Collaboration, "Evidence for the associated production of a vector boson (W, Z) and top quark pair in the dilepton and trilepton channels in pp collision data at $\sqrt{s}$=8 TeV collected by the ATLAS detector at the LHC." ATLAS-CONF-2014-038, 2014.

[256] ATLAS Collaboration, "Measurement of ZZ production in pp collisions at $\sqrt{s} = 7$ TeV and limits on anomalous ZZZ and $ZZ\gamma$ couplings with the ATLAS detector," *JHEP*, vol. 1303, p. 128, 2013. arXiv:hep-ex/1211.6096.

[257] ATLAS Collaboration, "Measurement of WZ production in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector," *Eur.Phys.J.*, vol. C72, p. 2173, 2012. arXiv:hep-ex/1208.1390.

[258] ATLAS Collaboration, "Measurement of W^+W^- production in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector and limits on anomalous WWZ and $WW\gamma$ couplings," *Phys.Rev.*, vol. D87, p. 112001, 2013. arXiv:hep-ex/1210.2979.

[259] ATLAS Collaboration, "Measurement of the WW cross section in pp collisions with the ATLAS detector and limits on anomalous gauge couplings," *Phys.Lett. B*, vol. 712, p. 289, 2012. arXiv:hep-ex/1203.6232.

[260] ATLAS Collaboration, "SM Higgs production cross sections at 8 TeV." https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageAt8TeV, 2013.

[261] ATLAS Collaboration, "SM Higgs Branching Ratios and Partial - Decay Widths." https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageBR2, 2013.

[262] ATLAS Collaboration, "SM Background MC samples and cross sections." https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/SusyMCSamples2, 2013.

[263] ATLAS Collaboration, "Search for anomalous production of events with same-sign dileptons and b jets in 14.3 fb of pp collisions at $\sqrt{s} = 8$ TeV with the atlas detector." ATLAS-CONF-2013-051, 2013.

[264] M. Baak, G. Besjes, D. Cote, A. Koutsman, J. Lorenz, *et al.*, "HistFitter software framework for statistical data analysis," 2014. arXiv:hep-ex/1410.1280.

[265] ATLAS Collaboration, "HistFitter," 2013. https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/SusyFitter.

[266] K. Belasco, K. Cranmer, *et al.*, "The RooStats Project," in *ACAT2010*, 2010.

[267] K. Cranmer, "HistFactory User Guide," 2013. https://twiki.cern.ch/twiki/pub/RooStats/WebHome/HistFactoryLikelihood.pdf.

[268] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics," *Eur.Phys.J.*, p. 1554, 2011. arXiv:hep-ph/1007.1727.

[269] A. Read, "Presentation of search results: The CL(s) technique," *J.Phys. G: Nuclear and Particle Physics*, vol. 28, p. 2693, 2002.

[270] "Luminosity uncertainties." https://indico.cern.ch/getFile.py/access?contribId=7&resId=1&materialId=slides&confId=192530.

[271] "Physics analysis tools, extended pile-up reweighting." https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/ExtendedPileupReweighting.

[272] O. Ducu and J. Maurer, "Search for heavy top bottom partners in same-sign lepton final states with atlas data," *Global Journal on Technology*, vol. 4, p. 298, 2013.

[273] ATLAS Collaboration, "Search for supersymmetry using events with three leptons, multiple jets, and missing transverse momentum in 13.0 fb^{-1} of pp collisions with the ATLAS detector at $\sqrt{s} = 8$ TeV." ATLAS-CONF-2012-151, 2012.

[274] G. Cowan, *Statistical Data Analysis*. Clarendon Press, 1998.

[275] ATLAS Collaboration, "Exclusion Limits for the Gtt Model." https://twiki.cern.ch/twiki/bin/view/AtlasPublic/CombinedSummaryPlots, 2013.

[276] CMS, "Search for new physics in events with same-sign dileptons and jets in pp collisions at $\sqrt{s} = 8$ TeV," *JHEP*, vol. 1401, p. 163, 2014. arXiv:hep-ph/1311.6736.

[277] O. Ducu, "Search for new physics in events with 2 same sign leptons, jets and missing transverse energy with ATLAS and CMS detectors." https://indico.in2p3.fr/event/8565/session/6/contribution/5, 2014.

[278] ATLAS Collaboration, "Evidence for Electroweak Production of $W^{\pm}W^{\pm}jj$ in pp Collisions at $\sqrt{s} =$ 8 TeV with the ATLAS Detector," *Phys. Rev. Lett.*, vol. 113, p. 141803, 2014.

[279] ATLAS Collaboration, "Search for anomalous production of prompt same-sign lepton pairs and pair-produced doubly charged Higgs bosons with $\sqrt{s} = 8$ TeV pp collisions using the ATLAS detector," 2014. arXiv:hep-ex/1412.0237.

[280] ATLAS Collaboration, "Search for microscopic black holes in a like-sign dimuon final state using large track multiplicity with the ATLAS detector," *Phys. Rev. D*, vol. 88, p. 072001, 2013. arXiv:hep-ex/1308.4075.

[281] ATLAS Collaboration, "Search for heavy long-lived multi-charged particles in pp collisions at $\sqrt{s} =$ 8 TeV using the ATLAS detector," 2015. Submitted to Eur.Phys.J.C, arXiv:hep-ex/1504.04188.

[282] O. Ducu *et al.*, "Search for strongly produced superpartners in final states with same-sign leptons or three leptons and jets: preparing for 2015 analyses." ATL-COM-PHYS-2015-329, 2015.

[283] C. Graphic Design, "LHC consolidations 2013-14." CERN-GE-1302027, 2013.

[284] "2015 commissioning outline." http://lhc-commissioning.web.cern.ch//.

[285] ATLAS Collaboration, "IBL Module Loading onto Stave and Quality Check." ATL-COM-INDET-2015-006.

[286] ATLAS Collaboration, "ATLAS Run-2 Detector Status." https://twiki.cern.ch/twiki/bin/view/AtlasPublic/ApprovedPlotsATLASDetector, 2015.

[287] "LHC Run-2 Luminosity." https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.

[288] ATLAS Collaboration, "Impact Parameter Resolution." IDTR-2015-007.

[289] ATLAS Collaboration, "Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} = 13$ TeV." ATL-PHYS-PUB-2015-015.

[290] ATLAS Collaboration, "Performance of missing transverse momentum reconstruction with the ATLAS detector in the first proton-proton collisions at $\sqrt{s} = 13$ TeV." ATL-PHYS-PUB-2015-027.

[291] ATLAS Collaboration, "Electron identification measurements in ATLAS using $\sqrt{s} = 13$ TeV data with 50 ns bunch spacing." ATL-PHYS-PUB-2015-041.

[292] ATLAS Collaboration, "MCP public plots in early Run-II: momentum calibration validation,reconstruction efficiency and isolation efficiency based on 6 pb^{-1}." ATL-COM-MUON-2015-030.

[293] O. Ducu *et al.*, "Search for strongly produced superpartners in final states with same-sign leptons or three leptons and jets in pp collisions at $\sqrt{s} = 13$ TeV (Supporting note)." ATL-COM-PHYS-2015-1150, 2015.

[294] E. Halkiadakis, G. Redlinger, and D. Shih, "Status and Implications of Beyond-the-Standard-Model Searches at the LHC," *Ann. Rev. Nucl. Part. Sci.*, vol. 64, 2014. rXiv:hep-ex/1411.1427.

[295] "Tevatron and LHC parton luminosity comparison plots." http://www.hep.ph.ic.ac.uk/~wstirlin/plots/plots.html.

[296] O. Ducu, J. Maurer, P. Pralavorio, *et al.*, "Run-2 prospects for top Yukawa coupling measurement in the ttH channel with two same-sign leptons." ATL-COM-PHYS-2015-479.

[297] E. Thompson *et al.*, "Quantum diaries, top quark." http://www.quantumdiaries.org/tag/top-quark/, 2014.

[298] ATLAS Collaboration, "Performance of boosted top quark identification in 2012 ATLAS data." ATLAS-CONF-2013-084.

[299] ATLAS Collaboration, "A search for $t\bar{t}$ resonances in the lepton plus jets final state with ATLAS using 14 fb^{-1} of pp collisions at $\sqrt{s}$=8 TeV." ATLAS-CONF-2013-052.

[300] ATLAS Collaboration, "Search for gluino-mediated stop and sbottom pair production in 3 events with b-jets and large missing transverse momentum." ATL-COM-PHYS-2015-319, 2015. ATLAS internal document.

[301] ATLAS Collaboration, "Measurements of the properties of the Higgs-like boson in the two photon decay channel with the ATLAS detector using 25 fb^{-1} of proton-proton collision data." ATLAS-CONF-2013-012, 2013.

[302] ATLAS Collaboration, "Measurements of the properties of the Higgs-like boson in the four lepton decay channel with the ATLAS detector using $25 fb^{-1}$ of proton-proton collision data." ATLAS-CONF-2013-013, 2013.

[303] CMS Collaboration, "Combination of standard model Higgs boson searches and measurements of the properties of the new boson with a mass near 125 GeV." CMS-PAS-HIG-13-005, 2013.

[304] ATLAS Collaboration, "Search for the associated production of the Higgs boson with a top quark pair in multi-lepton final states with the ATLAS detector." ATLAS-CONF-2015-006.

[305] ATLAS and CMS collaborations, "Combined Measurement of the Higgs Boson Mass in *pp* Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments," *Phys.Rev.Lett.*, vol. 114, p. 191803, 2015. arXiv:hep-ex/1503.07589.

[306] LEP Electroweak Working Group, "Precision Electroweak Measurements and Constraints on the Standard Model," 2010. arXiv:hep-ex/1012.2367.

[307] R. Barate *et al.*, "Search for the standard model Higgs boson at LEP," *Phys.Lett.*, vol. B565, p. 61, 2003. arXiv:hep-ex/0306033.

[308] T. Aaltonen *et al.*, "Combined search for the standard model Higgs boson decaying to a $b\bar{b}$ pair using the full CDF data set," *Phys.Rev.Lett.*, vol. 109, p. 111802, 2012. arXiv:hep-ex/1207.1707.

[309] ATLAS Collaboration, "Combined search for the standard model higgs boson in pp collisions at $\sqrt{s}$=7 TeV with the atlas detector," *Phys. Rev.*, vol. D86, p. 032003, 2012. arXiv:hep-ex/1207.0319.

[310] ATLAS Collaboration, "Study of the spin and parity of the Higgs boson in di-boson decays with the ATLAS detector." ATLAS-CONF-2015-008, 2015.

[311] ATLAS Collaboration, "Measurements of the Higgs boson production and decay rates and couplings using pp collision data at sqrt(s) = 7 and 8 TeV in the ATLAS experiment." ATLAS-CONF-2015-007, 2015.

[312] J. Adelman *et al.*, "Search for the Standard Model Higgs boson produced in association with top quarks in pp collisions at 8 TeV with the ATLAS detector at the LHC." ATL-COM-PHYS-2014-1471.

[313] ATLAS Collaboration, "Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in pp collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector," 2015. arXiv:hep-ex/1503.05066.

[314] ATLAS Collaboration, "Search for $H \to \gamma\gamma$ produced in association with top quarks and constraints on the Yukawa coupling between the top quark and the Higgs boson using data taken at 7 TeV and 8 TeV with the ATLAS detector," *Phys.Lett.*, vol. B740, p. 222, 2015. arXiv:hep-ex/1409.3122.

[315] A. Babar *et al.*, "Search for Standard Model Higgs boson production in association with a top quark pair in multilepton signatures with the ATLAS detector : Object and event selection." ATL-COM-PHYS-2014-053.

[316] A. Babar *et al.*, "Search for ttH in the multilepton final 1 state: backgrounds and their estimation." ATL-COM-PHYS-2014-222.

[317] A. Babar *et al.*, "Search for ttH in the multilepton final state: results." ATL-COM-PHYS-2014-1389.

[318] ATLAS Collaboration, "Measurements of the Higgs boson production and decay rates and coupling strengths using pp collision data at $\sqrt{s}$ = 7 and 8 TeV in the ATLAS experiment." ATLAS-CONF-2015-007.

[319] J. Adelman *et al.*, "Search for ttH and tH production in the diphoton channel with 20.3 fb^{-1} of data at 8 TeV and 4.5 fb^{-1} of data at 7 TeV taken with the ATLAS detector." ATL-COM-PHYS-2014-011.

[320] 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," 2015. arXiv:hep-ex/1412.8662.

[321] CMS Collaboration, "Search for a standard model Higgs boson produced in association with a top-quark pair and decaying to bottom quarks using a matrix element method," 2015. arXiv:hep-ex/1502.02485.

[322] LHC Higgs Cross Section Working Group Collaboration, S. Heinemeyer et al., "Handbook of LHC Higgs Cross Sections: 3. Higgs Properties." arXiv:1307.1347.

[323] https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageAt1314TeV.

[324] J. M. Campbell and R. K. Ellis, "$t\bar{t} + W$ production and decay at NLO," *JHEP*, vol. 1207, p. 052, 2012. arXiv:1204.5678.

[325] https://twiki.cern.ch/twiki/bin/view/LHCPhysics/TtbarNNLO.

[326] https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/PDFReweight.

[327] J. M. Campbell, R. K. Ellis, and R. H. Rontsch, "Single top production in association with a Z boson at the LHC," *Phys. Rev.*, vol. D87, p. 114006, 2013. arXiv:1302.3856.

[328] J. M. Campbell, R. K. Ellis, and C. Williams, "Vector boson pair production at the LHC," *JHEP*, vol. 1107, p. 018, 2011. arXiv:1105.0020.