

DISS. ETH NO. 26007

**SEARCH FOR SUPERSYMMETRY IN OPPOSITE SIGN
SAME FLAVOR LEPTON FINAL STATES WITH THE CMS
DETECTOR**

A THESIS SUBMITTED TO ATTAIN THE DEGREE OF
DOCTOR OF SCIENCES of ETH ZURICH
(Dr. sc. ETH Zürich)

PRESENTED BY
Minna Leonora Vesterbacka Olsson
Master of Science, Lund University
Born on 14 July 1989
Citizen of Sweden

ACCEPTED ON THE RECOMMENDATION OF
Prof. Dr. Rainer Wallny, examiner
Dr. Else Lytken, Senior Lecturer, co-examiner

2019

Abstract

This thesis presents searches for physics beyond the Standard Model (SM) of particle physics. The searches are performed using 35.9 fb^{-1} of proton–proton collision data collected with the CMS detector during 2016 at a center-of-mass energy of $\sqrt{s} = 13\text{ TeV}$, at the CERN Large Hadron Collider (LHC). The new physics phenomena targeted assume strong and electroweak production of Supersymmetric (SUSY) particles in the framework of minimally supersymmetric standard model. A powerful tool for searching for SUSY is by using final state scenarios containing two electrons or two muons of opposite electric charge.

The document contains a summary of the theoretical framework of the SM and SUSY, along with a comprehensive description of the CMS experiment at the LHC accelerator complex. This thesis is based on two analyses, the first targeting strong and electroweak production of superpartners that lead to the leptonic final state along with many jets and large missing transverse momentum. The other analysis is targeting the direct production of the superpartners of the SM electrons and muon, using a final state of only leptons and large missing transverse momentum. Common to these searches is the use of the missing transverse momentum, which is a crucial variable for searches for R -parity conserving SUSY.

The thesis contains a chapter on the performance of the missing transverse momentum reconstruction algorithm. The remainder of the thesis contains an overview of the search strategies and the estimation of the SM background processes specific to each SUSY production mode.

Since no excess of collision data is observed with respect to the predicted SM backgrounds in neither of the searches, a statistical interpretation of the results is performed to set upper limits in the production cross sections on the SUSY particles. These limits greatly extend the limits set using 8 TeV collision data during the LHC Run 1.

Zusammenfassung

In dieser Dissertation werden physikalische Prozesse jenseits des Standardmodells der Teilchenphysik (SM) gesucht. Die Analysen werden mit 35.9 fb^{-1} Proton Proton Kollisionsdaten durchgeführt, die 2016 mit dem CMS Detektor bei einer Schwerpunktssnergie von 13 TeV im LHC am CERN gesammelt wurden. Die neuen Physikphänomene zielen auf starke und elektroschwache Produktion von supersymmetrischen (SUSY) Teilchen im Rahmen des minimal SUSY Standardmodells ab. Leistungsfähige Suchstrategien nach SUSY Teilchen verwenden Endzustandsszenarien mit zwei Elektronen oder zwei Myonen entgegengesetzter elektrischer Ladung.

Dieses Manuskript enthält eine Zusammenfassung der theoretischen Modelle von SM und SUSY sowie eine umfassende Beschreibung des CMS Experiments am Large Hadron Collider (LHC) Beschleunigerkomplex. Die Arbeit basiert auf zwei Analysen: Die erste zielt auf die starke und elektroschwache Produktion von Superpartnern ab, die zu einem leptonischen Endzustand zusammen mit mehreren Jets und großem fehlendem transversalen Impuls führen. Die zweite Analyse untersucht die direkte Produktion der Superpartner von SM-Elektronen und -Myonen, wobei ein rein leptonischer Endzustand mit großem fehlendem transversalem Impuls verwendet wird. Beide Suchen haben die Verwendung des fehlenden transversalen Impulses gemein, der eine entscheidende Variable für die Suche nach *R*-Parity erhaltender SUSY ist.

Diese Dissertation enthält ein Kapitel über die Leistung des Rekonstruktionsalgorithmus des fehlenden transversalen Impulses sowie einen Überblick über die Suchstrategien und die Abschätzung der SUSY Produktionsmodus spezifischen SM Hintergrundprozesse.

Da bei keiner der Suchen ein Überschuss an Kollisionsdaten in Bezug auf die vorhergesagten SM Hintergrundprozesse beobachtet wird, wird eine statistische Interpretation der Ergebnisse durchgeführt, um Obergrenzen für die Produktionsquerschnitte der SUSY Teilchen festzulegen. Diese neuen Grenzwerte erweitern die während des LHC Run 1 unter Verwendung von 8 TeV Kollisionsdaten gemessenen Grenzwerte erheblich.

Acknowledgements

A great number of people deserve recognition for directly or indirectly providing me the support needed to complete this thesis. First and foremost, I am grateful for the mentorship of my advisor at ETH Zürich, Professor Rainer Wallny. I would like to thank Rainer for his support and for all the opportunities he has provided me during the past four years, and for pairing me up with the best post-doc I could ask for, Pablo Martinez Ruiz Del Arbol. Pablo has provided me the guidance and creativity needed to become a great physicist. I consider myself extremely fortunate to have learnt from someone as knowledgeable as Pablo, and his constant support and inspiration has improved many workdays. I owe an immense amount of gratitude to Zeynep Demiragli, for her leadership, her mentorship, and later on, also our friendship. There has been times during my PhD when I lacked motivation and energy. Without the tireless support and encouragement of Lesya Shchutska, Loukas Gouskos and Mariarosaria D'Alfonso the MET paper and Slepton paper would not have been published.

I was lucky to be surrounded by a great group of friends during my time in Geneva. I would especially like to thank Rickard, Sophie, Linnea and Andrew for all the great adventures, R1 breaks and late nights. This place would not have been the same without you, and I will cherish our friendships for life. To Stany, you are a great inspiration and have always been the best support. I am proud to call you my friend.

Finally, I want to express my endless gratitude to my family. To my older brother, Adam, I am thankful for your brotherly advice and your love. I know I always have your support, even though I might not always follow the path you would have wanted. To his wife, Linnéa, I am thankful for our common passion for science and our long friendship. Your energy and motivation has always been an inspiration to me, and I am looking forward to having a third doctor in the family :). To my father, Martin, the first doctor in the family, I am thankful that you early on sparked my interest for science. Even though I did not become the doctor you had intended me to, I know how extremely proud you are of me today. To my mother, Lillemor, I am thankful for your unconditional love and the close relationship we have had for 30 years, and will have for many more. Our weekly phonecalls have been the greatest support I could imagine during these past four years. I dedicate this work to my parents, as their undivided love and encouragement is the reason I can call myself a doctor today.

TABLE OF CONTENTS

	<u>page</u>
1 INTRODUCTION	9
I THEORETICAL FOUNDATIONS	11
2 THE STANDARD MODEL	13
2.1 Particles and interactions	14
2.2 The SM Lagrangian	16
2.3 The success and shortcomings of the SM	24
3 SUPERSYMMETRY	26
3.1 Problems of the SM	27
3.2 Minimal Supersymmetric Standard Model	28
3.3 SUSY breaking mechanisms	32
4 LEPTONIC SUSY SEARCHES	34
4.1 SUSY at the LHC	35
4.2 Simplified models	36
4.3 SUSY with opposite sign same flavor leptons	36
II EXPERIMENTAL SETUP	42
5 THE LARGE HADRON COLLIDER	43
5.1 The accelerator chain	44
5.2 Beam parameters	45
5.3 Coordinate system and kinematic variables	46
6 THE CMS EXPERIMENT	48
6.1 The CMS detector	49
6.2 The Magnet	50
6.3 The Tracker	50
6.4 The ECAL	51
6.5 The HCAL	53
6.6 The muon system	54
7 EVENT RECONSTRUCTION	56

7.1	The trigger system	57
7.2	Data reconstruction	59
7.3	Simulated events	59
8	PHYSICS OBJECTS	65
8.1	The Particle Flow Algorithm	66
8.2	Electrons	68
8.3	Muons	71
8.4	Jets	74
8.5	b -tagged jets	77
8.6	Isolated tracks	79
III	SEARCH METHODOLOGY	80
9	MISSING TRANSVERSE MOMENTUM	81
9.1	Performance of p_T^{miss} reconstruction algorithms	82
9.2	Calibration of p_T^{miss}	84
9.3	Event selection	85
9.4	p_T^{miss} filters	88
9.5	p_T^{miss} performance	90
9.6	Performance of PF p_T^{miss} algorithm	91
9.7	Performance of PUPPI p_T^{miss} algorithm	97
9.8	High pileup studies	101
9.9	Summary	102
10	EVENT SELECTIONS	103
10.1	Datasets	104
10.2	Triggers	104
10.3	Lepton pair selection	106
10.4	Jet and b -jets	107
10.5	p_T^{miss}	108
10.6	M_{T2}	108
10.7	$t\bar{t}$ likelihood discriminator	110
11	BACKGROUND ESTIMATIONS	112
11.1	Standard model processes	113
11.2	Flavor-symmetric background estimation	115
11.3	Drell-Yan background estimation	122
11.4	$Z+\nu$ background estimation	123
11.5	Rare background estimation	124
11.6	Fake lepton background estimation	124
12	STATISTICAL ANALYSIS	125
12.1	Signal models	126
12.2	Maximum likelihood	127
12.3	Hypothesis testing	128

12.4	Visualization of the interpretation	129
13	SYSTEMATIC UNCERTAINTIES	131
13.1	Nuisance parameters	132
13.2	Experimental uncertainties	132
13.3	Theoretical uncertainties	134
13.4	Uncertainties on background predictions	135
IV	ANALYSIS RESULTS	137
14	SEARCH FOR COLORED SUPERPARTNER PRODUCTION	138
14.1	Analysis strategy	139
14.2	Background estimation	140
14.3	Systematic uncertainties	147
14.4	Results	147
14.5	Interpretation	154
14.6	Summary	155
15	SEARCH FOR ELECTROWEAK SUPERPARTNER PRODUCTION	156
15.1	Analysis strategy	157
15.2	Background estimation	158
15.3	Systematic uncertainties	159
15.4	Results	160
15.5	Interpretation	161
15.6	Summary	163
16	SEARCH FOR DIRECT SLEPTON PRODUCTION	164
16.1	Analysis strategy	165
16.2	Background estimation	165
16.3	Systematic uncertainties	171
16.4	Results	173
16.5	Interpretation	174
16.6	Summary	179
17	CONCLUSIONS	180
APPENDIX		
APPENDIX A:	Simulated samples used for SUSY searches	182
APPENDIX B:	Simulated samples for p_T^{miss} performance study	186
APPENDIX C:	Fit functions for kinematic edge search	188
LIST OF TABLES		190

LIST OF FIGURES	194
REFERENCES	201

1 — INTRODUCTION

The Standard Model (SM) of particle physics is an exceptionally successful theory shown to accurately predict the behavior of the particles and forces that make up the most basic constituents of the Universe. It manages to correctly describe all of the known forces except for the gravitational. The SM predicts that massive particles of the theory acquire their mass by interacting with a scalar particle called the Higgs boson. On July 4, 2012 two collaborations at the Large Hadron Collider (LHC), the A Toroidal LHC Apparatus (ATLAS) and Compact Muon Solenoid (CMS), announced the discovery of a new boson of a mass of 125 GeV with properties similar to the SM Higgs boson. This boson had been theorized since the 1960's and its discovery has further grounded the SM as a leading theory that explains particles and interactions.

But in order for the SM to be considered a complete theory, it needs to be able to explain some outstanding features. Among these features, is why the lightest particle, the neutrino, and the heaviest particle, the top quark, span multiple orders of magnitude in mass. Furthermore, a vast portion of the Universe is predicted to consist of an as-of-yet undetected, non-luminous form of matter, known as dark matter. The current formulation of the SM is unable to provide an explanation of the origin of this dark matter. In order to account for these, and other, phenomena possible theoretical extensions of the SM are proposed, of which Supersymmetry (SUSY) is by far the theory that has attracted most popularity. The popularity originates from the theory's ability to solve many of the problems of the SM, by doubling the particle content of the SM.

This work presents a number of searches for SUSY particles, that all have in common that they produce two electrons or two muons of opposite electric charge. The analyses make use of the data from proton-proton (pp) collisions recorded by the CMS detector at the LHC in Geneva, Switzerland during 2016. The collision energy amounts to the unprecedented value of 13 TeV in center-of-mass, and the amount of data, the integrated luminosity, corresponds to 35.9 fb^{-1} .

This thesis is structured in four parts. The first part consists of three chapters that introduces the theoretical framework that compromises the foundation and motivation of this thesis. The SM of particle physics is presented, its successes and shortcomings are discussed, that leads to the motivation of an extension of the SM, SUSY, which is introduced in the subsequent chapter. The first part concludes with a summary of how searches for SUSY can be performed using hadron colliders.

The second part of the thesis contains a description of the experimental setup needed to conduct searches for SUSY. The LHC accelerator complex is introduced, followed by an overview of the layout of the CMS experiment. The means by which the data is collected is presented, along with a description of the algorithms that reconstruct the physics objects.

The third part of the thesis outlines the strategy by which searches are performed. Searches for SUSY, that often contain the production of particles that does not interact with the detector material, rely on a variable

known as missing transverse momentum. This variable is used to measure particles that escape detection, and is reconstructed as the negative vectorial sum of all the visible objects in the event. The first chapter of this part summarizes the excellent performance of the missing transverse momentum reconstruction algorithms used in CMS, which is an essential ingredient in any SUSY search. The rest of the part contains a description of how the searches for SUSY particles is performed, including the object selections, signal region construction and how the known SM processes that are backgrounds for the searches are estimated. Finally, the statistical analysis used to interpret the results is presented.

The final part of the thesis contains the actual searches for SUSY that I conducted as a PhD student at ETH Zürich. Three searches are presented, namely a search for colored superpartners, electroweak superpartners, and a search for direct production of the superpartners of the SM leptons. To date, the results provide the most stringent limits on the electroweak superpartners, and provide exclusion on a number of SUSY particles.

Part I

THEORETICAL FOUNDATIONS

2 — The Standard Model

A well understood underlying theory is essential for interpreting the result of an experiment. There are theories behind all the phenomena surrounding us, that can predict the trajectory of a baseball or the temperature of a cup of coffee after it has been left to cool down. For the physics of everyday life, classical mechanics and thermodynamics are remarkably successful. However, the physics is very different at velocities close to the speed of light, at different *energy scales* and for constituents smaller than atoms. In these regimes, other theories dictate the physical world. Particle physics describes the physics of the smallest known constituents of matter, and the theoretical framework is called the Standard Model (SM) of particle physics. The SM is an extremely successful theory that is able to correctly describe almost all known particles, forces and phenomena that the Universe is composed of. Although it is successful, there are phenomena that the SM is unable to explain, such as the origin of dark matter, the matter-antimatter asymmetry in the Universe, the unknown mechanism behind accelerating cosmic expansion, and the dynamics underlying the common force of gravity.

This chapter provides an overview of all the SM particles and an introduction to the theoretical framework. The convention is to formulate particle physics using a Lagrangian, in the framework of quantum field theory (QFT). From the Lagrangian measurable quantities such as the cross sections and branching ratios can be calculated. The chapter begins with an introduction of the concept of symmetries and their relationship to mathematical groups, that is crucial for writing down the SM Lagrangian. After the SM Lagrangian has been formulated, the chapter concludes with the motivations for an extension of the SM in order to account for the unexplained phenomena.

2.1 Particles and interactions

All known matter is built by fundamental particles, and more precisely of so called *fermions*. The fundamental interactions, or forces, that act between the fermions are mediated by *bosons*. These particles are categorized according to their spin. Fermions have half integer spin, whereas bosons have integer spin, and the various sorts of fermions and bosons will be introduced in the next section.

Particle content of the SM

For most people, the most familiar of the fundamental particles is the electron, e^- . Some people might also be familiar with the so called up and down quarks, u and d . These are the building blocks that make up all the matter surrounding us. The quarks are combined into protons and neutrons that in turn can combine into nuclei, that create atoms when combined with electrons. The electron and the up and down quark are collectively referred to as fermions, as they have the common feature of having half integer spin. The electron is a part of a larger *family* of particles called *leptons*, and the up and down quarks are members of the *quark* family. Both families of fermions come in three different generations, where each generation contains a new set of fermions, which are ordered by their mass. Further, the organization of the SM particles associates an electron, e^- , to its *electron neutrino*, ν_e , into *doublets*, and this pair is the first of the generations of the lepton family. The heavier muon lepton, μ^- , and its muon neutrino, ν_μ , is the second generation, and the heaviest tau lepton τ^- and its tau neutrino ν_τ is the third generation. The up and down quarks, u and d , are the first and lightest generation of quarks, followed by the second generation of quarks containing the heavier charm and strange quarks, c and s . The third generation contains the heaviest quarks, the top and bottom quarks, t and b . The difference between quarks and leptons is that quarks carry strong charge Y_s whereas leptons do not.

There exists a quantum number in the SM, the lepton number L , that needs to be conserved in particle interactions. For each of the fermions there is a corresponding antiparticle, doubling the population of the particle zoo. The antiparticle carries opposite electric charge Q to the particle, and is denoted either by the sign of the charge or by a bar over the particle name, i.e. the antiparticle of the electron e^- is e^+ (the positron), and the antiparticle of the up-quark u is \bar{u} (anti-up). Further, all the fermions come in two so called chirality states, denoted as *right* or *left*-handed, with a subscript of R or L . In the same way that the electric charge is the defining difference between a particle and its antiparticle, so is the so called *weak* charge different between *right* and *left*-handed particles. The third component of the weak isospin is denoted T_3 , and this quantity is used to relate the so called weak hypercharge Y_W to the electric charge Q through the relationship $Y_W = 2(Q - T_3)$. Neutrinos are a special case of this ordering. As they do not carry electric charge, the antiparticle of the neutrino is ill-defined. Further, the right handed chirality state does not carry weak charge. Table 2-1 facilitates the overview of the particles introduced so far.

Fundamental interactions

The other set of particles in the SM, the bosons, carry integer spin. Between the fermions introduced in Table 2-1 act fundamental forces. The forces are carried, or mediated, through gauge bosons which are responsible for forming more complex entities, like atoms, and all the more complicated objects that atoms comprise. There are four types of spin-1 force carriers, the gluons, the photon, the W^+/W^- bosons and the Z boson. The gluons, that in fact come in 8 different versions, mediate the strong force between particles with color charge, i.e. the quarks. The photon mediates the electromagnetic force between particles with electric charge, i.e. the leptons, the quarks and the charged gauge bosons. The W^+/W^- bosons and Z mediate the weak force between particles with isospin or weak hypercharge Y_W , such as the charged and neutral leptons, and the quarks. The gauge bosons and their properties are summarized in Table 2-2. Common to the gauge

Table 2-1. An overview of the fermions of the SM.

	Generation			Charge			
	First	Second	Third	Q	T_3	Y_W	Y_S
Charged leptons	electron	muon	tau				
	$m_e = 0.5 \text{ MeV}$	$m_\mu = 106 \text{ MeV}$	$m_\tau = 1.78 \text{ GeV}$				
	e_L	μ_L	τ_L	-1	$\frac{1}{2}$	-1	no
Neutral leptons	electron ν	muon ν	tau ν				
	$m_{\nu_e} < \text{eV}$	$m_{\nu_\mu} < \text{eV}$	$m_{\nu_\tau} < \text{eV}$				
	ν_e	ν_μ	ν_τ	0	$-\frac{1}{2}$	-1	no
Up-type quarks	up	charm	top				
	$m_u = 2.3 \text{ MeV}$	$m_c = 1.29 \text{ GeV}$	$m_t = 173 \text{ GeV}$				
	u_L	c_L	t_L	$+\frac{2}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	yes
Down-type quarks	down	strange	bottom				
	$m_d = 4.8 \text{ MeV}$	$m_s = 95 \text{ MeV}$	$m_b = 4.8 \text{ GeV}$				
	d_L	s_L	b_L	$-\frac{1}{3}$	$-\frac{1}{2}$	$\frac{1}{3}$	yes
	d_R	s_R	b_R	$-\frac{1}{3}$	0	$-\frac{2}{3}$	yes

Table 2-2. An overview of the gauge bosons of the SM.

Name	Symbol	Interaction	Electromagnetic charge	Mass
Z-boson	Z	Weak	0	91.2 GeV
W-boson	W^+/W^-	Weak	$+/-$	80.4 GeV
Photon	γ	Electromagnetic	0	0
Gluon	g	Strong	0	0

bosons listed in Table 2-2 is that they all are spin-1 bosons. In addition to the spin-1 bosons, the SM predicts a scalar (spin-0) boson, called the Higgs boson. This boson does not carry any force like the spin-1 bosons, but is responsible for giving masses to all the other bosons and the fermions, in a procedure known as the *Higgs mechanism* that will be introduced in Section 2.2.

This section serves as an introduction to the SM particles and their properties. What seems like an ad hoc organization of the masses and the couplings of the particles, in fact relies on a sound theoretical framework. In the next section, the SM Lagrangian is introduced, that will provide more rigor to this brief introduction of the particles of the SM..

2.2 The SM Lagrangian

The mathematical framework used to formulate the SM is quantum field theory (QFT), which treats the previously introduced particles as fields, represented by wave functions. QFT enables calculation of measurable probabilities, such as the probability for two particles to collide and produce another set of particles and the probability for a particle to decay into other particles. These probabilities make the foundation that is used in experimental high energy physics, by enabling for calculation of cross sections and branching fractions. The cross section of a process expresses the probability for two particles to collide and produce other particles. The branching fractions are the probabilities of a particle to decay into different particles, or “decay modes”. The fundamental forces, that are the starting points to calculate these probabilities, are written in terms of Lagrangian densities.

Symmetries and groups

Before the SM Lagrange density can be formulated, the concept of symmetries is crucial to introduce. In classical mechanics, the Lagrangian \mathcal{L}^1 is used to describe the dynamics of the system. For non-relativistic particles, it can simply be defined as:

$$\mathcal{L} = T - V \tag{2-1}$$

where T is the kinetic energy, and V is the potential energy. A transformation of a variable appearing in a Lagrangian \mathcal{L} , is called a symmetry of \mathcal{L} if it leaves the \mathcal{L} unchanged. Examples are the rotational, translational, reflectional and Lorentz transformations. The classical field theory of electricity and magnetism is left unchanged after any or all of those transformations.² Symmetries are referred to as either *global* or *local*. A symmetry is called global if there is no dependence on some spacetime coordinate x , or local if there is. The latter requirement is stronger.

One of the most important contributions to theoretical particle physics, proven by Emmy Noether [1], dictates that there is a conserved physical quantity associated to each local transformation that leaves the Lagrangian invariant. To put this important contribution in context, a Lagrangian symmetric under rotational transformation dictates that the angular momentum of the system is conserved. Similarly, the conservation of linear momentum is related to translational invariance, and energy conservation to time translational invariance. Symmetries are described through the mathematical structure of *groups*. A group is defined as a set of elements and a composition rule. The criteria are summarized as: the combination of two elements, E , should give another element, there exists an identity element I such that $E \cdot I = I \cdot E = E$, there exists an inverse E^{-1} such that $E \cdot E^{-1} = E^{-1} \cdot E = I$ and the composition rule (denoted by \cdot) is associative [2]. Groups enter in to the formulation of particle physics as they describe the property of carrying out a transformation on a physical system that leaves the system unchanged. The most simple example of a group is the one dimensional unitary group, or $U(1)$ group. The set of all complex phases of a wave function $U(\alpha) = e^{-i\alpha}$, with α being a real parameter, form a group. It is straight forward to verify that the criteria

¹ The term Lagrangian refers to the Lagrange density, $\int d^4x \mathcal{L}(x, t)$, but as it is obvious from the context what is meant, I will continue to refer to the Lagrange density as the Lagrangian.

² In reality, the statement is not harder to understand than the statement that a physics experiment should give the same outcome, regardless if we rotated the setup, if we are moving at some constant velocity or perform the experiment at another location.

listed above hold. Consider that some theory be invariant under

$$\Psi(\vec{x}, t) \rightarrow \Psi'(\vec{x}, t) = e^{-i\alpha} \Psi(\vec{x}, t) \quad (2-2)$$

with α being a constant (not depending on some \vec{x} or t)³. This is a global gauge transformation as $\Psi(\vec{x}, t)$ transforms the same everywhere, i.e. there is no dependence on a spacetime coordinate. If one instead considers that the α depends on some \vec{x} , $\alpha(\vec{x}, t)$, the theory needs to be invariant under

$$\Psi(\vec{x}, t) \rightarrow \Psi'(\vec{x}, t) = e^{-i\alpha(\vec{x}, t)} \Psi(\vec{x}, t). \quad (2-3)$$

A Lagrangian is normally defined using some derivatives, as in the case of the momentum operator $\vec{p}/2m = -\nabla/2m$ (in natural units). As is obvious from Equation 2-3, any derivative in the Lagrangian with a dependence on some \vec{x} will result in an additional term involving $\alpha(\vec{x}, t)$ from letting the derivative act on $e^{-i\alpha(\vec{x}, t)}$. This then leads to the invariance being spoiled. In order to recover the invariance, a trick is performed to counteract the introduction of this additional term. The idea is to introduce a field, A_μ that transforms as $A_\mu \rightarrow A'_\mu = A_\mu - \frac{\partial_\mu \alpha(\vec{x}, t)}{e}$, in the so called *covariant derivative*

$$\mathcal{D}_\mu = \partial_\mu - ieA_\mu. \quad (2-4)$$

The covariant derivatives replaces the derivatives in the Lagrangian, and it can easily be shown that the introduction of the field A_μ recovers the invariance. The result can be interpreted in the following way: Local phase invariance of a theory *requires* the existence of a field A_μ . More importantly, one can associate this A_μ field, which is a four-vector, to a vector, i.e. a spin-1 particle. Later in this chapter, we will see that this spin-1 field can be associated, in part, with the photon field. This example summarizes the principle at hand, by requiring local phase invariance of a group, the existence of an interaction field is required. This principle will be used to identify the interaction fields through the more complex group structure.

The local symmetry group of the SM is $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ and in the next section the electroweak forces, described by $SU(2)_L \otimes U(1)_Y$, will be introduced, followed by the strong force, described by the $SU(3)_C$ group. In the end, all the gauge bosons introduced in Table 2-2 will have been identified in this local symmetry group.

The electroweak theory

In the same way that the invariance of the theory under the $U(1)$ transformation required the existence of a field A_μ and an associated spin-1 gauge boson, an internal invariance under a set of transformations that form a so called $SU(2)$ group will require the existence of more spin-1 gauge bosons. The $SU(2)$ group can be called the electroweak $SU(2)$ invariance, and will be shown to have the associated gauge bosons W_μ^i with $i = 1, 2$ or 3 . In order to identify the gauge bosons W_μ^i , we have to write down the $SU(2)_L \otimes U(1)_Y$ Lagrangian. The starting point when writing down the SM Lagrangian is the Dirac Lagrangian,

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi \quad (2-5)$$

³ This might seem like a too rigorous explanation for a thesis on experimental high energy physics, but the following example helps understand the more advanced groups that make up the SM theory.

where the ψ is the fermion field, the γ^μ are the γ -matrices, the $\bar{\psi} = \gamma^0\psi$ is the conjugate fermion field and the m is the fermion mass [2]. The particles with spin- $\frac{1}{2}$, fermions, are put in doublets⁴

$$\Psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}. \quad (2-6)$$

Similarly to what was done in Section 2.2, the local invariance of the theory under the $SU(2)$ group is achieved by the addition of the term $i\bar{\Psi}\gamma_\mu \frac{\sigma_i}{2} W_\mu^i \Psi$. This term describes the interactions between three massive spin-1 fields W_μ^i and two spin- $\frac{1}{2}$ fields, where the W_μ^i transforms as $W_\mu^i \rightarrow W'_\mu^i = W_\mu^i + \partial_\mu a_i(x)$. Ignoring the mass term for a second, the Lagrangian can now be written as

$$\mathcal{L}_{SU(2)} = i\bar{\Psi}\gamma_\mu \partial^\mu \Psi + \bar{\Psi}\gamma_\mu \frac{\sigma_i}{2} (W^\mu)^i \Psi - \frac{1}{4} (W_{\mu\nu})^j (W^{\mu\nu})^j \quad (2-7)$$

where $(W_{\mu\nu})^j = \partial_\mu (W_\nu)^j - \partial_\nu (W_\mu)^j$. Equation 2-7 is thus $SU(2)$ invariant. This locally $SU(2)$ invariant Lagrangian can be unified with the locally $U(1)$ invariant Lagrangian:

$$\mathcal{L}_{U(1)} = -m\bar{\psi}\psi + \bar{\psi}\gamma_\mu (i\partial^\mu + g' \frac{Y}{2} B^\mu) \psi - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}. \quad (2-8)$$

The Y is the same as the subscript of $U(1)$ in $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$, and denotes the weak hypercharge Y_W that was introduced in Table 2-1. The combined $\mathcal{L}_{U(1)}$ and $\mathcal{L}_{SU(2)}$ looks like

$$\mathcal{L}_{SU(2) \otimes U(1)} = \bar{\Psi}\gamma_\mu (i\partial^\mu + g' \frac{Y_W}{2} B^\mu + g \frac{\sigma^i}{2} (W^\mu)^i) \Psi - \frac{1}{4} (W_{\mu\nu})^j (W^{\mu\nu})^j - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \quad (2-9)$$

where g denotes the coupling strengths of the three $(W^\mu)^i$ fields, and g' denotes the coupling strength of the field B^μ . The last two terms can be interpreted as the kinetic energies of the gauge fields B^μ and $(W^\mu)^i$.

At this point, the concept of symmetry breaking has to be introduced. Gauge symmetries can be *broken*, meaning that the ground state wave function breaks the symmetry. The electroweak symmetry breaking converts the four massless bosons $(W^\mu)^i$ ($i = 1, 2, 3$) and B^μ into the three massive bosons W^+ , W^- and Z^0 ⁵ that are the mediators of the weak interaction and a massless photon γ that is the mediator of the electromagnetic interaction, that were all listed in Table 2-2. The W bosons are linear combinations of the $(W^\mu)^1$ and $(W^\mu)^2$ fields, according to:

$$W^\pm = \frac{1}{\sqrt{2}} ((W^\mu)^1 \mp i(W^\mu)^2) \quad (2-10)$$

and the Z^μ and A^μ (the fields for the Z boson and γ) are linear combinations of the B^μ and W_3^μ :

$$\begin{pmatrix} A^\mu \\ Z^\mu \end{pmatrix} = \begin{pmatrix} \cos \theta_W & \sin \theta_W \\ -\sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} B^\mu \\ (W^\mu)^3 \end{pmatrix}. \quad (2-11)$$

⁴ These doublets could have also been denoted as a lepton doublet $l = \begin{pmatrix} e^- \\ \nu_e \end{pmatrix}$, or quark doublet $q = \begin{pmatrix} u \\ d \end{pmatrix}$. Recalling the ordering of the fermions in Table 2-1, one can see that the fermion doublets are in fact ordered according to the generations of the quark and lepton families.

⁵ Here the charge of the Z boson is explicitly stated, but in the rest of the thesis the Z^0 boson will just be denoted by Z .

This is the unification of the electromagnetic and weak forces, and this procedure was introduced by Sheldon L. Glashow, Abdus Salam and Steven Weinberg in the 60's [3–5]. The θ_W in the matrix is called the weak mixing angle, and is determined through the ratio of the coupling constants g and g' according to:

$$\tan \theta_W = \frac{g'}{g}. \quad (2-12)$$

At this point, it is worth pointing out that there are no mass terms for the boson or fermion fields present in Equation 2–9. If they were to be introduced "by hand", by adding mass terms associated to the boson and fermion fields, they would spoil the symmetry. Instead, it will be shown in Section 2.2 that the symmetry of the electroweak interaction needs to be spontaneously broken via the Higgs mechanism in order for the particles to acquire mass. But before the Higgs mechanism can be described, the Quantum Chromodynamics, QCD, the theory of the strong interactions, has to be introduced to complete the Lagrangian.

Quantum chromodynamics

The $SU(2) \otimes U(1)$ Lagrangian, $\mathcal{L}_{SU(2) \otimes U(1)}$ defined in the previous section is written in terms of fermion fields, ψ . It also involves the introduction of the gauge fields B^μ and W_i^μ and explains the electroweak interaction of fermions and bosons. Similarly, the strong interaction, that acts between quarks through the strong *color* charge, can be formulated as a Lagrangian. The QFT explaining the color ("chromo") interactions is called Quantum Chromodynamics [6–8]. Instead of using general fermion fields ψ to formulate the QCD Lagrangian, it suffices to write it only in terms of quark fields. It could be done in terms of general fermion fields, i.e. leptonic and quark fields, but the leptonic fields does not interact via the strong force and will thus render these terms zero. Requiring local gauge invariance under a set of transformations that form an $SU(3)$ group results in the following QCD Lagrangian [9]:

$$\mathcal{L}_{\text{QCD}} = \sum_q \bar{\psi}_{q,a} (i\gamma^\mu \partial_\mu \delta_{a,b} - g_s \gamma^\mu t_{a,b}^C \mathcal{A}_\mu^C) \psi_{q,b} - \frac{1}{4} F_{\mu\nu}^A G^{A\mu\nu}. \quad (2-13)$$

The first terms contain the interactions and kinematics of the quark fields $\psi_{j,a}$ where q indicates the flavor and a runs from 1 to $N_C = 3$ where N_C represents the number of colors. The \mathcal{A}_μ^C represent the gluon fields with the subscript C running from 1 to $N_C^2 - 1$, resulting in 8 different gluons, and the $t_{a,b}^C$ are 3×3 matrices and represent the generators of the $SU(3)$ group. The g_s is the strong coupling constant, and repeated indices are summed over. The last term is the gauge term that contains the field tensor $F_{\mu\nu}^A$ given by:

$$F_{\mu\nu}^A = \partial_\mu \mathcal{A}_\nu^A - \partial_\nu \mathcal{A}_\mu^A - g_s f_{ABC} \mathcal{A}_\mu^B \mathcal{A}_\nu^C, \quad [t^A, t^B] = i f_{ABC} t^C \quad (2-14)$$

The last term in the field tensor, the $g_s f_{ABC} \mathcal{A}_\mu^B \mathcal{A}_\nu^C$ has remarkable consequences for the theory, as it introduces *self-interaction* of the gauge bosons, that will be described in the following section.

Screening, anti-screening, asymptotic freedom and confinement

A result of the non-abelian nature of the QCD leads to the self-interaction of the gluons. This is in contrast to QED that is an abelian theory, and does not have a self-interaction term of the gauge boson. The immediate consequence of the self-interaction of gluons becomes evident when comparing the gauge boson exchange possible in QED and QCD. In QED, a photon exchange diagram can just simply be a photon exchanged between two electrons, or involve an effect by which the vacuum around an electric charge can be polarized by photons splitting into $e^+ e^-$ pairs that then recombine. The so-called coupling strength $\alpha_2(q^2)$, of the

$SU(2)$ theory, is related to the probability of a higher order loop to occur, can be expressed as[2]:

$$\alpha_2(q^2) = \frac{\alpha_2(\mu^2)}{1 + \frac{\alpha_2(\mu^2)}{3\pi} \ln \frac{\mu^2}{(-q^2)}} \quad (2-15)$$

where q^2 is the energy scale, and can be interpreted as a probe that can see more or less close to the origin. As the answer is dependent on the energy scale q^2 , the α_2 must be measured at some particular q^2 . As an example, the choice of q^2 can be denoted μ^2 . The sign in the denominator leads to the following physical effect. For very large q^2 , the $\alpha_2(q^2)$ becomes large. Assuming a negative charge at some origin, like an electron, surrounded by particle anti-particle pairs, emerging from the vacuum and annihilating back. The positively charged particle will be attracted to the negatively charged origin, and the negatively charged particle will be repelled. A small q^2 , corresponding to a probe at some distance, observes a shielding of the original charge due to the additional positive charges surrounding it, meaning that the charge observed from a distance *seems* smaller. On the other hand, for a large q^2 corresponding to a probe that gets closer to the origin, the shielding is less and the actual charge or the origin is more accurate. This effect is known as *screening*, and is introduced in this chapter in order to facilitate the introduction of the concept of *anti-screening*, which is the equivalent, but opposite, effect in QCD.

In QCD, a gluon exchange diagram has an additional contribution with respect to the photon exchange diagram in QED. This additional contribution is due to the self-interaction of the gluons that provides one more diagram involving the production of a gluon loop, in addition to a quark anti-quark loop. This diagram contributes more as the color charge of the gluon is larger than that of a quark. Additionally, the contribution is opposite to that in the QED loop case. This can be understood through the following example. Consider a quark of blue color charge at the origin. A gluon exchange from this quark to another quark (say one of anti-red color charge) would carry the blue charge and the anti-red charge. So instead more blue color charge is produced in the gluon cloud thus enhancing the blue color charge observed by a low energy probe. If one instead approaches the charge with a higher q^2 probe, the screened charge is penetrated and measures a lower net charge. This is the concept of anti-screening [2]. Mathematically, this can be represented by the QCD equivalence of Equation 2-15

$$\alpha_3(q^2) = \frac{\alpha_3(\mu^2)}{1 + \frac{\alpha_3(\mu^2)}{12\pi} (33 - 2n_f) \ln \frac{(-q^2)}{\mu^2}} \quad (2-16)$$

where n_f is the number of flavors, i.e. 6. In the limit of a large q^2 , the $\alpha_3(q^2)$ decreases.⁶ This is what is called *asymptotic freedom* [10, 11] and entails that gluons and quarks appear as almost free constituents within a bound object such as *within* a proton or a neutron.

Another peculiar feature of QCD is *color confinement*. It implies that colored particles can never be free, and are instead always contained within hadrons – colorless bound states of quarks. This can be understood in the following way, that follow the description given in [2]. Instead, if a quark is produced in a collision of two protons, the quark will move away in a direction with respect to the quark to which it was originally bound. Once the distance increases, instead of being released as a free quark, the energy available is used to

⁶ The α_3 , the α_2 and the α_1 (the last one has not been mentioned so far, but is the coupling strength associated to the $U(1)$ symmetry) are all dependent on the energy q^2 and have drastically different values. This will be treated more in depth in Chapter 3, where the introduction of Supersymmetric particles can manage to unify these three coupling constants at a certain energy scale.

produce a new formation of quarks in a colorless hadrons. Eventually, the original quark will have traveled and produced new hadrons on the way, in a process called *hadronization*, that produces showers of hadrons known as *jets*.

The color confinement also results in the fact that bound quark states need to be colorless. Hadrons are divided according to how the quarks are combined to form a bound state. Mesons are composed of two quarks of the same color charge and anti-color charge. The other alternative is a combination of three quarks of different colors, such as red-blue-green (or anti-red-anti-blue-anti-green), and this formation is known as baryons. Protons and neutrons are baryons, that consist of two up quarks and one down quark (proton), or one up quark and two down quarks (neutron). There is an additive quantum number associated to hadrons, the baryon number (B), which is defined as $B = \frac{1}{3}(n_q - n_{\bar{q}})$. For a baryon, consisting of three quarks, the baryon number is 1, whereas it is 0 for a meson⁷, consisting of a quark and an anti-quark. This quantity is conserved in SM processes.

The Higgs Mechanism

The outstanding issue after the formulation of the $SU(2) \otimes U(1)$ Lagrangian in Equation 2–9 is that of how the particles acquire mass. Simply adding mass terms for the fermion and boson fields would break the local gauge invariance. This problem, however, can be solved by the concept of *spontaneous symmetry breaking*. The point of departure for acquiring masses is to introduce a four component complex scalar doublet field Φ ,

$$\Phi = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \quad (2-17)$$

and write the $\mathcal{L}_{SU(2) \otimes U(1)}$ in terms of these doublets, instead of what was previously done, namely write it in terms of doublets of fermion fields, [2]. Equation 2–9 now becomes:

$$\mathcal{L}_{SU(2) \otimes U(1)} = \Phi^\dagger \left(ig' \frac{Y_W}{2} B_\mu + ig \frac{\sigma_i}{2} (W_\mu)^i \right)^\dagger \left(ig' \frac{Y_W}{2} B^\mu + ig \frac{\sigma^i}{2} (W^\mu)^i \right) \Phi + \rho^2 \Phi^\dagger \Phi - \lambda (\Phi^\dagger \Phi)^2. \quad (2-18)$$

The last two terms are known as the Higgs potential $-V(\Phi)$, and can be rewritten as a function of one of the spin-0 doublet fields, $V(\phi) = -\rho^2 |\phi|^2 + \lambda |\phi|^4$. The minimum of $V(\phi)$ can be calculated in the traditional way, namely by requiring $\frac{\partial V(\phi)}{\partial \phi} = 0$. The minimum is obtained for $\phi_{\min} = \sqrt{\frac{\rho^2}{2\lambda}} e^{i\phi}$, meaning that there exists a minimum for every ϕ , i.e. an infinite number of minima which lie on a circle with radius $\sqrt{\frac{\rho^2}{2\lambda}}$. The potential $V(\phi)$ is visualized in Figure 2-1, and has a sombrero shape with the circle of minima at the bottom. Spontaneous symmetry breaking occurs when one of the infinite possibilities of minima is chosen, similarly to how a bead would roll down the slope of the sombrero and settle in one of the positions around the brim of the hat. The spin-0 doublet at the minimum can be rewritten as:

$$\phi_{\min} = \begin{pmatrix} 0 \\ \sqrt{\frac{\rho^2}{2\lambda}} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}. \quad (2-19)$$

Where the v is just a redefinition for brevity. Evaluating the relevant terms in the Lagrangian in Equation 2–18

$$\phi^\dagger \left(ig' \frac{Y_W}{2} B_\mu + ig \frac{\sigma_i}{2} (W_\mu)^i \right)^\dagger \left(ig' \frac{Y_W}{2} B^\mu + ig \frac{\sigma^i}{2} (W^\mu)^i \right) \phi \quad (2-20)$$

⁷ And of course, the baryon number is also zero for leptons.

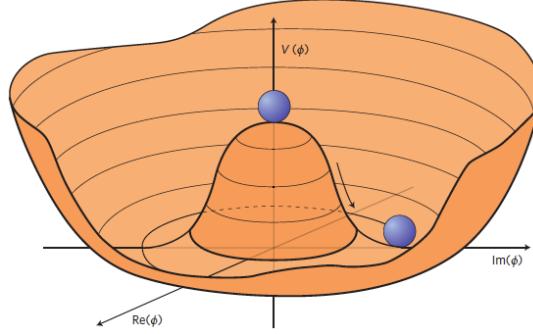


Figure 2-1. A prototypical effective ‘Sombrero’ potential that leads to ‘spontaneous’ symmetry breaking. The vacuum, i.e., the lowest-energy state, is described by a randomly-chosen point around the bottom of the brim of the hat [12].

with $Y_W = 1$ and using the minimum of the potential from Equation 2–19 yields:

$$\frac{1}{8} \left| \begin{pmatrix} g'B_\mu + gW_\mu^3 & g(W_\mu^1 - iW_\mu^2) \\ g(W_\mu^1 + iW_\mu^2) & g'B_\mu - gW_\mu^3 \end{pmatrix} \begin{pmatrix} 0 \\ v \end{pmatrix} \right|^2 \quad (2-21)$$

$$= \frac{1}{8} v^2 g^2 ((W_\mu^1)^2 + (W_\mu^2)^2) + \frac{1}{8} v^2 (g'B_\mu - gW_\mu^3)^2 \quad (2-22)$$

Remembering the identities of Equation 2–10, the first term can be written in terms of W^+ and W^- :

$$\left(\frac{1}{2} vg \right)^2 W_\mu^+ W^\mu - \quad (2-23)$$

The mass term of a charged boson in the Lagrangian is $m^2 W^+ W^-$, hence we can identify the $vg/2$ as the mass of the W boson! Similarly, remembering the identities in Equation 2–11, the mass associated to the Z boson field can be obtained:

$$m_Z = \frac{v}{2} \sqrt{g'^2 + g^2} \quad (2-24)$$

Since no terms of $A_\mu A^\mu$ appear, one can conclude that the mass associated to this photon field must be equal to zero, such that $m_\gamma^2 \cdot A_\mu A^\mu = 0$. The v that was previously introduced for brevity is known as the vacuum expectation value, or vev. To summarize, the gauge theory of massless gauge bosons acquired mass terms through spontaneous symmetry breaking of the Higgs potential. The means by which fermions acquire mass is called Yukawa interactions and will be explained in the subsequent section.

Yukawa interactions

As has been shown in the previous section, the Higgs mechanism is responsible for giving masses to the gauge bosons. In this section, the origin of the fermion masses will be explained. As the Higgs field in an $SU(2)$ doublet is available, it is possible to add an $SU(2)$ invariant interaction between fermions and the Higgs field to the Lagrangian. The explanation will for simplicity be done for the interaction between the first generation of leptons and the Higgs field, but it can be done in the same way for the interaction between all the generations of leptons and quarks and the Higgs field. The term to add to the Lagrangian to explain the interaction of the Higgs field with the first generation leptons is:

$$\mathcal{L}_{\text{int}} = y_f (\bar{L} \phi e_R^- + \phi^\dagger e_R^- l) \quad (2-25)$$

where the y_f is the so called Yukawa coupling for fermions, and $L = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L$. Following the same steps as before, the interaction term of the Lagrangian can be evaluated at some field

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix}. \quad (2-26)$$

with H being the Higgs boson. The \mathcal{L}_{int} then becomes:

$$\mathcal{L}_{\text{int}} = \frac{y_f v}{\sqrt{2}} (\bar{e}_L^- e_R^- + \bar{e}_R^- e_L^-) + \frac{y_f}{\sqrt{2}} (\bar{e}_L^- e_R^- + \bar{e}_R^- e_L^-) H \quad (2-27)$$

The $\frac{y_f v}{\sqrt{2}}$ in the first term has the form of expected for a fermion mass, thus we can identify

$$m_e = \frac{y_f v}{\sqrt{2}} \quad (2-28)$$

Rewriting the interaction Lagrangian with the electron mass, and simplify the $\bar{e}_L^- e_R^- + \bar{e}_R^- e_L^-$ term to $\bar{e}e$:

$$\mathcal{L}_{\text{int}} = m_e \bar{e}e + \frac{m_e}{v} \bar{e}e H \quad (2-29)$$

The second term indicates that there is a possible interaction between electrons and the Higgs field, with a coupling strength of m_e/v . This can be used to calculate the probability for, for example, an electron and a positron to radiate a Higgs boson, or for a Higgs boson to decay into an electron and a positron. A remarkable feature is that the coupling is *proportional* to the fermion mass, meaning that the interaction with the Higgs field is more probable for more massive fermions.

Equation 2–27 gave the interaction term of the Lagrangian in terms of a doublets of the first generation leptons. Of course, it can also be written in terms of quark doublets, or the other lepton generations. The quark doublets can be expressed as up-type and down-type, and have the form:

$$\begin{pmatrix} u \\ d' \end{pmatrix}_L \begin{pmatrix} c \\ s' \end{pmatrix}_L \begin{pmatrix} t \\ b' \end{pmatrix}_L \quad (2-30)$$

where the weak eigenstates are denoted with a prime. In order to go from the weak eigenstates (d', s', b') to their mass eigenstates the Cabibbo-Kobayashi-Maskawa (CKM) matrix is used [13], which is defined as:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \quad (2-31)$$

and the values are listed in [9]

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} 0.974 & 0.225 & 0.004 \\ 0.225 & 0.973 & 0.041 \\ 0.009 & 0.040 & 0.999 \end{pmatrix}. \quad (2-32)$$

The CKM matrix is close to unity, indicating that transitions between the quark generations is heavily suppressed. The CKM matrix can be incorporated in the Lagrangian through the following term that describes the coupling of the quarks to the W boson through a so called *charged* current:

$$\mathcal{L}_{CC} = \frac{g}{\sqrt{2}} W_\mu^+ (\nu_L \gamma^\mu e_L + V_{CKM} \bar{u}_L \gamma^\mu d_L) + \frac{g}{\sqrt{2}} W_\mu^- (\bar{e}_L \gamma^\mu \nu_L + V_{CKM} \bar{d}_L \gamma^\mu u_L). \quad (2-33)$$

The CC subscript stands for charged current. What is evident from Equation 2–33 is that the W^\pm only couple to the left-handed up-type and down-type quarks, and the charged and neutral leptons, indicating that there is a breaking of the parity \mathcal{P} and charge conjugation \mathcal{C} symmetries⁸. Instead, for this electroweak interaction, only the combined symmetry, the \mathcal{CP} , is conserved. To add to the complexity, for electroweak interactions it has been observed that in fact the \mathcal{CP} symmetry is violated, called \mathcal{CP} violation.

2.3 The success and shortcomings of the SM

The SM formulated in the previous sections provides a neat and comprehensible theory for the fundamental particles and interactions. All the fermions and the bosons listed in Tables 2-1 and 2-2 have been predicted and later discovered throughout the later half of the 20th century. Their properties have been measured to a very high precision. The latest discovery was that of the Higgs boson, by the ATLAS [14] and CMS [15] collaborations in 2012. Before then, there had not been a confirmation of the spontaneous breaking of the electroweak symmetry that makes fermions and bosons acquire mass. Its discovery was a great success that solidified the theoretical framework of the SM established in the 1970s. What has not been covered so far in this thesis, is the various shortcomings of the current formulation of the SM. In Table 2-2 the gauge bosons of the SM are summarized, and the interactions they represent. The gravitational interaction is not included in this table as the SM lacks a description of gravity, which renders the theory only able to describe three out of four fundamental interactions.

Further, the claim that the SM is able to explain all known phenomena only holds for the immediate world that surrounds us. Instead, there is overwhelming evidence from cosmological observations that the Universe largely consist of matter that the SM is not able to explain. In fact, only 5% of the energy and matter in the Universe is composed of the SM particles introduced in Tables 2-1 and 2-2. The remaining composition of the Universe is unknown, and is labeled *dark energy* and *dark matter* and make up 68% and 27% respectively [16–18]. The evidence of the existence of non-luminous matter, hence the word “dark”, is provided by data from weak and strong gravitational lensing by large scale structures [19, 20]. There are numerous theories that predict a particle compatible with the features of dark matter, namely massive and decoupled from the electromagnetic, weak and strong interactions. One of these theories is Supersymmetry (SUSY), which will be presented in the next chapter and constitute the motivation for this thesis.

An aesthetical problem of the theory is the need for extreme fine tuning of the SM in order to encompass the known masses of the particles. The problem is that m_H^2 receives large quantum corrections from the virtual effects of every particle that couples to the Higgs field [21]. The corrections are of the form:

$$\delta m_H^2 = -\frac{(y_f)^2}{8\pi^2} [\Lambda^2 + \mathcal{O}(\ln \Lambda)] \quad (2-34)$$

⁸ These symmetries have not been introduced previously as there has been no need for it. Charge conjugation transforms a particle to its anti-particle ($C|e_L^-\rangle = |e_L^+\rangle$), whereas parity transformation changes the handedness ($P|e_L^-\rangle = |e_R^-\rangle$). They are both conserved in electromagnetic and strong interactions.

where the y_f is the Yukawa coupling of the fermions which is proportional to the mass of the fermions, i.e. largest for the top quark. The Λ is a cutoff scale, that takes values of around the Planck scale.⁹ The mass of the Higgs boson is the sum of the tree-level mass and the sum of all loop-corrections, as shown in Equation 3–1, which is highly dependent on Λ . As the Higgs boson mass is at 125.09 ± 0.24 GeV [22], extreme fine tuning is needed in order for it to be stable. This problem is known as the hierarchy problem. As will be showed in the subsequent chapter, the additional particles that SUSY predicts can instead cancel the Λ^2 terms in Equation 3–1, thus stabilizing the Higgs mass without need for fine tuning.

The SM can not explain why there is more matter than anti-matter in the Universe, nor why neutrinos have masses. The list of problems of the SM can go on, but the take-home message of this chapter is that it is still a very successful theory, with immense predictive power. At this point, discarding the SM is not an option. Instead extensions of the SM, known as Beyond SM (BSM) theories that could provide answers to one or all of the above problems, are probed. In the next chapter, the most promising of the BSM theories, SUSY, is introduced, followed by the means by which one can probe the existence of it using high energy physics experiments.

⁹ The Planck-scale is defined through the Planck mass, $m_P = \sqrt{\frac{\hbar}{G}}$. With the gravitational constant G of $6.67 \times 10^{-11} \text{ m}^3 \text{ kg}^{-1} \text{ s}^2$ this corresponds to a value of the Planck-scale of around 10^{19} GeV. This denotes the scale at which descriptions of interaction in the SM fail.

3 — SUPERSYMMETRY

As is clear from the previous chapter, there is a need for a theory that can provide answers to the open questions of the SM. Many attempts to formulate theories beyond the SM have been made, such as theories involving extra hidden dimensions [23], technicolor [24] and grand unified theories with an $SU(5)$ invariance [25]. As providing a description of all the BSM theories on the market is out of the scope of this thesis, emphasis will be put on the most popular extension of the SM, namely *Supersymmetry* (SUSY) [26–28]. SUSY was formulated in the 1970s and is able to tackle many of the problems listed at the end of the previous chapter. Similar to the formulation of the SM, that introduces particles with partners of opposite electric charge or chirality, SUSY adds a final symmetry to the theory. This final symmetry is between fermions and bosons, and dictates that there exists a fermionic SUSY partner to each boson and vice versa.

This chapter contains a review of the known problems of the SM, and how SUSY could solve them, closely following the excellent overview given in [21]. An introduction to the so called Minimal Supersymmetric Standard Model (MSSM) is provided in Section 3.2. The formulation of the MSSM Lagrangian is given, along with a description of the procedure by which SUSY breaks to give masses of the SUSY particles different to that of their SM partners.

3.1 Problems of the SM

As summarized at the end of the last chapter, the SM is far from a complete theory as it is unable to provide explanations of phenomena such as the origin of dark matter and neutrino masses. The fact that it also relies on many parameters and requires extreme fine tuning in order to hold motivates the introduction of a more aesthetically pleasing extension of the model. The hierarchy problem of the SM and the problem with the non-converging coupling constants of the SM are introduced in the following.

The hierarchy problem

The so called hierarchy problem, that was introduced at the end of the last chapter, requires fine tuning in order to be able to stabilize the Higgs boson mass [21]. Recalling the problem, the m_H^2 receives enormous quantum corrections that arise from virtual loops of all the particles that couple to the Higgs field. If the Higgs field couples to a fermion f , corresponding to a term $y_f H \bar{f} f$ in the Lagrangian, and there are no chiral or local gauge symmetries, the correction to m_H^2 will be of the form

$$\delta m_{H,f}^2 = \frac{(y_f)^2}{8\pi^2} \left[-\Lambda^2 + 6m_f^2 \ln \frac{\Lambda}{m_f} \right] \quad (3-1)$$

This correction can be visualized in Figure 3-1. Λ is an arbitrary scale that can be interpreted as the energy scale at which new physics phenomena appear. The y_f , first introduced in Equation 2-25 is the strength at which the fermion couples to the Higgs field. As y_f is proportional to the fermion mass, the top quark will contribute most to Equation 3-1. The quadratic divergences introduced in Equation 3-1 through the scale Λ



Figure 3-1. One-loop quantum corrections to the Higgs squared mass parameter m_H^2 , due to a Dirac fermion f (left), and a scalar S (right).

are not desirable. They imply that there is extreme fine tuning at work that tame the divergences, spanning many orders of magnitude. Another, more aesthetically pleasing, solution to this problem is outlined below. The divergences can be canceled through the introduction of a new set of particles, or more precisely, a new bosonic version of each fermion. Assume the existence of a scalar version of a SM fermion, of mass m_s , that couples to the Higgs field with a Lagrangian term of $-\lambda_s |H|^2 |S|^2$. The corrections to the Higgs mass due to this new boson is of the form

$$\delta m_{H,s}^2 = \frac{\lambda_s}{16\pi^2} \left[\Lambda^2 - 2m_s^2 \ln \frac{\Lambda}{m_s} \right]. \quad (3-2)$$

and visualized in the right of Figure 3-1. As the correction factors to the Higgs mass are added up, and they are of opposite sign, these two terms will now neatly cancel the divergences, if the couplings are related according to $\lambda_s = \lambda_f^2$. The remaining contribution to the Higgs mass is then [29]:

$$\delta m_{H,s+f}^2 \approx \frac{\lambda_f^2}{4\pi^2} (m_s^2 - m_f^2) \ln \frac{\Lambda}{m_s} \quad (3-3)$$

As long as the masses of the new scalar bosons are not too far from that of their SM partners, the corrections to the Higgs mass are well-behaved.

Unification of forces

The coupling constants of the symmetry group of the SM $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$, α_1 , α_2 and α_3 were first introduced in Section 2.2. From their definitions, it is clear that they depend on the energy scale q^2 , and that their values differ from each other. Instead of having three distinct coupling constants, it is theorized that these couplings can instead converge to the same value at some energy scale, thus unifying the forces such that $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_{\text{GUT}}$ where the GUT stands for Grand Unified Theory. If one were to calculate the evolution of the couplings using renormalization group equations with two-loop effects, the coupling constants evolve as shown by dashed lines in Figure 3-2 [21]. Here the α^{-1} of the three couplings indeed meet each other but at different energy scales, thus not resulting in the desirable unification of all three forces.

As has already been hinted in the beginning of this chapter, SUSY introduces a new set of particles, thus modifying the slopes of the inverse of the coupling constants in Figure 3-2. By requiring that the three forces should meet, one can calculate at what energy scale that should be, and also the mass scale of the SUSY particles. This is visualized in Figure 3-2 by solid lines, where the scale at which the forces meet is the so called $m_{\text{GUT}} = 10^{16 \pm 0.3} \text{ GeV}$ yielding a value of $\alpha_{\text{GUT}}^{-1} = 25.7 \pm 1.7$ [30]. This unification of forces that would

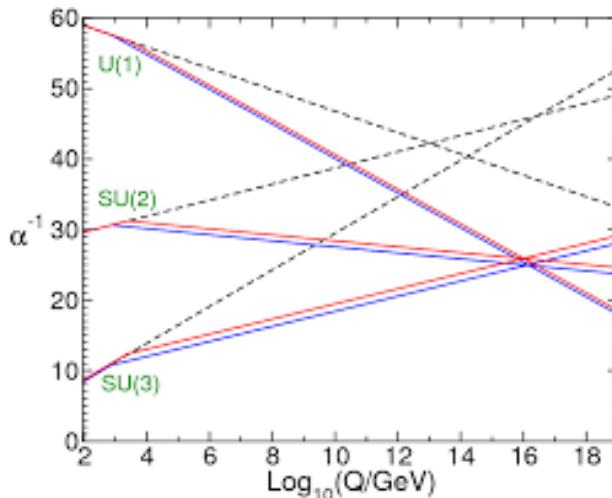


Figure 3-2. Two-loop renormalization group evolution of the inverse gauge couplings α^{-1} in the SM (dashed lines) and the MSSM (solid lines). In the MSSM case, the sparticle masses are treated as a common threshold varied between 750 GeV and 2.5 TeV, and $\alpha_3(m_Z)$ is varied between 0.117 and 0.120. [21]

arise in SUSY is indeed an attractive feature.

3.2 Minimal Supersymmetric Standard Model

In SUSY, each SM particle is associated with a supersymmetric partner, called a superpartner. Just like how the anti-particle of the electron is the positron that carries opposite electric charge, the superpartners of the SM particles carries spin that differ by $\frac{1}{2}$. Formally, this can be described by an operator Q that changes a SM bosonic state $|\text{boson}\rangle$ to a fermionic state $|\text{fermion}\rangle$, and vice versa.

$$Q |\text{fermion}\rangle = Q |\text{boson}\rangle, \quad Q |\text{boson}\rangle = Q |\text{fermion}\rangle \quad (3-4)$$

The operator Q and its hermitian conjugate Q^\dagger are fermionic operators (they carry spin $\frac{1}{2}$), implying that SUSY is a spacetime symmetry rather than an internal or gauge symmetry. Further, the operators need to

satisfy the following commutation and anti-commutation relations

$$\{Q, Q^\dagger\} = P^\mu \quad (3-5)$$

$$\{Q, Q\} = \{Q^\dagger, Q^\dagger\} = 0 \quad (3-6)$$

$$[P^\mu, Q] = [P^\mu, Q^\dagger] = 0 \quad (3-7)$$

with the four-momentum generator of spacetime translations P^μ , which is a bosonic operator [21]. Q and Q^\dagger form a supersymmetry algebra, *superalgebra*, and the fermions and bosons are arranged in the irreducible representations of this algebra. The representations are referred to as *supermultiplets*, and contain fermions and bosons, where the boson is the so called *superpartner* of the fermion in that supermultiplet. The multiplets contain an equal number of fermions and bosons, $n_{\text{fermion}} = n_{\text{boson}}$. Further, particles in the same supermultiplet must have the same gauge interaction, meaning that they would carry the same quantum numbers such as color charge, electric charge and the same value of the weak isospin coupling.

Particle content of the MSSM

The simplest implementation of SUSY is the MSSM, when the number of copies of the operators Q and Q^\dagger is 1.¹ The immediate implication is that the number of particles is doubled, by including a superpartner of each SM particle, with the exact same quantum numbers except the spin. The superpartners of the SM fermions are scalars (spin-0) and are denoted *scalar-fermions* or *sfermions* for short. The superpartners of the SM charged leptons are scalar leptons, *sleptons*, and denoted *selectrons*, *smuons* and *staus* respectively. Similarly, the superpartners of the quarks are called scalar-quarks, *squarks*. In the SM, only left-handed neutrinos exist, hence only superpartners to these left-handed neutrinos, *sneutrinos* are predicted. The superpartners of the SM gauge bosons have a postfix of *-ino* on the SM boson name. The superpartner of the SM gluon is the *glu-ino*. The SM particle symbols are denoted with a tilde in order to specify that they are the superpartner ($g \leftrightarrow \tilde{g}$). The SM fermions, their superpartners of the first generation, the SM gauge bosons and their superpartners are listed in Table 3-1. The defining feature of supersymmetric partners is that all other quantum numbers other than the spin are identical. One striking feature from Table 3-1 is the

Table 3-1. An overview of the superpartners of the SM bosons and fermions [21].

SM particle & SUSY partner	chirality	spin-0	spin- $\frac{1}{2}$	spin-1
leptons & sleptons	L	$(\tilde{\ell}_L, \tilde{\nu}_L)$	(ν_L, ℓ_L)	—
	R	$\tilde{\ell}_R$	ℓ_R	—
quarks & squarks	L	$(\tilde{u}_L, \tilde{d}_L)$	(u_L, d_L)	—
	R	\tilde{u}_R \tilde{d}_R	u_R d_R	—
gauge bosons & gauginos	—	—	\tilde{g}	g
	—	—	\tilde{W}	W
	—	—	\tilde{B}	B
higgs & higgsinos	—	(H_u^+, H_u^0)	$(\tilde{H}_u^+, \tilde{H}_u^0)$	—
	—	(H_d^0, H_d^-)	$(\tilde{H}_d^0, \tilde{H}_d^-)$	—

existence of more than one SM Higgs boson. SUSY *requires* the existence of two Higgs doublets, (H_u^+, H_u^0)

¹ In general, there could be a multiple of copies of the operators Q and Q^\dagger , up to 8 [31, 32].

and (H_d^+, H_d^0) , ordered in $SU(2)_L$ doublet complex scalar fields.² The neutral scalar Higgs boson that we know and have discovered would be a linear combination of the H_u^0 and H_d^0 . To each of these SM Higgs bosons is a superpartner, a *higgsino*, associated to it.

Another thing to notice in Table 3-1 is the SM gauge bosons. Following the discussion in the previous chapter, the usual W , Z and γ bosons are the mass eigenstates of the gauge eigenstates $W_{1,2,3}$ and B . For this reason, the superpartners of the W and B are the ones listed, the so called *wino* and *bino*. Similarly to how the W , Z and γ bosons are mixtures of $W_{1,2,3}$ and B , the winos and binos mix to form the mass eigenstates $\tilde{\chi}_1^\pm, \tilde{\chi}_2^\pm, \tilde{\chi}_1^0, \tilde{\chi}_2^0$ and so on. The ones with electric charge are called *charginos*, and the ones that are neutral, *neutralinos*. Only the mass eigenstates are observable via experiments. The mass eigenstates of the SUSY particles are listed in Table 3-2. For the heaviest of the fermions, the t and b quarks and the τ leptons, the superpartners

Table 3-2. An overview of the gauge eigenstates and the mass eigenstates of the SUSY particles. [21].

	gauge eigenstates	mass eigenstates
squarks	$\tilde{u}_L, \tilde{u}_R, \tilde{d}_L, \tilde{d}_R$	$\tilde{u}_L, \tilde{u}_R, \tilde{d}_L, \tilde{d}_R$
	$\tilde{c}_L, \tilde{c}_R, \tilde{s}_L, \tilde{s}_R$	$\tilde{c}_L, \tilde{c}_R, \tilde{s}_L, \tilde{s}_R$
	$\tilde{t}_L, \tilde{t}_R, \tilde{b}_L, \tilde{b}_R$	$\tilde{t}_1, \tilde{t}_2, \tilde{b}_1, \tilde{b}_2$
sleptons	$\tilde{e}_L, \tilde{e}_R, \tilde{\nu}_e$	$\tilde{e}_L, \tilde{e}_R, \tilde{\nu}_e$
	$\tilde{\mu}_L, \tilde{\mu}_R, \tilde{\nu}_\mu$	$\tilde{\mu}_L, \tilde{\mu}_R, \tilde{\nu}_\mu$
	$\tilde{\tau}_L, \tilde{\tau}_R, \tilde{\nu}_\tau$	$\tilde{\tau}_1, \tilde{\tau}_2, \tilde{\nu}_\tau$
gluino	\tilde{g}	\tilde{g}
gravitino	\tilde{G}	\tilde{G}
neutralinos	$\tilde{B}, \tilde{W}, \tilde{H}_u^0, \tilde{H}_d^0$	$\tilde{\chi}_1^0, \tilde{\chi}_2^0, \tilde{\chi}_3^0, \tilde{\chi}_4^0$
charginos	$\tilde{W}^\pm, \tilde{H}_u^\pm, \tilde{H}_d^\pm$	$\tilde{\chi}_1^\pm, \tilde{\chi}_2^\pm$

are mixed to form mass eigenstates. Sfermion mixing for the first two families is assumed to be negligible. If all other features than the spin is the same for the SM and SUSY particles, the masses would also be the same. But as the superpartners have not been discovered at the same masses as their SM partners, there must be some mechanism included that gives the SUSY particles different masses. As we are already accustomed to the broken symmetry of the electroweak theory, it is clear that SUSY, if it exists, is a broken symmetry.

R-parity

Recalling the baryon and lepton numbers (B and L) that were introduced in the previous chapter. In the SM, the baryon and lepton numbers need to be conserved in an interaction.³ The MSSM can allow for interactions that violate either baryon or lepton number, although no such processes have been observed. Instead, a new quantity is formulated, denoted R [34], defined using the baryon and lepton numbers, as well as the spin s :

$$R = 3(B - L) + 2s. \quad (3-8)$$

² This might seem like a bold prediction, requiring four more Higgs bosons, but it turns out that the lightest Higgs boson could indeed be compatible with the 125 GeV one discovered. [33]

³ Consider a decay of a muon: $\mu^-(L_e = 0, L_\mu = +1) \rightarrow e^-(L_e = +1, L_\mu = 0) + \bar{\nu}_e(L_e = -1, L_\mu = 0) + \nu_\mu(L_e = 0, L_\mu = +1)$, the lepton numbers before and after are the same $L_e = 0, L_\mu = +1$.

From this quantity, the R -parity, P_R , is defined as $P_R = (-1)^R$. For SM particles, P_R takes a value of +1, whereas for SUSY particles it takes a value of -1. The MSSM is defined such that R -parity is conserved in each interaction vertex, and the conservation leads to a couple of experimental implications. The first consequence is that SUSY particles are produced in pairs. The second consequence is that a SUSY particle must decay into another SUSY particle and a SM particle, when it is kinematically allowed. When it is not, i.e. when there are no lighter SUSY particles for a SUSY particle to decay into, this one particle is known as the *lightest supersymmetric particle*, or LSP. This feature of MSSM is what makes it a favorable theory in explaining the origin of dark matter. The LSP is electrically neutral, only interacts weakly with ordinary matter and is stable due to the R -parity conservation; hence it is an attractive candidate for dark matter. The MSSM is a R -parity conserving model (RPC), but there exist models that violate R -parity (RPV), that are lacking the features described above.

The MSSM Lagrangian

Similarly to what was done in the previous chapter, the theory can be written down in terms of a Lagrangian [29]. Neglecting for a second that MSSM must be a broken symmetry, the starting point for writing down the SUSY Lagrangian is:

$$\mathcal{L}_{\text{SUSY}} = \mathcal{L}_{\text{int.}} + \mathcal{L}_{\text{chiral.}} - V(\phi) \quad (3-9)$$

The first term represents the kinetic terms and the gauge interactions, that can be written as:

$$\mathcal{L}_{\text{int.}} = -\mathcal{D}^\mu \phi^{*i} \mathcal{D}_\mu \phi_i - i\psi^{\dagger i} \bar{\sigma}^\mu \mathcal{D}_\mu \psi_i - \frac{1}{4} F_{\mu\nu}^4 F^{\mu\nu a} - i\lambda^{\dagger a} \bar{\sigma}^\mu \mathcal{D}_\mu \lambda^a - \sqrt{2}g[(\phi^{*i} T^a \psi_i) \lambda^a + \lambda^{\dagger a} (\psi^{\dagger i} T^a \phi_i)]. \quad (3-10)$$

The indices i run over all chiralities and flavors, the a is a label for the gauge bosons, and the $\mathcal{D}_\mu \phi_i$, $\mathcal{D}_\mu \psi_i$ and $\mathcal{D}_\mu \lambda^a$ are the covariant derivatives written as:

$$\mathcal{D}_\mu \phi_i = \partial_\mu \phi_i + ig A_\mu^a (T^a \phi)_i, \quad (3-11)$$

$$\mathcal{D}_\mu \psi_i = \partial_\mu \psi_i + ig A_\mu^a (T^a \psi)_i, \quad (3-12)$$

$$\mathcal{D}_\mu \lambda^a = \partial_\mu \lambda^a - gf^{ABC} A_\mu^B A_\nu^C. \quad (3-13)$$

The T^a are the generators of the gauge group. The second line in Equation 3–10 contains the same field strength tensor as introduced in Equation 2–14, that yields the kinetic terms of the gauge fields. The second term in the $\mathcal{L}_{\text{SUSY}}$ is the MSSM Lagrangian for chiral fields $\mathcal{L}_{\text{chiral.}}$ that is written as

$$\mathcal{L}_{\text{chiral.}} = -\frac{1}{2} \frac{\partial^2 W}{\partial \phi_i \partial \phi_j} \psi_i \psi_j + \text{hermitian conjugate.} \quad (3-14)$$

The W in the chiral term is the so called *superpotential* that in MSSM takes the form:

$$W = \epsilon_{ij}(-\bar{L}_L^i h_L \bar{E}_L^C H_d^j - \bar{Q}_L^i h_D \bar{D}_L^C H_d^j + \bar{Q}_L^i h_U \bar{U}_L^C H_d^j + \mu H_u^i H_d^j) \quad (3-15)$$

with i, j being isospin indices, $\epsilon_{ij} = -\epsilon_{ji}$, h_L , h_D , h_U are Yukawa couplings and the μ is a Higgs mixing parameter. R -parity that was introduced in the previous section, is conserved in the superpotential of MSSM. The final term in the $\mathcal{L}_{\text{SUSY}}$ is the MSSM scalar potential, $V(\phi)$. It contains two contributions, the chiral contribution and the gauge contribution, that can be written in the following way:

$$F_i = \frac{\partial W}{\partial \phi_i}, \quad (3-16)$$

$$D^a = -g \phi^{*i} T_{ij}^a \phi_j, \quad (3-17)$$

respectively. The final expression for the potential is:

$$V(\phi) = F_i^* F_i + \frac{1}{2} D^a D^a \quad (3-18)$$

The $\mathcal{L}_{\text{SUSY}}$ will be the starting point for the next section, where the breaking of the SUSY is presented.

3.3 SUSY breaking mechanisms

As the superpartners have not been discovered at the same masses as their SM partners, SUSY is a broken symmetry. In order to accommodate for this required breaking of the symmetry while still maintaining the desirable feature of SUSY that cancels the divergences in Equation 3–1, a so called *soft* SUSY breaking scenario is considered. This is done by adding a new term, a soft SUSY breaking term, to the $\mathcal{L}_{\text{SUSY}}$ [35, 36]:

$$\mathcal{L} = \mathcal{L}_{\text{SUSY}} + \mathcal{L}_{\text{soft}} \quad (3-19)$$

The $\mathcal{L}_{\text{SUSY}}$ is the same as introduced previously, and preserves the supersymmetric interactions. The soft terms include the introduction of mass parameters that lift the mass degeneracy of the SM particles and their superpartners. As these parameters contribute to the radiative corrections discussed previously, they need to be less than a few TeV in order to still solve the hierarchy problem. Following the description in [21], it is evident that SUSY breaking should occur in a hidden sector of particles that have very little direct couplings to the visible sector chiral supermultiplets of the MSSM. The question is what these mediating interactions should be, that is the connection between the hidden sector and the visible sector. There are countless proposals to the design of the mediating interactions in the soft term on the market, of which the two most popular will be introduced.

SUSY breaking in SUGRA

The first proposal for what the mediating interactions are dates back to the beginning of the 1980s [37–41]. It proposes that the mediating force is gravity, hence the name SUGRA, short for SUperGRAvity theory, generalized to many different theories in various numbers of dimensions and involving N supersymmetries. mSUGRA means minimal supergravity theory, and is a special case of SUGRA with $N = 1$. If SUSY is broken in the hidden sector by a vacuum expectation value $\langle F \rangle$ then soft terms in the Lagrangian becomes

$$m_{\text{soft}} \approx \frac{\langle F \rangle}{M_P} \quad (3-20)$$

where M_P is the reduced Planck mass. For m_{soft} of order a few hundred GeV, the expected scale associated with the origin of SUSY breaking in the hidden sector should be around $\sqrt{\langle F \rangle} \approx 10^{10} - 10^{11}$ GeV. In mSUGRA the gravitino is supposed to be heavy (same order as m_{soft} , interacting very weakly, and the LSP is most likely a neutralino). Additionally, the mSUGRA formulation results in a great reduction of the number of parameters.

SUSY breaking in GMSB

Gauge mediated supersymmetry breaking [42–44], or GMSB for short, is a model that assumes the mediating interactions between the hidden sector and the visible sector to be ordinary electroweak and QCD gauge interactions with extra chiral supermultiplets. As opposed to SUGRA where the communication between the hidden sector and the visible sector happen through a mediator, in GMSB the hidden sector couple to a messenger sector instead of directly with the visible sector. In this scenario the soft term arise due to loop diagrams that include a messenger particle. The messengers are chiral supermultiplets that couple to the $\langle F \rangle$ and have the ordinary $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ interactions. The soft term is now

$$m_{\text{soft}} \approx \frac{\alpha_a}{4\pi} \frac{\langle F \rangle}{M_{\text{messenger}}} \quad (3-21)$$

where the $\frac{\alpha_a}{4\pi}$ stems from the ordinary gauge interactions and the $M_{\text{messenger}}$ is the mass of the messenger. The immediate effect of Equation 3–21 is that in order to yield a soft mass of the order of the weak scale, the scale of the SUSY breaking should be as low as $\sqrt{\langle F \rangle} \approx 10^4 - 10^5$ GeV, which is much lower than that obtained in mSUGRA. As an effect of this low energy scale, the gravitino is assumed to be very light (≈ 1 eV) and is thus always the LSP.

What might not obvious from the discussions throughout this chapter is that the unconstrained MSSM contains a huge number of free parameters (105 in addition to the 19 free parameters of the SM!) [45]. In order to be able to use the model in a reasonable way, one needs a way to reduce the number of parameters [46]. At hand are a couple of phenomenological constraints that can be used to bring down the number of parameters, that will be listed in the following. Recalling the \mathcal{CP} violation that was introduced in Section 2.2. There is some evidence for \mathcal{CP} violation, but the parameters of the MSSM can be reduced by assuming that there are no new sources of \mathcal{CP} violation. Further, an assumption is made that there are no so called flavor changing neutral currents, FCNC. Finally, the assumption that the soft SUSY breaking scalar masses are the same for the first and second generations [46]. By imposing these phenomenological constraints, the number of parameters are reduced to only 19. This model is known as phenomenological MSSM, or pMSSM for short.

4 — Leptonic SUSY searches

The shortcomings of the SM has led to the proposal of myriad BSM theories. Using the worlds most powerful hadron collider, the Large Hadron Collider, new particles could be produced that embody the SM extensions. As is clear from the previous chapter, SUSY is a very elegant theory that is able to provide solutions to almost all problems with the SM. The purpose of this thesis is thus to discover, or exclude the existence of, these SUSY particles.

In this chapter, the hypotheses of the MSSM concerning production and decay mechanisms in proton-proton (pp) collisions at the LHC are explored. A powerful tool for searching for particles that would give evidence for new physics is by looking for two leptons of same flavor and opposite charge. This chapter provides an introduction to the many SUSY scenarios that can be explored by using this final state, that enables for searching for the production of colored and electroweak superpartners as well as direct slepton production.

4.1 SUSY at the LHC

SUSY particles have the potential to be produced in pp collisions at the LHC. As detailed in Chapter 3, SUSY particles are produced in pairs. The cross sections for the various production modes can be calculated using the theoretical framework detailed in the previous chapter. For a center-of-mass energy of 13 TeV, which is what is used for the searches documented in this thesis, the cross sections calculated at NLO precision is visualized in Figure 4-1. As is clear from this visualization, the production of colored superpartners is greatly

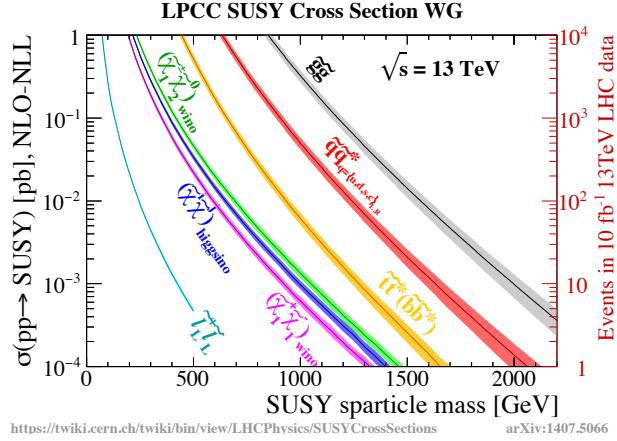


Figure 4-1. Theory cross sections for selected SUSY processes [47].

enhanced compared to the production of electroweak superpartners. In natural SUSY [48–50], the partner of the top quark is expected to be light and can thus reduce the fine tuning of the quadratic divergences to the Higgs boson mass. The breaking of SUSY, as introduced in Section 3.3, leading to mass differences between the SUSY particles and their SM partners, dictates that the MSSM particles should not be too massive. The implication is that the so called natural SUSY should have superpartner masses at the TeV scale while still conserve R-parity, giving an LSP and at least one colored low mass superpartner around 1 TeV. Historically, the proposed superpartner at the TeV scale has been the lightest top squark. After searches for top squarks at the CMS and ATLAS experiment, the top squark mass is excluded up to more than 1 TeV [51]. Instead, the attention is turned to light higgsinos, that provide a promising avenue for discovery of natural SUSY. Light higgsinos have the ability to solve the hierarchy problem as well as avoiding an unnatural fine-tuning of the SUSY particle masses [52, 53]. Searches for higgsinos have been performed prior to the ones at the LHC, at LEP, providing limits on the higgsino masses to around 100 GeV. After the Run I of the LHC at a center-of-mass energy of 8 TeV, [54, 55], the high expectations for a hint of natural supersymmetry were unfortunately not met. Additionally, the discovery of the Higgs boson adds constraints to the implementation of SUSY [56]. But the further increment of the center-of-mass energy during the Run II of the LHC has provided hope to further probe for TeV-scale SUSY.

4.2 Simplified models

The full SUSY models, such as the pMSSM introduced in the previous chapter, are still relatively difficult to calculate despite the fact that the number of parameters are reduced in the constrained theories. For this reason, the SUSY searches at the LHC are performed using so called simplified model spectra (SMS) [57, 58], which is a collection of hypothetical models of SUSY particle production as they may occur in pp collisions at the LHC. Another reason to use SMS, apart from the computational aspect, is to try to be as agnostic as possible on the specific models and instead focus more on the topologies. All the models treated in this thesis assume R-parity conservation, meaning that all models include pair production of SUSY particles, with subsequent decays leading to an LSP at the end of each leg.

In SMS, branching ratios of SUSY particles are often set to 100% while the SM branching ratios are those measured. In addition, different assumptions can be made on the SUSY particle masses. They can be either Majorana¹, or mass-degenerate. Further, if there are more than two SUSY particles produced in the model, its mass is dictated by the masses of the other SUSY particles. This is normally referred to as *mass splitting*, which denotes the difference between the mass of the SUSY particles produced in the hard scatter and the mass of the LSP, $\Delta m = m_{\text{hard scatter}} - m_{\text{LSP}}$. If the Δm is small, these scenarios are referred to as *compressed*, meaning that there is not much available energy for the decay products, where as in the *uncompressed* scenarios, enough energy is available to produce high p_T decay products. This latter scenario can also be referred to as *boosted*, as enough energy is available to give the decay products of, say, a vector boson a boost. Another effect of the uncompressed scenarios is the production of on-shell vector bosons that would have been too massive to be produced given the available energy in a compressed signal scenario. A case in which $m_{\text{LSP}} > m_{\text{hard scatter}}$ is kinematically forbidden.

All these simplifications are done to make the models easier to understand, but this approach has the drawback that it is more complicated to fit in to a theoretical framework. This is done because if no assumptions are made or simplifications are done, the SUSY mass spectrum would simply be too complicated to search for. The simplified models help us see where there might be hints of SUSY, and if there is, a thorough investigation of the full model can be performed in that phase space. The simplified scenarios are categorized according to the interaction of the SUSY particles produced in the hard scatter event. Colored SUSY scenarios involve the production of SUSY particles that interact via the strong force, such as the gluinos or bottom squarks. Electroweak SUSY scenarios involve the production of SUSY particles such as charginos, neutralinos and sleptons.

4.3 SUSY with opposite sign same flavor leptons

Two opposite sign same flavor leptons provide a powerful search tool for strong and electroweak SUSY. This particular final state can appear in many SUSY scenarios, either through the production of an on-shell Z-boson or through the direct or intermediate production of sleptons. The strength of this final state does not only lie in the ability to target many SUSY scenarios, but the SM backgrounds are relatively small and very well understood, which makes them ideal for various data-driven background estimation techniques. To put this into perspective, there are inclusive searches for strong SUSY production in all hadronic final states that are completely swamped by QCD and W+jets processes. On the other side of the spectrum, there are searches targeting electroweak SUSY production in multilepton final states that instead are low in

¹ A majorana fermion is a fermion that is its own antiparticle.

SM backgrounds, but the backgrounds from charge misidentification of leptons and the jets faking leptons pose a major challenge. The searches in this thesis represent the middle ground between these two extreme final state scenarios with the strength of targeting a whole range of SUSY production modes. The strategy deployed in this thesis is to define regions targeting both a high signal acceptance and purity, called signal regions (SRs). Throughout this thesis, this particular leptonic final state is referred to as “opposite sign, same flavor”, “opposite charge, same flavor”, or “ $e^+e^-/\mu^+\mu^-$ ” interchangeably.

This thesis is based on three results. The first result [59] is a general search for SUSY in strong and electroweak production modes. The second result [60] presents a search targeting one particular SUSY production mode, namely direct selectron and smuon production. As all work performed in big collaborations such as the CMS experiment, no paper can be published independently. Instead, the work of many people at different institutes is required, and all authors of the CMS experiment contribute to the collection of the data, validation of reconstruction algorithms, derivation of corrections, developing of analysis methods, to name a few. With this in mind, it is appropriate to mention that the first paper presented in this thesis relies on a collaboration of researchers at ETH Zürich, RWTH Aachen, University of California at San Diego, University of Oviedo and The Institute of Physics of Cantabria. I will highlight my contributions to the first paper, and cite the work of my collaborators when necessary. Since my main contribution to the first paper was in the search for chargino-neutralino production ($\tilde{\chi}_1^\pm-\tilde{\chi}_2^0$) and higgsino production, I will dedicate a large fraction of this thesis to a description of this search, and refer to it in the following as the “Electroweak superpartner search”. A search for colored superpartners, more precisely targeting gluino (\tilde{g}) and sbottom (\tilde{b}) production, will also be presented in this thesis, referred to in the following as “Colored superpartner search”. The second paper that this thesis is based on, is a search for directly produced selectrons (\tilde{e}_L, \tilde{e}_R) and smuons ($\tilde{\mu}_L, \tilde{\mu}_R$), referred to in the following as “Slepton search”. The slepton search has many similarities to the searches for colored and electroweak superpartners, but has some main differences that needs other SM background prediction techniques. What is common to the searches documented in [59] and [60], is the assumption of R -parity conserving models. These models thus predict one or more LSPs in the decay chain, that results in large missing transverse momentum (p_T^{miss}). This variable is very complex and relies on the accurate identification and energy measurement of all particles in the event. In order to have any ability to accurately differentiate large p_T^{miss} from SUSY particles, from other sources of p_T^{miss} such as due to mismeasured jets, from pileup or from neutrinos, highly performant p_T^{miss} reconstruction algorithms are needed. The third paper [61] is a study of the performance of this important discriminating variable. This paper summarizes the performance of the two most common p_T^{miss} reconstruction algorithms during the 2016 running of the LHC. My contributions to this paper are numerous, and the highlights are summarized in Chapter 9.

This chapter will contain a brief overview of the different SUSY production modes in the first section, followed by a description of the SM background processes common for the three searches. The order of the SUSY scenarios presented below is dictated by the cross section of the processes. The search for the colored superpartners is the most established one. This is because it has the highest production cross section, meaning that it has had higher sensitivity compared to the electroweak searches during previous iterations. The search for electroweak superpartners has a lower cross section associated to it, and many of the scenarios probed are for the first time presented in this paper. Lastly, the direct slepton search, which is the SUSY process with the lowest cross section, is presented.

Search for colored superpartners

The search for colored superpartners using opposite sign same flavor leptons can be done by targeting two production modes, that are presented in the following.

Gluino pair production

The first production mode assumes GMSB, as introduced in Section 3.3, a model that assumes strong production of a pair of gluinos (\tilde{g}) that each decays into a pair of quarks (u, d, s, c , or b) and the lightest neutralino, $\tilde{\chi}_1^0$, shown on the left of Figure 4-2. The leptons are a result of the decay of the on-shell Z-boson at the end of the decay chain, and the large p_T^{miss} is due to the gravitino (\tilde{G}).

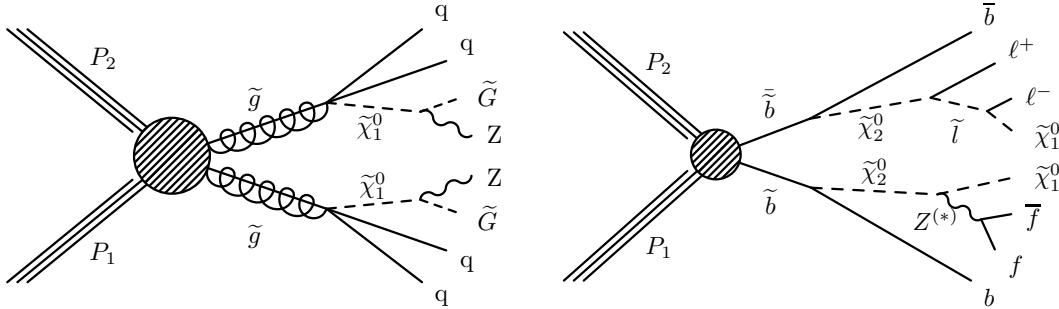


Figure 4-2. Diagrams for strong SUSY production. The gluino GMSB model targeted by the strong on-Z search is shown on the left, that contains one dilepton pair stemming from an on-shell Z boson decay. On the right is a diagram showing a model in which bottom squarks are pair produced with subsequent decays that contain at least one dilepton pair. This model features a characteristic edge shape in the $m_{\ell\ell}$ spectrum given approximately by the mass difference of the $\tilde{\chi}_2^0$ and $\tilde{\chi}_1^0$.

Sbottom pair production

The model involving the production of the superpartner of the bottom quark can also be targeted using opposite sign same flavor leptons. In this model, the \tilde{b} quarks decay to a bottom quark and $\tilde{\chi}_2^0$. Two assumptions are made for the decay of the $\tilde{\chi}_2^0$. In one case (upper half of the Feynman diagram in the right of Figure 4-2), the $\tilde{\chi}_2^0$ decay via a two-body decay via a real ℓ and is allowed if at least one ℓ is lighter than the mass difference of the neutralinos. This scenario would result in an endpoint expressed through

$$(m_{\ell\ell}^{\max})^2 = \frac{(m_{\tilde{\chi}_2^0}^2 - m_\ell^2)(m_\ell^2 - m_{\tilde{\chi}_1^0}^2)}{m_\ell^2} \quad (4-1)$$

where $m_{\tilde{\ell}}$ is the mass of the intermediate slepton, that in this work is considered to be either a \tilde{e} or a $\tilde{\mu}$. This results in the particular feature that the invariant mass of the two leptons has a triangular, “edge”, shape with an endpoint dictated by the $m_{\ell\ell}^{\max}$ in Equation 4-1. The search for this model will from now on be called the “edge”, and the model is referred to as the sbottom or slepton edge model. A potential excess in data compatible with this model would not only be a nice hint of SUSY, but the endpoint in the invariant mass spectrum would in fact give information on the mass hierarchy of the SUSY particles. For this reason, a fit is performed to determine the position of the edge, which will be discussed further when the analysis strategy for the colored SUSY searches is presented in Chapter 14. Further, the \tilde{b} and $\tilde{\chi}_2^0$ masses are free parameters, the mass of the $\tilde{\chi}_1^0$ is set to 100 GeV and the sleptons are assumed to be degenerate with the mass being average of the $\tilde{\chi}_2^0$ and $\tilde{\chi}_1^0$ masses.

The other decay mode of the $\tilde{\chi}_2^0$ considered is visualized in the lower part of the Feynman diagram on the right of Figure 4-2. In this model, the $\tilde{\chi}_2^0$ decays to a Z boson and the $\tilde{\chi}_1^0$ LSP. The Z boson can be on-shell or off-shell depending on the mass difference between the neutralinos, and can decay to any fermion pair allowed by the SM but in this work only the decay to an electron or muon pair is considered. The branching fractions of the two $\tilde{\chi}_2^0$ decay modes is considered to be 50% each, and the model is interpreted in context of the exclusion of \tilde{b} and $\tilde{\chi}_2^0$.

Search for electroweak superpartners

The SUSY models that assume electroweak superpartner production are presented in the following.

Chargino-Neutralino production

The $\tilde{\chi}_1^\pm$ - $\tilde{\chi}_2^0$ production is depicted in Figure 4-3. In this model, the $\tilde{\chi}_1^\pm$ is set to decay to a W boson and a $\tilde{\chi}_1^0$, which is the LSP, while the next-to-lightest neutralino, $\tilde{\chi}_2^0$, decays to a Z boson and $\tilde{\chi}_1^0$. The production cross sections for this model are computed in a limit of mass-degenerate wino $\tilde{\chi}_1^\pm$ and $\tilde{\chi}_2^0$, and light bino $\tilde{\chi}_1^0$. All the other SUSY particles are assumed to be heavy and decoupled.

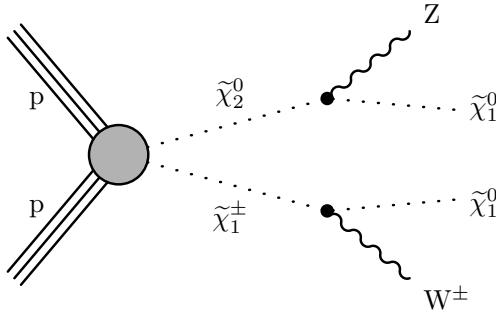


Figure 4-3. Diagram corresponding to the chargino-neutralino production with the $\tilde{\chi}_1^\pm$ and $\tilde{\chi}_2^0$ decaying into vector bosons (W and Z) and the LSP.

Higgsino production

The remaining two models considered in the first paper assume the production of $\tilde{\chi}_1^0$ - $\tilde{\chi}_1^0$ pairs in GMSB. For bino- or wino-like neutralinos, the neutralino pair production cross section is very small, and thus a specific GMSB model is considered, with four mass-degenerate higgsinos ($\tilde{\chi}_1^\pm$, $\tilde{\chi}_2^0$, and $\tilde{\chi}_1^0$) and a massless gravitino as the LSP [62–64]. In the production of any two of these higgsinos, $\tilde{\chi}_1^\pm$ or $\tilde{\chi}_2^0$ decays immediately to $\tilde{\chi}_1^0$ and low-momentum particles that do not impact the analysis, effectively yielding pair production of $\tilde{\chi}_1^0\tilde{\chi}_1^0$. The sum of the various higgsino production possibilities ($\tilde{\chi}_1^\pm\tilde{\chi}_2^0$, $\tilde{\chi}_1^0\tilde{\chi}_2^0$, $\tilde{\chi}_1^\pm\tilde{\chi}_1^\pm$, $\tilde{\chi}_1^\pm\tilde{\chi}_1^0$) describes the effective $\tilde{\chi}_1^0$ production mechanism. Hence, none of the higgsino production possibilities are explicitly shown in Figure 4-4, only the effective $\tilde{\chi}_1^0$ production is displayed. In the first model (left of Figure 4-4), the only allowed decay of the lightest neutralino is a prompt decay of the $\tilde{\chi}_1^0$ to a Z boson and a massless gravitino.

In the other model (right of Figure 4-4), the lightest neutralino is allowed to decay to a gravitino and either a Z boson or an SM-like Higgs boson, with a 50% branching fraction to each decay channel. The cross sections for higgsino pair production are computed in a limit of mass-degenerate higgsino states $\tilde{\chi}_2^0$, $\tilde{\chi}_1^\pm$, and $\tilde{\chi}_1^0$. Again, all the other SUSY particles are assumed to be heavy and decoupled. Following the convention of real mixing matrices and signed neutralino masses [65], the sign of the mass of $\tilde{\chi}_1^0$ ($\tilde{\chi}_2^0$) are set to +1 (-1). The

lightest two neutralino states are defined as symmetric (anti-symmetric) combinations of higgsino states by setting the product of the elements N_{i3} and N_{i4} of the neutralino mixing matrix N to $+0.5$ (-0.5) for $i = 1$ (2). The elements U_{12} and V_{12} of the chargino mixing matrices U and V are set to 1 [65].

Common to these models is that they produce an on-shell Z boson giving a pair of leptons of opposite charge and same flavor, in association with a SM boson, and two LSPs resulting in large p_T^{miss} . The SM boson is either a W, Z or a H, and their decay mode to jets (or b-tagged jets in the case of the H) is targeted in the search.

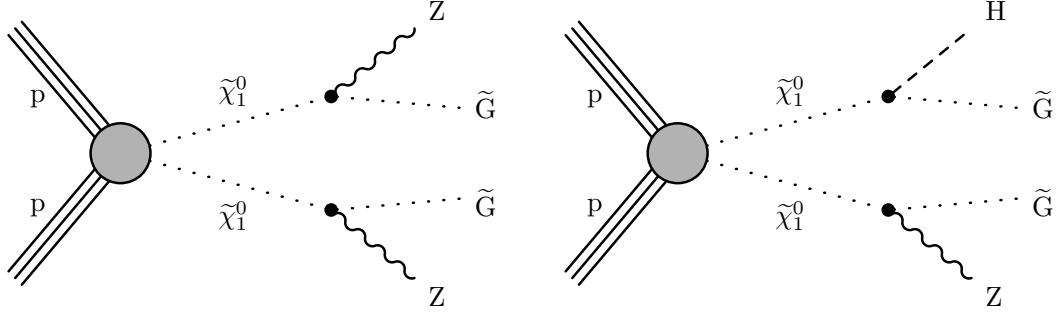


Figure 4-4. Diagrams corresponding to the neutralino-neutralino model of where the neutralinos are allowed to decay to a gravitino and a Z boson (left) and where the neutralinos are allowed to decay to a gravitino and a Z boson or a Higgs boson, with a 50% branching fraction to each decay channel (right).

Search for direct slepton production

Sleptons ($\tilde{e}_L, \tilde{\mu}_L, \tilde{\tau}_L, \tilde{e}_R, \tilde{\mu}_R, \tilde{\tau}_R$), are the superpartners of the charged left-handed and right-handed SM leptons.² At sufficiently heavy slepton masses, the sleptons undergo a two-body decay into one of the heavier neutralinos or a chargino, while direct decays to a neutralino LSP are favored for light slepton masses. The second paper that this thesis is based on is on a search for directly produced selectrons and smuons, under the assumption of direct decays $\tilde{\ell} \rightarrow \ell \tilde{\chi}_1^0$ with 100% branching ratio, as sketched in Figure 4-5. A search for slepton pair production has been performed with Run I 8 TeV data with the CMS experiment, [66], where slepton masses were excluded up to approximately 275 GeV. Figure 4-6 shows the increase in cross section of the left handed sleptons with the higher center-of-mass energy of 13 TeV, and is showing promising prospects on discovery or exclusion for this search.

² It is worth pointing out that the handedness does not refer to the helicity of sleptons (they are spin-0 particles!) but rather the handedness of the superpartners. The \tilde{e}_L is thus the superpartner of the left-handed electron and nothing else.

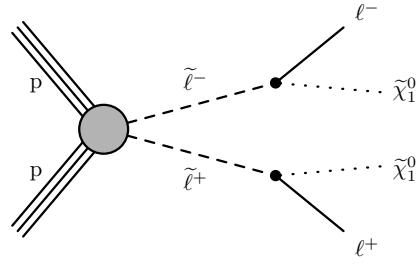


Figure 4-5. Diagram corresponding to the slepton model with two selectrons (smuons) directly produced and decay into electrons (muons) and a LSP

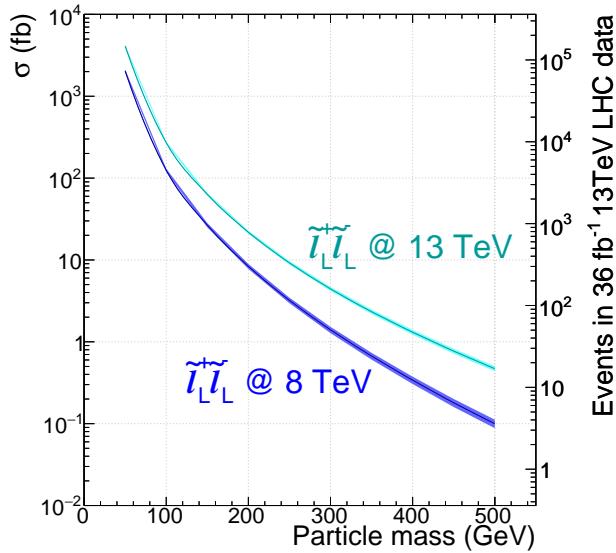


Figure 4-6. Graphs showing the cross sections for slepton pair production at a center-of-mass energy of 8 TeV and 13 TeV.

Part II

EXPERIMENTAL SETUP

5 — The Large Hadron Collider

This thesis uses data recorded with the Compact Muon Solenoid (CMS) experiment at the Large Hadron Collider (LHC) at the European Organization for Nuclear Research (CERN) near Geneva, Switzerland. The Large Hadron Collider is the worlds most powerful particle accelerator [67]. It is a circular superconducting accelerator that measures 27 km in circumference and is located 100 m underground in the same tunnel that hosted the Large Electron Positron (LEP) collider [68]. Around the LHC ring are four major particle physics experiments located. This chapter provides a short overview of the accelerator complex at CERN and the LHC.

5.1 The accelerator chain

The LHC ring is divided into 8 sectors separated by eight access points to the tunnel, named interaction points (IPs). The purpose of the LHC is to accelerate protons or heavy ions to center-of-mass energies of up to 14 TeV for protons and 2.76 TeV per nucleon for lead ions. The two rings in the LHC accelerate protons or heavy ions in opposite directions, to be collided at four of the eight IPs where huge particle detectors are located. The multi-purpose high luminosity experiments along the LHC ring are the CMS [69] and ATLAS[70] experiments located at IP5 and IP1, respectively. Both experiments collect data from pp, proton-ion and ion-ion collisions. The dedicated heavy ion collision detector ALICE (A Large Ion Collider Experiment) [71] is located at IP2, that collect the same data as ATLAS and CMS. The LHCb [72] experiment is dedicated to low luminosity B-physics and collect data from pp collisions at IP8.

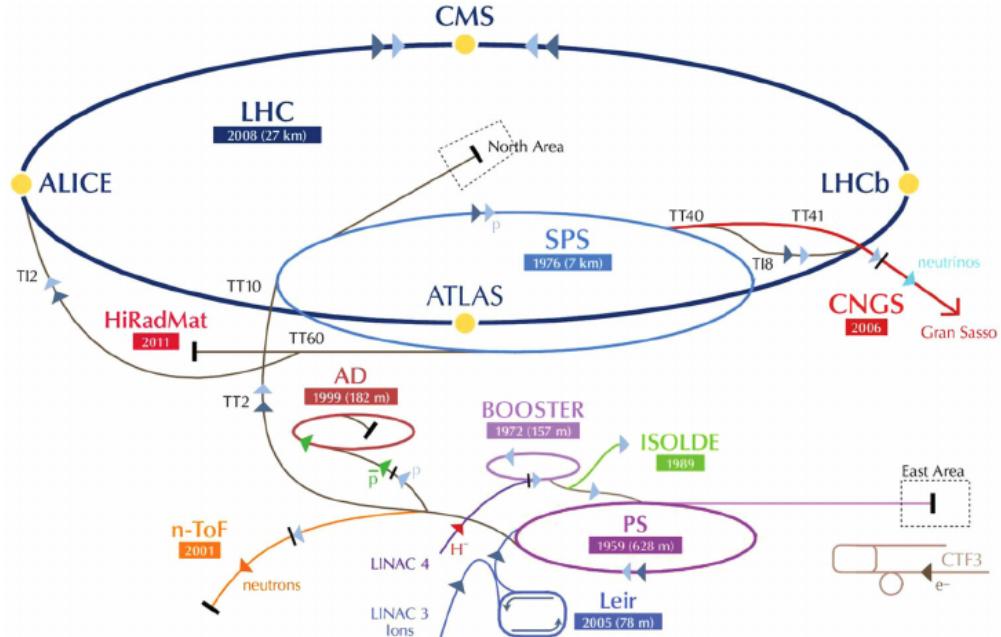


Figure 5-1. Sketch of the various accelerators and experiments hosted at CERN [73].

In order to get the protons to the center-of-mass energies quoted above, a long chain of circular and linear accelerators are needed. The protons start by being part of atoms in Hydrogen gas. Upon passing through an electric field, protons are separated from their electron and injected into the Linear Accelerator 2 (Linac 2). The Linac 2 accelerates the protons up to 50 MeV by passing them through alternating positive and negative cylindrical conductors charged by radiofrequency cavities. After the Linac 2, the protons are injected into a circular collider, the Proton Synchrotron (PS) Booster, where four superimposed synchrotron rings accelerate the protons to an energy of 1.4 GeV. The protons in the Booster are kept in a circular path by magnetic dipole fields. Following the PS Booster is the PS which is a circular accelerator with a circumference of 628 meters. The protons are accelerated by conventional magnets to an energy of 25 GeV, and they are kept on the circular path by 100 bending dipole magnets. The Super Proton Synchrotron (SPS) is the next accelerator for the protons and the last step before injection to the LHC. The SPS make use of the same technique as the PS but the larger circumference enables for the larger output energy of 450 GeV.

Up until now, the protons have been traveling in “bunches”, which are packets of hundred billion protons, that are separated in time by 25 ns (forming so called bunch trains). Upon injection into the LHC, the proton bunch trains are split in to two trains that enter the LHC in two opposing directions. The goal of the LHC

accelerator is to get these protons from 450 GeV to the maximum design energy of 7 TeV. This is achieved by letting the proton beams traverse radiofrequency (RF) cavities that are cooled down to 4.5 K using liquid Helium (up until this stage all accelerators have been operating in room temperature). The RF cavities provide a high frequency alternating electric field of 400.8 MHz. There are eight single-cell cavities per beam that produce the nominal voltage of 16 MV during storage that results in an energy gain per particle per turn is 485 keV. In addition to the RF cavities there are 1232 dipole magnets keeping the protons on the circular path. The dipole magnets measure 15 meters in length and weighing 30 tons each and are made of a niobium-titanium alloy [74]. The magnets are superconductors that can be operated at a temperature of 1.9 K and can reach a magnetic field of 8.33 T. In addition to the dipole magnets for bending the path of the protons, various other magnets such as decapole, sextupole and quadropole magnets are used for controlling or correcting the path. The accelerator chain is illustrated in Figure 5-1. Between 2010-2011 and during 2012, the LHC collided protons at a center-of-mass beam energy of $\sqrt{s} = 7 \text{ TeV}$ and $\sqrt{s} = 8 \text{ TeV}$ respectively, while the energy increased to $\sqrt{s} = 13 \text{ TeV}$ in 2015, 2016 and 2017.

5.2 Beam parameters

The number of events (N_{event}) that can be generated in a collision is dictated by an interplay between theoretical predictions and engineering capabilities, summarizes as

$$N_{\text{process}} = \mathcal{L}_{\text{int}} \sigma_{\text{process}} \quad (5-1)$$

where σ_{event} is the cross section of a particular process. The \mathcal{L}_{int} is the so-called integrated luminosity that is defined through the *instantaneous* luminosity. The instantaneous luminosity depends only on the beam parameters and is defined as:

$$L = \frac{N_b^2 n_b f_{\text{rev}} \gamma_r}{4\pi \epsilon_n \beta^*} F \quad (5-2)$$

where N_b is the number of particles per bunch, n_b the number of bunches per beam f_{rev} the revolution frequency of each bunch. Further, the γ_r is the relativistic gamma factor, ϵ_n is the normalized beam emittance, β^* is the β -function at the collision point and F a geometrical factor inversely proportional to the crossing angle of the two beams at the IP. The peak design LHC instantaneous luminosity is $\mathcal{L} = 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, and is reached for the beam parameters summarized in Table 5-1. By integrating the instantaneous luminosity L

Table 5-1. Beam parameters for beams in the LHC at injection and collision energy [67].

Parameter	Injection	Collision
Beam energy [GeV]	450	7000
Relativistic gamma factor (γ_r)	479.6	7461
Beam emittance (ϵ_n) [$\mu \text{ rad}$]	3.5	3.75
Half crossing angle [μrad]	± 160	± 142.5
β -function (β^*) [m]	18	0.55
Revolution frequency (f_{rev}) [Hz]		11245
Number of bunches (n_b)		2808
Particles per bunch (N_b)		1.15×10^{11}

(that has the unit of $\text{cm}^{-2} \text{ s}^{-1}$) over time, the result is the so called integrated luminosity, denoted by \mathcal{L}_{int} . The integrated luminosity has the unit cm^{-2} which is the inverse of the unit of cross section, which makes the resulting number of events in a process in Equation 5-1 unitless. The peak instantaneous stable luminosity reached by the LHC during 2016 was $1.527 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ and the total integrated luminosity greatly

surpassed the predictions and expectations. In the future High Luminosity run of the LHC that is foreseen for 2025, the so called HL-LHC, the expected instantaneous luminosity will be up to $7.5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$. The increase in instantaneous luminosity that the HL-LHC promises will grant the opportunity to achieve unprecedented precision of SM electroweak measurements, and a new window of exploration for physics beyond the SM. The main challenge will be to reduce effects from overlapping bunch crossings, and will be further discussed in Chapter 9.

5.3 Coordinate system and kinematic variables

The cylindrical shape of the multi-purpose high energy physics experiments, such as CMS and ATLAS, lend themselves well to the use of cylindrical coordinates. In these coordinates, the azimuthal angle ϕ is defined in the transverse $x - y$ plane perpendicular to the beam line. The polar angle θ (that strictly speaking belongs to the spherical coordinates) is also used, and is measured from the z -axis, as shown in Figure 5-2. The

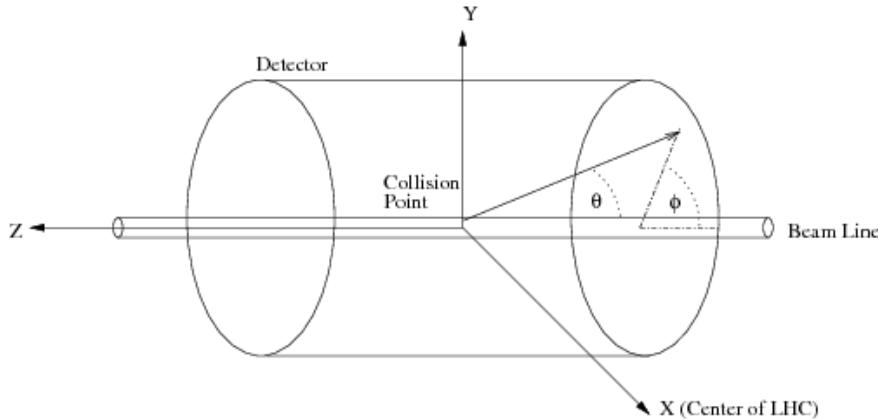


Figure 5-2. Illustration of the CMS coordinate system [75].

rapidity y is defined as:

$$y = \frac{1}{2} \ln \left(\frac{E + p_L}{E - p_L} \right) \quad (5-3)$$

with p_L being the momentum in the longitudinal direction. For relativistic particles it is common practice to describe the trajectory of the particles using the so called *pseudorapidity* η rather than y ¹, where η is defined as

$$\eta = -\ln \left[\tan \left(\frac{\theta}{2} \right) \right]. \quad (5-4)$$

This quantity will be heavily used in the next chapter to define the coverage of the various subdetectors, and beyond this chapter to define the invariant angular distance between particles as $(\Delta R)^2 = (\Delta\eta)^2 + (\Delta\phi)^2$. As the purpose of high energy particle collisions is to use the available momentum of the protons at velocities close to the speed of light to produce heavier particles, or *resonances*. A couple of essential kinematic variables used throughout the thesis will be identified using the kinematics of inelastic pp scattering. A momentum transfer scale of $\hat{s} = M^2$, is needed in order to produce a resonance of mass M [76]. Here $\hat{s} = x_1 x_2 s$ where $x_{1,2}$ are the fractions of the incoming hadrons momenta $p_{1,2}$ that partons 1 and 2 carry. The square root of s is the center-of-mass energy. The energy and the longitudinal momentum of a state produced by colliding

¹ This expression is obtained by rewriting the E in Equation 5-3 as $E = (p^2 + m^2)$ and identify $p_L/p = \cos\theta$.

two hadrons h_1 and h_2 is given by:

$$E = \frac{(x_1 + x_2)\sqrt{s}}{2}, p_L = \frac{(x_1 - x_2)\sqrt{s}}{2} \quad (5-5)$$

where the partons masses are neglected. Inserting the relations in Equation 5–5 into the definition of the rapidity in Equation 5–3, the following is obtained:

$$e^y = \sqrt{\frac{x_1}{x_2}}, x_{1,2} = \frac{M}{\sqrt{s}} e^{\pm y} \quad (5-6)$$

For a very large M , the two momentum fractions have to be very large with the result of $e^y \rightarrow 1$ i.e. $y \rightarrow 0$, called central rapidity. This implies that the decay products of a massive resonance from an interesting hard scatter event should end up in the central part of the detector.

The above argument, that most interesting physics events take place at central rapidities, motivates the design of a detector with higher precision in the central region. Additionally, the use of the *transverse* momentum p_T of a particle is preferred over the momentum p of a particle. Another key feature of high energy particle collisions is the momentum conservation in the transverse plane perpendicular to the beam axis. This conservation is motivated by the fact that the transverse momentum before the collision is well known (0). Further, any energy that is lost due to the production of particles that travel close to along the beam axis will not be detected. For these reasons, the conservation of momentum holds in the transverse plane only, instead of for all components of the momentum. In addition to the transverse momentum p_T , the transverse energy E_T and missing transverse momentum p_T^{miss} are commonly used.

Another key variable is the *invariant mass*, defined using kinematic and directional information of two particles, 1 and 2,

$$m_{1,2} = \sqrt{(E_1 + E_2)^2 - |\mathbf{p}_1 + \mathbf{p}_2|^2} \quad (5-7)$$

$$= \sqrt{2p_{T,1}p_{T,2}[\cosh(\eta_1 - \eta_2) - \cos(\phi_1 - \phi_2)]} \quad (5-8)$$

If the two particles are in fact the decay products of a resonance of mass M , the $m_{1,2}$ would coincide with the mass M . The invariant mass is thus a variable that can be used to *reconstruct* a resonance with two visible decay products. It will be used throughout this thesis as a means to reconstruct the Z boson using two leptons, and will be denoted with $m_{\ell\ell}$. If a value of $m_{\ell\ell}$ is close to the Z boson mass, the leptons will be referred to as 'compatible with the Z boson'.

Further, the transverse mass, M_T can be constructed using the transverse energy and momentum of two objects, often when one of the objects can not be detected. The definition of the M_T is

$$M_T = \sqrt{m_1^2 + m_2^2 + 2(E_{T,1}E_{T,2} - \mathbf{p}_{T,1}\mathbf{p}_{T,2})}. \quad (5-9)$$

This quantity will prove useful when reconstructing the mass of the W boson using a lepton and a neutrino.

Finally, it is worth mentioning that the practice in high energy physics is to set the speed of light c and the Plack constant \hbar to 1, i.e. $c = \hbar = 1$. This results in the energy being measured in the same unit as mass, which is eV (instead of eV/c^2).

6 — The CMS experiment

The CMS experiment is a multipurpose apparatus designed to detect particles from collisions of protons or heavy ions. It is located 100 m under ground at the fifth interaction point along the LHC ring (“Point 5”) in Cessy, France. The CMS experiment was designed to observe and study the properties of the Higgs boson, and other known Standard Model particles, as well as search for new physics. Apart from detecting particles from pp collisions, the CMS detector can also detect particles from heavy ion collisions.

This chapter provides a description of the main hardware components of the CMS experiment. A brief overview of the structure of the CMS detector and the coordinate system is given. The main features of the detector are introduced in Sections 6.3 - 6.6, that include the silicon tracking device closest to the beamline, the surrounding calorimeters and the 3.8T superconducting solenoid, followed by the outermost layer consisting of a muon detector. The data used in this thesis was collected during the 2016 running of the LHC, and this chapter will reflect the layout of CMS detector during this year.¹

¹ The tracker was upgraded after the 2016 running of the LHC with the inclusion of a new silicon pixel layer in the barrel.

6.1 The CMS detector

A full description of the CMS experiment can be found in [69] and a sketch is shown in Figure 6-1. This

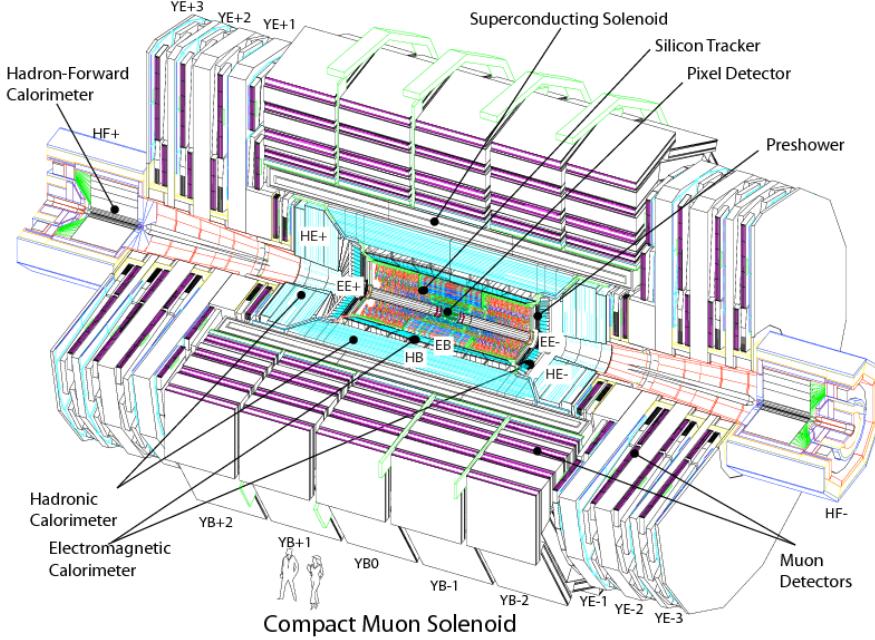


Figure 6-1. General view of the CMS detector. The major detector components are indicated, together with the acronyms for the various CMS construction modules [77].

multi-purpose apparatus is built around the LHC beam line with circular layers of subdetectors with increasing radii, resulting in a cylindrical shape with a diameter of 16 m and length of 22 m. The z -axis is defined to be the direction along the LHC beamline inside CMS, with the positive direction pointing towards the Jura mountains from Point 5. In 2016, the CMS experiment recorded data from pp collisions at $\sqrt{s} = 13$ TeV corresponding to 37.8 fb^{-1} of integrated luminosity, as seen in Figure 6-2, of which 35.9 fb^{-1} was declared to be good for analysis and used throughout this thesis. In the following sections, the magnet and the various subdetectors are introduced.

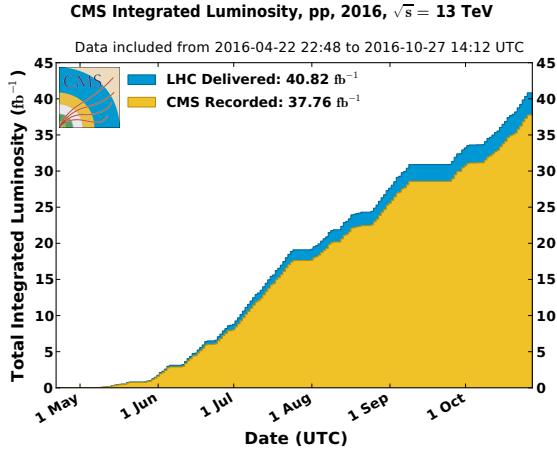


Figure 6-2. Cumulative measured luminosity versus day delivered. This measurement uses the best available offline measurement and calibrations [78].

6.2 The Magnet

The word “Solenoid” in Compact Muon Solenoid refers to the superconducting solenoid magnet that the experiment is built around. A strong magnetic field is a key feature at an experiment of this magnitude as it is essential for momentum resolution of charged particles in the tracker. The solenoid measures a length of 12.9 m and six meters in inner diameter. The large inner radius allows for the tracker and calorimeters to be fully contained within the solenoid. This design reduces the material between the calorimeters and the IP which is desireble for precise momentum measurements. The solenoid is composed of a conducting material, a NbTi alloy, subjected to a current of 19 kA that results in a magnetic field of 3.8 T. The magnet system includes a return yoke to keep the magnetic field lines homogeneous with the distance from the IP, and total weight including the solenoid is 11 000 tons [79].

6.3 The Tracker

The CMS inner tracker measures a length of 5.8 m and a diameter of 2.5 m, and the overview that will be presented in the following is visualized in Figure 6-3 and follows the description in [80]. In order to achieve a high efficiency in tagging b-quark jets, it is crucial that the tracker layers are as close to the LHC beamline as possible. Closest to the beamline is the smallest subdetector of the tracker, the silicon pixel tracker, consisting of a barrel (BPix) of three layers of radii 4.4, 7.3, 10.2 cm, and two pairs of endcap disks (FPix) at a distance of $|z| = 34.5$ cm and $|z| = 46.5$ cm away from the interaction point. These layers provide a three dimensional position measurement with a transverse coordinate position resolution of $10 \mu\text{m}$ and a longitudinal coordinate position resolution of $20 - 40 \mu\text{m}$. The total number of pixels in the BPix and FPix amounts to 66 million and cover in total an area of 1 m^2 . Outside of the pixel tracker are four subsystems of 9.3 million silicon micro strips.

The tracker inner barrel (TIB) is composed of four layers and cover a region of radius 20 cm to 55 cm. The pitch² is $80 \mu\text{m}$ for the first two layers of the TIB and $120 \mu\text{m}$ for the next two layers, and they are all

² The pitch is the distance between the p+ implants in the n type SiO₂ bulk.

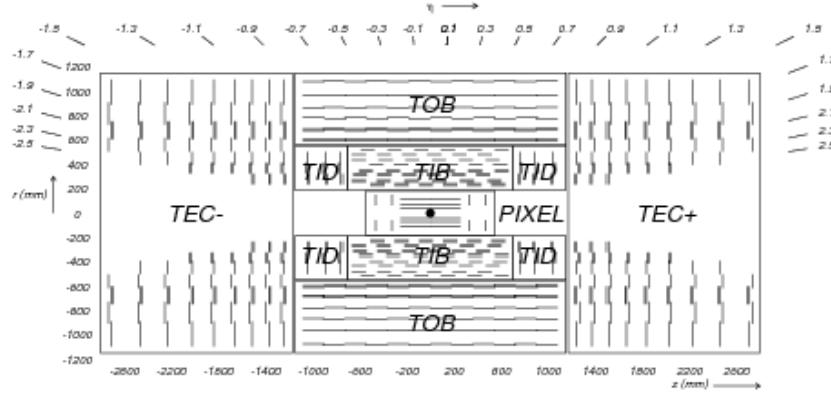


Figure 6-3. Schematic cross section through the CMS tracker in the r - z plane. Each line-element represents a detector module. Closely spaced double line-elements indicate back-to-back silicon strip modules, in which one module is rotated through a ‘stereo’ angle, so as to permit reconstruction of the hit positions in 3-D. Within a given layer, each module is shifted slightly in r or z with respect to its neighboring modules, which allows them to overlap, thereby avoiding gaps in the acceptance [80].

oriented parallel to the beamline [81]. The tracker inner disks (TID) complements the TIB with three disks on each side that extends to coverage in the $|z|$ direction to ± 118 cm. The TID modules are built using radially placed sensors forming a wedge shape with pitch ranging from 81 to $158\text{ }\mu\text{m}$.

Outside of the TIB, the tracker outer barrel (TOB) consisting of six barrel layers cover the region beyond 55 cm in radius and the same $|z|$ coverage as the TID along the beamline. The TOB has thicker strip sensors of $500\text{ }\mu\text{m}$ and the first four layers use strips with a pitch of $183\text{ }\mu\text{m}$ and the last two layers use strips with a pitch of $122\text{ }\mu\text{m}$. Finally, the last tracker subsystem is the tracker endcap (TEC) that consist of nine disks on each side of the TIDs and TOB, with a total coverage of the region $124 < |z| < 282$ cm. The TOB and TEC have a resolution of the position measurement ranging from $18 - 47\text{ }\mu\text{m}$. Putting all of this together, the inner tracker fully covers the region $|\eta| < 2.4$ and guarantees in total 9 hit measurements. The performance is quantified in terms of the resolution of the p_T for single muons, which results in 0.65 to 1.5% at $\eta = 0$ and 1 to 2% at $|\eta| = 1.6$ for muons of 10 and 100 GeV.

6.4 The ECAL

The choice of a high resolution CMS electromagnetic calorimeter (ECAL) is motivated by the physics potential that was foreseen for the discovery of the Higgs boson in 2012, as photons and electrons are key ingredients in at least three decay channels [82]. This motivation lead to the choice of the design of a hermetic, homogeneous, fine grained scintillating calorimeter consisting of 75,848 lead tungstate (PbWO_4) crystals. The crystals are arranged in a central barrel region ($|\eta| \leq 1.48$) have a crystal length of 23 cm and a front face of $2.2 \times 2.2\text{ cm}^2$. The crystals arranged in the two ECAL endcap regions (up to $|\eta| = 3.0$) measure 22 cm in length and a front face of $2.86 \times 2.86\text{ cm}^2$. In the barrel, the crystals are organized into 36 supermodules, each containing 1,700 crystals. In the endcaps, the crystals are organized into dees each containing 3662 crystals. Due to the challenges associated with fitting a subdetector in a confined volume, there is a slight gap between the end of the ECAL barrel and the endcaps where there is no crystals covering the region $1.4 \leq |\eta| \leq 1.6$. This is known as the ECAL transition region, and will be used in this thesis to reject events with leptons in this region. The layout of the supermodules and dees is presented in Figure 6-4. In addition to the barrel and endcaps, the ECAL also consists of a preshower detector which is based on lead absorber and silicon strip

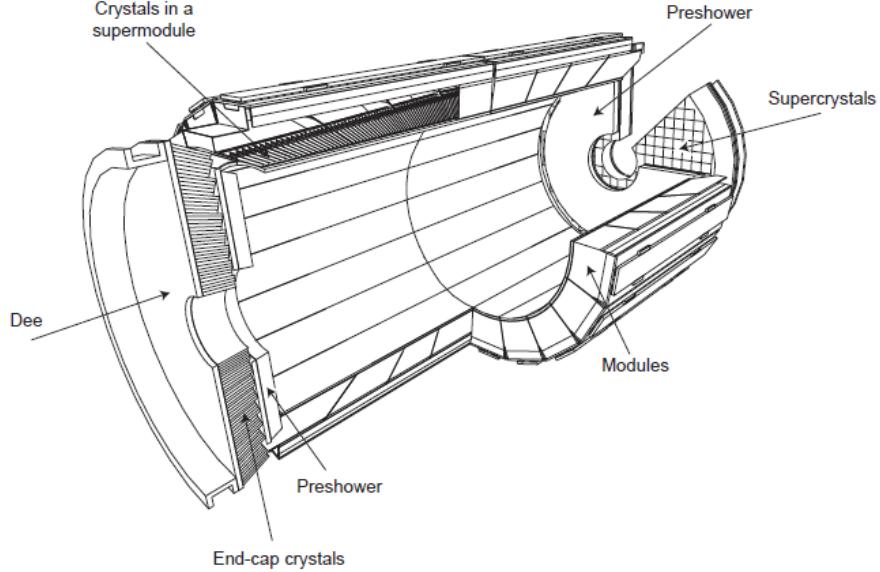


Figure 6-4. Layout of the CMS ECAL, showing the barrel supermodules, the two endcaps and the preshower detectors. The ECAL barrel coverage is up to $|\eta| = 1.48$, the endcaps extend the coverage to $|\eta| = 3.0$ and the preshower detector fiducial area is approximately $1.65 \leq |\eta| \leq 2.6$ [83].

sensors. The preshower covers a region of $1.65 \leq |\eta| \leq 2.6$ and is motivated by the ability to improve the differentiation of the $\pi^0 \rightarrow \gamma\gamma$ process from $h \rightarrow \gamma\gamma$. The lead tungstate scintillating crystals are of high density ($\rho = 8.28 \text{ g/cm}^3$) and have the nice feature of short radiation length (X_0) and small Molière radius (R_M), where these quantities are related according to

$$R_M = 0.0265 X_0 (Z + 1.2) \quad (6-1)$$

where Z is the atomic number. The energy absorbed in the crystals by the incoming electrons and photons is emitted as light, 80% of the time within 25 ns. The energy resolution obtained with this design has been quantified at beam tests and results in

$$\frac{\sigma_E}{E} = \frac{2.8\%}{\sqrt{E}} \oplus \frac{12\%}{E} \oplus 0.3\% \quad (6-2)$$

where the order of the terms are stochastic, noise and constant term respectively [83]. The light produced in the scintillating crystals is collected by avalanche photodiodes (APDs) in the barrel and by vacuum phototriodes (VPTs) in the endcaps. The APDs are sensitive to changes in temperature according to $-2.3\text{ }^\circ\text{C}$. In the very front end (VFE) cards the signals from the APDs are pre-amplified and shaped by an ASIC multi-gain pre-amplifier chip. Trigger towers are formed by 5×5 crystals, and the trigger primitives are generated from the summed amplitudes of these 25 crystals in the front end (FE) cards.

6.5 The HCAL

The hadronic calorimeter is located outside of the ECAL while still being, mostly, contained within the solenoid. In contrast to the ECAL, the HCAL has the purpose to identify a variety of particles, mainly hadronic jets but also help in identifying electrons, photons and muons with information from the ECAL and muon systems. The ECAL is designed to fully contain the electromagnetic shower induced by electrons or photons, and the HCAL, being the next layer, is designed to contain hadronic showers and fully absorb hadrons before the solenoid. As much larger depth of detector material is required compared to that of the ECAL, hadronic calorimetry is considered much more challenging. Additionally, the HCAL energy resolution is also worse compared to that of the ECAL due to intrinsic fluctuations.³

The HCAL layout consist of fours sections as shown in Figure 6-5. The barrel (HB) covers the $0 \leq |\eta| \leq 1.4$

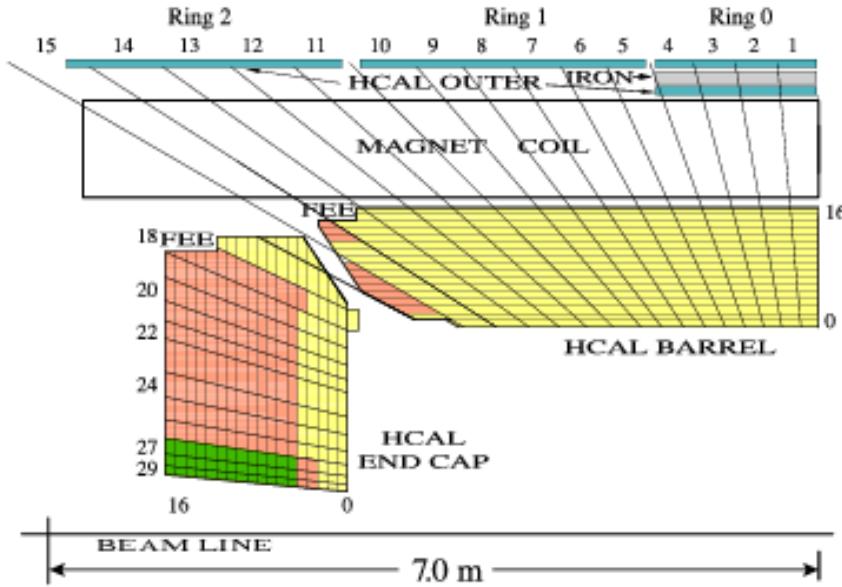


Figure 6-5. The HCAL tower segmentation for one-fourth of the HB, HO, and HE detectors. The numbers on top and on the left refer to the tower numbers. The numbers on the right and on the bottom (0-16) indicate the scintillator layers numbers inserted into slots in the absorber. The shading represents independent longitudinal readouts in the HB/HE overlap and the small angle regions [84].

range and the endcaps (HE) cover the $1.4 \leq |\eta| \leq 3$ range, and both are sampling calorimeters made by layers of brass absorber alternated with plastic scintillator. As the depth of the HB is limited to what can be contained within the solenoid, the thickness at $\eta = 0$ only covers 5.8 nuclear interaction lengths (λ_I) which is increased to $10\lambda_I$ at $|\eta| = 1.2$ ⁴. Only about 95% of hadrons above $p_T \geq 100$ GeV are fully contained in

³ Intrinsic fluctuations are a result of the significant incoming energy fraction being invisible, since it is deployed in processes like nuclear break-up.

⁴ The nuclear interaction length is the mean-free path that an incident hadron can travel in a medium before it is fully absorbed due to nuclear interaction. It is defined as $\frac{1}{\lambda_I} = \sigma_{inel} \frac{N_A \cdot \rho}{A}$, where σ_{inel} is the inelastic cross-section, ρ the density and A the atomic mass.

the HCAL, leading to a small fraction of hadrons making it through the solenoid. For this reason, an Outer HCAL (HO) is located outside of the solenoid to recover these hadrons, and uses the same plastic scintillator as HB and HE but the magnet material as the absorber. The HB, HE and HO sections all use wave-length shifting fibers to extract the scintillating light that is guided to hybrid photodiodes (HPDs).

The forward part of the HCAL (HF) covers the range $3 \leq |\eta| \leq 5$ and is located 11.1 m from the interaction point. The purpose of the HF is to improve the measurement of p_T^{miss} and to identify very forward jets, as those produced in vector boson fusion (VBF). The region covered by the HF is subject to the largest particle flux with radiation doses reaching 100 Mrad/year. For this reason, a different construction choice is made for this region, exploiting Cherenkov calorimetry technique. The HF is composed of steel absorber interspersed with quartz core and acrylic clad fibers in the longitudinal direction [85]. The fibers collect the Cherenkov light produced by the showers in the absorbers, and the light is sent to a photomultiplier tube (PMT).

6.6 The muon system

The final layer of the CMS detector is dedicated to the measurement of muons, and is the motivation behind the word “Muon” in Compact Muon Solenoid. The objective of this subsystem is to identify muons, trigger on muons and measure the momentum and charge of muons. A sketch of the experimental setup is shown in Figure 6-6 and the following descriptions follow closely that in [86]. A combination of three detector

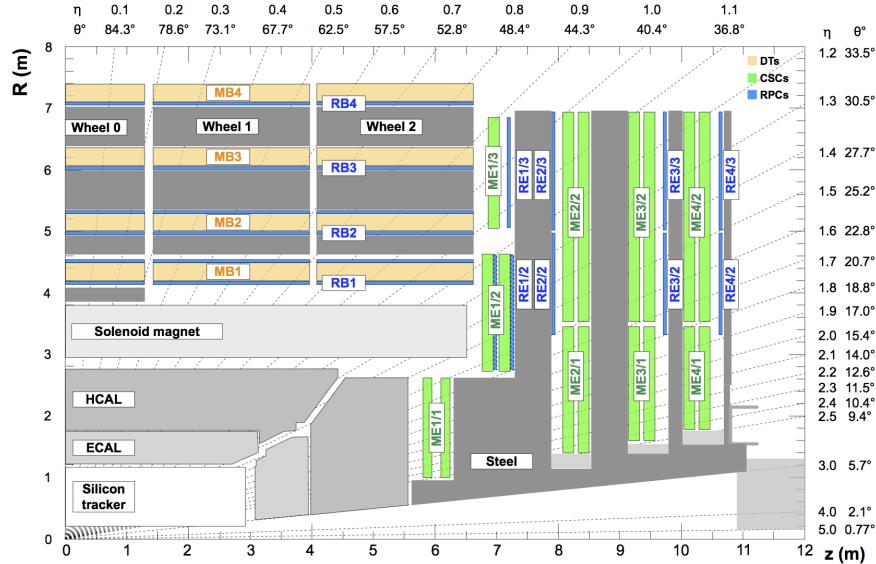


Figure 6-6. Cross section of a quadrant of the CMS detector with the axis parallel to the beam (z) running horizontally and the radius (R) increasing upward. The interaction point is at the lower left corner. The locations of the various muon stations and the steel flux-return disks (dark areas) are shown. The DTs are labeled MB (“Muon Barrel”) and the CSCs are labeled ME (“Muon Endcap”). RPCs are mounted in both the barrel and endcaps of CMS, where they are labeled RB and RE, respectively [86].

techniques is deployed that are motivated by the various expected experimental implications. Similarly to the subdetector systems described in the previous sections, the muon systems consists of a barrel and two endcaps, where the barrel is interleaved with layers of the steel flux-return yoke.

The detector technique deployed in both the barrel and endcaps is gas ionization particle detectors and each of the modules are commonly referred to as chambers. As the muon system is located outside of the solenoid, the effective magnetic field is diminished to below 0.4 T between the yoke segments in the barrel. This results in non-uniform magnetic field strengths and a reversion of the muon trajectory. In addition to the relatively low magnetic field in the barrel, the expected rate of muons is low, thus making the use of drift tubes (DT) a suitable choice. The barrel DT chambers cover a region of $|\eta| \leq 1.2$ and is split in 12 segments around ϕ . The DT chambers consist of drift cells. Each cell contains a gold plated stainless steel anode wire operating at 3.6 kV, surrounded by a gas mixture of 85% of Argon and 15% CO₂. This results in a drift time of 400 ns. Cathode plates on the sides of the cells operate at ± 1.8 kV. An incident muon will release electrons in the gas volume that will drift to the anode and produce avalanches in the region close to the wire where the field strength is increased. Four layers of parallel cells form a super layer (SL) and a chamber consists of three SLs where one layer is oriented orthogonally to the other two in order to enable precise measurement in both the $r - \phi$ and the $r - z$ directions.

The muon system endcaps cover a region ($0.9 \leq |\eta| \leq 2.4$) where the expected rate of muons and neutron background is much higher than that in the barrel. As this requires a faster response, Cathode Strip Chambers (CSCs) that are multiwire proportional counters are used, motivated by the shorter drift path than that of the DTs. Additionally, the CSCs can tolerate the higher magnetic field that the muon endcap regions is subjected to. Each endcap consist of four rings of chambers mounted on the face of the endcap steel disks. A CSC consist of 6 layers, and the cathode strips are aligned radially outwards while the anode wires run perpendicularly to the strips, allowing for a position measurement in $r - \phi$. Each layer has 80 cathode strips. The chambers are filled with a gas mixture of 50% CO₂, 40% Argon and 10% CF₄, and the anode wires operate at 2.9 kV to 3.6 kV depending on the ring. Interspersed in the barrel DT layers and endcap CSC layers are resistive plate chambers (RPCs). The purpose of these layers is to provide fast and independent triggering at lower p_T thresholds in the region $|\eta| \leq 1.6$. The RPCs are double gap chambers, where each gap consist of resistive Bakelite plates separated by a gas filled gap volume of the same thickness. When a charged particle crosses the RPC, the gas in the gap volumes is ionized and an avalanche is created by the high electric field (due to the application of a 9.6 kV voltage) and an image is induced that is picked up by the readout strips.

7 — Event Reconstruction

The basis of any LHC data analysis relies on the concept of an ‘event’. An event refers to a collision of at least two protons that results in the full readout of the CMS detector, and subsequently the combination of subdetector information that forms physics objects. The LHC delivers proton bunches at a rate of 40 MHz but only a small fraction of the protons in the bunches result in collisions of interest to the CMS physics program. In fact, if all collisions would lead to a full readout of the detector, there would not be enough bandwidth to readout the information, nor space to store it for offline analysis. With these limitations in mind, the CMS has developed a two-tier system to select events of physics interest, the Level 1 (L1) trigger and the high level trigger (HLT).

The trigger system and the computational infrastructure used for the event reconstruction is presented in this chapter. The data from pp collisions recorded by the CMS experiment is what is used to measure properties of SM particles and search for new physics. But in order to properly study detector effects and design searches for unknown physics, simulated events are used. Monte Carlo generated events are used, with inputs from the theoretical framework presented in Chapter 2. The interaction of particles with matter can also be simulated, which is crucial for validation of the performance of the detectors and the reconstruction algorithms. This chapter describes how the physics processes are generated and how their interaction with the detector material is simulated.

7.1 The trigger system

During 2016 pp run of the LHC, the LHC delivered proton bunches with a time separation of 25 ns, and a peak luminosity reaching the unprecedented value of $10^{-34}\text{cm}^{-2}\text{s}^{-1}$. At this peak luminosity, the pp interaction rate exceeds 1 GHz, as the mean of the number of interaction per bunch is 25, which is a rate impossible to readout with the technology to date. A two tier triggering system helps differentiate which of the 1 GHz contain interesting events, while discarding the rest. This enables a reduction of the rate down to 400 Hz, which is feasible for offline storage.

The Level 1 trigger

The L1 trigger is a hardware system that uses input from the calorimeters and the muon detectors to make a decision to keep the event or not [87]. Tracking information is not included at this step, as track reconstruction is too time consuming. Instead, information from the calorimeters and the muon detectors is combined into

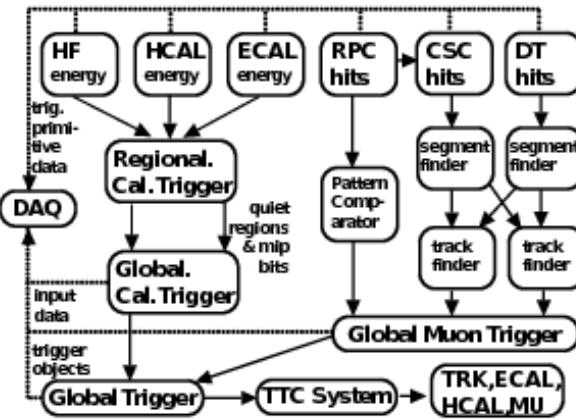


Figure 7-1. Chart showing the organization of the various components of any L1-accept. The calorimetry and muons systems work in parallel and are combined into a global trigger [87].

a global trigger, as illustrated in Figure 7-1. The L1 calorimeter trigger uses inputs like transverse energy and quality flags of the ECAL, HCAL and HF, in the form of Trigger Primitives (TP) from coarsely grouped trigger towers of the calorimeters. This information is provided to the regional calorimeter trigger (RCT), where it is combined to form $e\gamma$ candidates. The next step is the global calorimeter trigger (GCT), where jets are formed using the sum of transverse energy E_T , and the information on the pseudorapidity and τ veto from the RCT is used to label them as central, forward and τ jets. A crude calculation of the p_T^{miss} can also be performed at L1, by summing the x and y components of the transverse energy in quadrature and rotating the vector by 180° . The objects returned by the GCT, isolated and non-isolated $e\gamma$ candidates, central, forward and τ jets and p_T^{miss} , are passed on to the global trigger (GT) and the information is used, together with that of the muon triggers, to decide if an event is kept or not. At this time, there has also been a collection of information gathered from the muon system. Again, no information from the tracker is used at L1, but muons can still be more or less efficiently identified by the muon system. Using the CSC and DT track finders, the tracks of muons can be identified along with their p_T . This information, together with muon trigger candidate hits in the RPC, is sent to the global muon trigger (GMT). The information from the GCT and GMT is combined and a decision is made whether to keep the event or not, so called L1-accept. This trigger system has now brought down the rate of 1 GHz to 100 kHz, and the L1-accept is passed to all subdetectors that are read out and passed on to the HLT.

A way to further reduce the rates is to scale them down. Some processes with large cross sections, such as QCD, produces single photon events at rates higher than manageable, especially at low photon p_T . For this reason, the single photon triggers are *prescaled*, meaning only a fraction of the events are recorded, and the fraction is evolving with the luminosity during data taking. At analysis level, the events recorded with the prescaled triggers are scaled up according to whatever value they were prescaled with. In this thesis, these kinds of prescaled triggers are used when collecting the single photon sample used for the p_T^{miss} performance studies.

The high level trigger

While the L1 is completely hardware based and process the information from the detector underground in the experimental cavern, the HLT is both software and hardware based and located in computer farms on the ground level, running on 13,000 CPU cores. At this level, a more thorough object reconstruction is performed with the L1 information. The so called HLT path is a set of algorithms executed in a sequence of steps. As the tracking is more computing expensive, the first steps is to make a requirement on information from the calorimeters and the muon detectors, before performing the track reconstruction.

The basic idea is to enable triggering on high quality objects for offline reconstruction, while keeping the rates to a minimum. For this reason, variables such as the p_T and isolation of an object are used in different combination. The rate for triggering on a low p_T muon would be very high if no other requirements are imposed, but if one further requires a well isolated muon, this reduces the rate and possible misidentification of the muon. Conversely, as the production rate of higher p_T muons is lower, one can afford to only impose a p_T requirement and still keep the rate low. As will be seen later in the thesis, this is the reason for the use of the different HLT paths such as `HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL` (involving isolation requirement on both muons) and `HLT_Mu30_TkMu11` (involving higher p_T requirements and no isolation requirements), to ensure triggering on all possible events with interesting physics.

Further, HLT paths can involve objects like the Calo p_T^{miss} (computed using only calorimeter deposits), Particle Flow (PF) p_T^{miss} (computed using only PF jets), H_T (scalar sum of all jet p_T above a threshold) and $\cancel{H_T}$ (the missing H_T). The p_T^{miss} triggers are most sensitive to triggering on anomalous events where the large p_T^{miss} is originating from noise, beam halo or other sources, which will be discussed in Chapter 9. In order to keep low rates for these triggers, noise cleaning algorithms are applied at the HLT, and energy deposits associated to beam halo or HB/HE noise is removed from the energy sum and the calorimeter based p_T^{miss} is recomputed. The noise cleaning algorithms are fully efficient in Run 2 and reduce the rate by a factor of 2.5. Additionally, there is even the possibility to get a better estimation of the p_T^{miss} at HLT, by propagating the jet energy corrections (JECs) to the jets and to the computation of the PF p_T^{miss} .¹ This results in an improved efficiency and a further rate reduction. The HLT reduces the rate from the L1 output of 100 kHz to 1 kHz. The data that now have passed the L1 and HLT will be sent of the CERN Tier-0, described in the subsequent section. As the following reconstruction of the data is very computing intense and time consuming, there is the possibility to only use the cores to process some of the data immediately, while some is left for later reconstruction.

¹ The JECs will get a proper introduction in Section 8.4, and the p_T^{miss} calibration in Section 9.2.

7.2 Data reconstruction

The computing of the LHC experiments is made on a four tiered computing system. The tiers are ordered with descending order of importance. The largest tier, the Tier-0 is the CERN Data center, located at the CERN main site, and is the first line of tiers that the data collected by the detectors at the LHC experiment is processed at. The Tier-1 consists of 13 computing centers connected to the Grid located all over the world, and share the data reconstruction with the Tier-0, along with providing storage. The remaining computing tiers are the Tier-2 and Tier-3, which consists of computing resources at universities and institutes where analysis work is performed and stored. Once the CMS HLT has decided what events to keep, the information is sent to the Tier-0 at CERN, where the reconstruction of the event is performed. The CMS software, CMSSW, is a centrally maintained code base that reconstructs objects out of calorimeter deposits, hits in the muon chambers and tracks in the pixel and silicon trackers, and stores them in *Event Data Model* format (EDM).

Dataformats

The reconstruction of the collision data is processed at the Tier-0 and is stored in so-called RECO format, a format containing much of the detector information and is thus associated with very large event sizes of 1.2 MB. In order to make the storage and analysis as easy as possible, a set of data tiers are processed that keeps the event sizes to a minimum. The first step in the data tier format processing is the Analysis Object Data (AOD) which has discarded much of the RAW detector information not needed for analysis, and decreased the event size greatly to around 300 kB. The next step is the MINIAOD format, that was introduced for the Run 2 of the LHC. The MINIAOD contains all high level physics objects, all high level corrections for jets and p_T^{miss} , all particles returned by the Particle Flow algorithm, all generator truth information for simulation, and all trigger information. By only saving the above information needed for mainstream analyses, the event sizes are kept at 30-50 kB. In order to cope with the huge amounts of data collected during Run 2, and the various sets of MC samples specific to setup of each year, a further reduction in the event size is needed, and to this end the so called nanoAOD is developed, that manages to keep the event size at 1 kB. The MINIAOD data and MC samples are stored on the Tier-0 and Tier-1, and analysis specific frameworks are used to create subsets of samples containing the particular datasets needed. As the analyses presented in this thesis are based on different dileptonic datasets, the analysis specific framework used is aimed at picking data collected with dileptonic triggers and the various SM processes containing leptons, and store this in a format called Trees. These trees are used on analysis level for plotting, counting, fitting and statistical analysis.

7.3 Simulated events

Any search or measurement at the LHC are relying on simulated events, so called Monte Carlo (MC) generated events. Physical processes are simulated in a chain, starting from the foundations dictated by the theoretical framework as described in Chapter 2, followed by the decay, radiation and hadronization of the particles produced, and finally, the simulation of the interaction of the generated process with the detector material. Events are generated from both known SM processes to predict some backgrounds in the searches, and in order to validate the data-driven background prediction methods.

As the SUSY signals analyzed in this thesis are multiple, and include numerous assumptions on the masses of the SUSY particles, the full chain of generation for all signal scenarios would be too computationally heavy to be feasible for an experiment performing hundreds of measurement and searches simultaneously. For this reason, a faster simulation package is used for the SUSY signals on the level of the detector simulation step,

that greatly reduces the computational time by parametrizing the interactions. The gain in the reduction of the computational time is traded off with a slight decrease in precision, but is taken into account in the statistical analysis. The subsequent chapters contain the description of the three levels of simulation, and the software packages used.

Hard parton scattering

This section relies on the theoretical framework presented in Chapter 2. At hadron colliders, as opposed to lepton colliders, the energy available in the collisions is distributed over the three valence quarks of the hadrons, the sea quarks and the gluons. In order to properly model the possible interactions, parton distribution functions (PDFs) are used, that dictates the probability of finding a parton within a proton, with a certain energy fraction x of the proton. Further, there are several ways of calculating the matrix element (ME). The square of the MEs are the transition probabilities of the gluons or (anti-)quarks in to the physical process of interest. The different packages utilizes different order of the strong coupling constant α_s , and are choice of package depends on the requested precision. These packages are known as Monte Carlo (MC) generators, which utilizes the MC random sampling technique.

Parton distribution functions

Parton distribution functions explain the probability to find a parton in a proton with some given energy fraction. Due to the non-perturbative nature of partons, i.e. that partons can not be freely observed, the parton distribution functions can not be computed from first principles, however, they can determined by fitting observables to experimental data. One way of determining the PDFs is through the so-called NNPDF3.0, that uses Neural Networks (NN) to model datasets collected by the ATLAS, CMS, HERA-II and LHCb experiments [88]. The NNPDF3.0 tool is developed to be used for searches and measurements during the LHC Run 2, and is used for all generated samples used in this thesis. To demonstrate the behavior of the parton distributions, the NNPDF3.1 NNLO are displayed in Figure 7-2. The NNPDF3.1 is a successor of NNPDF3.0, and although it is not used for the MC in this thesis it lends itself well for a demonstration. As can be seen in Figure 7-2, the red curve, corresponding to the gluon PDF scaled down by a factor of 10, is dominating at the low energies, for both resolution scales. The corresponding interpretation is that the LHC pp collisions are dominated by low energy gluon induced processes, with a sub-dominant contribution of valence quarks at higher energy fractions. This feature is explaining the necessity of the very high collision rate of 40 MHz, as most of the interactions are between low-energy carrying gluons, which do not have enough energy to creat “interesting” physics events.

Hard scattering processes

The generation of a physics process relies on the calculation of the matrix elements, which states the transition probabilities of the partons involved in the interaction to particles. The MEs can be calculated in many different ways, and each calculation corresponds to a MC generator, with the most significant difference being the order of α_s used. Leading order (LO) corresponds to the first order in α_s . In this thesis, a generator that uses LO ME calculation is Madgraph [90], which includes contributions with higher jet multiplicities via MLM [91] multileg matching and is used to simulate e.g. single γ events. Higher order calculations, such as next-to LO (NLO) and next-to-next-to LO (NNLO) include higher orders in α_s and are therefore more accurately describing the physical process. Madgraph can also be combined with a NLO generator, MC@NLO, into MADGRAPH5_AMC@NLO, and is used when generating e.g. the $WZ \rightarrow 3l1\nu$ process. A full NLO generator is POWHEG [92, 93], that is used to simulate e.g. the $t\bar{t} \rightarrow 2l2\nu$ process. The complete list of samples used in this thesis is presented in Appendix A.

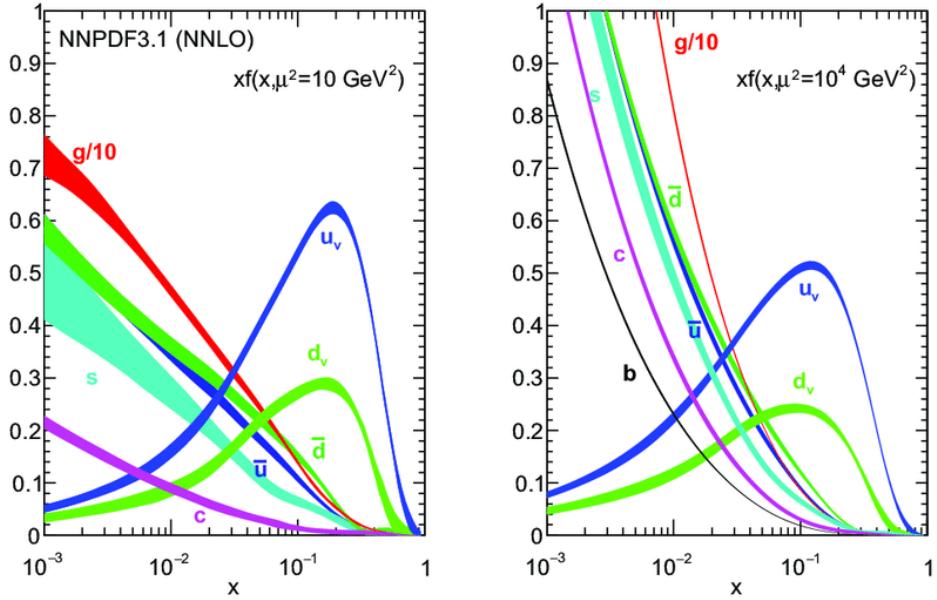


Figure 7-2. The NNPDF3.1 NNLO PDFs, evaluated at two resolution scales; $\mu^2 = 10 \text{ GeV}^2$ (left) and $\mu^2 = 10^4 \text{ GeV}^2$ (right) [89].

Multiparton interactions and parton showers

Due to the busy nature of hadron collisions, partons other than those involved in the hard scatter interact. This is known as Multiparton interactions (MPI). Initial state radiation (ISR) and final state radiation (FSR) are, as the names suggest, an emission of a gluon or a γ , either before or after the main interaction in the event. Apart from both types of radiation are adding a photon to the event, ISR and FSR result in slight kinematic differences. To illustrate this, we assume a Drell-Yan production, $Z \rightarrow e^+e^-$, with ISR and FSR. In the case of the ISR, the photon is radiated by the partons before the production of the Z boson, with a result of carrying off some of the available energy in the collision. Conversely, in FSR, the photon is radiated off of the electron or positron decayed from the Z boson, carrying off some of the energy of that lepton. So if one would reconstruct the invariant mass of the leptons, it would not completely form the Z boson mass of 91 GeV. Instead, one would have to construct the invariant mass of the leptons and the photon in order to recover the Z boson mass. The simulation of both the MPI and ISR and FSR is done with PYTHIA [94].

Hadronization

Colored quarks and gluons produced in the hard scatter event can not exist freely due to color confinement. The mechanism of transforming the colored partons to colorless particles is known as hadronization. The hadronization step is challenging, as is all QCD calculations, and a phenomenological approach is taken in generators such as PYTHIA 8.2 [94]. The Lund string model [95] is the basis of the phenomenological approach, where the QCD field lines can be interpreted as being compressed into a tube like structure, a string. The fragmentation in this model can be viewed as starting in a middle and spreading outwards by repeated breaks of these strings, forming new quark anti-quark pairs.

Decay

Finally, PYTHIA is also taking care of the production of resonances and the decay of these unstable particles into stable particles. All resonances are decayed sequentially as part of the hard process, and so the total cross-section as calculated by PYTHIA is dependent upon the available decay channels of the resonance, with the effect that not including a decay channel will decrease the cross-section accordingly. Conversely, particle decays are performed after hadronization, and changing the decay channels of a particle will not affect the total cross-section.

Detector simulation

Now that the simulation of the physics processes in hadron collisions have been described, one final step remains in order to be able to compare the data to the simulation, namely how the produced particles interact with the matter in the detector. GEANT4 [96] is a simulation toolkit used to describe the interaction of particles with matter and can simulate everything from tracking of particles bending in a magnetic field to the response of detector components. In GEANT4, a detailed model of CMS is implemented that is taking care of the ionization, multiple scattering, and nuclear interactions and outputs a data format that is similar to that of the collision data. It is this step of the MC simulation that is most computationally heavy, and for this reason, a faster, simpler package is used to simulate the many SUSY scenarios, with the trade off of less precision. Worth noting is that if there would be any hints of SUSY in a particular scenario, a full simulation would be performed.

Event weights

In order to be able to compare the simulated data to the real collision data, each MC sample need to be reweighted according to a so called event weight. The weight takes the form of

$$w = \frac{\sigma \cdot \mathcal{L}}{N} \quad (7-1)$$

where σ is the cross section of that process, \mathcal{L} is the integrated luminosity of the collision data the simulation is compared to, and N is the number of generated events. This weight is then multiplied to each simulated event. From the above equation, one can see that for a process with large cross section, the more simulated events are needed to keep the weight to a desirable small value. The motivation to keep the weight as small as possible, is the statistical error on the MC sample, which goes as the square root of the sum of squared event weights. So the larger the N , the smaller the w and thus the smaller the statistical error.

Pileup

As described in Chapter 5, the LHC delivers more than one pp interaction per bunch crossing in order to increase chances for successful interactions. The motivation for this originates from the PDFs discussed previously in this chapter, and the fact that most pp collisions at the LHC produce not so interesting low mass QCD interactions with little physical interest. The proton bunches that collide are in fact spread out in space, leading to the hard interaction taking place at slightly different places each time. A quantity used frequently throughout this thesis is the *primary vertex* (PV), that indicates where the hard interaction took place. This position is reconstructed using information from the tracking system, and is useful when reconstructing various physics objects.

The number of interactions per bunch crossing is described by a Poissonian distribution, and has a mean of 27 interactions during the 2016 datataking, and is depicted in Figure 7-3 [78]. This multiple pp interactions per bunch crossing, known as pileup, will be discussed in the chapter on the performance of the p_T^{miss} algorithms. The reason for this is the challenges that pileup pose on the experiment in efficiently distinguishing the PV from pileup vertices, and results in inefficiencies in the energy assignment of various physics objects. Specifically, the pileup has a significant effect on the reconstruction of the p_T^{miss} . As the multiple pp collisions per bunch crossing depends on the beam conditions, and can vary over time between LHC fills, the pileup distribution varies significantly over time. For this reason, it is difficult to properly model the pileup in simulation, which results in the necessity of a correction of the modelled number of vertices in simulation to match the distribution in data. This is known as "pileup reweighting" and is used for all simulated samples. The two searches presented in this thesis use a pileup reweighting technique that reweights the total number of true interaction in the event. The p_T^{miss} performance study presented in Chapter 9 use a reweighting

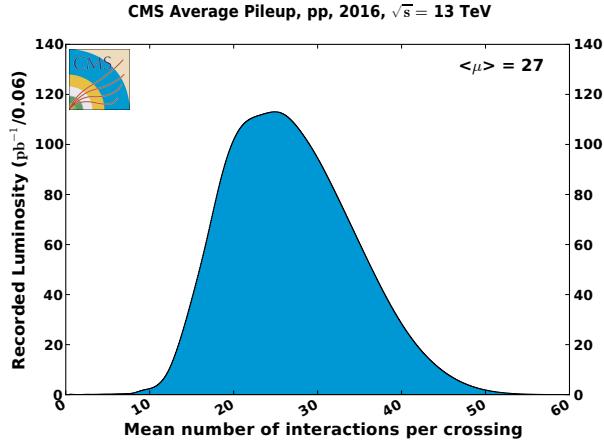


Figure 7-3. Mean number of interactions per bunch crossing for the 2016 pp run at 13 TeV. The cross section is taken to be 80 mb [78].

technique that reweights the number of reconstructed vertices. The slight mismodeling of the pileup in the simulated signal samples is taken into account as a systematic uncertainty in the signal interpretation, which will be further discussed later.

8 — PHYSICS OBJECTS

This chapter contains an overview of the reconstruction of physics objects produced in pp collisions in CMS. The Particle Flow algorithm (PF) is used in this thesis, which is an algorithm that combines subdetector information to form physics objects.

The chapter offers a top down structure that starts by describing the highest level reconstruction algorithm, the Particle Flow algorithm, and later the description on how the inputs for the PF algorithm are calculated. I have paid special attention to those objects that play a crucial role in the analysis pursued in this thesis. The particular choices used in the analysis in terms of reconstruction and identification parameters are also listed and justified.

8.1 The Particle Flow Algorithm

General purpose high energy particle physics detectors are built on the principle of layers of sub-detectors around the beam axis. The calorimeter layers and muon systems are designed for particles to get completely absorbed or produce hits in them. Intuitively, one could think that each layer would lend it self useful for reconstructing a certain type of particle: electrons and photons could be reconstructed from electromagnetic showers in the ECAL, hadrons reconstructed from hadronic showers in the HCAL, jets from combined calorimeter signals, muons from hits in the muon system. However, the idea behind the PF algorithm is to instead optimally combine the information from the different sub-detector layers. The following is a simplified description of the PF algorithm that closely follow that described in [97].

The design of the CMS detector as described in Chapter 6 has proven to be well-suited for PF reconstruction of the physics objects. Reasons behind this statement include the excellent muon detectors of CMS that provide efficient and pure identification of muons. The combination of a fine-grained tracker and a strong magnetic field effectively measures charged particle tracks that make up $\sim 65\%$ of the jet energy. Additionally, the excellent resolution and the segmentation of the ECAL provides separation of energy deposits from particles in hadronic jets and accounts for $\sim 25\%$ of the jet energy. The segmentation of the HCAL allows to differentiate charged from neutral hadrons, and subsequently the measurement of the remaining $\sim 10\%$ of the jet energy.

Summary of the PF algorithm

The PF algorithm is summarized in the following sections. The summary contains some detailed information on the reconstruction of the electrons and muons that will be covered later in this chapter.

Track extrapolation

Tracks from the pixel and strip tracker are extrapolated from the last measured hit in the tracker to the calorimeters. The track is linked to a cluster in the calorimeters according to some boundary conditions, and if several clusters are linked to a track, the cluster with the smallest distance to the track is kept. Tracks are also extrapolated to the muon detector.

Muon identification

In the second step of the algorithm muons are identified by a set of requirements on global/tracker muons, as described in Section 8.3. The tracks and clusters associated to these muons are not considered in the rest of the algorithm.

Electron/photon identification

After muons have been identified, the electron and isolated photon identification follow. Electron candidates are seeded from a track that is fitted using the Gaussian sum filter (GSF) algorithm (that will be described more in detail in Section 8.2), if there are no additional tracks linked to that ECAL cluster. Photon candidates are seeded from ECAL superclusters with E_T greater than 10 GeV and no link to a GSF track. The tracks and clusters are then removed from the remainder of the algorithm.

Hadron identification

At this point, muons, electrons and isolated photons have been identified and their corresponding tracks and clusters removed. Charged hadrons (π^\pm , K^\pm , protons etc.), neutral hadrons (K_L^0 , neutrons etc.) and non-isolated photons (from π^0 decays) are identified. Neutral hadrons and photons are identified by ECAL and HCAL clusters not linked to any track. More precisely, within the tracker volume, an ECAL cluster give rise to a photon whereas a HCAL cluster give rise to a neutral hadron. Outside of the tracker volume it is not possible to differentiate charged from neutral hadrons, and therefore ECAL and HCAL clusters that are linked are interpreted as charged or neutral hadrons, whereas only an ECAL cluster is identified as a photon.

Remaining identification

The remaining HCAL clusters can be linked to the remaining tracks (that are not already linked to another HCAL cluster), that subsequently can be linked to the remaining ECAL clusters. The total energy of the ECAL and HCAL clusters is then used to determine the *calibrated calorimetric energy* under a single charged hadron hypothesis. The sum of the track momenta is also computed and compared to the calibrated calorimetric energy. If the calibrated calorimetric energy is larger than the sum of the track momenta, this can be interpreted as a photon or a neutral hadron. If the calibrated calorimetric energy is compatible with the sum of the track momenta, this can be interpreted as a charged hadron. If the calibrated calorimetric energy is smaller than the sum of the track momenta, this can be interpreted as a muon.

To summarize; electrons, muons, photons, charged hadrons and neutral hadrons have been identified by the PF algorithm. These *PF candidates* are subsequently used to reconstruct jets, p_T^{miss} and τ lepton candidates.

Once the PF candidates, and higher level objects such as jets, have been reconstructed, these can be used to select candidate events in the search for SUSY. As has been mentioned in the Chapter 4, two leptons of opposite charge and same flavor are expected in all the studied signal processes, either from the decay of an on-shell or off-shell Z boson, through a pair production of sleptons, or through a decay chain containing a slepton. There are several SM background processes that can result in opposite sign same flavor leptons, and common to these are that they include a decay of a Z boson, virtual photon or a W boson. Although these bosons decay democratically to each lepton generation, only the first and second lepton generations are considered in this thesis. In this work, dielectron (ee) and dimuon ($\mu\mu$) events are categorized as same flavor (SF) and events containing an electron and a muon ($e\mu$) are categorized as opposite flavor (OF). The reason to exclude the third lepton generation, the τ^\pm leptons and their corresponding ν_τ 's, is due to the challenging final state. The τ lepton, being the most massive of the leptons, decay hadronically $\sim 65\%$ of the time. As an effect, including this decay mode in the various signal scenarios would result in the introduction of overwhelming backgrounds from multijet QCD processes, that in turn would not contribute to improve the sensitivity of the searches. However, the search for direct stau production in semi-leptonic and hadronic final states has been performed by CMS [98, 99], where the challenging background due to τ leptons is taken into account. In the following chapters, the reconstruction, identification and isolation of electrons and muons are presented.

8.2 Electrons

Electrons are reconstructed by associating ECAL clusters with a track reconstructed in the silicon detector. The following contains a description of the clustering of the ECAL deposits, the electron track and the combination of the two, and is roughly based on [100]. Finally, the identification and isolation requirements imposed on the electrons to be categorized as signal events or used for background estimation techniques are presented.

Electron reconstruction

Bremsstrahlung causes electrons to lose energy when interacting with detector material. Depending on the amount of tracker material before the ECAL, the electrons can lose more or less energy due to radiation of photons. For this reason, the energy of the measured electrons will be underestimated if one does not take into account the radiated photons.

Clustering

Clusters of ECAL crystals are constructed for the above reason, with the aim to collect the energy deposits of the crystals surrounding a seed crystal. Two clustering algorithms are deployed that group crystals together into supercrystals (SC) following a set of threshold criteria on the E_T of the seeded crystal, and taking into account the differing crystal layouts of the ECAL barrel and endcap. Additionally, PF clusters are reconstructed by aggregating crystals surrounding a seed, following thresholds set on the electronic noise levels.

Seeding

The electron seed, the first step in the electron track reconstruction, can either be an ECAL-based seed or a tracker-based seed. As the naming suggests, the former seed is based on information from the SC that is extrapolated back to the vertex. The latter is a part of the PF reconstruction algorithm. The tracker-based procedure has advantages over the ECAL-based procedure for low p_T electrons. As these electrons bend significantly in the magnetic field, the radiated photons are spread out and can not all be contained in a SC. This results in the extrapolation of the track starting from the SC in the ECAL-based approach to be sub-optimal, as not all radiated energy is contained in it. Additionally, for electrons in jets, the SC position and energy is affected by the other particle contributions, making the ECAL-based approach inefficient. To determine the tracker-based seed, the Kalman Filter (KF) track reconstruction can be used, but is sub-optimal when there is radiative losses of the electrons in the tracker material. In the cases where a lot of bremsstrahlung took place, the KF algorithm will not be able to follow the change in the curvature of the electron trajectory as a result of the energy lost. As a result, the hits can not be collected, or they are of bad quality. To recover the trajectory of these bremsstrahlung electron, the KF tracks are refitted using the Gaussian sum filter (GSF) algorithm. The GSF algorithm takes into account that the bremsstrahlung electron energy loss is non-Gaussian, and model the bremsstrahlung energy loss by weighted sums of Gaussians instead of just one Gaussian used in the KF algorithm. The χ^2 of the KF and GSF fitting, together with energy matching of the ECAL and tracker are used in an MVA to select the tracker seed as the electron seed.

Tracking

The ECAL-based seed and tracker-based seed are merged into a unique collection and are submitted to the full electron tracking with twelve GSF components. The track is built using the combinatorial KF method which starts at the electron seed and proceeds iteratively in the next layer of the tracker, and the hits are collected.

Track and cluster association

Electron candidates are constructed by associating a GSF track to an ECAL cluster. ECAL-seeded electrons are now by construction associated to the ECAL cluster used to determine the ECAL-based seed. The tracker-seeded electrons are associated to the aforementioned PF clusters.

Electron identification

The rationale behind the electron identification is to efficiently differentiate prompt electrons from the PV, electrons from photon conversions, jets misidentified as electrons, and electrons from b-quark decays. As there are many discriminating variables available, the electron identification can be performed by either a sequential set of selections on these variables, or combine them in a MVA for improved discrimination. The electron identification used in these searches use the MVA identification technique, and below follow a description of the variables used in the analysis.

$\sigma_\eta \sigma_\eta$

The shower shape variable $\sigma_\eta \sigma_\eta$ is a measure of the spread of the electron shower along the η direction. It is defined as $\sigma_\eta \sigma_\eta = \frac{\sum (\eta_i - \bar{\eta})^2 w_i}{\sum w_i}$ where the sums run over the 5×5 matrix of crystals around the highest E_T crystal of SC, and w_i denotes a weight that is logarithmically dependent on the contained energy. This variable exhibits a sharp peak for prompt electrons, whereas for misidentified electrons it is more spread, making it an ideal discrimination variable.

$$|\Delta\eta| = |\eta_{SC} - \eta_{in}^{extrap}|$$

This quantity denotes the separation between the SC energy-weighted η position and the track η extrapolated from the innermost track position and direction to the point of closest approach (PCA) to the SC. The $|\Delta\eta|$ increases with the amount of bremsstrahlung.

$$|\Delta\phi| = |\phi_{SC} - \phi_{in}^{extrap}|$$

The azimuthal separation between the SC energy-weighted ϕ position and the track ϕ extrapolated from the innermost track position and direction to the PCA to the SC.

H/E

The ratio H/E , where E is the energy of the SC, and H the sum of the HCAL tower energies within $\Delta R = \sqrt{((\Delta\eta)^2 + (\Delta\phi)^2)} = 0.15$, is used to estimate the energy leakage into the HCAL. A well-identified electron would be expected to have a low H/E owing to the high X_0 of the CMS detector, thereby containing the EM showering before it reaches the HCAL. Hadron fakes instead exhibit a larger H/E .

$|1/E - 1/p|$

This quantity expresses an energy-momentum matching requirement using the SC energy, E , and the track momentum, p , at the PCA to the track vertex. The requirement helps reject backgrounds from hadronic activity where the spread of the E is not localized, resulting in a low E/p . It is also efficient in rejecting backgrounds where a $\pi^0 \rightarrow \gamma\gamma$ decay occurs in the close vicinity of a charged hadron, resulting in a very high E/p ratio.

d_{xy} and d_z

The transverse and longitudinal distance between the electron track and the PV. The smaller the d_{xy} and d_z , the more likely the electron is of coming from the hard interaction.

Missing hits

Missing hits are used to suppress electrons from photon conversions. As photon conversions take place close to the beampipe or the pixel system, missing hits can occur due to the large change in curvature of the electron trajectory. At most, one missing hit is allowed for an accepted trajectory candidate. Additionally, to avoid including hits from converted bremsstrahlung photons in the reconstruction of primary electron tracks, an increased χ^2 penalty is applied to trajectory candidates with one missing hit.

Conversion veto

In order to reject secondary electrons produced in the conversion of photons in the tracker material, a vertexing algorithm is used. The hits in the tracker from the converted photon are fit to a common vertex using the well-defined topological constraint that tracks from conversions have virtually the same tangent at the conversion vertex in both the (r, ϕ) and (r, z) planes. The converted photon candidates are rejected according to the χ^2 probability of the fit.

Electron isolation

An isolation variable, the mini-isolation is used which features a shrinking cone-size with increasing p_T of the lepton. Thus, the cone size in which the PF particles are summed to calculate the relative isolation is no longer constant, but a function of the p_T of the lepton

$$R = \frac{10.}{\min[\max(p_T, 50), 200]} .$$

For p_T values below 50 GeV, this leads to a constant cone size of 0.2. For p_T values between 50 GeV and 200 GeV, the cone size shrinks from 0.2 to 0.05 at which it remains for higher p_T leptons. Corrections to the isolation are applied by subtracting the average energy density ρ from the effective geometrical area of the lepton's isolation cone. The variable cone size is taken into account for this correction.

Table 8-1 summarizes the identification, isolation and impact parameter (IP)¹ criteria imposed on the electrons in order for them to be considered in the analysis. The MVA trained discriminator is optimized on electrons from prompt W-boson decays in $t\bar{t}$ versus leptons stemming from so-called “fakes”² in $t\bar{t}$. The working point used for the MVA identification discriminator corresponds to a “tight” value, developed at the end of 2016. The actual MVA cut value depends on the lepton p_T and $|\eta|$. In each $|\eta|$ bin the lower value is used for electrons with $p_T > 25$ GeV while the cut decreases linearly from the upper to the lower value for p_T between 15 and 25 GeV. Additionally, conversion rejection cuts are applied.

Table 8-1. Electron selection criteria.

cut	value
Identification	
MVA	Tight 2016 Working point
Conversion rejection	
number of lost hits	0
conversion veto	pass
Isolation	
mini Isolation	< 0.1
Impact parameter	
d_{xy}	< 5 mm
d_z	< 10 mm
SIP3D	< 8

8.3 Muons

The calorimeters of CMS are efficiently stopping electrons, photons, charged and neutral hadrons, resulting in muons (and of course neutrinos) being the only particles reaching the muon systems. The muon track reconstruction and identification described below is based on [101].

¹ The impact parameter is defined as the distance between the primary vertex and the tracks at their closest point of approach.

² “Fake” leptons are mostly semi-leptonic b-quark decays where $b \rightarrow cW \rightarrow c\ell\nu$

Muon reconstruction

Inner tracks and tracks in the muon systems are used as input for the muon-track reconstruction, that can be categorized in the following three ways:

Standalone-muon track

DT and CSC hits are clustered and the track segments formed are used as seeds for the fitting. DT, CSC and RPC hits are used in the final fitting to reconstruct a standalone-muon track.

Tracker-muon track

An “inside-out” approach is used for the tracker-muon tracks. If an inner track has $p_T > 0.5 \text{ GeV}$ and $p > 2.5 \text{ GeV}$, then it is extrapolated to the muon system. If the extrapolated track and at least one track in a muon segment have an absolute difference in their position in x -coordinates less than 3 cm or the ratio of this distance to the uncertainty is less than 4, then tracks are considered to be matched.

Global-muon track

In contrast to the tracker-muon track, an “outside-in” approach is used for global muon tracks. A standalone muon track is matched to an inner track, if the position and momentum of the two tracks are compatible. The two tracks are combined and fitted using KF technique to form a global muon track. The global muon track procedure improves the momentum resolution for muons of $p_T > 200 \text{ GeV}$ compared to a tracker-muon track only fit. This is because large p_T muons have a higher probability to reach more than one muon segment, which is required for the global muon reconstruction. Lower p_T muons benefit from the tracker-muon reconstruction as only one muon segment is required in the matching. However, most muons are reconstructed as global muons, tracker muons or as both. The standalone muon tracks have worse momentum resolution and can more often pick up cosmic muons.

Muon identification

From an experimental point of view, muons are much easier to measure than electrons. As the radiated power due to Bremsstrahlung is much lower for muons than for electrons³ there are less ambiguities in the muon reconstruction than in the electron reconstruction. Additionally, the unlikely process of photon conversion to two muons make a conversion veto redundant. Instead, the different variables used for the muon identification is listed below:

³ The radiated power due to Bremsstrahlung goes as m^{-4} , meaning that it is suppressed for more massive particles such as muons and protons.

Reconstruction type

A muon can be identified using the various muon-track reconstruction categorization, i.e. require standalone muon track, tracker-muon track, or global-muon track. PF muon identification requirement can also be used imposed.

Fraction of valid tracker hits

Fraction of hits from inner tracker layers that the muon traverses.

Segment compatibility

The muon segment compatibility is computed by propagating the tracker track to the muon system, and evaluating both the number of matched segments in all stations and the closeness of the matching in position and direction

Kink-finding

A kink-finding algorithm splits the tracker track into two separate tracks at several places along the trajectory. For each split the algorithm makes a comparison between the two separate tracks, with a large χ^2 indicating that the two tracks are incompatible with being a single track.

Position match

The tracker and standalone tracks are matched according to their position.

Normalized χ^2

Requirement on the global fit required to have goodness-of-fit per degree of freedom.

 d_{xy} and d_z

Requirements on the transverse and longitudinal distance between the tracker track and the location of the PV. This cut is applied to suppress events with cosmic muons, tracks from pile-up or muons originating from in-flight decays.

SIP_{3D}

Requirements on the 3D significance of the IP is further used to identify prompt muons.

Muon isolation

The isolation variable used for electrons, the mini-isolation, is also used for muons.

There are several identification categorizations developed in CMS for various analysis needs, with working points aimed to more or less efficiently suppress fakes from punch through hadrons while keeping a high muon identification efficiency. In this work, the medium muon ID is used, which is optimized for prompt muons and for muons from heavy flavor decay. The definition of the medium muon ID is summarized in Table 8-2, along with the muon isolation requirements and IP selection criteria.

Table 8-2. Muon selection criteria.

cut	value
Loose Muon ID	
PF Muon ID	True
Is Global OR Tracker Muon	True
Medium Muon ID	
Loose Muon ID	True
Global Muon	True
Fraction of valid tracker hits	> 0.8
Normalized global-track χ^2	< 3
Tracker-Standalone position match χ^2	< 12
Kick-finder	< 20
Tight Segment compatibility	> 0.451
Isolation	
mini isolation	< 0.2
Impact parameter	
d_{xy}	< 5 mm
d_z	< 10 mm
SIP3D	< 8

8.4 Jets

The PF candidates found through the PF algorithm described in Section 8.1 are used as input in a clustering algorithm for hadronic jets. A variety of clustering algorithms are on the market, with different optimizations for high or low p_T jets or different shapes, with different computation speeds. In the following the ingredients needed for well-measured jets are described, starting with pileup mitigation, followed by the sequential jet clustering algorithms, and finally the identification and corrections needed.

Charged hadron subtraction

The clustering algorithms are efficient in forming jets while keeping the contamination of effects from pileup and underlying events at a minimum. During the clustering, there is a first line of defense against pileup called Charged Hadron Subtraction (CHS). CHS is a type of particle-by-particle pileup subtraction performed on the PF candidates, that assigns charged hadrons to the primary vertex or pileup vertices, using tracking information. Charged hadrons clearly associated to a pileup vertex are removed.

Jet clustering algorithms

The clustering algorithm most commonly used today is the so-called anti- k_t algorithm [102], which is a member of a class of sequential recombination algorithms used at hadron colliders. The idea behind these sequential clustering algorithms is to utilize the distance d_{ij} between particles i and j , the distance d_{iB} between particle i and the beam B and the transverse momenta of the particles $k_{ti,j}$ when forming a jet. The distance parameters d_{ij} and d_{iB} are defined as:

$$d_{ij} = \min(k_{ti}^{2a}, k_{tj}^{2a}) \times \frac{(\Delta_{ij}^2)}{R^2}, \quad (8-1)$$

$$d_{iB} = k_{Ti}^a, \quad (8-2)$$

where $\Delta_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$, η and ϕ the pseudorapidity and azimuth distance between the particles, and R the radius parameter of the jet cone. The parameter a determines the type of clustering algorithm. The algorithm follows these steps: compute all distances d_{ij} and d_{iB} , and find the smallest one. If smallest is a d_{ij} , combine (sum four momenta) the two particles i and j , update distances, proceed to find the next smallest distance. If smallest is a d_{iB} , remove particle i and call it a jet. The following steps are repeated until all particles are clustered into jets.

The predecessors to the anti- k_t algorithms are recovered for $a = 1$ (k_t algorithm [103]) or $a = 0$ (Cambridge/Aachen (CA) algorithm [104]). Setting the $a = -1$ is the basis of the anti- k_t algorithm. The choice of a negative a is motivated by the following reasoning. Considering an event with a few hard particles and many soft particles. The distance parameter d_{i1} between the hard particle 1 and the various soft particles is dominated by the p_T of the hard particle, as a result of the negative exponent that makes the d_{ij} small for a large p_T hard particle 1, and larger for equally separated soft particles. As a result, the soft particles will more likely cluster with hard ones before they cluster with another soft particle. A further consequence is that soft particles do not modify the shape of the jet whereas hard ones do, meaning that the jet boundary is resilient to soft radiation. This is a desired feature, also known as infrared and collinear (IRC) safety. Namely that neither soft emissions or collinear splittings should not change the boundary of jets. Comparing the shape of jets clustered with CA, k_t and anti- k_t algorithms respectively, the anti- k_t algorithm results in more conical jets. In this work, jets are clustered with the anti- k_t algorithm with a radius parameter of 0.4 (AK4).

Jet identification

Identification criteria are imposed on the PF jets in order to suppress noise contributions from the calorimeters. The selections are based on relative energy fractions carried by the PF candidates with respect to their total jet energy, and the number of PF candidates in a jet. When tracking information is available ($|\eta| < 2.4$) additional cuts are applied on the charged candidates. The Loose Jet ID working point is used in this work and its definition is summarized in Table 8-3.

Table 8-3. Definition of Loose Jet ID working point.

Variable	$ \eta \leq 2.7$	$2.7 < \eta \leq 3$	$ \eta > 3$
Neutral hadron fraction	< 0.99	< 0.98	-
Neutral EM fraction	< 0.99	> 0.01	< 0.90
Number of constituents	> 1	-	-
Number of neutrals	-	> 2	> 10
Additional cuts for $ \eta < 2.4$			
Charged hadron fraction	> 0	> 0	> 0
Charged multiplicity	> 0	> 0	> 0
Charged EM fraction	< 0.99	> 0.99	< 0.90

Jet calibration

As jets are complex objects consisting of highly energetic quarks and gluons that built on input from several subdetector, the correct energy assignment to jets is a challenge. Therefore, a set of jet energy corrections (JECs) are determined to account for various effects and applied to jets in data and simulation. There are three different types of corrections, which are listed below and follow the description outlined in [105].

Pileup corrections

The so-called pileup offset corrections or L1 corrections are determined to reduce the effect of pileup on the jet energy. The corrections are determined in a simulated QCD dijet events, with and without pileup overlayed, before and after CHS applied. The corrections are parametrized as a function of a set of jet related quantities such as the jet area A_i , jet pseudorapidity η , jet p_T and the energy density ρ .⁴

⁴ The energy density ρ is defined as the median of the transverse momentum of the jets over their area, $\rho = \text{median}(p_{T,i}/A_i)$

MC-truth corrections

The MC-truth corrections, also known as L2L3 MC-truth corrections, are based on the comparison of a particle level jet and its corresponding reconstructed version. The corrections are derived from a QCD dijet simulated sample as a function of jet p_T and η .

Residual corrections

The residual corrections, or L2L3 correction, are corrections to be applied on data only. The relative, or η dependent, correction is determined with QCD dijet events, where a tag jet in the barrel region is compared to a probe jet with no η restriction of similar p_T . The absolute correction is determined in γ +jets, $Z(ee/\mu\mu)$ +jets or multijet events, where a well-measured γ , Z boson or jet is compared to the recoiling jet.

8.5 *b*-tagged jets

Heavy flavor jet identification algorithms are used in CMS as a means to select events containing top quark decays or SM bosons decaying through b -quarks, and the following description of the algorithms follow that introduced in [106]. Recalling the CKM matrix introduced in Section 2.2, the top quark decay almost exclusively to bottom quarks. The bottom quark, on the other hand, have less options due to the suppression embedded in the V_{cb} and V_{ub} elements of the CKM matrix. The suppressed transition results in the relatively long lifetime ($\sim 10^{-12}$ s) and consequently a non negligible flight distance (few millimeters). As a result, this displacement leads to the possibility to reconstruct an additional vertex where the b hadron decay takes place. This makes for a perfect experimental handle for identifying b -quark jets, known as “ b -tagging”.

The additional vertex where the decay takes place is called the secondary vertex (SV), and is reconstructed from the displaced tracks that are characterized by their IP. The SV reconstruction algorithm used in CMS in the LHC Run 2 is the inclusive vertex finder (IVF) algorithm. This algorithm uses as input all reconstructed tracks in the event with $p_T \geq 0.8$ GeV and with longitudinal IP ≤ 0.3 cm. The IVF algorithm is seeded by tracks with a three dimensional IP greater than $50\text{ }\mu\text{m}$ and the significance of the two dimensional IP (IP divided by its uncertainty) greater than 1.2. After this, the track clustering is performed by associating the seed track to any other track by imposing requirements on the distance at the point of closest approach and the angle between the tracks. The resulting track clusters are fitted using an adaptive vertex fitter. To resolve the track ambiguity when a track can be both associated to the PV and a SV, a track is discarded from the SV if it is more compatible with the PV. Following this track arbitration, the SV position is refitted. After the IVF algorithm is performed, the vertices found are used as input to the combined secondary vertex tagger version 2 (CSVv2), where in addition to the IVF vertex, two tracks per jet is required separated from the jet axis by ΔR less than 0.3.

The CSVv2 algorithm contains a step of training that is performed in the three independent vertex categories listed below.

- RecoVertex: a jet that contains more than one SV.
- PseudoVertex: no SV is found but at least two tracks with 2D IP significance larger than 2 and the combined invariant mass not compatible with a K_S^0 .
- NoVertex: when none of the above categories are filled.

The variables combined in the algorithms are:

- SV 2D flight distance significance: the significance of the 2D flight distance of the SV with least uncertainty on its flight distance for jets in the RecoVertex category.
- Number of SVs: the number of SVs for jets in the RecoVertex category.
- Track η_{rel} : the track η relative to the jet axis for the track with the highest 2D IP significance for jets in the RecoVertex and PseudoVertex categories.
- Corrected SV mass: the corrected mass of the SV with the smallest uncertainty on its flight distance for jets in the RecoVertex category or the invariant mass obtained from the total summed four-momentum vector of the selected tracks for jets in the PseudoVertex category.
- Number of tracks from SV: the number of tracks associated with the SV for jets in the RecoVertex category or the number of selected tracks for jets in the PseudoVertex category.
- SV energy ratio: the energy of the SV with least uncertainty on its flight distance over the energy of the total summed four momentum vector of the selected tracks.
- $\Delta R(SV, jet)$: defined as the distance between the flight direction of the SV with least uncertainty on its flight distance and the jet axis for jets in the RecoVertex category, or the distance between the total summed four-momentum vector of the selected tracks for jets in the PseudoVertex category.
- 3D IP significance of first four tracks: the 3D IP significances of the four tracks with the highest 2D IP significance.
- Track $p_{T,rel}$: the track p_T relative to the jet axis, i.e. the track momentum perpendicular to the jet axis, for the track with the highest 2D IP significance.
- $\Delta R(track, jet)$: the distance between the track and the jet axis for the track with the highest 2D IP significance.
- Track $p_{T,rel}$ ratio: the track p_T relative to the jet axis divided by the magnitude of the track momentum vector for the track with the highest 2D IP significance.
- Track distance: the distance between the track and the jet axis at their point of closest approach for the track with the highest 2D IP significance.
- Track decay length: the distance between the PV and the track at the point of closest approach between the track and the jet axis for the track with the highest 2D IP significance.
- Summed tracks E_T ratio: the E_T of the total summed four-momentum vector of the selected tracks divided by the transverse energy of the jet.
- $\Delta R(\text{summed tracks}, \text{jet})$: the ΔR between the total summed four momentum vector of the tracks and the jet axis.
- The number of selected tracks.
- The jet p_T and η .

The above discriminating variables are combined in a neural network and the final discriminator is shown in Figure 8-1.

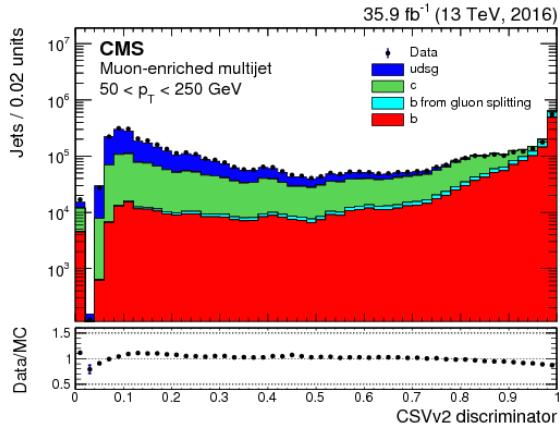


Figure 8-1. The CSVv2 discriminator for jets in a muon-enriched multijet sample. The simulated contributions of each jet flavor are shown with different colors [106].

8.6 Isolated tracks

Isolated tracks will be used in this thesis as a means to improve the signal selection efficiency and background rejection. The idea is to remove events that seem to contain more than two leptons. Instead of removing events with a well-identified third lepton, a looser criteria can be imposed, in the form of an isolated track. These tracks are defined using charged PF candidates with different requirements depending on the flavor. PF electrons and PF muons, are required to pass $p_T > 5 \text{ GeV}$, $|dz| < 0.1 \text{ cm}$, as well as being well associated to the PV, and track isolation cuts of $\text{isolation}/p_T < 0.2$ and $\text{isolation} < 8 \text{ GeV}$. The track isolation sum is computed from all charged PF candidates within a cone of $\Delta R < 0.3$, requiring them to pass $|dz| < 0.1 \text{ cm}$ with respect to the PV. Charged PF hadrons are required to pass $p_T > 10 \text{ GeV}$, $|dz| < 0.1 \text{ cm}$, be well associated to the PV, and track isolation cuts of $\text{isolation}/p_T < 0.1$ and $\text{isolation} < 8 \text{ GeV}$. The track isolation is computed in the same way as for PF leptons above.

Part III

SEARCH METHODOLOGY

9 — Missing transverse momentum

At the forefront of the SUSY physics program are searches where R-parity is conserved, resulting in one or more lightest SUSY particles that have no SM particles to decay into. The result is LSPs that are unable to interact with the detector material and thus escape detection. The existence of such particles can be inferred by the momentum imbalance in the transverse plane, \vec{p}_T^{miss} , with its magnitude denoted p_T^{miss} . In R -parity conserving SUSY models, the p_T^{miss} provides an excellent search tool for SUSY. But other sources can contribute to a large momentum imbalance. Any process with a leptonically decaying W-boson produces a neutrino that escape the detector similarly as the LSP. Additionally, as jets are complex objects to measure, and their energy are corrected through JECs, any over- or under-measurement in the jet energy will result in a false imbalance which would register as p_T^{miss} . In order to perform a SUSY search where R -parity is conserved, a deep understanding of the p_T^{miss} object is needed to distinguish the p_T^{miss} originating from LSPs from SM processes. A challenge for reconstructing physics objects is differentiating tracks from the primary vertex with tracks from overlapping bunch crossings in multiple pp collisions (pileup).

The SUSY searches presented in this thesis contain all the above listed sources of p_T^{miss} . All models are R -parity conserving, thus providing an LSP. The SM backgrounds contains large p_T^{miss} stemming from neutrinos from leptonically decaying $t\bar{t}$ and due to mismeasured jets in Z+jets events. In order to be able to use this discriminating variable to differentiate the SUSY signal from other sources of p_T^{miss} , a detailed study is performed to assess the performance of the p_T^{miss} reconstruction algorithm used in this thesis. The performance of the p_T^{miss} algorithm using a novel pileup mitigating technique, PUPPI [107], is also evaluated.

This chapter presents a detailed study of the performance of two commonly used p_T^{miss} reconstruction algorithms that was originally published in [61]. This chapter highlights the contributions I made to this publication, and the important connection a robust p_T^{miss} algorithm has in the SUSY searches presented in this thesis. The chapter concludes with a specific study on the performance of the algorithms under extreme pileup conditions, as is expected in the High Luminosity phase of the LHC.

9.1 Performance of p_T^{miss} reconstruction algorithms

In collision events, the transverse momentum of the partons is small compared to the energy available in the center-of-mass, and does not depend on their longitudinal energy. Thus, an assumption can be made, that the initial transverse momentum of the system formed by the partons is zero. As an effect of this assumption, if particles escape detection, a transverse momentum inequilibrium is created. The final states containing one or more neutrinos therefore result in a significant missing energy corresponding to the vectorial sum of the neutrino momenta. When no neutrinos are created in the event, all missing transverse momentum is due to detector inefficiencies and reconstruction issues. Therefore, final states without neutrinos are ideal for the study of the performance of the missing transverse momentum reconstruction algorithms originating from detector effects. The p_T^{miss} is defined as the magnitude of the negative vectorial sum of the particles in the event

$$p_T^{\text{miss}} = |\vec{p}_T^{\text{miss}}| = \left| - \sum \vec{p}_T \right| \quad (9-1)$$

In CMS, two algorithms for the p_T^{miss} reconstruction are used, which will be presented in the following section. The first one is Particle Flow p_T^{miss} (PF p_T^{miss}) and the second one uses a pileup mitigation technique, 'PileUp Per Particle Identification' (PUPPI) to reconstruct the p_T^{miss} .

Particle Flow p_T^{miss} reconstruction

The first reconstruction algorithm is PF p_T^{miss} , which is the magnitude of the negative of the vectorial sum of all PF candidates in an event:

$$\text{PF } p_T^{\text{miss}} = \left| - \sum_{i \in \text{PF}} \vec{p}_{T,i} \right| \quad (9-2)$$

As will be shown in the following, the PF p_T^{miss} algorithm is highly performant and is therefore used in the majority of CMS analyses.

PUPPI p_T^{miss} reconstruction

The PUPPI p_T^{miss} algorithm uses the 'pileup per particle identification' method [107]. This method has been developed to reduce the dependence of pileup on physics objects.

PUPPI algorithm

This section contains a summary of the PUPPI algorithm, which follows that documented in [107]. The idea is to estimate how likely the PF candidates are to be originating from pileup, and reweight the particle four-momentum accordingly with a weight, w_i , close to 1 if the candidate is from the hard scatter and close to 0 for particles from pileup. The procedure to calculate the w_i starts with defining a value α_i for each particle,

$$\alpha_i = \log \sum_{\substack{j \in \text{event} \\ j \neq i}} \left(\frac{p_{T,j}}{\Delta R_{ij}} \right)^2 \times \Theta(\Delta R_{ij} - R_{min}) \times \Theta(R_0 - \Delta R_{ij}), \quad (9-3)$$

where Θ is the Heaviside step function. The α of the i -th particle is thus depending on the p_T of the surrounding particles, and the distance between them in $\eta - \phi$ space, defined as the cone ΔR_{ij} . Only particles within some R_0 around particle i are considered. Surrounding particles j are discarded that are within some minimum radius R_{min} close to the particle i , to reduce the effect from collinear splittings. When the particle i is from the hard scatter, the surrounding particles tend to be close in ΔR because of the collinear singularity of the parton shower, resulting in a relatively larger α_i . On the other hand, a wider separation in ΔR is expected for particles originating from pileup, as they should have no correlation with the direction of particle i , resulting in a smaller value for α_i . The p_T of the j -th particles is also used in the calculation of α_i , and the

characteristic of this variable is that it is generally softer for particles originating from pileup, yielding the desired smaller value of α_i , and the opposite for when the particle is from the hard scatter.

Now that the α_i is defined, the question of what particles should be summed over arises. For this, two regions are used, reflecting the design of the detector; the central region ($|\eta| \leq 2.4$), in which the tracking can distinguish charged tracks from the primary vertex from charged tracks from pileup vertices, and the forward region ($|\eta| > 2.4$) where this discrimination is not possible. Where tracking is available, the PF algorithm can provide the following PF candidates; neutral particles, charged hadrons from the primary vertex and charged hadrons from pileup vertices. This results in two different computations of α_i , one in the central region where tracking is available α_i^C and one in the forward region α_i^F , namely

$$\alpha_i^C = \log \sum_{\substack{j \in \text{Ch, PV} \\ j \neq i}} \left(\frac{p_{T,j}}{\Delta R_{ij}} \right) \times \Theta(\Delta R_{ij} - R_{min}) \times \Theta(R_0 - \Delta R_{ij}), \quad (9-4)$$

$$\alpha_i^F = \log \sum_{\substack{j \in \text{event} \\ j \neq i}} \left(\frac{p_{T,j}}{\Delta R_{ij}} \right) \times \Theta(\Delta R_{ij} - R_{min}) \times \Theta(R_0 - \Delta R_{ij}). \quad (9-5)$$

The α_i^C is defined with sum over all charged hadrons from the PV, whereas the α_i^F is the sum over all particles in the event, since no tracking can be used to differentiate that hadrons from pileup or the PV. The difference between these two computations is that in the central case, a particle j originating from pileup is discarded from the event, whereas this distinction can not be done in the forward region. The w_i is a weight that is used to rescale the four momentum of particle i , and is calculated using the α_i as an input. The calculation of w_i is done by introducing the following quantity:

$$\chi_i^2 = \Theta(\alpha_i - \bar{\alpha}_{\text{PU}}) \times \frac{(\alpha_i - \bar{\alpha}_{\text{PU}})^2}{\sigma_{\text{PU}}^2} \quad (9-6)$$

where the $\bar{\alpha}_{\text{PU}}$ and σ_{PU} are used to characterize the distributions on an event-by-event basis, and defined as

$$\bar{\alpha}_{\text{PU}}^\eta = \text{median}\{\alpha_{i \in \text{Ch, PU}}^\eta\} \quad (9-7)$$

and

$$\sigma_{\text{PU}}^\eta = \text{RMS}\{\alpha_{i \in \text{Ch, PU}}^\eta\}. \quad (9-8)$$

The super script η is C for central or F for forward regions, indicating what region is used for the computation. As can be seen in Equation 9-6, the χ_i^2 distribution quantifies how much the α_i value is fluctuating from the pileup median $\bar{\alpha}_{\text{PU}}$. Any value of α_i below the $\bar{\alpha}_{\text{PU}}$ is considered pileup like, and due to the definition involving the Heaviside function, these values will result in a χ_i^2 of 0. Conversely, large values of α_i that are far from the $\bar{\alpha}_{\text{PU}}$ will result in a large χ_i^2 . Finally, the w_i is defined by

$$w_i = F_{\chi^2, \text{NDF}=1}(\chi_i^2) \quad (9-9)$$

with $F_{\chi^2, \text{NDF}=1}$ being the cumulative distribution function of the χ^2 distribution. As a result, whenever the χ_i^2 is 0, the final w_i is 0, whenever the χ_i^2 is large, the final w_i is 1, and all values of χ_i^2 in between results in a fractional weight between 0 and 1.

PUPPI p_T^{miss} reconstruction

Now that the PUPPI particle weights have been defined, the actual rescaling of the particles and how they enter the p_T^{miss} calculation will be covered. For each event, the value of α_i^η is computed for all charged pileup candidates, and the corresponding median and RMS distributions $\bar{\alpha}_{\text{PU}}^\eta$ and $\bar{\sigma}_{\text{PU}}^\eta$. Where tracking is available, i.e. $|\eta| \leq 2.4$, the particles originating from pileup can be easily distinguished, and those receive a weight of 0, and are completely discarded in the remainder of the calculation, whereas the charged hadrons from the primary vertex receive a weight of 1. The weights w_i of the remaining particles are calculated, and the four-momenta of these particles is rescaled by this w_i . The charged hadrons from the primary vertex and the rescaled remaining particles used to reinterpret the event, in the context of jet clustering, or in this case, the p_T^{miss} calculation, according to:

$$\text{PUPPI } \vec{p}_T^{\text{miss}} = - \sum_{i \in PF} w_i \times \vec{p}_{T,i} \quad (9-10)$$

9.2 Calibration of p_T^{miss}

As the p_T^{miss} reconstruction is dependent on the accurate measurement of all the reconstructed physics objects, any inefficiency in the reconstruction or minimum energy or p_T thresholds will bias the energy scale of the p_T^{miss} . As described in Section 8.4, the energy of the jets are corrected with JECs. If these corrections are not taken into account in the computation of the p_T^{miss} , there will be a significant bias and imbalance in the event. Therefore, the energy scale of p_T^{miss} is improved by propagating the correction of the p_T of the jets, $\vec{p}_{T,\text{jet}}^{\text{corr}}$ to p_T^{miss} in the following way:

$$\text{Type I } \vec{p}_T^{\text{miss}} = \vec{p}_T^{\text{miss}} - \sum_{\text{jets}} (\vec{p}_{T,\text{jet}}^{\text{corr}} - \vec{p}_{T,\text{jet}}) \quad (9-11)$$

In the rest of the thesis, the “Raw p_T^{miss} ” is the uncorrected p_T^{miss} , and the corrected p_T^{miss} , commonly known as the “Type-I p_T^{miss} ”, will be referred to as just p_T^{miss} . As jets are complex objects to measure, the choice of the jets whose corrections should be taken into account is a question of optimization, and subject to future improvement. The motivation behind the 15 GeV p_T threshold of the jets is to reduce the contribution of jets from pileup. As will be shown in the next section, this choice of the p_T threshold gives a response very close to unity.

Further ambiguity can arise if a jet is very close to a reconstructed muon, or resembles an electron or a photon. If a muon reconstructed using the outer tracking system overlaps with a jet, its four-momentum is subtracted from the four-momentum of the jet, and the JECs appropriate for the modified jet momentum is used in the p_T^{miss} calculation. Jets are reconstructed from energy deposited in both the HCAL and ECAL, with various fractions of the energy in each calorimeter. In order to not correct jets that are in fact an electron or a photon, i.e. with a large electromagnetic (EM) energy fraction, a choice is made to only correct jets with an EM fraction of less than 90%. The choice of 90% has been providing a well calibrated p_T^{miss} object but can be improved, as will be shown in the next section. The p_T^{miss} relies on the accurate measurement of the reconstructed physics objects, namely muons, electrons, photons, hadronically decaying τ leptons, jets, and unclustered energy (UE). The factorization of the p_T^{miss} into these physics objects, and the variation of each object’s momentum within its scale and resolution uncertainties provide a good estimate of the uncertainty that each of the objects contributes to the p_T^{miss} .

In the rest of this chapter, the uncertainty in the p_T^{miss} is evaluated by comparing the recalculated p_T^{miss} to the nominal, not varied, p_T^{miss} . As already hinted, the uncertainty related to the energy measurement of the jets is the dominant uncertainty in these measurements. These uncertainties are split into those relating to

the JES uncertainties, which are up to 3% (12%) for jets inside (outside) the tracker acceptance, and the JER uncertainties that range between 5–20%. More information on the JES and JER uncertainties is given in Chapter 13. A subdominant uncertainty is that related to the measurement of the muon energy scale, which amounts to 0.2%, and electron and photon energy scale, which amounts to 0.6% (1.5%) in the barrel (endcap). The UE uncertainty is evaluated based on the momentum resolution of each PF candidate, which depends on the type of the candidate. A detailed description of the PF candidate calibration can be found in [80, 97, 108]. The p_T measurement of PF charged hadrons is dominated by the tracker resolution. For PF neutral hadrons, the p_T resolution is dominated by the resolution of the HCAL. The ECAL resolution dominates the PF photon p_T measurement, whereas HF intrinsic resolution dominates that for the PF particles in the HF.

9.3 Event selection

Dilepton and single photon samples are used to study the p_T^{miss} response and resolution. These samples are chosen as they are dominated by the DY process where no genuine p_T^{miss} from neutrinos is expected, and serves as a good tool to measure the performance of the p_T^{miss} originating from detector inefficiencies or jet mismeasurements. One of the major backgrounds in the searches presented in this thesis, is large p_T^{miss} due to jet mismeasurement in the DY process. In Chapter 11 the technique by which this background is predicted is presented. It relies on the assumption that the p_T^{miss} spectrum is the same in Z+jets as in γ +jets, and solely due to the hadronic jets recoiling off of the boson. Therefore, these events are used to validate this assumption, in terms of p_T^{miss} tails, response and resolution.

Dilepton event samples

The dilepton samples are subdivided into two categories based on the flavor of the lepton, namely $Z \rightarrow \mu^+ \mu^-$ and $Z \rightarrow e^+ e^-$ samples. The dileptonic datasets are presented in Table 9-1. The events for the $Z \rightarrow \mu^+ \mu^-$

Table 9-1. Datasets used for the p_T^{miss} performance study in $Z \rightarrow ll$ events.

Dilepton datasets used for p_T^{miss} performance study
Dielectron samples
/DoubleEG/Run2016B-03Feb2017_ver2-v2/MINIAOD
/DoubleEG/Run2016(C-G)-03Feb2017-v1/MINIAOD
/DoubleEG/Run2016H-03Feb2017_ver2-v1/MINIAOD
/DoubleEG/Run2016H-03Feb2017_ver3-v1/MINIAOD
Dimuon samples
/DoubleMuon/Run2016B-03Feb2017_ver2-v2/MINIAOD
/DoubleMuon/Run2016(C-G)-03Feb2017-v1/MINIAOD
/DoubleMuon/Run2016H-03Feb2017_ver2-v1/MINIAOD
/DoubleMuon/Run2016H-03Feb2017_ver3-v1/MINIAOD

and $Z \rightarrow e^+ e^-$ samples are recorded using dimuon and dielectron triggers, as presented in Table 9-2 that select events where the p_T of the two leading leptons are above asymmetric thresholds. The simulated samples used are presented in Appendix B. Candidate events are required to have both the leading (subleading) lepton p_T greater than 25(20) GeV and an invariant mass in the range of 80 to 100 GeV, compatible with the mass of the Z boson. In order to have a pure sample of dilepton events originating from Drell–Yan production, a veto is applied on any event containing a third lepton of $p_T > 20$ GeV. The spectrum of the Z boson transverse momentum, historically referred to as q_T instead of p_T , is shown in Figure 9-1 where only the statistical uncertainty in the simulated samples is considered as the dilepton energy resolution is very good.

Table 9-2. Triggers used for the p_T^{miss} performance study in $Z \rightarrow ll$ events.

Dilepton triggers used for p_T^{miss} study	
Dimuon and single muon triggers	
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_v*	
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_v*	
HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_v*	
HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ_v*	
HLT_Mu27_TkMu8_v*	
HLT_Mu30_TkMu11_v*	
HLT_Mu50_v*	
HLT_IsoTkMu22_v*	
HLT_IsoMu27_v*	
HLT_IsoMu24_v*	
HLT_IsoMu22_v*	
Dielectron triggers	
HLT_Ele17_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v*	
HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v*	
HLT_DoubleEle33_CaloIdL_GsfTrkIdVL_v*	
HLT_DoubleEle33_CaloIdL_GsfTrkIdVL_MW_v*	
HLT_Ele27_eta2p1_WPLoose_Gsf_v*	
HLT_Ele27_WPTight_Gsf_v*	
HLT_Ele35_WPLoose_Gsf_v*	

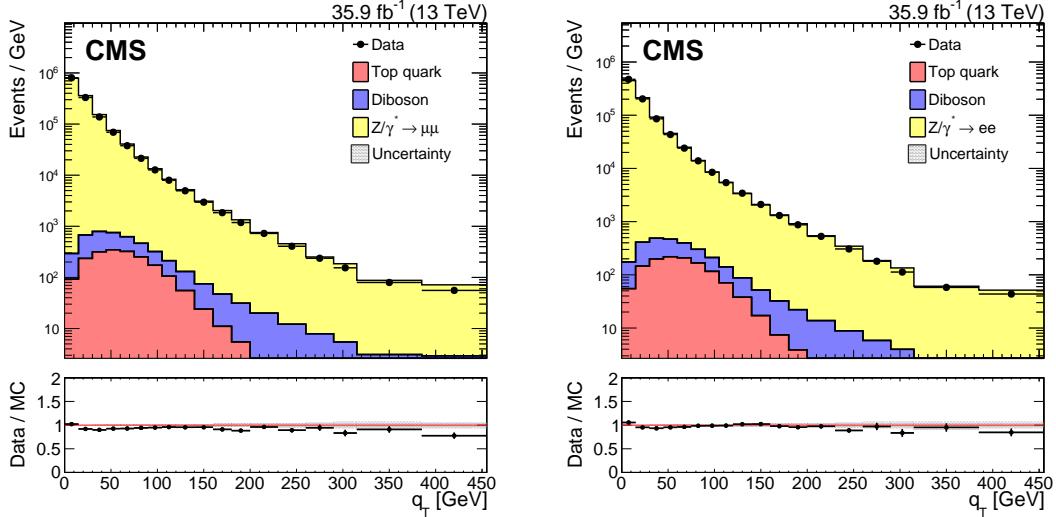


Figure 9-1. Z boson q_T in $Z \rightarrow \mu^+\mu^-$ (left) and $Z \rightarrow e^+e^-$ (right) samples. The Diboson contribution corresponds to processes with two electroweak bosons produced in the final state. The Top quark contribution corresponds to the top anti-top pair and single top production processes. The band in the data to simulation ratio corresponds to the statistical uncertainty in simulated samples [61].

Single-photon event sample

The events in the single-photon sample are presented in Table 9-3, and are selected using a set of prescaled isolated single-photon triggers with varying thresholds, presented in Table 9-4. The simulated samples used are presented in Appendix B. The p_T thresholds of the triggers are 30, 50, 75, 90, 120, and 165 GeV, and the first five triggers had different L1 accept rate (prescale) during the data taking periods, following the luminosity. Candidate events are weighted based on the prescale values of the triggers. One tight ID photon

Table 9-3. Datasets used for the p_T^{miss} performance study in single γ events.

Datasets used for p_T^{miss} study
/SinglePhoton/Run2016B-03Feb2017_ver2-v2/MINIAOD
/SinglePhoton/Run2016(C-G)-03Feb2017-v1/MINIAOD
/SinglePhoton/Run2016H-03Feb2017_ver2-v1/MINIAOD
/SinglePhoton/Run2016H-03Feb2017_ver3-v1/MINIAOD

Table 9-4. Triggers used for the p_T^{miss} performance study in single γ events.

Single photon triggers used for p_T^{miss} study
HLT_Photon30_R9Id90_HE10_IsoM_v*
HLT_Photon50_R9Id90_HE10_IsoM_v*
HLT_Photon75_R9Id90_HE10_IsoM_v*
HLT_Photon90_R9Id90_HE10_IsoM_v*
HLT_Photon120_R9Id90_HE10_IsoM_v*
HLT_Photon165_R9Id90_HE10_IsoM_v*

with $p_T > 50$ GeV is required and events are vetoed that contain leptons of $p_T > 20$ GeV. At least one jet with $p_T \geq 40$ GeV is required to ensure that the candidate events, i.e. a jet recoiling off of the photon, are selected. To match the trigger conditions, the leading photon is further required to have the ratio of the energy deposited in a 3×3 crystal region of the ECAL, centered around the crystal containing an energy deposit greater than all of its immediate neighbors, to the energy of the entire deposit of the photon greater than 0.9. The photon q_T spectrum is shown in Figure 9-2. Similarly to Figure 9-1 only the statistical uncertainty in the simulated samples is considered as the photon energy resolution is very good.

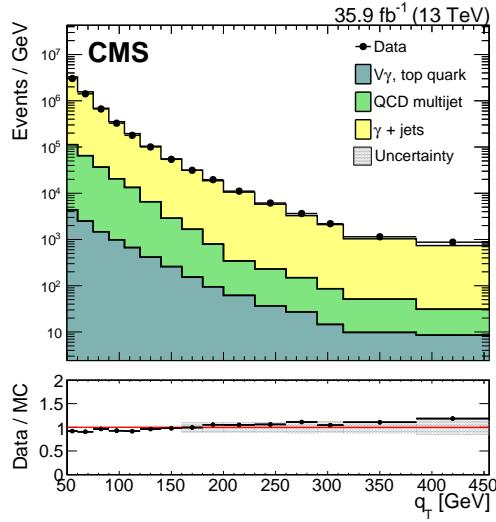


Figure 9-2. Distribution of the photon q_T in the single-photon sample. The $V\gamma + \text{Top}$ quark contribution corresponds to the $Z\gamma$, $W\gamma$, top anti-top pair and single top production processes. The band in the data to simulation ratio corresponds to the statistical uncertainty in the simulated samples [61].

9.4 p_T^{miss} filters

Non-physical, anomalous high- p_T^{miss} events can arise because of a variety of reconstruction failures or malfunctioning detectors. In the p_T^{miss} study presented in this chapter, and in the analyses covered in this thesis, these anomalous p_T^{miss} events are filtered away on an event basis. In the following, the various filters are presented and the underlying source of the anomalous p_T^{miss} .

HCAL filters

The geometrical patterns of HPD or RBX channels as well as the pulse shape and timing information are utilized by various HCAL barrel and endcap (HBHE) algorithms to identify and eliminate noise. These filter algorithms operate both in “noise filtering” and “event filtering” modes. In the noise filtering mode, the anomalous energy deposits are removed from the event reconstruction; in the event filtering mode, the bunch crossing is removed from the data set. In addition, there is an isolation-based noise filter that utilizes a topological algorithm, where energy deposits in HCAL and ECAL are combined and compared with measurements from the tracker to identify isolated anomalous activity in HBHE.

ECAL filters

One large source of anomalous p_T^{miss} signals can be created if a hadron hits the front-end electronics of the ECAL SCs, creating a large false signal. Additionally, anomalously high energetic deposits in the SCs, and the lack of information for channels that have nonfunctioning readout electronics, are removed through dedicated noise filters. During the data taking relevant for this thesis, five ECAL endcap SCs produced large, anomalous pulses, leading to spurious p_T^{miss} . These crystals were removed from the readout, and their energies were not considered. Furthermore, in about 0.7% of ECAL towers, the crystal-by-crystal information is not available. The trigger primitive (TP) information, however, is still available, and is used to estimate the energy. The TP information saturates above 127.5 GeV. Events with a TP close to saturation in one of these crystals are removed.

Beam halo filter

A final large source of anomalous large p_T^{miss} is due to so called machine-induced backgrounds or beam halo, meaning the production of muons when beam protons undergo collisions upstream of the detector. The characteristics of beam halo particles is that they travel parallel to the beamline. If they leave energy deposits in the calorimeters, those will be along a line with constant ϕ . Similarly, any interaction of the beam halo particles in the CSC, will be in line with the calorimeter deposits. The filter exploits information from both the CSC and the calorimeters, and an example event display for a beam halo event is shown in Figure 9-3, where collinear hits in the CSC are visible.

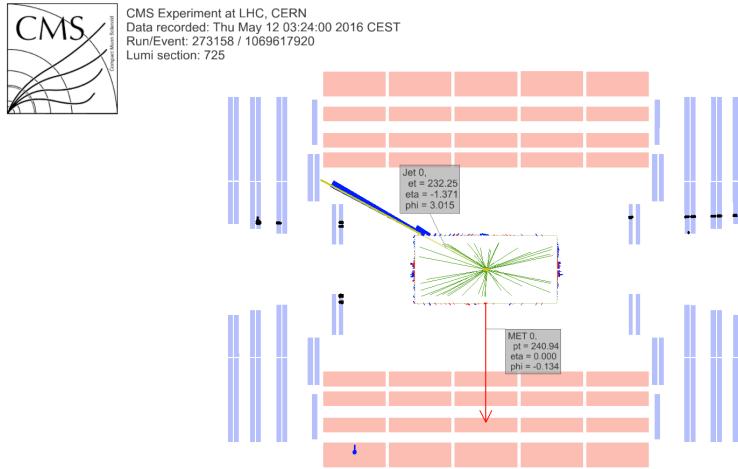


Figure 9-3. Event display for a beam halo event with collinear hits in the CSC (black), p_T^{miss} of $\sim 241 \text{ GeV}$, and a jet with $p_T = \sim 232 \text{ GeV}$ [61].

Reconstruction filters

During the 2016 data taking of Run 2 of the LHC, an additional source of anomalous p_T^{miss} events was uncovered. Large anomalous p_T^{miss} arose as an effect of a misreconstruction issue of muons during the muon-tracking iteration step. If there is a high p_T and low-quality track reconstructed, it could contribute to p_T^{miss} by either adding a misreconstructed PF muon, or as a misreconstructed PF charged hadron to the p_T^{miss} calculation. The relative p_T uncertainty of the track p_T , determined by the Tune-P algorithm [101], is used to identify misreconstructed muons and charged hadrons. An event with an identified misreconstructed muon or a charged hadron are rejected using a dedicated filter.

To visualize the need of these event filters, dijet and monojet events are used. Figure 9-4 (left) shows a comparison of the p_T^{miss} before and after the application of the event filters for the dijet sample, where the events with large p_T^{miss} are found to mostly be due to electronic noise in the calorimeters. Figure 9-4 (right) shows a comparison of the jet ϕ distribution before and after the application of the event filters for the monojet sample, where the excess of events with jet $\phi \approx 0$ or $\phi \approx \pi$ are typical for beam halo events.

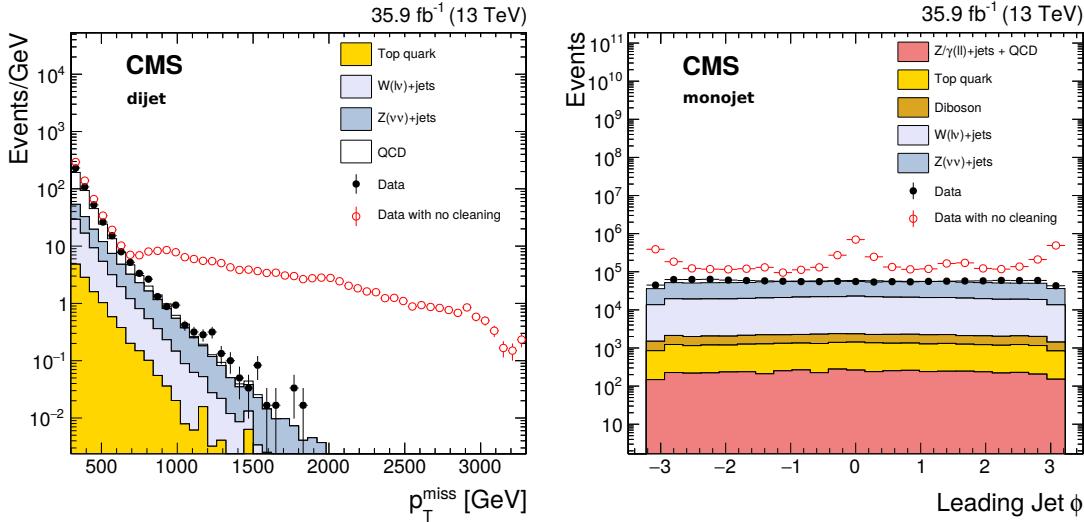


Figure 9-4. The p_T^{miss} (left) and jet ϕ (right) distributions for events passing the dijet (left) and monojet (right) selection with the event filtering algorithms applied, including that based on jet identification requirements (filled markers), without the event filtering algorithms applied (open markers), and from simulation (solid histograms) [61].

9.5 p_T^{miss} performance

After having covered the p_T^{miss} reconstruction algorithms, their calibration and the filtering of anomalous p_T^{miss} events, the main part of the performance study will be presented. As has already been conveyed, the p_T^{miss} is a very sensitive observable, and is relying on accurate object reconstruction and efficient sub-detectors. Therefore it is a great tool for monitoring the detector and reconstruction performance during data-taking. The performance of the p_T^{miss} can be summarized in three parts. By monitoring the data and simulation of the p_T^{miss} in events with no genuine p_T^{miss} from neutrinos, any anomalies will show up as p_T^{miss} tails, or as an overall disagreement between the data and the simulation. The p_T^{miss} response is crucial for validating the JECs introduced in Section 8.4, or any issues related to the muon, electron or photon energy scale. The p_T^{miss} resolution is valuable for monitoring how the performance degrades as a function of some variable, such as the number of vertices. Initially, the studies presented in this thesis were performed during the 2016 data taking, and was crucial in uncovering various reconstruction inefficiencies, and was later refined and summarized in a publication.

p_T^{miss} performance using hadronic recoil

A well-measured Z/γ boson provides a unique event axis and a precise momentum scale. Such events should have little or no genuine p_T^{miss} , and the hadronic recoil is projected onto the axis of this well measured boson, as illustrated in Figure 9-5. Formally, the hadronic recoil (u) is defined as the vector p_T sum of all PF candidates except for the vector boson (or its decay products in the case of the Z boson decay). The assumption of momentum conservation in the transverse plane imposes $\vec{q}_T + \vec{u}_T + \vec{p}_T^{\text{miss}} = 0$. The hadronic recoil can be split into its parallel and perpendicular components with respect to the boson axis, and these quantities, u_{\parallel} and u_{\perp} , are used to study the p_T^{miss} response and resolution. The p_T^{miss} response is defined as $-\langle u_{\parallel} \rangle / \langle q_T \rangle$ where $\langle \rangle$ indicates the mean. As the u_{\parallel} by definition should be aligned with the axis of the well measured boson, this quantity is studied as it gives a measure of how well balanced the event is. Ideally, if all the jets that make up the hadronic recoil and the boson are perfectly measured, the u_{\parallel} and the q_T would be of the same magnitude and perfectly back to back, and the response would be 1. But as the experimental

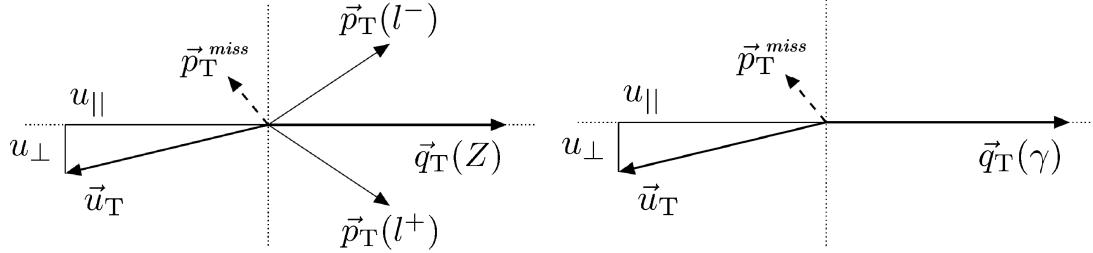


Figure 9-5. Illustration of the Z boson (left) and photon (right) event kinematics in the transverse plane. The vector \vec{u}_T denotes the vectorial sum of all particles reconstructed in the event except for the two leptons from the Z decay (left) or the photon (right) [61].

situation is far from ideal, any mismeasurement of the jets would give rise to p_T^{miss} , and this can be shown as a non-unity response.

The p_T^{miss} resolution is assessed as the RMS of the $u_{||} + q_T$ and u_{\perp} distributions, and are denoted by $\sigma(u_{||})$ and $\sigma(u_{\perp})$, respectively. The $u_{||} + q_T$ is studied as it shows the degradation of the p_T^{miss} resolution along the axis of the boson. For example, the $\sigma(u_{\perp})$ evaluated as a function of the q_T shows the degradation of the resolution with the q_T of the boson. The u_{\perp} on the other hand, is a quantity that reflects isotropic effects, such as pileup. As pileup does not have a preferred direction it will thus show up everywhere in the event and not just along the axis of the boson. Therefore the u_{\perp} is useful to study the degradation of the resolution as a function of number of vertices.

9.6 Performance of PF p_T^{miss} algorithm

The p_T^{miss} in selections with no genuine p_T^{miss} from neutrinos is shown in Figure 9-6, and exhibits a good agreement between the data and the simulation. The lower pads of Figure 9-6 show the magnitude of the uncertainties in these events, where the p_T^{miss} resolution is dominated by the resolution of the hadronic activity, ranging up to 10–15% for the jet momentum resolution [105]. Since the momentum resolution for leptons (photons) σ_{p_T}/p_T of 1–4% (1–3%) [101, 108] is sub-dominant to the resolution of the hadronic activity, these uncertainties are not taken into account in the ratios, and the final uncertainty shown in the figures include uncertainties in the JES, the JER, and the UE, added in quadrature. The increase in the uncertainty band around 40 GeV is related to the JES and the JER sources in events with at least one jet and no genuine p_T^{miss} . For higher values of p_T^{miss} , where processes with genuine p_T^{miss} , e.g. top quark background, become more dominant, the uncertainty decreases. These distributions display a very similar shape of the p_T^{miss} spectrum in the three channels studied ($Z \rightarrow \mu^+\mu^-$, $Z \rightarrow e^+e^-$ and $\gamma+$ jets), confirming that the p_T^{miss} spectrum in $\gamma+$ jets events can be used to estimate the p_T^{miss} spectrum in the $Z \rightarrow l^+l^-$ signal events, which is the foundation of one of the major background prediction techniques used in this thesis.

Distributions of $u_{||} + q_T$ and u_{\perp} in $Z \rightarrow \mu^+\mu^-$, $Z \rightarrow e^+e^-$ and $\gamma+$ jets events are shown in Figure 9-7. The kinematic definition of $u_{||}$ dictates that for processes with no genuine p_T^{miss} , $u_{||}$ is balanced with the boson q_T . Therefore, the vectorial sum of $u_{||}$ and q_T results in a symmetric distribution, centered at zero; any deviations from this behavior imply imperfect calibration of p_T^{miss} . In events with genuine p_T^{miss} , due to the presence of the neutrinos, $u_{||}$ and q_T are not balanced, leading to an asymmetric distribution as can be seen as the Top quark and Electroweak contribution in the distributions. Due to the assumed isotropic nature of the energy fluctuations of the detector noise and underlying event, the u_{\perp} distribution is symmetric with a mean value of 0.

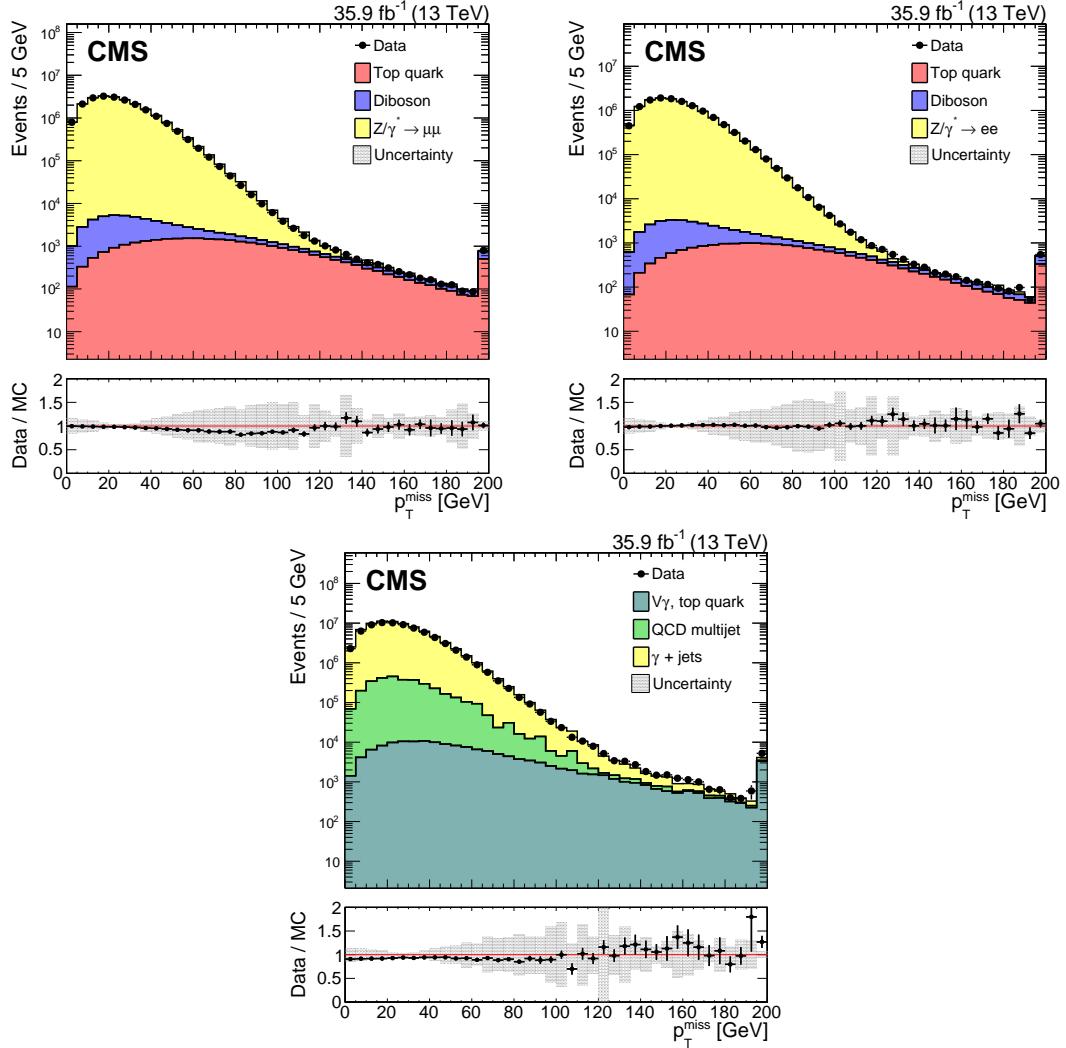


Figure 9-6. The p_T^{miss} for events passing the dimuon (left), dielectron (right) and single photon (bottom) selections, in data (black markers) and simulation (solid histograms). The lower bands show the data to simulation ratio with the systematic uncertainties due to the JES, the JER, and variations in the UE are added in quadrature [61].

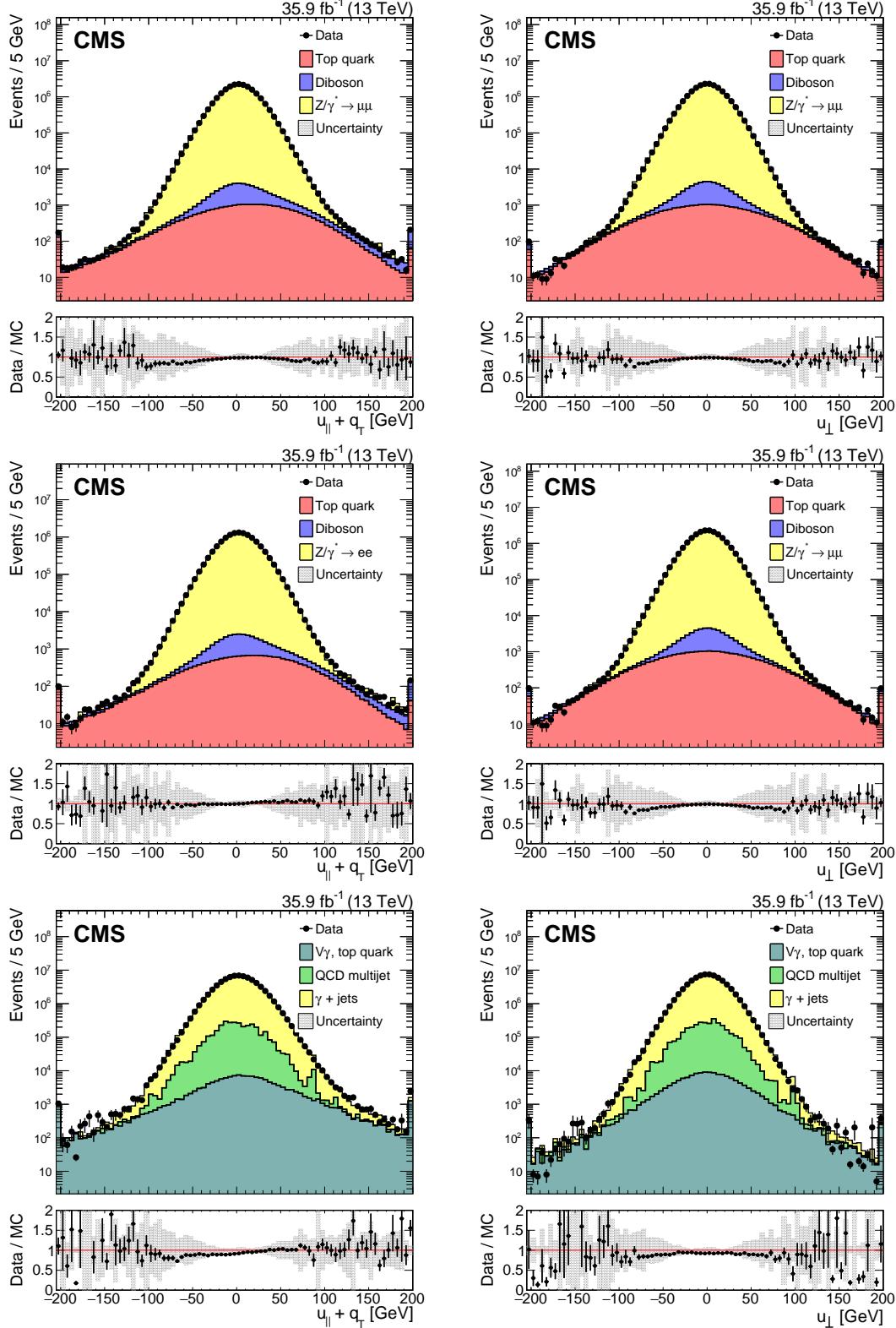


Figure 9-7. Distribution of $u_{\parallel} + q_T$ (left) and u_{\perp} (right) components of the hadronic recoil, in data (filled markers) and simulation (solid histograms), in the $Z \rightarrow \mu^+\mu^-$ (top), $Z \rightarrow e^+e^-$ (middle), and $\gamma + \text{jets}$ (bottom) samples [61].

Figure 9-8 shows the p_T^{miss} response as a function of q_T , in data and simulation, in $Z \rightarrow \mu^+ \mu^-$, $Z \rightarrow e^+ e^-$ and photon events. Two types of response are displayed. On the left is the response of the uncorrected (raw) p_T^{miss} , in which the JECs have not been propagated to the p_T^{miss} . The effect of correcting the p_T^{miss} , the so called “Type-I” correction, is shown on the right, where the response is now much closer to unity, proving that the corrected jets results in a well balanced event. The underestimation of the hadronic response observed at smaller $q_T \leq 100$ GeV is due to the significant contribution of the uncalibrated component of p_T^{miss} , which mainly consists of jets with $p_T < 15$ GeV and unclustered particles. The response of p_T^{miss} is found to agree between all three samples within 2%.

The residual response difference between the samples stems from the different mechanism used to disambiguate muons, electrons, and photons from jets used in the correction of the p_T^{miss} , as discussed in Section 9.2. Simulation studies has shown that in the case of electrons and photons, a small fraction ($\lesssim 10\%$) of jets survive the disambiguation criteria yet overlap with prompt electrons and photons. As a result, these jets wrongly contribute to the p_T^{miss} calibration, leading to a 1–2% lower response in the electron and photon channels. The uncorrected p_T^{miss} response reveals a perfect agreement between the electron and muon events, further proving that the disagreement in the corrected response is due to the propagation of the JECs and the disambiguation criteria of what jets to correct.

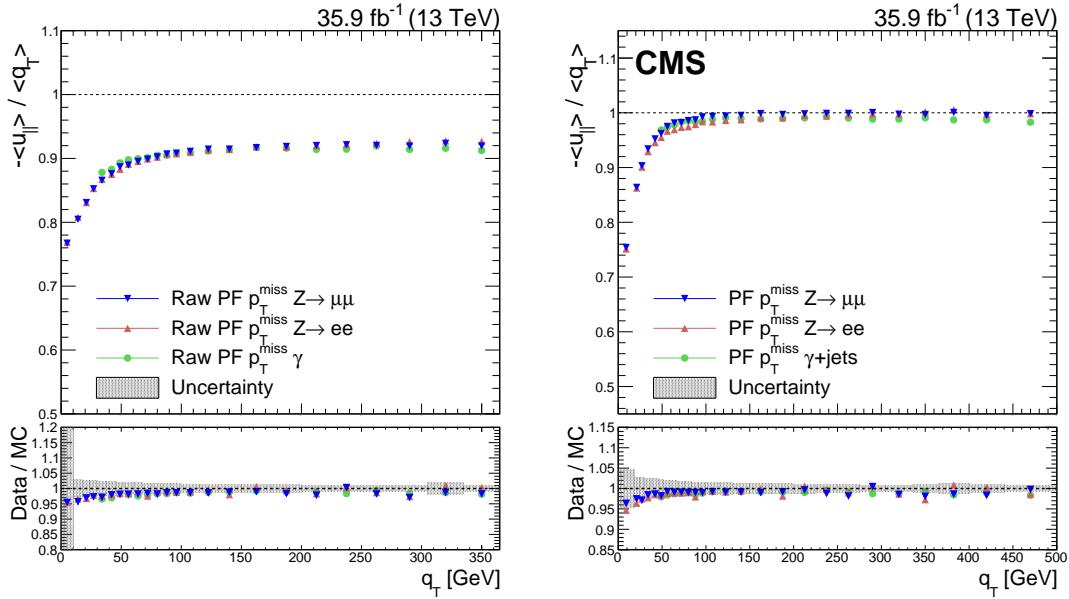


Figure 9-8. Upper panels: Response of the Raw (left) and Type 1 corrected (right) p_T^{miss} in data in $Z \rightarrow \mu^+ \mu^-$, $Z \rightarrow e^+ e^-$, and $\gamma + \text{jets}$ events. Lower panels: Ratio of the p_T^{miss} response in data and simulation. The band corresponds to the systematic uncertainties due to the JES, the JER, and variations in the UE added in quadrature, estimated from the $Z \rightarrow e^+ e^-$ sample [61].

The resolution of the p_T^{miss} is evaluated as a function of the boson q_T , the number of vertices, and the $\sum E_T$, which is the scalar p_T sum of all PF candidates. Figure 9-9 shows the resolution as a function of q_T . In order to compare the resolution of p_T^{miss} consistently across the samples, the resolution in each sample is corrected for the differences observed in the response, with a resulting negligible impact on the results. Due to the isotropic nature of energy fluctuations stemming from detector noise and the underlying event, the dependence of the resolution of u_{\perp} on q_T is smaller than for u_{\parallel} . For $q_T > 200$ GeV, the p_T^{miss} resolution is $\approx 13\%$ and $\approx 9\%$, for the u_{\parallel} and u_{\perp} , respectively.

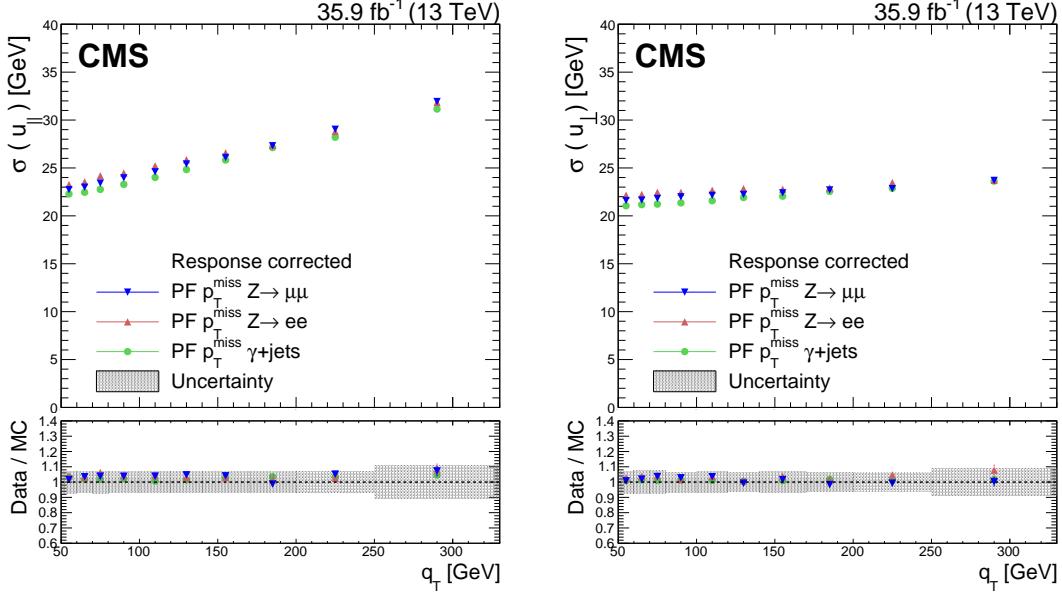


Figure 9-9. Resolution of the u_{\parallel} and u_{\perp} components of the hadronic recoil as a function of q_T in $Z \rightarrow \mu^+\mu^-$, $Z \rightarrow e^+e^-$, and $\gamma+\text{jets}$ events [61].

Figure 9-10 displays the resolution as a function of N_{vtx} . There is a significant dependence on N_{vtx} , since only pileup mitigation techniques are deployed to the PF jets and not the PF p_T^{miss} algorithm. The resolution is parametrized as:

$$f(N_{\text{vtx}}) = \sqrt{\sigma_c^2 + \frac{N_{\text{vtx}}}{0.70}\sigma_{\text{PU}}^2}, \quad (9-12)$$

where σ_c is the resolution term induced by the hard scattering interaction and σ_{PU} is the resolution term induced on average by an additional pileup interaction. The factor 0.70 accounts for the vertex reconstruction efficiency [109]. Results of the parametrization for the u_{\parallel} and u_{\perp} components are given in Table 9-5, where the uncertainties displayed for both components are obtained from the fit, and for simulation the JES, the JER, and UE uncertainties are added in quadrature. Each additional pileup vertex is found to degrade the resolution of each component by 3.8–4.0 GeV.

Table 9-5. Parametrization results of the resolution curves for the u_{\parallel} and u_{\perp} components as a function of N_{vtx} . The parameter values for σ_c are obtained from data and simulation, and the values for σ_{PU} are obtained from data, along with a ratio R_{PU} of data and simulation.

Process	$\sigma_c(\text{data})[\text{GeV}]$	$\sigma_c(\text{MC})[\text{GeV}]$	$\sigma_{\text{PU}}(\text{data})[\text{GeV}]$	$R_r = \sigma_{\text{PU}}(\text{data})/\sigma_{\text{PU}}(\text{MC})$
u_{\parallel} component				
$Z \rightarrow \mu^+\mu^-$	13.9 ± 0.07	11.9 ± 1.53	3.82 ± 0.01	0.95 ± 0.04
$Z \rightarrow e^+e^-$	14.6 ± 0.09	12.0 ± 1.09	3.80 ± 0.02	0.95 ± 0.03
$\gamma + \text{jets}$	12.2 ± 0.10	10.2 ± 1.98	3.97 ± 0.02	0.97 ± 0.05
u_{\perp} component				
$Z \rightarrow \mu^+\mu^-$	10.3 ± 0.08	8.58 ± 2.20	3.87 ± 0.01	0.97 ± 0.04
$Z \rightarrow e^+e^-$	10.7 ± 0.10	8.71 ± 1.76	3.89 ± 0.01	0.96 ± 0.03
$\gamma + \text{jets}$	9.04 ± 0.11	6.93 ± 2.70	3.94 ± 0.01	0.97 ± 0.04

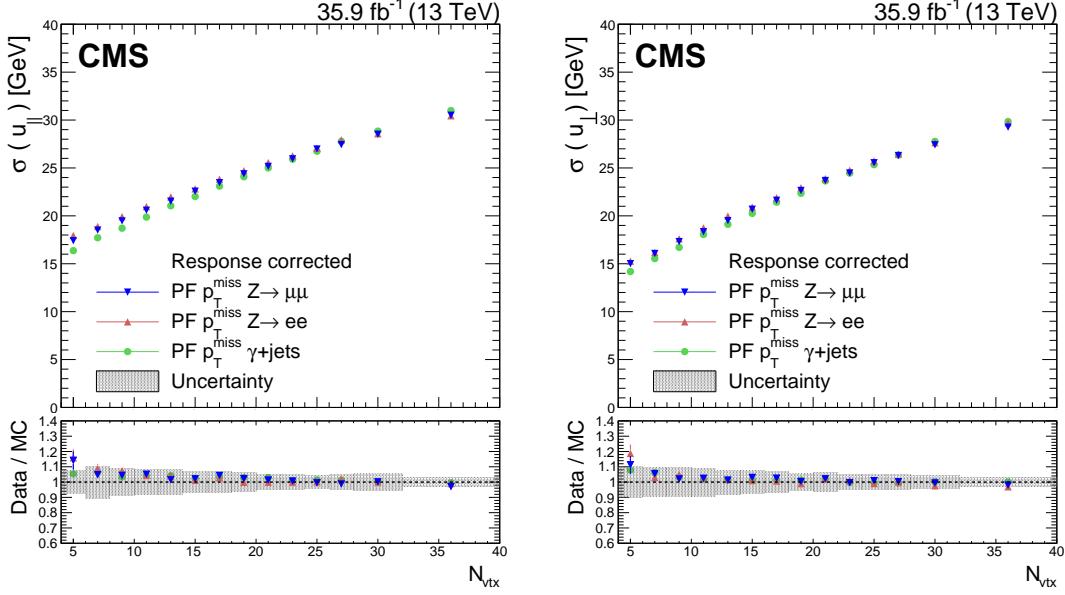


Figure 9-10. Resolution of the u_{\parallel} and u_{\perp} components of the hadronic recoil as a function of the reconstructed vertices in $Z \rightarrow \mu^+\mu^-$, $Z \rightarrow e^+e^-$, and $\gamma + \text{jets}$ events [61].

Lastly, Figure 9-11 shows the resolution of u_{\parallel} and u_{\perp} as a function of the scalar p_T sum of all PF candidates ($\sum E_T$). The resolution measured in different samples, and in data and simulation, are found to be in good agreement. The relative p_T^{miss} resolution improves with increasing $\sum E_T$, driven by the amount of the activity in the calorimeters. The resolution in different samples is parametrized as:

$$\sigma_{u_{\perp}, u_{\parallel}} = \sigma_0 + \sigma_s \sqrt{\sum E_T}, \quad (9-13)$$

where σ_0 is the resolution term induced by intrinsic detector noise and σ_s is the stochastic resolution term. Results of the parametrization for the u_{\parallel} and u_{\perp} components are given in Table 9-6, where the uncertainties are the same as in Table 9-5.

Table 9-6. Parametrization results of the resolution curves for u_{\parallel} and u_{\perp} components as a function of the scalar p_T sum of all PF candidates. The parameter values for σ_0 are obtained from data and simulation, whereas the σ_s are obtained from data along with the ratio R_s , the ratio of data and simulation.

Process	$\sigma_0(\text{data})[\text{GeV}]$	$\sigma_0(\text{MC})[\text{GeV}]$	$\sigma_s[\text{GeV}^{1/2}]$	$R_s = \sigma_s(\text{data})/\sigma_s(\text{MC})$
u_{\parallel} component				
$Z \rightarrow \mu^+\mu^-$	1.98 ± 0.07	0.85 ± 2.45	0.64 ± 0.01	0.95 ± 0.11
$Z \rightarrow e^+e^-$	2.18 ± 0.09	0.19 ± 2.90	0.64 ± 0.01	0.92 ± 0.11
$\gamma + \text{jets}$	1.85 ± 0.09	0.94 ± 2.52	0.64 ± 0.01	0.96 ± 0.11
u_{\perp} component				
$Z \rightarrow \mu^+\mu^-$	-1.63 ± 0.06	-1.72 ± 2.53	0.68 ± 0.01	0.99 ± 0.11
$Z \rightarrow e^+e^-$	-1.42 ± 0.08	-1.98 ± 2.95	0.69 ± 0.01	0.96 ± 0.12
$\gamma + \text{jets}$	-1.16 ± 0.08	-1.31 ± 2.53	0.68 ± 0.01	0.98 ± 0.11

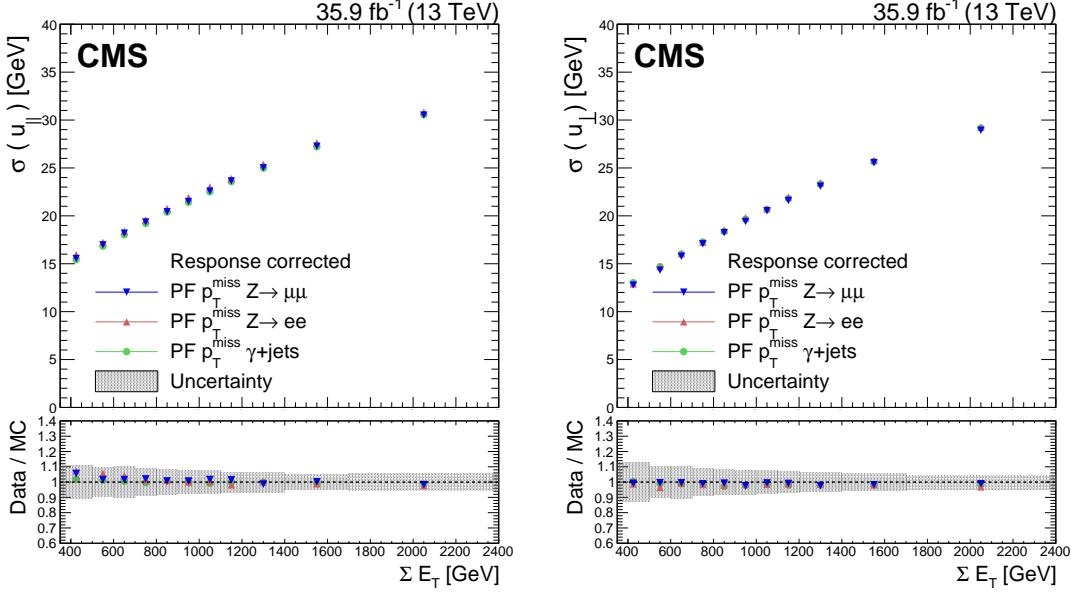


Figure 9-11. Resolution of the u_{\parallel} and u_{\perp} components of the hadronic recoil as a function of the scalar p_T sum of all PF candidates (lower row), in $Z \rightarrow \mu^+\mu^-$, $Z \rightarrow e^+e^-$ and γ +jets events [61].

9.7 Performance of PUPPI p_T^{miss} algorithm

The PUPPI p_T^{miss} distributions in the dilepton samples are shown in Figure 9-12. The data distributions are modeled well by the simulation, in both the muon and the electron channels. Similar to the case of PF p_T^{miss} , the p_T^{miss} resolution in these events is dominated by the resolution of the hadronic activity, but the PUPPI-weighted PF candidates yield a much improved resolution for jets compared to the PF case. This is also reflected in the uncertainty shown in the figures, which includes the uncertainties due to jet energy scale and resolution, and the energy scale of the unclustered particles. The distributions in $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ events of the vectorial sum $u_{\parallel} + q_T$ and of u_{\perp} using PUPPI p_T^{miss} , are shown in Figure 9-13.

Following the same arguments as in the PF p_T^{miss} case, in events with no genuine p_T^{miss} , the vectorial sum of u_{\parallel} and q_T is symmetric around zero, whereas for processes with genuine p_T^{miss} an asymmetric behavior is observed. The distribution of u_{\perp} is symmetric around zero. Simulation describes data well for all distributions. Figure 9-14 shows the PUPPI p_T^{miss} response as a function of q_T , extracted from data and simulation in $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ events. The response reaches unity for $Z \rightarrow \mu^+\mu^-$ events at a boson p_T of 150 GeV; while for PF p_T^{miss} the response is close to unity at 100 GeV. The slower rise of the response to unity is due to the removal of PF candidates that are wrongly associated with pileup interactions by the PUPPI algorithm. Similarly to PF p_T^{miss} , there is no response correction for the UE for PUPPI p_T^{miss} , which results in an underestimated response for low q_T . The response of p_T^{miss} is found to agree between the different samples within 2%.

In Figure 9-15, the resolution of the PUPPI p_T^{miss} algorithm is overlayed with the same results obtained using PF p_T^{miss} . Compared to the case of PF p_T^{miss} , the result obtained with the PUPPI algorithm show a reduced dependence on the number of pileup interactions. The resolution in different samples is parametrized using Equation (9-12), and the result of the parameterization are given in Table 9-7. Each additional pileup interaction is found to degrade the resolution of each component by up to 2 GeV. This resolution degradation corresponds to half of what is observed in the case of PF p_T^{miss} .

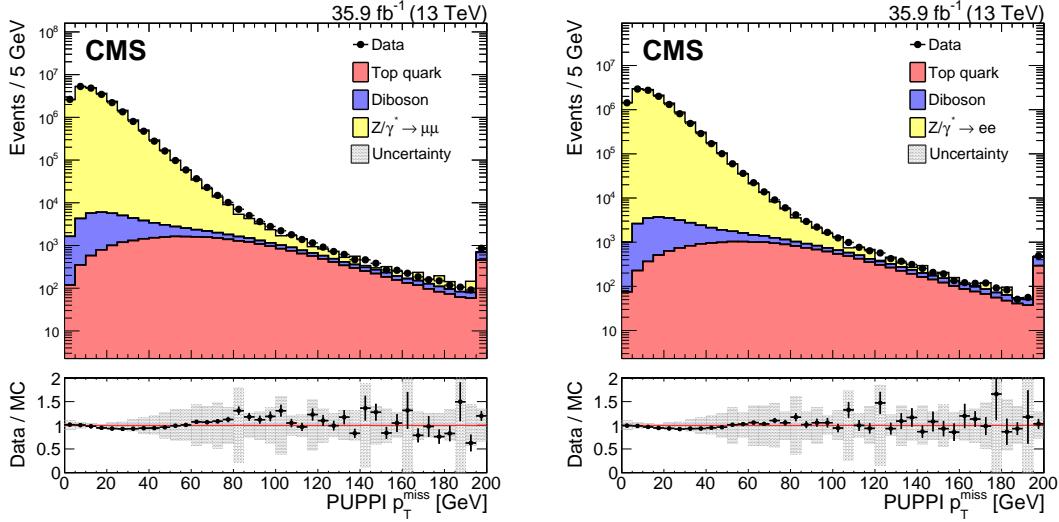


Figure 9-12. Upper panels: Distributions of PUPPI p_T^{miss} in $Z \rightarrow \mu^+ \mu^-$ (left) and $Z \rightarrow e^+ e^-$ (right) events [61].

Table 9-7. Parameterization results of the resolution curves for PUPPI u_{\parallel} and u_{\perp} components as a function of N_{vtx} . The parameter values for σ_c are obtained from data and simulation, and the values for σ_{PU} are obtained from data, along with the ratio R_{PU} of data and simulation.

Process	$\sigma_c(\text{data})[\text{GeV}]$	$\sigma_c(\text{MC})[\text{GeV}]$	$\sigma_{\text{PU}}(\text{data})[\text{GeV}]$	$R_{\text{PU}} = \sigma_{\text{PU}}(\text{data})/\sigma_{\text{PU}}(\text{MC})$
u_{\parallel} component				
$Z \rightarrow \mu^+ \mu^-$	18.9 ± 0.05	17.5 ± 0.74	1.93 ± 0.02	0.97 ± 0.11
$Z \rightarrow e^+ e^-$	18.9 ± 0.06	17.4 ± 0.80	1.94 ± 0.03	0.98 ± 0.12
u_{\perp} component				
$Z \rightarrow \mu^+ \mu^-$	14.2 ± 0.04	13.6 ± 0.59	1.78 ± 0.01	0.97 ± 0.09
$Z \rightarrow e^+ e^-$	14.3 ± 0.05	13.6 ± 0.59	1.80 ± 0.02	0.96 ± 0.09

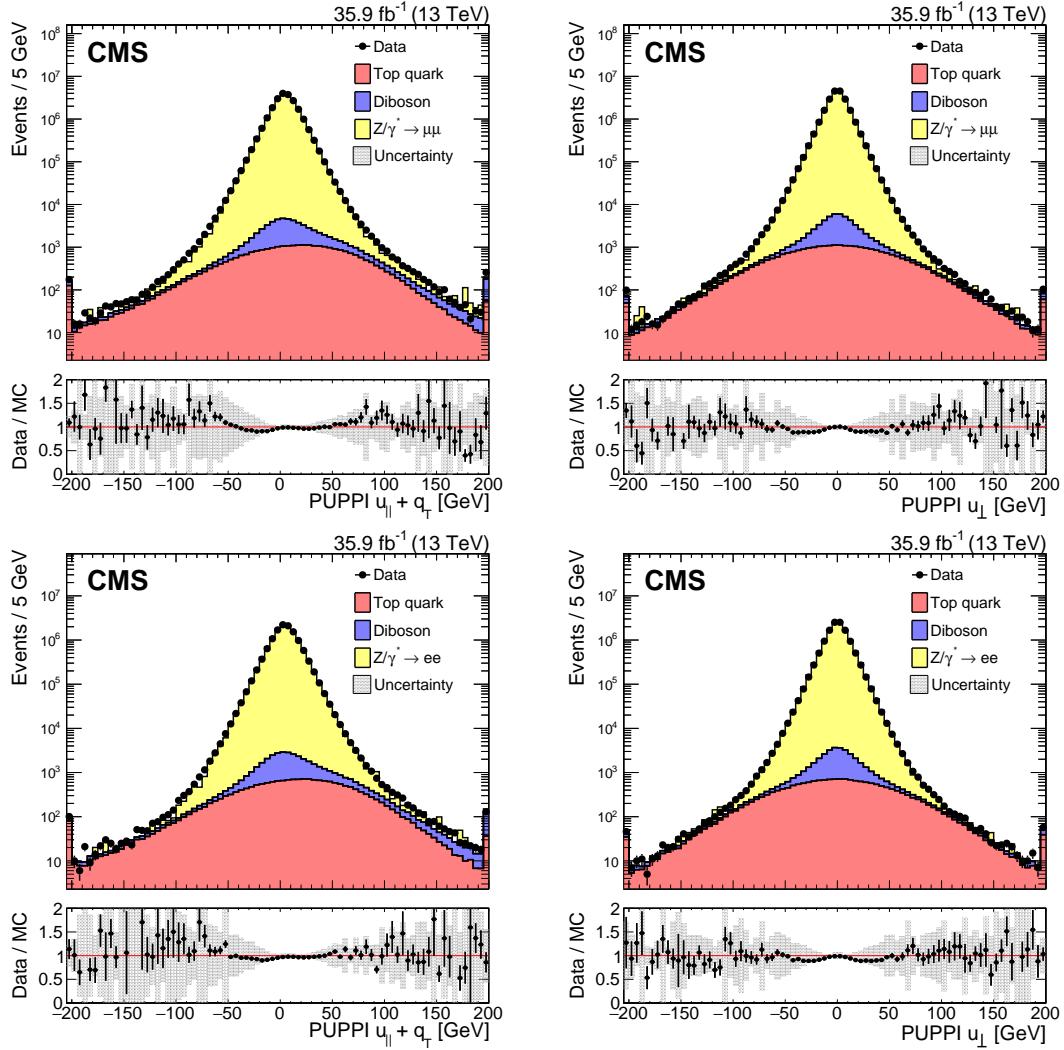


Figure 9-13. Upper panels: Distributions of the $u_{\parallel} + q_T$ and u_{\perp} components of the hadronic recoil, in data (filled markers) and simulation (solid histograms), for the $Z \rightarrow \mu^+\mu^-$ (upper) and $Z \rightarrow e^+e^-$ (lower) events [61].

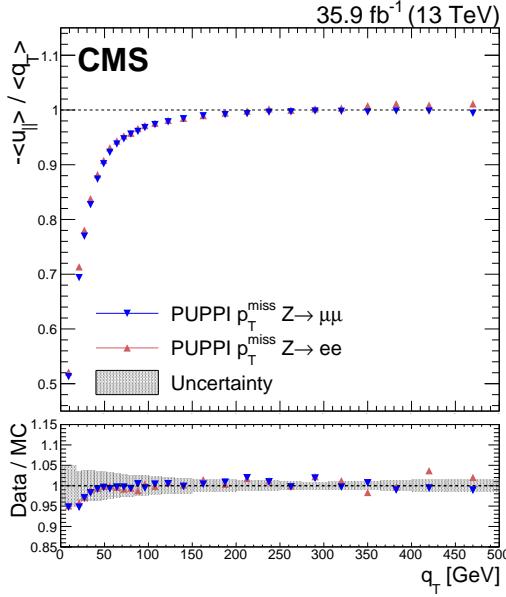


Figure 9-14. Response of PUPPI p_T^{miss} , defined as $-\langle u_{||} \rangle / \langle q_T \rangle$, in data in $Z \rightarrow \mu^+ \mu^-$ and $Z \rightarrow e^+ e^-$ events.

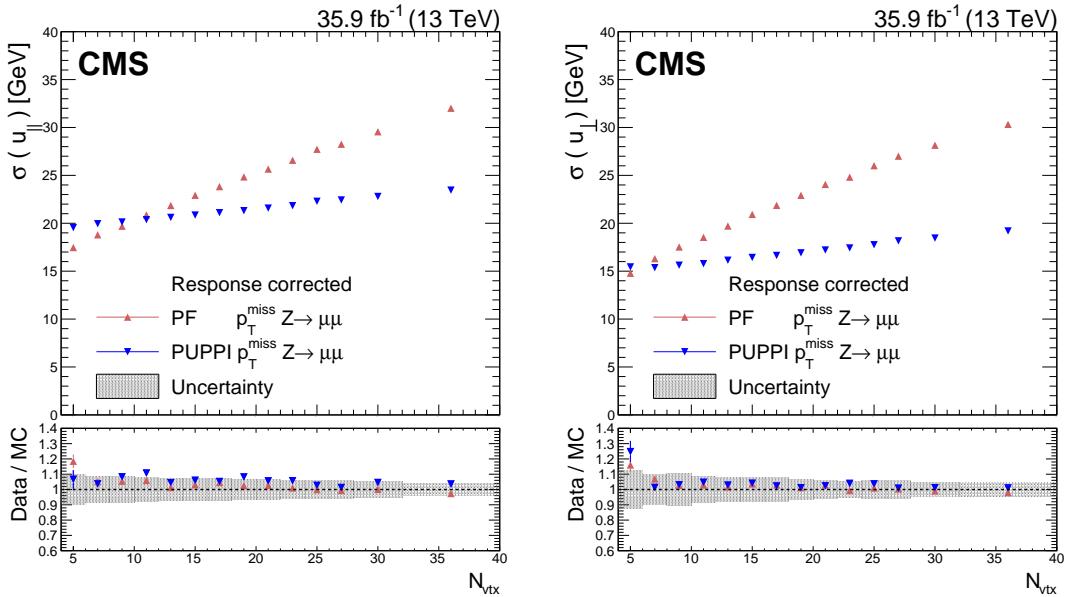


Figure 9-15. PUPPI and PF p_T^{miss} resolution of $u_{||}$ (left) and u_{\perp} (right) components of the hadronic recoil as a function of N_{vtx} , in $Z \rightarrow \mu^+ \mu^-$ events [61].

9.8 High pileup studies

By 2025, the LHC will be upgraded with a goal to increase the integrated luminosity by a factor of 10 beyond the original design. The new design, known as High Luminosity LHC (HL-LHC), will pose a major challenge due to the unprecedented increase of pileup expected (~ 140) in the collision events per bunch crossing. In preparation, LHC delivered a special pp collisions data with conditions similar to the ones expected at the HL-LHC. The "high pileup" data set, as it will be referred in what follows, corresponds to an integrated luminosity of 5 pb^{-1} . The bunch setup for the data was three isolated bunches of an average pileup of 70-100, and 2×48 bunch trains, corresponding to an average pileup of 35-50. Dedicated simulated samples were produced with similar conditions.

Using the high pileup data set and the dedicated MC simulations, the performance of the PF p_T^{miss} and PUPPI p_T^{miss} algorithms are studied in dimuon samples. The p_T^{miss} resolution of the $u_{||}$ and u_{\perp} components of the hadronic recoil as a function of the number of reconstructed vertices is shown in Figure 9-16. The results obtained from the high pileup data is overlayed with the ones obtained from the nominal data and found to be in agreement within the statistical uncertainties. Furthermore, the PUPPI p_T^{miss} is found to have more stable resolution across the full range.

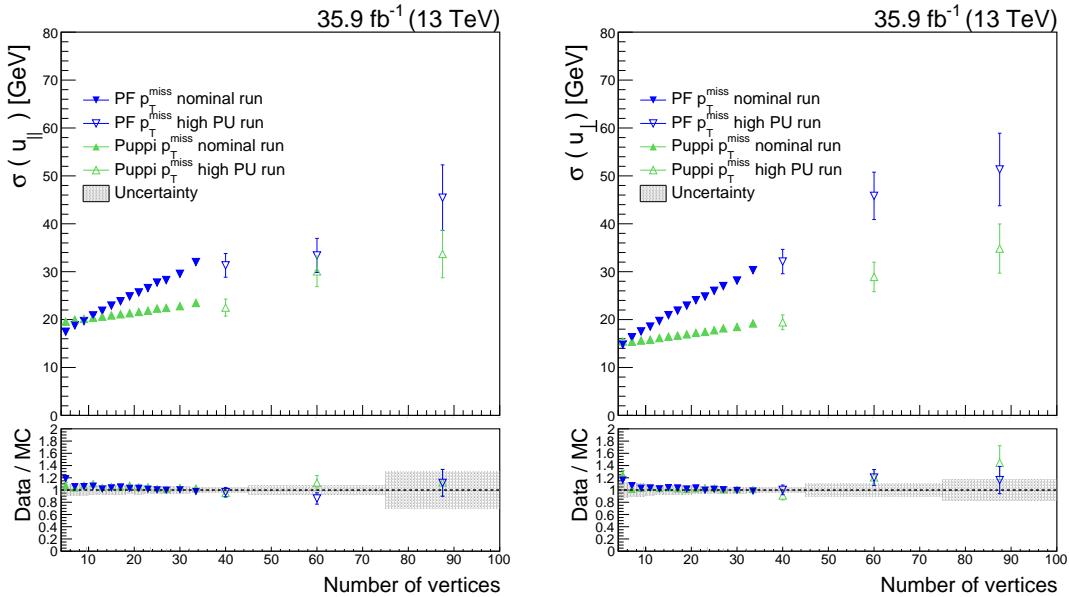


Figure 9-16. Upper panels: PUPPI and PF p_T^{miss} resolution of $u_{||}$ (left) and u_{\perp} (right) components of the hadronic recoil as a function of N_{vtx} , in $Z \rightarrow \mu^+ \mu^-$ events. The blue (green) markers correspond to the PF (PUPPI) p_T^{miss} reconstruction algorithm, with filled (open) markers for the nominal run (high pileup run).

9.9 Summary

The performance of the two p_T^{miss} reconstruction algorithms has been presented. The p_T^{miss} is a complex object that relies on the correct measurement of all the visible particles in an event. It is extremely sensitive to detector inefficiencies and mismeasured jets. Despite the inherent difficulties associated to the measurement of this complex variable, the PF and PUPPI algorithms provide a well-calibrated p_T^{miss} object. Both are proven to be ideal discriminating variables for SUSY searches where p_T^{miss} from potential LSPs need to be distinguished from p_T^{miss} from neutrinos and jet resolution effects. Further, the calibration of the p_T^{miss} , by propagating the JECs, provides a well balanced object compared to the uncorrected p_T^{miss} . Numerous p_T^{miss} filters are deployed that reduce the p_T^{miss} from sources that create large anomalous p_T^{miss} , such as beam halo or non functioning readout electronics in the calorimeters. The response and resolution of the p_T^{miss} is determined using events with an identified leptonically decaying Z boson or an isolated photon, which are measured in data and are in agreement with the expectations from simulation. Especially the PUPPI p_T^{miss} algorithm is showing a promising robustness against pileup, and will be a crucial algorithm for searches and measurements during the Run 3 and High Luminosity phases of the LHC. Both the PF and PUPPI p_T^{miss} algorithms demonstrate high performance and are ideal as discriminating variables in SUSY searches.

10 — Event selections

The p_T^{miss} variable, that has just been thoroughly described in the previous chapter, is the natural ingredient for any R -parity conserving SUSY search. What is special with the searches presented in this thesis, is that in addition to just searching for event with large p_T^{miss} , the choice two leptons of opposite charge and same flavor provides a large signal sensitivity and discrimination from SM backgrounds. This selection enables searches for both strongly and electroweakly produced SUSY, and for directly produced sleptons. The common selections are the large p_T^{miss} and the two electrons or muons. In addition, slightly different requirements on the hadronic activity in the events, and whether the lepton pair is compatible with a Z boson or not, are imposed to target the different signal models.

This chapter contains an overview of the datasets and triggers used for these analysis and the physics objects and selections applied to define the signal regions. In addition to the basic selections on leptons, jets, b-jets and p_T^{miss} , more specific variables such as the M_{T2} variable and a $t\bar{t}$ likelihood discriminator are presented. These variables are efficient in rejecting and categorizing SM background and are used in the design of the searches.

10.1 Datasets

As the searches presented in this thesis has the commonality that two same flavor opposite sign leptons are produced, naturally, the dielectron and dimuon streams of 13 TeV pp collision data are used. Additionally, an electron-muon datasets is used to collect events in a control regions needed for the background prediction, and JetHT and MET datasets are used to estimate the trigger efficiencies. All data samples are summarized in Table 10-1.

Table 10-1. Datasets used in the strong, electroweak and slepton searches and p_T^{miss} study

Signal events
/DoubleEG/Run2016B-03Feb2017_ver2-v2/MINIAOD
/DoubleEG/Run2016(C-G)-03Feb2017-v1/MINIAOD
/DoubleEG/Run2016H-03Feb2017_ver2-v1/MINIAOD
/DoubleEG/Run2016H-03Feb2017_ver3-v1/MINIAOD
/DoubleMuon/Run2016B-03Feb2017_ver2-v2/MINIAOD
/DoubleMuon/Run2016(C-G)-03Feb2017-v1/MINIAOD
/DoubleMuon/Run2016H-03Feb2017_ver2-v1/MINIAOD
/DoubleMuon/Run2016H-03Feb2017_ver3-v1/MINIAOD
Datasets for background prediction
/MuonEG/Run2016B-03Feb2017_ver2-v2/MINIAOD
/MuonEG/Run2016(C-G)-03Feb2017-v1/MINIAOD
/MuonEG/Run2016H-03Feb2017_ver2-v1/MINIAOD
/MuonEG/Run2016H-03Feb2017_ver3-v1/MINIAOD
/JetHT/Run2016B-03Feb2017_ver2-v2/MINIAOD
/JetHT/Run2016(C-G)-03Feb2017-v1/MINIAOD
/JetHT/Run2016H-03Feb2017_ver2-v1/MINIAOD
/JetHT/Run2016H-03Feb2017_ver3-v1/MINIAOD
/MET/Run2016B-03Feb2017_ver2-v2/MINIAOD
/MET/Run2016(C-G)-03Feb2017-v1/MINIAOD
/MET/Run2016H-03Feb2017_ver2-v1/MINIAOD
/MET/Run2016H-03Feb2017_ver3-v1/MINIAOD

10.2 Triggers

The trigger selection for the searches presented in this thesis is driven by the requirement of at least two leptons at the HLT. Due to the changes in instantaneous luminosity, different dilepton triggers were active at different times and with varying prescales. This results in the need for a variety of triggers with slightly different requirements. Non isolated double lepton paths are included to increase the efficiency in events with large dilepton system p_T . Triggers with an isolation requirement on the leptons enable for the recording of lower p_T leptons. The p_T requirements are asymmetric and depend on the flavor composition of the triggers. Supporting triggers, with requirements on the jet H_T or online p_T^{miss} , are used for the study of the trigger efficiencies used for background prediction techniques, taken from hadronic events. Additionally, triggers with requirements on the presence of an electron and a muon, are used to collect a sample dominated by $t\bar{t}$ events for the same purpose. All signal and supporting triggers are documented in Table 10-2.

Table 10-2. Triggers used in the strong, electroweak and slepton searches. The first section are the triggers used in the signal regions, while the supporting triggers are used for the calculation of the trigger efficiencies of the signal triggers and for control regions.

Signal triggers
Dimuon triggers
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_v*
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_v*
HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_v*
HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ_v*
HLT_Mu27_TkMu8_v*
HLT_Mu30_TkMu11_v*
Dielectron triggers
HLT_Ele17_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v*
HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v*
HLT_DoubleEle33_CaloIdL_GsfTrkIdVL_v*
HLT_DoubleEle33_CaloIdL_GsfTrkIdVL_MW_v*
Supporting triggers
HLT_Mu8_TrkIsoVVL_Ele17_CaloIdL_TrackIdL_IsoVL_v*
HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_v*
HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v*
HLT_Mu17_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_v*
HLT_Mu23_TrkIsoVVL_Ele8_CaloIdL_TrackIdL_IsoVL_v*
HLT_Mu23_TrkIsoVVL_Ele8_CaloIdL_TrackIdL_IsoVL_DZ_v*
HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_v*
HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v*
HLT_Mu30_Ele30_CaloIdL_GsfTrkIdVL_v*
HLT_Mu33_Ele33_CaloIdL_GsfTrkIdVL_v*
HLT_PFHT125_v*
HLT_PFHT200_v*
HLT_PFHT250_v*
HLT_PFHT300_v*
HLT_PFHT350_v*
HLT_PFHT400_v*
HLT_PFHT475_v*
HLT_PFHT600_v*
HLT_PFHT650_v*
HLT_PFHT800_v*
HLT_PFHT900_v*

Several MC event generators are used to simulate the background and signal processes in this analysis, with the different parts of the generators introduced in Section 7.3. The simulation is normalized to luminosity using cross sections from <https://twiki.cern.ch/twiki/bin/viewauth/CMS/SummaryTable1G25ns>. The PYTHIA8 [110] package is used for parton showering, hadronization and underlying event simulation with the tune CUETP8M1, as described in Section 7.3. The various simulated samples used are presented in Appendix A.

10.3 Lepton pair selection

The principle behind the main background prediction method, that will be presented in the next chapter, relies on the lepton flavor symmetry of the W decay. For this reason, the identification and isolation requirements of the leptons are chosen so that they are as similar as possible between the flavors. This principle is reflected in the selections of the trigger requirement of the leptons of $p_T > 23, 17, 12,$ and 8 GeV depending on the exact path. Full efficiency for any of these values is reached at a p_T of $25(20) \text{ GeV}$ for the leading (trailing) lepton, which is the requirement on the leptons for the control region. One specific selection of this analysis is that not only the electrons are rejected if they appear in the transition region between the barrel and the endcap, but also the muons. The reason for this is that the flavor symmetric background is taken from $e\mu$ events, thus the necessity of having symmetric cuts not only on the efficiency but also on the fiducial regions. For this reason any lepton within the $|\eta|$ region of 1.4 to 1.6 is rejected, to keep the reconstruction of electrons and muons as similar as possible.

Since there are events with multiple lepton pairs, it is important to define an unambiguous way of selecting the “relevant” opposite-sign, same-flavor lepton pair. The implemented algorithm selects the two highest p_T leptons which are fully identified and that have a distance between them of 0.1 in ΔR . This is to say, there is no cross-cleaning or prioritization between electrons and muons applied, and non-identified leptons (including the ECAL transition region) do not enter in the consideration of the lepton pair selection.

The invariant mass, introduced in Section 5.3 of the leptons can be used to define the signal regions for the different searches. In the searches where an on-shell production of a Z boson is expected, naturally the signal region is defined to have the lepton pair compatible with the Z boson mass. This is the case for the electroweak superpartner search, and the colored superpartner search for gluinos. On the other hand, the colored SUSY search with a sbottom induced process and an intermediate slepton decay result in a lepton pair that is not compatible with the Z boson mass. This search instead utilizes bins in $m_{\ell\ell}$ outside of the Z mass window of 86-96 GeV. Finally, the search for direct slepton production includes two leptons that are not compatible with the Z mass. To suppress the Drell-Yan background, a generous veto on the Z boson mass of 76-106 GeV is applied in the signal region. All the signal and control region criteria on the leptons is summarized in Table 10-4.

Table 10-3. Lepton kinematic criteria.

Colored and Electroweak On-Z SR lepton selection	
Flavor	$e^+e^-/\mu^+\mu^-$
Leading p_T	$> 25 \text{ GeV}$
Subleading p_T	$> 20 \text{ GeV}$
$ \eta $	< 2.4 and $\exists [1.4, 1.6]$
$m_{\ell\ell}$	$\in [86, 96] \text{ GeV}$
Edge SR lepton selection	
Flavor	$e^+e^-/\mu^+\mu^-$
Leading p_T	$> 25 \text{ GeV}$
Subleading p_T	$> 20 \text{ GeV}$
$ \eta $	< 2.4 and $\exists [1.4, 1.6]$
$m_{\ell\ell}$	bins from 20 to 400+ GeV
Slepton SR lepton selection	
Flavor	$e^+e^-/\mu^+\mu^-$
Leading p_T	$> 50 \text{ GeV}$
Subleading p_T	$> 20 \text{ GeV}$
$ \eta $	< 2.4 and $\exists [1.4, 1.6]$
$m_{\ell\ell}$	$\exists [76, 106] \text{ GeV}$
CR lepton selection	
Flavor	$e^+e^-/\mu^+\mu^-/e^\pm\mu^\mp$
Leading p_T	$> 25 \text{ GeV}$
Subleading p_T	$> 20 \text{ GeV}$
$ \eta $	< 2.4 and $\exists [1.4, 1.6]$

10.4 Jet and b-jets

The jets used throughout this thesis are clustered from PF objects using anti- k_t algorithm with a distance parameter of 0.4, after excluding charged hadrons originating from pileup, as described in Section 8.4. Further, the jets are corrected with JECs, as described in Section 8.4. Identification of jets originating from b quark decays is done using the CSV v2 algorithm, introduced in Section 8.5. Further, the scalar sum of all jet p_T is referred to as H_T and is used to design some of the signal regions. The colored and electroweak superpartner searches have a requirement on the $|\Delta\phi|$ between the leading jets and the p_T^{miss} to be larger than 0.4 to reduce contamination from fake p_T^{miss} events that can be aligned with the jets. The p_T of the jets used for the control regions used for the background prediction methods is 35 GeV and the jets are required to be within $|\eta| < 2.4$. The jets are required to be separated from selected leptons by 0.4 in ΔR . In the slepton search no hadronic activity is expected, therefore a veto on jets of p_T greater than 25 GeV and $|\eta| < 2.4$ is imposed.

Table 10-4. Jet kinematic criteria.

Colored and electroweak SR jet selection	
AK4 jets ≥ 2	
p_T	$> 35 \text{ GeV}$
$ \eta $	< 2.4
$ \Delta\phi(\text{jet}_{1,2}, p_T^{\text{miss}}) $	> 0.4
b-tagged jets ≥ 0	
CSVv2 discriminator	Medium WP
p_T	$> 25 \text{ GeV}$
$ \eta $	< 2.4
Slepton SR jet selection	
AK4 jets = 0	
p_T	$> 25 \text{ GeV}$
$ \eta $	< 2.4
CR jet selection	
AK4 jets ≥ 2	
p_T	$> 35 \text{ GeV}$
$ \eta $	< 2.4

10.5 p_T^{miss}

Common to the various searches presented in this thesis is the existence of LSPs, as only R-parity conserving models are considered. The LSPs are massive and escape detection, and would manifest themselves through the imbalance in the transverse plane as p_T^{miss} . For this reason, all signal regions exploit the p_T^{miss} variable either explicitly through a binning in the variable, or implicitly in the M_{T2} variable that uses it as an input. The "Type-I corrected" p_T^{miss} is used, with all p_T^{miss} filters applied to remove anomalous p_T^{miss} events. Chapter 9 contains a thorough overview of the p_T^{miss} reconstruction algorithms and their performance.

10.6 M_{T2}

The leptonic M_{T2} variable is used to define all signal regions. It is a generalization of the transverse mass for pair-produced particles which decay into visible and invisible objects, as described in [111, 112]. The variable is efficient in differentiating SM $t\bar{t}$ from the signal final state scenarios, as it has a clear endpoint at the W boson mass for SM $t\bar{t}$. The idea behind this powerful variable is summarized below. Starting off with the leptonic decay of a top quark,

$$t \rightarrow W^+ b \rightarrow l^+ \nu b. \quad (10-1)$$

From a theoretical point of view, it is straight forward to construct the transverse mass (m_T) of the W boson using the kinematic variables of the lepton and the neutrino and energy and momentum conservation in the transverse plane,

$$M_T = \sqrt{m_l^2 + m_\nu^2 + 2(E_T^l E_T^\nu - \vec{p}_T^l \vec{p}_T^\nu)} \quad (10-2)$$

where the \vec{p}_T^l and \vec{p}_T^ν are the transverse momenta of the lepton and the neutrino, and E_T^l and E_T^ν the transverse energy of the same. In this case, the m_T has an endpoint very close to the W boson mass. In the case of dileptonic $t\bar{t}$,

$$pp \rightarrow t\bar{t} \rightarrow W^+ b + W^- \bar{b} \rightarrow l^+ \nu b + l^- \bar{\nu} \bar{b} \quad (10-3)$$

the situation gets more complicated. First of all, the pairing of the leptons and neutrinos that form the W bosons is not known. Further, the neutrinos escape detection and the momenta they carry is just collectively summed up in the p_T^{miss} . But let us disregard these experimental difficulties for a second. If one instead assume that the correct pairing is known, then one could create an upper bound dictated by the W boson mass according to

$$M_W^2 \geq \max\{M_T^2(\vec{p}_T^{l+}, \vec{p}_T^{\nu}), M_T^2(\vec{p}_T^{l-}, \vec{p}_T^{\bar{\nu}})\}. \quad (10-4)$$

Meaning that in the case of dileptonic $t\bar{t}$, if the lepton pairing is right the both transverse masses found would have endpoints very close to the W boson mass. In order to cope with the problem of not knowing the p_T of the neutrinos, one can instead scan over all possible p_T^{miss} partitions:

$$M_W \geq \min_{\vec{p}_T^{\text{miss}} = \vec{p}_{T1}^{\text{miss}} + \vec{p}_{T2}^{\text{miss}}} \left(\max\{M_T(\vec{p}_T^{l_1}, \vec{p}_{T1}^{\text{miss}}), M_T(\vec{p}_T^{l_2}, \vec{p}_{T2}^{\text{miss}})\} \right). \quad (10-5)$$

This is the formal definition of the M_{T2}

$$M_{T2} = \min_{\vec{p}_T^{\text{miss}} = \vec{p}_{T1}^{\text{miss}} + \vec{p}_{T2}^{\text{miss}}} \left(\max\{M_T(\vec{p}_T^{l_1}, \vec{p}_{T1}^{\text{miss}}), M_T(\vec{p}_T^{l_2}, \vec{p}_{T2}^{\text{miss}})\} \right). \quad (10-6)$$

For dileptonic $t\bar{t}$, the M_{T2} will be very close to the W boson mass, whereas for signal scenarios, where the leptons are paired with the momentum of the LSPs, which will result in values much larger than the W boson mass. The M_{T2} variable is used to define all signal regions in all searches, as it effectively reduce contributions from SM $t\bar{t}$ by imposing a cut around 80 GeV. The strength of this variable can be visualized in Figure 10-1, for a baseline signal region selection that is used for the search for colored superpartner production, in simulation. The M_{T2} requirement of 80 GeV is shown with a dashed line to indicate the reduction of SM backgrounds achieved above 80 GeV. As can be seen in Figure 10-1 the M_{T2} for dileptonic $t\bar{t}$ does not have a perfect endpoint at 80 GeV. This has to do with the possible mismeasurements of the jets, and other instrumental effects, that spoils the distribution and extends the $t\bar{t}$ contribution beyond the W mass endpoint.

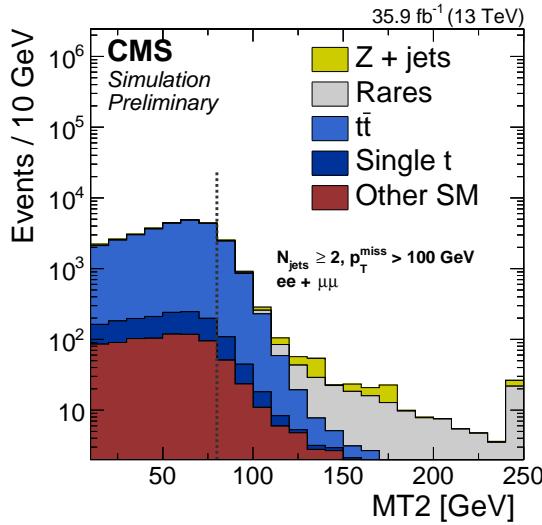


Figure 10-1. The M_{T2} distribution after requiring a minimum of 2 jets and $p_T^{\text{miss}} > 100$ GeV in addition to an opposite sign, same flavor lepton pair. The cut value of $M_{T2} = 80$ GeV that is used in most parts of the analysis is indicated by a grey dashed line [113].

10.7 $t\bar{t}$ likelihood discriminator

The main SM background in the search for sbottom pair production in the slepton edge model is $t\bar{t}$. The M_{T2} variable introduced above is doing a great job in reducing the backgrounds from $t\bar{t}$. However, a further reduction of the backgrounds and increase in sensitivity can be achieved by further controlling the $t\bar{t}$ background. To this end, a likelihood discriminant is developed that categorizes how $t\bar{t}$ -like an event is, and divide the events in $t\bar{t}$ -like and non- $t\bar{t}$ -like.

The idea is to use four, mostly uncorrelated, variables characteristic of the $t\bar{t}$ process, extract their shapes from fits in an opposite flavor control regions, and interpret each of the shapes as PDFs describing the probability of one event to be $t\bar{t}$ -like. To calculate this likelihood the probability density functions of the four observables are determined by maximum likelihood fits in the opposite flavor control sample in the same kinematic region as the same flavor SR. The four variables are:

- p_T^{miss}
- dilepton p_T
- $|\Delta\phi|$ between the leptons
- $\Sigma m_{\ell b}$

The $\Sigma m_{\ell b}$ is the sum of the invariant masses of the two lepton and b-tagged jet systems, and should have an endpoint at $2\sqrt{m_t^2 - m_W^2}$ for events resulting from top quarks. The $m_{\ell b}$ is found by calculating all possible invariant mass combinations between the leptons and the jets in the event. The combination that gives the smallest value of $m_{\ell b}$ is kept and the procedure is repeated for the remaining leptons and jets until the smallest value is found for that combination. The $\Sigma m_{\ell b}$ is the summation of the minimized $m_{\ell b}$'s. A priority in the calculation is given to the jets that are b-tagged, as these more accurately reflect the $t\bar{t}$ environment that is targeted. This means that if one or more b-tagged jets are present in the event, the $m_{\ell b}$ between the leptons and the b-tagged jets is minimized first, and then the remaining jets are considered for the minimization of the second leptons $m_{\ell b}$. The fit function for the p_T^{miss} spectrum is a sum of two exponentials. The $|\Delta\phi|$ is fitted with a second-order polynomial. A Crystal-Ball function is used to fit the dilepton p_T and $\Sigma m_{\ell b}$ spectra. The four variables in OF data and the fits are shown in Figure 10-2. For each SF event passing in the SR requirements, a likelihood function is constructed by multiplying the evaluation of each of the pdfs. The negative logarithm is then defined as the likelihood discriminant and denoted by NLL. The working point of $t\bar{t}$ -like and non- $t\bar{t}$ -like is chosen at an efficiency corresponding to roughly 95%, which translates to a value of NLL of 21.

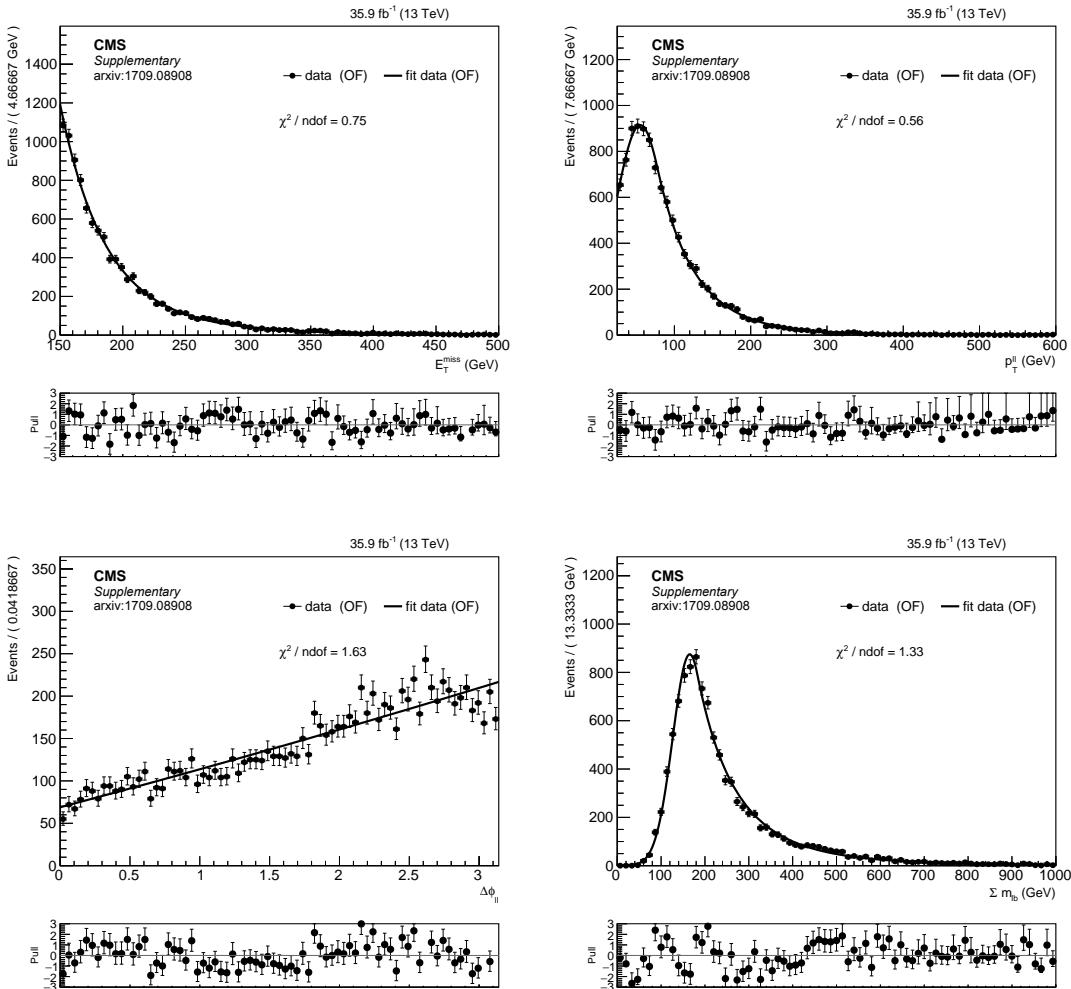


Figure 10-2. PDFs for the four input variables to the likelihood discriminant: p_T^{miss} (top left), dilepton p_T (top right), $|\Delta\phi|$ between the leptons (bottom left), and $m_{\ell b}$ (bottom right) [113].

11 — Background Estimations

The SM background processes in this search are well known and allows for the use of data-driven background estimation techniques. Further, some backgrounds are estimated from simulation with designated control region to validate the normalization.

The background composition differs depending on if the search involves the production of an on-shell Z boson selection (such as for the colored and electroweak superpartner searches) or not (such as for the Edge and direct slepton search). This results in, on the one hand, a significant contribution from Drell-Yan, on the other hand, a suppression thereof. In the signal region where Drell-Yan is the dominant background, a data-driven technique is used that predicts the p_T^{miss} spectrum in Z + jets events by the p_T^{miss} spectrum in γ +jets events. In the slepton search the Drell-Yan is heavily suppressed by stringent cuts on p_T^{miss} , number of jets and M_{T2} , and is thus estimated using simulation. What is common to all the searches is the contribution from processes containing one or more top quarks or W bosons that decay leptonically. These processes are referred to in the following as flavor symmetric, due to the flavor symmetry of the W boson decay. This background is estimated using a data-driven technique. Diboson processes, such as WZ and ZZ contribute in most SRs if one or two of the leptons are out of acceptance or are not well reconstructed. Together with the $t\bar{t}Z$ process, these processes are referred to as "Z+ ν " backgrounds, as they all contain prompt leptons and p_T^{miss} from a neutrino.

This chapter contains an overview of the SM background processes present in the searches, and is followed by a description on the background estimation techniques used to estimate them. All figures without citation are taken from the CMS Analysis Note AN-16-482.

11.1 Standard model processes

First, the physical processes that constitute the backgrounds in these searches are presented. Of course, there are many more SM processes produced at the LHC, such as the overwhelming QCD multijet background, but these and other processes are effectively suppressed by requiring the baseline selection of two leptons of opposite charge and same flavor. First, the SM processes that result in two of these leptons will be presented, followed by the technique used to estimate them.

Top related processes

Leptonically decaying top anti-top pair production provides a major background in all searches documented in this thesis, as the leptons are of opposite charge. Additionally, single top production can also result in this signature if one of the jets is misidentified as a lepton. Both processes are depicted in Figure 11-1 and involve a leptonically decaying W boson, and these backgrounds are predicted using the flavor symmetry of the W decay, described in the subsequent chapter. These backgrounds, together with others that will be presented later in this section, are referred to as Flavor Symmetric (FS) in this thesis. As these backgrounds all contain the production of a leptonically decaying W boson, they can be heavily suppressed by a cut at the W mass of the M_{T2} distribution.

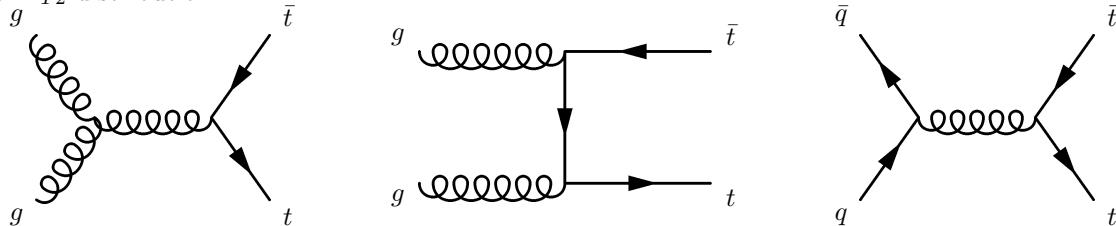


Figure 11-1. The leading order diagrams of $t\bar{t}$ production, through gluon fusion (left and middle) and quark-antiquark annihilation (right).

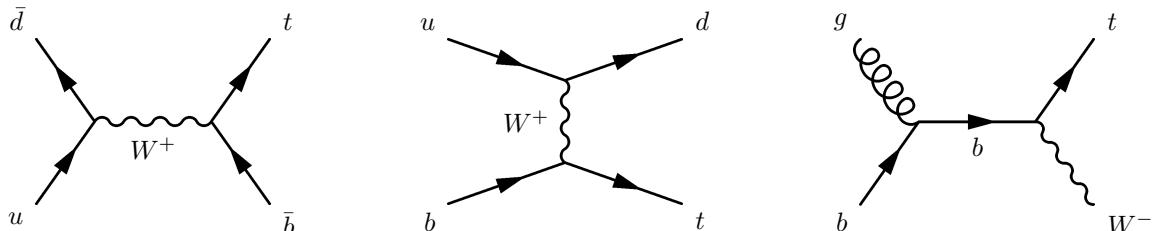


Figure 11-2. Single top quark production through s-channel (left), t-channel (middle) and in association with a W boson (right).

Drell-Yan

The large cross section Drell-Yan process, where a Z boson decays to two leptons, is a major background in the search for electroweak superpartners and the search for gluino production. Both these signal scenarios involve the production of an on-shell Z boson, and thus require that the lepton pair should be compatible with a Z boson. The first order Drell-Yan process is shown in Figure 11-3. This process contains no production of neutrinos, with the result that the p_T^{miss} is solely due to jet resolution and detector effects. This process will in the following be referred to as DY, DY+jets, Drell-Yan or Z+jets.

As was thoroughly described in the chapter on the p_T^{miss} performance, Chapter 9, the DY is very similar to the single photon production in the sense that it does not contain any real p_T^{miss} from neutrinos. Therefore, the p_T^{miss} contribution from DY can be estimated through the p_T^{miss} from single photon production. This

data-driven technique is referred to as " p_T^{miss} template method" and profits from the high statistical power of the single γ process. In the direct slepton production search, a Z boson veto is applied that heavily suppresses the DY, and the very minor contribution from DY is taken from simulation.

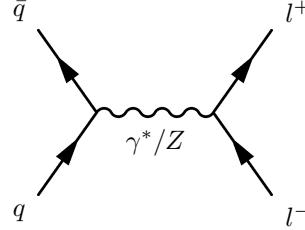


Figure 11-3. Leading order DY production.

Diboson production

Diagrams for diboson production are shown in Figure 11-6, where V_1 and V_2 are either W and W, Z and Z, or W and Z. The WW process, where both W bosons decay leptonically, is flavor symmetric and is estimated using the flavor symmetric background prediction method described in the subsequent chapters. The ZZ process and WZ process, when one Z boson decays to charged leptons and one Z boson decay to neutrinos, are dominant backgrounds in the direct slepton search as they fulfill the criteria of no hadronic activity and large p_T^{miss} from neutrinos. If both bosons decay leptonically, the WZ process can result in opposite sign same flavor pairs and enter the signal regions of all searches as a background if one of the leptons is out of η or p_T acceptance.

The ZZ and WZ are referred to as "Z + ν backgrounds" in this thesis and are estimated through simulation with translation factors derived from dedicated control regions. As the strong, electroweak and slepton searches have different kinematic features, the translation factors for these processes are derived separately for the three searches.

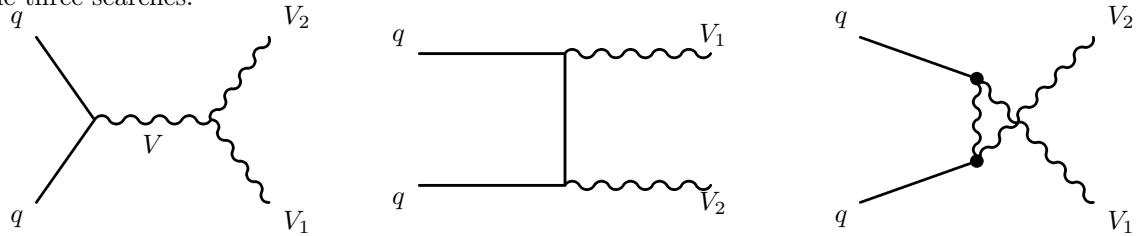
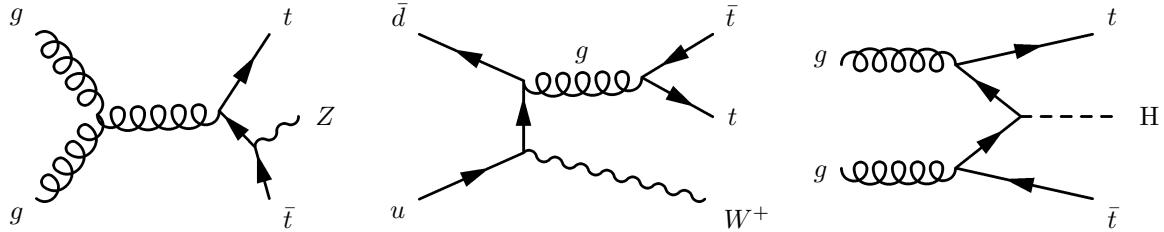
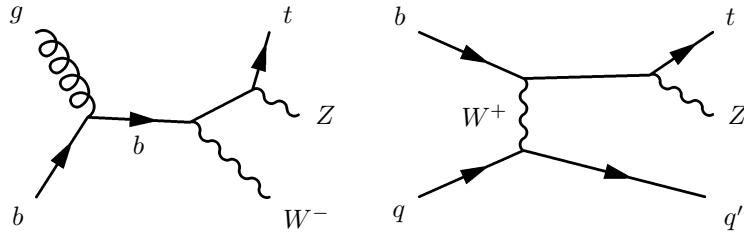


Figure 11-4. Leading order diboson production through s -channel (left), t -channel (middle) or u -channel (right).

Rare backgrounds

The processes categorized as rare are named according to their defining feature of having a low cross section. These processes are $t\bar{t}Z$, $t\bar{t}W$, $t\bar{t}H$, tWZ and tZq and have the commonality that they all can produce at least two leptons, jets and some p_T^{miss} from neutrinos, making them end up in the various signal regions, although with a very low rate.

Figure 11-5. Leading order $t\bar{t}Z$ (left), $t\bar{t}W$ (middle) and $t\bar{t}H$ (right) production.Figure 11-6. Leading order tWZ (left) and tZq (right) production.

11.2 Flavor-symmetric background estimation

Physical processes that feature decays into same flavor and opposite flavor lepton pairs with equal probability are comprised in this background estimation. This is, mainly the $t\bar{t}$ process, but also other processes such as WW , or $t\bar{t}W$ production.

The idea behind the background prediction technique is visualized in Figure 11-7 where a selection of same flavor and opposite flavor leptons is applied in an Edge-like signal region in simulation, showing that the $t\bar{t}$ spectrum has the same shape and magnitude in both selections. Other processes such as WZ , ZZ and $t\bar{t}Z$

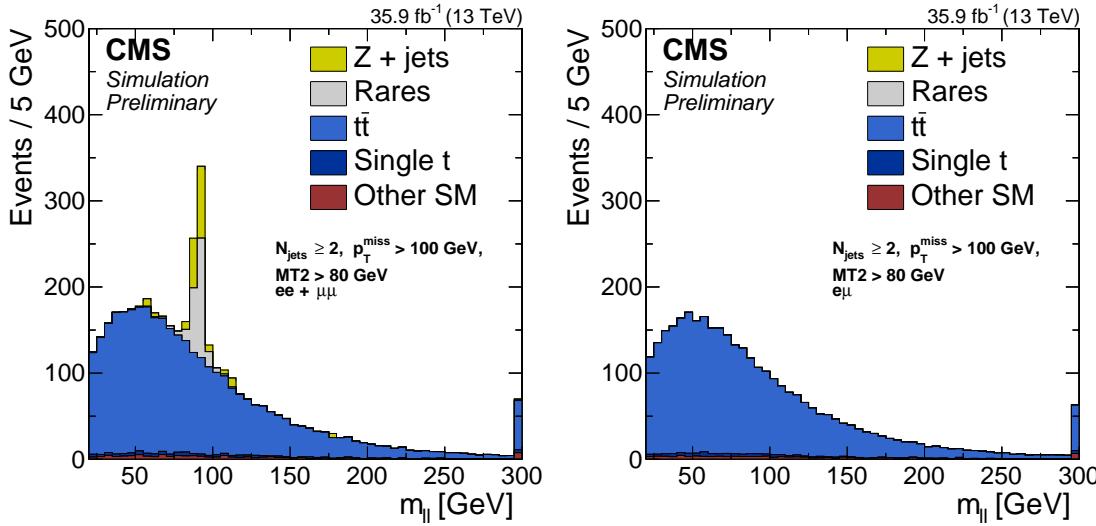


Figure 11-7. The dilepton invariant mass distribution after requiring a baseline SR definition for the Edge search in addition to an opposite sign, SF (left) or opposite sign OF (right) lepton pair [59].

have some flavor symmetric component and some component stemming from the decay of the Z boson. These processes are treated separately in Section 11.4 and are referred to as "Z+ ν " processes. The FS background

prediction in SF events utilizes the number of OF events as follows:

$$\begin{aligned} N_{ee} &= R_{ee/OF} \times N_{OF} \\ N_{\mu\mu} &= R_{\mu\mu/OF} \times N_{OF} \\ N_{SF} &\equiv N_{ee} + N_{\mu\mu} = R_{SF/OF} \times N_{OF} \end{aligned} \quad (11-1)$$

As dictated by physics, the ratios R should simply be $R_{ee/OF} = R_{\mu\mu/OF} = 0.5$ and $R_{SF/OF} = 1$. But since it is impossible to have a perfect detector, especially in such a large scale experiment, the ratios are slightly off from 0.5 and 1. This is due to trigger, reconstruction and identification efficiencies that are not identical for detected electrons and muons. So instead of just using the assumptions on the ratios to be 0.5 or 1, the ratios are instead measured to take into account these differences in efficiencies. In the rest of this section, a derivation on how the number of SF events can be estimated is presented. The following naming scheme is used: An epsilon without an upper index includes all efficiencies (trigger, selection and reconstruction), an epsilon with "R" as a superscript indicates that reconstruction and selection efficiencies are applied, an epsilon with "T" as a superscript indicates that trigger efficiencies are applied while the superscript "hard" stands for the quantity on particle level. An assumption is made on the reconstruction and selection efficiencies for the two leptons in the event to be uncorrelated, i.e. $\epsilon_{\ell\ell} = \epsilon_\ell \cdot \epsilon_\ell$. The ratio of efficiencies for muons to electrons is basically saying how well muons are measured with respect to electrons given certain trigger and reconstruction selection. The quantity can be expressed in the following way:

$$\begin{aligned} r_{\mu/e} &= \frac{\epsilon_\mu}{\epsilon_e} \approx \sqrt{\frac{\epsilon_{\mu\mu}^R \epsilon_{\mu\mu}^T}{\epsilon_{ee}^R \epsilon_{ee}^T}} = \sqrt{\frac{N_{\mu\mu}}{N_{ee}}} \\ r_{\mu/e} &\approx r_{\mu/e}^R \sqrt{\frac{\epsilon_{\mu\mu}^T}{\epsilon_{ee}^T}} \end{aligned} \quad (11-2)$$

From this, the number of ee and $\mu\mu$ events can be estimated in the following way:

$$\begin{aligned} N_{ee} &= \epsilon_{ee}^T N_{ee}^R = \epsilon_{ee}^T (\epsilon_e^R)^2 N_{ee}^{hard} = \frac{1}{2} \epsilon_{ee}^T (\epsilon_e^R)^2 N_{OF}^{hard} = \frac{1}{2} \epsilon_{ee}^T \frac{\epsilon_e^R}{\epsilon_\mu^R} N_{OF}^R \\ &= \frac{1}{2} \frac{1}{r_{\mu/e}^R} \frac{\epsilon_{ee}^T}{\epsilon_{e\mu}^T} N_{OF} = \frac{1}{2} \frac{1}{r_{\mu/e}} \frac{\sqrt{\epsilon_{ee}^T \epsilon_{\mu\mu}^T}}{\epsilon_{e\mu}^T} N_{OF} \end{aligned} \quad (11-3)$$

$$\begin{aligned} N_{\mu\mu} &= \epsilon_{\mu\mu}^T N_{\mu\mu}^R = \epsilon_{\mu\mu}^T (\epsilon_\mu^R)^2 N_{\mu\mu}^{hard} = \frac{1}{2} \epsilon_{\mu\mu}^T (\epsilon_\mu^R)^2 N_{OF}^{hard} = \frac{1}{2} \epsilon_{\mu\mu}^T \frac{\epsilon_\mu^R}{\epsilon_e^R} N_{OF}^R \\ &= \frac{1}{2} r_{\mu/e}^R \frac{\epsilon_{\mu\mu}^T}{\epsilon_{e\mu}^T} N_{OF} = \frac{1}{2} r_{\mu/e} \frac{\sqrt{\epsilon_{ee}^T \epsilon_{\mu\mu}^T}}{\epsilon_{e\mu}^T} N_{OF} \end{aligned} \quad (11-4)$$

As the final destination for this derivation is the prediction of SF yields, the above equalities can be combined to give an estimate of the N_{SF} :

$$\begin{aligned} N_{SF} &= \frac{1}{2} \left(r_{\mu/e} + \frac{1}{r_{\mu/e}} \right) \frac{\sqrt{\epsilon_{ee}^T \epsilon_{\mu\mu}^T}}{\epsilon_{e\mu}^T} N_{OF} \\ &= \frac{1}{2} \left(r_{\mu/e} + \frac{1}{r_{\mu/e}} \right) R_T N_{OF} = R_{SF/OF} N_{OF} \end{aligned} \quad (11-5)$$

This parameterization underlines the advantage of using the combined SF sample compared to the separate $e^\pm e^\mp$ and $\mu^\pm \mu^\mp$ samples. While $R_{\ell\ell/OF}$ is directly affected by the differences in reconstruction and trigger efficiencies by the factors $r_{\mu/e}$ or $r_{\mu/e}^{-1}$, these differences partially cancel out in $R_{SF/OF}$. In order to get to a final value for $R_{SF/OF}$, two approaches have been taken. The first is a direct measurement of the ratio $R_{SF/OF}$ in data in a control region enriched in $t\bar{t}$, while the second consists of the separated estimation of the $r_{\mu/e}$ and R_T factors. This section contains a description of the direct measurement method, followed by a description of the factorization method.

Direct measurement method

The direct measurement of the $R_{SF/OF}$ is the ratio of SF to OF events in a $t\bar{t}$ enriched control region. This region is defined orthogonal to the signal regions in terms of jets and p_T^{miss} . It lies just outside them in order to not be affected by a large extrapolation. The exact selections for the $t\bar{t}$ control region are:

- the same lepton selection
- exactly two jets
- p_T^{miss} between 100 and 150 GeV
- $70 \geq m_{\ell\ell} \geq 110$ GeV excluded

Even though all signal regions are defined with a M_{T2} cut, the choice is made to not define the $t\bar{t}$ CR with a M_{T2} requirement. The reason for this is the significant reduction of the $t\bar{t}$ background and thus the available statistics that the M_{T2} requirement would lead to. A large mass window around the Z boson mass is excluded to avoid the contamination with Drell-Yan backgrounds. The measurement consist of counting the number of ee, $\mu\mu$ and SF events in this control region, and taking the ratio of this number to the number of OF events in the same region, to construct the $R_{ee/OF}$, $R_{\mu\mu/OF}$ and $R_{SF/OF}$ separately.

The results are shown in Table 11-1 and the data and the simulation agree within 2% for the combination of both flavors. If the flavors are treated separately, the agreement between the data and the simulation is between 1 and 3%. As a closure test, the ratio of $R_{SF/OF}$ in the control and signal region (transfer factor) on simulation is studied. It is found to be compatible with unity within 2% for the combination of both flavors and for each flavor separately. To evaluate possible dependencies when extrapolating from control to

Table 11-1. Observed event yields in the control region and the resulting values for $R_{SF/OF}$, $R_{ee/OF}$, and $R_{\mu\mu/OF}$ for both data and MC. The transfer factor is defined as the ratio of $R_{SF/OF}$ in the signal region divided by $R_{SF/OF}$ in the control region.

	N_{SF}	N_{OF}	$R_{SF/OF} \pm \sigma_{\text{stat}}$
Data	13438	12138	1.107 ± 0.014
MC	13290	12189	1.090 ± 0.005
	N_{ee}	N_{OF}	$R_{ee/OF} \pm \sigma_{\text{stat}}$
Data	4976	12138	0.410 ± 0.007
MC	4852	12189	0.398 ± 0.003
	$N_{\mu\mu}$	N_{OF}	$R_{\mu\mu/OF} \pm \sigma_{\text{stat}}$
Data	8462	12138	0.697 ± 0.010
MC	8438	12189	0.692 ± 0.004

signal region and determine a systematic uncertainty, the ratio $R_{SF/OF}$ is studied by relaxing each of the control region variables and check the evolution of the $R_{SF/OF}$ as a function of the same variable, in data and simulation. The data shows some fluctuations due to the limited statistics in the control region, but no

clear trends are observed. These studies are used to assign a systematic uncertainty as the variation of the R_{SF}/OF that would be sufficient to cover the differences between the data and the simulation. It is found that a variation 4% is enough cover any potential fluctuations and is assigned as a systematic uncertainties.

Factorization method

The second method to predict the number of SF events in the SR is called the Factorization method and explicitly takes into account the trigger and reconstruction efficiencies of the leptons. In order to calculate R_{SF}/OF according to Equation 11-7, $r_{\mu/e}$ and R_T are measured in data.

Measurement of $r_{\mu/e}$

The measurement of $r_{\mu/e}$ is performed in a high statistics DY enriched control region summarized below:

- the same lepton selection
- more than or exactly two jets
- p_T^{miss} below 50 GeV
- $60 \geq m_{\ell\ell} \geq 120$ GeV

A small dependency on $r_{\mu/e}$ is found as a function of the second lepton p_T is observed, especially at low p_T where the muons are reconstructed at a higher efficiency than electrons. In order to correct for this trend, a parameterization of this dependency is performed. The following parameterization as a function of the p_T of the second lepton was chosen:

$$r_{\mu/e} = a + \frac{b}{p_T} \quad (11-6)$$

a and b are constants that are determined in a fit to data. The fit for this parameterization is shown in Figure 11-8. The determined parameters are stated in Table 11-2. Since the effect is mainly at very low lepton p_T

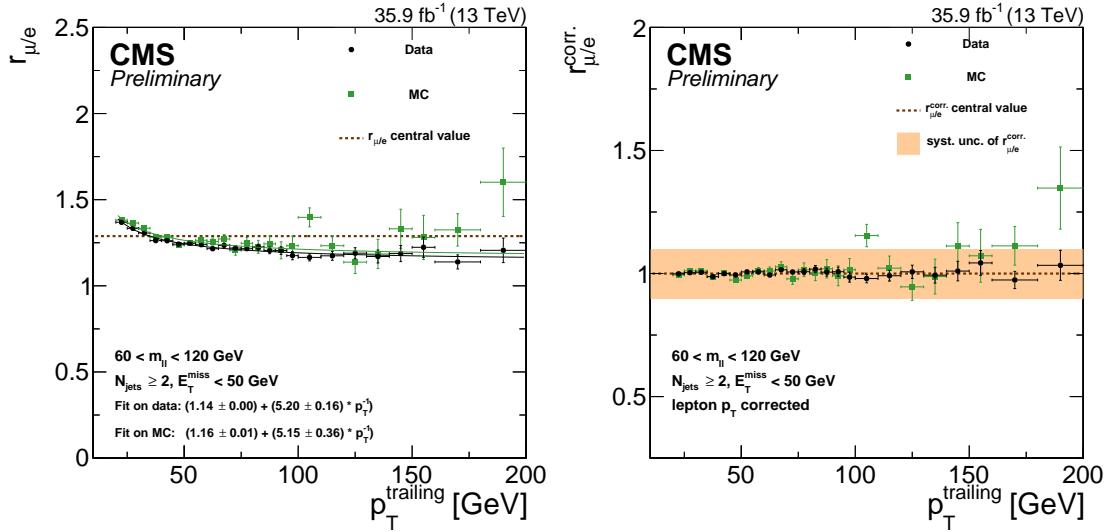


Figure 11-8. The $r_{\mu/e}$ dependency on the p_T of the second lepton for data and MC. The plots show the values for $r_{\mu/e}$ before (left) and after (right) the parameterization on the second lepton p_T is propagated to the dielectron events. The central value in the left plot indicates the $r_{\mu/e}$ value that would be obtained without the parameterization. The fit values of the parameterization are shown in the left plot.

Table 11-2. Result of the fit of $r_{\mu/e}$ as a function of the p_T of the trailing lepton in the DY control region. The same quantities derived from simulation are shown for comparison. Only statistical uncertainties are given.

	a	b
Data	1.14 ± 0.01	5.20 ± 0.16
MC	1.16 ± 0.01	5.15 ± 0.36

and less than 5% of the events fall into this category, a flat systematic uncertainty of 10% is chosen to cover this remaining trend instead of doing a multi-variable fit. An uncertainty band indicates this 10% uncertainty in Figure 11-8 (right) that shows the corrected $r_{\mu/e}^{corr.}$ for the second lepton p_T . Further dependencies of $r_{\mu/e}^{corr.}$ on important observables are studied and no significant trends are observed. The chosen 10% systematic uncertainty is sufficient to cover any remaining trends in these observables. In addition to the studies on the dependence of $r_{\mu/e}^{corr.}$ on different variables, the dependency on the $0.5(r_{\mu/e}^{corr.} + 1/r_{\mu/e}^{corr.})$, which is the factor that actually goes into the prediction of the N_{SF} as seen in Equation 11-7, on some SR variables is also performed. This is especially important as the control region where the $r_{\mu/e}$ is measured is far from the signal regions. In Figure 11-9 the results of some of these studies can be observed where the results on data are shown in black and simulation in green. The central value observed on data is indicated by a dashed line. To determine the systematic uncertainty the fit parameters a and b are varied by its statistical uncertainty and the full prediction is shifted up and down by the 10% systematic uncertainty from the studies above. The relative change of the mean value is used as the systematic uncertainty of the method. The quantity $0.5(r_{\mu/e}^{corr.} + 1/r_{\mu/e}^{corr.})$ is especially stable with respect to p_T^{miss} and M_{T2} , as can be seen in Figure 11-9, validating the extrapolation from the DY control region to the signal region.

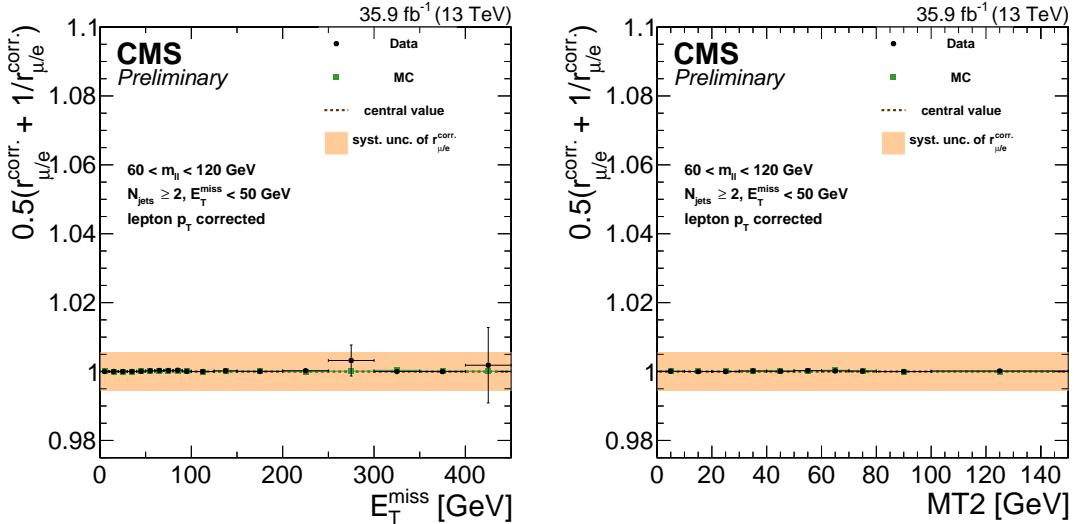


Figure 11-9. The $0.5(r_{\mu/e}^{corr.} + 1/r_{\mu/e}^{corr.})$ dependency on p_T^{miss} (left) and M_{T2} (right) for data and MC after the parameterization on the second lepton p_T is propagated to the dielectron events. The uncertainty introduced by shifting the fit parameters by its statistical uncertainty and the full prediction by 10% is indicated by the orange band.

Measurement of R_T

The second ingredient in the Factorization method for predicting the N_{SF} in the signal region is the R_T , which is a measure of the trigger efficiencies. The R_T is defined through the trigger efficiencies, ϵ_{ll}^T , according to:

$$R_T = \frac{\sqrt{\epsilon_{ee}^T \epsilon_{\mu\mu}^T}}{\epsilon_{e\mu}^T} \quad (11-7)$$

The trigger efficiencies are measured using a data sample collected with Particle Flow H_T triggers, in which dilepton events are selected. This way, the efficiencies of the lepton triggers can be studied by comparing how many of the dilepton events in the H_T triggered data sample was also picked up by the dilepton triggers. The PF H_T triggers used have thresholds between 125 GeV and 900 GeV, and the dilepton triggers studied are listed in Table 10-2 in Section 10.2. The dilepton trigger efficiency is calculated as the fraction of events in the lepton unbiased sample that also passes the dilepton triggers for the given flavor combination:

$$\epsilon_{ll}^T = \frac{ll \text{ selection} \cap H_T \text{ trigger} \cap ll \text{ trigger}}{ll \text{ selection} \cap H_T \text{ trigger}}. \quad (11-8)$$

No particular requirement is applied to the data sample, except all events in the signal regions or in the $t\bar{t}$ control region used for the direct measurement of $R_{SF/OF}$ are excluded. In terms of number of jets and p_T^{miss} , this translated to $N_{\text{jets}} \geq 2$ and $p_T^{\text{miss}} \geq 100$ GeV, and in order to be fully efficient in terms of the H_T triggers, a requirement on the offline $H_T \geq 200$ GeV is imposed. The result of the trigger efficiencies measurement in data and simulation is shown in Table 11-3. The number of events is much larger in simulation as no H_T cross triggers are required. The efficiencies in data are about 96% for dielectron triggers, 95% for dimuon triggers and 91% for electron-muon triggers. This leads to a value for R_T of 1.052, from the equality in Equation 11-8. The efficiencies for dimuon and electron-muon on MC are 2-3% higher, but $R_T = 1.045$ is close to the value on data. A systematic uncertainty of 3% is chosen and assigned to each trigger efficiency, which corresponds to the maximal deviation between the efficiencies in data and simulation. This results in a total uncertainty of about 4% on R_T after error propagation of the individual uncertainties on the efficiencies. Similarly to the direct measurement of $R_{SF/OF}$ and $r_{\mu/e}$, the dependence on some signal region variables is studied, and is shown in Figure 11-10. No significant dependency of R_T on any event property is observed and the chosen systematic uncertainty is sufficient to cover any fluctuation in data and MC. The results

Table 11-3. Trigger efficiency values for data and MC with OS, $p_T > 25(20)$ GeV and $H_T > 200$ GeV.

Data			MC			
nominator	denominator	$\epsilon_{\text{trigger}} \pm \sigma_{\text{stat}}$	nominator	denominator	$\epsilon_{\text{trigger}} \pm \sigma_{\text{stat}}$	
ee	12070	12584	0.959±0.002	111787	116654	0.958±0.001
$\mu\mu$	8741	9230	0.947±0.002	190067	194687	0.976±0.001
e μ	2437	2690	0.906±0.006	43069	46520	0.926±0.000
R_T	1.052 ± 0.043		1.045 ± 0.041			

from the direct and factorization methods for the FS background prediction have comparable uncertainties and central values, and can thus be combined using the weighted average. For the factorization method, no constant result for the $R_{SF/OF}$ can be given, due to the parameterization of $r_{\mu/e}$ that has to be applied on an event-to-event basis. Instead, the factorization method is applied in each SR where the obtained prediction is divided by the number of OF events to get an $R_{SF/OF}$ factor for the factorization method in this region. This factor is combined with $R_{SF/OF}$ from the direct measurement method by using the weighted average. The $R_{SF/OF}$ values obtained by the direct measurement are the same for all signal regions as they are derived in

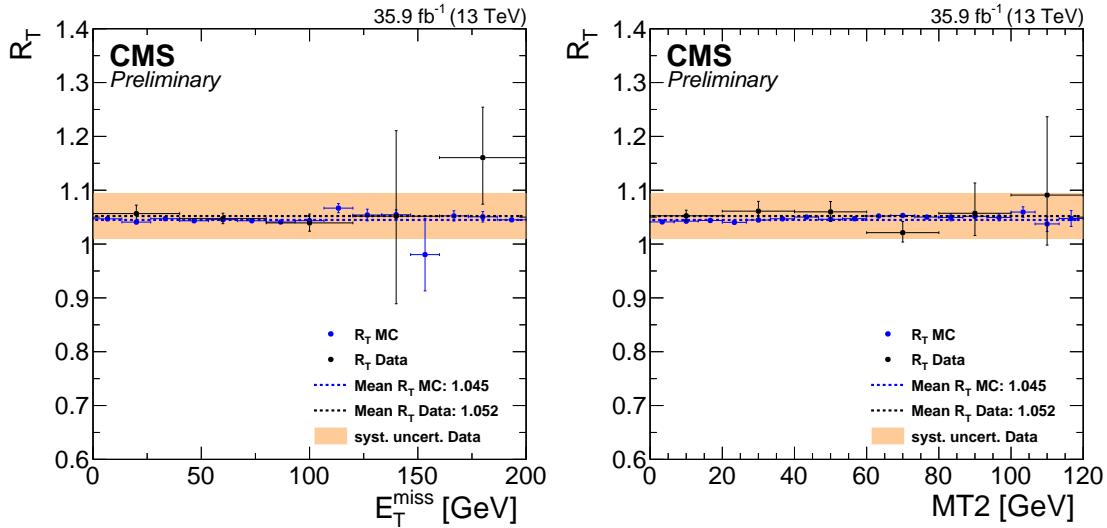


Figure 11-10. Dependency of the R_T ratio on the p_T^{miss} (left) and M_{T2} (right), for data and MC. The systematic uncertainty of about 4% on R_T is indicated by the orange band.

one control region, and are given in Table 11-1. The different treatment of the FS background estimation needed for each analysis strategy will be described in Chapters 14, 15 and 16.

11.3 Drell-Yan background estimation

The Drell-Yan process, i.e. Z+jets process, enter all signal regions in the searches presented in this thesis. The p_T^{miss} in the Z+jets process is solely due to the hadronic recoil and more precisely due to mismeasured jets. A more detailed treatment of the p_T^{miss} in these events is presented in the chapter on p_T^{miss} performance in Chapter 9. What can be seen there is that the shape of the p_T^{miss} spectrum in Z+jets is the same as in γ +jets events, with a difference in the normalization due to the higher cross section associated with the γ +jets process. This similarity in the shape of the p_T^{miss} spectrum is exploited in this thesis, by estimating the Z+jets in the signal region from a data sample of γ +jets in the signal region.

A single photon data sample is collected using single photon triggers with p_T thresholds of 22–165 GeV. The data collected with the prescaled triggers is corrected by the associated prescale values for each data taking period and run. As there are differences between the mass of the two bosons, the γ p_T spectrum is reweighted to match the Z spectrum in each signal region to account for this difference. An orthogonal control region with p_T^{miss} 50–100 GeV is used to normalize the γ +jets sample to match the Z+jets sample. Although suppressed by the lower cross sections and a lepton veto, electroweak processes with real p_T^{miss} such as $W\gamma$ can contribute significantly to the p_T^{miss} tails in the single γ sample. The contributions from electroweak processes are subtracted from the photon p_T^{miss} spectrum using simulation. The modeling of the simulated electoral processes is validated in an orthogonal control region containing one muon and one photon. There is good agreement observed in this control region and a conservative 30% systematic uncertainty is assigned to the electroweak subtraction procedure.

As the M_{T2} variable is used to define all signal regions, the M_{T2} has to be constructed in the γ +jets data sample. But since two visible objects are needed when calculating the M_{T2} , a special treatment is needed since there is only one photon in the events. The $M_{T2}(\ell\ell)$ can be emulated in the γ +jets sample by a simulated decay of the photon to two leptons, and where the $M_{T2}(\ell\ell)$ is calculated with the two decayed leptons as the visible objects. This decay is done by assuming the mother particle has the mass of a Z boson and the p_T of the photon reconstructed from data. The angular distribution of the leptons is accounted for by assuming a scenario where the direction of the spin of the mother particle is sampled from a distribution that is uniformly distributed in $1 + \cos^2(\theta)$, where θ is the polar angle in the reference frame in which the mother particle is at rest. After the photon is decayed, the same p_T and η requirements that are applied to the Z+jets events are applied to the decay products from the photon. $M_{T2}(\ell\ell)$ is constructed using these leptons, and the same cut is applied as in each signal region. Finally, the p_T distribution is reweighted as described previously.

The final systematic uncertainty assigned to this method is stemming from four sources. These sources are the statistical uncertainty of the γ +jets data sample, a closure method performed in simulation, the normalization performed in the low p_T^{miss} bin and the electroweak subtraction procedure. The closure test in simulation is evaluated separately in each of the signal regions. The systematic uncertainty is taken as the larger of the MC statistical uncertainty or non-closure and varies between 10 – 80% for the various signal regions. The statistical uncertainty from the normalization of the template prediction in data in each signal region (using p_T^{miss} between 50 – 100 GeV) as a systematic uncertainty on the prediction. This uncertainty ranges from 7 – 30% depending on the signal region. The uncertainty on the electroweak subtraction procedure is chosen to be 30% as described previously in this section. The different treatment of the Drell-Yan background needed for each analysis strategy will be described in Chapters 14, 15 and 16.

As mentioned in Section 4.3, the search for colored and electroweak superpartners was done in collaboration with researchers at five institutes. The p_T^{miss} template method described in this section summarizes the contribution from the group at University of California at San Diego. Although it is originally the intellectual

property of researchers at UCSD, it is still highlighted and described in this thesis as it plays a significant role in the searches for SUSY with opposite sign same flavor leptons. Worth pointing out is that significant contributions was made to the method by the group of ETH Zürich, for example through the idea of decaying the photon to two visible objects. As the Drell-Yan process is highly suppressed in the search for direct slepton production, the p_T^{miss} template method is thus not used in this particular search.

11.4 $Z+\nu$ background estimation

The group of SM backgrounds called $Z+\nu$ processes have in common that they all produce a dilepton pair compatible with the Z boson and one or more neutrinos from a Z or W decay. These processes are not strictly flavor symmetric, as more SF lepton pairs are produced than OF lepton pairs, and they are not DY processes where the p_T^{miss} is stemming from instrumental effects. There are many SM processes like this, and the highest cross section ones that contribute significantly to the various signal regions are WZ , ZZ and $t\bar{t}Z$. The WZ produce three charged leptons at particle level and one neutrino and is through the third lepton veto in the signal region heavily suppressed. However, if the third lepton is for some reason not reconstructed, this process enter the signal regions. The ZZ process produce two charged leptons and two neutrinos, and thus enter the signal regions through the large p_T^{miss} from the neutrinos. This process contribute significantly to the slepton signal region, as to first order no jets are produced.

The $t\bar{t}Z$ process produce two leptons compatible with a Z boson, and jets and p_T^{miss} from the decay of the top quarks. Although this is not a high cross section process, it still enters the on- Z signal regions. All these processes are estimated from simulation, but as the signal region phase space is low in statistics and not always well modeled, it is important to validate the simulation in dedicated orthogonal control regions. The control region validation for the three different processes is presented in the following, and with slightly different treatment depending on the signal region.

$WZ \rightarrow 3l\nu$ background

Since the signal models considered should only contain a final state with two leptons, a veto is applied on events containing more than two leptons, more precisely in the form of additional isotracks detailed in Section 8.6. As these additional isotracks are required to have a $p_T > 5$ GeV and be within $|\eta| < 2.4$, there is still a possibility for these leptons to not be reconstructed if they fall out of acceptance, and thus the event is not vetoed. This kind of background is dominated by the $WZ \rightarrow 3l\nu$ process. The reason why the events fail the third lepton veto is due to the low p_T leptons that fail the p_T requirement of 5 GeV. What is also worth noting is that the larger fraction of these events have a third lepton originating from a Z boson. In the cases where the lost lepton is from the Z boson, the resulting lepton pair contribution is FS, meaning SF lepton pairs are produced at the same rate as OF lepton pairs. This contribution is covered by the data-driven FS background prediction technique presented in Section 11.2. On the other hand, if the lost lepton is originating from the decay of the W boson, the leptons from this background process is dominantly same flavor. In order to treat these backgrounds in a correct way and avoid double counting contributions, the leptons from the WZ process is estimated from simulation only if both leptons are originating from a Z boson decay.

To validate that the simulation is properly modeling the p_T^{miss} tails where the signal regions are defined, an orthogonal control region is constructed by inverting the third lepton veto. By requiring 3 well identified leptons, the region is orthogonal to the signal region. The rest of the requirements are summarized below, where the idea is to obtain a region pure in WZ contribution while keeping the definition close to the signal region definition. As the signal regions for the slepton search differ from the rest of the searches in terms of the jet requirement, the validation of the WZ simulation is done in two separate control regions. One control

region is closer to the on-Z searches and require more than 2 jets and the other one is closer to the slepton signal region and require no jets. From these different control regions, transfer factors are derived for the different searches, that are later applied to the simulation in the respective signal regions. The control regions and the final transfer factors are introduced for the various analysis strategies presented in Chapters 14, 15 and 16.

$ZZ \rightarrow 2l2\nu$ background

The second type of $Z+\nu$ background in this search is stemming from the production of two Z bosons decaying to two leptons and two neutrinos. Although this process is suppressed by the jet requirement in the on-Z signal regions, it can still contribute as it is a process containing a production of a leptonically decaying Z boson. In the slepton search, on the other hand, the contribution is enhanced as to first order no jets are produced. But since the slepton signal region is designed with a veto on leptons compatible with a Z boson, this process can only enter the signal region if the Z boson is off-shell and produce leptons with an invariant mass far away from the on-shell Z boson mass. The simulation is validated for both scenarios separately, and transfer factors are derived for the on-Z and slepton searches separately. Similarly as to the $WZ \rightarrow 3l\nu$ background, the control regions and the final transfer factors are introduced for the various analysis strategies presented in Chapters 14, 15 and 16.

11.5 Rare background estimation

A final category of Standard Model backgrounds are rare processes, such as $t\bar{t}W$, $t\bar{t}H$, tWZ and tZq . These processes are neither flavor symmetric nor have all their p_T^{miss} from instrumental effects. Instead they can all result in final states with more than two or more leptons, jets and neutrinos, and thus fall in to the various signal regions that require two or more jets. Triboson processes, such as WWZ , WZZ and ZZZ contribute in the signal regions of the direct slepton search. What is common to all these processes is that they have very low cross sections. The contribution is taken from simulation and a conservative systematic uncertainty of 50% is assigned. The simulated samples and the cross sections of these processes are summarized in Appendix A.

11.6 Fake lepton background estimation

The attentive reader could now point out that a class of backgrounds is missing, namely that due to so called fake leptons. Due to inefficiencies in reconstruction procedures, or due to faulty detector signals, lepton candidates can be reconstructed that in fact does not correspond to a genuine lepton. Additionally, other real particles, such as light-flavor quarks or gluon jets may be reconstructed as a light lepton. In order for events with one fake lepton to enter the signal region, they have to be reconstructed along with a real lepton.¹ The standard model background of $W+\text{jets}$, with the W decaying leptonically and one of the jets faking a lepton could potentially lead to a signal like event. However, as this process involves the flavor democratic leptonic decay of the W boson, this contribution will in the end be flavor symmetric, and thus taken into account by the flavor symmetric background prediction method.

¹ The probability for two fake leptons in one event is just too small to be considered a sizeable background.

12 — Statistical Analysis

The primary goal of a search for new physics, is to *discover* new physics through an excess of events compatible with a particular model. If no excess of events is observed, the goal is shifted from a *discovery* to an *exclusion*, where the results are used to set exclusion limits on the models. As mentioned in the introduction of the thesis, the opposite sign lepton final state is a powerful search tool as it can target many SUSY production modes.

This chapter contains a short recap of the signal models probed in the thesis, followed by an introduction to the statistical analysis used for the limit setting. Concepts like the maximum likelihood and hypothesis testing are covered, and how they contribute to the limit setting procedure.

12.1 Signal models

The SUSY signal models treated in this thesis are introduced in Section 4. Depending on the model, exclusion limits are set on either one SUSY particle, represented as a 1D band, or two SUSY particles represented in a 2D plane of one SUSY mass parameter versus the other. As the SUSY mass spectrum is potentially very complex, exclusion limits can only be determined for one or two SUSY particles for each model, specified under the "Free parameters" column in Table 12-1. Additionally, assumptions has to be made on the other SUSY particles. One common assumption is that all SUSY particles other than those interpreted in are heavy and decoupled. Other assumptions, such as on the masses of the SUSY particles present in the decay chains are summarized under the "Other parameters" column in Table 12-1. In general, the production cross

Table 12-1. Summary of SUSY models and the sparticle masses that are constrained.

Signal model	Free parameters	Other parameters	Targeted by region:
GMSB	$m_{\tilde{g}}, m_{\tilde{\chi}_1^0}$	$m_{\tilde{\chi}_1^0} = 1 \text{ GeV}$	Strong on-Z (6 SRs)
Edge	$m_{\tilde{b}}, m_{\tilde{\chi}_2^0}$	$m_{\tilde{\chi}_1^0} = 100 \text{ GeV}, m_{\tilde{\ell}} = 0.5(m_{\tilde{\chi}_2^0} + m_{\tilde{\chi}_1^0})$	Edge SRs (18 SRs)
Wino	$m_{\tilde{\chi}_1^\pm}, m_{\tilde{\chi}_2^0}, m_{\tilde{\chi}_1^0}$	$m_{\tilde{\chi}_1^\pm} = m_{\tilde{\chi}_2^0}$	VZ SRs (4 SRs)
Higgsino	$m_{\tilde{\chi}_1^0}$	$m_{\tilde{\chi}_1^\pm} \approx m_{\tilde{\chi}_2^0} \approx m_{\tilde{\chi}_1^0}, m_{\tilde{\chi}_1^0} = 1 \text{ GeV}$	VZ/ZH SRs (7 SRs)
Slepton	$m_{\tilde{\ell}}, m_{\tilde{\ell}_L}, m_{\tilde{\ell}_R}$	-	Slepton SRs (4 SRs)
Selectron	$m_{\tilde{e}}, m_{\tilde{e}_L}, m_{\tilde{e}_R}$	-	Slepton SRs (4 SRs)
Smuon	$m_{\tilde{\mu}}, m_{\tilde{\mu}_L}, m_{\tilde{\mu}_R}$	-	Slepton SRs (4 SRs)

sections are computed at NLO plus next-to-leading-log (NLL) precision [47, 114, 115]. Two possible scenarios are assumed for the mixing of the binos, winos and higgsinos to form the $\tilde{\chi}_1^\pm$ and $\tilde{\chi}_2^0$. The higgsino case, where the $\tilde{\chi}_1^\pm$ and $\tilde{\chi}_2^0$ are mainly composed of higgsinos, the production cross sections are computed at NLO plus next-to-leading-log (NLL) precision in a limit of mass-degenerate higgsino $\tilde{\chi}_1^\pm$, $\tilde{\chi}_2^0$, and $\tilde{\chi}_1^0$ [116–118]. In the other case, the wino case, where the $\tilde{\chi}_1^\pm$ and $\tilde{\chi}_2^0$ are mainly composed of winos and binos, the production cross sections are computed at NLO plus next-to-leading-log (NLL) precision in a limit of mass-degenerate wino $\tilde{\chi}_2^0$ and $\tilde{\chi}_1^\pm$, light bino $\tilde{\chi}_1^0$ [116–118]. For the sleptons, the production cross sections are computed at NLO plus next-to-leading-log (NLL) precision for any single generation of left- or right-handed selectrons and smuons [116, 119]. For all these cases, all the other sparticles assumed to be heavy and decoupled.

The statistical analysis is outlined in the following section. The goal of the section is to introduce the CL_s method used for interpretation, which is done through introducing a couple of essential concepts. The model parameter estimation is first introduced, followed by a description of the hypothesis testing, which is performed in order to determine if a model should be rejected or accepted. The section is followed by a description of the various sources of systematic uncertainties. A more detailed discussion of the statistical methods can be found in [120].

12.2 Maximum likelihood

As outlined in [120], the outcome of an experiment can be modeled as a set of random variables x_1, \dots, x_n whose distribution takes into account both theoretical and experimental effects. Theory and detector effects affecting the measurement are described according to *nuisance parameters* $\theta_1, \dots, \theta_k$. The *likelihood* function is the overall probability distribution function evaluated for the observations x_1, \dots, x_n and can be written as:

$$\mathcal{L} = f(x_1, \dots, x_n, \theta_1, \dots, \theta_k) \quad (12-1)$$

For N independent measurements, such as events in a high energy physics collision, the likelihood function can be written as:

$$\mathcal{L} = \prod_{i=1}^N f(x_1^i, \dots, x_n^i, \theta_1, \dots, \theta_k) \quad (12-2)$$

From this, the *extended likelihood* can be defined as:

$$\mathcal{L}' = P(N, \theta_1, \dots, \theta_k) \prod_{i=1}^N f(x_1^i, \dots, x_n^i, \theta_1, \dots, \theta_k) \quad (12-3)$$

with $P(N, \theta_1, \dots, \theta_k)$ the distribution of N . The distribution is Poissonian and defined as:

$$P(N, \theta_1, \dots, \theta_k) = \frac{\lambda(\theta_1, \dots, \theta_k)^N e^{-\lambda(\theta_1, \dots, \theta_k)}}{N!} \quad (12-4)$$

where λ is the number of expected events in a bin containing contributions from both signal and background processes, such that $\lambda = \mu s + b$. Here b denotes the number of background events, s the number of signal events, and the μ represents the signal strength scaling that is 0 for a null hypothesis (null hypothesis) and 1 for a nominal signal hypothesis (test hypothesis). Putting all this together, the extended likelihood for a Poissonian process looks like:

$$\mathcal{L}'(\vec{x}; s, b, \vec{\theta}) = \frac{(s+b)^N e^{-(s+b)}}{N!} \prod_{i=1}^N \left(f_s P_s(x_i; \vec{\theta}) + f_b P_b(x_i; \vec{\theta}) \right) \quad (12-5)$$

where $\vec{x} = x_1, \dots, x_n$ and $\vec{\theta} = \theta_1, \dots, \theta_n$, and the f_s and f_b are the fraction of signal and background events, defined as:

$$f_s = \frac{s}{s+b} \quad (12-6)$$

$$f_b = \frac{b}{s+b} \quad (12-7)$$

The P_s and P_b are the PDFs for signal and background. Now Equation 12-5 becomes:

$$\mathcal{L}(\vec{x}; s, b, \vec{\theta}) = \frac{e^{-(s+b)}}{N!} \prod_{i=1}^N \left(s P_s(x_i; \vec{\theta}) + b P_b(x_i; \vec{\theta}) \right) \quad (12-8)$$

The $\mathcal{L}(\vec{x}; s, b, \vec{\theta})$ clearly a function of the nuisance parameters, and the values of $\vec{\theta}$ that maximize the likelihood, the *maximum likelihood* are said to fit the observation best.

12.3 Hypothesis testing

Hypothesis testing is the question whether some observed data sample is compatible with one theory model or another alternative one. Now that the maximum likelihood has been introduced, the frequentist test of a hypothesis can be performed. The level of agreement between the test hypothesis and the observation is determined through the definition of a test statistic $q(\mu)$, which is a ratio of maximized likelihoods:

$$q(\mu) = -2 \ln \frac{\mathcal{L}(\text{data}|\mu, \hat{\theta}(\mu))}{\mathcal{L}(\text{data}|\hat{\mu}, \hat{\theta}(\mu))}, \quad \hat{\mu} \leq \mu \quad (12-9)$$

The μ in the numerator is a fixed number and only the nuisance parameters θ are allowed to vary to obtain a maximum of the likelihood, where the $\hat{\theta}(\mu)$ is the value that maximizes \mathcal{L} . In the denominator both the signal strength μ and the nuisance parameters are allowed to vary to maximize the likelihood, which happens at the $\hat{\mu}$ and $\hat{\theta}$, that are known as maximum likelihood estimators. The $\hat{\mu}$ is constrained by an upper bound, μ , providing the test statistic a one-sided upper limit. The $q(\mu)$ is one number, that can be calculated for the observed data. If the experiment is repeated multiple times, the $q(\mu)$ would take on different values, spread according to a probability density function. Moreover, the $q(\mu)$ would take different values depending on the hypothesis H that is being tested.

The concept of p -value can be introduced through the probability density functions of $q(\mu)$,

$$p = \int_{q(\mu)^{\text{obs}}}^{\text{inf}} P(q(\mu)|H) dq(\mu). \quad (12-10)$$

A p -value is the probability, assuming H_0 (the null hypothesis) to be true, of getting a value of the test statistic as result of our test at least as extreme as the observed test statistic. In this context, seeing a large enough discrepancy in the data from the nulls prediction leads to a low p -value, the rejection of the null, and a discovery. With a model for the null and another for the expected data, the expected p -value can be quantified. An analysis with a lower p -value represents more convincing evidence to rule out the null, and therefore an analysis with a lower expected p -value can be considered more sensitive. The common practice in high energy physics, is to relate the p -value into a quantile of a unit Gaussian to express the significance level Z , defined as:

$$Z = \phi^{-1}(1 - p) \quad (12-11)$$

with the ϕ being the cumulative distribution function of a standard normal distribution. The area of the tail starting at an upward fluctuation of Z standard deviations from the mean of the Gaussian random variable, should be equal to the p -value, as visualized in Figure 12-1. A signal with a significance level of 3σ is referred to in the community as *evidence*, and corresponds to a p -value of $< 1.35 \times 10^{-3}$. A signal with a significance level of 5σ is referred to in the community as *observation*, and corresponds to a p -value of $< 2.9 \times 10^{-7}$. The above p -values will not be used in this thesis, as it has already been hinted that no evidence for new physics was found. Instead, the so-called one-sided modified frequentist confidence level CL_s [121] is used to set limits on the physics models. The probability density functions P of the $q(\mu)$ under the background only hypothesis and the background+signal hypothesis are used to define the CL_b and CL_{s+b} :

$$\text{CL}_b = P(q(\mu) \geq q(\mu)^{\text{obs.}} | H_0) \quad (12-12)$$

$$\text{CL}_{s+b} = P(q(\mu) \geq q(\mu)^{\text{obs.}} | H_1) \quad (12-13)$$

The practice in high energy physics at the LHC is to set upper limits on μ through the CL_s , defined as:

$$\text{CL}_s = \frac{\text{CL}_{s+b}}{\text{CL}_b}. \quad (12-14)$$

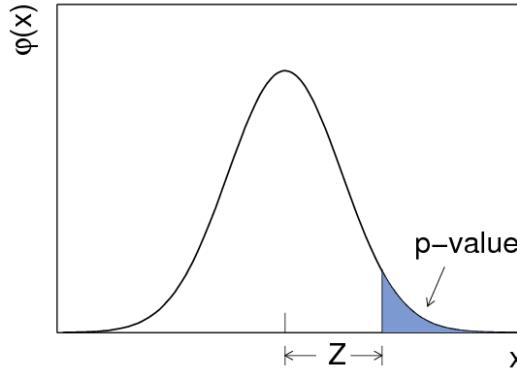


Figure 12-1. The standard normal distribution $\phi(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)$ showing the relation between the significance Z and the p -value [121].

The CL_{s+b} could in principle also be used to set limits on the signal strength, but it turns out that the CL_s is less affected by under-fluctuating backgrounds. The choice of CL_s thus gives more conservative limits. A statistical tool is used that calculates the probabilities of the maximized likelihood ratios using simulated pseudo-experiments where the pseudo-data is randomly drawn from Poisson distributions of the mean of the predictions and nuisances parameters according to their pdfs, and an asymptotic formulation [121] is used, that takes into account the systematic and statistical uncertainties. The common practice in SUSY is to calculate the results at 95% CL_s which is defined as the μ that produces $CL_s = 0.05$. This means that exclusion of a given model can be claimed if the observed CL_s is less than 0.05 for a signal strength $\mu > 0$. By deriving a number of events that correspond to the upper limit of the CL_s test statistic $q(\mu)$ that reach a value of 0.05, and this number of events corresponds the upper limit on the cross-section.

12.4 Visualization of the interpretation

In most searches presented in this thesis, the statistical analysis will be performed as a function of two SUSY particles, most commonly as a function of the SUSY particle initiating the decay chain and the LSP, see the free parameters column in Table 12-1. The visualization of the upper limit on the cross-section will be shown as a colored 2D map in the plane of the exclusion limits, with a color band that spans from blue to red, corresponding to lower to higher values of the upper limit on the cross-section. This leads to two possible scenarios can be interpreted in the following way. If the cross-section is lower than the theoretical cross-section in a particular mass scenario, one can declare that this given point with the acceptances and cross-section is excluded. If the cross-section is higher, however, then the signal could still be present without it being noticed in the results. There is a solid black line that illustrate the observed limit on the SUSY particle masses assuming theory cross sections. The dashed red lines are the equivalent expected limit. The mass points are fully connected if the observed or expected upper limit on the μ is found to be 1. This means that the mass scenarios under the lines are excluded. The effect of varying the full experimental uncertainty $\pm 1\sigma$ (or in some cases also $\pm 2\sigma$) is visualized with thin red dashed lines, around the solid red line. Similarly, the thin black solid lines around the thicker solid black line visualize the effect of a variation of $\pm 1\sigma$ of the theoretical uncertainty on SUSY production cross section.

In the search for higgsinos, the results will not be interpreted in terms of masses of two SUSY particles, instead, the mass of the Gravitino LSP is set to 1 GeV. Instead of having the upper limit on the cross section be displayed as a 2D color map as in the interpretation as a function of two masses, the upper limit is now visualized on the y -axis, with the x -axis representing the mass of the $\tilde{\chi}_1^0$. The observed limit is

again visualized as a solid dashed lines, but in this 1D interpretation the dashed line represent the expected limit. The variation of the experimental uncertainties by $\pm 1\sigma$ and $\pm 2\sigma$ is visualized as a green and yellow band. The theoretical cross section is visualized as a red line, with the thickness of the line representing the theoretical uncertainty. A model in this 1D interpretation can be considered excluded below the mass where the theoretical and observed limits intersect.

13 — Systematic uncertainties

When estimating the number of background and signal events that enter the signal regions, they are associated systematic effects arising from various different sources. These sources can be either experimental or theoretical. Experimental uncertainties arise from detector effects, faulty reconstruction algorithms, and uncertainties on observables that are used to categorize the events. Another experimental uncertainty is that related to the luminosity measurement, which affects the number of predicted signal and background events. Theoretical uncertainties account for limited knowledge of the underlying theoretical models that calculate the cross section, proton PDFs and QCD scales, as described in Chapter 2.

In the previous chapter, the procedure of how the results can be interpreted in terms of signal models was presented. This procedure relies on an appropriate treatment of these systematic uncertainties. This chapter contains a brief introduction on how the systematic uncertainties enter the statistical analysis used to interpret the results. The sources of experimental and theoretical uncertainties evaluated in the signal simulation, are described. The chapter concludes with a summary of the systematic uncertainties assigned to the various background prediction techniques.

13.1 Nuisance parameters

The effect of systematic uncertainties are incorporated into the likelihood in Equation 12–5 by means of nuisance parameters, represented by θ_k with k running over all the sources of uncertainties [120]. Nominal values of nuisance parameters are random variables distributed according to some probability distribution function, PDF, denoted $\rho(\theta_k)$. Almost all uncertainties treated in this thesis are uncertainties that affect the normalization of a process. These uncertainties are modeled using nuisance parameters with log-normal probability densities, defined as

$$\rho(\theta) = \frac{1}{\sqrt{2\pi} \ln \kappa} \frac{1}{\theta} \exp \left(-\frac{(\ln(\theta/\tilde{\theta}))^2}{2(\ln \kappa)^2} \right). \quad (13-1)$$

The κ represents the spread of the distribution, and $\tilde{\theta}$ is the best estimate of the nuisance θ . The log-normal PDF is preferably used in cases with large uncertainties, as the distribution has a longer tail in comparison with a Gaussian PDF. This avoids the problem of negative parameter values obtained from a Gaussian probability density, as $\rho(\theta) \rightarrow 0$ for $\theta \rightarrow 0$. The uncertainties that are not modeled using nuisance parameters with log-normal probability densities are instead modeled using nuisance parameters with gamma distribution. The gamma distribution is useful for describing variables bounded at one side, for example $0 < X < \infty$ and is used for the systematic uncertainties associated to the data-driven background prediction techniques.

In the following sections the sources of the experimental and theoretical uncertainties are explained. The actual magnitude of the uncertainties will be described separately for each search, in Chapters 14, 15 and 16.

13.2 Experimental uncertainties

The sources of experimental systematic uncertainties dominant in the analyses this thesis is based on are listed below. They are all modeled using nuisance parameters with log-normal probability densities, and are assumed to be fully correlated. Due to the different kinematic features present in the slepton search, additional systematic uncertainties are introduced. The systematic uncertainties specific to the slepton search are indicated in the description.

Integrated luminosity

The uncertainty associated to the measurement of the integrated luminosity delivered by the LHC to the CMS experiment in 2016 amounts to 2.5% [122].

Pileup modeling

The number of pp collisions per bunch crossings vary strongly with time. Therefore, the distribution in simulation is difficult to model properly as only one distribution of pileup interactions is simulated. This is done by drawing a number of so-called minimum-bias events from a simulated sample according to the input distribution, which clearly does not reflect the time varying pileup in data. To this end, the number of true pileup interactions is reweighted to match the number of interactions in data by dividing the two quantities. The systematic uncertainty associated to this modeling is calculated with two different procedures for the first published paper [59] and the second published paper [60]. The appropriate procedure to estimate the effect of the pileup modeling in simulation is to vary the total inelastic cross section, 68.6 mb, up and down by 5% and evaluate the effect the variation has on the normalization [123]. However, at the time of publication of the first paper, the search for strong and electroweak SUSY, the pileup distribution in simulation was not properly modeled. For this reason, the normal procedure to estimate the impact of pileup on the analysis

could not be performed, and an alternative approach was chosen. Instead, the signal acceptance was evaluated in high- and low-pileup regimes separately and the difference was treated as the magnitude of the effect.

Lepton uncertainties

The electron and muon selection efficiency amounts to 2% per lepton, after varying the p_T and η dependent scale factors. The total uncertainty of 5% is taking into account both leptons. An additional uncertainty arise from the modeling of the lepton efficiency in the fast simulation used for signal. The uncertainty related to the modeling of the dileptonic triggers is measured according to the procedure introduced in Section 11.2.

Jet energy uncertainties

The uncertainty in the jet energy scale (JES) is assessed by shifting the jet energy correction factors for each jet by one standard deviation up and down and recalculating the kinematic quantities such as the p_T^{miss} , M_{T2} . The uncertainty due to jet energy resolution (JER), assessed by comparing the nominal smearing of the jet's p_T to a smearing performed with a Gaussian of the nominal width plus the uncertainty on the width, is found to be negligible.

b-tag modeling

The uncertainties in the *b* tagging efficiency and mistag probability are measured on independent control samples by varying the *b*-tagging scale factors up and down according to their uncertainties [106].

p_T^{miss} uncertainties

The unclustered energy, the electron and muon energy scales are varied up and down by one standard deviation and propagated to the p_T^{miss} . The maximal difference between the nominal p_T^{miss} and the up or down variations are taken as the magnitude of the uncertainty. To account for possible mismodeling of the p_T^{miss} resolution in fast simulation, the signal yields using the nominal reconstructed p_T^{miss} is compared to the generator level p_T^{miss} . The average of these p_T^{miss} quantities is taken as the central value, and half of the difference between the values as the uncertainty.

MC statistics

The number of generated events depends on the model and the specific point in the 2D SUSY mass-scan and varies between 10 000 and 50 000. Hundreds of SUSY signal mass scenarios are generated and for each scenario a decent sample size is required. But as an effect of limited available computation power and storage, the limited sample sizes can result in a large systematic uncertainty related to the statistical power of the simulated samples in some parts of the phase space.

13.3 Theoretical uncertainties

The theoretical systematic uncertainties are listed below. They are all modeled using nuisance parameters with log-normal probability densities, and are assumed to be fully correlated.

Factorization and renormalization scales

The factorization and renormalization scale uncertainties, also known as QCD scale uncertainties, fall under the category of theory uncertainties, that are not accounting for experimental effects. Instead, the QCD scale uncertainties estimate the effects due to missing corrections from higher order perturbative QCD calculations. The factorization scale, μ_F and renormalization scale, μ_R , are halved or doubled and the changes are propagated to the p_T^{miss} . A special treatment of the evaluation of the QCD scale uncertainties is performed for the slepton search. The motivation behind the special treatment is the jet veto that defines the slepton signal region. The so-called Stewart-Tackmann prescription on the treatment of theory uncertainties in a jet binned analysis is followed [124]. Following the Stewart-Tackmann prescription, that suggests that scale variations in exclusive fixed-order predictions can underestimate the perturbative uncertainty for the cross sections. The underestimation can be due to the cancellations between the perturbative corrections, leading to large K factors and those that induce logarithmic sensitivity to the jet-bin boundary. To account for this underestimation and to properly assess the uncertainty due to varying the renormalization and factorization scales, it is suggested to first evaluate the cross section in an inclusive N -jet selection and then use this to compute the uncertainty in the exclusive N -jet selection from the difference:

$$\sigma_{N\text{jets}} = \sigma_{\geq N\text{jets}} - \sigma_{\geq (N+1)\text{jets}} \quad (13-2)$$

In the slepton analysis, this N is of course 0, and the uncertainty on the cross section in the 0 jet selection is obtained after propagating the errors from the $\sigma_{\geq N}$ and $\sigma_{\geq N+1}$ in quadrature:

$$\delta_{0\text{jet}}^2 = \delta_{\geq 0\text{jet}}^2 + \delta_{\geq 1\text{jet}}^2 \quad (13-3)$$

Parton distribution function (PDF)

The second type of theory uncertainties are those associated to the choice of parton distribution function. The procedure is to reweight the simulation samples with 100 PDF replicas [125]. This uncertainty is evaluated only in the slepton search. The reason to not perform it for the strong and electroweak search is that in these searches an additional uncertainty due to modeling of initial state radiation (ISR) is included. This is already incorporated in the empirical uncertainty in the modeling of ISR as described in Chapter 2, and we therefore do not apply a dedicated uncertainty in signal acceptance from PDF variations. The Madgraph signal simulation is reweighted according to an initial state radiation reweighting procedures, which differ for strongly and electroweakly produced signals. For the strongly produced signal model, the reweighting is based on the number of ISR jets, and the uncertainty is taken as half of the magnitude of the weights. For electroweakly produced signals, the reweighting is based on the Z p_T agreement with simulation, as measured in dilepton events, and the full magnitude of the weights is taken as the uncertainty.

13.4 Uncertainties on background predictions

The sources of systematic uncertainties listed so far are evaluated on the signal simulation. The SM background estimation techniques also have systematic uncertainties assigned to them. The sources and the magnitude of the uncertainties have been partly been mentioned in Chapter 11, but will be summarized here for clarity.

Flavor Symmetric backgrounds

The uncertainties associated to the FS background estimation techniques stem from several sources. The uncertainty assigned to the direct measurement of the $R_{\text{SF}/\text{OF}}$ is 4%. The factorization method uses both the $r_{\mu/e}$ and R_T as input. The $r_{\mu/e}$ is assigned a 10% uncertainty after simulation studies. Similarly, the trigger efficiencies for the dielectron, dimuon and electron muon triggers are chosen to be 3%, and when propagated to the R_T results in 4%. As the method is a data-driven technique, the largest limitation is the number of data events in the OF control region used to predict the number of SF events in the signal region. This background is modeled using the gamma distribution, as this one is useful for describing variables bounded at one side. An additional 30% uncertainty is assigned to the FS prediction method in the on-Z searches, that will be described in more detail in Chapter 14.

Drell-Yan background

For the other data-driven background prediction method, again, the statistics of the data sample is the limiting factor. The p_T^{miss} template method predicts the p_T^{miss} in Z+jets events by the p_T^{miss} in γ +jets events. The systematic uncertainty assigned to the method is determined by a closure test in simulation, using γ +jets to predict the yield of Z+jets in each analysis bin. An uncertainty is assigned from the results of the closure test as the larger of the difference between the γ +jets prediction and the Z+jets yield for each p_T^{miss} region or the simulation statistical uncertainty. The final values range from 10 – 80% depending on the region, with large values mainly driven by low statistical power in the control regions.

The search for direct slepton production is designed to keep the DY contribution to a minimum. Therefore the use of a data-driven method to predict this almost negligible background is redundant. Instead, the DY is predicted using simulation and a conservative 50% uncertainty is assigned to the method.

Z+ ν background

As will be thoroughly described in Chapters 14, 15 and 16, the treatment of the Z+ ν and slightly different in the three types of searches. In the searches for electroweak and strong superpartner production, the transfer factors derived in the control regions to correct the simulation have large statistical uncertainties due to the sparsely populated control regions. Conservative systematic uncertainties are assigned to the transfer factors, of 30% for the WZ and t \bar{t} Z backgrounds, and 50% for the ZZ background.

For the WZ and ZZ backgrounds in the search for direct slepton production, on the other hand, have a different treatment of the uncertainties. As the control regions where the transfer factors have a higher statistical power, assigning a systematic uncertainty of the same size as the statistical uncertainty will not be sufficient. Instead, an uncertainty on the transfer factor is taken as the statistical uncertainty from the control sample. In addition to this uncertainty, experimental and theoretical uncertainties as listed previously in this chapter are taken into account. A full description of the systematic uncertainties taken into account for these backgrounds are given in Section 16.3 of Chapter 16.

Rare backgrounds

Any additional SM process that is not taken into account by the previously described methods are estimated from simulation. These backgrounds have a conservative 50% uncertainty assigned to them.

Part IV

ANALYSIS RESULTS

14 — Search for colored superpartner production

Equipped with the theoretical foundation, the experimental apparatus and the physics objects, the actual searches for SUSY can be performed. The searches are divided in to three categories; the search for colored superpartners, the search for electroweak superpartners, and, finally, the search for direct slepton production. The first two searches are published in [59] and the third search is published in [60].

During Run 1, the CMS collaboration reported an excess in the slepton edge model that targets sbottom production [126]. The local significance of the excess was 2.6 standard deviations. In addition, the ATLAS collaboration reported an excess in the GMSB gluino production model [127] with the Run 1 dataset. The local significance of this excess was 3.0 standard deviations. Effort has been put into confirming these excesses using the Run 2 dataset.

The first chapter summarizes the for the search for colored superpartners in final states with two same flavor opposite sign leptons. The specific search regions designed to target \tilde{b} and \tilde{g} production modes are presented, together with the SM background estimation techniques deployed. The results and the statistical interpretation of the results conclude the chapter. All figures without citation are taken from the CMS Analysis Note AN-16-482.

14.1 Analysis strategy

Pairs of electrons or pairs of muons are selected using the lepton selection criteria outlined in Section 10.3. The offline p_T threshold for jets identified according to the requirements listed in Section 10.4 is set to 35 GeV, and is kept at 25 GeV for b-tagged jets. A baseline selection of more than two jets and p_T^{miss} greater than 100 GeV is imposed. For all signal regions introduced below, a requirement is imposed that the two jets with the highest p_T have a separation in ϕ from the p_T^{miss} of at least 0.4, in order to reduce the Drell–Yan contribution. The search for colored superpartners are split in to the On-Z signal region, targeting gluino production, and the Edge signal region, targeting sbottom production.

On-Z signal region

The defining feature of the search for gluinos in the context of two opposite sign, same flavor leptons is the production of an on-shell Z boson. For this reason, SR events are selected to be compatible with a Z boson, i.e. having $m_{\ell\ell} \in [86, 96]$ GeV. Next, selections are made on the level of hadronic activity, by defining three SRs, “SRA” requiring 2-3 jets, “SRB” requiring 4-5 jets and “SRC” requiring more than 6 jets. Further, these three SRs are split into two categories by requiring either one or more b-tagged jets, or veto b-tagged jets. The two categories have requirements on the M_{T2} of ≥ 80 GeV for the b-veto region and ≥ 100 GeV for the b-tagged region to suppress contribution from $t\bar{t}$. For the lower jet multiplicity SRs (SRA and SRB) requirements on the H_T are imposed, ≥ 500 GeV for b-veto region and ≥ 200 GeV for b-tagged region. Finally, the 6 SRs are binned in p_T^{miss} . There are fewer p_T^{miss} bins in the highest jet multiplicity bins as they are already low in statistics. The final SRs targeting the GMSB gluino induced process are summarized in Table 14-1.

Table 14-1. Summary of the colored SUSY On-Z SR.

Region	Colored SUSY On-Z SR, $86 \geq m_{\ell\ell} \geq 96$				
	N_{jets}	$N_{\text{b-jets}}$	H_T [GeV]	M_{T2} [GeV]	p_T^{miss} [GeV]
SRA (b-veto)	2-3	= 0	≥ 500	≥ 80	100-150, 150-250, ≥ 250
SRB (b-veto)	4-5	= 0	≥ 500	≥ 80	100-150, 150-250, ≥ 250
SRC (b-veto)	≥ 6	= 0	—	≥ 80	100-150, ≥ 150
SRA (b-tag)	2-3	≥ 1	≥ 200	≥ 100	100-150, 150-250, ≥ 250
SRB (b-tag)	4-5	≥ 1	≥ 200	≥ 100	100-150, 150-250, ≥ 250
SRC (b-tag)	≥ 6	≥ 1	—	≥ 100	100-150, ≥ 150

Edge signal region

The production of sbottoms with an intermediate slepton decay giving an edge shape in the invariant mass is called the “edge” search. This search has an explicit Z boson veto, making the dominant SM background that of $t\bar{t}$. In order to deal with this background, first a requirement on the M_{T2} of ≥ 80 GeV is applied. The $t\bar{t}$ likelihood introduced in Section 10.7 is then used to further categorize how $t\bar{t}$ -like an event is, and divide the events in $t\bar{t}$ -like and non- $t\bar{t}$ -like. The edge SR is defined with a $t\bar{t}$ -like category corresponding to a discriminator value of less than 21, and non- $t\bar{t}$ -like with values larger than 21.

As discussed in Section 4.3, any potential excess would show up as an edge shape in the invariant mass spectrum of the two leptons. For this reason, the signal region is binned in the $m_{\ell\ell}$ to be able reflect the position of the edge. The edge SR definitions is summarized in Table 14-2.

Table 14-2. Summary of the Edge SR.

Region	Edge SR				
	N_{jets}	p_T^{miss} [GeV]	M_{T2} [GeV]	NLL	$m_{\ell\ell}$ [GeV] (excluding 86-96)
t <bar>t}-like</bar>	≥ 2	≥ 150	≥ 80	≤ 21	[20, 60, 86, 96, 150, 200, 300, 400+]
non-t <bar>t>-like</bar>	≥ 2	≥ 150	≥ 80	> 21	[20, 60, 86, 96, 150, 200, 300, 400+]
edge fit	≥ 2	≥ 150	≥ 80	-	> 20

In addition to the counting experiment, a fit is performed in this baseline region to search for the position of the edge in the $m_{\ell\ell}$ spectrum. The procedure to extract the edge position is briefly introduced in the following. Although it is not the main focus of the thesis, nor has any original contribution from the author of this thesis, it is included for completeness and historical reasons, as it provides a means by which the Run 1 excess can be confirmed or rejected.

The likelihood has three components, a component for the FS background, a component for backgrounds containing a Z boson (DY and Z+ ν backgrounds), and finally a component for the signal shape. The FS component is modeled with a Crystal-Ball function, the Z background component is modeled by a Breit-Wigner function and the signal component is modeled by a triangular shape convolved with a Gaussian distribution. A first fit is carried out separately for electrons and muons in the DY enriched control region to extract the shape of the backgrounds with a Z boson. A second fit is then carried out in the SR using an extended unbinned maximum likelihood fit, which is performed simultaneously for the dilepton $m_{\ell\ell}$ distribution in the dielectron, dimuon, and electron-muon flavor final state. More information on the fit functions and the fitting procedure is given in Appendix C and in [59].

14.2 Background estimation

Opposite sign same flavor lepton pairs can arise from many sources, as explained in Chapter 11. Dominant backgrounds in the On-Z search for gluinos is contributions from Drell-Yan, whereas flavor symmetric processes such as dileptonic tt} is dominant in the Edge search for sbottoms. The specific treatment of the various background prediction techniques introduced in Chapter 11 is summarized in the following sections.

Flavor symmetric background

FS backgrounds enter both in the On-Z search for gluinos and the Edge search for sbottoms. The treatment of the FS background in the two cases is summarized in the following.

Edge search

The FS background is dominant in the Edge search. The resulting background estimates for FS backgrounds in the signal region of the Edge search are summarized in Table 14-3. It can be observed that the event-by-event reweighting in the factorization method yields smaller $R_{\text{SF/OF}}$ values for higher mass bins (especially in the tt} like selection). This makes sense since $m_{\ell\ell}$ is correlated to the p_T of the leptons and higher masses usually correspond to higher lepton p_T and thus smaller reweighting factors. Overall, the $R_{\text{SF/OF}}$ factors from the factorization method are a bit smaller than the factor from the control region method but still agree within their uncertainties. While FS background prediction method described above is clearly simple and clean, its main drawback is the limited statistical power in the scarcely populated kinematic regions.

Table 14-3. Resulting estimates for flavor-symmetric backgrounds in the Edge SR. Given is the observed event yield in OF events (N_{OF}), the estimate in the SF channel using the event-by-event reweighting of the factorization method ($N_{SF}^{factorization}$), $R_{SF/OF}$ for the factorization method ($R_{SF/OF}^{factorization}$), $R_{SF/OF}$ obtained from the direct measurement ($R_{SF/OF}^{direct}$), $R_{SF/OF}$ when combining this results from direct measurement and factorization methods ($R_{SF/OF}^{combined}$), and the combined final prediction (N_{SF}^{final}). Statistical and systematic uncertainties are given separately.

$m_{\ell\ell}$ [GeV]	N_{OF}	$N_{SF}^{factorization}$	$R_{SF/OF}^{factorization}$	$R_{SF/OF}^{direct}$	$R_{SF/OF}^{combined}$	N_{SF}^{final}
ttbar like						
20-60	264	$289.1^{+18.3}_{-17.3} \pm 14.2$	1.10 ± 0.05	1.11 ± 0.01	1.10 ± 0.04	$290.9^{+18.5}_{-17.4} \pm 9.3$
60-86	164	$179.1^{+14.5}_{-13.4} \pm 8.7$	1.09 ± 0.05	1.11 ± 0.01	1.10 ± 0.03	$180.5^{+14.7}_{-13.6} \pm 5.7$
96-150	160	$173.5^{+14.3}_{-13.2} \pm 8.2$	1.08 ± 0.05	1.11 ± 0.01	1.10 ± 0.03	$175.5^{+14.4}_{-13.3} \pm 5.5$
150-200	67	$72.4^{+10.0}_{-8.8} \pm 3.3$	1.08 ± 0.05	1.11 ± 0.01	1.09 ± 0.03	$73.3^{+10.1}_{-8.9} \pm 2.3$
200-300	43	$46.1^{+8.2}_{-7.0} \pm 2.1$	1.07 ± 0.05	1.11 ± 0.01	1.09 ± 0.03	$46.9^{+8.3}_{-7.1} \pm 1.4$
300-400	17	$18.2^{+5.6}_{-4.4} \pm 0.8$	1.07 ± 0.05	1.11 ± 0.01	1.09 ± 0.03	$18.5^{+5.7}_{-4.4} \pm 0.6$
>400	4	$4.3^{+3.4}_{-2.0} \pm 0.2$	1.07 ± 0.05	1.11 ± 0.01	1.09 ± 0.03	$4.3^{+3.4}_{-2.1} \pm 0.1$
non ttbar like						
20-60	3	$3.2^{+3.1}_{-1.8} \pm 0.1$	1.07 ± 0.05	1.11 ± 0.01	1.09 ± 0.03	$3.3^{+3.2}_{-1.8} \pm 0.1$
60-86	3	$3.2^{+3.1}_{-1.7} \pm 0.1$	1.07 ± 0.05	1.11 ± 0.01	1.09 ± 0.03	$3.3^{+3.2}_{-1.8} \pm 0.1$
96-150	6	$6.5^{+3.9}_{-2.6} \pm 0.3$	1.08 ± 0.05	1.11 ± 0.01	1.09 ± 0.03	$6.6^{+3.9}_{-2.6} \pm 0.2$
150-200	5	$5.4^{+3.6}_{-2.3} \pm 0.2$	1.08 ± 0.05	1.11 ± 0.01	1.09 ± 0.03	$5.5^{+3.7}_{-2.4} \pm 0.2$
200-300	3	$3.2^{+3.1}_{-1.7} \pm 0.1$	1.07 ± 0.05	1.11 ± 0.01	1.09 ± 0.03	$3.3^{+3.2}_{-1.8} \pm 0.1$
300-400	3	$3.2^{+3.1}_{-1.7} \pm 0.1$	1.07 ± 0.05	1.11 ± 0.01	1.09 ± 0.03	$3.3^{+3.2}_{-1.8} \pm 0.1$
>400	1	$1.1^{+2.4}_{-0.9} \pm 0.0$	1.06 ± 0.05	1.11 ± 0.01	1.09 ± 0.03	$1.1^{+2.5}_{-0.9} \pm 0.0$

On-Z search

The search for colored superpartners in the On-Z search regions suppress the $t\bar{t}$ process by selecting events with a dilepton pair compatible with a Z boson. Further, very low yields of the FS backgrounds is obtained with a cut on M_{T2} . Of course, reducing the SM backgrounds to a minimal is desirable, but one still wants to make sure that the backgrounds are properly predicted and do not suffer from too large statistical errors. In order to cope with the reduced OF statistics in the SR of the On-Z search, a slightly adapted method of predicting the FS background is proposed. The idea is to exploit that the On-Z searches only have signal regions where the leptons are compatible with a Z boson, and the rest of the invariant mass spectrum is free to use to measure the OF events in the signal region. This lever arm to extend the OF control region can be used together with $f_{m\ell\ell}$, which is the ratio of OF events in the On-Z region over the number of events in the extended $m_{\ell\ell}$ region,

$$f_{m\ell\ell} = \frac{N_{OF}(86 < m_{\ell\ell} < 96)}{N_{OF}(m_{\ell\ell} > 20)} \quad (14-1)$$

This consideration allows to extend the $m_{\ell\ell}$ (from $86 < m_{\ell\ell} < 96$ GeV to $m_{\ell\ell} > 20$ GeV) window from which the estimation is taken by a large fraction. By implementing this approach in extending the OF control region, the simplified formula from before then becomes

$$N_{SF} = N_{OF}^{ext.m_{\ell\ell}} \cdot R_{SF/OF} \cdot f_{m\ell\ell}, \quad (14-2)$$

While $R_{SF/OF}$ can be assumed to be the same number as for the strong search, the factor $f_{m\ell\ell}$ has to be measured. The $f_{m\ell\ell}$ is measured in all On-Z signal regions and the results are shown in Figure 14-1. A central value of $f_{m\ell\ell} = 0.065 \pm 0.02$ is chosen as it is compatible over all signal regions. The magnitude of

the uncertainty is chosen to cover the differences in central values observed in simulation. The statistical uncertainties on the data points are larger but there is agreement within the assigned systematic uncertainty.

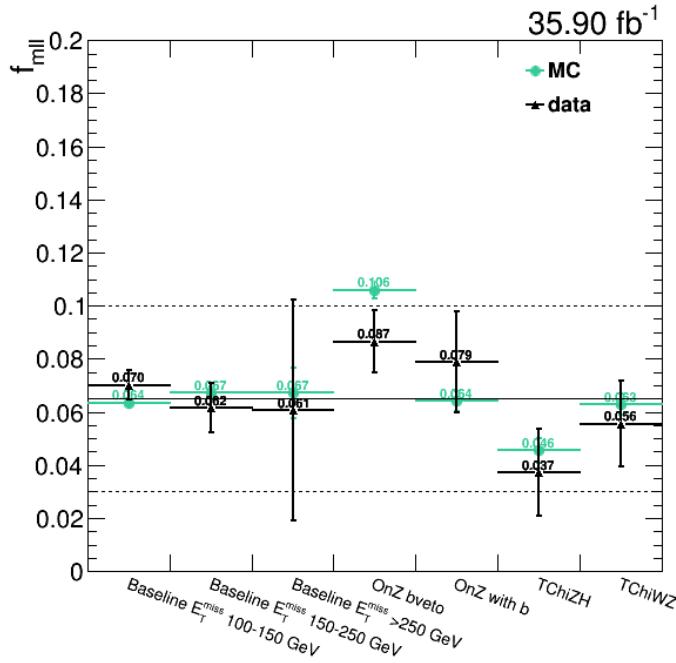


Figure 14-1. The f_{mll} evaluated in each On-Z signal region.

Drell-Yan

The Drell-Yan background is estimated through the p_T^{miss} templates method, and have slightly different treatment depending on the search.

Edge search

In the edge search, the signal regions are split in bins of $m_{\ell\ell}$ to increase the sensitivity to a large range of sbottom masses. Drell-Yan is a dominant process in the searches where no Z boson veto is imposed, but it can also contribute in the $m_{\ell\ell}$ bins around the Z boson mass. For this reason, the Drell-Yan contribution at the Z boson mass is estimated from a single γ sample, and the contribution outside of the Z peak is estimated through a translation factor. The translation factor is called $R_{\text{out/in}}$ and is the ratio of SF lepton events in a $m_{\ell\ell}$ signal region bin (20-60, 60-86, 96-150, 150-200, 200-300, 300-400 and > 400 GeV) over the SF lepton events with a $m_{\ell\ell}$ in 86-96 GeV. This factor is measured in a Drell-Yan enriched control region constructed with at least two jets and $p_T^{\text{miss}} < 50$ GeV. A requirement of $M_{T2} > 80$ GeV is imposed to keep maintain the kinematic features of the signal regions. In order to further purify the Drell-Yan contribution in this region, the number OF lepton events are removed from the number of SF lepton events, thereby reducing the contribution from $t\bar{t}$. The measured values of $R_{\text{out/in}}$ in the edge signal regions is shown in Table 16-5.

Table 14-4. Measured values for $R_{\text{out/in}}$ for data and MC in the different signal regions of the edge search.

$m_{\ell\ell}$ [GeV]	Data			MC	
	$N_{\text{in}}: 4295 \pm 65$			$N_{\text{in}}: 3954 \pm 62$	
	N_{out}	$R_{\text{out/in}}$	N_{out}	$R_{\text{out/in}}$	
20-60	229 ± 15	0.053 ± 0.027	138 ± 12	0.035 ± 0.018	
60-86	551 ± 23	0.128 ± 0.064	453 ± 21	0.115 ± 0.058	
96-150	671 ± 26	0.156 ± 0.078	644 ± 25	0.163 ± 0.082	
150-200	74 ± 9	0.017 ± 0.017	55 ± 8	0.014 ± 0.014	
200-300	52 ± 9	0.012 ± 0.012	51 ± 8	0.013 ± 0.013	
300-400	22 ± 5	0.005 ± 0.005	19 ± 5	0.005 ± 0.005	
>400	23 ± 5	0.006 ± 0.006	27 ± 5	0.007 ± 0.007	

On-Z search

The prediction from the p_T^{miss} template method is summarized in Table 14-5. The predicted number of events in the On-Z signal regions are given with a decomposition of the magnitude of the systematic uncertainties from the four sources considered.

Table 14-5. Summary of template predictions with systematic uncertainties added in quadrature in the strong On-Z signal regions together with the individual systematic uncertainties from each source.

SR	p_T^{miss} [GeV]	Prediction	Closure	Normalization	Statistical	EWK sub.
SRA (b-veto)	50-100	208.5 ± 16.1	0.0	15.3	5.0	0.0
	100-150	13.6 ± 3.1	2.7	1.0	1.1	0.3
	150-250	2.5 ± 0.9	0.6	0.2	0.4	0.4
	250+	3.3 ± 2.4	0.9	0.2	2.2	0.4
SRA (b-tag)	50-100	92.2 ± 10.4	0.0	10.0	2.8	0.0
	100-150	8.2 ± 2.1	1.6	0.9	0.9	0.3
	150-250	1.2 ± 0.5	0.3	0.1	0.2	0.4
	250+	0.5 ± 0.3	0.1	0.1	0.2	0.2
SRB (b-veto)	50-100	130.1 ± 12.8	0.0	12.1	4.1	0.0
	100-150	12.8 ± 2.4	1.5	1.2	1.3	0.2
	150-250	0.9 ± 0.4	0.1	0.1	0.2	0.3
	250+	0.4 ± 0.2	0.1	0.1	0.1	0.2
SRB (b-tag)	50-100	37.9 ± 6.7	0.0	6.5	1.9	0.0
	100-150	7.7 ± 3.1	0.9	1.3	2.7	0.2
	150-250	4.0 ± 3.3	0.7	0.7	3.2	0.3
	250+	0.1 ± 0.1	0.1	0.1	0.1	0.2
SRC (b-veto)	50-100	23.8 ± 5.5	0.0	5.2	1.9	0.0
	100-150	1.2 ± 0.4	0.2	0.3	0.3	0.1
	150+	0.1 ± 0.1	0.1	0.1	0.1	0.1
SRC (b-tag)	50-100	9.9 ± 3.7	0.0	3.5	1.4	0.0
	100-150	0.1 ± 0.5	0.1	0.1	0.1	0.5
	150+	0.0 ± 0.3	0.0	0.0	0.0	0.3

WZ → 3lν background

The WZ → 3lν background is large in the On-Z search for gluinos, when one of the prompt leptons is lost due to acceptance, as described in Section 11.4. The contribution is estimated from simulation, with a transfer factor derived in a WZ control region. The WZ control region designed for the On-Z searches is defined as

- three tight ID leptons of any flavor or charge with $p_T \geq 25$ GeV for the leading and $p_T \geq 20$ GeV for the subsequent leptons
- $p_T^{\text{miss}} < 60$ GeV
- veto b-tagged jets of $p_T > 25$ GeV
- more than two jets of $p_T > 35$ GeV
- $|\Delta\phi(\text{jet}_{1,2}, p_T^{\text{miss}})| > 0.4$

The result is presented in Table 14-6, where the WZ signal simulation is compared to the simulated backgrounds and the observed yields that enter this control region. A transfer factor of 1.06 is derived as the data with simulated backgrounds subtracted over the signal process simulation. As the statistical uncertainty is dominant in this control region, due to the low number of data yields, a flat systematic uncertainty of 30% is chosen. This is chosen instead of evaluating the effect of varying jet energy scale and resolution, and varying the pdf and scale, as these uncertainties are subdominant to the statistical uncertainty. The transfer factor is presented in Table 16-3, using the MC samples summarized in Appendix A. As can be seen in

Table 14-6. Transfer factor derived in the three lepton control region for the On-Z search.

On-Z 3 lepton CR	
signal MC	116.0 ± 3.23
bkg. MC	50.6 ± 2.9
data	164
data-bkg.	123.4 ± 13.1
(data-bkg.)/sig.	1.06 ± 0.12

Figure 14-2, the p_T^{miss} tails are well modeled in this control region.

ZZ → 2l2ν background

The ZZ → 2l2ν process, introduced in Section 11.4 is a sub-dominant background in the On-Z search, as this signal region requires more than two jets. This sub-dominant contribution is estimated from simulation, and a transfer factor is derived in a four lepton control region that targets the ZZ → 4l process. The four lepton control region is defined as

- four tight ID leptons of any flavor or charge with $p_T \geq 25$ GeV for the leading and $p_T \geq 20$ GeV for the subsequent leptons
- the leptons that form the best Z candidate is required to have $m_{\ell\ell}$ within 86 to 96 GeV.
- the leptons that form the other Z candidate is required to have $m_{\ell\ell}$ greater than 20 GeV.
- $|\Delta\phi(\text{jet}_{1,2}, p_T^{\text{miss}})| > 0.4$.

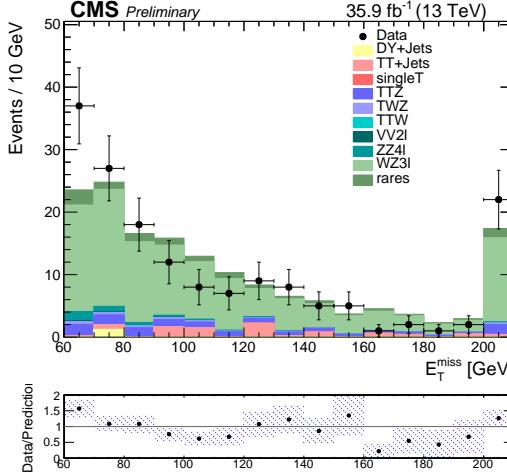


Figure 14-2. The p_T^{miss} (left) and the M_{T2} in the three lepton control region in data and MC.

At the time of publication, not all the ZZ processes were generated¹ making the derivation of the translation factor particularly important as some processes are missing. The effect of the missing samples in simulation is reflected in Figure 14-3, where the observed data obviously contain all physical processes resulting in four leptons. The transfer factor that is used to correct for the missing simulation is presented in Table 14-7. As

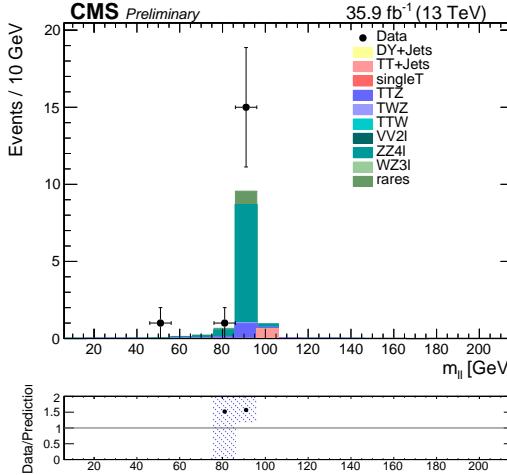


Figure 14-3. The invariant mass of the two lepton pair forming the best Z candidate in the four lepton control region in data and simulation.

the four lepton sample in this control region is very small, the large statistical uncertainty is taken as the total uncertainty on the method.

¹ Only the `/ZZTo4L_13TeV_powheg_pythia8` was available (as opposed to in the slepton search where all production modes of the ZZ processes listed in Appendix A was available).

Table 14-7. Transfer factor derived in the four lepton control region for On-Z searches.

On-Z 4 lepton CR	
signal MC	7.7 ± 0.2
bkg. MC	1.9 ± 0.2
data	15
data-bkg.	13.1 ± 3.9
(data-bkg.)/sig.	1.71 ± 0.5

t̄Z background

The t̄Z process contains two leptons from the Z boson decay, and more leptons and neutrinos depending on the decays of the top and anti-top quarks. If some of these leptons are lost due to p_T or η acceptance, the process can enter the On-Z signal regions. If the top and anti-top quarks decay hadronically, there is no p_T^{miss} due to neutrinos in the decays, and thus would the process not be present in the signal regions. The t̄Z control region is designed by inverting the third lepton veto and using the following requirements

- three tight ID leptons of any flavor or charge with $p_T \geq 25$ GeV for the leading and $p_T \geq 20$ GeV for the subsequent leptons
- the leptons that form the best Z candidate is required to have $m_{\ell\ell}$ within 86 to 96 GeV.
- exactly two b-tagged jets of $p_T > 25$ GeV.
- $p_T^{\text{miss}} > 30$ GeV
- $|\Delta\phi(\text{jet}_{1,2}, p_T^{\text{miss}})| > 0.4$.

The invariant mass constructed with the leptons most compatible with the Z boson mass is shown in Figure 14-4. There is a slight disagreement between the simulation and the data, and to account for this

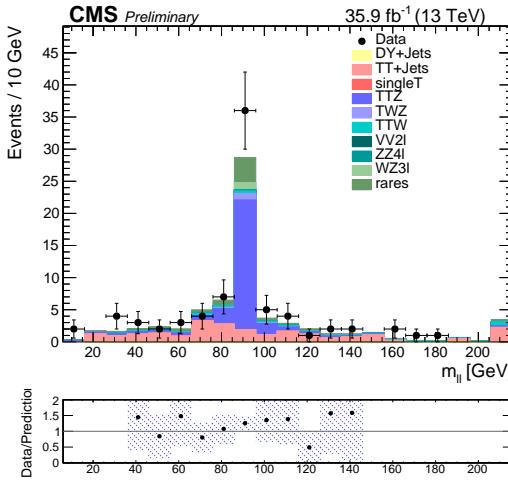


Figure 14-4. The invariant mass of the two lepton pair forming the best Z candidate in the t̄Z control region in data and simulation.

a transfer factor is derived that is summarized in Table 14-8. The final translation factor is 1.4 with a conservative systematic uncertainty that is chosen from the statistical uncertainty on the data sample.

Table 14-8. Transfer factor derived in the $t\bar{t}Z$ control region

$t\bar{t}Z$ ICR	
signal MC	20.2 ± 0.4
bkg. MC	8.5 ± 1.1
data	36
data-bkg.	27.5 ± 6.1
(data-bkg.)/sig.	1.4 ± 0.3

14.3 Systematic uncertainties

The sources of the systematic uncertainties for the signal simulation used in the search for colored superpartners is summarized in Table 14-9. The uncertainties affect the overall normalization of the process, and the nuisance parameters obey a log-normal distribution. The statistical errors on the predicted number of signal events is uncorrelated across the bins, while all other uncertainties are considered correlated across the search regions.

Table 14-9. Systematic uncertainties taken into account for the signal yields and their typical values.

Source of uncertainty	Uncertainty (%)
Integrated luminosity	2.5
Lepton reconstruction and isolation	5
Fast simulation lepton efficiency	4
<i>b</i> -tag modeling	0–5
Trigger modeling	3
Jet energy scale	0–5
ISR modeling	0–2.5
Pileup	1–2
Fast simulation p_T^{miss} modeling	0–4
Renormalization//factorization scales	1–3
MC statistical uncertainty	1–15

14.4 Results

Once all the background prediction techniques are in place, it is time to compare the predictions to the observed yield in the signal region kinematic variables. The p_T^{miss} is the observable which provide the strongest discriminator the various SUSY signals considered and the SM backgrounds. To visualize the results, the p_T^{miss} of the stacked predicted SM backgrounds are overlayed with the observed data and the final uncertainties, including statistical and systematic components, are represented as shaded band. In addition to the main observable, the p_T^{miss} , multiple other observables are scrutinized as a means to ascertain that significant discrepancies between the data and the simulation are not present in the signal region. As can be seen, there is a good agreement between the predicted backgrounds and the observed data, indicating the absence of a

significant excess. In absence of an excess in data, the subsequent section will contain the interpretation of the results in terms of models of new physics.

On-Z search

The results of the On-Z search for gluinos is presented in Table 14-10 and Table 14-11, and visualized in Figure 14-5. The largest background is stemming from Drell-Yan process in all signal regions and p_T^{miss} regions. The data is generally consistent with the SM prediction and no significant excess is observed, well within the systematic uncertainties.

Table 14-10. Standard model background predictions compared to observed yield in the On-Z signal regions (SRA-SRB) [59].

SRA (b-veto)			
p_T^{miss} [GeV]	100–150	150–250	> 250
DY+jets	13.6±3.1	2.5±0.9	3.3±2.4
FS bkg.	0.4 ^{+0.3} _{-0.2}	0.2 ^{+0.2} _{-0.1}	0.2 ^{+0.2} _{-0.1}
Z + ν	0.8±0.3	1.4±0.4	2.4±0.8
Total background	14.8±3.2	4.0±1.0	5.9±2.5
Data	23	5	4
SRA (b-tag)			
p_T^{miss} [GeV]	100–150	150–250	> 250
DY+jets	8.2±2.1	1.2±0.5	0.5±0.3
FS bkg.	2.3±0.8	1.7 ^{+0.7} _{-0.6}	0.1 ^{+0.2} _{-0.1}
Z + ν	1.9±0.4	2.0±0.5	1.8±0.6
Total background	12.4±2.3	4.9±1.0	2.5±0.7
Data	14	7	1
SRB (b-veto)			
p_T^{miss} [GeV]	100–150	150–250	> 250
DY+jets	12.8±2.3	0.9±0.3	0.4±0.2
FS bkg	0.4 ^{+0.3} _{-0.2}	0.4 ^{+0.3} _{-0.2}	0.1 ^{+0.2} _{-0.1}
Z + ν	0.3±0.1	0.7±0.2	1.2±0.4
Total background	13.6±2.4	2.0±0.5	1.6±0.4
Data	10	4	0
SRB (b-tag)			
p_T^{miss} [GeV]	100–150	150–250	> 250
DY+jets	7.7±3.2	4.0±3.4	0.1±0.1
FS bkg.	1.4 ^{+0.6} _{-0.5}	1.1 ^{+0.5} _{-0.4}	0.2 ^{+0.2} _{-0.1}
Z + ν	2.0±0.5	2.3±0.6	1.0±0.3
Total background	11.1±3.3	7.4 ^{+3.5} _{-3.4}	1.3 ^{+0.4} _{-0.3}
Data	10	5	0

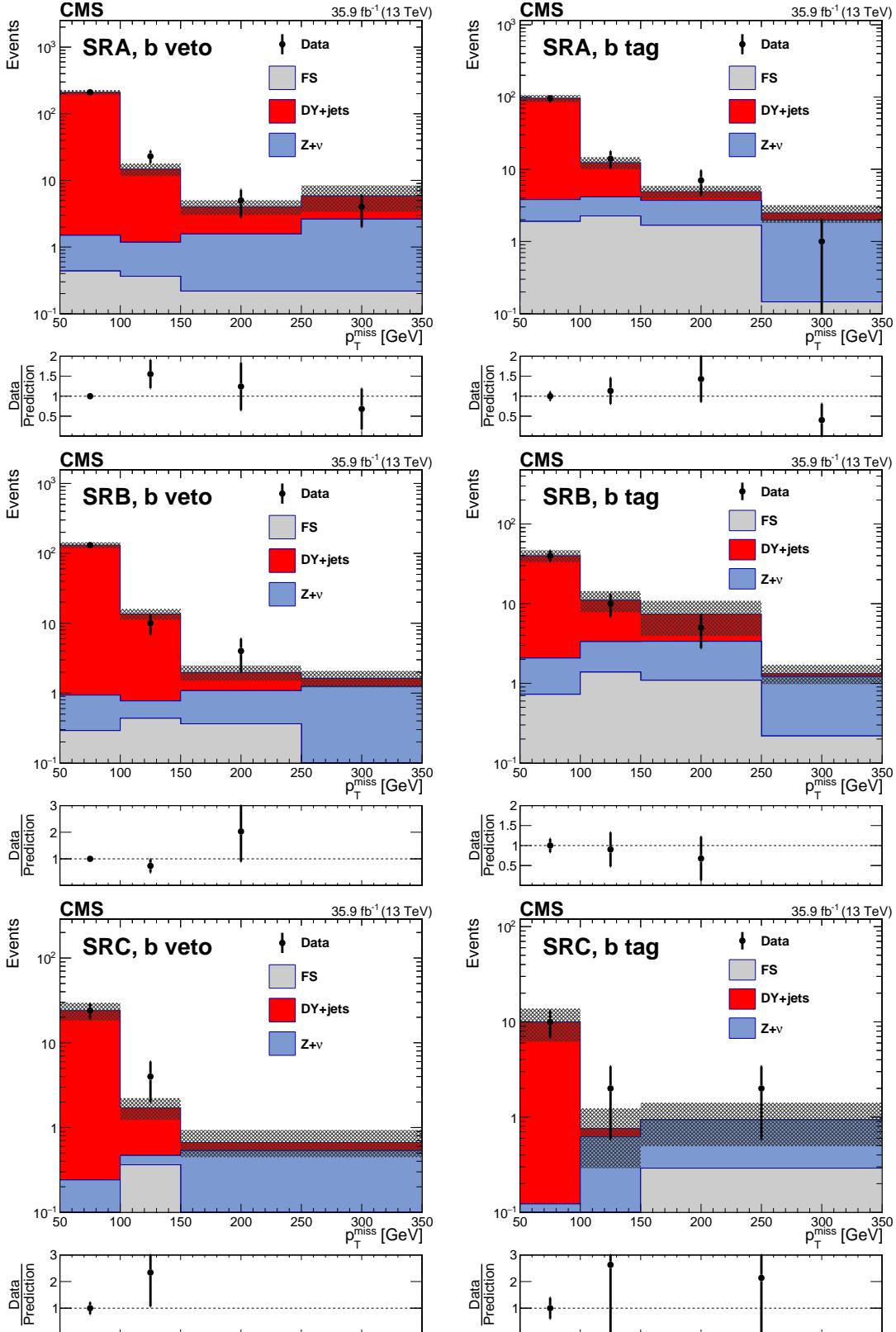


Figure 14-5. The p_T^{miss} distribution is shown for data compared to the background prediction in the on-Z SRs with no b-tagged jets (left) and at least 1 b-tagged jet (right). The rows show SRA (upper), SRB (middle), and SRC (lower) [59].

Table 14-11. Standard model background predictions compared to observed yield in the strong On-Z signal regions (SRC) [59].

SRC (b-veto)		
p_T^{miss} [GeV]	100–150	> 150
DY+jets	1.2±0.4	0.1±0.1
FS bkg.	0.4 ^{+0.3} _{-0.2}	0.1 ^{+0.2} _{-0.1}
Z + ν	0.1±0.1	0.5±0.2
Total background	1.7±0.5	0.7 ^{+0.3} _{-0.2}
Data	4	0
SRC (b-tag)		
p_T^{miss} [GeV]	100–150	> 150
DY+jets	0.1±0.4	0.0±0.3
FS bkg.	0.0 ^{+0.1} _{-0.0}	0.3±0.2
Z + ν	0.6±0.2	0.6±0.2
Total background	0.8±0.5	0.9 ^{+0.5} _{-0.4}
Data	2	2

Edge search

The edge search is binned in the invariant mass of the SF leptons, instead of in p_T^{miss} . Seven $m_{\ell\ell}$ bins are further split into two categories according to the $t\bar{t}$ likelihood discriminant, resulting in a total of 14 signal regions. The predicted SM backgrounds are compared to the observed data in Table 14-12 and summarized in Figure 14-7 and 14-6.

Table 14-12. Predicted and observed yields in each bin of the edge search counting experiment. The uncertainties shown include both statistical and systematic sources [59].

$m_{\ell\ell}$ range [GeV]	FS bkg.	DY+jets	Z + ν	Total background	Data
tt-like					
20–60	291 ⁺²¹ ₋₂₀	0.4±0.3	1.4±0.5	293 ⁺²¹ ₋₂₀	273
60–86	181 ⁺¹⁶ ₋₁₅	0.9±0.7	8.8±3.4	190 ⁺¹⁶ ₋₁₅	190
96–150	176 ⁺¹⁵ ₋₁₄	1.1±0.9	6.0±2.4	182 ⁺¹⁶ ₋₁₅	192
150–200	73 ⁺¹⁰ ₋₉	0.1±0.1	0.4±0.2	74 ⁺¹⁰ ₋₉	66
200–300	46.9 ^{+8.4} _{-7.3}	< 0.1	0.3±0.1	47.3 ^{+8.4} _{-7.3}	42
300–400	18.5 ^{+5.7} _{-4.5}	< 0.1	< 0.1	18.6 ^{+5.7} _{-4.5}	11
> 400	4.3 ^{+3.4} _{-2.1}	< 0.1	< 0.1	4.5 ^{+3.4} _{-2.1}	4
Not-tt-like					
20–60	3.3 ^{+3.2} _{-1.8}	0.7±0.5	1.4±0.5	5.3 ^{+3.3} _{-1.9}	6
60–86	3.3 ^{+3.2} _{-1.8}	1.6±1.3	6.9±2.7	11.8 ^{+4.4} _{-3.5}	19
96–150	6.6 ^{+3.9} _{-2.6}	1.9±1.5	6.8±2.7	15.3 ^{+5.0} _{-4.1}	28
150–200	5.5 ^{+3.7} _{-2.4}	0.2±0.3	0.7±0.3	6.4 ^{+3.7} _{-2.4}	7
200–300	3.3 ^{+3.2} _{-1.8}	0.2±0.2	0.5±0.2	3.9 ^{+3.2} _{-1.8}	4
300–400	3.3 ^{+3.2} _{-1.8}	< 0.1	0.2±0.1	3.5 ^{+3.2} _{-1.8}	0
> 400	1.1 ^{+2.5} _{-0.9}	< 0.1	0.4±0.2	1.6 ^{+2.5} _{-0.9}	5

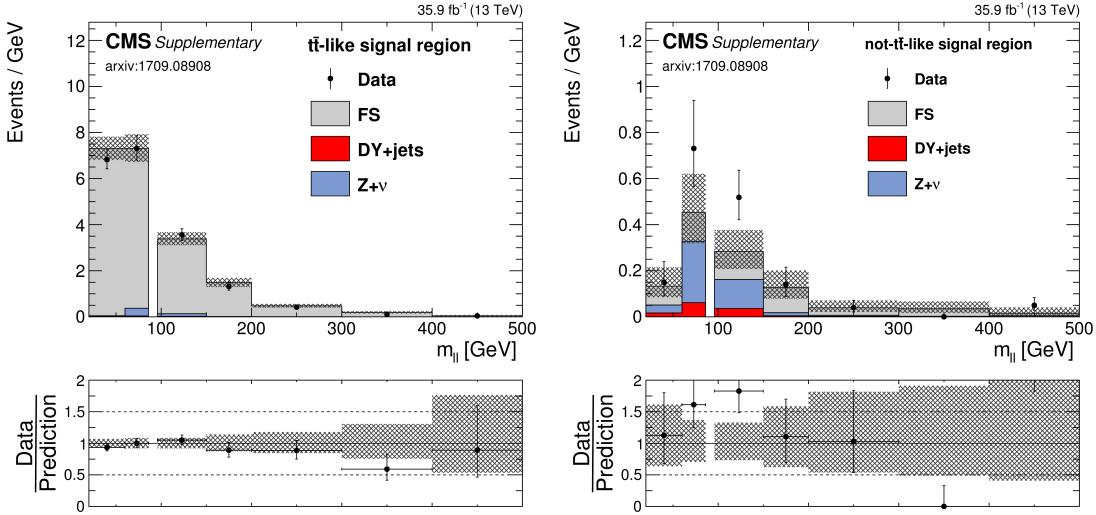


Figure 14-6. $m_{\ell\ell}$ distribution in the $t\bar{t}$ -like (left) and non- $t\bar{t}$ -like edge signal regions. The number of observed events, shown as black data points, is compared to the total background estimate, shown as a blue line with a blue uncertainty band. The non flavor symmetric background component from instrumental p_T^{miss} is indicated as a green area while the non flavor symmetric background with neutrinos is shown as a violet area [59].

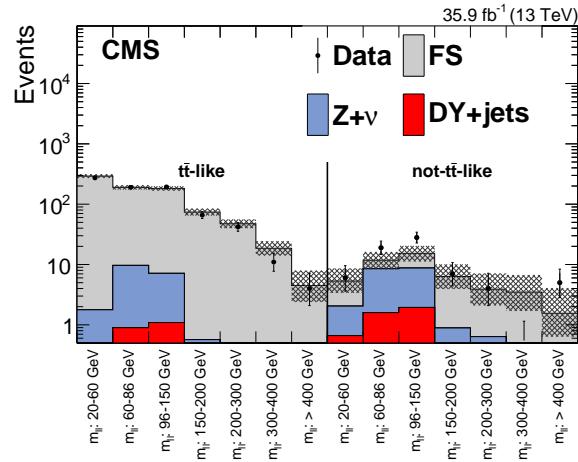


Figure 14-7. Results of the counting experiment of the edge search. For each SR, the number of observed events, shown as black data points, is compared to the total background estimate. The hashed band shows the total uncertainty in the background prediction, including statistical and systematic sources [59].

The main SM background in the edge search is due to the flavor symmetric processes. As the uncertainty on the flavor symmetric background prediction technique is driven by the statistical uncertainty in the number of events in the OF control sample, this becomes the dominant uncertainty in the high-mass and non-tt-like regions, where the event counts are low. There is good agreement between prediction and observation for all SRs. There is a small deviation observed in the non-tt-like region for masses between 96 and 150 GeV, with an excess corresponding to a local significance of 2.0 standard deviations.

Edge fit results

The invariant mass of the leptons and the results of the kinematic fit are shown in Figure 14-8, with the fit results presented in Table 14-13. When evaluating the signal hypothesis in the baseline signal region, a signal yield of 61 ± 28 events is obtained, with a fitted edge position of $144.2^{+3.3}_{-2.2}$ GeV. This is in agreement with the upwards fluctuations in the signal region in $m_{\ell\ell}$ 96-150 GeV and corresponds to a local significance of 2.3 standard deviations. To estimate the global p -value [128] of this result, the test statistic,

$$q(\mu) = -2 \ln \frac{\mathcal{L}(\text{data}|\mu, \hat{\theta}(\mu))}{\mathcal{L}(\text{data}|\hat{\mu}, \hat{\theta}(\mu))} \quad (14-3)$$

introduced in Chapter 12 is used. The test statistics is evaluated on data and compared to the respective quantity on a large sample of background-only pseudo-experiments where the edge position can have any value. The resulting p -value is interpreted as the one-sided tail probability of a Gaussian distribution and corresponds to an excess in the observed number of events compared to the SM background prediction with a global significance of 1.5 standard deviations.

Table 14-13. Results of the unbinned maximum likelihood fit for event yields in the edge fit SR, including the Drell–Yan and FS background components, along with the fitted signal contribution and edge position. The fitted value for $R_{\text{SF}/\text{OF}}$ and the local and global signal significances in terms of standard deviations are also given. The uncertainties account for both statistical and systematic components.

Drell–Yan yield	191 ± 19
FS yield	768 ± 24
$R_{\text{SF}/\text{OF}}$	1.07 ± 0.03
Signal yield	61 ± 28
$m_{\ell\ell}^{\text{edge}}$	$144.2^{+3.3}_{-2.2}$ GeV
Local significance	2.3 s.d.
Global significance	1.5 s.d.

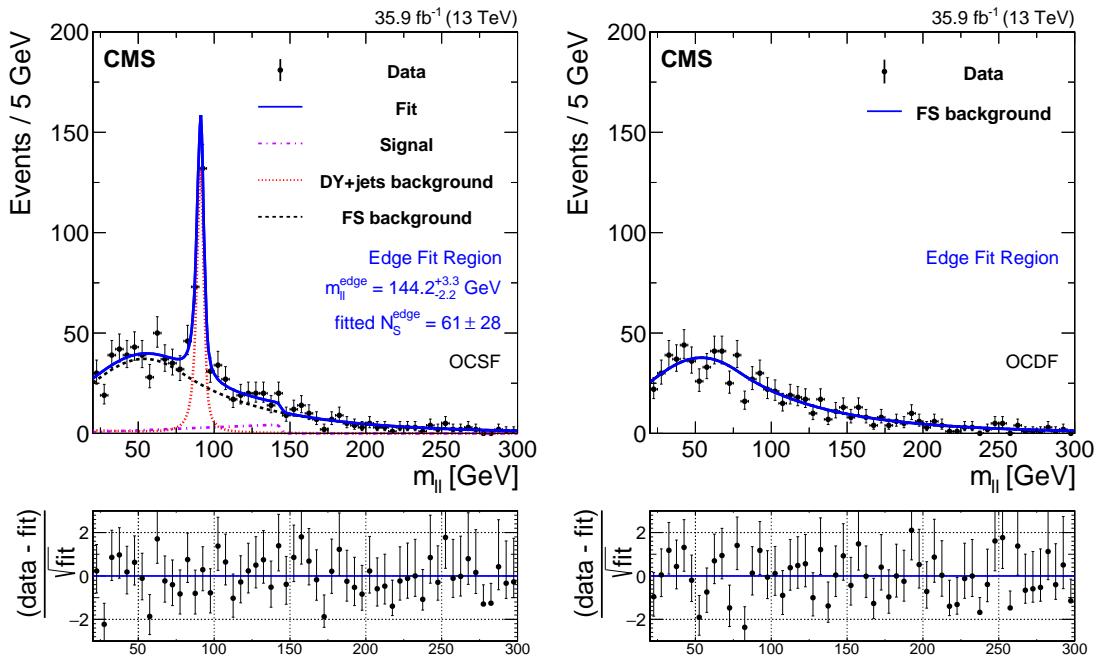


Figure 14-8. Fit of the dilepton invariant mass distributions to the signal-plus-background hypothesis in the “Edge fit” SR from Table 14-2, projected on the same flavor (left) and opposite flavor (right) event samples. The fit shape is shown as a solid blue line. The individual fit components are indicated by dashed and dotted lines. The FS background is shown with a black dashed line. The Drell-Yan background is displayed with a red dotted line. The extracted signal component is displayed with a purple dash-dotted line. The lower panel in each plot shows the difference between the observation and the fit, divided by the square root of the number of fitted events [59].

14.5 Interpretation

The search for colored superpartners is interpreted in terms of exclusion of gluino and sbottom masses. The gluino GMSB model leads to a signature containing at least six jets in the final state when one of the Z bosons decays leptonically and the other decays hadronically. Therefore, most of the sensitivity of the On-Z search is provided by the high jet multiplicity SRs. All of the On-Z strong production SRs are considered, however, to set limits in this model, as signal jets can be lost due to p_T or η acceptance, or not be well reconstructed. The expected and observed limits are presented in Figure 14-9 as a function of the \tilde{g} and $\tilde{\chi}_1^0$ masses. Gluino masses up to 1500–1770 GeV depending on the mass of $\tilde{\chi}_1^0$ are excluded. This represents an improvement of around 500 GeV compared to the previously published CMS result [126].

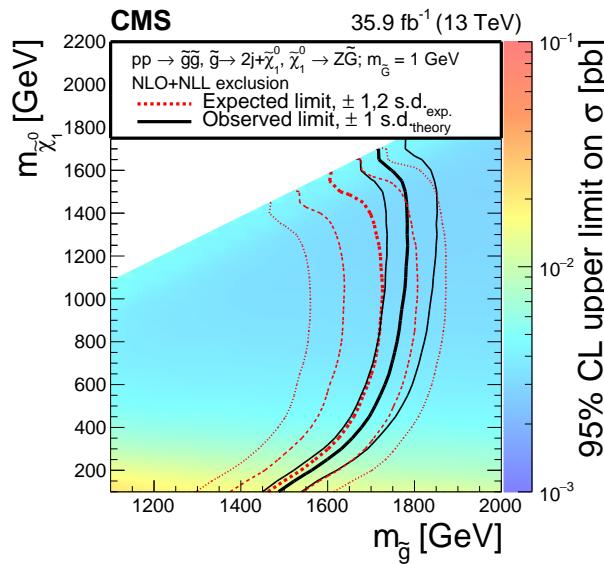


Figure 14-9. Cross section upper limit and exclusion contours at 95% CL for the gluino GMSB model as a function of the \tilde{g} and $\tilde{\chi}_1^0$ masses, obtained from the results of the strong production On-Z search. The region to the left of the thick red dotted (black solid) line is excluded by the expected (observed) limit. The thin red dotted curves indicate the regions containing 68 and 95% of the distribution of limits expected under the background-only hypothesis. The thin solid black curves show the change in the observed limit due to variation of the signal cross sections within their theoretical uncertainties [59].

The Edge search is interpreted using the slepton edge model that targets the direct sbottom production, combining the seven $m_{\ell\ell}$ bins and the two $t\bar{t}$ likelihood regions. In Figure 14-10 is the exclusion contour shown as a function of the \tilde{b} and $\tilde{\chi}_2^0$ masses. Sbottom masses are excluded up to around 980–1200 GeV, depending on the mass of $\tilde{\chi}_2^0$, extending previous exclusion limits in the same model by 400–600 GeV. A decrease of the sensitivity is observed for those models where the $\tilde{\chi}_2^0$ mass is in the range 200–300 GeV. The $m_{\ell\ell}$ distribution for these models has an edge in the range 100–200 GeV, and most of the signal events fall either into the SRs with the highest background prediction or in the range $86 < m_{\ell\ell} < 96$ GeV, which is excluded in this search to reduce Drell–Yan contributions. The observed limit in this regime is weaker than the expected one due to the deviation in the not- $t\bar{t}$ -like, 96–150 GeV mass bin. For high $\tilde{\chi}_2^0$ masses, the majority of signal events fall into the highest mass bins, which are nearly background free. This results in increased sensitivity for these mass points. A weaker observed limit is observed in the highest $m_{\ell\ell}$, non- $t\bar{t}$ likelihood bin, where 5 events are

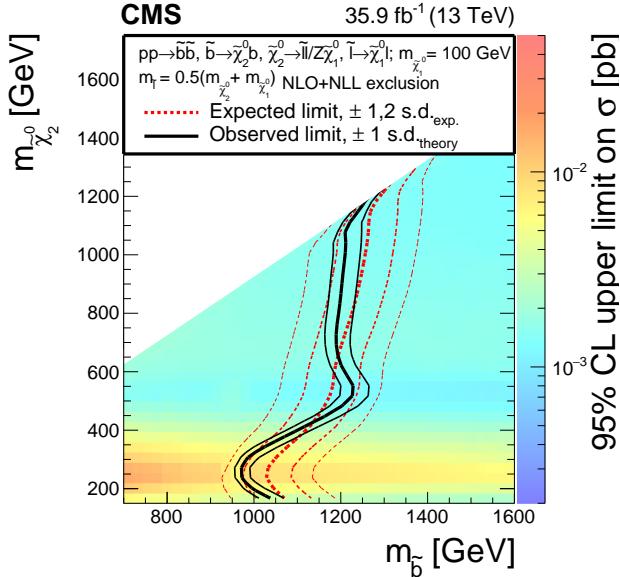


Figure 14-10. Cross section upper limit and exclusion contours at 95% CL for the slepton edge model as a function of the b and $\tilde{\chi}_2^0$ masses, obtained from the results of the Edge search. The region to the left of the thick red dotted (black solid) line is excluded by the expected (observed) limit. The thin red dotted curves indicate the regions containing 68 and 95% of the distribution of limits expected under the background-only hypothesis. The thin solid black curves show the change in the observed limit due to variation of the signal cross sections within their theoretical uncertainties [59].

observed and 1.6 expected. The not- $t\bar{t}$ -like $m_{\ell\ell}$ bin of 300–400 GeV contains 0 observed events compared to an expectation of 3.5 events, yielding the stronger observed limit for the $\tilde{\chi}_2^0$ masses of about 500 GeV.

14.6 Summary

In this chapter, a search for colored superpartners in events with opposite sign, same flavor leptons, jets, and missing transverse momentum has been presented. As this search had a significant excess in Run 1 in the direct sbottom production model, it was of uttermost importance to provide an accurate thorough analysis with the Run 2 dataset. For this reason five institutes provided three different frameworks for cross-checks, validation and synchronization, where I was responsible for one of the frameworks.

The first search for colored superpartners is a search for gluinos that produce a cascade decay involving an on-shell Z boson. The search is designed to target a dilepton invariant mass ($m_{\ell\ell}$) compatible with the Z boson. The other search for colored superpartners assume direct production of sbottom quarks, and the search is designed to target a kinematic edge in the distribution of $m_{\ell\ell}$ as a result of a decay of an intermediate slepton. By comparing the observation to estimates for SM backgrounds obtained from data control samples, no statistically significant evidence for a signal has been observed. The reported excesses from Run 1 are thus not confirmed.

Gluino masses below 1500–1770 GeV have been excluded, depending on the neutralino mass, extending the exclusion limits derived from the previous CMS publication by almost 500 GeV. Bottom squark masses below 980–1200 GeV have been excluded, depending on the mass of the second neutralino. These extend the previous CMS exclusion limits in the same model by 400–600 GeV.

15 — Search for electroweak superpartner production

This chapter is devoted to the search for $\tilde{\chi}_1^\pm \tilde{\chi}_2^0$ and $\tilde{\chi}_1^0 \tilde{\chi}_1^0$ that give rise to a final state of two opposite sign same flavor leptons. Models of electroweak superpartner production are attractive avenues for discovery of new physics under the assumption that other new physics particles are heavy and decoupled and therefore not possible to produce at the LHC. The success of the many colored SUSY searches performed by the CMS and ATLAS experiments, one of which is just presented in Chapter 14, has been pushing the limits on colored superpartners. This success is shifting the attention of the community towards the electroweak SUSY scenarios, that are still not fully excluded.

The search for electroweak superpartners using opposite sign same flavor final states has many similarities to the search for colored superpartners. This is reflected in the estimations of the SM background, that are almost identical. Instead, in this chapter, I highlight my most significant contribution to this search, namely the design of the signal regions in the search for charginos, neutralinos and higgsinos, the validation of the background estimation techniques in these regions, the results and the statistical analysis. All figures without citation are taken from the CMS Analysis Note AN-16-482.

15.1 Analysis strategy

Pairs of electrons or pairs of muons are selected using the lepton selection criteria outlined in Section 10.3. The offline p_T threshold for jets identified according to the requirements listed in Section 10.4 is set to 35 GeV, and is kept at 25 GeV for b-tagged jets. The electroweak superpartner searches presented in Section 4.3 have two defining features. The first one is the production of an on-shell Z boson. The second is the production of a vector boson (W or Z) or a Higgs boson that decay hadronically, to first order to two jets. The production of $\tilde{\chi}_1^\pm$ and $\tilde{\chi}_2^0$ is targeted by a signal region called "VZ" SR. The production of mass degenerate higgsinos, that decay immediately to two $\tilde{\chi}_1^0$'s, is targeted by two different SRs depending on the branching fraction assumed for the decay of the $\tilde{\chi}_1^0$. On the one hand, a branching fraction of 100% is assumed for the decay $\tilde{\chi}_1^0 \rightarrow Z\tilde{G}$ ("ZZ"). In this case the Z boson decay hadronically to light jets or b-tagged jets. The VZ SR is designed to pick up the cases where the Z boson decays to light jets. The other assumption on the branching fraction is 50% of $\tilde{\chi}_1^0 \rightarrow Z\tilde{G}$ and 50% of $\tilde{\chi}_1^0 \rightarrow H\tilde{G}$. The so called "ZH" SR is designed by exploiting the most frequent H boson decay mode, $H \rightarrow b\bar{b}^1$, while still being sensitive to the less frequent decay mode of the $Z \rightarrow b\bar{b}^2$. For all signal regions introduced below, a requirement is imposed that the two jets with the highest p_T have a separation in ϕ from the p_T^{miss} of at least 0.4, in order to reduce the Drell–Yan contribution.

VZ signal region

The idea behind the design of the VZ SR is to target a hadronically decaying vector boson. The Z boson that is the decay product of the $\tilde{\chi}_2^0$ is easily selected by requiring two leptons of opposite sign and same flavor compatible with the Z boson. A naive first step to design this signal region is to try to reconstruct the invariant mass of two jets and require it to be close to the W boson mass. But the existence of more than two jets in the event makes the signal region definition more difficult. It was shown that the largest discrimination between the signal and SM processes was to impose a cut on the invariant mass of two jets to be less than 110 GeV. The choice of 110 GeV was chosen after an optimization, and is large enough to not only target the W boson mass (80.4 GeV), but also the slightly larger Z boson mass (91.2 GeV) [9]. As many SM processes result in more jets than two, a choice on what jets to reconstruct the invariant mass with must be made. The reasoning behind the choice of jets is as follows. A W boson with enough p_T ³ will result in the decay products to propagate in the direction of the mother particle. This results in the decay products of the W boson to be more collimated, and is commonly known as boost. This can happen in the case of large $\tilde{\chi}_2^0$ and $\tilde{\chi}_1^\pm$ masses, where there is enough available momentum to give the Z and W bosons a boost. In the absence of hints of SUSY, we want to set as stringent limits as possible on the different SUSY particle masses. Targeting exclusion of the more massive $\tilde{\chi}_2^0$ and $\tilde{\chi}_1^\pm$, the final state is pushed to this more boosted scenarios. For this reason, if there are more jets than two in the event, the boosted signal feature is targeted by constructing the invariant mass using the jets that are *closest in $\Delta\phi$* .

Further, a veto on b-tagged jets is applied motivated by on the one hand the favored W boson decay mode to light quarks, and on the other hand the suppression of $t\bar{t}$. Finally, as any electroweak SUSY would show up in the p_T^{miss} tails, the VZ SR is binned in p_T^{miss} . Similarly to the signal regions targeting the on-Z colored superpartner production, the two jets with the highest p_T are required to be separated in ϕ from the p_T^{miss} of

¹ 58% [129]

² Z decay to down-type quarks is only 15.2% [9]

³ Enough p_T for a boson V to be considered "boosted" is dictated by $p_T^V \geq \frac{2m_V}{R}$ where R is the radius parameter of the jet clustering algorithm.

at least 0.4. The final VZ SR that targets electroweak production of $\tilde{\chi}_2^0$ and $\tilde{\chi}_1^\pm$ (VZ) or higgsino production (ZZ) is summarized in Table 15-1.

Table 15-1. Summary of the electroweak VZ SR.

Electroweak VZ SR					
Region	N_{jets}	$N_{\text{b-jets}}$	M_{T2} [GeV]	m_{jj} (closest $\Delta\phi$) [GeV]	p_T^{miss} [GeV]
VZ	≥ 2	$= 0$	≥ 80	≤ 110	[100, 150, 250, 350+]

ZH signal region

The electroweak higgsino production has many similarities to the $\tilde{\chi}_2^0$ and $\tilde{\chi}_1^\pm$ targeted by the VZ SR. The similarities lies in the production of an on-shell Z boson, providing the opposite sign same flavor leptons and large p_T^{miss} from the LSPs. Under the assumption that the $\tilde{\chi}_1^0$ can decay with a 50% probability to a Z boson or a H boson and a \tilde{G} LSP, the SR is designed to target the H boson decaying to a pair of bottom quarks. For this, the events in the signal region are required to have exactly two b-tagged jets, that form an invariant mass of less than 150 GeV. As requiring two b-tagged jets increases the existence of SM $t\bar{t}$ to enter the SR, this contribution is reduced by a requirement on a M_{T2} variable. This variable is different than the pure leptonic variable used in the previous SR definitions. Instead this $M_{T2}(\ell b\ell b)$ is defined by pairing each lepton with a b-tagged jet, and all combinations of M_{T2} is calculated and the smallest value is used. This variable has an endpoint at the top quark mass, and to reduce the contribution from $t\bar{t}$ a SR requirement on > 200 GeV is imposed. The ZH SR definition is summarized in Table 16-1.

Table 15-2. Summary of the Electroweak ZH SR.

Electroweak ZH SR					
Region	N_{jets}	$N_{\text{b-jets}}$	m_{bb} [GeV]	$M_{T2}(\ell b\ell b)$ [GeV]	p_T^{miss} [GeV]
ZH	≥ 2	$= 2$	≤ 150	≥ 200	[100, 150, 250+]

15.2 Background estimation

The SM background processes present in the search for electroweak superpartners is very similar to those in the on-Z search for colored superpartners presented in Chapter 14, and are predicted in the same way.

Flavor symmetric background

The flavor symmetric prediction is identical to that presented in Section 14.2.

Z+ ν background

Both the VZ and the ZH signal regions are defined by selecting two leptons that are compatible with an on-shell Z boson. Similarly to the on-Z search for gluinos presented in the previous chapter, a major background stems from ZZ, WZ or t \bar{t} Z, where one of the leptons is not reconstructed. The same approach is taken as described in the previous chapter, that predicts this background from simulation and extracts a transfer factor from a dedicated control region. The same control regions that are described in Section 14.2 are used, which results in the same transfer factors.

Drell–Yan background

The p_T^{miss} template method, presented in Section 11.3, is used to predict the p_T^{miss} spectrum due to instrumental effects in the electroweak signal regions. The prediction from the p_T^{miss} template method is summarized in Table 15-3. The predicted number of events in the electroweak signal region are given with a decomposition of the magnitude of the systematic uncertainties from the four sources considered.

Table 15-3. Summary of template predictions with systematic uncertainties added in quadrature in the strong and electroweak on-Z signal regions together with the individual systematic uncertainties from each source.

SR	p_T^{miss} [GeV]	Prediction	Closure	Normalization	Statistical	EWK sub.
VZ	50-100	773.2 ± 31.9	0.0	29.9	11.1	0.0
	100-150	29.3 ± 4.4	3.2	1.1	2.2	1.8
	150-250	2.9 ± 2.1	0.7	0.1	0.4	1.9
	250-350	1.0 ± 0.7	0.2	0.1	0.2	0.6
	350+	0.3 ± 0.3	0.1	0.1	0.1	0.3
ZH	50-100	76.7 ± 9.4	0.0	9.1	2.4	0.0
	100-150	2.9 ± 2.4	2.3	0.3	0.4	0.2
	150-250	0.3 ± 0.2	0.1	0.1	0.1	0.2
	250+	0.1 ± 0.1	0.1	0.1	0.1	0.1

15.3 Systematic uncertainties

The sources of the systematic uncertainties for the signal simulation used in the search for electroweak superpartners is summarized in Table 15-4. The uncertainties affect the overall normalization of the process, and the nuisance parameters obey a log-normal distribution. The statistical errors on the predicted number of signal events is uncorrelated across the bins, while all other uncertainties are considered correlated across the search regions.

Table 15-4. Systematic uncertainties taken into account for the signal yields and their typical values.

Source of uncertainty	Uncertainty (%)
Integrated luminosity	2.5
Lepton reconstruction and isolation	5
Fast simulation lepton efficiency	4
b-tag modeling	0–5
Trigger modeling	3
Jet energy scale	0–5
ISR modeling	0–2.5
Pileup	1–2
Fast simulation p_T^{miss} modeling	0–4
Renormalization//factorization scales	1–3
MC statistical uncertainty	1–15

15.4 Results

The search for electroweak superpartners is done with two search regions. The VZ signal region is targeting the $\tilde{\chi}_1^\pm$ - $\tilde{\chi}_2^0$ production decaying to a W boson, a Z boson and LSP $\tilde{\chi}_1^0$. This signal region is also sensitive to the production of higgsinos that decay directly to a $\tilde{\chi}_1^0$ - $\tilde{\chi}_1^0$ pair, that decay to two Z bosons and a gravitino LSP. The same production mode is targeted with the ZH signal region, but in this case a the $\tilde{\chi}_1^0$ is assumed to decay democratically to a Z boson and a Higgs boson. Good agreement is observed in all signal regions, presented in Table 15-5 and Table 15-6 and visualized in Figure 15-1.

Table 15-5. Predicted and observed event yields are shown for the electroweak VZ SR, for each p_T^{miss} bin defined in Table 15-1. The uncertainties shown include both statistical and systematic sources.

p_T^{miss} [GeV]	VZ			
	100–150	150–250	250–350	>350
DY+jets	29.3±4.4	2.9±2.0	1.0±0.7	0.3±0.3
FS	11.1±3.6	3.2±1.1	0.1 ^{+0.2} _{−0.1}	0.1 ^{+0.2} _{−0.1}
Z + ν	14.5±4.0	15.5±5.1	5.0±1.8	2.2±0.9
Total background	54.9±7.0	21.6±5.6	6.0±1.9	2.5±0.9
Data	57	29	2	0

Table 15-6. Predicted and observed event yields are shown for the electroweak on-Z SR (ZH), for each p_T^{miss} bin defined in Table 16-1. The uncertainties shown include both statistical and systematic sources.

p_T^{miss} [GeV]	ZH		
	100–150	150–250	>250
DY+jets	2.9±2.4	0.3±0.2	0.1±0.1
FS	4.0±1.4	4.7±1.6	0.9±0.4
Z + ν	0.7±0.2	0.6±0.2	0.3±0.1
Total background	7.6±2.8	5.6±1.6	1.3±0.4
Data	9	5	1

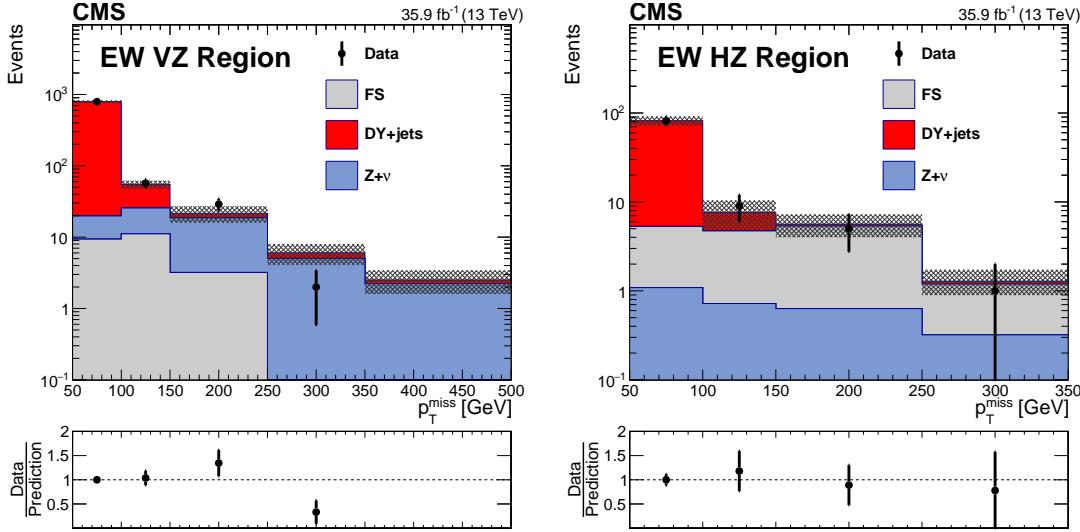


Figure 15-1. The p_T^{miss} distribution is shown for data compared to the background prediction in the on-Z VZ (left) and ZH (right) SRs. The p_T^{miss} template prediction for each SR is normalized to the first bin of each distribution, and therefore the prediction agrees with the data by construction [59].

15.5 Interpretation

The search for electroweak superpartners is interpreted using the models described in Section 4. For the model of $\tilde{\chi}_1^\pm \tilde{\chi}_2^0$ production with decays to a Z boson and a W boson, the VZ SR provides almost all of the sensitivity. Figure 15-2 shows the cross section upper limits and the exclusion lines at 95% CL, as a function of the $\tilde{\chi}_1^\pm$ (or $\tilde{\chi}_2^0$) and $\tilde{\chi}_1^0$ masses. The $\tilde{\chi}_1^\pm$ masses are probed between approximately 160 and 610 GeV, depending on the mass of $\tilde{\chi}_1^0$. The observed yields being smaller than predicted in the two highest p_T^{miss} bins of the VZ SR results in the observed limit being stronger than expected. This result extends the observed exclusion using 8 TeV data by around 300 GeV in the mass of $\tilde{\chi}_1^\pm$ [126].

For the model of $\tilde{\chi}_1^0 \tilde{\chi}_1^0$ production with decays to ZZ, the VZ region contains most of the signal, but the HZ SR accepts the events where the Z boson decays to $b\bar{b}$. The limit is shown in Figure 15-3 as a function of the $\tilde{\chi}_1^0$ mass. $\tilde{\chi}_1^0$ masses are probed up to around 650 GeV. The observed limit is stronger than the expected due to the deficit of observed events in the high- p_T^{miss} bins of the VZ SR. This result extends the observed limit by around 300 GeV compared to the result using 8 TeV data [126]. For the model of $\tilde{\chi}_1^0 \tilde{\chi}_1^0$ production with decays to HZ, the HZ SR dominates the expected limit. The maximal branching fraction to the HZ final state is 50%, achieved when $\tilde{\chi}_1^0$ decays with 50% probability to either the Z or Higgs boson. This scenario includes a 25% branching fraction to the ZZ topology, targeted by the VZ signal region. Limits are set on the 50% branching fraction model in Figure 15-3 using these assumptions and considering the signal contributions from both the ZZ and HZ topologies, targeted by both the VZ and ZH signal regions. In this mixed decay model, $\tilde{\chi}_1^0$ masses are probed up to around 500 GeV. The observed limit at high masses is dominated by the same effect as in the pure ZZ topology. For masses below 200 GeV, the events from the HZ topology alone give an expected exclusion that is 2–5 times more stringent than those from the ZZ topology alone, while for higher masses, the two topologies yield expected limits that are similar to within 30%. The previous exclusion limit using 8 TeV data is extended by around 200 GeV.

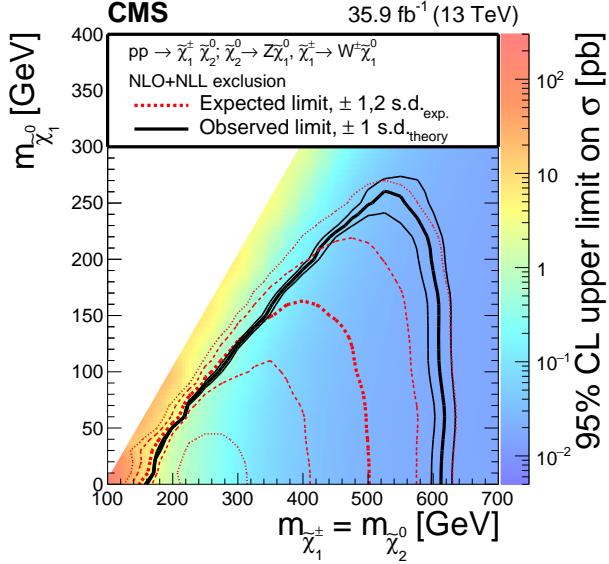


Figure 15-2. Cross section upper limit and exclusion contours at 95% CL for the model of $\tilde{\chi}_1^\pm \tilde{\chi}_2^0$ production, where the exclusion is expressed as a function of the $\tilde{\chi}_1^\pm$ (equal to $\tilde{\chi}_2^0$) and $\tilde{\chi}_1^0$ masses, obtained using the on-Z VZ signal region. The region under the thick red dotted (black solid) line is excluded by the expected (observed) limit. The thin red dotted curves indicate the regions containing 68 and 95% of the distribution of limits expected under the background-only hypothesis. The thin solid black curves show the change in the observed limit due to variation of the signal cross sections within their theoretical uncertainties [59].

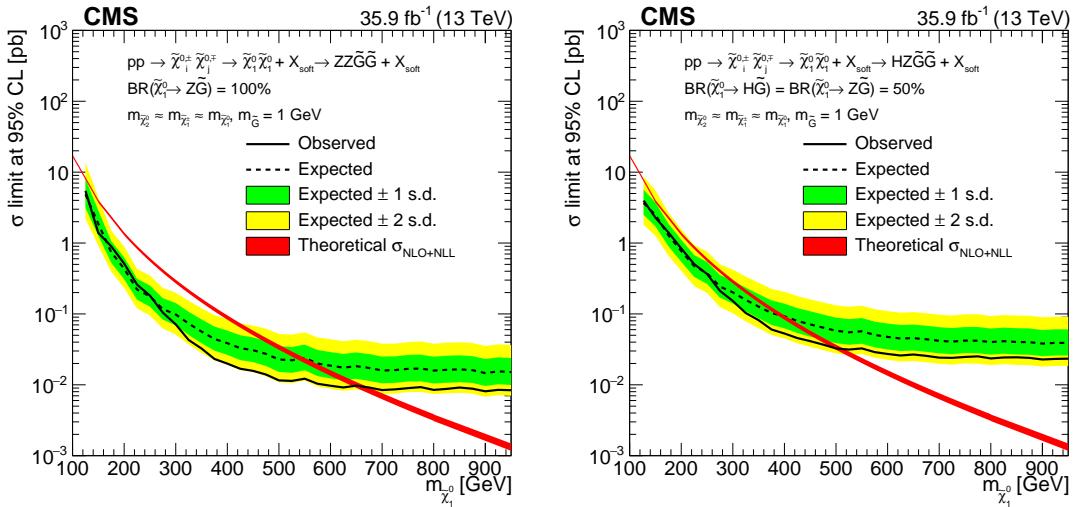


Figure 15-3. Cross section upper limit and exclusion lines at 95% CL, as a function of the $\tilde{\chi}_1^0$ mass, for the search for electroweak production in the ZZ topology (left) and with a 50% branching fraction to each of the Z and Higgs bosons (right). The red band shows the theoretical cross section, with the thickness of band representing the theoretical uncertainty in the signal cross section. Regions where the black dotted line reaches below the theoretical cross section are expected to be excluded. The green (yellow) band indicates the region containing 68 (95)% of the distribution of limits expected under the background-only hypothesis. The observed upper limit on the cross section is shown with a solid black line [59].

15.6 Summary

A search for electroweak superpartners in events with opposite sign, same flavor leptons, jets, and missing transverse momentum has been presented. Searches are performed for signals with a dilepton invariant mass ($m_{\ell\ell}$) compatible with the Z boson. By comparing the observation to estimates for SM backgrounds obtained from data control samples, no statistically significant evidence for a signal has been observed.

The search for electroweak superpartners with an on-shell Z boson has been interpreted in multiple simplified models. For chargino-neutralino production, where the neutralino decays to a Z boson and the lightest supersymmetric particle (LSP) and the chargino decays to a W boson and the LSP, chargino masses are probed in the range 160–610 GeV. In a GMSB model of neutralino-neutralino production decaying to ZZ and LSPs, neutralino masses are probed up to around 650 GeV. Assuming GMSB production where the neutralino has a branching fraction of 50% to the Z boson and 50% to the Higgs boson, neutralino masses are probed up to around 500 GeV. Compared to published CMS results using 8 TeV data, these extend the exclusion limits by around 200–300 GeV depending on the model.

16 — Search for direct slepton production

Finally, the production mode with the lowest cross section associated to it is treated. Direct production of selectrons and smuons provide a very clean final state that makes for a simple but effective search design. To first order, no hadronic activity is expected, giving a great handle to reducing SM background processes by imposing a jet veto. A 100% branching ratio of sleptons to leptons and the $\tilde{\chi}_1^0$ LSP is assumed, that allows for a search with large p_T^{miss} requirements.

In any big collaboration such as CMS, no researcher can solely claim the responsibility of the results of an analysis. Instead, all analyses rely on the support of thousands of colleagues in building the experiment, collecting the data, validating the data, deriving corrections and so on. While keeping the importance of collaboration in mind, one can still acknowledge individual contributions. As opposed to the searches presented in the two previous chapters that were performed in collaboration with four other institutes, the search for direct slepton production was performed solely by the author of this thesis.

This chapter summarizes the analysis strategy, the various background prediction techniques and concludes by setting limits on a selectrons and smuons. All figures without citation are taken from the CMS Analysis Note AN-17-170.

16.1 Analysis strategy

The searches presented so far have the commonality of requiring at least two jets in the final state. The search for direct slepton production differs in the fact that no hadronic activity is expected. The slepton SR is therefore designed to veto jets within $|\eta| = 2.4$ and of $p_T > 25 \text{ GeV}$. Further, the slepton SR is designed using the leptonic M_{T2} variable, and the leading lepton p_T required to be above 50 GeV . These specific selections result in different sources of SM backgrounds compared to the colored and electroweak searches presented in the previous chapters. In the search for sleptons, limits are set on the selectrons and smuon production in dielectron and dimuon final states separately. The reason for this is because a priori, the mass of the selectrons and smuons have no reason to be the same, in the same way as the mass of the electrons and muons are not the same. For this reason, the SR is defined for dielectron and dimuon events separately, and combined.

Table 16-1. Summary of the direct slepton production SR.

Slepton SR				
Region	Flavor	N_{jets}	$M_{T2} [\text{GeV}]$	$p_T^{\text{miss}} [\text{GeV}]$
Slepton	$\mu^\pm \mu^\mp + e^\pm e^\mp$	= 0	≥ 90	[100, 150, 225, 300+]
Selectron	$e^\pm e^\mp$	= 0	≥ 90	[100, 150, 225, 300+]
Smuon	$\mu^\pm \mu^\mp$	= 0	≥ 90	[100, 150, 225, 300+]

16.2 Background estimation

Flavor symmetric background

The main FS background in the slepton search is stemming from WW production. This is in contrast to the general strong and electroweak search, where the main background is due to $t\bar{t}$. The $t\bar{t}$ is reduced in the slepton search as all jets are vetoed, making the WW the main FS contribution. As the $t\bar{t}$ enriched control region that is used to measure the $R_{\text{SF/OF}}$ in the direct measurement method is very similar yet orthogonal to the Edge signal region, an extrapolation of this factor to the signal region is easily validated. The problem arise in the slepton search, where the $t\bar{t}$ control region is far from the slepton signal region that has a veto on jets. For this reason, additional checks are performed to validate the various factors of the FS background prediction method as a function of number of jets. As can be seen in Figure 16-1, no trend is observed in the $r_{\mu/e}$ and the $0.5(r_{\mu/e}^{\text{corr.}} + 1/r_{\mu/e}^{\text{corr.}})$ variables in the low jet multiplicities, indicating that there is no problem to extrapolate the results to the slepton SR.

However, a trend is observed for the $R_{\text{SF/OF}}$ from the direct measurement in the 0 jet bin. This can be seen in Figure 16-2. This trend in the first jet bin could indeed cause some worry. This increase in the number of same flavor events in the 0 jet bin is due to the DY process, which is present in the $t\bar{t}$ enriched control region after relaxing the jet requirement. This statement is further motivated by the fact that the discrepancy show up in data and not in MC where only FS processes are included. But since the signal region is designed with a cut of $M_{T2} \geq 90 \text{ GeV}$, the $R_{\text{SF/OF}}$ can instead be validated in the jet veto case by requiring a $M_{T2} \geq 40 \text{ GeV}$ which is reducing a large fraction of DY contribution while maintaining reasonable statistics. Figure 16-3 shows the $R_{\text{SF/OF}}$ as a function of number of jets in the $t\bar{t}$ control region, but with a requirement that the events that have 0 jets must have a M_{T2} within $40 - 90 \text{ GeV}$ (the upper requirement is added to avoid unblinding), and a good agreement between data and simulation is obtained, indicating that the extrapolation to the 0 jet SR is fine as long as the SR is defined with a M_{T2} requirement.

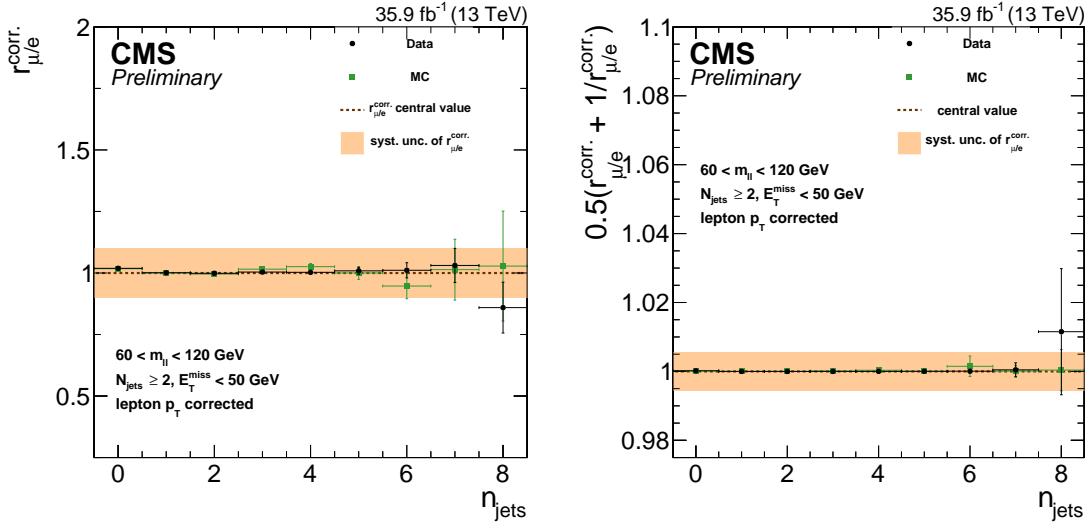


Figure 16-1. The $r_{\mu/e}$ (left) and $0.5(r_{\mu/e}^{\text{corr.}} + 1/r_{\mu/e}^{\text{corr.}})$ (right) as a function of the number of jets for data and MC. No significant trend is observed in the lower jet multiplicities.

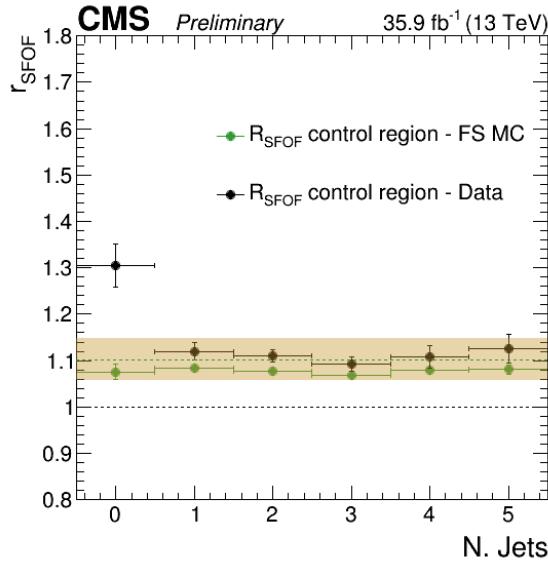


Figure 16-2. $R_{\text{SF/OF}}$ as a function of number of jets derived in a $t\bar{t}$ enriched region after relaxing the 2 jet requirement. The MC simulation takes into account all FS processes. In order to not unbind the signal region, events with are excluded.

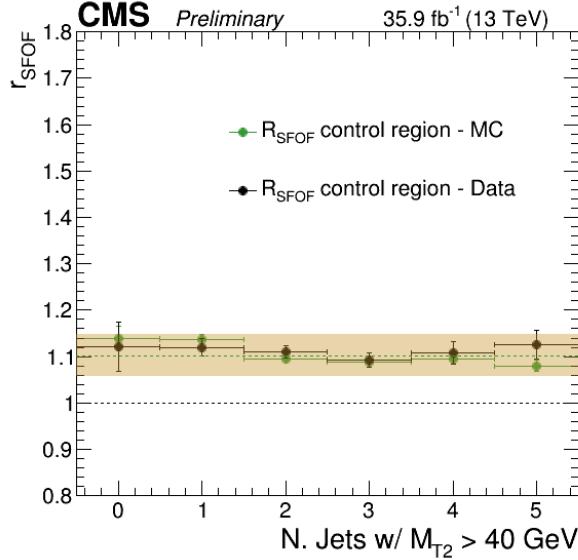


Figure 16-3. $R_{SF/OF}$ as a function of number of jets derived in a $t\bar{t}$ enriched control region after relaxing the 2 jet requirement and requiring that in a 0 jet event, a requirement of the M_{T2} to be within $40 - 90 \text{ GeV}$ is imposed. The MC simulation takes into account all flavor symmetric and non-flavor symmetric processes such as DY.

Further, the slepton search is the only search presented in this thesis that has interpretations in the dielectron and dimuon channels separately. The reason behind this is that there is no a priori reason why the selectrons and smuons should have the same masses, so grouping them together and only interpret in the terms of same-flavor sleptons is not enough.¹ But since the factorization method utilizes the measurement of $r_{\mu/e}$ and R_T , that uses a mixture of measurements for electrons and muons. This means that the $r_{\mu/e}$ or R_T can not be used to predict the N_{ee} or $N_{\mu\mu}$ separately. Instead only the R_{ee}/OF and $R_{\mu\mu}/\text{OF}$ ratios from the direct measurement listed in Table 11-1 can be used to predict the N_{ee} or $N_{\mu\mu}$. The price to pay by only using one of the methods instead of the combined one is that one can not profit from the reduced systematic uncertainty that one gets in the combination of the two methods. But, as might have already been noted, the drawback of the FS background prediction method is the large statistical errors associated to the poorly populated OF event bins in the SR. So a reduction in the systematic uncertainty as a result of a combination of the two methods is still a small reduction compared to the large statistical error. Table 16-2 summarizes the resulting background estimates for FS backgrounds in the slepton signal regions, split into the SF leptons, dielectron and dimuon signal regions.

¹ Of course, the interpretation in selectrons and smuons separately is also desirable as it gives separate entries in the PDG booklet.

Table 16-2. Resulting estimates for flavor-symmetric backgrounds in the Slepton search.

p_T^{miss} [GeV]	N_{OF}	$N_{SF}^{\text{factorization}}$	$R_{\text{SF}/\text{OF}}^{\text{factorization}}$	$R_{\text{SF}/\text{OF}}^{\text{direct}}$	$R_{\text{SF}/\text{OF}}^{\text{combined}}$	N_{SF}^{final}
SF lepton SR						
100-150	88	$97^{+12.5}_{-11.5}$	1.08 ± 0.07	1.11 ± 0.01	1.09 ± 0.01	96^{+13}_{-12}
150-225	14	$15.4^{+5.7}_{-4.5}$	1.08 ± 0.07	1.11 ± 0.01	1.09 ± 0.01	$15.3^{+5.6}_{-4.5}$
225-300	4	$4.4^{+3.4}_{-2.4}$	1.08 ± 0.07	1.11 ± 0.01	1.09 ± 0.01	$4.4^{+3.6}_{-2.3}$
>300	1	$1.1^{+2.6}_{-1.1}$	1.07 ± 0.07	1.11 ± 0.01	1.09 ± 0.01	$1.1^{+2.5}_{-1.0}$
Dielectron SR						
100-150	88	-	-	0.41 ± 0.01	0.41 ± 0.01	$36.1^{+6.6}_{-6.3}$
150-225	14	-	-	0.41 ± 0.01	0.41 ± 0.01	$5.7^{+2.5}_{-2.1}$
225-300	4	-	-	0.41 ± 0.01	0.41 ± 0.01	$1.6^{+1.5}_{-1.1}$
>300	1	-	-	0.41 ± 0.01	0.41 ± 0.01	$0.41^{+1}_{-0.5}$
Dimuon SR						
100-150	88	-	-	0.70 ± 0.01	0.70 ± 0.01	$61.3^{+9.1}_{-8.5}$
150-225	14	-	-	0.70 ± 0.01	0.70 ± 0.01	$9.8^{+3.9}_{-3.2}$
225-300	4	-	-	0.70 ± 0.01	0.70 ± 0.01	$2.8^{+2.4}_{-1.7}$
>300	1	-	-	0.70 ± 0.01	0.70 ± 0.01	$0.7^{+1.7}_{-0.8}$

WZ $\rightarrow 3l\nu$ background

The three lepton control region to target WZ in the slepton search is defined as

- three tight ID leptons of any flavor or charge with $p_T > 25$ GeV for the leading and $p_T > 20$ GeV for the subsequent leptons
- $p_T^{\text{miss}} > 70$ GeV
- the leptons that form the best Z candidate is required to have $m_{\ell\ell}$ within 76 to 106 GeV.
- $M_T > 50$ GeV, where the M_T is constructed with the lepton that does not form the best Z candidate, and is thus assumed to be originating from the W decay.
- Jet veto ($p_T > 25$ GeV)

In Figure 16-4, the p_T^{miss} and the M_{T2} constructed using the leptons from the Z boson decay, in the three lepton control region. The agreement in these signal region variables gives confidence that the WZ process is well modeled. From this region, a transfer factor is derived that is applied in the slepton signal region to account for the slight difference in normalization observed in the three lepton control region. An additional systematic uncertainty is assigned to this method, by letting the jet energy scale corrections vary up and down in the computation of the p_T^{miss} , and the difference in the yield in the background and WZ MC is propagated to the error on the final scale factor. The resulting transfer factor is 1.06 ± 0.06 , where the 6% is the statistical uncertainty on the data in the control region which is taken as a systematic uncertainty on the method. Further, systematic uncertainties such as the variations of the jet energy scale and resolution, as well as PDF and scale variations is evaluated for this method, which will be further explained in Section 13. The transfer factor is presented in Table 16-3, using the MC samples summarized in Table A-3 in Appendix A.

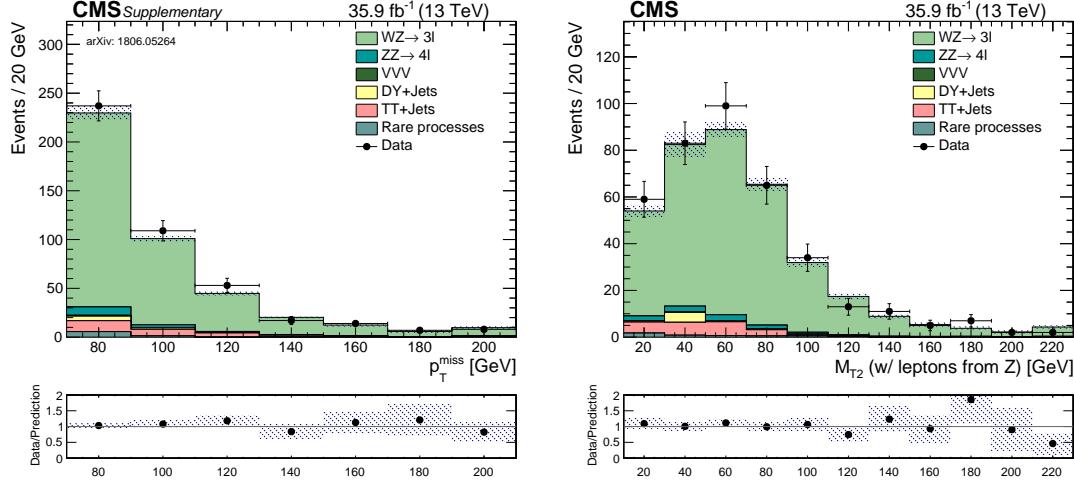


Figure 16-4. The p_T^{miss} (left) and the M_{T2} in the three lepton control region in data and MC [60].

Table 16-3. Transfer factor derived in the three lepton control region for slepton search.

Slepton 3 lepton CR	
signal MC	368.46 ± 6.01
bkg. MC	52.93 ± 4.90
data	445
data-bkg.	392.07 ± 21.66
(data-bkg.)/sig.	1.06 ± 0.06

ZZ → 2l2ν background

In the slepton search, the ZZ → 2l2ν process is dominant, especially in the high p_T^{miss} signal regions. This process can enter the signal region if the leptonically decaying Z boson is off-shell, producing a lepton pair that escapes the Z boson veto. This process is estimated from simulation, after validating the simulation in a four lepton control region, enriched in the ZZ* → 4l process. This control region enables for the validation of the modeling of the lepton pair from the off-shell Z boson, by treating the other two leptons that are compatible with a good Z candidate as neutrinos. The p_T of the leptons forming the best Z candidate is added to the p_T^{miss} , so to mimic the p_T^{miss} produced if these charged leptons were replaced with neutrinos. The agreement between the data and the simulation is checked in variables such as the p_T^{miss} and the M_{T2} recomputed with the remaining leptons and the emulated p_T^{miss} . The four lepton control region used for the slepton search is defined as

- four tight ID leptons of any flavor or charge with $p_T \geq 25 \text{ GeV}$ for the leading and $p_T \geq 20 \text{ GeV}$ for the subsequent leptons
- the leptons that form the best Z candidate is required to have $m_{\ell\ell}$ within 76 to 106 GeV.
- the leptons that form the other Z candidate is required to have $m_{\ell\ell}$ within 50 to 130 GeV.
- veto events with jets ($p_T \geq 25 \text{ GeV}$)

A NNLO/NLO k-factor is applied to the $ZZ \rightarrow 2l2\nu$ and $ZZ \rightarrow 4l$ simulation, which is provided as a function of the generator level p_T , mass and $\Delta\phi$ of the diboson system. The diboson p_T dependent k-factor is eventually applied to the $ZZ \rightarrow 2l2\nu$ prediction in the signal region, and an uncertainty on the shape is derived as the difference between the distributions after applying the p_T dependent k-factor and no k-factor, after both distributions are normalized to their areas. The distributions of the generator p_T and mass of the diboson systems in the slepton signal region is presented in Figure 16-5, after the respective k-factor has been applied. Further information on the k-factors is documented in Appendix A. A four lepton control

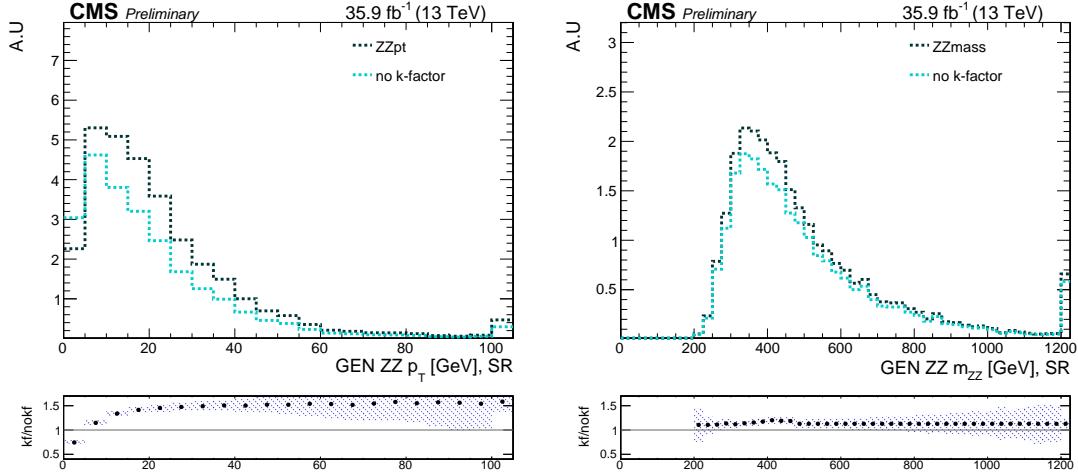


Figure 16-5. Distributions of the generator p_T , mass and $\Delta\phi$ of the diboson system, in the signal region, with and without the respective k-factor applied, in ZZTo2L2Nu MC.

region is constructed, which will be used to estimate the $ZZ \rightarrow 2l2\nu$ background. The four lepton signature is very pure and provides a good data/MC agreement. Figure 16-6 shows the invariant mass of the best Z candidate and the M_{T2} , in data and MC. In Table 16-4 the transfer factor of the four lepton control region is summarized, where the signal simulation is the $ZZ \rightarrow 4l$ process summarized in Appendix A (here all production modes are included as opposed to in the on-Z search where many of the samples were not generated at the time of the publication). A transfer factor of 0.94 is obtained with a statistical uncertainty of 7%, and this 7% is used as a systematic uncertainty on the method. As the slepton signal region is defined with a selection of leading leptons of p_T greater than 50 GeV, this differs from the control region definition of p_T greater than 25 GeV to maintain the statistical power of the sample.

Table 16-4. Transfer factor derived in the four lepton control region for slepton search.

Slepton 4 lepton CR	
signal MC	184.7 ± 1.2
bkg. MC	5.7 ± 1.7
data	179
data-bkg.	173.4 ± 13.5
(data-bkg.)/sig.	0.94 ± 0.07

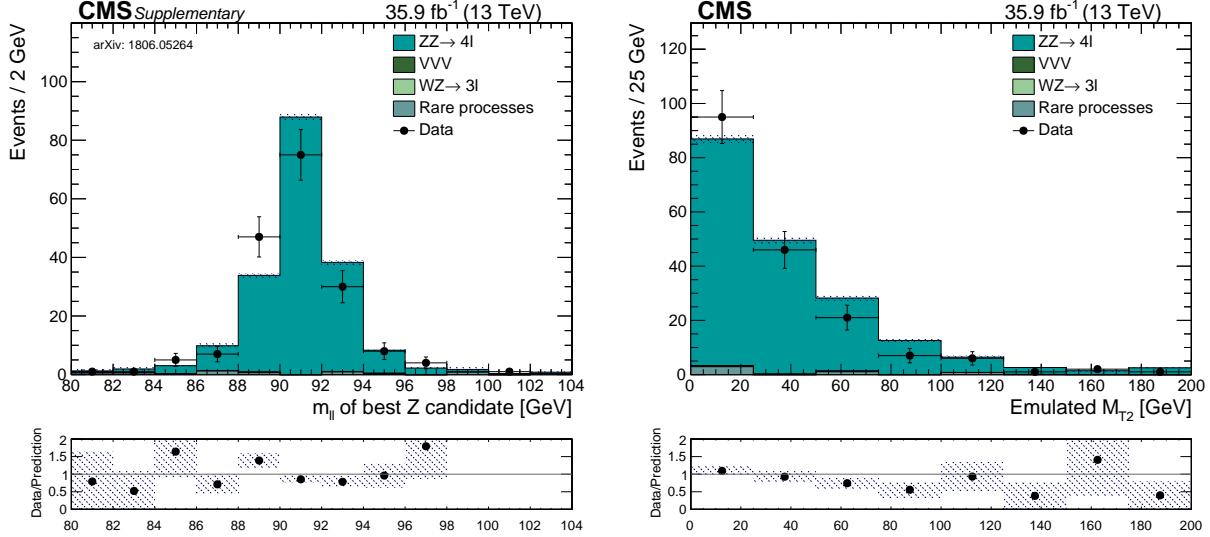


Figure 16-6. The invariant mass of the two lepton pair forming the best Z candidate in the four lepton control region (left). The emulated M_{T2} is the M_{T2} formed with the p_T^{miss} that has the p_T of the leptons forming the best Z candidate added to it, and the remaining two leptons (right).

Drell-Yan

The Drell-Yan is a subdominant background in the slepton search, as it is heavily suppressed by the Z boson veto, the jet veto, the large M_{T2} and p_T^{miss} requirements. For this reason the background is predicted from simulation, in order to avoid large systematic uncertainties due to the poorly populated p_T^{miss} bins. Similarly to the Drell-Yan treatment in the edge search, an $R_{\text{out/in}}$ method is used to transfer the prediction of events on the Z peak to outside of the Z window. The difference is that the events in the Z window is taken from simulation. As the slepton signal region is not binned in the invariant mass like the edge search is, the $R_{\text{out/in}}$ only needs to be computed twice, $R_{\text{out/in}}$ above and below the Z window.

Table 16-5. Measured values for $R_{\text{out/in}}$ for data and MC in the different signal regions of the slepton search.

m_{ee} [GeV]	N_{in}	N_{out}	$R_{\text{out/in}}$
20-76	9096 ± 96	634 ± 26	0.069 ± 0.05
106+	9096 ± 96	628 ± 27	0.069 ± 0.05

16.3 Systematic uncertainties

Systematic uncertainties for diboson backgrounds

The SM contribution from WZ and ZZ processes is taken directly from MC, corrected with a transfer factor and with additional systematic uncertainties propagated. A summary of the sources and the magnitude of the uncertainties considered for ZZ and WZ processes is shown in Table 16-6 and Table 16-7. For the ZZ an additional uncertainty, stemming from the application of the k-factor, is propagated as the difference in shape observed between applying and not applying the generator diboson p_T k-factor.

Table 16-6. List of systematic uncertainties taken into account for the ZZ background prediction method.

Source of uncertainty	Uncertainty (%)
Lepton reconstruction and isolation	5
Trigger modeling	3
Jet energy scale	0.5-2
Renormalization/factorization scales	2-4
PDF	3-17
k-factor shape	1-3
ZZ CR transfer factor	7

Table 16-7. List of systematic uncertainties taken into account for the WZ background prediction method.

Source of uncertainty	Uncertainty (%)
Lepton reconstruction and isolation	5
Trigger modeling	3
Jet energy scale	0.5-6
Renormalization/factorization scales	3-7
PDF	0.5-11
WZ CR transfer factor	6

Systematic uncertainties for signal

The magnitude of the systematic uncertainties in search for direct slepton production is summarized in Table 16-8. In addition to the sources of systematic uncertainties discussed in Chapter 13, additional p_T^{miss} uncertainties are evaluated for the signal simulation in the slepton analysis. As the main physics objects in the slepton analysis are the leptons and the p_T^{miss} , it is crucial to evaluate how the p_T^{miss} changes if the energy of one of the other objects in the event varies. The muon energy scale, the electron energy scale and the unclustered energy scale is varied up and down by 1σ , and the effect is propagated to the p_T^{miss} . This results in a quite sizeable effect, especially in the muon case where the uncertainty is up to 20%, driven by a few signal events with very high p_T muons.

Table 16-8. List of systematic uncertainties taken into account for the signal yields.

Source of uncertainty	Uncertainty (%)
Integrated luminosity	2.5
Lepton reconstruction and isolation	5
Fast simulation electron efficiency	1–2.5
Fast simulation muon efficiency	1–3
Trigger modeling	3
Jet energy scale	1–15
Pileup	0.5–7
Fast simulation p_T^{miss} modeling	0.5–20
Unclustered energy shifted p_T^{miss}	0.5–8
Muon energy scale shifted p_T^{miss}	0.5–20
Electron energy scale shifted p_T^{miss}	0.5–4
Renormalization/factorization scales	1–11
PDF	3
MC statistical uncertainty	0.5–20

16.4 Results

Finally, the predicted background compared to the observed yields in the signal regions designed to target direct slepton production is presented. As opposed to the results previously presented, the signal regions is not only categorized by SF lepton pairs, but is further categorized in dielectron and dimuon pairs. The observed number of events in data in the SR are compared with the stacked SM background estimates as shown in Figure 16-7 (SF events), and summarized in Table 16-9 for SF events and in Table 16-10 for dielectron and dimuon events, separately. The M_{T2} shape of the stacked SM background estimates, the observed data and three signal scenarios are also shown in Figure 16-7, for SF events, with all SR selection applied except the M_{T2} requirement.

Table 16-9. The predicted SM background contributions, their sum and the observed number of SF events in data. The uncertainties associated with the background yields stem from statistical and systematic sources.

p_T^{miss} [GeV]	Same flavor events			
	100–150	150–225	225–300	≥ 300
FS bkg.	96^{+13}_{-12}	$15.3^{+5.6}_{-4.5}$	$4.4^{+3.6}_{-2.3}$	$1.1^{+2.5}_{-1.0}$
ZZ	13.5 ± 1.5	9.78 ± 1.19	2.84 ± 0.56	1.86 ± 0.12
WZ	6.04 ± 1.19	2.69 ± 0.88	0.86 ± 0.45	0.21 ± 0.20
DY+jets	$2.01^{+0.39}_{-0.23}$	$0.00 + 0.28$	$0.00 + 0.28$	$0.00 + 0.28$
Rare processes	0.69 ± 0.44	0.68 ± 0.47	$0.00 + 0.20$	0.05 ± 0.12
Total prediction	118^{+13}_{-12}	$28.4^{+5.9}_{-4.8}$	$7.9^{+3.7}_{-2.4}$	$3.2^{+2.6}_{-1.1}$
Data	101	31	7	7

Table 16-10. The predicted SM background contributions, their sum and the observed number of dielectron (upper) and dimuon (lower) events in data. The uncertainties associated with the yields stem from statistical and systematic source.

Dielectron events				
p_T^{miss} [GeV]	100–150	150–225	225–300	≥ 300
FS bkg.	$36.1^{+6.6}_{-6.3}$	$5.7^{+2.5}_{-2.1}$	$1.6^{+1.5}_{-1.1}$	$0.41^{+1}_{-0.5}$
ZZ	5.17 ± 0.68	3.79 ± 0.58	1.18 ± 0.31	0.69 ± 0.07
WZ	2.65 ± 0.68	1.16 ± 0.45	0.39 ± 0.33	0.21 ± 0.20
DY+jets	$0.98^{+0.14}_{-0.15}$	$0.00 + 0.28$	$0.00 + 0.28$	$0.00 + 0.28$
Rare processes	0.02 ± 0.14	0.26 ± 0.21	$0.00 + 0.11$	0.06 ± 0.04
Total prediction	$45^{+6.7}_{-6.4}$	$11.0^{+2.6}_{-2.3}$	$3.2^{+1.6}_{-1.2}$	$1.4^{+1.1}_{-0.6}$
Data	45	10	2	2

Dimuon events				
p_T^{miss} [GeV]	100–150	150–225	225–300	≥ 300
FS bkg.	$61.3^{+9.1}_{-8.5}$	$9.8^{+3.9}_{-3.2}$	$2.8^{+2.4}_{-1.7}$	$0.70^{+1.7}_{-0.8}$
ZZ	8.33 ± 0.99	5.98 ± 0.80	1.67 ± 0.42	1.17 ± 0.10
WZ	3.40 ± 0.91	1.53 ± 0.73	0.47 ± 0.30	$0.00 + 0.06$
DY+jets	$1.03^{+0.33}_{-0.14}$	$0.00 + 0.28$	$0.00 + 0.28$	$0.00 + 0.28$
Rare processes	0.66 ± 0.41	0.42 ± 0.35	$0.00 + 0.16$	$0.00 + 0.11$
Total prediction	$75^{+9.2}_{-8.7}$	$17.7^{+4.1}_{-3.4}$	$4.8^{+2.5}_{-1.8}$	$1.9^{+1.7}_{-0.8}$
Data	56	21	5	5

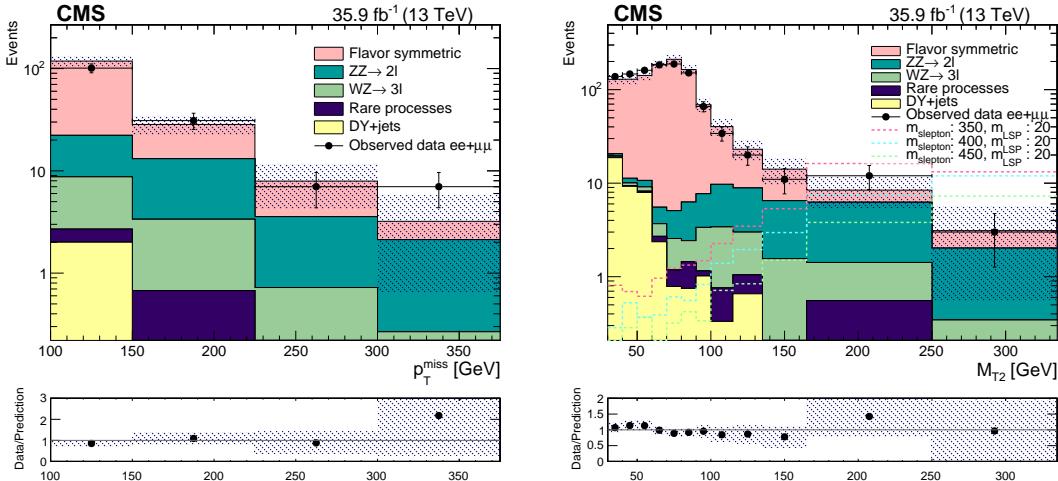


Figure 16-7. The p_T^{miss} (left) distribution for the resulting SM background yields estimated in the slepton SF analysis SR overlayed with the observed data as black points. The M_{T2} distribution (right) in the same signal region with three signal scenarios overlayed [60].

16.5 Interpretation

Upper limits on the direct slepton pair production cross section are displayed in Figure 16-8 for three scenarios: assuming the existence of both flavor mass degenerate left- and right-handed sleptons, for only left-handed sleptons, and for only right-handed sleptons. Similarly, the limits on direct selectron and smuon production

are displayed in Figures 16-9 and 16-10, respectively. The Figures 16-8 16-10 also show the 95% CL exclusion contours, as a function of the $\tilde{\ell}$ and $\tilde{\chi}_1^0$ masses. Note that the cross section at a given mass for right-handed sleptons is expected to be about one third of that for left-handed sleptons. The analysis probes slepton masses up to approximately 450, 400, or 290 GeV, assuming both left- and right-handed, left-handed only, or right-handed sleptons, and a massless LSP. For models with high slepton masses and light LSPs the sensitivity is driven by the highest p_T^{miss} bin. The sensitivity is reduced at higher LSP masses due to the effect of the lepton acceptance. In the case of selectrons (smuons), the limits corresponding to these 3 scenarios are 350, 310 and 250 GeV (310, 280, and 210 GeV). Since the dimuon data yield in the highest p_T^{miss} bin is somewhat higher than predicted, the observed limits in this channel are weaker than expected in the absence of signal. These results improve the previous 8 TeV exclusion limits by 100–150 GeV in the slepton mass [66].

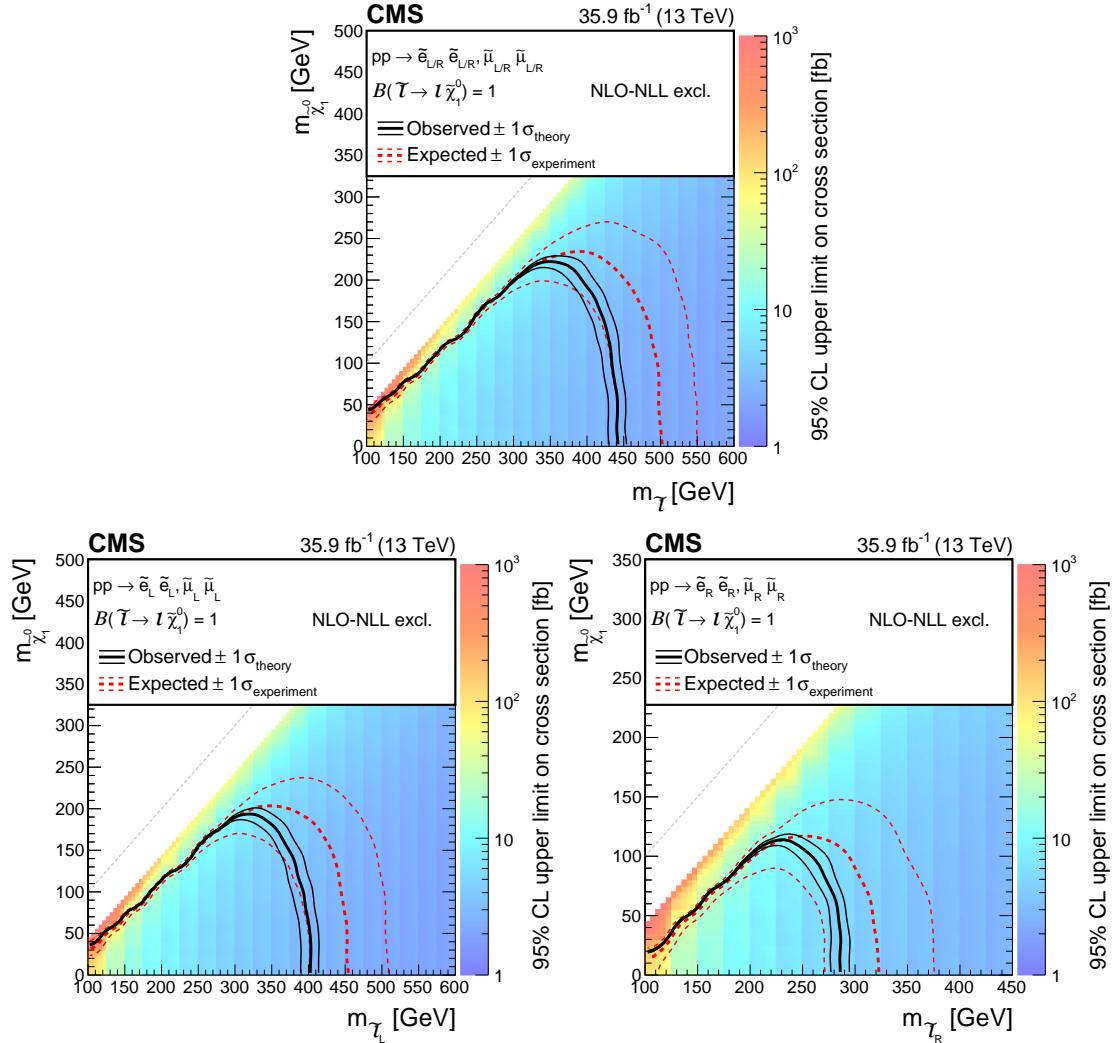


Figure 16-8. Cross section upper limit and exclusion contours at 95% CL for direct slepton production of two flavors, selectrons and smuons, as a function of the $\tilde{\chi}_1^0$ and $\tilde{\ell}$ masses, assuming the production of both left- and right-handed sleptons (upper) or production of only left- (lower left) or right-handed (lower right). The region under the thick red dotted (black solid) line is excluded by the expected (observed) limit. The thin red dotted curves indicate the regions containing 95% of the distribution of limits expected under the background-only hypothesis. The thin solid black curves show the change in the observed limit due to variation of the signal cross sections within their theoretical uncertainties [60].

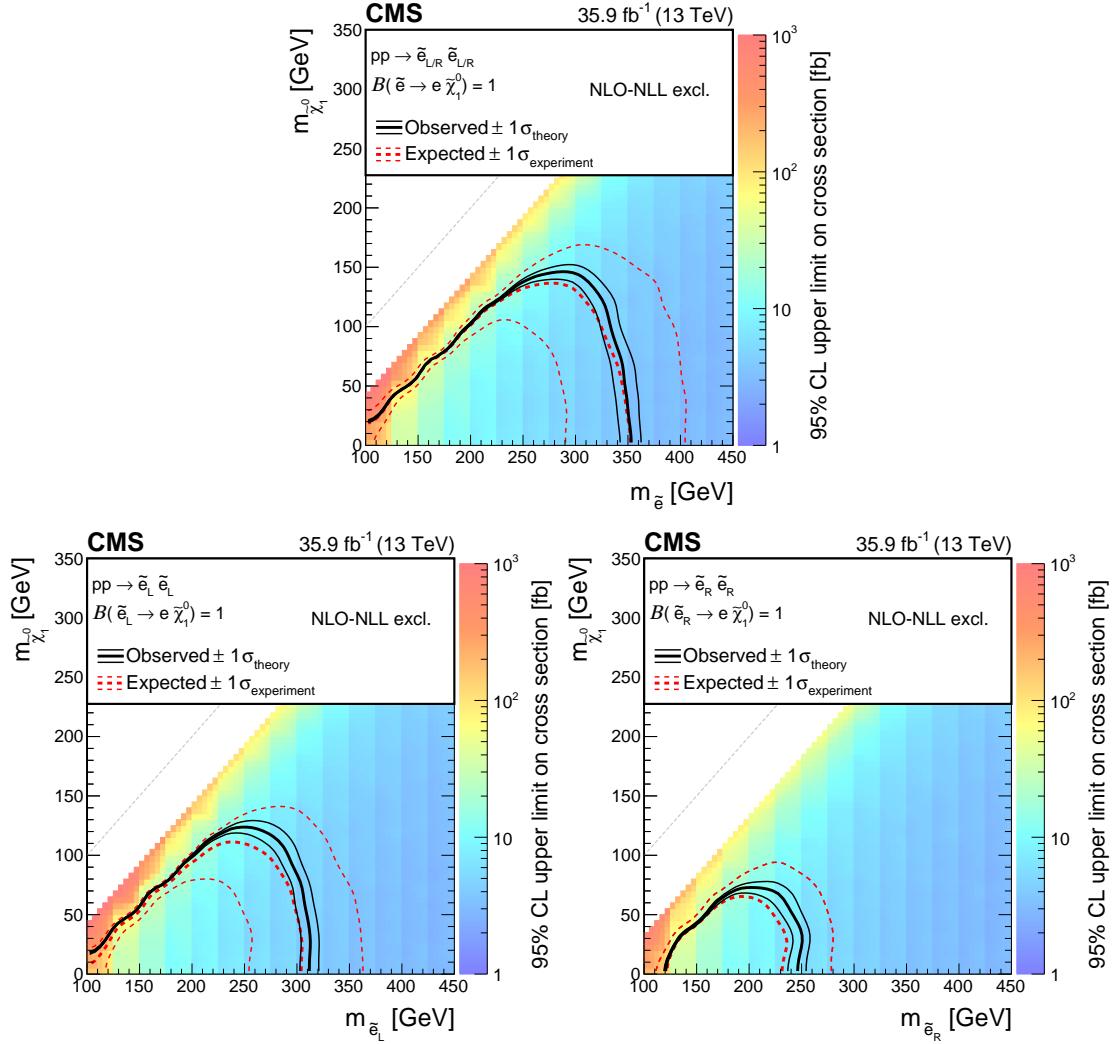


Figure 16-9. Cross section upper limit and exclusion contours at 95% CL for direct selectron production as a function of the $\tilde{\chi}_1^0$ and $\tilde{\ell}$ masses, assuming the production of both left- and right-handed selectrons (upper), or production of only left- (lower left) or right-handed (lower right) selectrons. The region under the thick red dotted (black solid) line is excluded by the expected (observed) limit. The thin red dotted curves indicate the regions containing 95% of the distribution of limits expected under the background-only hypothesis. For the right-handed selectrons, only the $+1\sigma$ expected line (thin red dotted curve) is shown as no exclusion can be made at -1σ . The thin solid black curves show the change in the observed limit due to variation of the signal cross sections within their theoretical uncertainties [60].

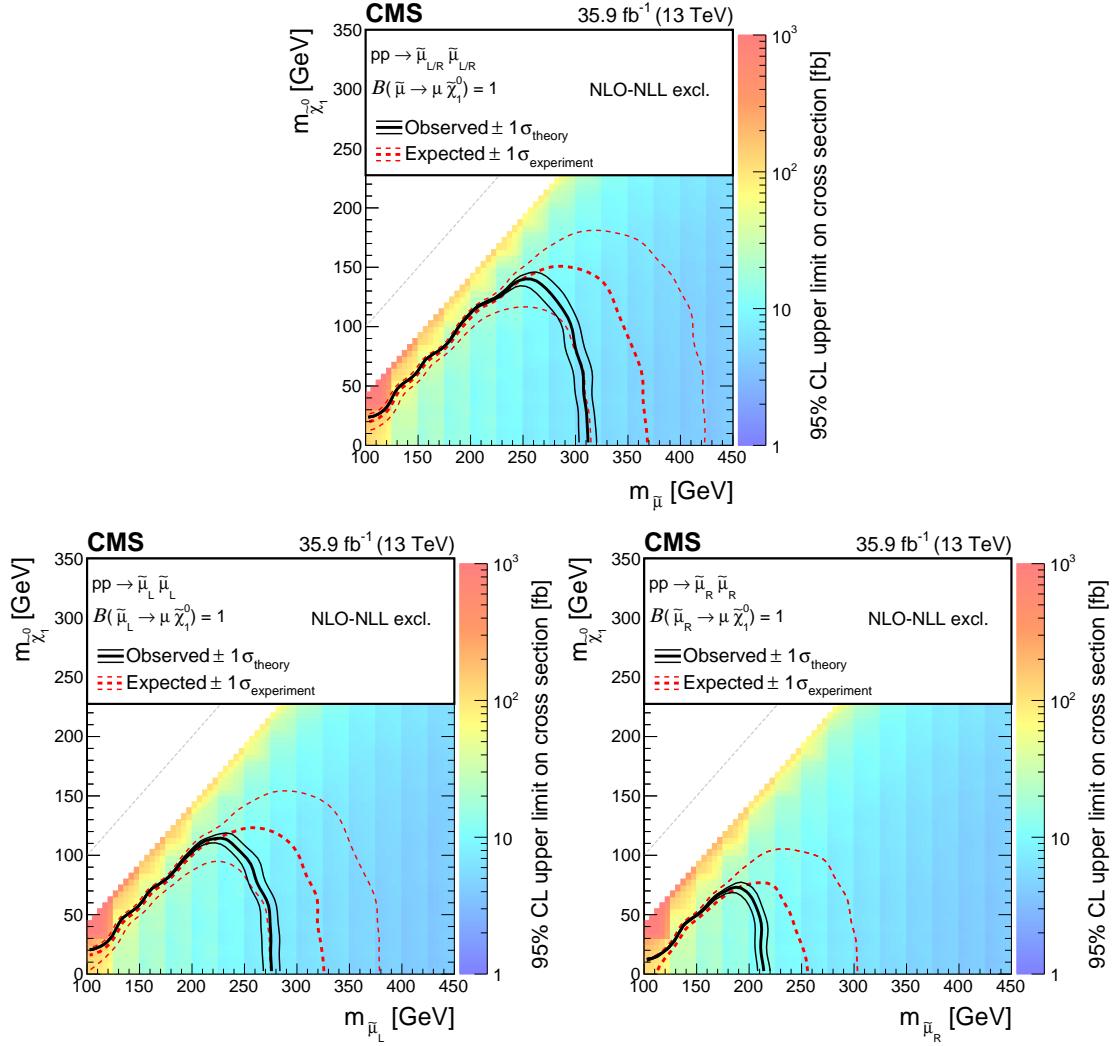


Figure 16-10. Cross section upper limit and exclusion contours at 95% CL for direct smuon production as a function of the $\tilde{\chi}_1^0$ and $\tilde{\ell}$ masses, assuming the production of both left- and right-handed smuons (upper), or production of only left- (lower left) or right-handed (lower right) smuons. The region under the thick red dotted (black solid) line is excluded by the expected (observed) limit. The thin red dotted curves indicate the regions containing 95% of the distribution of limits expected under the background-only hypothesis. For the right-handed smuons, only the $+1\sigma$ expected line (thin red dotted curve) is shown as no exclusion can be made at -1σ . The thin solid black curves show the change in the observed limit due to variation of the signal cross sections within their theoretical uncertainties [60].

16.6 Summary

This chapter has outlined a search for direct slepton (selectron or smuon) production, in events with opposite sign, same flavor leptons, no jets, and missing transverse momentum. Observations are in agreement with expectations from the SM within the statistical and systematic uncertainties. Slepton masses up to 290, 400 and 450 GeV respectively are excluded at 95% confidence level, assuming a massless LSP. Exclusion limits are also provided assuming a massless LSP and right-handed only, left-handed only and right-and left-handed single flavor production scenarios, excluding selectron (smuon) masses up to 250, 310 and 350 GeV (210, 280 and 310 GeV), respectively. These results improve the previous exclusion limits measured by the CMS experiment at a center-of-mass energy of 8 TeV by 100-150 GeV in slepton masses.

17 — Conclusions

The theoretical motivations behind Supersymmetry (SUSY) as an extension of the Standard Model of particle physics (SM) are numerous. The introduction of a new set of SUSY particles, superpartners to the SM particles, has the ability to solve almost all of the outstanding problems with the current formulation of the SM. The overwhelming evidence for the existence of non-luminous matter, dark matter, could be explained by SUSY, that predicts a massive particle that does not interact with ordinary matter through any of the known fundamental forces.

These SUSY particles have the potential of being produced in proton-proton collisions, and indirectly detected by large high energy physics experiments. The ATLAS and CMS experiments, at the LHC at CERN, have conducted searches for SUSY particles during the first operational run of the LHC. This thesis presents a set of searches for SUSY particles at the unprecedented center-of-mass energy of 13 TeV. The strategy deployed in this thesis is to look for events containing two electrons or two muons of opposite electric charge, in association with jets and large missing transverse momentum. This particular final state is a powerful tool, as it is sensitive to many types of SUSY particles.

This thesis has presented searches for SUSY particles produced electroweak and strong production modes, originally published in [59] and [60]. The integrated luminosity of the data sample analyzed corresponds to 35.9 fb^{-1} , collected by the CMS detector at a center-of-mass energy of 13 TeV, during the 2016 running of the LHC. During the Run 1 of the LHC, both CMS and ATLAS reported modest excesses in the strong SUSY production modes. With the increase in center-of-mass energy and luminosity that the Run 2 offered, the first part of this thesis work consisted of confirming these excesses. With the 35.9 fb^{-1} , it became clear that the SM backgrounds model the data very well, thus rejecting the hypothesis that there would be a hint of strongly produced SUSY.

The search continued, now on the hunt for electroweak SUSY. The superpartners of the SM gauge bosons, the charginos and neutralinos, and the higgsinos, are much lower cross section processes, but have great theoretical features if they existed. For example, a light higgsino can solve the hierarchy problem of the SM. Two searches were designed, one targeting chargino and neutralino production, and one targeting higgsino production. My most significant contribution to the search for electroweak superpartners was the design of the signal region that targeted massive electroweak superpartner by selecting boosted decay products. The search for charginos and neutralinos pushed the exclusion of these superpartners by up to 300 GeV, and the search for higgsinos by up to 200 GeV.

My final, and most significant, contribution to the quest for finding physics beyond the SM, was done in the context of superpartners of the SM charged leptons. The direct production of selectrons and smuons is done at a very low rate at hadron colliders, but has the advantage of resulting in a very clean final state, consisting of only two electrons or muons of opposite charge and large p_T^{miss} from the neutralino LSP. Again,

there was no hint of SUSY in the data, but the exclusion limits were pushed by up to 200 GeV with respect to the Run 1 exclusion.

At the time of writing this thesis, the LHC Run 2 data taking has just ended. The analysis strategies exploited in the searches presented in this thesis will be further developed and incorporated in the future versions of the analysis using the full Run 2 dataset. Although the results presented in this thesis has provided impressive exclusion limits for a number of SUSY models, there is still immense discovery potential for the lower cross section SUSY models. The increase in instantaneous luminosity foreseen during the Run 3 of the LHC, and the subsequent High Luminosity LHC, will provide larger data samples which is crucial to probe the rarer SUSY processes. This thesis has provided a summary of the performance the p_T^{miss} reconstruction algorithms during current and future data taking conditions, in the context of pileup. The algorithm development, especially the PUPPI algorithm, has been shown in this thesis to be a promising algorithm to cope with the increase in pileup and the challenging environment of the HL LHC.

APPENDIX A

Simulated samples for SUSY searches

Several MC event generators are used to simulate the background and signal processes in this analysis, with the different parts of the generators introduced in Section 7.3. The simulation is normalized to luminosity using cross sections from <https://twiki.cern.ch/twiki/bin/viewauth/CMS/SummaryTable1G25ns>. The PYTHIA8 [110] package is used for parton showering, hadronization and underlying event simulation with the tune CUETP8M1, as described in Section 7.3. In Tables A-1, A-2, A-3 and A-4, the various MC simulations used for the different background prediction methods are listed.

Higher order corrections

For the $WZ \rightarrow l^+l^-l\nu$ and $WZ \rightarrow l^+l^-2q$ processes, a NLO to NNLO correction factor of 1.109 is applied [130]. For the $qq \rightarrow ZZ$ process, a QCD NLO to NNLO correction factor is applied as a function of generator-level p_T of the diboson system.

Table A-1. Simulated SM datasets used for the flavor symmetric (FS) background prediction. All samples are of the MINIAOD data format and of the version RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV_v6/ and Tune is short for the pythia8 tune CUETP8M1.

Process	Dataset	σ (pb)
ttbar		
$t\bar{t} \rightarrow l^+ \nu b + l^- \bar{\nu} \bar{b}$	/TTTo2L2Nu_TuneCUETP8M2_ttHtranche3_13TeV-powheg-pythia8	$831.76 \times 0.1086^2 \times 9$
$t\bar{t} \rightarrow l^- \bar{\nu} + \text{jets}$	/TTJets_SingleLeptFromTbar_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	182.2
$t\bar{t} \rightarrow l^+ \nu + \text{jets}$	/TTJets_SingleLeptFromT_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	182.2
Single Top		
$W^+ \rightarrow t\bar{b}$	/ST_s-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8	3.36
$\bar{b} \rightarrow \bar{t} W^+$	/ST_tW_antitop_5f_NoFullyHadronicDecays_13TeV-powheg	11.7
$b \rightarrow t W^-$	/ST_tW_top_5f_NoFullyHadronicDecays_13TeV-powheg	11.7
$q\bar{b} \rightarrow q'\bar{t}$	/ST_t-channel_antitop_4f_inclDecays_13TeV-powhegV2-madspin-pythia8	124.0
$qb \rightarrow q't$	/ST_t-channel_top_4f_inclDecays_13TeV-powhegV2-madspin-pythia8	208.0
Diboson (FS)		
$WW \rightarrow l^+ \nu l^- \bar{\nu}$	/WWTo2L2Nu_13TeV-powheg-pythia8	$(118.7 - 3.974) \times 0.1086^2 \times 9$
$gg \rightarrow WW \rightarrow l^+ \nu l^- \bar{\nu}$	/GluGluWWTo2L2Nu_MCFM_13TeV	$(3.974 \times 0.1086^2 \times 9 \times 1.4$
$gg \rightarrow H \rightarrow WW$	/GluGluHToWWTo2L2Nu_M125_13TeV_powheg_JHUGen_pythia8	1.002
WW	/WW_DoubleScattering_13TeV-pythia8	1.617
WW	/WpWpJJ_EWK-QCD_TuneCUETP8M1_13TeV-madgraph-pythia8	0.037
$q\bar{q} \rightarrow l\nu\gamma$	/WGToLNuG_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	405.3
Triboson (FS)		
WWW	/WWW_Tune_13TeV-amcatnlo-pythia8	0.209
$WW\gamma$	/WWG_Tune_13TeV-amcatnlo-pythia8	0.215
Rare (FS)		
$t\bar{t}W$	/TTWJetsToLNu_Tune_13TeV-amcatnlo-madspin-pythia8	0.204
$t\bar{t}W$	/TTWJetsToQQ_Tune_13TeV-amcatnloFXFX-madspin-pythia8	0.406
$t\bar{t}H$	/ttHToNonbb_M125_TuneCUETP8M2_ttHtranche3_13TeV-powheg-pythia8	0.215
VH	/VHToNonbb_M125_13TeV_amcatnloFXFX_madspin_pythia8	0.952
$ttt\bar{t}$	/TTTT_TuneCUETP8M2T4_13TeV-amcatnlo-pythia8	0.009
$W+\text{jets}$	/WJetsToLNu_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	61527

Table A-2. Simulated SM datasets used for the ZZ to 4 lepton control regions. All samples are of the MINIAOD data format and of the version `RunIIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV_v6/` and Tune is short for the pythia8 tune CUETP8M1. The k-factor referred to is specified in Subsection 17

Process	Dataset	σ (pb)
$ZZ \rightarrow 4l$		
$ZZ \rightarrow 4l$	<code>/ZZTo4L_13TeV_powheg_pythia8</code>	$1.256 \times$ k-factor
$gg \rightarrow H \rightarrow ZZ$	<code>/GluGluHToZZTo4L_M125_13TeV_powheg2_JHUgenV6_pythia8</code>	0.013
$qq \rightarrow H \rightarrow ZZ$	<code>/VBF_HToZZTo4L_M125_13TeV_powheg2_JHUgenV6_pythia8</code>	0.001
$gg \rightarrow ZZ \rightarrow 4e$	<code>/GluGluToContinToZZTo4e_13TeV_MCFM701_pythia8</code>	0.001586×2.3
$gg \rightarrow ZZ \rightarrow 4\mu$	<code>/GluGluToContinToZZTo4mu_13TeV_MCFM701_pythia8</code>	0.001586×2.3
$gg \rightarrow ZZ \rightarrow 4\tau$	<code>/GluGluToContinToZZTo4tau_13TeV_MCFM701_pythia8</code>	0.001586×2.3
$gg \rightarrow ZZ \rightarrow 2e2\tau$	<code>/GluGluToContinToZZTo2e2tau_13TeV_MCFM701_pythia8</code>	0.003194×2.3
$gg \rightarrow ZZ \rightarrow 2e2\mu$	<code>/GluGluToContinToZZTo2e2mu_13TeV_MCFM701_pythia8</code>	0.003194×2.3
$gg \rightarrow ZZ \rightarrow 2\mu2\tau$	<code>/GluGluToContinToZZTo2mu2tau_13TeV_MCFM701_pythia8</code>	0.003194×2.3
$ZZ \rightarrow 2l2\nu$		
$ZZ \rightarrow 2l2\nu$	<code>/ZZTo2L2Nu_13TeV_powheg_pythia8</code>	$0.564 \times$ k-factor
$gg \rightarrow ZZ \rightarrow 2e2\nu$	<code>/GluGluToContinToZZTo2e2nu_13TeV_MCFM701_pythia8</code>	0.001720×2.3
$gg \rightarrow ZZ \rightarrow 2\mu2\nu$	<code>/GluGluToContinToZZTo2mu2nu_13TeV_MCFM701_pythia8</code>	0.001720×2.3
Others		
$ZZ \rightarrow 2l2q$	<code>/ZZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8</code>	3.28
ZZZ	<code>/ZZZ_Tune_13TeV_amcatnlo_pythia8</code>	0.0139
VH	<code>/VHToNonbb_M125_13TeV_amcatnloFXFX_madspin_pythia8</code>	0.952
$q\bar{q} \rightarrow l^+l^-\gamma$	<code>/ZGTo2LG_TuneCUETP8M1_13TeV_amcatnloFXFX_pythia8</code>	123.9

Table A-3. Simulated SM datasets used for the WZ control regions. All samples are of the MINIAOD data format and of the version `RunIIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV_v6/` and Tune is short for the pythia8 tune CUETP8M1.

Process	Dataset	σ (pb)
$WZ \rightarrow l^+l^-\nu\nu$	<code>/WZTo3LNu_TuneCUETP8M1_13TeV-powheg_pythia8</code>	4.429×1.109
$WZ \rightarrow l^+l^-2q$	<code>/WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8</code>	5.595×1.109

Table A-4. Various non flavor symmetric processes. All samples are of the MINIAOD data format and of the version

`RunIIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV_v6/` and `Tune` is short for the `pythia8` tune `CUETP8M1`.

Process	Dataset	σ (pb)
Drell-Yan		
$Z/\gamma^* \rightarrow l^+l^- (50)$	<code>/DYJetsToLL_M-50_Tune_13TeV-madgraphMLM-pythia8</code>	1921.8×3
$Z/\gamma^* \rightarrow l^+l^- (10 - 50)$	<code>/DYJetsToLL_M-10to50_Tune_13TeV-madgraphMLM-pythia8</code>	18610
Various non FS		
$t\bar{t}Z$	<code>/TTZToLL_M-1to10_TuneCUETP8M1_13TeV-madgraphMLM-pythia8</code>	0.049
$t\bar{t}Z$	<code>/TTZToLLNuNu_M-10_Tune_13TeV-amcatnlo-pythia8</code>	0.253
$t\bar{t}Z$	<code>/TTZToQQ_Tune_13TeV-amcatnlo-pythia8</code>	0.530
tZq	<code>/tZq_ll_4f_13TeV-amcatnlo-pythia8</code>	0.076
tWZ	<code>/ST_tWll_5f_L0_13TeV-MadGraph-pythia8</code>	0.011
WWZ	<code>/WWZ_Tune_13TeV-amcatnlo-pythia8</code>	0.165
WZZ	<code>/WZZ_Tune_13TeV-amcatnlo-pythia8</code>	0.056

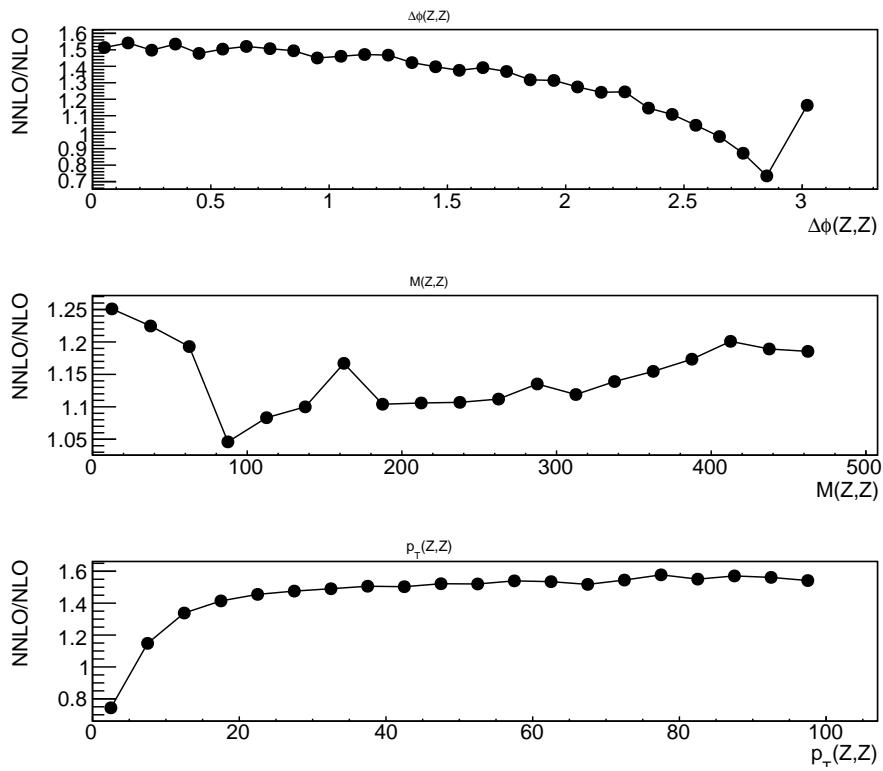


Figure A-1. QCD NNLO/NLO k factors for the $qq \rightarrow ZZ$ process in generator level variables of the diboson system.

APPENDIX B

Simulated samples for p_T^{miss} performance study

Tables A-5 and A-6 contain all the simulated samples used for the p_T^{miss} performance studies.

Table A-5. Simulated SM datasets used for the p_T^{miss} performance study with final states containing two leptons.

Process	Dataset	σ (pb)
DY		
Drell-Yan	/DYJetsToLL_M-50_Tune_13TeV-amcatnloFXFX	6025.2
Top		
$t\bar{t} \rightarrow l^+ \nu b + l^- \bar{\nu} \bar{b}$	/TTTo2L2Nu_TuneCUETP8M2_ttHtranche3_13TeV-powheg-pythia8	$831.76 \times 0.1086^2 \times 9$
$W^+ \rightarrow t\bar{b}$	/ST_s-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8	3.36
$\bar{b} \rightarrow \bar{t}W^+$	/ST_t-channel_antitop_4f_inclusiveDecays_13TeV-powheg	80.95
$qb \rightarrow q't$	/ST_t-channel_top_4f_inclusiveDecays_13TeV-powheg	136.02
$b \rightarrow tW^-$	/ST_tW_top_5f_NoFullyHadronicDecays_13TeV-powheg	11.7
EWK		
$ZZ \rightarrow 2l2q$	/ZZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	3.28
$ZZ \rightarrow 2l2\nu$	/ZZTo2L2Nu_13TeV_powheg_pythia8	0.564
$ZZ \rightarrow 4l$	/ZZTo4L_13TeV_powheg_pythia8	1.256
$WZ \rightarrow l^+ l^- l\nu$	/WZTo3LNu_TuneCUETP8M1_13TeV-powheg-pythia8	4.429×1.109
$WZ \rightarrow l^+ l^- 2q$	/WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	5.595×1.109
$WW \rightarrow l^+ \nu l^- \bar{\nu}$	/WWTo2L2Nu_13TeV-powheg_pythia8	$(118.7 - 3.974) \times 0.1086^2 \times 9$
WW	/WWW_Tune_13TeV-amcatnlo-pythia8	0.209
WWZ	/WWZ_Tune_13TeV-amcatnlo-pythia8	0.165
WZZ	/WZZ_Tune_13TeV-amcatnlo-pythia8	0.055
ZZZ	/ZZZ_Tune_13TeV-amcatnlo-pythia8	0.014

Table A-6. Simulated SM datasets used for the p_T^{miss} performance study with final states containing one photon.

Process	Dataset	σ (pb)
$\gamma + \text{jets}$		
$\gamma + \text{jets}$	/GJets_HT-40To100_Tune_13TeV_madgraphMLM	20730
$\gamma + \text{jets}$	/GJets_HT-100To200_Tune_13TeV_madgraphMLM	9226
$\gamma + \text{jets}$	/GJets_HT-200To400_Tune_13TeV_madgraphMLM	2300
$\gamma + \text{jets}$	/GJets_HT-400To600_Tune_13TeV_madgraphMLM	277
$\gamma + \text{jets}$	/GJets_HT-600ToInf_Tune_13TeV_madgraphMLM	93.4
QCD		
QCD	/QCD_HT-100to200_Tune_13TeV_madgraphMLM	2.799×10^7
QCD	/QCD_HT-200to300_Tune_13TeV_madgraphMLM	17.35×10^5
QCD	/QCD_HT-300to500_Tune_13TeV_madgraphMLM	3.668×10^5
QCD	/QCD_HT-500to700_Tune_13TeV_madgraphMLM	2.937×10^4
QCD	/QCD_HT-700to1000_Tune_13TeV_madgraphMLM	6524
QCD	/QCD_HT-1000to1500_Tune_13TeV_madgraphMLM	1064
QCD	/QCD_HT-1500to2000_Tune_13TeV_madgraphMLM	119.9
QCD	/QCD_HT-2000toInf_Tune_13TeV_madgraphMLM	25.2
$V\gamma + \text{top quark}$		
$W + \text{jets}$	/WJetsToLNu_HT-100to200_Tune_13TeV_madgraphMLM	1345×1.21
$W + \text{jets}$	/WJetsToLNu_HT-200to400_Tune_13TeV_madgraphMLM	359.7×1.21
$W + \text{jets}$	/WJetsToLNu_HT-400to600_Tune_13TeV_madgraphMLM	48.9×1.21
$W + \text{jets}$	/WJetsToLNu_HT-600to800_Tune_13TeV_madgraphMLM	12.1×1.21
$W + \text{jets}$	/WJetsToLNu_HT-800to1200_Tune_13TeV_madgraphMLM	5.50×1.21
$W + \text{jets}$	/WJetsToLNu_HT-1200to2500_Tune_13TeV_madgraphMLM	1.33×1.21
$W + \text{jets}$	/WJetsToLNu_HT-2500toInf_Tune_13TeV_madgraphMLM	0.03×1.21
$W\gamma$	/WGToLNuG_13TeV-amcatnloFXFX	489
$Z\gamma$	/ZGTo2LG_13TeV-amcatnloFXFX	117.9
$Z\gamma$	/ZGTo2NuG_13TeV-amcatnloFXFX	28.1
$\text{top} + \gamma$	/TTGJets_Tune_13TeV_amcatnloFXFX_madspin	4621
$\text{top} + \gamma$	/TGJets_Tune_13TeV_amcatnlo_madspin	2.97

APPENDIX C

Fit functions for kinematic edge search

In addition to a cut and count approach, a fit is performed to the mass spectrum to search for a kinematic edge. More information on the fit can be found in CMS analysis note AN-16-482. The fit is carried out simultaneously to the $m_{\ell\ell}$ distribution in the dielectron, dimuon, and opposite flavor final state. The fit consists of three separate components: the FS backgrounds, the Drell–Yan backgrounds and the signal, fitted to their corresponding shapes.

Model for FS backgrounds

The FS background is described using a Crystal-Ball function: $\mathcal{P}_{FSCB}(m_{\ell\ell})$:

$$\mathcal{P}_{FSCB}(m_{\ell\ell}) = \begin{cases} \exp\left(-\frac{(m_{ll}-\mu_{FSCB})^2}{2\sigma_{FSCB}^2}\right) & \text{if } \frac{m_{ll}-\mu_{FSCB}}{\sigma_{FSCB}} < \alpha_{FS}, \\ A(B + \frac{m_{ll}-\mu_{FSCB}}{\sigma_{FSCB}})^{-n_{FS}} & \text{if } \frac{m_{ll}-\mu_{FSCB}}{\sigma_{FSCB}} > \alpha_{FS}, \end{cases} \quad (\text{A-1})$$

where

$$A = \left(\frac{n}{|\alpha_{FS}|}\right)^{n_{FS}} \exp\left(-\frac{|\alpha_{FS}|^2}{2}\right) \quad \text{and} \quad B = \frac{n_{FS}}{|\alpha_{FS}|} - |\alpha_{FS}|. \quad (\text{A-2})$$

Because of the requirement that the function and its derivative both be continuous, the FS model is left with four independent parameters plus the normalization. While the three samples (dielectron, dimuon and opposite flavor) have different normalizations, they all use the same FS shape parameters:

$$\vec{p}_{FS} = (\mu_{FSCB}, \sigma_{FSCB}, \alpha_{FS}, n_{FS})$$

Model for Drell-Yan backgrounds

The shape for backgrounds containing a Z is based on the sum of an exponential function and a convolution of a Breit-Wigner $\mathcal{P}_{BW}(m_{ll}; m_Z, \sigma_Z)$ (with physical mean m_Z and width σ_Z) and a double-sided Crystal-Ball function $\mathcal{P}_{DSCB}(m_{\ell\ell})$:

$$\mathcal{P}_{DSCB}(m_{\ell\ell}) = \begin{cases} A_1(B_1 - \frac{m_{ll}-\mu_{CB}}{\sigma_{CB}})^{-n_1} & \text{if } \frac{m_{ll}-\mu_{CB}}{\sigma_{CB}} < -\alpha_1 \\ \exp\left(-\frac{(m_{ll}-\mu_{CB})^2}{2\sigma_{CB}^2}\right) & \text{if } -\alpha_1 < \frac{m_{ll}-\mu_{CB}}{\sigma_{CB}} < \alpha_2 \\ A_2(B_2 + \frac{m_{ll}-\mu_{CB}}{\sigma_{CB}})^{-n_2} & \text{if } \frac{m_{ll}-\mu_{CB}}{\sigma_{CB}} > \alpha_2 \end{cases}$$

where

$$A_i = \left(\frac{n_i}{|\alpha_i|} \right)^{n_i} \cdot \exp \left(-\frac{|\alpha_i|^2}{2} \right) \quad \text{and} \quad B_i = \frac{n_i}{|\alpha_i|} - |\alpha_i|.$$

The full description is therefore

$$\mathcal{P}_{DY}(m_{\ell\ell}) = f_{exp} \mathcal{P}_{exp}(m_{\ell\ell}) + (1 - f_{exp}) \int \mathcal{P}_{DSCB}(m_{\ell\ell}) \mathcal{P}_{BW}(m_{\ell\ell} - m') dm'.$$

The model is fit separately for electrons and muons in a control region enriched in backgrounds containing Z bosons in order to extract the shape, which is then used with all parameters fixed (except the normalization) in the signal region. The CR used to extract the shape is the same as the one used to extract $R_{out/in}$, i.e. contains events with two or more jets, and an upper p_T^{miss} cut at 50 GeV to improve DY purity. The electron and muon models use different values for the double-sided Crystal Ball function, and the relative normalization (f_{exp}) between the exponential and the convoluted function. They do, however, use the same value for m_Z^{pdg} and σ_Z^{pdg} (fixed to PDG values). The full parameter set in the fit for electrons is

$$\vec{p}_Z^e = (\mu_{CB}^{ee}, \sigma_{CB}^{ee}, \alpha_1^{ee}, \alpha_2^{ee}, n_1^{ee}, n_2^{ee}, f_{exp}^{ee}, \mu_{exp}^{ee})$$

and for muons

$$\vec{p}_Z^\mu = (\mu_{CB}^{\mu\mu}, \sigma_{CB}^{\mu\mu}, \alpha_1^{\mu\mu}, \alpha_2^{\mu\mu}, n_1^{\mu\mu}, n_2^{\mu\mu}, f_{exp}^{\mu\mu}, \mu_{exp}^{\mu\mu})$$

Signal Model

The signal component is based on an edge model for two subsequent two-body decays with endpoint $m_{\ell\ell}^{edge}$, including a model for the dilepton mass resolution, $\sigma_{\ell\ell}$:

$$\mathcal{P}_S(m_{\ell\ell}) = \frac{1}{\sqrt{2\pi}\sigma_{\ell\ell}} \int_0^{m_{\ell\ell}^{edge}} y \cdot \exp \left(-\frac{(m_{\ell\ell} - y)^2}{2\sigma_{\ell\ell}^2} \right) dy$$

The function describes a triangle convoluted with a Gaussian to account for the mass resolution. The electron and muon models share the edge position $m_{\ell\ell}^{edge}$, but use their own value for σ (which comes from the model for backgrounds containing a Z). The parameters are therefore

$$\vec{p}_S^e = (m_{ll}^{edge}, \sigma_{CB}^{ee})$$

for electrons and

$$\vec{p}_S^\mu = (m_{ll}^{edge}, \sigma_{CB}^{\mu\mu})$$

for muons.

LIST OF TABLES

Table	<u>page</u>
2-1 An overview of the fermions of the SM.	15
2-2 An overview of the gauge bosons of the SM.	15
3-1 An overview of the superpartners of the SM bosons and fermions [21].	29
3-2 An overview of the gauge eigenstates and the mass eigenstates of the SUSY particles. [21].	30
5-1 Beam parameters for beams in the LHC at injection and collision energy [67].	45
8-1 Electron selection criteria.	71
8-2 Muon selection criteria.	74
8-3 Definition of Loose Jet ID working point.	76
9-1 Datasets used for the p_T^{miss} performance study in $Z \rightarrow ll$ events.	85
9-2 Triggers used for the p_T^{miss} performance study in $Z \rightarrow ll$ events.	86
9-3 Datasets used for the p_T^{miss} performance study in single γ events.	87
9-4 Triggers used for the p_T^{miss} performance study in single γ events.	87
9-5 Parametrization results of the resolution curves for the u_{\parallel} and u_{\perp} components as a function of N_{vtx} . The parameter values for σ_c are obtained from data and simulation, and the values for σ_{PU} are obtained from data, along with a ratio R_{PU} of data and simulation.	95
9-6 Parametrization results of the resolution curves for u_{\parallel} and u_{\perp} components as a function of the scalar p_T sum of all PF candidates. The parameter values for σ_0 are obtained from data and simulation, whereas the σ_s are obtained from data along with the ratio R_s , the ratio of data and simulation.	96
9-7 Parameterization results of the resolution curves for PUPPI u_{\parallel} and u_{\perp} components as a function of N_{vtx} . The parameter values for σ_c are obtained from data and simulation, and the values for σ_{PU} are obtained from data, along with the ratio R_{PU} of data and simulation.	98
10-1 Datasets used in the strong, electroweak and slepton searches and p_T^{miss} study	104
10-2 Triggers used in the strong, electroweak and slepton searches. The first section are the triggers used in the signal regions, while the supporting triggers are used for the calculation of the trigger efficiencies of the signal triggers and for control regions.	105

10-3 Lepton kinematic criteria.	107
10-4 Jet kinematic criteria.	108
11-1 Observed event yields in the control region and the resulting values for $R_{SF/OF}$, $R_{ee/OF}$, and $R_{\mu\mu/OF}$ for both data and MC. The transfer factor is defined as the ratio of $R_{SF/OF}$ in the signal region divided by $R_{SF/OF}$ in the control region.	117
11-2 Result of the fit of $r_{\mu/e}$ as a function of the p_T of the trailing lepton in the DY control region. The same quantities derived from simulation are shown for comparison. Only statistical uncertainties are given.	119
11-3 Trigger efficiency values for data and MC with OS, $p_T > 25(20)$ GeV and $H_T > 200$ GeV.	120
12-1 Summary of SUSY models and the sparticle masses that are constrained.	126
14-1 Summary of the colored SUSY On-Z SR.	139
14-2 Summary of the Edge SR.	140
14-3 Resulting estimates for flavor-symmetric backgrounds in the Edge SR. Given is the observed event yield in OF events (N_{OF}), the estimate in the SF channel using the event-by-event reweighting of the factorization method ($N_{SF}^{factorization}$), $R_{SF/OF}$ for the factorization method ($R_{SF/OF}^{factorization}$), $R_{SF/OF}$ obtained from the direct measurement ($R_{SF/OF}^{direct}$), $R_{SF/OF}$ when combining this results from direct measurement and factorization methods ($R_{SF/OF}^{combined}$), and the combined final prediction (N_{SF}^{final}). Statistical and systematic uncertainties are given separately.	141
14-4 Measured values for $R_{out/in}$ for data and MC in the different signal regions of the edge search.	143
14-5 Summary of template predictions with systematic uncertainties added in quadrature in the strong On-Z signal regions together with the individual systematic uncertainties from each source.	143
14-6 Transfer factor derived in the three lepton control region for the On-Z search.	144
14-7 Transfer factor derived in the four lepton control region for On-Z searches.	146
14-8 Transfer factor derived in the $t\bar{t}Z$ control region	147
14-9 Systematic uncertainties taken into account for the signal yields and their typical values.	147
14-10 Standard model background predictions compared to observed yield in the On-Z signal regions (SRA-SRB) [59].	148
14-11 Standard model background predictions compared to observed yield in the strong On-Z signal regions (SRC) [59].	150
14-12 Predicted and observed yields in each bin of the edge search counting experiment. The uncertainties shown include both statistical and systematic sources [59].	150
14-13 Results of the unbinned maximum likelihood fit for event yields in the edge fit SR, including the Drell–Yan and FS background components, along with the fitted signal contribution and edge position. The fitted value for $R_{SF/OF}$ and the local and global signal significances in terms of standard deviations are also given. The uncertainties account for both statistical and systematic components.	152

15-1	Summary of the electroweak VZ SR.	158
15-2	Summary of the Electroweak ZH SR.	158
15-3	Summary of template predictions with systematic uncertainties added in quadrature in the strong and electroweak on-Z signal regions together with the individual systematic uncertainties from each source.	159
15-4	Systematic uncertainties taken into account for the signal yields and their typical values.	160
15-5	Predicted and observed event yields are shown for the electroweak VZ SR, for each p_T^{miss} bin defined in Table 15-1. The uncertainties shown include both statistical and systematic sources.	160
15-6	Predicted and observed event yields are shown for the electroweak on-Z SR (ZH), for each p_T^{miss} bin defined in Table 16-1. The uncertainties shown include both statistical and systematic sources.	160
16-1	Summary of the direct slepton production SR.	165
16-2	Resulting estimates for flavor-symmetric backgrounds in the Slepton search.	168
16-3	Transfer factor derived in the three lepton control region for slepton search.	169
16-4	Transfer factor derived in the four lepton control region for slepton search.	170
16-5	Measured values for $R_{\text{out/in}}$ for data and MC in the different signal regions of the slepton search.	171
16-6	List of systematic uncertainties taken into account for the ZZ background prediction method.	172
16-7	List of systematic uncertainties taken into account for the WZ background prediction method.	172
16-8	List of systematic uncertainties taken into account for the signal yields.	173
16-9	The predicted SM background contributions, their sum and the observed number of SF events in data. The uncertainties associated with the background yields stem from statistical and systematic sources.	173
16-10	The predicted SM background contributions, their sum and the observed number of dielectron (upper) and dimuon (lower) events in data. The uncertainties associated with the yields stem from statistical and systematic source.	174
A-1	Simulated SM datasets used for the flavor symmetric (FS) background prediction. All samples are of the MINIAOD data format and of the version RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_Trane and Tune is short for the pythia8 tune CUETP8M1.	183
A-2	Simulated SM datasets used for the ZZ to 4 lepton control regions. All samples are of the MINIAOD data format and of the version RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV_v6/ and Tune is short for the pythia8 tune CUETP8M1. The k-factor referred to is specified in Subsection 17184	
A-3	Simulated SM datasets used for the WZ control regions. All samples are of the MINIAOD data format and of the version RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV_v6/ and Tune is short for the pythia8 tune CUETP8M1.	184
A-4	Various non flavor symmetric processes. All samples are of the MINIAOD data format and of the version RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TracheIV_v6/ and Tune is short for the pythia8 tune CUETP8M1.	185

A-5 Simulated SM datasets used for the p_T^{miss} performance study with final states containing two leptons. 186

A-6 Simulated SM datasets used for the p_T^{miss} performance study with final states containing one photon. 187

LIST OF FIGURES

<u>Figure</u>	<u>page</u>
2-1 A prototypical effective ‘Sombrero’ potential that leads to ‘spontaneous’ symmetry breaking. The vacuum, i.e., the lowest-energy state, is described by a randomly-chosen point around the bottom of the brim of the hat [12].	22
3-1 One-loop quantum corrections to the Higgs squared mass parameter m_H^2 , due to a Dirac fermion f (left), and a scalar S (right).	27
3-2 Two-loop renormalization group evolution of the inverse gauge couplings α^{-1} in the SM (dashed lines) and the MSSM (solid lines). In the MSSM case, the sparticle masses are treated as a common threshold varied between 750 GeV and 2.5 TeV, and $\alpha_3(m_Z)$ is varied between 0.117 and 0.120. [21]	28
4-1 Theory cross sections for selected SUSY processes [47].	35
4-2 Diagrams for strong SUSY production. The gluino GMSB model targeted by the strong on-Z search is shown on the left, that contains one dilepton pair stemming from an on-shell Z boson decay. On the right is a diagram showing a model in which bottom squarks are pair produced with subsequent decays that contain at least one dilepton pair. This model features a characteristic edge shape in the $m_{\ell\ell}$ spectrum given approximately by the mass difference of the $\tilde{\chi}_2^0$ and $\tilde{\chi}_1^0$	38
4-3 Diagram corresponding to the chargino-neutralino production with the $\tilde{\chi}_1^\pm$ and $\tilde{\chi}_2^0$ decaying into vector bosons (W and Z) and the LSP.	39
4-4 Diagrams corresponding to the neutralino-neutralino model of where the neutralinos are allowed to decay to a gravitino and a Z boson (left) and where the neutralinos are allowed to decay to a gravitino and a Z boson or a Higgs boson, with a 50% branching fraction to each decay channel (right).	40
4-5 Diagram corresponding to the slepton model with two selectrons (smuons) directly produced and decay into electrons (muons) and a LSP	41
4-6 Graphs showing the cross sections for slepton pair production at a center-of-mass energy of 8 TeV and 13 TeV.	41
5-1 Sketch of the various accelerators and experiments hosted at CERN [73].	44
5-2 Illustration of the CMS coordinate system [75].	46
6-1 General view of the CMS detector. The major detector components are indicated, together with the acronyms for the various CMS construction modules [77].	49

6-2	Cumulative measured luminosity versus day delivered. This measurement uses the best available offline measurement and calibrations [78].	50
6-3	Schematic cross section through the CMS tracker in the r - z plane. Each line-element represents a detector module. Closely spaced double line-elements indicate back-to-back silicon strip modules, in which one module is rotated through a ‘stereo’ angle, so as to permit reconstruction of the hit positions in 3-D. Within a given layer, each module is shifted slightly in r or z with respect to its neighboring modules, which allows them to overlap, thereby avoiding gaps in the acceptance [80].	51
6-4	Layout of the CMS ECAL, showing the barrel supermodules, the two endcaps and the preshower detectors. The ECAL barrel coverage is up to $ \eta = 1.48$, the endcaps extend the coverage to $ \eta = 3.0$ and the preshower detector fiducial area is approximately $1.65 \leq \eta \leq 2.6$ [83].	52
6-5	The HCAL tower segmentation for one-fourth of the HB, HO, and HE detectors. The numbers on top and on the left refer to the tower numbers. The numbers on the right and on the bottom (0-16) indicate the scintillator layers numbers inserted into slots in the absorber. The shading represents independent longitudinal readouts in the HB/HE overlap and the small angle regions [84].	53
6-6	Cross section of a quadrant of the CMS detector with the axis parallel to the beam (z) running horizontally and the radius (R) increasing upward. The interaction point is at the lower left corner. The locations of the various muon stations and the steel flux-return disks (dark areas) are shown. The DTs are labeled MB (“Muon Barrel”) and the CSCs are labeled ME (“Muon Endcap”). RPCs are mounted in both the barrel and endcaps of CMS, where they are labeled RB and RE, respectively [86].	54
7-1	Chart showing the organization of the various components of any L1-accept. The calorimetry and muons systems work in parallel and are combined into a global trigger [87].	57
7-2	The NNPDF3.1 NNLO PDFs, evaluated at two resolution scales; $\mu^2 = 10 \text{ GeV}^2$ (left) and $\mu^2 = 10^4 \text{ GeV}^2$ (right) [89].	61
7-3	Mean number of interactions per bunch crossing for the 2016 pp run at 13 TeV. The cross section is taken to be 80 mb [78].	64
8-1	The CSVv2 discriminator for jets in a muon-enriched multijet sample. The simulated contributions of each jet flavor are shown with different colors [106].	79
9-1	Z boson q_T in $Z \rightarrow \mu^+\mu^-$ (left) and $Z \rightarrow e^+e^-$ (right) samples. The Diboson contribution corresponds to processes with two electroweak bosons produced in the final state. The Top quark contribution corresponds to the top anti-top pair and single top production processes. The band in the data to simulation ratio corresponds to the statistical uncertainty in simulated samples [61].	86
9-2	Distribution of the photon q_T in the single-photon sample. The $V\gamma + \text{Top}$ quark contribution corresponds to the $Z\gamma$, $W\gamma$, top anti-top pair and single top production processes. The band in the data to simulation ratio corresponds to the statistical uncertainty in the simulated samples [61].	88
9-3	Event display for a beam halo event with collinear hits in the CSC (black), p_T^{miss} of $\sim 241 \text{ GeV}$, and a jet with $p_T = \sim 232 \text{ GeV}$ [61].	89
9-4	The p_T^{miss} (left) and jet ϕ (right) distributions for events passing the dijet (left) and monojet (right) selection with the event filtering algorithms applied, including that based on jet identification requirements (filled markers), without the event filtering algorithms applied (open markers), and from simulation (solid histograms) [61].	90

9-5 Illustration of the Z boson (left) and photon (right) event kinematics in the transverse plane. The vector \vec{u}_T denotes the vectorial sum of all particles reconstructed in the event except for the two leptons from the Z decay (left) or the photon (right) [61].	91
9-6 The p_T^{miss} for events passing the dimuon (left), dielectron (right) and single photon (bottom) selections, in data (black markers) and simulation (solid histograms). The lower bands show the data to simulation ratio with the systematic uncertainties due to the JES, the JER, and variations in the UE are added in quadrature [61].	92
9-7 Distribution of $u_{\parallel}+q_T$ (left) and u_{\perp} (right) components of the hadronic recoil, in data (filled markers) and simulation (solid histograms), in the $Z \rightarrow \mu^+\mu^-$ (top), $Z \rightarrow e^+e^-$ (middle), and $\gamma+\text{jets}$ (bottom) samples [61].	93
9-8 Upper panels: Response of the Raw (left) and Type 1 corrected (right) p_T^{miss} in data in $Z \rightarrow \mu^+\mu^-$, $Z \rightarrow e^+e^-$, and $\gamma+\text{jets}$ events. Lower panels: Ratio of the p_T^{miss} response in data and simulation. The band corresponds to the systematic uncertainties due to the JES, the JER, and variations in the UE added in quadrature, estimated from the $Z \rightarrow e^+e^-$ sample [61].	94
9-9 Resolution of the u_{\parallel} and u_{\perp} components of the hadronic recoil as a function of q_T in $Z \rightarrow \mu^+\mu^-$, $Z \rightarrow e^+e^-$, and $\gamma+\text{jets}$ events [61].	95
9-10 Resolution of the u_{\parallel} and u_{\perp} components of the hadronic recoil as a function of the reconstructed vertices in $Z \rightarrow \mu^+\mu^-$, $Z \rightarrow e^+e^-$, and $\gamma+\text{jets}$ events [61].	96
9-11 Resolution of the u_{\parallel} and u_{\perp} components of the hadronic recoil as a function of the scalar p_T sum of all PF candidates (lower row), in $Z \rightarrow \mu^+\mu^-$, $Z \rightarrow e^+e^-$ and $\gamma+\text{jets}$ events [61].	97
9-12 Upper panels: Distributions of PUPPI p_T^{miss} in $Z \rightarrow \mu^+\mu^-$ (left) and $Z \rightarrow e^+e^-$ (right) events [61].	98
9-13 Upper panels: Distributions of the $u_{\parallel}+q_T$ and u_{\perp} components of the hadronic recoil, in data (filled markers) and simulation (solid histograms), for the $Z \rightarrow \mu^+\mu^-$ (upper) and $Z \rightarrow e^+e^-$ (lower) events [61].	99
9-14 Response of PUPPI p_T^{miss} , defined as $-\langle u_{\parallel} \rangle / \langle q_T \rangle$, in data in $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ events.	100
9-15 PUPPI and PF p_T^{miss} resolution of u_{\parallel} (left) and u_{\perp} (right) components of the hadronic recoil as a function of N_{vtx} , in $Z \rightarrow \mu^+\mu^-$ events [61].	100
9-16 Upper panels: PUPPI and PF p_T^{miss} resolution of u_{\parallel} (left) and u_{\perp} (right) components of the hadronic recoil as a function of N_{vtx} , in $Z \rightarrow \mu^+\mu^-$ events. The blue (green) markers correspond to the PF (PUPPI) p_T^{miss} reconstruction algorithm, with filled (open) markers for the nominal run (high pileup run).	101
10-1 The M_{T2} distribution after requiring a minimum of 2 jets and $p_T^{\text{miss}} > 100$ GeV in addition to an opposite sign, same flavor lepton pair. The cut value of $M_{T2} > 80$ GeV that is used in most parts of the analysis is indicated by a grey dashed line [113].	109
10-2 PDFs for the four input variables to the likelihood discriminant: p_T^{miss} (top left), dilepton p_T (top right), $ \Delta\phi $ between the leptons (bottom left), and $m_{\ell b}$ (bottom right) [113].	111
11-1 The leading order diagrams of $t\bar{t}$ production, through gluon fusion (left and middle) and quark-antiquark annihilation (right).	113

11-2 Single top quark production through s-channel (left), t-channel (middle) and in association with a W boson (right).	113
11-3 Leading order DY production.	114
11-4 Leading order diboson production through <i>s</i> -channel (left), <i>t</i> -channel (middle) or <i>u</i> -channel (right).	114
11-5 Leading order $t\bar{t}Z$ (left), $t\bar{t}W$ (middle) and $t\bar{t}H$ (right) production.	115
11-6 Leading order tWZ (left) and tZq (right) production.	115
11-7 The dilepton invariant mass distribution after requiring a baseline SR definition for the Edge search in addition to an opposite sign, SF (left) or opposite sign OF (right) lepton pair [59].	115
11-8 The $r_{\mu/e}$ dependency on the p_T of the second lepton for data and MC. The plots show the values for $r_{\mu/e}$ before (left) and after (right) the parameterization on the second lepton p_T is propagated to the dielectron events. The central value in the left plot indicates the $r_{\mu/e}$ value that would be obtained without the parameterization. The fit values of the parameterization are shown in the left plot.	118
11-9 The $0.5(r_{\mu/e}^{corr.} + 1/r_{\mu/e}^{corr.})$ dependency on p_T^{miss} (left) and M_{T2} (right) for data and MC after the parameterization on the second lepton p_T is propagated to the dielectron events. The uncertainty introduced by shifting the fit parameters by its statistical uncertainty and the full prediction by 10% is indicated by the orange band.	119
11-10 Dependency of the R_T ratio on the p_T^{miss} (left) and M_{T2} (right), for data and MC. The systematic uncertainty of about 4% on R_T is indicated by the orange band.	121
12-1 The standard normal distribution $\phi(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)$ showing the relation between the significance Z and the p -value [121].	129
14-1 The f_{mll} evaluated in each On-Z signal region.	142
14-2 The p_T^{miss} (left) and the M_{T2} in the three lepton control region in data and MC.	145
14-3 The invariant mass of the two lepton pair forming the best Z candidate in the four lepton control region in data and simulation.	145
14-4 The invariant mass of the two lepton pair forming the best Z candidate in the $t\bar{t}Z$ control region in data and simulation.	146
14-5 The p_T^{miss} distribution is shown for data compared to the background prediction in the on-Z SRs with no b-tagged jets (left) and at least 1 b-tagged jet (right). The rows show SRA (upper), SRB (middle), and SRC (lower) [59].	149
14-6 $m_{\ell\ell}$ distribution in the $t\bar{t}$ -like (left) and non- $t\bar{t}$ -like edge signal regions. The number of observed events, shown as black data points, is compared to the total background estimate, shown as a blue line with a blue uncertainty band. The non flavor symmetric background component from instrumental p_T^{miss} is indicated as a green area while the non flavor symmetric background with neutrinos is shown as a violet area [59].	151

14-7 Results of the counting experiment of the edge search. For each SR, the number of observed events, shown as black data points, is compared to the total background estimate. The hashed band shows the total uncertainty in the background prediction, including statistical and systematic sources [59].	151
14-8 Fit of the dilepton invariant mass distributions to the signal-plus-background hypothesis in the “Edge fit” SR from Table 14-2, projected on the same flavor (left) and opposite flavor (right) event samples. The fit shape is shown as a solid blue line. The individual fit components are indicated by dashed and dotted lines. The FS background is shown with a black dashed line. The Drell-Yan background is displayed with a red dotted line. The extracted signal component is displayed with a purple dash-dotted line. The lower panel in each plot shows the difference between the observation and the fit, divided by the square root of the number of fitted events [59].	153
14-9 Cross section upper limit and exclusion contours at 95% CL for the gluino GMSB model as a function of the \tilde{g} and $\tilde{\chi}_1^0$ masses, obtained from the results of the strong production On-Z search. The region to the left of the thick red dotted (black solid) line is excluded by the expected (observed) limit. The thin red dotted curves indicate the regions containing 68 and 95% of the distribution of limits expected under the background-only hypothesis. The thin solid black curves show the change in the observed limit due to variation of the signal cross sections within their theoretical uncertainties [59].	154
14-10 Cross section upper limit and exclusion contours at 95% CL for the slepton edge model as a function of the \tilde{b} and $\tilde{\chi}_2^0$ masses, obtained from the results of the Edge search. The region to the left of the thick red dotted (black solid) line is excluded by the expected (observed) limit. The thin red dotted curves indicate the regions containing 68 and 95% of the distribution of limits expected under the background-only hypothesis. The thin solid black curves show the change in the observed limit due to variation of the signal cross sections within their theoretical uncertainties [59].	155
15-1 The p_T^{miss} distribution is shown for data compared to the background prediction in the on-Z VZ (left) and ZH (right) SRs. The p_T^{miss} template prediction for each SR is normalized to the first bin of each distribution, and therefore the prediction agrees with the data by construction [59].	161
15-2 Cross section upper limit and exclusion contours at 95% CL for the model of $\tilde{\chi}_1^\pm \tilde{\chi}_2^0$ production, where the exclusion is expressed as a function of the $\tilde{\chi}_1^\pm$ (equal to $\tilde{\chi}_2^0$) and $\tilde{\chi}_1^0$ masses, obtained using the on-Z VZ signal region. The region under the thick red dotted (black solid) line is excluded by the expected (observed) limit. The thin red dotted curves indicate the regions containing 68 and 95% of the distribution of limits expected under the background-only hypothesis. The thin solid black curves show the change in the observed limit due to variation of the signal cross sections within their theoretical uncertainties [59].	162
15-3 Cross section upper limit and exclusion lines at 95% CL, as a function of the $\tilde{\chi}_1^0$ mass, for the search for electroweak production in the ZZ topology (left) and with a 50% branching fraction to each of the Z and Higgs bosons (right). The red band shows the theoretical cross section, with the thickness of band representing the theoretical uncertainty in the signal cross section. Regions where the black dotted line reaches below the theoretical cross section are expected to be excluded. The green (yellow) band indicates the region containing 68 (95)% of the distribution of limits expected under the background-only hypothesis. The observed upper limit on the cross section is shown with a solid black line [59].	162

16-1 The $r_{\mu/e}$ (left) and $0.5(r_{\mu/e}^{corr.} + 1/r_{\mu/e}^{corr.})$ (right) as a function of the number of jets for data and MC. No significant trend is observed in the lower jet multiplicities.	166
16-2 $R_{SF/OF}$ as a function of number of jets derived in a $t\bar{t}$ enriched region after relaxing the 2 jet requirement. The MC simulation takes into account all FS processes. In order to not unbind the signal region, events with are excluded.	166
16-3 $R_{SF/OF}$ as a function of number of jets derived in a $t\bar{t}$ enriched control region after relaxing the 2 jet requirement and requiring that in a 0 jet event, a requirement of the M_{T2} to be within $40 - 90$ GeV is imposed. The MC simulation takes into account all flavor symmetric and non-flavor symmetric processes such as DY.	167
16-4 The p_T^{miss} (left) and the M_{T2} in the three lepton control region in data and MC [60].	169
16-5 Distributions of the generator p_T , mass and $\Delta\phi$ of the diboson system, in the signal region, with and without the respective k-factor applied, in ZZTo2L2Nu MC.	170
16-6 The invariant mass of the two lepton pair forming the best Z candidate in the four lepton control region (left). The emulated M_{T2} is the M_{T2} formed with the p_T^{miss} that has the p_T of the leptons forming the best Z candidate added to it, and the remaining two leptons (right).	171
16-7 The p_T^{miss} (left) distribution for the resulting SM background yields estimated in the slepton SF analysis SR overlayed with the observed data as black points. The M_{T2} distribution (right) in the same signal region with three signal scenarios overlayed [60].	174
16-8 Cross section upper limit and exclusion contours at 95% CL for direct slepton production of two flavors, selectrons and smuons, as a function of the $\tilde{\chi}_1^0$ and $\tilde{\ell}$ masses, assuming the production of both left- and right-handed sleptons (upper) or production of only left- (lower left) or right-handed (lower right). The region under the thick red dotted (black solid) line is excluded by the expected (observed) limit. The thin red dotted curves indicate the regions containing 95% of the distribution of limits expected under the background-only hypothesis. The thin solid black curves show the change in the observed limit due to variation of the signal cross sections within their theoretical uncertainties [60].	176
16-9 Cross section upper limit and exclusion contours at 95% CL for direct selectron production as a function of the $\tilde{\chi}_1^0$ and $\tilde{\ell}$ masses, assuming the production of both left- and right-handed selectrons (upper), or production of only left- (lower left) or right-handed (lower right) selectrons. The region under the thick red dotted (black solid) line is excluded by the expected (observed) limit. The thin red dotted curves indicate the regions containing 95% of the distribution of limits expected under the background-only hypothesis. For the right-handed selectrons, only the $+1\sigma$ expected line (thin red dotted curve) is shown as no exclusion can be made at -1σ . The thin solid black curves show the change in the observed limit due to variation of the signal cross sections within their theoretical uncertainties [60].	177

16-10	Cross section upper limit and exclusion contours at 95% CL for direct smuon production as a function of the $\tilde{\chi}_1^0$ and $\tilde{\ell}$ masses, assuming the production of both left- and right-handed smuons (upper), or production of only left- (lower left) or right-handed (lower right) smuons. The region under the thick red dotted (black solid) line is excluded by the expected (observed) limit. The thin red dotted curves indicate the regions containing 95% of the distribution of limits expected under the background-only hypothesis. For the right-handed smuons, only the $+1\sigma$ expected line (thin red dotted curve) is shown as no exclusion can be made at -1σ . The thin solid black curves show the change in the observed limit due to variation of the signal cross sections within their theoretical uncertainties [60].	178
A-1	QCD NNLO/NLO k factors for the $q\bar{q} \rightarrow ZZ$ process in generator level variables of the diboson system.	185

REFERENCES

- [1] E. Noether, “Invariante variationsprobleme,” *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, vol. 1918, pp. 235–257, 1918. [Online]. Available: <http://eudml.org/doc/59024>
- [2] G. Kane, *Modern elementary particle physics: explaining and extending the standard model; 2nd ed.* Cambridge: Cambridge University Press, 2017. [Online]. Available: <https://cds.cern.ch/record/2244793>
- [3] S. L. Glashow, “Partial Symmetries of Weak Interactions,” *Nucl. Phys.*, vol. 22, pp. 579–588, 1961.
- [4] A. Salam, “Weak and Electromagnetic Interactions,” *Conf. Proc.*, vol. C680519, pp. 367–377, 1968.
- [5] S. Weinberg, “A Model of Leptons,” *Phys. Rev. Lett.*, vol. 19, pp. 1264–1266, 1967.
- [6] M. Y. Han and Y. Nambu, “Three Triplet Model with Double SU(3) Symmetry,” *Phys. Rev.*, vol. 139, pp. B1006–B1010, 1965, [,187(1965)].
- [7] M. Gell-Mann, “Symmetries of baryons and mesons,” *Phys. Rev.*, vol. 125, pp. 1067–1084, 1962.
- [8] O. W. Greenberg, “Spin and Unitary Spin Independence in a Paraquark Model of Baryons and Mesons,” *Phys. Rev. Lett.*, vol. 13, pp. 598–602, 1964.
- [9] M. Tanabashi *et al.*, “Review of Particle Physics,” *Phys. Rev.*, vol. D98, no. 3, p. 030001, 2018.
- [10] D. J. Gross and F. Wilczek, “Ultraviolet Behavior of Nonabelian Gauge Theories,” *Phys. Rev. Lett.*, vol. 30, pp. 1343–1346, 1973, [,271(1973)].
- [11] H. D. Politzer, “Reliable Perturbative Results for Strong Interactions?” *Phys. Rev. Lett.*, vol. 30, pp. 1346–1349, 1973, [,274(1973)].
- [12] J. Ellis, M. K. Gaillard, and D. V. Nanopoulos, “A Historical Profile of the Higgs Boson. An Updated Historical Profile of the Higgs Boson,” no. arXiv:1504.07217. KCL-PH-TH-2015-20. LCTS-2015-10. CERN-PH-TH-2015-098, p. 22 p, Apr 2015, comments: 22 pages, 5 figures, update of arXiv:1201.6045, to be published in the volume. [Online]. Available: <https://cds.cern.ch/record/2012465>
- [13] N. Cabibbo, “Unitary symmetry and leptonic decays,” *Phys. Rev. Lett.*, vol. 10, pp. 531–533, Jun 1963. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.10.531>
- [14] 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.*, vol. B716, pp. 1–29, 2012.
- [15] S. Chatrchyan *et al.*, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC,” *Phys. Lett.*, vol. B716, pp. 30–61, 2012.

- [16] G. Bertone, D. Hooper, and J. Silk, “Particle dark matter: Evidence, candidates and constraints,” *Phys. Rept.*, vol. 405, pp. 279–390, 2005.
- [17] M. Aguilar *et al.*, “First Result from the Alpha Magnetic Spectrometer on the International Space Station: Precision Measurement of the Positron Fraction in Primary Cosmic Rays of 0.5350 GeV,” *Phys. Rev. Lett.*, vol. 110, p. 141102, 2013.
- [18] P. J. E. Peebles and B. Ratra, “The Cosmological constant and dark energy,” *Rev. Mod. Phys.*, vol. 75, pp. 559–606, 2003, [,592(2002)].
- [19] A. Refregier, “Weak gravitational lensing by large scale structure,” *Ann. Rev. Astron. Astrophys.*, vol. 41, pp. 645–668, 2003.
- [20] J. A. Tyson, G. P. Kochanski, and I. P. Dell’Antonio, “Detailed mass map of CL0024+1654 from strong lensing,” *Astrophys. J.*, vol. 498, p. L107, 1998.
- [21] S. P. Martin, “A Supersymmetry primer,” pp. 1–98, 1997, [Adv. Ser. Direct. High Energy Phys.18,1(1998)].
- [22] K. Olive, “Review of particle physics,” *Chinese Physics C*, vol. 40, no. 10, p. 100001, oct 2016. [Online]. Available: <https://doi.org/10.1088%2F1674-1137%2F40%2F10%2F100001>
- [23] N. Arkani-Hamed, S. Dimopoulos, and G. R. Dvali, “The Hierarchy problem and new dimensions at a millimeter,” *Phys. Lett.*, vol. B429, pp. 263–272, 1998.
- [24] C. T. Hill and E. H. Simmons, “Strong dynamics and electroweak symmetry breaking,” *Phys. Rept.*, vol. 381, pp. 235–402, 2003, [Erratum: Phys. Rept.390,553(2004)].
- [25] H. Georgi and S. L. Glashow, “Unity of All Elementary Particle Forces,” *Phys. Rev. Lett.*, vol. 32, pp. 438–441, 1974.
- [26] J. Wess and B. Zumino, “A Lagrangian Model Invariant Under Supergauge Transformations,” *Phys. Lett.*, vol. 49B, p. 52, 1974.
- [27] P. Fayet, “Supergauge Invariant Extension of the Higgs Mechanism and a Model for the electron and Its Neutrino,” *Nucl. Phys.*, vol. B90, pp. 104–124, 1975.
- [28] H. P. Nilles, “Supersymmetry, Supergravity and Particle Physics,” *Phys. Rept.*, vol. 110, pp. 1–162, 1984.
- [29] L. Pape and D. Treille, “Supersymmetry facing experiment: Much ado (already) about nothing (yet),” *Rept. Prog. Phys.*, vol. 69, pp. 2843–3067, 2006.
- [30] U. Amaldi, W. de Boer, and H. Furstenau, “Comparison of grand unified theories with electroweak and strong coupling constants measured at LEP,” *Phys. Lett.*, vol. B260, pp. 447–455, 1991.
- [31] R. Haag, J. T. Lopuszanski, and M. Sohnius, “All Possible Generators of Supersymmetries of the s Matrix,” *Nucl. Phys.*, vol. B88, p. 257, 1975, [,257(1974)].
- [32] S. R. Coleman and J. Mandula, “All Possible Symmetries of the S Matrix,” *Phys. Rev.*, vol. 159, pp. 1251–1256, 1967.
- [33] J. Ellis, “Supersymmetric Fits after the Higgs Discovery and Implications for Model Building,” *Eur. Phys. J. C*, vol. 74, no. arXiv:1312.5426. KCL-PH-TH-2013-46. LCTS-2013-31. CERN-PH-TH-2013-309,

p. 2732. 11 p, Dec 2013, comments: 11 pages, 11 figures, invited article in a special edition of EPJC (European Journal of Physics C) on. [Online]. Available: <https://cds.cern.ch/record/1638463>

- [34] G. R. Farrar and P. Fayet, “Phenomenology of the Production, Decay, and Detection of New Hadronic States Associated with Supersymmetry,” *Phys. Lett.*, vol. 76B, pp. 575–579, 1978.
- [35] R. Kaul and P. Majumdar, “Cancellation of quadratically divergent mass corrections in globally supersymmetric spontaneously broken gauge theories,” *Nuclear Physics B*, vol. 199, p. 3658, 05 1982.
- [36] D. J. H. Chung, L. L. Everett, G. L. Kane, S. F. King, J. D. Lykken, and L.-T. Wang, “The Soft supersymmetry breaking Lagrangian: Theory and applications,” *Phys. Rept.*, vol. 407, pp. 1–203, 2005.
- [37] A. H. Chamseddine, R. L. Arnowitt, and P. Nath, “Locally Supersymmetric Grand Unification,” *Phys. Rev. Lett.*, vol. 49, p. 970, 1982.
- [38] R. Barbieri, S. Ferrara, and C. A. Savoy, “Gauge Models with Spontaneously Broken Local Supersymmetry,” *Phys. Lett.*, vol. 119B, p. 343, 1982.
- [39] L. E. Ibanez, “Locally Supersymmetric SU(5) Grand Unification,” *Phys. Lett.*, vol. 118B, pp. 73–78, 1982.
- [40] L. J. Hall, J. D. Lykken, and S. Weinberg, “Supergravity as the Messenger of Supersymmetry Breaking,” *Phys. Rev.*, vol. D27, pp. 2359–2378, 1983.
- [41] J. R. Ellis, D. V. Nanopoulos, and K. Tamvakis, “Grand Unification in Simple Supergravity,” *Phys. Lett.*, vol. 121B, pp. 123–129, 1983.
- [42] M. Dine and W. Fischler, “A Phenomenological Model of Particle Physics Based on Supersymmetry,” *Phys. Lett.*, vol. 110B, pp. 227–231, 1982.
- [43] C. R. Nappi and B. A. Ovrut, “Supersymmetric Extension of the SU(3) x SU(2) x U(1) Model,” *Phys. Lett.*, vol. 113B, pp. 175–179, 1982.
- [44] M. Dine and A. E. Nelson, “Dynamical supersymmetry breaking at low energies,” *Phys. Rev. D*, vol. 48, no. hep-ph/9303230. SCIPP-93-03. UCSD-PTH-93-05, pp. 1277–1287. 28 p, Mar 1993. [Online]. Available: <https://cds.cern.ch/record/248065>
- [45] S. Dimopoulos and D. W. Sutter, “The Supersymmetric flavor problem,” *Nucl. Phys.*, vol. B452, pp. 496–512, 1995.
- [46] A. Djouadi *et al.*, “The Minimal supersymmetric standard model: Group summary report,” in *GDR (Groupement De Recherche) - Supersymetrie Montpellier, France, April 15-17, 1998*, 1998.
- [47] C. Borschensky, M. Krmer, A. Kulesza, M. Mangano, S. Padhi, T. Plehn, and X. Portell, “Squark and gluino production cross sections in pp collisions at $\sqrt{s} = 13, 14, 33$ and 100 TeV,” *Eur. Phys. J.*, vol. C74, no. 12, p. 3174, 2014.
- [48] S. Dimopoulos and G. F. Giudice, “Naturalness constraints in supersymmetric theories with nonuniversal soft terms,” *Phys. Lett.*, vol. B357, pp. 573–578, 1995.
- [49] R. Barbieri and D. Pappadopulo, “S-particles at their naturalness limits,” *JHEP*, vol. 10, p. 061, 2009.
- [50] M. Papucci, J. T. Ruderman, and A. Weiler, “Natural SUSY Endures,” *JHEP*, vol. 09, p. 035, 2012.

- [51] A. M. Sirunyan *et al.*, “Search for direct production of supersymmetric partners of the top quark in the all-jets final state in proton-proton collisions at $\sqrt{s} = 13$ TeV,” *JHEP*, vol. 10, p. 005, 2017.
- [52] Z. Han, G. D. Kribs, A. Martin, and A. Menon, “Hunting quasidegenerate Higgsinos,” *Phys. Rev.*, vol. D89, no. 7, p. 075007, 2014.
- [53] G. F. Giudice, T. Han, K. Wang, and L.-T. Wang, “Nearly Degenerate Gauginos and Dark Matter at the LHC,” *Phys. Rev.*, vol. D81, p. 115011, 2010.
- [54] S. Kraml, “SUSY status after one year of LHC,” 2012.
- [55] C. Autermann, “Experimental status of supersymmetry after the LHC Run-I,” *Prog. Part. Nucl. Phys.*, vol. 90, pp. 125–155, 2016.
- [56] A. Arbey, M. Battaglia, A. Djouadi, F. Mahmoudi, and J. Quevillon, “Implications of a 125 GeV Higgs for supersymmetric models,” *Phys. Lett.*, vol. B708, pp. 162–169, 2012.
- [57] J. Alwall, P. Schuster, and N. Toro, “Simplified Models for a First Characterization of New Physics at the LHC,” *Phys. Rev.*, vol. D79, p. 075020, 2009.
- [58] D. Alves, “Simplified Models for LHC New Physics Searches,” *J. Phys.*, vol. G39, p. 105005, 2012.
- [59] A. M. Sirunyan *et al.*, “Search for new phenomena in final states with two opposite-charge, same-flavor leptons, jets, and missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV,” *JHEP*, vol. 03, p. 076, 2018.
- [60] ——, “Search for supersymmetric partners of electrons and muons in proton-proton collisions at $\sqrt{s} = 13$ TeV,” *Phys. Lett.*, vol. B790, pp. 140–166, 2019.
- [61] ——, “Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13$ TeV using the CMS detector,” *JINST*, vol. 14, no. 07, p. P07004, 2019.
- [62] K. T. Matchev and S. D. Thomas, “Higgs and Z boson signatures of supersymmetry,” *Phys. Rev. D*, vol. 62, p. 077702, 2000.
- [63] P. Meade, M. Reece, and D. Shih, “Prompt decays of general neutralino NLSPs at the Tevatron,” *JHEP*, vol. 05, p. 105, 2010.
- [64] J. T. Ruderman and D. Shih, “General neutralino NLSPs at the early LHC,” *JHEP*, vol. 08, p. 159, 2012.
- [65] P. Z. Skands *et al.*, “SUSY Les Houches accord: Interfacing SUSY spectrum calculators, decay packages, and event generators,” *JHEP*, vol. 07, p. 036, 2004.
- [66] V. Khachatryan *et al.*, “Searches for electroweak production of charginos, neutralinos, and sleptons decaying to leptons and W, Z, and Higgs bosons in pp collisions at 8 TeV,” *Eur. Phys. J.*, vol. C74, no. 9, p. 3036, 2014.
- [67] O. S. Brning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, *LHC Design Report*, ser. CERN Yellow Reports: Monographs. Geneva: CERN, 2004. [Online]. Available: <https://cds.cern.ch/record/782076>
- [68] “Technical proposal: L3,” Tech. Rep. CERN-LEPC-83-5. LEPC-P-4, 1983. [Online]. Available: <https://cds.cern.ch/record/297266>

- [69] S. Chatrchyan *et al.*, “The CMS Experiment at the CERN LHC,” *JINST*, vol. 3, p. S08004, 2008.
- [70] W. W. Armstrong *et al.*, “ATLAS: Technical proposal for a general-purpose p p experiment at the Large Hadron Collider at CERN,” 1994.
- [71] “ALICE: Technical proposal for a large ion collider experiment at the CERN LHC,” 1995.
- [72] *LHCb : Technical Proposal*, ser. Tech. Proposal. Geneva: CERN, 1998. [Online]. Available: <https://cds.cern.ch/record/622031>
- [73] J. Haffner, “The CERN accelerator complex. Complexe des accrateurs du CERN,” Oct 2013, general Photo. [Online]. Available: <http://cds.cern.ch/record/1621894>
- [74] D. Boussard and T. Linnecar, “The LHC Superconducting RF System,” *Adv. Cryog. Eng.*, vol. 45A, pp. 835–844, 2000.
- [75] M. Schott and M. Dunford, “Review of single vector boson production in pp collisions at $\sqrt{s} = 7$ TeV,” *Eur. Phys. J.*, vol. C74, p. 2916, 2014.
- [76] G. Dissertori, “LHC Detectors and Early Physics,” in *Proceedings, 65th Scottish Universities Summer School in Physics: LHC Physics (SUSSP65): St. Andrews, UK, August 16-29, 2009*, 2010, pp. 197–230.
- [77] S. Chatrchyan *et al.*, “Commissioning of the CMS Experiment and the Cosmic Run at Four Tesla,” *JINST*, vol. 5, p. T03001, 2010.
- [78] “CMS Luminosity Measurements for the 2016 Data Taking Period,” CERN, Geneva, Tech. Rep. CMS-PAS-LUM-17-001, 2017. [Online]. Available: <http://cds.cern.ch/record/2257069>
- [79] *The CMS magnet project: Technical Design Report*, ser. Technical Design Report CMS. Geneva: CERN, 1997. [Online]. Available: <https://cds.cern.ch/record/331056>
- [80] S. Chatrchyan *et al.*, “Description and performance of track and primary-vertex reconstruction with the CMS tracker,” *JINST*, vol. 9, no. 10, p. P10009, 2014.
- [81] M. D’Alfonso *et al.*, “Validation tests of the CMS TIB/TID structures,” 2009.
- [82] D. J. A. Cockerill, “The CMS Electromagnetic Calorimeter at the LHC,” in *Proceedings, 34th International Conference on High Energy Physics (ICHEP 2008): Philadelphia, Pennsylvania, July 30-August 5, 2008*, 2008.
- [83] S. Chatrchyan *et al.*, “Energy Calibration and Resolution of the CMS Electromagnetic Calorimeter in pp Collisions at $\sqrt{s} = 7$ TeV,” *JINST*, vol. 8, p. P09009, 2013, [JINST8,9009(2013)].
- [84] S. Abdullin *et al.*, “Design, performance, and calibration of CMS hadron-barrel calorimeter wedges,” *Eur. Phys. J.*, vol. C55, no. 1, pp. 159–171, 2008.
- [85] *The CMS hadron calorimeter project: Technical Design Report*, ser. Technical Design Report CMS. Geneva: CERN, 1997, the following files are from [ja href=](#). [Online]. Available: <https://cds.cern.ch/record/357153>
- [86] S. Chatrchyan *et al.*, “Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV,” *JINST*, vol. 7, p. P10002, 2012.
- [87] V. Khachatryan *et al.*, “The CMS trigger system,” *JINST*, vol. 12, no. 01, p. P01020, 2017.

- [88] R. D. Ball *et al.*, “Parton distributions for the LHC Run II,” *JHEP*, vol. 04, p. 040, 2015.
- [89] R. Ball *et al.*, “Parton distributions from high-precision collider data,” *Eur. Phys. J.*, vol. C77, no. 10, p. 663, 2017.
- [90] 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*, vol. 07, p. 079, 2014.
- [91] M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani, “Matching matrix elements and shower evolution for top-quark production in hadronic collisions,” *JHEP*, vol. 01, p. 013, 2007.
- [92] C. Oleari, “The POWHEG-BOX,” *Nucl. Phys. Proc. Suppl.*, vol. 205-206, p. 36, 2010.
- [93] S. Alioli, P. Nason, C. Oleari, and E. Re, “NLO single-top production matched with shower in POWHEG: s - and t -channel contributions,” *JHEP*, vol. 09, p. 111, 2009, [Erratum: *JHEP* **02** (2010) 011].
- [94] 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.*, vol. 191, pp. 159–177, 2015.
- [95] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, “Parton Fragmentation and String Dynamics,” *Phys. Rept.*, vol. 97, pp. 31–145, 1983.
- [96] S. Agostinelli *et al.*, “GEANT4: A Simulation toolkit,” *Nucl. Instrum. Meth.*, vol. A506, pp. 250–303, 2003.
- [97] A. M. Sirunyan *et al.*, “Particle-flow reconstruction and global event description with the CMS detector,” *JINST*, vol. 12, no. 10, p. P10003, 2017.
- [98] “Search for supersymmetry in events with tau leptons and missing transverse momentum in proton-proton collisions at $\sqrt{s}=13$ TeV,” CERN, Geneva, Tech. Rep. CMS-PAS-SUS-17-002, 2017. [Online]. Available: <https://cds.cern.ch/record/2297162>
- [99] A. M. Sirunyan *et al.*, “Search for supersymmetry in events with a τ lepton pair and missing transverse momentum in proton-proton collisions at $\sqrt{s} = 13$ TeV,” *JHEP*, vol. 11, p. 151, 2018.
- [100] V. Khachatryan *et al.*, “Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $s = 8$ TeV,” *JINST*, vol. 10, no. 06, p. P06005, 2015.
- [101] A. M. Sirunyan *et al.*, “Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV,” *JINST*, vol. 13, no. 06, p. P06015, 2018.
- [102] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- k_t jet clustering algorithm,” *JHEP*, vol. 04, p. 063, 2008.
- [103] G. P. Salam, “Towards Jetography,” *Eur. Phys. J.*, vol. C67, pp. 637–686, 2010.
- [104] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, “Better jet clustering algorithms,” *JHEP*, vol. 08, p. 001, 1997.
- [105] V. Khachatryan *et al.*, “Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV,” *JINST*, vol. 12, no. 02, p. P02014, 2017.

- [106] A. M. Sirunyan *et al.*, “Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV,” *JINST*, vol. 13, no. 05, p. P05011, 2018.
- [107] D. Bertolini, P. Harris, M. Low, and N. Tran, “Pileup Per Particle Identification,” *JHEP*, vol. 10, p. 059, 2014.
- [108] V. Khachatryan *et al.*, “Performance of Photon Reconstruction and Identification with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV,” *JINST*, vol. 10, no. 08, p. P08010, 2015.
- [109] “2017 tracking performance plots,” Apr 2017. [Online]. Available: <https://cds.cern.ch/record/2290524>
- [110] T. Sjöstrand, S. Mrenna, and P. Skands, “PYTHIA 6.4 physics and manual,” *JHEP*, vol. 05, p. 026, 2006.
- [111] C. G. Lester and D. J. Summers, “Measuring masses of semiinvisibly decaying particles pair produced at hadron colliders,” *Phys. Lett.*, vol. B463, pp. 99–103, 1999.
- [112] A. Barr, C. Lester, and P. Stephens, “A variable for measuring masses at hadron colliders when missing energy is expected mt_2 the truth behind the glamour,” *Journal of Physics G: Nuclear and Particle Physics*, vol. 29, no. 10, pp. 2343–2363, sep 2003. [Online]. Available: <https://doi.org/10.1088%2F0954-3899%2F29%2F10%2F304>
- [113] “Search for new physics in final states with two opposite-sign, same-flavor leptons, jets, and missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV,” CERN, Geneva, Tech. Rep. CMS-PAS-SUS-16-034, 2017. [Online]. Available: <https://cds.cern.ch/record/2257293>
- [114] A. Kulesza and L. Motyka, “Threshold resummation for squark-antisquark and gluino-pair production at the LHC,” *Phys. Rev. Lett.*, vol. 102, p. 111802, 2009.
- [115] ——, “Soft gluon resummation for the production of gluino-gluino and squark-antisquark pairs at the LHC,” *Phys. Rev. D*, vol. 80, p. 095004, 2009.
- [116] W. Beenakker, M. Klasen, M. Kramer, T. Plehn, M. Spira, and P. M. Zerwas, “Production of Charginos, Neutralinos, and Sleptons at Hadron Colliders,” *Phys. Rev. Lett.*, vol. 83, p. 3780, 1999.
- [117] B. Fuks, M. Klasen, D. Lamprea, and M. Rothering, “Gaugino production in proton-proton collisions at a center-of-mass energy of 8 TeV,” *JHEP*, vol. 10, p. 081, 2012.
- [118] ——, “Precision predictions for electroweak superpartner production at hadron colliders with RESUMINO,” *Eur. Phys. J. C*, vol. 73, p. 2480, 2013.
- [119] B. Fuks, M. Klasen, D. R. Lamprea, and M. Rothering, “Revisiting slepton pair production at the Large Hadron Collider,” *JHEP*, vol. 01, p. 168, 2014.
- [120] L. Lista, “Practical Statistics for Particle Physicists,” in *Proceedings, 2016 European School of High-Energy Physics (ESHEP2016): Skeikampen, Norway, June 15-28 2016*, 2017, pp. 213–258.
- [121] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, “Asymptotic formulae for likelihood-based tests of new physics,” *Eur. Phys. J.*, vol. C71, p. 1554, 2011, [Erratum: *Eur. Phys. J.*C73,2501(2013)].
- [122] “CMS Luminosity Measurements for the 2016 Data Taking Period,” CERN, Geneva, Tech. Rep. CMS-PAS-LUM-17-001, 2017. [Online]. Available: <http://cds.cern.ch/record/2257069>

- [123] A. M. Sirunyan *et al.*, “Measurement of the inelastic proton-proton cross section at $\sqrt{s} = 13$ TeV,” *JHEP*, vol. 07, p. 161, 2018.
- [124] I. W. Stewart and F. J. Tackmann, “Theory Uncertainties for Higgs and Other Searches Using Jet Bins,” *Phys. Rev.*, vol. D85, p. 034011, 2012.
- [125] J. Butterworth *et al.*, “PDF4LHC recommendations for LHC Run II,” *J. Phys.*, vol. G43, p. 023001, 2016.
- [126] V. Khachatryan *et al.*, “Search for Physics Beyond the Standard Model in Events with Two Leptons, Jets, and Missing Transverse Momentum in pp Collisions at $\text{sqrt}(s) = 8$ TeV,” *JHEP*, vol. 04, p. 124, 2015.
- [127] G. Aad *et al.*, “Search for supersymmetry in events containing a same-flavour opposite-sign dilepton pair, jets, and large missing transverse momentum in $\sqrt{s} = 8$ TeV pp collisions with the ATLAS detector,” *Eur. Phys. J.*, vol. C75, no. 7, p. 318, 2015, [Erratum: Eur. Phys. J.C75,no.10,463(2015)].
- [128] E. Gross and O. Vitells, “Trial factors for the look elsewhere effect in high energy physics,” *Eur. Phys. J.*, vol. C70, pp. 525–530, 2010.
- [129] D. de Florian *et al.*, “Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector,” 2016.
- [130] M. Grazzini, S. Kallweit, D. Rathlev, and M. Wiesemann, “ $W^\pm Z$ production at hadron colliders in NNLO QCD,” *Phys. Lett.*, vol. B761, pp. 179–183, 2016.