National and Kapodistrian University of Athens
Physics Department
Nuclear and Particle Physics division

A search for Supersymmetry in the compressed mass spectrum in events with three soft leptons and missing transverse momentum with the CMS experiment at the LHC

Papavergou Ioanna

ioannapapaver@gmail.com

Supervisors:
Konstantinos Vellidis, Professor of Physics

October 2, 2022

Acknowledgements

I am deeply grateful to all the people that contributed to the completion of my PhD.

I would like to thank my supervisor Costas Vellidis for the opportunity he gave me to work at CERN, one of the largest and most exciting scientific environments of our times. He also gave me guidance and support and tough me how to deep dive into difficult and compelling projects and I want to thank him for that as well. Additonally, I want to thank Paris Sphicas for the opportunity he gave me to be part of the NKUA CMS team and the AMVA4NewPhysics. He was a great inspiration to me.

I am grateful to the whole CMS SUSY Soft Leptons team and specifically Cristina Botta and Marco Peruzzi for the incredible collaboration, the interesting discussions and all their help and support.

It was a great privilege to be part of the AMVA4NewPhysics ITN. It helped to me to expand my research interests, collaborate with some brilliant scientists, travel and gain valuable experience.

I would like to thank the NKUA CMS group for all the interesting discussions. Thank you George Karathanasis, Pantelis Kontaxakis and George Melachrinos for all the long meeting that always turned into something useful. Special thanks to Manos Vourliotis. Our excellent collaboration taught lot.

I am deeply grateful to Dimitris for the invaluable support, his encouragement and all our endless discussions.

Finally, I want to thank my parents and my sister that encourage me from an early age to pursue my dreams, always try hard and achieve my goals. Their limitless support throughout those years was of great importance. I would like to dedicate this thesis to my dad who showed the beauty of trying to answer nature's mysteries.

Abstract

The first part of this dissertation presents a brief overview of the Standard Model of particle physics with emphasis on its shortcomings and the motivations behind the theory of Supersymmetry. The latter is a theory that links the fermions and the bosons and provides a solution to many of the open problems of modern physics. The principles and the phenomenology of Supersymmetry are briefly described along with the overview of the exclusion limits on the supersymmetric phase space obtained from the direct searches in the collider experiments.

The LHC is expected to be able to produce supersymmetric particles through its proton proton collisions. In case of this event the CMS detector is expected to collect such events. A description of the LHC and the CMS detector are presented in the second part of the dissertation.

Not all of the collision events produced at the LHC are interesting for physics searches and therefore only a small fraction of them is saved for offline analysis. The third part of the document describes the CMS trigger that is used to collect the interesting physics events and the offline algorithms that is used for the physics object reconstruction. Additionally, this third part of the thesis presents the service work that was conducted in the context of this PhD.

The fourth part describes the search for Supersymmetry in events with soft leptons and missing transverse energy in the final state that was conducted as the main physics analysis of this thesis. These signatures are typical of supersymmetry scenarios with a small mass splitting between the lightest and the next-to-lightest supersymmetric particles, also referred to as compressed mass spectra. The search was conducted with the full Run 2 data collected by the CMS detector in proton proton collision at center of mass energy of 13 TeV. No significant deviation from the Standard Model expectation is observed. Therefore, the results are interpreted in the context of various supersymmetric models and upper limits are set on the masses of the relevant supersymmetric particles.

The final part of the dissertation highlights the basic concepts of machine learning and its application in high energy physics. A supersymmetric sensitivity study in the compressed mass spectrum that was performed using machine learning algorithms is presented in this part. The study targeted events with 3 soft leptons and missing transverse momentum in the final stated and was performed with the 2016 CMS data. The results of the machine learning analysis are compared to the respective results obtained with the "baseline" method.

Contents

Chapter 1

The Standard Model and its extension to Supersymmetry

The Standard Model (SM) of particle physics offers an incredibly accurate description of the fundamental particles and their interactions. Its theoretical development was completed over the decade of 1970 and it has many successes to its name, with a plethora of experimental validations. Its most recent triumph was the discovery of the Higgs boson in 2012, at the Large Hadron Collider (LHC), from the Compact Muon Solenoid (CMS) and the ATLAS Collaboration, where ATLAS stands for "A Toroidal LHC ApparatuS". However, there is observational evidence that cannot be accommodated within the theory of the SM and therefore, new theories Beyond the SM (BSM) have been developed. Supersymmetry (SUSY) is one of those BSM theories that aims to address questions that the SM has left unanswered.

This first chapter of the thesis presents a brief introduction to the main concepts of the SM and SUSY and it is based on Ref. [1, 2]. The structure of the chapter is the following: Sections 1.1-1.2 describe the basic concepts of the SM and some of its experimental validation and shortcomings. Section 1.3 highlights the motivation of the SUSY theory and Section 1.4 presents the derivation of the general SUSY Lagrangian. The Minimal SUSY SM and the chargino and neutralino mass spectra are discussed in Section 1.5 and 1.6 respectively. In Section 1.7, the current status of the exclusion limits on the SUSY parameter space, as derived by collider searches is presented. Section 1.8 discusses the motivation for SUSY searches in regions where the mass difference between the lightest and next to lightest SUSY particles in the decay chain is low. This scenario corresponds to the compressed mass spectrum. An overview of the SUSY searches conducted at CMS in the compressed mass spectra is given in 1.9. Finally, Section 1.10 discusses the status of indirect SUSY searches in precision measurement experiments.

1.1 Standard Model particles and interactions

The central feature of the SM is that it provides a unified theory for matter and force carrier particles. It is comprised of fundamental, spin-1/2, matter particles called fermions and fundamental, spin-1, force-carrying particles called gauge bosons.

Fermions are divided into quarks (q) or leptons based on their properties and their interactions with bosons. Leptons can be electrically charged (ℓ) or neutral (ν) while all quarks carry electric charge. There are 3 generations of fermions, each consisting of one charged and one neutral lepton and two quarks. The electron, electron neutrino, up and down quark are collectively called the first generation. The second generation consists of the muon, muon neutrino, charm and strange quarks and the third generation consists of the tau, tau neutrino, top and bottom quarks. The quarks are divided into up-type and down-type indicating that the quarks of the second and the third generation share some of their quantum numbers with the quarks of the first generation. The masses of the charged leptons and the quarks increase with the generation number. For example $m_\mu \approx 200\ m_e$ and $m_\tau \approx 3500\ m_e$. Table 1.1 presents the 12 fundamental fermions, their electric charges and their masses.

	Leptons			Quarks		
	Particle	El. Charge	mass [GeV]	Particle	El. Charge	mass [GeV]
1st Gen	ν_e	0	~ 0	u	$+2/3$	0.005
	e	-1	0.0005	d	$-1/3$	0.003
2nd Gen	ν_μ	0	~ 0	c	$+2/3$	1.3
	μ	-1	0.106	s	$-1/3$	0.1
3rd Gen	ν_τ	0	~ 0	t	$+2/3$	174
	τ	-1	1.78	b	$-1/3$	4.5

Table 1.1: The 12 fundamental fermions divided into leptons and quarks, in three generations.

The SM vector bosons are the gluon (g), the photon (γ), the weak charged ($W^\pm$) and neutral current (Z) bosons. The gluon is the mediator of the strong force and carries a unique property called color charge. Quarks are the only fermions carrying the color charge and therefore, they are the only ones that couple to gluons. The electromgnetic force is carried by the photon and it is acted upon charged fermions, with a strength of about 20 times lower than the strong force for two protons in a nucleus. The weak force is mediated by the massive charged current W and neutral current Z bosons. All 12 fermions carry the charge of the weak interaction, called weak isospin, and hence they all interact weakly. The weak force is about 7 orders of magnitude weaker than the electromagnetic force for two protons in a nucleus.

The additional fourth force observed in nature is gravity and it is by far the weakest of the fundamental forces. It is about 36 orders of magnitude weaker than the electromagnetic force of two protons in a nucleus. It is responsible for the interaction of massive macroscopic objects and it is believed to be mediated by a massless, spin-2 particle called graviton. However, such a particle has never been observed. Table 1.2 presents the four fundamental forces and some of their properties.

Force	Boson	Mass [GeV]
Strong	g	0
Electromagnetic	γ	0
Weak	$W^{\pm}$	80.4
	Z	91.2
Gravity	G(?)	0

Table 1.2: The fundamental forces and their mediator.

The final piece of the SM is the Higgs boson which was discovered at CERN in 2012 [3, 4]. The Higgs mass was measured to be $m_H \equiv 125\,\text{GeV}$ and it differs from all other SM gauge bosons because it is a spin–0 scalar particle. It has a very special role, providing the mechanism by which the other SM particles acquire mass.

The SM is a local gauge invariant Quantum Field Theory (QFT) under the symmetry group SU(3) × SU(2) × U(1), which defines the interactions between the particles. Quantum Electrodynamics (QED) and Quantum Chromodynamics (QCD) are the gauge invariant QFTs describing electromagnetic and strong interactions respectively. The electroweak theory (EWK) provides a unified gauge invariant QFT of the electromagnetic and the weak interactions.

The QED Lagrangian density is invariant under a local gauge transformation of the group U(1). This invariance yields interaction of spin-1/2 fermions with electromagnetic fields, mediated by massless photons.

In the weak interactions the local gauge invariance of the Lagrangian under the symmetry group SU(2), results in 3 fields and thus 3 massless gauge bosons ($W^{\pm}$, Z) that correspond to the 3 generators of the group. It has been found that the weak interaction mediators are massive. This leads to the idea that there is an additional mechanism that gives masses to the W and Z bosons. This is the Higgs mechanism and will be described later in this section. It has been observed that the weak charged-current interaction has a left-handed (LH) structure in the sense that it couples to LH chiral fermions and right-hand (RH) chiral antifermions and not vise versa, indicating parity violation. For this reason the symmetry group is referred to as $\mathrm{SU(2)_L}$. The LH particles are organised into weak isospin doublets containing fermions that differ by unit charge. The RH particles and LH antiparticles are organised in weak isospin singlets.

The non-Abelian character of the SU(2) weak isospin symmetry group leads to gauge boson self-coupling with 3-line (ZW^+W^-, γW^+W^-) and 4-line ($W^+W^-W^+W^-$, ZZW^+W^-, $\gamma\gamma W^+W^-$, γZW^+W^-) vertices called trilinear and quartict gauge boson couplings. These couplings give rise to multi-boson production processes and their measurement is among the stringent test of the symmetry structure of the SM.

The unification of the weak and the electromagnetic QFT into one gauge theory is achieved by requiring its Lagrangian to be invariant under a $\mathrm{SU(2)_L} \times \mathrm{U(1)_Y}$ gauge symmetry. The Lagrangian consists of one charged and one neutral part. Due to the invariance under the gauge transformation three W and one B massless gauge boson fields arise. The third W component and the B field mix to form the photon and the neutral-current weak boson. The B gauge field couples to a new kind of charge, called weak hypercharge. This is why in the EWK unification the U(1) electromagnetic gauge symmetry is replaced by $\mathrm{U(1)_Y}$ gauge symmetry. The weak hypercharge can be expressed as a linear combination of the electromagnetic charge and the third component of the weak isospin following the relation $\Upsilon = 2(Q - I_3)$.

In the strong interaction the local gauge invariance of the QCD Lagrangian under the SU(3) group, gives rise to 8 gauge fields that correspond to the 8 massless gluons. Gluons are characterized by a unique charge, as mentioned earlier, that is called color and can be r, g or b. This is why the QCD symmetry group is referred to as $\mathrm{SU}(3)_c$. Only particles that carry color can couple to gluons. Interaction terms between gluons and quarks and gluon self-interaction terms appear in the QCD Lagrangian.

Object with colour charge are only observed confined into singlet states and objects with non-zero colour cannot propagate as free particles. This is known as the hypothesis of colour confinement and can explain the lack of free quarks observation. This is believed to originate from the fact that gluons carry colour charge and therefore interact with each other through triple (ggg) or quardruple (gggg) couplings.

The Higgs sector of the SM is constructed by 2 complex scalar fields in an isospin douplet with Y=1. One of the scalars is neutral and the other is electrically charged.

$$\phi(x) = \begin{pmatrix} \phi^+(x) \\ \phi^0(x) \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{1.1}$$

and the Higgs potential is of the form

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda(\phi^\dagger \phi)^2 \tag{1.2}$$

For $\lambda > 0$ and $\mu^2 < 0$ the potential has infinite set of degenerate non-zero minima satisfying $\phi^\dagger\phi = -\mu^2/2\lambda = \upsilon^2/2$.

The choice of one of the non-zero vacuum expectation states breaks the symmetry of the group and this is known as the spontaneously symmetry breaking.

The symmetry group $SU(2)_L \times U(1)_Y$ is spontaneously broken via the Higgs mechanism [5, 6] and the weak bosons acquire their mass. However, after the symmetry breaking the neutral photon should remain massless. Therefore, the vacuum expectation state should break $SU(2)_L \times U(1)_Y$ and remain invariant under $U(1)_{EM}$. The vacuum expectation state is chosen such that $\phi_1 = \phi_2 = \phi_4 = 0$ and $\phi_3 = \upsilon$, with quantum numbers $Q = 0$, $Y = 1$ and $I_3 = 1/2$ and so the field expanded about this minimum should be

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \upsilon + h(x) \end{pmatrix} \tag{1.3}$$

where h(x) is the Higgs field. The theoretical prediction of the weak bosons masses arise by substituting Equation 1.3 into the gauge invariant Lagrangian under the $SU(2)_L \times U(1)_Y$ group. The mass of the W boson is determined by the coupling constant g_W of $SU(2)_L$ gauge interactions and the vacuum expectation value of the Higgs field υ as $m_W = \frac{1}{2} g_W \upsilon$ and $m_Z = \frac{cos\theta_W}{m_W}$, where θ_W is the weak mixing angle. By using the measured vales of m_W and g_W the vacuum expectation value of the Higgs field is $\upsilon = 246$ GeV. The mass of the neutral scalar Higgs boson is given by $m_H = \sqrt{2\lambda}\upsilon$.

The massive fermions may also acquire their masses by the Higgs mechanism through the spontaneous symmetry breaking of the $SU(2)_L \times U(1)_Y$ gauge group. The fermionic mass term arises from the combination of the Higgs doublet and the LH fermionic doublet in a $SU(2)_L \times U(1)_Y$ invariant Lagrangian. After the spontaneous symmetry breaking the mass term that represents the coupling of the fermionic field to the Higgs field through its non-zero vacuum expectation value arises. The fermionic mass depends on the Higgs vacuum expectation value and the Yukawa coupling g_f following the expression

$$m_f = \frac{\upsilon g_f}{\sqrt{2}} \tag{1.4}$$

The Yukawa coupling is not predicted by the Higgs mechanism but it is dictated by the observed fermionic masses.

Additionally, the non-linear character of the Higgs potential leads to self-coupling terms with 3-line vertices (hhh) and 4-line vertices (hhhh) which will lead to di-Higgs production processes with very low cross sections. The measurement of these couplings will provide stringent test of the structure of the Higgs sector and therefore the observation of those processes is among the milestones of the HL-LHC.

1.2 Standard Model validation and shortcomings

The SM has successfully described myriad of particles and processes. The theoretical work of the SM was largely completed over the 1970's and the particle discoveries were finalised with the Higgs boson in 2012.

The theoretical predictions of the SM have been tested in a wide variety of experiments during the last decades. In most cases the experimental observations and the SM predictions were found to be consistent and precision measurements of SM parameters have been made in different experiments and found to be in good agreement. A selection of the SM tests are highlighted in Subsection 1.2.1 and some of the issues that the SM has failed to address are presented in Subsection 1.2.2.

1.2.1 Validation of the Standard Model

The internal consistency of the SM has been checked with the measurements of a large set of SM cross sections. Figure 1.1 shows the good agreement between the SM predicted and measured cross section from CMS at 7, 8 and 13 TeV Runs.

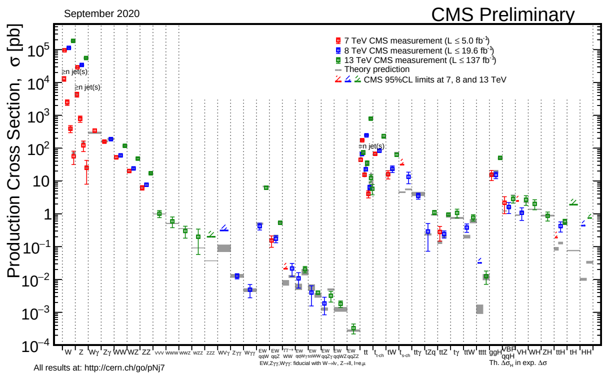

Figure 1.1: The SM production cross sections for various processes as measured CMS experiments at the LHC. Figure from Reference [7].

The discovery of the Higgs boson in 2012 is one of the most striking validation tests of the SM [3, 4]. The measurements by ATLAS and CMS Collaborations relay on the two decay channels $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$. The $H \to \gamma\gamma$ search is performed on a narrow peak over a smoothly falling background, in the invariant mass distribution of 2 high p_T photons. The background arises from prompt $\gamma\gamma$, γ+jets and di-jet events. The $H \to ZZ^* \to 4\ell$ search is performed on a narrow mass peak over a small continuous background which is mostly dominated by $q\bar{q} \to ZZ^*$ and $gg \to ZZ^*$ events. Figure 1.2 summarizes all measurements of the Higgs boson mass, including the individual and combined Run 1 measurements and the Run 2 measurement by ATLAS and CMS for both the $\gamma\gamma$ and the $ZZ^* \to 4\ell$ channels.

After the discovery of the Higgs boson, the last piece of the theory that was missing observation, some of the fundamental parameters of the SM became redundant thanks to the internal gauge symmetry of the theory. One of the most important is the mass of the W boson, which can be calculated at tree

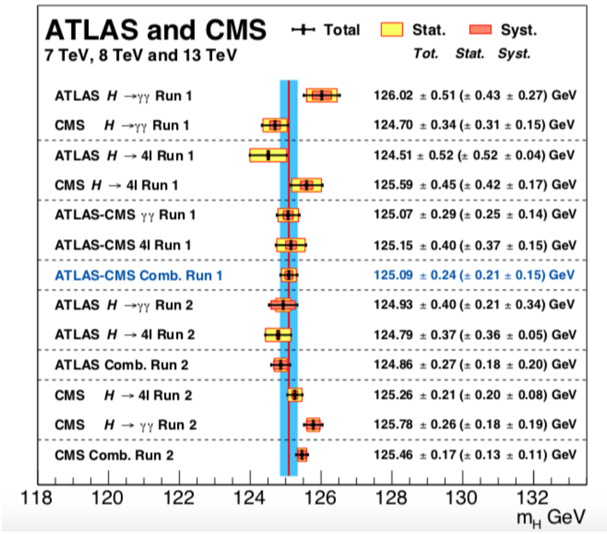

Figure 1.2: Summary of the CMS and ATLAS mass measurements in the $\gamma\gamma$ and ZZ channels in Run 1 and Run 2. Figure from Reference [8].

level from very precisely measured observables: the fine structure constant α, the Fermi constant G_F, and the mass of the Z boson. Loop corrections to the W boson mass can be calculated using additional observables which are also well constrained experimentally: the top quark mass and the Higgs boson mass. Therefore, the theory now provides a very accurate expectation for the W boson mass which can be tested with precise measurements down to the level of few MeV.

The W boson was discovered in 1983 at CERN [9, 10] and its mass measurement has been studied at the e^+e^- LEP experiments, the $p\bar{p}$ Tevatron experiments and at the LHC. These experiments will be briefly discussed in Sec. 1.7.

In the hadron collider searches the on-shell W boson is characterised by high p_T charged lepton and missing energy from its decay. The mass of the W boson is derived by comparing the distributions of the reconstructed W transverse mass, charged lepton p_{T}, and neutrino p_{T} to simulation for several values of M_W. The first precision measurement of the W boson mass was provided by the LEP experiments with data collected during its two runs [11]. This measurement was dominant in the world average value until 2007, when CDF at the Tevatron collider made its first precision measurement. The CDF and DØ Collaborations at Tevatron have made the measurement both with Run 1 and with part of the Run 2 data [12, 13]. In 2017 the ATLAS Collaboration provided a measurement of M_W with data collected during the 7 TeV Run [14]. Figure 1.3 presents the measurements of the W boson mass. The combination of those measurements accounting for the correlations between them yield the world average $M_W = 80.379 \pm 0.012$ GeV, which is also shown in Fig. 1.3.

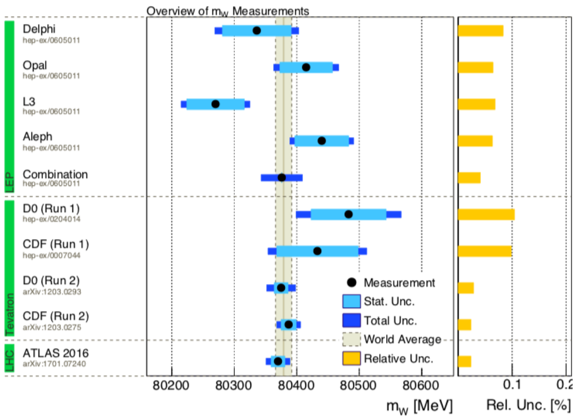

Figure 1.3: Overview of selected measurements of M_W , including the most precise measurements from LEP, Tevatron and the LHC. The figure is taken from [15]

The CDF II experiment at Fermilab Tevatron collider has recently published a measurement of the W boson mass [16]. This new result was obtained with data collected in $p\bar{p}$ collisions at a center-of-mass energy of 1.96 TeV and corresponding to 8.8 fb^{-1} of integrated luminosity. The data sample consists of approximately 4 million W boson candidates. The new measurement finds the W boson mass to be $M_W = 80433.5 \pm 9.4$ MeV (stat+syst uncertainty). The precision of the measurement exceeds all the previous measurements of the W boson mass and is in significant tension of 7.0 σ with the SM expectation of 80357 ± 6 MeV [17]. The latter suggests the possibility of improvements to the SM calculation or of extensions to the SM.

1.2.2 Shortcomings of the Standard Model

The SM of particle physics is a rigorous theory, incredibly accurate in its predictions, and widely validated in many collider and non-collider experiments. Despite its predictive power and its validity at energies up to the electroweak scales, there is observational evidence, that the SM fails to explain. This leads to theories stating that the SM is only a realization of a more fundamental, larger framework in our energy reach. In this concept the SM is not an ultimate theory of nature, but rather an effective theory, that successfully describes experimental observations up to the accessible energies.

The SM of particle physics contains 26 free parameters, namely 12 fermion masses including neutrinos, 3 gauge couplings, 2 parameters related to the Higgs potential (μ, λ), 1 CP-violating phase of QCD (typically set to 0), 3 flavor mixing angles + 1 CP-violating phase in the quark sector, and 3 flavor mixing angles + 1 CP-violating phase in the neutrino sector. These parameters are not emerging from theoretical principles but they are chosen to match the observation. The SM is unable to provide a good explanation for the observed mass pattern within a single fermionic generation, depicted in Fig. 1.4, or even why the fermions come in three generations.

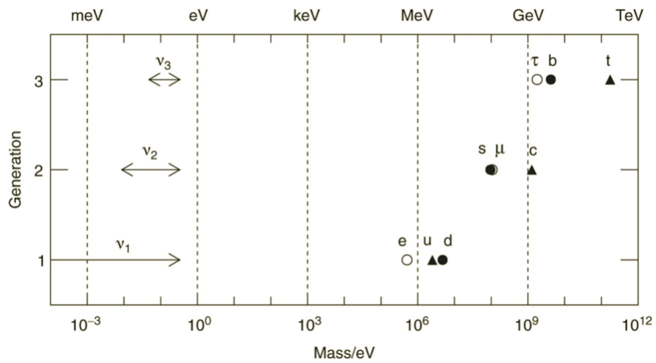

Figure 1.4: Fermionic masses of the 3 lepton and quarks generations. The neutrino masses are shown as ranges assuming the hierarchy of $(m_1 < m_2 < m_3)$ and assuming upper limits on the sum of neutrino masses from cosmological constants. Figure from Reference [1].

One subtle point of modern physics is gravity, that is described by the theory of general relativity in the framework of a classical gauge theory in the 4-dimensional space-time, but it is a non-renormalizble QFT. The weak strength of gravity compared to the other SM forces makes its role insignificant in microscopic interactions and therefore the predictive power of the SM is saved in almost all cases. Exceptions to this are extreme cases of strong gravitational effects such as the black holes where the SM cannot provide a good description without the theory of quantum gravity.

The rest of the section discusses open questions of the modern particle physics, that the SM fails to address, starting from the issues related to macrocosmic observations such as the nature of dark matter and dark energy, and the matter-antimatter asymmetry. The finely-tuned Higgs boson mass, which is known as the "Hierarchy problem" (see below), is an important open issue of the SM and it is discussed at the end of the section.

Dark Matter (DM)

As early as the 1930's, it was realised that the matter we see around us, is only a very small fraction of the total amount of matter in the Universe. The remainder is made up of matter, that does not interact electromagnetically, and thus called Dark Matter (DM). The most direct evidence for DM, is coming from the rotation curves of spiral galaxies, like the Milky Way. The rotation curve depicts the orbital circular velocity of the stars, as a function of the radial distance from the galactic center. The circular velocity is expected to follow $u(r) \sim \sqrt{\frac{GM(r)}{r}}$ from Newton's law of motion and of the gravitational force, where $M(r)$ is the total mass within r. One would expect the radial velocity to decrease with r, following $u(r) \sim \frac{1}{\sqrt{r}}$, however the rotation curves, like the one in Fig. 1.5, show that $u(r) = const.$ at higher r and thus $M(r) \sim r$. This implies that the mass content of the galaxy increases, and the circular velocity of the stars remains almost constant with r. This observation suggests, that there is a significant contribution to the matter of the galaxy that is non-luminous.

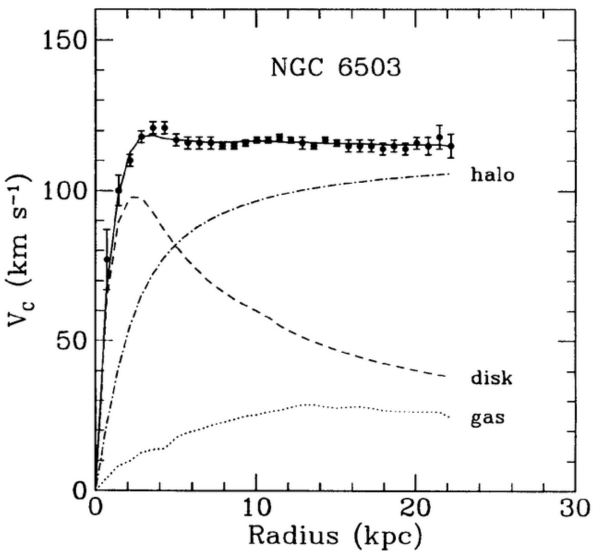

Figure 1.5: Rotation curves of NGC 6503 galaxy. The dotted curve represents the contribution of gas, the dashed curve the contribution of disk and dashed-dotted line the contribution of dark matter. Figure from Reference [18].

Further evidence for the DM is provided by the precision measurements of the cosmic microwave background (CMB) anisotropies, observed by satellite experiments like COBE [19], WMAP [20] and Planck [21]. Detailed analysis of the CMB anisotropies yield that the 5% of the total energy-matter density of the Universe, is in the form baryonic matter, 27% of it is in the form of cold (non-relativistic) dark matter, and the rest of it is in the form of "Dark Energy" (DE). This constitutes the standard model of cosmology, also called ΛCDM, where Λ refers to the non-zero cosmological constant associated with the Dark Energy, and CDM stands for Cold DM. Additionally, DM is required to explain the formation of gravity-initiated structures in the Universe from the very small structures in the early time to the large scale structures at later times.

Our understanding of cosmology and structure formation, together with the study of detailed simulations, provide insights on the nature of DM. ΛCDM favours cold DM, in the form a weakly interacting massive particle (WIMP), that can correctly predict the cosmological structure. Such particles arise naturally from extensions of the SM, such as the R-parity conserving supersymmetry (SUSY) discussed below. WIMPs are searched in collider experiments, expected to be directly observed through their production during the collisions. Complementary WIMP searches are conducted in direct DM experiments, in which the signal is the energy deposit from the DM and detector material interaction (LUX [22], XENON1T [23]). In indirect DM searches, the signal is SM particles produced by DM annihilating or decaying, in astrophysical sources (KM3NET [24], IceCube [25]).

In the context of ΛCDM the 68% of the total energy-mass budget of the Universe, is attributed to the DE. The term appeared in the bibliography in 1998, after the discovery of the accelerating expansion of the Universe. The simplest explanation of DE, is that it is a fundamental energy of space denoted as the cosmological constant Λ. The latter is predicted by Einstein's theory of general relativity to have a gravitational effect. The true nature of DE is yet to be understood. The solution to the open problem of the DE nature may be given by some modification of the classical theory of gravity.

Matter-antimatter asymmetry

This asymmetry refers to the imbalance of normal baryonic matter and antimatter in the observed Universe[1]. It is now clear that the Universe was hot in its early stages, and both matter and antimatter were present, interacting via pair production and annihilation. When the particle energies became too small for pair production to occur, almost all matter-antimatter pairs annihilated, and only a very small amount of baryonic matter survived. This procedure is called baryogenesis and the relic baryonic matter is responsible for the amount of the matter in our current Universe.

According to Shakarov [26] there are three conditions that should be fullfilled for the baryogenesis to occur: 1) baryonic number violation, so that transitions from system with $B = 0$ to systems with $B \neq 0$ is allowed, 2) CP violation, to allow for processes to happen at different rates for matter and antimatter and 3) interaction to occur out of thermal equilibrium. The baryonic-asymmetry interaction rate, should be lower than the expansion rate of the Universe, allowing particles and antiparticles not to be in thermal equilibrium. This would decrease the occurrence of their pair-annihilations.

Baryonic number violation is not predicted in SM of particle physics, and thus theories beyond the standard model, should be adopted to explain the mechanism of matter-antimatter asymmetry.

[1] In astrophysics and cosmology the term baryonic matter refers to both fermionic and baryonic matter.

Grand Unification

The coupling constants of the three SM interactions have similar strengths at the electroweak scale ($q^2 \sim 100\,\mathrm{GeV}$), $\alpha^{-1} \approx 128$, $\alpha_W^{-1} \approx 30$, $\alpha_s^{-1} \approx 9$, and they are running with energy (q). In QED, only the fermionic loops contribute to the photon self-energy, and the coupling strength α, increases with q. In strong interactions there are contribution to the renormalization of the theory both from fermionic, and gluon-gluon self interactions loops, resulting in decreasing α_S with energy. The coupling of the weak interaction, α_W, also decrease with energy, due to the weak boson self-interaction.

The running couplings tend to converge at very high values, but they do not meet exactly at the same point. In 1970's Georgi and Glashow suggested that SM symmetries, can be accommodated within a larger SU(5) symmetry [27], depicted in Fig. 1.6 (a). This is called the Grand Unification Theory (GUT) and brings together the running coupling strengths, although not-exactly at the same point, at very large energy scales of $\sim 10^{15}$ GeV.

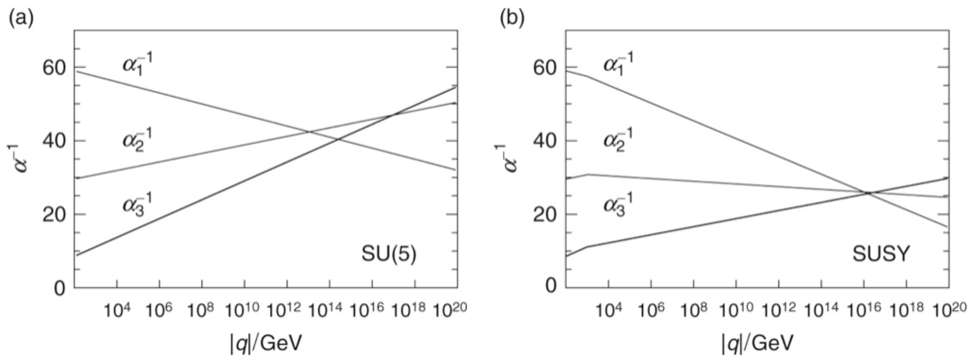

Figure 1.6: A representation of the running coupling constants (a) in the SU(5) GUT and (b) in a supersymmetric extension of SU(5) with new particles and masses around 1 TeV. $\alpha_1 \equiv \alpha_{QED}$, $\alpha_2 \equiv a_W$ and $\alpha_3 \equiv \alpha_s$

If new particles from theories beyond the standard model are accounted for, the running coupling strengths would change, due to additional loop contributions. For example, SUSY particles at mass scale $\Lambda_{SUSY} = 1\,\mathrm{TeV}$, would modify the running of the couplings, and make them converge into a single point at $q \sim 10^{16}\,\mathrm{GeV}$, shown in Fig. 1.6 (b). The convergence suggests, that the 3 interactions of the SM, are a low energy manifestation of some more general, yet-unknown, unified theory, which in higher energy scales, may even include gravity. However, it still remains unclear how quantum theory and general relativity should be unified into a consistent theory.

The hierarchy problem

Each QFT of the SM, applies the concept of renormalisation, with finite results obtained from higher order corrections. A correction to the Higgs mass squared (m_H^2) from a loop containing a fermion, with mass m_f, is shown in Fig. 1.7 (a). The Higgs coupling to the fermion, enters the Lagrangian with a term $-\lambda_f \bar{f} H f$, and the correction to the Higgs mass is of the form

$$\Delta m_H^2 = -\frac{| \lambda_f |^2}{8\pi^2} \Lambda^2 + ... \tag{1.5}$$

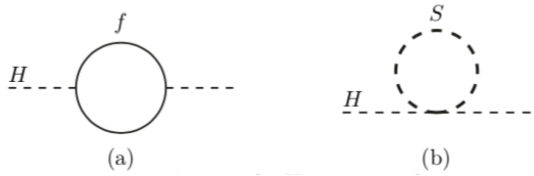

Figure 1.7: One-loop quantum corrections to the Higgs squared mass parameter, due to (a) a fermion f, and (b) a scalar S [2].

where Λ is the energy cutoff, up to which the extended theory of the SM is valid, such as the Planck mass scale ($\sim 10^{19}$ GeV).

At the Planck scale, the correction to m_H^2 is $\sim$ 30 orders of magnitude larger than the value of the Higgs mass observed (125 GeV). This suggests, that there are some "unnatural" cancellations between the higher order corrections and hence the hierarchy problem is also referred to as "naturalness problem".

Furthermore, there are contributions to the Higgs mass from bosonic loops shown in Fig. 1.7 (b). The coupling of the bosons with mass m_S to the Higgs boson enters the Lagrangian with a term $-\lambda_S |H|^2 |S|^2$ and the correction to the Higgs mass is:

$$\Delta m_H^2 = \frac{\lambda_S}{16\pi^2} \Lambda^2 + ... \tag{1.6}$$

The hierarchy problem can be solved either by accepting an extremely finely-tuned, coincidental, mechanism, spanning about 30 orders of magnitude, or by accepting that a yet unknown phenomenon can naturally cancel out the additional terms in the Higgs mass correction.

Comparing equations 1.5 and 1.6, one notices that both fermionic and bosonic corrections are proportional to Λ^2, and that the fermionic contribution has a negative sign while the bosonic is positive. This suggests that the contributions can cancel out each other, if for each fermion loop there exists a boson loop, and vice versa, and the couplings to the Higgs field satisfy the relation $\lambda_S = 2|\lambda_f|^2$.

Supersymmetry (SUSY) provides a natural solution to the hierarchy problem, by relating bosons and fermions and providing terms in the Higgs mass correction that cancel out each other. The residual corrections are logarithmicaly dependant on the particle mass and Λ. If the corrections to m_H^2 are to be kept naturally small ($\Delta m_H^2 < m_H^2$), the bosonic and fermionic partners should have small mass splittings of ≤ 1 TeV, motivating the theory of EWK-scale (light) SUSY.

There are a number of alternative theories suggested to address the problem of fine tuning over the years. Historically the first theory was called technicolour [28] and suggested a new strongly interacting particle with mass in the scale of TeV. In this theory the scalar states are not elementary but bounded states of fermion-antifermion [29]. In addition to technicolour, other theories have been proposed to explain the problem of fine-tuning more recently. Such an example is the idea that the gravitational scale is as low as a few TeV and the ultraviolet cut-off is close to the weak scale, therefore, the fine-tuning problem then evaporates [30].

1.3 Motivation for Supersymmetry and basic concepts

The theory of SUSY developed in the 1970s [31, 32, 33] and it is formulated in an extension of the Minkowski spacetime, known an superspace. SUSY is defined by the chiral transformations of Equations 1.8-1.10, as the fundamental symmetry of super-space. It has to be a broken symmetry which, after breaking, leaves two exact symmetries of ordinary space-time, namely the Lorentz symmetry of rotations and boosts and the Poincare symmetry of translations. SUSY doubles the spectrum of all particles and antiparticles by introducing their super-partners via the supercharge operator shown in Equation 1.7, a boson super-partner for each elementary fermion and a fermion super-partner for each elementary boson of the SM.

The SUSY transformations are performed with a fermionic spinor operator Q with spin 1/2

$$Q|fermion\rangle = |boson\rangle, \quad Q|boson\rangle = |fermion\rangle \tag{1.7}$$

The properties of Q and $Q^{\dagger}$, also known as *supercharge operators*, can be summarised in

$$\{Q, Q^{\dagger}\} = P^{\mu} \tag{1.8}$$
$$\{Q, Q\} = \{Q^{\dagger}, Q^{\dagger}\} = 0 \tag{1.9}$$
$$[P^{\mu}, Q] = [P^{\mu}, Q^{\dagger}] = 0 \tag{1.10}$$

where P^{μ} is the 4 momentum generator of spacetime translation.

Stable SUSY particle states contain both fermionic and bosonic states which are called *supermultiplets* and contain the same number of bosonic and fermionic degrees of freedom. The super-partners of the same supermultiplet transform to each other by linear combinations of Q and $Q^\dagger$, and differ by 1/2 unit in spin. If SUSY is not a broken symmetry, all supermultiplet members have the same mass, given that the squared mass operator $-P^2$ commutes with Q and $Q^\dagger$. Additionally, super-partners have common charge, color and weak isospin given that SUSY generators commute with gauge transformation generators. Therefore, SUSY retains all gauge symmetries of the SM.

The simplest supermultiplets are composed of a single Weyl fermion and two real, spin 0, scalar fields assembled into one complex scalar field. These supermultiplets are called *chiral* and the super-partners of the SM quarks and leptons are called "squarks"/"sleptons" or collectively "sfermions". Since the left- and right-handed (LH, RH) fermions have different gauge transformation properties in the SM, each state has its own complex scalar superpartner. The sfermions are denoted as $\widetilde{l}_L$, $\widetilde{l}_R$, $\widetilde{\nu}_l$, $\widetilde{q}_L$, $\widetilde{q}_R$, where the L and R refers to the helicity of the Weyl fermions. The super-partners $\widetilde{f}_L$ and $\widetilde{f}_R$ are scalar bosons and they are expected to appear in mixed mass eigenstates $\widetilde{f}_1$ and $\widetilde{f}_2$. Table 1.3 presents the simplest chiral supermultiplets.

The Higgs boson also resides in a chiral supermultiplet. The structure of SUSY theory requires the existence of 2 Higgs supermultiplets with $Y = \pm 1/2$, one that gives mass to up-type quarks H_u and one that gives mass to the down type quarks H_d. The weak isospin doublet of H_u is (H_u^+, H_u^0) and the one of H_d is (H_d^0, H_d^-). The Higgs super-partners are called higgsinos, they are denoted as $\widetilde{H}_u$ and $\widetilde{H}_d$ and they have spin equal to 1/2.

Names	Superfields	spin 0	spin 1/2
squarks-quarks	Q	$(\widetilde{u}_L\ \widetilde{d}_L)$	$(u_L\ d_L)$
(×3 families)	$\bar{u}$	$\widetilde{u}_R^*$	$u_R^\dagger$
	$\bar{d}$	$\widetilde{d}_R^*$	$d_R^\dagger$
sleptons-leptons	L	$(\widetilde{\nu}\ \widetilde{e}_L)$	$(\nu\ e_L)$
(×3 families)	$\bar{e}$	$\widetilde{e}_R^*$	$e_R^\dagger$
Higgs-Higgsinos	H_u	$(H_u^+\ H_u^0)$	$(\widetilde{H}_u^+\ \widetilde{H}_u^0)$
	H_d	$(H_d^0\ H_d^-)$	$(\widetilde{H}_d^0\ \widetilde{H}_d^-)$

Table 1.3: The chiral supermultiplets.

The second simplest supermultiplet is called *gauge* and it is composed of one spin-1 vector boson and a Weyl fermion with spin-1/2 called "gaugino". Color gauge interactions are mediated by spin-1 gluon whose spin-1/2 superpartner is called gluino ($\widetilde{g}$). The electroweak spin-1 gauge mediators ($W^\pm$, W^0, B^0) are associated to spin-1/2 super-partners called winos and bino and denoted as $\widetilde{W}^\pm$, $\widetilde{W}^0$, $\widetilde{B}^0$. The mixture of $\widetilde{W}^0$ and $\widetilde{B}^0$ are the zino $\widetilde{Z}^0$ and photino $\widetilde{\gamma}$. The gauge supermultiplets are summarized in Tab. 1.4. If gravity is included, the spin-2 predicted graviton belongs to a supermultiplet with a spin-3/2, masseless superpartner called gravitino.

Names	spin 1/2	spin 1
gluinos-gluons	$\widetilde{g}$	g
wino-W boson	$\widetilde{W}^{\pm}\ \widetilde{W}^{0}$	$W^{\pm}\ W^{0}$
bino-B boson	$\widetilde{B}^{0}$	B^{0}

Table 1.4: The gauge supermultiplets

The chiral and gauge supermultiplets make up the Minimal Supersymmetric Standard Model (MSSM). The lack of SUSY particle observation implies that the symmetry is broken and thus the masses of the SUSY particles differ from that of their SM partners. In order for the broken SUSY to provide a solution to the hierarchy problem, the theory is believed to be "softly" broken. This implies that the largest mass scale associated to SUSY breaking, should not be unnaturally large compared to the EWK scale and the mass splitting between the SM and SUSY partners, should not be more that 1TeV [34, 35].

1.4 General SUSY Lagrangian

This subsection follows Ref. [2] and aims at presenting the general structure of the SUSY field theory without describing in great detail the derivations. The spacetime metric adopted here is $\eta_{\mu\nu} = diag(-1, +1, +1, +1)$ in accordance with Reference [2].

The SUSY Lagrangian will contain the chiral and gauge supermultiplets, their interaction terms and the soft breaking term.

Lets start from the free chiral supermultiplet part, the simplest action can be expressed as

$$S = \int d^4x(\mathcal{L}_{Scalar} + \mathcal{L}_{Fermion}) \tag{1.11}$$

where $\mathcal{L}_{Scalar} = -\partial^{\mu}\phi^{*}\partial_{\mu}\phi$ and $\mathcal{L}_{Fermion} = i\psi^{\dagger}\bar{\sigma}^{\mu}\partial_{\mu}\psi$, where ϕ denotes the scalar bosonic field and ψ the fermionic field.

A simple transformation that turns the bosonic field into something involving the fermionic field ψ, can be of the form

$$\delta\phi = \epsilon\psi, \quad \delta\phi^{*} = \epsilon^{\dagger}\psi^{\dagger} \tag{1.12}$$

where ϵ is an infinitesimal anti-commuting Weyl fermion object that parametrizes the SUSY transformation.

In order for the action 1.11 to be invariant under SUSY transformation the $\delta\mathcal{L}_{Scalar} = -\delta\mathcal{L}_{fermion}$ relation should hold. For this to happen the transformation of the fermionic field should be linear in ϵ and should contain $\partial_\mu\phi$, so the SUSY transformation of the fermionic field is

$$\delta\psi = -i(\sigma^\mu\epsilon^\dagger)\partial_\mu\phi, \quad \delta\psi^\dagger = i(\epsilon\sigma^\mu)\partial_\mu\phi^* \tag{1.13}$$

The transformations 1.12 and 1.13 result in $\delta S = 0$ and so the action is invariant under the SUSY transformation.

The last piece for proving that the transformation described previously is a SUSY transformation, is to show the closure of the theory or in other words show that the commutator of SUSY transformations, parametrized by ϵ_1 and ϵ_2, generates another symmetry of the theory. Such a derivation shows that the fermionic and bosonic terms of the commutators vanish if the equation of motion of the bosonic field is enforced ($\bar{\sigma}^\mu\partial_\mu\psi = 0$), and thus the symmetry only closes on-shell and not off-shell. This is resolved by introducing an auxiliary, non propagating, complex scalar field F and a new term ($\mathcal{L}_{Auxiliary} = F^*F$) in the action of the free chiral SUSY Lagrangian 1.11.

The auxiliary field can be transformed as

$$\delta F = -i\epsilon^\dagger\bar{\sigma}^\mu\partial_\mu\psi, \quad \delta F^* = i\partial_\mu\psi^\dagger\bar{\sigma}^\mu\epsilon \tag{1.14}$$

and the additional terms of ϵF in $\delta\psi$ and $\epsilon^\dagger F^*$ in $\delta\psi^\dagger$ are added in the fermionic transformation 1.13. These terms in $\delta\mathcal{L}_{Fermion}$ cancel out $\delta\mathcal{L}_{Auxiliary}$ and thus $\delta S = 0$ and the theory is invariant under the SUSY transformation. Additionally, one can show for the theory that

$$[\delta\epsilon_1, \delta\epsilon_2]X = i(-\epsilon_1\sigma^\mu\epsilon_2 + \epsilon_2\sigma^\mu\epsilon_1^\dagger)\partial_\mu X \tag{1.15}$$

where $X = \psi\psi^\dagger$, $\phi\phi^*$, FF^*. The commutator of the SUSY transformations, parametrized by ϵ_1 and ϵ_2, gives the derivative of the original field without imposing any equation of motion, and thus the symmetry closes off-shell.

In summary, the free Lagrangian density of the chiral supermultiplet labeled by the index i, consists of the scalar, the fermionic and the auxiliary part and it is given by:

$$\mathcal{L}_{Free} = -\partial^\mu\phi^{*i}\partial_\mu\phi_i + i\psi^{\dagger i}\bar{\sigma}^\mu\partial_\mu\psi_i + F^{*i}F_i \tag{1.16}$$

The most general interaction chiral supermultiplet Lagrangian is

$$\mathcal{L}_{Int} = (-\frac{1}{2}W^{ij}\psi_i\psi_j + W^iF_i) + c.c \tag{1.17}$$

where W^{ij} and W^i are polynomials of first and second degree in the scalar fields ϕ and ϕ^*. Requiring $\mathcal{L}_{Int}$ to be invariant under SUSY transformation the *superpotential* is introduced

$$W = \frac{1}{2}M^{ij}\phi_i\phi_j + \frac{1}{6}y^{ijk}\phi_i\phi_j\phi_k \tag{1.18}$$

where M^{ij} is the symmetric mass matrix for fermionic field and y^{ijk} is a Yukawa coupling of a scalar fiend and two fermionic fields. The superpotential is a holomorphic function of the field ϕ, in the sense that it only depends on fields of the chiral supermultiplet and not their complex conjugate. The terms W^{ij} and W^i can be expressed in terms of the superpotential as

$$W^{ij} = \frac{\delta^2 W}{\delta\phi_i\delta\phi_j}, \quad W^i = \frac{\delta W}{\delta\phi_i} \tag{1.19}$$

The auxiliary field F can be eliminated from the Lagrangian by using the equations of motion $F_i = -W_i^*$ and $F^{*i} = -W^i$ and the chiral supermultiplet Lagrangian containing the free and the interaction part can now be written as

$$\mathcal{L}_{Chiral} = -\partial^\mu\phi^{*i}\partial_\mu\phi_i + i\psi^{\dagger i}\bar{\sigma}^\mu\partial_\mu\psi_i - \frac{1}{2}(W^{ij}\psi_i\psi_j + W^*_{ij}\psi^{\dagger i}\psi^{\dagger j}) - W^iW^*_i \tag{1.20}$$

The last term is the scalar potential of the theory, $V(\phi,\phi^*)$, so the full Lagrangian density can be expressed as

$$\begin{aligned}\mathcal{L}_{Chiral} =& -\partial^\mu\phi^{*i}\partial_\mu\phi_i - V(\phi,\phi^*) + i\psi^{\dagger i}\bar{\sigma}^\mu\partial_\mu\psi_i - \frac{1}{2}M^{ij}\psi_i\psi_j - \frac{1}{2}M^*_{ij}\psi^{\dagger i}\psi^{\dagger j}\\ & -\frac{1}{2}y^{ijk}\phi_i\psi_j\psi_k - \frac{1}{2}y^*_{ijk}\phi^{*i}\psi^{\dagger j}\psi^{\dagger k}\end{aligned} \tag{1.21}$$

The scalar potential can be written in terms of the superpotential as

$$\begin{aligned}V(\phi\phi^*) = & M^*_{ik}M^{kj}\phi^{*i}\phi_j + \frac{1}{2}M^{in}y^*_{jkm}\phi_i\phi^{*j}\phi^{*k}\\ & + \frac{1}{2}M^*_{in}y^{jkn}\phi^{*i}\phi_j\phi_k + \frac{1}{4}y^{ijn}y^*_{klm}\phi_i\phi_j\phi^{*k}\phi^{*l}\end{aligned} \tag{1.22}$$

A vector supermultiplet consists of a massless boson field A^α_μ and a two-component Weyl gaugino λ^α, where α runs over the generators of the gauge group. Under the gauge transformation the fields transform as

$$\begin{aligned}&A^\alpha_\mu \to A^\alpha_\mu + \partial_\mu\Lambda^\alpha + gf^{abc}A^b_\mu\Lambda^c\\ &\lambda^\alpha \to gf^{abc}\lambda^b\Lambda^c\end{aligned} \tag{1.23}$$

Λ^α is an infinitesimal gauge transformation, g is the gauge coupling and f^{abc} are the totally antisymmetric structure constants. The Lagrangian density of the free gauge supermultiplet is

$$\mathcal{L}_{Gauge} = -\frac{1}{4}F^\alpha_{\mu\nu}F^{\mu\nu}_\alpha + i\lambda^{\dagger\alpha}\bar{\sigma}^\mu\nabla_\mu\lambda^\alpha + \frac{1}{2}D^\alpha D^\alpha \tag{1.24}$$

where $F^\alpha_{\mu\nu}$ is the Yang-Mills strength and $\nabla_\mu\lambda^\alpha$ is the covariant derivative of the gaugino field

$$\begin{aligned} F^\alpha_{\mu\nu} &= \partial_\mu A^\alpha_\nu - \partial_\nu A^\alpha_\mu + gf^{abc}A^b_\mu A^c_\nu \\ \nabla_\mu\lambda^\alpha &= \partial_\mu\lambda^\alpha + gf^{abc}A^b_\mu\lambda^c \end{aligned} \tag{1.25}$$

D^α is a bosonic auxiliary field with no kinetic term, introduced to make the symmetry consistent off-shell, in analogy with the F auxiliary field introduced in the chiral supermultiplet Lagrangian density.

Now we can obtain the general SUSY Lagrangian density that contains both the chiral and gauge supermultiplets and their interaction terms. In order to have a general SUSY Lagrangian which is gauge invariant, the partial derivatives in Eq. 1.21 must be replaced by the covariant derivative $\nabla_\mu = \partial_\mu - igT^\alpha A^a_\mu$.

$$\partial_\mu\phi_i \to \nabla_\mu\phi_i = \partial_\mu\phi_i - igA^\alpha_\mu(T^\alpha\phi)_i \tag{1.26}$$

$$\partial_\mu\phi^{i*} \to \nabla_\mu\phi^{i*} = \partial_\mu\phi^{i*} + igA^\alpha_\mu(\phi^i T^\alpha)^i \tag{1.27}$$

$$\partial_\mu\psi_i \to \nabla_\mu\psi_i = \partial_\mu\psi_i - igA^\alpha_\mu(T^\alpha\psi)_i \tag{1.28}$$

This achieves the coupling of the vector bosons of the gauge supermultiplet (A^α_μ) to the scalars and fermions of the chiral supermultiplet.

The general SUSY Lagrangian density is

$$\begin{aligned} \mathcal{L}_{SUSY} =& \mathcal{L}_{Chiral} + \mathcal{L}_{Gauge} \\ & -\sqrt{2}g(\phi^* T^\alpha\psi)\lambda^\alpha - \sqrt{2}g\lambda^{\dagger\alpha}(\psi^\dagger T^\alpha\phi) + g(\phi^* T^\alpha\phi)D^\alpha \end{aligned} \tag{1.29}$$

Where, T^α is the generator of the gauge transformation, $\mathcal{L}_{Chiral}$ is presented in Eq. 1.21 but here the partial derivatives are replaced by the covariant derivatives, and $\mathcal{L}_{Gauge}$ is presented in Eq. 1.24. The last three terms of Eq. 1.29 correspond to the coupling of the fermionic gaugino λ^α and the bosonic auxiliary field D^α, to the scalars (ϕ, ϕ^*) and fermions (ψ, $\psi^\dagger$) of the chiral supermultiplet.

The two terms of the Lagrangian density that contain the gauge auxiliary field are $\frac{1}{2}D^\alpha D^\alpha$ from the free gauge part (Eq. 1.24) and $g(\phi^* T^\alpha\phi)D^\alpha$ interaction term (Eq. 1.29). The resulting equation of motion of the gauge auxiliary field is $D^\alpha = -g(\phi^* T^\alpha\phi)$. This implies that the gauge auxiliary field, similarly to the scalar auxiliary fields (F_i), can be expressed in terms of the scalar fields, and thus the scalar potential can be written as

$$V(\phi,\phi^*) = F^{*i}F_i + \frac{1}{2}\sum_\alpha D^\alpha D^\alpha = W^*_i W^i + \frac{1}{2}\sum_\alpha g^2_\alpha(\phi^* T^\alpha\phi)^2 \tag{1.30}$$

The first term is called "F-term" and it is fixed by the Yukawa coupling and fermionic mass while the second term is called "D-term" and it is fixed by the gauge interactions. The sum accounts for the different gauge couplings a group can have.

The interaction vertices as described by the superpotential are depicted in Figures 1.8 and 1.9. The arrows point in the direction of propagation for ϕ and ψ fields and opposite direction for ϕ^* and $\psi^\dagger$. The dashed lines correspond to bosonic fields and the solid lines to the fermionic fields. The vertices shown in Figure 1.8 are all determined by the dimensionless parameters y^{ijk}. The first two vertices correspond to the term $\frac{1}{2}y^{ijk}\phi_i\psi_j\psi_k$, $\frac{1}{2}y^*_{ijk}\phi^{i*}\psi^\dagger_j\psi^\dagger_k$ which are the Yukawa interaction and its complex conjugate from equation 1.21. The third vertex shows the scalar interactions described by the term $\frac{1}{4}y^{ijn}y^*_{klm}\phi_i\phi_j\phi^{*k}\phi^{*l}$ in equation 1.22.

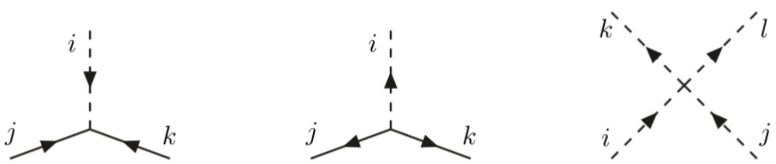

Figure 1.8: Dimensionless interaction vertices corresponding to terms of the equations 1.21 and 1.22 [2].

The Figure 1.9 shows supersymmetric vertices with coupling dimensions of $[mass]$ and $[mass]^2$. The vertices from left to right correspond to the (scalar)^3 interactions determined by the mass parameter and the Yukawa coupling as described in term $\frac{1}{2}M^{in}y^*_{jkn}\phi_i\phi^{j*}\phi^{k*}$ and $\frac{1}{2}M^*_{in}y^{jkn}\phi^{i*}\phi_j\phi_k$ of equation 1.22. The fermionic propagators described by the terms $\frac{1}{2}M^{ij}\psi_i\psi_j$ and $\frac{1}{2}M^*_{ij}\psi^{i\dagger}\psi^{j\dagger}$ in equation 1.21 and the scalar propagator described by the term $M^*_{ik}M^{kj}\phi^{i*}\phi_j$ in equation 1.22.

Figure 1.10 depicts the gauge interaction vertices. The curly lines correspond to the gauge boson fields and the curly and solid lines overlaid correspond to the gaugino fields. The vertices (a) and (b) correspond to gauge boson interactions represented in term $\frac{1}{4}F^a_{\mu\nu}F^{\mu\nu}_a$ in equation 1.24. These gauge interactions are similar to the gluon and electroweak boson self-interaction vertices of the SM. The

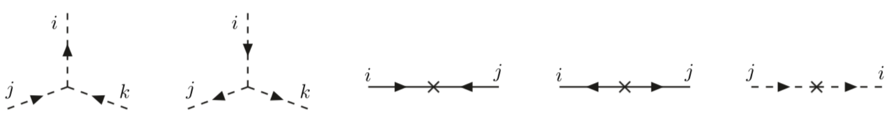

Figure 1.9: Dimensionfull interaction vertices described by terms of the equations 1.21 and 1.22 [2].

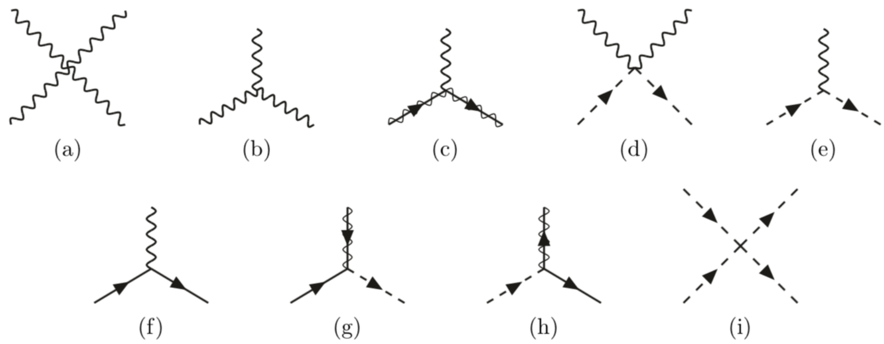

Figure 1.10: SUSY gauge interaction vertices [2].

vertices (c), (d), (e), (f) present the interactions of the gauge bosons and fermions and scalar fields in equations 1.25 and 1.26-1.28. The vertex (g) corresponds to the term $\sqrt{2}g(\phi^* T^a \psi)\lambda^a$ of equation 1.29 and represents the coupling of the gaugino to the chiral fermion and complex scalar field of. Respectively, the vertex (h) is the complex conjugate of (g) and represents the term $\sqrt{2}g\lambda^{a\dagger}(\psi^\dagger T^a \phi)$ from the same equation. Finally, the vertex (i) shows the scalar quartic interaction vertex determined by the gauge couplings. These interactions correspond to the term $\frac{1}{2}\sum_a g_a^2(\phi^* T^a \phi)^2$ of equation 1.30.

1.5 Minimal Supersymmetric Standard Model Extension

The chiral and gauge supermultiplet described in subsection 1.3 define the MSSM. This is a $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge theory that is softly broken and its superpotential is given by

$$W_{MSSM} = \bar{u}y_u QH_u + \bar{d}y_d QH_d + \bar{e}y_e LH_d + \mu H_u H_d \tag{1.31}$$

where H_u, H_d, Q, L, $\bar{u}$, $\bar{d}$, $\bar{e}$ are the chiral superfields, y_u, y_d, y_e are dimensionless Yukawa couplings and μ is the SUSY Higgs boson mass parameter.

The top quark, bottom quark and tau lepton are the heaviest fermions in the SM and therefore it is often approximated that only the (3,3) family components of each of Yukawa coupling are important. With this approximation the MSSM superpotential can be written as a function of the third fermionic family and the Higgs field, in terms of separate weak isospin components ($Q_3 = (tb)$, $L_3 = (\nu_\tau \tau)$, $H_u = (H_u^+ H_u^0)$, $H_d = (H_d^0 H_u^-)$, $\bar{u}_3 = \bar{t}$, $\bar{d}_3 = \bar{b}$, $\bar{e}_3 = \bar{\tau}$) as

$$
\begin{aligned}
W_{MSSM} \approx y_t(\bar{t}tH_u^0 + \bar{t}bH_u^\dagger) - y_b(\bar{b}tH_d^- - \bar{b}bH_d^0) \\
-y_\tau(\bar{\tau}\nu H_d^- - \bar{\tau}\tau H_d^0) + \mu(H_u^\dagger H_d^- - H_u^0 H_d^0)
\end{aligned}
\tag{1.32}
$$

The vertices presented in Figure 1.11 correspond to the first part of the first term in equation 1.32, for the coupling of the top quark to the Higgs or higgisno and the top squark. Vertex (a) corresponds to the top quark coupling to the neutral scalar Higgs and to the top anti-quark, vertex (b) corresponds to the coupling of the LH top squark to the neutral Higgsino and top quark and vertex (c) shows the coupling of the top anti-squark to the higgsino and the top quark. Similar vertices arise for the second part of the first term in equation 1.32 substituting H_u^0 with $H_u^\dagger$ and h_L with $-b_L$ with tildes where is appropriate.

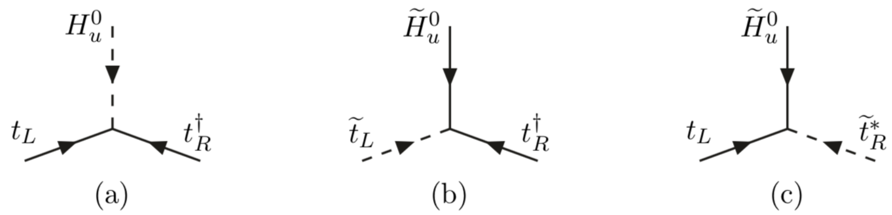

Figure 1.11: Top quark Yukawa couplings presented in equation 1.31 [2].

Three of the scalar quartic interactions, of the last term of equation 1.22, with strength proportional to y_t are shown in Figure 1.12.

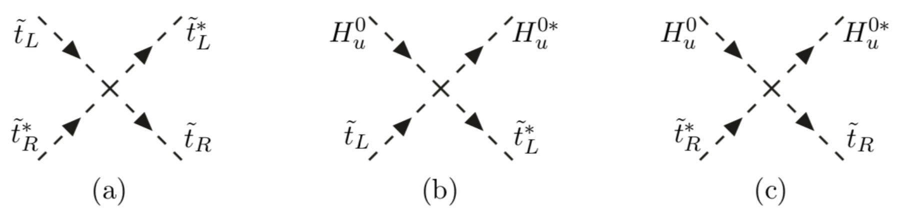

Figure 1.12: Some of the interactions of (scalar)^4 with strength proportional to y_t^2 [2].

The vertices in Figure 1.13 depict the couplings of (squark, quark) pair to gluino (vertex (a)) and of the (squark, quark), (slepton, lepton) and (Higgs, Higgsino) to Winos and Bino (vertices (b) and (c)). For each of these diagrams there is another diagram with the arrows reversed. It should be noted that Winos couple only to LH squarks and sleptons and that (slepton, lepton) and (Higgs, Higgsino) pairs do not couple to gluino.

Equation 1.31 describes the minimal superpotential that is sufficient to produce a viable model. Additional gauge invariant and holomorphic, in the chiral

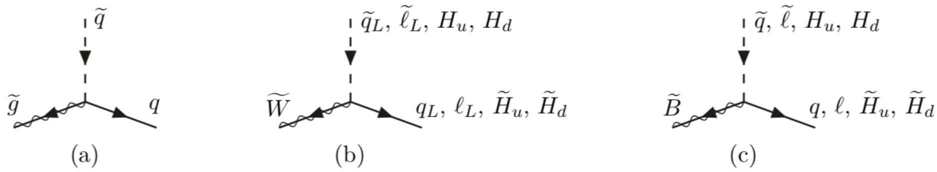

Figure 1.13: Couplings of gluino, wino and bino to the MSSM pairs of (scalar, fermion)

superfields, terms can be written in the W_{MSSM}, but they are ignored because they violate leptonic (L) and baryonic (B) number conservation.

The non-observed proton decay demonstrates the most obvious experimental constraint for non-violating L and B conservation. A new symmetry can be added in the MSSM in order to eliminate the L and B violating terms in the superpotential. This symmetry is called R-parity and has eigenvalues

$$P_R := (-1)^{3(B-L)+2s} \tag{1.33}$$

where s denotes the spin of the particle. R-parity is a Z_2 [2] symmetry and transforms differently the fields of the same supermultiplet. All the SM particles have $P_R = +1$ while SUSY particles have $P_R = -1$. The conservation of R-parity implies that the R-product in every interaction vertex should be $+1$. Therefore, every R-parity conserving SUSY interaction vertex must involve an even number of SUSY particles. This means that every R-parity conserving SUSY particle must decay to one SUSY particle and one or more SM particles. Furthermore, the last product of the decay chain of an R-parity conserving SUSY particle is called the lightest SUSY particle (LSP). The LSP is stable, electrically neutral, interacts only weakly with ordinary matter, and is a good DM candidate.

SUSY is realised to be a broken symmetry of nature since none of its predicted particles is yet observed. In order to maintain its ability to solve the hierarchy problem, the SUSY breaking should be soft and the mass differences between SM and SUSY partners should be $\sim$ 1TeV.

The exact breaking mechanism of the theory is not yet known and thus an adhoc "soft" breaking mechanism is considered. The most general "soft" breaking terms that can be added in the SUSY Lagrangian density are

$$\mathcal{L}_{soft} = -(\frac{1}{2}M_\alpha \lambda^\alpha \lambda^\alpha + \frac{1}{6}\alpha^{ijk}\phi_i\phi_j\phi_k + \frac{1}{2}b^{ij}\phi_i\phi_j) + c.c - (m^2)^i_j \phi^{j*}\phi_i \tag{1.34}$$

where M_α are the gaugino masses for each gauge group, b^{ij} and $(m^2)^i_j$ are the scalar mass terms and α^{ijk} is the trilinear scalar coupling.

[2] A Z_n symmetry leaves invariant a plane figure after a rotation of 2π/n radians.

The vertices in Figure 1.14 show the soft SUSY breaking terms of equation 1.34. Vertex (a) corresponds to the gaugino mass term $\frac{1}{2}M_a\lambda^a\lambda^a$, vertex (b) corresponds to the scalar squared mass m^2 term $(m^2)^i_j\phi^{j*}\phi_i$, vertex (c) corresponds to the term $\frac{1}{2}b^{ij}\phi_i\phi_j$ and vertex (d) corresponds to the term $\frac{1}{6}a^{ijk}\phi_i\phi_j\phi_k$.

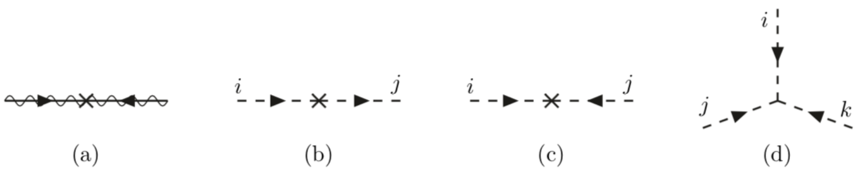

Figure 1.14: Soft SUSY breaking terms from equation 1.34 [2].

In the case of the MSSM Eq. 1.34 becomes

$$\begin{aligned}\mathcal{L}^{MSSM}_{Soft} =& -\frac{1}{2}(M_3\widetilde{g}\widetilde{g} + M_2\widetilde{W}\widetilde{W} + M_1\widetilde{B}\widetilde{B} + c.c)\\ & -(\widetilde{\bar{u}}\alpha_u\widetilde{Q}H_u - \widetilde{\bar{d}}\alpha_d\widetilde{Q}H_d - \widetilde{\bar{e}}\alpha_e\widetilde{L}H_d + c.c)\\ & -\widetilde{Q}^\dagger m^2_Q\widetilde{Q} - \widetilde{L}^\dagger m^2_L\widetilde{L} - \widetilde{\bar{u}}m^2_{\bar{u}}\widetilde{\bar{u}}^\dagger - \widetilde{\bar{d}}m^2_{\bar{d}}\widetilde{\bar{d}}^\dagger - \widetilde{\bar{e}}m^2_{\bar{e}}\widetilde{\bar{e}}^\dagger\\ & -m^2_{H_u}H^*_uH_u - m^2_{H_d}H^*_dH_d\\ & -(bH_uH_d + c.c)\end{aligned} \tag{1.35}$$

where M_3, M_2 and M_1 are the gluino, wino and bino mass terms respectively, m_Q, m_L, $m_{\bar{u}}$, $m_{\bar{d}}$ and $m_{\bar{e}}$ are the mass matrices for the squarks and sleptons, the coefficients α_u, a_d and α_e are the trilinear coupling matrices between the Higgs fields and the sfermions and $m^2_{H_u}$, $m^2_{H_d}$ and b are scalar mass parameters contributing to the Higgs potential. It is expected that the mass terms are of the order of ~ 1 TeV.

The soft breaking mechanism of SUSY results in 4 physical Higgs states denoted as h, H, A, $H^\pm$ together with their 4 super-partners. These states are mixing with the gauginos and give rise to neutralinos ($\widetilde{\chi}^0_i$) and charginos ($\widetilde{\chi}^\pm_i$), that will be described in more detail in Sec. 1.6

The "soft" MSSM breaking mechanism adds a large amount (105) of free parameters in the model. The interpretation of the SUSY searches is made feasible by considering approaches such as the minimal supergravity (mSUGRA) or Constraint Minimal Supersymmetric Standard Model (CMSSM) that adds constraints and reduces the number of free parameters of the model to 5. The mSUGRA is a local SUSY model incorporating gravity at the GUT scale, which then mediates the soft global SUSY breaking at the electroweak scale [36]. It mainly excludes parameter space in squark and gluino production, which are the processes with the maximum production cross section at LHC. The free parameters of mSUGRA are

- m_0: common sfermion mass at GUT scale
- $m_{1/2}$: common gaugino mass in GUT scale
- A_0: common trilinear coupling in GUT scale
- $tan\beta$: the ratio of the vacuum expectation values of the two Higgs doublets
- $sign(\mu)$

Another approach is the phenomenological-MSSM (pMSSM) model [37], which uses experimental data to eliminate free parameters that are highly constrained and thus reduces the number of free parameters to 19.

Additionally, simplified models [38, 39] are considered for the interpretations in most of the SUSY searches conducted at the LHC and probing natural SUSY. These model are designed to involve only a few new particles produced during pp collisions and their interactions, while the rest of the particles are integrated out. The simplified models are limits of more general SUSY theories. They are described by a small number of parameters, directly related to the observables from the collider experiments, such as the masses of the new particles, the cross sections and the branching fractions (BF). A generic signature of a simplified model contains a number of SM particles and the LSP SUSY particles at the final state and assumes 100% BF for the decay of the superpartners. The LSP interacts weakly and therefore it escapes detection leading to missing energy in the final state measured on the transverse plane to the beam, also refereed to as MET. In the simplified models the two SUSY particles produced through the hard scattering are assumed to be mass-degenerate. Since there is no observation of SUSY particle yet, assumptions on their masses and the mass differences between the pair produced SUSY particles and the LSP mass should be made. Therefore the mass of the pair produced SUSY particle and the mass splitting between the latter and the LSP are two useful parameters of the models. A mass configuration of pair produced SUSY particle and mass splitting is called mass point. When the mass splitting that is probed by the model is small (< 50 GeV) the SUSY mass spectra is called compressed. Both pMSSM and simplified models are considered for the interpretation of the SUSY analysis that will be presented in Chap. 4.

1.6 The Neutralinos and Charginos mass spectra

The neutral higgsinos ($\widetilde{H}_u^0$, $\widetilde{H}_d^0$) and neutral electroweak gauginos (also referred to as electroweakinos) ($\widetilde{W}^0$ and $\widetilde{B}$) are combined and form 4 mass eigenstates called neutralinos, denoted as $\widetilde{\chi}_i^0$ or $\widetilde{N}_i$. The charged higgsinos ($\widetilde{H}_u^+$ and $\widetilde{H}_d^-$) and charged electroweak gauginos ($\widetilde{W}^{\pm}$) couple to form 2 mass eigenstates called charginos, denoted as $\widetilde{\chi}_j^{\pm}$ or $\widetilde{C}_j$. By convention the mass eigenstate indices follow accenting mass order. In the R-parity conserving models the $\widetilde{\chi}_1^0$ is the lightest neutral SUSY particle.

The mixed states are characterized by a number of parameters. For charginos these are the wino mass parameter M_2, the higgsino mass parameter μ, and $tan\beta$, and for neutralinos these are the same parameters in addition to the bino mass parameter M_1.

The chargino mass matrix is

$$\mathcal{M}_{\widetilde{\chi}^\pm} = \begin{pmatrix} M_2 & \sqrt{2}m_W sin\beta \\ \sqrt{2}m_W cos\beta & \mu \end{pmatrix} \tag{1.36}$$

where the off-diagonal terms arise from gauge coupling of winos and higgsinos.

The mass-squared chargino eigenvalues arise from the diagonalization of the symmetric $\mathcal{M}_{\widetilde{\chi}^\pm}\mathcal{M}^T_{\widetilde{\chi}^\pm}$ matrix and are

$$m^2_{\widetilde{\chi}^\pm_{1/2}} = \frac{1}{2}[M_2^2 + \mu^2 + 2m_W^2 \mp \sqrt{(M_2^2 + \mu^2 + 2m_W^2)^2 - 4(\mu M_2 - m_W^2 sin2\beta)}] \tag{1.37}$$

The neutralino mass eigenvalues are obtained by the diagonalization of the neutralino mass matrix

$$\mathcal{M}_{\widetilde{\chi}^0} = \begin{pmatrix} M_1 & 0 & -cos\beta sin\theta_W m_Z & sin\beta sin\theta_W m_Z \\ 0 & M_2 & cos\beta cos\theta_W m_Z & -sin\beta \cos\theta_W m_Z \\ -cos\beta sin\theta_W m_Z & cos\beta \cos\theta_W m_Z & 0 & -\mu \\ sin\beta sin\theta_W m_Z & -sin\beta cos\theta_W m_Z & -\mu & 0 \end{pmatrix} \tag{1.38}$$

The $\mathcal{M}_{\widetilde{\chi}^0}$ defines 4 mass eigenstates described by 3 mass parameters (M_1, M_2 and μ). The bino-like LSP arises when $M_1 \ll M_2, \mu$ and the wino-like LSP when $M_2 \ll M_1, \mu$. In the bino LSP scenario the mass eigenvalues of the lightest and next to lightest neutralinos are

$$m_{\widetilde{\chi}_1^0} = M_1 - \frac{m_Z^2 sin^2\theta_W (M_1 + \mu sin2\beta)}{\mu^2 - M_1^2} \tag{1.39}$$

$$m_{\widetilde{\chi}_2^0} = M_2 - \frac{m_W^2 (M_2 + \mu sin2\beta)}{\mu^2 - M_2^2} \tag{1.40}$$

In the case of a higgsino LSP which arises when $\mu \ll M_1, M_2$, the masses of the two lightest neutralinos can be approximated by

$$m_{\widetilde{\chi}_{1,2}^0} = |\mu| + \frac{m_Z^2 (sgn(\mu) \mp sin2\beta)(\mu \pm M_1 cos^2\theta_W \pm M_2 sin^2\theta_W)}{2(\mu \pm M_1)(\mu \pm M_2)} \tag{1.41}$$

In the above equation the M_1 and M_2 mass parameters were taken to be real and positive by convention and the μ mass parameter was taken to be real with sign ± 1 [2].

1.7 Direct SUSY searches in collider experiments

SUSY may manifest itself in collider searches, in produced sparticles that decay to lighter SM particles and SUSY particles. In this thesis we focus on R-parity conserved scenarios, in which the sparticles are always produced in pairs and the final states of the decay contain the LSPs and some SM particles. The decay patterns and hence the signature of each process, depend on the sparticles that have been paired produced.

Based on the nature of the collider's beam, the SUSY searches conducted in different experiments vary. For example, the proton-proton colliders produce interactions at higher center-of-mass energies than those in the electron-positron colliders. In the proton-proton interactions the QCD-mediated processes cross section is large, resulting in higher sensitivity in squarks and gluinos. The latter are expected to have high mass and momentum and decay hadronically, resulting in final states with energetic jets and MET due to the LSPs. Depending on the multiplicity of the leptons in the event, the searches can be characterized as 0-lepton, 1-lepton, 2-leptons etc, targeting final states with the respective number of leptons, jets and MET. The most dominant SM background in searches with no leptons in the final states arises from QCD, while as the multiplicity of the leptons increases the QCD background gets suppressed and the SM backgrounds from W and Z bosons decays become more important. Figure 1.15 illustrates the LHC SUSY production cross section at Next-to-Leading-Order + Next-to-Leading-Logarithm (NLO+NLL) perturbation + resummation accuracy, at a center of mass energy $\sqrt{s} = 13\,\text{TeV}$.

The dominant production mechanism of SUSY particles is expected to be strong production of first and second generation squarks and gluinos, through the productions $pp \to \widetilde{g}\widetilde{g},\ \widetilde{q}\widetilde{q},\ \widetilde{g}\widetilde{q}$. The 2 generations of squarks are considered mass degenerate. The cross section for third generation squarks is smaller by almost 2 orders of magnitude and the electroweakino production cross section is more than 2 orders of magnitude lower than the colored SUSY particles. Slepton pair production processes have the smallest predicted production cross section at the $\sqrt{s} = 13\,\text{TeV}$ LHC.

SUSY at the electroweak scale is motivated as a solution to the hierarchy problem. In this context, the most relevant terms for SUSY phenomenology arise from the interplay between the masses of the third generation squarks and the Yukawa coupling of the top quark to the Higgs boson. Significant mixing is expected between $\widetilde{t}_L$ and $\widetilde{t}_R$ and the mass eigenstate $\widetilde{t}_1$ is expected to be the lightest squark.

Due to the lack of experimental evidence for SUSY particles, upper limits are set to restrict the SUSY parameter space. Currently the strongest limits on the SUSY parameter space are coming from CMS and ATLAS, the two general purpose detectors of LHC. However, numerous direct SUSY searches have been conducted in colliders like LEP and Tevatron, prior to the LHC era.

The electron-positron LEP collider at CERN, operated from 1989 to 2000 and its experiments have conducted searches for new physics, exploiting the advantage of clean experiment environment due to the nature of the beams. The experiments at LEP collected integrated luminosity of ~ 230 pb^{-1} and operated at maximum center of mass energy of $\sqrt{s} = 209$ GeV.

SUSY searches have also been conducted extensively by the CDF and DØ experiments of the proton-antiproton Tevatron collider at FermiLab. The Tevatron collider operated from 1992 to 1996 at a centre-of-mass energy of $\sqrt{s} = 1.8$ TeV and from 2001 to 2011 at a centre-of-mass energy of $\sqrt{s} = 1.96$ TeV. The data collected by the two experiments correspond to integrated luminosities of around ~ 10 fb^{-1} for CDF and ~ 11 fb^{-1} for DØ.

The CMS and ATLAS experiments at the LHC started their operation in 2010, simultaneously with the proton-proton beam operation of the collider. Detailed description of the LHC Runs and integrated luminosities corresponding to the collected datasets is presented in Sec. 2.1.2.

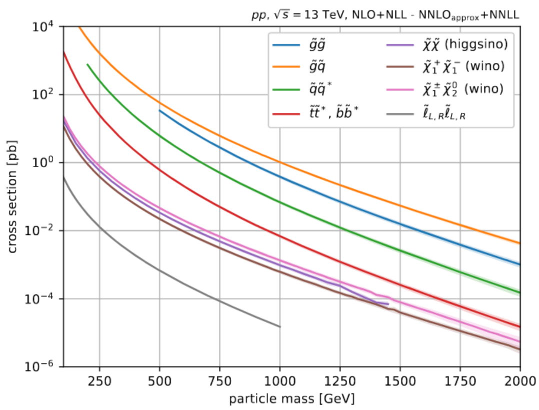

Figure 1.15: The predicted production cross section for different SUSY particles as a function of their mass.

1.7.1 SUSY exclusion limits from collider experiments

This subsection presents an overview of the exclusion limits on the SUSY parameter space, which are set for R-parity conserving models by the collider experiments. The first part of the subsection focuses on the masses of gluinos and squarks. The charginos and neutralinos mass limits are described next, and lastly the limits on the slepton masses are highlighted. Throughout the section limits set by older experiments at LEP and Tevatron are presented alongside the latest limits from the CMS Collaboration. Limits reported by the ATLAS Collaboration are compatible with those of CMS. Therefore, they are not described in detail but instead, references to the corresponding ATLAS publications have been added.

Gluinos and squarks mass limits

Assuming R-parity conservation, the decay modes and the typical signatures that searches are focused on depend on the assumed mass hierarchy between the squark and the gluino mass. Table 1.7.1 presents the typical final state signatures and the decay modes of first and second generation squarks and gluinos for different mass hierarchy scenarios. Symbol X in the typical signatures column denotes the additional initial or final state radiation jets or cascades, and MET. Assuming gluinos to be heavier than squarks, pair produced squarks will predominantly decay to a quark and a neutralino. This will typically be searched for in signatures involving two jets and MET, with potential extra jets stemming from initial state or final state radiation. If the gluino mass is lower than the squark mass, the pair-produced gluino decays to a pair of quark-antiquark and a neutralino that will lead to 4 jets and MET together with extra jets in the final state. When the gluino and the squark masses are similar, associated production of squark and gluino is possible. The squark will decay to quark and neutralino while the gluino will decay to a pair of quark-antiquark and neutralino, resulting in a final state of 3 or more jets and MET.

Mass Hierarchy	Production	Decay	Typical signature
$m_{\widetilde{q}} \ll m_{\widetilde{g}}$	$\widetilde{q}\,\widetilde{q}$, $\widetilde{q}\,\widetilde{\bar{q}}$	$\widetilde{q} \to q\widetilde{\chi}^0$	≥ 2 jets + MET + X
$m_{\widetilde{q}} \approx m_{\widetilde{g}}$	$\widetilde{q}\,\widetilde{g}$, $\widetilde{\bar{q}}\widetilde{g}$	$\widetilde{q} \to q\widetilde{\chi}^0$ $\widetilde{g} \to q\bar{q}\widetilde{\chi}^0$	≥ 3 jets + MET + X
$m_{\widetilde{q}} \gg m_{\widetilde{g}}$	$\widetilde{g}\,\widetilde{g}$	$\widetilde{g} \to q\bar{q}\widetilde{\chi}^0$	≥ 4 jets + MET + X

Table 1.5: Decay modes and typical signatures of first and second generation squarks and gluinos based on their mass hierarchy. Table taken from [8].

For the gluino masses, Tevatron has set lower limits at $\sim 310\,\mathrm{GeV}$ assuming the mSUGRA model for all squark masses, or $\sim 390\,\mathrm{GeV}$ if $m_{\widetilde{q}} \approx m_{\widetilde{g}}$ [40, 41]. These limits have been further constrained by LHC experiments in the framework of simplified models for the interpretations.

The gluino mass limits are set by studying three different decay modes of the gluino-gluino pair production. The first mode is the $\widetilde{g} \to q\bar{q}\widetilde{\chi}^0_1$ where the gluino decays to first or second generation quarks resulting in light jets, and neutralinos in the final state. The limits are calculated on the $m_{\widetilde{g}}$-$m_{\widetilde{\chi}_1^0}$ plane and the lowest value is set at $m_{\widetilde{g}} \geq 2\,\text{TeV}$ for massless neutralino, while no limit can be set for $m_{\widetilde{\chi}_1^0} > 1.2$ TeV. The second decay mode considered is the $\widetilde{g} \to b\bar{b}\widetilde{\chi}_1^0$ leading to 4 b-jets and 2 neutralinos in the final state. From the study of this decay mode, gluino masses below 2.3 TeV are excluded when the neutralino is considered massless, while no limit can be set for $m_{\widetilde{\chi}_1^0} > 1.5\,\text{TeV}$. Finally, the decay mode of $\widetilde{g} \to t\bar{t}\widetilde{\chi}_1^0$ that can lead in different final state signatures is studied. For the latter decay mode the estimated lowest limit is $m_{\widetilde{g}} \geq 2.25\,\text{TeV}$ for masseless neutralino, while no limit can be set for $m_{\widetilde{\chi}_1^0} > 1.3$ TeV. The limits on the gluino mass are presented in [42] for the CMS Collaboration and in [43] for the ATLAS Collaboration.

The Fig. 1.16 presents the CMS mass limits at 95% CL plots obtained with the Run–2 dataset that corresponds to luminosity of 137 fb^{-1}. The mass limit for simplified model of gluino pair production and decay to pairs of light, top and bottom quarks and the LSP are presented in the top left, top right and bottom plot respectively. All the summary plots show for comparison the respective limit obtained with integrated luminosity of 36 fb^{-1} in light gray. The solid lines show the observed and the dashed the expected limits. The curves with different colors (red, blue green, teal) correspond to different analyses. The paper number of every analysis is mentioned in the corresponding label.

Regarding the squark mass limits, Tevatron has set lowest limits for the first and second generation at $m_{\widetilde{q}} \geq 380\,\text{GeV}$ for all gluino masses or $m_{\widetilde{q}} \geq 390\,\text{GeV}$ if $m_{\widetilde{g}} \approx m_{\widetilde{q}}$, assuming the mSUGRA model for the interpretation [40, 41].

The CMS Collaboration has set limits on the squark mass assuming simplified models, and the pair produced squark to decay as $\widetilde{q} \to q\widetilde{\chi}_1^0$. Assuming mass degeneracy for the first and second squark generation, the estimated limit is $m_{\widetilde{q}} \geq 1.75\,\text{TeV}$. In the case of no mass degeneracy between the generations, the limit for the production of a single squark is found $m_{\widetilde{q}} \geq 1.3\,\text{TeV}$ and no limit can be set for $m_{\widetilde{\chi}_1^0} > 600$ GeV [45]. The respective ATLAS results are presented in references [43].

Limits on the third generation squark mass have been set both by the Tevatron and the LHC experiments. The pair produced bottom squarks can decay through $\widetilde{b} \to b\widetilde{\chi}_1^0$ giving rise to b-jets and MET in the final state. The Tevatron limit is $m_{\widetilde{b}} \geq 247\,\text{GeV}$ for massless neutralino [46, 47]. The CMS upper limit is $m_{\widetilde{b}} \geq 1.25\,\text{TeV}$ for massless neutralino while no limit can be placed on direct bottom squark production for neutralino masses of $m_{\widetilde{\chi}_1^0} > 700\,\text{GeV}$ [45]. The respective ATLAS results are presented in reference [48].

Top squarks decay predominantly via 2-body decays $\widetilde{t} \to t\widetilde{\chi}_1^0$ and $\widetilde{t} \to b\widetilde{\chi}^{\pm}$. If these decay modes are not kinematically allowed, top squarks can decay through a 2 body decay to $\widetilde{t} \to c\widetilde{\chi}_1^0$ or $\widetilde{t} \to bf\bar{f}'\widetilde{\chi}_1^0$. The latter is a 4 body decay with an off-shell exchanged particle, if $m_{\widetilde{t}}$-$m_{\widetilde{\chi}^0} > m_b$, or a 3 body decay ($\widetilde{t} \to bW\widetilde{\chi}_1^0$ or

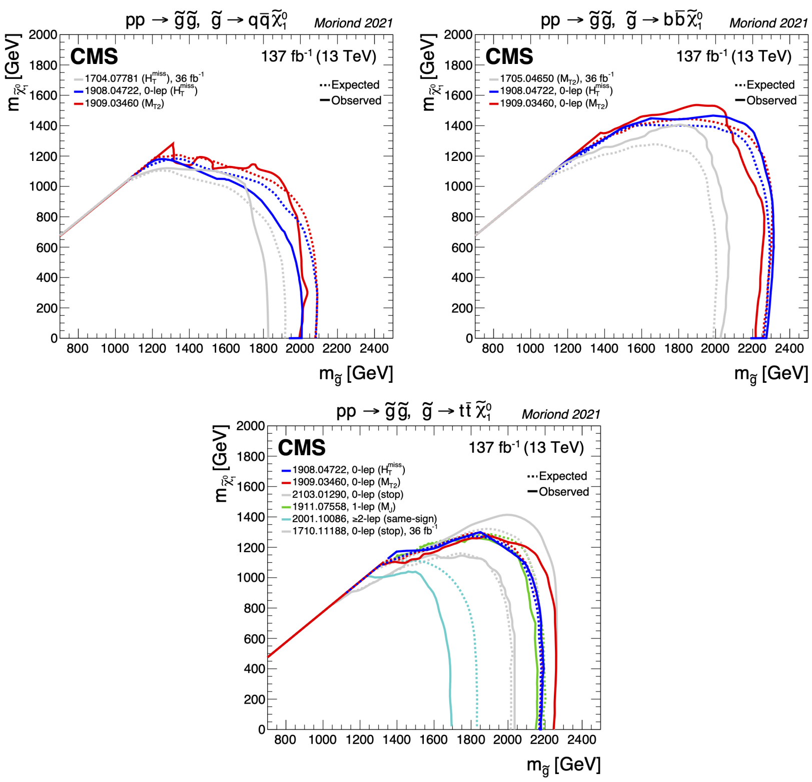

Figure 1.16: The CMS summary mass limits plots at 95% CL [44]. The solid (dashed) lines correspond to the observed (median expected) limits. The limits are set for simplified models of gluino pair production and decay to pairs of light flavour quarks and the LSP (top left), simplified model of gluino pair production and decay to pairs of bottom quarks and LSP (top right) and simplified model of gluino pair production and decay to pairs of top quarks and LSP (bottom).

$\tilde{t} \rightarrow b\ell\tilde{\ell})$, if it is kinematically allowed for the exchanged particle to be on-shell. The LEP Collaboration has set limits on the top squark mass at $m_{\tilde{t}} \geq 96\,\text{GeV}$ in the $\tilde{t} \rightarrow c\tilde{\chi}^0_1$ decay mode [49].

The limit set by the Tevatron on the top squark mass is $m_{\tilde{t}} \geq 235\,\text{GeV}$ for $m_{\tilde{\nu}} < 50\,\text{GeV}$, considering the decay chain of $\tilde{t} \rightarrow b\ell\tilde{\nu}$ and $m_{\tilde{t}} \geq 180$ GeV for $m_{\tilde{\chi}^0} < 95$ GeV, considering the decay mode $\tilde{t} \rightarrow c\tilde{\chi}^0$ [46, 50].

The LHC experiments have set limits on the top squark mass using simplified models and considering the different decay modes. For the case of $\tilde{t} \rightarrow t\tilde{\chi}^0_1$ the CMS limit is $m_{\tilde{t}} \geq 1.2\,\text{TeV}$ for massless neutralino and no limit can be set for $m_{\tilde{\chi}^0_1} > 600\,\text{GeV}$ [51]. For the decay mode of $\tilde{t} \rightarrow b\tilde{\chi}^{\pm} \rightarrow bW^{\pm}\tilde{\chi}^0_1$, the CMS limit is $m_{\tilde{t}} \geq 1.18\,\text{TeV}$ for massless neutralino and no limit can be set for $m_{\tilde{\chi}^0_1} \geq 550\,\text{GeV}$. In addition the CMS limit for the decay modes $\tilde{t}\tilde{t} \rightarrow tbW\tilde{\chi}^0\tilde{\chi}^0$ is $m_{\tilde{t}} \geq 1\,\text{TeV}$ and can be set for $m_{\tilde{\chi}^0_1} > 500\,\text{GeV}$. These results are presented in ref. [51] from the

CMS Collaboration, and respectively in ref. [52] from ATLAS.

The plots in Figure 1.17 show the 95% CL CMS mass limits for the simplified model of bottom squark pair production and decay to bottom quark and LSP on the left and simplified models of top squark pair production and decay to top quark and LSP on the right.

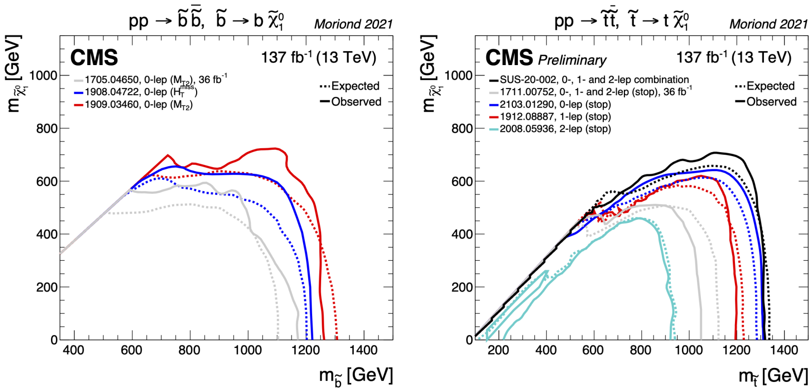

Figure 1.17: The CMS summary mass limits plots at 95% CL [44]. The solid (dashed) lines correspond to the observed (median expected) limits. The limits are set for simplified models of bottom squark pair production and decay to top quarks and the LSP (left plot) and simplified model of bottom squark pair production and decay to top quarks and LSP (right plots).

Chargino and neutralino mass limits

As summarized in Sec. 1.3, charginos are mixed states of charged winos and higgsinos and they can decay via 2-body decays, $\widetilde{\chi}^{\pm} \rightarrow \widetilde{f}\bar{f}'$, or 3-body decays $\widetilde{\chi}^{\pm} \rightarrow f\bar{f}\widetilde{\chi}^{0}$ through a virtual W boson. Neutralinos are mixed states of neutral wino, higgsino and bino.

The LEP experiments have set limits on chargino mass at $m_{\widetilde{\chi}^{\pm}} \geq 103.5\,\text{GeV}$ [53], considering fully hadronic, semi-leptonic and fully leptonic decay modes.

The LHC production cross section of pair-produced electoweakinos is more than 2 orders of magnitude smaller than that of the colored SUSY particles, as depicted in Fig 1.15. The LHC experiments study the chargino pair production and decay to lepton, neutrino and neutralino, mediated by slepton as $\widetilde{\chi}_1^{\pm}\widetilde{\chi}_1^{\pm}(\rightarrow \nu\ell\widetilde{\ell}\widetilde{\nu}) \rightarrow \ell\ell\nu\nu\widetilde{\chi}_1^0\widetilde{\chi}_1^0$. The CMS Collaboration has set a limit on the chargino mass assuming simplified models at $m_{\widetilde{\chi}_1^{\pm}} \geq 800\,\text{GeV}$ for massless $\widetilde{\chi}^0$, and no limit can be set for $m_{\widetilde{\chi}_1^0} > 350\,\text{GeV}$ [54]. The ATLAS result is presented in reference [55].

The case of chargino-neutralino production and decay to charge multileptons and MET in the final state, have been studied by CMS, and the results have been interpreted in the context of simplified models. The first scenario is the decay through sleptons $\widetilde{\chi}_1^{\pm}\widetilde{\chi}_2^0 \to \ell\ell\ell\nu\widetilde{\chi}_1^0\widetilde{\chi}_1^0$ or through sneutrinos $\widetilde{\chi}_1^{\pm}\widetilde{\chi}_2^0 \to \ell\nu\nu\nu\widetilde{\chi}_1^0\widetilde{\chi}_1^0$. These searches assume that $m_{\widetilde{\chi}_1^{\pm}} = m_{\widetilde{\chi}_2^0}$ and the mass of the slepton is between $m_{\widetilde{\chi}_1^{\pm}}$ and $m_{\widetilde{\chi}_1^0}$. The limit of $m_{\widetilde{\chi}_1^{\pm}}(m_{\widetilde{\chi}_2^0}) \geq 1.3$ TeV has been set for massless $\widetilde{\chi}_1^0$ and no limit can be set for $m_{\widetilde{\chi}_1^0} > 800\,\text{GeV}$. The CMS results are presented in ref. [56] and the ATLAS results in ref. [57].

The second decay scenario for the associated chargino-neutralino production is the 2-body decay to W, Z or H boson and $\widetilde{\chi}_1^0$. CMS has set limits for the decay mode of $\widetilde{\chi}_1^{\pm} \to W^{\pm}\widetilde{\chi}_1^0$ and $\widetilde{\chi}_2^0 \to Z\widetilde{\chi}_1^0$ or $\widetilde{\chi}_2^0 \to H\widetilde{\chi}_1^0$. For the WZ channel the limit is $m_{\widetilde{\chi}_1^{\pm}}(m_{\widetilde{\chi}_2^0}) \geq 750\,\text{GeV}$ and no limit can be set for $m_{\widetilde{\chi}_1^0} > 350\,\text{GeV}$ [58], while for the WH channel $m_{\widetilde{\chi}_1^{\pm}}(m_{\widetilde{\chi}_2^0}) \geq 500\,\text{GeV}$ for massless neutralino, and no limit can be set for $m_{\widetilde{\chi}_1^0} > 350\,\text{GeV}$ [59]. The summary limit plots from CMS, for both the WZ and the WH scenarios are presented in Figure 1.18. The ATLAS results are presented in ref. [60].

Higgsinos are a desirable target for LHC due to their contribution at tree-level to the Higgs boson mass matrix. The higgsino mass is controlled by the higgsino mass parameter μ as described in Section 1.6, which is expected to be near the weak scale due to naturalness. Direct higgsino searches suffer from low production rates and they are traditionally searched for in final states with 3 leptons (3ℓ) and MET that originate from the decay of more massive electroweakinos to W/Z bosons. The sensitivity of 3ℓ+MET searches is larger in regions where the $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0}$ is of the order of (m_W, m_Z) and decreases as the $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0}$ decreases because the intermediate gauge bosons goes off-shell and the emitted leptons become very soft to be detected. These difficult-to-target phase space corresponds to the electroweakino spectrum where natural SUSY is expected [61].

The main subject of this thesis is the SUSY search that targets the electroweakino productions in the compressed mass spectra. The strategy for targeting the very compressed spectra exploits final states with 2ℓ+MET and an initial state radiation that will give boost to the final state of the event. Additionally, final states with 3ℓ+MET are used to target higher $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0}$. The motivation for compressed mass SUSY searches is presented in the next section 1.8 and an overview of the compressed mass SUSY searches conducted at CMS is presented in Section 1.9. The analysis strategy is described in detail in the dedicated Chapter 4.

The left plot in Fig. 1.18 shows the summary exclusion limits obtained by CMS in the context of simplified models, from analyses targeting signatures of associated chargino-neutralino production and decay to W and Z bosons and LSP. The exclusion power of the curves decreases towards the limit of $m_{\widetilde{\chi}_1^0} = m_{\widetilde{\chi}_2^0}$. In the region of low $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0}$ the decay products of the electroweakino decays are very soft and the region posses difficulties in detection due to low event acceptance and high SM background contamination. The black line in the left plot of Figure 1.18 shows the exclusion power of the physics analysis of this thesis, with the full

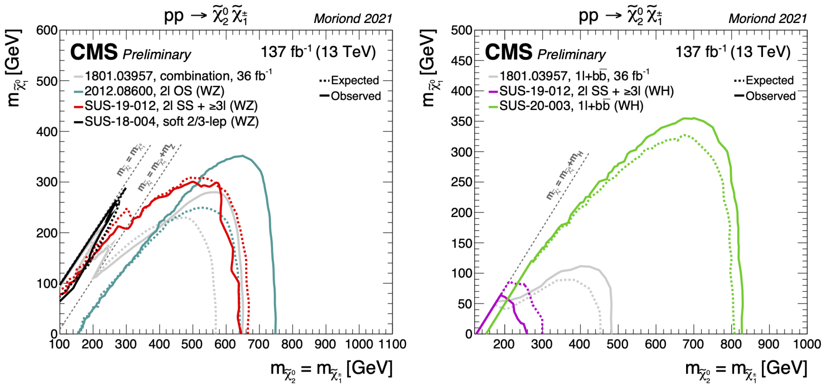

Figure 1.18: Mass limits at 95% CL obtained by CMS in the context of simplified models. The solid (dashed) lines correspond to the observed (median expected) limits [44]. On the left a summary plot of the CMS analyses targeting signatures of associated chargino-neutralino production and decay to W and Z bosons and LSP. On the right a summary plot of the analyses results that target chargino-neutralino production and decay via W and H bosons and LSP. In both cases the NLSPs are assumed to be degenerate in mass.

.

Run 2 data and its complementary with the rest of the electroweakino searches conducted by other CMS groups.

Charged sleptons mass limits

Sleptons may decay to their SM leptonic partner and the LSP. The LEP experiments have set limits on allowed masses for this decay mode. The limits on the slepton masses depend on the mass difference between the slepton and the LSP. The limit on the smuon was set at $m_{\widetilde{\mu}} \geq 94\,\mathrm{GeV}$ for $\Delta M_{(\widetilde{\mu}-\widetilde{\chi}^0_1)} > 10\,\mathrm{GeV}$, on the selectron at $m_{\widetilde{e}} \geq 100\,\mathrm{GeV}$ for $m_{\widetilde{\chi}^0_1} < 85\,\mathrm{GeV}$ and for the stau $m_{\widetilde{\tau}} \geq 93\,\mathrm{GeV}$ for $\Delta M_{(\widetilde{\tau}-\widetilde{\chi}^0_1)} > 7\,\mathrm{GeV}$. [62]

The LHC production cross section for slepton is almost two orders of magnitude lower than the electroweakino production cross section and thus a large amount of data is needed to surpass the LEP limits. The CMS and ATLAS Collaborations assume simplified models and the slepton decays to their SM partner and the LSP. The CMS Collaboration has set limits on the smuon mass at $m_{\widetilde{\mu}} > 320\,\mathrm{GeV}$ and no limit has been set for $m_{\widetilde{\chi}^0_1} > 150\,\mathrm{GeV}$. The lowest limit on the selectron mass is set at $m_{\widetilde{e}} \geq 350\,\mathrm{GeV}$ and no limit was set for $m_{\widetilde{\chi}^0_1} > 150\,\mathrm{GeV}$ [63]. The ATLAS results are presented in reference [55]. The stau masses up to 150 GeV have been excluded from CMS for massless LSP [64]. The ATLAS result on the stau mass limit is presented in reference [65].

1.8 Motivation for compressed mass SUSY searches

The R-parity conserving SUSY models predict that the lightest SUSY particle (LSP) is stable, electrically neutral, interacts only weakly with ordinary matter and it is massive (WIMP). It is therefore a very attractive DM candidate, as described in Section 1.5. The absence of observed SUSY WIMP at the collider searches can be interpreted as an indication that SUSY particles have very large masses. This motivates SUSY searches that target signatures with significant MET due to the LSP and jets and/or leptons with high p_T. The lack of evidence for SUSY particles led to strong constraints on the SUSY parameter space, as discussed in Subsection 1.7.1.

The lack of SUSY particle observation could alternatively mean that SUSY signal lays in regions of the parameter space that are difficult to be explored, such as the compressed mass spectrum, in which the mass difference of the pair-produced SUSY particle (NLSP) and the LSP is small ($\Delta m_{\text{NLSP-LSP}}$ <50 GeV). In this scenario most of the decay's energy and momentum is carried by the LSP while the visible particles in the final state have low momentum and are called soft. Events with such characteristics can be distinguished from bulk SM processes by requiring a jet with large p_{T} arising from initial state radiation (ISR) that leads to a large boost of the SUSY particle pair and thus large MET. Signatures in the compressed mass spectrum, with soft jets and MET in the final state suffer from the presence of very high SM backgrounds arising from QCD multijet events. Requiring soft leptons in the final state reduces significantly the QCD background and improves the sensitivity to new physics.

Natural SUSY suggests that the mass differences between the superpartners and their corresponding SM particles must not be too large. The superpartners are expected to have masses at the electroweak scales [34, 66]. This suggests that there should be a lightest stable electroweakino and at least one colored SUSY particle with mass of approximately below 1 TeV. It is usually assumed that the lightest colored superpartner is the top squark. Light higgsinos and potentially light top squarks can have compressed mass spectra and can lead to signatures with soft leptons and moderate to high MET, providing a window to natural SUSY [61, 67]. The stringent limits on light higgsino mass prior to LHC era have been set by LEP [68, 69] which excluded $m_{\widetilde{\chi}_2^0} = m_{\widetilde{\chi}_1^\pm} < 103.5$ GeV for a mass splitting of $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0} = 3$ GeV.

Special attention must be paid in the design of the analyses that target the compressed SUSY mass spectrum. The reconstruction and selection of very soft leptons accompanied with moderate MET is difficult due to the lower trigger efficiency and the high non-prompt SM background. Additionally, SUSY signatures in the compressed mass spectrum may be very similar to background arising from SM processes like the case of the top "corridor" analysis that will be described in the next section. Therefore, dedicated triggers and analysis methods are required to cope with the specificities of the phase space.

The SUSY search of this thesis targets signatures of pair-produced charginos and neutralinos or collectively called electroweakinos, decaying to 2 or 3 leptons and moderate to high MET. Additionally, the search is sensitive to light top squark pair production and decay via 4-body decays. In the top squark signature the $\widetilde{t}$ and the LSP are assumed to be nearly mass degenerate. This is typical of the so-called “co-annihilation region”, which reproduces the observed DM relic density and allows for dark matter to be provided solely by the SUSY LSP [70].

1.9 Overview of compressed mass SUSY searches at CMS

There are a number of SUSY searches conducted at CMS targeting the compressed mass spectrum and complement the SUSY parameter space at small $\Delta m_{\text{NLSP-LSP}}$. Here only a brief overview is presented. The SUSY searches can be broadly divided to those with full hadronic and leptonic final states.

Starting from the hadronic final states, the pair-production and decay of third generation squarks to quarks and $\widetilde{\chi}_1^0$ is a search that focuses on $\widetilde{t} \to c\widetilde{\chi}_1^0$ [71]. This study considers a compressed mass scenario for the top squark decay that can arise when the mass splitting, $\Delta m_{\widetilde{t}}\text{-}m_{\widetilde{\chi}_1^0}$, is below the mass of the W boson. In these cases the decay process of $\widetilde{t} \to t\widetilde{\chi}_1^0, t \to bW$ is suppressed because the top quark and the W boson must be virtual. Therefore the $\widetilde{t} \to c\widetilde{\chi}_1^0$ and not the $\widetilde{t} \to t\widetilde{\chi}_1^0$ is considered. Both bottom and top squark pair productions were studied in the context of simplified models. The search was conducted with the 2016 dataset collected at $\sqrt{s} = 13$ TeV, that corresponds to integrated luminosity of 35.9 fb^{-1}. It targets signatures with 2-4 jets, large MET (>250 GeV) and no leptons in the final state. In the compressed regime of $\Delta m_{\widetilde{t}\text{-}\widetilde{\chi}_1^0} < m_W$ the search excludes $m_{\widetilde{t}} \leq 510$ GeV for mass differences of $\Delta m_{\widetilde{t}\text{-}\widetilde{\chi}_1^0} < 10$ GeV.

An additional SUSY search with hadronic final state is that looking for events with soft hadronically decaying τ leptons and large MET induced by a highly energetic initial state radiation jet [72]. This event signature is consistent with direct or indirect tau slepton production ($\widetilde{\tau}$), through $\widetilde{\chi}_1^{\pm}\widetilde{\chi}_1^{\pm}$ or $\widetilde{\chi}_1^{\pm}\widetilde{\chi}_2^0$. The decay chains are $pp \to \widetilde{\tau}\widetilde{\tau}j$ or $pp \to \widetilde{\chi}_1^{\pm}\widetilde{\chi}_1^{\pm}j \to \widetilde{\tau}\widetilde{\tau}\nu_\tau\nu_\tau j \to \tau\widetilde{\chi}_1^0\tau\widetilde{\chi}_1^0\nu_\tau\nu_\tau j$ and $pp \to \widetilde{\chi}_1^{\pm}\widetilde{\chi}_2^0 j \to \widetilde{\tau}\tau j$. The search focuses on mass differences of $\Delta m_{\widetilde{\tau}\text{-}\widetilde{\chi}_1^0} < 50$ GeV. These scenarios are motivated by the DM coannihilation models[3] and can account for the observed relic DM density. The search was conducted with the data set collected during 2016 and 2017 and correspond to integrated luminosity of 77.2 fb^{-1}. The dominant SM prompt background arises from W+jets and Z+jet events

[3] In the coannihilation model the $\widetilde{\chi}_1^0$ interacts with another SUSY particle and this results in the production of SM particles

which contain genuine hadronic τ leptons, jet and real MET, and "fake" QCD background which arises from misidentified jets. The search sets 95% CL upper limits on the sum of $\widetilde{\chi}_1^{\pm}$, $\widetilde{\chi}_2^0$ and $\widetilde{\tau}$ production cross section for $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0} = 50$ GeV resulting in a lower limit of 290 GeV on the mass of the $\widetilde{\chi}_1^{\pm}$.

In the case of the leptonic final states, there are a number of compressed mass SUSY searches. A SUSY search that targets the top quark decay via 4-body or $\widetilde{\chi}_1^{\pm}$ mediated modes in single lepton final state was conducted with the 2016 data of total integrated luminosity of 35.9 fb^{-1} [73]. This search targets the compressed mass spectra in which the pair-produced $\widetilde{t}$ can decay either directly or through $\widetilde{\chi}_1^{\pm}$ into $bf\bar{f}'\widetilde{\chi}_1^0$. The final state of the search contains jets, MET and 1 lepton and the dominant SM background arises from W+jets and $\mathrm{t\bar{t}}$ processes. For the signal selection two analysis methods are implemented: a sequential selection based on discriminating variables and a multivariate technique. Upper limits on the top squark mass at 95% confidence level are set and reach up to 560 GeV, depending on the $\Delta m_{\widetilde{t}\text{-}\widetilde{\chi}_1^0}$ and the decay mode.

A top "corridor" analysis has been conducted and presented during Moriond 2021 [74]. This analysis targets the region of the parameter space in which the kinematics of the pair-produced top quark and top squark are very similar due to the mass difference of the top squark and the LSP being very close to the top mass ($m_{\widetilde{t}}$-$m_{\widetilde{\chi}_1^0} \approx m_t$). Therefore in this search, the 99% of the SM background arises from $\mathrm{t\bar{t}}$ events. The analysis is looking for signatures with 2 jets, large MET and 2 leptons of opposite sign, exploiting a simplified model for the $\widetilde{t} \to t\widetilde{\chi}_1^0$ decay. A parametric Deep Neural Network is used for the signal-background discrimination and the full top corridor phase space is excluded. The parametric Deep Neural Network is described in more detail in Sec. 6.1.2. In this analysis the top squark masses $m_{\widetilde{t}}$ of 145-295 GeV are excluded for neutralino masses from 0 to 100 GeV, with a mass differences of $\Delta m_{\widetilde{t}\text{-}\widetilde{\chi}_1^0}$ in the window of 30 GeV around the mass of the top quark.

Additionally, a SUSY search in vector boson fusion (VBF) topology with 0- or 1-lepton final state is presented in Ref. [75]. The analysis was conducted with the 2016 pp collisions data, at $\sqrt{s} = 13$ TeV and integrated luminosity of 35.9 fb^{-1}. This analysis is searching for produced $\widetilde{\chi}_2^0$-$\widetilde{\chi}_1^{\pm}$ or pair-produced $\widetilde{\chi}_1^{\pm}$ followed by $\widetilde{\chi}_1^{\pm} \to \ell^{\pm}\nu\widetilde{\chi}_1^0$, $\widetilde{\chi}_2^0 \to \ell^{\pm}\ell^{\mp}\widetilde{\chi}_1^0$ via light slepton or virtual SM boson. The search considers the scenario of "lightest-slepton" model where the NLSP is a $\widetilde{\ell}$ and the "WZ" model where sleptons are too heavy and the chargino and neutralino decays proceed via W^* and Z^* bosons. The high p_{T} oppositely-directed jets create recoil effect that facilitates the detection of MET in the event and the identification of the soft decay products in compressed-spectrum scenarios. The final states are characterised by one or no soft leptons, hard jets and MET from the neutrinos and the LSPs. The results are interpreted in the context of R-parity conserving MSSM model of pure electroweak VBF production of charginos and neutralinos. The models assume bino-like $\widetilde{\chi}_1^0$ and wino-like $\widetilde{\chi}_2^0$ and $\widetilde{\chi}_1^{\pm}$, with $m_{\widetilde{\chi}_2^0} = m_{\widetilde{\chi}_1^{\pm}}$.

Electroweakino masses, considering $m_{\widetilde{\chi}_1^{\pm}} = m_{\widetilde{\chi}_2^0}$, are excluded for a compressed mass spectrum of $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0}$ =1-30 GeV and assuming $\mathcal{B}(\widetilde{\chi}_2^0 \to \ell\widetilde{\ell} \to \ell\ell\widetilde{\chi}_1^0) = 1$ and $\mathcal{B}(\widetilde{\chi}_1^{\pm} \to \nu\widetilde{\ell} \to \nu\ell\widetilde{\chi}_1^0) = 1$. The $m_{\widetilde{\chi}_1^{\pm}} = m_{\widetilde{\chi}_2^0}$ masses are excluded up to 112 GeV for $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0} = 1$ in both models and up to 215 (175) GeV for $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0} = 30$ in the "lightest-sleptons" ("WZ") model, at 95% CL.

The SUSY search in events with soft leptons and MET in the final state, that will be presented in this chapter, is an update and extension of the 2016 analysis [76]. The search is targeting the pair-produced electroweakinos in the compressed mass spectra. Additionally, it has discovery potential for pair-produced light top squark, nearly mass degenerate with the LSP, that decay leptonically via 4-body decays. The signal models that were used for the interpretation of the results are the wino/bino, the higgsino and the light top squark pair production and will be described in detail in Section 4.2.

The strategy of the 2016 analysis was to select event with 2 soft leptons with opposite electric charge, moderate to high MET and one initial state radiation jet in the signal, that enhances the MET in the final state. The dominant prompt SM arises from $t\bar{t}$, W+jets, Z/$\gamma*$+jets processes. Non-prompt (also called "fake") leptons background from jet missidentification or heavy flavour decays is also an important SM background, due to the low p_{T} requirements on the leptons. The 2016 soft opposite sign 2ℓ analysis excluded at 95% confidence level wino-like $\widetilde{\chi}_1^{\pm}/\widetilde{\chi}_2^0$ masses up to 230 GeV for $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0} = 20$ GeV. In the context of the higgsino-like model, $\widetilde{\chi}_1^{\pm}/\widetilde{\chi}_2^0$ masses were excluded up to 168 GeV for the same mass difference. For the light top squark model, top squark masses up to 450 GeV were excluded for a mass difference of 40 GeV relative to the lightest neutralino. The lowest mass difference between the NLSP and the LSP for the wino/bino and higgsino scenarios was 5 GeV, at which $\widetilde{\chi}_1^{\pm}$ masses up to 165 GeV were excluded in the wino/bino model and up to 146 GeV in the higgsino simplified model.

1.10 Indirect and global SUSY searches

Constraints on the SUSY parameter space are obtained not only from direct collider searches, but also from indirect searches at low-energy experiments and from astrophysical observations. Some examples of indirect SUSY constraints that will be highlighted in this section are the rare B-meson decays measurement, the muon anomalous magnetic moment measurement and astrophysical measurement of the relic DM density. Further, global searches that combine constraints from direct and indirect SUSY searches will be briefly discussed at the end of the section.

The $B_s^0 \to \mu^+\mu^-$ and $B_d^0 \to \mu^+\mu^-$ decays are very rare due to the transition between quarks of different generations (CKM suppressed). The flavour changing

EWK neutral current interactions are forbidden, thus the rare B-meson decays are further suppressed due to higher order transitions. Figure 1.19 show the Feynman diagrams of the $B_s^0 \to \mu^+\mu^-$ decay with higher-order flavour changing neutral current processes allowed in the SM.

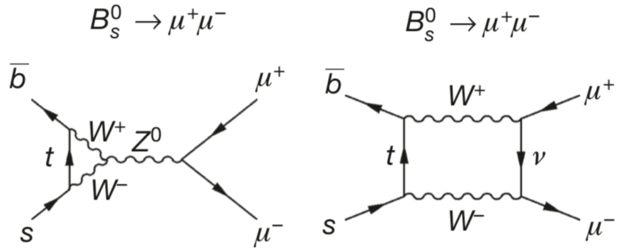

Figure 1.19: Feynman diagrams of the higher order flavor changing neutral current processes for the $B_s^0 \to \mu^+\mu^-$ decay allowed in the SM [77].

The SM prediction of the BF for those rare B-meson decays, accounting for higher-order electromagnetic and strong interaction effects, are calculated to be $BF(B_s^0 \to \mu^+\mu^-)_{SM} = (3.66 \pm 0.23) \times 10^{-9}$ and $BF(B^0 \to \mu^+\mu^-)_{SM} = (1.06 \pm 0.09) \times 10^{-10}$.

Theories beyond the standard model predict that such decays can be mediated by new SUSY particles. A BF measurement of the rare B-meson decays, that deviates significantly from its theoretical SM prediction, would give insight on how the SM should be extended. In 2015, the CMS and LHCb Collaborations combined data corresponding to $25fb^{-1}$ and 3 fb^{-1} of integrated luminosity, respectively, from Run 1 and reported the observation of the $B_s^0 \to \mu^+\mu^-$ decay and evidence for $B^0 \to \mu^+\mu^-$ decay [77]. The measured BF are $BF(B_s^0 \to \mu^+\mu^-) = (2.8^{+0.7}_{-0.6}) \times 10^{-9}$ and $BF(B^0 \to \mu^+\mu^-) = (3.9^{+1.6}_{-1.4}) \times 10^{-10}$, with both statistical and systematic uncertainties included. The statistical significances obtained are 6.2σ for the $B_s^0 \to \mu^+\mu^-$ and 3.2σ for the $B^0 \to \mu^+\mu^-$ decay mode, both compatible with the SM prediction, allowing for constraints to be set on the BSM theories. These results were updated in 2017 by LHCb for integrated luminosity of 4.4 fb^{-1} [78]. The Collaboration reported a statistical significance of 7.8σ for the $B_s^0 \to \mu^+\mu^-$ excess and an upper limit for the $BF(B^0 \to \mu^+\mu^-) < 3.4 \times 10^{-10}$, in agreement with the SM prediction.

Indirect constraints on the SUSY parameter space are also provided from the measurements of the muon anomalous magnetic moment. The magnetic moment of a Dirac muon is expressed in terms of the gyromagetic ratio and the intrinsic angular momentum $\mathbf{S}$ as $\boldsymbol{\mu} = g_\mu \frac{q_\mu}{2m_\mu}\boldsymbol{S}$ and the predicted value for the gyromagnetic ratio, is $g_\mu = 2$. Small deviations from $g_\mu = 2$ occur through quantum loop effects and are parametrised by the anomalous magnetic moment $\alpha_\mu \equiv \frac{g_\mu - 2}{2}$, which is accurately predicted within the SM and can be measured with high precision. A deviation between the SM predicted and the observed value could be a sign for New Physics.

The experiment E821 at the Brookhaven National Laboratory performed a measurement of $\alpha_\mu^{exp} = 116592091(63) \times 10^{-11}$ by studying μ^+ and μ^- in a constant external magnetic field while they circulate a storage ring [79]. More recently the Muon g-2 Collaboration at the Fermi National Accelerator laboratory measured the muon anomalous magnetic moment to be $\alpha_\mu^{exp} = 116592040(54) \times 10^{-11}$ [80]. The SM prediction of the muon anomalous magnetic moment is $\alpha_\mu^{SM} = 116591830(48) \times 10^{-11}$. The difference between experimental and theoretical value is $\Delta\alpha_\mu^{exp-SM} = 261(79) \times 10^{-11}$ and the error accounts both for the experimental value and the theory prediction. The experimental to theoretical prediction difference is about 3.3 times the 1σ combined error.

SUSY particle loops are good candidates to explain the $\Delta\alpha_{mu}^{(exp-SM)}$. The SUSY contributions to α_μ can span a wide range of possibilities, depending on the SUSY parameters. Generic models have been studied to illustrate how SUSY can contribute to α_μ [81] and it was found that for large values of $tan\beta$

$$\alpha_\mu^{SUSY} = 130 \times 10^{-11} \; (\frac{100GeV}{m_{SUSY}})^2 \; tan\beta \tag{1.42}$$

Accounting for the observed $\Delta\alpha_\mu^{(exp-SM)}$ the relation between m_{SUSY} and $tan\beta$ should be

$$m_{SUSY} = 71\sqrt{tan\beta} \tag{1.43}$$

and so for large values of $tan\beta$ ($\sim 3 - 40$) the SUSY masses should be in the range of 120-500 GeV. This SUSY mass range implies that the SUSY particles should be detected in the LHC experiments, however the lack of evidence of those particles introduces tension between the direct SUSY searches and the observations from low energy experiments. This will be discussed in more detail, later in this subsection.

The astrophysical measurements of the relic DM density from 2015 Planck data [82] evidence that $\Omega_{DM}h^2 = 0.11$. The R-parity conserving SUSY theories predict the existence of the lightest stable neutralino (LSP) which is a very promising DM particle candidate. The SUSY DM density that is to be estimated by the LHC experiments should be in accordance with the astrophysical observations. The Planck data constrain the theories of the SUSY particle interactions that are searched for in the LHC experiments.

Apart from direct searches in the collider experiment and indirect constraints from low energy experiments, global SUSY searches provide an additional interpretation of the allowed SUSY parameter space. These global fits combine experimental constraints from direct and indirect SUSY searches at the LHC, from direct and indirect DM searches, from astrophysical DM density calculation, from electroweak precision and flavour physics observables and from Higgs experimental observables.

Before the LHC data the global fits were mostly driven by the indirect SUSY constraints, while the CMS and ATLAS results increased the importance of direct SUSY constraints on the global searches. There have been multiple global SUSY studies adopting different SUSY constraint models. For example, global fits of constrained minimal supersymmetric models (CMSSM) analyze the impact of the ATLAS and CMS Run 1 data, incorporating constraints by other experiments such as precision electroweak and flavor measurements, relic DM density and direct DM searches results. The analysis pushes the masses of the first and second squark generation and the gluino masses beyond 2 TeV while the best fit value for the stop mass is found at ~ 1 TeV [83, 84]. Figure 1.20 presents the mass spectrum for the global best-fit CMSSM point from Ref. [83]. It demonstrates that the only light super-partners are the lightest stop, the lightest two neutralinos and the lightest chargino, which is almost degenerate in mass with the $\widetilde{\chi}_2^0$.

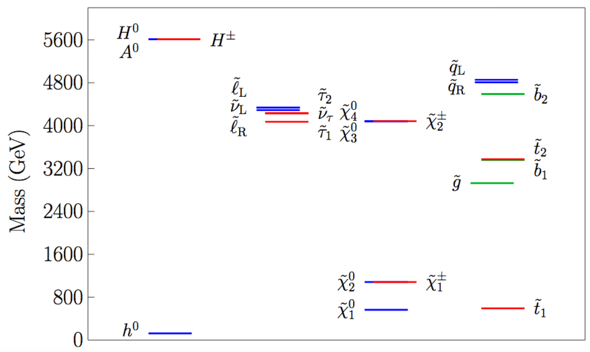

Figure 1.20: Sparticle mass spectrum of the CMSSM best-fit point from [83]

Additional global fits are performed on less constrained SUSY models like the pMSSM [85, 86, 87]. Figure 1.21 presents the pMSSM11 best-fit mass spectra and the decay paths with BF $> 5\%$ from Ref. [87]. The particle spectra for the best-fit points for the pMSSM11 concluded that the first and second slepton generations are heavier than the 3rd squark generation, which may lie within the reach of future LHC runs. Additionally, they find light and almost degenerate $\widetilde{\chi}_1^0$, $\widetilde{\chi}_2^0$ and $\widetilde{\chi}_1^\pm$.

The global fits are mostly impacted from the Higgs, electroweak and heavy flavour related constraints and from the muon anomalous magnetic moment constraint.

In summary, SUSY is postulated to be a softly broken symmetry of Nature whose breaking scale cannot deviate much from the electroweak scale, if it is to solve the hierarchy problem. The CMS and ATLAS Collaborations have yielded lower limits on the SUSY particle masses and no signal has been found in low energy indirect SUSY searches.

The first and second generation squark masses are excluded up to ~ 1.3 TeV,

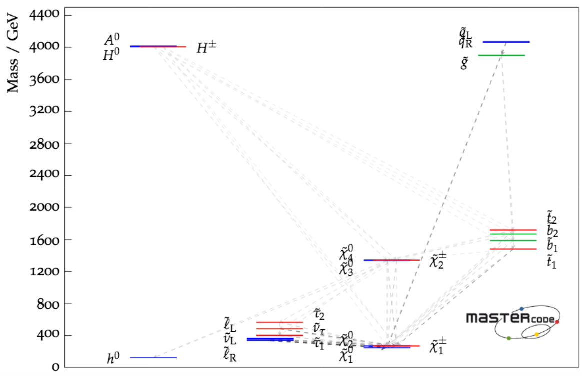

Figure 1.21: Sparticle spectra for the best-fit points for the pMSSM11 and decay paths with BF $> 5\%$ from Ref. [87].

the chargino (neutralino) masses are excluded up to $\sim 600 - 700\,\mathrm{GeV}$ for LSP masses $< 300\,\mathrm{GeV}$ and slepton masses up to $320 - 350\,\mathrm{GeV}$ for LSP of $< 150\,\mathrm{GeV}$. The gluino masses are excluded up to $\sim 2 - 2.3\,\mathrm{TeV}$ for LSP masses $< 1.2 - 1.5\,\mathrm{TeV}$.

The prospects from future experiments are that at least one of the gluino or stop particles is expected to be detectable at High Luminosity LHC (HL-LHC). Light higgsinos are expected to be accessible through compressed searches at the HL-LHC. If gluinos and electroweakinos cannot be accessible at the LHC in the very compressed regime, new facilities such as e^+e^- linear colliders would suffice to detect natural higgsinos.

The current SUSY exclusion limits are set in the space of SUSY particle masses only, assuming gauge couplings inherited from the SM. To extend the SUSY parameter space searches and include scans of the couplings down to very low values, new signatures should be directly looked for in the future, such as decays involving long-lived SUSY particles. Such signatures are already being searched at the LHC but with limited efficiency. A full exploration of the new parameter space requires advanced measurement techniques, putting new technological challenges in the LHC new physics search program.

Chapter 2

LHC and the CMS detector

The previous chapter discussed the very important role of experimental observations in the establishment of the particle physics. The experimental data guide and constrain theoretical physics models. In collider experiments, high density and large energy of the colliding particles benefit the production of massive particle. This instructs the need for accelerating machines that can reach very high energies, such as those reached at LHC, and collisions that result in very rich events.

In Section 2.1 the LHC layout, its performance during the Run 1 and Run 2 of data taking and the prospect and upgrades foreseen towards the HL-LHC are presented. The CMS detector layout, its subsystem technologies and the upgrades that each subsystem will undergo for the HL-LHC are highlighted in Section 2.2.

2.1 Large Hadron Collider

The LHC is the largest and most powerful particle accelerator ever built, operating at the European Organization for Nuclear Research (CERN) laboratory, near Geneva, Switzerland [88, 89, 90]. It is installed in a 27 km ring of superconducting magnets, 100 m beneath the surface of the earth, located inside the tunnel that previously hosted the LEP collider. The LHC was designed to collide two high energy proton beams with a maximum center of mass energy of 14 TeV, but it can also collide heavy ion beams in a wide range of atomic numbers with energies of up to 5 TeV per nucleon. The machine consists of two parallel rings hosting counter-rotating proton beams that are accelerated to nearly the speed of light and collide at the 4 interaction points. There are multiple magnets in different sizes and varieties, aiming at keeping the beams in circular orbit, focusing and squeezing them before collision. More specifically there are 1232 dipole magnets 15 m in length which bend the beams and 392 quadrupole magnets, each 5–7 m long, which focus the beams. Just prior to collision, magnets are used to "squeeze" the particles closer together to increase the chances of collisions

2.1.1 LHC Layout and Accelerator Systems

The LHC layout follows the LEP tunnel geometry. It has 8 arcs and 8 straight sections which can serve as experimental locations. Only 4 out of the 8 straight sections are currently occupied with detectors. The two general purpose detectors, ATLAS and CMS, are located at diametrically opposed straight sections at interaction points 1 and 5 respectively. The two smaller experiments, ALICE and LHCb are located at point 2 and point 8, together with LHC's beam injection points shown in green circles in Figure 2.1. The injection kick occurs in the vertical plane to the beams and the injection beams arriving at the LHC from below the reference plane. Points 3 and 7 contain two collimation systems each and point 4 contains RF systems that accelerate the beam particles. The beams are extracted from the machine at point 6 which contains the beam dump insertion.

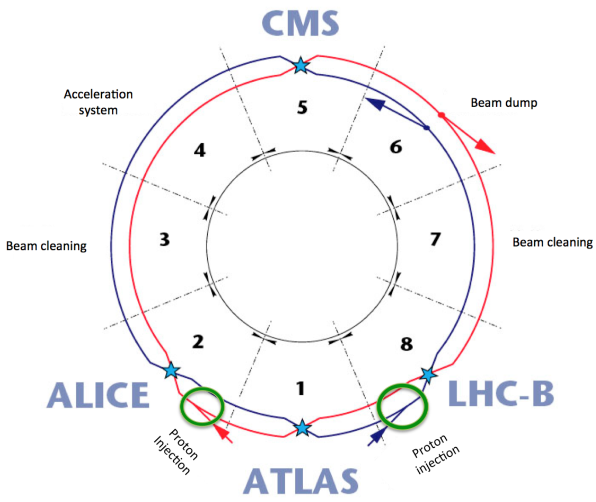

Figure 2.1: Layout of the LHC, showing separation into octants. The four experiments, the beam injection points, the acceleration system, the beam cleaning and the beam dump locations are shown. Taken from Ref. [91]

The accelerator complex of LHC consists of a collection of machines that progressively accelerate the proton beams at higher energies as shown in Fig. 2.2. The source of the proton beam is a bottle of hydrogen gas. An electric field is used to strip hydrogen atoms of their electrons and yield protons. The first accelerator in the chain is Linac2 (replaced by Linac4 in 2020) and accelerates the protons to energy of 50 MeV (160 MeV by Linac4). Then the proton beam is injected into the Proton Synchrotron Booster which accelerates protons to 1.4 GeV. The next in the accelerator chain is the Proton Synchrotron (PS) which pushes the proton beam up to 25 GeV and the Super Proton Synchrotron (SPS) where the beam is further accelerated up to 450 GeV. Then the protons are split and injected in the LHC beam pipes where they circulate for about 20 minutes before they reach their final energy. Protons are aggregated in bunches containing 1.15×10^{11} protons each. The bunches are separated in intervals of 25 ns.

The accelerator complex includes the Antiproton Decelerator and the Online Isotope Mass Separator (ISOLDE) facility, the Compact Linear Collider test area, as well as the neutron time-of-flight facility (nTOF). Linac 3 and LEIR are used for the lead ion production and acceleration.

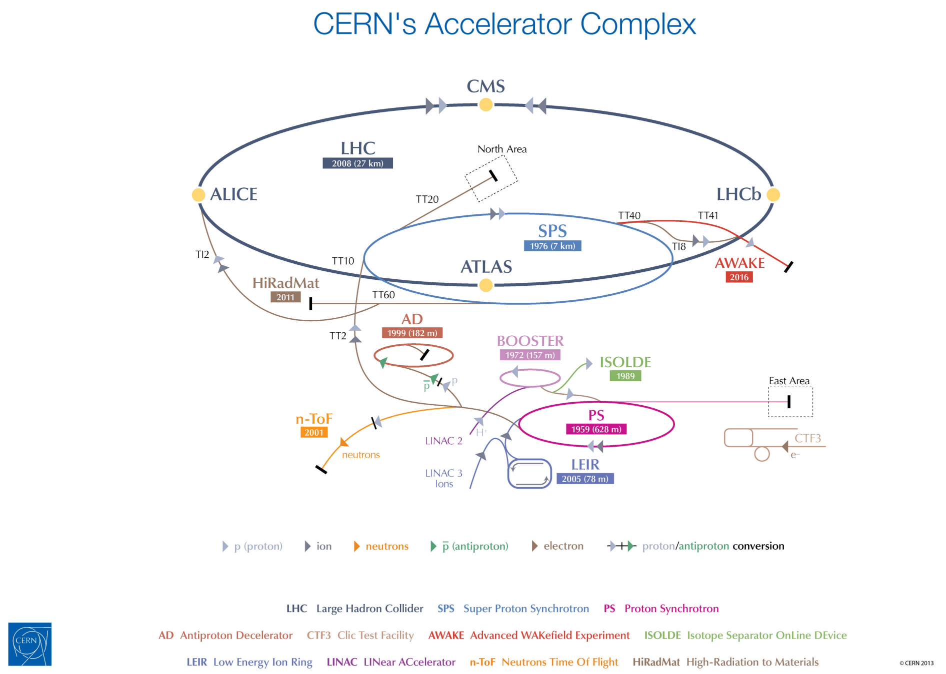

Figure 2.2: The CERN accelerator complex. Taken from Ref. [92].

2.1.2 LHC Luminosity and Performance

The instantaneous rate of interactions for a given pp interaction cross section is given by

$$\frac{dN}{dt} = \mathcal{L}\sigma \tag{2.1}$$

where $\mathcal{L}$ is the instantaneous luminosity. The machine's instantaneous luminosity depends only on the beam parameters and can be written as

$$\mathcal{L} = \frac{N_b^2 n_b f \gamma}{4\pi \epsilon_n \beta *} F \tag{2.2}$$

- N_b is the number of particles per bunch (nominal value 1.15 10^{11});
- n_b is the number of bunches per beam (nominal value 2808 for bunch crossings every 25ns);
- f is the revolution frequency (11.2 kHz);
- γ is the proton energy in units of the proton mass;
- β^* is the amplitude function. Quantifies how narrow the beam is. The lower the β^* the narrower the beam (nominal value 0.55 m);
- ϵ_n is the transverse normalised emittance of the beam. Low emittance beam is a beam where the particles are confined to a small distance and have nearly the same momentum. This results in higher likelihood for particles to interact and thus to higher luminosity. It has units of length and it is measured on the plane transverse to the beam;
- F is a correction factor for the crossing angle of the two beams (nominal value 0.85);

The above results in the nominal (design) instantaneous luminosity of the LHC $\mathcal{L} \sim 10^{34}$ s^{-1} cm^{-2} [88].

The integral of instantaneous luminosity over time is called integrated luminosity and it is a measure of the collected data size. Integrated luminosity is an important parameter that directly relates to the number of observed events and quantifies the performance of an accelerator. Fig. 2.3 shows the integrated luminosity delivered by the LHC and collected by the CMS experiment, during the Run 1 (2010-2012) and the Run 2 (2015-2018) data-taking for pp collisions at 13 TeV.

The Run 1 of data taking lasted from 2010 to 2012. For the first year of this run LHC operated at a center-of-mass energy of $\sqrt{s}$ = 7 TeV and CMS collected integrated luminosity of $\sim$ 5.59 fb^{-1}. For the second part of Run 1, the center-of-mass energy was increased by 1 TeV and the experiment collected integrated luminosity of $\sim$ 21.79 fb^{-1}. The LHC machine and the detectors were repaired and upgraded during a Long Shutdown (LS) 1 between 2013 and 2015. The data taking period from 2015 to 2018 is called Run 2. During that period LHC operated at 13 TeV and the integrated luminosity collected by CMS is $\sim$ 150.76 fb^{-1}. Currently the LHC machine undergoes its LS2 since 2018, for upgrades and maintenance while a Run 3 is designed for 2022-2024. After Run 3 a LS3 is programmed during which the machine will undergo major upgrades in order to prepare for HL-LHC.

An *event* at LHC is identified as a collision between two protons of the same bunch crossing that produces a hard scattering of interest. These collisions are accompanied by multiple soft pp collisions in the same or adjacent bunch crossings

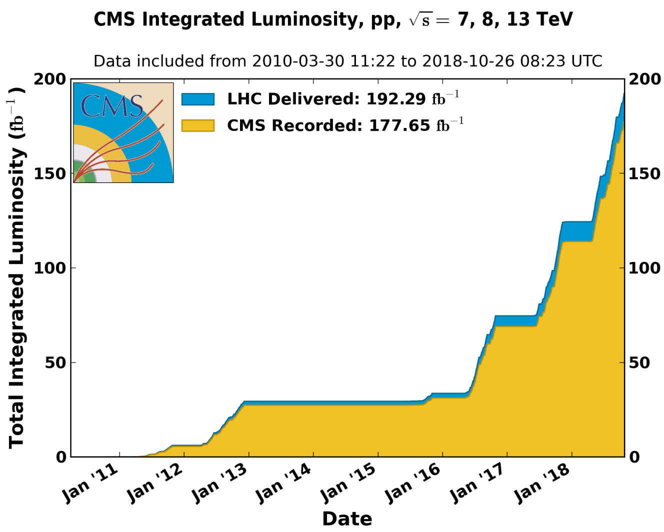

Figure 2.3: Cumulative delivered and recorded integrated luminosity versus time for 2010-2012 and 2015-2018 (pp data only). Taken from Ref. [93].

called Pile-Up (PU). The number of the additional interactions rises with the instantaneous luminosity.

The number of interactions per crossing for each bunch crossing is given by

$$\mu = \mathcal{L}_{bunch}\sigma_{tot} \tag{2.3}$$

where $\mathcal{L}_{bunch}$ is the instantaneous luminosity per bunch and σ_{tot} is the total pp cross-section which is 100 mb. The mean value of interactions per crossing (μ) correspond to the mean of Poisson distribution of the number of interactions per crossing for each bunch. Figure 2.4 shows the distribution of the average number of interactions per crossing for pp collisions in Run 2. The overall mean values and the pp inelastic cross section are also shown. The plot uses only data that passed the "golden" certification (i.e. all CMS sub-detectors were flagged to be on for any kind of usage in physics analysis).

2.1.3 High Luminosity LHC

The upgrade of the LHC and the full exploitation of its capabilities is of major importance for the discovery of new physics and for the in-depth study of observed phenomena, as described in Section 1.7. In order to increase its discovery potential, LHC will undergo a major upgrade in the 2020s to increase its collision rate. The goal of the HL-LHC program is to increase the instantaneous luminosity by a factor of 5 beyond the LHC design luminosity, reaching 5-10 10^{34}cm^{-2} s^{-1}. The number of events per bunch-crossing (PU) will increase up to about 200.

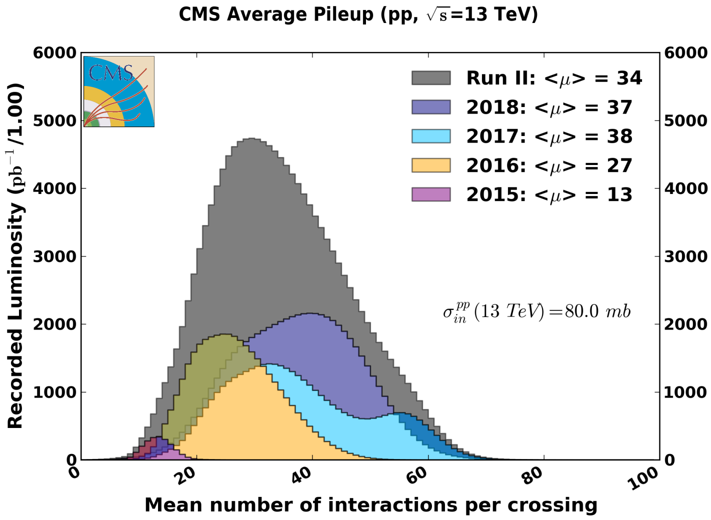

Figure 2.4: Interactions per crossing (PU) for 2015-2018 data taking.

With this performance the total integrated luminosity is expected to increase by a factor of 10, up to 3000 fb^{-1} after 10 years of operation starting from 2027-28 [94].

The HL-LHC developments require 10 years of studying, testing and optimizations. The goal is to have the main hardware of HL-LHC installed and the machine's configuration commissioned during the LS3. The upgraded configuration will rely on cutting edge technology such as 12 T superconducting magnets, very compact and ultra precise phase control superconducting beam cavities and new technologies for beam collimation. Concerning the beam upgrade, the $\beta*$ amplitude function will be reduced to 0.15 m for luminosity upgrade.

2.2 Compact Muon Solenoid Detector

The CMS detector is one of two largest general-purpose devices (together with ATLAS) at the LHC and it operates at interaction point 5. It has a length of 28.7 m, diameter of 15 m and it weights 14 ktonnes. CMS is designed to measure stable particles produced during the beam collisions, except from neutrinos that only interact weakly and escape detection. The unstable particles can be reconstructed due to the accurate energy and momentum measurements and the good spatial resolution of the detector. The main goal of the detector was the study of the electroweak breaking through the discovery and study of the Higgs boson and BSM searches at the TeV scales like natural supersymmetry. However, searches conducted at CMS are not restricted to those fields. A plethora of analyses are performed, exploiting the full range of pp and heavy-ion collision data.

The detector's main requirements for a fruitful physics program can be summarised as:

- Good muon identification, momentum resolution, angle and charge determination and good dimuon mass resolution;
- Good charged particle momentum resolution and reconstruction efficiency in the inner tracker;
- Good electromagnetic energy resolution and wide geometric coverage. Good diphoton and dielectron mass resolution;
- Good hadronic energy resolution and wide geometric coverage of the hadronic calorimeters;

As mentioned above, CMS will undergo major upgrades during its LS3 (Phase-2 upgrade) in order to cope with the HL-LHC conditions. The factors that should be taken into account for the CMS Phase-2 upgrade are the higher luminosity, the possible detector degradations and changes to the trigger, most notably the increases in L1 latency and rate.

2.2.1 CMS Layout

The detector's geometry is specifically designed for optimal particle detection and maximum reconstruction performance. It is cylindrical, with its axis along the beam line and the collision region centered at its geometrical center. Its central region is called barrel and the disks that cover its forward regions are called endcaps. The detector's materials are placed in cocentrical layers, each optimised for specific purpose. The interaction of the colliding particles with the detector materials result in electric signal. This signal is measured, digitised and finally analysed by computers.

In the innermost part of the detector a tracker device is installed. The tracker is a cylinder of 5.8 m long and diameter of 2.6 m. Its outer part consists of 10 layers of silicon strip detector, while its inner part that surrounds the interaction point, is made of 4 layers of silicon pixel detectors. Its purpose is the identification of charged particles and the measurement of their trajectories, momentum and charges. Outside the tracker the electromagnetic and hadronic calorimeters are located sequentially. The ECAL is designed to detect electrons and photons, while the HCAL is designed to detect jets of hadrons.

The precise momentum measurement of high-energy charged particles requires large bending power by a strong magnetic field. The main feature of the CMS detector is a superconducting solenoid magnet that generates a 3.8 T, nearly homogeneous magnetic field, parallel to the beam line. The large magnetic field provides strong bending on the muon tracks, before they enter the muon chambers. The solenoid is 13 m long, it has a diameter of 6 m and accommodates the tracker and the calorimeters within its volume.

Outside the solenoid, the muon system is integrated in the iron return yoke frame which confines the magnetic field outside the solenoid to allow for momentum measurement with the muon detectors, as well as to protect the detector electronics. The muon detection utilizes the gas ionization technology and its purpose is to identify and measure the momentum and signs of the muons.

Figure 2.5 presents the CMS layout.

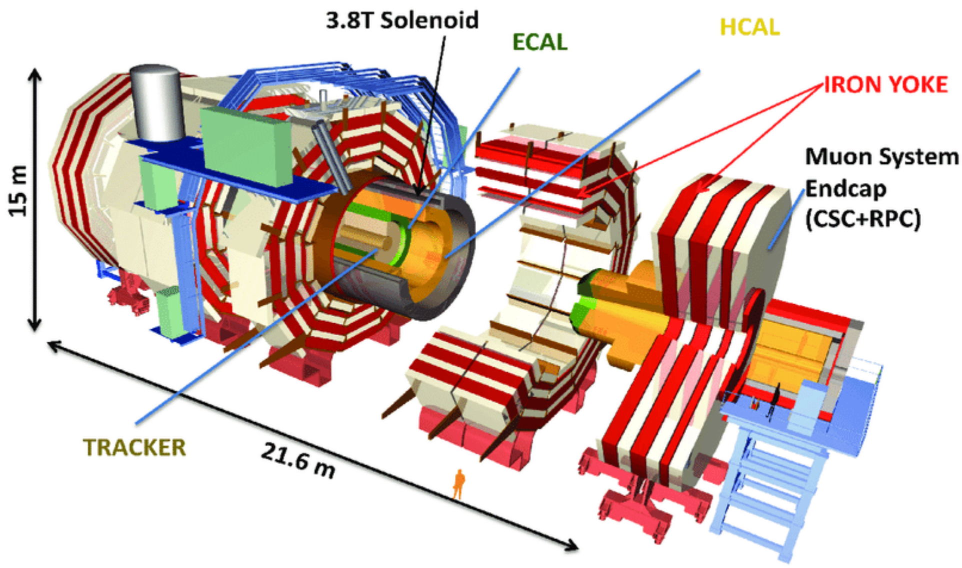

Figure 2.5: Schematic view of the CMS detector. Taken from Ref. [95]

The coordinate system adopted by CMS has the origin centered at the nominal collision point in the heart of the experiment. The y-axis is perpendicular to the beam line pointing upwards, x-axis is pointing radially inwards towards the center of the detector and z-axis is along the beam line pointing from point 5 to the point 4. The azimuthal angle ϕ is measured from x-axis in the x-y plane and the radial coordinate r is measured in the x-y plane. Pseudorapidity is a very important measure in collider physics experiments, it is defined as $\eta = -tan(\frac{\theta}{2})$ and it used to indicate the polar angle of the particle. $E_T = \sqrt{p_T^2 + m^2}$ and $p_T = \sqrt{p_x^2 + p_y^2}$ are the energy and momentum measured on the x-y plane transverse to the z-axis.

The detector has almost full solid angle coverage. The central barrel region covers up to $|\eta| < 1.48$ and the two endcaps cover up to $|\eta| < 3$. In the very forward regions of HCAL the coverage of pseudorapidity is extended to $\eta = 5$.

A detailed description of the CMS layout can be found in Ref. [96]. The following subsections discuss the above mentioned CMS detector components in more detail following loosely Ref. [96, 97].

2.2.2 Inner Tracking System

The innermost part of the CMS detector is the tracking system. Its purpose is to provide precise and efficient measurement of charged particle trajectories through the ionisation they produce along their path. Charged particles traversing the tracker induce electron-hole pairs, which create measurable currents that are digitized. The resulted “hits” are grouped into tracks using advanced pattern recognition algorithms and reconstruct a trajectory per particle. The origin or "vertex" and the direction of flight of the particle are also indicated by the reconstructed trajectory.

The tracker is required to have high granularity, fast response and high radiation and age resistance. Therefore, silicon technology was used for the construction of the tracker. The tracker was designed to operate without loss of efficiency up to an integrated luminosity of 500 fb^{-1} and an average PU of less than 50 collisions per bunch crossing. The solenoid magnet covers fully the tracker, whose total sensitivity area is 200 m^2 and its acceptance expands up to $|\eta| < 2.5$.

The part of the tracker that surrounds cylindrically the interaction point is the pixel detector. It consists of 3 detector layers in the barrel at radii 4.4, 7.3 and 10.2 cm and two endcap layers at ± 34.5 cm and ± 46.5 cm. There are 66^6 pixels of size 100×150 μm^2 covering a region of 1 m^2.

After the end of 2016 data taking period the pixel detector was replaced with an upgraded version called CMS Phase-1 pixel detector [98, 99]. It consists of 4 concentric barrel layer at radii 2.9, 6.8, 10.9, and 16.0 cm. Its endcaps consist of 3 disks of pixel modules located at ± 29.1 and ± 39.6 and ± 51.5 cm from the center of the detector. There are 1856 sensor modules, covering a volume of 1.9 m^2. Each module consists of a sensor with 160×416 pixels connected to 16 readout chips. The CMS Phase-1 pixel detector delivers 4 high precision space tracking points covering up to $|\eta| < 2.5$ range, improved pattern recognition and track reconstruction and also redundancy to mitigate hit losses.

The silicon strip detector is located outside the pixel detector at radius between 20-116 cm. The barrel part ($|\eta| < 1.5$) consists of 10 layers of silicon micro-strips and the endcaps ($1.5 < |\eta| < 2.5$) have 3 inner mini disks and 9 outer disks. The silicon tracker can be divided into:

1. The Tracker Inner Barrel/Disks (TIB/TID): 4 barrel layers of micro-strip sensors covering up to $|z| < 65$ cm and 3 disks at each end covering $65 < |z| < 120$ cm. Both TIB and TID cover $20 < r < 55$ cm. Each micro-strip sensor is 320 μm thick, configured parallel to the beam in the barrel and

radially in the endcap disks. TIB/TID provide up to 4 r-ϕ measurements per charged particle trajectory

2. The Tracker Outer Barrel (TOB): It surrounds the TIB/TID. Extends in radius from $55 < r < 116$ cm and within $|z| < 118$ cm. It consists of 6 barrel layers of thick micro-strip sensors 500 μm

3. The Tracker EndCaps (TEC): They cover the region $124 < |z| < 282$ cm and $22.5 < |r| < 116$ cm. Each TEC is composed of 9 disks with silicon micro-strip detectors, 320-500 μm thick, configured in the radial direction. They provide up to 9 ϕ measurements per trajectory

The configuration of the TIB/TID and TOB and TED is presented in Figure 2.6.

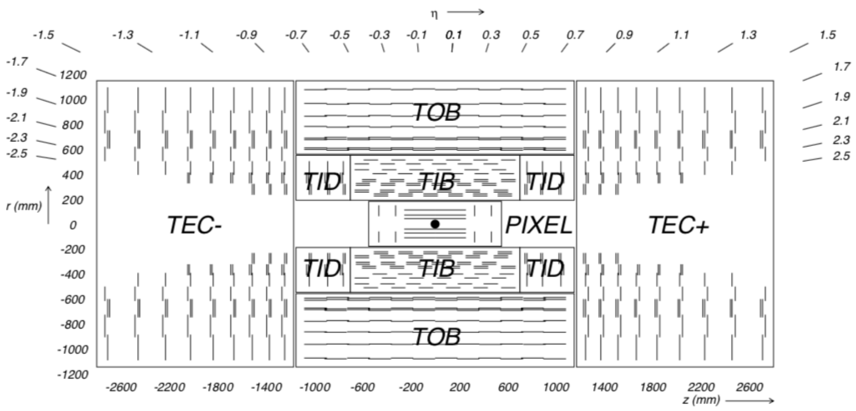

Figure 2.6: Schematic cross section through the CMS tracker. Each line represents a detector module. Taken from Ref. [97]

The tracker transverse momentum resolution for high momentum tracks (100 GeV) in the central region ($|\eta| < 1.6$) is 1-2% and degrades at higher η. For lower momentum tracks the transverse momentum resolution of the tracker is dominated by multiple scattering. The transverse and longitudinal impact parameter resolution is 10 μm for high momentum tracks while it reduces at lower momentum due to multiple scattering.

Tracker Upgrades for HL-LHC

The current pixel and silicon tracker will not be able to survive the harsh conditions of HL-LHC therefore, a new, upgraded, radiation hard, silicon tracker will be installed in the center of CMS. The layout of the Phase-2 CMS tracker will have the barrel, made out of cylindrical layers and the endcap parts, made out of discs. The innermost barrel region (IT) will be made out of 4 layers of pixel detectors and will provide 3-dimensional hit coordinates and excellent vertex identification. Six layers of the outer tracker (OT) detector will follow the IT barrel. The OT will consist of modules with two silicon sensors separated by few millimeters. The inner three layers of the OT barrel will comprise modules made of one pixelated sensor and one micro-strip sensor and the outer three layers will be instrumented with modules with two micro-strip sensors. The endcap discs will be also made of pixelated sensors and micro-strip sensors.

The pixel Phase 2 tracker will consist of very small pixels ($\sim$ factor of 6 smaller from the current configuration) allowing for robust resolution and radiation tolerance. Additionally the pixel tracker will be greatly extended in the forward region, enhancing the performance of the very forward physics signals [100]. The OT will feature increased radiation tolerance and higher granularity for efficient tracking and good track separation. It will be able to provide information at the Level-1 trigger allowing for good trigger rates without losses of signal [101].

2.2.3 Calorimeter detectors

The calorimeters are widely used in particle detectors for energy measurement of charged and neutral particles through total absorption in a block of matter. Particles like electrons, positrons, photons and hadrons enter the calorimeters, interact with the detector material and produce showers of particles that travel until they are fully absorbed. High granularity of the calorimeters is required for the accurate localization of energy deposits, which are combined with further sub-detector information for particle identification.

Electromagnetic (EM) showers occur when a high energy electron or photon enters the calorimeter. The photon will interact with the detector material via e^+e^- pair production while the electron and positron will interact with the detector material through bremsstrahlung radiation. An important variable of the EM showers is the radiation length X_0, it defines the average length an electron has to travel to reduce its initial energy by $1/e$ or the 7/9 of the mean free path of a photon. The shower's depth depends on the initial particle's energy as:

$$X = X_0 \frac{ln(\frac{E_0}{E_C})}{ln2} \tag{2.4}$$

where E_0 is the initial particle energy and E_C is the critical energy at which the ionization and bremsstrahlung rates are equal. If the energy of the shower is smaller than the critical energy, the shower stops. Another important variable is the Molière radius (R_M) which defines the cylinder in which the 90% of the EM shower's energy is contained and characterises the width of the shower. Both the radiation length and the Molière radius depend on the detector's material.

Hadronic showers arise from the interaction of hadrons originating from quarks and gluons produced in the hard scattering with the detector material. In the calorimeter volume, hadronic jets consist of long-lived mesons (pions and kaons) and stable baryons (protons and neutrons), occasionally including neutrinos and muons from heavy-flavour decays. The neutral pions decay promptly to two photons which lead to EM showers, therefore a hadronic jet typically has a small component in the EM calorimeter as well. The more energetic the initial particles the higher the shower's particle multiplicity. The size of the hadronic showers is large, therefore a large volume of high density hadronic calorimeters is required in order to completely contain the hadronic shower. Similarly to the radiation length of the EM shower, the hadronic shower is characterised by the interaction length (λ_I) which is the mean free path between nuclear collisions.

The CMS calorimeter is installed around the tracking system and its bigger part is enclosed in the solenoid. It consists of two parts, the Electromagnetic calorimeters (ECAL) and the Hadronic calorimeters (HCAL). The CMS calorimeter system is almost fully hermetic covering pseudorapidity up to $|\eta| < 5$ in order to provide a reliable measurement of the missing energy.

Electromagnetic Calorimeter

The ECAL is almost a hermetic homogeneous [1] detector layer that surrounds the tracker [96, 97, 102]. It absorbs and measures the energy of electrons, positrons and photons produced during collisions. The ECAL extends up to $|\eta| < 3$ and consists of 75,848 lead tungstate crystals. Lead tungstate was chosen for the detector's material because it has short radiation length ($X_0 = 0.89$ cm) and Molière radius ($R_M = 2.2$ cm) and high density (8.28 g/cm^3) and therefore assures fine granularity and a compact calorimeter. The quality and scintillation properties of the lead tungstate allow for high radiation resistance and fast response of the calorimeter.

As electrons and photons pass through the ECAL, their electromagnetic shower results in cascades giving rise to scintillation in the crystals. Avalanche photodiods (APDs) are used as photodetectors in the barrel and Vacuum phototriods are used in the endcaps (VPT). The lateral size of the crystals is $\sim 1R_M$ and hence 90% of the EM shower can be contained within a single crystal. The number of scintillation photons that will be emitted by the crystals and the amplification of the photodetectors depend on the temperature. A cooling system is used to keep the temperature of the system constant.

The ECAL barrel is at 129 $cm <$ r $<$ 175 cm, extends in $\eta < 1.479$ and consists of 61,200 crystals. The shape of the crystals is tapered with a front cross section of 2.2×2.2 cm^2 and 2.6×2.6 cm^2 at the rare face. Each crystal has a length of 23 cm and a radiation length of 25.8 X_0. The ECAL endcaps extend up to $1.479 < |\eta| < 3$ and contain 7,324 crystals each. The crystals have

[1] The entire volume is sensitive and contributes a signal

length of 22 cm, rare face cross section $30 \times 30\ mm^2$ and front face cross section $28.62 \times 28.62\ mm^2$. An ECAL Preshower (ES) detector is installed in front of the ECAL endcaps in pseudorapidity $1.653 < |\eta| < 2.6$. It is a sampling calorimeter [2] with two layers of silicon microstrip detectors followed by two lead absorber planes. It improves the spatial resolution of energy measurement and benefits the detection of $\pi^0 \to \gamma\gamma$ which can potentially be confused with a single photon. The geometry of the CMS ECAL is shown in Fig. 2.7.

The energy resolution of the ECAL has been measured in test beams to be [97]

$$(\frac{\sigma}{E})^2 = (\frac{2.8\%}{\sqrt{E[GeV]}})^2 + (\frac{12.0\%}{E[GeV]})^2 + (0.3\%)^2 \tag{2.5}$$

The first term in Eq. 2.5 is the stochastic term and parametrises the intrinsic energy fluctuations of the shower. The second term is the noise term and accounts for electronic and digitization noise or energy fluctuations from external to the shower sources. The last term is the constant term and accounts for calibration errors or leakage of the EM shower.

Hadronic Calorimeter

The HCAL is a sampling calorimeter that consists of alternating layers of brass or stainless steel as absorber and plastic scintillator or quartz fiber tiles as sensitive material that measures the energy deposit. Its main purpose is the absorption and energy measurement of strongly interacting particles. These particles create hadronic showers in the brass layer and induce detectable light in the scintillator which is guided by embedded wavelength-shifting to readouts.

HCAL is divided into 4 regions: the Hadronic Barrel (HB) and the Hadronic Endcaps (HE) that surround the ECAL and are located inside the superconducting solenoid. The Hadronic Outer calorimeter (HO) is placed just outside the solenoid, complements the HB and acts as an energy tail catcher. The Hadronic Forward calorimeter (HF) is located 11.2 m away from the interaction point and covers high pseudorapidity ranges $3.0 < |\eta| = 5.0$. The geometry of the HCAL is shown in Fig. 2.7.

The HB has a length of 9 m with inner diameter 6 m and it extends up to $|\eta| < 1.4$. It consists of 36 azimuthal wedges, each containing 14 layers of brass absorber plates, aligned parallel to the beam and two stainless steel layers in its inner and outer part for structural strength. There are 17 scintillator layers, interspersed between the stainless steel and brass material. HB is segmented in $2,304$ towers with size $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$. The total absorber thickness at $\eta = 0$ is 5.82 λ_I and increase up to 10.6 λ_I at $|\eta| = 1.3$. The HE covers

[2] The material that produces the particle shower is distinct from the material that measures the deposited energy

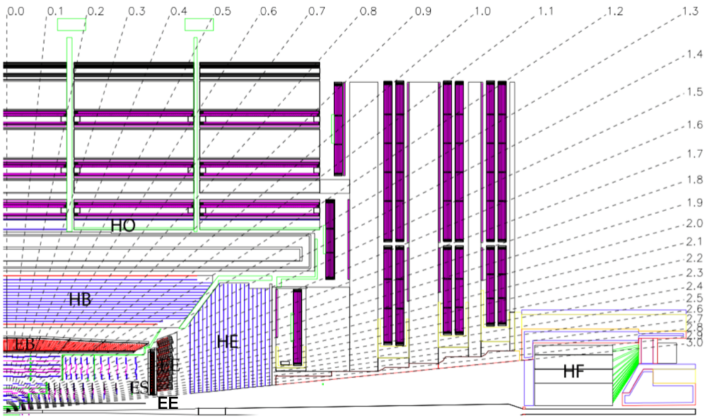

Figure 2.7: Cross section of one quadrant of the CMS calorimeter system showing the EB, ES, EE, HB, HE, HO, HF. Taken from Ref. [97]

the pseudorapidity region of $1.3 < |\eta| < 3$, it is divided into 18 ϕ sectors, and it is composed of 79 mm thick brass absorber plates with a 9 mm gap for the scintillator material. In the $\eta - \phi$ plane it consists of $2,304$ towers with varying sizes. The total interaction length of the HE is $\sim 10\lambda_I$.

The HO is placed in the central region of CMS, covering $|\eta| < 1.3$, just outside of the solenoid, in front of the first barrel muon detector layer. It is used to identify particles from energy tails in the hadronic showers, that passed the HB and the solenoid. HO is placed in 5 rings along the z axis, following the geometry of the barrel muon detector. Each ring has 12 ϕ sectors, separated by 75 mm stainless steel which holds together the successive layers of iron yoke and muon chambers. HO extends the HB absorption to approximately $10\lambda_I$.

The HF is located in the very forward region beyond the muon endcaps, at $|z| = 11.2$m, covering pseudorapidity ranges $3.0 < |\eta| < 5.2$. It is used to cover the full hermiticity of the CMS and collect even the softest, most forward, collision products, which will contribute to accurately reconstruct the $E_{\mathrm{T}}^{\mathrm{miss}}$ of the events. Quartz fibers are used as the active material and stainless steel as the absorber material of the HF.

HCAL and ECAL barrels were exposed to beams of electrons, pions, protons and muons and yield a hadronic energy resolution of [103]

$$(\frac{\sigma}{E})^2 = (\frac{85\%}{\sqrt{E[GeV]}})^2 + (7.0\%)^2 \tag{2.6}$$

the noise term is found to be negligible. Similar energy resolution was found in the endcaps.

Calorimeter Upgrades for HL-LHC, the HGCal

The conditions of the HL-LHC will be challenging for the longevity and the performance of the detector. These challenges affect particularly the high pseudorapidity regions where the radiation levels are significantly higher than in the central regions. The HL-LHC upgrades dictate re-examination of the ability of the active material and electronics of the CMS calorimeters to operate at very high luminosities and PU [100].

Radiation damage is not an issue for the detector material of the EB, however, its electronics must be upgraded to meet the Level-1 readout rate and latency requirements. For the HB, some portions of the detector active material will need to be replaced, while there is no need for upgrades of the photodetector and electronics. The HO will see relatively little dose and the HF is designed to withstand the full HL-LHC integrated luminosity, without significant loss of performance. No upgrade is planed for these subsystems.

To address the challenge of high radiation and PU the CMS collaboration will replace the ECAL and HCAL endcaps ($1.5 < |\eta| < 3$) with a High Granularity Calorimeter (HGCal) [104]. The design of the HGCal incorporates an electromagnetic section, which starts near the front face of the current EE and consists of 28 layers of tungsten and copper plates interleaved with silicon sensors as an active material for a radiation length of 26 X_0 and 1.5 λ_I. The hadronic section has 12 layers of brass and copper plates and silicon sensors as an active material for an interaction length of 3.5 λ_I. A backing-hadronic (BH) calorimeter made of brass absorber and plastic scintillator will be added at the end of HGCal, adding 5 interaction lengths λ_I for the full shower containment.

The active material of the HGAal will be in hexagonal silicon sensors with $\sim$ 1-0.5 cm^2 cells and highly segmented plastic scintillators in the BH. The high granularity of the system will allow for 3D topology measurements of shower energy and precise time-stamping. HGCal will facilitate particle flow type calorimetry that will enhance the particle identification, the PU rejection and the jet energy resolution.

2.2.4 Muon Detectors

Muons are characterized by long lifetime (c$\tau_\mu \approx$ 660m) in their rest frame [8], large mass compared to electrons and no colour charges. Taking into account the boost of the muon from the rest frame to the transverse plane of the lab frame, the mean decay length becomes $c\tau_\mu x p_\mu / m_\mu$, i.e. in the range from tens to thousands of km for the typical transverse momenta of the muons produced in CMS. Therefore, the muons can be treated as stable particles for all practical purposes.

For a charged particle to initiate an EM shower in an absorbing material and become a minimum ionising particle (MIP), its energy must exceed the critical energy in the material. After that point, energy loss by bremsstrahlung, leading to a shower, begins to dominate over energy loss by ionisation by which no shower can start. From the Bethe-Bloch theory of energy loss, the critical energy scales with the square of the mass of the particle. Therefore, the critical energy for the muons is 40,000 higher than that for the electrons in the same material. Thus only extremely energetic muons, at the TeV scale, have a small chance to shower in the calorimeters of CMS. The electrons, on the other hand, produced in CMS with energies orders of magnitude higher than the critical energy in the absorbing materials of the EM calorimeter, always shower in it. The same argument applies also to other charged particles such as pions, kaons, and protons.

The muon system is the outermost component of the CMS detector, placed outside the solenoid, interplaced with the iron yoke layers, forming a cylindrical geometry. Ideally only muons and neutrinos, that escape detection, reach the muon system. The muon detection is based on gas ionization. The charged particles enter the gas chamber, ionize the gas atoms and cause electric signals. These signals are called "hits" and are associated with well-defined locations in the detector. Hits in the muon stations are combined with tracker information for muon identification and momentum and charge measurement. The layout of the muon detector technologies is shown in Fig. 2.9. Different detector technologies have been adopted in the barrel and the endcaps, to account for the different conditions.

Drift tubes (DT) are used in the central region, where the particle fluxes are low and the magnetic field is nearly uniform. The barrel muon system covers up to $|\eta| < 1.2$, it is segmented into 5 wheels along the direction of the beam and each wheel is divided into 12 ϕ sectors, as depicted in Fig. 2.8. The muon barrel consists of 4 layers (stations) of a total of 250 DT chambers. The first 3 DT layers contain 12 planes of DT cells each, 8 of them measure coordinates in the r-ϕ plane and 4 of them measure in the r-z plane. The outermost DT layer measures only in the r-ϕ with 8 planes of DT cells.

The endcaps are characterized by strongly non-uniform magnetic field and higher particle flux compared to the barrel. The robust Cathodic Strip Chambers (CSC) are used to equip these forward regions, covering a pseudorapidity range of $0.9 < |\eta| < 2.4$. There are 468 radiation resistance CSCs in total, with fast response and fine segmentation. Grouped in 4 stations in each endcap the chambers are positioned perpendicular to the beam line, interspersed between the iron yoke layers. The anode wires and the cathode strips of the CSCs are perpendicular to each other and provide 2 coordinate measurements for each muon passing through the volume of the chamber. As depicted in Fig. 2.9, there are 6 layers of Resistive Plate Chambers (RPC) employed in the barrel and 4 layers in each endcap, covering pseudorapidity up to $|\eta| < 1.9$ and an additional fifth layer covering $1.2 < |\eta| < 1.6$. The single point resolution varies from $80 - 120$ μm

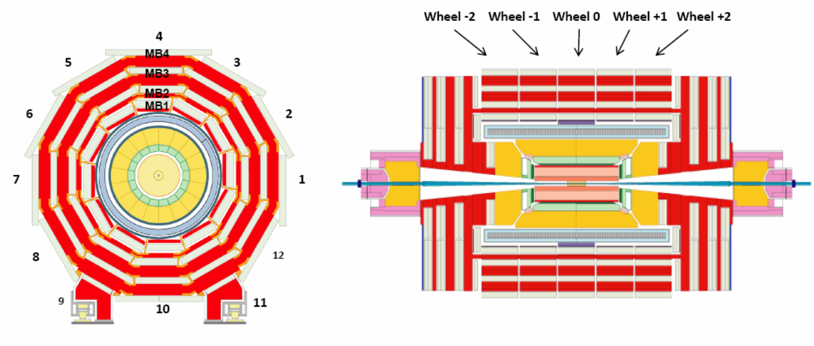

Figure 2.8: CMS schematic view on the r-ϕ plane showing the 12 ϕ sectors (left), schematic view on the r-z plane showing the 5 wheels (right). Taken from Ref. [105]

in DTs to $40 - 150\ \mu m$ in CSCs and $0.8 - 1.4\ cm$ in RPCs [106]. The DTs and CSCs exhibit better spatial resolution of the muon track measurements while RPCs show a faster response. RPCs deliver highly accurate time tagging that can be used for triggering purposes. All the muon detectors are used collectively to achieve an optimal resolution in both spatial and time measurements.

Muon System Upgrades for the HL-LHC

The major LHC upgrades towards the HL-LHC, discussed in 2.1.3, motivates the muon systems upgrade in order to cope with the increased luminosity and the high PU. The muon gas detectors are essential for the precise and accurate muon identification and measurement. Upgrades on the electronics will ensure high performance during the harsh HL-LHC conditions. The DT electronics will be replaced in order to improve radiation tolerance and increase the trigger rate capabilities. Additionally, the CSC front end electronics will be replaced to account for the increased occupancy and the larger L1 rates.

The muon identification will become more challenging in the forward region of the detector, due to the high particle rate and the low magnetic bending. These regions will be enhanced with additional new detectors, which will add measured hits in the muon tracks and provide robust track reconstruction. Two layers of Gas Electron Multiplier (GEM) detectors will be added in the forward region in the first 2 muon stations and cover the pseudorapidity range of $1.6 < |\eta| < 2.4$. In these stations the bending angle is large and the momentum determination is most effective. The additional detectors will increase the path length by 15-40 cm in each station. Additionally, in the 3rd and 4th muon stations, improved RPCs (iRPCs) will be installed and provide background reduction in triggering with a very precise time tagging that can be used for PU mitigation. Figure 2.9 shows the new CMS muon system, including the GEMs and iRPCs. The smaller size of the HGCal compared to the current calorimeter endcaps, will allow for the muon system to extend to higher η regions. A small but precise muon detector will be installed at the back of HGCal, in front of the first layer of the existing muon endcap. This will be a 6 GEM layer detector (ME0 in Fig. 2.9) and will extend the muon system coverage up to $|\eta| = 2.8$

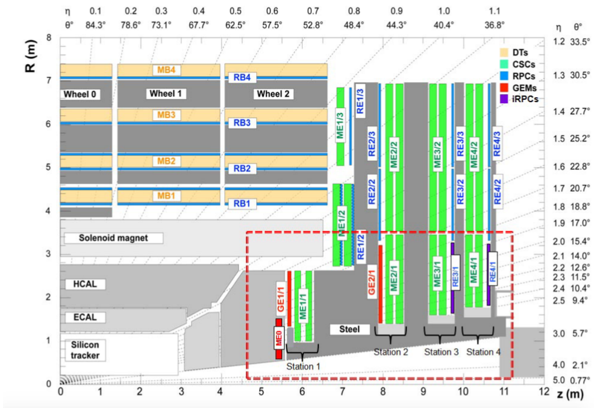

Figure 2.9: View of a quarter of the CMS system. The new forward muon detectors for Phase-II are contained within the red dashed box. The GEMs are indicated in red and the iRPCs in dark blue. Taken from Ref. [100]

Chapter 3

CMS Triggering and Object Reconstruction

The LHC collides proton bunches at a rate of 40 MHz, spaced 25 ns apart. Considering a storage size of ~ 1 MB per bunch crossing event, the LHC collision frequency would result to data output of 40 TB per second. However, not all the collision events contain interesting information for physics searches and due to storage restrictions, only a small fraction of those is saved for offline analysis. It is necessary to "trigger" on events with interesting signatures for the CMS physics program and filter out the unwanted events. The filtering of the interesting events is performed by the CMS trigger system which consists of two stages: the Level-1 Trigger (L1 Trigger) and the High Level Trigger (HLT).

The structure of this chapter is the following: Section 3.1 describes the CMS L1 Trigger, its performance during Run 2 and the plans for its upgrade towards HL-LHC. The HLT and the Data Acquisition system are presented in Section 3.2 while Section 3.3 elaborates on the offline physics object reconstruction by describing the Particle Flow algorithm and the transition from channel-based signals to combined entities. The reconstruction of physics objects such as leptons, hadronic jets and missing transverse energy, that will be used in physics analyses, is detailed at the end of the section.

Three projects are presented in this chapter and concern the L1 dimuon trigger performance, the study of the Kalamn Muon Track Finder algorithm performance on the low-mass dimuon reconstruction and Phase-2 studies of the topological $\tau \to \mu\mu\mu$ L1 trigger. The projects are presented in dedicated sub-section following the description of the firmware and the Phase-2 upgrade.

3.1 The Level-1 Trigger

The L1 Trigger system is instrumented with custom-designed hardware processors, and runs event selection algorithms using information from the CMS subdetectors. It takes a decision within 3.8 μs and selects up to 100 kHz of interesting events, out of the 40 MHz rate it receives. The CMS L1 Trigger system has been largely upgraded between 2015-2016, motivated by the increase of the LHC center of mass energy from 8 to 13 TeV. The L1 Trigger is composed of 2 parts: the L1 calorimeter trigger and the L1 muon trigger. The output of the two subsystems is collected by the Global Trigger (GT) which takes the final decision on the event. The decision is made based on $\sim$ 300 event selection algorithms that depend on kinematic quantities, position, isolation and quality of the event's objects. The selection algorithms, also referred to as L1 Trigger seeds, are executed in parallel for the final trigger decision. The structure of the CMS L1 Trigger is presented in Fig. 3.1

The L1 calorimeter trigger consists of two layers. The Layer-1 receives, calibrates and sorts the local energy deposits, which are called "trigger primitives" (TP) and are sent from ECAL, HCAL and HF. The Layer-2 uses the TPs to reconstruct physics objects. The input to Layer-1 is organized into trigger towers (TT) that correspond to $\Delta\eta \times \Delta\phi$ of 0.087×0.087 each. Every TT encodes energy deposits at a specific position in the calorimeters. A Time Multiplexed Trigger (TMT) architecture is used and allows for the information of all the TT in the event to be received by the Layer-2. There are no regional boundaries in the object reconstruction and the full granularity is exploited when the energy deposits are computed. The output of the TMT nodes are collected in the de-multiplexing (DeMux) node, sorted and sent to GT [108].

The L1 muon trigger consists of 3 regional muon reconstructing algorithms [109, 110].

- The Barrel Muon Track Finder (BMTF) receives DT TPs and RPC hits from $|\eta| < 0.83$. The TPs and hits are combined in "superprimitives" in TwinMux;
- The Overlap Muon Track Finder (OMTF) receives uncombined DT TPs and RPC hits transmitted from TwinMux, together with CSC TPs. The TPs and hits delivered to the OMTF cover the range from $0.83 < |\eta| < 1.24$;
- The Endcap Muon Track Finder (EMTF) takes as input CSC TPs and RPC hits from the forward pseudorapidity regions of $1.24 < |\eta| < 2.4$, through CPPF;

The muon track finder algorithms measure the charge, transverse momentum and angle of the L1 muon candidates, and assign to them a quality bit based on the reconstruction fit. The Global Muon Trigger (GMT) receives up to 36 L1

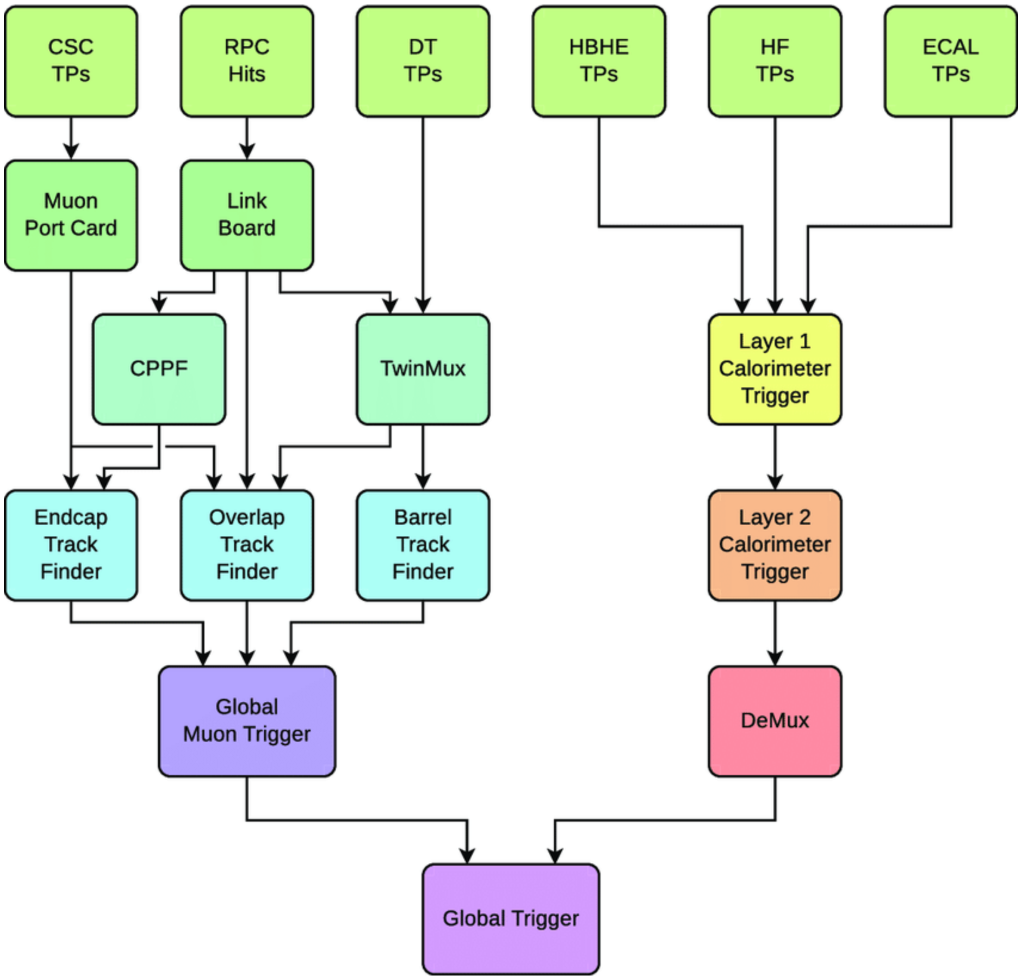

Figure 3.1: Diagram of the CMS L1 Trigger system. The L1 muon trigger (left) and the L1 calorimeter trigger (right) receive input trigger primitives and hits from the different sub-detectors. Their outputs are combined in the GT (bottom) which takes the final trigger decision. Labels in the diagram correspond to trigger primitives (TPs), concentration preprocessing and fan-out (CPPF) and de-multiplexing card (DeMux). Taken from Ref. [107]

muon candidates from each track finder. The GMT sorts the candidates in transverse momentum, assigns to them a quality bit based on the p_T resolution and the number of hits and removes the duplicates. Additionally, the GMT receives spatial coordinates for each candidate in the muon stations and extrapolates the track back to the interaction point. The extrapolation corrections are derived from simulation as a function of p_T, η, ϕ and charge and are stored in look up tables (LUT). The corrected coordinates are propagated to the GT and improve the performance of the L1 trigger seeds that rely on the invariant mass or the spatial coordinate difference between multiple muons. The improvement on the dimuon mass reconstruction at L1 with the coordinate extrapolation to the vertex will be discussed in more details later in this section.

A maximum of 8 muon candidates are chosen based on a combination of quality and transverse momentum metric and are sent to the GT for the final L1 Trigger decision. The GT collects L1 muons and calorimeter objects and executes every selection algorithm in parallel in order to make the final trigger decision.

3.1.1 The L1 Calorimeter Trigger Performance

Electrons and photons are indistinguishable to the L1 trigger and therefore they are referred to as e/γ objects. The e/γ trigger efficiency is measured with the tag-and-probe method on $Z \to ee$ events. The tag-and-probe method exploits the fact that the leading leg ("tag") of the Z decay, triggers the event, and the subleading leg ("probe") can be used to measure the unbiased trigger efficiency [111]. The left plot of Fig. 3.2 shows the L1 e/γ trigger efficiency for Run 2, as a function of the offline reconstructed electron E_T for thresholds of 30 and 40 GeV.

The hadronically decaying τ_h leptons are reconstructed to one or multiple charged or neutral pions that produce energy clusters. The performance of the L1 τ algorithm is measured in Run 2 data on $Z \to \tau_\mu \tau_h$ events using the tag-and-probe method. The efficiency shown in the middle plot of Fig. 3.2 is measured as a function of the offline reconstructed p_T of the τ_h, for three L1 E_T thresholds and reaches the plateau of 100%.

The efficiency of the L1 jet triggers is measured inclusively in η using an independent data sample, collected with a single-muon selection algorithm, is shown in the right plot of Fig. 3.2. It shows a sharp turn-on and high efficiency for a number of thresholds used in Run 2.

3.1.2 The L1 Muon Trigger Performance

The efficiency of the L1 muon trigger was measured in data with the tag-and-probe technique with offline reconstructed muons from $Z \to \ell\ell$ events. Figure 3.3 shows the inclusive in η efficiency as a function of the reconstructed probe muon p_T, for a L1 p_T threshold of 22 GeV. The integrated efficiency is higher than 90% for the specific $L1$ p_T threshold and it reaches 95% at the plateau.

The GT combines information from the L1 calorimeter trigger and the GMT. Its large processing power allows the implementation of a menu with sophisticated, analysis-targeted, L1 trigger seeds. The next paragraph describes one type of analysis-targeted selection algorithm which aims at triggering on the low-mass dimuon resonances.

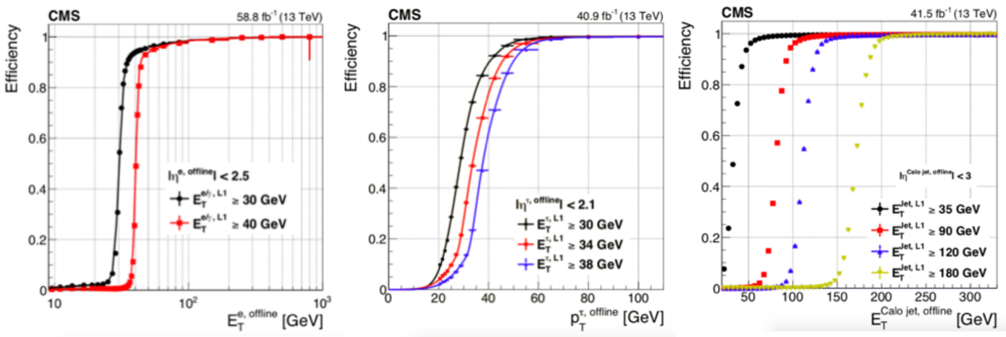

Figure 3.2: The L1 e/γ trigger efficiency as a function of the offline reconstructed electron E_T for thresholds of 30 and 40 GeV (left); the L1 τ trigger efficiency as a function of the offline reconstructed τ lepton p_T for L1 τ E_T thresholds of 30, 34 and 38 GeV (middle); the L1 jet trigger efficiency as a function of reconstructed calorimeter jet E_T, for L1 jet E_T thresholds of 35, 90, 120 and 180 GeV (right). Taken from Ref. [107].

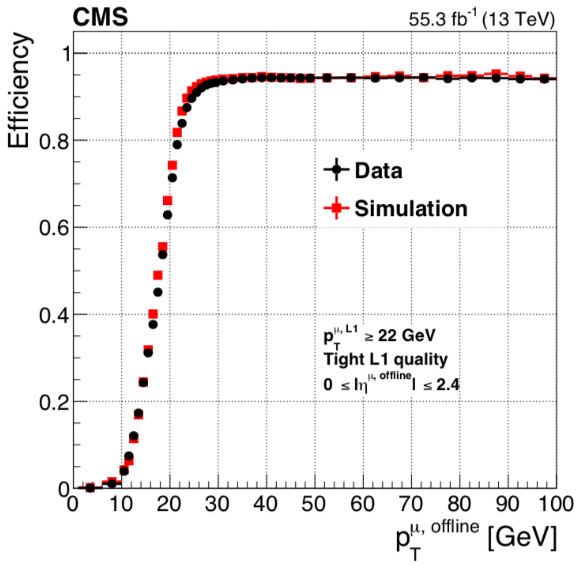

Figure 3.3: The single-muon L1 Trigger efficiency, for 2018 data and simulation as a function of probe muon p_T, for all reconstructed muons in the CMS acceptance. The efficiency is measured with the tag-and-probe method. Take from Ref. [107].

3.1.3 The L1 Dimuon Trigger Performance

The p_T thresholds of the usual L1 dimuon trigger seeds are not well adapted to record low mass resonances. These thresholds are typically 15 GeV on the leading muon and 5 GeV on the subleading muon, so they only select very boosted low-mass dimuon resonances. Dedicated triggers are used to collect low-mass dimuon pairs which can potentially be sensitive to new physics. The study of the L1 dimuon mass resolution was performed using the 2017 MuOnia data sample. The MuOnia data sample is collected using L1 trigger seeds that require $3 < m_{\mu\mu} < 9$ GeV and $5 < m_{\mu\mu} < 17$GeV. They seed a combination of HLT paths with low dimuon invariant mass threshold, designed to collect muons from the decay of the Φ resonance, and medium dimuon invariant mass thresholds, dedicated to collect muons from the decay of Υ meson.

For the offline calculation of the dimuon invariant mass only events with at least two L1 and two offline reconstructed muons are considered. The L1 muon candidates are sorted in descending p_T and only the leading and subleading muons, with opposite sign, within $|\eta| < 2.4$ are kept in the selection. The reconstructed muons are matched to the two L1 muon candidates by imposing a cut on the angular distance of the objects as $\Delta R^{min} = \sqrt{\Delta\phi^2 + \Delta\eta^2} < 0.3$.

Figure 3.4 shows the offline and the L1 $m_{\mu\mu}$ spectra, with and without extrapolation of the L1 muon track at the interaction point. The muon track extrapolation to the vertex has been added online in the GMT during 2017 and brings a clear improvement in the L1 dimuon invariant mass resolution. The 9.46 GeV Υ meson peak can be isolated quite distinctly after the track extrapolation to the nominal vertex. The L1 mass spectrum appears shifted compared to the offline due to p_T offsets designed to make the L1 muon trigger 90% efficient at a given p_T threshold. The results of the study are presented in [107] and were used successfully by many b-physics triggers. A recent example of a successful low-mass trigger is the 5.6 σ observation of $B_s^0 \to \mu^+\mu^-$ decay [112].

Phase-2 upgrade of the Level-1 Muon Trigger

The Phase-2 upgrade of the L1 trigger is designed to benefit from the HL-LHC sub-detector upgrades. It is planned to use state of the art techniques, that have been used for offline reconstruction. The L1 trigger rate will increase from 100 kHz to 750 kHz and the latency from 3.8 μs to 12.5 μs.

The Phase-2 muon trigger system will be substantially upgraded with respect to the legacy system. It will exploit inputs from new detectors, L1 tracks from the tracker track, new interfaces, new muon track finders and topological object reconstruction algorithms. At the same time, it will maintaining its regional division into Barrel, Overlap, and Endcap track finders. The signal from the DT, CSC, RPC and from the GEMs and iRPCs will be added to reconstruct standalone muons and measure their p_T. In the barrel region the Kalman Barrel Muon Track Finder (KBMTF) algorithm will replace the legacy BMTF algorithm. The new track finder employs an approximate implementation of the Kalman Filter (KF) algorithm [113] for the track reconstruction. The KBMTF has the advantage of reconstructing muon tracks with and without vertex constraints, allowing

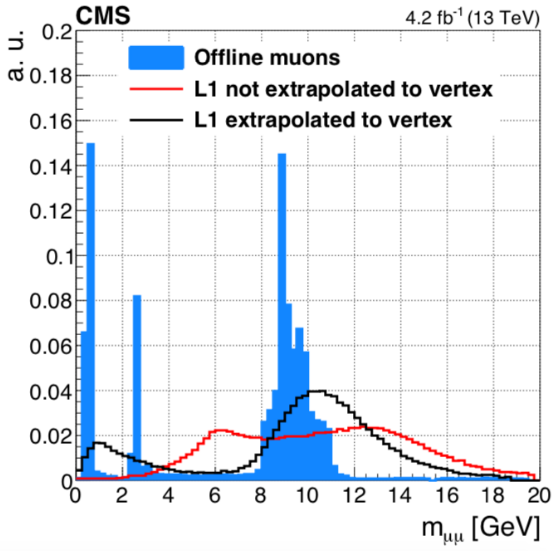

Figure 3.4: The $m_{\mu\mu}$ spectrum of offline muons, compared to the L1 spectrum with and without track extrapolation to the vertex. The L1 spectrum appears shifted compared to the offline, due to p_T offsets designed to make the L1 muon trigger 90% efficient at a given L1 threshold [107].

for displaced muons to be triggered efficiently. The performance of the KBMTF in reconstructing low-mass dimuon pairs, that originate from the decay of a J/ψ meson was studied. The KBMTF algorithm and the performance study will be discussed in more detail in the next sub-section. The OMTF is planned to use an improved version of the naive Bayes classifier, which it currently uses for the muon p_T assignment. The EMTF++ will replace the current EMTF and will combine information from the CSCs, RPCs, GEMs and iRPCs. A deep neural network will be used for the p_T assignment, with and without vertex constraint. The availability of the tracker tracks at L1 will allow for another L1 muon category to be reconstructed. The propagation of the L1 tracks to the muon detectors and their association with muon stubs from at least 2 muon stations, lead to the reconstruction of the TkMuStub L1 objects. Those L1 muon objects are assigned the tracker track p_T, which has significantly better resolution than the measurement of standalone muons. Additionally, these objects allow the tagging of very soft muons, that only leave hits in the first muon stations, closest to the interaction region, and would not be reconstructed as standalone muons. Also, muon reconstruction from regions with limited muon station coverage is possible with the new TkMuStub L1 object [113].

3.1.4 The KBMTF Performance

The BMTF is an efficient algorithm used for the identification of muon tracks and the measurement of their transverse momentum during Phase 1. It uses information from two muon stations and vertex extrapolation, and assigns momentum to the muon track using LUT. The goal for HL-LHC is to include information from more than two muon station in the muon track reconstruction, in order to improve the momentum resolution of the algorithm. Additionally, it will support momentum assignment without vertex constraint in order to trigger on displaced muon tracks.

The KBMTF was proposed in 2017, for the upgrade of the L1 barrel muon track finder algorithm, towards the HL-LHC. It was developed, studied and implemented in firmware during the end of Run 2 and it will be the main muon track finder algorithm in barrel, already from the beginning of Run 3 data taking.

The KBMTF track reconstruction is an iterative process which starts from the outer muon station. The algorithm uses the track parameters and their uncertainties at the outer station, neglects the muon energy losses in the muon stations, and predicts the segment parameters and their uncertainties in the next station. The predicted track parameters are propagated to the next station, the closest measured stub is selected and the parameters of the track are updated based on the values and uncertainties of the measurement. The same procedure is repeated up to the innermost muon station where the track parameters are saved without vertex constrain. The track is further extrapolated to the vertex, taking into account energy losses, and the vertex constrained track parameters are also saved. Overall the KF algorithm exploits the measurements in all detector stations, provides a vertex unconstrained measurement for displaced muons, and a vertex constrained measurement for muons originating from the beamspot. It therefore satisfies all the requirements for the upgrade of the muon trigger in the barrel. The sketch in Figure 3.5 shows a slice of the CMS barrel with a muon track traversing the detector and leaving hits in four muon stations.

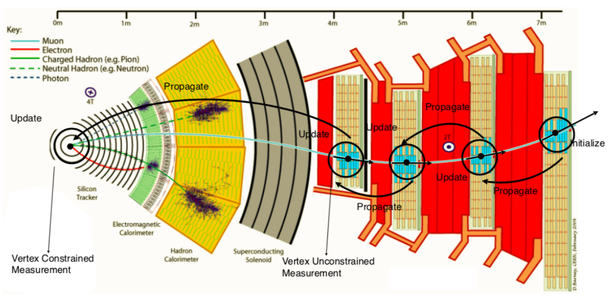

Figure 3.5: A muon track traversing the CMS detector. The Kalman Filter track finder is illustrated, starting from the outermost muon station, propagating inwards and updating the track parameters at each station. Taken from Ref. [113].

The track parameters update is a matrix manipulation process which results in the Kalman gain matrix. The Kalman gain was studied in simulation and found that it can be precalculated for different values of the curvature and different combination of the hits. Therefore, the Kalman filter propagates the track from station to station and updates its parameters based on the precalculated Kalman gain.

The KBMTF firmware was commissioned in the CMS L1 trigger, in the end of the 2018 data taking period, into the same FPGA card with the legacy BMTF. The two algorithms run in parallel, the BMTF was used for trigger and the KBMTF was read out in the DAQ for each collected event. With the implementation during Run 2, real data were collected and the performance of the algorithm was studied. The KBMTF will be deployed online as the default barrel muon track finder algorithm in Run 3.

The KBMTF performance was studied and compared to that of the legacy BMTF algorithm. The study focuses on the L1 efficiency of reconstructing low mass dimuon pairs, that originate from the decay of J/ψ mesons. The two muon legs from a J/ψ decay are expected to be in small proximity due to the high momentum of the mother particle. For the performance study a J/$\psi \to \mu\mu$ Monte Carlo sample with a $p_{T\ Gen}^{J/\psi} > 8$ GeV cut was used. The sample was generated using Pythia 8 event generator [114] and GEANT 4 [115] for the event simulation in the detector. The sample simulates the Run 2 conditions and it has a flat PU of 28 to 62.

The KBMTF and BMTF efficiencies to reconstruct the J/ψ dimuon pairs, were measured as the ratio of the number of the L1 reconstructed J/ψ events over the number of generated J/ψ events.

$$eff_{J/\psi} = \frac{N_{L1}^{J/\psi\ candidates}}{N_{gen}^{J/\psi}} \tag{3.1}$$

The generator level muons that originate from the J/ψ decay, are required to be within the barrel ($|\eta| < 0.83$) and have $p_T^{\mu,\ Gen} > 6$ GeV. Additionally, they are required to have a dimuon mass within the J/ψ mass window ($3.05 < M_{\mu\mu} < 3.15$ GeV) and opposite electric charge. The L1 muon candidates are selected such that they both have $p_T^{L1} > 5$ GeV, $|\eta| < 0.83$. Additionally, the L1 candidates are matched to the generator level muons from J/ψ, within a cone of $\Delta R_{min} < 0.35$. The L1 muon candidates are required to be within the same dimuon mass window as the generator level pair and to have opposite electric charge.

The plots in Figure 3.6 show the BMTF and KBMTF efficiencies of reconstructing the J/ψ dimuon pair as a function of $\Delta\eta$, $\Delta\phi$ and ΔR between the two generator level muons. It is observed that both efficiencies decrease as the angular distance between the two generator level legs of the J/ψ increases. This translates into lower efficiency of the L1 track finder algorithms to reconstruct muons that originate from low p_T J/ψ.

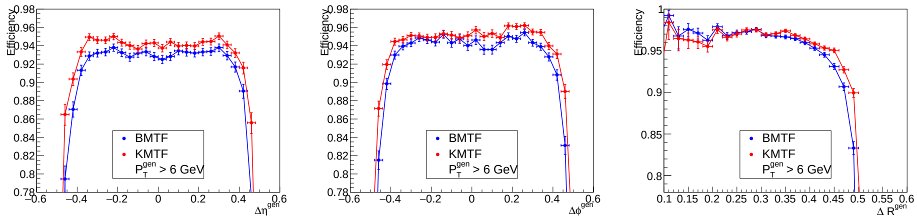

Figure 3.6: L1 BMTF and KBMTF efficiency as a function of $\Delta\eta^{Gen}$ (left), $\Delta\phi^{Gen}$ (middle) and ΔR^{Gen} (right). The two generator level muons are selected such that they originate from the decay of a J/ψ meson.

The decrease in the L1 efficiency can be explained by the fact that the J/ψ p_T approaches the p_T threshold of the generator level muons. In this region, the BMTF and KBMTF efficiencies follow a sigmoid function and therefore the efficiency decreases gradually. This can be observed also in figure 3.7.

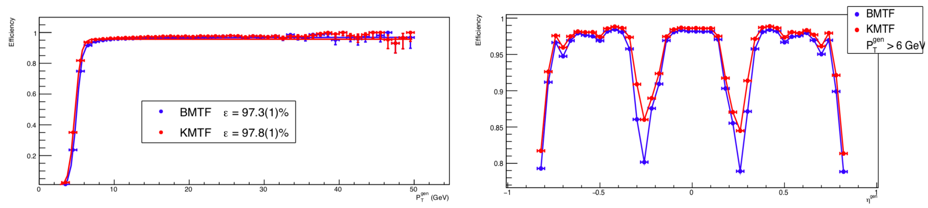

Figure 3.7: L1 BMTF and KBMTF efficiency in reconstructing L1 muons that originate from the decay of J/ψ mesons, as a function of the generator level J/ψ p_T (left) and generator level η (right). The efficiency vs p_T^{Gen} curve has been fitted with a Crystalball function.

The study shows that the new KBMTF algorithm has very good efficiency in reconstructing muons that originate from the decay of J/ψ mesons. The performance of the new algorithm is very similar to that of the legacy BMTF algorithm. The study was conducted during the developments of the KBMTF firmware and emulator and contributed in the efforts for the commissioning of the algorithm at the end of Run 2.

3.1.5 The topological $\tau \to \mu\mu\mu$ L1 Trigger for Phase-2

The lepton flavor violation (LFV) [116, 117] is an established fact in the neutral lepton sector, since the observation of the neutrino oscillations [118]. The latter shows that neutrinos have non-zero mass and sizeable mixing among their flavors. Evidence from solar [119, 120], atmospheric [121] and reactor neutrino experiments [122] established a pattern for the neutrino mass differences and the mixing angles which is described by the PMNS matrix. However, the mixing between different flavours of charged leptons, has not been observed yet. The SM predicted branching ratio of charge LFV (cLFV) is very small ($\mathcal{O}(10^{-40} - 10^{-50})$) and any experimental observation of such a phenomenon would be a clear evidence for new physics. Therefore, the area of searches for cLFV is very active [117, 123, 124].

The cLFV is searched in channels like $\mu \to e\gamma$ and $\mu \to 3e$ in precision measurement experiments [125, 126, 127]. The decay channels of $K_L^0 \to e^{\pm}\mu^{\mp}$ [128] and $K^+ \to \pi^+\mu^+e^-$ [129] are investigated in dedicated experiments that use kaon beams while the decay channels of $B^0 \to e^{\pm}\mu^{\mp}$ [130] and $D^0 \to e^{\pm}\mu^{\mp}$ [131] are searched at the LHCb. The cLFV decays of the Z boson were studied at LEP [132, 133] and recently the ATLAS collaboration has set an upper bound on the $Z \to e^{\pm}\mu^{\mp}$ mode [134]. Finally, limits on the cLFV decay of the τ lepton have been set by e^+e^- experiments such as Belle and BaBar but also from the LHCb, CMS and ATLAS collaborations .

In the performance study presented here, we are interested in the cLFV decay of $\tau \to \mu\mu\mu$. The strongest limit on its branching ratio was obtained by the Belle collaboration to be $\mathcal{BR}(\tau \to \mu\mu\mu) < 2.1 \times 10^{-8}$ [135]. The CMS collaboration calculated the upper limit on the branching ratio $\mathcal{BR}(\tau \to \mu\mu\mu) < 8.0 \times 10^{-8}$ [136], using the 2016 dataset which corresponds to integrated luminosity of 33.2 fb^{-1}. The CMS upper limit on the branching ratio accounts both for the τ production through W boson decays and through heavy flavor hadron decays which are more prominent in the endcaps.

The L1 trigger system is going through developments and improvements in order to achieve the goals of the HL-LHC physics program. The Phase-2 upgrades of the L1 trigger system utilize technological advances to enhance the physics object selectivity. The high efficiency of the L1 muon trigger is of great importance for the Phase-2 upgrade. The L1 muon trigger capabilities will be extended and tracks from the tracker will be used together with other trigger objects for the muon reconstruction. The latter will lead in L1 muon p_{T} resolution improvements close to the level of the offline reconstruction and extension of the muon p_{T} thresholds to lower values. The combination of the L1 tracker tracker (TTTracks) with muon tracks or stubs will allow for improved p_{T} and spacial resolution of the L1 muons and the deployment of trigger paths with low p_{T} thresholds and high muon efficiency. These improvements will extend the CMS reach in signatures with soft leptons, multi-lepton trigger paths in which the targeted muons are in close-proximity and track-based isolation trigger requirements.

The L1 TTTracks will be combined with L1 muon tracks or stubs and produce the track-correlated muons. This is a new type of muons that can be introduced due to the addition of the L1 tracker in the Phase-2 L1 trigger. This new type of L1 muons will have improved efficiency and resolution compared to the Phase-1 L1 muon objects. The tracker will provide the L1 track reconstruction and the muon systems will be used for the muon identification. The track-correlated muon will be produced in the Global Muon Trigger (GMT) unit of Phase-2, which will be interfaced with the L1 tracker track finder.

The matching of the L1 TTTrack to a combination of muon stubs lead to the L1 track-correlated muon called TkMuStub. The L1 TTTracks p_{T} measurement is assign to the TkMuStub muon trigger object. This allows for tagging low p_{T} muons that only leave hits in the muon stations closest to the interaction region and would not be reconstructed by the muon track finding algorithms. This is particularly important for triggering on physics signatures with soft muons such the $\tau \rightarrow \mu\mu\mu$.

Part of the physics program of the CMS collaboration for the Phase-2 upgrade aims at improving the sensitivity in the $\tau \rightarrow \mu\mu\mu$ channel and therefore a dedicated L1 trigger is important. The study presented in this sub-section focuses in the L1 Trigger performance of a dedicated $\tau \rightarrow \mu\mu\mu$ trigger where the τ lepton originates from the decay of a W boson. The study is complementary to those looking at $\tau \rightarrow \mu\mu\mu$ in which the τ lepton originates from heavy flavor hadron decays. While more τ leptons are produced from heavy flavor hadron decays, the τ leptons from the W decays tend to have higher p_T and are typically isolated from other primary vertex decay products. Therefore, they tend to have lower background contamination and acceptance in all the regions of the CMS detector. The drawback of the $W \rightarrow \tau\nu \rightarrow \mu\mu\mu\nu$ channel is the low production cross section of the W boson. The total of maximum expected events to be produced in the full integrated luminosity of 300 fb^{-1} of the HL-LHC operation is 1230, almost 2 order of magnitude lower than the expected $\tau \rightarrow \mu\mu\mu$ events from heavy flavour hadron decays. The above number has been calculated using the latest CMS measurement of the W production cross section [137] and the Belle upper limit on the $\tau \rightarrow \mu\mu\mu$ branching ratio [135].

For the L1 algorithm performance study described here, a dedicated WToTauTo3Mu Monte Carlo sample with PU 200 was used as signal. Additionally, a minimum bias sample [1] with PU 200 was used for the rate measurement. Both samples were produced with PYTHIA 8 generator [114] with the Phase-2 running conditions simulated.

The signal sample was produced with a filter at generator level such that the muons with the highest and second highest p_T in the event (leading and subleading) have $p_T > 2.5$ GeV, and they are both within $|\eta| < 2.9$. The generator level filter efficiency is 72%. Three generator level muons are selected such that they originate from a τ or a μ, within $|\eta| < 2.4$ and they are stable. The leading muon should have $p_T(\ell_1) > 5$ GeV, the subleading muon should have $p_T(\ell_2) > 3$ GeV and the trailing muon should have $p_T(\ell_3) > 2.5$ GeV. The acceptance of $\tau \rightarrow \mu\mu\mu$ events, at generator level, after applying all the selection cuts is 36%.

[1] A minimum bias sample contains events that are selected with a "loose" trigger that accepts a large fraction of the overall inelastic cross section

Regarding the L1 objects, different combinations of TkMuStubs and TT-Tracks are used. The combinations are grouped into "pure" algorithms that use 3 same type L1 objects such as 3 TkMuStubs or 3 TTTracks, and "combined" algorithms that use a mixture of TkMuStubs and TTTracks such as 2 TkMuStubs + 1 TTTrack or 1 TkMuStub + 2 TTTracks.

For the L1 τ candidate reconstruction, kinematic and geometric quantities are exploited. The selection cuts on these quantities serve both for the signal selection and the PU background rejection.

The square invariant mass of the τ candidate after omitting the muon mass is computed as

$$\begin{aligned} M_\tau^2 = 2p_{\mathrm{T}1}p_{\mathrm{T}2}[(cosh(\eta_1-\eta_2)-cos(\phi_1-\phi_2))]+ \\ 2p_{\mathrm{T}2}p_{\mathrm{T}3}[(cosh(\eta_2-\eta_3)-cos(\phi_2-\phi_3))]+ \\ 2p_{\mathrm{T}3}p_{\mathrm{T}1}[cosh(\eta_3-\eta_1)-cos(\phi_3-\phi_1)] \end{aligned} \tag{3.2}$$

The computation of the τ candidate invariant mass depends only on the p_{T}, η and ϕ of the three L1 muons and therefore can be easily implemented in the firmware. The proposed algorithms will use the square mass for the τ mass cut because the square root function is not available in the firmware.

The invariant mass of the three TkMuStaubs and 3 TTTracks, matched to the generator level muon from the τ decay, in the signal sample are presented in Figure 3.8

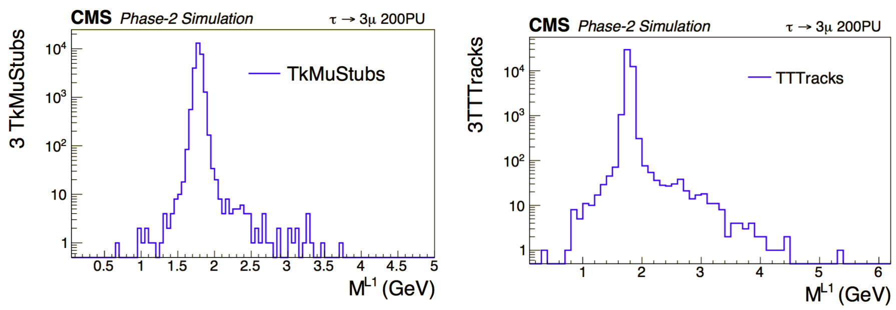

Figure 3.8: The invariant mass of the 3 TkMuStubs (left) and the 3 TTTracks (right).

From the plots in Fig. 3.8 it is apparent that the τ candidate invariant mass distribution has lower background when the 3 TkMuStub L1 objects are used for the reconstructions. Additionally, we conclude that the mass window of [1.5, 2.2] GeV is the appropriate selection in order to reduce the background and retain the signal muons from the τ decay.

Additionally, the three muons from the τ decay are expected to be in close proximity and therefore the ΔR between the pairs is small. The plots in Figure 3.9 show the ΔR distribution of all the pairs.

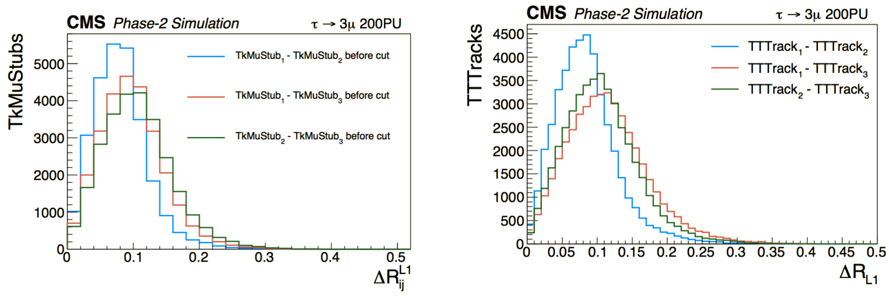

Figure 3.9: The ΔR distribution of the 3 TkMuStubs (left) and the 3 TTTracks (right).

From the plots it is concluded that the ΔR of the signal muons is always lower than 0.3. Therefore a cut on the ΔR of all L1 pairs at 0.3 is applied in order to reduce the background and lead to lower L1 rates.

In addition, the muon tracks are expected to originate from the same vertex. The plots in Figure 3.10 show the dZ_0^{max} distribution for the case of 3 TkMuStubs and 3 TTTracks. In both cases the distribution drop rapidly as the dZ_0^{max} increases. In the 3 TkMuStubs case the background is reduced compared to the 3 TTTracks case, as expected. It was found that a L1 cut at $dZ_0^{max} < 1$ cm is effective in reducing background and retaining a high L1 trigger efficiency.

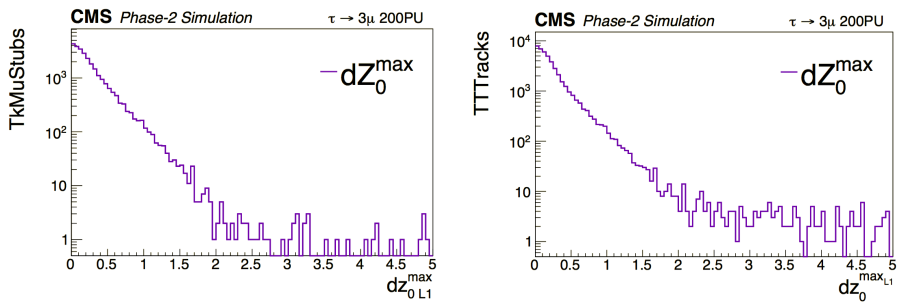

Figure 3.10: The dZ_0^{max} distribution of the 3 TkMuStubs (left) and the 3 TTTracks (right).

The selection requirements on the L1 objects are summarized below.

- The invariant mass of the 3 objects to be within the τ window of $1.5 < M_{3L1} < 2.2$ GeV;
- The ΔR of all the L1 pairs to be $\Delta R < 0.3$. This cut is effective in rejecting background from PU while the muons from τ are expected to be in close proximity;

- The 3 L1 objects are required to originate from the same vertex by applying a cut of $\Delta z_0^{max} < 1$ cm;

For the efficiency measurement the L1 muon candidates are matched to the generator level muons using $\Delta p^{min}_{T(L1-Gen)}/p_{T(Gen)} < 0.05$ within a cone of $\Delta R_{(L1-Gen)} < 0.2$. For the case of the "combined" algorithms, as a first step of the efficiency measurement, the L1 TTTracks are matched to the generator level muons with $\Delta R_{(TTTrack-Gen)} < 0.2$ and $\Delta p^{min}_{T(TTTrack-Gen)}/p_{T(Gen)} < 0.05$ cuts. As a next step the TkMuStubs (either 1 or 2 depending on the "combined" algorithm) are matched to the generator level matched L1 TTTracks, from step one, with $\Delta p_{T(TkMuStub-TTTrack\ matched)} < 0.02$ GeV within a cone of $\Delta R_{(TkMuStub-TTTrack\ matched)} < 0.2$. A cut on the $\Delta p_{T(TkMuStubs-TTTrack)} > 0.02$ is applied for all the combinations of TTTracks and TkMuStubs of the "combined" algorithms, in order to avoid double-counting of the same TTTrack.

The isolation of the L1 reconstructed τ candidate from L1 TTTracks is motivated by 2 factors: the expectation that the τ candidate should be isolated from other decay products, and that isolation will reduce the background that arises from jet. A simple algorithm for the isolation of the L1 reconstructed τ candidate was developed. The steps of the algorithm can be easily implemented in firmware and are the following:

- Find the coordinates of the L1 reconstructed τ candidate using the centroid formula: $C : (\eta, \phi) = (\frac{\eta_1+\eta_2+\eta_3}{3}, \frac{\phi_1+\phi_2+\phi_3}{3})$;
- Find the $\Delta R^{max}_{(L1,C)}$ between the centroid and the 3 selected L1 candidates that pass the kinematic and geometric cuts;
- Within the $\Delta R^{max}_{(L1,C)}$ cone apply $\sum_{i=0}^{N\ TTTracks} p_{T(i)} - \sum_{j=0}^{3\ L1} p_{T(j)} < 5$ GeV;

The performance of the "pure" and "combined" L1 Triggers was studied with and without isolation on the L1 τ candidate. The L1 τ efficiency as a function of the p_T of the generator level τ is shown in Figure 3.11.

The efficiency plateau is $\sim$ 50% for the 3 TkMuStubs algorithm without isolation, it gradually increases with the replacement of TkMuStub with TTTrack and reaches $\sim$ 85% for the 3 TTTracks algorithm. The application of the isolation reduces the efficiency plateau by $\sim 5 - 10\%$ in all algorithms.

The rate was measured with a Phase-2 Minimum Bias sample with 200 PU. In this sample the background arises from soft and hard QCD interactions. The rate is measured with the formula

$$Rate = \frac{N_{MC}(p_T^{max} > Threshold)}{N_{MC}^{Tot.}} * N_{bunches} * f_{LHC} \qquad (3.3)$$

where $N_{bunches} = 2760$ and $f_{LHC} = 11.246$ kHz. The L1 τ candidates are selected by applying the kinematic and geometric cuts mentioned above. The rates of the L1 "pure" and "combined" algorithms is shown in Figure 3.12

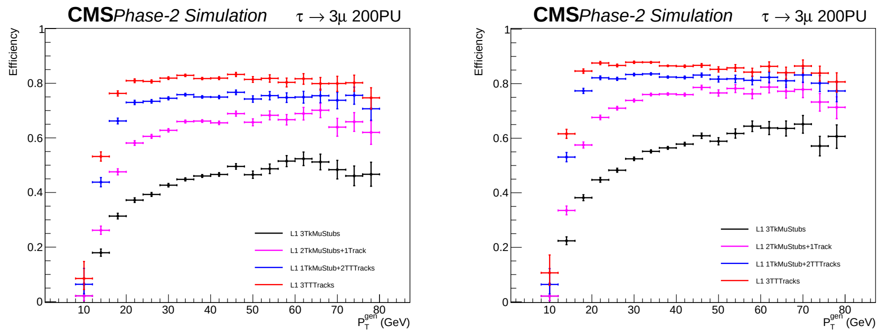

Figure 3.11: The L1 τ candidate efficiency as a function of of the p_T of the generator level τ, with (left) and without (right) the TTTrack isolation. The 4 curves present the efficiency of the 4 different L1 algorithms namely: 3 TkMuStubs (black), 2 TkMuStubs + 1 TTTrack (magenta), 1 TkMuStub + 2 TTTracks (blue) and 3 TTTracks (red).

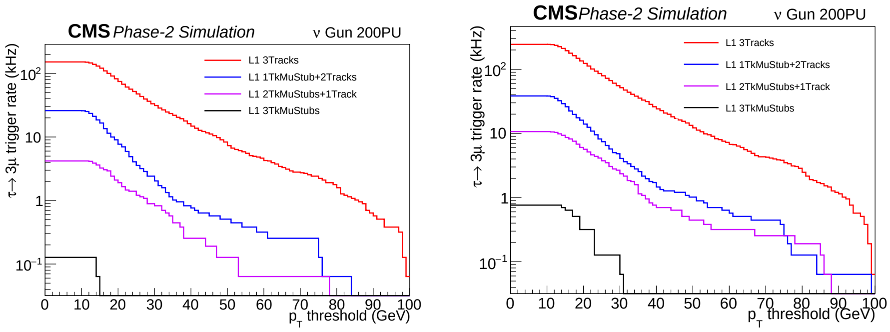

Figure 3.12: L1 τ candidate rate as a function of L1 τ candidate p_T threshold, with (left) and without (right) the TTTrack isolation on the L1 τ candidate. The 4 curves present the rate of the 4 different L1 algorithms namely: 3 TkMuStubs (black), 2 TkMuStubs + 1 TTTrack (magenta), 1 TkMuStub + 2 TTTracks (blue) and 3 TTTracks (red).

The rate is found to be negligible for the 3 TkMuStubs algorithm and non-affordable for the 3 TTTracks algorithms, with or without isolation. The "combined" algorithms give manageable rate of $\sim 10-40$ kHz without L1 τ candidate isolation and even lower around $\sim 4-25$ kHz with the L1 τ candidate isolation, when no threshold is applied on the p_T of the L1 τ candidate. The ROC curves shown in Figure 3.13, depict the rate as a function of generator level τ p_T at which the L1 efficiency is $\geq 90\%$ of the plateau, for multiple L1 p_T thresholds.

One of the "combined" L1 algorithms or a combination of both can be easily implemented in the GMT, with a p_T threshold on the L1 reconstructed τ at $\sim 20-22$ GeV, yielding an efficiency of $\sim 70\%$ and a rate of less than 10 kHz.

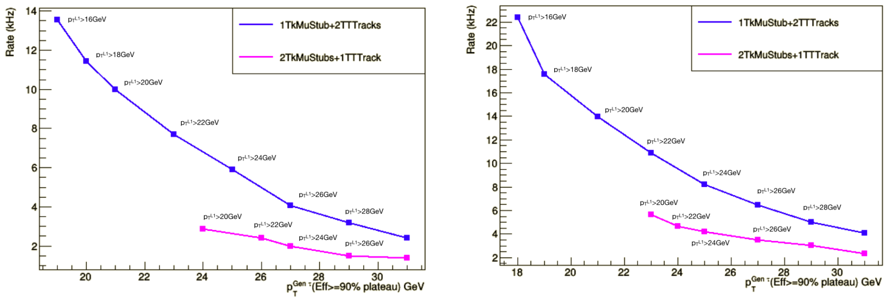

Figure 3.13: The ROC curves show the L1 rate with(left) and without (right) the TTTrack isolation, as a function of generator level τ p_T at which the L1 efficiency is $\geq 90\%$ of the plateau, for multiple L1 τ candidate p_T thresholds. The ROC curves are shown for the "combined" algorithm of 2 TkMuStubs + 1 TTTrack (magenta) and 1 TkMuStub + 2 TTTracks (blue).

The algorithms use simple variables like η, ϕ and p_T, are global in η and can run as complementary triggers to the one that targets $\tau \to \mu\mu\mu$ from heavy flavor decays in the endcaps. The "combined" algorithms, if implemented, will result in a maximum number of expected $W \to \tau \to \mu\mu\mu$ events of $\sim 300 - 360$, after the full HL-LHC operation. The maximum number of events has been estimated assuming the upper limit on the branching ratio of $\mathcal{BR}(\tau \to \mu\mu\mu) = 2.1 \times 10^{-8}$ as measured by the Belle collaboration and the generator level acceptance of 35% as mentioned above. An additional 10% yield should be expected from the $Z \to \tau\mu\mu$ channel assuming the same acceptance and efficiency. Such a L1 efficiency would not be possible with the already existing triple-muon triggers because the L1 TTTracks will not be available in the GT.

3.2 The High Level Trigger and Data Acquisition

The HLT task is to further reduce the event rate to 1 kHz. To achieve this, all events that pass the L1 Trigger are sent to a computer farm of approximately 30k CPU cores, known as the Event Filter Farm. This is located in a dedicated room at the surface of the CMS cavern. The HLT has access to the full detector readout, including the tracker information, and runs a lighter version of the offline event reconstruction.

An important concept of the HLT data processing is the "path", which is a set of algorithms steps, running in predefined order, both for reconstruction of physics objects and for selection. Each path is a sequence of steps of increasing complexity. If an event is accepted by at least one path it is stored, otherwise it is discarded.

Events accepted by the HLT are transferred to another software process called the storage manager, under the supervision of the Data Acquisition system (DAQ). The data events are first stored locally on disk, into multiple primary datasets (PD). Each PD is fed by a number of logically coherent trigger paths. When operating, CMS records about 80 TB of data every day ($\sim 80 \times 10^3$ seconds per day $*$ 1 kHZ $*$ 1 MB the raw event size), which corresponds to approximately $\sim 80 \times 120 \sim 9.6$ PB per run period (one run period is almost 1/3 of the full calendar year). The recorded events are reconstructed with offline algorithms that will be presented in the next section. The post-processing phase results in data stored in NanoAOD format based on the ROOT framework [138]. A world wide computing farm with three-tier structure is used for the processing and storage of the data.

3.3 Physics Object Reconstruction

The CMS sub-detector components record digital information in the form of binary hits and energy deposits. These measurements are combined to form the signatures of stable particles and provide an estimate of their energy, momentum, trajectory and particle type. This is referred to as physics object reconstruction and it is important for detailed event reconstruction. The physics analysis that will be presented in Chapter 4 uses multiple physics objects such as muon, electrons, jets and MET induced from not detected neutrinos and possibly the SUSY particles in the final state. These physics objects are reconstructed using the Particle Flow (PF) algorithm [139]. The following section describes the PF algorithm steps and the physics objects reconstruction methods.

3.3.1 The Particle Flow Algorithm

The PF algorithm identifies and reconstructs all stable particles in an event. As stable, are characterized the particles that have low or zero probability to decay at the collision point, or in flight within the detector volume. These are further used to reconstruct composite objects such as jets, hadronically decaying taus, MET and interaction vertices. The number of reconstructed particles in each pp collision provides a full event description.

The PF algorithm is structured in three steps. Firstly, it takes as input *hits* from the tracking detectors, *energy deposits* from the calorimeters and *muon hits* from the muon systems, and it reconstructs tracks and energy deposit clusters which are collectively referred to as PF elements. In the second step, the algorithm performs spatial correlation of the 3 sub-detector findings with the link algorithm. Lastly, particle hypotheses are inferred and derived objects such as jets, MET and vertices are computed. The three steps of the PF algorithm are described in the following subsections.

Tracks and clustering of energy deposit

The inner tracker track reconstruction relies on the Combinatorial Track Finding (CTF) algorithm and it can be decomposed into 3 logical steps [140, 141]. Firstly, the seed generation is performed by combining pairs or triplets of pixel hits in neighboring layers. These combined hits are considered as potential tracks (also referred to as seeds). The second step is the track finding based on a KF patter recognition. The seeds formed in the previous step are extrapolated outwards, to the neighboring tracker layers. Compatible hits are assigned to the track. The track is updated with the new hits and its parameters are recomputed in every stage of the KF method. The updated tracks are kept or discarded based on a quality bit assigned by the algorithm. Ambiguities with tracks sharing hits are resolved in favour of the trajectories with the best quality. The third step of the CTF algorithm is the final track fit, in which the trajectories are refit to the full set of hits again with a KF. The algorithm estimates five track parameters: the distance on the transverse plane between the origin and the point of closest approach between the track and the beam axis (called impact point), d_0; the separation of the track from the collision point on the beam axis, z_0; the azimuthal angle of the track at the impact point, ϕ_0; the polar angle of the track θ; and the transverse momentum, p_T. The CTF algorithm is applied 3 or 4 times iteratively to improve the track reconstruction efficiency, and minimise the number of falsely reconstructed tracks. Each iteration of the CTF algorithm starts from hits that were not associated with the highest quality track in the previous iteration.

The energy clustering algorithm is performed separately in each sub-detector: ECAL barrel and endcaps, HCAL barrel and endcaps and the 2 preshower layers. In HF no clustering is performed, the EM and hadronic components of each cell are directly giving rise to HF clusters. The clusters are formed through grouping energy deposits in the calorimeter cells. The first step of the clustering algorithm is the seed generation, which starts from cells with the maximum energy deposit among their neighboring cells[2] that exceeds a threshold. Starting from the seeds, the clusters are formed by aggregating cells with at least a corner in common with a cell already in the cluster and an energy deposit that exceeds the expected noise by a factor of 2 [139].

The Link Algorithm

The link algorithm performs a spatial correlation of the PF elements (tracks and energy clusters) from different sub-detectors, which were reconstructed in the previous step. It considers only PF element pairs of the nearest neighbours on the (η, ϕ) plane. If two PF elements are linked, the algorithm defines a distance between them. The link algorithm produces PF blocks of elements, associated by a link through common elements.

[2] Cells located in proximity, sharing a side or a corner with the seeding cell

The tracks from the central tracker and the calorimeter energy deposit clusters are linked through the extrapolation of the last measured hit in the tracker, to the two layers of the preshower, the ECAL and the HCAL. The track is linked to a cluster if its extrapolated position is within the energy cluster area. The distance between the extrapolated track position and the cluster position in the (η, ϕ) plane is the link distance. Only the link with the smallest distance is kept.

A link between ECAL and HCAL energy deposit cluster is established when a cluster position in the more granular calorimeter (preshower or ECAL) is within the cluster envelope of the less granular calorimeter (ECAL or HCAL). The link distance is defined in a similar manner to that of the track-cluster link.

Muon detectors' hits are linked together to form standalone muons or they link to inner tracker tracks to form tracker or global muons. The muon reconstruction algorithms will be described in subsection 3.3.3.

The identification and reconstruction sequence in each PF block proceeds with the following order. Firstly, muon candidates are identified and reconstructed and the corresponding PF elements are removed from the PF block. The electron identification and reconstruction follows. Energetic and isolated photons are identified in the same step. The corresponding tracks and ECAL or preshower clusters are excluded from further consideration. When all blocks have been processed and all particles have been identified the global event description becomes available.

3.3.2 Vertices

Particles produced at the same place during the pp collisions, have their tracks originating from a common point called vertex. Every bunch crossing gives rise to a vertex at which the hard process of the event happens, called primary vertex (PV), and other vertices from PU interactions. Vertices which originate from tracks of unstable particles that decay in the detector due to their long life-time are called secondary vertices (SV).

The vertices are reconstructed from a set of available reconstructed tracks following a deterministic annealing algorithm approach [142, 143, 140]. The vertex with the highest track p_T^2 sum is associated to the PV. Additionally, the PV needs to lie within 24 cm in the z direction and 2 cm in transverse direction from the nominal interaction point, where the proton beams cross.

The distance of a track from the PV is measured in terms of the impact parameter (IP) along the z axis (d_z) or on the xy plane (d_{xy}). The 3-dimensional IP is estimated as $IP_{3D} = \sqrt{d_z^2 + d_{xy}^2}$. The significance of IP_{3D}, defined as the ratio of IP_{3D} over its uncertainty, $IP_{3D}/\Delta(IP_{3D})$, is an effective handle for the promptness of an object.

3.3.3 Muons

The muon reconstruction uses information from the inner tracker, which provides precise momentum resolution, and from the muon detectors [144]. High purity in the muon reconstruction is achieved, due to the calorimeters absorbing additional particles.

The reconstruction starts locally in each muon sub-detector, and then information from the tracker and the muon sub-detectors are combined to reconstruct the muon track. The high level muon physics object collection is composed of 3 types:

- *Standalone muon*: DT, CSC and RPC hits are reconstructed from the digitized signal. The DT and CSC hits are clustered to form segments (or 'stubs'), which are used for pattern recognition in the muon detectors. DT and CSC segments and RPC hits are gathered to reconstruct a standalone muon track.
- *Global muon*: each standalone muon is matched to an inner tracker track if the parameters of the two are compatible after propagation to a common surface.
- *Tracker muon*: each inner tracker track with $p_T > 0.5$ GeV is extrapolated to the muon stations. If at least one segment matches the track, a tracker muon track is reconstructed.

About 99% of the muons produced in the CMS acceptance, are reconstructed as global or tracker muons. If two reconstructed muons share the same inner tracker track they are merged into one single object.

The global muon reconstruction is more efficient for higher p_T muons that penetrate more than one muon stations. The tracker muon reconstruction algorithm is particularly useful for low p_T muons that may not leave enough hits in the muon chambers. However, punch through hadrons [3] may also be misreconstructed as tracker muons.

In the analysis carried out for this thesis, the muons are PF reconstructed as global or tracker muons.

3.3.4 Soft MVA Muon ID

The algorithms described in the previous subsection are combined to provide robust and efficient muon reconstruction [145]. A selection based on various muons identification variables is applied by the physics analyses in order to provide the

[3]Highly energetic charged hadrons overcoming HCAL and leaving hits in the innermost muon stations, mimicking low p_T muons

desired balance between identification efficiency and purity. The most common muon identification algorithms used by the CMS physics analyses are:

- *Loose muon ID*: global or tracker muon identified as muon by the PF event reconstruction. This ID is designed to be highly efficient for prompt muons and for muons arising from heavy and light quark decays.
- *Medium muon ID*: loose ID muon with additional track-quality and muon-quality requirements. This ID is designed to be highly efficient for prompt muons and for muons from heavy quark decays.
- *Tight muon ID*: global muon identified as muon in a PF event reconstruction. The candidate is required to have a χ^2/d.o.f of the track fit of less than 10, at least one muon chamber hit included in the global muon track fit and at least two muon segments from the muon stations to be matched with the tracker track in order to suppress hadronic punch-through and muon decays in flight. In addition the tracker track must have at least 10 inner-tracker hits, a $|d_{xy}|$ of less than 2 mm and a d_z of less than 5 mm. With these cuts the rate of muons from decays in flight or from PU is significantly reduced. The tight muon ID is widely used by the CMS physics analyses for prompt muon identification.
- *Soft muon ID*: global muon identified as muon in a PF event reconstruction. The tracker track must have at least 5 tracker layers with hits to guarantee a good p_T measurement and at least one measurement in a pixel layer, in order to suppress muons from decays in flight. Bad quality tracks are rejected and a loose compatibility with the PV is required by $|d_{xy}| < 0.3$ cm and $d_z < 20$ cm. The soft muon ID is optimal for low p_T (< 10 GeV/c) muons and the selection is used mostly in B-physics analyses in CMS.

During the PhD work of this thesis, a new soft multivariate (MVA) muon ID for Run 3 was developed. The goal of the soft MVA muon ID is to use multivariate methods for the classification of signal and background muons. It is aiming at reducing the PU rate and retaining high signal identification efficiency.

A soft MVA muon ID is developed from the CMS collaboration specifically for the needs of the analysis targeting the $B_s\mu\mu$ decays. This soft MVA ID, hereafter denoted as B_s soft MVA ID, aims at reducing the fake rate from hadrons for low p_{T} muons originating from B hadron. The full description of the pre-selection and the training of the boosted decision tree (BDT) are described in the internal CMS analysis note AN-2016/178. It should be noted that the B_s soft MVA ID was developed using muons originating from B_s decays as signal and applying a tight pre-selection on the reconstructed training muons. Namely, the particles are required to be global muons with a χ^2/d.o.f of the track fit of less than 10, at least one valid muon hit, at least one muon segments from the muon stations to be matched with the tracker track, at least one valid hit in the pixel tracker, at least 5 tracker layers with me measurements, and high-purity tracks. The B_s soft MVA ID brought a 50% reduction in the identification of background muons as signal with a 10% reduction in the signal identification efficiency. However, it is desirable to develop a more general soft MVA ID for the muons, without applying the tight preselection requirements, such that it can be used more widely by the analyses in the collaboration.

For the purpose of the current study a JPsiToMuMu Monte Carlo, generated with the PYTHIA8 generator [114] and simulated with the 2018 conditions was used. The generator level muons have p_T within the range of [0, 100] GeV.

A very loose pre-selection on the reconstructed muons is applied. Specifically, the reconstructed muons are required to be global or tracker muons within the CMS acceptance of η [-2.4, 2.4]. Additionally, the barrel muons are required to have $p_T > 3.5$ and the muons in the endcaps $p_T > 2$ GeV in order to reject very soft background muons from PU. An upper bound at 100 GeV is applied on the reconstructed muons, as the soft muon MVA ID is not targeted for the identification of such high p_T leptons. Finally, the track is required to pass the high purity requirement in order to reject tracks with bad quality.

A random forest and a gradient boosting classifier were trained, optimized and tested. The latter machine learning models are described in detail in Sec. 6.6. In the end, the one with the highest area under the ROC curve was used for the object classification.

The two classifiers were trained with signal muons from the J/ψ decays and background muons from PU and punch-through.

The low level variables that are used for the cut-based definition of the medium, tight and soft muon ID were used as the training input variables for the machine learning models. The list of the training variables is shown below:

- p_{T}
- η
- χ^2/d.o.f of track fit.
- Tracker-Standalone position match.
- Track kink: the extrapolated states from the two halves of the tracker tracks are compared to get a χ^2 at each tracker layer.
- Segment compatibility: a weight is assigned if a station has been crossed and penalize if a station did not have a matching segment. It depends on the "depth" of the muon passage. Station weight is reduced for stations with badly matched segments.

- Number of valid hits in μ chambers.
- Number of μ segments in μ stations.
- Number of hits in pixel tracker.
- Number of tracker layers with hits.
- Fraction of valid tracker hits.
- Tracker track matched with at least one muon segment.
- Number of pixel layers with hits.

The distributions of the training variables are shown in Fig. 3.14 for the signal and the background, normalized to the same integral.

The background distributions presented in Fig. 3.14 are p_{T} and η reweighted in order to avoid exploiting the differences in the p_{T} and η distributions of the signal and background in the training of the model. The weighting factors are computed by dividing the number of signal and background objects in p_{T} and η bins and applied as sample weights on the background objects. The 2D plots in Fig. 3.15 show the number of signal and background objects in p_{T} and η bins. The η range is divided into 4 bins and a variable binning is applied in p_{T}. Namely, bins of 2 GeV for the p_{T} range of [2, 20] GeV, bins of 15 GeV for the p_{T} range of (20,50] and bins of 25 GeV for the p_{T} range of (50, 100] GeV. The above binning scheme is applied in order to avoid empty bins.

The total of signal and background objects after the preselection are $\sim 500 \times 10^3$ and $\sim 80 \times 10^3$ respectively. From those objects the 60% is used for training and the 40% is used for the testing of the classifier.

A random forest with maximum depth of 16 [4], minimum samples leaf [5] of 3, minimum samples split [6] of 6 and 250 estimators was trained. Additionally, a gradient boosting classifier with learning rate [7] of 0.3, maximum depth [8] equal to 7, minimum samples split of 5 and 69 estimators was trained. The parameters of the models also referred to as hyperparameters, were chosen after a random search optimization. The definitions of the hyperparameters mentioned above are also explained in detail in Subsection 6.6.2.

The plot in Fig. 3.16 shows the ROC curves of the random forest and the gradient boosting classifiers. The ROC curves depict the performance of a algorithms for multiple thresholds on the classifiers output. The x axis shows the

[4] The maximum depth of the individual trees limits the number of nodes in it

[5] The minimum training data point or samples required at every leaf node for a node splitting to be considered

[6] The minimum number of training data points required for an intermediate node split to be performed

[7] The step size for every iteration while moving towards the minimization of the loss function

[8] The maximum depth of the individual trees which limits the number of nodes in them

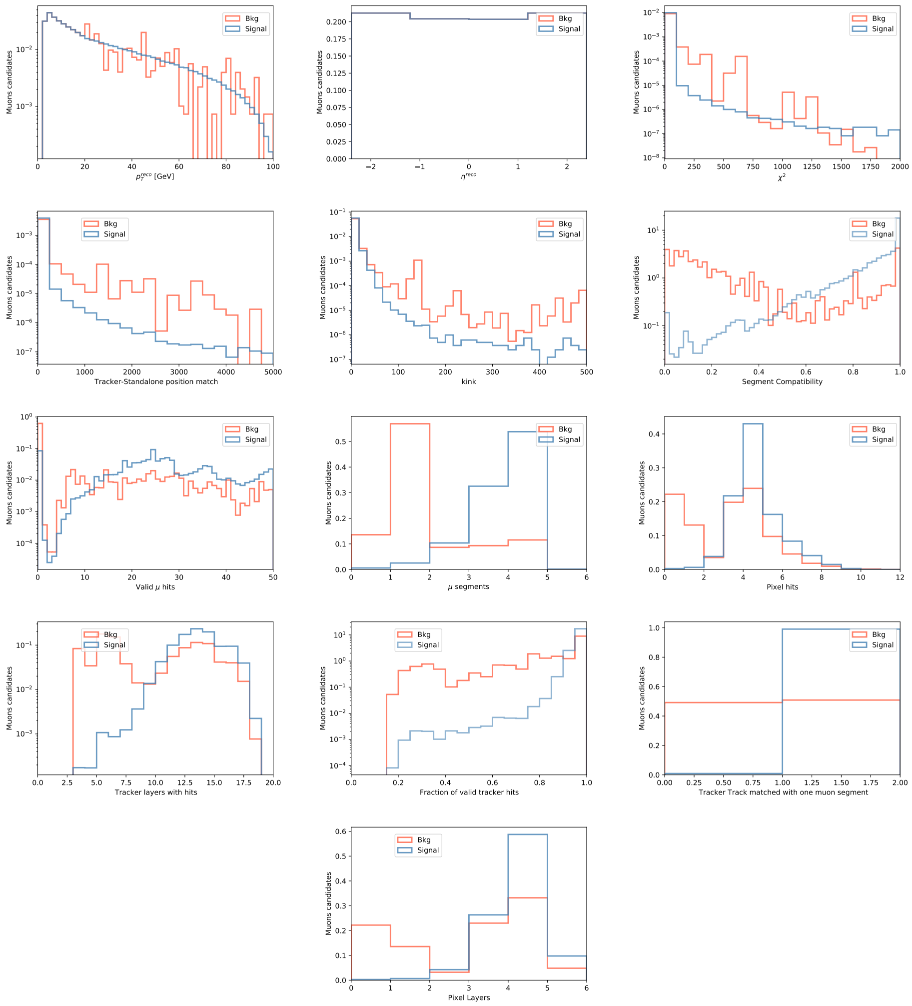

Figure 3.14: The training variables used for the development of the soft MVA ID. The plots in the first row show p_{T} (left), η (middle), χ^2/d.o.f of track fit (right). The plots in the second row show the Tracker-Standalone position match (left), track kink (middle), segment compatibility (right). The plots in the third row show the number of valid muon hits in the muon chambers (right), the number of muon segments in the muon chambers (middle), the number of hits in the pixel tracker (right). The plots in the fourth row show the number of tracker layers with hits (left), the fraction of valid tracker layers (middle), the boolean of the tracker track being matched with one muon segment (right). The plot in the bottom row shows the number of pixel layers with hits.

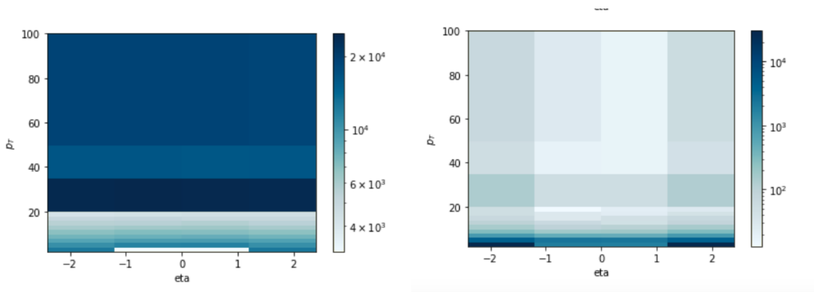

Figure 3.15: Signal (right) and background (left) objects in p_{T} and η bins. The 2D plots are used to calculate the p_{T}-η reweighting.

false positive rate which is the proportion of signal muons identified as background muons. The y axis shows the true positive rate which is the proportion of signal objects identified as such. The blue dot represents the true positive rate and the false positive rate of the cut-based soft muon ID.

The true positive rate is estimated as the number of signal objects passing the ID selection over the total signal objects. The false positive rate is estimated as the background objects passing the ID selection over the total background objects.

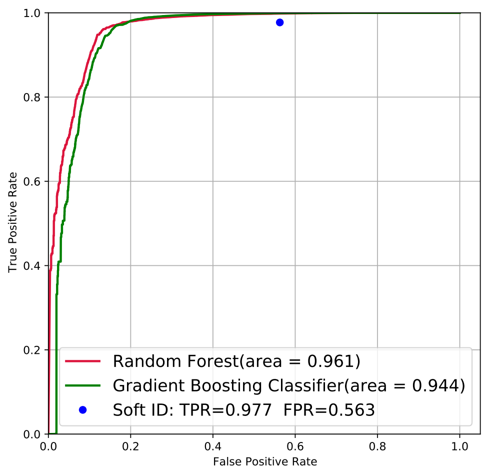

Figure 3.16: The ROC curves of the random forest and the gradient boosting classifier. The blue point is the true positive rate and the false positive rate of the cut-based soft muon ID.

The random forest has area under ROC curve equal to 0.96 and the gradient boosting classifier has area under ROC curve equal tot 0.94. Therefore the random forest is chosen as the most efficient classifier and will be used for the final object classification.

In order to validate the random forest training the confusion matrix is evaluated on the testing sample. The confusion matrix presented in Fig. 3.17 is a summary of the prediction results of the classification. It shows the rate at which background objects are identified correctly in the top diagonal box and the true positive rate in the bottom diagonal box. The off-diagonal boxes show the false positive rate (top row) and the rate at which signal objects are identified as background, also referred to as false negative rate (bottom row). Therefore a confusion matrix of a healthy training should have high values close to unit in its diagonal, and low values close to 0 in its off diagonal. From the matrix in Fig. 3.17 it is concluded that the random forest predicts correctly the signal and background objects with high rates, while the rate it misidentifies an object is low.

The right plot in Fig. 3.17 shows the comparison between training and testing sample on the random forest's output distribution. The good agreement between the training and the testing distribution validates a robust training with no overfitting. The model is able to reproduce with the testing sample, the distribution of the training sample.

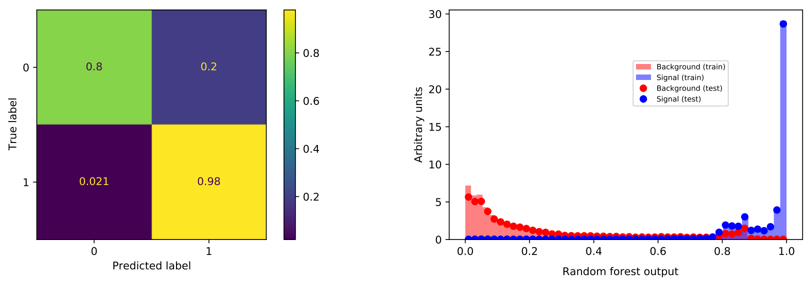

Figure 3.17: The random forest confusion matrix calculated on the testing sample (left). The random forest output distribution on the training and the testing sample (right).

The most important variables for the training of the random forest are the segment compatibility, the fraction of the valid tracker hits and the number of valid hits in the muon chambers. This is to be expected as the distributions of the signal and background in those variables show significant discrimination in the plots in Fig. 3.14.

The curves in Fig. 3.18 show the true positive rate and the false positive rate for multiple thresholds of the random forest output. The green dashed line represents the working point with true positive rate equal to that of the cut-based soft muon ID. The purple dashed line shows the working point with false positive rate equal to that of the cut-based soft muon ID. The thresholds values of the two working points and the true positive and false negative rates are summarized in Tab. 3.1.

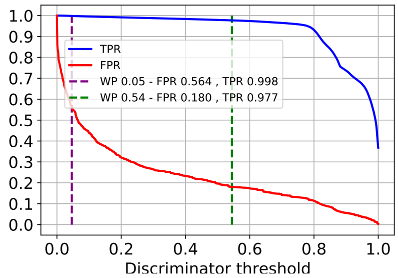

Figure 3.18: True positive rate (TPR) and false positive rate (FPR) for multiple thresholds of random forest output. The green dashed line is the working point with true positive rate equal to that of the cut-based soft muon ID. The purple dashed line is the working point with false positive rate equal to that of the cut-based soft muon ID.

Working Point	True Positive Rate	False Positive Rate
0.05	0.99	0.56
0.54	0.97	0.18

Table 3.1: Working points, true positive rate and false positive rate of the random forest classifier.

The 0.54 working point is chosen as a threshold for the classifier's output. This working point has a 2% lower true positive and significantly lower false positive rate than the 0.05 working point.

The true positive rate and false positive rate were estimated in p_T and η bins for the 0.54 working point. The Fig. 3.19 presents the comparison of the true positive rate and the false positive rate of the cut-based and the MVA soft muon IDs.

For the very low values of muon p_T (<6 GeV) the soft MVA muon ID has 20-50% lower false positive rate with a maximum reduction in the signal efficiency of 25% at the very low p_T bins. The false positive and true positive rates of the two muon IDs is similar at higher p_T. The soft muon MVA ID reduces significantly the rate at which background soft muons are identified as signal with a small reduction in the signal efficiency.

3.3.5 Electrons and Photons

The electron PF reconstruction, combines inner tracker tracks and ECAL energy deposit clusters. There are 2 sources of identification inefficiency for electrons. Firstly, the charged particles tracks in the tracker and the ECAL energy cluster

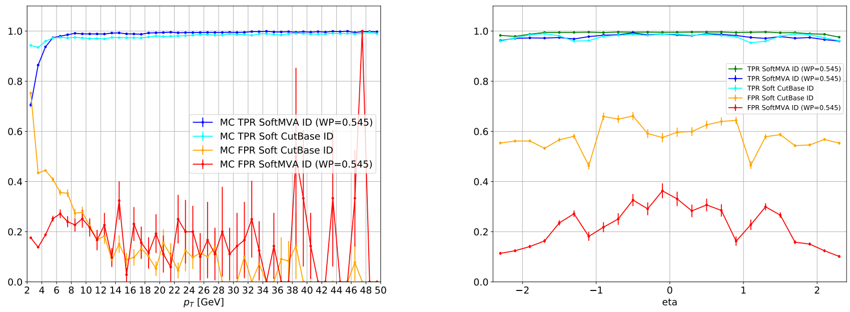

Figure 3.19: The true positive rate and the false positive rate of the cut-base and the MVA soft muon IDs. In p_{T} (left) and η (right) bins.

can also be produced by other types of particles. Secondly, the thick tracker material can cause soft electrons to lose significant fraction of their energy through bremsstrahlung before reaching ECAL. Due to the strong magnetic field, the energy loss can lead to kinks in the trajectory of the electron. This can result in failed matching between the track and the ECAL energy deposits, or electron absorption before showering.

For a reliable electron reconstruction, all the radiated energy should be gathered and combined with the electron's energy into superclusters (SC), along the ϕ direction. The PF electron track reconstruction, is seeded from the results of the KF track finding procedure. In the case of bremsstrahlung radiation, the pattern recognition may be inefficient, therefore a preselection based on the number of hits and the fit χ^2 is applied and the selected tracks are re-fit with a Gaussian-sum filter (GSF). The PF electron candidates are reconstructed by linking SC to the GSF tracks. An ECAL-based electron candidate is built when the linking starts from the SC and extrapolates to the GSF track. A tracker-based electron is reconstructed when the linking starts from the GSF track towards the ECAL SC. The PF electron momentum is estimated by combining the tracker and the ECAL SC observables, while the charge assigned to the track is estimated by the sign of the GSF track curvature.

Isolated photons are reconstructed in the PF algorithm together with the electrons. The reconstruction is seeded from ECAL SC with $p_T > 10$ GeV and no GSF track should be associated to the energy cluster.

3.3.6 Jets

The hadronic jets are signatures of quarks and gluons produced in the parton hard scatterings during the pp collisions. It can be seen as a shower of collimated particles and it can be reconstructed by clustering particles together to form geometrical cones. The jets in CMS are reconstructed with an anti-k_T algorithm [146] implemented in the FASTJET package [147]. There are 3 possible ways of jet reconstruction:

- Calorimeter jets are reconstructed by energy deposit sums in the HCAL and geometrically corresponding ECAL towers;
- Jet-plus-Track jets are reconstructed by calorimeter jets with improved energy response and resolution by incorporating inner tracker information;
- PF jet is reconstructed by clustering all the reconstructed PF candidates

The 4-momentum of the jet is obtained by summing the momentum of its constituents. The reconstructed jets need to be calibrated in order to have the correct energy scale. This is achieved by applying the Jet Energy Corrections (JEC). JEC corrects for the energy offsets coming from PU, the detector response to hadrons and residual differences between data and simulation as a function of η and p_T [148].

3.3.7 Missing Transverse Energy

As described in Section 1.7, stable SUSY particles or neutrinos produced during the pp collisions, do not interact with the detector material and therefore they escape detection. Their trace can be inferred from the transverse momentum imbalance of all visible particles in the final state. This is denoted as MET. The raw MET is defined as the opposite of the vectorial sum of the transverse momenta of all the visible particles and jets of the event

$$\vec{E}_T^{miss}(raw) = -\sum_{i=1}^{N_{visible}} \vec{p}_T^{\,i} \tag{3.4}$$

However, the raw MET is systematically different from the true MET. In order to account for this, the raw MET is corrected. In the analysis carried out in this thesis, type-I correction is applied on MET. This correction propagates the JEC to the MET calculation. Therefore the type-I corrected MET is defined as

$$\vec{E}_T^{miss}(type-I\ corr) = \vec{E}_T^{miss}(raw) + \sum_{i}^{jets} \vec{p}_T^{\,i} - \sum_{i} \vec{p}_T^{\,iJEC} \tag{3.5}$$

3.3.8 b Jet Tagging

Jets arising from gluons, up, down or strange quarks are referred to as light-flavour jets, while those originating from bottom or charm quarks are called heavy-flavour jets. The efficient identification of the mother quark or gluon of the jet is very important for the physics searches conducted at CMS. Tagging algorithms based on multivariate analysis (MVA) approaches, have been developed for jet identification in CMS. The output of the MVA algorithm is a metric of the probability of the jet in question to be of a certain kind.

In the scope of this thesis, only the b-jet tagging is used. The lifetime of the b-flavoured hadrons is approximately 1.5 ps and thus their decays lead to displaced tracks which reconstruct a SV [149]. The displacement of the track with respect to the PV is characterized by the IP described in 3.3.2. Additionally, b-jets have larger mass than the light jets and almost 20% of their decays include a muon or an electron.

In the physics analysis of the current thesis, b-jets are identified by the medium working point (WP) of the DeepCSV tagging algorithm. This MVA algorithm uses a deep neural network trained with 50 million jets. The deep neural network is trained using as input variables the IP of the tracks, properties of the reconstructed vertices and the absence or presence of leptons. The DeepCSV b tagging efficiency at the medium WP is 68% and the misidentification probability is 1% [149].

Chapter 4

Physics Data Analysis

Physics analyses study events with different particle content and kinematic properties at their final states, called "signatures". There can be multiple physics processes leading to the same signatures. The number and the identities of the final state particles are referred to as "topology". A SUSY physics analysis falls in the class of the so-called searches for new physics.

A search typically starts from applying selection requirements on the recorded events and on their objects in order to increase the purity of the sample in the process of interest. The selection starts from the trigger level and more complex cuts are applied progressively offline. In a signature-specific SUSY search, such as the one presented here, a signal model of the new physics signature is adopted. This is in contrast to the model-independent searches, where events with a given topology are explored for deviations from the SM predictions, without adopting any model for such deviations, e.g. the production and decay rate of a new physics process.

Usually, SM processes have the same signature with that of the signal model and therefore they are considered as the background of the search. The contribution of the background processes is predicted with dedicated estimation methods, which should be validated. An overprediction or underprediction of the background expectation in a SUSY search can lead to a false negative or false positive result, respectively. Additionally, kinematic regions where the signal contribution is expected to be maximum, are defined and called "Search Regions" (SR). The expected number of events from signal and background processes is called "yields" and they are subject to statistical and systematic uncertainties. The assessment of the systematic uncertainties of the search is a key component of every physics analysis.

This chapter presents a SUSY search conducted with the full Run 2 data collected by CMS. The search targets signatures of electroweakino pair production and decay to final states with soft leptons (two or three) and MET. Section 4.1 is an introduction to the search and the upgrades with respect to the previous iteration of the analysis which was conducted with the 2016 data. Sec. 4.2

describes the signal models used for the interpretations of the results, Sec. 4.3, 4.4, 4.5 presents the data set and simulation samples, the triggers and the offline object definition and event reconstruction adopted for the analysis. In Sec. 4.6 the event selection for maximum signal acceptance and SM background rejection in the SR is highlighted and Sec. 4.7 elaborates on the methods used for the SM background prediction. The systematic uncertainties are presented in Sec. 4.8.

4.1 Introduction to compressed mass SUSY search with soft leptons and MET with the CMS Run 2 data

The extension of the 2016 analysis of electroweakino pair production and decay to soft leptons and MET, to the full Run 2 CMS data and total integrated luminosity up to 137 fb^{-1} is the main focus of this thesis. The ATLAS collaboration has published similar results in events with two opposite sign soft leptons, using the full data-set collected during 2016-2018 LHC running period [150].

Major upgrades have been incorporated in the analysis strategy compared to the 2016 analysis. The most important are:

- The soft 3ℓ final state is added in the search signatures and is looked for together with the 2ℓ final states. The 3ℓ final state is expected to add to the sensitivity of the search in higher $\Delta m_{\tilde{\chi}^0_2\text{-}\tilde{\chi}^0_1}$ due to acceptance effects.
- The lower selection boundary on the 2ℓ invariant mass was relaxed from 4 to 1 GeV. Re-optimized signal selection is applied for higher sensitivity to the lower $\Delta m_{\tilde{\chi}^0_2\text{-}\tilde{\chi}^0_1}$ values. The extension to lower $\Delta m_{\tilde{\chi}^0_2\text{-}\tilde{\chi}^0_1}$ was possible due to simulated samples extending to low $M(\ell\ell)$ becoming available.
- Improvements on the non-prompt background prediction, which is the dominant background of the search, are employed
- Refinements on the signal modelling

The full Run 2 data analysis is targeting the production of $\tilde{\chi}^0_1\tilde{\chi}^0_2$ or $\tilde{\chi}^\pm_1\tilde{\chi}^0_2$ and their decay to 2 or 3 soft leptons via a Z^* boson (and the W^* boson in the 3ℓ final state). The Feynamn diagrams of the processes are presented in Figures 4.1 and 4.2. At least one pair of muons or electrons or muon-electron with opposite sign is required in the final state, depending on the signal model, together with at least one jet which in the signal arises from initial state radiation and boosts the final state, inducing moderate to significant missing transverse energy.

The SR are defined such that the maximum possible signal acceptance and lower background contamination are achieved. The SM background is estimated in dedicated control region presented in Sec. 4.7. The systematic uncertainties, described in Section 4.8, affect the signal and background prediction and are incorporated as nuisance parameters in the fit. The final results are extracted with a binned maximum likelihood fit of the expected signal and background from all the SR and the control regions to data. A hypothesis test is performed with the frequentist approach that employs the CLs prescription and upper limits at 95% CL on the expected and observed production cross section are set [151, 152, 153, 154]. The statistical methods, the results and the conclusions of the search are presented in the next chapter (Chapter. 5).

4.2 Signal models

The analysis is probing the direct pair production of electroweakinos $\tilde{\chi}_1^{\pm}$ $\tilde{\chi}_2^0$ or $\tilde{\chi}_1^0$ $\tilde{\chi}_2^0$. In R-parity conserving SUSY, the lightest neutralino $\tilde{\chi}_1^0$ is the lightest supersymmetric particle. The heavier electroweakinos can decay to soft leptons via off-shell bosons. The final state of the decay will contain the soft leptons and missing transverse energy from the LSPs (and the neutrinos when W^* also decays leptonically in the $\tilde{\chi}_1^{\pm}\tilde{\chi}_2^0$ case) as shown in Figure 4.1.

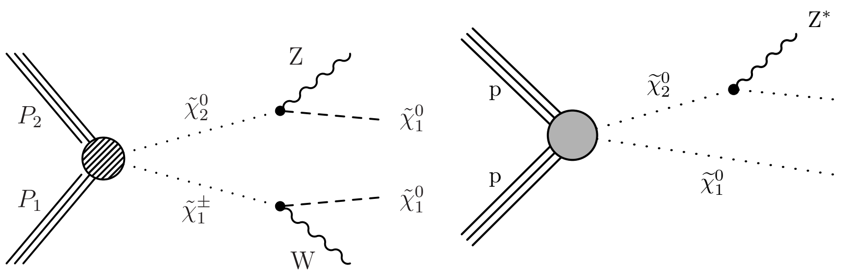

Figure 4.1: Diagrams for electroweakino production in the case of $\tilde{\chi}_2^0\tilde{\chi}_1^{\pm}$ which decay as $\tilde{\chi}_2^0 \rightarrow Z^*\tilde{\chi}_1^0$ and $\tilde{\chi}_1^{\pm} \rightarrow W\tilde{\chi}_1^0$ (left) and the $\tilde{\chi}_2^0$ $\tilde{\chi}_1^0$ production where $\tilde{\chi}_2^0 \rightarrow Z^*\tilde{\chi}_1^0$ (right).

The electroweakino mass eigenstates are a mixture of wino bino and higgsino, as described in Section 1.6. In the MSSM SUSY model the masses of the electroweakinos are parametrized in terms of the bino, wino and higgsino mass parameters M_1, M_2, μ. Different scenarios for the hierarchy of the mass parameters are considered for the interpretation of the results.

- In the Higgsino scenario $\mu << M_1$, M_2 is assumed. In this case $\tilde{\chi}_1^0$, $\tilde{\chi}_2^0$, $\tilde{\chi}_1^{\pm}$ are a triplet of higgsino-like light compressed states. The mass splitting between the NLSP and the LSP is approximated by the $m_W^2/min(M_1, M_2)$. These scenarios are theoretically motivated by natural SUSY in which μ is near the electroweak scale while M_1 and M_2 can be larger.
- In the Wino/Bino scenario $M_1 < M_2 << \mu$ is assumed. In this scenario $\tilde{\chi}_2^0$ and $\tilde{\chi}_1^{\pm}$ are wino dominated while $\tilde{\chi}_1^0$ is bino dominated, therefore the LSP

mass is $M_{\widetilde{\chi}_1^0} \approx M_1$ and the NLSPs are mass degenerate with $m_{\widetilde{\chi}_2^0}, m_{\widetilde{\chi}_1^\pm} \approx M_2$. In this scenario a bino-like LSP is a good thermal relic DM candidate, depleted from early Universe through coannihlation processes and matches the observed relic DM density [155]. This scenario has typically larger cross section than the higgsino scenario.

The Higgsino and the TChiWZ simplified models are assumed for the interpretation of the Higgsino and Wino/Bino scenarios results respectively [156, 157]. These simplified models extend the SM with $\widetilde{\chi}_1^0$, $\widetilde{\chi}_2^0$, $\widetilde{\chi}_1^\pm$ states. The branching fractions of $\mathcal{B}(\widetilde{\chi}_2^0 \to Z^*\widetilde{\chi}_1^0) = 100\%$ and $\mathcal{B}(\widetilde{\chi}_1^\pm \to W^*\widetilde{\chi}_1^0) = 100\%$ are assumed and the mass of the new states are free parameters of the model. Additionally, a phenomenological MSSM (pMSSM) model has been used for the interpretation of the Higgsino scenario results.

The model of light top squark pair production and decay either via a 4-body decay or via $b\widetilde{\chi}_1^\pm \to W^*\widetilde{\chi}_1^0 \to \ell\nu\widetilde{\chi}_1^0$ to 2 fermions and LSPs is considered as an additional interpretation of the result. The Feynman diagrams of the decays are shown in Figure 4.2. The W^* decay to two soft leptons which are not expected to be of the same flavor, while the b quarks are not sufficiently energetic and thus cannot be reconstructed. The T2bff and and T2bW simplified models are used for the interpretation of the two decay modes [158].

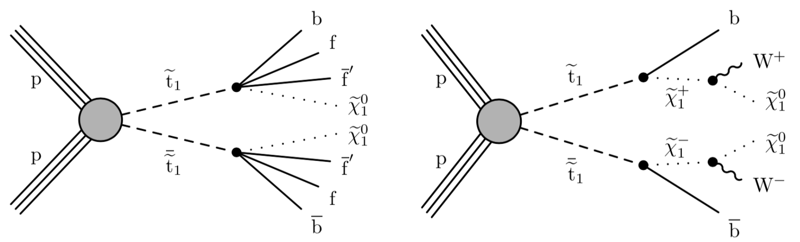

Figure 4.2: Diagrams of the top squark pair production and decay via 4-body decay (left) and via $b\widetilde{\chi}_1^\pm$ and further $\widetilde{\chi}_1^\pm$ decay as $\widetilde{\chi}_1^\pm \to W^*\widetilde{\chi}_1^0$ (right).

4.2.1 Higgsino and Wino/Bino simplified models

In the Wino/Bino scenario the $\widetilde{\chi}_2^0$ and $\widetilde{\chi}_1^\pm$ eigenstates are dominantly Wino-like. The $m_{\widetilde{\chi}_1^\pm} = m_{\widetilde{\chi}_2^0}$ is automatically satisfied as each mass state is given by the Wino mass parameter M_2 and $\widetilde{\chi}_1^0$ is Bino-like. The TChiWZ model that is used for the interpretation of the Wino/Bino scenario, assumes $m_{\widetilde{\chi}_1^\pm} = m_{\widetilde{\chi}_2^0}$. In the Higgsino simplified model the assumption on relation between the chargino and the two neutralino masses is $m_{\widetilde{\chi}_1^\pm} = \frac{1}{2}(m_{\widetilde{\chi}_2^0} - m_{\widetilde{\chi}_1^0})$.

In the TChiWZ model the $\widetilde{\chi}_2^0\widetilde{\chi}_1^\pm$ processes are considered, while in the Higgsino simplified model both $\widetilde{\chi}_2^0\widetilde{\chi}_1^\pm$ and $\widetilde{\chi}_2^0\widetilde{\chi}_1^0$ productions are assumed. Each process of the Higgsino model is included in different sample. The $\widetilde{\chi}_2^0\widetilde{\chi}_1^\pm$ sample will be referred to as "N2C1 Higgsino" and the $\widetilde{\chi}_2^0\widetilde{\chi}_1^0$ will be referred to as "N2N1 Higgsino".

The production cross section and decay modes depend on the composition of the electroweakinos in each model. In the Wino/Bino scenario the produced $\widetilde{\chi}_1^\pm$, $\widetilde{\chi}_2^0$ are dominantly Wino and the production cross section is computed accordingly with reference values from [159]. The production cross section for the Wino/Bino model is computed at NLO precision with respect to the coupling constant, for the matrix element, plus next-to-leading-log (NLL) precision, for the soft gluon resummation of the initial state, assuming mass degenerate Wino $\widetilde{\chi}_2^0$, $\widetilde{\chi}_1^\pm$ and light Bino $\widetilde{\chi}_1^0$ and all the other SUSY particles are assumed to be heavy and decoupled [160, 161]. As an example, the production cross section of $\widetilde{\chi}_2^0\,\widetilde{\chi}_1^\pm$ for $m_{\widetilde{\chi}_2^0} = m_{\widetilde{\chi}_1^\pm} = 150\,\text{GeV}$ is 5.2 pb. The production cross sections for the N2N1 and N2C1 Higgsino models are computed at NLO plus NLL precision in a limit of mass-degenerate higgsino $\widetilde{\chi}_2^0$, $\widetilde{\chi}_1^\pm$, and all the other sparticles assumed to be heavy and decoupled. The production cross section in the Higgsino scenario for the sparticle mass of 150 GeV is 1.2 pb, almost 4 times lower than in the Wino/Bino scenario. A grid of signal samples is generated with varying $m_{\widetilde{\chi}_2^0}$ and $m_{\widetilde{\chi}_1^0}$, with $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0} \in [1,\ 3,\ 5,\ 7.5,\ 10,\ 15,\ 20,\ 30,\ 40,\ 50]\,\text{GeV}$ for $m_{\widetilde{\chi}_2^0}$ in the range of 100–300 GeV.

Signal re-weighting factors

The $\mathcal{B}(\widetilde{\chi}_2^0 \to Z^*\widetilde{\chi}_1^0)$ and $\mathcal{B}(\widetilde{\chi}_2^0 \to W^*\widetilde{\chi}_1^0)$ are fixed at 100% in the Wino/Bino and Higgsino models. In the signal samples used in the analysis, Z^* always decay leptonically, while there is no restriction for the W^* decay. The branching ratio of the $\mathcal{B}(Z^* \to ll)$ and $\mathcal{B}(W^* \to \ell\nu)$ depend on the invariant mass of the off-shell vector bosons. For the mass splittings considered in the analysis ($\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0} <$ 50 GeV), individual decay modes of W^*/Z^* are subject to phase space suppression due to massive decay products. The branching ratios of the Z^* and W^* leptonic decay modes are calculated as a function of the mass splitting of the NLSP and the LSP, for both the electroweakino simplified models, where the $\widetilde{\chi}_2^0$ and $\widetilde{\chi}_1^\pm$ decays are computed with PYTHIA, and with SDecay module of SUSYHIT 1.5a which provides a coherent calculation of the decays including higher order corrections and taking into account second and third generation fermion masses [162]. The effect of the loop-induced radiative correction on the $\widetilde{\chi}_2^0$ decay were found to be negligible in the $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0}$ range of the analysis, and therefore the $\widetilde{\chi}_2^0 \to \widetilde{\chi}_1^0\gamma$ decays are vetoed in the branching fraction calculation. The SUSYHIT and PYTHIA results of the $\widetilde{\chi}_1^\pm \to W^*\widetilde{\chi}_1^0$ and $\widetilde{\chi}_2^0 \to Z^*\widetilde{\chi}_1^0$ branching fractions for the various decay modes of the W^* and the Z^*, respectively, are shown in Figure 4.3 and are used to derive appropriate per event re-weighting factors that are applied on the TChiWZ and Higgsino events. For the case of W^* where the hadronic channel is open, this is taken under consideration in the re-weighting factor estimation. The impact of the re-weighting on the signal acceptance is of the order of $\sim 10\%$. It should be noted that the analysis sensitivity extends down to $\Delta m_{\tilde{\chi}_2^0\text{-}\tilde{\chi}_1^0} = 3$ GeV.

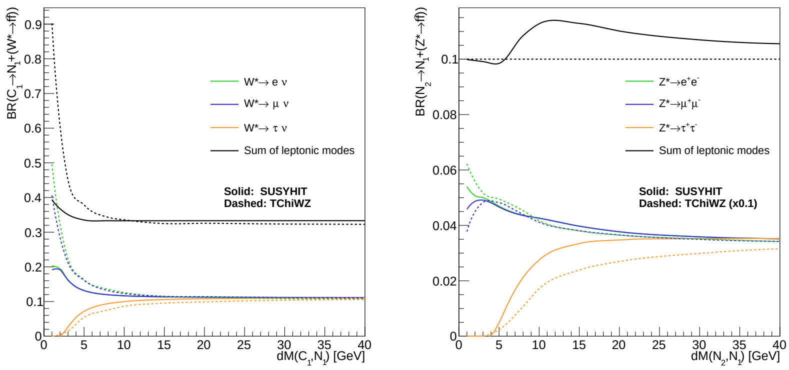

Figure 4.3: The branching ratio of $\tilde{\chi}_1^{\pm}$ decays as a function of $\Delta m_{\tilde{\chi}_2^0\text{-}\tilde{\chi}_1^0}$ (left) and the branching ratio of $\tilde{\chi}_2^0$ decays as a function of $\Delta m_{\tilde{\chi}_2^0\text{-}\tilde{\chi}_1^0}$ (right). The solid line shows the branching ratio computed with SUSYHIT and the dashed lines the branching ratios computed with PYTHIA.

A key handle for electorweakino signal discrimination from SM background is the invariant mass of the 2 soft leptons with same flavors and opposite signs. The observable $m_{\ell\ell}$ has a kinematic endpoint at the splitting value of the NLSP and LSP mass eigenvalues ($m_{\tilde{\chi}_2^0} - m_{\tilde{\chi}_1^0}$) and it is used in the final fit for the limit calculation as it will be described in Section 5.2.

In the simplified model the $\tilde{\chi}_2^0$ decay is done by PYTHIA according to pure phase space. In the full MSSM model the differential decay rate is given by

$$\frac{d\Gamma_{\tilde{\chi}_2^0 \to \tilde{\chi}_1^0 \ell\bar{\ell}}}{dm} = C m_{\ell\ell} \frac{\sqrt{m_{\ell\ell}^4 - m_{\ell\ell}^2(\mu^2 + M^2) + (\mu M)^2}}{(m_{\ell\ell}^2 - m_Z^2)} [-2m_{\ell\ell}^4 + m_{\ell\ell}^2(2M^2+\mu^2)+(\mu M)^2] \tag{4.1}$$

In the formula C is the normalization constant, $M(\ell\ell)$ is the 2ℓ invariant mass, $\mu = m_{\tilde{\chi}_2^0} - m_{\tilde{\chi}_1^0}$ and $M = m_{\tilde{\chi}_2^0} + m_{\tilde{\chi}_1^0}$ and $m_{\tilde{\chi}_1^0}$, $m_{\tilde{\chi}_2^0}$ are the signed eigenvalues of the NLSP and the LSP neutralinos [163]. The signed mass eigenvalues enter the expression only through squares of sums and differences and thus the expression depends on the sign of $m_{\tilde{\chi}_1^0} \times m_{\tilde{\chi}_2^0}$

The distribution of the 2ℓ invariant mass depends on the relative sign of the signed mass eigenvalues of $\tilde{\chi}_1^0$ and $\tilde{\chi}_2^0$. The product $m_{\tilde{\chi}_2^0} \times m_{\tilde{\chi}_1^0}$ can be positive or negative in the Wino/Bino scenario and negative in the Higgsino scenario. Figure 4.4 shows the differential decay rate for the scenarios of $m_{\tilde{\chi}_2^0} \times m_{\tilde{\chi}_1^0} > 0$ in blue and $m_{\tilde{\chi}_2^0} \times m_{\tilde{\chi}_1^0} < 0$ in red and the case where the matrix element is assumed flat in magenta. The latter represents the default simulated sample used in the analysis and it is corrected by a re-weighting factor for each scenario. The plot on the left shows the distributions of theoretical calculations (line) compared to simulations where the two scenarios of $m_{\widetilde{\chi}_2^0} \times m_{\widetilde{\chi}_1^0}$ signs are considered. The $M(\ell\ell)$ distribution of the pure phase space (PYTHIA) decay is re-weighted by the theory lineshapes of the two scenarios. The right plot shows the good agreement between the phase space only simulation re-weighted compared to simulation which accounts for the sign of $m_{\widetilde{\chi}_2^0} \times m_{\widetilde{\chi}_1^0}$.

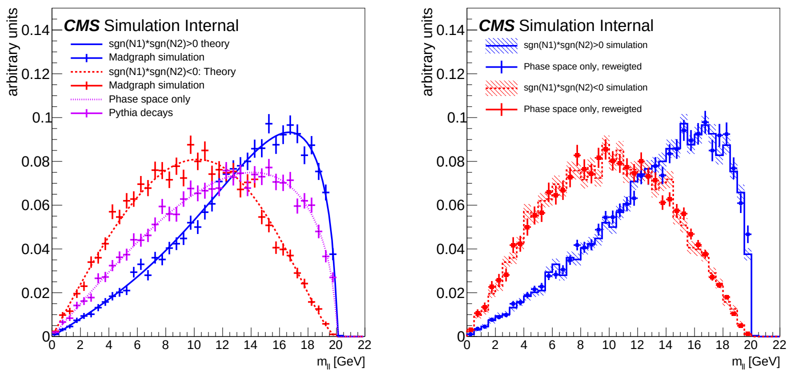

Figure 4.4: The left plot shows differential decay rate for the decay $\widetilde{\chi}_2^0 \to \widetilde{\chi}_1^0 \ell\bar{\ell}$ as a function of $M(\ell\ell)$. The lines correspond to theoretical calculations and the points correspond to Monte Carlo simulations with statistical errors. The red line/points show decay rate as a function of $M(\ell\ell)$ when the full matrix element is included in the calculation of the decay rate and $m_{\widetilde{\chi}_2^0} \times m_{\widetilde{\chi}_1^0} < 0$ while the blue line/points show the case of $m_{\widetilde{\chi}_2^0} \times m_{\widetilde{\chi}_1^0} > 0$. The magenta line/points show the case where a flat matrix element is considered. The right plot presents the decay rate as a function of the $M(\ell\ell)$ where the theoretical calculations are used to re-weight the phase space only simulation in both scenarios and compared to genuine simulation.

Both the refinement of the W/Z BR and the reweighting of the $M(\ell\ell)$ to account for the full calculation of the $\widetilde{\chi}_2^0 \to \widetilde{\chi}_1^0 \ell\bar{\ell}$ decay rate are new additions to the analysis.

4.2.2 Higgsino pMSSM model

Additionally a pMSSM model has been used for the interpretation of the Higgsino scenario. The pMSSM is a projection of the full MSSM model of 105-free parameters to 18 free parameters. The pMSSM model assumes no flavour changing neutral currents, no new sources of CP violation and first and second generation universality. The effects of non electroweakino parameters are decoupled and the parameters that are left to vary are the higgsino μ, bino M_1 and wino M_2 mass parameters. The effect of the variation of $tan\beta$ was found to be negligible and thus it is fixed at 10. The parameters are further reduced to a 2-dimentional grid by assuming $M_2 = 2M_1$. Therefore in the pMSSM Higgsino interpretation the $\mu - M_1$ values are scanned to cover a range of $m_{\widetilde{\chi}_2^0}$ of $100 - 250$ GeV and $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0}$ in the range of $5 - 50$ GeV.

Figure 4.5 (left) shows the $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0}$ as a function of M_1 and μ. For low $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0}$ down to 4 GeV, M_1 goes up to 1.2 TeV. The μ parameter is restricted by the production cross section, which is shown in Fig. 4.5 (right) as a function of M_1 and μ. The cross section has been calculated at NLO-NLL accuracy at $\sqrt{s} = 13$ TeV. The plot in Figure 4.5 shows that the cross section nearly constant with M_1 and decreases with μ. The analysis is expected to be sensitive up to $\mu \sim 200$ GeV. As a safety margin an additional factor of 2 for the cross section, which yields $\mu \sim 240$ GeV was desired. The analysis is more sensitive in the leptonic decays of $\widetilde{\chi}_2^0$ therefore the most relevant production modes out of $\widetilde{\chi}_2^0\widetilde{\chi}_1^0$, $\widetilde{\chi}_2^0\widetilde{\chi}_1^{\pm}$, $\widetilde{\chi}_1^{\pm}\widetilde{\chi}_1^0$ and $\widetilde{\chi}_1^{+}\widetilde{\chi}_1^{-}$ are $\widetilde{\chi}_2^0\widetilde{\chi}_1^{\pm}$ and $\widetilde{\chi}_2^0\widetilde{\chi}_1^0$.

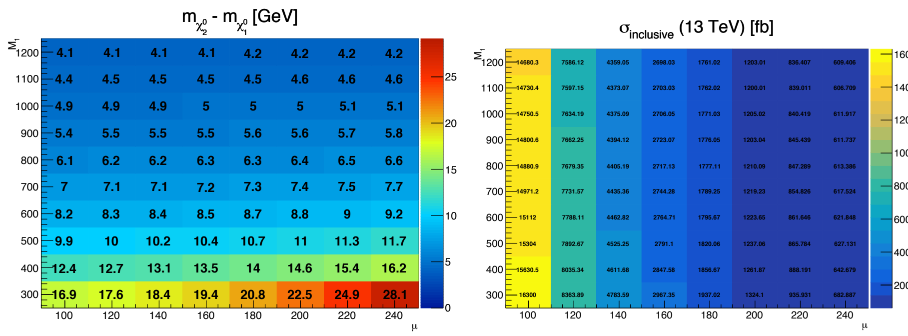

Figure 4.5: The neutralino mass splitting (left) and production cross section (right) as a function of the Higgsino and the Bino mass parameters of the pMSSM model.

4.2.3 Top squark decay simplified models

For the interpretation of the top squark anti-squark pair production and decay to top quark and neutralino, two simplified models are considered. The first is denoted as T2Bff and considers the scenario of the top squark NLSP decaying via a 4-body decay to bottom quark, a pair of fermion-antifermion and a neutralino, as shown in the left diagram of Fig. 4.2. The second model is denoted as T2BW. For these signals to enter the sensitivity reach of the analysis, the $\Delta m_{\widetilde{t}-\widetilde{\chi}_1^0}$ must be small. In this model each pair-produced top squark decays to a b quark and a chargino which subsequently decays to a W boson and an LSP with a branching fraction of 100% as presented in the right diagram of Fig. 4.2. In the T2BW model the relation between the SUSY particle masses assumed is $m_{\widetilde{\chi}_1^{\pm}} = \frac{1}{2}(m_{\widetilde{t}} + m_{\widetilde{\chi}_1^0})$. In order for these signal to enter the sensitivity reach of the analysis, the $\Delta m_{\widetilde{t}-\widetilde{\chi}_1^0}$

is small. The mass grid that is covered for these stop signal extends in the ranges of $\Delta m_{\tilde{t}-\tilde{\chi}_1^0}$ in the range of 10-80 GeV and $m_{\tilde{t}}$ in the range of of 200-1000 GeV.

4.3 Data sets and MC samples

The analysis uses centrally produced NanoAODv6 samples both for data and simulation. It exploits the full data set from Run 2 LHC pp collisions period which correspond to integrated luminosities of 35.9 fb^{-1} in 2016, 41.5 fb^{-1} in 2017 and 59.7 fb^{-1} in 2018. The data events used in the analysis have been collected in the MET and DoubleMu primary datasets with specific trigger algorithms that will be described in Section 4.4. The list of the data sets is presented in Table A.1 in Appendix A.1.

Simulated samples are used for the prediction of the SM background. The $t\bar{t}$, W+jets and Drell-Yan (DY) processes are simulated at LO with MadGraph_aMC@NLO event generator [164, 165] using NNPDF3.0 LO parton distribution functions (PDFs) [164]. Other SM background processes such as diboson, single top quark and rare processes are produced with MadGraph_aMCNLO or POWHEG v2.0 [166, 167, 168] at NLO, using the same PDF mentioned earlier. Showering and hadronization is done by PYTHIA 8 [114] and the CMS detector simulation by the Geant4 package [115]. A

For the signal simulation, pair-produced SUSY particles are generated at LO with MadGraph_aMC@NLO and their decay is simulated in PYTHIA 8. The full CMS detector simulation with Geant4was done for the simplified model samples of TChiWZ and Higgsino (FullSIM). The pMSSM Higgsino, T2Bff and T2bW sample were produced using the simplified detector simulation with the FastSIM package [169]. A detailed list of the signal model names and the corresponding sample names are presented in Table A.2 of Appendix A.1.

4.4 Triggers

The signature that is searched for contains soft leptons and medium to high MET. It is important to apply those criteria in the online event selection, at trigger level, in order to maximize the purity of the collected data sample in signal-like events. The trigger algorithms employed for the online event selection play a very important role in the core design of the analysis. Two HLT trigger algorithms are used for maximum event acceptance and efficiency in the ranges of the offline p_T^{miss}[1] that is of interest to the search. The HLT trigger algorithms relay on the the presence of high p_T^{miss} or the presence of lower p_T^{miss} and two

[1]The p_T^{miss} denotes the amplitude of the negative vector sum of the transverse momentum p_{T} of all PF particle candidates of the event excluding the muons.

muons. Therefore, two broad categories of event are defined based on their offline value of p_T^{miss}, namely the high MET and low MET.

The efficiency measures the probability for a physics object passing the selection requirements at trigger level to also be selected offline by the dedicated isolation and ID requirements. The efficiency is measured both in data and simulation using the tag-and-probe technique and correction scale factors are applied on the simulation.

4.4.1 HLT trigger algorithms

The events with offline $p_T^{miss} > 200$ GeV fall in the high MET region and are probed by an inclusive $p_T^{miss} > 120$ GeV, (denoted as `PFMETNoMu120` in the trigger name) HLT trigger algorithm. The inclusive p_T^{miss} HLT trigger requires missing transverse hadronic energy [2] greater than 120 GeV (denoted as `PFMHTNoMu120` in the trigger name) and transverse hadronic energy [3] greater than 60 GeV (denoted as `PFHT60` in the trigger name). The H_T requirement is applied in the algorithm only in 2017 and 2018 and not in 2016. The pure p_T^{miss} HLT algorithm is seeded at L1 by seeds that require at least 100 or 110 GeV of p_T^{miss} and at least 60 GeV of H_T. At Level 1, jets (MET) are reconstructed (is calculated) based on the calorimeter information only without including the muons.

Additionally, events with lower p_T^{miss}, in the low MET bin, are probed by a double-μ plus MET HLT trigger, seeded by double-μ and MET L1 trigger seeds. The HLT algorithm requires raw $p_T^{miss} > 50$ [4] and two muons with $p_{\mathrm{T}}^{ll} > 3$ GeV and opposite electric charge (sign). In addition, it requires the invariant di-muon mass to be in the range of $3.8 < m_{ll} < 56$ GeV and an upper cut of 0.5 cm on the distance of closest approach (DCA), which is the smallest 3-dimensional distance of the 2 muon tracks. In 2017 and 2018 the DCA requirement was substituted with a $\Delta z < 0.2$ cm cut, which is the distance between the vertices of origin of the two muons on the z-axis. The Δz cut was found to have higher efficiency for triggering on the prompt muons (originating from the primary vertex) of the analysis. To account for the mismatch of L1 p_T^{miss} and the HLT raw p_T^{miss} a p_T^{miss} cut was added in the 2017 and 2018 HLT double-mu plus MET trigger algorithm. Events with offline raw $p_T^{miss} > 125$ GeV and $125 < p_T^{miss} < 200$ GeV are accepted in the low MET region. Given that the analysis targets very soft muons ($p_{\mathrm{T}} < 30$ GeV) the raw p_T^{miss} and p_T^{miss} are approximately equivalent.

[2]The missing transverse hadronic energy MH_T is defined as the amplitude of the negative vector sum of the transverse momentum p_{T} of all the PF jets with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 5$ that pass tight identification criteria (denoted as `IDTight` in the trigger name) based on their components and charged/neutral hadronic/electromagnetic energy fractions. Its calculation excludes the muons in the event.

[3]The transverse hadronic energy H_T is defined as the scalar p_{T} sum of all reconstructed PF jets with $p_{\mathrm{T}} > 30$ GeV and $|\eta| < 2.5$.

[4]The raw missing transverse energy is the PF p_T^{miss} without excluding the muons

Due to the harsher PU and some issues in prescaling, an additional requirement on the jets $p_{\mathrm{T}} > 60$ GeV was introduced in the 2017 L1 seed. This cut has no effect in the offline analysis since a much higher $H_T > 100$ GeV cut is applied offline.

In order to control the SM background arising from WZ processes it is important to remove the upper cut on the invariant di-muon mass. For the event selection in the WZ enriched region a double-μ HLT trigger algorithm that requires $p_{\mathrm{T}} > 17$ GeV and $p_{\mathrm{T}} > 8$ GeV on the muons and the tracker track isolation (`TrkIso`) to pass the very loose WP (VVL) was used.

In 2016 the double-μ plus MET HLT trigger algorithm was added on the 20^{th} of June, therefore the trigger did not collect the full luminosity of the Run period. The total integrated luminosity it collected is 33.2 fb^{-1}. In 2017 the double-μ plus MET trigger was not available in the first Run thus the integrated luminosity collected is 36.7 fb^{-1} while in 2018 it collected a bit less than 59.2 fb^{-1} due to an accidental disabling of the L1 seed algorithm. The pure p_T^{miss} HLT trigger algorithm was included for the full Run periods in all three years and the full integrated luminosity has been collected.

Table 4.1 presents the HLT trigger algorithm paths used for the event selection in the SR and the SM background control region (CR), in the two MET regions. In addition the collected integrated luminosity per year from each path is presented. The symbols `PFMET` and `PFMHT` shown in the table correspond to MET and $\mathrm{MH_T}$ respectively, recontacted with the PF algorithm described in Chapter 3. The symbol `TrkIso` refers to the tracker-measured isolation.

Region	Offline range	HLT path (year : luminosity)
SR, CR	$125 <$ raw p_T^{miss} && $125 < p_T^{miss} < 200\,\mathrm{GeV}$	`HLT_DoubleMu3_PFMET50` (2016 : 33.2 fb^{-1}) `HLT_DoubleMu3_DZ_PFMET50_PFMHT60` (2017 : 36.7 fb^{-1}, 2018 : 59.2 fb^{-1})
SR, CR	$p_T^{miss} \geq 200\,\mathrm{GeV}$	`HLT_PFMETNoMu120_PFMHTNoMu120_IDTight` (2016 : 35.9 fb^{-1}) `HLT_PFMETNoMu120_PFMHTNoMu120_IDTight` (2017 : 41.4 fb^{-1}, 2018 : 59.7 fb^{-1}) `HLT_PFMETNoMu120_PFMHTNoMu120_IDTight_PFHT60` (2017 : 41.4 fb^{-1}, 2018 : 59.7 fb^{-1})
Low MET WZ CR	$125 <$ raw p_T^{miss} && $125 < p_T^{miss} < 200\,\mathrm{GeV}$	`HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ` (2016 : 35.9 fb^{-1}) `HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8` (2017 : 36.7 fb^{-1}, 2018 : 59.7 fb^{-1})

Table 4.1: List of HLT path applied to the 2016, 2017 and 2018 data set. "CR" means all the control regions but the WZ low MET CR in which the double-μ trigger is used and it is presented in the last row of the table.

4.4.2 Trigger efficiencies

The efficiency of offline selected events to have also been selected by the online HLT trigger algorithms is measured both in data and simulation. The efficiency is measured separately for the different trigger cuts as the ratio of events that pass the full HLT trigger selection over the number of events passing the full HLT trigger selection but the cut whose efficiency is being measured. The simulated events are corrected with data-to-simulation efficiency scale factors (SF), to account for simulation or reconstruction inefficiencies.

High MET region

For the efficiency measurement of the pure p_T^{miss} HLT trigger algorithm, the SingleMuon data set [5] is used. The events are required to pass the unprescaled lowest threshold Iso(Tk)Mu HLT paths[6] and an offline $H_T > 100$ GeV cut, in accordance with the offline analysis event selection. The efficiency is measured as the fraction of events passing the pure p_T^{miss} HLT path over the total number of events of the above selection. Figure 4.6 presents the efficiency as a function of the p_T^{miss} for the three years. In 2016 the plateau is reached at $p_T^{miss} = 200$ GeV and it is $\sim 97\%$ both in data and simulation. However, for the other two years the offline selection is not enough to ensure that the events of the analysis fall on the plateau of the efficiency. For 2017 and 2018 the efficiency is larger than 80% at $p_T^{miss} = 200$ GeV and reaches the plateau at $\sim 98\%$ both for data and simulation at $p_T^{miss} \sim 230$ GeV. It is crucial that the parametriztion of the efficiency turn-on curve and plateau describes well the points. The parametrtization is done with an error function in terms of p_T^{miss}.

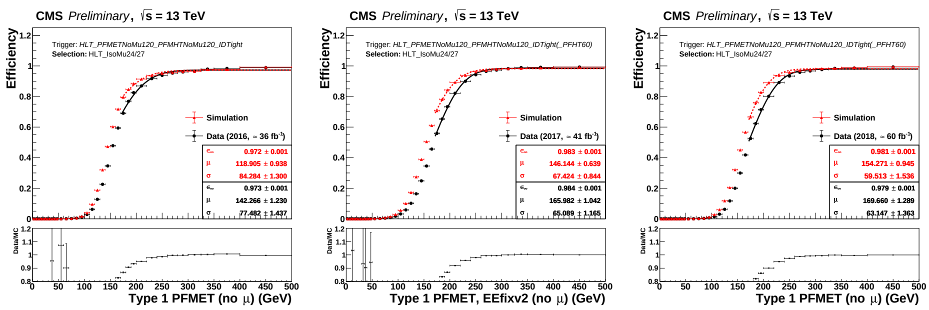

Figure 4.6: The pure p_T^{miss} HLT trigger algorithm efficiencies as function of offline p_T^{miss} measured in both data and simulation in 2016 (left), 2017 (middle) and 2018 (right).

Low MET region

[5]A SingleMuon primary data set consists of events collected by different single muon trigger
[6]Where `Iso(Tk)Mu` stands for (tracker-) calorimeter-measured isolation.

The efficiency of the double-μ plus MET HLT trigger algorithm is measured separately for every selection criteria. Therefore, there is the leptonic part ($\mu_1,\ \mu_2$), the hadronic part (MET) and the DCA or Δz part. The inefficiency of invariant mass requirement was found to be negligible in all three years and it is omitted. An example of the dimuon invariant mass efficiency, measured in 2018 data and simulation, is presented in Figure 4.7.

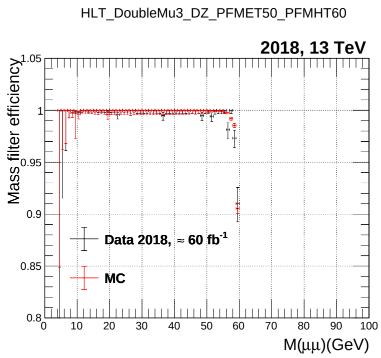

Figure 4.7: The di-muon invariant mass efficiency as a function of $m_{\mu\mu}$ for 2018.

The total efficiency can be factorised as

$$\epsilon = \epsilon_{\mu_1}\epsilon_{\mu_2} \times \epsilon_{DCA/\Delta z} \times \epsilon_{MET} \tag{4.2}$$

The leptonic part efficiency for each muon leg have been computed with the tag-and-probe method described in Subsection 3.1.1. Events are required to pass the lowest threshold unprescaled Iso(Tk)Mu single muon HLT trigger paths. The tag muon is matched to the object that fired the trigger and passed the p_T^{miss} requirement of the L1 trigger algorithm that seeded the HLT path. The efficiency is the fraction of the probe muons passing the last muon filter of the HLT path over all the probe muons. The leptonic part of the efficiency is measured in bins of $p_{\mathrm{T}\ \mu}$ and η.

The efficiency maps in Fig. 4.8 show the leptonic efficiency as a function of muon p_{T} and η in data and simulation for 2016, 2017 and 2018.

It should be noted that in 2016 the DCA part is included in the lepton efficiency computation. This means that in order to avoid double counting the DCA efficiency, equation 4.2 should be modified according to

$$\epsilon = \frac{\epsilon_{\mu_1}\epsilon\mu_2}{\epsilon_{DCA}} \times \epsilon_{p_T^{miss}} \tag{4.3}$$

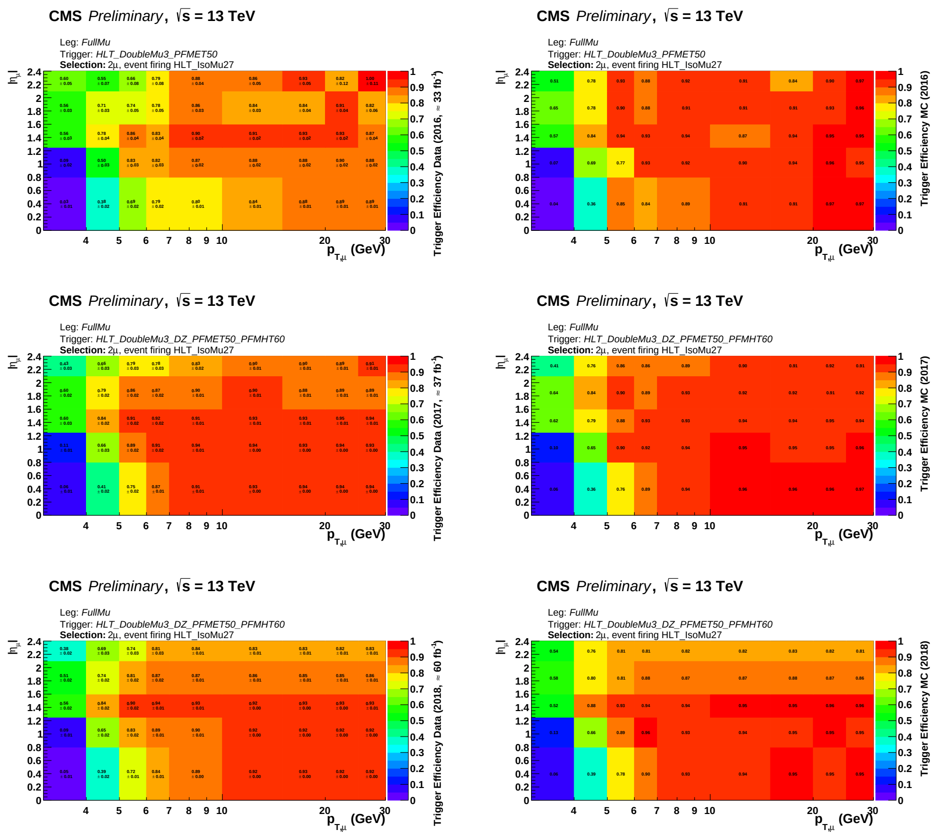

Figure 4.8: The leptonic part of the double-μ plus MET efficiency measured in data (left) and simulation (right) as a function of p_{T} and η of the muons. The plots from top to bottom show the efficiency measured in 2016, 2017 and 2018 respectively. The 2016 maps include sizeable DCA inefficiency measured on data.

The 2016 DCA part of the efficiency is also measured with the tag-and-probe method. Events are required to pass a double-μ HLT trigger with loose tracker track isolation requirements, that is seeded by a L1 double-muon seed. The tag and the probe muons have to pass the identification criteria of the analysis, they are required to be of opposite electric charge and in the invariant mass range of $4 < m_{ll} < 56$ GeV. The efficiency is parametrized as function of the η of the two muons as shown in Fug. 4.9. The muons are chosen such that $\eta_{\mu_1} > \eta_{\mu_2}$ hence the upper left half of the map is completely empty. Due to the kinematic cuts on the muon p_T at 17 and 8 GeV on the leading and subleading respectively, in combination with the invariant mass not the whole η_{μ_1}-η_{μ_2} plane can be probed. For cases where η_{μ_1}-$\eta_{\mu_2} > 2.5$ a flat 91% and 96% efficiency is used for data and simulation as the average value of the efficiency over the range of validity of the map. As mentioned above, in 2017 and 2018 the DCA cut of the trigger was replaced with a Δz cut. The efficiency was measures with the tag-and-probe method as described for 2016 and found to be very stable across the η_{μ_1}-η_{μ_2} plane. As a result a flat number of the average over all the bins is used. The Δz efficiency both in 2017 and 2018 on data was found to be 99% and on the simulation was found to be 99%.

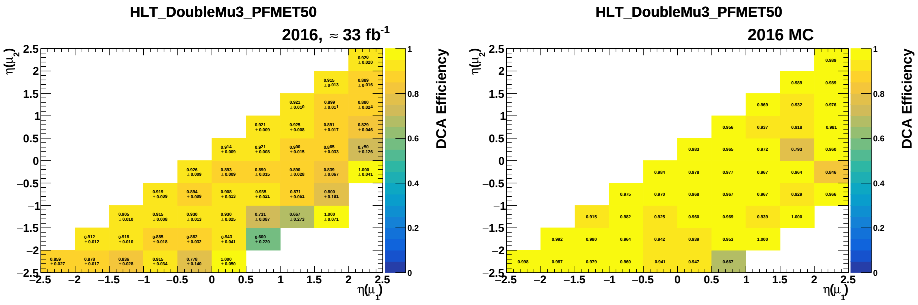

Figure 4.9: The 2016 DCA efficiency maps for data (left) and simulation (right). The DCA efficiency is measured as a function of η_{μ_1} and η_{μ_2}.

The MET part of the efficiency is also measured with the tag-and-probe method. The events are required to have at least two muons matching the ID criteria of the analysis and at least one jet with $p_{\mathrm{T}} > 30$ GeV. A muon p_{T} cut is applied as $10 < p_{\mathrm{T}} < 50$ GeV in order to probe soft muons but also retain good statistics. The two muons are required to have opposite electric charge and invariant mass in the range $4 < M(\ell\ell) < 56$ GeV. The events should pass double-μ HLT triggers with Δz and mass requirements. In the numerator the double-μ plus MET HLT path is required to have fired. The efficiency is measured as a function of raw p_T^{miss} and p_T^{miss}. The efficiency map shown in Fig. 4.10 presents the ϵ_{MET} measured in 2018 data and simulation as a function of raw p_T^{miss} and p_T^{miss}. The plots show that the efficiency becomes stable for ranges of raw $p_T^{miss} > 125$ GeV and $p_T^{miss} > 125$ GeV where the offline cut of the analysis is set.

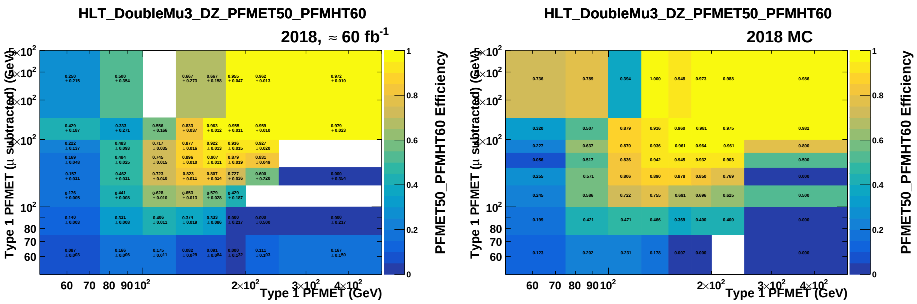

Figure 4.10: The 2018 combined L1 and HLT MET efficiency maps, ϵ_{MET}, with respect to the offline analysis criteria. The efficiency maps are given as a function of raw p_T^{miss} and p_T^{miss}.

4.4.3 Scale factors and uncertainties

The efficiency of the 2016 pure p_T^{miss} HLT trigger algorithm, is very similar for data and simulation, $\sim 97\%$ on the plateau, which is reached around 200 GeV. The simulation efficiency on the turn-on is a few per cent higher compared to data but this has very minor effect on the analysis as most of the offline selected events fall in the plateau. The trigger SF are estimated as the ratio of the efficiency measured in data over the efficiency measured in simulation and was found to be very close to unity for 2016. In 2017 and 2018 the efficiency plateau is reached a bit after 200 GeV and it is $\sim 98\%$ both for data and simulation. From the parametrization of the shape of the efficiency turn-on, SFs that deviate $\sim 10\%$ from unity have been estimated for this region and are applied to correct for the simulation efficiency overestimation.

For the double-μ plus MET HLT trigger algorithm the SF are factorized similarly to the efficiency factorization as

$$SF = \frac{(\epsilon_{\mu_1}\epsilon_{\mu_2} \times \epsilon_{\mathrm{DCA}/\Delta\mathrm{Z}} \times \epsilon_{\mathrm{MET}})_{\mathrm{Data}}}{(\epsilon_{\mu_1}\epsilon_{\mu_2} \times \epsilon_{\mathrm{DCA}/\Delta\mathrm{Z}} \times \epsilon_{\mathrm{MET}})_{\mathrm{MC}}} \tag{4.4}$$

The SF are applied to all FullSIM simulated events where the trigger is emulated and required in the event selection. For the FastSIM signal simulation only the numerator of eq. 4.4 is used to rescale the simulated events to the data efficiency as the HLT trigger algorithms are not emulated in the simulation. In 216 the total double-μ plus MET efficiency is measured to be $\sim 80\%$ in data and a bit higher in simulation. In data the leptonic part without the DCA inefficiency is $\sim 95\%$, the DCA part brings a $\sim 10\%$ inefficiency and the MET part efficiency is $\sim 95\%$. The data-to-simulation SF were measured to deviate by $\sim 5-10\%$ form unity. In 2017 and 2018 the data efficiency inclusively was measured to be $\sim 85\%$. The increased efficiency compared to 2016 is due to the substitution of the DCA requirement with the Δz. The simulation efficiency was measured to be slightly higher compared to data and the 2017 and 2018 SF are found to range from $\sim 0.9-0.95$.

The uncertainties on the efficiency account both for statistical and systematic effects and are propagated to the rest of the analysis as systematic uncertainties. The values of the uncertainties were suggested by the trigger group that measured the SF, based on the typical data/MC differences and the statistical uncertainty. Each muon leg of the double-μ plus p_T^{miss} HLT trigger algorithm is assigned a 2% uncertainty. For the MET efficiency a 2% uncertainty is assigned on the plateau and 5% uncertainty is assigned on the turn-on both for the double-μ plus p_T^{miss} and the pure p_T^{miss} triggers. The same logic is followed for all the years.

4.5 Event reconstruction and object definition

The events are reconstructed with the PF algorithm as explained in detail in Section 3.3.1. The standard physics objects are used as provided by the CMS Physics Object Group (POG).

4.5.1 Electrons

Electrons Tight Identification Selection

The electrons are build from the association of ECAL clusters to GSF tracks, described in 3.3.5. The criteria that an electron should fulfill in order to pass the tight identification (Tight ID) requirements of the analysis are presented in Table 4.2.

Cut	Electron Tight ID
p_T	> 5 GeV
$\|\eta\|$	≤ 2.5
No missing pixel hit	✓
No conversion vertex association	✓
3D impact parameter (IP_{3D})	< 0.01cm
3D impact parameter significance ($\sigma_{IP_{3D}}$)	< 2.0
2D impact parameter ($\|d_{xy}\|$)	< 0.05 cm
z-axis vertex projection ($\|d_z\|$)	< 0.1 cm
Absolute isolation (Iso_{abs})	< 5 GeV
Relative isolation (Iso_{rel})	< 0.5
Deep CSV veto	custom loose WP
Electron MVA ID	tight WP

Table 4.2: Isolation and identification criteria imposed on electron candidates in order to pass the Tight ID selections. The criteria are the same across the years. The electron MVA ID and the Deep CSV have different WPs and trainings depending on the year. The $IP_{3D}, \sigma_{IP_{3D}}, d_{xy}$ and d_z are described in 3.3.2.

The analysis targets the compressed region of the SUSY parameter space where the leptons are soft, the lowest boundary in p_T on the electrons is 5 GeV. In the SR an upper cut at 30 GeV is applied while in some SM background CR the upper p_T cut is relaxed. All these will be described in detail in Section 4.6 and 4.7. The electrons are required to be within the ECAL acceptance of $|\eta| < 2.5$. The electron track should not lack any pixel hit and should not be associated to a $\gamma^* \rightarrow e^+e^-$ conversion vertex. The "promptness" of the tight electron is achieved with the cuts applied on the IP_{3D} and its significance. The absolute isolation is defined as the energy sum deposited around the electron within a cone of $\Delta R < 0.3$ excluding the electrons energy. This is required to be lower than 5 GeV while the relative isolation ($Iso_{abs}/p_T(e)$) should be lower than 0.5. The relative isolation efficiency was chosen such that it ensure high purity of "good" electrons in the lower p_{T} region while for $p_{\mathrm{T}} > 10$ GeV the absolute isolation becomes dominant. Electrons are vetoed against being included in a b-jet. This is achieved by applying a veto on electrons associated with a b-tagged jet passing the loose WP of the Deep CSV which is 0.1522 for 2017 and 0.1241 for 2018. For 2016, the official loose Deep CSV ID has been loosened to 0.4 in order to match the tight ID efficiency of the other years.

Additionally, electrons must be identified according to the Electron MVA ID. It should be noted that different MVA trainings are used per year following the official recommendations: In 2016 and 2018 the MVANoIso94XV2 was used while in 2017 the MVANoIso94XV1 training was used. The Tight WP values were centrally provided by the SUSY Physics Analysis Group (PAG) down to $p_{\mathrm{T}} = 10$ GeV. For the needs of the analysis the WP were extended to the very low p_{T} regions such that they ensure continuity in the electron Tight MVA ID efficiency curve.

Electrons efficiency and scale factors

The electron reconstruction and identification efficiencies are measured with the tag-and-probe method and data-to-simulation SF are derived and applied to simulated events in the SR and SM background CR, to harmonize the selection efficiency in data and simulation. The efficiency is measured in η bins and it is factorized in the reconstruction and identification/isolation efficiency as:

$$\epsilon(e) = \epsilon(e)_{\mathrm{Reco}} * \epsilon(e)_{\mathrm{Tight\ ID\ |\ Reco}} \tag{4.5}$$

The reconstruction efficiency was measured centrally in CMS by the EGamma POG down to 10 GeV. The SF were measured to be stable and very close to unity in all the years and η regions except from the difficult regions of the barrel-endcap gaps at $|\eta| \sim 1.5$. For the lower p_{T} regions (5-10 GeV) the SFs of the last provided p_{T} bin of 10-20 GeV are used. This decision is justified by the observed stability of the SF across the provided p_{T} bins. The uncertainties of the SF are also provided by POG and are of the order of a few per cent.

As an example, the 2018 electron reconstruction SFs as provided by the EGamma POG, in η bins down to $p_{\mathrm{T}} = 10$ GeV are shown in Figure 4.11

The Tight ID efficiency is measured with the tag-and-probe method on electrons from the Z boson decay. A DY sample is used to measure the efficiency in simulation and the SingleElectron dataset to measure the efficiency in data. The efficiency values are extracted using a double Voigtian distribution [7] for the signal fit and the CMS shape function which is an exponential multiplied by an

[7] The Voigtian distribution is the convolution of a Lorentz and a Gaussian distribution

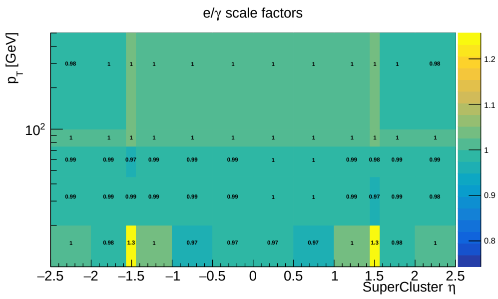

Figure 4.11: The EGamma POG electron reconstruction SF for the $10 < p_\mathrm{T} < 500$ GeV range computed using the full integrated luminosity of 2018.

error function, for the background. The systematic uncertainties were estimated by using alternative functions for the signal (double Gaussian) and background (exponential) fit. The statistical uncertainty is included and it has been found to be negligible. The tight isolation criteria of the electron Tight ID causes electrons to fail the ID requirements due to their own radiation at low di-electron mass. This takes away part of their energy leading to a secondary bump in the low di-electron invariant mass. Therefore the usage of a peak+bump function is necessary for the signal fit. Additionally, the efficiency measurement at very low p_T is challenging due to the very large background arising from $W + jets$ processes. The difficulty in modelling the SM background in these regime leads to large systematic uncertainties. Additional large statistical uncertainty arises from the small signal peak on top of the large background.

The plots in Figure 4.12 show the electron Tight ID efficiency measured in 2016 data and simulation, in the barrel ($|\eta| < 1.5$) and endcap ($|\eta| > 1.5$). The large uncertainty fluctuations in the lower p_T bins is attributed to the difficulty of the fits described above. The ratio plots show data-to-simulation scale factors and their uncertainty.

The efficiency starts at low values, around $\sim 40\%$ for the low p_T electrons, due to the challenges described above and it reaches the plateau of $\sim 50\%$ in the endcap and $\sim 60\%$ in the barrel. There is a general data/simulation overefficiency of $\sim 5\%$ in the endcap and $\sim 10\%$ in the barrel with $\sim$ 5-10% uncertainty. In 2017 and 2018 the data-to-simulation SF are measured to be $\sim$ 0.85-0.90 with uncertainty up to $\sim 10\%$ in the barrel and $\sim$ 1.05-1.20 with uncertainty of $\sim$ 5-20% in the endcap.

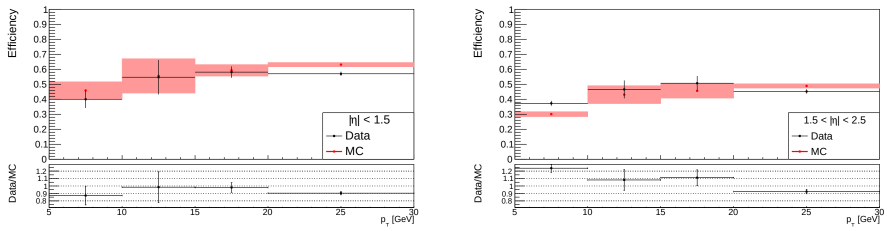

Figure 4.12: The electron efficiencies of the Tight ID measured in 2016 data and simulation as a function of p_{T} in the barrel (left) and the endcap (right). The uncertainty bands contain both statistical and systematic uncertainties for the data and the same for the simulation. The large fluctuations of the uncertainty observed for $p_{\mathrm{T}} < 20$ GeV is attributed to the difficulty of the fits a the low p_{T}. The data-to-simulation scale factors are shown at the bottom of each plot.

4.5.2 Muons

Muons Tight Identification Selection

The muons can either be reconstructed as "global" muons or "tracker" muons. These definitions are described in Section 3.3.3. Table 4.3 summarizes the muon Tight ID selection requirements.

Cut	Muon Tight ID
p_T	$> 5(3.5)$ GeV
$\|\eta\|$	≤ 2.5
Loose ID	✓
Soft ID	✓
3D impact parameter (IP_{3D})	< 0.01cm
3D impact parameter significance ($\sigma_{IP_{3D}}$)	< 2.0
2D impact parameter ($\|d_{xy}\|$)	< 0.05 cm
z-axis vertex projection ($\|d_z\|$)	< 0.1 cm
Absolute isolation (Iso_{abs})	< 5 GeV
Relative isolation (Iso_{rel})	< 0.5
Deep CSV veto	custom loose WP

Table 4.3: Isolation and identification criteria imposed on muon candidates for the Tight ID selection. The criteria are the same between all the years. The Deep CSV has different WPs and trainings depending on the year

In accordance with the electron Tight ID selection, the muon must be soft hence the lowest boundary of the p_{T} is set at 5 GeV or 3.5 GeV when it is possible. This enhances the sensitivity of the search in the very compressed regions. Section 4.6 contains a dedicated discussion on this topic. The reconstructed muons are required to be within the tracker acceptance $|\eta| < 2.4$. The IP_{3D}, isolation and b-tag jet veto requirements are the same for the muons and the electrons. The muon candidates are required to pass the Loose muon ID and the Soft muon ID selection cuts which are defined by the muon POG.

The "Loose muon" is a global or a tracker muon. The "Soft muon" is a tracker track matched with at least one muon segment. It has more than 5 strip tracker hits and at least one pixel tracker hit. The upper cut of $\chi^2/ndof$ of the inner tracker fit is 1.8 and the d_{xy} and $|d_z|$ upper cuts are 3 cm and 30 cm respectively.

Muons efficiency and scale factors

The muon efficiency is measured as a function of η and the boundary between barrel and endcap is set at $|\eta| = 1.2$. The total selection and reconstruction efficiency is factorized into three components: the reconstruction efficiency of a track in the inner tracker, the identification efficiency of a track as a Loose muon candidate and the efficiency of the remaining selection criteria of the tight muon related to the identification, isolation and impact parameter.

$$\epsilon(\mu) = \epsilon(\mu)_{\text{Tracking}} * \epsilon(\mu)_{\text{Loose ID} \mid \text{Tracking}} * \epsilon(\mu)_{\text{Tight ID} \mid \text{Loose ID}} \tag{4.6}$$

The factor of the tracking efficiency is provided centrally in the CMS by the Tracking POG and is applied to muons according to the Muon POG recommendations. The tracking SF are very close to unity and therefore, no correction is needed for this part of the efficiency. The Loose ID scale factors are also compatible with unity in most cases with small deviations ($\sim$ 1-2%) in phase space regions where the muon reconstruction is difficult, such as very low p_{T} muons in the barrel that can barely reach the muon system or very forward η regions. The Loose muon ID SF are measured on the J/ψ peak for $p_{\text{T}} < 20$ GeV and on the Z peak for higher p_{T}. The statistical and systematic uncertainties, provided by the POG, are combined and the total uncertainty is of the order of a few percent.

The plot in Fig. 4.13 presents the 2D maps of the muon Loose ID SF measured in 2018. The SF are very similar across the thee years.

The component of the muon Tight ID selection efficiency $\epsilon(\mu)_{\text{Tight ID} \mid \text{Loose ID}}$ is measured in muons from Z boson decays which are expected to provide higher purity than muons from J/ψ production, due to the strict requirements on the impact parameter and the muon isolation. The procedure of the measurement is similar to the one of tight electron efficiencies.

The data and simulation efficiencies measured in 2016, 2017 and 2018 are shown in the plots of Fig. 4.14. The uncertainty bands contain both statistical and systematic uncertainties and the ratio plots show the data-to-simulation SFs and their uncertainty.

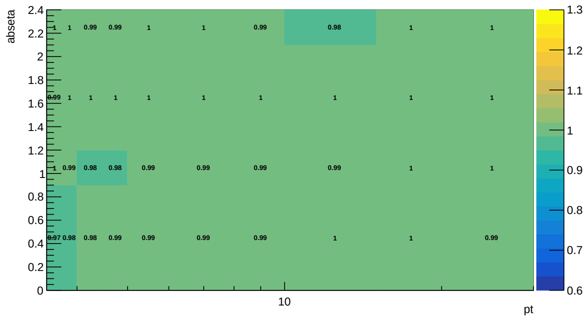

Figure 4.13: The Loose muon ID scale factors measured on the J/ψ peak for the $3.5 < p_T < 20$ GeV range and the Z peak for $p_T > 20$ GeV, computed using the full integrated luminosity of 2018.

The 2017 muon Tight ID efficiency measured in simulation matches the one measured in data to a large extent, leading to SF very close to unit. In 2016 and 2018 the simulation efficiencies were found to be lower than in data. The uncertainties were computed the same way as the corresponding electron Tight ID SF uncertainties. They follow the same trend, i.e. decreasing with the p_T of the lepton.

4.5.3 Lepton LooseNotTight ID selection

The design of the loose lepton identification selection that is used in the data-driven non-prompt background estimation, is a very delicate part of the analysis and special attention was given to it. The loose lepton identification is defined as the OR between the Tight ID and the LooseNotTight ID. The selection criteria of the LooseNotTight ID are presented in Table 4.4.

$$\text{Loose ID} = \text{Tight ID} \mid\mid \text{LooseNotTight ID} \tag{4.7}$$

The LooseNotTight ID can be thought of as an extra requirement that leptons failing the Tight ID selection should pass in order to pass the loose selection. The design of the LooseNotTight ID is motivated by its importance in the non-prompt or fake lepton background prediction with the data-driven tight-to-loose method. The loose lepton selection has been tuned to balance the jet flavour dependence of the method and ensure a good closure of the method. The non-prompt or fake lepton background prediction method and the loose lepton selection will be described in more detail in 4.7.3.

It should be noted that the loose lepton selection is only used in the data-driven non-prompt background prediction, it is not part of the event selection

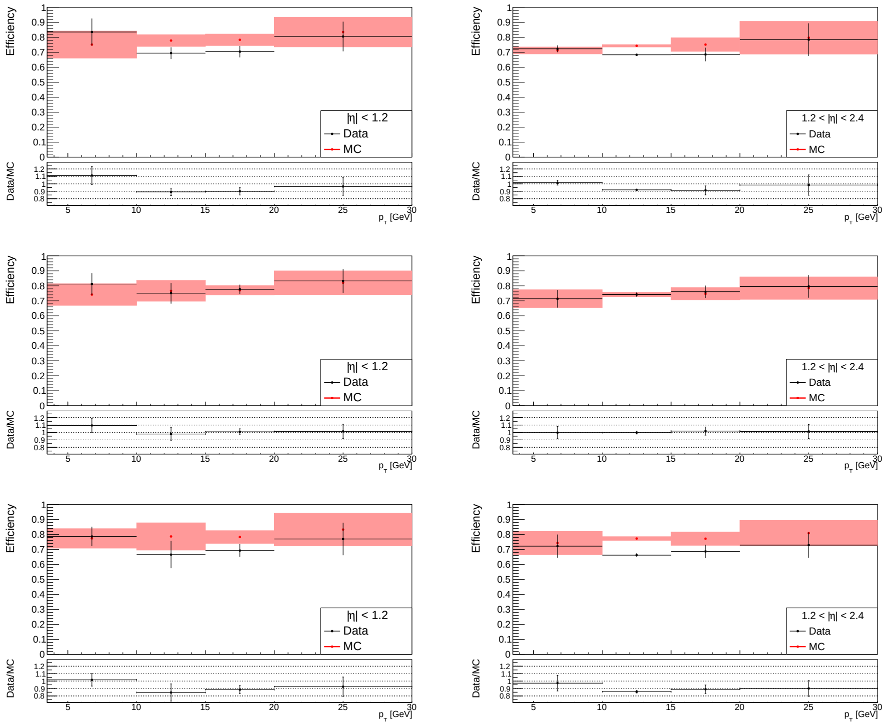

Figure 4.14: Muon Tight ID efficiency measured in data (black points) and simulation (red points) in 2016 (top row) and 2017 (middle row) and 2018 (last row). The left column presents the efficiency measured in the barrel and the right column the efficiency measured in the endcap. The uncertainty is both statistical and systematic for the data and the same for the simulation. The data-to-simulation SF are shown in the ratio plots.

applied on MC samples and therefore there is not needed for efficiency and SF calculation.

4.5.4 Jets, b-tag jets and missing transverse momentum

Jets are reconstructed using the anti-kT algorithm [170] with a distance parameter $R = 0.4$. Every jet is required to have a transverse momentum of at least 25 GeV, be located within the tracker acceptance ($|\eta| < 2.4$), and satisfy the recommended ID requirements as defined for each year by the JetMET POG [171, 172, 173].

A jet cleaning is applied by removing the closest jets to the LooseNotTight lepton, provided they share a common PF candidate, in order to avoid double counting of jets and leptons.

LooseNotTight ID Cut	Electron	Muon
p_{T}	≥ 5	$\geq 5(3.5)$
$\lvert\eta\rvert$	≤ 2.4	
Electron MVA ID	custom ID	-
No missing pixel hit	✓	-
Loose ID	-	✓
Soft ID	-	✓
IP_{3D} [cm]	< 0.0175	
$\sigma_{IP_{3D}}$	< 2.5	
d_{xy} [cm]	< 0.05	
d_z [cm]	< 0.1	
Iso_{Abs}	$< 20.0\ GeV + \frac{300\ GeV^2}{p_T}$	
Iso_{Rel}	< 1.0	
Deep CSV veto	-	

Table 4.4: List of all the selection criteria of the LooseNotTight ID leptons

The b jets are tagged using the medium working point of the DeepCSV discriminant as described in 3.3.8. The medium DeepCSV WP was derived by the b-tagging and vertexing POG and described in [174]. The numerical value of the cut is different for each year of data taking.

Jet energy scale corrections are applied using the standard procedure and tools that are recommended by the CMS JetMET POG, both on data and simulation [175].

The MET is reconstructed based on PF object reconstruction, and the Type-I corrections are applied to it. A detailed discussion on the MET reconstruction and Type-I corrected MET can be found in 3.3.7.

Event filters are designed to remove data events with anomalously high values of p_T^{miss}. These can arise from multiple sources, such as detector noise, reconstruction inefficiencies or beam related effects (e.g. beam halos). The recommended event filters are used according to the official CMS recommendations [176]

4.6 Event selection

The search is looking for signatures with 2 or 3 soft leptons and moderate or high MET in the final state, that originate from the decay of the pair-produced electroweakinos or the top squarks. The HLT trigger algorithms that are used for the online event selection play a key role in the design of the analysis. There are two broad p_T^{miss} regions based on the acceptance of the HLT trigger algorithms, as described in Section 4.4. The search regions (SR) of the analysis are defined based on the HLT trigger algorithm scheme and are divided into the so-called

MET bins. The MET bins have been chosen such that high and stable efficiency of the online selection is ensured.

In the 2ℓ-electroweakino (ewk) SR there are 4 MET bins namely:

- Low MET bin: raw p_T^{miss} >125 GeV and 125<p_T^{miss} <200 GeV
- Medium MET bin: 200<p_T^{miss} <240 GeV
- High MET bin: 240< p_T^{miss} < 290 GeV
- Ultra High MET bin: p_T^{miss} >290 GeV

In the tri-lepton SR (3ℓ-ewk SR) the medium, high and ultra high MET bins are merged into one due to lower event yields.

For the search targeting signature with 2 leptons and MET from the top squark decay (2ℓ-stop SR) 4 MET bins are defined by increasing the upper bounds in the medium, high and ultra high MET bins of the 2ℓ-ewk SR by 50 GeV. The reason behind this is that the event selection in the 2ℓ-stop SR is looser than in the 2ℓ-ewk SR as it will be explained in Section 4.6, thus the MET bin boundaries had to be adjusted in order to achieve maximum sensitivity.

- Low MET bin: raw $p_T^{miss} > 125$ GeV and $125 < p_T^{miss} < 200$ GeV
- Medium MET bin: $200 < p_T^{miss} < 290$ GeV
- High MET bin: $290 < p_T^{miss} < 340$ GeV
- Ultra High MET bin: $p_T^{miss} > 340$ GeV

It should be noted that this MET binning scheme is new compared to the 2016 analysis. The reason is that in the 2016 iteration the MET binning optimization was performed based on the raw p_T^{miss} and not on the p_T^{miss}. Given that p_T^{miss} tends to be higher than the raw p_T^{miss} the p_T^{miss} boundaries of the MET bins should move to higher values in order to maintain the sensitivity, hence the MET bins boundaries were shifted by 40 GeV with respect to the 2016 analysis. However, in the 2016 analysis there were only 3 MET bins, and by increasing all the upper boundaries by 40 GeV, the low MET bin was expanded. This bin is seeded by the double-μ plus p_T^{miss} HLT trigger and can accommodate neither very low $M(\ell\ell)$ selection down to 1 GeV (this will be discussed in more detail in 4.6.1), nor events with electrons. Therefore, instead of increasing the low MET bin upper boundary by 40 GeV, the medium MET bin of $200 < p_T^{miss} < 240$ GeV was added and seeded by the pure p_T^{miss} HLT trigger algorithm. This led to the new analysis MET binning scheme presented above.

In the low MET bins the lepton flavor requirement, as instructed by the trigger, is that of (at least) two opposite sign (OS) muons. In the medium, high and ultra high MET bins of the 2ℓ-ewk SR and the high MET bins of 3ℓ-ewk SR, the requirement of one or at least one opposite sign and same flavour (SF) pair of leptons is applied respectively. All leptonic pairs are accepted in the medium, high and ultra high MET bins of the 2ℓ-stop SR SR.

In every MET bin further binning on a signal-to-background discriminating variable is performed in order to increase the sensitivity of the search. In the 2ℓ-ewk SR, which targets two soft leptons coming from the Z^* decay, the invariant mass of the leptonic pair is expected to have an endpoint according to the mass difference of the $\widetilde{\chi}_1^0$ and $\widetilde{\chi}_2^0$. Similarly, in the 3ℓ-ewk SR the leptonic pair from the Z^* decay is targeted and the Z^* mass is bounded by the mass difference of the $\widetilde{\chi}_1^0$ and $\widetilde{\chi}_2^0$. Therefore, the dilepton invariant mass (minimum 2ℓ invariant mass in the case of the 3ℓ-ewk SR) of a same flavor and opposite sign pair is used as binning variable in the ewk SR. The $M(\ell\ell)$ binning used is $(4,\ 10,\ 20,\ 30,\ 50)$ GeV in the low MET bin and an additional low $M(\ell\ell)$ bin of 1-4 GeV is added in the medium, high and ultra high MET bins. This will be discussed in greater detail in 4.6.1. In the 2ℓ-stop SR the invariant mass of the two soft leptons is not expected to have an endpoint as the two leptons are not originating from the same boson, thus the binning is applied on the leading lepton transverse momentum $p_T(\ell 1)$ with boundaries $(5,\ 8,\ 12,\ 16,\ 20,\ 25,\ 30)$ GeV. A lower $p_T(\ell 1)$ bin of 3.5-5 GeV is added in the medium, high, ultra high MET bins.

Figure 4.15 presents a graphical representation of the 2ℓ- and 3ℓ-ewk SR binning in the four MET bins and further in $M(\ell\ell)$ in 2ℓ-ewk SR, $M(\ell\ell)^{min}_{SFOS}$in 3ℓ-ewk SR and $p_T(\ell 1)$ in the 2ℓ-stop SR.

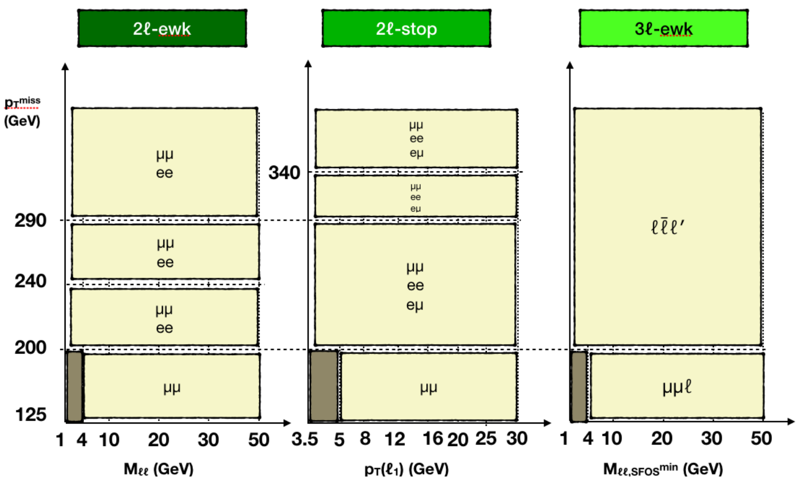

Figure 4.15: Grid plot showing the SR binning in p_T^{miss} and $M(\ell\ell)$ (in 2ℓ-ewk SR), $M(\ell\ell)^{min}_{SFOS}$(in 3ℓ-ewk SR) and in $p_T(\ell 1)$ (in 2ℓ-stop SR). The lepton flavour requirements applied offline in every bin, and indicated by the triggers, are shown.

4.6.1 Baseline and extended event selection

The events selection applied in the low MET bins of the SRs is presented in Table 4.5. The definition of the 2ℓ SRs (first and second column) slightly differs between the stop and ewk SR. The third column shows the requirements for the 3ℓ-ewk SR.

Criterion	2ℓ-stop SR	2ℓ-ewk SR	3ℓ-ewk SR
Triggers (Section 4.4, data and bkg MC)	✓	✓	✓
N_ℓ	=2	=2	=3
1 OS pair	✓	✓	✓
1 SF pair	✓	✓	✓
$p_{\mathrm{T}}(\ell_i)(i=1,2,3)$ [GeV]		(5,30)	
$\|\eta\|$		< 2.5	
$M(\ell\ell)$ ($M(\ell\ell)^{min}_{SFOS}$in 3ℓ-ewk SR) [GeV]		(4, 50)	
$M(\ell\ell)$ ($M(\ell\ell)^{min}_{SFOS}$in 3ℓ-ewk SR) [GeV]		veto (9, 10.5) (Υ veto)	
$M^{\max}_{\ell\ell,AFAS}$ [GeV]	–	–	< 60
$p_{\mathrm{T}}^{\ell\ell}$ [GeV]	> 3		–
$M_{\mathrm{T}}(\ell_i, p_T^{miss})$ $(i=1,2)$ [GeV]	–	< 70	–
raw p_T^{miss} [GeV]		> 125	
p_T^{miss} [GeV]		> 125	
H_T [GeV]		> 100	
p_T^{miss}/H_T	(2/3, 1.4)		–
Jet ID tight WP for leading jet	✓	✓	–
$N_{\mathrm{b\ jets}}(p_{\mathrm{T}} > 25\ \mathrm{GeV})$=0	✓	✓	✓
$m_{\tau\tau}$ [GeV]	veto (0,160)		–

Table 4.5: List of the selection criteria that events must satisfied in order to be selected in the low MET SR. AFAS stands for all flavor all signs.

In the previous iteration of the analysis the lowest $M(\ell\ell)$ boundary was set at 4 GeV dictated by the available simulation samples, the trigger requirements and the background rejection considerations. The goal of the current analysis is to extend the sensitivity to signal models in the very compressed mass regions of $\Delta m_{\widetilde{\chi}^0_2\text{-}\widetilde{\chi}^0_1} < 5$ GeV. Changes and additions concerning the available simulated samples and the analysis event selection cuts were applied in order to extend the sensitivity to lower $M(\ell\ell)$.

The previously available simulated samples were produced with $M(\ell\ell)$ >4 or 5 GeV. The newest available VVTo2L2Nu, ZZTo4, DYJetsToLL samples have a lower $M(\ell\ell)$ thresholds of 1 or even 0.1 GeV in the case of WZTo3LNu. The double-μ plus p_T^{miss} HLT path used in the low MET bin, triggers on dimuon pairs within the invariant mass range of $4 < M(\ell\ell) < 56$ GeV. Therefore, it is not possible to relax the lowest $M(\ell\ell)$ boundary in this MET bin. However, there is no $M(\ell\ell)$ restriction in the pure p_T^{miss} HLT path and thus the low $M(\ell\ell)$ extension will be performed in the medium, high, ultra high MET bins. In the offline events selection the lowest $M(\ell\ell)$ boundary is relaxed from 4 to 1 GeV in the higher MET bins in order to recover the efficiency of signal with small $\Delta m_{\tilde{\chi}_2^0\text{-}\tilde{\chi}_1^0}$. Together with the $M(\ell\ell)$ relaxation a J/ψ veto that excludes $3 < M(\ell\ell) < 3.2$ GeV is applied.

The lepton ID definition, presented in Section 4.5 instructs that only the muon p_T can be relaxed to 3.5 GeV while for the electrons the lowest p_T boundary can be 5 GeV. Lowering the muon p_T to 3.5 GeV in the medium, high and ultra high MET bins increases the acceptance for the low $\Delta m_{\tilde{\chi}_2^0\text{-}\tilde{\chi}_1^0}$ signal. Figure 4.16 (right) shows the normalized signal $M(\ell\ell)$ distribution in the SR for $p_T^{miss} > 200$ GeV, after applying all the SR event selection cuts and relaxing the $M(\ell\ell)$ and muon p_T cuts.

Finally, a $\Delta R_{\ell\ell} > 0.3$ should be applied when the $M(\ell\ell)$ is extended to 1 GeV. Figure 4.16 (left) shows $\Delta R_{\ell\ell}$ normalised distribution for multiple signal mass splittings in $p_T^{miss} > 200$ GeV. It shows that signal with very small mass splitting peaks at low $R_{\ell\ell}$. This could result in issues with lepton self-vetoing in the isolation cone ($\Delta R = 0.3$).

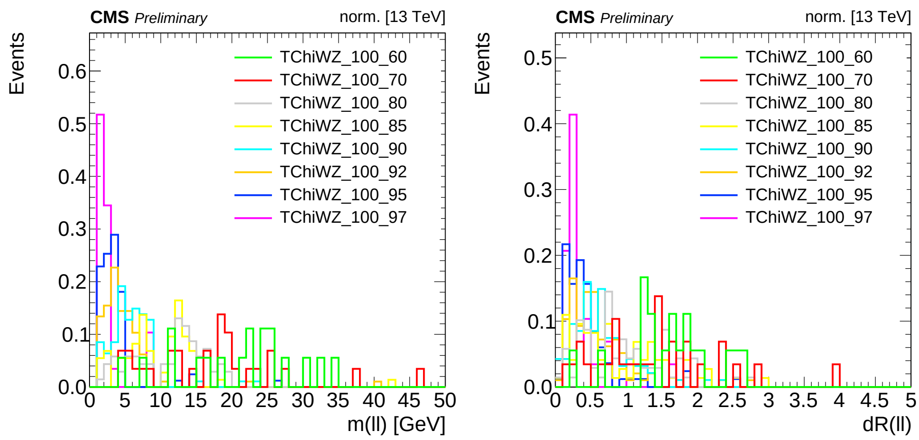

Figure 4.16: Normalized signal $M(\ell\ell)$ (left) and $\Delta R_{\ell\ell}$ (right) distributions for the medium and high MET bins after applying all SR cuts. The $M(\ell\ell)$ cut was relaxed from 5 GeV to 1 GeV and the (sub)leading muon p_T cut was relaxed from 5 GeV to 3.5 GeV

With the relaxation of the $M(\ell\ell)$ signal events populate the lowest $M(\ell\ell)$ bins while almost no background events enter those bins. This is more pronounced as p_T^{miss} increases, as the SM prompt and non-prompt backgrounds are suppressed from the p_T^{miss} cut and the J/ψ veto, and cannot populate the very low $M(\ell\ell)$ region. Relaxing the muon p_T cut on top of the low $M(\ell\ell)$ requirement in the 2ℓ-SR increases the signal acceptance but it also allows more non-prompt events to enter the higher $M(\ell\ell)$ of the SR. In the 3ℓ-SR the increase of the non-prompt or fake background is not significant. Therefore, the degradation of the sensitivity expected in the the higher $\Delta m_{\tilde{\chi}_2^0\text{-}\tilde{\chi}_1^0}$ of the 2ℓ-SR caused by the background increase, will be cancelled out by the 3ℓ-SR which mostly affects the intermediate and high $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0}$ regions.

The "nominal" signal event yields with the $M(\ell\ell) > 4$ GeV and the corresponding signal yields with the relaxation of $M(\ell\ell) > 1$ GeV, $p_T(\ell 1) > 3.5$ GeV and the additional $\Delta R_{\ell\ell} > 0.3$ cut are shown in Table 4.6 for 200 $< p_T^{miss} < 250$ GeV and in Table 4.7 for $p_T^{miss} > 250$ GeV. The signal is the TCHIWZ in the 2ℓ-ewk SR.

Signal point	$M(\ell\ell) > 4$Gev	$M(\ell\ell) > 1$ GeV	$M(\ell\ell) > 1$ GeV $p_T(\mu) > 3.5$ GeV	$M(\ell\ell) > 1$ GeV $p_T(\mu) > 3.5$ GeV $\Delta R_{\ell\ell} > 0.3$
TCHIWZ 100/40	45.07 ± 13.01	45.07 ± 13.01	48.83 ± 13.54	45.07 ± 13.01
TCHIWZ 100/90	72.76 ± 15.88	97.01 ± 18.33	128.19 ± 21.07	107.40 ±19.29
TCHIWZ 100/5	20.83 ± 8.50	59.03 ± 14.32	90.28 ± 17.70	69.44 ± 15.53

Table 4.6: From left to right : the signal yields after the full 2ℓ-ewk SR event selection with $M(\ell\ell) > 4$ GeV, signal yields after subsequently relaxing the $M(\ell\ell)$ and the $p_T(\mu)$ cuts and applying a $\Delta R_{\ell\ell} > 0.3$ cut. The signal yields correspond to 200 $< p_T^{miss} < 250$ GeV. The uncertainties are statistical only. The first number in the signal name denotes the mass of $\widetilde{\chi}_2^0/\widetilde{\chi}_1^{\pm}$ and the second number denotes the $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0}$.

Signal point	$M(\ell\ell) > 4$Gev	$M(\ell\ell) > 1$ GeV	$M(\ell\ell) > 1$ GeV $p_T(\mu) > 3.5$ GeV	$M(\ell\ell) > 1$ GeV $p_T(\mu) > 3.5$ GeV $\Delta R_{ll} > 0.3$
TCHIWZ _100_60	3.76 ± 3.76	3.76 ± 3.73	7.51 ± 5.31	7.51 ± 5.31
TCHIWZ _100_90	79.69 ± 16.62	90.08 ± 17.67	124.72 ± 20.79	103.94 ± 18.98
TCHIWZ _100_95	13.89 ± 6.94	100.69 ± 18.70	128.47 ± 21.12	72.92 ± 15.91

Table 4.7: From left to right : the signal yields after the full 2ℓ-ewk SR SR event selection with $M(\ell\ell) > 4$ GeV, signal yields after subsequently relaxing the $M(\ell\ell)$ and the $p_T(\mu)$ cuts and applying a $\Delta R_{ll} > 0.3$ cut. The signal yields correspond to $p_T^{miss} > 250$ GeV. The uncertainties are statistical only.

Table 4.8 presents all the SR event selection cuts modifications that are applied for the extension to lower $M(\ell\ell)$ in the medium, high and ultra high MET bins of the SR.

The requirement of one same flavor lepton pair is relaxed in the medium, high and ultra high MET bins of the 2ℓ-stop SR because the leptons in the final state are not expected to originate from a Z* decay.

A general note on the plots that will appear in the rest of this subsection, Figures 4.17-4.21, show the distribution of the variables that were used in order to select the SUSY signal and reject SM background in the SR. For the production of the plots, all the SR cuts are applied except the one depicted. The plots are done with the 2018 simulation, the signal (TCHIWZ) is scaled by a factor of

Criterion	2ℓ-stop SR	2ℓ-ewk SR	3ℓ-ewk SR
$p_T(\mu_i)(i = 1, 2, 3)$ [GeV]		> 3.5	
$M(\ell\ell)$ ($M(\ell\ell)^{min}_{SFOS}$in 3ℓ-ewk SR) [GeV]		(1, 50)	
$M(\ell\ell)$ ($M(\ell\ell)^{min}_{SFOS}$in 3ℓ-ewk SR) [GeV]		veto (3, 3.2) (J/ψ veto)	
$\Delta R_{\ell\ell} > 0.3$	✓	✓	✓
1 SF pair	–	✓	✓

Table 4.8: List of modification to event selections criteria in order to select events in the medium, high, ultra MET regions where the selection is relaxed to lower $M(\ell\ell)$.

10, and three mass points with $\Delta m_{\tilde{\chi}^0_2\text{-}\tilde{\chi}^0_1} = 5$ GeV, $\Delta m_{\tilde{\chi}^0_2\text{-}\tilde{\chi}^0_1} = 10$ GeV and $\Delta m_{\tilde{\chi}^0_2\text{-}\tilde{\chi}^0_1} = 40$ GeV are shown. The signal name in the legend of the plot, for example TChi175/5, should be translated as the TChiWZ model with $m_{\tilde{\chi}^0_2} = m_{\tilde{\chi}^\pm_1} = 175$ GeV and $\Delta m_{\tilde{\chi}^0_2\text{-}\tilde{\chi}^0_1} = 5$ GeV. The non-prompt background has been estimated with the semi-DD method that will be described in Sec. 4.7. The plots are made in an inclusive MET bin for $p_T^{miss} > 125$ GeV and $p_T^{miss} > 125$ GeV, using the appropriate triggers in every MET range (double-μ plus p_T^{miss} up to p_T^{miss} <200 GeV and pure p_T^{miss} for p_T^{miss} >200 GeV).

The first part of the Table 4.5 and the Table 4.8 list the event selection cuts based on the *lepton quantities*. The requirements on the number, flavor and charge of the leptons are in line with the topology of the signal hypothesis under study.

Leptons are required to have $p_T < 30$ GeV; this was identified as the p_T value below which the analysis is more sensitive in excluding the benchmark models in the compressed mass regions. The lepton IDs, presented in Section 4.5, are defined for electrons down to 5 GeV and for muons down to 3.5 GeV. However, due to trigger requirements, the lowest offline muon p_T threshold is 5 GeV in the low MET bin, where the double-μ plus p_T^{miss} trigger is applied, and 3.5 GeV in the medium, high and ultra high MET bins, where the pure p_T^{miss} trigger is applied. The $\Delta m_{\tilde{\chi}^0_2\text{-}\tilde{\chi}^0_1}$ is proportional to the leptons p_T, the higher the $\Delta m_{\tilde{\chi}^0_2\text{-}\tilde{\chi}^0_1}$ the higher the lepton p_T. The analysis targets the compressed mass region therefore the lepton p_T should remain low. The lepton p_T distribution of the leading and subleading lepton are illustrated in Fig. 4.17. Comparing the distributions of the different signal mass points one can observe the dependence of the $\Delta m_{\tilde{\chi}^0_2\text{-}\tilde{\chi}^0_1}$ on the lepton p_T.

Further requirements on the event are applied and are listed in the second part of Table 4.5. These cuts are related to *background rejection* and the *ISR requirement* for the case of the 2ℓ-SR.

The transverse mass between a lepton and MET is defined as

$$M_T(\ell_i, p_T^{miss}) = \sqrt{2p_T(\ell_i)p_T^{miss}(1 - cos[\Delta\phi(\ell_i - p_T^{miss})])} \quad i = (1, 2) \tag{4.8}$$

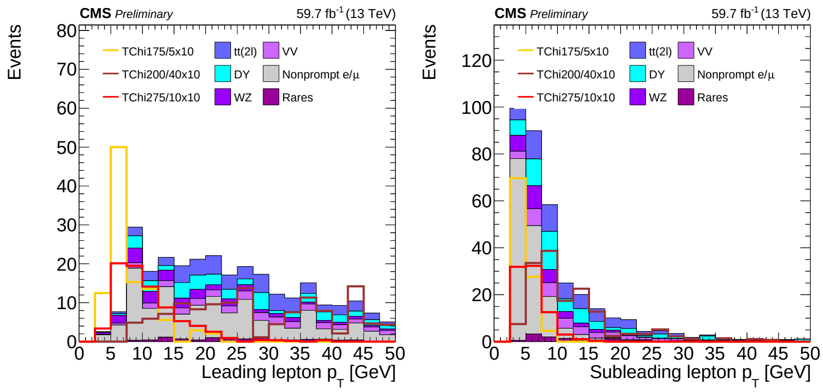

Figure 4.17: The leading lepton p_{T} (left) and subleading lepton p_{T} (right) distributions in the 2ℓ-ewk SR when all the SR cuts but the lepton p_{T} cuts are applied. The plots are done with 2018 simulation, in an inclusive MET bin with the appropriate trigger application. The non-prompt background is estimated with the semi-DD method.

The cut of $M_{\mathrm{T}}(\ell_i, p_T^{miss}) < 70\,\mathrm{GeV}$ $(i = 1, 2)$ is found to be effective in reducing the $\mathrm{t\bar{t}}$ background because in the signal the leading lepton is typically aligned with the LSP. The TChiWZ and SM background distribution of the M_{T} is shown in Fig. 4.18. All the SR selection cuts are applied except the M_{T} cuts on both leptons. This cut is not effective in discriminating the stop signal as shown in Fig. 4.18 bottom row, and therefore it is relaxed in the 2ℓ-stop SR selection. Additionally, the M_{T} cut is relaxed in the 3ℓ-ewk SR SR in order to increase the statistics.

The MET binning is discussed in detail in the beginning of Section 4.6. The lowest boundary p_T^{miss} and raw p_T^{miss} is set at 125 GeV. The TChiWZ and SM background distributions of p_T^{miss} and raw p_T^{miss} for a threshold of 125 GeV are shown in Fig. 4.19. For raw $p_T^{miss} > 125\,\mathrm{GeV}$ and $125 < p_T^{miss} < 200\,\mathrm{GeV}$ the double-μ plus p_T^{miss} trigger is applied while for $p_T^{miss} > 200\,\mathrm{GeV}$ the pure p_T^{miss} trigger is applied together with the low $M(\ell\ell)$ extension setup.

An $H_T > 100$ GeV applied both in the 2ℓ and 3ℓ-SR, suppresses background events with low hadronic activity. This cut is also driven by the HLT trigger cut of PFMHT at 60 GeV.

The selection of $(2/3) < (p_T^{miss}/H_T) < 1.4$ in the 2ℓ final state is proven to suppress very effectively QCD events. The QCD events cannot be simulated with sufficient statistics in order to have some acceptance after the analysis selection cuts, therefore there is no QCD simulated samples included in the plots. The p_T^{miss}/H_T distribution is shown in Fig. 4.20 (left). The leading jet in the 2ℓ-SR is required to pass the tight WP of the jet ID as defined by the Jet/MET POG. The requirements on the tight Jet ID WP removes jets from calorimetric noise as well

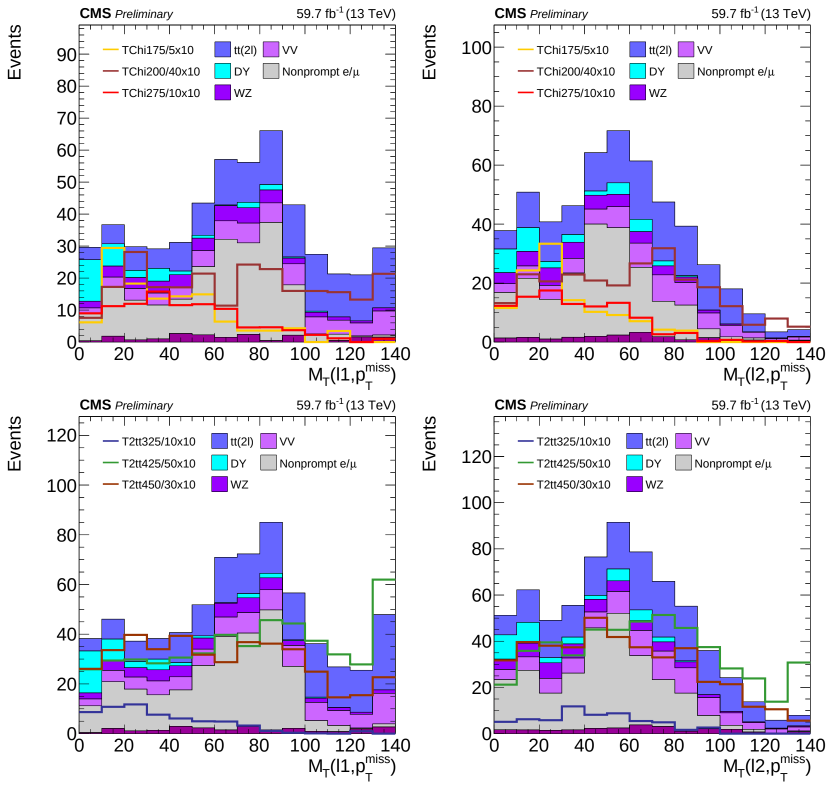

Figure 4.18: The $M_\mathrm{T}(\ell_1, p_T^{miss})$ (left) and $M_\mathrm{T}(\ell_2, p_T^{miss})$ (right) distributions in the SR when all the SR cuts are applied except the $M_\mathrm{T}(\ell_i, p_T^{miss})$. Top row shows the distribution in the 2ℓ-ewk SR and bottom row shows the 2ℓ-stop SR. The plots are made with 2018 simulation, in an inclusive MET bin with the appropriate trigger application. The non-prompt background is estimated with the semi-DD method.

as jets from misreconstructed leptons. The tight WP requirement in combination with the $H_T > 100$ GeV cut can only be satisfied by events with at least one ISR jet. The reason behind this is that the LSP does not interact strongly so no photon or gluon FSR can arise from it. Additionally, the emitted SM particles and their decay products are off-shell and soft due to the small ΔM requirement for the signal, therefore, FSR jets that can arise from W^* hadronic decays will be soft and rejected by the $H_T > 100$ GeV cut. The $(2/3) < (p_T^{miss}/H_T) < 1.4$ cut and the requirement of the leading jet to pass the tight WP are not applied in the 3ℓ-ewk SR in order to retain some statistics in it.

Events containing b-jets as defined in Section 4.5 are vetoed in order to suppress $\mathrm{t\bar{t}}$ background. The b-jet selection contains a lower bound on the jet p_T at 25 GeV, thus events with softer b-jets, such as the signal events from top squark decays are still retained. The distribution of the number of jets in the 2ℓ-ewk SR

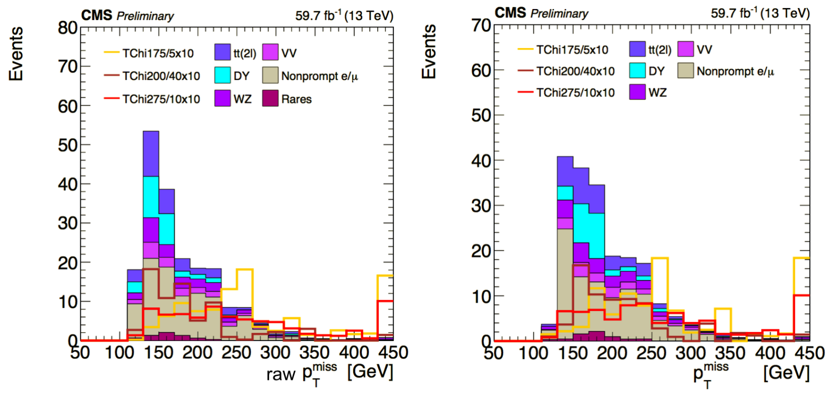

Figure 4.19: The raw p_T^{miss} (left) and p_T^{miss} (right) distributions in the SR when all the 2ℓ-ewk SR SR cuts are applied except except on the variables shown in the plots. The plots are made with 2018 simulation, in an inclusive MET bin with the appropriate trigger application. The non-prompt background is estimated with the semi-DD method.

is shown in Fig. 4.20 (left).

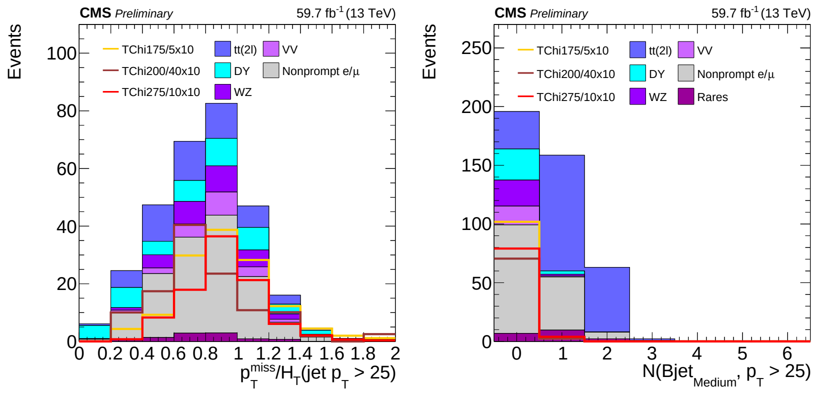

Figure 4.20: The p_T^{miss}/H_T (left) and the number of b-jets (right) distributions in the SR when all the SR cuts are applied, except on the variables shown in the plots. The plots are made with 2018 simulation, in an inclusive MET bin with the appropriate trigger application. The non-prompt background is estimated with the semi-DD method.

A veto is applied on the $0 \leq m_{\tau\tau} < 160$ GeV to reduce contamination from $Z\gamma^* \rightarrow \tau\tau$ DY background events which is an important background for the search. In Z+jet events, in which Z decays to two τ leptons and they further decay leptonically as $\tau \rightarrow \ell\nu\nu$, Z recoils against energetic jet ($H_T > 100$ GeV) and therefore it is boosted. The τ leptons are highly relativistic travelling collimated in the Z direction, and the lepton and the 2 neutrinos from each τ decay are collinear with the τ. In these decays the flight direction of the final lepton is close to the original flight direction of the Z boson. The $\tau\tau$ invariant mass can be reconstructed as was done in [61, 177] and it can be used as a handle to reject $Z \to \tau\tau$ background events. The sum of the neutrinos transverse energy can be written as

$$\vec{p}_T^{miss} = \xi_1 \vec{p}_T(\ell 1) + \xi_2 \vec{p}_T(\ell 1) \tag{4.9}$$

where ξ_i is a scale factor relating the transverse momentum of the neutrinos to the $p_T(\ell_i)$ of the lepton (collinearity) from the ith τ decay. Additionally one can write

$$-\vec{H}_T = (1+\xi_1)\vec{p}_T(\ell 1) + (1+\xi_2)\vec{p}_T(\ell 2) \tag{4.10}$$

Solving the system of the equation, the scale factors ξ_1 and ξ_2 can be calculated and the τ four momentum can be derived. The reconstructed τ momenta are further used to estimate the $m_{\tau\tau}$. For 2ℓ events from $Z \to \tau\tau$ decays it is expected that $\xi_i > 0$ and $m_{\tau\tau}$ peaks near the Z boson mass. For these events, the neutrino momentum vector will usually point in between the two lepton momentum vectors in the transverse plane. For background events where the MET of the event arises from heavy SM particles (t, W) decays, the lepton and MET (from neutrinos) directions are uncorrelated and the $\vec{p}_T^{miss}$ may point away, or even backwards, from one or both leptons so that $\xi_i < 0$. This is also the case for the signal where MET arises from the LSP. In these cases $m_{\tau\tau}$ can take non physical negative values. Figure 4.21 illustrates the $m_{\tau\tau}$ distribution in TChiWZ and background. If the p_T^{miss} is not lying in-between the two soft leptons $\xi_i < 0$. This case corresponds to τ leptons having opposite direction than their daughters (soft leptons). We set the reconstructed invariant mass to be negative for these cases ($m_{\tau\tau} = -|m_{\tau\tau}|$)

The $m_{\tau\tau}$ cut is not applied in the 3ℓ-ewk SR in order to retain some statistics in this SR.

4.7 Methods for background estimation

The SM background that enters the 2ℓ and 3ℓ SR can be broadly classified into four major categories:

- **Prompt 2ℓ processes:**
  $Z/\gamma \to \tau\tau$ Drell-Yan (DY) and dileptonic $t\bar{t}$ pair decay are the most relevant processes leading to two prompt leptons in the final state. Both processes contain real p_T^{miss}, though at different proportions. In $t\bar{t}$ the prompt leptons are produced in a subsequent leptonic decay of the W boson that arises from the $t \to bW$ decay. The leptonic W decay gives rise to an undetectable neutrino, which constitutes the real p_T^{miss} in this process. Given the minimal p_T^{miss} requirement at 125 GeV, direct DY production of e^+e^- or $\mu^+\mu^-$

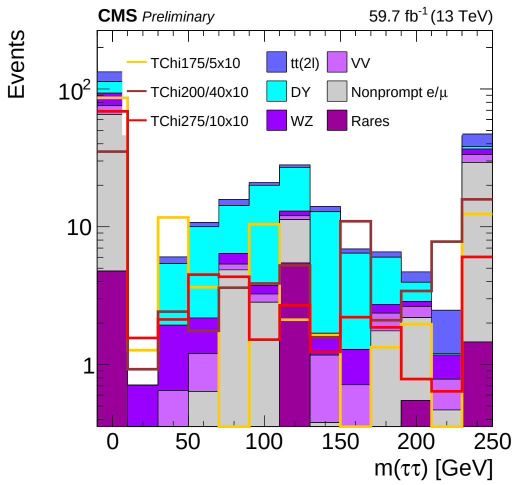

Figure 4.21: The $m_{\tau\tau}$ distributions in the SR when all the SR cuts are applied, except on the variable shown in the plot. The plots are made with 2018 simulation, in an inclusive MET bin with the appropriate trigger application. The non-prompt background is estimated with the semi-DD method. The first and the last bin are underflow and overflow respectively.

pairs is negligible because these processes would not contain any real p_T^{miss}. However, if neutral boson (Z or γ) decay to $\tau^+\tau^-$, the individual τ leptons can further decay to an e or μ and the associated ν's, which lead to real p_T^{miss} in the event. Therefore, the aforementioned decay via a τ lepton pair poses a relevant background component in the 2ℓ SRs.

- **Prompt diboson processes:**
  Diboson production (VV) enters both the 2ℓ and the 3ℓ-ewk SR. In the 2ℓ final state these events arise predominantly from WW production. Further contributions come from Wγ but they are generally small. Background events arising from WZ and ZZ production pose the most important background component in the 3ℓ-ewk SR. In the plots of this thesis the WZ process where both bosons decay leptonically is shown in green histograms and is labelled as 'WZ', while the rest of diboson processes (WW, ZZ, WZ except the fully leptonic one) are represented with violet histograms and labeled as 'VV'

- **Backgrounds from non-prompt and fake leptons:**
  Even though the tight ID lepton selection is designed to accept prompt leptons and reject non-prompt and fake (not real leptons) candidates, there is a residual inefficiency giving rise to a distinct background component. These events are mostly coming from W+jets, t+jets, $t\bar{t}$ +jets and DY+jets processes. The former two contribute only to the 2ℓ SRs while the latter two can also contribute in the 3ℓ-ewk SR. For example, in the W+jets case,

the jet that accompanies the W boson can be misreconstructed as a (fake) lepton or can contain a real lepton (but non-prompt). If it forms an OS pair with the prompt lepton from the W boson decay and satisfies the criteria outlined in Table 4.5 and/or Table 4.8, then the event passes the tight ID.

- **Rare processes:**
  Previously unmentioned SM processes leading to only minor contributions in the SR are collectively referred to as *rare*. These comprise the production of tW ttV, ttH, tZq, tWZ, VVV and processes involving the conversion of a photon to a pair of electrons.

For every one of the dominant prompt and non-prompt/fake SM background a dedicated control region (CR) is designed. The CR are defined such that they are orthogonal to the SR and enriched in the associated SM background process. Each CR is split into two MET bins in order to match the event categorization employed in the SR:

- Low MET: raw $p_T^{miss} > 125\,\text{GeV}$ and $125 < p_T^{miss} < 200$ GeV;
- High MET: $p_T^{miss} > 200$ GeV.

More specifically, the following regions are defined to constrain and verify the modeling of the dominant prompt SM backgrounds:

- A DY and a $t\bar{t}$ CR with very high purity in their associated background and negligible signal contribution.
- One region designed to be enriched in trilepton events from WZ processes. In this region there is a non-negligible signal contribution. This region is referred to as the WZ-enriched.
- One region with moderate purity targeting 2ℓ from diboson processes, VV, is referred to as the VV validation region (VR).
- A 2ℓ same-sign (SS) CR comprising events with same-sign leptons is used to validate the data-driven non-prompt/fake background prediction method and further constraint this background. The SS CR is defined only for p_T^{miss} > 200 GeV (high MET bin) and cannot be extended to lower p_T^{miss} due to an opposite sign requirement on the lepton pairs of the dimuon+p_T^{miss} trigger.

The $M(\ell\ell)$ distributions of the DY CR, $t\bar{t}$ CR, WZ-enriched region and SS CR are included in the signal extraction procedure, which is based on a maximum likelihood fit of the data, while the VV VR is only used to assess the data-simulation agreement and estimate the corresponding background normalization uncertainty.

It is worth noting that for the best description of each of the prompt backgrounds in their dedicated CR, the non-prompt/fake background in these regions has been estimated with the data-driven (DD) method. In the DD method, data events containing at least one lepton that fails the tight selection but passes the loose lepton identification selection (application region), are weighted by a transfer factor which depends on the probability of a lepton that passes the loose lepton identification requirements to pass the tight ID selection. This probability is refer to as the fake rate. The method that was used for the fake rate measurement in data, is described in more details in section 4.7.3. The complete formula of the transfer factor depends not only on the fake rate but also on the prompt rate of the leptons, which is the efficiency of prompt leptons to pass the tight ID selection and it is measured in simulation. The prompt rate is included in the transfer factor formula in order to account for the prompt contamination in the application region. For the non-prompt/fake background estimation in the $t\bar{t}$ CR, VV CR and WZ-enriched region the prompt rate was measured in leptons originating from Z boson. A dedicated prompt rate was measured on leptons originating purely from $Z/\gamma \rightarrow \tau\tau$ decays to be used in the DY CR, as the prompt contamination in the application region of DY found to be purely from $Z/\gamma \rightarrow \tau\tau$ decays.

4.7.1 Prompt dilepton processes

The DY and $t\bar{t}$ processes comprise the dominant prompt SM background processes in the 2ℓ SR. Their contribution is estimated from simulated events with the normalization assessed in the dedicated CR. By construction, the CR phase space is similar to that of the SR but enriched in the associated background process and negligible signal contribution. A detailed discussion on the DY and $t\bar{t}$ CR follows.

DY control region

DY events can lead to a prompt 2ℓOS signature whose presence in the SR is significantly suppressed by the lower MET cut of $p_T^{miss} \geq 125$ GeV. However, there are still two ways that DY events can enter the SR: either by mismeasurement of jets giving rise to p_T^{miss} or by $Z/\gamma \rightarrow \tau_\ell\tau_\ell$ decays, where τ_ℓ denotes a τ lepton that further decays to either an electron or a muon and neutrinos. The former case is usually negligible, while the latter can lead to a sizable background contribution. The DY contribution in the SR can be reduced by the $m_{\tau\tau}$ cuts described in 4.6.1. This selection cut is efficient in rejecting events in the $m_{\tau\tau}$ region where the $Z/\gamma \rightarrow \tau\tau$ events are expected. For the DY CR definition the $m_{\tau\tau}$ cut is inverted with respect to the SR and the allowed range is $0 \leq m_{\tau\tau} < 160$ GeV. This also ensures orthogonality with the SR. Additionally, no upper cut on the lepton p_T is applied to enrich the statistics of the CR.

The DY contribution is estimated from simulated events and its normalization is extracted from the dedicated CR. The selection criteria of the DY CR are summarized in Table 4.9.

DY CR
Triggers (Section 4.4)
$p_T^{miss}$and raw $p_T^{miss} > 125\,\text{GeV}$
$(2/3) < (p_T^{miss}/H_T) < 1.4$
$H_T > 100$ GeV
Tight jet ID for ISR jet
$N_{\text{b jets}} = 0$ (Medium WP)
$N_\ell = 2$
OS pair
SF pair (low MET)
$p_T(\mu(e)) > 5$ (5) GeV (low MET)
$p_T(\mu(e)) > 3.5$ (5) GeV (high MET)
$p_T^{\ell\ell} > 3$ GeV
$4(1) < M(\ell\ell) < 50$ GeV (low(high) MET)
$\Delta R_{\ell\ell} > 0.3$ (high MET)
J/ψ veto: $M(\ell\ell) < 3$ GeV or 3.2 GeV$< M(\ell\ell)$ (high MET)
Υ veto: $M(\ell\ell) < 9$ GeV or 10.5 GeV$< M(\ell\ell)$
$m_T(\ell_i, p_T^{miss}) < 70$ GeV $(i = 1, 2)$
$0 \leq m_{\tau\tau} < 160$ GeV

Table 4.9: Event selection criteria for the DY CR, enriched in $Z/\gamma \to \tau_\ell\tau_\ell$ decays. The invertion of the $m_{\tau\tau}$ cut ensures orthogonality to the SR.

The DY CR is split into two MET regions as described in the beginning of 4.7. The purity of the low MET (High MET) DY CR is 69% (72%) as obtained from simulated events in 2016 and is similar in the other years. The pre-fit plot of the $M(\ell\ell)$ distribution in the DY CR combining the data and simulation of the three years is shown at the top row of Figure 4.22 in the low (left) and high (right) MET bins. Given the moderate statistics, the pre-fit agreement of data and simulation seems good, both in the low as well as the higher MET region. The $M(\ell\ell)$ distribution is used in the final fit which determines the uncertainty assigned on the normalization of DY. In the pre-fit plots only the statistical uncertainty is included and thus the total uncertainty band and the statistical uncertainty band coincide.

The post-fit plots of the $M(\ell\ell)$ distribution in the DY CR are shown in the bottom row of Fig. 4.22 in the low (left) and high (right) MET bins. In the post-fit plots both statistical and systematic uncertainties are included and the total uncertainty band is plotted. A scale factor of ~ 1.4 is found between the pre-fit and post-fit normalization of the DY in the low MET bin and ~ 1.3 in the high MET bin when data and simulation of the three years are combined.

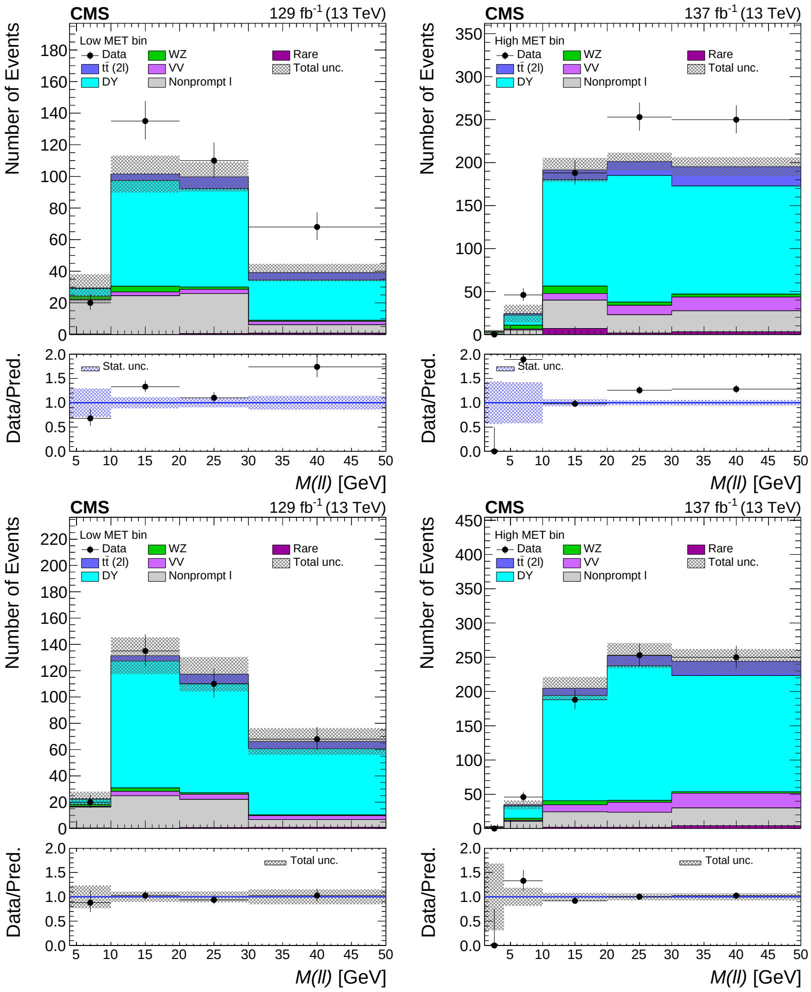

Figure 4.22: The $M(\ell\ell)$ distribution is shown for the low (left) and high (right) MET bins for the DY CR. The top row shows shows the pre-fit plots and the bottom row shows the post-fit plots.

Dilepton $t\bar{t}$ control region

The leptonic W boson decays in $t\bar{t}$ events can yield two prompt leptons in the final state. This SM background component is significantly suppressed in the SR by vetoing events that contain at least one b-tagged jet. The residual contribution is estimated from simulated events which are normalized to data in a dedicated CR which is defined in Table 4.10. Orthogonality to the search region selection is ensured by requiring at least one b-tagged jet with $p_T > 25$ GeV. Additionally, the upper cut on the M_T between the MET and the leptons and the upper p_T cut on the lepton p_T's are removed in order to increase the statistics.

This $t\bar{t}$ CR is also split in two MET bins according to the raw p_T^{miss} and p_T^{miss}

$\mathrm{t\bar{t}}$ control region
Triggers (Section 4.4)
$p_T^{miss}$and raw $p_T^{miss} > 125$ GeV
$(2/3) < (p_T^{miss}/H_T) < 1.4$
$H_T > 100$ GeV
Tight jet ID for ISR jet
$N_{\mathrm{b\ jets}} \geq 1$ (Medium WP)
$N_\ell = 2$
OS pair
SF pair (low MET)
$p_{\mathrm{T}}(\mu(e)) > 5$ (5) GeV (low MET)
$p_{\mathrm{T}}(\mu(e)) > 3.5$ (5) GeV (high MET)
$p_{\mathrm{T}}^{\ell\ell} > 3$ GeV
$4(1) < M(\ell\ell) < 50$ GeV (low(high) MET)
$\Delta R_{\ell\ell} > 0.3$ (high MET)
J/ψ veto: $M(\ell\ell) < 3$ or 3.2 GeV< $M(\ell\ell)$ (high MET)
Υ veto: $M(\ell\ell) < 9$ or 10.5 GeV< $M(\ell\ell)$
$m_{\tau\tau} < 0$ or 160 GeV< $m_{\tau\tau}$

Table 4.10: Event selection criteria for dilepton events of the $\mathrm{t\bar{t}}$ CR.

criteria. Plots in the top row in Fig. 4.23 show the $M(\ell\ell)$ distribution in the $\mathrm{t\bar{t}}$ CR pre-fit in the low (left) and high (right) MET bins. In the pre-fit plots only the statistical uncertainty is included and thus the total uncertainty band and the statistical uncertainty band coincide. The bottom row of the same Figure shows the $\mathrm{t\bar{t}}$ CR plots after the normalization has been fixed by the fit or post-fit plots. In these both statistical and systematic uncertainties are included and only the total uncertainty band is plotted. The data/simulation agreement seems quite good in general. The purity of the low-MET (high-MET) region is 90% (78%) as obtained from simulated events in the 2016 analysis and remains at this level for the other years as well. The $M(\ell\ell)$ distributions enter the final simultaneous fit with its data-to-simulation normalization factor and its uncertainty freely floating.

A scale factor of $\sim 0.9(1)$ is found between the pre-fit and post-fit normalization of the $\mathrm{t\bar{t}}$ in the low(high) MET.

4.7.2 Prompt diboson processes

The SM diboson production can yield final states with two or more prompt leptons. In the case of 2ℓ-ewk SR and 2ℓ-stop SR the most important diboson background stems from WW events while the dominant multiboson process in the 3ℓ-ewk SR arises from the WZ production. A dedicated validation region

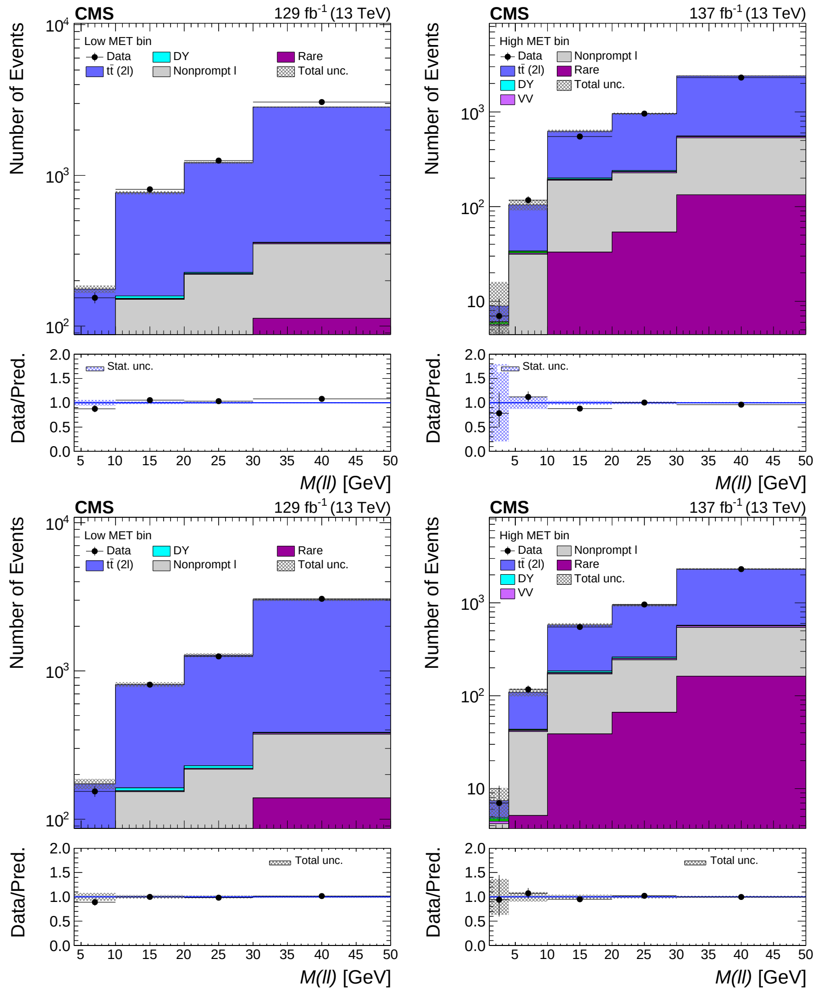

Figure 4.23: The $M(\ell\ell)$ distribution is shown for the low (left) and high (right) MET bins for the $t\bar{t}$ CR. The top row shows shows the pre-fit plots and the bottom row shows the post-fit plots.

(VR) enriched in VV processes, is used to compare the simulation to the data and validate the prediction. A dedicated CR is defined for the normalization of the WZ process. Below, the VV and WZ enriched regions are described in detail. The former is defined by explicitly requiring 2ℓ in the final state search, while the latter is defined with the 3ℓ selection. The VV VR and the WZ CR are described below.

VV validation region

Double boson, VV, production accounts for the mixture of WW, WZ (not fully leptonic) and ZZ events, where the processes are given in descending contributing order. The event selection criteria that define the VV VR are presented in Table 4.11. The selection is based on the 2ℓ SR selection, inverting the M_{T} requirement and requiring a high p_{T} leading lepton ($p_{\mathrm{T}} > 30$ GeV) in order to ensure orthogonality to the SR.

VV validation region
Triggers (Section 4.4)
$p_T^{miss}$and raw $p_T^{miss} > 125$ GeV
$(2/3) < (p_T^{miss}/H_T) < 1.4$
$H_T > 100\,\mathrm{GeV}$
Tight jet ID for ISR jet
$N_{\mathrm{b\ jets}} = 0$ (Loose WP)
$N_\ell = 2$
OS pair
SF pair (low MET)
$p_{\mathrm{T}}(\ell_1) > 30$ GeV
$p_{\mathrm{T}}(\ell_2) > 5$ (5) GeV (low MET)
$p_{\mathrm{T}}(\mu_2(e_2)) > 3.5$ (5) GeV (high MET)
$p_{\mathrm{T}}^{\ell\ell} > 3$ GeV
$4(1) < M(\ell\ell) < 50$ GeV (low(high) MET)
$\Delta R_{\ell\ell} > 0.3$ (high MET)
J/ψ veto: $M(\ell\ell) < 3$ or 3.2 GeV$< M(\ell\ell)$ (high MET)
Υ veto: $M(\ell\ell) < 9$ or 10.5 GeV$< M(\ell\ell)$
$M_{\mathrm{T}}(\ell_i, p_T^{miss}) > 90$ GeV ($i = 1$ *or* 2)
$m_{\tau\tau} < 0$ or 160 GeV$< m_{\tau\tau}$

Table 4.11: Event selection criteria for dilepton events entering the VV VR.

Events passing the selection are categorized into two MET bins. The VV VR is designed to be as close as possible to the SR phase space. The purity of this selection is $\sim$ 20%($\sim$ 30%)in the low (high) MET selection in 2016 (and comparable to that in the other years), which is why the VR is not used in the final fit. Instead a systematic uncertainty of 50% arising from the maximum data-to-simulation discrepancy in the VV VR is assigned to the VV background.

Figure 4.24 illustrates the pre-fit $M(\ell\ell)$ distribution in the VV VR low and high MET bins. In the plots only the statistical uncertainty is included and thus the total uncertainty band and the statistical uncertainty band coincide.

WZ-enriched region

In order to assess and normalize the prediction of the SM WZ (fully leptonic) production, which is important in the 3ℓ-ewk SR, a WZ-enriched region is employed in two MET bins. The event selection is summarized in Table 4.12 and ensures a

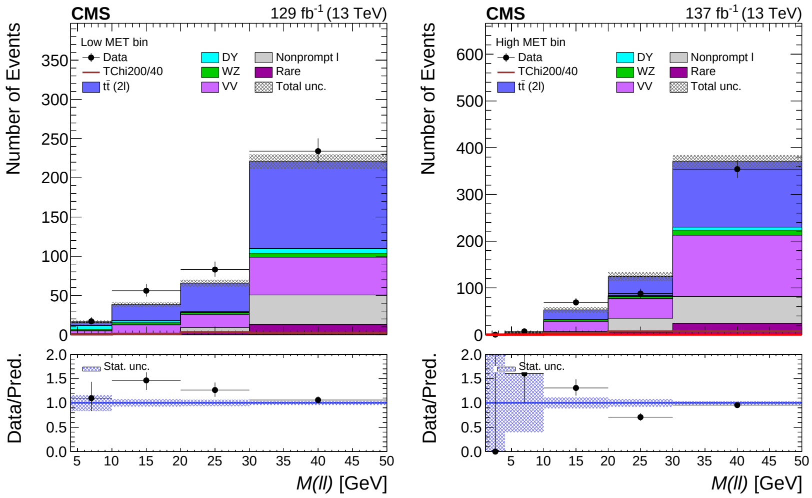

Figure 4.24: The $M(\ell\ell)$ distribution is shown for the low (left) and high (right) MET bins, pre-fit, in the VV VR.

purity in SM WZ events of $\sim 80\%$ ($\sim 90\%$) for the low (high) MET bin of 2016. Similar purities are achieved in the other years as well. The selection criteria are designed based on the 3ℓ-ewk SR selection with some important differences.

Firstly, the upper bound of the invariant 2ℓ mass is removed, such that events with $M(\ell\ell) > 50$ GeV can enter the WZ-enriched region. The is justified by the fact that in the SM WZ production the $M(\ell\ell)$ is expected to be close to the pole mass of the Z boson at 91.2 GeV. However, the double-μ plus p_T^{miss} trigger used in the low MET SR bin, has an upper bound on the dimuon invariant mass at 60 GeV and therefore, cannot be used for triggering events in the WZ CR. As a result, a pure double-μ trigger has been used for the online event selection in the low MET bin of the WZ-enriched region. In the high MET bin the pure p_T^{miss} HLT paths, described in Section 4.4, are used similarly to the other CRs.

The WZ-enriched region is defined such that the leading lepton has $p_{\mathrm{T}} >$ 30 GeV, which is motivated by the fact that the final state leptons of SM WZ events are generally harder than the signal soft leptons. This requirement ensures orthogonality between the WZ-enriched region and the SR. The subleading and trailing leptons should have $p_{\mathrm{T}} > 10$ GeV (15 GeV if electrons in the high MET bin). The offline lepton p_{T} requirements should also follow the online selection. The online-p_{T} thresholds of the pure double-μ paths are at 17 and 8 GeV, therefore the offline selection requires at least one muon with $p_{\mathrm{T}} > 20$ GeV and one muon with $p_{\mathrm{T}} > 10$ GeV in the low MET bin.

The low $M(\ell\ell)$ range of the WZ-enriched region has some non-negligible contribution from the signal points with intermediate and higher mass splittings ($30 - 40$ GeV) with respect to the overall WZ contribution. The effect of this at-

WZ-enriched region
Triggers (Section 4.4)
$p_T^{miss}$and raw $p_T^{miss} > 125$ GeV
$H_T > 100$ GeV
$N_{\mathrm{b\ jets}} = 0$ (Medium WP)
$N_\ell = 3$
One OSSF pair
$p_\mathrm{T}(\ell_1) > 30$ GeV
At least two μ, one of which with $p_\mathrm{T} > 20$ GeV (low MET)
$p_\mathrm{T}(\mu_i(e_i)) > 10$ (15) GeV ($i = 2, 3$)
4(1)$<M(\ell\ell)^{min}_{SFOS}$(low(high) MET)
$\Delta R_{\ell\ell} > 0.3$ (high MET)
J/ψ veto: $M(\ell\ell) < 3$ or 3.2 GeV$< M(\ell\ell)$ (high MET)
Υ veto: $M(\ell\ell) < 9$ or 10.5 GeV$< M(\ell\ell)$

Table 4.12: Selection criteria for 3ℓ events entering the WZ enriched control region.

tribute will be discussed in Sec. 5.2 and the corresponding plots will be presented. Therefore, the $1 < M(\ell\ell) < 30$ GeV region of the WZ-enriched region contributes to the sensitivity of the analysis and is defined as “WZ-like selection SR”. The $30 < M(\ell\ell)$ GeV region, which includes the majority of the WZ process yields, is defined as “WZ CR”. The plots in Fig. 4.25 show the $M(\ell\ell)$ distribution in the WZ-enriched region, pre-fit (post-fit) in the top (bottom) row. The WZ-enriched region enters the final fit and the data-to-simulation normalization factor for the WZ background as well as its systematic uncertainty is left to be determined by the final fit.

4.7.3 Non-prompt leptons

The terms fake or non-prompt refer to leptons that are non isolated or they are produced away from the primary vertex or they are mimicking the real leptons. For electrons, the main sources of non-prompt or fake leptons arise from photon conversions, semi-leptonic heavy flavour decays and misidentification of charged hadrons. In the case of the muons, the non-prompt background originates mostly from in-flight decays of mesons (charged kaons or pions) and semi-leptonic heavy flavour decays. A smaller contribution comes from punch-through jets. The non-prompt or fake background will be collectively called non-prompt background for the rest of the thesis.

The main non-prompt background in the 2ℓ SR arises from W+jet events in which one prompt lepton is coming from the W boson and one non-prompt from the jets. In the 3ℓ-ewk SR final state the main source of non-prompt background is the dileptonic $\mathrm{t\bar{t}}$ +jet process, where two prompt leptons come from the leptonic

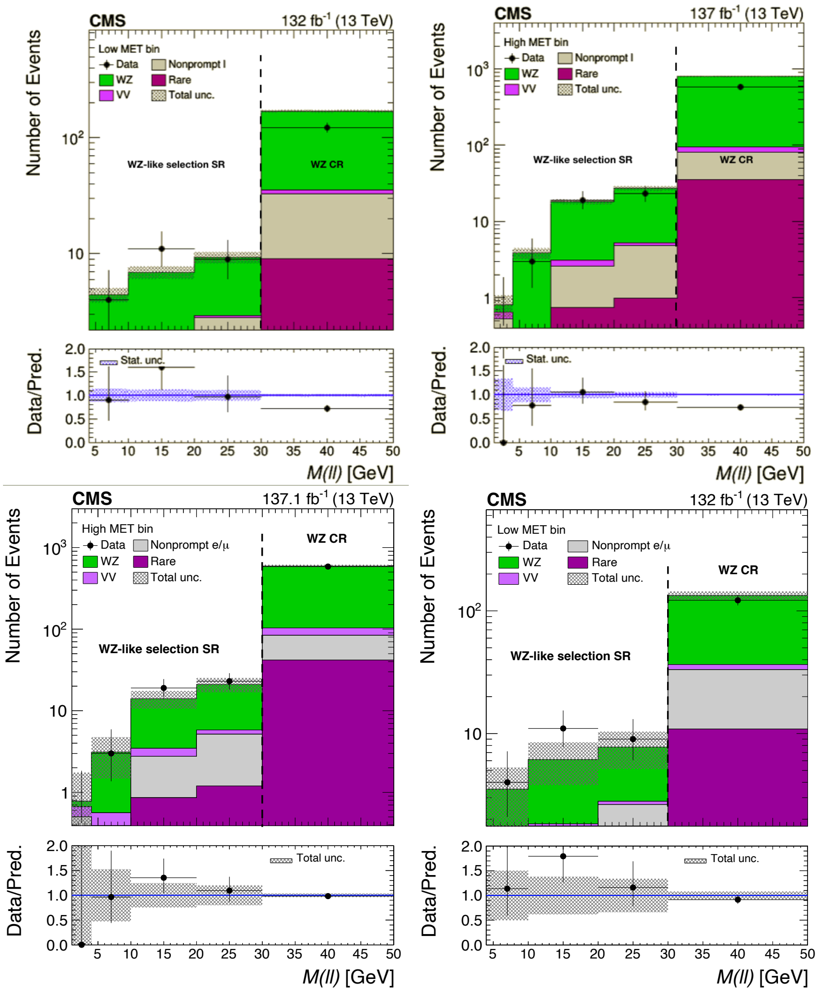

Figure 4.25: The $M(\ell\ell)$ distribution is shown for the low (left) and high (right) MET bins for the WZ-enriched region. The top row shows the pre-fit and the bottom row shows the post-fit plots.

W decays from the two top quarks and one non-prompt lepton comes from the jets.

The background from non-prompt leptons is evaluated with the DD "tight-to-loose" method [178], which is widely used in the CMS analyses. The method is based on the definition of three independent regions. The first is the SR described in Sec. 4.6, in which the non-prompt background yield needs to be estimated. The second is called "application region" (AR) and it is a non-prompt enriched region with the same kinematic cuts as the SR but dominated by events with at least one lepton that fails the tight ID selection while passes the loose identification selection. The extrapolation from the AR to the SR is done by weighting the AR events with a transfer factor. The latter depends on the probability of a non-prompt lepton to pass the tight ID selection. This missidentification probability is called fake rate and it is measured in data as a function of η and p_{T} in a QCD enriched region called “measurement region" (MR). The definition of the AR and MR and the details of the fake-rate measurement in data will be described in the following subsections.

Application region

The AR is defined with the kinematic cuts of the SR, employing the same MET binning scheme of 4 MET bins in the 2ℓ-ewk AR and 2ℓ-stop AR and 2 MET bins in the 3ℓ-ewk AR. However, the SR requirement of both leptons passing the tight ID selection is inverted. At least one of the leptons should fail the tight identification and isolation requirements, ensuring that the AR is enriched in non-prompt leptons and that it is orthogonal to the SR. The plots in Fig. 4.26-4.28 show the 2ℓ-ewk AR, 2ℓ-stop AR and 3ℓ-ewk AR respectively, comparing data to simulation. The MET binning is that of the respective SR.

The discrepancies between the data and simulation distributions, observed especially in the low MET bins of the AR, arise from missing MC background (eg. missing QCD). Semi-leptonic decays in jets can give a significant amount of non-prompt lepton pairs due to the large multijet cross sections, while MET comes mostly from jet mismeasurement and thus it is low, because in such events there are no real weak bosons producing energetic neutrinos. Due to analysis selection cuts, it is difficult to have a QCD simulation with sufficient statistics, to get some events in the analysis regions. Therefore the non-prompt background in the SR is estimated with methods that use the data. In the low MET bin a pure DD method in which the AR data events are weighted by the transfer factors for the estimation of the non-prompt background in the SR, is used. Therefore, the method relies exclusively on data for the non-prompt background estimation and the missing simulation sample does not affect the method in any way. A detailed example of how the DD method is used in the low MET bin of 2ℓ-ewk AR for the non-prompt background estimation, is described in Appendix B.1.

From the 3ℓ-ewk AR plots and the medium, high and ultra high MET bins of the 2ℓ-ewk AR it is apparent that the statistical power of the data is limited and as a result the DD method cannot be used for the fake background prediction in those SR bins. In order to overcome the restriction due to low data statistics, an alternative method is used for the non-prompt background estimation. Non-prompt simulation templates are scaled to data in the low and high 3ℓ-ewk AR and the medium, high and ultra high MET bins of the 2ℓ-ewk AR. The transfer factors are applied to the scaled simulation templates in order to estimate the non-prompt background in the SR. This is called the semi-DD method for the non-prompt background estimation. The scaling of the non-prompt simulation

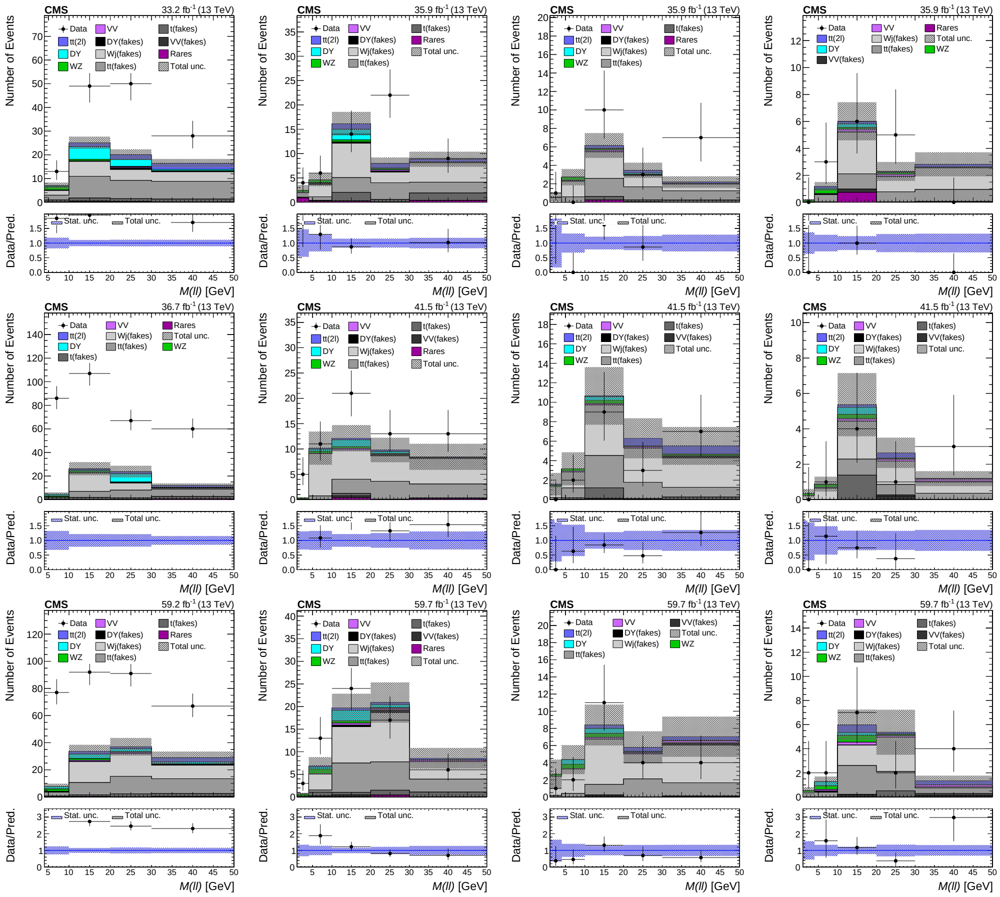

Figure 4.26: The $M(\ell\ell)$ distribution in 2ℓ-ewk AR plots in 2016 (top), 2017 (middle) and 2018 (bottom) data and simulation. From left to right the plots show the low, medium, high and ultra high MET bins.

templates to data is done with the semi-DD scale factor (semi-DD SF). These are measured separately in the three years, in sidebands (1LooseNotTightLepton, 2LooseNotTightLeptons, 3LooseNotTightLeptons in the 3ℓ search) of the medium, high and ultra high MET bins of the 2ℓ-ewk AR and the low and high MET bins of the 3ℓ-ewk AR. The semi-DD scale factors are defined as:

$$\text{Semi-dd SF} = \frac{\text{AR data } - \text{AR prompt simulation}}{\text{AR non prompt simulation}} \tag{4.11}$$

Tables 4.13-4.15 present the semi-DD SF, calculated in the 2ℓ-ewk AR, 2ℓ-stop AR and 3ℓ-ewk AR sidebands.

The plots in Fig. 4.29-4.31 present the $M(\ell\ell)$ and p_{T} distributions in the medium, high and ultra high MET bins of the 2ℓ-ewk AR and 2ℓ-stop AR respectively and the minimum $M(\ell\ell)$ distribution of a same flavor and opposite

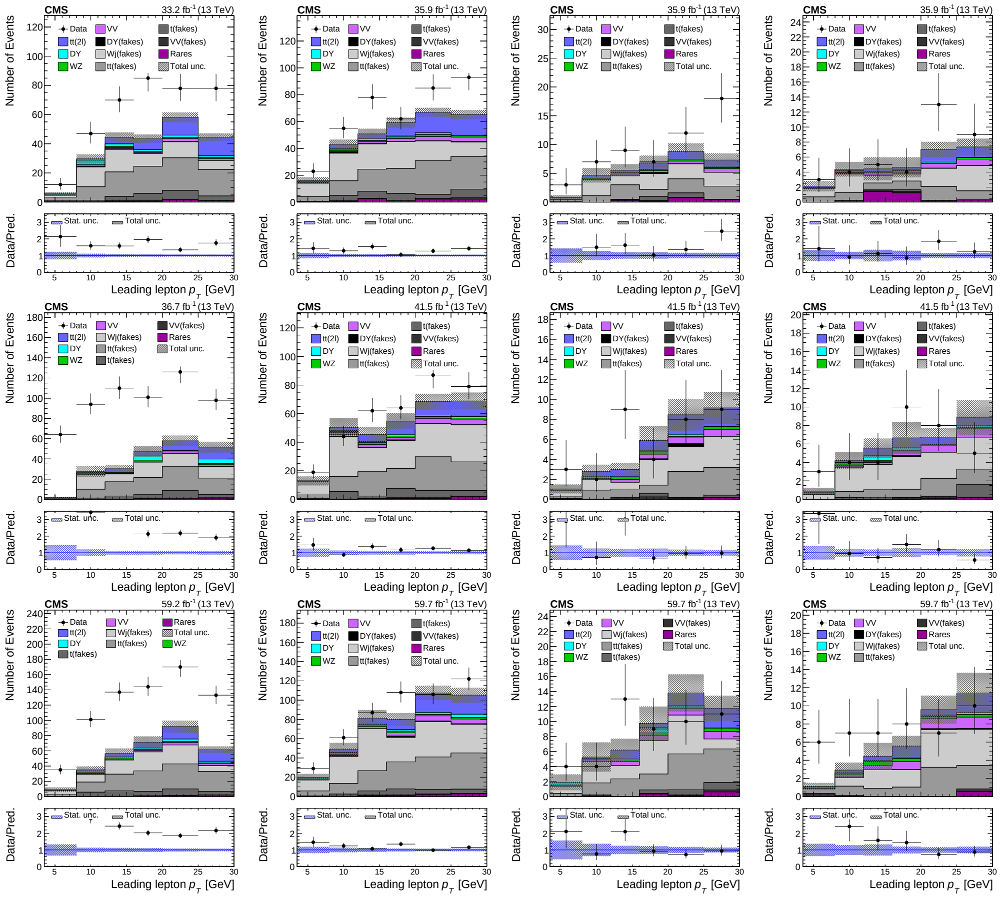

Figure 4.27: The leading lepton p_{T} distribution in the 2ℓ-stop AR, in 2016 (top), 2017 (middle) and 2018 (bottom) data and simulation. From left to right the plots show the low, medium, high and ultra high MET bins.

Semi-dd Scale Factors	2ℓ ewk 200-240 GeV	2ℓ ewk 240-290 GeV	2ℓ ewk >290 GeV	2ℓ stop 200-240 GeV	2ℓ stop 240-290 GeV	2ℓ stop >290 GeV	3ℓ 125-200 GeV	3ℓ >200 GeV
1 LooseNotTight	1.70	1.28	1.21	1.34	1.97	1.39	1.66	1.31
2 LooseNotTight	0.55	3.21	0.95	1.71	1.34	1.33	1.64	0.63
3 LooseNotTight	-	-	-	-	-	-	6.12	0.00

Table 4.13: The 2016 semi-DD SF estimated in the 2ℓ-ewk AR (left), 2ℓ-stop AR (middle) and the 3ℓ-ewk AR sidebands (right).

sign lepton pair in the 3ℓ-ewk AR low and high MET bins. The non-prompt background in simulation has been scaled to data with the semi-DD SF.

Even after the scaling of the non-prompt simulation templates to data, there are low $M(\ell\ell)$ bins which are empty or with very small non-prompt yield, such

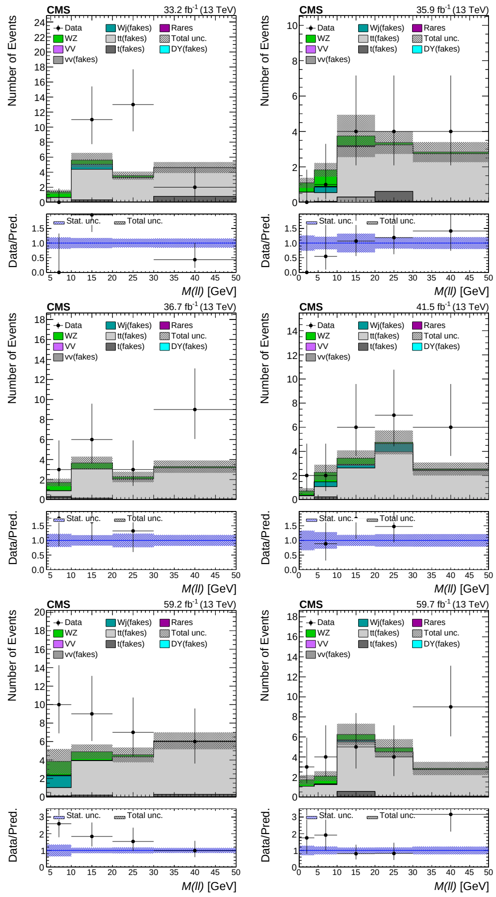

Figure 4.28: The minimum $M_{\ell\ell}$ of a same flavor and opposite sign lepton pair in 3ℓ-ewk AR plots. The left column shows the low and the right column the high MET bin, in 2016 (top), 2017 (middle) and 2018 (bottom) data and simulation.

Semi-dd Scale Factors	2ℓ ewk			2ℓ stop			3ℓ	
	200-240 GeV	240-290 GeV	>290 GeV	200-240 GeV	240-290 GeV	>290 GeV	125-200 GeV	>200 GeV
1 LooseNotTight	1.49	0.79	0.89	1.15	1.43	1.09	1.22	1.72
2 LooseNotTight	2.69	0.27	0.46	1.48	0.34	0.63	3.22	1.86
3 LooseNotTight	-	-	-	-	-	-	0.00	0.96

Table 4.14: The 2017 semi-DD SF estimated in the 2ℓ-ewk AR (left), 2ℓ-stop AR (middle) and the 3ℓ-ewk AR sidebands (right).

as the $M(\ell\ell)$ 1-4 GeV in the medium MET bin in 2017. For the non-prompt background estimation in the SR the non-prompt simulated yields, scaled to data, will be multiplied by transfer factors and this can lead to SR $M(\ell\ell)$ bins with negative non-prompt yield as will be discussed below. In order to overcome this restriction and exploit the full statistical power of the non-prompt simulation, the

Semi-dd Scale Factors	2ℓ ewk 200-240 GeV	240-290 GeV	>290 GeV	2ℓ stop 200-240 GeV	240-290 GeV	>290 GeV	3ℓ 125-200 GeV	>200 GeV
1 LooseNotTight	1.19	0.78	0.94	1.23	0.72	1.63	1.81	1.09
2 LooseNotTight	0.79	0.51	3.08	1.06	3.47	0.64	1.96	1.87
3 LooseNotTight	-	-	-	-	-	-	0.99	2.89

Table 4.15: The 2018 semi-DD SF estimated in the 2ℓ-ewk AR (left), 2ℓ-stop AR (middle) and the 3ℓ-ewk AR sidebands (right).

$M(\ell\ell)$ shape of the non-prompt simulation in the 2ℓ-ewk AR will be considered in an inclusive MET>200 GeV bin. This "merging" of the MET bins will not be applied in the 2ℓ-stop AR bins, which have higher statistics and therefore they do not result in SR bins with negative non-prompt yield.

The $M(\ell\ell)$ shape invariance in MET was studied by relaxing the lepton ID selection and comparing the shapes of the non-prompt simulation across the 2ℓ-ewk AR MET bins. The left plot in Figure 4.32 shows the $M(\ell\ell)$ distribution of the non-prompt simulation in different AR MET bins. The right plot shows the ratio of the medium MET $M(\ell\ell)$ distribution over the average of the $M(\ell\ell)$ distributions in the other two MET bins. A linear function $f(x) = p_0+p_1\times(\bar{x}-x)$ is fitted to the points, where $\bar{x}$ is the weighted mean of the distribution and it is equal to 18.1 in this case. The fit results in $c_0 = p_0 + p_1 \times \bar{x} = 1.010 \pm 0.15$ for the constant term and $c_1 = p_1 = 0.004 \pm 0.007$ for the slope. The plot of the $M(\ell\ell)$ shapes and the compatibility of the constant term of the fit with unity within the statistical uncertainty, ensures the invariance of the $M(\ell\ell)$ with MET. The slope term is compatible with zero but its uncertainty is used to assign a dedicated systematic uncertainty to the non-prompt background in the medium, high and ultra high MET bins of the 2ℓ-ewk SR, to account for the $M(\ell\ell)$ shape extrapolation between the MET bins.

The MET inclusive non-prompt simulation templates are weighted by a rate factor which accounts for their normalisation to the non-prompt simulation of every MET bin. The rate factor is defined as the ratio

$$\text{Rate factor} = \frac{\text{non-prompt simulation in MET bins}}{\text{non-prompt simulation in merged MET bin}} \tag{4.12}$$

The total weight factor applied on the non-prompt simulation of 2ℓ-ewk AR medium, high and ultra high MET bins, includes the semi-DD scale factor for the normalisation of the non-prompt simulation templates of every AR MET bin to the data, and the rate factor for the normalisation of the MET inclusive AR non-prompt simulation templates to the non-prompt simulation of every MET bin. The total weight factor is estimated in the AR sidebands and it is calculated as

$$\text{Total weight factor} = \text{semi-DD SF} * \text{Rate factor} \tag{4.13}$$

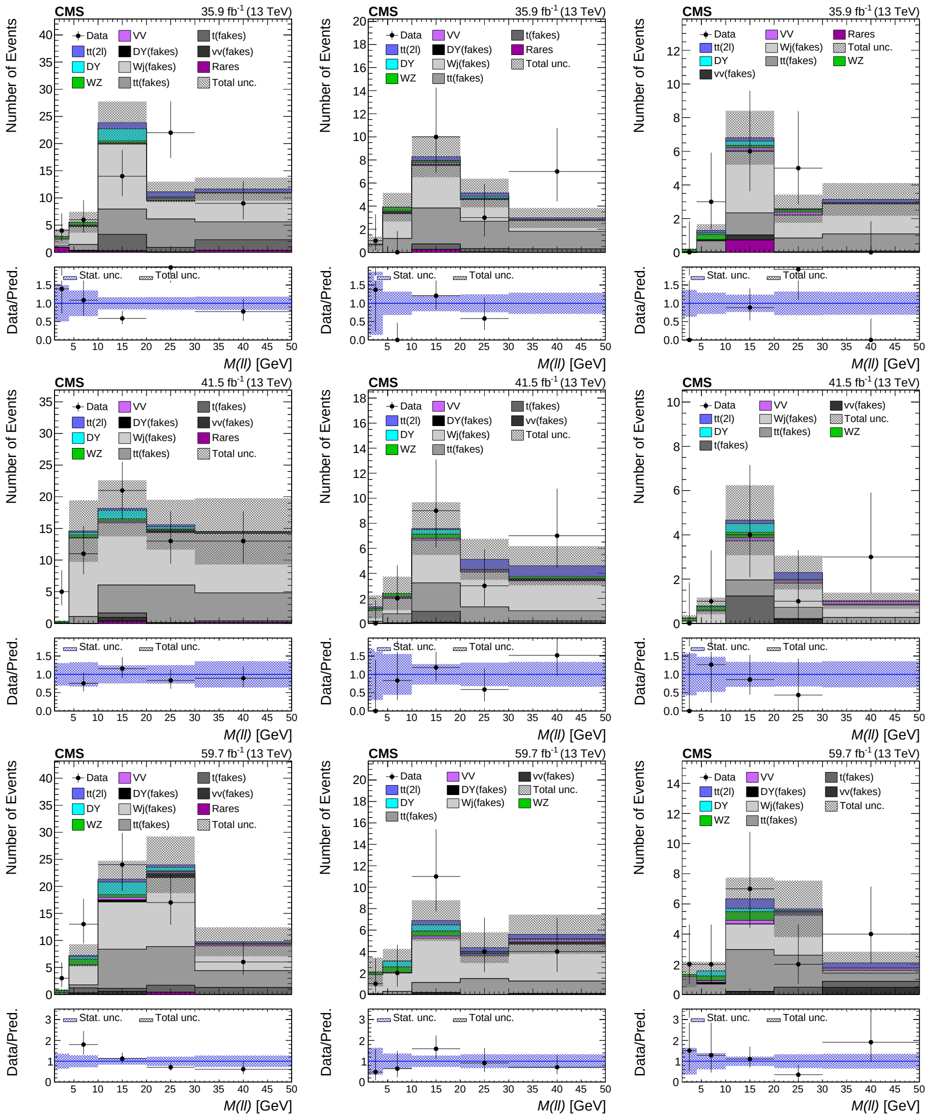

Figure 4.29: The $M(\ell\ell)$ distribution with non-prompt simulation scaled to data with the semi-DD scale factor, in the 2ℓ-ewk AR. The leftmost column plots show the medium MET bin, the middle column shows the high MET bin and the rightmost column shows the ultra high MET bin. From top to bottom the 2016, 2017 and 2018 plots are shown.

Table 4.16 presents the total weight factors estimated for the 2ℓ-ewk AR, medium, high and ultra high MET bins.

Figure 4.33 illustrate the 2ℓ-ewk AR in the 3 MET bins and the three years separately. The non-prompt simulation templates are taken from an inclusive MET bin (MET$>$ 200 GeV) and weighted by the total weight factor.

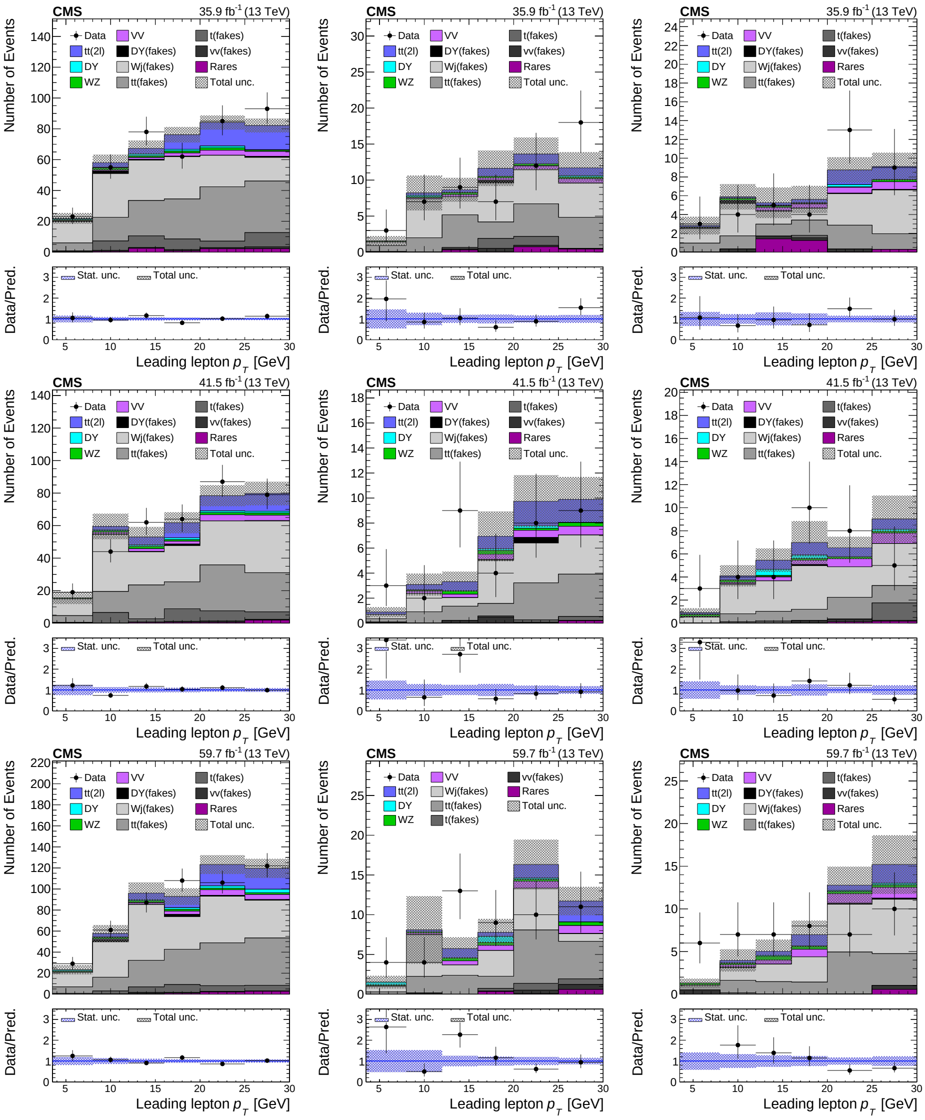

Figure 4.30: The leading lepton p_{T} distribution with non-prompt simulation scaled to data with the semi-DD scale factor, in the 2ℓ-stop AR. The leftmost column plots show the medium MET bin, the middle column shows the high MET bin and the rightmost column shows the ultra high MET bin. From top to bottom the 2016, 2017 and 2018 plots are shown.

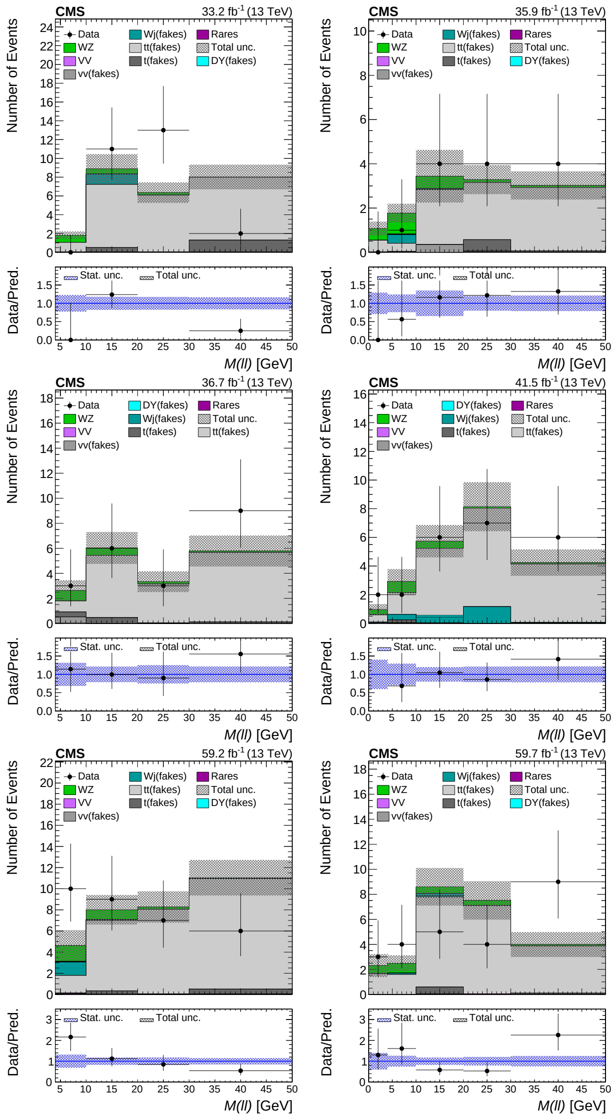

Figure 4.31: The minimum $M(\ell\ell)$ of a same flavor and opposite sign lepton pair distribution with non-prompt simulation scaled to data with the semi-DD scale factor, in the 3ℓ-ewk AR. The left column plots show the low MET bin and the right column shows the high MET bin. From top to bottom the 2016, 2017 and 2018 plots are shown.

It must be noticed that using the non-prompt simulation templates in the inclusive MET bin, with the total weight factor applied, reduces significantly the statistical uncertainty of the non-prompt prediction in the AR. Additionally, bins which previously had zero MC fake prediction (such as the $M(\ell\ell)$ 1-4 GeV bin in medium MET AR Figure 4.29) now have a positive MC fake prediction.

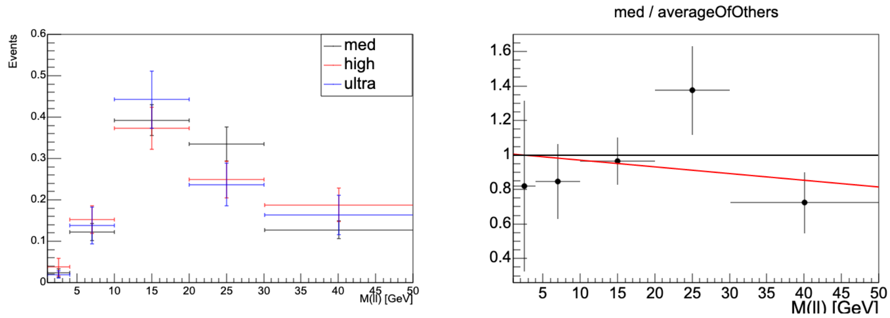

Figure 4.32: The $M(\ell\ell)$ shape of simulated non-prompt events in the medium, high and ultra high MET bins of the AR (left). The ratio of medium MET $M(\ell\ell)$ over the average $M(\ell\ell)$ of the high and ultra high MET and a linear fit to the points (right).

Total weight factor	2ℓ ewk 200-240 GeV			240-290 GeV			>290 GeV		
	2016	2017	2018	2016	2017	2018	2016	2017	2018
1 LooseNotTight	0.96	0.78	0.69	0.33	0.28	0.22	0.21	0.11	0.12
2 LooseNotTight	0.37	1.42	0.43	0.43	0.09	0.17	0.18	0.07	0.39

Table 4.16: The total weight factors estimated in the medium, high and ultra high MET bins of 2ℓ-ewk AR sidebands.

Measurement region

The probability of a non-prompt lepton to pass the tight ID, also referred to as fake rate, is measured in data, in a QCD enriched region called measurement region (MR). The fake rate is measured in p_T and η bins [8], to disentangle any p_T and η dependency, separately for muons and electrons. It is defined by requiring one lepton passing the loose identification selection and one jet with $p_T \geq 30$ GeV, separated from the lepton by a $\Delta R \geq 0.7$. For the measurement of the muon fake rate, events are selected online with prescaled single muons triggers with no isolation requirements. The `HLT_Mu3_PFJet40` is used for muons with $p_T < 10\,\text{GeV}$ and the `HLT_Mu8` for muons with $10 \leq p_T \leq 30\,\text{GeV}$. To regulate the bias between the two triggers, an offline PF Jet cut at 50 GeV is applied. For the electron fake rate, a combination of prescaled `HLT_PFJet` triggers is used. An offline PF Jet cut is applied at $p_T \geq 40\,\text{GeV}$.

[8] barrel $0 < \eta < 1.2$, endcap $1.2 < \eta < 2.4$

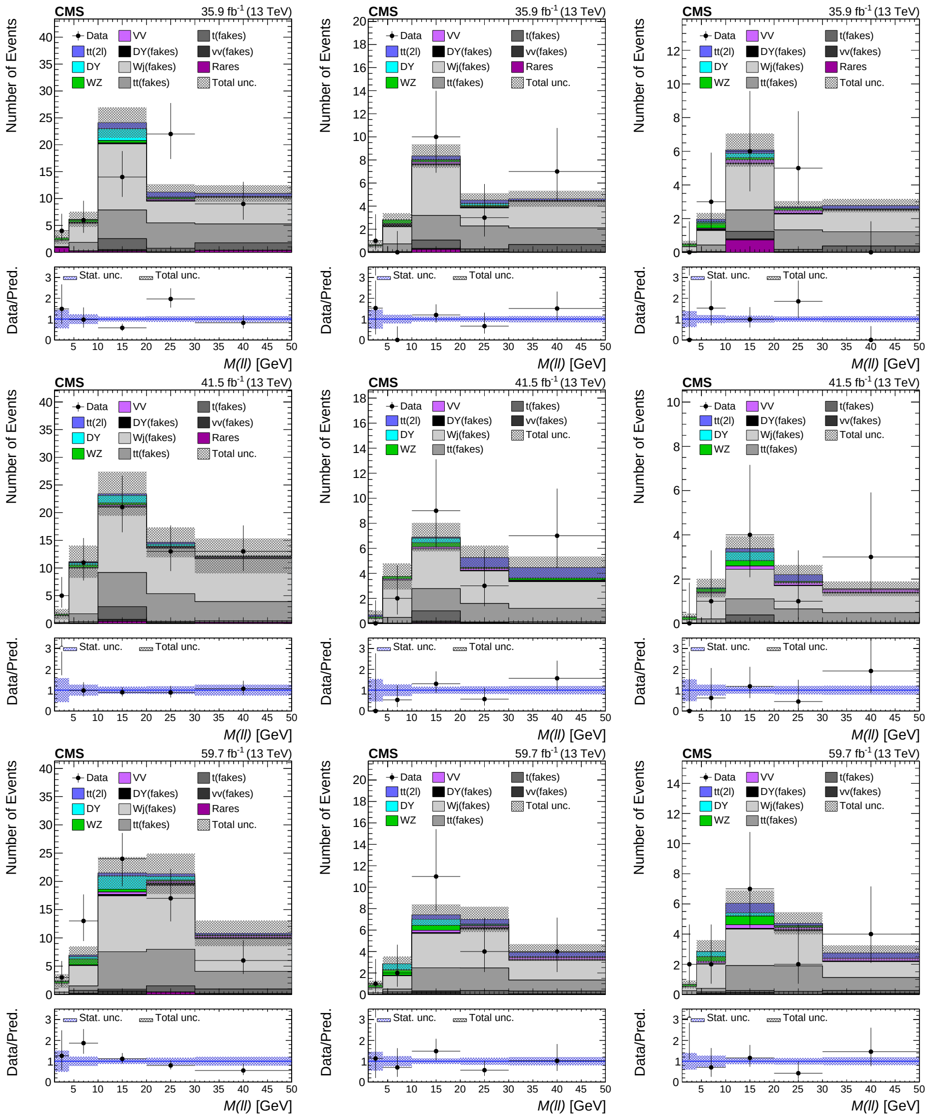

Figure 4.33: The $M(\ell\ell)$ distribution in the 2ℓ-ewk AR. The plots from left to right present data and simulation in medium, high and ultra high MET bins and from top to bottom in 2016, 2017 and 2018. The non-prompt simulated templates are merged in MET and scaled with the total weight factor.

Tuning of the lepton loose identification and the closure of the method

An important point of the DD non-prompt background prediction is that the fake rate is measured on QCD data events and applied on W+jets and $t\bar{t}$ events. Therefore, the jet flavour composition of the MR and the AR may differ and this can affect the prediction of the non-prompt background. As described briefly in 4.5.3, the Loose ID selection is a delicate aspect of the analysis as it affects the closure of the DD non-prompt background methods. In order to minimize the jet flavour dependency of the fake rate the LooseNotTight ID definition is tuned. In the end, the OR of the LooseNotTight ID and the Tight ID is used as lepton Loose ID selection. The tuning of the Loose ID was performed based on the agreement of the fake rate of jet originating from different flavor quarks (b-jets, c-jets and light jets), in QCD simulation.

It was found that removing the b tag veto and imposing an upper cut on the Iso_{Rel} at 1, in the LooseNotTight ID definition, could minimize the flavour dependency of the fake rate for muons. Due to the different sources of fake leptons in the case of electrons, an additional custom tuning of the SUSY Electron MVA ID, on top of the b tag cut removal and the $Iso_{Rel} \geq 1$, was needed. The loose identification definition cuts are collectively presented in Table 4.4. The results of the study are illustrated in Figures 4.34. The plots are made with 2018 QCD simulated events in the MR. The study had very similar results for 2016 and 2017. The event selection requires a lepton from b or c or light jets and an away jet with $p_{\mathrm{T}} \geq 30\,\mathrm{GeV}$ separated from the lepton with a $\Delta R \geq 0.7$ cut. Additionally, the events should pass the online event selection of `HLT_Mu3_PFJet40` (left), `HLT_Mu8` (middle) and electron PFJet triggers (right), separately in the barrel (top) and endcap (bottom). Good agreement is observed between the different jet flavour fake rates, both for muons and electrons, after applying the tuned Loose ID selection.

A closure test is performed by applying the fake rate measured in QCD simulation, on the non-prompt simulation events of the AR, and compare the result to the yields of the non-prompt simulation events that pass the Tight ID selection of the SR. The closure test is shown in Figure 4.35. The plots show kinematic variable distributions, inclusive in MET, in 2016 (left), 2017 (middle) and 2018 (right). The error band in the ratio plot shows the statistical uncertainty propagated to the ratio of non-prompt simulation over the AR non-prompt simulation weighted by the QCD MC fake rate. The maximum non-closure found to be $\sim 40\%$ for all the three years and it is used as a systematic uncertainty on the non-prompt background prediction in the final fit.

Measurement of the fake rate in data

An important aspect of measuring the fake rate in multijet data events is the contamination from prompt leptons, mostly originating from W and Z production in association with jets. The transverse mass (M_{T}) of leptons and MET is used as a discriminating variable to subtract such contamination, due to its different shape for the two processes. A slightly modified definition of M_{T} is used in which the lepton p_{T} is substituted with a constant value in order avoid bias in the measurement from correlations between the fake rate, that depends on p_{T}, and the fitted variable. The distribution of M_T^{fix} for QCD, W+jets and Z+jets events

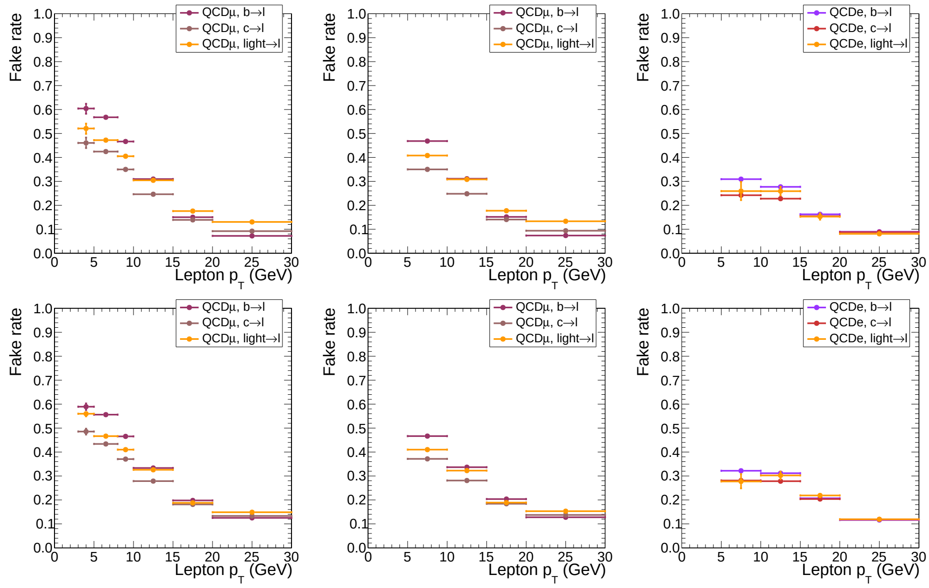

Figure 4.34: Fake rate plots made with 2018 QCD simulated events. The applied selection requires a lepton from b or c or light jets and an away jet with $p_{\mathrm{T}} \geq 30\,\mathrm{GeV}$ separated from the lepton with a $\Delta R \geq 0.7$ cut. Top (bottom) row shows the fake rate in the barrel (endcap). Plots in the left, middle and right column illustrate the fake rate measured with events that pass the online selection of `HLT_Mu3_PFJet40`, `HLT_Mu8` and electron `HLT_PFJet` respectively.

is illustrated in Fig. 4.36. The Z+jets processes are suppressed due to the MR selection requirement of exactly one lepton in the event. The M_T^{fix} is defined as:

$$M_T^{fix}(\ell, p_T^{miss}) = \sqrt{2 \cdot 35\ GeV\ p_T^{miss}(1 - cos\Delta\phi)} \tag{4.14}$$

For the prompt contamination subtraction three alternative methods are used. These methods are widely used in CMS SUSY analyses.

In the first method the QCD and V+jets simulation is normalised to the data by fitting M_T^{fix} templates on events that pass the tight ID selection (the numerator of the fake rate). A tight cut of $M_T^{fix} < 20$ GeV is applied, in order to profit from the difference in the M_T^{fix} distributions of the W/Z + jets and the QCD events. The fake rate is measured in p_{T} and η bins, by subtracting the residual prompt contamination from both the numerator and the denominator.

The second method relies on unfolding the QCD FR from two distinct regions of M_T^{fix}: one for small (S) values (0-20 GeV) and one for large values (L) (70-120 GeV). The procedure relies on two measurements of the fake rate in data, one in the S region (f_S) and one in the L region (f_L). Assuming the fake rate to be independent of M_T^{fix}, and taking the ratio of V+jets events expected in the

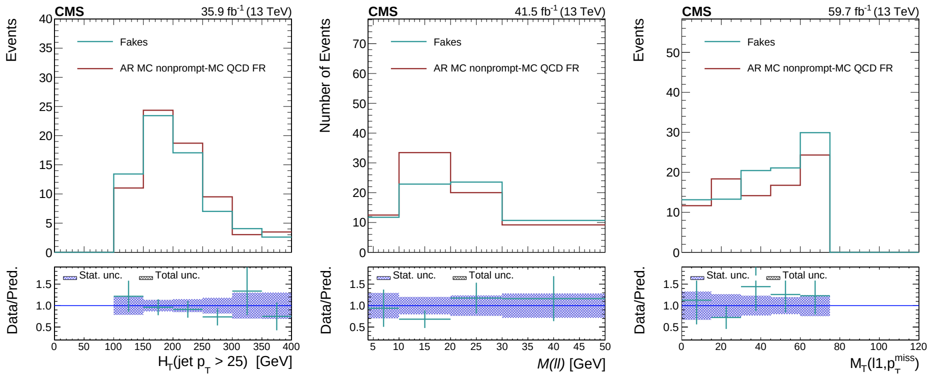

Figure 4.35: Closure tests of the fake rate method: comparison of non-prompt simulated events passing the Tight ID selection of the SR (blue line) to the prediction obtained by applying the fake rate (measured in QCD MC) on the AR simulated events (red line). The error band in the ratio plot shows the statistical uncertainty propagated to the ratio of non-prompt simulation over the AR non-prompt simulation weighted by the QCD MC fake rate. The plot on the left shows the jet H_T distribution in 2016, the middle plot shows the 2ℓ invariant mass in 2017 and the rightmost plot shows the $M_T(\ell_1\text{-}p_T^{miss})$ in 2018. The plots are inclusive in MET.

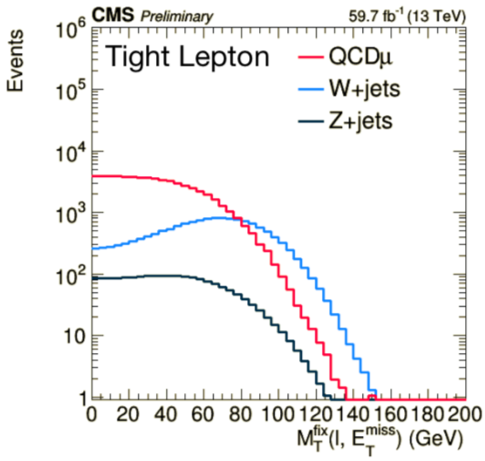

Figure 4.36: The M_T^{fix} distribution in the measurement region, plotted with leptons passing the tight ID selection.

two regions ($N^S_{V+jets}/N^L_{V+jets}$) from the simulation it is possible to unfold the QCD fake rate in p_{T} and η bins from:

$$f^{QCD} = \frac{f_S - f_L \cdot r^{SL}_{V+jets}}{1 - r^{SL}_{V+jets}} \tag{4.15}$$

where r^{SL}_{V+jets} is given by

$$r^{SL}_{V+jets} = (\frac{N^S_{V+jets}}{N^L_{V+jets}})/(\frac{N^S_{Data}}{N^L_{Data}}) \tag{4.16}$$

The third method of the prompt contamination subtraction relies on simultaneous fit of M_T^{fix} templates for the passing probes (Loose and Tight ID) and failing probes (LooseNotTight ID) events, using QCD and V+jets in p_{T} and η bins. After fitting both only the QCD MC templates are used for the fake rate measurement.

Within uncertainties, the three methods of the prompt subtraction agree well and they also agree with the fake rate measured in simulation. For the final measurement of the fake rate in data a combination of the three methods is used by taking as central value the weighted average and as uncertainty band the envelope of the three uncertainty bands. The results of the different methods are presented in Fig. 4.37. The top row shows the fake rate measured in the barrel and the bottom row shows the fake rate measured in the endcap. The leftmost, middle and rightmost columns illustrate the fake rate measured on muons with the `HLT_Mu3_PFJet40` election, muon with the `HLT_Mu8` selection and electron with the `HLT_PFJet` online selection respectively.

The final fake rate measurement in data after the prompt subtraction with the combination of the three methods is shown in Fig. 4.38. The plots show the fake rate measurements in data compared to that measured in simulation. The left column shows the fake rate in barrel and the right column shows the fake rate in endcap. The top row shows the muon fake rate measure and the bottom row shows the electron fake rate. The dashed line in the muon fake rate plots indicates where the switch between the `HLT_Mu3_PFJet40` and `HLT_Mu8` online trigger selection takes place. As an example only the 2016 fake rate plots are shown. The results are very similar for the rest of the years.

Tight-to-loose method

In the tight-to-loose method the non-prompt background is calculated in the SR by applying the transfer factor on the AR non-prompt events.

In the DD tight-to-loose method (also used in the 2016 analysis), all SR (tight-tight denoted as N_{PP}) and AR (at least one not tight, denoted as N_{PF}, N_{FP} and N_{FF}) data events, which can be directly measured, are used to determine the number of events with prompt-not prompt leptons (denoted as N_{01}, N_{10}) and not-prompt-not-prompt leptons (N_{00}), which cannot be directly measured.

The probability of one non-prompt lepton to pass the tight ID selection, the fake rate (f), is measured in data (described above), the probability of a non-prompt lepton to fail the Tight ID selection is 1-f. The probability of a prompt

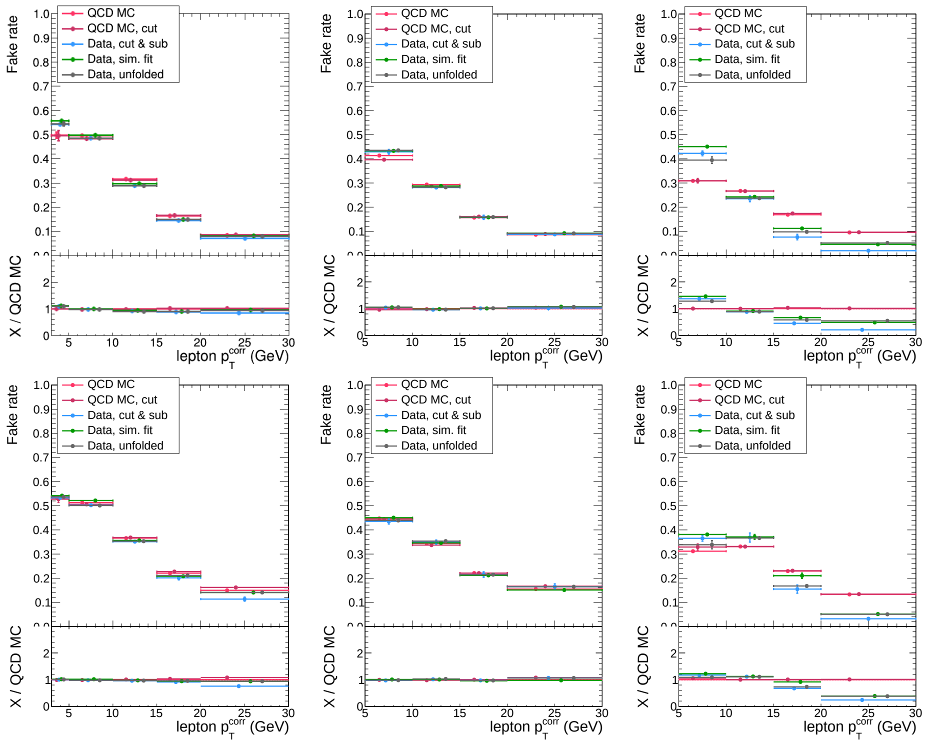

Figure 4.37: The fake rate measured with the three prompt subtraction methods as described in the text. The top row shows the fake rate measured in the barrel and the bottom row shows the fake rate measured in the endcap. The leftmost, middle and rightmost columns illustrate the fake rate measured with the `HLT_Mu3_PFJet40` in 2016, `HLT_Mu8` in 2017 and `HLT_PFJet` electron 2018 online selection.

lepton to pass the Tight ID selection is called the prompt rate (p) while the probability of a prompt lepton to fail the Tight ID is 1-p. The prompt rates (PR) as measured in simulation for prompt leptons from Z or W or τ are shown in Fig. 4.39.

In an example of events with only one lepton, the number of events with tight lepton is $N_P = pN_1 + fN_0$ and the number of events with no tight lepton is $N_F = (1-p)N_1 + (1-f)N_0$. Extending this to events with two leptons, the number of events with two tight leptons can be written as:

$$N_{PP} = p_1 p_2 N_{11} + p_1 f_2 N_{10} + f_1 p_2 N_{01} + f_1 f_2 N_{00} \tag{4.17}$$

The number of events with one tight lepton can be written as:

$$N_{PF} = p_1(1-p_2)N_{11} + p_1(1-f_2)N_{10} + f_1(1-p_2)N_{01} + (1-f_1)f_2 N_{00} \tag{4.18}$$

$$N_{FP} = (1-p_1)p_2 N_{11} + (1-p_1)f_2 N_{10} + (1-f_1)p_2 N_{01} + (1-f_1)f_2 N_{00} \tag{4.19}$$

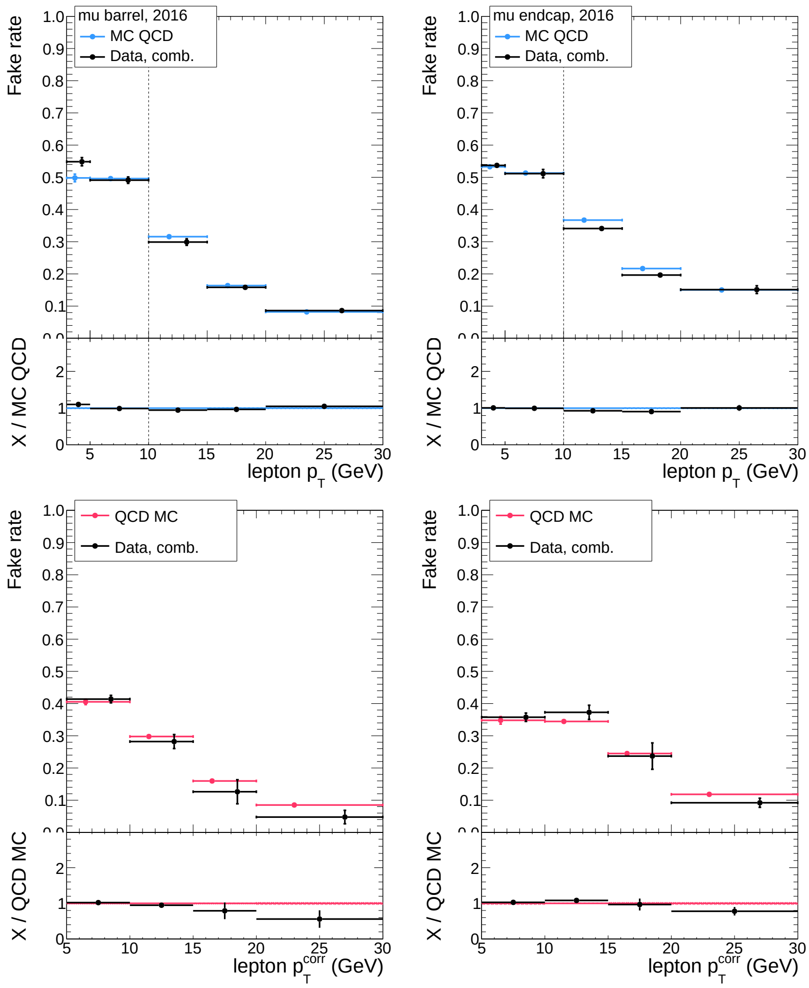

Figure 4.38: Fake rate measurement in data after prompt contamination subtraction with the combination of the three methods (black), compared to FR measured in QCD simulation (blue for muon, red for electrons). The left (right) column shows the fake rate measure in the barrel (endcap). The top row illustrates the muon fake rate and the bottom row the electron fake rate.

Finally the number of events with no tight lepton is:

$$\begin{aligned} N_{FF} = {} & (1-p_1)(1-p_2)N_{11} + (1-p_1)(1-f_2)N_{10} \\ & +(1-f_1)(1-p_2)N_{01} + (1-f_1)(1-f_2)N_{00} \end{aligned} \tag{4.20}$$

Eq. 4.17-4.20 can be written in matrix form as

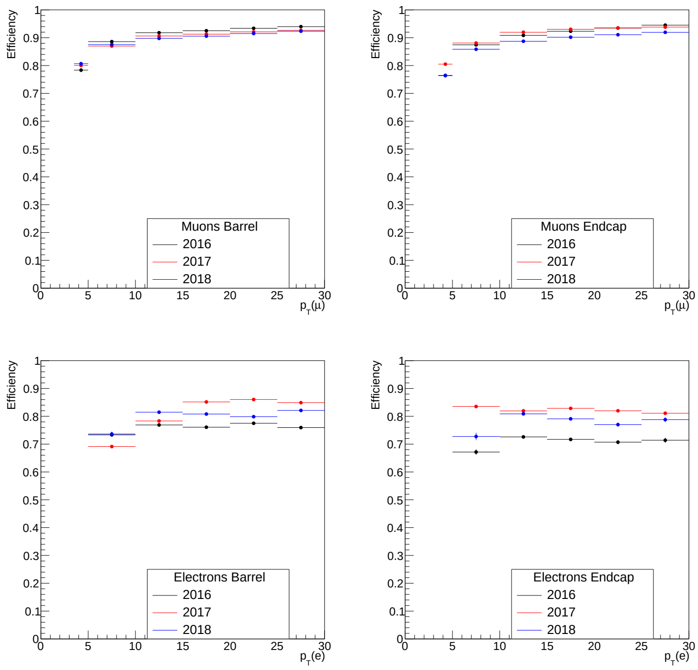

Figure 4.39: The prompt rates in p_{T} and η bins overlaid for the three years. Top row shows the prompt rate measure for muons and the bottom row for electors. The left column shows the prompt rate in barrel and the right column in endcap.

$$\begin{pmatrix} N_{PP} \\ N_{PF} \\ N_{FP} \\ N_{FF} \end{pmatrix} = \begin{pmatrix} p_1p_2 & p_1f_2 & f_1p_2 & f_1f_2 \\ p_1(1-p_2) & p_1(1-f_2) & f_1(1-p_2) & f_1(1-f_2) \\ (1-p_1)p_2 & (1-p_1)f_2 & (1-f_1)p_2 & (1-f_1)f_2 \\ (1-p_1)(1-p_2) & (1-p_1)(1-f_2) & (1-f_1)(1-p_2) & (1-f_1)(1-f_2) \end{pmatrix} \begin{pmatrix} N_{11} \\ N_{10} \\ N_{01} \\ N_{00} \end{pmatrix} \tag{4.21}$$

The number of events with non-prompt leptons in the SR (tight-tight) can be estimated by

$$N_{PP}^{bkg9} = p_1f_2N_{10} + f_1p_2N_{01} + f_1f_2N_{00} \tag{4.22}$$

The N_{10}, N_{01} and N_{00} can be estimated by inverting the matrix 4.21 and using all the SR events (two-tight) and AR events (at least 1 LooseNotTight) – N_{PP}, N_{PF}, N_{FP} and N_{FF}. The inverted equation is

$$\begin{pmatrix} N_{11} \\ N_{10} \\ N_{01} \\ N_{00} \end{pmatrix} = \frac{1}{(p_1-f_1)(p_2-f_1)} \begin{pmatrix} X & X & X & X \\ -(1-f_1)(1-p_2) & (1-f_1)p_2 & f_1(1-p_2) & -f_1p_2 \\ -(1-p_1)(1-f_2) & (1-p_1)f_2 & p_1(1-f_2) & -p_1f_2 \\ (1-p_1)(1-p_2) & -(1-p_1)p_2 & -p_1(1-p_2) & p_1p_2 \end{pmatrix} \begin{pmatrix} N_{PP} \\ N_{PF} \\ N_{FP} \\ N_{FF} \end{pmatrix} \tag{4.23}$$

The N_{11} is not relevant for the calculation events with at least one non-prompt lepton, as can be deduced from Eq. 4.22, therefore the corresponding line has been left out of the calculations.

From Eq. 4.23 the N_{10}, N_{01} and N_{00} are estimated as

$$\begin{aligned} N_{10} = \frac{1}{(p_1-f_1)(p_2-f_2)} \cdot [&-(1-f_1)(1-p_2)N_{PP} + (1-f_1)p_2N_{PF} \\ &+f_1(1-p_2)N_{FP} - f_1p_2N_{FF}] \end{aligned} \tag{4.24}$$

$$\begin{aligned} N_{01} = \frac{1}{(p_1-f_1)(p_2-f_2)} \cdot [&-(1-p_1)(1-f_2)N_{PP} + (1-p_1)f_2N_{PF} \\ &+p_1(1-f_2)N_{FP} - p_1f_2N_{FF}] \end{aligned} \tag{4.25}$$

$$\begin{aligned} N_{00} = \frac{1}{(p_1-f_1)(p_2-f_2)} \cdot [&-(1-p_1)(1-p_2)N_{PP} + (-1)(1-p_1)p_2N_{PF} \\ &+(-1)p_1(1-p_2)N_{FP} - p_1p_2N_{FF}] \end{aligned} \tag{4.26}$$

[9] Events with two tight leptons, at least one of them is non-prompt

From Eq. 4.22 and 4.24-4.26 the weights to be applied on the N_{PP}, N_{PF}, N_{FP} and N_{FF} events are calculated as

$$\begin{aligned} W_{PP} = \frac{1}{(p_1 - f_1)(p_2 - f_2)} \cdot [(-1)p_1 f_2 (1 - f_1)(1 - p_2) \\ +(-1) f_1 p_2 (1 - p_1)(1 - f_2) + (-1) f_1 f_2 (1 - p_1)(1 - p_2)] \end{aligned} \tag{4.27}$$

$$\begin{aligned} W_{PF} &= \frac{1}{(p_1 - f_1)(p_2 - f_2)} \cdot [p_1 f_2 (1 - f_1) p_2 \\ &\quad + f_1 p_2 (1 - p_1) f_2 + (-1) f_1 f_2 (1 - p_1) p_2] \\ &= \frac{1}{(p_1 - f_1)(p_2 - f_2)} \cdot [p_1 f_2 (1 - f_1) p_2] \end{aligned} \tag{4.28}$$

$$\begin{aligned} W_{FP} &= \frac{1}{(p_1 - f_1)(p_2 - f_2)} \cdot [p_1 f_2 f_1 (1 - p_2) \\ &\quad + f_1 p_2 p_1 (1 - f_2) + (-1) f_1 f_2 p_1 (1 - p_2)] \\ &= \frac{1}{(p_1 - f_1)(p_2 - f_2)} \cdot [f_1 p_2 p_1 (1 - f_2)] \end{aligned} \tag{4.29}$$

$$\begin{aligned} W_{FF} &= \frac{1}{(p_1 - f_1)(p_2 - f_2)} \cdot [(-1) p_1 f_2 f_1 p_2 \\ &\quad + f_1 p_2 (-1) p_1 f_2 + (-1) f_1 f_2 p_1 p_2] \\ &= \frac{1}{(p_1 - f_1)(p_2 - f_2)} \cdot [(-1) f_1 f_2 p_1 p_2] \end{aligned} \tag{4.30}$$

The weight factors defined in eq. 4.27-4.30 are also called transfer factors. This is the full DD tight-to-loose method that is used for the non-prompt background estimation in the low MET bin of the 2ℓ ewk and stop SR, the WZ enriched region and all the CR bins.

In the medium, high and ultra high MET bins of 2ℓ ewk and stop SR and the two MET bins of the 3ℓ SR the non-prompt background is estimated with the semi-DD method. It is noted that in the 2ℓ ewk AR medium high and ultra MET bins, the non-prompt simulation is taken from an inclusive MET bin and scaled by the total weight factor, which account for the normalization to data and the non-prompt background of every MET bin. In brief, this method uses the AR non-prompt simulation shapes scaled to data and weighted by the transfer factors, for the estimation of the non-prompt events in the SR. In the semi-DD method only the AR events of N_{PF}, N_{FP} and N_{FF} are used. The SR N_{PP} cannot be used in this method, since the SR simulation cannot be used as a template to be normalized to data. This is equivalent with assuming that $p > f$.

The equation 4.21 is significantly simplified by assuming that the contribution of prompt leptons failing the tight selection $(1-p)$ can be neglected with respect to the contribution of non-prompt leptons failing the tight selection $(1-f)$. Therefore $(1-p_1=0$ and $1-p_2=0)$. The simplified matrix equation is:

$$\begin{pmatrix} N_{PP} \\ N_{PF} \\ N_{FP} \\ N_{FF} \end{pmatrix} = \begin{pmatrix} p_1p_2 & p_1f_2 & f_1p_2 & f_1f_2 \\ 0 & p_1(1-f_2) & 0 & f_1(1-f_2) \\ 0 & 0 & (1-f_1)p_2 & (1-f_1)f_2 \\ 0 & 0 & 0 & (1-f_1)(1-f_2) \end{pmatrix} \begin{pmatrix} N_{11} \\ N_{10} \\ N_{01} \\ N_{00} \end{pmatrix} \tag{4.31}$$

The transfer factors to be applied on the AR N_{PF}, N_{FP} and N_{FF} are:

$$W_{PP} = 0 \tag{4.32}$$

$$W_{PF} = \frac{f_2}{1-f_2} \tag{4.33}$$

$$W_{FP} = \frac{f_1}{1-f_1} \tag{4.34}$$

$$W_{FF} = \frac{-f_1f_2}{(1-f_1)(1-f_2)} \tag{4.35}$$

The semi-DD method keeps the robustness from measuring the fake rates and the non-prompt simulation normalisation directly and purely from data, while using a minimum of information from simulation to reduce bin-by-bin statistical fluctuations.

To summarize, the DD method is used for the non-prompt background estimation in regions with large number of data events, such as the low MET bins of the 2ℓ ewk and stop SR, the WZ enriched region and all the MET bins of the DY and $t\bar{t}$ CRs. The semi-DD method is used in bins with low data yield such as the medium, high and ultra high MET bins of the 2ℓ-stop SR and both low and high MET bins of the 3ℓ-ewk SR. In the medium, high and ultra high MET bins of the 2ℓ-ewk SR the non-prompt background is estimated using the semi-DD method with MET-inclusive non-prompt simulation templates. In addition, the semi-DD method is used in the same-sign control region, defined above, which is used as a CR in the final fit to constrain better the fake background in the SR.

The same-sign control region

The semi-DD method is validated by comparing its performance to that of the DD method, in a dedicated CR. For this purpose a non-prompt enriched region is defined by using the 2ℓ-stop SR event selection in $p_T^{miss} > 200\,\text{GeV}$, where the opposite sign requirement is modified to same sign. This region can only be defined for $p_T^{miss} > 200\,\text{GeV}$ due to an opposite sign requirement of the double-μ plus p_T^{miss} online selection in low MET bin.

The comparison of the two methods is presented in Fig. 4.40. The plots show the SS CR, combining the three years, the non-prompt background is estimated with the DD method (left) and the semi-DD method (right). The comparison shows that the methods agree within their statistical uncertainty. In addition, a significant statistical uncertainty reduction is observed in the semi-DD fakes estimation method.

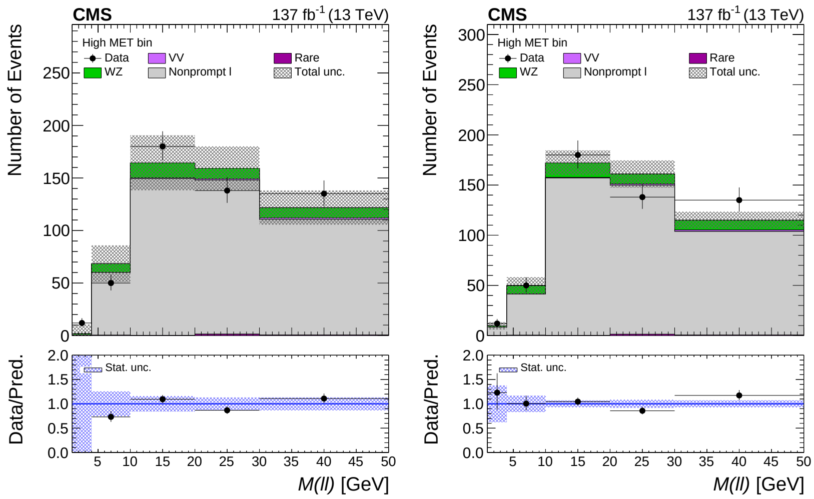

Figure 4.40: The pre-fit distribution of the $M(\ell\ell)$ variable in the high MET bin of the SS CR. The non-prompt background is estimated with the DD method in the left plot and the semi-DD method in the right plot. Uncertainties include only the statistical components.

The SS CR is used in the final fit to better constraint the contribution of the non-prompt background in the SR. The post-fit $M(\ell\ell)$ distribution in the SS CR is presented in Fig. 4.41. A scale factor of 0.94 is estimated for each year between the pre-fit and post-fit normalization of the non-prompt background prediction.

4.7.4 Other backgrounds

Other SM processes leading to signatures with 2ℓ or 3ℓ, such as tW, ttV, ttH, tZq, tWZ, VVV processes and processes involving the conversion of a photon, play a minor role in the background composition of the SR of the analysis. These processes are gathered together in the “Rares” category. This background component is estimated from simulated events and a 50% systematic uncertainty is assigned to its normalization.

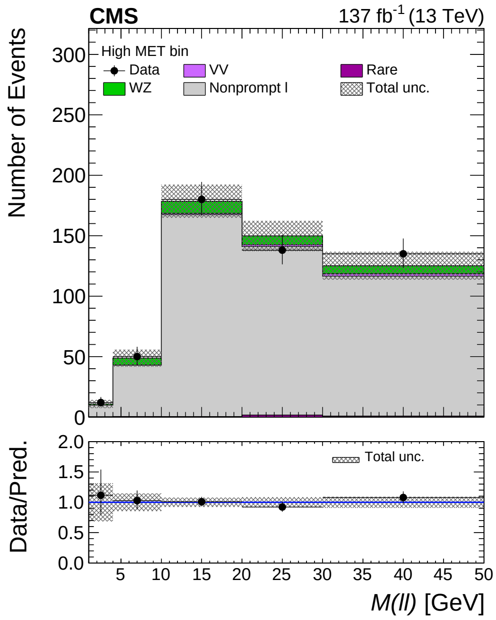

Figure 4.41: The post-fit distribution of the $M(\ell\ell)$ variable in the high MET bin of the SS CR. Uncertainties include both the statistical and systematic components.

4.8 Systematic uncertainties

The prediction of signal and background events of the SR suffers systematic uncertainties that can arise from various sources. Experimental uncertainties arise from the limited detector response and resolution and inefficiencies in the event reconstruction algorithms. Theoretical uncertainties account for the limited theoretical knowledge of important quantities for the modeling of signal and background processes. Other uncertainties are connected to the performance of the methods used for the background estimation.

Systematic uncertainties are applied either as normalisation uncertainties that affect the overall normalization of the process in question or as shape uncertainties which affect both the acceptance and shape of the kinematic distribution. The latter constitutes a recalculation of the yields taking into account a reduced or enhanced acceptance of the process and migration of events across the search bins.

This subsection covers all the the systematic uncertainties that were considered in the compressed mass SUSY search and their affect on the final prediction.

The **jet energy scale** (JEC) corrections account for any differences in the jet energy measurement between data and simulation. Small variations of the jet energy scale corrections are performed according to the official CMS recommendations. The variations of the predictions are included as a shape uncertainty in the final fit. They are applied on all the simulated processes, in all regions of the analysis. This JEC uncertainty results in 1-10% uncertainty on the final prediction.

Dedicated **b-tagging** scale factors are estimated and applied on all simulated events in all regions in order to harmonize the performance of the b-tagging algorithm between the data and the simulation. The scale factors are varied by their uncertainties and the results are included as shape uncertainties in the final fit. The b-tag uncertainty affects the final prediction by 1-4%.

Dedicated weights are applied on the simulation samples to correct the **number of vertices** (PU) distribution according to the one measured in data. The number of interactions per bunch crossing is estimated from the total inelastic cross section, which was measured to be $\sigma_{Tot\ Inel.} = 69.2\ mb$, with an uncertainty of 4.6% [179]. The total inelastic cross section uncertainty is propagated to the PU weights and their variations are used as a shape uncertainty, applied on all simulation, in all regions, uncorrelated across years. The post-fit effect on the final prediction is 1-2%.

The uncertainty on the **luminosity** measurements is applied as a flat number on all simulation, in all regions. The luminosity uncertainties are 2.5%, 2.3% and 2.5% for the 2016 [180], 2017 [180] and 2018 [181] respectively.

The impact of **ISR** is important in the compressed region of parameter space, as it is required to boost the final state and induce high MET. Differences in the ISR distribution between data and simulation is a source of systematic uncertainty. Year uncorrelated studies have been performed to address whether any additional corrections need to be applied to improve the ISR modelling. According to the SUSY PAG recommendations, the ISR modeling should be treated differently in 2016 and 2017/2018. For 2016 a p_{T} dependent reweighting should be applied on the ISR modeling. The deviation from unity is taken as a shape systematic uncertainty and its post-fit effect on the yields is 5%. On the other hand, for 2017 and 2018 the ISR modeling corrections are found to be very small (less than one percent difference from unity) therefore, a flat 1% systematic uncertainty is used.

In order to correct the simulated events for potential mismodeling of the **trigger** algorithm, data-to-simulation scale factors are applied to all the FullSIM events. The uncertainties of the trigger scale factors are included as shape uncertainties on the simulated events, in all regions, uncorrelated for every year. The uncertainties are further computed separately for the low and medium, high and ultra MET regions due to the different triggers and applied independently in the 2ℓ regions and in the 3ℓ regions in order to account for the additional lepton in the latter. The low MET WZ enriched region is treated separately from the rest of the 3ℓ regions due to the pure double-μ trigger algorithm, applied for the event selection in it. The pre-fit per-event uncertainty is estimated to be at most 5%. The post-fit effect of the uncertainty in the final prediction is 2-9%.

Scale factors are used to harmonize the performance of the **lepton ID** selection between data and simulation. The object reconstruction scale factors have been centrally produced, whereas the scale factors of the SOS Tight ID have been computed internally within the analysis group with the tag-and-probe method. The efficiency measurements and scale factors together with the per-bin uncertainties are presented in Section 4.5. The uncertainties on the lepton ID scale factors are applied as shape uncertainties on all the simulated events on the leading and subleading leptons in the 2ℓ region and additionally on the trailing lepton in the 3ℓ regions. The scale factors and uncertainties have been calculated separately for every year. The pre-fit effect on the total prediction doesn't exceed 5% and the post-fit effect on the final prediction is 2-9%.

During 2016 and 2017 a gradual shift in the timing of the ECAL trigger primitives towards early values was observed (prefiring). This caused a large fraction of electromagnetic objects with $|\eta| > 2.5$ to be associated to previous bunch crossing. Correction factors are applied to account for this and their uncertainty (1%) is propagated to the final result.

Regarding the prompt background, normalization and shape uncertainties are applied. The **DY, $\mathbf{t\bar{t}}$ and WZ (fully leptonic)** which are the dominant backgrounds for the 2ℓ and 3ℓ searches respectively, are estimated from simulated events whose prediction is validated to the data in the dedicated CR as described in Section 4.7. The $M(\ell\ell)$ distributions of these CR are included in the final fit together with the SR and an unconstrained scale factor is included as a nuisance parameter in the maximum likelihood fit, to correct the normalization of the simulation yields of each process to match the data. The normalisation of these processes are left to float freely and decided by the fit, independently for every year and every MET bin. The post-fit uncertainties vary from 15-35% in DY, $\sim 15\%$ in $\mathrm{t\bar{t}}$ and 12-27% in WZ.

The modeling of the **VV** background in the 2ℓ final state is validated in a dedicated VR and a flat systematic uncertainty of 50% is assigned to its prediction. In the 3ℓ final state, VV appears only in the form of ZZ and Z/γ^* and a flat uncertainty of 50% is assigned to their prediction.

In order to fully account for the uncertainty on the **rare** processes and given their minor contribution in the SR, a conservative flat systematic uncertainty of 50% has been assigned on this background prediction.

The residual **non-prompt** lepton background is estimated via the DD (in the CRs and the 2ℓ ewk and stop low MET SRs) and a semi-DD (for the rest of the SRs) methods. The performance of the tight-to-loose method is studied with a closure test presented in Section 4.7 and a normalization systematic uncertainty of 40% is assigned to the prediction of the non-prompt background. The latter is further constrained by the SS CR which is included in the final fit, binned in $M(\ell\ell)$ in a single MET bin ($p_T^{miss} > 200\,\mathrm{GeV}$). The post-fit effect of the non-prompt background uncertainty is 5%.

An additional shape uncertainty is applied to the non-prompt background to account for potential disagreements between data and simulation templates used in the ARs for the semi-DD method. The post-fit effect of this shape uncertainty on the non-prompt background is 8%. One extra shape uncertainty is applied on the non-prompt background yields of the SRs for which the corresponding AR distributions are merged across MET bins. The uncertainty accounts for minor shape discrepancies across MET bins. The post-fit effect of the latter shape uncertainties on the non-prompt background is approximately 5%.

Systematic uncertainties are additionally applied on the FastSIM samples (T2Bff, T2bW, Higgsino pMSSM). The triggers are not simulated for the FastSIM samples, so the simulations are corrected with the trigger efficiency measured in data instead of applying the trigger scale factors. The trigger efficiency uncertainties are taken into account.

In addition, the difference in the p_T^{miss} reconstruction between FullSIM and FastSIM is taken into account as an extra systematic uncertainty on the FastSIM sample and varies in the range of 1-10% post-fit.

All other centrally produced FastSIM related SF (b-tagging, JEC) are also applied, taking their uncertainty into account. Finally, for all the signal samples, the SUSY theory uncertainty on the signal cross section due to the variation of the parton density functions is included in the $\pm 1\sigma$ curves on the limit scans, and amounts to 3.5–8.5%.

All the systematic uncertainties presented above, are included as nuisance parameters in the maximum likelihood fit to the data. The dominant prompt background uncertainties are included as log-uniform distributed parameters, whereas all other uncertainties are included as log-normal distributed parameters. The log-normal is a continuous probability distribution of a random variable whose logarithm is normally distributed. The distribution is characterized by a parameter k and affects the expected yields in a multiplicative way. A positive $+1\sigma$ variation corresponds to yield scaled by k with respect to the nominal yield, while a negative -1σ variation correspond to a scaling by 1/k. The log-uniform is a uniform distribution between 1/k and k. The prompt background uncertainties are modeled with a log-uniform distribution and are left to float freely in the fit. The pre-fit values of the allowed variation range are set to be large in order to include the value determined by the fit. Theses normalisation are not constrained by a Gaussian distribution around a pre-fit value as it would be if they were modeled as log-normal nuisance parameter but instead they have equal probability to get calibrated at the proper value within the given range.

Table 4.17 presents the list of the systematic uncertainties, their post-fit effect on the yields, their treatment either as normalization or as shape uncertainties and the processes they are applied on. The uncertainties are broadly categorized into three groups. The first part of the table present the uncertainties that are applied on all simulated processes, the second part shows the uncertainties applied only on background processes and the last part of the table contains the uncertainties that are applied exclusively on the signal processes.

Source	Post-fit effect	Treatment	Non-prompt	MC background	MC Signal
Luminosity	2.3-2.5%	Norm.	–	✓	✓
PU	1-2%	Shape	–	✓	✓
JEC	1-10%	Shape	–	✓	✓
b-tagging	1-4%	Shape	–	✓	✓
Trigger SF	2-9 %	Shape	–	✓	✓
Lepton SF	2-9%	Shape	–	✓	✓
ISR	1-5%	Shape	–	✓	✓
Prefiring	1%	Shape	–	✓	✓
DY norm.	15-35%	Norm.	–	DY	–
$t\bar{t}$ norm.	15%	Norm.	–	$t\bar{t}$	–
WZ norm.	12-27%	Norm.	–	WZ	–
VV norm.	2%	Norm.	–	VV	–
Rares norm.	0.2%	Norm.	–	Rares	–
Non-prompt norm.	5%	Norm.	✓	–	–
Non-prompt MC template agreement to data	8%	Shape	✓	–	–
Non-prompt MET merging	5%	Shape	✓	–	–
Signal cross section	3.5-.5%	Shape	–	–	✓
FastSim corrections	1-10%	Shape	–	–	FastSim signal

Table 4.17: Systematic uncertainties and their post-fit effect on the yields. The first group shows the uncertainties applied on all simulated processes, the second group shows the uncertainties applied on background processes and the third group shows the uncertainties applied on the signal processes.

Chapter 5

Statistical interpretation and results

The search strategy for optimally selecting the events which are relevant to the search and rejecting the SM background in the SR has been described together with the analysis binning method on sensitive variables. Additionally, the methods for the prompt and non-prompt background prediction have been discussed in detail. Histograms of continuous observables such as the $M(\ell\ell)$ or the leading lepton p_{T} of each process expectation and the data are used to extract a result on the compatibility of the observed data with the hypothesis tests. To this end this is a binned shape analysis.

The main question in a search for new physics is weather the observed data can be explained purely by the SM background or if there is some space for contribution of new phenomena. These two scenarios comprise the background only or null hypothesis and the signal+background hypothesis respectively and they are denoted as H_0 and H_μ. The interpretations of the results constitutes a hypothesis testing which evaluate the compatibility of the two hypotheses with the observed data.

This chapter discusses the statistical interpretation in Sec. 5.1, the results of the search in the context of multiple SUSY models in Sec. 5.2 and the conclusions in Sec. 5.3.

5.1 Statistical Interpretation

5.1.1 Binned profile likelihood

In a SR the number of observed events in all the SR bins n are $\chi_1, ..., \chi_n$. In each bin i the observed data χ_i are distributed according to Poisson probability density function (pdf) is

$$f_i(\chi_i|\mu) = \frac{(\mu S_i + B_i)^{\chi_i}}{\chi_i!} e^{-(\mu S_i + B_i)} \tag{5.1}$$

In the above expression S_i and B_i denote the expected signal and total background contribution in the i^{th} bin of the SR. The variable μ is the signal strength modifier and it is the parameter of interest. It quantifies the likelihood that a given signal is compatible with the observation. It is equal to 1 when the expected signal events and the number of signal events from the maximization of the likelihood are equal, or 0 for the null hypothesis. Assuming that the bins of the SR are statistically independent, the likelihood function can be constructed as the product of Poisson distributions for each bin.

$$\mathcal{L}(\chi|\mu) = \prod_{i=1}^{n\ bins} f_i(\chi_i|\mu) = \prod_{i=1}^{n\ bins} \frac{(\mu S_i + B_i)^{\chi_i}}{\chi_i!} e^{-(\mu S_i + B_i)} \tag{5.2}$$

In this way the information of all the SR bins is combined.

As discussed in Sec. 4.8 the predicted signal and background (S_i and B_i) are subject to systematic uncertainties which must be included in the likelihood in the form of nuisance parameter θ_{ij}, where i stands for the SR bin and j the nuisance uncertainty and takes values from 1 up to the maximum number of systematic uncertainties in every bin, m. The expected signal and background should now be written as a function of the nuisance parameters as $S_i(\theta)$ and $B_i(\theta)$. Additionally, these nuisance parameters are random variable themselves and therefore they are distributed according to a pdf $f(\theta_{ij}^0|\theta_{ij})$, where θ_i^0 is the default value of the nuisance parameter. The uncertainties that are applied as multiplicative factors such as the normalization uncertainties, take the form of log-normal distributions with width equal to the size of the uncertainty. The shape uncertainties, such as the trigger scale factor uncertainty are treated as Gaussian distributed. Finally the profile likelihood can be written as

$$\mathcal{L}(data|\mu,\theta) = \prod_{i=1}^{n\ bins} f_i(\chi_i|\mu) = \prod_{i=1}^{n\ bins} \frac{(\mu S_i + B_i)^{\chi_i}}{\chi_i!} e^{-(\mu S_i + B_i)} \times \prod_{i=1}^{n\ bins} \prod_{j=1}^{m} f(\theta_{ij}^0|\theta_{ij}) \tag{5.3}$$

The hypothesis testing relies on finding the parameters μ and θ the maximize $\mathcal{L}$ for the null and the signal+background hypothesis separately. The hypothesis that results in higher $\mathcal{L}$ value is better suited to explain the observed data.

5.1.2 Parameter estimation

The parameter values that maximise $\mathcal{L}$ can be found by solving

$$\frac{\partial \mathcal{L}}{\partial \lambda_i} = 0 \tag{5.4}$$

where λ_i denotes a free parameter (μ or θ). The values that maximize the likelihood are called maximum likelihood (ML) estimators and are denoted as $\hat{\mu}$ and $\hat{\theta}$. The estimators are random variables and they are distributed according to a pdf around their true value. The expectation value $\bar{\hat{\lambda}}_i$ and its variance $\hat{\sigma}^2$ can be computed once the functional form of $\hat{\lambda}_i$ is derived.

The default value of the nuisance parameter is estimated from physics principles and are referred to as prior or pre-fit values (θ^0). The maximum likelihood calculations yield a posterior or post-fit value of the parameter ($\hat{\theta}$) which is constrained by the maximum likelihood procedure if it is not the same as the pre-fit value.

5.1.3 Hypothesis Testing

The next step is to infer the compatibility of the observed data with the two hypothesis H_0 and H_μ. For simplicity H_μ is treated as a single hypothesis for a given parameter μ, however it should be noted that every value of μ corresponds to different H_μ because the nuisance parameters change in the fit when μ changes. A test statistic q_μ is constructed as the ratio of the 2 profile likelihoods [1] and will be used to conclude which of the two hypothesis is more likely to happen.

$$q_\mu = -2ln(\frac{\mathcal{L}(data|\mu, \hat{\hat{\theta_\mu}})}{\mathcal{L}(data|\hat{\mu}, \hat{\theta})}) \tag{5.5}$$

In order to compute q_μ, $\mathcal{L}$ must be maximised twice. The maximization of the numerator will yield the ML estimator $\hat{\hat{\theta}}_\mu$ for a given μ and data. The maximization of the denominator will yield $\hat{\theta}$ and $\hat{\mu}$ that correspond to the global maximum of the likelihood. The range of values for the signal strength is [0, μ], the lower bound is motivated by the assumption that the presence of signal cannot reduce the rate of the background. The upper bound is imposed by hand in order to guarantee that upwards fluctuations of data will not be interpreted as the presence of signal.

The observed value of the test statistic can be calculated for a given signal strength q_μ^{obs} and the nuisances that best describe the experimental data, $\hat{\theta}_0^{obs}$ and

[1]The profile likelihood is defined as $\mathcal{L}(data|\mu\hat{\theta}) = max_\theta \mathcal{L}(data|\mu, \theta)$

$\hat{\theta}_{\mu}^{obs}$ for the two hypotheses, can be found by the maximization of the numerator and the denominator.

The test statistic q_μ is also a random variable and it is distributed according to a pdf. This pdf depends on the hypothesis in question which further depend on the signal strength and can be written as $g = g(q_\mu|H_\mu)$. In order to construct the pdf of the test statistic toy MC pseudo-data are generated assuming a signal with μ in H_μ and $\mu = 0$ in H_0. For the toy simulation production the nuisance parameters $\hat{\theta_{\mu}^{obs}}$ and $\hat{\theta}_{0}^{obs}$ are set to the values obtained by the fit on the observed data. From the pdf distributions the p-values for a given efficiency α can be estimated as:

$$p_\mu = P(q_\mu \geq q_\mu^{obs}|H_\mu) \Rightarrow p_\mu = \int_{q_\mu^{obs}}^{\infty} g(q_\mu|\mu, \hat{\theta_\mu}^{obs})dq_\mu \equiv CL_{S+B} \tag{5.6}$$

$$1 - p_{\mu=0} = P(q_\mu \geq q_\mu^{obs}|H_0) \Rightarrow 1 - p_0 = \int_{q_\mu^{obs}}^{\infty} g(q_\mu|\mu = 0, \hat{\theta_{\mu=0}}^{obs})dq_\mu \equiv CL_B \tag{5.7}$$

The p-values quantify the probability for H_μ to reproduce the data and $1 - p_{\mu=0}$ shows the probability that the observed test statistic is compatible with the background only hypothesis. Then the ratio can be defined as

$$CL_s(\mu) = \frac{p_\mu}{1 - p_{\mu=0}} \tag{5.8}$$

This is the frequentist-like confidence interval [154, 152]. This is the final ingredient needed to calculate the probability of a given signal+background hypothesis with a certain value of signal strength to be compatible with the observed data or not. The $CL_s(\mu)$ is compared conventionally in particle physics, with 0.05. If $CL_s < 0.05$ for $\mu > 0$ the signal model with μ is excluded with 95% CL. The μ_{UL} values for which $CL_\mu = 0.05$ are called the upper limit and they are usually propagated to the cross section as $\sigma_{UL} = \mu_{UL}\sigma_{H_\mu}$. This is what will be presented in the color map of the limit plots in the next section.

In SUSY searches as the one of this thesis, the masses of the SUSY particles are free parameters. Therefore the upper limit is evaluated for every mass point. An asymptotic approximation is used to reduce the computational effort. This relies on a simplified test-statistic which is built from μ, $\hat{\mu}$ and σ [2] as $q_\mu = \frac{(\mu-\hat{\mu})^2}{\sigma^2}$.

For the estimation of the background-only expectation upper limit, the data fit is performed conditionally for a $\mu = 0$ and the nuisance parameters are set to the values obtained in the fit. With this at hand a set of pseudo-data that follow Poisson distribution is produced. For each one of them the CL_s and μ_{UL} are calculated.

[2] The standard deviation of $\hat{\mu}$

5.2 Results

The custom HiggsCombine software tool is employed for the maximum likelihood fit and the limit estimation. The package provides various statistical techniques available in RooFit/RooStat [182]. It takes as input the datacards form, for every SR and CR of the analysis. The datacards are simple text files, containing the signal and background yields, the systematic uncertainties and the names of the templates with the distribution of the variables that will be used for the fitting.

The results are extracted from the simultaneous binned maximum likelihood fit of the expected signal and background prediction to the Run 2 dataset, from all the CR and SR of the search. The normalization of the prompt background processes in the CR is left to float freely in the fit. The distributions that enter the maximum likelihood fit is the $M(\ell\ell)$ for the 2ℓ-ewk SR, DY and $\mathrm{t\bar{t}}$ CR, the $M(\ell\ell)^{min}_{SFOS}$for the 3ℓ-ewk SR and WZ-enriched region and the leading lepton p_{T} for the 2ℓ-stop SR. In the prompt background CR, 2ℓ-ewk SR and 3ℓ-ewk SR the distributions are comprised by 4 bins in the low-MET bins and 5 bins in the rest of the MET bins. In the 2ℓ-stop SR the leading lepton p_{T} distribution is comprised by 6 bins. The SR and CR bins used for the ML fit are presented in Tab. 5.1. The non-prompt background is estimated with the DD method in the low MET bins of the 2ℓ SR and with the semi-DD method in the medium, high and ultra high MET bins of the 2ℓ SR and all the MET bins of the 3ℓ SR. The systematic are included in the fit as nuisance parameters.

Regions	Medium/High/Ultra MET	Low MET
2ℓ-stop SR × 4 MET bins	6 × leading lepton p_{T}	6 × leading lepton p_{T}
2ℓ-ewk SR × 4 MET bins	$5 \times M(\ell\ell)$	$4 \times M(\ell\ell)$
3ℓ-ewk SR × 2 MET bins	$5 \times M(\ell\ell)^{min}_{SFOS}$	$4 \times M(\ell\ell)^{min}_{SFOS}$
CR SS × 1 MET bin	$5 \times M(\ell\ell)$	–
CR DY × 2 MET bins	$5 \times M(\ell\ell)$	$4 \times M(\ell\ell)$
CR $\mathrm{t\bar{t}}$ × 2 MET bins	$5 \times M(\ell\ell)$	$4 \times M(\ell\ell)$
WZ enriched region × 2 MET bins	$5 \times M(\ell\ell)^{min}_{SFOS}$	$4 \times M(\ell\ell)^{min}_{SFOS}$

Table 5.1: Summary of SR and the CR bins included in the ML fit.

The post-fit SR plots are presented in Figs. 5.1-5.3. The SR yields from the various backgrounds and the data are presented in Tables 5.2–5.2. The yields correspond to the post-fit results (background-only), extracted from the maximum likelihood fit to the data. The uncertainties include both statistical and systematic components. No significant deviation from the SM prediction is observed in the data.

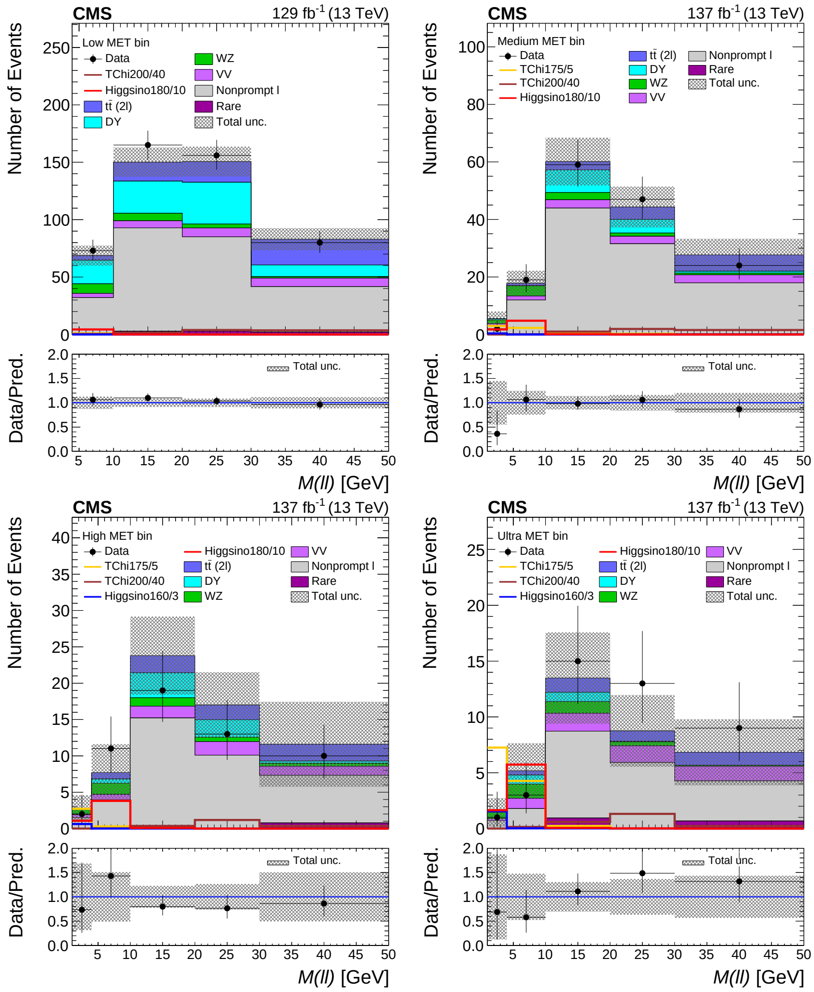

Figure 5.1: The $M(\ell\ell)$ distribution in the 2ℓ-ewk SR. The top row shows the low (left) and medium (right) MET bins and the bottom row shows the high (left) and ultra high (right) MET bin. Both the statistical and systematic components are incorporated in the depicted uncertainty bands. The signal distributions that are overlaid on the plot represent the TCHIWZ and simplified Higgsino models for the positive $M(\ell\ell)$ reweighting.

5.2.1 Nuisance parameters

In order to better understand the effect of the ML fit on the nuisance parameters, the plots in Figure 5.4 report the pulls (middle column) and impacts (rightmost column) of the nuisance parameters in the background-only fit.

The pull is the ratio $(\hat{\theta} - \theta_0)/\Delta\theta$ where $\hat{\theta}$ is the post-fit value of the nuisance parameter, θ_0 is the pre-fit value of the nuisance parameter and $\Delta\theta$ is the pre-fit uncertainty on the nuisance parameter. The asymmetric error bars show the

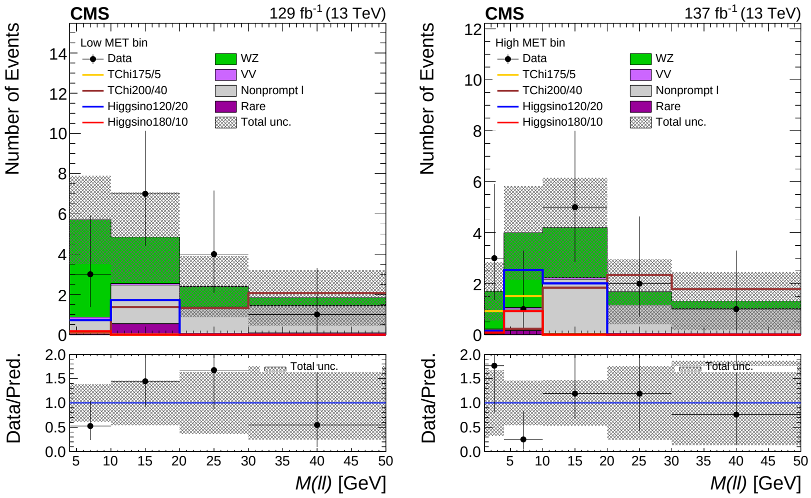

Figure 5.2: The minimum $M(\ell\ell)$ distribution of a same flavor and opposite sign lepton pair in the 3ℓ-ewk SR. The left plot shows the low and the right plot the high MET bin. Both the statistical and systematic components are incorporated in the depicted uncertainty bands. The signal distributions that are overlaid on the plot represent the TChiWZ and simplified Higgsino models for the positive $M(\ell\ell)$ reweighting.

p_T^{miss} [GeV]	$M(\ell\ell)$ [GeV]	$t\bar{t}$	DY	VV	WZ	Rare	Non-prompt	Total bkg	Data
125–200	4–10	4.0 ± 2.0	20.6 ± 5.2	3.7 ± 2.4	8.3 ± 2.6	$0.28^{+0.72}_{-0.27}$	31.9 ± 5.6	68.7 ± 8.7	73
	10–20	16.5 ± 4.2	28.0 ± 6.2	6.2 ± 3.2	6.5 ± 2.3	2.8 ± 2.1	90.1 ± 9.3	151 ± 13	165
	20–30	18.0 ± 4.4	36.3 ± 7.1	7.8 ± 3.5	3.5 ± 1.7	2.9 ± 2.1	82.1 ± 8.9	151 ± 13	156
	30–50	22.4 ± 4.9	10.2 ± 3.7	7.4 ± 3.5	1.3 ± 1.0	2.1 ± 1.8	39.6 ± 6.2	82.9 ± 9.6	80
200–240	1–4	$0.11^{+0.33}_{-0.10}$	$0.37^{+0.72}_{-0.36}$	$0.7^{+1.1}_{-0.7}$	1.3 ± 1.0	$0.04^{+0.23}_{-0.03}$	3.0 ± 2.0	5.5 ± 2.5	2
	4–10	$0.75^{+0.90}_{-0.74}$	$0.15^{+0.50}_{-0.14}$	$1.4^{+1.5}_{-1.4}$	3.5 ± 1.7	$0.14^{+0.39}_{-0.13}$	11.9 ± 3.6	17.8 ± 4.4	19
	10–20	2.9 ± 1.7	7.9 ± 3.4	2.9 ± 2.2	2.5 ± 1.4	1.2 ± 1.2	42.8 ± 6.8	60.1 ± 8.3	59
	20–30	4.3 ± 2.1	4.7 ± 2.6	2.6 ± 2.0	1.1 ± 1.0	$0.27^{+0.54}_{-0.26}$	31.3 ± 5.8	44.3 ± 7.1	47
	30–50	5.7 ± 2.4	$0.6^{+1.0}_{-0.6}$	2.8 ± 2.1	$0.63^{+0.70}_{-0.62}$	$0.35^{+0.65}_{-0.34}$	17.6 ± 4.4	27.7 ± 5.6	24
240–290	1–4	< 0.02	< 0.1	$0.43^{+0.88}_{-0.42}$	0.8 ± 0.8	< 0.07	1.5 ± 1.3	2.7 ± 1.9	2
	4–10	$0.9^{+1.2}_{-0.9}$	$0.57^{+0.97}_{-0.56}$	$0.8^{+1.1}_{-0.8}$	1.5 ± 1.1	$0.3^{+2.6}_{-0.3}$	3.7 ± 2.0	7.7 ± 3.9	11
	10–20	2.4 ± 1.6	3.4 ± 2.3	1.6 ± 1.6	1.2 ± 0.9	$0.3^{+1.3}_{-0.3}$	14.9 ± 4.0	23.8 ± 5.4	19
	20–30	2.0 ± 1.5	2.4 ± 1.9	1.9 ± 1.7	$0.61^{+0.67}_{-0.60}$	$0.03^{+0.45}_{-0.02}$	10.1 ± 3.3	17.0 ± 4.5	13
	30–50	2.3 ± 1.7	$0.32^{+0.73}_{-0.31}$	$1.2^{+1.4}_{-1.1}$	$0.40^{+0.53}_{-0.39}$	$0.8^{+4.6}_{-0.7}$	6.6 ± 2.7	11.6 ± 5.8	10
> 290	1–4	< 0.02	< 0.1	$0.18^{+0.65}_{-0.17}$	$0.57^{+0.65}_{-0.56}$	< 0.01	$0.70^{+0.88}_{-0.69}$	1.5 ± 1.3	1
	4–10	$0.38^{+0.64}_{-0.37}$	$0.8^{+1.1}_{-0.8}$	$0.9^{+1.2}_{-0.9}$	1.3 ± 1.0	$0.12^{+0.44}_{-0.11}$	1.7 ± 1.3	5.2 ± 2.5	3
	10–20	1.3 ± 1.2	$0.8^{+1.2}_{-0.8}$	1.6 ± 1.6	1.05 ± 0.89	$0.9^{+1.4}_{-0.9}$	7.8 ± 2.9	13.5 ± 4.1	15
	20–30	$0.9^{+1.0}_{-0.9}$	$0.06^{+0.28}_{-0.05}$	$1.5^{+1.6}_{-1.5}$	$0.3^{+0.50}_{-0.34}$	< 0.09	5.9 ± 2.5	8.8 ± 3.2	13
	30–50	1.2 ± 1.1	< 0.1	$1.3^{+1.5}_{-1.3}$	$0.09^{+0.24}_{-0.08}$	$0.7^{+1.2}_{-0.7}$	3.6 ± 2.0	6.8 ± 3.0	9

Table 5.2: Observed and predicted yields extracted from the maximum likelihood fit, in the 2ℓ-ewk SR. The uncertainties include both the statistical and systematic components.

post-fit uncertainty divided by the pre-fit uncertainty. Therefore, parameters with error bars lower than ± 1 are constrained by the fit.

The impact shows how much would the signal strength, denoted as $\hat{r}$, changes if the nuisance parameter is varied by $\pm 1\sigma$ θ. The plot shows the shift $\Delta\hat{r}$ induced

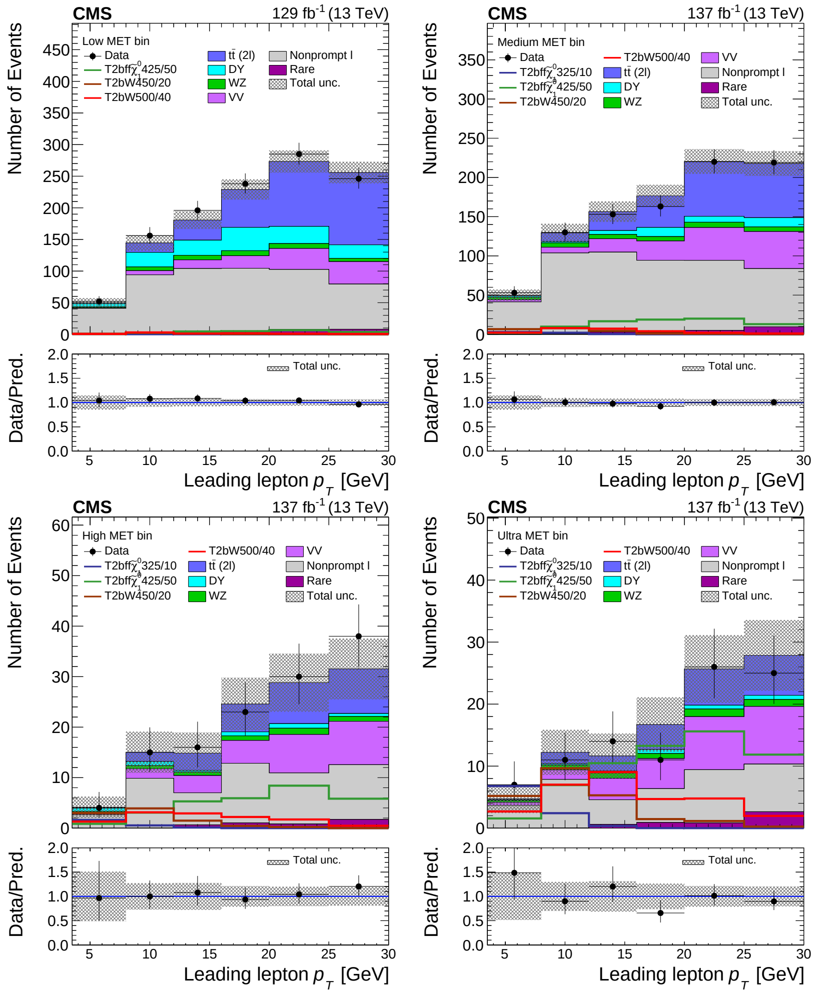

Figure 5.3: The leading lepton p_{T} distribution in the 2ℓ-stop SR. The top row shows the low (left) and medium (right) MET bins and the bottom row shows the high (left) and ultra high (right) MET bin. Both the statistical and systematic components are incorporated in the depicted uncertainty bands. The signal distributions that are overlaid on the plot represent the T2Bff and T2BW simplified.

when the nuisance parameter is fixed to its post-fit $\pm 1\sigma$ values and the fit is redone with N-1 parameters. This is a measurement of the correlation between the nuisance parameters and the signal strength and determines which nuisances have the largest effect on its uncertainty. The direction of the shaded uncertainty bands indicate correlation/anticorrelation of the nuisance with the signal strength.

Nuisance parameters that include the lnU suffix in their naming, presented in light gray, represent the normalisation of the prompt SM background processes which are modeled as log-uniform uncertainties and are left to float in the fit.

p_T^{miss} [GeV]	$M(\ell\ell)_{SFOS}^{min}$[GeV]	VV	WZ	Rare	Non-prompt	Total bkg	Data
125–200	4–10	$0.18^{+0.54}_{-0.17}$	4.8 ± 1.9	$0.08^{+0.38}_{-0.07}$	$0.61^{+0.83}_{-0.60}$	5.7 ± 2.2	3
	10–20	$0.08^{+0.35}_{-0.07}$	2.3 ± 1.3	$0.5^{+1.0}_{-0.5}$	1.9 ± 1.4	4.9 ± 2.2	7
	20–30	$0.03^{+0.23}_{-0.02}$	1.0 ± 1.0	$0.07^{+0.35}_{-0.06}$	1.3 ± 1.2	2.4 ± 1.5	4
	30–50	$0.01^{+0.13}_{-0.01}$	$0.39^{+0.55}_{-0.38}$	$0.08^{+0.37}_{-0.07}$	1.4 ± 1.2	1.8 ± 1.4	1
> 200	1–4	$0.01^{+0.18}_{-0.01}$	1.5 ± 1.0	$0.03^{+0.20}_{-0.02}$	$0.18^{+0.44}_{-0.17}$	1.7 ± 1.2	3
	4–10	$0.05^{+0.34}_{-0.04}$	2.9 ± 1.4	$0.16^{+0.47}_{-0.15}$	$0.85^{+0.99}_{-0.84}$	4.0 ± 1.8	1
	10–20	$0.06^{+0.32}_{-0.05}$	2.0 ± 1.2	$0.05^{+0.26}_{-0.04}$	2.1 ± 1.5	4.2 ± 2.0	5
	20–30	< 0.002	$0.52^{+0.60}_{-0.51}$	$0.06^{+0.29}_{-0.05}$	1.1 ± 1.1	1.7 ± 1.3	2
	30–50	< 0.002	$0.31^{+0.46}_{-0.30}$	$0.03^{+0.23}_{-0.02}$	1.0 ± 1.0	1.3 ± 1.1	1

Table 5.3: Observed and predicted yields extracted from the maximum likelihood fit, in the 3ℓ-ewk SR. Uncertainties include both the statistical and systematic components.

p_T^{miss} [GeV]	$M(\ell\ell)_{SFOS}^{min}$[GeV]	VV	WZ	Rare	Non-prompt	Total bkg	Data
125–200	4–10	$0.13^{+0.47}_{-0.12}$	2.6 ± 1.4	$0.31^{+0.67}_{-0.30}$	$0.49^{+0.70}_{-0.48}$	3.5 ± 1.8	4
	10–20	$0.14^{+0.47}_{-0.13}$	4.3 ± 1.8	$0.47^{+0.83}_{-0.46}$	1.2 ± 1.1	6.1 ± 2.3	11
	20–30	$0.17^{+0.51}_{-0.16}$	5.0 ± 2.0	$0.50^{+0.85}_{-0.49}$	2.1 ± 1.5	7.8 ± 2.6	9
> 200	1–4	$0.16^{+0.56}_{-0.15}$	$0.11^{+0.29}_{-0.10}$	$0.06^{+0.33}_{-0.05}$	$0.44^{+0.66}_{-0.43}$	$0.78^{+0.97}_{-0.77}$	0
	4–10	$0.22^{+0.60}_{-0.21}$	2.6 ± 1.4	$0.10^{+0.38}_{-0.09}$	$0.24^{+0.59}_{-0.23}$	3.1 ± 1.6	3
	10–20	$0.7^{+1.1}_{-0.7}$	10.6 ± 2.8	$0.9^{+1.1}_{-0.9}$	1.9 ± 1.4	14.0 ± 3.4	19
	20–30	$0.7^{+1.0}_{-0.7}$	15.2 ± 3.3	$1.2^{+1.3}_{-1.2}$	4.0 ± 2.0	21.0 ± 4.2	23

Table 5.4: Observed and predicted yields extracted from the maximum likelihood fit, in the WZ-enriched region. Uncertainties include both the statistical and systematic components.

p_T^{miss} [GeV]	$p_T(\ell_1)$ [GeV]	$t\bar{t}$	DY	VV	WZ	Rare	Non-prompt	Total bkg	Data
125–200	3.5–8	1.2 ± 1.2	5.2 ± 3.1	$1.0^{+1.2}_{-1.0}$	1.4 ± 1.1	$0.06^{+0.27}_{-0.05}$	41.0 ± 6.3	49.9 ± 7.2	52
	8–12	15.0 ± 4.0	22.9 ± 5.9	6.6 ± 3.1	6.0 ± 2.1	$0.96^{+0.99}_{-0.95}$	93.1 ± 9.4	144 ± 12	156
	12–16	31.8 ± 5.9	24.0 ± 6.1	13.7 ± 4.5	7.2 ± 2.4	2.8 ± 1.7	101.3 ± 9.9	180 ± 14	196
	16–20	59.9 ± 8.0	36.9 ± 7.5	19.8 ± 5.5	7.9 ± 2.5	4.2 ± 2.1	100.2 ± 9.8	229 ± 16	238
	20–25	103 ± 11	27.2 ± 6.5	33.2 ± 7.1	7.7 ± 2.5	7.5 ± 2.8	95.0 ± 9.5	273 ± 18	285
	25–30	114 ± 11	21.4 ± 5.7	35.5 ± 7.3	5.1 ± 2.0	8.0 ± 2.8	71.5 ± 8.3	256 ± 17	246
200–290	3.5–8	1.1 ± 1.0	1.7 ± 1.5	2.8 ± 2.1	2.9 ± 1.4	$0.04^{+0.20}_{-0.03}$	41.3 ± 6.6	49.9 ± 7.3	53
	8–12	11.0 ± 3.3	1.6 ± 1.5	7.3 ± 3.3	5.6 ± 2.0	$0.43^{+0.65}_{-0.42}$	103 ± 10	129 ± 12	130
	12–16	24.1 ± 4.9	5.0 ± 2.6	17.1 ± 5.0	5.5 ± 2.0	2.9 ± 1.7	102 ± 10	156 ± 13	153
	16–20	40.3 ± 6.3	11.7 ± 4.2	24.7 ± 6.1	5.6 ± 2.0	2.4 ± 1.6	92.0 ± 9.8	177 ± 14	163
	20–25	69.9 ± 8.3	7.6 ± 3.4	41.9 ± 7.9	6.7 ± 2.2	5.0 ± 2.2	89.3 ± 9.7	220 ± 16	220
	25–30	69.0 ± 8.3	11.8 ± 4.1	47.3 ± 8.4	5.9 ± 2.0	9.6 ± 3.1	74.2 ± 8.9	218 ± 16	219
290–340	3.5–8	$0.15^{+0.35}_{-0.14}$	$0.67^{+0.90}_{-0.66}$	$0.34^{+0.72}_{-0.33}$	$0.29^{+0.44}_{-0.28}$	< 0.05	2.7 ± 1.7	4.1 ± 2.1	4
	8–12	1.9 ± 1.4	$0.8^{+1.1}_{-0.8}$	1.9 ± 1.7	$0.64^{+0.67}_{-0.63}$	$0.01^{+0.11}_{-0.01}$	9.9 ± 3.2	15.0 ± 4.1	15
	12–16	3.4 ± 1.8	$0.33^{+0.61}_{-0.32}$	3.4 ± 2.3	0.69 ± 0.69	$0.64^{+0.85}_{-0.63}$	6.4 ± 2.6	14.8 ± 4.1	16
	16–20	5.5 ± 2.3	$0.8^{+1.1}_{-0.8}$	4.5 ± 2.6	0.91 ± 0.80	1.0 ± 1.0	11.8 ± 3.5	24.6 ± 5.2	23
	20–25	8.1 ± 2.8	$0.9^{+1.2}_{-0.9}$	7.6 ± 3.4	1.24 ± 0.93	$0.82^{+0.89}_{-0.81}$	10.1 ± 3.2	28.8 ± 5.8	30
	25–30	8.8 ± 2.9	$0.58^{+0.97}_{-0.57}$	8.6 ± 3.6	0.96 ± 0.81	1.7 ± 1.3	10.8 ± 3.4	31.5 ± 6.0	38
> 340	3.5–8	$0.12^{+0.37}_{-0.11}$	$0.14^{+0.51}_{-0.13}$	$0.48^{+0.86}_{-0.47}$	$0.29^{+0.46}_{-0.28}$	< 0.03	3.7 ± 2.0	4.7 ± 2.3	7
	8–12	1.8 ± 1.3	$0.22^{+0.59}_{-0.21}$	1.5 ± 1.5	0.78 ± 0.75	$0.02^{+0.12}_{-0.01}$	7.8 ± 2.9	12.2 ± 3.6	11
	12–16	2.4 ± 1.5	$0.31^{+0.63}_{-0.30}$	3.5 ± 2.3	0.87 ± 0.78	$0.60^{+0.79}_{-0.59}$	4.0 ± 2.0	11.6 ± 3.6	14
	16–20	4.0 ± 2.0	$0.64^{+0.89}_{-0.63}$	4.9 ± 2.7	0.80 ± 0.75	$0.90^{+0.93}_{-0.89}$	5.5 ± 2.5	16.7 ± 4.4	11
	20–25	5.8 ± 2.3	$0.62^{+0.95}_{-0.61}$	8.6 ± 3.6	1.22 ± 0.93	$0.84^{+0.92}_{-0.83}$	8.6 ± 3.0	25.7 ± 5.5	26
	25–30	6.5 ± 2.5	$0.7^{+1.0}_{-0.7}$	9.3 ± 3.7	1.12 ± 0.88	2.6 ± 1.6	7.7 ± 2.9	27.9 ± 5.7	25

Table 5.5: Observed and predicted yields extracted from the maximum likelihood fit, in the 2ℓ-stop SR. Uncertainties include both the statistical and systematic components.

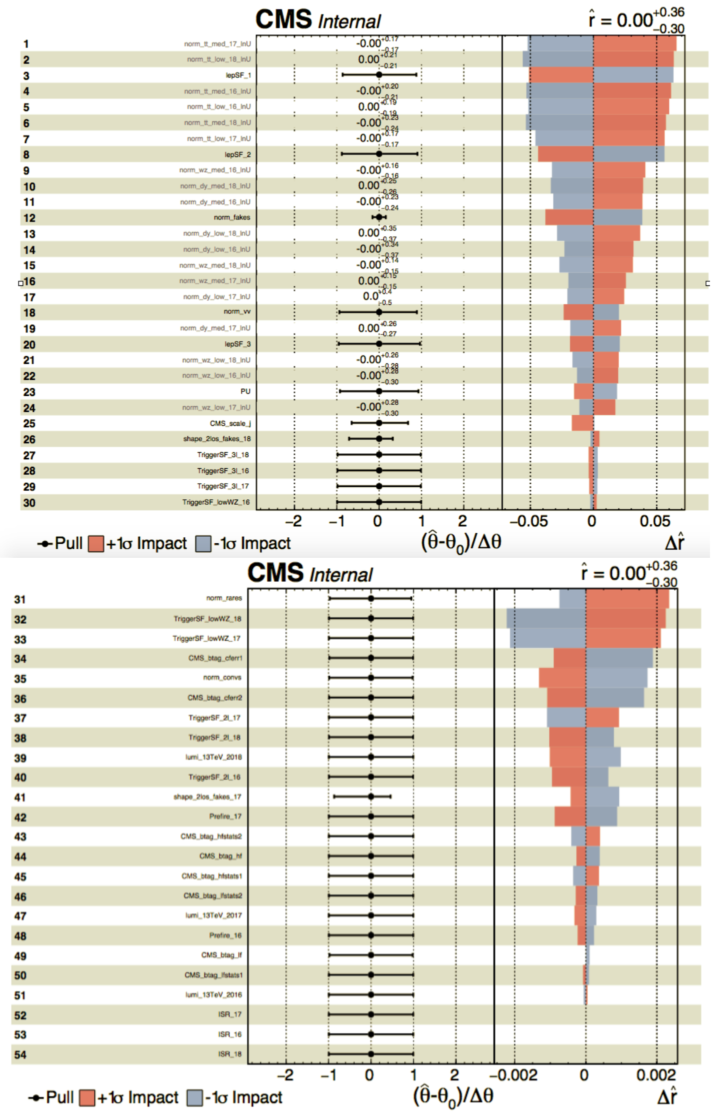

Figure 5.4: The pulls and the impact of the nuisance parameters on the signal strength. The first column presents the name of the nuisance parameters. The ones presented in light grey with the lnU suffix correspond to the prompt background normalization and are left to float in the fit. The middle column shows the pulls and the leftmost column shows the impact of the nuisance parameters on the signal strength.

From the plots one can conclude that the non-prompt background normalization gets largely constrained, from 40% pre-fit to 6% post-fit, due to the SS CR which is included in the fit. Additionally, all other prompt backgrounds get constrained from their dedicated CR. The normalization of the prompt $t\bar{t}$ background and the lepton SF have the largest impact on the signal strength in the range of 5-6%.

5.2.2 Interpretations

No significant deviation from the SM expectation is observed. Therefore, the results of the analysis are interpreted as constraints on the simplified Wino/Bino, Higgsino, T2Bff and T2BW models and the pMSSM Higgsino model. Table 5.1 presents a summary of the number of SR and CR that are included in the fit in each interpretation. Hypothesis test is performed using the frequentist approach and the CLs prescription is implemented. Limits at 95% CL are set on the production cross section for sparticle pairs as a function of the their masses.

The DY, $t\bar{t}$ and SS CR and the WZ enriched region are added in the ML fit to constrain the normalization of the respective SM processes as described in the previous section. Table 5.6 presents the SR and CR included in the ML fit in the various interpretations.

Regions	TChiWZ	Higgsino	T2Bff/T2BW
2ℓ-stop SR	–	–	✓
2ℓ-ewk SR	✓	✓	–
3ℓ-ewk SR	✓	✓	–
CR SS	✓	✓	✓
CR DY	✓	✓	✓
CR $t\bar{t}$	✓	✓	✓
WZ enriched Region	✓	✓	✓

Table 5.6: Summary of SR and the CR included in the ML fit for every interpretation.

Simplified TChiWZ model

The exclusion limit on the production cross section as a function of the $m_{\widetilde{\chi}_2^0}$ and $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0}$ is shown in Figure 5.5 for the TChiWZ model. In this model the $\widetilde{\chi}_1^{\pm}$ and $\widetilde{\chi}_2^0$ are pure Winos and the $\widetilde{\chi}_1^0$ LSP is Bino. Considering the full matrix element of the $\widetilde{\chi}_2^0$ decay, the relative sign of the neutralinos mass eigenvalues leads to $M(\ell\ell)$ distributions variations as described in Sec. 4.2. Therefore, two limit plots for the two signal $M(\ell\ell)$ reweighting scenarios are presented.

The exclusion line connects the signal points for which the signal strength, defined as the ratio of number of post-fit signal events over the number of pre-fit signal events, is equal to 1 and thus the cross section of the signal point predicted by the fit is equal to the cross section of the model. Points on the 2D mass plane are excluded when the fitted cross section is lower than the model cross section.

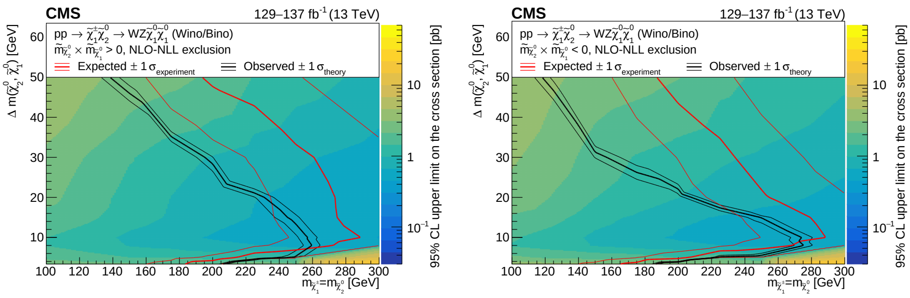

Figure 5.5: The exclusion limit on the production cross section in the TChiWZ model. The black line shows the 95% CL observed exclusion limit and the variation bands (thin black lines) correspond to the cross section uncertainty of the simplified TChiWZ model. The red line presents the 95% CL expected exclusion limit on the production cross section with the band (thin lines) covering 68% of the limits in the absence of signal. The $m_{\widetilde{\chi}_2^0} \times m_{\widetilde{\chi}_1^0} > 0$ $M(\ell\ell)$ spectrum reweighting scenario is shown in the left plot and the $m_{\widetilde{\chi}_2^0} \times m_{\widetilde{\chi}_1^0} < 0$ in the right plot.

Most of the exclusion power of the search in the low $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0}$ region arises from the medium/high/ultra MET bins of the 2ℓ-ewk SR enhanced by the extension to lower $M(\ell\ell)$ and lepton p_{T}. The exclusion power at higher $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0}$ originates from the low MET SR and the 3ℓ-ewk SR.

The exclusion limit reaches $m_{\widetilde{\chi}_2^0} = 200$ GeV for $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0} = 3$ GeV. At lower $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0}$ values the sensitivity of the search is reduced due to low acceptance and reconstruction efficiency of the very soft leptons. Figure 5.6 presents the 2016 result. Even though there cannot be an even comparison to the 2016 result due to the major differences in analysis method, it can be noticed that the sensitivity of the search is now extended by 2 GeV downwards in $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0}$.

The lower $M(\ell\ell)$ and lepton p_{T} selection together with the MET binning, increase the analysis acceptance in the $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0} \sim 10$ GeV. The TChiWZ signal points with $m_{\widetilde{\chi}_2^0} < 280$ GeV and $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0}$ are excluded and the sensitivity is improved by almost 120 GeV with respect to 2016 analysis.

The 3ℓ-ewk SR enhances the sensitivity at higher $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0}$ and extends the exclusion limit to $m_{\widetilde{\chi}_2^0} = 190$ GeV for $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0} = 40$ GeV. It is worth mentioning that the 2016 analysis, did not include the 3ℓ-ewk SR.

From the exclusion limit plots in Fig.5.5, the observed limit is weaker than the expected in intermediate and high $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0}$ ranges. The discrepancy between the two is almost 2σ and the reason behind this can be factorised into two effects:

- Firstly, an underprediction of the SM background observed in intermediate and high $M(\ell\ell)$ values in the most sensitive MET bins of the 2ℓ-ewk SR.

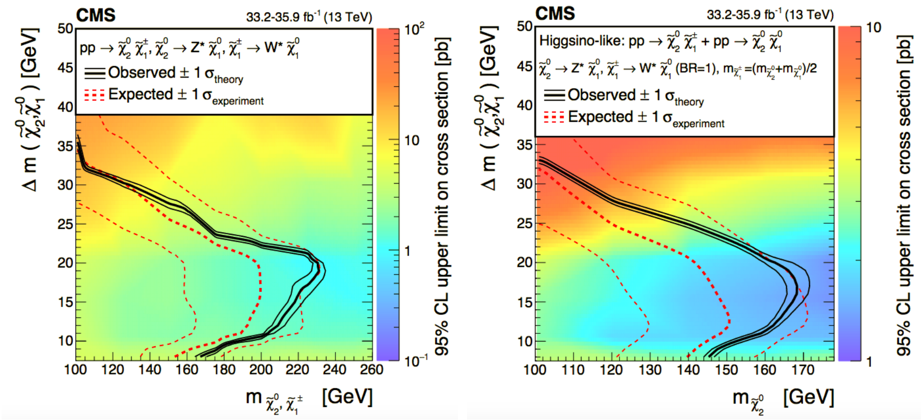

Figure 5.6: The 2016 exclusion limit on the production cross section in the TChiWZ (left)and simplified Higgsino (right) model. The black line shows the 95% CL observed exclusion limit and the variation bands (thin black lines) correspond to the cross section uncertainty of the simplified TChiWZ model. The red line presents the 95% CL expected exclusion limit on the production cross section with the band (thin lines) covering 68% of the limits in the absence of signal [76].

This can be seen in Tab. 5.2, ultra MET bin (p_T^{miss} >290) in the $M(\ell\ell)$ bin of [20,30] GeV and in Tb.5.2 low MET bin $M(\ell\ell)$ bin of [10,20] GeV and [20,30] GeV. This effect accounts for the 1σ discrepancy

- The additional 1σ discrepancy between the observed and expected exclusion limit originates from the WZ enriched region, which is included in the fit and has a small signal contribution in the low $M(\ell\ell)$ bins as described in 4.7.2. The plots in Fig. 5.7 present the $M(\ell\ell)^{min}_{SFOS}$ distribution of a same flavor and opposite sign lepton pair in the WZ enriched region, pre-fit (left) and post-fit (right). The pre-fit plot shows only the statistical uncertainty while the post-fit plot illustrates both statistical and systematic uncertainties. The fit fixes the WZ normalization based on the higher $M(\ell\ell)$ bin which is the most statistically enriched. In the WZ enriched region pre-fit $M(\ell\ell)$ distribution, the SM background prediction in the largest $M(\ell\ell)$ bin is higher than the observed data, while in the lower $M(\ell\ell)$ bins some underprediction is observed within the statistical uncertainty. Due to variations in the data/simulation agreement across the $M(\ell\ell)$ bins, the fit driven by the last $M(\ell\ell)$ bin, pulls the normalization of the WZ down. Therefore, the residual data/simulation disagreement in the lower $M(\ell\ell)$ bins is transferred by the fit to signal points with intermediate $\Delta m_{\tilde{\chi}_2^0\text{-}\tilde{\chi}_1^0}$. This results to the lower observed limit in this $\Delta m_{\tilde{\chi}_2^0\text{-}\tilde{\chi}_1^0}$ region. It was checked that by fixing the WZ normalization to its post-fit value and repeating the fit reduces the discrepancy between the expected and observed exclusion limit to 1σ

Additionally, the shift on the $M(\ell\ell)$ distribution due to the negative mass reweighting for the case of $m_{\widetilde{\chi}_2^0} \times m_{\widetilde{\chi}_1^0} < 0$ causes the upper limit of the predicted signal cross section from the fit to be higher, leading to a smaller range of excluded mass parameters in this scenario. The excess has a maximum local significance of 2.4σ for the signal mass point with $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0} = 40\,\text{GeV}$ and $m_{\widetilde{\chi}_2^0} = 140\,\text{GeV}$.

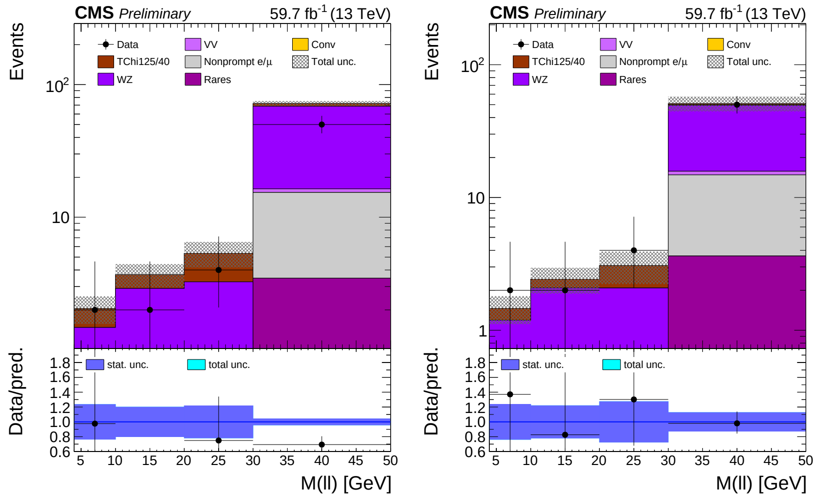

Figure 5.7: The minimum $M(\ell\ell)$ distribution of a same flavor opposite sign lepton pair in the WZ enriched region, low MET bin 2018, pre-fit (left) and post-fit (right). The uncertainties in the pre-fit plot are statistical only, the systematic uncertainties are included in the post-fit plot.

Simplified Higgsino model

The top plot in Fig. 5.8 presents the exclusion limit on the production cross section for the simplified Higgsino model. The chargino and neutralino masses follow the $m_{\widetilde{\chi}_1^\pm} = \frac{1}{2}(m_{\widetilde{\chi}_2^0} - m_{\widetilde{\chi}_1^0})$ assumption, also described in Sec. 4.2. The simplified Higgsino model includes both neutralino pair production and neutralino-chargino production. The $M(\ell\ell)$ distribution is reweighted according to the $m_{\widetilde{\chi}_2^0} \times m_{\widetilde{\chi}_1^0} < 0$ which is the only possible scenario for higgsino.

The analysis excluded $m_{\widetilde{\chi}_2^0} < 150$ GeV at $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0} = 3$ GeV and up to $m_{\widetilde{\chi}_2^0} = 210$ GeV at $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0} = 7$ GeV which is an extension of ~ 65 GeV compared to the 2016 result, presented in Fig. 5.6. The negative $M(\ell\ell)$ reweighting modifies the $M(\ell\ell)$ distribution and it has the tendency to make the higher $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0}$ limit weaker compared to positive or no reweighting. Therefore, comparing the new observed exclusion with that of the 2016 analysis which was obtained without applying any $M(\ell\ell)$ reweighting, one can notice a decrease in the sensitivity at

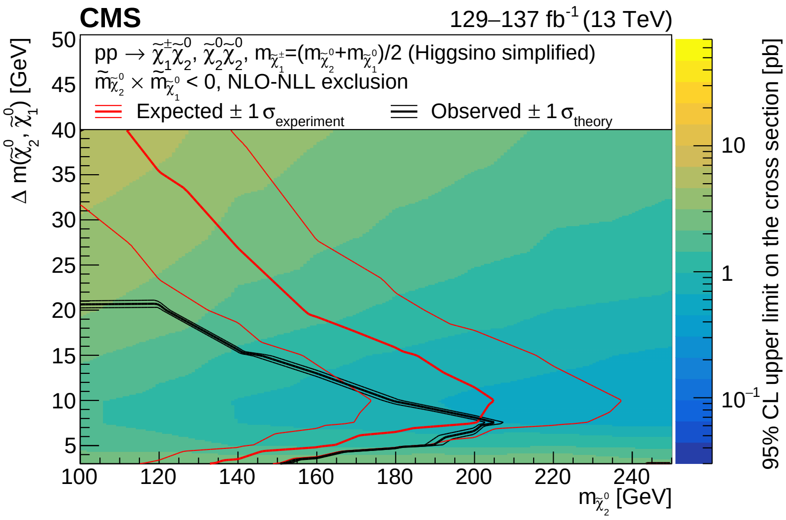

Figure 5.8: The exclusion limit on the production cross section in the simplified Higgsino model. The black line shows the 95% CL observed exclusion limit and the variation bands (thin black lines) correspond to the cross section uncertainty of the Higgsino model. The red line present the 95% CL expected exclusion limit on the production cross section with the band (thin lines) covering 68% of the limits in the absence of signal.

the higher $\Delta m_{\tilde{\chi}_2^0\text{-}\tilde{\chi}_1^0}$ regions. However, comparing the sensitivity of the search at $\Delta m_{\tilde{\chi}_2^0\text{-}\tilde{\chi}_1^0}$=30 GeV, it is noticed that it has been improved by $\sim$25 GeV .

Simplified T2Bff and T2BW model

Additionally, the results of the analysis are interpreted in terms of simplified T2Bff and T2BW models with some changes in the event selection with respect to the one applied for the electroweakino search. For the stop search the m_T mass cut is relaxed and no same-flavor and opposite sign lepton pair requirement is applied because the leptons are expected to be produced from two different bosons. Additionally the MET binning is slightly different. The upper boundaries of the bins are shifted by 50 GeV due to increased statistics. Finally, the leading lepton p_{T} is used as the binning variable. The plots in Fig. 5.9 show the exclusion limit on the production cross section of the T2Bff simplified model on the left and the T2BW simplified model on the right.

There is a major improvement with respect to the 2016 analysis by almost 50 GeV in $m_{\tilde{t}}$ at $\Delta m_{\tilde{\chi}_2^0\text{-}\tilde{\chi}_1^0}$ = 40 GeV.

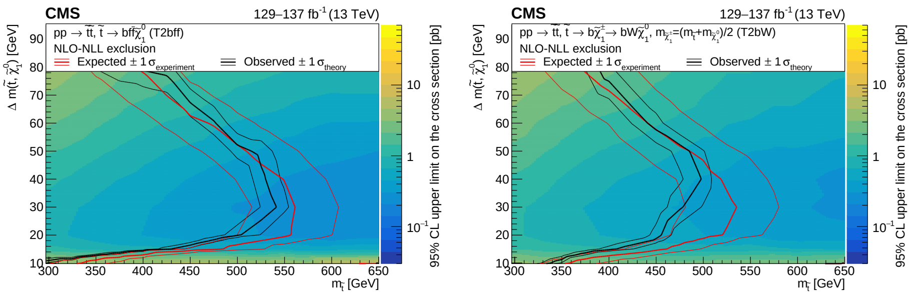

Figure 5.9: The exclusion limit on the production cross section in the simplified T2Bff model (left) and T2BW model (right). The black line shows the 95% CL observed exclusion limit and the variation bands (thin black lines) correspond to the cross section uncertainty of the models. The red line presents the 95% CL expected exclusion limit on the production cross section with the band (thin lines) covering 68% of the limits in the absence of signal.

pMSSM Higgsino

Lastly, the results are interpreted in term of the pMSSM Higgsino model described in Sec. 4.2. The 95% exclusion limit on the production cross section is presented in Fig. 5.10 on the higgsino-bino mass parameter plane (μ–M_1).

In the pMSSM Higgsino model, larger μ values roughly correspond to larger masses for the parent supersymmetric particles. Larger values of the M_1 parameter correspond to smaller values of the mass difference $\Delta m_{\tilde{\chi}^0_2\text{-}\tilde{\chi}^0_1}$. This explains the inverted behaviour of the exclusion limit compared to the one obtained with the simplified model showed in Fig. 5.8.

5.3 Summary and Conclusions

A search for SUSY in events with two or three soft leptons and p_T^{miss} in the compressed mass spectrum, with the full Run 2 dataset is presented. These signatures arise from models of electrowekinos decaying to the lightest neutralinos via leptonic decays. In the R-parity conserving models, considered in this search the lightest neutralino is an attractive DM candidate. Natural SUSY predicts light electorweakinos that can have compressed mass spectrum and lead to signatures with soft leptons.

The search presented in this chapter is an extension of the 2016 dataset analysis, with multiple upgrades in the analysis method. The analysis has been optimised to cover $M(\ell\ell)$ range down to 1 GeV and the signal selection methods have been re-optimised for higher sensitivity in the low $\Delta m_{\tilde{\chi}^0_2\text{-}\tilde{\chi}^0_1}$ regions. The method for the non-prompt background prediction, which is the dominant background in

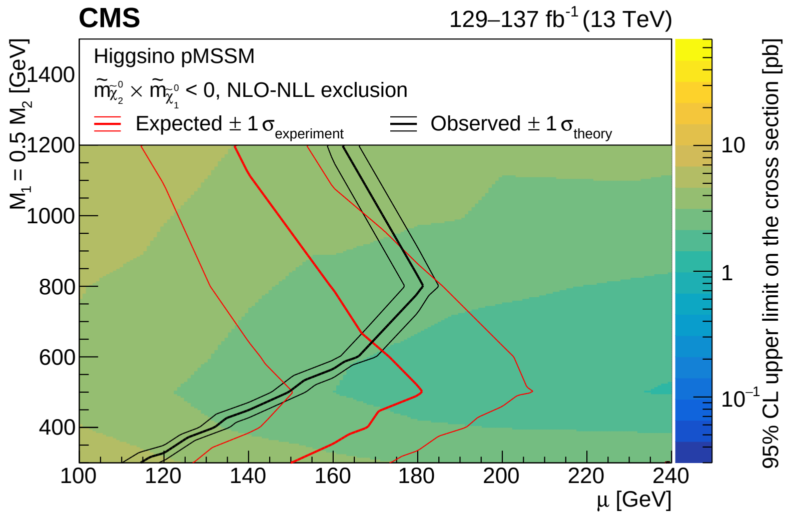

Figure 5.10: The exclusion limit on the production cross section in the pMSSM Higgsino model. The black line shows the 95% CL observed exclusion limit and the variation bands (thin black lines) correspond to the cross section uncertainty of the model. The red line present the 95% CL expected exclusion limit on the production cross section with the band (thin lines) covering 68% of the limits in the absence of signal.

the SR, has been significantly improved and the search for 3ℓ final states is added in the analysis strategy.

The results are interpreted in terms of simplified wino/bino and Higgsino models, stop production models and a pMSSM Higgsino model. A hypothesis test is performed and upper limits at 95% CL on the signal production cross section are set. All the SR and prompt background CR and the WZ enriched region are included in the ML fit. The systematic uncertainties are included as nuisance parameters in the fit.

The simplified wino-bino model in which the NLSP $\widetilde{\chi}_2^0$ and $\widetilde{\chi}_1^{\pm}$ are produced and decay following the $\widetilde{\chi}_2^0\ \widetilde{\chi}_1^{\pm} \to Z^*W^* \to \widetilde{\chi}_1^0\ \widetilde{\chi}_1^0$ is considered. In this model, pure wino production cross section is assumed and mass differences of $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0}$ less than 50 GeV are explored. The $\widetilde{\chi}_2^0\ \widetilde{\chi}_1^{\pm}$ masses are excluded up to 280 GeV for $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0}$ of 10 GeV, at 95 % CL.

In the simplified higgsino model interpretation the $\widetilde{\chi}_2^0$ and $\widetilde{\chi}_1^{\pm}$ masses are assumed to follow $m_{\widetilde{\chi}_1^{\pm}} = \frac{1}{2}(m_{\widetilde{\chi}_2^0} m_{\widetilde{\chi}_1^0})$. The model includes both neutralino pair production and neutralino-chargino production. The excluded masses reach up to 210 GeV for $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0}$ of 7 GeV and 150 GeV for a highly compressed scenario with $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0}$ of 3 GeV.

In the pMSSM higgsino model, the limits are presented in the plane of the higgsino-bino mass parameters μ-M_1 and the higgsino mass parameter μ is excluded up to 170 GeV, for a bino mass parameter M_1 of 600 GeV. As the M_1 increases, the mass splitting $\Delta m_{\widetilde{\chi}^0_2\text{-}\widetilde{\chi}^0_1}$ decreases. For $M_1 = 800 GeV$, the μ is excluded up to 180 GeV.

Finally, the results are interpreted in terms of two mass-degenerate stop benchmark models, T2bff and T2bW. Top squarks with masses below 540 (480) GeV are excluded for the T2bff (T2bW) top squark decay model, with a $\Delta m_{\widetilde{t},\widetilde{\chi}^0_1}$ mass splitting at 30 GeV.

The prospects of the SUSY search in the compressed mass spectra involve efforts in multiple directions. The group has developed a new parametric signal extraction method which uses the $M(\ell\ell)$ as a discriminating variable with optimized binning for every mass-spitting scenario. This new binning method uses the theoretical signal shape which has a kinematic end point at the signal splitting. The optimal binning is chosen such that there are equal signal quantiles and a minimum number of background event i every bin, while the relative background importance of the various processes are considered. The first results of the method show significant improvement on the TChiWZ exclusion limits up to 30%. Additionally, the group has extended the search for long-lived SUSY particles that can evade detection due to the event selection and reconstruction algorithms that target the prompt decays.

CHAPTER 6

Sensitivity study with machine learning methods

Machine Learning (ML) [183] is currently one of the most innovative technologies in data analysis. It allows computational systems to learn features of the data and thus improve their analysis based on patterns. The objective of ML methods is to learn how to make decisions with the minimum human intervention and they are used in multiple industry applications including health care, autonomous cars, finance and more. They are also widely used in science and high energy physics (HEP) is not an exception. These algorithms have been extensively used in HEP for the past two decades and they have brought improvements in data analyses, in event and object reconstruction techniques, in online event selection and in many more HEP fields that will be described in later sections.

The current chapter of the thesis presents the results of a SUSY sensitivity study in the compressed mass spectrum, performed using ML algorithms. The search targets events with 3 soft leptons and p_T^{miss} in the final state. The structure of the chapter is the following: Sec. 6.1 presents some of the most important and basic concepts on ML and an overview of the ML applications in HEP and in SUSY specifically. Sec. 6.2 introduces the signature this study focuses on. The object and event selection are described in Sec. 6.3 and Sec.6.4. Sec. 6.5 and Sec. 6.6 discuss the training features and the ML algorithms that are examined, respectively. The results are presented in Sec. 6.7 and they are compared to results obtained when using the "baseline" approach, where the shape of a kinematic variable is used for the maximum-likelihood fit. Finally, the conclusions and a discussion on the results are presented in Sec. 6.8.

6.1 Basic ML concepts

The two main objectives of modern HEP are precision measurements of SM processes and searches for BSM physics. Both are typically characterised by faint signal and large background in events involving physics objects that are difficult to reconstruct. The ML algorithms are the state-of-the-art tools used for signal-to-background discrimination, event and object reconstruction, online event selection and more.

The most commonly used ML algorithms in HEP have been the boosted decision trees (BDT) and the neural networks (NN). Deep learning has gained a lot of interest among the HEP society in the last decade. Deep learning algorithms are based on multi-layer NN which together with powerful computational resources and tools has brought significant improvements in the training algorithms and their discrimination power. Providing a detailed description of deep learning algorithms is out of the scope of this thesis. This chapter describes data analysis techniques using simple machine learning algorithms and therefore the focus will be on them.

6.1.1 Basic ML concepts

Supervised and unsupervised ML

The ML algorithms can be divided into two broad categories of *supervised* and *unsupervised* learning. The majority of practical ML algorithms, as well as the ML algorithms used in the sensitivity study presented in this chapter, use supervised learning, which is based on function approximation. In supervised learning the input to the algorithm is a set of n input data samples (i.e. data events in the case of HEP) denoted as x_i, $i = 1, ..., n$, each one accompanied by a label y_i. The set of x_i, y_i is the training sample. The algorithm is used to infer a mapping function f from the input to the output:

$$f : x_i \to y_i \tag{6.1}$$

The goal is for the algorithm to approximate f such that given a new input data set, without labels, it can predict y with good accuracy.

It is a common practice to split the available input dataset into the training sample and the testing samples. The training sample is used to define the model that will be able to make predictions based on the inferred function. The model is dynamically corrected during the training, when the prediction it made is wrong, based on the known label of the input. The fraction of the times the function f predicted correctly the label y_i of x_i is called *accuracy* of the model. The training process continues until the model reaches a desired level of accuracy. The testing sample is used to measure the performance of the model on input data that were not used for its training.

In supervised training process the function f is expressed in terms of weights w as f_w. The loss function is constructed for each data sample i as $\mathcal{L}(f_w(x_i, y_i))$ and quantifies the goodness of fit. This is a metric of how well the approximated function describes the truth target. The weights are iteratively adjusted during the training in order to minimize the average loss function over the training set.

Generalization is an important aspect of a good ML predictive model and refers to its ability to generalize from the data it was trained on to any given data. It allows the model to make good predictions on new data that it hasn't been trained on. When a model learns every detail of the training data it looses its ability to make good predictions on new data. This is called *overfitting* and it is problematic because the model picks up features specific to the training dataset like statistical fluctuations, that cannot be extended to other dataset. On the other hands, a model can be *underfitted* when it cannot predict well on its training data neither on new data. Underfitting can be a result of a short training period. Ideally, a good ML model should be selected at the point between overfitting and underfitting.

In the unsupervised models, the input data has no label available. The goal of the algorithm is to recover underlying structure or dependencies in the dataset. A typical use of unsupervised learning is the data clustering in which data with common properties are clustered into groups. An alternative unsupervised learning is the association rule learning processes in which a rule that describes large portion of the data is learned [184]. No unsupervised learning algorithms are used in the study, therefore it is not further discussed in this section.

Classification and Regression

A typical example of supervised learning, very commonly used in HEP, is the data classification. A classification algorithm estimates the mapping function f from the input data x to categorical output labels y. The simplest case is the binary classification in which the input data are categorized into two groups A and B, or in signal and background in the HEP problems. The sensitivity study of this chapter is a binary classification problem. Popular algorithms used for binary classification problems are the BDTs, Naive Bayes classifiers, random forest and NN. All these algorithms will be described in more detail in Sec. 6.6

The regression algorithms are used to predict continuous output variables. An example of regression ML algorithms widely used in HEP are the object tagger that will be discussed in Sec. 6.1.2. Examples of algorithms usually used in regression problems are: linear regression, NN, nearest neighbours, lasso regression, Gaussian processes and others.

6.1.2 ML in HEP

Over the past two decades the HEP society has been using ML algorithms for the collection and the analysis of its large and complex datasets. This subsection discusses a collection of the most important HEP challenges that have been benefited by the use of ML algorithms.

Event triggering

The large luminosity at the LHC and large production cross-section result in classes of particles produced in large abundance. Due to space restrictions not all of the events can be stored for offline analysis. Therefore it is becoming important to perform real time analysis and accurate object reconstruction at the online trigger level. ML algorithms are used to perform reconstruction and analysis on the fly and improve the event selection efficiency already at the L1 trigger level. A BDT was implemented in the the EMTF L1 algorithm that was used for the Run2 data collection [185]. For the upgrade of the Level 1 trigger, described in Sec. 3.1, both the OMTF and the EMTF++ algorithms are designed to use ML for the p_T assignment of the muon candidate. In addition, a dedicated BDT classifier will be used for the identification of the electromagnetic object at the HGCAL. The BDT will reject PU induced background while achieving good signal efficiency [113].

Event Simulation

Simulated events are extensively used in HEP, for new physics searches and discoveries. It becomes more computationally expensive to produce accurate simulations as the complexity of the LHC data increases. The most time and resources consuming step of the simulation is the detector response to the traversing particle, which can take up to several minutes for a single event. Fast simulation can decrease this time but it suffers from low accuracy. Recently, generative adversarial networks which learn to sample from feature distributions by minimizing the distance between the generated and the real distribution, are used to produce simulated events. The first results show high accuracy in simulation and reduction in the CPU time needed for the production [186]. The simulation production with generative adversarial networks presents a potential solution to the large number of simulated events required by the physics program of the experiments for the HL-LHC.

Jet identification

The reconstruction of the hadronic jets is an important ingredient for most of the physics analysis conducted at the LHC experiments. Jets originating from heavy quarks can be distinguished from other jets due to their characteristic secondary vertex, associated with the long lifetime of heavy-flavoured hadrons. The identification of the initial quark of the jet is called jet-tagging. Deep learning algorithms have been used for bottom jet tagging, exploiting the characteristics of the jets.

The most recent deep learning algorithms developed by the CMS collaboration to provide multi-class classification for jet tagging are the DeepCSV [187], the DeepFlavor [188] and the DeepJet [189] taggers. The DeepCSV uses a deep NN and it can classify the jets into 4 categories, namely b-jets, c-jets, light (u, d, s) and jets from gluons. The DeepFlavor uses a larger input sample together with deeper network with convolutional layers. The DeepJet is the updated version of the DeepFlavor and incorporates a deep NN with convolutional layers and Long Short-Term Memory recurrent layer [1] performing additional classification between the light flavor jets. All of them result in high b-tagging efficiency and low mis-tagging rate. Additionally they can be used for generic heavy flavor jet tagging due to their ability for multi-class classification.

Event Classification

ML algorithms have been used extensively by the LHC experiments for signal-to-background classification. One of the first and foremost application of ML in HEP was on the Higgs boson discovery analyses preformed by CMS [3] and ATLAS [4] collaborations. Both analyses use a BDT for the classification of small signal over large and smoothing falling background. More recently, the search for associated production of H and top quark-antiquark pair, where H decays to bottom quarks, conducted with the 2016 and 2017 CMS data, implemented ML algorithm for signal-to-background classification. The search provides the first evidence of the production with an observed significance of 3.9σ [190].

One of the first attempts to use ML algorithms in CMS SUSY analyses is the search for top squark pair production in events with single μ or electron, jets and high MET and M_T [191]. The analysis was performed on the 2012 CMS data and multiple BDTs classifiers were optimised in mass bins in order to gain sensitivity in a range of signal kinematics.

With the advent of deep learning more sophisticated ML tools that offer new opportunities were developed. A new structure of ML classifiers that include physics parameters, like mass, in their input variables together with the measured features was introduced. This algorithm is called parametric NN [192] and it can interpolate between the physics parameters values and therefore replace the need for training multiple classifiers.

The parametric NN have been used in two recent SUSY searches conducted on the full Run 2 CMS dataset. The first is the search for electroweak production of $\widetilde{\chi}^{\pm}$ and $\widetilde{\chi}^{0}$ [56]. The analysis targets events with three or more leptons, up to 2 hadronically decaying τ or 2 leptons of the same charge. The results are interpreted in terms of several simplified models covering a wide range of $\widetilde{\chi}^{\pm}$ and

[1] A Long Short-Term Memory is an artificial recurrent neural network with feedback connections that can process not only single data points but also sequences of data

$\widetilde{\chi}^0$ production and decay scenarios. The analysis uses a parameteric NN to target several models with large SM background contamination.

The other example of SUSY analysis that uses the parametric NN tool is the top "corridor" that was highlighted also in Sec. 1.9. The masses of the top squarks and the neutralinos are introduced to the parametric deep NN as input together with other observables. The NN is trained on signal events with top squark decaying to top quark and neutralino and $t\bar{t}$ events as the SM background.

6.2 Introduction to SUSY sensitivity study with ML algorithms

The SUSY sensitivity study is conducted as part of a deliverable project of the Horizon 2020-funded Innovative Training Network named AMVA4NewPhysics. All the research activity of the training network can be found in [193]. The objective of the study is to explore the use of ML methods on signature specific searches for new physics. More specifically, the current chapter discusses the sensitivity reach of a SUSY search with and without the use of ML tools.

The study focuses on compressed mass SUSY signatures with three soft leptons and p_T^{miss} in the final state. The leptons arise from pair produced, mass degenerate $\widetilde{\chi}_1^{\pm}$-$\widetilde{\chi}_2^0$ that decay to off-shell W and Z bosons, depicted in the left diagram in Fig. 4.1. The bosons decay leptonicaly to three soft leptons and a neutrino together with $\widetilde{\chi}_1^0$ which both account for the p_T^{miss} of the event. This is one of the decay channel of the main analysis workflow as described in Sec. 4.1. The mass of $\widetilde{\chi}_2^0$ and $\widetilde{\chi}_1^{\pm}$ are considered degenerate and the $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0}$ is small. For the purpose of the study and the training of the ML algorithm the specific signal mass point with $M_{\widetilde{\chi}_2^0} = M_{\widetilde{\chi}_1^{\pm}} = 225$ GeV and $\Delta m_{\widetilde{\chi}_2^0\text{-}\widetilde{\chi}_1^0} = 20$ GeV is considered.

The sensitivity study is performed with 2016 simulation with the setup of the 2016 baseline analysis [76]. Components such as trigger and lepton scale factors are extrapolated directly from the 2016 analysis and a detailed discussion on their derivation is out of the scope of this chapter. A description of their calculation can be found in the analysis summary note [194]. Therefore, the analysis steps and the results of this chapter should not be compared directly to the full Run-2 baseline analysis of Chap. 4.

All the SM processes that result in final states with soft leptons and p_T^{miss} or can mimic such a signature, contribute to the background of the search. The dominant prompt SM background arises from diboson (WZ) leptonic decays. Additionally, the non-prompt or fake background arises from events with leptons that are non isolated or they are produced away from the primary vertex or they are hadrons mimicking the leptons. The definitions of the non-prompt backgrounds are given in Sec. 4.7.3. In the 3ℓ final state the main source of non-prompt background is the $t\bar{t}$ process. The background processes of the $t\bar{t}$ and

W+jets and the signal sample are simulated in leading-order with Madgraph5 event generator [195]. For diboson and single top processes next-to-leading-order Madgraph5 aMCatNLO [196] and Powhegv1.0 [166, 167, 168] generators are used. Showering and hadronization is done by Pythia [114] package and the detector simulation by the GEANT [115] package. For the TChiWZ simplified signal model a fast detector simulation is used [197].

The sensitivity reach of the search is obtained with two methods; implementing ML algorithm for the classification of the signal and the background events, and using the baseline approach in which a sensitive kinematic variable is used for the event binning. Both methods are performed with the same setup and the 2016 simulation. This allows for a fair comparison between the two results.

6.3 Object definition

The standard physics objects are used as provided by the CMS POG. The leptons are reconstructed using the PF algorithm discussed in Sec. 3.3.1 and their p_T and η are required to be inside the trigger acceptance within the boundaries of the tracker. For the leading lepton this is translated to $p_T > 5$ GeV and $|\eta| < 2.5$. An upper bound at 30 GeV on the leading lepton p_T is applied, similarly to the baseline analysis. The motivation for this requirement is described in Sec.4.6.1. The subleading and trailing electrons(muons) are required to have $p_T > 5(3.5)$ GeV.

The muons should pass the Loose and Soft ID [111]. Prompt electrons are identified using a MVA discriminant obtained from the training of a boosted decision tree using the ECAL shower shape and the track quality as input features. This electron MVA discriminant was developed prior to the one used in the baseline full Run–2 analysis of Chap. 4. The loose WP as employed by the $H \to ZZ \to 4\ell$ analysis [198] is used for identifying electrons with $p_T < 10$ GeV and a tighter WP is used for electrons with $p_T > 10$ GeV. Additionally, the electron track should not lack any pixel hit and should not be associated to a conversion vertex.

Events with a jet with $p_T > 25$ GeV, that were tagged as b-jets by passing the loose WP of the Combined Secondary Vertex (CSV) tagger [199] are vetoed in event selection.

One important aspect of performing data analysis with ML algorithms is to have an adequate dataset that allows for robust training and validation of the model. The event and object selection requirements applied in the baseline analysis are designed to discard the background events and non-prompt leptons and maintain the signal. This results in a very small final dataset, not sufficient for training a robust classifier. Therefore, the event and object selection requirements that are not related to trigger requirements or the definition of the phase space searched by the analysis (such as the upper p_T cut of the leptons) will be relaxed in the ML study. Regarding the object identification the selection cuts presented in the second block of Tab. 4.2 and 4.3 are relaxed. Specifically, the IP_{3D}, $\sigma_{IP_{3D}}$, d_{xy}, d_z, and the cuts on the absolute and relative Isolation. These requirement refer to the "promptness" of the leptons, therefore it is expected that their relaxation will affect the number of non-prompt or fake leptons that enter the SRs.

6.4 Event selection

The HLT trigger algorithms that were used for the event selection play a major role in the design of the SR, similarly to what is discussed in Sec. 4.4. A pure p_T^{miss} trigger algorithm is used for the selection of events with offline $p_T^{miss} > 200$ GeV, and a double-μ plus p_T^{miss} is used in the SR with lower offline p_T^{miss} in the range of 125–200 GeV. The collected integrated luminosity of the pure p_T^{miss} trigger is 35.9 fb^{-1} and of the double-μ plus p_T^{miss} is 33.2 fb^{-1} because it became available for datataking after the 20^{th} of June in 2016.

The selected events have exactly three leptons with at least one pair of same flavor and opposite sign leptons. This is motivated from the signal signature in which the two final leptons originate from the decay of the off-shell Z boson. According to the lepton flavor requirements of the trigger algorithms, at least two muons are required in the the events of the low MET SR bin and no requirement of the lepton flavor is applied in the events of the high MET SR bin. The minimum dilepton invariant mass is required to be within the range of 4–50 GeV.

All the event selection cuts that veto resonances, presented in Tab.4.5, are relaxed in order to retain the background events that would otherwise be rejected.

In summary the selection criteria applied in the 3ℓ SRs of the sensitivity study with ML are presented in Table 6.1

All the prompt and non-prompt backgrounds are taken from simulation which are corrected by scale factors. The signal and background simulated events are weighted by the cross-section of the process, the luminosity of the corresponding MET SR bin, the trigger scale factors, the lepton scale factors and the PU weights for 2016. The lepton reconstruction and selection efficiency was measured with the tag-and-probe method and the scale factors are extracted by comparing the data and simulation efficiency [194]. The trigger efficiency, that a trilepton event selected with the analysis criteria would fire one of the triggers, was evaluated and the SF were estimated by data-MC comparison for the two MET SR bins separately. Additionally, a normalization scale factor of 1.2 is applied on the WZ prompt background. This scale factor was measured by comparing the data and simulation in a dedicated WZ CR.

Th number of signal and background events that pass the event selection are shown in Table 6.2.

Criterion	3ℓ-ewk SR
Triggers	✓
N_{lep}	=3
OS pair	✓
SF pair	✓
$p_{\mathrm{T}}^{\ell 1} > 5$ GeV	✓
$p_{\mathrm{T}}^{\ell} < 30$ GeV	✓
$p_{\mathrm{T}}^{\ell 2} > 5$ if e (3.5 if μ) GeV	✓
$p_{\mathrm{T}}^{\ell 3} > 5$ if e (3.5 if μ) GeV	✓
$\lvert\eta\rvert < 2.5$	✓
$4 < M_{\ell\ell,SFOS}^{min} < 50\,\mathrm{GeV}$	✓
$p_T^{miss} > 125\,\mathrm{GeV}$	✓

Table 6.1: List of the event selection criteria in the 3ℓ SR of the sensitivity study with ML algorithms.

Samples	**Number of events**	
	Low MET bin	High MET bin
TChiWZ 225/20	4475	8659
WZ (prompt)	22	24
ZZ (prompt)	6	2
$t\bar{t}$(fakes)	126	64
t(fakes)	13	3
tW(fakes)	11	3
W+Jets(fakes)	13	33
DY+jets(fake)	6	12
Total Background	197	141

Table 6.2: The number of background and signal events that pass the event selection cuts.

All the background processes are merged into one inclusive background "pool" which is further divided randomly in half, into training and testing samples. The signal is divided into training and testing with the same random procedure.

6.5 Training Variables

Multiple kinematic variables such as transverse momenta, η and ϕ, invariant masses of leptonic pairs, vector and scalar sum of the transverse momenta of the leptons, and others are studied in the early stages of the study. The variables described in Table 6.3 are found to be the most discriminating among the tested ones and are chosen for the classifier's training. The distributions of the variables in the two MET bins are shown in Fig. 6.1 and 6.2.

Kinematic variable	Short description
p_T^{miss}	The missing energy on the plane transverse to the beam
$p_T^{3\ell}$	The vector sum of the transverse momenta of the 3 leptons
$L_T^{3\ell}$	The scalar sum of the transverse momenta of the 3 leptons
$p_z^{3\ell}$	The sum of the momenta of the 3 leptons on the z axis (the beam axis)
$M_{\ell\ell,SFOS}^{min}$	The minimum invariant mass of a same-flavour, opposite-sign leptonic pair
$\Delta\text{R}_{\ell\ell}^{max}$	The maximum separation of all pairs of final-state leptons in the (η, ϕ) space
$\Delta\phi_{\ell-E_T^{miss}}^{max}$	The minimum separation between any lepton and the missing transverse momentum vector in the (r, ϕ) space.

Table 6.3: Kinematic variables used for the training of the classifier. ΔR is defined as $\sqrt{\Delta\eta^2 + \Delta\phi^2}$.

Figure 6.3 presents the correlation matrix of the training variables for background. The matrix is produced with simulated events that populate the Low MET SR bin. The correlations in the High MET bin are similar. The most significant correlation is observed between the $L_T^{3\ell}$ and the $p_T^{3\ell}$ variables as expected. This is attributed to the fact that they are the vector and scalar sum of the same quantities (the lepton transverse momenta).

6.6 ML algorithm for signal and background classification

The Scikit-learn library [200] for python [201] is used for the scope of this study. Multiple classifiers were trained namely, a Naive Bayes classifier, a Gradient Boosting Tree, a Multi-layer Perceptron NN and a Random Forest. The most efficient was chosen and used for the classification of the testing sample events.

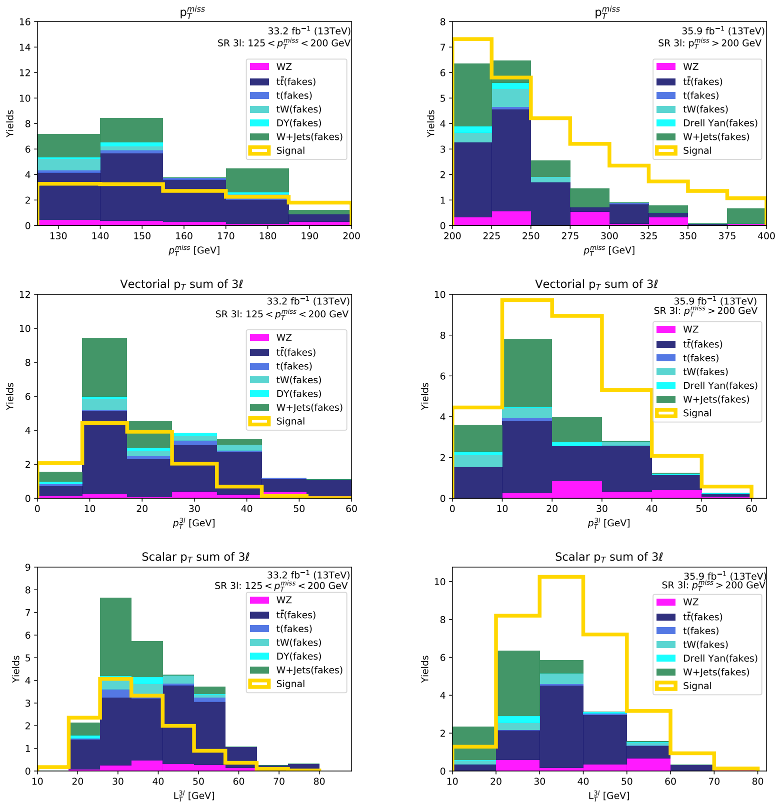

Figure 6.1: The left(right) column shows the variable in the Low(High) MET SR bin. The top row shows the p_T^{miss}, the middle row shows the vector sum of the p_{T} of the three leptons ($p_T(3\ell)$) and the bottom row shows the scalar sum of the p_{T} of the three leptons ($L_T(3\ell)$).

6.6.1 K-Fold cross validation

The K-Fold cross validation is a method to evaluate the ML algorithm with the minimum training data "wasting". It is a resampling procedure, with one parameter K. Instead of splitting the training sample into training and validating sets in order to validate the algorithm, the K-Fold cross validation method splits the training dataset into K smaller sets. The model is trained K times using the K-1 sets in each iteration. The resulted model is validated on the remaining set. In the end the mean of the K trained models is used for the final predictions. For the study the K-Fold cross validation with 5 folds is used for the training of all the classifiers. Figure 6.4 presents the splitting of the dataset when the K-Fold cross validation with 5 folds is used.

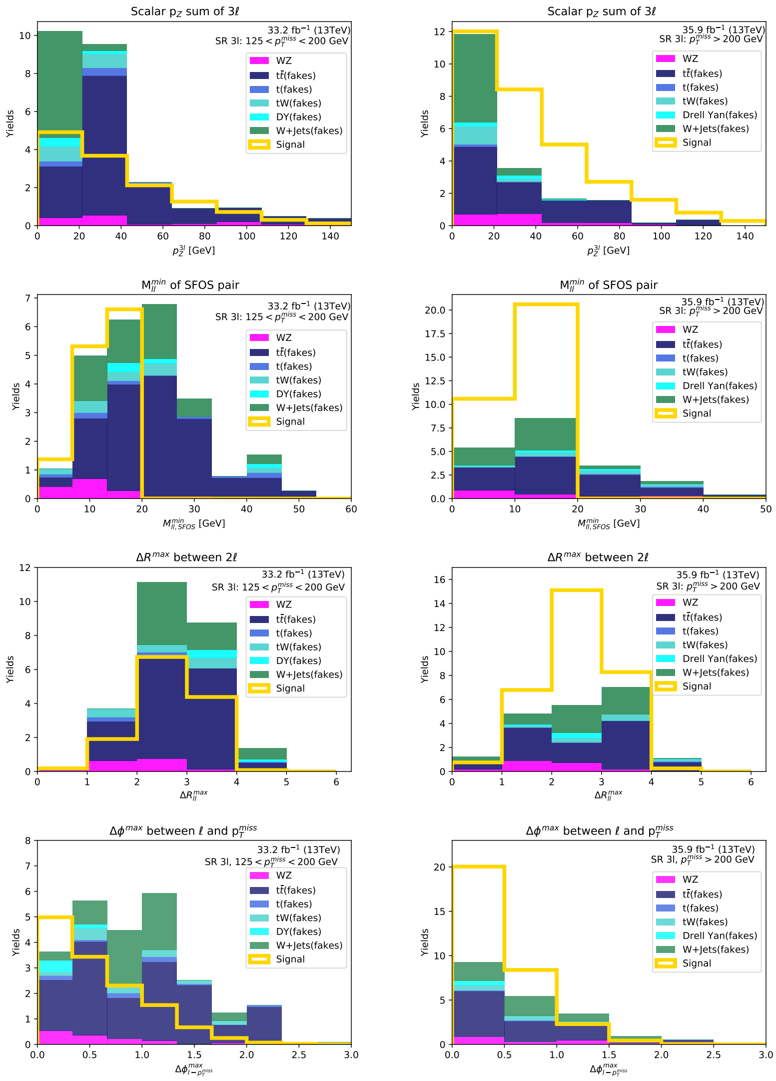

Figure 6.2: The left(right) column shows the variable in the Low(High) MET SR bin. The top row shows the scalar sum of p_Z of the three leptons, the second row shows minimum dilepton mass of a same flavor and opposite sign leptonic pair, the third row shows the minimum separation between any lepton and the missing transverse momentum vector in the (r, ϕ) plane and the bottom row shows maximum separation of all pairs of final state leptons, in the (η,ϕ) plane.

6.6.2 Algorithm choice and optimisation

The parameters of the classifiers are tuned for maximum efficiency. The optimal parameters of the models are found with the grid search method. This generates

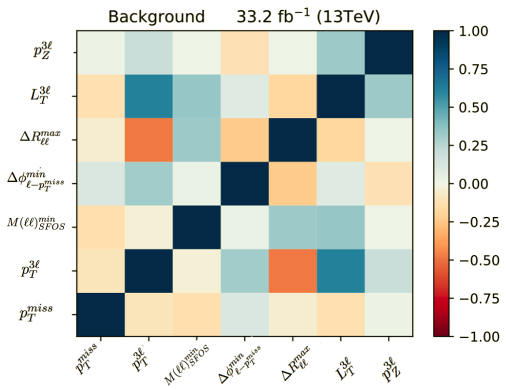

Figure 6.3: The correlation matrix evaluated on the background evets in the Low MET SR bin

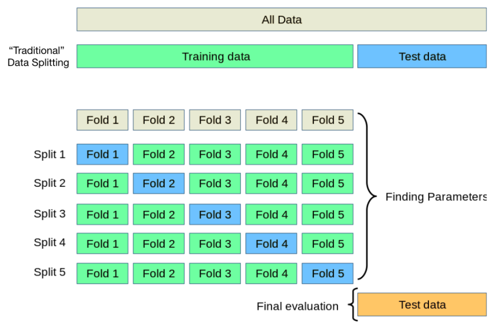

Figure 6.4: A schematic representation of the K-Fold cross validation training data splitting.

a set of hyperparameters candidates [2] from a grid of values specified by the user. The optimal parameters are chosen based on a figure of merit decided by the user. Every model is fitted separately and all the possible combinations of parameter values are evaluated. The best combination of parameters is chosen. The grid search for every classifier is performed using the area under receiver operating characteristic (ROC) curve as a figure of merit, with the K-Fold cross validation with 5 folds implemented.

The ROC curve illustrates the performance of a ML algorithm and pictures the proportion of signal events identified as such over the total signal events (true positive rate) versus the ratio of background events categorised as signal over the total amount of actual background events (false positive rate).

The ROC curves of the four trained ML models are presented after the short description of the classifiers and their properties.

[2]The hyperparameters are refereed to the parameters of the classifiers such as min sample leaf, loss function etc.

Naive Bayes classifier

The Naive Bayes is one of the simplest and fastest supervised classification algorithm. It is based on the Bayes theorem with the assumption that the training features are independent. This assumption simplifies largely the computations and justifies the name of the classifier. The Bayes theorem states that

$$P(y|x_1, ..., x_n) = \frac{P(x_1, ..., x_n|y)P(y)}{P(x_1, ..., x_n)} \tag{6.2}$$

where y is the class corresponding to either signal or background in our binary classification problem, $x_1, ..., x_n$ are the input features values. $P(y|x_1, ..., x_n)$ is the posterior probability of a class given a set of feature values, $P(x_1, ..., x_n|y)$ is the likelihood of a feature for a given class, and $P(y)$ and $P(x_1, ..., x_n)$ are the prior probabilities of a class and the features respectively.

The naive assumption of independent input features results in

$$P(x_i|y, x_1, .., x_{i-1}, x_{i+1}, ..., x_n) = P(x_i|y) \tag{6.3}$$

and Eq. 6.2 can be written as

$$P(y|x_1, ..., x_n) = \frac{P(y)\prod_{i=1}^{n} P(x_i|y)}{P(x_1, ..., x_n)} \tag{6.4}$$

The posterior probability for the different classes is calculated and the one with the highest probability is selected. The maximum posterior probability is called maximum aposteriori estimation $MAP(y)$. The denominator in Eq. 6.4, $P(x_1, ..., x_n)$ is a constant normalization term and can be dropped. Therefore $MAP(y)$ is written as

$$MAP(y) = max(P(y)\prod_{i=1}^{n} P(x_i|y)) \tag{6.5}$$

During the training process of the Naive Bayes classifier the prior probability of every class and the likelihood of the feature given a class are calculated. The prior probability of every class, $P(y)$, is measured as the frequency of the instances that belong to the same class divided by the total number of instance. The likelihood or conditional probability of the features given the class is measured as the frequency of each value for a given class, divided by the frequency of instances that belong to that class. The prediction for new data is made by estimating the $MAP(y)$.

In this study the likelihood $P(x_i|y)$ is assumed to be a Gaussian. Therefore, the mean and the standard deviation for each input feature for every class are stored together with the prior probability of every class.

In spite of the over-simplified assumptions, Naive Bayes classifiers have performed well in many real-world problems such as in spam filtering.

Ensemble methods

The ensemble methods combine the prediction of several simple estimators and provide improved robustness and generalization over a single estimator. These simple estimators are sometimes called weak learners and in the case of this study are decision trees.

There are two broad categories of ensembling that will be discussed in this section

1. *Boosting* is based on the sequential building of estimators from simple decision trees. In this study the gradient boosting classifier (GBC) is used.
2. *Averaging* is based on building independent estimators and average their prediction. The combined estimators are expected to demonstrate reduced variance compared to the single decision tree. A random forest classifier (RFC) that is based on average ensemble method is used in this study.

Both averaging and boosting methods are built from the decision trees. These are predictive models that learn by simple decision rules. They can output categorical or numerical prediction for classification or regression problems respectively. A schematic representation of a classification decision tree is shown in Fig. 6.5. The objective of the model is to go from the input observables x_i to predictions about their target values or labels (S,B). The tree starts from the root node and performs a sequence of binary splits of data in every node, according to decision rules that set cuts on the training variables x_i. Each event of the dataset ends up in one of the terminal leaf nodes and classified as signal S or background B.

The decision rules are based on specific algorithms and recursively generate the structure of the decision tree. At each step the variable that best splits the set of data is chosen. The quality of the separation is measured with metrics of impurity such as the Gini index [203]. In the case of a binary classification problem the Gini index depends on the probabilities of assigning a random event in class A or in class B, therefore it is a measure of event missclassification.

$$Gini\ Index = p_s p_b + p_b p_s \tag{6.6}$$

where p_s and p_b are the purity of signal and background respectively. The purity of signal can be written also as $p = p_s = 1 - p_b$. Therefore the Gini Index can be expressed as $Gini\ Index = 2p(1-p)$. The result of the decision learning process is the minimisation of the Gini Index and thus the impurity.

Gradient Boosted classifier

The GBC is one of the most powerful techniques for building a predictive model. In the boosting technique the weights of training misclassified events are

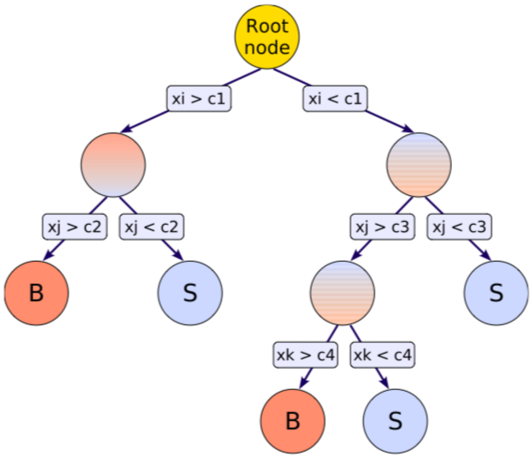

Figure 6.5: Schematic representation of a simple decision tree for S-B classification. Image source: [202].

increased and new trees are added. This is an iterative process and is repeated as many times as the number of boosting stages. The additional trees are trained on the set of data that were misclassified by the prior model while the weights of the previous trees are left unchanged. The new decision trees are fit to correct the residual impurities by prior trees. The models are fit using a differentiable loss function, that quantifies the difference between the real value and the prediction. The optimization is based on the minimization of the gradient loss, hence the name of the model.

For the purposes of this study two GBC were used for the classification of the events in the two SR MET bins. The parameters of the models also referred to as hyperparameters are presented in Table 6.4 and discussed below.

Hyperparameters	**Low** p_T^{miss}	**High** p_T^{miss}
splitting criterion	Friedman MSE	
loss function	deviance	
max features	0.5	
min sample leaf	1	
min sample split	2	
number of decision trees	200	
sub-samples	0.5	
max tree depth	6	1
learning rate	0.02	0.2

Table 6.4: The values of the hyperparameters of the two GBC used for event classification in the two SRs.

The Friedman MSE is a metric for measuring quality of the split by using the mean square error with improvement score by Friedman [204]. This splitting requirement depends on the mean squared errors which is the average squared difference between the estimated values and the actual value and a global weight which denotes the probability of a class in every node after the split.

The loss function has to be differentiable and it is used for the optimization of the classifier. The impurity minimization is performed by applying gradient descent to reduce the loss. The deviance or logistic regression function is used in this study. This is a sigmoid function that takes real numbers and map them into 0 and 1. It is similar to the linear regression in the sense that input values are combined linearly using coefficent values in order to predict the output value. However, it differs from the linear regression in that its output is modeled in binary values.

The maximum features hyperparameter denotes the number of features to be considered when looking for the best split. A float value means that the integer of number of features $\times$ max features or in this case 3 features are used when searching for the best split.

The minimum sample leaf refers to the minimum training data point (also referred to as samples in ML) required at every leaf node for a node splitting to be considered. The minimum sample split defines the minimum number of samples required for an intermediate node split to be performed. The number of decision trees denotes the number of boosting stages to be performed and the sub-sample hyperparameter reflects the fraction of samples to be used for fitting individual decision trees. The learning rate defines the step size for every iteration while moving towards the minimization of the loss function. Finally, the maximum tree depth denotes the maximum depth of the individual trees and essential limits the number of nodes in them. The learning rate and the maximum tree depth differs in the two classifier due to the difference in the sample statistics in the two SR bins.

Random Forest classifier

The RFC is based on the averaging ensemble method and it is built from individual decision trees. Each of the constituents decision trees is built from a random sample drawn from the training set and during the training the best split is found exploiting a random subset of half of the input features. The prediction is given as the average prediction of the individual classifiers. The randomness of the model decreases its variance and prevents from overfitting. Essentially its main difference from the gradient boosting classifier is that the constituents of the ensemble are built independently.

The hyperparameters of the RFC models that were used are presented in Table 6.5

Hyperparameters	Low p_T^{miss}	High p_T^{miss}
Splitting function	Gini index	
max features	0.5	
min sample leaf	1	
min sample split	2	
number of trees in the forest	200	
max depth of the tree	1	3

Table 6.5: The values of the hyperparameters of the two RFC used for event classification in the two SRs.

Multi-layer Perceptrons

The NN is another very popular ML algorithm, widely used for classification and regression. It consists of the input variables, the layers of neurons or hidden layers, and the output which is the prediction of the model. Similarly to the previously described ML algorithms, NN are trained to learn how to make predictions.

The NN typically consist of multiple layers of connected neurons, each receiving the output of the previous layer. Each neuron transform the outputs of the previous layer by applying a weighted linear summation, followed by a non-linear activation function. The weight are updated and adjusted during the training through an optimization process, in order to make an accurate prediction. The goal is to build a function that maps the input to the output. A loss function is used to measure the difference between the true and predicted value and it is minimized with gradient descent during the training. The gradient of the loss function is evaluated at every point of the parameter space and it is descending towards the minimum. This is evaluated with the backpropagation method.

The NN training process can be briefly summarized in the following steps:

1. Forward propagation through the layers of the network in order to evaluate the loss function.

2. Backward propagation during which the gradient of the loss function is evaluated in every neuron.

3. Take a step down the gradient and update the weight in every neuron according to the slope of the loss function at this point. The step is equal to the learning rate of the model.

In the sensitivity study of this chapter, multi-layer perceptrons (MLP) are used. MLP are a category of feed forward, fully connected NN, that are trained with the backpropagation method and use the rectified linear unit activation function called RelU. The RelU function returns $f(x) = max(0, x)$ where x is the input of the neuron.

Two MLPs are trained for the Low and High MET SR bins and they both consist of 2 hidden layers. The first MLP contains 8 neurons and the second 6 neurons in the first hidden layer. The second hidden layer consists of 1 neuron in both MLPs. The maximum number of iterations in both MLPs is set to 200 and a regularization term of 0.00001 to avoid overfitting is used.

6.6.3 Performance studies

The plots in Fig. 6.6 show the comparison between the ROC curves of the four trained classifiers. In both MET SR bins the GBC is found to be the most efficient and it is used for the event classification. The curves presented in the plots are the mean of the 5-Fold cross validation process of the optimized classifiers

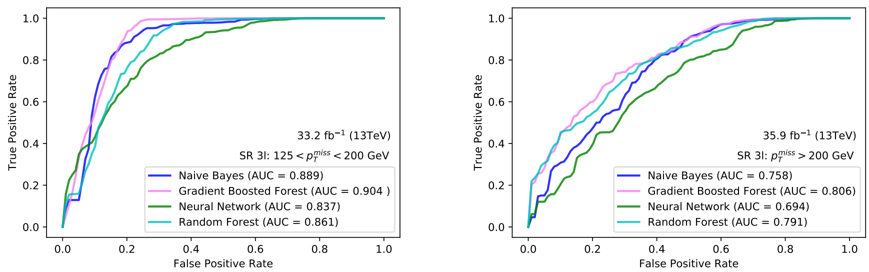

Figure 6.6: ROC curves in the two MET SR bins (Left: Medium MET bin, 125 $< p_T^{miss} <$ 200 GeV; Right: High MET bin, $p_T^{miss} >$ 200 GeV). The curves represent the mean value of 5-fold cross validation. The different curves show the performance of the four classifiers.

The importance of the features is computed as decrease in the impurity of the node weighted by the probability to reach the node. The latter probability is estimated by the number of samples that reach the node divided by the total number of samples. The higher the value the more important the feature. The training feature of the study in order of importance for the Low and High MET GBC are shown in Table 6.6 and Table 6.7

6.7 Results

This section presents the method for the estimation of the systematic uncertainties and the templates used in the final fit for the exclusion limit calculation. Additionally, the corresponding steps of the baseline approach to be compared to the ML study are discussed. The asymptotic limit of the signal strength, and the significance of the signal mass point are presented here. In addition to the signal mass point that was used for training of the classifier, a new point of $M_{\tilde{\chi}_2^0} = M_{\tilde{\chi}_1^\pm}$

Feature	Importance
$M_{\ell\ell,SFOS}^{min}$	21%
$\Delta\phi^{min}(\ell - p_T^{miss})$	15%
$p_T^{3\ell}$	14%
$\Delta R_{\ell\ell}^{max}$	13%
p_T^{miss}	12%
$p_Z^{3\ell}$	12%
$H_T^{3\ell}$	10%

Table 6.6: The training features of the Low MET GBC in order of importance.

Feature	Importance
$M_{\ell\ell,SFOS}^{min}$	18%
$\Delta\phi^{min}(\ell - p_T^{miss})$	17%
$p_Z^{3\ell}$	14%
$\Delta R_{\ell\ell}^{max}$	14%
$H_T^{3\ell}$	13%
$p_T^{3\ell}$	12%
p_T^{miss}	11%

Table 6.7: The training features of the High MET GBC in order of importance.

= 250 GeV and $\Delta M_{\tilde{\chi}_2^0-\tilde{\chi}_1^0}$ = 40 GeV is used for testing and the sensitivity for this mass point is computed. The purpose of the test is to investigate the power of the classifier on different mass points and the gain on the sensitivity of the search, if any.

It should be noted that the sensitivity of the baseline method that is presented here, is not the official result of the CMS sensitivity on the compressed electroweak SUSY channel with the 3ℓ in the final state. The official CMS results of the full Run–2 analysis are presented in Chapter 4.

Signal and Background templates

The HiggsCombine software tool was employed for the maximum likelihood fit and the limit calculation. In this sensitivity study the SR data is not used for the limit calculation, therefore only the expected limit was calculated. The GBC output and the $M_{\ell\ell,SFOS}^{min}$ distribution are used as the binning variable of the signal and the background in the ML study and the baseline study respectively. The plots in Fig. 6.7 show the shapes of the GBC outputs for the signal and the total background. In this plots the GBC output is computed when training and testing is done on the same signal mass point ($M_{\tilde{\chi}_2^0} = M_{\tilde{\chi}_1^\pm}$ = 225 GeV and $\Delta M_{\tilde{\chi}_2^0-\tilde{\chi}_1^0}$ = 20 GeV). The number of bins used in the signal and background distributions is limited to three in order to minimise the impact of our ignorance of the distributions' shape uncertainties.

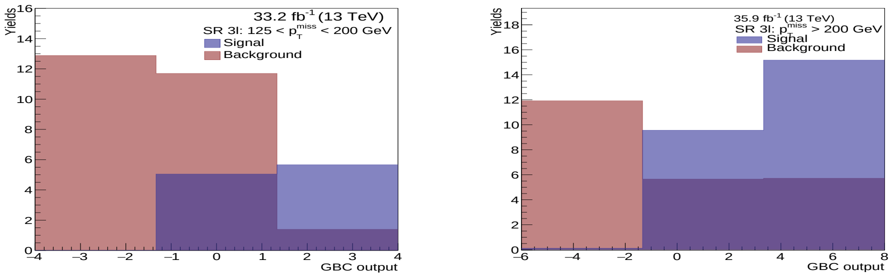

Figure 6.7: GBC output distribution of signal mass point $\mathrm{M}_{\tilde{\chi}_2^0} = \mathrm{M}_{\tilde{\chi}_1^\pm} = 225$ GeV and $\Delta\mathrm{M}_{\tilde{\chi}_2^0 - \tilde{\chi}_1^0} = 20$ GeV and total background. Low MET bin, 125< $p_T^{miss} < 200$ GeV (left); High MET bin, $\mathrm{p}_T^{miss} > 200$ GeV (right)

The most important systematic uncertainties of the baseline analysis are extrapolated to the ML study. These uncertainties are related to the detector response and inefficiencies in the event reconstruction. A summary of the uncertainties is presented below:

- Uncertainty on the 2016 luminosity measurement is 2.5% [205].
- A 50% uncertainty is assigned on the di-boson background normalization. The 50% accounts for the data-simulation discrepancy agreement checked in dedicated di-boson region.
- A 100% systematic uncertainty is applied on the non-prompt background because it is estimated purely by simulation and 100% uncertainty is also applied on the rare SM processes such as the tW
- On the trigger efficiency measurement a flat 5% is applied
- On the lepton scale factor a 6% is applied

The above systematic uncertainties are included in the fit as log-normal distributed nuisance parameters. Additionally, two shape uncertainties were considered, namely the uncertainty on the b-tagging scale factor and the JEC. For the shape uncertainties, the GBC output was calculated by varying the b-tagging scale factor and the JEC up and down by one standard deviation both in signal and backgrounds.

Asymptotic limit on the signal strength

The limit relies on an asymptotic approximation of the distribution of the test-statistics, based on profile likelihood ratio, under a signal plus background hypothesis, in order to compute the Confidence Level of background only hypothesis

(CL_S). Section 5.1 presents a detailed discussion on the steps of the method. As a binning variable in the final fit the $M^{mis}_{\ell\ell,SFOS}$ was used in the case of the baseline-like analysis, while the GBC output is used in the case of the ML study. The uncertainties mentioned earlier were implemented in the two different cases.

A difference that should be noted between the baseline and the ML method is the method for the non-prompt background estimation. The baseline analysis uses a data-driven tight-to-loose method for the non-prompt background estimation. This provides a better description and a robust control of the background. In the ML study the non-prompt background is taken from the simulation and a large systematic uncertainty is assigned to account for its missmodelling.

The a-priori expected significance was calculated as the ratio of profile likelihoods, one in which the signal strength is set to 0 and the other in which it is free to float. The significance is given by equation 6.7:

$$Significance = -2ln[\frac{\mathcal{L}(r=0,\hat{\theta}_0)}{\mathcal{L}(r=\hat{r},\hat{\theta})}] \tag{6.7}$$

The results of the fit together with the estimated significance are shown in Tables 6.8, 6.9.

Figure of Merit	ML approach	Baseline approach
Asymptotic limit r	< 0.90	< 1.67
68 % CL	[1.53, 1.69]	[1.10, 2.66]
95 % CL	[0.35, 2.72]	[0.78, 4.13]
Significance	2.8	1.40

Table 6.8: Asymptotic limit, one- and two-standard deviation confidence level and the a-priori expected significance, calculated in the multi-variate and baseline-like analysis for the same signal mass point $M\widetilde{\chi}^0_2 = M\widetilde{\chi}^\pm_1$ =225 GeV and $\Delta M_{\widetilde{\chi}^0_2-\widetilde{\chi}^0_1} =$ 20 GeV

Figure of Merit	ML approach	Baseline approach
Asymptotic limit r	< 2.21	< 2.32
68 % CL	[1.30, 4.13]	[1.45, 3.91]
95 % CL	[0.86, 6.70]	[0.98, 6.29]
Significance	1.38	1.20

Table 6.9: Asymptotic limit, one- and two-standard deviation confidence level and the a-priori expected significance, calculated in the multi-variate and baseline-like analysis for the same signal mass point $M\widetilde{\chi}^0_2 = M\widetilde{\chi}^\pm_1$ =250 GeV and $\Delta M_{\widetilde{\chi}^0_2-\widetilde{\chi}^0_1} =$ 40 GeV, when the classifier has been trained on the $M\widetilde{\chi}^0_2 = M\widetilde{\chi}^\pm_1$ =225 GeV and $\Delta M_{\widetilde{\chi}^0_2-\widetilde{\chi}^0_1} = 20$ GeV signal mass point.

6.8 Discussion

The section presented the study of ML algorithms for the problem of the signal and background classification in the context of model specific searches for BSM physics. The SUSY search focuses on the electroweakino leptonic decay in the compressed mass spectrum scenario in final state with three soft leptons and missing transverse momentum. Throughout the study, the results of the MVA approach are compared to the respective baseline-like study. The background samples used in both cases are the same and all the scale factors and most of the systematic uncertainties have been extrapolated from the baseline-like to the multi-variate study. The study is presented in the public document of the AMVA4NewPysics ITN, at the research and innovation participant portal [206].

It was of our interest to compare the MVA results to the baseline-like results in order to estimate the gain in sensitivity, if any. It was found that the multi-variate method is more powerful than the baseline-like approach, providing a more stringent limit along with a higher a-priori expected significance. The results show that the MVA discriminant brings an improvement in the sensitivity of the baseline-like approach by a factor of 2 on significance and of 1.8 on the signal strength exclusion limit, when the training and the testing is done on the same signal mass point. The two methods, show similar performance when the MVA discriminant is tested on a different signal mass point indicating that the method is biased towards the signal mass point that it was trained on. Future work in this direction involves training on multiple signal mass points or use of signal masses as parameters in the training (parametrised raining) in order to develop a more universal result. An alternative approach, of training on multiple signal mass points could possibly lead to a more universal result. This method requires high resources and a lot of time for training and therefor it is a sub-optimal solution.

The Monte Carlo background samples used in the MVA study are the same samples used in the baseline analysis in which event number does not necessarily need to be large. In the case of an MVA study large samples are needed in order to achieve an unbiased and trustworthy training. The limitation on the number of the background events in the official samples instructed the loosening of the event selection, by removing the lepton Iso, IP3D and Sip3D cuts. The latter enlarged the phase space and enabled the increase on the sample statistics. However, the optimal solution would be the usage of larger background samples, where the tight selection cuts could be applied and the surviving events would be enough for the purposes of the MVA study. The latter solution would require a large sample production campaign that was out of the scope and the timeline of the study.

To conclude, with the advance of the machine learning tools the MVA methods are progressively used more widely in BSM searches and in high energy physics in general. The MVA methods have been used for online event triggering, event simulation, jet identification and event classification in multiple analyses. They are expected to bring great improvements in all of the aforementioned fields. However it should be noted that the MVA approaches need special care and treatment in order to avoid biases and achieve a robust and trustworthy result.

Bibliography

[1] M. Thomson. *Modern particle physics.* Cambridge University Press, New York, 2013.

[2] S. P. Martin. A supersymmetry primer. *Advanced Series on Directions in High Energy Physics*, page 1–98, Jul 1998.

[3] S. Chatrchyan et al. Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC. *Phys. Lett. B*, 716:30–61, 2012.

[4] G. Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B*, 716:1–29, 2012.

[5] P. W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964.

[6] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964.

[7] Summary of CMS cross section measurements.

[8] P. A. Zyla et al. Review of Particle Physics. *PTEP*, 2020(8):083C01, 2020.

[9] M. Banner et al. Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider. *Phys. Lett. B*, 122:476–485, 1983.

[10] G. Arnison et al. Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540$ GeV. *Phys. Lett. B*, 122:103–116, 1983.

[11] S. Schael et al. Electroweak Measurements in Electron-Positron Collisions at W-Boson-Pair Energies at LEP. *Phys. Rept.*, 532:119–244, 2013.

[12] T. Aaltonen et al. Precise measurement of the W-boson mass with the CDF II detector. *Phys. Rev. Lett.*, 108:151803, 2012.

[13] V. Mukhamedovich Abazov et al. Measurement of the W Boson Mass with the D0 Detector. *Phys. Rev. Lett.*, 108:151804, 2012.

[14] M. Aaboud et al. Measurement of the W-boson mass in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *Eur. Phys. J. C*, 78(2):110, 2018. [Erratum: Eur.Phys.J.C 78, 898 (2018)].

[15] Jens Erler and Matthias Schott. Electroweak Precision Tests of the Standard Model after the Discovery of the Higgs Boson. *Prog. Part. Nucl. Phys.*, 106:68–119, 2019.

[16] T Aaltonen, S Amerio, D Amidei, A Anastassov, A Annovi, J Antos, G Apollinari, Jeffrey Appel, T Arisawa, A Artikov, Jonathan Asaadi, W Ashmanskas, B Auerbach, A Aurisano, Fatin Azfar, William Badgett, Taran Bae, A Barbaro-Galtieri, and V Barnes. High-precision measurement of the w boson mass with the cdf ii detector. *Science (New York, N.Y.)*, 376:170–176, 04 2022.

[17] P. A. Zyla et al. Review of Particle Physics. *PTEP*, 2020(8):083C01, 2020.

[18] K. G. Begeman, A. H. Broeils, and R. H. Sanders. Extended rotation curves of spiral galaxies: dark haloes and modified dynamics. *Monthly Notices of the Royal Astronomical Society*, 249(3):523–537, 04 1991.

[19] G. F. Smoot. Cobe observations and results. *Conference on 3K cosmology*, 1999.

[20] C. L. Bennett, D. Larson, J. L. Weiland, N. Jarosik, G. Hinshaw, N. Odegard, K. M. Smith, R. S. Hill, B. Gold, M. Halpern, and et al. Nine-year wilkinson microwave anisotropy probe WMAP observations: Final maps and results. *The Astrophysical Journal Supplement Series*, 208(2):20, Sep 2013.

[21] N. Aghanim, Y. Akrami, F. Arroja, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, and et al. Planck2018 results. *Astronomy Astrophysics*, 641:A1, Sep 2020.

[22] D. S. Akerib, X. Bai, S. Bedikian, E. Bernard, A. Bernstein, A. Bolozdynya, A. Bradley, D. Byram, S.B. Cahn, C. Camp, and et al. The large underground xenon LUX experiment. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 704:111–126, Mar 2013.

[23] E. Aprile, J. Aalbers, F. Agostini, M. Alfonsi, F. D. Amaro, M. Anthony, B. Antunes, F. Arneodo, M. Balata, and et al. The XENON1T dark matter experiment. *The European Physical Journal C*, 77(12), Dec 2017.

[24] A. Margiotta. The KM3NeT deep-sea neutrino telescope. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 766:83–87, Dec 2014.

[25] M. G. Aartsen, M. Ackermann, J. Adams, J.A. Aguilar, M. Ahlers, M. Ahrens, D. Altmann, K. Andeen, T. Anderson, I. Ansseau, and et al. The IceCube neutrino observatory: instrumentation and online systems. *Journal of Instrumentation*, 12(03):P03012–P03012, Mar 2017.

[26] A. D. Sakharov. Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe. *Sov. Phys. Usp.*, 34(5):392–393, 1991.

[27] H. Georgi and S. L. Glashow. Unity of all elementary-particle forces. *Phys. Rev. Lett.*, 32:438–441, Feb 1974.

[28] L. Susskind. Dynamics of spontaneous symmetry breaking in the Weinberg-Salam theory. *Phys. Rev. D*, 20:2619–2625, Nov 1979.

[29] I. Aitchison. *Supersymmetry in Particle Physics: An Elementary Introduction.* Cambridge University Press, 2007.

[30] I. Antoniadis, N. Arkani-Hamed, S. Dimopoulos, and G. R. Dvali. New dimensions at a millimeter to a Fermi and superstrings at a TeV. *Phys. Lett. B*, 436:257–263, 1998.

[31] P. Fayet. Supersymmetry and weak, electromagnetic and strong interactions. *Physics Letters B*, 64(2):159 – 162, 1976.

[32] P. Fayet. Spontaneously broken supersymmetric theories of weak, electromagnetic and strong interactions. *Physics Letters B*, 69(4):489 – 494, 1977.

[33] H.P. Nilles. Supersymmetry, supergravity and particle physics. *Physics Reports*, 110(1):1 – 162, 1984.

[34] S. Dimopoulos and G. F. Giudice. Naturalness constraints in supersymmetric theories with nonuniversal soft terms. *Phys. Lett. B*, 357:573–578, 1995.

[35] M. Papucci, T. J. Ruderman, and A. Weiler. Natural SUSY Endures. *JHEP*, 09:035, 2012.

[36] A. H. Chamseddine, R. L. Arnowitt, and P. Nath. Locally Supersymmetric Grand Unification. *Phys. Rev. Lett.*, 49:970, 1982.

[37] A. Djouadi et al. The Minimal supersymmetric standard model: Group summary report. In *GDR (Groupement De Recherche) - Supersymetrie*, 12 1998.

[38] J. Alwall, P. Schuster, and N. Toro. Simplified Models for a First Characterization of New Physics at the LHC. *Phys. Rev. D*, 79:075020, 2009.

[39] D. Alves. Simplified Models for LHC New Physics Searches. *J. Phys. G*, 39:105005, 2012.

[40] T. Aaltonen et al. Inclusive Search for Squark and Gluino Production in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$-TeV. *Phys. Rev. Lett.*, 102:121801, 2009.

[41] V. M. Abazov et al. Search for squarks and gluinos in events with jets and missing transverse energy using 2.1 fb^{-1} of $p\bar{p}$ collision data at $\sqrt{s} = 1.96$-TeV. *Phys. Lett. B*, 660:449–457, 2008.

[42] A. M. Sirunyan et al. Search for supersymmetry in proton-proton collisions at 13 TeV in final states with jets and missing transverse momentum. *JHEP*, 10:244, 2019.

[43] Search for squarks and gluinos in final states with jets and missing transverse momentum using 139 fb^{-1} of $\sqrt{s}$ =13 TeV pp collision data with the ATLAS detector. Technical Report ATLAS-CONF-2019-040, CERN, Geneva, Aug 2019.

[44] CMS. Supersymmetry physics results twiki page. https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsSUS.

[45] A. M. Sirunyan et al. Searches for physics beyond the standard model with the M_{T2} variable in hadronic final states with and without disappearing tracks in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Eur. Phys. J. C*, 80(1):3, 2020.

[46] T. Aaltonen et al. Search for the Production of Scalar Bottom Quarks in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV. *Phys. Rev. Lett.*, 105:081802, 2010.

[47] V. Mukhamedovich Abazov et al. Search for Scalar Bottom Quarks and Third-Generation Leptoquarks in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV. *Phys. Lett. B*, 693:95–101, 2010.

[48] G. Aad et al. Search for bottom-squark pair production in pp collision events at $\sqrt{s} = 13$ TeV with hadronically decaying τ-leptons, b-jets and missing transverse momentum using the ATLAS detector. 3 2021.

[49] LEP2 SUSY Working Group. Combined LEP stop and sbottom results 183-208 GeV - ALEPH, DELPHI, L3, and OPAL experiments. note LEPSUSYWG/04-02.1.

[50] V. Mukhamedovich Abazov et al. Search for Pair Production of the Scalar Top Quark in the Electron + Muon Final State. *Phys. Lett. B*, 696:321–327, 2011.

[51] A. M. Sirunyan et al. Search for direct top squark pair production in events with one lepton, jets, and missing transverse momentum at 13 TeV with the CMS experiment. *JHEP*, 05:032, 2020.

[52] G. Aad et al. Search for new phenomena with top quark pairs in final states with one lepton, jets, and missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *JHEP*, 04:174, 2021.

[53] Lep2 SUSY Working Group. Combined LEP Charginos results, up to 208 GeV for large m0. note LEPSUSYWG/01-03.1.

[54] A. M. Sirunyan et al. Searches for pair production of charginos and top squarks in final states with two oppositely charged leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV. *JHEP*, 11:079, 2018.

[55] G. Aad et al. Search for electroweak production of charginos and sleptons decaying into final states with two leptons and missing transverse momentum in $\sqrt{s} = 13$ TeV pp collisions using the ATLAS detector. *Eur. Phys. J. C*, 80(2):123, 2020.

[56] Search for electroweak production of charginos and neutralinos in proton-proton collisions at $\sqrt{s} = 13$ TeV. Technical report, CERN, Geneva, 2021.

[57] M. Aaboud et al. Search for electroweak production of supersymmetric particles in final states with two or three leptons at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector. *Eur. Phys. J. C*, 78(12):995, 2018.

[58] A. M. Sirunyan et al. Search for supersymmetry in final states with two oppositely charged same-flavor leptons and missing transverse momentum in proton-proton collisions at $\sqrt{s} = 13$ TeV. *JHEP*, 04:123, 2021.

[59] Search for chargino-neutralino production in final states with a Higgs boson and a W boson. Technical report, CERN, Geneva, 2021.

[60] M. Aaboud et al. Search for chargino and neutralino production in final states with a Higgs boson and missing transverse momentum at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Rev. D*, 100(1):012006, 2019.

[61] Z. Han, G. D. Kribs, A. Martin, and A. Menon. Hunting quasidegenerate Higgsinos. *Phys. Rev. D*, 89(7):075007, 2014.

[62] Lep2 SUSY Working Group. Combined LEP Selectron/Smuon/Stau Results, 183-208 GeV. note LEPSUSYWG/04-01.1.

[63] A. M. Sirunyan et al. Search for supersymmetric partners of electrons and muons in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Phys. Lett. B*, 790:140–166, 2019.

[64] A. M. Sirunyan et al. Search for direct pair production of supersymmetric partners to the τ lepton in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Eur. Phys. J. C*, 80(3):189, 2020.

[65] Search for direct stau production in events with two hadronic tau leptons in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector. Technical Report ATLAS-CONF-2019-018, CERN, Geneva, May 2019.

[66] S. Dimopoulos and H. Georgi. Softly broken supersymmetry and su(5). *Nuclear Physics B*, 193(1):150–162, 1981.

[67] G. F. Giudice, T. Han, K. Wang, and L. Wang. Nearly Degenerate Gauginos and Dark Matter at the LHC. *Phys. Rev. D*, 81:115011, 2010.

[68] A. Heister et al. Search for charginos nearly mass degenerate with the lightest neutralino in e+ e- collisions at center-of-mass energies up to 209-GeV. *Phys. Lett. B*, 533:223–236, 2002.

[69] J. Abdallah et al. Searches for supersymmetric particles in e+ e- collisions up to 208-GeV and interpretation of the results within the MSSM. *Eur. Phys. J. C*, 31:421–479, 2003.

[70] C. Balázs, M. Carena, and C. E. M. Wagner. Dark matter, light top squarks, and electroweak baryogenesis. *Phys. Rev. D*, 70:015007, Jul 2004.

[71] A. M. Sirunyan et al. Search for the pair production of third-generation squarks with two-body decays to a bottom or charm quark and a neutralino in proton–proton collisions at $\sqrt{s} = 13$ TeV. *Phys. Lett. B*, 778:263–291, 2018.

[72] A. M. Sirunyan et al. Search for Supersymmetry with a Compressed Mass Spectrum in Events with a Soft τ Lepton, a Highly Energetic Jet, and Large Missing Transverse Momentum in proton-proton Collisions at $\sqrt{s} =$ TeV. *Phys. Rev. Lett.*, 124(4):041803, 2020.

[73] A. M. Sirunyan et al. Search for top squarks decaying via four-body or chargino-mediated modes in single-lepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV. *JHEP*, 09:065, 2018.

[74] A. Tumasyan et al. Combined searches for the production of supersymmetric top quark partners in proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$. *Eur. Phys. J. C*, 81(11):970, 2021.

[75] A. M. Sirunyan et al. Search for supersymmetry with a compressed mass spectrum in the vector boson fusion topology with 1-lepton and 0-lepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV. *JHEP*, 08:150, 2019.

[76] A. M. Sirunyan et al. Search for new physics in events with two soft oppositely charged leptons and missing transverse momentum in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Phys. Lett. B*, 782:440–467, 2018.

[77] V. Khachatryan et al. Observation of the rare $B_s^0 \to \mu^+\mu^-$ decay from the combined analysis of CMS and LHCb data. *Nature*, 522:68–72, 2015.

[78] R. Aaij et al. Measurement of the $B_s^0 \to \mu^+\mu^-$ branching fraction and effective lifetime and search for $B^0 \to \mu^+\mu^-$ decays. *Phys. Rev. Lett.*, 118(19):191801, 2017.

[79] G. W. Bennett et al. Final Report of the Muon E821 Anomalous Magnetic Moment Measurement at BNL. *Phys. Rev. D*, 73:072003, 2006.

[80] B. Abi et al. Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm. *Phys. Rev. Lett.*, 126(14):141801, 2021.

[81] A. Czarnecki and W. J. Marciano. The Muon anomalous magnetic moment: A Harbinger for 'new physics'. *Phys. Rev. D*, 64:013014, 2001.

[82] P. A. R. Ade et al. Planck 2015 results. XIII. Cosmological parameters. *Astron. Astrophys.*, 594:A13, 2016.

[83] P. Athron et al. Global fits of GUT-scale SUSY models with GAMBIT. *Eur. Phys. J. C*, 77(12):824, 2017.

[84] O. Buchmueller et al. The CMSSM and NUHM1 after LHC Run 1. *Eur. Phys. J. C*, 74(6):2922, 2014.

[85] P. Athron et al. A global fit of the MSSM with GAMBIT. *Eur. Phys. J. C*, 77(12):879, 2017.

[86] K. J. de Vries et al. The pMSSM10 after LHC Run 1. *Eur. Phys. J. C*, 75(9):422, 2015.

[87] E. Bagnaschi et al. Likelihood Analysis of the pMSSM11 in Light of LHC 13-TeV Data. *Eur. Phys. J. C*, 78(3):256, 2018.

[88] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock. *LHC Design Report.* CERN Yellow Reports: Monographs. CERN, Geneva, 2004.

[89] O. Sim Brüning, P. Collier, P Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock. *LHC Design Report.* CERN Yellow Reports: Monographs. CERN, Geneva, 2004.

[90] LHC Machine. *JINST*, 3:S08001, 2008.

[91] J. Bauer and T. Muller. Prospects for the Observation of Electroweak Top Quark Production with the CMS Experiment. 03 2021.

[92] F. Marcastel. CERN's Accelerator Complex. La chaîne des accélérateurs du CERN. Oct 2013. General Photo.

[93] CMS Collaboration Luminosity Public Results.

[94] G. Apollinari, O. Brüning, T. Nakamoto, and Lucio Rossi. High Luminosity Large Hadron Collider HL-LHC. *CERN Yellow Rep.*, (5):1–19, 2015.

[95] E. Focardi. Status of the cms detector. *Physics Procedia*, 37:119–127, 12 2012.

[96] G. L. "Bayatian and others". CMS Physics: Technical Design Report Volume 1: Detector Performance and Software. 2006.

[97] S. Chatrchyan et al. The CMS Experiment at the CERN LHC. *JINST*, 3:S08004, 2008.

[98] CMS Technical Design Report for the Pixel Detector Upgrade. 9 2012.

[99] The CMS Phase-1 Pixel Detector Upgrade. 12 2020.

[100] D. Contardo, J. Klute, M.and Mans, L. Silvestris, and J Butler. Technical Proposal for the Phase-II Upgrade of the CMS Detector. Technical Report CERN-LHCC-2015-010. LHCC-P-008. CMS-TDR-15-02, Geneva, Jun 2015.

[101] S. Roy Chowdhury. The Phase-2 Upgrade of the CMS Outer Tracker. Technical Report CMS-CR-2020-052, CERN, Geneva, Feb 2020.

[102] CMS: The electromagnetic calorimeter. Technical design report. 12 1997.

[103] CMS Collaboration. Performance of the cms hadron calorimeter with cosmic ray muons and lhc beam data. *Journal of Instrumentation*, 5(03):T03012–T03012, Mar 2010.

[104] C. Ochando. HGCAL: A High-Granularity Calorimeter for the endcaps of CMS at HL-LHC. *J. Phys.: Conf. Ser.*, 928(1):012025. 4 p, 2017.

[105] CMS Collaboration. Performance of the cms drift-tube chamber local trigger with cosmic rays. *Journal of Instrumentation*, 5(03):T03003–T03003, Mar 2010.

[106] A. M. Sirunyan et al. Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV. *JINST*, 13(06):P06015, 2018.

[107] A. M. Sirunyan et al. Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s} = 13$ TeV. *JINST*, 15(10):P10017, 2020.

[108] M. Baber et al. Development and testing of an upgrade to the CMS level-1 calorimeter trigger. *JINST*, 9(01):C01006, 2014.

[109] M. Jeitler, J. Lingemann, D. Rabady, H. Sakulin, and A. Stahl. Upgrade of the cms global muon trigger. *IEEE Transactions on Nuclear Science*, 62(3):1104–1109, 2015.

[110] J. Fulcher, J. Lingemann, D. Rabady, T. Reis, and H. Sakulin. The new global muon trigger of the cms experiment. *IEEE Transactions on Nuclear Science*, 64(6):1467–1473, 2017.

[111] S. Chatrchyan et al. Performance of CMS Muon Reconstruction in *pp* Collision Events at $\sqrt{s} = 7$ TeV. *JINST*, 7:P10002, 2012.

[112] A. M. Sirunyan et al. Measurement of properties of $\mathrm{B}^0_\mathrm{s} \to \mu^+\mu^-$ decays and search for $\mathrm{B}^0 \to \mu^+\mu^-$ with the CMS experiment. *JHEP*, 04:188, 2020.

[113] The Phase-2 Upgrade of the CMS Level-1 Trigger. Technical Report CERN-LHCC-2020-004. CMS-TDR-021, CERN, Geneva, Apr 2020. Final version.

[114] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands. An introduction to PYTHIA 8.2. *Comput. Phys. Commun.*, 191:159–177, 2015.

[115] S. Agostinelli et al. GEANT4–a simulation toolkit. *Nucl. Instrum. Meth. A*, 506:250–303, 2003.

[116] William J. M., Toshinori M., and J. Michael R. Charged lepton flavor violation experiments. *Annual Review of Nuclear and Particle Science*, 58(1):315–341, 2008.

[117] A. de Gouvea and P. Vogel. Lepton Flavor and Number Conservation, and Physics Beyond the Standard Model. *Prog. Part. Nucl. Phys.*, 71:75–92, 2013.

[118] J. et al. Beringer. Review of particle physics. *Phys. Rev. D*, 86:010001, Jul 2012.

[119] J. Hosaka et al. Solar neutrino measurements in super-Kamiokande-I. *Phys. Rev. D*, 73:112001, 2006.

[120] B. Aharmim et al. Combined Analysis of all Three Phases of Solar Neutrino Data from the Sudbury Neutrino Observatory. *Phys. Rev. C*, 88:025501, 2013.

[121] Y. Ashie et al. A Measurement of atmospheric neutrino oscillation parameters by SUPER-KAMIOKANDE I. *Phys. Rev. D*, 71:112005, 2005.

[122] S. Abe et al. Precision Measurement of Neutrino Oscillation Parameters with KamLAND. *Phys. Rev. Lett.*, 100:221803, 2008.

[123] S. Mihara, J.P. Miller, P. Paradisi, and G. Piredda. Charged lepton flavor–violation experiments. *Annual Review of Nuclear and Particle Science*, 63(1):531–552, 2013.

[124] R. H. Bernstein and P. S. Cooper. Charged Lepton Flavor Violation: An Experimenter's Guide. *Phys. Rept.*, 532:27–64, 2013.

[125] A. M. Baldini et al. Search for the lepton flavour violating decay $\mu^+ \to \mathrm{e}^+\gamma$ with the full dataset of the MEG experiment. *Eur. Phys. J. C*, 76(8):434, 2016.

[126] U. Bellgardt et al. Search for the Decay $\mu^+ \to \mathrm{e}^+e^+e^-$. *Nucl. Phys. B*, 299:1–6, 1988.

[127] W. H. Bertl et al. A Search for muon to electron conversion in muonic gold. *Eur. Phys. J. C*, 47:337–346, 2006.

[128] D. Ambrose et al. New limit on muon and electron lepton number violation from $K_0(L) \to \mu^+e^+$ decay. *Phys. Rev. Lett.*, 81:5734–5737, 1998.

[129] A. Sher et al. An Improved upper limit on the decay $K^+ \to pi^+\mu^+e^-$. *Phys. Rev. D*, 72:012005, 2005.

[130] R. Aaij et al. Search for the lepton-flavour violating decays $\mathrm{B}^0_{(s)} \to e^\pm\mu^\mp$. *JHEP*, 03:078, 2018.

[131] R. Aaij et al. Search for the lepton-flavour violating decay $D^0 \to e^\pm\mu^\mp$. *Phys. Lett. B*, 754:167–175, 2016.

[132] R. Akers et al. A Search for lepton flavor violating Z0 decays. *Z. Phys. C*, 67:555–564, 1995.

[133] P. Abreu et al. Search for lepton flavor number violating Z0 decays. *Z. Phys. C*, 73:243–251, 1997.

[134] G. Aad et al. Search for the lepton flavor violating decay $\mathrm{Z}{\to}\mathrm{e}\mu$ in pp collisions at $\sqrt{s}$ TeV with the ATLAS detector. *Phys. Rev. D*, 90(7):072010, 2014.

[135] K. Hayasaka et al. Search for Lepton Flavor Violating Tau Decays into Three Leptons with 719 Million Produced Tau+Tau- Pairs. *Phys. Lett. B*, 687:139–143, 2010.

[136] A. M. Sirunyan et al. Search for the lepton flavor violating decay $\tau \to 3\mu$ in proton-proton collisions at $\sqrt{s} = 13$ TeV. *JHEP*, 01:163, 2021.

[137] Measurement of inclusive W and Z boson production cross sections in pp collisions at sqrt(s)=13 TeV. 2015.

[138] M. Peruzzi, G. Petrucciani, and A. Rizzi. The NanoAOD event data format in CMS. *J. Phys. Conf. Ser.*, 1525(1):012038, 2020.

[139] A. M. Sirunyan et al. Particle-flow reconstruction and global event description with the CMS detector. *JINST*, 12(10):P10003, 2017.

[140] S. Chatrchyan et al. Description and performance of track and primary-vertex reconstruction with the CMS tracker. *JINST*, 9(10):P10009, 2014.

[141] P. Azzurri. Track Reconstruction Performance in CMS. *Nucl. Phys. B Proc. Suppl.*, 197:275–278, 2009.

[142] E. Chabanat and N. Estre. Deterministic Annealing for Vertex Finding at CMS. 2005.

[143] T. Speer, K. Prokofiev, R. Frühwirth, W. Waltenberger, and P. Vanlaer. Vertex Fitting in the CMS Tracker. Technical Report CMS-NOTE-2006-032, CERN, Geneva, Feb 2006.

[144] S. Chatrchyan et al. The Performance of the CMS Muon Detector in proton-proton Collisions at $\sqrt{s} = 7$ TeV at the LHC. *JINST*, 8:P11002, 2013.

[145] S. Chatrchyan et al. Performance of CMS Muon Reconstruction in pp Collision Events at $\sqrt{s} = 7$ TeV. *JINST*, 7:P10002, 2012.

[146] M. Cacciari, G. P. Salam, and G. Soyez. The anti-k_t jet clustering algorithm. *JHEP*, 04:063, 2008.

[147] M. Cacciari, G. P. Salam, and G. Soyez. FastJet User Manual. *Eur. Phys. J. C*, 72:1896, 2012.

[148] V. Khachatryan et al. Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV. *JINST*, 12(02):P02014, 2017.

[149] A. M. Sirunyan et al. Identification of heavy-flavour jets with the cms detector in pp collisions at 13 tev. *JINST*, 13(05):P05011, 2018.

[150] G. Aad et al. Searches for electroweak production of supersymmetric particles with compressed mass spectra in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector. *Phys. Rev. D*, 101(5):052005, 2020.

[151] T. Junk. Confidence level computation for combining searches with small statistics. *Nucl. Instrum. Meth. A*, 434:435, 1999.

[152] A. L. Read. Presentation of search results: the CL_s technique. *J. Phys. G*, 28:2693, 2002.

[153] G. Cowan, K. Cranmer, E. Gross, and O. Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J. C*, 71:1554, 2011. [Erratum: Eur. Phys. J.C73,2501(2013)].

[154] ATLAS and CMS Collaborations, LHC Higgs Combination Group. Procedure for the lhc higgs boson search combination in summer 2011. Technical Report ATL-PHYS-PUB/2011-11, CMS NOTE 2011/005, 2011.

[155] J. Edsjo and P. Gondolo. Neutralino relic density including coannihilations. *Phys. Rev. D*, 56:1879–1894, 1997.

[156] W. Beenakker, M. Klasen, M. Krämer, T. Plehn, M. Spira, and P. M. Zerwas. Production of charginos, neutralinos, and sleptons at hadron colliders. *Phys. Rev. Lett.*, 83:3780–3783, Nov 1999.

[157] S. Chatrchyan et al. Interpretation of Searches for Supersymmetry with Simplified Models. *Phys. Rev. D*, 88(5):052017, 2013.

[158] R. Gröber, Margarete M. Mühlleitner, E. Popenda, and A. Wlotzka. Light Stop Decays: Implications for LHC Searches. *Eur. Phys. J. C*, 75:420, 2015.

[159] LHC SUSY Cross Section Working Group. NLO-NLL wino-like chargino-neutralino (N2C1) cross sections, 2017.

[160] B. Fuks, M. Klasen, D. R. Lamprea, and M. Rothering. Gaugino production in proton-proton collisions at a center-of-mass energy of 8 TeV. *JHEP*, 10:081, 2012.

[161] B. Fuks, M. Klasen, D. R. Lamprea, and M. Rothering. Precision predictions for electroweak superpartner production at hadron colliders with RESUMMINO. *Eur. Phys. J. C*, 73:2480, 2013.

[162] A. Djouadi, M. M. Muhlleitner, and M. Spira. Decays of supersymmetric particles: The Program SUSY-HIT (SUspect-SdecaY-Hdecay-InTerface). *Acta Phys. Polon. B*, 38:635–644, 2007.

[163] U. De Sanctis, T. Lari, S. Montesano, and C. Troncon. Perspectives for the detection and measurement of supersymmetry in the focus point region of mSUGRA models with the ATLAS detector at LHC. *Eur. Phys. J. C*, 52:743–758, 2007.

[164] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014.

[165] R. Frederix and S. Frixione. Merging meets matching in MC@NLO. *JHEP*, 12:061, 2012.

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

[167] S. Alioli, P. Nason, C. Oleari, and E. Re. NLO single-top production matched with shower in POWHEG: s- and t-channel contributions. *JHEP*, 09:111, 2009. [Erratum: JHEP 02, 011 (2010)].

[168] E. Re. Single-top Wt-channel production matched with parton showers using the POWHEG method. *Eur. Phys. J. C*, 71:1547, 2011.

[169] S. Abdullin, P. Azzi, F. Beaudette, P. Janot, and A. Perrotta. The fast simulation of the CMS detector at LHC. *J. Phys. Conf. Ser.*, 331:032049, 2011.

[170] M. Cacciari and G. P. Salam and G. Soyez. The anti-ktjet clustering algorithm. *Journal of High Energy Physics*, 2008(04):063–063, apr 2008.

[171] JetMET Physics Object Group. Jet Identification for the 13 TeV data Run2016.

[172] JetMET Physics Object Group. Jet Identification for the 13 TeV data Run2017.

[173] JetMET Physics Object Group. Jet Identification for the 13 TeV data Run2018.

[174] Btag Physics Object Group. Recommendation for Using b-tag Objects in Physics Analyses.

[175] JetMET Physics Object Group. Recommended Jet Energy Corrections and Uncertainties For Data and MC.

[176] JetMET Physics Object Group. MET Filter Recommendations for Run II.

[177] H. Baer, A. Mustafayev, and X. Tata. Monojet plus soft dilepton signal from light higgsino pair production at LHC14. *Phys. Rev. D*, 90(11):115007, 2014.

[178] CMS Collaboration. Performance of methods for data-driven background estimation in SUSY searches. Technical Report CMS-PAS-SUS-10-001, CERN, Geneva, 2010.

[179] CMS Collaboration. Measurement of the inelastic proton-proton cross section at $\sqrt{s} = 13$ TeV. *JHEP*, 2018:161, 2018.

[180] CMS Collaboration. CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13$ TeV. Technical Report CMS-PAS-LUM-17-004, CERN, Geneva, 2018.

[181] CMS Collaboration. CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV. Technical Report CMS-PAS-LUM-18-002, CERN, Geneva, 2019.

[182] A. L. Read. Presentation of search results: the CL_s technique. *Journal of Physics G: Nuclear and Particle Physics*, 28(10):2693, 2002.

[183] T. Hastie, R. Tibshirani, and J. Friedman. *The elements of statistical learning: data mining, inference and prediction.* Springer, 2 edition, 2009.

[184] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová. Machine learning and the physical sciences. *Rev. Mod. Phys.*, 91(4):045002, 2019.

[185] D. E. Acosta, A. W. Brinkerhoff, E. L. Busch, A. M. Carnes, I. Furic, S. Gleyzer, K. Kotov, J. Low, A. Madorsky, J. T. Rorie, B. Scurlock, and W. Shi. Boosted Decision Trees in the Level-1 Muon Endcap Trigger at CMS. Technical report, CERN, Geneva, Oct 2017.

[186] R. Di Sipio, M. Faucci Giannelli, S. Ketabchi Haghighat, and S. Palazzo. DijetGAN: A Generative-Adversarial Network Approach for the Simulation of QCD Dijet Events at the LHC. *JHEP*, 08:110, 2019.

[187] Heavy flavor identification at CMS with deep neural networks. Mar 2017.

[188] CMS Phase 1 heavy flavour identification performance and developments. May 2017.

[189] New Developments for Jet Substructure Reconstruction in CMS. Jul 2017.

[190] Measurement of $\mathrm{t\bar{t}H}$ production in the $\mathrm{H} \to \mathrm{b\bar{b}}$ decay channel in $41.5\,\mathrm{fb}^{-1}$ of proton-proton collision data at $\sqrt{s} = 13\,\mathrm{TeV}$. Technical report, CERN, Geneva, 2019.

[191] S. Chatrchyan et al. Search for Top-Squark Pair Production in the Single-Lepton Final State in pp Collisions at $\sqrt{s} = 8$ TeV. *Eur. Phys. J. C*, 73(12):2677, 2013.

[192] P. Baldi, K. Cranmer, T. Faucett, P. Sadowski, and D. Whiteson. Parameterized neural networks for high-energy physics. *Eur. Phys. J. C*, 76(5):235, 2016.

[193] A. Stakia, T. Dorigo, G. Banelli, D. Bortoletto, A. Casa, P. de Castro, C. Delaere, J. Donini, L. Finos, M. Gallinaro, A. Giammanco, A. Held, F. Jiménez Morales, G. Kotkowski, S. P. Liew, F. Maltoni, G. Menardi, I. Papavergou, A. Saggio, B. Scarpa, G. C. Strong, C. Tosciri, J. Varela, P. Vischia, and A. Weiler. Advances in multi-variate analysis methods for new physics searches at the large hadron collider. *Reviews in Physics*, 7:100063, 2021.

[194] Search for new physics in the compressed mass spectra scenario using events with two soft opposite-sign leptons and missing transverse momentum at $\sqrt{s} = 13$ TeV. Technical report, CERN, Geneva, 2016.

[195] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer. MadGraph 5 : Going Beyond. *JHEP*, 06:128, 2011.

[196] P. Nason. A New method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 11:040, 2004.

[197] S. Abdullin, P. Azzi, F. Beaudette, P. Janot, and A. Perrotta. Fast Simulation of the CMS Detector at the LHC. Technical report, CERN, Geneva, Dec 2010.

[198] Studies of Higgs boson production in the four-lepton final state at $\sqrt{s} = 13$ TeV. Technical report, CERN, Geneva, 2016.

[199] Identification of b quark jets at the CMS Experiment in the LHC Run 2. Technical report, CERN, Geneva, 2016.

[200] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[201] G. Van Rossum and F. L. Drake Jr. *Python reference manual*. Centrum voor Wiskunde en Informatica Amsterdam, 1995.

[202] A. Hocker et al. TMVA - Toolkit for Multivariate Data Analysis. 3 2007.

[203] Y. Coadou. Boosted Decision Trees and Applications. *EPJ Web Conf.*, 55:02004, 2013.

[204] J. H. Friedman. Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29:1189–1232, 2000.

[205] CMS Collaboration. CMS luminosity measurements for the 2016 data taking period. Technical Report CMS-PAS-LUM-17-001, CERN, Geneva, 2017.

[206] AMVA4NewPhysics. D2.5 – public report on the performance of algorithms for lhc searches of new physics. https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5c1fb67c7&appId=PPGMS.

Appendix A

Datasets and MC samples

A.1 Datasets and MC samples

2016	2017	2018
Campaign 25 Oct 2019		
-	-	DoubleMuon_Run2018A
DoubleMuon_Run2016B	DoubleMuon_Run2017B	DoubleMuon_Run2018B
DoubleMuon_Run2016C	DoubleMuon_Run2017C	DoubleMuon_Run2018C
DoubleMuon_Run2016D	DoubleMuon_Run2017D	DoubleMuon_Run2018D
DoubleMuon_Run2016E	DoubleMuon_Run2017E	-
DoubleMuon_Run2016F	DoubleMuon_Run2017F	-
DoubleMuon_Run2016G	-	-
DoubleMuon_Run2016H	-	-
-	-	MET_Run2018A
MET_Run2016B	MET_Run2017B	MET_Run2018B
MET_Run2016C	MET_Run2017C	MET_Run2018C
MET_Run2016D	MET_Run2017D	MET_Run2018D
MET_Run2016E	MET_Run2017E	-
MET_Run2016F	MET_Run2017F	-
MET_Run2016G	-	-
MET_Run2016H	-	-

Table A.1: List of all the data sets used for the compressed SUSY search with two or three soft leptons and p_T^{miss}, presented in Chapter 4

Model name	Sample Name
TChiWZ	SMS-TChiWZ_ZToLL_mZMin-0p1_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
N2C1 Higgsino	SMS-N2C1-higgsino_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
N2N1 Higgsino	SMS-N2N1-higgsino_TuneCP2_13TeV-madgraphMLM-pythia8
pMSSM Higgsino	MSSM-higgsino_no1l_2lfilter_TuneCP2_13TeV-madgraphMLM-pythia8
T2bff	SMS-T2tt_dM-10to80_2Lfilter_TuneCP2_13TeV-madgraphMLM-pythia8
T2bW	SMS-T2bW_X05_dM-10to80_2Lfilter_mWMin-0p1_TuneCP2_13TeV-madgraphMLM-pythia8

Table A.2: List of signal models and sample names.

Symbolic sample name	Sample name	cross section [pb]
TTJets_Dilepton	TTJets_DiLept_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	87.32
TTJets_SingleLeptonFromT	TTJets_SingleLeptFromT_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	182.2
TTJets_SingleLeptonFromTbar	TTJets_SingleLeptFromTbar_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	182.2
DYJetsToLL_M1to5_HT70to100	DYJetsToLL_M-1to5_HT-70to100_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	638
DYJetsToLL_M1to5_HT100to200	DYJetsToLL_M-1to5_HT-100to200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	476
DYJetsToLL_M1to5_HT200to400	DYJetsToLL_M-1to5_HT-200to400_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	119.1
DYJetsToLL_M1to5_HT400to600	DYJetsToLL_M-1to5_HT-400to600_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	11.57
DYJetsToLL_M1to5_HT600toInf	DYJetsToLL_M-1to5_HT-600toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	3.458
DYJetsToLL_M5to50_HT100to200	DYJetsToLL_M-5to50_HT-100to200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	224.2
DYJetsToLL_M5to50_HT200to400	DYJetsToLL_M-5to50_HT-200to400_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	37.20
DYJetsToLL_M5to50_HT400to600	DYJetsToLL_M-5to50_HT-400to600_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	3.581
DYJetsToLL_M5to50_HT600toInf	DYJetsToLL_M-5to50_HT-600toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1.124
DYJetsToLL_M50_HT70to100	DYJetsToLL_M-50_HT-70to100_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	209.5
DYJetsToLL_M50_HT100to200	DYJetsToLL_M-50_HT-100to200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	181.3
DYJetsToLL_M50_HT200to400	DYJetsToLL_M-50_HT-200to400_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	50.41
DYJetsToLL_M50_HT400to600	DYJetsToLL_M-50_HT-400to600_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	6.983
DYJetsToLL_M50_HT600to800	DYJetsToLL_M-50_HT-600to800_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1.681
DYJetsToLL_M50_HT800to1200	DYJetsToLL_M-50_HT-800to1200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.7753
DYJetsToLL_M50_HT1200to2500	DYJetsToLL_M-50_HT-1200to2500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.1862
DYJetsToLL_M50_HT2500toInf	DYJetsToLL_M-50_HT-2500toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.004384
WJetsToLNu_HT70to100	WJetsToLNu_HT-70To100_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1595
WJetsToLNu_HT100to200	WJetsToLNu_HT-100To200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1627
WJetsToLNu_HT200to400	WJetsToLNu_HT-200To400_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	435.2
WJetsToLNu_HT400to600	WJetsToLNu_HT-400To600_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	59.18
WJetsToLNu_HT600to800	WJetsToLNu_HT-600To800_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	14.58
WJetsToLNu_HT800to1200	WJetsToLNu_HT-800To1200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	6.656
WJetsToLNu_HT1200to2500	WJetsToLNu_HT-1200To2500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1.608
WJetsToLNu_HT2500toInf	WJetsToLNu_HT-2500ToInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.0388
WZTo3LNu_mllmin01	WZTo3LNu_mllmin01_NNPDF31_TuneCUETP8M1_13TeV_powheg_pythia8	62.17
WZTo2L2Q	WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	5.60
WZTo1L1Nu2Q	WZTo1L1Nu2Q_13TeV_amcatnloFXFX_madspin_pythia8	10.71
ZZTo2L2Q	ZZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	4.04
ZZTo4L_M1toInf	ZZTo4L_M-1toInf_13TeV_powheg_pythia8	13.92
VVTo2L2Nu_M1toInf	VVTo2L2Nu_M-1toInf_TuneCUETP8M1_13TeV_amcatnloFXFX_madspin_pythia8	12.33
WpWpJJ	WpWpJJ_EWK-QCD_TuneCUETP8M1_13TeV-madgraph-pythia8	0.03711
WWDoubleTo2L	WWTo2L2Nu_DoubleScattering_13TeV-pythia8	0.1729
WWToLNuQQ	WWToLNuQQ_13TeV-powheg	43.53
TGJets_lep	TGJets_leptonDecays_13TeV_amcatnlo_madspin_pythia8	1.018
TTGJets	TTGJets_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8	3.76
WGToLNuG_amcatnlo	WGToLNuG_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia	511.2
ZGTo2LG	ZGToLLG_01J_5f_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	131.3
T_tWch_noFullyHad	ST_tW_top_5f_NoFullyHadronicDecays_13TeV-powheg$_TuneCUETP8M1$	19.55
T_sch_lep	ST_s-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8_TuneCUETP8M1	3.6806
T_tch	ST_t-channel_top_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1	136.02
TBar_tch	ST_t-channel_antitop_4f_InclusiveDecays_TuneCP5_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1	80.95
TBar_tWch_noFullyHad	ST_tW_antitop_5f_NoFullyHadronicDecays_13TeV-powheg$_TuneCUETP8M1$	19.55
TTWToLNu	TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	0.2043
TTZToLLNuNu	TTZToLLNuNu_M-10_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.2529
TTZToLLNuNu_m1to10	TZToLLNuNu_M-10_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.2529
TZQToLL	tZq_ll_4f_13TeV-amcatnlo-pythia8	0.0758
tWll	ST_tWll_5f_LO_13TeV-MadGraph-pythia8	0.01123
WWW_ll	WWW_4F_DiLeptonFilter_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.007201
WWZ	WWZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.1651
WZZ	WZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.05565
ZZZ	ZZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.01398

Table A.3: List of all the 2016 MC samples for simulated backgrounds. The first column shows the symbolic names the second column shows the MC sample name and the last column presents the cross section in pb.

Symbolic sample name	Sample name	cross section [pb]
TTJets_Dilepton	TTJets_DiLept_TuneCP5_13TeV-madgraphMLM-pythia8	87.32
TTJets_SingleLeptonFromT	TTJets_SingleLeptFromT_TuneCP5_13TeV-madgraphMLM-pythia8	182.2
TTJets_SingleLeptonFromTbar	TTJets_SingleLeptFromTbar_TuneCP5_13TeV-madgraphMLM-pythia8	182.2
DYJetsToLL_M1to4_HT70to100	DYJetsToLL_M-1to4_HT-70to100_TuneCP5_13TeV-madgraphMLM-pythia8	638.0
DYJetsToLL_M1to4_HT100to200	DYJetsToLL_M-1to4_HT-100to200_TuneCP5_13TeV-madgraphMLM-pythia8	476.0
DYJetsToLL_M1to4_HT200to400	DYJetsToLL_M-1to4_HT-200to400_TuneCP5_13TeV-madgraphMLM-pythia8	85.91
DYJetsToLL_M1to4_HT400to600	DYJetsToLL_M-1to4_HT-400to600_TuneCP5_13TeV-madgraphMLM-pythia8	8.203
DYJetsToLL_M1to4_HT600toInf	DYJetsToLL_M-1to4_HT-600toInf_TuneCP5_13TeV-madgraphMLM-pythia8	2.442
DYJetsToLL_M4to50_HT100to200	DYJetsToLL_M-4to50_HT-100to200_TuneCP5_13TeV-madgraphMLM-pythia8	202.8
DYJetsToLL_M4to50_HT200to400	DYJetsToLL_M-4to50_HT-200to400_TuneCP5_13TeV-madgraphMLM-pythia8	53.70
DYJetsToLL_M4to50_HT400to600	DYJetsToLL_M-4to50_HT-400to600_TuneCP5_13TeV-madgraphMLM-pythia8	5.660
DYJetsToLL_M4to50_HT600toInf	DYJetsToLL_M-4to50_HT-600toInf_TuneCP5_13TeV-madgraphMLM-pythia8	1.852
DYJetsToLL_M50_HT100to200	DYJetsToLL_M-50_HT-100to200_TuneCP5_13TeV-madgraphMLM-pythia8	174.0
DYJetsToLL_M50_HT200to400	DYJetsToLL_M-50_HT-200to400_TuneCP5_13TeV-madgraphMLM-pythia8	53.26
DYJetsToLL_M50_HT400to600	DYJetsToLL_M-50_HT-400to600_TuneCP5_13TeV-madgraphMLM-pythia8	7.582
DYJetsToLL_M50_HT600to800	DYJetsToLL_M-50_HT-600to800_TuneCP5_13TeV-madgraphMLM-pythia8	1.882
DYJetsToLL_M50_HT800to1200	DYJetsToLL_M-50_HT-800to1200_TuneCP5_13TeV-madgraphMLM-pythia8	0.8728
DYJetsToLL_M50_HT1200to2500	DYJetsToLL_M-50_HT-1200to2500_TuneCP5_13TeV-madgraphMLM-pythia8	0.2079
DYJetsToLL_M50_HT2500toInf	DYJetsToLL_M-50_HT-2500toInf_TuneCP5_13TeV-madgraphMLM-pythia8	0.003764
WJetsToLNu_HT100to200	WJetsToLNu_HT-100To200_TuneCP5_13TeV-madgraphMLM-pythia8	1.632
WJetsToLNu_HT200to400	WJetsToLNu_HT-200To400_TuneCP5_13TeV-madgraphMLM-pythia8	477.4
WJetsToLNu_HT400to600	WJetsToLNu_HT-400To600_TuneCP5_13TeV-madgraphMLM-pythia8	67.29
WJetsToLNu_HT600to800	WJetsToLNu_HT-600To800_TuneCP5_13TeV-madgraphMLM-pythia8	14.95
WJetsToLNu_HT800to1200	WJetsToLNu_HT-800To1200_TuneCP5_13TeV-madgraphMLM-pythia8	6.137
WJetsToLNu_HT1200to2500	WJetsToLNu_HT-1200To2500_TuneCP5_13TeV-madgraphMLM-pythia8	1.253
WJetsToLNu_HT2500toInf	WJetsToLNu_HT-2500ToInf_TuneCP5_13TeV-madgraphMLM-pythia8	0.009582
WZTo3LNu_mllmin01	WZTo3LNu_mllmin01_NNPDF31_TuneCP5_13TeV_powheg_pythia8	62.17
WZTo2L2Q	WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	5.60
WZTo1L1Nu2Q	WZTo1L1Nu2Q_13TeV_amcatnloFXFX_madspin_pythia8	10.71
ZZTo2L2Q	ZZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	3.28
ZZTo4L_M1toInf	ZZTo4L_M-1toInf_TuneCP5_13TeV_powheg_pythia8	13.74
VVTo2L2Nu_M1toInf	VVTo2L2Nu_M-1toInf_TuneCP5_13TeV_amcatnloFXFX_madspin_pythia8	14.75
WpWpJJ	WpWpJJ_EWK-QCD_TuneCP5_13TeV-madgraph-pythia8	0.04914
WW_DPS	WW_DoubleScattering_13TeV-pythia8_TuneCP5	1.921
WWToLNuQQ	WWToLNuQQ_NNPDF31_TuneCP5_13TeV-powheg-pythia8	43.53
TGJets_lep	TGJets_leptonDecays_TuneCP5_13TeV-madgraph-pythia8	1.018
TTGJets	TTGJets_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8	4.09
WGToLNuG	WGToLNuG_TuneCP5_13TeV-madgraphMLM-pythia8	466.1
ZGTo2LG	ZGToLLG_01J_5f_TuneCP5_13TeV-amcatnloFXFX-pythia8	55.78
T_tWch_noFullyHad	ST_tW_top_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8	19.55
T_sch_lep	ST_s-channel_4f_leptonDecays_TuneCP5_13TeV-madgraph-pythia8	9.704
T_tch	ST_t-channel_top_4f_InclusiveDecays_TuneCP5$_{1}3TeV - powheg - madspin - pythia8$	136.02
TBar_tch	ST_t-channel_antitop_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8	80.95
TBar_tWch_noFullyHad (Rare)	ST_tW_antitop_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8	19.55
TTZToLLNuNu_amc	TTZToLLNuNu_M-10_TuneCP5_13TeV-amcatnlo-pythia8	0.2529
TZQToLL	tZq_ll_4f_ckm_NLO_TuneCP5_13TeV-madgraph-pythia8	0.07358
tWll	ST_tWll_5f_LO_TuneCP5_PSweights_13TeV-madgraph-pythia8	0.01123
WWW_ll	WWW_4F_DiLeptonFilter_TuneCP5_13TeV-amcatnlo-pythia8	0.007201
WWZ	WWZ_TuneCP5_13TeV-amcatnlo-pythia8	0.1651
WZZ	WZZ_TuneCP5_13TeV-amcatnlo-pythia8	0.05565
ZZZ	ZZZ_TuneCP5_13TeV-amcatnlo-pythia8	0.01398

Table A.4: List of all the 2017 MC samples for simulated backgrounds. The first column shows the symbolic names the second column shows the MC sample name and the last column presents the cross section in pb.

Symbolic sample name	Sample name	cross section [pb]
TTJets_Dilepton	TTJets_DiLept_TuneCP5_13TeV-madgraphMLM-pythia8	87.32
TTJets_SingleLeptonFromT	TTJets_SingleLeptFromT_TuneCP5_13TeV-madgraphMLM-pythia8	182.2
TTJets_SingleLeptonFromTbar	TTJets_SingleLeptFromTbar_TuneCP5_13TeV-madgraphMLM-pythia8	182.2
DYJetsToLL_M1to4_HT70to100	DYJetsToLL_M-1to4_HT-70to100_TuneCP5_13TeV-madgraphMLM-pythia8	624.5
DYJetsToLL_M1to4_HT100to200	DYJetsToLL_M-1to4_HT-100to200_TuneCP5_13TeV-madgraphMLM-pythia8	484.3
DYJetsToLL_M1to4_HT200to400	DYJetsToLL_M-1to4_HT-200to400_TuneCP5_13TeV-madgraphMLM-pythia8	85.78
DYJetsToLL_M1to4_HT400to600	DYJetsToLL_M-1to4_HT-400to600_TuneCP5_13TeV-madgraphMLM-pythia8	8.203
DYJetsToLL_M1to4_HT600toInf	DYJetsToLL_M-1to4_HT-600toInf_TuneCP5_13TeV-madgraphMLM-pythia8	2.465
DYJetsToLL_M4to50_HT70to100	DYJetsToLL_M-4to50_HT-70to100_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	145.4
DYJetsToLL_M4to50_HT100to200	DYJetsToLL_M-4to50_HT-100to200_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	202.8
DYJetsToLL_M4to50_HT200to400	DYJetsToLL_M-4to50_HT-200to400_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	53.70
DYJetsToLL_M4to50_HT400to600	DYJetsToLL_M-4to50_HT-400to600_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	5.660
DYJetsToLL_M4to50_HT600toInf	DYJetsToLL_M-4to50_HT-600toInf_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8	1.852
DYJetsToLL_M50_HT100to200	DYJetsToLL_M-50_HT-100to200_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	174.0
DYJetsToLL_M50_HT200to400	DYJetsToLL_M-50_HT-200to400_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	53.26
DYJetsToLL_M50_HT400to600	DYJetsToLL_M-50_HT-400to600_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	7.582
DYJetsToLL_M50_HT600to800	DYJetsToLL_M-50_HT-600to800_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	1.882
DYJetsToLL_M50_HT800to1200	DYJetsToLL_M-50_HT-800to1200_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	0.8728
DYJetsToLL_M50_HT1200to2500	DYJetsToLL_M-50_HT-1200to2500_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	0.2079
DYJetsToLL_M50_HT2500toInf	DYJetsToLL_M-50_HT-2500toInf_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	0.003764
WJetsToLNu_HT100to200	WJetsToLNu_HT-100To200_TuneCP5_13TeV-madgraphMLM-pythia8	1.632
WJetsToLNu_HT200to400	WJetsToLNu_HT-200To400_TuneCP5_13TeV-madgraphMLM-pythia8	477.4
WJetsToLNu_HT400to600	WJetsToLNu_HT-400To600_TuneCP5_13TeV-madgraphMLM-pythia8	67.29
WJetsToLNu_HT600to800	WJetsToLNu_HT-600To800_TuneCP5_13TeV-madgraphMLM-pythia8	14.95
WJetsToLNu_HT800to1200	WJetsToLNu_HT-800To1200_TuneCP5_13TeV-madgraphMLM-pythia8	6.137
WJetsToLNu_HT1200to2500	WJetsToLNu_HT-1200To2500_TuneCP5_13TeV-madgraphMLM-pythia8	1.253
WJetsToLNu_HT2500toInf	WJetsToLNu_HT-2500ToInf_TuneCP5_13TeV-madgraphMLM-pythia8	0.009582
WZTo3LNu_mllmin01	WZTo3LNu_mllmin01_NNPDF31_TuneCP5_13TeV_powheg_pythia8	62.17
WZTo2L2Q	WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	5.60
ZZTo2L2Q	ZZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	3.28
ZZTo4L_M1toInf	ZZTo4L_M-1toInf_TuneCP5_13TeV_powheg_pythia8	13.74
VVTo2L2Nu_M1toInf	VVTo2L2Nu_M-1toInf_TuneCP5_13TeV_amcatnloFXFX_madspin_pythia8	14.75
WpWpJJ	WpWpJJ_EWK-QCD_TuneCP5_13TeV-madgraph-pythia8	0.04914
WW_DPS	WW_DoubleScattering_13TeV-pythia8_TuneCP5	1.921
WWToLNuQQ	WWToLNuQQ_NNPDF31_TuneCP5_13TeV-powheg-pythia8	43.53
TGJets_lep	TGJets_leptonDecays_TuneCP5_13TeV-madgraph-pythia8	1.018
TTGJets	TTGJets_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8	4.09
WGToLNuG	WGToLNuG_TuneCP5_13TeV-madgraphMLM-pythia8	466.1
ZGTo2LG	ZGToLLG_01J_5f_TuneCP5_13TeV-amcatnloFXFX-pythia8	55.78
T_tWch_noFullyHad	ST_tW_top_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8	19.55
T_sch_lep	ST_s-channel_4f_leptonDecays_TuneCP5_13TeV-madgraph-pythia8	9.704
T_tch	ST_t-channel_top_4f_InclusiveDecays_TuneCP5$_13TeV-powheg-madspin-pythia8$	136.02
TBar_tch	ST_t-channel_antitop_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8	80.95
TBar_tWch_noFullyHad (Rare)	ST_tW_antitop_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8	19.55
TTWToLNu_fxfx	TTWJetsToLNu_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8	0.2043
TTZToLLNuNu_amc	TTZToLLNuNu_M-10_TuneCP5_13TeV-amcatnlo-pythia8	0.2529
TZQToLL	tZq_ll_4f_ckm_NLO_TuneCP5_13TeV-madgraph-pythia8	0.07358
tWll	ST_tWll_5f_LO_TuneCP5_PSweights_13TeV-madgraph-pythia8	0.01123
WWW_ll	WWW_4F_DiLeptonFilter_TuneCP5_13TeV-amcatnlo-pythia8	0.007201
WWZ	WWZ_TuneCP5_13TeV-amcatnlo-pythia8	0.1651
WZZ	WZZ_TuneCP5_13TeV-amcatnlo-pythia8	0.05565
ZZZ	ZZZ_TuneCP5_13TeV-amcatnlo-pythia8	0.01398

Table A.5: List of all the 2018 MC samples for simulated backgrounds. The first column shows the symbolic names the second column shows the MC sample name and the last column presents the cross section in pb.

Appendix B

The data driven and the semi-data driven methods

B.1 The data driven and the semi-data driven methods

This appendix describes the DD and the semi-data driven tight-to-loose methods used in the analysis for the non-prompt background prediction in the SR.

The number of non-prompt events that pass the selection of two tight ID leptons in the SR is given by:

$$N_{PP}^{bkg} = p_1 f_2 N_{10} + f_1 p_2 N_{01} + f_1 f_2 N_{00} \tag{B.1}$$

and the weights to be applied on the N_{PP}, N_{PF}, N_{FP} and N_{FF} are given by Eq. 4.27, 4.28, 4.29 and 4.30 which are described in Sec. 4.7.3.

The first step of the DD method is to define the AR, as described in Sec. 4.7.3. Figure B.1 depicts the 2016 2ℓ-ewk AR, low (left) and ultra high (right) MET bin. The non-prompt background in the plots is taken from simulation. In the ultra high MET bin the MC non-prompt $M(\ell\ell)$ shape is taken from an inclusive MET (MET> 200 GeV) bin normalised to the MC non-prompts template of the ultra high MET bin with the rate factor described in 4.7.3 equation 4.12.

Table B.1 presents the yields of the MC backgrounds in the 2ℓ-ewk AR in the low and ultra high MET bins.

The weights of equations 4.27-4.30 are applied on the SR and the AR data events and the non-prompt background in the SR is estimated. The plots in Fig. B.2 show the low (left) and ultra (right) MET bins of 2016 2ℓ-ewk SR, the non-prompt background is estimated with the data-driven method.

Table B.2 presents the background yields in the low and ultra MET bin of 2ℓ-ewk SR 2016. The non-prompt background yields have been estimated with the DD method.

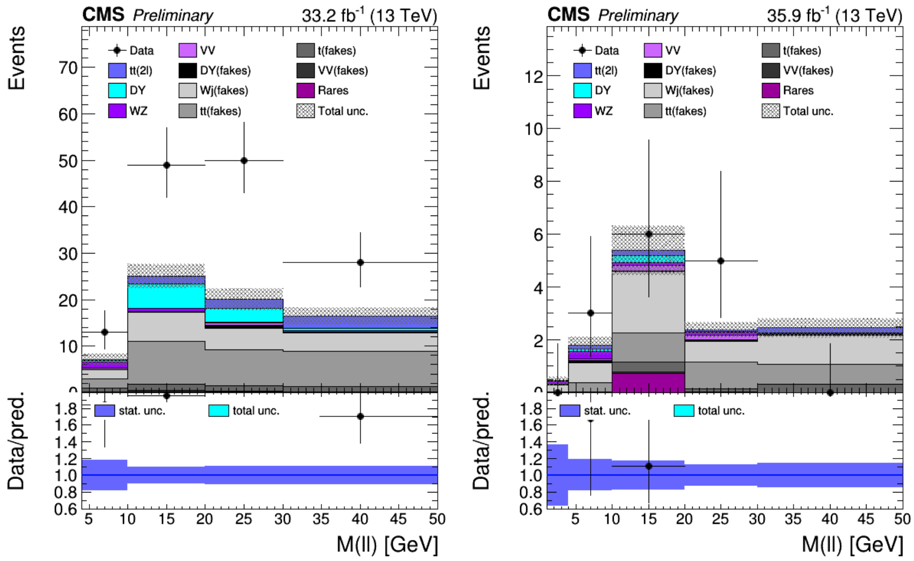

Figure B.1: The $M(\ell\ell)$ distribution in the 2ℓ-ewk AR low MET bin 2016 (left) and in the 2ℓ-ewk AR ultra MET bin 2016 (right). The non-prompt background is taken from simulation, in the ultra high MET bin, the MC non-prompt $M(\ell\ell)$ shape is taken from an inclusive MET (MET$>$ 200 GeV) bin normalised to the MC non-prompt template of the ultra high MET bin.

Background	Yield and stat uncertainty	
	Low MET bin	Ultra MET bin
tt(2l)	6.38 $\pm$ 0.75	0.58 $\pm$ 0.24
DY	9.16 $\pm$ 1.98	0.40 $\pm$ 0.24
WZ	2.41 $\pm$ 0.24	0.71 $\pm$ 0.14
VV	1.12 $\pm$ 0.14	0.50 $\pm$ 0.09
Rares	0.41 $\pm$ 0.26	0.75 $\pm$ 0.73
DY (non prompt)	0.81 $\pm$ 0.60	0.16 $\pm$ 0.07
W (non prompt)	16.80 $\pm$ 2.84	5.24 $\pm$ 0.66
tt (non prompt)	26.43 $\pm$ 1.56	3.19 $\pm$ 0.25
t (non prompt)	4.41 $\pm$ 1.07	0.84 $\pm$ 0.21
VV (non prompt)	0.73 $\pm$ 0.32	0.08 $\pm$ 0.04
Total bkg	68.67 $\pm$ 4.09	12.46$\pm$ 1.11
Data	140	14

Table B.1: Yields of prompt and non-prompt MC backgrounds and the data events in the low and ultra MET bins of the 2016 2ℓ-ewk AR.

From the fake yields per $M(\ell\ell)$ bin shown in Table B.3 it can be seen that the DD method can be used in the low MET ewk SR bin, where the statistics of the data is sufficient. On the contrary, in the ultra high MET bin the DD method

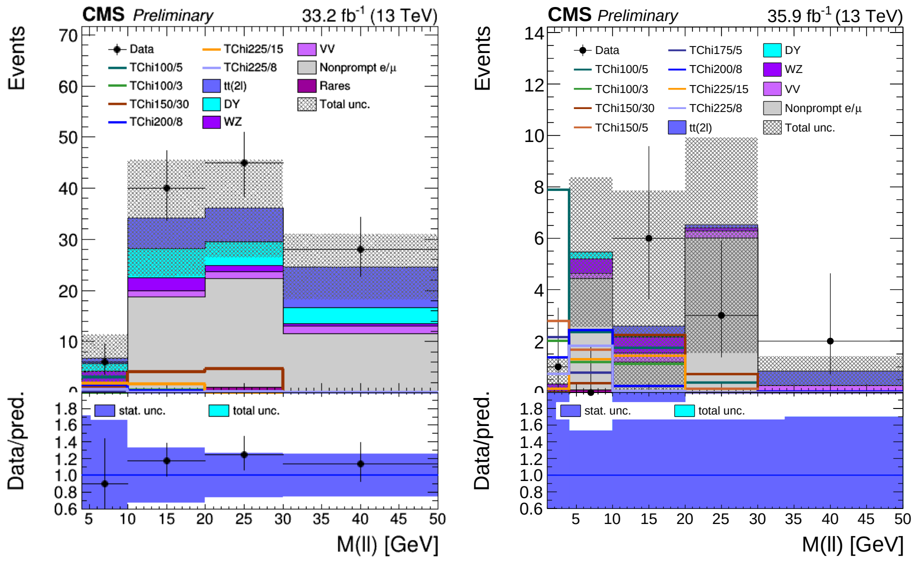

Figure B.2: The $M(\ell\ell)$ distribution of 2ℓ-ewk SR 2016 low (left) and ultra high (right) MET bin. The non-prompt background is estimated with the DD method.

Process	yield and stat uncertainty	
	Low MET bin	Ultra MET bin
tt(2l)	21.40 ± 1.40	1.08 ±0.33
DY	15.03 ± 2.72	0.30 ± 0.25
WZ	7.20 ± 0.42	1.58 ± 0.22
VV	4.76 ± 0.28	1.14 ± 0.14
Rares	1.38 ± 0.54	0.11 ± 0.10
Non-prompt	51.69 ± 16.52	9.81 ± 7.90
Total bkg	101.46 ± 16.82	114.03 ± 7.91
Data	119	12

Table B.2: Yields of prompt MC, data-driven non-prompt process and the data events in the low and ultra MET bins of the 2016 2ℓ-ewk SR.

results in bins with negative non-prompt bin content. It is worth mentioning that in the case of $M(\ell\ell)$ 1-4 GeV, the total (prompt and non-prompt) bin content gets negative, as the bin is populated by prompt WZ with yield 0.28 ± 0.10 and non-prompt background with yield -1.04 ± 1.04.

The solution to this problem is the semi-dd method for the fakes estimation described in Sec. 4.7.3. In the semi-dd method the shape of the non-prompt background is taken from simulation which is normalised to data. Normalising the MC to data means that N_{PP}, N_{PF}, N_{FP} and N_{FF} should be normalised to the respective data. However the N_{PP} MC of the SR cannot be normalised to the

$M(\ell\ell)$ bin [GeV]	data driven fake yield and stat uncertainty	
	Low MET bin	Ultra MET bin
1-4	-	-1.04 $\pm$ 1.04
4-10	0.48 $\pm$4.66	4.33 $\pm$ 2.89
10-20	18.50 $\pm$ 11.21	1.17 $\pm$ 5.27
20-30	21.18 $\pm$ 9.34	6.00 $\pm$ 4.98
30-50	11.54 $\pm$ 6.18	-0.66 $\pm$ 0.53

Table B.3: Yields of the DD non-prompt background per $M(\ell\ell)$ bin in the low and ultra MET bins of the 2016 2ℓ-ewk SR.

SR data events. Therefore only the AR events can be used for the non-prompt background estimation in the semi-dd method. The AR is divided into two independent sidebands with 1 LooseNotTight (LNT) lepton, 2 LooseNotTight leptons or 3 LooseNotTight leptons in case of 3 lepton final state, and the MC templates are normalised to the data of every sideband.

Figure B.3 shows the two sidebands of the 2ℓ-ewk AR ultra high MET bin.

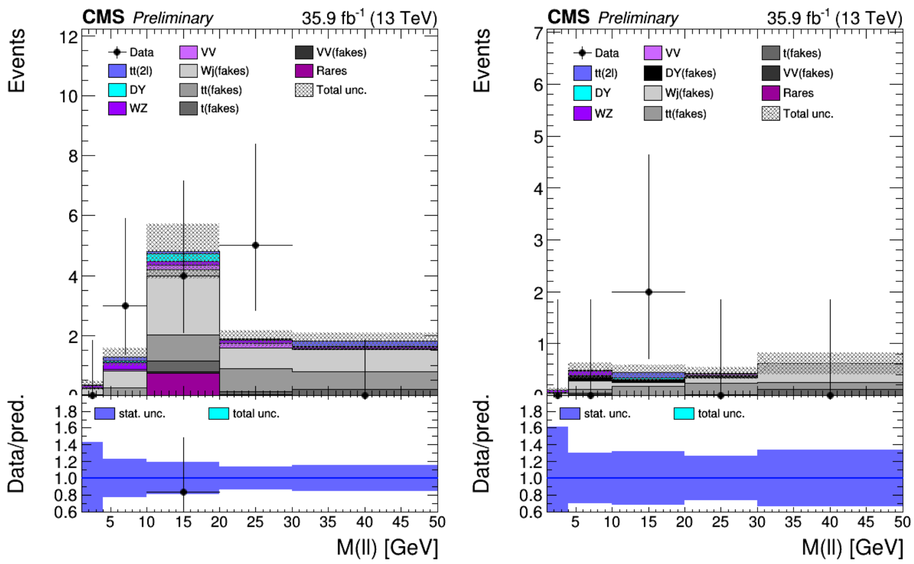

Figure B.3: The $M(\ell\ell)$ distribution in the 2016 2ℓ-ewk AR sidebands of the ultra MET bin. The left plot shows the 1LNT AR sideband and the right plot shows the 2 LNT AR sideband.

Table B.4 presents the background yields in the 2ℓ-ewk AR sidebands of the ultra high MET bin 2016.

Process	yield and stat uncertainty	
	1LNT	2LNT
tt(2l)	0.47 ± 0.22	0.11 ± 0.11
DY+jets	0.36 ± 0.24	0.05 ± 0.03
WZ	0.57 ± 0.13	0.14 ± 0.06
VV	0.42 ± 0.08	0.08 ± 0.03
Rares	0.73 ± 0.73	0.02 ±0.02
DY+jets (non prompt)	0.01 ±0.00	0.17 ± 0.08
W+jets (non prompt)	4.49 ± 0.61	0.73 ± 0.24
tt (non prompt)	2.58 ±0.22	0.62 ± 0.11
t (non prompt)	0.71 ± 0.9	0.14 ± 0.09
VV (non prompt)	0.07 ± 0.04	0.01 ± 0.01
Total bkg	10.38 ± 1.06	2.08 ± 0.32
Data	12	2

Table B.4: The MC yields per process in the 2ℓ-ewk AR sidebands.

The next step of the method is the estimation of the normalization scale factors in order to scale the MC non-promp $M(\ell\ell)$ shape to the data. The normalization scale factors are called semi-dd scale factors and are defined in Sec. 4.7.3, Eq. 4.11. For the cases of the 2ℓ-ewk AR medium/high/ultra high MET bins, where the AR MC non-prompt are taken from an inclusive MET bin, the total weight factor to be applied on the MC non-prompt templates is estimated using the equation 4.13. The total weight factor incorporates the rate factor, for the normalisation of the MET inclusive MC non-prompt templates to the MC non-prompt of every MET bin, and the semi-dd scale factor for the normalisation of the MC non-prompt to the data in every MET bin.

The semi-dd SFs for the MC non-prompt normalisation to data in the 2ℓ-ewk AR ultra MET bin 2016 are presented in Table 4.13. The semi-dd SF, estimated using equation 4.11, are

$$1LNT\ SF = \frac{12 - 2.55}{7.86} = 1.206 \tag{B.2}$$

$$2LNT\ SF = \frac{2 - 0.41}{1.68} = 0.946 \tag{B.3}$$

The plots in Fig. B.4 show the 2ℓ-ewk AR sidebands of the ultra high MET bin 2016, with the MC non-prompt normalised to data. The statistical uncertainty is that of the non-prompt simulation.

The last step of the semi-dd method is the application of the transfer factors on the scaled MC non-prompt background. In the semi-dd method the AR non-prompt simulation is used for the estimation of the non-prompt background in the SR. The weight to be applied on the N_{PF}, N_{FP} and N_{FF} events are given in Eq. 4.32-4.35.

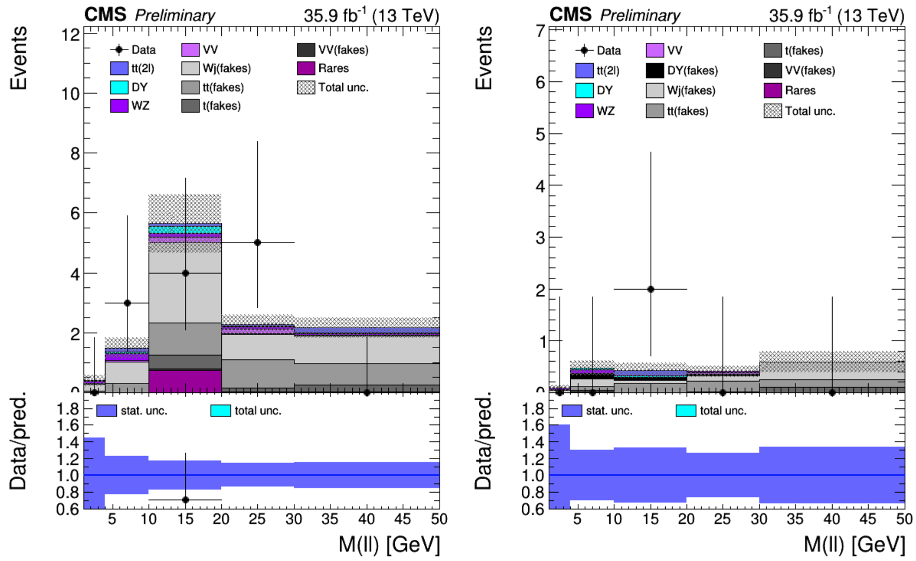

Figure B.4: The $M(\ell\ell)$ distribution in the 2ℓ-ewk AR sidebands of the ultra MET bin 2016. The MC non-prompt background is weighted by the total weight factor.

More specifically, in the 1LNT sideband the non-prompt simulated events that have been scaled with the 1LNT semi-dd SF, are weighed by a factor of $f/(1-f)$, while in the 2LNT side bands the respective events are weighted by a transfer factor of $-f_1 f_2/(1-f1)(1-f2)$.

The result of the application of the transfer factor on the normalised MC non-prompt background is shown in Fig. B.5. The plot shows the $M(\ell\ell)$ distribution in the 2ℓ-ewk SR ultra high MET bin 2016. The non-prompt background has been estimated with the semi-dd method as described above, and the statistical uncertainty is the uncertainty of the non-prompt simulation.

The final number of the non-prompt yields in the 2ℓ-ewk SR as estimated by the semi-dd method is 7.21 ± 0.82.

Comparing Figure B.5 to the left plot of Figure**??** it is noticed that with the semi-dd method the statistical uncertainty of the non-prompt estimation is significantly reduced. Pathological cases where the non-prompt background fluctuated to negative values are now corrected and all the SR bins have positive fake bin content.

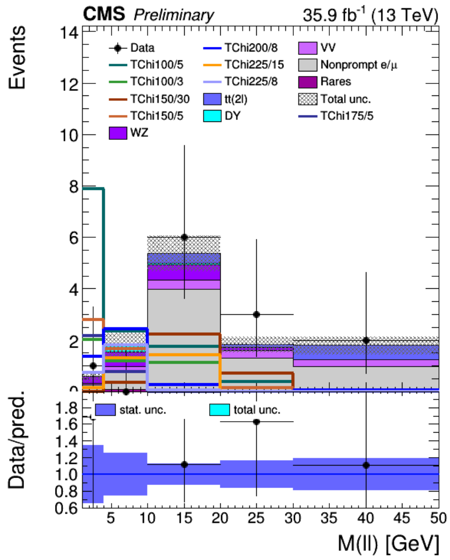

Figure B.5: The $M(\ell\ell)$ distribution in the 2ℓ-ewk SR ultra MET bin 2016. The non-prompt background has been estimated with the semi-dd method.